The 13th National Conference on Computing and Information Technology

Proceedings of NCCIT 2017
The 13th National Conference on Computing and Information Technology
6th - 7th July 2017
at Arnoma Grand Hotel Bangkok, Thailand
www.nccit.net

Faculty of Information Technology
King Mongkut's University of Technology North Bangkok

Rajamangala University of Technology Thanyaburi

บทความวิจัย

การประชุมทางวิชาการระดับชาติด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศ ครั้งที่ 13
6-7 กรกฎาคม 2560
โรงแรม อโนมา แกรนด์ กรุงเทพฯ

คณะเทคโนโลยีสารสนเทศ
มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี

TABLEOF CONTENTS

Message from KMUTNB President...……...……ii
Message from RMUTT President....…….....iv
Message from RMUTTGeneral Chair...…….....vi
Message fromKMUTNBGeneral Chair…….....x
Conference Organizers...……..…….....xiv
Conference Organization Committee..……..xvi
Steering Committee..…..…….....xvii
Technical Program Committee...…..…….....xviii
Technical Program Contents.....xx
Keynote Speakers...…….....xxxvi
Regular Papers..…….....1
Author Index...…….....843

สารจากอธิการบดี
มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

ตามนโยบายของรัฐบาลในการขับเคลื่อนเศรษฐกิจด้วยเทคโนโลยีสู่ยุคประเทศไทย 4.0 หรือยุคเศรษฐกิจดิจิทัลคอมพิวเตอร์และเทคโนโลยีสารสนเทศเป็นกลไกสำคัญในการขับเคลื่อนยุคเศรษฐกิจดิจิทัล ซึ่งต้องอาศัยพื้นฐานจากระบบคอมพิวเตอร์และเทคโนโลยีสารสนเทศในการสนับสนุนการจัดการด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศ ต้องมีความทันสมัยสามารถตอบสนองต่อความต้องการของประชาชนและประเทศชาติ และมีคุณภาพไปสู่ความเป็นสากล ทั้งทางด้านการจัดการเรียนการสอนด้านอุตสาหกรรมและธุรกิจเอกชน ประเทศต่าง ๆ ทั่วโลกมีการลงทุนทรัพยากรเพื่อการวางโครงสร้างและปัจจัยพื้นฐานด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศ เพื่อให้การศึกษาของประชาชนในชาติพร้อมพัฒนานวัตกรรมการเรียนการสอน ที่มุ่งเน้นการนำเทคโนโลยีสารสนเทศมาใช้เป็นประโยชน์ได้อย่างเป็นรูปธรรมในการจัดการศึกษาทุกระดับ และเทคโนโลยีสารสนเทศจะช่วยให้การบริหารงานและการตัดสินใจต่างๆ เป็นไปอย่างรวดเร็วและมีประสิทธิภาพเพิ่มผลผลิตและผลในอุตสาหกรรมและธุรกิจได้กำไรทางการค้าที่สูง

การจัดงานประชุมวิชาการด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศอย่างต่อเนื่องทุก ๆ ปี เป็นเรื่องที่น่ายินดี โดยผู้จัดมุ่งหวังที่จะให้อาจารย์ นักวิจัย และนักศึกษา ที่มีผลงานวิจัยที่ค้นพบใหม่ ๆ ได้มานำเสนอและเผยแพร่สู่สาธารณะเป็นเวลาที่ทุกท่านจะได้มาพบปะแลกเปลี่ยนความรู้กัน แสวงหาความร่วมมือกันทั้งภาครัฐต่อรัฐ และภาครัฐกับเอกชน ทั้งนี้เป็นการส่งเสริมการพัฒนาศักยภาพ ยกระดับคุณภาพการวิจัยด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศ ที่จะนำพาประเทศไปสู่การแข่งขันในภูมิภาคและสากลต่อไป

ขอขอบคุณคณะกรรมการเครือข่ายหน่วยงานวิชาการ คณะกรรมการจัดงานทุกส่วนโดยเฉพาะคณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือที่มีส่วนสนับสนุนการจัดการประชุมทางวิชาการครั้งนี้ ขออวยพรให้การประชุมทางวิชาการครั้งนี้ประสบความสำเร็จบรรลุวัตถุประสงค์ที่ตั้งไว้ทุกประการ



(ศาสตราจารย์ ดร.สุชาติ เซี่ยงฉิน)
อธิการบดี มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

MESSAGE FROM KMUTNB PRESIDENT

An important government policy is to move forward Thailand to Digital Economy era so called “Thailand 4.0”. Computer and Information Technology is a key mechanism in Digital Economy, which relies on the basis of computer systems and information technology support. Computer and Information Technologies are important and must be well prepared to provide better respond to modern needs of the people and the nation with quality to be universal. Every country around the world has invested resources to structure and fundamentals of Computer and Information Technology to provide public education in national development and innovative teaching. The focus on Computer and Information Technology is a tangible benefit of education at all levels. As for industries and private businesses, Computer and Information Technology allows management and decision tasks to be a fast and effective way increasing productivity and profitability of trade.

The Conference on Computer and Information Technology has been continued for many years with good responses from the research communities. The organizers aim to provide a stage for the exchange of knowledge from university lecturers, researchers, and students with new research findings presented and published in the conference proceedings. It also provides a platform to seek for academic cooperation among the researchers joining the conference from both the state government and the private sector. This would improve the quality of research in Computer and Information Technology, taking the country to be able to compete both in regional and international.

I would like to thank everyone including the NCCIT collaborative universities, all co-organizers, especially the staff members of the Faculty of Information Technology, King Mongkut’s University of Technology North Bangkok to organize and manage this conference. I wish the conference achieves its expected.



(Professor Dr.-Ing.habil.Suchart Siengchin)

President, King Mongkut’s University of Technology North Bangkok

สารจากอธิการบดี
มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี

ในยุคของการเปลี่ยนผ่านทางดิจิทัลไปสู่ ประเทศไทย 4.0 เป็นที่ยอมรับกันโดยทั่วไปว่า การพัฒนาคุณภาพของทุนมนุษย์ในชาติได้กลายเป็นเงื่อนไขสำคัญต่อพัฒนาการของประเทศในปัจจุบันนี้ ท่ามกลางการเปลี่ยนแปลงทางด้านเทคโนโลยี สภาพเศรษฐกิจระหว่างประเทศที่มีการแข่งขันสูง สภาพปัญหาสังคมวัฒนธรรม และปัญหาสิ่งแวดล้อมของโลก การพัฒนาทางด้านกำลังคนของประเทศที่เข้มแข็งทั้งร่างกาย จิตใจ และสติปัญญา เพื่อพร้อมที่จะเผชิญกับปัญหาและความเปลี่ยนแปลงอย่างมีวิจารณญาณ ไม่ว่าจะในสังคมเมืองไปจนถึงสังคมระดับฐานราก จะเป็นหลักประกันความมั่นคงโดยรวมของประเทศในระยะยาว ในการนี้เทคโนโลยีดิจิทัล นับว่าเป็นกลไกสำคัญประการหนึ่ง ในการสนับสนุนให้มีการจัดการด้านการศึกษาที่ดีและมีคุณภาพไปสู่ความเป็นสากล ดังจะเห็นได้จากทุกประเทศที่เจริญแล้วต่างมีการลงทุนทรัพยากรเพื่อการวางโครงสร้างและปัจจัยพื้นฐานด้านเทคโนโลยีสารสนเทศ เพื่อการศึกษาของประชาชนในชาติพร้อมพัฒนานวัตกรรมการเรียนการสอน ที่มุ่งเน้นการนำเทคโนโลยีสารสนเทศมาใช้เป็นประโยชน์ได้อย่างเป็นรูปธรรมในการจัดการศึกษาทุกระดับ จุดมุ่งหมายสำคัญก็เพื่อยกระดับคุณภาพของทุนมนุษย์ในชาติให้ก้าวไกล ทั่วถึง และเท่าเทียมกันนั่นเอง

ขอขอบคุณหน่วยงาน คณะกรรมการจัดงาน และคณะบุคคลที่มีส่วนสนับสนุนในการจัดการประชุมทางวิชาการครั้งนี้ ขออวยพรให้การประชุมทางวิชาการครั้งนี้ประสบความสำเร็จบรรลุวัตถุประสงค์ทุกประการ



(รองศาสตราจารย์ ดร.ประเสริฐ ปิ่นปฐมรัฐ)
อธิการบดีมหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี

MESSAGE FROM RMUTT PRESIDENT

In the Digital Transformation to Thailand 4.0, it is generally accepted that the development of a nation stems from technical advancements and this has become the main key factor which dictates the development of any country. Many issues can affect development such as international economics, highly competitive markets, social and cultural differences, and global environmental problems. Thus, strengthening a country's capability to gain knowledge both physically and mentally provides the ability to deal with any critical issues, not only in an urban society but also surrounding countryside societies too. To ensure the overall stability of a country in the long term, digital technology is considered one of the most important mechanisms to support the management of educational quality and to provide the ability to continually improve it. This can be seen from many civilized countries that have invested resources and fundamental infrastructures for the deployment of Digital Technology to their education system. The development of innovative teaching and learning that focuses on bringing Information Technology to the forefront of education benefits the population as a whole. The main goal is to improve the quality of life and progress evenly and equally to a better future.

I would like to say a special thank to everybody involved in this conference, from partners to stakeholders, without you this would not be possible. And I hope this conference provides a good opportunity for all your voices to be heard.



(Associate Professor Dr.Prasert Pinpathomrath)
President, Rajamangala University of Technology Thanyaburi

สารจากประธานจัดงาน
มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี

หนังสือเล่มนี้เป็นเอกสารรวมเล่มรายงานผลวิจัยจากการประชุมวิชาการระดับชาติด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศ ครั้งที่ 13(NCCIT2017) ซึ่งจัดขึ้นที่โรงแรมแกรนด์ อโนมา กรุงเทพ ในวันที่ 6 – 7 กรกฎาคม 2560 ซึ่งจัดร่วมกับการประชุมวิชาการระดับนานาชาติด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศ ครั้งที่ 13 (The 13th International Conference on Computing and Information Technology: IC^{2}IT2017) เป็นการจัดร่วมกันของมหาวิทยาลัยเครือข่ายทั้งในประเทศและต่างประเทศ

สำหรับปีนี้ NCCIT2017 ได้รับบทความวิจัยจำนวน 265 ฉบับ จากสถาบันทั่วประเทศ โดยถูกส่งให้ผู้ทรงคุณวุฒิพิจารณาจำนวน 3 ท่าน และบทความที่ได้รับการตอบรับจากผู้ทรงคุณวุฒิอย่างน้อย 2 ใน 3 ท่าน ผลงานวิจัยที่ได้รับการตอบรับมีจำนวน 137 ผลงาน คิดเป็นร้อยละ 51.70 ที่นำเสนอปากเปล่าและนำบรรจุในเอกสารรวมเล่มรายงานผลวิจัยงานประชุมวิชาการ NCCIT2017 โดยผลงานวิจัยที่มีคุณภาพดีเยี่ยมจำนวนหนึ่งจะถูกคัดเลือกเพื่อพิจารณาลงพิมพ์ในวารสารเทคโนโลยีสารสนเทศ (ISSN 1685-8573) ซึ่งเป็นวารสารวิชาการที่อยู่ในฐานข้อมูล Thailand Journal Citation Index (TCI) กลุ่มที่ 1

คณะกรรมการจัดงานประชุม NCCIT2017 ขอขอบคุณนักวิจัยทุกท่านที่ส่งบทความวิจัยเพื่อพิจารณาทั้งผู้ที่ได้รับการตอบรับและผู้ที่ไม่ได้รับการตอบรับในการนำเสนอบทความวิจัย ขอขอบคุณคณะกรรมการพิจารณาบทความที่เสียสละเวลาอันมีค่าเพื่อพิจารณาบทความอย่างเข้มข้นได้มาตรฐาน ทั้งนี้ เพื่อให้ได้ผลงานวิจัยที่มีคุณภาพในการลงพิมพ์ในเอกสารรวมเล่มรายงานผลวิจัย และการนำเสนอในวันที่ประชุมวิชาการ NCCIT2017

ขอขอบพระคุณ รองศาสตราจารย์ ดร.ประเสริฐ ปิ่นปฐมรัฐ อธิการบดีมหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี สำหรับการสนับสนุนการประชุมวิชาการ NCCIT2017

ขอขอบคุณมหาวิทยาลัยเครือข่ายในประเทศทั้ง 11 แห่ง ที่ให้การสนับสนุนเป็นอย่างดีทั้งด้านการเงินและบุคลากร ในการร่วมเป็นกรรมการเครือข่าย กรรมการดำเนินงาน และกรรมการพิจารณาบทความ ได้แก่ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ คณะวิทยาการสารสนเทศ มหาวิทยาลัยมหาสารคาม คณะวิทยาศาสตร์ มหาวิทยาลัยอุบลราชธานี มหาวิทยาลัยราชภัฏกาญจนบุรี คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏนครปฐม สำนักวิทยบริการและเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีราชมงคลกรุงเทพ สำนักวิทยบริการและเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี คณะเทคโนโลยีและสิ่งแวดล้อมมหาวิทยาลัยสงขลานครินทร์ วิทยาเขตภูเก็ต คณะสถิติประยุกต์ สถาบันบัณฑิตพัฒนบริหารศาสตร์ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยราชภัฏเพชรบุรี และ คณะศิลปศาสตร์และวิทยาศาสาตร์ มหาวิทยาลัยราชภัฏศรีสะเกษ

ขอขอบคุณเป็นพิเศษสำหรับ รศ.ดร.อรุณี อินทรไพโรจน์ อาจารย์นิติ วิทยาวิโรจน์ ดร.วิเชษฐ์ พลายมาศ ผศ.ดร.อุไรวรรณ อินทร์แหยม ผศ.ดร.ณัทณรงค์ จตุรัส อาจารย์ปองพล นิลพฤกษ์ ผู้บริหารและเจ้าหน้าที่

สำนักวิทยบริการและเทคโนโลยีสารสนเทศ คณะกรรมการดำเนินงาน คณาจารย์จากคณะวิทยาศาสตร์และเทคโนโลยี คณะวิศวกรรมศาสตร์ คณะครุศาสตร์อุตสาหกรรม คณะบริหารธุรกิจ มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี

หวังเป็นอย่างยิ่งว่า NCCIT จะได้ให้โอกาสที่ดีสำหรับคณาจารย์ นักวิจัย นักวิชาการ และนักศึกษา ได้นำเสนอผลงานเป็นการพัฒนาตนเองไปสู่ความเป็นนักวิจัยที่เข้มแข็งยิ่งๆ ขึ้นไป

(นายพงศ์พิชญ์ ต่วนภูษา)
รองอธิการบดี มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี
ประธานจัดงานประชุมวิชาการ NCCIT2017

MESSAGE FROM RMUTT GENERAL CHAIR

This volume contains the papers of the 13thNational Conference on Computing and Information Technology (NCCIT2017) held at Arnoma Hotel Bangkok, Thailand, on July 6th – 7th, 2017. Traditionally, it is organized in conjunction with the International Conference on Computing and Information Technology (IC2IT), which is supported by many universities both in Thailand and overseas.

This year, the program committee of NCCIT2017 received 265 submissions from Thai institutions. Each submission was reviewed by at least 2, mostly 3 members of the program committee to avoid contradictory results. On these judgments, the committee decided to accept 137 papers (51.70% accepted rate) for oral presentation and inclusion in the conference proceedings of NCCIT2017. The papers from NCCIT2017 with excellent quality will be selected for consideration and hopefully published in the Information Technology Journal (ISSN 1685-8573), which is included in a database of academic journals in Thailand Journal Citation Index (TCI) Group 1.

A special thanks is given to Associate Professor Dr.Prasert Pinpathomrath, RMUTT President, for his support of NCCIT from the first year onward, and for providing us with the necessary resources from RMUTT.

I would like to also thank the 11 co-operative universities for supporting both in finance and personnel, which are Faculty of Information Technology, King Mongkut's University of Technology North Bangkok (KMUTNB), Faculty of Information Science Mahasarakham University (MSU), Faculty of Science, UbonRatchathani University (UBU),KanchanaburiRajabhat University (KRU), Faculty of Science and Technology, NakhonPathomRajabhat University (NPRU), Office of Academic Resources and Information Technology, Rajamangala University of Technology Thanyaburi (RMUTT), Office of Academic Resources and Information Technology, Rajamangala University of Technology Krung Thep(RMUTK), Faculty of Technology and Environment, Prince of SongklaUniversity Phuket Campus (PSU), Faculty of Applied Statistics, National Institute of Development Administration (NIDA), Faculty of Information Technology, Phetchaburi Rajabhat University (PRU), and Faculty of Arts and Sciences, SrisaketRajabhat University (SSRU).

A very special and warm thank you also goes to Assoc.Prof.Dr.Arunee Intrapairoj, Mr.Niti Witthayawiroj,Dr.Wichet Plaimart, Asst.Prof.Dr.Uraiwan Inyaem, Asst.Prof.Dr.Natnarong Jaturat, Mr.Pongpon Nilaphruek, the Committee from multi-faculty RMUTT, management and staffs of the Office of Academic Resources and Information TechnologyRMUTTand the

supportive research team; the proceedings could not have been completed in a timely manner and correct format, without the painstaking work. We would like to also thank all authors for their submissions and the members of the program committee for their great work and valuable time.

Hopefully, NCCIT will provide good opportunities for faculty, researchers, scholars and students to present their work to develop themselves to become qualified researchers.



(Pongpith Tuenpusa)
Vice President, Rajamangala University of Technology Thanyaburi
NCCIT2017 General Chair

สารจากประธานเครือข่าย มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

กิจกรรมเครือข่ายการประชุมวิชาการด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศเริ่มต้นตั้งแต่ปี พ.ศ. 2548 ด้วยความร่วมมือของหน่วยงานเครือข่ายวิชาการจัดงานประชุมวิชาการ NCCIT และ IC^{2}IT เป็นประจำทุกปีมา นับเป็นเวลาร่วม 13 ปี โดยมีวัตถุประสงค์เพื่อเป็นเวทีสำหรับการนำเสนอผลงานวิจัยด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศ ของอาจารย์ นักศึกษา และนักวิจัย จากสถาบันต่างๆ ทั้งในประเทศและต่างประเทศได้พบปะกันทั้งภาครัฐและภาคธุรกิจ เป็นการเผยแพร่ผลงานวิจัยด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศ จากทฤษฎีสู่งานประยุกต์เพื่อขับเคลื่อนสังคม ที่ผ่านมางานประชุมวิชาการ NCCIT และ IC^{2}IT จะมีคณาจารย์ นักศึกษา นักวิจัย ให้ความสนใจส่งผลงานวิจัยเป็นจำนวนมาก โดยเฉพาะอย่างยิ่งงานประชุมวิชาการ IC^{2}IT ได้รับการตีพิมพ์โดยสำนักพิมพ์ Springer-Verlagซึ่งอยู่ในฐานข้อมูล Scopus, ISI Proceedings และฐานข้อมูลนานาชาติอื่นๆ ที่มีชื่อเสียงเป็นที่ยอมรับระดับนานาชาติ ส่วนงานวิจัยภาษาไทย NCCIT ที่มีคุณภาพระดับดีเด่นจะได้รับการขยายเพิ่มเติมและลงพิมพ์ในวารสารเทคโนโลยีสารสนเทศ ซึ่งอยู่ในฐานข้อมูลTCI กลุ่มที่ 1

ขอขอบคุณนักวิจัยทุกท่านสำหรับการเข้าร่วมนำเสนอผลงานใหม่ๆ และมีคุณภาพสูง ทั้งนี้มีบทความจำนวนมากที่ไม่ได้รับการตอบรับเข้าร่วมนำเสนอ เนื่องจากความเข้มข้นของการพิจารณาจาก คณะกรรมการที่ต้องการรักษาคุณภาพและมาตรฐานงานวิจัย NCCIT ขอขอบคุณคณะกรรมการพิจารณาบทความที่เสียสละเวลาอันมีค่าเพื่อให้ได้ผลงานวิจัยที่มีคุณภาพในการลงพิมพ์ในเอกสารรวมเล่มรายงานผลวิจัย และการนำเสนอในที่ประชุมวิชาการ NCCIT2017

ขอขอบคุณมหาวิทยาลัยเครือข่ายในประเทศไทยทั้งสิบเอ็ดแห่งที่ให้การสนับสนุนเป็นอย่างดีทั้งด้านการเงิน และบุคลากรในการร่วมเป็นกรรมการเครือข่าย กรรมการดำเนินงาน และกรรมการพิจารณาบทความ ซึ่งปีพ.ศ. 2560 หน่วยงานเครือข่าย 11 แห่ง ประกอบด้วยหน่วยงานจากสถาบันอุดมศึกษา ได้แก่ 1. คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ (KMUTNB) 2. คณะวิทยาการสารสนเทศ มหาวิทยาลัยมหาสารคาม (MSU) 3. คณะวิทยาศาสตร์ มหาวิทยาลัยอุบลราชธานี (UBU)4. มหาวิทยาลัยราชภัฏกาญจนบุรี (KRU) 5. คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏนครปฐม (NPRU) 6. สำนักวิทยบริการและเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีราชมงคลกรุงเทพ (RMUTK)7. สำนักวิทยบริการและเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี (RMUTT) 8. คณะเทคโนโลยีและสิ่งแวดล้อม มหาวิทยาลัยสงขลานครินทร์ วิทยาเขตภูเก็ต (PSU) 9. คณะสถิติประยุกต์ สถาบันบัณฑิตพัฒนบริหารศาสตร์ (NIDA) 10. คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยราชภัฏเพชรบุรี และ 11. คณะศิลปศาสตร์และวิทยาศาสาตร์ มหาวิทยาลัยราชภัฏศรีสะเกษ

สำหรับปีนี้มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรีให้เกียรติรับเป็นเจ้าภาพในการจัดงาน NCCIT2017 ขอบคุณรองศาสตราจารย์ ดร.ประเสริฐ ปิ่นปฐมรัฐ อธิการบดี ที่สนับสนุน และขอขอบคุณอาจารย์พงศ์พิชญ์

ต่วนภูยา รองอธิการบดี ประธานจัดงาน NCCIT2017 อาจารย์นิติ วิทยาวิโรจน์ผู้อำนวยการสำนักวิทยบริการและเทคโนโลยีสารสนเทศและทีมงานของมหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี ซึ่งทำหน้าที่เป็นเจ้าภาพในการจัดงานได้อย่างมีประสิทธิภาพ

ขอขอบพระคุณ ศาสตราจารย์ ดร. สุชาติ เชี่ยงฉิน อธิการบดีมหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ สำหรับการสนับสนุนการประชุมวิชาการ NCCIT อย่างต่อเนื่อง ขอขอบคุณผู้ช่วยศาสตราจารย์ ดร.สุนันทา สดสี นางสาววัชรีวรรณ จิตต์สกุล และทีมงานจากคณะเทคโนโลยีสารสนเทศมหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ ที่ได้ประสานงานกับนักวิจัยและผู้ที่เกี่ยวข้องเพื่อจัดรวบรวม ทำเล่มรายงานผลงานวิจัยสำหรับงานประชุมให้สำเร็จในเวลาที่เหมาะสม ขอขอบคุณผู้ช่วยศาสตราจารย์จีระศักดิ์ นำประดิษฐ์ และนายอมรศักดิ์ อมรธนานันท์ ซึ่งเป็นผู้พัฒนาและดูแลเว็บไซต์ของงานประชุมเป็นอย่างดีขอขอบคุณคณาจารย์และเจ้าหน้าที่ของมหาวิทยาลัยเครือข่ายที่ให้การสนับสนุนด้านบุคคลากรและงบประมาณ

ท้ายนี้ผมหวังเป็นอย่างยิ่งว่างานประชุมวิชาการ NCCIT จะเป็นเวทีการนำเสนอผลงานวิจัยและนำไปสู่การพัฒนาตนเองเพื่อเป็นนักวิจัยดีมีคุณภาพสูงยิ่ง ๆ ขึ้นไป NCCIT จะเป็นแพลตฟอร์มสำหรับการแลกเปลี่ยนความรู้ ในด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศและจะสร้างแรงบันดาลใจให้นักวิจัยเพื่อสร้างความคิดใหม่การคิดค้นนวัตกรรม และพบพันธมิตรใหม่ ๆเป็นเครือข่ายการวิจัยสู่การพัฒนาชาติอย่างยั่งยืนต่อไป



(รองศาสตราจารย์ ดร. พยุง มีสัจ)
คณบดีคณะเทคโนโลยีสารสนเทศ มจพ.
ประธานเครือข่ายงานประชุมวิชาการ NCCIT

MESSAGE FROM KMUTNB GENERAL CHAIR

Our academic activities on Computer and Information Technology conferences started in 2005 with the cooperation of academic networking organizations to held NCCIT and IC^{2}IT conferences every year. It has been 13 years of the conferences to provide a platform for the presentation of research in computer and information technology to teachers, students and researchers from various institutions, both domestically and internationally. The researchers can meet together both the government and the business sectors. Research papers to contribute in the conferences are in the main areas and trends on computer and IT from theory to applications for social mobilization. There are a number of faculty members, researchers, and researchers, as usual, attending the NCCIT and IC^{2}IT conferences. The IC^{2}IT papers have been published by Springer-Verlag, which is indexed in the Scopus database as well as ISI proceedings and other famous international databases. For NCCIT, outstanding papers will be expanded to publishin the Information Technology Journal, which is indexed in TCI Group 1.

I would like to thanks to all the researchers for submitting their new research works with high quality to publish in the NCCIT proceedings. It is regret that there are many articles not been accepted to present due to the concentration of reviews from the technical committee to maintain the quality and standards of NCCIT.In addition, I would like to give a special Thanks to the technical committee members for considering articles that sacrifice valuable time to obtain quality research resulting in the publication of the research papers presented at NCCIT 2017.

With great helps from the elevenuniversity partners in Thailand for their financial support and staff to become a network committee, executive director committee, and technical program committee. In the year 2017, main networking organizations from higher education institutions are1)Faculty of Information Technology, King Mongkut's University of Technology North Bangkok (KMUTNB), 2)Faculty of Information Science Mahasarakham University (MSU), 3)Faculty of Science, UbonRatchathani University (UBU), 4)KanchanaburiRajabhat University (KRU), 5)Faculty of Science and Technology, NakhonPathomRajabhat University (NPRU), 6)Office of Academic Resources and Information Technology, Rajamangala University of Technology Thanyaburi (RMUTT), 7)Office of Academic Resources and Information Technology, Rajamangala University of Technology Krung Thep(RMUTK), 8)Faculty of Technology and Environment, Prince of SongklaUniversity Phuket Campus (PSU), 9)Faculty of Applied Statistics, National Institute of Development Administration (NIDA), 10)Faculty of Information Technology, Phetchaburi Rajabhat University (PRU), and 11) Faculty of Arts and Sciences, SrisaketRajabhat University (SSRU).

For this year, Rajamangala University of Technology Thanyaburiis honored to host the NCCIT 2017. I would like to thanks Associate Professor Dr.Prasert Pinprathomrat, RMUTT President, in support of hosting the conference. In addition, I would like to give big thanks to Mr.Pongpith Tuenpusa, RMUTT Vice President, General Chair of NCCIT2017 and Mr.Niti Wittayawirote, Director of Office of Academic Resource and Information Technology, as well as his team for doing such a good job on the event management. It serves as the host of the event effectively.

I would like to thank Professor Dr.-Ing.habil.Suchart Siengchin, President of King Mongkut's University of Technology North Bangkok (KMUTNB) for his support of NCCITand for providing us with the necessary resources from KMUTNB. A very special thank goes to Assistant Professor Dr.Sunantha Sodsee, Ms.Watchareewan Jitsakul, and academic support team from Faculty of Information Technology, King Mongkut's University of Technology North Bangkok,for collaborating, collecting and making quality proceedings to publish on time. Thanks to Assistant Professor Jeerasak Numpradit, and Mr.Armornsak Armornthananun, the developer and webmaster of the conference.

Finally, I hope that the NCCIT will continue to serve community as a platform for research, leading to self-development to set high quality research standard. It will be a forum for knowledge exchange in computer and informationtechnology. It will inspire researchers for creating new ideas, innovation discovery, and finding new partners as well as to build up a research network for national sustainable development.



(Associate Professor Dr. Phayung Meesad)

Dean, Faculty of Information Technology

King Mongkut's University of Technology North Bangkok

NCCIT 2017 General Chair

CONFERENCE ORGANIZERS

King Mongkut's University of Technology North Bangkok, Thailand
Faculty of Information Technology

Rajamangala University of Technology Thanyaburi, Thailand
Office of Academic Resources and Information Technology

FernUniversität in Hagen, Germany

Chemnitz University, Germany

Oklahoma State University, USA

Edith Cowan University, Australia

Hanoi National University of Education, Vietnam

Gesellschaft für Informatik (GI) , Germany

Mahasarakham University, Thailand
Faculty of Information Technology

Ubon Ratchathani University, Thailand
Faculty of Science

Kanchanaburi Rajabhat University, Thailand

Nakhon Pathom Rajabhat University, Thailand
Faculty of Scienceand Technology

CONFERENCE ORGANIZERS



Phetchaburi Rajabhat University, Thailand

Faculty of Information Technology



Rajamangala University of Technology Krungthep, Thailand

Center of Academic Resources and Information Technology



Prince Of Songkla University, Phuket Campus, Thailand

Faculty of Technology and Environment

National Institute of Development Administration, Thailand

Department of Information Technology and Systems Management, School of Applied Statistics





Sisaket Rajabhat University, Thailand

Council of IT Deans of Thailand (CITT) , Thailand



CONFERENCE ORGANIZATION COMMITTEE

General Chair:	Mr.Pongpith Tuenpusa, RMUTT
General Chair:	Assoc.Prof.Dr.Phayung Meesad, KMUTNB

STEERING COMMITTEE

Assoc.Prof.Dr.Utit Inprasit, UBU

Assoc.Prof.Dr.Piya Kovintavewat, NPRU

Assoc.Prof.Dr.Pun Thongchumnum, PSU

Asst.Prof.Dr.Sujin Butdisuwan, MSU

Asst.Prof.Dr.Photjanee Sukchovna, KRU

Asst.Prof.Dr.Veerachai Khonchoho, PBRU

Asst.Prof.Dr.Sutep Tongngam, NIDA

Asst.Prof.Dr.Sunantha Sodsee, KMUTNB

Dr.Thawatchai Sarawong, RMUTK

Dr.Jedsada Chatree, SSKRU

TECHNICAL PROGRAM COMMITTEE

Adisak Intana, PSU
Adisak Sangsongfa, KMUTNB
Alisa Kongthon, NECTEC
Annupan Rodtook, RU
Ariya Namvong, KKU
Arunee Inttrapairoj, RMUTT
Aziz Nanthaamornphong, PSU
Burasakorn Yoosooka, RMUTT
Charun Sanrach, KMUTNB
Chatchawin Namman, UBU
Chatklaw Jareanpon, MSU
Chitsutha Soomlek, KKU
Choochart Haruechaiyasa, NECTEC
Dech Thammasiri, NPRU
Gamgarn Somprasertsri, MSU
Jamornkul Laokietkul, Chandra
Jantima Polpinij, MSU
Jaree Thongkam, MSU
Jaturapith Krohkaew, RMUTT
Jiratta Phuboon-Ob, MSU
Jirawat Tanthong, PSU
Julaluk Watthananon, RMUTT
Kairung Hengproprohm, NPRU
Kanjana Thongglin, PSU
Keeratiburt Kanchanasatian, RMUTT
Khachakrit Liamthaisong, IT MSU
Khongthep Boonmee, RMUTT
Kiattisak Yochanang, KMUTNB
Kitipoom Vipahasna, RMUTT
Kittiwann Nimkerdphol, RMUTT
Korawit Prutsachainimmit, PSU
Kornchawal Chaipah, KKU
Kriengsak Treeprapin, UBU
Kuntpong Woraratpanya, KMITL
Kwankamon Dittakan, PSU
Mahasak Ketcham, KMUTNB
Maleerat Sodanil, KMUTNB
Manasawee Kaenampornpan, MSU
Meechai Lohakan, Kmutnb
Meta Sirikul, RMUTT
Metee Pigultong, RMUTT
Montean Rattanasiriwongwut, KMUTNB
Nachayapat Rodprayoon, RMUTT
Nachirat Rachburee, RMUTT
Nadh Ditcharoen, UBU
Nalinpat Porrawatpreyakorn, KMUTNB
Narong Phothi, KMUTNB
Narongrit Masusai, MSU
Natnarong Jaturat, RMUTT
Nattapong Tongtep, PSU
Nattavee Utakrit, KMUTNB
Nawaporn Wisitpongphan, KMUTNB
Niti Witthayawiroj, RMUTT
Nongluk Promthong, RMUTT
Noppasak Tantisattayanon, RMUTR
Nuttachot Promrit, SU
Ohm Sornil, NIDA
Padej Phomasakha Na Sakolnakorn, CU
Panana Tangwannawit, PCRU
Panida Songram, MSU
Paranya Palwisut, NPRU
Parinya Sanguansat, PIM
Passakon Prathombutr, NECTEC
Patharawut Saengsiri, TISTR

TECHNICAL PROGRAM COMMITTEE

Patrapee Sunantapot, RMUTT
Phatthanaphong Chomphuwiset, MSU
Phayung Meesad, KMUTNB
Piyanan Pannim, RMUTT
Piyawad Kasabai, MSU
Pongpisit Wuttidittachotti, KMUTNB
Pongpon Nilaphruek, RMUTT
Prachyanun Nilsook, KMUTNB
Prapas Thongrak, RMUTT
Prasong Praneetpolgrang, SU
Pudsadee Boonrawd, KMUTNB
Rungsan Suwannahong, RMUTT
Rungtip Charoensak, MSU
Sajjaporn Waijanya, SU
Sakchai Tangwannawit, KMUTNB
Saowalak Arampongsanuwat, KMUTNB
Siranee Nuchitprasitchai, MTU
Sirapat Boonkrong, KMUTNB
Sithiporn Kerdsumang, kmutnb
Sompong Valuvanathorn, UBU
Sucha Smanchat, KMUTNB
Suchart Khummanee, MSU
Sujin Butdisnwan, MSU
Suksaeng Kukanok, IBASOFT
Sumitra Nuanmeesri, RSU
Sumonta Kasemvilas, KKU
Sunantha Sodsee, KMUTNB
Supakit Nootyaskool, KMITL
Supoj Hengpraprohm, NPRU
Surapong Auwatanamongkol, NIDA
Sureerut Inmor, RMUTT
Tanapon Jensuttiwetchakul, KMUTNB
Tanasai Sucontphunt, NIDA
Thawatchai Sahapong, RMU
Thawatchai Sarawong, RMUTK
Thippaya Chintakovid, KMUTNB
Thongchai Chuachan, Surindra RU
Uraiwan Inyaem, RMUTT
Vatinee Nuipian, KMUTNB
Veerachai Khonchoho, Phetchaburi RU
Wachirawut Thamviset, KKU
Waraphan Sarasureeporn, RMUTT
Watchareewan Jitsakul, KMUTNB
Wattana Punlumjeak, RMUTT
Weena Janratchakool, RMUTT
Werayut Pimpaporn, KU
Wichet Plaimart, RMUTT
Winai Nadee, PSU
Wongkot Sriurai, UBU
Youdthachchai Lertworaprachaya, RMUTT

TECHNICAL PROGRAM CONTENTS

Thursday 6th July 2017	
08:00 – 09:00	**Registration**
09:00 – 09:15	Report (in Thai) by **Assoc.Prof.Dr.Phayung Meesad** *Dean of Faculty of Information Technology,* *King Mongkut's University of Technology North Bangkok* Report (in Thai) by **Mr.Pongpith Tuenpusa** *Vice President* *Rajamangala University of Technology Thanyaburi* Report by **Prof.Dr.-Ing.habil.Herwig Unger** *IC²IT2017 Technical Program Chair,* *Fern University in Hagen, Germany*
09:15 – 09:30	Welcome Speech by **Assoc.Prof.Dr.Prasert Pinpathomrat** *President of Rajamangala University of Technology Thanyaburi*
09:30 – 09:45	Opening Ceremony by **Prof.Dr.-Ing.habil.Suchart Siengchin** *President of King Mongkut's University of Technology North Bangkok*
09:45 – 10:45	Invited Keynote Speech by: **Prof.Dr.Jie Lu** • Associate Dean in Research in the Faculty of Engineering and Information Technology at the University of Technology Sydney (UTS).
	Topic: Fuzzy Transfer Learning for Prediction and Decision Support
10:45 – 11:00	*Coffee Break*
11:00 – 12:00	Invited Keynote Speech by: **Prof.Dr.Hanno Lefmann** • Professor for Theoretical Computer Science and Information Security at Chemnitz University of Technology, Germany.
	Topic: Testing and Estimating Graph Parameters for Monotone Properties
12:00 – 13:00	*Lunch*
13:00 – 18:00	**Paper Presentation**
18:00 – 22:00	**Welcome Dinner**

Thursday 6th July 2017		
ROOM 1: Data Science and Machine Learning		**Page**
13.00 – 13.20 NCCIT2017-265	**Optimization Problems for Planthopper Algorithm** *Kaewjai Apornpisarn,Phayung Meesad and Vatinee Nuipian*	1
13:20 – 13:40 NCCIT2017-005	**The Comparison of Data Mining Techniques for Predicting Student Success in Vocation Education** *Pattanaphong Donratand Jaree Thongkam*	8
13:40 – 14:00 NCCIT2017-033	**The Suspicious Transactions Forecasting using Neural Network** *Narroup Rak-ngam and Narongdech Keeratipranon*	14
14.00 - 14.20 NCCIT2017-217	**Ensemble for Highly Imbalanced Text Classification base on Deep Learning with Feature Selection and Data Balancing Techniques** *Supathep Satiman,Phayung Meesad and Piyanoot Voraboot*	20
14:20 – 14:40 NCCIT2017-211	**NER Detection from Thai Tweets Related to Natural Disaster** *Maleerat Sodanil and Yuttana Lungkatoong*	26
14:40 – 15:00 NCCIT2017-047	**Predicting Student's Data for Performance by Data Mining** *Chotika Iamprik and Kwanruetai Sudadet*	32
15:00 – 15:20	*Coffee Break*	
15:20 – 15:40 NCCIT2017-065	**Forecasting Number of Dengue Patients Using Time Series Technique** *Trirat Soemsap,Sureechay Meemoo,Patcharee Oankaew and Pornvimon Klongsangson*	38
15:40 – 16:00 NCCIT2017-070	**Analysis of Affecting Factors to Customer Reviews using Opinion Mining** *Pattaraporn Chotnasiaw,Wararat Songpan,Somjit Arch-int and Saiyan Saiyod*	44
16:00 – 16:20 NCCIT2017-075	**Predictive Analytic for Student Dropout in High Vocational Certificate Using Data Mining Technique** *Sakkarin Phupanna,Paweena Wanchai,Wararat Songpan,and Nunnapus Benjamas*	51
16:20 – 16:40 NCCIT2017-111	**Drug Searching for Targeted Cancer Therapy by Centrality Depth First Search on Cell Signal Transduction Pathway** *Massaya Samnienggam and Tanasanee Phienthrakul*	57
16:40 – 17:00 NCCIT2017-119	**Potential Customer Recommendation on Twitter for Local Woven Fabric Products** *Nichakorn Pankong, SaowalakArampongsanuwat and Narisara Nakmaetee*	63
17:00 – 17:20 NCCIT2017-131	**Perfume Classification using Electronic Nose** *Panida Lorwongtrakool,Chaloemchai Lowongtrakool and Nitinan Mata*	70
17:20 – 17:40 NCCIT2017-230	**Cluster and Analysis Based on Multiple Intelligences Theory Using Data Mining** *Narongdech Rattananonsathien,Apichat Buatoom and Monsawan Ploymookda*	76
18:00 – 22:00	***Welcome Dinner and Best Paper Award Presentation***	

Thursday 6th July 2017		
ROOM 2: Data Network and Communication		**Page**
13:00 – 13:20 NCCIT2017-013	**Distance Estimation Through Mobile Sensing** *Puttipong Leakkaw and Sooksan Panichpapiboon*	82
13:20 – 13:40 NCCIT2017-058	**A System of Real-time Bus Tracking on Mobile Phone** *Suthat Banchong,Buncha Joompol and Jantima Polpinij*	88
13:40 – 14:00 NCCIT2017-086	**Gateway Application for Smartphone Ad hoc Network and Internet** *Watcharathorn .Laikhunakhorn, Tawatchai Kovicsakul,Tananop Thongthavorn and Sumet Prabhavat*	94
14.00 - 14.20 NCCIT2017-221	**Path Selection Algorithm for Multipath TCP in Software-Defined Networking** *Jirasak Jullawat,Kiattikun Kawila,Kulit Na Nakorn and Kultida Rojviboonchai*	100
14:20 – 14:40 NCCIT2017-227	**Accessories for wheelchair by using Internet of Things technology** *Nuttakit Vatcharatiansakul, Sithikorn Charoensub and Panwit Tuwanut*	106
14:40 – 15:00 NCCIT2017-257	**An Auto-scaling Mechanism in Docker System** *Teerapat Khunpech and Sucha Smanchat*	113
15:00 – 15:20	*Coffee Break*	
15.20 - 15.40 NCCIT2017-228	**An Architecture for Data Storage System on Multiple Cloud Service Providers** *Manutchai Preechahan and Sucha Smanchat*	121
15:40 – 16:00 NCCIT2017-241	**Applying Information Storage and Retrieval Technique for Auditing Information Technology Security Policy Enforcement** *Ruttasiri Tessiri and Nakornthip Prompoon*	126
16:00 – 16:20 NCCIT2017-256	**Temperature effect on Transmission Power Control for Wireless Sensor Networks** *Wilawan Rukpakavong and Sirikunya Nilpanich*	132
16:20 – 16:40 NCCIT2017-008	**Anti-Theft Motorcycle System Using Face Recognition via Cloud Computing Under Concept on Internet of Things** *Patcharaon Petchsamutr and Mahasak Ketcham*	137
16.40 – 17:00 NCCIT2017-061	**AES: Performance Evaluation of Symmetric Encryption** *Luxsanan Ploywattanawong and Sirapat Boonkrong*	144
17:00 – 17:20 NCCIT2017-067	**Comparison of Integer Factorization Using Fermat and Pollard Rho Algorithm** *Parinya Natho and Sirapat Boonkrong*	150
17:20 – 17.40 NCCIT2017-068	**An IoT Hub System for Named Data Networking** *Thongchai Chuachan,Pattarapat Waidee and Sirimongkon Meesiri*	156
18:00 – 22:00	***Welcome Dinner and Best Paper Award Presentation***	

Thursday 6th July 2017		
ROOM 3: Human-Computer Interface and Image Processing, Data Science and Machine Learning		**Page**
13:00 – 13:20 NCCIT2017-125	**Emotional Speech Recognition Using Context-Dependent Variation** *Nithaya Sakolpisit,Sippawit Saewang and Kuntpong Woraratpanya*	163
13:20 – 13:40 NCCIT2017-123	**Emotional Speech Recognition using Two-level Classification** *Sippawit Saewang,Nithaya Sakolpisit and Kuntpong Woraratpanya*	169
13:40 – 14:00 NCCIT2017-152	**Random Number Gird Password** *Wiriya Lawanpreda and Nuttanont Hongwarittorrn*	175
14:00 – 14:20 NCCIT2017-179	**Using Deep Learning for Printed Thai Characters Recognition** *Pongkun Pankerd, Tuangsit Tirabutr, Thotsapon Tamduang and Sunee Pongpinigpinyo*	181
14:20 – 14:40 NCCIT2017-180	**Rectifying Chromatic Aberration using Guided Image Filtering** *Pichamon Sinlalab, Porawat Visutsak and San Ratanasanya*	187
14:40 – 15:00 NCCIT2017-182	**Improvement on Guided Image Filter for Chromatic Aberration Detection** *Yuttasak Jinaporn, Porawat Visutsak and San Ratanasanya*	193
15:00 – 15:20	*Coffee Break*	
15:20 – 15:40 NCCIT2017-247	**A Video-Based People Counting System in indoor spaces** *Manoch Pracha and Sirichai Triamlumlerd*	199
15:40 – 16:00 NCCIT2017-097	**A Comparison of Image Retrieval and Classification based on Shape Feature** *Sukanchalika Boonmatham,Supranee Horma,Soawanee Prachayagringkai,Pornsak preelakha,Jedsada Boonsiri,Wannapa Phopli and Maleerat Sodanil*	205
16:00 – 16:20 NCCIT2017-055	**Automatic Gem Classification Using Color-based Features** *Pattamaporn Kengkarnrob and Saowaluk Watanapa*	211
16:20 – 16:40 NCCIT2017-091	**Comparison of Algorithm's Performance based on Color and Texture in Image Retrieval** *Prayoot Surasena,Supathep Satiman,Ekkalak Chimjan,Sutat Gammanee,Chaovalit Somboonpattanakit,Ittikon Thammachantuek and Maleerat Sodanil*	217
16:40 – 17:00 NCCIT2017-113	**Classification of Thai Herbal Leaf images based on Local Texture Feature** *Phuttasak Somchai and Sompong Valuvanathon*	223
17:00 – 17:20 NCCIT2017-133	**Rule Based Diagnostic Application for Thai Herbal Medicines** *Chayaporn Kaensar*	229
17:20 – 17:40 NCCIT2017-157	**Development of Oil Palm Disease Diagnosis System by using Decision Tree** *Wongkot Sriurai,Phongphet Thaporn and Watcharin Nampanya*	235
18:00 – 22:00	***Welcome Dinner and Best Paper Award Presentation***	

Thursday 6th July 2017		
ROOM 4: Information Technology and System Engineering		**Page**
13:00 – 13:20 NCCIT2017-007	**System Verify Product Details in Textile Museum by Smartphone Using QR Code Technology** *Nattavadee Hongboonmee and Sittipon Prajan*	241
13:20 – 13:40 NCCIT2017-014	**The Management Information System for Universal Design in Thailand: Thai-Udee** *Suraiya Charornreh,Hasuenee Masalong,Porntip Boontem and Nattapong Tongtep*	247
13:40 – 14:00 NCCIT2017-017	**Supplier Selection Using Integrated Multi-Criteria Decision Making and MCGP : A Case Study of NTN Trading Thailand Company Limited** *Nittaya Phoda,Sirapat Chiewchanwattana and Khamron Sunat*	253
14:00 – 14:20 NCCIT2017-064	**Realism Travel on Mobile Phone: Isan Buddhist Diocese of Mahasarakham** *Pratya Sutee,Jetsada Meela and Jantima Polpinij*	261
14:20 – 14:40 NCCIT2017-150	**The Empirical Evidence on Software Maintainability in Open Source Projects** *Thanyarat Kitpanich and Aziz Nanthaamornphong*	267
14:40 – 15:00 NCCIT2017-148	**Modern Fortran Transformation Rules for UML Sequence Diagrams** *Anawat Leatongkam and Aziz Nanthaamornphong*	274
15:00 – 15:20	*Coffee Break*	
15:20 – 15:40 NCCIT2017-098	**Information System for Quality Assurance in Program Level based on Cloud Computing Case Study : Rambhai Barni Rajabhat University** *Nitat Ninchawee and Wandee Chokchueypattanakit*	281
15:40 – 16:00 NCCIT2017-167	**Personal Software Process Improvement: AppDoSI Web-based Application** *Nuansri Denwattana and Apisit Saengsai*	287
16:00 – 16:20 NCCIT2017-105	**Stock-out Causes and Multiple-Criteria ABC Analysis:A Case Study of Retail Businesses** *Wanvisa Krutthai,Chatchai Leenawong and Bunthit Watanapa*	293
16:20 – 16:40 NCCIT2017-120	**The Adoption of Facebook vs Line for Classroom Communication: A Case Study of Faculty of Computer Science, Ubon Ratchathani Rajabhat University** *Charnsak Srisawatsakul and Waransanang Boontarig*	299
18:00 – 22:00	***Welcome Dinner and Best Paper Award Presentation***	

Thursday 6th July 2017		
ROOM 5: Learning Science		**Page**
13:00 – 13:20 NCCIT2017-027	**The Development of WBI by Project-Based Learning Techniques for Competency-Based Lesson Plan on Blended Learning by MIAP of Information and Communication Technology Course** *Phongdanai Jittavisuttikul and Krich Sintanakul*	305
13:20 – 13:40 NCCIT2017-029	**The Development of the Web Quest Online System** *Pattamavadee Kroeiram,and Jaree Thongkam*	311
13:40 – 14:00 NCCIT2017-141	**Yoga for Good Health : An iOS Application for Personal Yoga Training***Siripa Jantabon and Worasit Choochaiwattana*	318
14:00 – 14:20 NCCIT2017-259	**Ubiquitous Learning Model on Cloud Systems for Thai Education 4.0** *Nattaya Jitpongtae,Nivet Chirawivhitchai and Prasong Praneetpolgrang*	324
14:20 – 14:40 NCCIT2017-260	**Key Features in New Digital Media Affecting Student Perception and Thai Education Processes** *Boonchom Sudjit and Prasong Praneetpolgrang*	331
14.40 – 15.00 NCCIT2017-012	**Learning Effectiveness for Generation Z Using Micro Learning** *Tareenichar Leepeeravitit,Kwan Sitathani and WichianChutimaskul*	338
15:00 – 15:20	*Coffee Break*	
15:20 – 15:40 NCCIT2017-085	**Mobile Media Learning in Basic Chinese based on Pinyin** *Suwimon Huangkasem,Dechawut Wanichsan,Pornrudee Sittipong,Narumon Orachon and Rujika Booncherd*	344
15:40 – 16:00 NCCIT2017-090	**The development of the Computer Multimedia Instruction on compositions in visual arts for Information Technology undergraduate students of Nakhon Pathom Rajabhat University** *Salyapong Wichaidit and Kunlaya Charoenmongkonvilai1*	350
16:00 – 16:20 NCCIT2017-100	**The Development of WBI for Competency-Based Teaching Plan MIAP Learning in Circuit Description and Microcomputer Maintenance Using PjBL with Blended Learning** *Winai Phengpinyo and Krich Sintanakul*	355
16:20 – 16:40 NCCIT2017-145	**Text Readability Assessment using Loanword Extraction from Thai Texts** *Burhan Wanglem and Nattapong Tongtep*	361
18:00 – 22:00	***Welcome Dinner and Best Paper Award Presentation***	

Friday 7th July 2017	
08:00 – 09:00	Registration
09:00 – 10:00	Invited Keynote Speech by: **Prof.Dr.Prabhas Chongstitvatana** • Professor in the Department of Computer Engineering, Chulalongkorn University.
	Topic: Machine Learning Trends for Big Data
10:00 – 10:20	*Coffee Break*
10:20 – 12:00	Paper Presentation
12:00 – 13:00	*Lunch*
13:00 – 16:40	Paper Presentation

Friday 7th July 2017		
ROOM 1: Data Science and Machine Learning		**Page**
10:20 – 10:40 NCCIT2017-163	**Cartoonized Facial Expression Classification using Histogram of Oriented Gradient (HOG)** *Orawan Liabsiri and Saowaluk Watanapa*	367
10:40 – 11:00 NCCIT2017-215	**Ensemble Feature Selection for Sparse Data** *Wannapa Phopli,Sukanchalika Boonmatham and Phayung Meesad*	373
11:00 – 11:20 NCCIT2017-189	**Vowel Thai Speech Extraction for Robot Speech Recognition** *Mananchaya Amornpalang and Saiyan Saiyod*	379
11:20 – 11:40 NCCIT2017-196	**A Business Intelligence System for Selling Plan in Real Estate** *Donruedee Rapisuwan and Sakchai Tangwannawit*	385
11:40 – 12:00 NCCIT2017-188	**Shredded Document Reconstruction using Linear Prediction** *Tanasanee Phienthrakul*	391
12:00 – 13:00	*Lunch*	
13:20 – 13:40 NCCIT2017-018	**A System of Thai Text Document Retrieval by Speech** *Paphonput Sopon and Jantima Polpinij*	397
13:40 – 14:00 NCCIT2017-244	**Analysis of Scapular Data Using K-mean Clustering** *Porntheera Chanta and Wisan Tangwongcharoen*	403
14:00 – 14:20 NCCIT2017-004	**An analysis of Twitter in the passing of His Majesty King Bhumibol Adulyadej** *Araya Pudtal and Sukree Sinthupinyo*	409
14:20 – 14:40 NCCIT2017-006	**An Intelligent System for Discover Computer Skill** *Sinapas Juanngarm and Sorawit Yotalai*	415
14:40 – 15:00 NCCIT2017-248	**Enhancing Performance of Rule-based Predicting Models in Data Mining** *Duangkamol Sodalee and Jaree Thongkam*	421
15:00 – 15:20	*Coffee Break*	

Friday 7th July 2017		
ROOM 1: Data Science and Machine Learning		**Page**
15:20 – 15:40 NCCIT2017-234	**Factor Analysis of Chronic Kidney Disease for Diabetics Using Data Mining Techniques** *Krisada Khrueawan,Monlica Wattana,Chakchai So-In and Nunnapus Benjamas*	427
15:40 – 16:00 NCCIT2017-022	**Travelling Planning for Maximize Satisfaction using Particle Swarm Optimization and Fuzzy Set** *Tanagrit Chansaeng,Pita Jarupunphol and Narongdech Keeratipranon*	433
16:00 – 16:20 NCCIT2017-024	**Prediction on Two-Dimensional Gel Electrophoresis (2D-Gel) using Data Mining and Feature Selection Techniques** *Theanchai Ounoi,Samlee Mankhetkorn and Kitsana Waiyamai*	439
16:20 – 16:40 NCCIT2017-172	**Data customization for ETL Used Hybrid Schema Techniques** *Supawadee Klommit and Sakchai Tangwannawit*	447
16:40 – 17:00 NCCIT2017-181	**Applying Fuzzy Association Rules Mining in the Context of Thai Strength Families** *Phinyawat Sookgaew and Noppamas Pukkhem*	453
17:00 – 17:20 NCCIT2017-194	**Decision Model for Helpdesk System using Data Mining Technique** *Thanchanok Naowaratkul and Sakchai Tangwannawit*	459
17:20 – 17:40 NCCIT2017-220	**The Introduction of Music by Relation Analysis between Brainwave and Music while Human Sleeping.** *Wattanachai Wattanasingha,Chayakorn kongtad and Tosaporn saiyim*	465

Friday 7th July 2017		
ROOM 2: Human-Computer Interface and Image Processing, Data Science and Machine Learning, Data Network and Communication		**Page**
10:20 – 10:40 NCCIT2017-057	**Pickup Truck Model Classification from Front Diagonal View Using Histogram of Oriented Gradient** *Taranop Sairoon and Saowaluk Watanapa*	471
10:40 – 11:00 NCCIT2017-060	**Evaluation of haze removal of Tuned Tri-threshold Fuzzy Intensification Operators Algorithm with No-Reference Image Quality Assessment** *Pattanaphong Pothipassa*	477
11:00 – 11:20 NCCIT2017-069	**Analysis of Plant Disease on Leaves Application on Android using Support Vector Machine** *Thiphawan Sawaengmee, Supansa Najaitrong, Nitit Wangno and Khamron Sunat*	483
11:20 – 11:40 NCCIT2017-080	**Optical Music Recognition Using Template Matching and Amplitude Segmentation Methods** *Pratchaya Dokmaingam, Sirivimon Voraditpatthana and Tosaporn Saiyim*	490
11:40 – 12:00 NCCIT2017-237	**BreathingTraining Game** *Nuttanont Hongwarittorrn*	497
12:00 – 13:00	*Lunch*	
13:00 – 13:20 NCCIT2017-023	**Bus Facilitation Application for Blind** *Thearasak Phaladisailoed,Tanawat Gajaseni and Thanisa Numnonda*	503
13:20 – 13:40 NCCIT2017-232	**Additional Function Design and Development of Walking Stick for Supporting a Blind Man** *Panjarat Hanpanich and Suchart Rommaneeyarak*	509
13:40 – 14:00 NCCIT2017-099	**Estrus Prediction in Dairy Cows using Fuzzy Logic** *ChanidaTawornrach,SaiyanSaiyod,NgammijArch-Int,Chakchai So-In and AnanChanlun*	515
14:00 – 14:20 NCCIT2017-122	**Forecasting for Inventory Stock-out Reduction Case Study: Pharmaceutical Company** *Angsuma Thongthiew,Chartchai Leenawong and Bunthit Watanapa*	521
14:20 – 14:40 NCCIT2017-136	**The Development Customer Segmentation and Associations Algorithm Base on K-Means and FP-Growth** *Supitcha Saengchat and Montean Rattanasiriwongwut*	527
14:40 – 15:00 NCCIT2017-032	**Package Tour Generation From Online Data** *Nichamon Uraisriphong,Warisara Thumwirat,Bundit Thanasopon and Ponrudee Netisopakul*	534
15:00 – 15:20	*Coffee Break*	

Friday 7th July 2017		
ROOM 2: Human-Computer Interface and Image Processing, Data Science and Machine Learning, Data Network and Communication		**Page**
15:20 – 15:40 NCCIT2017-039	**Sarcasm Classification in Twitter using Probability of Tweets** *Kasidech Tapang and Janjao Mongkolnavin*	540
15:40 – 16:00 NCCIT2017-139	**BI System Development to Support Decision of Educational Administrators***Warunee Tamkoo and Kitsana Waiyamai*	546
16:00 – 16:20 NCCIT2017-149	**Introducing the Document Flow of the Official Document by Text Mining***Guntapich Anantasak and Nattavee Utakrit*	551
16:20 – 16:40 NCCIT2017-251	**Isarn Dialect Part of Speech Tagging** *Tanakul Pakkawesa andPusadee Seresangtakul*	557
16:40 – 17:00 NCCIT2017-240	**PL-USSR: Physical Location of Wireless Access Point with Considering Unacceptable Signal Strength Range** *Pratchaya Jaisudthi and Surachet Yuwarat*	564
17:00 – 17:20 NCCIT2017-236	**Factors effecting Cloud ERP Adoption in Manufacturing Firms in Thailand** *Pisut Norkaew,Watchapon Suksakul and Bundit Thanasopon*	569
17:20 – 17:40 NCCIT2017-115	**Adaptive Packet Filter with Probability Statistic Analysis in Intrusion Prevention System** *Weerayuth Sukkhamduang and Saiyan Saiyod*	575

Friday 7th July 2017		
ROOM 3: Information Technology and System Engineering		**Page**
10:20 – 10:40 NCCIT2017-142	**J-Talk: An iOS Application for Communication in Japanese** *Dunyarat kawviset and Worasit Choochaiwattana*	581
10:40 – 11:00 NCCIT2017-147	**Development of Thai Herb Recommendation System for Healthcare on Android Phone** *Nattavadee Hongboonmee and Chatsuda Ninsakul*	587
11:00 – 11:20 NCCIT2017-151	**Software Metrics Calculation System for Testing with Business Intelligence** *Siriwan Moolmao and Nattavee Utakrit*	593
11:20 – 11:40 NCCIT2017-202	**The Development of Information System for Fleet Maintenance Using Customer Retention Technique** *Phatchareeya Meetha and Sakchai Tangwannawit*	600
11.40 – 12.00 NCCIT2017-041	**A Development of Data Warehouse in Risk Assessment for Disaster Management** *Tanaporn Panrungsri and Esther Sangiamkul*	606
12:00 – 13:00	*Lunch*	
13:00 – 13:20 NCCIT2017-016	**Automatic Course Scheduling System applying Ontology and Rule-based***Burasakorn Yoosooka,Pongpon Nilaphruek and Khongthep Boonme*	612
13:20 – 13:40 NCCIT2017-048	**Airport Rail Link Transportation Application for the Hearing Impaired (NextStop)** *Arnon Chonrawut and Trithep Ratanapipop*	618
13:40 – 14:00 NCCIT2017-049	**An Applied of Indoor Positioning System for Local-based Knowledge Management** *Chisanupong Sooksawang,Thanakit Tantithanakorn and Boonprasert Surakratanasakul*	624
14:00 – 14:20 NCCIT2017-054	**Model-based Approach for Long Method Bad Smell Detection** *Theraphat Satheanphong and Songsak Rongviriyapanich*	630
14:20 – 14:40 NCCIT2017-104	**Information System Life Cycle Management using Social Network***Chatsayam Philam,Nawaporn Wisitpongphan and Nalinpat Porrawatpreyakorn*	637
14:40 – 15:00 NCCIT2017-081	**Ontology Development for Semantic Search of Organic Vegetable Cultivation** *Worapoj Suwanpipob,Ngamnij Arch-Int,Somjit Arch-Int and Paweena Wanchai*	643
15:00 – 15:20	*Coffee Break*	

Friday 7th July 2017		
ROOM 3: Information Technology and System Engineering		**Page**
15:20 – 15:40 NCCIT2017-093	**A Key Success Factors for Distributed Software Development** *Sudarat Worawetwichai,Tawa Khampachua and Nawaporn Wisitpongphan*	650
15:40 – 16:00 NCCIT2017-074	**Car Reservation and Tracking System Using Google Map Case Study Thailand Post Co., Ltd.** *Unchalee Rungthian and Maleerat Sodanil*	657
16:00 – 16:20 NCCIT2017-109	**Comparison of Vulnerabilities in Content Management Systems: Joomla, Wordpress and Drupal** *Pratchaya Jaisudthi,Kitti Mala and Sonchai Boonluea*	663
16:20 – 16:40 NCCIT2017-255	**Applied Google Map on Rabies in Humans and Animals** *Paravee Pattarakavin and Krerk Piromsopa*	667
16.40 – 17.00 NCCIT2017-015	**The Comparison of Data Classification Efficiency by using ENSEMBLE Technique for Analysis of Factors Affecting on the National Educational Test Results of Vocational Education (V-NET)** *Puttida Sakulviriyakitkul,Charun Sanrch and Thanrat Sintanakul*	674
17.00 – 17.20 NCCIT2017-243	**Comparison Study of NETPIE with other IoT Platforms** *Khwanchanok Srimool,Thapanee Chayakul,Aimaschana Niruntasukrat, Panita Pongpaibool and Sukumal Kitisin*	680

Friday 7th July 2017		
ROOM 4: Information Technology and System Engineering		**Page**
10:20 – 10:40 NCCIT2017-262	**Information Technology Competency Assessment Framework during Agile Software Development in Public Sector** *Tawa Khampachua and Somkid Saelee*	686
10:40 – 11:00 NCCIT2017-140	**Find Nearby Hospitals in Bangkok System on The Android Operating System** *Jarunan Boonwittayalikit and Montean Rattanasiriwongwut*	692
11:00 – 11:20 NCCIT2017-146	**Android Application Prototype Development for Supporting Self-care Behavior in Diabetic Patient)** *Rungsan Suwannahong and Sureerut Inmor*	699
11:20 – 11:40 NCCIT2017-101	**Personalized Nutrient Recommendation System for Diabetic Patient Using Ontology and Rule of Inference Approach** *Pattanee Arboonngam,Ngamnij Arch-int,Somjit Arch-int and Sukanya Khayantham*	706
11:40 – 12:00 NCCIT2017-183	**Ontology-based Health News Summarization** *Pannapa Nuanplord and Maleerat Sodanil*	713
12:00 – 13:00	*Lunch*	
13.00 – 13.20 NCCIT2017-153	**A Suitable Knowledge Management Website's Model for Officer in Domestic Supply and Procurement Division,Electricity Generating Authority of Thailand** *Parisa Pattarathanawat and Nattavee Utakrit*	719
13:20 – 13:40 NCCIT2017-159	**An Application of Risk Management Techniques with IT Project Monitoring and Controlling System** *Pongsakorn Naklamai and Nattavee Uttakrit*	725
13:40 – 14:00 NCCIT2017-165	**A Development of Decision Support System to Assign Jobs to the IT Technical Support** *Amnat Gongsuk and Nattavee Utakrit*	731
14:00 – 14:20 NCCIT2017-166	**Asean alphabet handwriting application on mobile device** *Jiruttigarn Gunteemoon,Supaluck Mongkhonrat,Phet Imtongkhum and Chakchai So-In*	737
14:20 – 14:40 NCCIT2017-037	**Automatic Report Generation using Report Name and Data Structure Extraction** *Trirat Pat-iam and Nattapong Tongtep*	744
14:00 – 14:20 NCCIT2017-103	**Payment system using NFC and RFID: A case in school cooperatives** *Pichayapa Maprajuab,Tanawit Tattakittiya,Suthasinee Iamsa-at and Khamron Sunat*	749
15:00 – 15:20	*Coffee Break*	
15:20 – 15:40 NCCIT2017-201	**A prototype of coaching student activities based on Adaptive Modality** *Phonepasuerth Bounnaphonh,Panchit Longpradit,Pramote Tongchin and Veerachai Khonchoho*	756
15:40 – 16:00 NCCIT2017-128	**Home Security System using Raspberry PI through Facebook Messenger** *Dechawut Wanichsan,Komsak Ponpruksa and Yanee Lamjong*	761

Friday 7th July 2017		
ROOM 5: Information Technology and System Engineering, Learning Science and Computer Education		**Page**
10:20 – 10:40 NCCIT2017-176	**Personal Medical Record for Laboratory Data Exchange System** *Patipat Tumsangthong,Wisit Wongwilai,Weerachai Chansud, Nithipat Wongchingchai and Asamaporn Chatrattikorn*	767
10:40 – 11:00 NCCIT2017-174	**Development of local knowledge information management system integrated with Nakhon Pathom tourist routes in spatial data format on geographic information systems.** *Paisan Simalaotao and Ubonrat Sirisukpoca*	773
11:00 – 11:20 NCCIT2017-178	**Audio book management application for the blind on mobile device** *Kingkanjana Khotkaew,Donlaya Suriyawong,Phet Imtongkhum and Chakchai So-In*	779
11:20 – 11:40 NCCIT2017-184	**Alzheimer aiding application on early stage on mobile device** *Chonnika Punjunsing,Piyaporn Apisuwan and Chakchai So-In*	785
11:40 – 12:00 NCCIT2017-185	**A 3D Simulation Game for Improving Knowledge of Vocabulary in the House** *Chutipon Blumrungjit,Preecha Kangngulam,Dechawut Wanichsan and Auemporn Rungsiri*	792
12:00 – 13:00	*Lunch*	
13:00 – 13:20 NCCIT2017-108	**Automatic Exam Analysis for The Different Learner System** *Sutthikarn Bojukrapan and Piyasuda Tanloet*	798
13:20 – 13:40 NCCIT2017-239	**The Development of Field Trip Learning Management System** *Sukita Boonraum and Somkid Saelee*	804
13:40 – 14:00 NCCIT2017-242	**Identifying Prototype of Automated Appliance Control System from User Behavior** *Teera Siriteerakul,Petchpailin Saeauy,Payu Poonsri and Sirikul Siriteerakul*	810
14:00 – 14:20 NCCIT2017-117	**The Jigsaw: An Augmented Reality Periodic Table for Supporting Collaborative Learning** *Watcharapong Thorat,Dechawut Wanichsan,Phollawat Wattanasiriyanon and Watcharee Waratchareeyakul*	814
14:20 – 14:40 NCCIT2017-124	**Game Hanuman burnt Longka to promote the study of Thai literature** *Natchapol Surarattanagon,Patcharee Chareonlap and Nathee Yongyut*	820
14:40 – 15:00 NCCIT2017-130	**The Design of Instructional Systems Based on Dick and Carey Model By Using 2D via Cloud Computing in the Course of the Information Systems Analysis and Design. For Teaching Profession. Computer Education Major Nakorn Pathom Rajaphat University** *Charinthorn Aumgri,Naphong Wannapiroon,Wassana Yaemprai,Paveena Chamking and Siwalai Chinchua*	826
15:00 – 15:20	*Coffee Break*	
15:20 – 15:40 NCCIT2017-214	**Mathematical model for the development of monitoring systems for solar water heating systems.** *Pongpith Tuenpusa,Niti Witthayawiroj,Sathaporn Thongwik and Suriya Sukarin*	832
15:40 – 16:00 NCCIT2017-203	**Mobile Application to Reading Drug Labels for Visually Impairment Using NFC Technology** *Natruja Suwannapak and Pudsadee Boonrawd*	837

KEYNOTE SPEAKERS



Professor Dr. Jie Lu
University of Technology Sydney

About the Speaker:

Professor Jie Lu is the Associate Dean in Research in the Faculty of Engineering and Information Technology at the University of Technology Sydney (UTS).

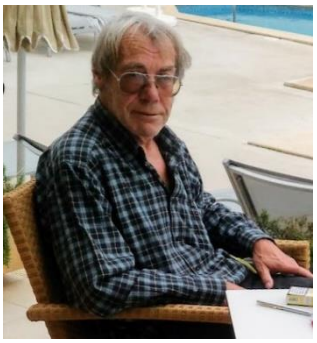
She is also the Director of the Decision Systems and e-Service Intelligence (DeSI) Research Laboratory in the Centre for Quantum Computation & Intelligent Systems. Her main research interests lie in the area of decision support systems, recommender systems, prediction and early warning systems, fuzzy transfer learning, concept drift and web-based e-service intelligence. She has published six research books and 400 papers in refereed journals and conference proceedings.

She has won seven Australian Research Council (ARC) discovery grants and 10 other research grants in the last 15 years. She received the first UTS Research Excellent Medal for Teaching and Research Integration in 2010 and other awards. She serves as Editor-In-Chief for Knowledge-Based Systems (Elsevier) and Editor-In-Chief for International Journal on Computational Intelligence Systems (Atlantis), and has delivered many keynote speeches at international conferences.

Topic: Fuzzy Transfer Learning for Prediction and Decision Support

Abstract

This presentation highlights the value of fuzzy transfer learning methods and related algorithms for handling complex prediction problems in rapidly-changing data distribution and data-shortage situations. It provides a framework for utilizing previously-acquired knowledge to predict new but similar problems quickly and effectively by using fuzzy set techniques. It systematically presents developments in fuzzy set-based transfer learning methods for prediction, including fuzzy transfer learning-based prediction framework, fuzzy domain adaptation, fuzzy cross-domain adaptation, and in particular, cross-domain adaptive fuzzy inference system, and their respective applications in prediction and decision support. This presentation demonstrates the successful use of fuzzy techniques in facilitating the incorporation of approximation and expressiveness of data uncertainties within knowledge transfer, machine learning and data-driven decision support systems.



Professor Dr. Hanno Lefmann
Chemnitz University of Technology

About the Speaker:

Hanno Lefmann studied mathematics and physics at the University of Bielefeld, Germany, where he got his diploma in mathematics in 1982. He continued to be with the University of Bielefeld until 1992 and got there in mathematics his PhD in 1985 as well as his habilitation in 1992.

In 1992 he moved to the University of Dortmund at the Computer Science department, Germany, where he has been until 2000. He had visiting positions at IBM Scientic Center in Heidelberg, Germany, in 1989, and, within in the United States, at Georgia Institute of Technology as well as at Emory University both in Atlanta in 1989/1990 and at the University of Idaho in Moscow in 1992, and in 1999 at Chemnitz University of Technology, Germany.

In 2000 he became a full professor for theoretical computer science and information security at Chemnitz University of Technology, Germany.

Topic: Testing and Estimating Graph Parameters for Monotone Properties

Abstract

In the last years, there has been considerable interest in finding constant-time randomized sampling algorithms that estimate whether a given structure satisfies some property or does not, or to approximate the value of some numerical function associated with this structure. We focus in this talk on the graph case and consider algorithms that have the ability to query whether any chosen pair of vertices in the input graph is adjacent or not. We consider subsets P of the set G of all finite graphs, which are are closed under isomorphism; those are called graph properties. This includes all nontrivial monotone and hereditary graph properties, which are inherited by subgraphs and by induced subgraphs, respectively. A typical example of a

monotone property is P = Forb(F) for a graph family F, i.e., the class of all graphs that do not contain a graph F ∈ F as a subgraph. Elements of Forb(F) are said to be F-free.

A graph property P is said to be testable if, for every $\varepsilon > 0$, there exist $qP = qP(\varepsilon)$, the query complexity, and a randomized algorithm TP, the tester, which performs at most qP queries in the input graph, satisfying the following property. For an n-vertex input graph Γ, the algorithm TP distinguishes with probability at least 2/3 between the cases in which Γ satisfies P (i.e., $dist(\Gamma, Forb(F)) \leq \varepsilon$) and in which Γ is ε-far (i.e., $dist(\Gamma, Forb(F)) > \varepsilon$) from satisfying P, that is, no graph obtained from Γ by the addition or removal of at most $\varepsilon n2/2$ edges satisfies P.

For any monotone graph property P, Alon, Shapira and Sudakov gave a natural algorithm that computes the distance from the induced sampled graph to P. However, one disadvantage of this approach is that its analysis relies heavily on stronger versions of the Szemer´edi Regularity Lemma. Therefore, their algorithm has a query complexity of order at least TOWER(poly(1/ε)), which is a tower of twos of height polynomial in 1/ε.

Similarly, a function $z : G \rightarrow Q$ from the set G of graphs into the set of rational numbers is called a graph parameter if it is invariant under permutations of vertices. A graph parameter $z : G \rightarrow Q$ is estimable if for every $\varepsilon > 0$ and every large graph Γ, the value of z(Γ) can be approximated up to an additive error of ε by an algorithm that only has access to a subgraph of Γ induced by a uniformly at random chosen set of $qz = qz(\varepsilon)$ vertices, the sample complexity.

We survey some recent results in this area. Amongst others, we introduce the notion of f-recoverable graph properties P for some function $f : (0, 1] \rightarrow Q$. For such a graph property P the graph parameter dist(Γ,P) can be estimated within an additive error of ε with sample complexity $2^{poly(f(\varepsilon/6)/\varepsilon)}$. We obtain similar results for the graph parameter, which, for a graph family F, counts the number of F-free subgraphs of an input graph Γ.



Professor Dr. Prabhas Chongstitvatana
Chulalongkorn University

About the Speaker:

Prabhas Chongstitvatana is a professor in the department of computer engineering, Chulalongkorn University. He earned B.Eng. in Electrical Engineering from Kasetsart University, Thailand in 1980 and Ph.D. from the department of artificial intelligence, Edinburgh University, U.K. in 1992. His research involves robotics, evolutionary computation, computer architecture, bioinformatics. He is a lifetime member of Thailand Engineering Institute, senior member of Thai Academy of Science and Technology, senior adviser of Thai Robotics Society, founding member of Thai Embedded System Association and IEEE Robotics and Automation Society. He was the president of ECTI Association of Thailand 2012-2013. He has been awarded "National Distinguished Researcher" by National Research Council of Thailand in 2009. His current interest is in building Quantum computer.

Topic: Machine Learning Trends for Big Data

Abstract

One of the most powerful technique to create artificial intelligence is Machine Learning. Recently, there are several achievements by Machine Learning that compete with human experts successfully. With these encouragements, a lot of interest and investment have been poured into this field with promising results. In this talk, I will explain the basic technique of Machine Learning, especially Deep Learning and Convolutional Neural Network. This form of Machine Learning can digest huge amount of data and creates mathematical models that can be used to perform classification and prediction tasks. The application for this ability to learn from large data has been exploit in many projects. I will illustrate a number of examples and discuss the current trend.



Professor Dr. Armin R. Mikler
University of North Texas

About the Speaker:

In 1997, Professor Mikler joined the Department of Computer Science at the University of North Texas with a PhD from Iowa State University. With the help of four courageous undergraduate students, he established the Network Research Laboratory (NRL), and with it, UNT's first Beowulf Cluster to facilitate complex simulations in support of the group's research on Computer Network Protocols and Distributed Systems. In addition to the inaugural group of students, who completed their MS theses under Dr. Mikler's guidance, the laboratory attracted many graduate students with interest in experimental design of protocols and algorithms for large distributed computing infrastructures.

In 2004, Dr. Mikler started to gradually move into a new field of research, which was motivated by the need to facilitate advances in the field of Public Health and Epidemiology through computational methods. He established the Computational Epidemiology Research Laboratory (CERL) with focus on the development of computational methodology to model and simulate the spread of diseases and the design and analysis of bioemergencyresponse plans. Together with colleagues in Biology and Geography, Dr. Mikler establishedthe truly interdisciplinary Center for Computational Epidemiology and Response Analysis (CeCERA)after receiving federal funds from the US Department of Health and Human Services.

Today, CeCERAis the home of over 15 PhD students who are conducting research in a variety of areas related toComputational Epidemiology, Ecology, Social Network Analysis, and High Performance Computing under Dr. Mikler's mentorship. Recent graduates of his research group are using their expertise inComputational Epidemiology as faculty members at different universities and as researchers atNational Laboratories. Dr. Mikler’s research on response plan design and analysis is supported by theTexas Department of State Health Services (DSHS), the

National Science Foundation (NSF), and theNational Institutes of Health (NIH). He has supervised over 30 PhD and MS theses and has publishedover 70 research articles related to a range of topics, including distributed systems, networking,computational epidemiology, and response plan design and analysis.

Topic: Computational Intelligence and Big Data Challenges in Computational Epidemiology and Population Health

Abstract

The development of computational approaches to solving problems in Public Health has commenced only about two decades ago and has led to the emergence of the field of Computational Epidemiology. Its primary goal is to provide health researchers with computational tools that facilitate the prediction and analysis of the progression of diseases in time and space. Whether modeling an Influenza epidemic in Germany or the spread of Dengue Fever in Thailand, computational models must be developed, which are informed by data from disparate sources. The design, implementation, and execution of such models represents a significant scientific challenge as it is often difficult to validate their fidelity. Further, data availability and representation across different geographic regions is inconsistent at best, which necessitates the design of region specific models. In this tutorial, we will introduce Computational Epidemiology as a datacentric example of computational intelligence for which Big Data challenges will have to be addressed by today's computer and information scientists. We will exemplify the types of data that will have to be integrated to inform model development in an effort to provide computational tools to researchers and practitioners in the domain of Population Health.

การหาค่าความเหมาะสมสำหรับขั้นตอนวิธีแบบเพลี้ยกระโดด
Optimization Problems for Planthopper Algorithm

แก้วใจ อาภรณ์พิศาล (Kaewjai Apornpisarn)[1] พยุง มีสัจ (Phayung Meesad)[2] และวาทินี นุ้ยเพียร(Vatinee Nuipian)[3]

[1]คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏนครปฐม

[2]ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

[3]ภาควิชาคอมพิวเตอร์ศึกษา คณะครุศาสตร์อุตสาหกรรม มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

[1]kaewjai@npru.ac.th, [2]phayung.m@it.kmutnb.ac.th, [3]vatinee.n@fte.kmutnb.ac.th

บทคัดย่อ

การหาค่าความเหมาะสมมีการประยุกต์ใช้งานอย่างกว้างขวางในงานด้านวิศวกรรม ด้านวิทยาศาสตร์ ด้านบริหาร และอื่นๆ วิธีการคำนวณเชิงวิวัฒนาการจึงถูกนำมาใช้ในการหาค่าที่เหมาะสมแบบเมตาฮิวริสติก สำหรับงานวิจัยนี้ผู้วิจัยจึงนำเสนอการหาค่าความเหมาะสมสำหรับขั้นตอนวิธีแบบเพลี้ยกระโดด ที่เกิดจากแนวคิดของพฤติกรรมการหาอาหารแบบเป็นกลุ่ม ด้วยการเคลื่อนที่เพื่อเปลี่ยนตำแหน่ง ซึ่งเป็นการเข้าใกล้แหล่งอาหารคือ การลู่เข้าสู่คำตอบ และจากการเปรียบเทียบขั้นตอนวิธีการค้นหาแบบนกคุเหว่า แบบค้างคาว และแบบแมลงหวี่มีข้อจำกัดในเรื่องการกำหนดพารามิเตอร์เป็นค่าคงที่และความเร็วของการลู่เข้าสู่คำตอบ ผู้วิจัยจึงทำการเปรียบเทียบกับ 3 ขั้นตอนวิธีข้างต้น จากผลการทดสอบของการหาค่าต่ำสุดของฟังก์ชันทางคณิตศาสตร์ 4 ฟังก์ชันคือ ฟังก์ชัน Sphere ฟังก์ชันไกรแวงก์ ฟังก์ชันโรเซนบร็อก และฟังก์ชันแอกค์ลีย์ พบว่าขั้นตอนวิธีแบบเพลี้ยกระโดดให้ค่าความเหมาะสมเฉลี่ยที่ดีสุดในฟังก์ชันไกรแวงก์ ($\bar{x}$=0.000076) และฟังก์ชันแอกค์ลีย์ ($\bar{x}$=0.023046) ดังนั้นวิธีที่นำเสนอสามารถแก้ไขปัญหาการหาค่าความเหมาะสมที่สุดได้เทียบเคียงกับวิธีอื่น และดีกว่าวิธีอื่นในบางฟังก์ชัน

คำสำคัญ: การหาค่าความเหมาะสม การคำนวณเชิงวิวัฒนาการ ฮิวริสติก ขั้นตอนวิธีแบบเพลี้ยกระโดด

Abstract

Optimization has been widely applied for various applications such as engineering, science, management, etc. Evolutionary computation has been in focus to the optimization problems in a heuristics way. This research aims to present the planthopper algorithm (PA) as an algorithm for finding the optimal values. The proposed algorithm is inspired by the food searching habit of the planthopper. To evaluate the proposed algorithm, we tested on optimization of benchmark mathematic functions, comparing to the Cuckoo Search Algorithm, Bat Algorithm and Fruit Fly Algorithm. Those algorithms have limitations on the speed of convergence into solutions and the determination of parameters as constant. In the problem of finding the minimum of 4 mathematical functions, the function is the Sphere function, Griewank function, Rosenbrock function and Ackley function. The Planthopper algorithm gives the best average fitness in the Griewank function ($\bar{x}$=0.000076) and the Ackley function ($\bar{x}$=0.023046). The results showed that the proposed algorithm could solve the optimization problems comparable to other existing algorithms, even better in some problems.

Keyword:Optimization, Evolutionary computation, Heuristic, Planthopper Algorithm

1. บทนำ

การหาค่าความเหมาะสมที่สุด (Optimization) เป็นวิธีการที่ใช้ในการหาคำตอบที่ดีที่สุดของปัญหาภายใต้เงื่อนไขหรือข้อจำกัดที่กำหนดขึ้น [1] ถือว่าเป็นสิ่งที่ช่วยในการแก้ปัญหา

ด้านวิทยาการคอมพิวเตอร์ งานด้านปัญญาประดิษฐ์ การวิจัยดำเนินงาน และสาขาอื่นๆ ที่เกี่ยวข้องได้เป็นอย่างดี [2]

ขั้นตอนวิธีแบบเมตาฮิวริสติกที่เหมาะสมที่สุด (Metaheuristic optimization algorithm) เกิดจากการเฝ้าสังเกตพฤติกรรมในธรรมชาติที่สนใจ และนำมาสกัดเป็นขั้นตอนวิธี (Algorithm) [3] เพื่อแก้ไขปัญหาที่มีความซับซ้อนได้อย่างรวดเร็ว คุณลักษณะหลัก 2 ลักษณะของการแก้ปัญหาคือ ระยะ Intensification เป็นการค้นหาผลลัพธ์ที่ดีและคัดเลือกคู่แข่งหรือผลลัพธ์ และระยะ Diversification เป็นการรับรองขั้นตอนวิธีที่ค้นหาพื้นที่ให้มีประสิทธิภาพมากขึ้น [4] โดยจุดประสงค์ในการพัฒนาคือ ต้องการหาคำตอบที่รวดเร็วเพื่อแก้ไขปัญหาขนาดใหญ่ เช่น ปัญหาด้านวิศวกรรม ปัญหาการจัดเส้นทางการขนส่ง และปัญหาการวางแผนทรัพยากรให้มีประสิทธิภาพมากขึ้น โดยมีหลายขั้นตอนวิธี เช่น จีเนติก (Genetic Algorithm: GA) ซึ่งเป็นการค้นหาแบบฝูงอนุภาค (Particle Swarm Optimization: PSO) หรือการล่าเหยื่อของฝูงสัตว์ [5-6] ตัวอย่างเช่น ขั้นตอนวิธีการหาค่าเหมาะสมแบบฝูงมด (Ant Colony Optimization: ACO) เป็นการหาอาหารในเส้นทางที่สั้นที่สุดด้วยการปล่อยฟีโรโมน [7] ขั้นตอนวิธีแบบหิ่งห้อย (Firefly Algorithm: FA) เป็นการดึงดูดกันของหิ่งห้อยโดยอาศัยการกระพริบแสง [8] ขั้นตอนวิธีการค้นหาแบบนกดุเหว่า (Cuckoo Search : CS) เป็นการนำไข่ไปให้นกชนิดอื่นฟักและรูปแบบการบินของผีเสื้อและแมลงวัน [9] ขั้นตอนวิธีแบบค้างคาว (Bat Algorithm: BA) เป็นการล่าเหยื่อโดยใช้เสียงสะท้อน [10] ขั้นตอนวิธีแบบแมลงหวี่ (Fruit Fly Optimization Algorithm: FOA) เป็นการหาอาหารของแมลงหวี่โดยใช้กลิ่นของอาหาร [11] ซึ่งขั้นตอนวิธีการค้นหาแบบนกดุเหว่ามีข้อจำกัดในเรื่องของการกำหนดค่าความเป็นไปได้ในการทิ้งรังและ Step size มีค่าคงที่ หากมีการกำหนดค่าทั้งสองใหม่จะทำให้ประสิทธิภาพลดลง และไม่สามารถหาคำตอบที่ดีที่สุดได้ ขั้นตอนวิธีแบบค้างคาว มีข้อจำกัดในเรื่องของความเร็วของการลู่เข้าสู่คำตอบลดลง หลังจากรันขั้นตอนวิธีในระยะเริ่มต้นไปแล้ว และขั้นตอนวิธีแบบแมลงหวี่ มีข้อจำกัดในเรื่องของการลู่เข้าสู่คำตอบช้า

งานวิจัยนี้จึงนำเสนอการหาค่าความเหมาะสมสำหรับขั้นตอนวิธีแบบเพลี้ยกระโดด ซึ่งเป็นการหาค่าความเหมาะสมที่ได้จากการศึกษาพฤติกรรมการหาอาหารแบบเป็นกลุ่ม โดยการเปรียบเทียบกับขั้นตอนวิธีการค้นหาแบบนกดุเหว่า แบบค้างคาว และแบบแมลงหวี่ ด้วยการใช้ฟังก์ชันทางคณิตศาสตร์ (Numerical optimization) เพื่อทดสอบประสิทธิภาพในการหาค่าความเหมาะสม

2. วรรณกรรมที่เกี่ยวข้อง

2.1 การหาค่าความเหมาะสม (Optimization)

ปัญหาการหาค่าความเหมาะสม สามารถแสดงได้ดังนี้ [12]

กำหนดให้ ฟังก์ชัน $f : A \rightarrow \boldsymbol{R}$ แมปจากค่าบางส่วนในเซต A ไปเป็นค่าจริง ต้องการค้นหาค่า x_0 ในเซต A ซึ่งหากค่า $f(x_0) \leq f(x)$ สำหรับค่า x ทั้งหมดในเซต A เรียกว่าการหาค่าต่ำสุด (Minimization) และหากค่า $f(x_0) \geq f(x)$ สำหรับค่า x ทั้งหมดในเซต A เรียกว่าการหาค่าสูงสุด (Maximization)

การหาค่าความเหมาะสมมีหลายวิธี ดังนี้ ขั้นตอนวิธีค้นหาค่าความเหมาะสมโดยมีพื้นฐานจากวิธีการทางคณิตศาสตร์ ได้แก่ การโปรแกรมเชิงเส้น (Linear programming) ใช้กับปัญหาซึ่งความสัมพันธ์กันระหว่างตัวแปรทุกตัวเป็นแบบเชิงเส้น การโปรแกรมเลขจำนวนเต็ม (Integer programming) เป็นโปรแกรมเชิงคณิตศาสตร์ที่มีข้อจำกัดของตัวแปรว่าจะต้องเป็นเลขจำนวนเต็มที่ไม่มีค่าเป็นลบ การโปรแกรมเชิงกำลังสอง (Quadratic programming) เป็นฟังก์ชันกำลังสองที่มีข้อจำกัดของปัญหาสมการหรืออสมการเชิงเส้น และการโปรแกรมไม่เป็นเส้นตรง (Nonlinear programming) เป็นโปรแกรมเชิงคณิตศาสตร์ที่ใช้หาคำตอบ กรณีที่ตัวแปรตัดสินใจมีความสัมพันธ์กันลักษณะที่ไม่เป็นเส้นตรง จะมีฟังก์ชันอย่างน้อยที่สุด 1 ฟังก์ชันไม่เป็นเส้นตรง เป็นต้น และการหาค่าความเหมาะสมที่มีพื้นฐานจากวิธีการเชิงสามัญสำนึก (Heuristics) ได้แก่ ขั้นตอนวิธีมีมีติก (Memetic algorithm) จีเนติก (Genetic Algorithm: GA) ขั้นตอนวิธีการหาค่าเหมาะสมแบบฝูงมด (Ant Colony Optimization) ขั้นตอนวิธีแบบหิ่งห้อย (Firefly Algorithm) [13] ขั้นตอนวิธีการค้นหาแบบนกดุเหว่า (Cuckoo Search) ขั้นตอนวิธีแบบค้างคาว (Bat Algorithm) และขั้นตอนวิธีแบบแมลงหวี่ (Fruit Fly Algorithm) เป็นต้น

2.2 ขั้นตอนวิธีการค้นหาแบบนกดุเหว่า (Cuckoo Search: CS)

เป็นขั้นตอนวิธีที่อาศัยหลักการธรรมชาติการวางไข่ของนกดุเหว่า โดยจะทำการวางไข่ลงไปในรัง ซึ่งรังที่จะทำการวางไข่จะถูกสุ่มเพื่อเลือกรังที่มีคุณภาพสูง เพราะจะทำให้ตัวอ่อนมีอัตราการรอดชีวิตสูงตามไปด้วย จึงมีผลทำให้ได้นกรุ่นถัดไป และขั้นตอนการบินหารังใหม่นั้นจะใช้หลักการของ Levy Flight ที่มีการเคลื่อนที่ไปได้ไกลกว่าปกติ [9]

2.3 ขั้นตอนวิธีแบบค้างคาว (Bat Algorithm: BA)

เป็นขั้นตอนวิธีที่มีแนวคิดจากพฤติกรรมการหาอาหารของค้างคาว โดยการปล่อยคลื่นเสียง [10]

2.4 ขั้นตอนวิธีแบบแมลงหวี่ (Fruit Fly Optimization Algorithm: FOA)

เป็นขั้นตอนวิธีที่มีแนวคิดจากพฤติกรรมในการค้นหาแหล่งอาหาร ซึ่งเป็นสัตว์ที่มีประสาทสัมผัสในการมองเห็นและรับกลิ่นที่ดีมาก ในการค้นหาแหล่งอาหารจะมีกระบวนการที่สำคัญอยู่ 2 กระบวนการ คือ กระบวนการแรกจะใช้ประสาทสัมผัสในการับกลิ่นของแหล่งอาหาร และกระบวนที่สองจะเป็นการใช้ประสาทสัมผัสในการมองเห็น โดยจะบินไปยังแหล่งอาหารเพื่อทำการค้นหาอาหาร [11]

2.5 พฤติกรรมของเพลี้ยกระโดด (Planthopper)

เพลี้ยกระโดดอยู่ในอันดับ Hemiptera มีวงจรชีวิต 3 ระยะ คือ ไข่ ตัวอ่อน และตัวเต็มวัย เป็นแมลงที่สามารถกระโดดได้อย่างรวดเร็ว และพลังมากที่สุดในบรรดาแมลงทั้งหมด กลไกของการกระโดดคล้ายหนังสติ๊กที่เคลื่อนไหวอย่างรวดเร็วของขาหลังซ้ายและขวาของตัวมันเอง ประสิทธิภาพในการกระโดดวัดจากระยะทาง (s) ดังสมการที่ 1

$$s = v cos\theta\ (2vsin\theta)/g \quad (1)$$

กำหนดให้ v คือ ความเร็วที่ take-off และ θ คือ มุมที่ใช้ take-off องศา g คือ ความเร่งเนื่องจากความโน้มถ่วง (9.81 m/s^2) [14]

จากสมการที่ 1 นำ $vcos\theta$ คูณเข้าไปในวงเล็บ จะได้

$$s = (2v^2 cos\theta sin\theta)/g \quad (2)$$

เนื่องจาก $2cos\theta sin\theta = sin2\theta$

$$s = (v^2 sin2\theta)/g \quad (3)$$

3. การพัฒนาขั้นตอนวิธีแบบเพลี้ยกระโดด

3.1 ขั้นตอนวิธีแบบเพลี้ยกระโดด มีขั้นตอนดังต่อไปนี้

ขั้นตอนแรก กำหนดพารามิเตอร์เริ่มต้น ประกอบด้วยรอบการทำซ้ำ (maxgen) จำนวนประชากรเพลี้ยกระโดด (sizepop) มุมที่ใช้ take-off (angle) และค่า Body mass

ขั้นตอนที่ 2 กำหนดตำแหน่งของเพลี้ยกระโดดเริ่มต้น

ขั้นตอนที่ 3 สร้างคำตอบใหม่โดยการปรับค่า Body Mass ค่าความเร็ว ระยะทาง และตำแหน่งของเพลี้ยกระโดดแต่ละตัว โดยใช้สมการ

$$v = v_{min} + r(v_{min} - v_{max}) \quad (4)$$

กำหนดให้ v คือ ความเร็วที่ take-off v_{min} คือ ความเร็วต่ำสุดที่ take-off มีค่าเท่ากับ 0 v_{max} คือ ความเร็วต่ำสุดที่ take-off มีค่าเท่ากับ 5 และใช้สมการ (3) โดยกำหนดให้ $r \in [0,1]$ คือเวกเตอร์จากการสุ่ม (Random Vector) ได้มาจากการแจกแจงแบบสม่ำเสมอ (Uniform Distributions) และ มุมที่ใช้ take-off มีค่าเท่ากับ 45 องศา

การปรับตำแหน่งของเพลี้ยกระโดด

$$x_{new} = (x_{old} - x_{best}) + s \quad (5)$$

กำหนดให้ x_{new} คือ ตำแหน่งใหม่ของเพลี้ยกระโดดที่ได้จากการปรับระยะทาง x_{old} คือตำแหน่งเดิมของเพลี้ยกระโดด x_{best} คือตำแหน่งที่ดีแบบ global และ s คือระยะทางที่เกิดขึ้นจากการกระโดด

ขั้นตอนที่ 4 ถ้า r ที่สุ่มขึ้นมีค่ามากกว่า Body mass ให้เลือกคำตอบที่ดีที่สุดตามฟังก์ชันวัตถุประสงค์ และสร้างคำตอบวงแคบเฉพาะที่ (Local Solution) รอบ ๆ คำตอบที่ดีที่สุด

ขั้นตอนที่ 5 การแทนค่าตำแหน่งของเพลี้ยกระโดดลงไปในฟังก์ชันจุดประสงค์

$$F = fitness(x_{new}) \quad (6)$$

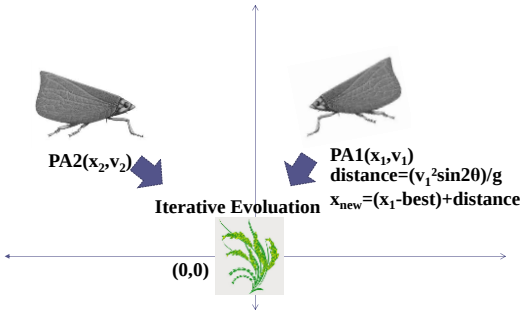
ขั้นตอนที่ 6 คือ การตรวจสอบตำแหน่งของเพลี้ยกระโดดร่วมกับฟังชันจุดประสงค์

$$[fmin\ best] = min\ (Fitness) \quad (7)$$

ขั้นตอนที่ 7 คือ การเก็บค่า โดยทำการบันทึกตำแหน่ง และคำตอบที่ดีที่สุดไว้

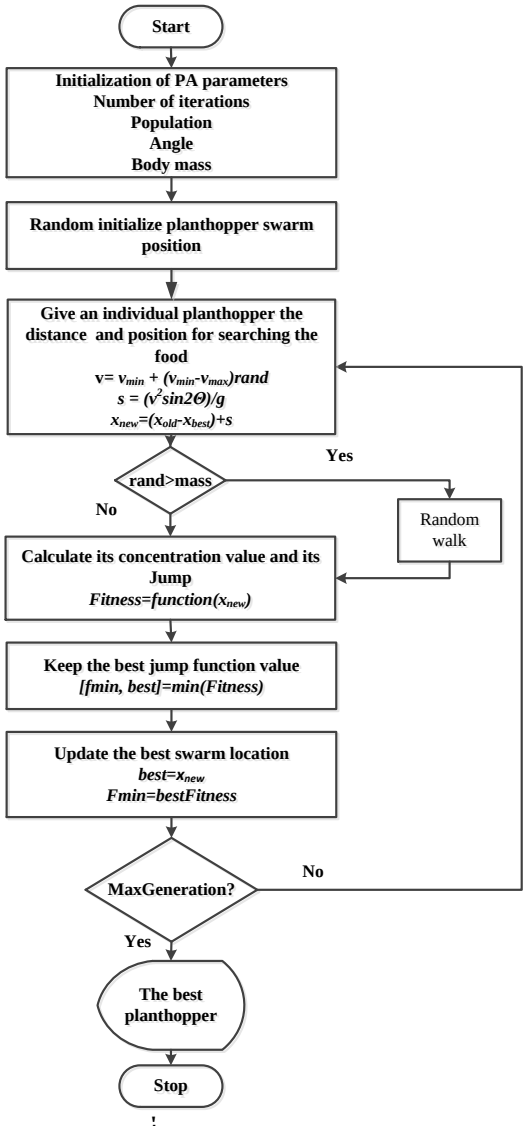
$best = x_{new}$
$fmin = bestFitness$

ขั้นตอนที่ 8 คือ การทำซ้ำ โดยเข้าสู่กระบวนการทำซ้ำในขั้นตอนที่ 3 ถึง 7 จนกว่าจะพบค่าที่เข้าใกล้ 0 มากที่สุด ซึ่งเป็นการลู่เข้าสู่คำตอบที่ดีที่สุด



ภาพที่ 1 กระบวนการการค้นหาอาหารของเพลี้ยกระโดด

ขั้นตอนวิธีแบบเพลี้ยกระโดด แสดงดังภาพที่ 2



ภาพที่ 2 ผังงานของ Planthopper Algorithm

Planthopper algorithm

Objective function $f(x), x = (x_1, \dots, x_d)^T$
Initialize the planthopper population $x_i (i = 1,2 \dots n)$
Initialize take-off angle, body mass r
While (t<Max number of iterations)
  Generate new solutions by adjusting distance, and updating velocities and locations/solutions
  If (rand>r)
    Select a solution among the best solutions
    Generate a local solutions around the selected best solution
  end if
Rank the Planthopper and find the current best
End while
Post-process results and visualization

ภาพที่ 3: Pesudo code of the Planthopper Algorithm (PA)

3.2 ทดสอบประสิทธิภาพ

การวัดประสิทธิภาพของขั้นตอนวิธีแบบเพลี้ยกระโดด สามารถวัดจากการใช้ฟังก์ชันทางคณิตศาสตร์จำนวน 4 ฟังก์ชัน คือ ฟังก์ชัน Sphere ฟังก์ชันไกรแวงก์ ฟังก์ชันโรเซนบร็อก และฟังก์ชันแอกค์ลีย์ ดังรายละเอียดต่อไปนี้

3.2.1 ฟังก์ชัน Sphere

$$f(x) = \sum_{i=1}^{n} (i \cdot x_i^2) \quad (8)$$

ฟังก์ชัน Sphere เมื่อ $-100 \geq x_i \leq 100$ สำหรับคำตอบของฟังก์ชันนี้คือ $x = 0$ ที่ทำให้ $f(x) = 0$

3.2.2 ฟังก์ชันไกรแวงก์ (Griewank function)

$$f(x) = \frac{1}{4000} \sum_{i=1}^{n} x_i^2 - \prod_{i=1}^{n} \cos\left(\frac{x_i}{\sqrt{i}}\right) + 1 \quad (9)$$

ฟังก์ชันไกรแวงก์ เมื่อ $-600 \geq x_i \leq 600$ สำหรับคำตอบของฟังก์ชันนี้คือ $x_i = 0$ ที่ทำให้ $f(x_i) = 0$

3.2.3 ฟังก์ชันโรเซนบร็อก (Rosenbrock function)

$$f(x) = \sum_{i=1}^{n-1} [100((x_{i+1} - x_i^2))^2 + (1 - x_i)^2] \quad (10)$$

ฟังก์ชันโรเซนบร็อก เมื่อ $-2.048 \geq x_i \leq 2.048$ สำหรับคำตอบของฟังก์ชันนี้คือ $x_i = 0$ ที่ทำให้ $f(x_i) = 0$

3.2.4 ฟังก์ชันแอกค์ลีย์ (Ackley function)

$$f(x) = -20 \cdot \exp(-0.2 \cdot \sqrt{\frac{1}{n} \sum_{i=1}^{n} x_i^2}) - \exp(\frac{1}{n} \sum_{i=1}^{n} \cos(2\pi x_i)) + 20 + e \quad (11)$$

ฟังก์ชันแอกค์ลีย์ เมื่อ $-32.768 \geq x_i \leq 32.768$ สำหรับคำตอบของฟังก์ชันนี้คือ $x_i = 0$ ที่ทำให้ $f(x_i) = 0$

4. ผลการดำเนินงานวิจัย

ขั้นตอนวิธีแบบเพลี้ยกระโดดได้รับการพัฒนาด้วยโปรแกรม MATLAB เพื่อทดสอบฟังก์ชันทางคณิตศาสตร์ทั้ง 4 ฟังก์ชัน โดยทำการเปรียบเทียบการลู่เข้าสู่คำตอบกับการค้นหาแบบนกดุเหว่า แบบค้างคาว และแบบแมลงหวี่ตามลำดับ

4.1 กำหนดพารามิเตอร์

กำหนดพารามิเตอร์ที่ใช้ในการทดสอบกับขั้นตอนวิธีทั้ง 4 แบบคือ จำนวนของกลุ่มประชากรเท่ากับ 50 และจำนวนรอบการทำซ้ำเท่ากับ 50 รอบ พารามิเตอร์ตัวอื่นที่มีลักษณะเฉพาะตามขั้นตอนวิธี เช่น ขั้นตอนวิธีการค้นหาแบบนกดุเหว่า ต้องแทนค่าความเป็นไปได้ในการทิ้งรัง = 0.25 และขั้นตอนวิธีค้างคาวต้องกำหนดความถี่ $fmin = 0$ และ $fmax = 5$ ความดังของเสียง = 0.5 และ อัตราการปล่อยพัลส์ (Pulse rate) = 0.5

4.2 ผลการทดสอบ

ส่วนที่ 1 การทดสอบการลู่เข้าสู่คำตอบ เมื่อทำการปรับเปลี่ยนมุมที่ใช้ในการกระโดด (θ)

ตารางที่ 1: ค่าเฉลี่ยการลู่เข้าสู่คำตอบ เมื่อมุมที่ใช้ในการกระโดดเปลี่ยนไป

θ	sphere		griewank		Rosenbrock		ackley	
	ค่าเฉลี่ย	Stdev	ค่าเฉลี่ย	Stdev	ค่าเฉลี่ย	Stdev	ค่าเฉลี่ย	Stdev
5	0.03709	0.04505	0.00020	0.00090	1.18141	1.96590	0.01702	0.06584
10	0.08858	0.06183	0.00029	0.00099	0.54436	0.73204	**0.00657**	0.02879
15	0.13185	0.12590	0.00023	0.00079	0.51377	0.97066	0.03639	0.11029
20	0.04938	0.09829	0.00027	0.00090	0.79798	1.55527	0.03485	0.11226
25	**0.01521**	0.03353	0.00006	0.00019	0.58326	1.16782	0.05669	0.14519
30	0.06382	0.04571	0.00012	0.00049	0.62289	1.16840	0.01587	0.06339
35	0.05989	0.08287	0.00107	0.00029	0.76894	1.04739	0.02482	0.08282
40	0.04701	0.05898	0.00019	0.00080	0.50223	0.88674	0.01217	0.04198
45	0.06507	0.06329	0.00008	0.00024	0.57737	0.73259	0.02913	0.09475
50	0.02920	0.04604	0.00025	0.00084	**0.16647**	0.24043	0.02387	0.07732
55	0.08216	0.11400	0.00016	0.00067	0.42992	0.55836	0.01017	0.03809
60	0.06926	0.06250	0.00012	0.00042	0.74512	1.02619	0.05032	0.11848
65	0.07717	0.08683	0.00028	0.00115	0.54675	1.10845	0.07719	0.17090
70	0.02858	0.05334	**0.00002**	0.00007	0.34221	0.60958	0.01330	0.04680
75	0.06858	0.10446	0.00037	0.00148	0.40926	0.33406	0.01448	0.05282
80	0.11030	0.07776	0.00002	0.00008	0.45831	0.77406	0.03394	0.10529
85	0.07055	0.08081	0.00068	0.00152	0.35271	0.46694	0.04916	0.13334
90	0.05716	0.02840	0.00018	0.00068	0.69796	0.93650	0.02269	0.05753

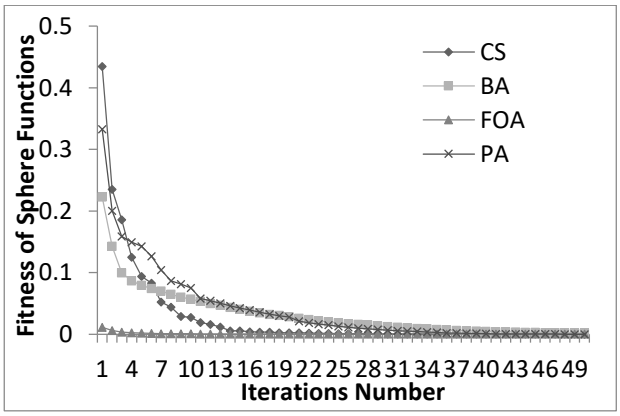
ตารางที่ 1 เป็นผลการทดสอบการลู่เข้าสู่คำตอบ เมื่อทำปรับเปลี่ยนมุมที่ใช้ในการกระโดด (θ) พบว่า ค่าเฉลี่ยของการลู่เข้าสู่คำตอบมีค่าที่ดี เมื่อปรับเปลี่ยนมุมที่ใช้ในการกระโดดให้เหมาะสมกับแต่ละฟังก์ชัน

ส่วนที่ 2 จากการทดสอบจำนวน 20 ครั้ง ผลการทดสอบด้วยฟังก์ชันทางคณิตศาสตร์ทั้ง 4 ฟังก์ชันดังตารางที่ 2

ตารางที่ 2: เปรียบเทียบผลลัพธ์ที่น้อยที่สุดของฟังก์ชันมาตรฐาน

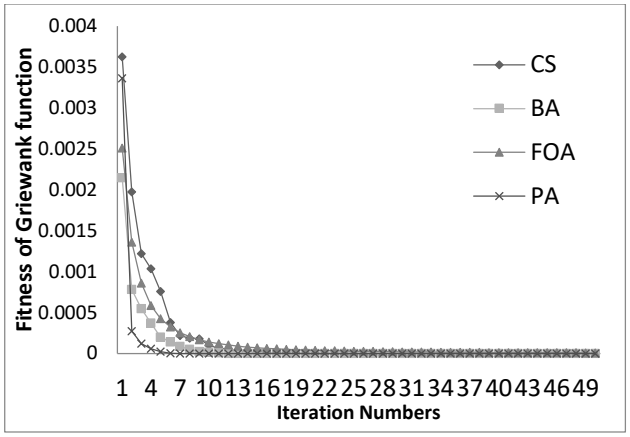
ฟังก์ชัน	ขั้นตอนวิธี	CS	BA	FOA	PA
Sphere	ค่าเฉลี่ย	0.02792	0.032755	**0.000683**	0.04025
	StdDev	0.075389	0.040887	0.001815	0.067782
Griewank	ค่าเฉลี่ย	0.000196	0.000087	0.000161	**0.000076**
	StdDev	0.000614	0.000330	0.000415	0.000476
Rosenbrock	ค่าเฉลี่ย	**0.219771**	1.34574	0.699449	0.518115
	StdDev	0.595578	0.82274	0.051182	1.139391
Ackley	ค่าเฉลี่ย	0.042266	0.041904	0.067097	**0.023046**
	StdDev	0.112516	0.107428	0.096404	0.06525

ตารางที่ 2 เป็นผลการเปรียบเทียบผลลัพธ์ของขั้นตอนวิธี 4 ประเภทได้แก่ ขั้นตอนวิธีการค้นหาแบบนกดุเหว่า แบบค้างคาว แบบแมลงหวี่ และแบบเพลี้ยกระโดด จากตารางจะเห็นได้ว่า ขั้นตอนวิธีแบบเพลี้ยกระโดดให้ค่าความเหมาะสมที่ดีที่สุดในส่วนของฟังก์ชันไกรแวงก์ได้ค่าความเหมาะสมเฉลี่ยเท่ากับ 0.000076 และฟังก์ชันแอกค์ลีย์ได้ค่าความเหมาะสมเฉลี่ยเท่ากับ 0.02346 รองลงมาเป็นฟังก์ชัน โรเซนบร็อกได้ค่าความเหมาะสมเฉลี่ยเท่ากับ 0.518115 และฟังก์ชัน Sphere ได้ค่าความเหมาะสมเฉลี่ยเท่ากับ 0.04025



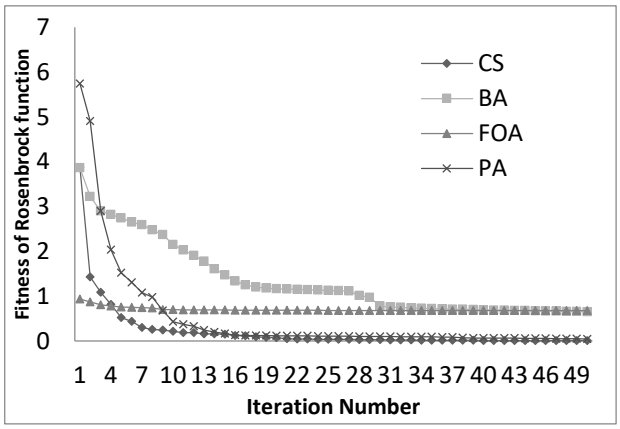
ภาพที่ 4: การเปรียบเทียบประสิทธิภาพของ CS BA FOA และ PA ด้วยฟังก์ชัน Sphere

จากภาพที่ 4 การลู่เข้าสู่คำตอบของฟังก์ชัน Sphere พบว่า ขั้นตอนวิธีแบบแมลงหวี่สามารถหาค่าความเหมาะสมได้ดีกว่าขั้นตอนวิธีแบบเพลี้ยกระโดด



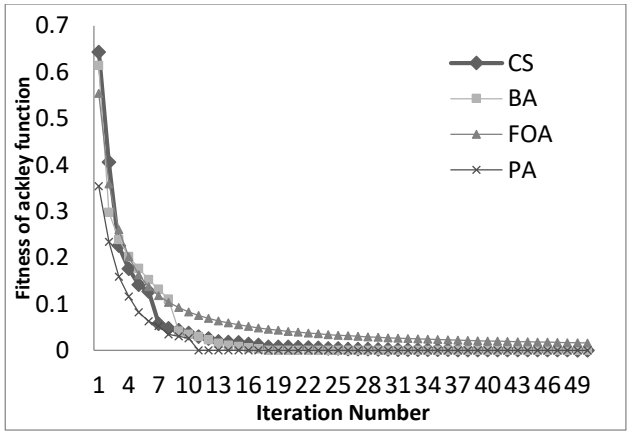
ภาพที่ 5: การเปรียบเทียบประสิทธิภาพของ CS BA FOA และ PA ด้วยฟังก์ชันไกรแวงก์

จากภาพที่ 5 การลู่เข้าสู่คำตอบของฟังก์ชันไกรแวงก์ พบว่า ขั้นตอนวิธีแบบเพลี้ยกระโดดสามารถหาค่าความเหมาะสมได้ดีกว่าขั้นตอนวิธีการค้นหาแบบนกดุเหว่า แบบค้างคาว และแบบแมลงหวี่



ภาพที่ 6: การเปรียบเทียบประสิทธิภาพของ CS BA FOA และ PA ด้วยฟังก์ชันโรเซนบร็อก

จากภาพที่ 6 การลู่เข้าสู่คำตอบของฟังก์ชันโรเซนบร็อกพบว่า ขั้นตอนวิธีการค้นหาแบบนกดุเหว่าสามารถหาค่าความเหมาะสมได้ดีกว่าขั้นตอนวิธีแบบเพลี้ยกระโดด แบบค้างคาว และแบบแมลงหวี่



ภาพที่ 7: การเปรียบเทียบประสิทธิภาพของ CS BA FOA และ PA ด้วยฟังก์ชันแอกค์ลีย์

จากภาพที่ 7 การลู่เข้าสู่คำตอบของฟังก์ชันแอกค์ลีย์ พบว่า ขั้นตอนวิธีแบบเพลี้ยกระโดดสามารถหาค่าความเหมาะสมได้ดีกว่าขั้นตอนวิธีการค้นหาแบบนกดุเหว่า แบบค้างคาว และแบบแมลงหวี่

5. สรุป

งานวิจัยนี้นำเสนอขั้นตอนวิธีแบบเพลี้ยกระโดด (Planthopper Algorithm: PA) เป็นขั้นตอนวิธีสำหรับการหาค่าความเหมาะสมที่สุดแบบที่ได้รับแรงบันดาลใจจากพฤติกรรมการหาอาหารแบบเป็นกลุ่ม ซึ่งจากการเปรียบเทียบกับขั้นตอนวิธีการค้นหาแบบนกดุเหว่า แบบค้างคาว และแบบแมลงหวี่ ด้วยฟังก์ชันทางคณิตศาสตร์ 4 ฟังก์ชันคือ ฟังก์ชัน Sphere ฟังก์ชันไกรแวงก์ ฟังก์ชันโรเซนบร็อก และฟังก์ชันแอกค์ลีย์ พบว่า ขั้นตอนวิธีแบบเพลี้ยกระโดดสามารถหาค่าความเหมาะสมได้เทียบเคียงกับขั้นตอนวิธีอื่น และดีกว่าขั้นตอนวิธีอื่นในบางฟังก์ชัน

การวิจัยต่อในอนาคตจำเป็นต้องเพิ่มประสิทธิภาพการหาค่าความเหมาะสมโดยเพิ่มเติมพารามิเตอร์ เช่น อัตราการตาย จำนวนเพศผู้เพศเมีย การจากฝูงไป การเข้าฝูงใหม่ เป็นต้น

เอกสารอ้างอิง

[1] Rai, Deepak, and Kirti Tyagi. “Bio-inspired optimization techniques: a critical comparative study.” *ACM SIGSOFT Software Engineering Notes*, 2013, pp. 1-7.

[2] ปารเมศ ชุติมา, เพ็ญพักตร์ ปิ่นภุมรีร์ม และพัฒนพงษ์ แสงหัตวัฒนา. “การประยุกต์ใช้เมมเมติกอัลกอริธึมสำหรับปัญหาการ

จัดลำดับสายการประกอบแบบผลิตผสมที่มีหลายวัตถุประสงค์ในระบบผลิตแบบทันเวลาพอดี." *การประชุมวิชาการข่ายงานวิศวกรรมอุตสาหการ*. 2550, หน้า 72-77.

[3] สุพจน์ เจาชัยเจริญกุล และเดชา พวงดาวเรือง, "เทคนิคการหาค่าเหมาะที่สุดแบบอภิศึกษาสำนึกแนวใหม่: การค้นหาแบบแก็ง," *การประชุมวิชาการด้านการวิจัยดำเนินงานแห่งชาติประจำปี 2556*, หน้า 233-240.

[4] C. Blum, and A. Roli. "Metaheuristics in combinatorial optimization: Overview and conceptual comparison." *ACM Computing Surveys*, 2003, pp.268-308.

[5] สุภกิจ นุตยะสกุล, "การคำนวณเชิงวิวัฒนาการระหว่าง เจเนติกอัลกอริทึม กับ พาติคอลสวอมออปติมันเซชัน" *Journal of Information Science and Technology*, Vol.2, No.2, pp.13-22.

[6] R. Eberhart and J. Kennedy, "A new optimizer using particle swarm theory," *Proceedings of the Sixth International Symposium on Micro Machine and Human Science* (MHS'95), 1995, pp. 39-43.

[7] Dorigo, M., & Di Caro, G. "Ant colony optimization: a new meta-heuristic." *Proceedings of Evolutionary Computation*, 1999, pp. 1470-1477.

[8] I. Fister, I. Fister Jr, X.-S. Yang, and J. Brest, "A comprehensive review of firefly algorithms," *Swarm and Evolutionary Computation*, vol. 13, pp. 34-46.

[9] X.S. Yang and S. Deb, "Cuckoo search via Levy flights," *World Congress on Nature & Biologically Inspired Computing* (NaBIC 2009), 2009, pp. 210-214.

[10] X.S. Yang, "A new metaheuristic bat-inspired algorithm," Nature inspired cooperative strategies for optimization (NICSO 2010): Springer, 2010, pp. 65-74.

[11] B. Xing , and W.J. Gao. "Fruit Fly Optimization Algorithm," *Innovative Computational Intelligence: A Rough Guide to 134 Clever Algorithms*. Springer International Publishing, 2014, pp. 167-170.

[12] "Mathematical optimization". Retrieved May 1, 2017, from https://en.wikipedia.org/wiki/Mathematical_optimization

[13] I. Fister, X.S. Yang, and J. Brest. "A comprehensive review of firefly algorithms." *Swarm and Evolutionary Computation*, 2013, pp. 34-46.

[14] B, Malcolm. "Jumping mechanisms in flatid planthoppers (Hemiptera, Flatidae)." *The Journal of experimental biology* 217.14, 2014, pp.2590-2600.

การเปรียบเทียบเทคนิคเหมืองข้อมูลในการพยากรณ์ความสำเร็จการศึกษาของนักเรียนระดับประกาศนียบัตรวิชาชีพ

The Comparison of Data Mining Techniques for Predicting Student Success in Vocation Education

พัฒนพงษ์ คลรัตน์ (Pattanaphong Donrat)[1] และจารี ทองคำ (Jaree Thongkam)[2]
สาขาเทคโนโลยีสารสนเทศ คณะวิทยาการสารสนเทศ มหาวิทยาลัยมหาสารคาม
[1]s2s9890@gmail.com, [2]jaree.thongkam@gmail.com

บทคัดย่อ

งานวิจัยนี้มีจุดประสงค์เพื่อเปรียบเทียบประสิทธิภาพของเทคนิคเหมืองข้อมูลในการสร้างแบบจำลองเพื่อพยากรณ์การสำเร็จการศึกษาในสถานศึกษาอาชีวศึกษา ซึ่งแบบจำลองสามารถนำไปใช้ในการแนะนำการเข้าศึกษาต่อของนักเรียนระดับมัธยมตอนต้นเพื่อศึกษาต่อในระดับอาชีวศึกษาเอกชน ข้อมูลถูกรวบรวมจากสถานศึกษาอาชีวศึกษาเอกชนจังหวัดกาฬสินธุ์ในปีการศึกษา 2557-2558 ข้อมูลประกอบด้วย 1,266 ข้อมูล 4 เทคนิคที่เป็นที่นิยมและมีประสิทธิภาพในการสร้างแบบจำลอง คือ C4.5, Random Forest, Random Tree และ Reduced Error Pruning Tree (REP Tree) โดยใช้หลักการ 10-fold cross validation ในการแบ่งกลุ่มข้อมูลเป็นชุดข้อมูลเรียนรู้และชุดข้อมูลทดสอบ และวัดประสิทธิภาพการพยากรณ์ของแบบจำลองด้วยค่าความถูกต้อง (Accuracy) ค่าความแม่นยำ (Precision) และค่าความระลึก (Recall) จากการศึกษาพบว่าแบบจำลอง C4.5 มีประสิทธิภาพในการพยากรณ์ด้วยค่าความถูกต้อง (Accuracy) สูงถึง 87.73% ค่าความแม่นยำ (Precision) 87.5% และค่าความระลึก (Recall) 87.7%

คำสำคัญ : แบบจำลองการพยากรณ์ความสำเร็จการศึกษาประกาศนียบัตรวิชาชีพ ต้นไม้การตัดสินใจเหมืองข้อมูล

Abstract

This research is to compare the efficiency of data mining techniques in modeling for predicting the Graduate in Private Vocational Education Institutes. The models could guide the study of junior high school students to be educated at the Private vocational. Data were collected from Private Vocational Education Institutes in Kalasin Province from 2014-2015. The data consist of 1,266 records. Four efficient techniques including C4.5, Random Forest, Random Tree and Reduced Error Pruning Tree (REP Tree) are used for building the prediction models. 10-fold cross validation is applied to segment the data into training and testing sets. Moreover, accuracy, precision and recall are used as the criteria for selecting the most effective model. According to the research, the experimental results showed that C4.5 is the most efficient prediction model which has accuracy up to 87.73%, precision 87.50%, and recall 87.70%, respectively.

Keyword: Predicting Graduation Model, Vocational Education Certificate, Decision Tree, Data Mining.

1. บทนำ

ปัจจุบันประเทศไทยนั้นอยู่สภาวะขาดแคลนแรงงานฝีมือ กระทรวงศึกษาธิการจึงได้มีนโยบายในปีพุทธศักราช 2554 การเปลี่ยนแปลงสัดส่วนของผู้เรียนอาชีวศึกษาต่อผู้เรียนสายสามัญจากเดิม 40:60 เป็น 60:40 ภายในปีพุทธศักราช 2561 ซึ่งเป็นปีสิ้นสุดของการปฏิรูปการศึกษาทศวรรษที่ 2 [1]

การเพิ่มจำนวนผู้สำเร็จการศึกษาตามระยะเวลาของหลักสูตรของอาชีวศึกษา ซึ่งหลักสูตรอาชีวศึกษาแบ่งออกได้เป็น 2 หลักสูตรคือ 1) หลักสูตรประกาศนียบัตรวิชาชีพ (ปวช.)

มีระยะเวลาการศึกษาตามหลักสูตร 3 ปี 2) หลักสูตรประกาศนียบัตรวิชาชีพชั้นสูง (ปวส.) มีระยะเวลาการศึกษาตามหลักสูตร 2 ปี ซึ่งการจัดให้ผู้เรียนสำเร็จการศึกษาตามระยะเวลาของหลักสูตรนั้น เป็นตัวบ่งชี้ตัวหนึ่งที่แสดงว่าสถานศึกษาได้จัดการศึกษาอย่างมีคุณภาพและประสิทธิภาพ ตามตัวบ่งชี้ของการประกันคุณภาพการศึกษาทั้งการประกันคุณภาพภายในและการประกันคุณภาพภายนอก [1]

สถานศึกษาอาชีวศึกษาเอกชน [2] ก็เป็นหน่วยงานหนึ่งในการจัดการศึกษาอาชีวศึกษาที่มีบทบาท สำคัญในการให้บริการการศึกษาทางด้านวิชาชีพเช่นเดียวกับสถานศึกษาอาชีวศึกษารัฐบาลซึ่งล้วนแต่มีบทบาทและความสำคัญอย่างยิ่งต่อการพัฒนาประเทศเพราะเป็นการศึกษาที่จัดเตรียมบุคคลให้มีอาชีพเป็นหลักในอนาคต และช่วยให้ผู้ที่มีอาชีพอยู่แล้วมีความก้าวหน้าในอาชีพนั้นเพิ่มขึ้น

ในการนี้ได้มีนักวิจัยที่นำเอาเหมืองข้อมูล ซึ่งเป็นกระบวนการในการค้นหาความรู้จากข้อมูล มาใช้ในการวิเคราะห์ เช่น เยาวภา ภารสำเร็จและคณะ [3] ได้ทำการเปรียบเทียบ 3 เทคนิค ในเหมืองข้อมูล ประกอบด้วย C4.5, Naïve Bayes และ k-Nearest Neighbour เพื่อใช้ในการจำแนกปัจจัยที่ส่งผลต่อระดับผลการเรียนของนักศึกษา ผลการเปรียบเทียบพบว่าเทคนิค C4.5 ให้ค่าประสิทธิภาพสูงที่สุด 73.55% ภรัณยา ปาลวิสุทธิ์ [4] พบว่าเทคนิค Random Forest มีค่าความแม่นยำร้อยละ 87.15 สามารถพยากรณ์ได้ดีกว่า J48 ID3 LMT และ CART ส่วน Kalmegh [5] พบว่า Random Tree มีค่าความถูกต้อง 100% ซึ่งมากกว่า REP Tree และ Simple Cart ในการจำแนกข่าวอินเดีย Kittipol Wisaeng [6] พบว่า REP Tree ให้ค่าความถูกต้องที่ 92.87% ในการจำแนก UCI Repository ได้แก่ Nursery, Iris, Anneal, Shuttle_trn, Voting, Waveform และ Sick

จะเห็นว่าเทคนิค Decision Tree เป็นเทคนิคที่มีประสิทธิภาพในการจำแนก ดังนั้นงานวิจัยนี้ได้ใช้ 4 เทคนิคในกลุ่มของ Decision Tree ที่มีประสิทธิภาพในการสร้างแบบจำลองเพื่อการพยากรณ์การสำเร็จการศึกษาในสถานศึกษาอาชีวศึกษาเอกชนจังหวัดกาฬสินธุ์ คือ C4.5, Random Forest, Random Tree และ Reduced Error Pruning Tree (REP Tree) คณะผู้วิจัยได้ใช้หลักการ 10-fold cross validation ในการแบ่งกลุ่มข้อมูลเป็นชุดข้อมูลเรียนรู้ และชุดข้อมูลทดสอบ และวัดประสิทธิภาพการพยากรณ์ของแบบจำลองด้วยค่าความถูกต้อง (Accuracy) ค่าความแม่นยำ (Precision) และค่าความระลึก (Recall) ในการแนะนำการเข้าศึกษาต่อของนักเรียนระดับมัธยมตอนต้นเพื่อศึกษาต่อในระดับอาชีวศึกษาเอกชนจังหวัดกาฬสินธุ์

2. งานวิจัยที่เกี่ยวข้อง

2.1 สถานศึกษาอาชีวศึกษาเอกชนในจังหวัดกาฬสินธุ์

สถานศึกษาอาชีวศึกษาเอกชนสังกัดสำนักงานคณะกรรมการส่งเสริมการศึกษาเอกชน (สช.) และมีหน่วยงานต้นสังกัดคือสำนักงานเขตพื้นที่การศึกษาประถมศึกษากาฬสินธุ์ เขต1 เขต2 และเขต3 ตามพื้นที่ที่สถานศึกษาตั้งอยู่ แต่ใช้หลักสูตรของสำนักงานคณะกรรมการการอาชีวศึกษา (สอศ.) จนมาถึงวันที่ 12 กุมภาพันธ์ 2559 เว็บไซต์ราชกิจจานุเบกษาได้เผยแพร่คำสั่งหัวหน้าคณะรักษาความสงบแห่งชาติที่ 8/2559 เรื่องการบริหารจัดการรวมสถานศึกษาอาชีวศึกษาภาครัฐและภาคเอกชน โดยให้รวมสถานศึกษาทั้งภาครัฐและเอกชนไว้ด้วยกันภายใต้สังกัดอาชีวศึกษาจังหวัดกาฬสินธุ์ (อศจ.กาฬสินธุ์)

2.2 เทคนิคในเหมืองข้อมูล

เทคนิคในเหมืองข้อมูลได้ถูกนำมาใช้ในการสร้างแบบจำลองกันอย่างแพร่หลายและมีประสิทธิภาพ เช่น C4.5, Random Forest, Random Tree และ Reduced Error Pruning Tree (REP Tree) ซึ่งสามารถอธิบายได้ดังต่อไปนี้

1. เทคนิค C4.5 [7] เป็นเทคนิคในการสร้างต้นไม้การตัดสินใจพัฒนาโดย J. Ross Quinlan ในปี 1993 โดยนำเอา ID3 มาปรับปรุงให้มีความสามารถมากขึ้นใช้วิธีการ Information Gain เพิ่มเติมการจัดการกับข้อมูล, ตัวเลข, ข้อมูลที่ขาดไปและไม่สมบูรณ์ และการ Prune ด้วยการแทนกิ่ง (Branch) ที่ไม่ช่วยในการตัดสินใจด้วย Leaf Node ที่ตัดสินใจได้ดีกว่า การแบ่งของ Tree ในการทำงานขั้นตอนแรกคล้ายกับการทำงานด้วย ID3 คือต้องหา Info และ Gain ออกมาก่อนซึ่งมีนักวิจัยหลายท่านได้นำเอาเทคนิคนี้มาใช้ในการพยากรณ์

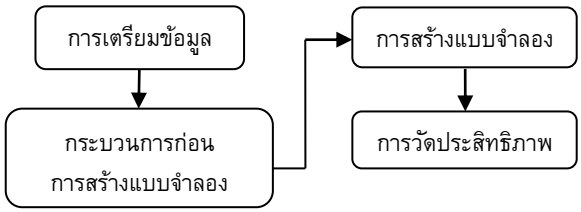
2. เทคนิค Random Forest [8] เป็นเทคนิคการสุ่มเลือกใช้ข้อมูลและคุณลักษณะ Decision Tree ซึ่งถูกสร้างจากการนำข้อมูลไปสุ่มเลือกตัวอย่างแบบเลือกแล้วใส่กลับ (Sampling with Replacement) แล้วนำมาสร้างเป็น Tree ซึ่งจะมีตัวอย่างส่วนหนึ่งที่ไม่ถูกเลือก ซึ่งข้อมูลส่วนนี้เรียกว่า Out-of-Bag (OOB) จะถูกนำมาใช้ในการทดสอบ Decision Tree วิธีการดังกล่าวนี้เรียกว่า Bagging ผลลัพธ์ที่ได้อย่างอิสระจาก Decision Tree ในแต่ละต้นถูกนำมาคิดเป็นผลการโหวต ผลโหวตที่มากที่สุดจะใช้ระบุสถานะของคลาส เทคนิค Random Forest ไม่จำเป็นต้องมีข้อมูลทดสอบ เพื่อประมาณความผิดพลาดเพราะข้อมูล OOB นั้นถูก นำมาใช้ทดสอบ Decision Tree นั้นแล้ว

3. เทคนิค Random Tree [9] คือ เทคนิคที่ใช้ในการจำแนกหมวดหมู่เช่นเดียวกับ C4.5 โดยมีหลักการสร้าง Tree จากการสุ่ม Tree หลายๆ แบบในแต่ละโหนดแล้วเลือกมาประมวลผลโดยไม่ใช้การ Prune และเนื่องจากจำนวนของ Tree เพิ่มขึ้นอย่างรวดเร็ว ซึ่งยากแก่การแก้ปัญหา การสร้าง Tree ที่เป็นไปได้ทั้งหมด ส่วนประกอบสามารถสร้างชุดสุ่มของ Tree ออกมาจากการกระจายชุดต่างๆของ Tree Random Tree เป็นการสุ่มวาดที่สุ่มจากชุดของ Tree ที่เป็นไปได้ ในบริบทนี้ "สุ่ม" หมายความว่า Tree ในชุดของ Tree แต่ละมีโอกาสเท่าเทียมกันของการเป็นตัวอย่าง วิธีที่บอกนี้ก็คือว่าการกระจายของ Tree คือ "ชุด" Tree สุ่มแบบต่อเนื่อง (CRT) เป็นแบบสุ่ม Tree จริง T_c และมีระยะในการเดินทางสั้นที่สุด Tree ชุดย่อยประกอบด้วย Tree ที่มี n จุด แต่ละองค์ประกอบของโครงสร้างที่แท้จริงคือราก ซึ่งมีนักวิจัยจำนวนมากได้นำเทคนิค Random Tree มาใช้ในการจำแนก

4. เทคนิค Reduced Error Pruning Tree (REP Tree) [9] คือ เทคนิคที่ใช้ regression tree logic และสร้าง Tree หลายๆ ต้นที่แตกต่างกัน หลังจากนั้นก็เลือกที่ดีที่สุดจาก Tree ที่สร้างทั้งหมดมาเป็นตัวแทนของ Tree ทั้งหมด ในการตัดกิ่งใช้ค่า mean square error ในการพยากรณ์เป็นพื้นฐานการวัด Reduced Error Pruning Tree (REP Tree) เป็น Decision Tree ที่มีการเรียนรู้และสร้างแบบจำลองอย่างรวดเร็วบนพื้นฐานของ Information gain หรือ reducing the variance และตัดกิ่งโดยใช้การลดข้อผิดพลาดในการตัดแต่ใช้ได้เฉพาะตัวแปรที่เป็นตัวเลขเท่านั้น มีนักวิจัยจำนวนมากได้นำเทคนิค REP Tree มาใช้ในการจำแนก

3. ขั้นตอนวิธีการดำเนินการวิจัย

ในการทำวิจัยคณะผู้วิจัยได้มีแบ่งวิธีการดำเนินการวิจัยออกเป็น 4 ขั้นตอนหลัก 1) การเตรียมข้อมูล 2) กระบวนการก่อนการสร้างแบบจำลอง 3) การสร้างแบบจำลอง และ 4) การวัดประสิทธิภาพแบบจำลอง ดังภาพที่ 1



ภาพที่ 1 The process of data mining

3.1 การเตรียมข้อมูล

การเตรียมข้อมูลของการศึกษาค้นคว้าอิสระครั้งนี้ ได้ข้อมูลมาจากการสำรวจข้อมูลจำนวนนักเรียนจากสถานศึกษาอาชีวศึกษาเอกชนในจังหวัดกาฬสินธุ์ ประจำปีการศึกษา 2557-2558 จำนวน 1,266 ข้อมูล จากสถานศึกษา 12 แห่ง ดังตารางที่ 1

ตารางที่ 1 จำนวนนักเรียนประจำปีการศึกษา 2557-2558

รายชื่อสถานศึกษา	2557	2558
1. วิทยาลัยเทคโนโลยีพณิชยการกาฬสินธุ์	176	197
2. วิทยาลัยเทคโนโลยีกมลาไสย	38	35
3. วิทยาลัยเทคโนโลยีเพิ่มพูนบริหารธุรกิจ	21	19
4. วิทยาลัยอาชีวศึกษาลำปาว	13	14
5. วิทยาลัยเทคโนโลยีสหัสขันธุ์	15	16
6. วิทยาลัยเทคโนโลยีสามชัย	9	11
7. วิทยาลัยอาชีวศึกษาไท-เทคกาฬสินธุ์	12	15
8. วิทยาลัยอาชีวศึกษาไทยเทคอีสาน	23	12
9. วิทยาลัยเทคโนโลยีสมเด็จพณิชยการ	132	268
10. วิทยาลัยเทคโนโลยีพัฒนบัณฑิต	18	10
11. วิทยาลัยอาชีวศึกษาไทยเทคเอเชีย	130	82
12. วิทยาลัยเทคโนโลยีณัชวินทร์	-	-
รวม	587	679
รวมทั้งสิ้น	1,266	

3.2 การทำกระบวนการก่อนการสร้างแบบจำลอง

ในขั้นตอนนี้ผู้วิจัยได้ทำการแปลงข้อมูลและทำการวิเคราะห์ปัจจัย

1. การแปลงข้อมูล (Data transformation) จากข้อมูลที่ได้จากการสำรวจข้อมูลจำนวนนักเรียนจากสถานศึกษาอาชีวศึกษาเอกชนในจังหวัดกาฬสินธุ์ ประจำปีการศึกษา 2557-2558 จำนวน 12 แห่ง มีทั้งหมด 20 ตัวแปร ดังตารางที่ 2

ตารางที่ 2 ตัวแปรในงานวิจัย

ชื่อตัวแปร	คำอธิบายตัวแปร	ชนิดของตัวแปร
Sex	เพศ	Nominal
Age	อายุ	Numeric
Number of siblings	จำนวนพี่น้อง	Numeric
Parent Status	สถานภาพครอบครัว	Nominal
Father Career	อาชีพบิดา	Nominal
Mother Career	อาชีพมารดา	Nominal
Village	หมู่บ้าน	Nominal
Tumbol	ตำบล	Nominal
Amphoe	อำเภอ	Nominal
Province	จังหวัด	Nominal
Thai	เกรดเฉลี่ยกลุ่มภาษาไทย	Numeric
Math	เกรดเฉลี่ยกลุ่มคณิตศาสตร์	Numeric
Science	เกรดเฉลี่ยกลุ่มวิทยาศาสตร์	Numeric
Social	เกรดเฉลี่ยกลุ่มสังคมศึกษา ศาสนา และวัฒนธรรม	Numeric
Health	เกรดเฉลี่ยกลุ่มสุขศึกษา พลศึกษา	Numeric
Art	เกรดเฉลี่ยกลุ่มศิลปะ	Numeric
Career and Technology	เกรดเฉลี่ยกลุ่มการงานอาชีพและเทคโนโลยี	Numeric
Foreign language	เกรดเฉลี่ยกลุ่มภาษาต่างประเทศ	Numeric
Type of course	ประเภทวิชาที่จบ	Nominal
Graduate	จบการศึกษา	Nominal

2. การวิเคราะห์ปัจจัยเป็นการคัดเลือกแอตทริบิวต์ที่สามารถเป็นตัวแทนของกลุ่มแอตทริบิวต์เพื่อลดจำนวนแอตทริบิวต์ในการพยากรณ์ซึ่งงานวิจัยนี้ได้นำเทคนิค Gain Ratio Attribute Eval มาใช้ในการวิเคราะห์ปัจจัยผลการวิเคราะห์ทำการตัดแอตทริบิวต์ที่มีผลต่อการพยากรณ์น้อยออกจำนวน 7 แอตทริบิวต์คือ Tumbol, Amphoe, Age, Career and Technology, Thai, Number of siblings และ Province คงเหลือ 13 แอตทริบิวต์ ดังนี้ Father Career, Mother Career, Village, Parent Status, Sex, Type of course, Science, Math, Social, Art, Foreign language, Health และ Graduate(Class)

3.3 การสร้างแบบจำลอง

การสร้างแบบจำลองพยากรณ์การสำเร็จการศึกษาของผู้ที่สนใจเข้าศึกษาต่อสถานศึกษาอาชีวศึกษาเอกชนในจังหวัดกาฬสินธุ์ เพื่อช่วยในการตัดสินว่าเมื่อเข้าศึกษาต่อแล้วจะสามารถสำเร็จการศึกษาได้หรือไม่ ด้วยตัวแปรทั้ง 20 ตัวแปรที่ได้จากการสำรวจข้อมูลจำนวนนักเรียนในปีการศึกษา 2557-2558 จำนวน 1,266 ข้อมูล แบ่งเป็น 2 คลาส คือ สำเร็จการศึกษาจำนวน 812 ข้อมูลและไม่สำเร็จการศึกษาจำนวน 454 ข้อมูล ด้วยการทำเหมืองข้อมูลและเทคนิคที่นำมาใช้ในการสร้างแบบจำลองมีจำนวน 4 เทคนิคคือ

1. เทคนิค C4.5
2. เทคนิค Random Forest
3. เทคนิค Random Tree
4. เทคนิค REP Tree

3.4 การวัดประสิทธิภาพของแบบจำลอง

ในการวัดประสิทธิภาพของแบบจำลองนั้น ได้มีการใช้เทคนิคแบบ 10-fold cross validation โดยจะทำการแบ่งข้อมูลออกเป็น 10 ชุดเท่าๆกัน จากนั้นจะทำการทดสอบทั้งหมด 10 รอบ โดยในแต่ละรอบจะใช้ข้อมูล 1 ชุดเป็นชุดทดสอบและอีก 9 ชุดที่เหลือเป็นชุดฝึกสอน ในรอบต่อไปก็ใช้ชุดข้อมูลถัดไปเป็นชุดทดสอบจนครบทั้ง 10 ชุดข้อมูล ซึ่งข้อมูลทั้งหมดจำนวน 1266 ข้อมูล แบ่งออกเป็น 10 ชุดข้อมูล ซึ่งคิดเป็นอัตราข้อมูลทดสอบต่อข้อมูลฝึก เป็นอัตราส่วน 10:90

การวิเคราะห์ประสิทธิภาพ คือ การวัดประสิทธิภาพการทำงานในแต่ละขั้นตอนวิธี สามารถวัดได้จากผลของการจำแนกกลุ่มข้อมูล โดยค่าของผลลัพธ์ที่ได้จากการจำแนกคือ ค่า True Positive (TP) ค่า True Negative (TN) ค่า False Positive (FP) ค่า False Negative (FN) และค่าความถูกต้อง (Accuracy) คือ ค่าที่แบบจำลองสามารถพยากรณ์ความสำเร็จการศึกษาหรือไม่สำเร็จได้ถูกต้อง ต่อนักเรียนทั้งหมด ดังสมการที่ 1

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (1)$$

ค่าความแม่นยำ (Precision) คือ ค่าที่แบบจำลองสามารถพยากรณ์นักเรียนที่สำเร็จการศึกษาได้ถูกต้อง ต่อจำนวนนักเรียนที่พยากรณ์สำเร็จการศึกษาทั้งหมด ดังสมการที่ 2

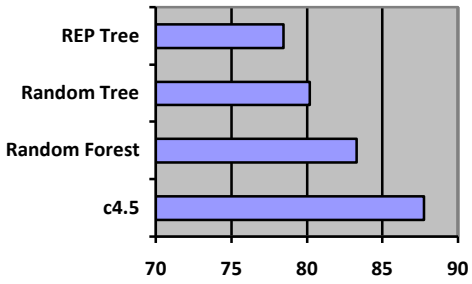
$$Precision = \frac{TP}{TP + FP} \quad (2)$$

และค่าความระลึก (Recall) คือ ค่าที่แบบจำลองสามารถพยากรณ์นักเรียนที่สำเร็จการศึกษาได้ถูกต้อง ต่อจำนวนนักเรียนที่สำเร็จการศึกษาจริงทั้งหมด ดังสมการที่ 3

$$\mathrm{Re}call = \frac{TP}{TP + FN} \quad (3)$$

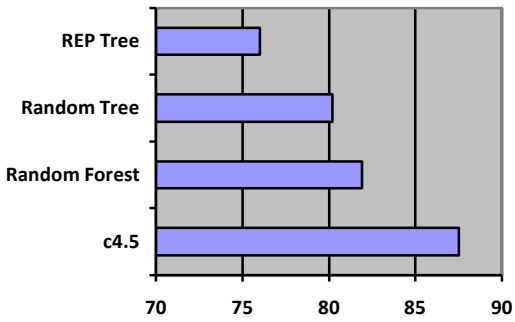
4. ผลการดำเนินการวิจัย

ในการศึกษาประสิทธิภาพของแบบจำลองนี้คณะผู้วิจัยได้นำเอาโปรแกรม WEKA เวอร์ชั่น 3.9.1 มาเป็นเครื่องมือมาใช้ในการสร้างแบบลองด้วยเทคนิค C4.5, Random Forest, Random Tree และ REP Tree โดยใช้ค่าพารามิเตอร์เริ่มต้นของโปรแกรม WEKA ในงานวิจัยนี้คณะผู้วิจัยได้นำเอา 10 fold cross-validation มาใช้ในการแบ่งข้อมูลออกเป็นชุดเรียนรู้ และชุดทดสอบ และได้วัดประสิทธิภาพของแบบจำลองด้วย Accuracy, Precision และ Recall ผลการทดลองสามารถแสดงได้ดัง ภาพที่ 2, ภาพที่ 3 และภาพที่ 4



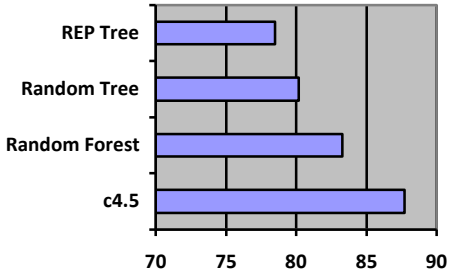
ภาพที่ 2 ผลการวัดประสิทธิภาพของแบบจำลองด้วย Accuracy

ภาพที่ 2 แสดงให้เห็นค่า Accuracy ของแบบจำลองได้แก่ เทคนิค C4.5 ให้ค่า Accuracy มากที่สุด 87.73% ต่อมา เทคนิค Random Forest ให้ค่า Accuracy 83.29% เทคนิค Random Tree ให้ค่า Accuracy 80.20% เทคนิค REP Tree ให้ค่า Accuracy 78.45% ตามลำดับ



ภาพที่ 3 ผลการวัดประสิทธิภาพของแบบจำลองด้วย Precision

ภาพที่ 3 แสดงให้เห็นค่า Precision ของแบบจำลองได้แก่ เทคนิค C4.5 ให้ค่า Precision มากที่สุด 87.5% ต่อมา เทคนิค Random Forest ให้ค่า Precision 81.9% เทคนิค Random Tree ให้ค่า Precision 80.2% เทคนิค REP Tree ให้ค่า Precision 76.0% ตามลำดับ



ภาพที่ 4 ผลการวัดประสิทธิภาพของแบบจำลองด้วย Recall

ภาพที่ 4 แสดงให้เห็นค่า Recall ของแบบจำลองได้แก่ เทคนิค C4.5 ให้ค่า Recall มากที่สุด 87.7% ต่อมา เทคนิค Random Forest ให้ค่า Recall 83.3% เทคนิค Random Tree ให้ค่า Recall 80.2% เทคนิค REP Tree ให้ค่า Recall 78.5% ตามลำดับ

5. สรุป

ในการศึกษาและเปรียบเทียบแบบจำลองที่เหมาะสมสำหรับการพยากรณ์การสำเร็จการศึกษาในสถานศึกษาอาชีวศึกษาเอกชนจังหวัดกาฬสินธุ์ โดยแบบจำลองที่ใช้ในการเปรียบเทียบ 4 เทคนิคได้แก่ C4.5, Random Forest, Random Tree และ REP Tree ผลการทดลองพบว่าเทคนิค C4.5 มีความเหมาะสมมากที่สุดในการพยากรณ์การสำเร็จการศึกษาในสถานศึกษาอาชีวศึกษาเอกชนจังหวัดกาฬสินธุ์ โดยการวัดประสิทธิภาพการทำงานของแบบจำลองด้วยค่า Accuracy ได้ 87.73% ค่า Precision ได้ 87.5% และค่า Recall ได้ 87.7% เพราะ C4.5 สามารถจำแนกตัวแปรชนิด nominal ได้อย่างเหมาะสม

เอกสารอ้างอิง

[1] ข่าวการศึกษา. การเพิ่มสัดส่วนสายอาชีพเป็น 60% สายสามัญเหลือ 40%. Available: http://www.unigang.com/Article/5170,(2559, 28 สิงหาคม).

[2] วันฉัตร ทิพย์มาศ, "ปัจจัยที่มีผลต่อการตัดสินใจเข้าศึกษาต่อวิทยาลัยอาชีวศึกษาเอกชนของนักเรียนในเขตภาคใต้ตอนบน," *วารสารศึกษาศาสตร์,* vol. 24, กุมภาพันธ์-สิงหาคม 2556.

[3] เยาวภา ภารสำเร็จ, จิรัฏฐา ภูบุญอบและวิรัตน์ พงษ์ศิริ, "การเปรียบเทียบอัลกอริทึมเหมืองข้อมูลเพื่อวิเคราะห์ปัจจัยที่ส่งผลต่อระดับผลการเรียนของนักศึกษา," *วารสารวิทยาศาสตร์และเทคโนโลยี* มหาวิทยาลัยมหาสารคาม, 2556.

[4] ภรัณยา ปาลวิสุทธิ์ "การเพิ่มประสิทธิภาพเทคนิคต้นไม้ตัดสินใจบนชุดข้อมูลที่ไม่สมดุล โดยวิธีการสุ่มเพิ่มตัวอย่างกลุ่มน้อยสำหรับข้อมูล การเป็นโรคติดอินเทอร์เน็ต " *วารสารเทคโนโลยีสารสนเทศ,* vol. 12, pp. 54-63, 2559.

[5] S. Kalmegh, "Analysis of WEKA Data Mining Algorithm REPTree, Simple Cart and RandomTree for Classification of Indian News," *International Journal of Innovative Science, Engineering & Technology,* vol. 2, pp. 438-446, 2015.

[6] K. Wisaeng, "A Comparison of Decision Tree Algorithms For UCI Repository Classification," *International Journal of Engineering Trends and Technology,* vol. 4, pp. 3393-3397, 2013.

[7] J. Han and M. Kamber, *Data Mining Concepts and Techniques*, Second Edition ed.: Morgan Kaufmann, 2006.

[8] L. Breiman, "Random Forest," *Machine Learning,* vol. 45, pp. 5-32, 2001.

[9] *Weka*. Available: http://www.cs.waikato.ac.nz/ml/index.html,(2016, 3 December).

การพยากรณ์ธุรกรรมที่มีเหตุอันควรสงสัยโดยใช้โครงข่ายประสาทเทียม
The Suspicious Transactions Forecasting using Neural Network

นฤป รักงาม (Narroup Rak-ngam)[1] และณรงค์เดช กีรติพรานนท์ (Narongdech Keeratipranon)[2]
สาขาวิชาวิศวกรรมคอมพิวเตอร์ และโทรคมนาคม
วิทยาลัยนวัตกรรมด้านเทคโนโลยีและวิศวกรรมศาสตร์ มหาวิทยาลัยธุรกิจบัณฑิตย์
[1]thailandhsn@hotmail.com, [2] narongdech.ken@dpu.ac.th

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาโมเดลในการประเมินหาธุรกรรมที่มีเหตุอันควรสงสัยในการฟอกเงิน ซึ่งการฟอกเงินเป็นความผิดทางอาญาที่ร้ายแรงส่งผลกระทบรุนแรงต่อทั้งระบบเศรษฐกิจ และ สังคม อาชญากรฟอกเงินได้ทำการปลอมแปลงและปกปิดเงินที่ได้มาโดยผิดกฎหมาย ดังนั้น องค์กรธุรกิจ และหน่วยงานต่างๆ มีความพยายามอย่างยิ่งในการต่อต้านการทุจริตทางด้านการเงิน เพื่อยับยั้งและป้องกันไม่ให้องค์กรเข้าไปมีส่วนเกี่ยวข้องกับการกระทำที่มิชอบด้วยกฎหมายหรือฝ่าฝืนระเบียบข้อบังคับของทางการ แต่ปัญหาคือในแต่ละวันมีธุรกรรมจำนวนมาก และมีการใช้ดุลพินิจของเจ้าหน้าที่ว่าธุรกรรมใดเป็นธุรกรรมที่มีเหตุอันควรสงสัย ในงานวิจัยนี้นำเสนอ ระบบที่สามารถพยากรณ์ธุรกรรมที่มีเหตุอันควรสงสัย โดยพิจารณาจากธุรกรรมที่เกิดร่วมกับองค์ประกอบต่าง ๆ เช่น ลักษณะพฤติการณ์, ข้อมูลส่วนบุคคล ด้วยวิธีโครงข่ายประสาทเทียม โดยใช้ข้อมูลธุรกรรมการเงินที่เกิดขึ้นในระหว่างเดือนกันยายน – ตุลาคม 2559 ผลการทดลองพบว่าระบบให้ค่าความถูกต้องในการพยากรณ์อยู่ในระดับที่น่าพอใจประมาณ 90 %

คำสำคัญ: ฟอกเงิน, โครงข่ายประสาทเทียม, ธุรกรรมที่มีเหตุอันควรสงสัย

Abstract

The purpose of the study was presented the model for identify the transaction that could related to money laundering. Money laundering is the serious criminal case which has severe affect to the economy and social. The criminal counterfeit and manipulate the financial statement for their illegal money. Therefore business organization and other institutions put a lot of efforts to oppose the financial fraud. This could prevent the organization to involve in the act that against the law or violate the rules. The problem is how could we identified the suspicious transactions. In this research, we would like to present the process to forecast the suspicious transactions by considering the transactions with other components such as behavioral characteristics and personal information. We use the artificial neural network system to identify the suspicious transactions. We use the transactions log that happened between September - October 2016 and we found that the system could identify the suspicious transactions is highly effective at a good level.

Keyword: Money Laundering, Artificial Neural Network, Suspicious Transaction

1. บทนำ

ปัจจุบันปัญหาการฟอกเงินถือเป็นปัญหาที่มีความสำคัญกระทบต่อสังคมและประเทศชาติ ในช่วงปีพ.ศ.2542-2549 สำนักงานป้องกันและปราบปรามการฟอกเงิน (ปปง.) ของไทยได้เปิดเผยว่า มีการยึดทรัพย์สินที่เกี่ยวข้องกับความผิดฐานฟอกเงิน เป็นจำนวนถึง 4,124.8 ล้านบาท ในราชอาณาจักร (ประมาณ 117,850,000 เหรียญสหรัฐฯ ในเวลานั้น) [10] ซึ่งการกระทำผิดในรูปแบบต่างๆ เพิ่มขึ้นอย่างต่อเนื่อง โดยกลุ่มอาชญากรรมหรือผู้กระทำผิดมีพฤติกรรมที่หลีกเลี่ยงโดยพัฒนารูปแบบการทำธุรกรรมให้ซับซ้อนเพื่อทำให้การตรวจสอบเส้นทางการเงินเป็นไปได้ยากมากขึ้น

ซึ่งปัญหาของผู้ที่มีหน้าที่ตรวจสอบ คือ จำนวนธุรกรรมที่เกิดขึ้นมีจำนวนมาก ดังนั้นจึงต้องพัฒนาระบบเพื่อช่วยในการตรวจสอบวิเคราะห์ข้อมูล โดยระบบดังกล่าวนั้นมีประสิทธิภาพเหมือนตาข่ายที่ใช้กรองคัดแยกธุรกรรมที่น่าสงสัย กับ ธุรกรรมปกติ โดยเป้าหมายสำคัญในการวิจัย คือ การสร้างระบบคัดกรองธุรกรรมโดยเกณฑ์ที่ใช้ในการคัดกรองนั้น จะต้องมีประสิทธิภาพเหมาะสมกับกลุ่มประเภทลูกค้า และรูปแบบการให้บริการของแต่ละผู้ให้บริการ และสามารถรองรับธุรกรรมในอนาคตได้

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 การฟอกเงิน (Money Laundering)

ฟอกเงิน หมายถึง การนำเงินหรือทรัพย์สินที่ได้มาจากการกระทำความผิดหรือได้มาโดยไม่ชอบด้วยกฎหมาย มาเปลี่ยนสภาพให้เป็นเงินหรือทรัพย์สินที่ได้มาอย่างถูกต้อง ซึ่งจะมีผลกระทบต่อความน่าเชื่อถือ ทำให้นักลงทุนไม่มีความมั่นใจ เพราะฐานะเศรษฐกิจบิดเบือนไม่สะท้อนความเป็นจริง

ธุรกรรม หมายถึง กิจกรรมที่เกี่ยวกับการทำนิติกรรม สัญญา หรือการดำเนินการใดๆกับผู้อื่นทางการเงิน ทางธุรกิจ หรือการดำเนินการเกี่ยวกับทรัพย์สิน [1]

ธุรกรรมที่มีเหตุอันควรสงสัย หมายความว่า ธุรกรรมที่เกี่ยวข้องหรืออาจเกี่ยวข้องกับการกระทำผิด ไม่ว่าจะเป็นธุรกรรมเพียงครั้งเดียวหรือหลายครั้ง

ประเด็นที่สำคัญคือ การพิจารณาประการหนึ่งคือ ความสงสัยหรือไม่สงสัย ในบางกรณีอาจเป็นเรื่องซับซ้อน ขึ้นอยู่กับสภาวะหรือพฤติการณ์แวดล้อม แล้วยังเป็นเรื่องของความคิดและประสบการณ์ในการทำงาน [5]

แนวทางการพิจารณารูปแบบของธุรกรรมที่มีเหตุอันควรสงสัย

1. ธุรกรรมที่มีลักษณะซับซ้อนผิดปกติ เมื่อเทียบกับการทำธุรกรรมประเภทนั้นๆ

2. มูลค่าในการทำธุรกรรมที่เกิดขึ้นไม่สอดคล้องกับข้อมูลรายได้หรือฐานะทางเศรษฐกิจของลูกค้า

3. การทำธุรกรรมที่มี จำนวนเงินใกล้เคียงกับขั้นต่ำที่กฎหมายกำหนด เพื่อหลีกเลี่ยงการถูกรายงานตามกฎหมาย

4. ธุรกรรมที่เกี่ยวข้องกับการกระทำผิดมูลฐาน

ด้วยการทำธุรกรรมที่มีเหตุอันควรสงสัย ซึ่งจะต้องรายงานสำนักป้องกันและปราบปรามการฟอกเงินหรือไม่นั้น ควรพิจารณาถึงข้อเท็จจริงและพฤติการณ์แวดล้อมองค์ประกอบต่างๆ ด้วยเหตุนี้นักวิจัยจึงศึกษาปัจจัยที่มีผลต่อการวิเคราะห์ธุรกรรมโดยอาศัยเทคนิคการจำแนกข้อมูล

2.2 การจำแนกประเภทข้อมูล

การจำแนกประเภทข้อมูล [2] เป็นกระบวนการสร้างตัวแบบโดยการจัดการข้อมูลให้อยู่ในกลุ่มที่กำหนด เพื่อแสดงให้เห็นความแตกต่างระหว่างคลาส (Class) และเพื่อทำนายว่าข้อมูลหนึ่งๆ ควรจัดอยู่ในคลาสใด ซึ่งในหลายงานวิจัยได้ใช้เทคนิคการจำแนกข้อมูลเพื่อใช้ในการพยากรณ์ [2,4] เป็นต้น โดยในงานวิจัยนี้ใช้โครงข่ายประสาทเทียม ซึ่งอ้างอิงจากงานวิจัย [2,3] โดยเปรียบเทียบผลลัพธ์ที่ได้จะมีความถูกต้องมากกว่าวิธีการอื่น

2.2.1 Neural Network

โครงข่ายประสาทเทียม (Artificial Neural Network) เป็นโมเดลทางคณิตศาสตร์ แนวคิดเริ่มต้นมาจากหลักการทำงานของระบบประสาททางชีววิทยาและเป็นวิทยาการที่พยายามเลียนแบบการทำงานของสมองมนุษย์เพื่อใช้ในการพยากรณ์เหตุการณ์จากข้อมูลที่มีอยู่ โครงข่ายประสาทเทียมประกอบด้วย 2 ขั้นตอน คือ

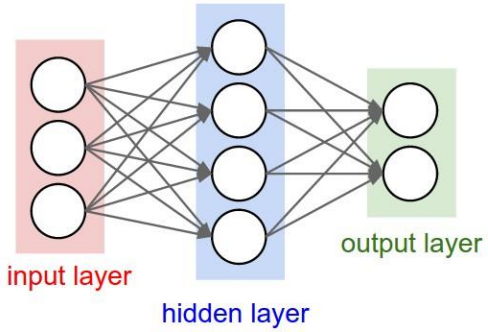
1. ขั้นตอนของกระบวนการเรียนรู้ จะมีการใช้ชุดข้อมูลตัวอย่าง (Training Data) สำหรับให้โครงข่ายประสาทเทียมได้เรียนรู้

2. ขั้นตอนของกระบวนการทดสอบ จะมีการใช้ชุดข้อมูลทดสอบ (Testing data) สำหรับทดสอบโครงข่ายประสาทเทียม ว่าผลลัพธ์การเรียนรู้อยู่ในระดับใด

โครงสร้างของโครงข่ายประสาทเทียมจะประกอบด้วยชั้นของข้อมูลเบื้องต้น 3 ชั้น ได้แก่ ชั้นอินพุต (Input layer) ชั้นซ่อนเร้น (Hidden layer) และชั้นแสดงผล (Output layer) ดังภาพที่ 1

ซึ่งการเรียนรู้ของโครงข่ายเป็นการเรียนรู้แบบป้อนไปข้างหน้า (Feed Forward Neural Networks) โดยที่ข้อมูลที่จะนำมาเรียนรู้จะถูกส่งเข้ามาที่ชั้นรับข้อมูล ซึ่งทุกๆหน่วยของชั้นรับข้อมูลจะทำการกระจายข้อมูลไปยังทุกๆหน่วยของชั้นซ่อน และทำการคำนวณเพื่อให้ได้ผลลัพธ์ ที่ชั้นผลลัพธ์ จากนั้นจะ

ตรวจสอบว่าผลลัพธ์ที่คำนวณได้มีค่าใกล้เคียงกับเป้าหมายที่กำหนดไว้หรือไม่ หากมีค่าไม่ใกล้เคียงกันก็จะเป็นส่วนที่ผิดพลาด ซึ่งจะต้องนำค่าผิดพลาดมาใช้ในการปรับการเรียนรู้ โดยการเรียนรู้แบบแพร่ย้อนกลับ (Back Propagation) นั้นเป็นการแพร่ความผิดพลาดจากชั้นผลลัพธ์ไปที่ชั้นซ่อน และทำการปรับค่าน้ำหนัก จากนั้นทำการแพร่กลับไปที่ชั้นรับข้อมูลเพื่อทำการปรับค่าน้ำหนักเช่นกัน การทำงานนี้จะทำซ้ำไปเรื่อยๆ จนกระทั่งได้ผลลัพธ์ใกล้เคียงกับผลลัพธ์ที่ตั้งเป้าหมายไว้ก็จะหยุดการเรียนรู้ [6]



ภาพที่ 1 โครงสร้างของโครงข่ายประสาทเทียม

2.3 การศึกษางานวิจัยที่เกี่ยวข้อง

มีงานวิจัยที่เกี่ยวข้องกับการพยากรณ์ธุรกรรมที่ผิดปกติ ในการเดินบัญชีรูปแบบต่าง ๆ โดยใช้ปัญญาประดิษฐ์ ในงานวิจัยของ J.A Adeyiga และคณะ [7] ได้ทดลองการค้นหาธุรกรรมที่ผิดปกติของ e-banking โดยประยุกต์ใช้โครงข่ายประสาทเทียม ซึ่งผลลัพธ์ที่ได้จากการศึกษาพบว่าโครงข่ายประสาทเทียมสามารถประยุกต์ใช้ในการตรวจสอบธุรกรรมที่ผิดปกติได้ดี เพราะมีความสามารถในการเรียนรู้รูปแบบ จากนั้นใช้ความรู้ที่เรียน ประเมินธุรกรรมใหม่ที่เข้ามาว่าผิดปกติหรือไม่ ซึ่งสอดคล้องกับงานวิจัยของ Raghavendra Patidar และ Lokesh Sharma [8] ได้ประยุกต์ใช้เทคนิคโครงข่ายประสาทเทียม ในการตรวจสอบธุรกรรมทุจริตของบัตรเครดิตซึ่งผลลัพธ์ที่ได้น่าพอใจ ผลการวิเคราะห์งานวิจัย 2 งานนี้ ได้ผลสรุปว่าโครงข่ายประสาทเทียมมีความสามารถในการพยากรณ์และได้ผลลัพธ์ที่แม่นยำเป็นที่น่าพอใจ ได้ศึกษางานวิจัยของ Jun Tang และ Jian Yin [9] ได้ทำการศึกษาการปรับปรุงระบบการคัดกรองธุรกรรมที่ต้องสงสัยในการฟอกเงินโดยใช้เทคนิค support vector machine (SVM) ในการปรับปรุงประสิทธิ์ภาพกฎเกณฑ์การคัดกรองธุรกรรมให้ดีขึ้น โดยผลงานวิจัยนี้กล่าวถึงโมเดลที่ได้พัฒนาขึ้นนั้นสามารถเรียนรู้รูปแบบธุรกรรมที่น่าสงสัยใหม่ ๆ สามารถลดค่า False Positive Rate ได้ดี

3. วิธีการดำเนินงานวิจัย

ข้อมูลในงานวิจัยนี้เป็นข้อมูลธุรกรรม ที่อยู่ในช่วงเดือนกันยายน ถึง เดือนตุลาคม ปี 2559 โดยผ่านการกรองข้อมูลด้าน Financial เบื้องต้น

3.1 การเก็บข้อมูล

ในการศึกษานี้ได้ใช้ข้อมูลการเดินบัญชีที่เกิดขึ้นระหว่างเดือน กันยายน-ตุลาคม 2559 โดยกรองเฉพาะธุรกรรมที่เข้าเงื่อนไข 2 ตัว ที่ได้อธิบายในตารางที่ 1

ตารางที่ 1 : เงื่อนไขในการกรองธุรกรรมที่ใช้ในงานวิจัย

	เงื่อนไข	จำนวนเงิน	วัน
1	มีเงินเข้า และ ออกในทันที โดยผลต่างไม่เกิน 5 %	1 แสน	ภายใน 1 วัน
2	มีการทำธุรกรรม เงินสด เข้าหรือออก โดยมียอดเงินใกล้เคียงกับรายงานธุรกรรม	1.80 – 1.99 ล้านบาท	ภายใน 1 วัน

ข้อมูลธุรกรรมที่ติดเงื่อนไข มีจำนวน 894 รายการ (300 บัญชี) ในขั้นตอนปกติ Investigator ต้องทำการวิเคราะห์ธุรกรรม โดยดูความสอดคล้องจากข้อมูลด้าน Financial ประกอบกับข้อมูล Non-Financial แต่ในงานวิจัยนี้จะแบ่งข้อมูลออกเป็น 2 ชุด คือ ข้อมูลสำหรับเรียนรู้ 268 รายการ (100 บัญชี) และ ข้อมูลสำหรับทดสอบ มี 625 รายการ (200 บัญชี) เพื่อใช้ทดสอบระบบงาน

ตารางที่ 2 : ข้อมูลชุดแรกด้าน Financial

Field	Detail
Total Deposit	ยอดรวมในการฝากของบัญชี/เดือน
Total Withdrawal	ยอดรวมในการถอนของบัญชี/เดือน
ความถี่ในการฝาก	จำนวนครั้งในการฝาก/เดือน
ความถี่ในการถอน	จำนวนครั้งในการถอน/เดือน

ตารางที่ 3 : ข้อมูลรายละเอียด Non-Financial

Field	Detail
Account Type	ประเภทบัญชี
Occupation	ลักษณะอาชีพ
Salary	รายได้เฉลี่ยต่อเดือน
Watch List	มีชื่อในฐานข้อมูลที่เฝ้าระวัง
Nationality	สัญชาติ/แหล่งที่มาของรายได้
Behavior Diagnosing	พฤติกรรมในการใช้จ่าย (ค่า Standard Deviation)
Age	ช่วงอายุ

คุณลักษณะของตัวแปรแต่ละชนิด

Account Type ในงานวิจัยใช้การประเมินความเสี่ยงของแต่ละธุรกรรมที่ให้บริการของบัญชีแต่ละประเภท โดยแนวทางในการประเมินความเสี่ยง เช่น การทำธุรกรรมแบบ Face to Face , บัญชีมีวัตถุประสงค์ที่ชัดเจน การโอนเปลี่ยนมือ หรือ มีจำนวนเงิน/จำนวนครั้งในการใช้บริการ

ตารางที่ 4 : ความเสี่ยงด้านการฟอกเงินในแต่ละผลิตภัณฑ์

ประเภทบริการ	ระดับความเสี่ยง
เงินออมทรัพย์/กระแสรายวัน	สูง
ตราสารทางการเงิน / เงินฝากประจำ / บัตรเครดิต	กลาง
สินเชื่อ / ประกันชีวิต	ต่ำ

Occupation ในงานวิจัยได้แบ่งระดับความเสี่ยงของลักษณะอาชีพตามประกาศสำนักงาน ปปง. เรื่อง แนวทางการกำหนดปัจจัยเพื่อพิจารณาความเสี่ยงของลูกค้าเกี่ยวกับอาชีพ โดยแบ่งความเสี่ยงเกี่ยวกับอาชีพออกเป็น 3 ระดับดังนี้

ตารางที่ 5 : ความเสี่ยงอาชีพ

กลุ่ม	รายละเอียด
1	นักการเมือง / ข้าราชการผู้บริหารระดับสูง (ผู้มีสถานะภาพทางการเมือง)
2	9 อาชีพที่เฝ้าระวัง ตามประกาศ ปปง.
3	อาชีพอื่น ๆ

Nationality ในงานวิจัยได้กำหนดให้พื้นที่หรือประเทศตามมาตรการทางการเงินเกี่ยวกับการฟอกเงิน (Financial Action Task Force: FATF) ที่กำหนดให้เป็นพื้นที่หรือประเทศที่ไม่มีการใช้หรือประยุกต์ใช้มาตรฐานสากลด้านการป้องกันปราบปรามการฟอกเงินและการต่อต้านการสนับสนุนทางการเงินแก่การก่อการร้ายอย่างเพียงพอ เป็นพื้นที่หรือประเทศที่มีความเสี่ยงสูง

ตารางที่ 6 : ความเสี่ยงด้านพื้นที่เสี่ยง

กลุ่ม	รายละเอียด
1	ประเทศความเสี่ยงสูง ตามกำหนด FATF
2	ประเทศไทย
3	ประเทศอื่น ๆ ที่ไม่อยู่ในกลุ่ม 1 และ 2

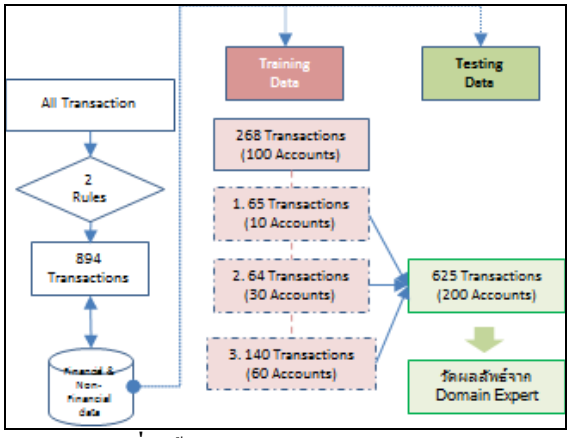
Age เพราะแผนการออมการลงทุนและการใช้จ่ายเงินในแต่ละช่วงอายุ จะมีรูปแบบที่แตกต่างกัน ในงานวิจัยนี้จึงได้แบ่งช่วงอายุออกเป็น 5 ช่วงอายุ ให้เหมาะสมกับรายได้ และ การใช้จ่าย

ตารางที่ 7 : การแบ่งช่วงอายุของบุคคล

กลุ่ม	รายละเอียด
1	วัยก่อนทำงาน (0 – 21 ปี)
2	ช่วงวัยเริ่มทำงาน (21- 30 ปี)
3	วัยเริ่มสร้างครอบครัว (31- 40 ปี)
4	วัยการงานมั่นคง (41- 55 ปี)
5	วัยเกษียณอายุ (55 ปีขึ้นไป)

3.2 วิธีการทดลอง

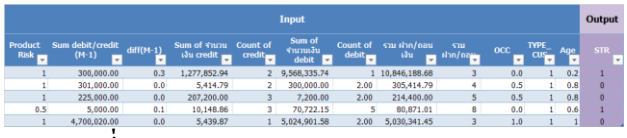
ขั้นตอนการพยากรณ์ด้วยเทคนิคโครงข่ายประสาทเทียม ในงานวิจัยนี้ โดยแบ่งข้อมูลของชุดเรียนรู้ออกเป็น 3 ชุด จาก 100 บัญชี เป็น 10 , 30 และ 60 ตามลำดับ ตามภาพที่ 2



ภาพที่ 2 ขั้นตอนการฝึกด้วยโครงข่ายประสาทเทียม

3.2.1 ขั้นตอนการทดลอง

ในงานวิจัยนี้ได้พัฒนาแบบจำลองของโครงข่ายประสาทเทียมโดยใช้ซอฟต์แวร์ Matlab โดยทำการฝึกด้วยข้อมูลที่เตรียมไว้สำหรับการเรียนรู้ ดังภาพที่ 3 แสดง Input File ที่ใช้สำหรับการ Training



ภาพที่ 3 แสดง Input File (Training.xls) บนโปรแกรม Excel

โดยข้อมูลจะมี input อยู่ทั้งหมด 12 node โดยมี output 1 node สำหรับจำนวน hidden Layer นั้นได้ใช้วิธี Trial & Error ทดลองเพื่อหาค่า Performance ที่ดีที่สุด ดัง ตารางที่ 8

ตารางที่ 8 : ผลการทดลอง ปรับค่า Hidden Layer

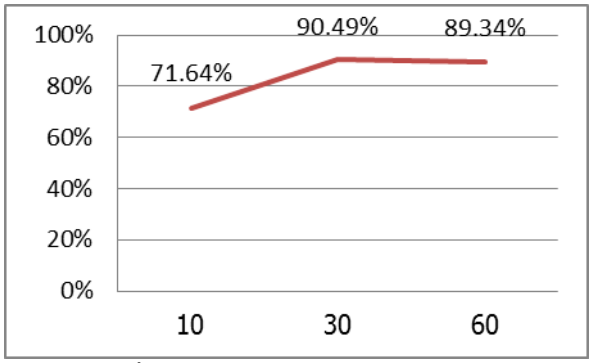
จำนวน Hidden Layer	*Best Validation Performance*
1	0.094302
5	0.22213
10	0.0007789
15	0.11735

จากตารางที่ 8 จะเห็นได้ว่า เมื่อใช้ hidden layer 10 node จะได้ค่า performance น้อยที่สุด ซึ่งเข้าใกล้ค่า goal มากที่สุด (ความผิดพลาดน้อยที่สุด) ซึ่งถูกเลือกมาใช้ในงานวิจัยนี้

เมื่อทำการฝึกเสร็จแล้วจึงนำโมเดลที่ฝึกมาทำการพยากรณ์ด้วยข้อมูลที่เตรียมไว้ จากนั้นนำผลการพยากรณ์มาเปรียบเทียบกับผลลัพธ์ที่ให้ผู้เชี่ยวชาญวิเคราะห์

4. ผลลัพธ์และการอภิปรายผล

การประเมินผลการวัดประสิทธิภาพของงานวิจัยนี้ โดยการนำเข้าข้อมูลทดสอบ ที่มีผลลัพธ์ที่ผ่านการวิเคราะห์จาก Domain Experts แล้ว จากกราฟในภาพที่ 4 แสดงให้เห็นว่าข้อมูลที่ใช้ในการเรียนรู้ 65 รายการ (10 บัญชี) จะมีความถูกต้องอยู่ที่ 71.64 % และ ถ้าปริมาณข้อมูลที่ใช้ในการเรียน 64 รายการ (30 บัญชี) และ 140 รายการ (60 บัญชี) ความแม่นยำจะเพิ่มขึ้นเป็น 90.49% และ 89.34 % ตามลำดับ ซึ่งผลการพยากรณ์เป็นที่น่าพอใจ



ภาพที่ 4 ผลการวิเคราะห์ผ่านซอฟแวร์ MATLAB

ตรวจสอบผลพบว่าส่วนมากที่ผิด เป็นมีลักษณะที่ ระบบพยากรณ์ว่าเป็น suspicious transactions เนื่องจากเป็นรูปแบบการเดินเงินที่ผิดปกติ แต่เป็นการโอนเงินไปมาระหว่างบัญชีของตัวเองในธนาคารเดียวกัน เพื่อจุดประสงค์การบริหารจัดการ หรือ เหตุผลอื่น ๆ ดังนั้น Domain Experts จึงวิเคราะห์ว่า ธุรกรรมดังกล่าว ไม่เป็นธุรกรรมที่มีเหตุอันควรสงสัย

5. สรุป

ในงานวิจัยนี้มีวัตถุประสงค์เพื่อนำเสนอระบบที่ช่วยวิเคราะห์ธุรกรรมที่น่าสงสัย โดยหลักการของโครงข่ายประสาทเทียม จากผลที่ได้ทำให้ทราบว่าถ้ามีจำนวนข้อมูลที่ใช้ในการเรียนรู้ปริมาณมาก จะส่งผลให้เปอร์เซ็นต์ ความถูกต้องสูงขึ้น แต่ในระดับหนึ่งนอกเหนือจากปริมาณข้อมูลที่ใช้ในการ

เรียนรู้แล้วนั้นคุณภาพของข้อมูลที่ใช้ในการจำแนกก็ส่งผลต่อการจำแนกข้อมูล

ซึ่งในครั้งต่อไป ผู้วิจัยอาจเพิ่มขั้นตอนการศึกษาคุณลักษณะของข้อมูลก่อน ว่าได้ข้อมูลตัวแปรที่มีคุณภาพมีความสามารถในการจำแนกข้อมูลและอาจจะเพิ่มตัวแปรที่มีผลต่อการวิเคราะห์ เช่น ข้อมูล CDD (customer due diligence) , Associated Account หรือ ช่องทางการใช้บริการ/สาขา เป็นต้น มาเป็นปัจจัยที่ช่วยในการวิเคราะห์ เพื่อนำไปพยากรณ์ข้อมูลในมิติต่าง ๆ และได้ผลลัพธ์ที่ดีขึ้น

เอกสารอ้างอิง

[1] พระราชบัญญัติป้องกันและปราบปรามการฟอกเงิน พ.ศ. 2542.

[2] นิภาพร ชนะมาร กนกวรรณ อรุณรัชฎารมย์ และ พรรณี สิทธิเดช, “การทำนายภาวะทารกขาดออกซิเจนระหว่างคลอดโดยใช้เทคนิคการทำเหมืองข้อมูล”, *The 9th National Conference on Computing and Information Technology,* 2013.

[3] สุภโชค เรืองศรี, “การเปรียบเทียบประสิทธิภาพระหว่างต้นไม้ตัดสินใจและโครงข่ายประสาทเทียมสำหรับคาดการณ์ช่วงปริมาณการใช้กระดาษภายในแผนกฝ่ายการผู้โดยสาร”, *The 9th National Conference on Computing and Information Technology,* 2013.

[4] วรรณวิทย์ คล้ายบุญส่ง และ สมชาย ปราการเจริญ, “การพยากรณ์กำลังคนที่ใช้ในการบำรุงรักษาซอฟต์แวร์ประเภท Web Based ด้วย Neural Network” , *The 5th National Conference on Computing and Information Technology,* 2009.

[5] นิพันธ์ อักษรกาญจน์, “ปัญหาธุรกรรมที่มีเหตุอันควรสงสัยตามพระราชบัญญัติป้องกันและปราบปรามการฟอกเงิน พ.ศ.2542” , เอกสารวิชาการ,หลักสูตร ผู้พิพากษาผู้บริหารในศาลขั้นต้น ,2012.

[6] กนิษฐา ตั้งไทยวัญ และ ณรงค์เดช กีรติพรานนท์, “การจำแนกข้อมูลพื้นที่ภาพถ่ายดาวเทียมด้วยโครงข่ายประสาทเทียมแบบเพอร์เซฟตรอนหลายชั้น”, *The 38th Electrical Engineering Conference*, 2015.

[7] J.A Adeyiga, J.O. Ezike, A. Omotosho and W. Amakulor, “A neural network based model for detecting irregularities in e-Banking transactions”, *African Journal of Computer and ICTs* , 2011.

[8] Raghavendra Patidar and Lokesh Sharma, “ Credit Card Fraud Detection Using Neural Network” , *International Journal of Soft Computing and Engineering , vol. 1,* 2011.

[9] Jun Tang and Jian Yin, “ Developing an intelligent data discriminating system of anti-money laundering based on SVM” , *International Conference on Machine Learning and Cybernetics, vol. 9,* 2005.

[10] ชมเกตุ งามไกวัล พรรณชฎา ศิริวรรณบุศย์ ฐนันดร์ศักดิ์ บวรนันทกุล และ อุนิษา เลิศโตมรสกุล, “ปัญหาการฟอกเงินในธนาคารต่างชาติในประเทศไทยที่ส่งผลกระทบต่อความเชื่อมั่นของนักลงทุน”, *Journal of Thai Justice System , Vol.1* , 2016.

[11] ประกาศสำนักงานป้องกันและปราบปรามการฟอกเงิน “ แนวทางในการกำหนดปัจจัยเพื่อพิจารณาความเสี่ยงของลูกค้าเกี่ยวกับอาชีพ” , 2017

วิธีการแบบรวมกลุ่มเพื่อจำแนกประเภทข้อความไม่สมดุลสูงด้วยเทคนิคการเรียนรู้เชิงลึกร่วมกับการคัดเลือกคุณลักษณะและการปรับสมดุลข้อมูล

Ensemble for Highly Imbalanced Text Classification base on Deep Learning with Feature Selection and Data Balancing Techniques

ศุภเทพ สติมั่น (Supathep Satiman)[1] พยุง มีสัจ (Phayung Meesad) [2] และปิยนุช วรบุตร (Piyanoot Voraboot)[3]

[1,3]สาขาวิชาคอมพิวเตอร์ธุรกิจ คณะบริหารธุรกิจและการจัดการ มหาวิทยาลัยราชภัฏอุบลราชธานี

[2]ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

supathep.s@ubru.ac.th, phayung.m@it.kmutnb.ac.th, piyanoot.v@ubru.ac.th

บทคัดย่อ

ข้อมูลไม่สมดุลยังเป็นปัญหาสำคัญที่ส่งผลกระทบต่อประสิทธิภาพในเรียนรู้ด้วยขั้นตอนวิธีมาตรฐานดั้งเดิม โดยเฉพาะอย่างยิ่งข้อมูลที่อยู่ในรูปแบบของข้อความ มีลักษณะเป็นข้อมูลที่มีมิติสูง ยิ่งส่งผลให้ความท้าทายในการจำแนกประเภทข้อความมากขึ้น งานวิจัยนี้มุ่งเน้นการแก้ปัญหาการจำแนกประเภทข้อความที่ไม่สมดุลสูง 3 ส่วนหลัก คือ 1) การแก้ไขที่ระดับข้อมูล 2) การแก้ปัญหาด้วยการคัดเลือกคุณลักษณะข้อมูล และ 3) การแก้ปัญหาที่ระดับขั้นตอนวิธี โดยนำเสนอวิธีการแบบรวมกลุ่มซึ่งรวมเทคนิคการเรียนรู้เชิงลึกร่วมกับเทคนิคการคัดเลือกคุณลักษณะ โดยการกรองคุณลักษณะด้วยค่าน้ำหนักอัตราส่วนมาตรฐานเกน และการปรับสมดุลข้อมูลโดยการสุ่มลดข้อมูลในคลาสหลักด้วยวิธีบูทสแทรป เพื่อจัดการปัญหาข้อมูลที่อยู่ในรูปแบบของข้อความที่ไม่สมดุลสูง ซึ่งผลการวิจัยแสดงให้เห็นว่า วิธีการที่นำเสนอสามารถจัดการปัญหาและให้ประสิทธิภาพในการจำแนกประเภทดีกว่าขั้นตอนวิธีมาตรฐานโดยทั่วไป

คำสำคัญ: การประมวลผลข้อความ ข้อมูลไม่สมดุลสูง
การจำแนกประเภทข้อความ การเรียนรู้เชิงลึก

Abstract

Imbalanced data set has caused an essential problem of the classification performance to learning by traditional classifier algorithm. Especially the data in text form is high dimensional data, which is a great challenge for text classification. The paper focuses on the study three main aspects to solve this problem: 1) data-level, 2) feature selection and 3) algorithmic-level. Thus, we propose a novel ensemble method, which combines deep learning with feature selection and data balancing. For feature selection, information gain ratio based weighted is used to filter attributes. Then, under-sampling based on bootstrapping techniques is performed in process of data balancing. The experimental results show that the proposed method is better than traditional classifier algorithms. It can handle classification problems by solving the highly imbalanced on text data.

Keyword: Text Processing, Highly Imbalanced Data
Text Classification, Deep Learning

1. บทนำ

ข้อมูลไม่สมดุล (Imbalanced Data) เป็นลักษณะของข้อมูลที่เกิดขึ้นโดยธรรมชาติ คือ มีการกระจายตัวของข้อมูลในแต่ละคลาสไม่เท่ากันหรือการกระจายแบบเบ้ (Skewed Distribution) โดยประกอบไปด้วย กลุ่มข้อมูลที่มีจำนวนมากหรือคลาสหลัก (Majority) และกลุ่มข้อมูลที่มีจำนวนน้อย หรือคลาสรอง (Minority) [1] โดยลักษณะข้อมูลไม่สมดุลนี้ส่งผลกระทบต่อประสิทธิภาพการเรียนรู้ของเครื่อง (Machine Learning) เมื่อนำมาเรียนรู้และจำแนกประเภทข้อมูล ทำให้การตัดสินใจเกิดความเอนเอียง โดยให้ความสำคัญกับคลาสหลักมากกว่าคลาสรอง [2] นักวิจัยได้นำเสนอการแก้ปัญหาความไม่สมดุลข้อมูลไว้ทั้งในระดับข้อมูล (Data Level) และระดับขั้นตอนวิธี (Algorithmic Level) โดยการแก้ปัญหาดังกล่าวยังพบข้อเสีย เช่น การเรียนรู้ข้อมูลที่ไม่ครอบคลุมกับข้อมูลจริง

หรือเกิดการสูญหายของข้อมูลที่มีคุณลักษณะสำคัญ ส่วนการแก้ปัญหาในระดับขั้นตอนวิธีนั้น เป็นการปรับขั้นตอนวิธีให้มีความยืนหยุ่นและเหมาะสมกับข้อมูลที่ไม่สมดุล แต่การแก้ปัญหาในระดับนี้ ยังจำกัดเฉพาะการเรียนรู้กับชุดข้อมูลที่ถูกจัดเก็บแบบมีโครงสร้างเป็นส่วนใหญ่ ทำให้ข้อมูลบางลักษณะ เช่น ชุดข้อมูลที่ไม่มีโครงสร้าง (Unstructured Data) หรือชุดข้อมูลที่มีมิติสูง (High Dimensional Data) นั้น อาจส่งผลต่อการลดทอนประสิทธิภาพในการเรียนรู้ [3] ซึ่งหนึ่งในเทคนิคการเรียนรู้ที่กำลังได้รับความสนใจและถูกนำมาใช้ในการเรียนรู้ข้อมูลในลักษณะนี้ คือการเรียนรู้เชิงลึก

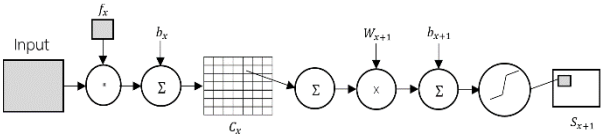
การเรียนรู้เชิงลึก (Deep Learning) เป็นเทคนิคที่ได้รับความสนใจในวงกว้าง เป็นเทคนิคการเรียนรู้แบบมีผู้สอน (Supervise Learning) ที่สามารถเรียนรู้และจดจำรูปแบบข้อมูลได้อย่างมีประสิทธิภาพ ในช่วงที่ผ่านมานักวิจัยได้นำการเรียนรู้เชิงลึกมาประยุกต์ใช้ทั้งในงานด้านการทำเหมืองข้อความทางการแพทย์และชีววิทยา [4] การวิเคราะห์อารมณ์จากข้อความ [5] และการนำเทคนิคการเรียนรู้เชิงลึกมาแก้ปัญหาข้อมูลไม่สมดุล [6] [7] ซึ่งจากผลจากการวิจัยพบว่า เทคนิคการเรียนรู้เชิงลึกสามารถเพิ่มประสิทธิภาพการเรียนรู้ข้อมูลที่มีความซับซ้อน มีมิติข้อมูลสูง และมีจำนวนมหาศาล ได้ดีกว่าขั้นตอนวิธีมาตรฐานโดยทั่วไป

จากปัญหาที่กล่าวมาข้างต้น งานวิจัยนี้จึงนำเสนอการเรียนรู้แบบใหม่โดยใช้วิธีการแบบรวมกลุ่มด้วยเทคนิคการเรียนรู้เชิงลึก ร่วมกับการลดมิติข้อมูลและการปรับความสมดุลข้อมูล เพื่อค้นหาขั้นตอนวิธีที่เหมาะสมกับลักษณะข้อมูล และแก้ปัญหาการการเรียนรู้ให้มีประสิทธิภาพที่ดีขึ้น

2. วรรณกรรมที่เกี่ยวข้อง

2.1 เทคนิคการเรียนรู้เชิงลึก (Deep Learning techniques)

การเรียนรู้เชิงลึกเป็นเทคนิคในกลุ่มโครงข่ายประสาทเทียม (Artificial Neural Network) ที่มีโครงสร้างขนาดใหญ่ ประกอบไปด้วยนิวรอนและชั้นซ่อนจำนวนมาก โดยนำหลักการแนวคิดการทำงานเลียนแบบการทำงานของสมองมนุษย์ที่มีความซับซ้อน สามารถเรียนรู้และจดจำรูปแบบข้อมูล (Pattern Recognition) [8] ได้อย่างมีประสิทธิภาพ โดยจุดเริ่มต้นกรอบแนวคิดการเรียนรู้เชิงลึก (Deep Learning Framework) คือ โครงข่ายประสาทเทียมแบบคอนโวลูชันนัล (Convolutional Neural Networks: CNNs) มีหลักการในการลดจำนวนข้อมูลก่อนการนำไปประมวลผล ดังนั้นเทคนิคจึงถูกนำมาใช้กับลักษณะข้อมูลที่มีมิติสูง มีความซับซ้อนและมีความใกล้เคียงกัน หรือข้อมูล 2 มิติ (two-dimensional data) เช่น ภาพนิ่ง หรือภาพเคลื่อนไหว โดย CNNs มีแนวคิดหลัก 2 ประการ คือ ลดการเชื่อมต่อระหว่างชั้นซ่อนและอินพุต โดยการเชื่อมต่อด้วยเซตย่อยขนาดเล็กเท่านั้น หรือที่เรียกว่า feature map และแนวคิดอีกประการ คือ การลดความซับซ้อนในการคำนวณภาพ โดยที่คุณสมบัติของภาพไม่เปลี่ยนแปลง [9] การทำงานของโครงข่ายประสาทเทียมแบบคอนโวลูชันนัล แสดงดังภาพที่ 1



ภาพที่ 1 : การบวนการเลือกกลุ่มข้อมูลตัวอย่างย่อยของโครงข่ายประสาทเทียมแบบคอนโวลูชันนัล

กระบวนการเริ่มต้นของ CNNs ประกอบไปด้วย การนำข้อมูลนำเข้า (Input) มารวมกัน ด้วยการสอนร่วมกับตัวกรอง)Filter (f_x และการเพิ่ม Bias b_x ในการสร้าง Convolution layer C_x โดย Sub-sampling จะประกอบไปด้วยค่าผลรวม)Summing) ของบริเวณใกล้เคียง (Neighborhoodในกรณี) (รูปภาพ เพื่อนบ้านคือ4 pixels(จากนั้นให้ค่าถ่วงน้ำหนักด้วย scalar W_{x+1} และเพิ่ม bias b_{x+1} และนำไปคำนวณผ่าน Sigmoid Function เพื่อนำไปสร้าง small feature map S_{x+1}

2.2 การปรับสมดุลข้อมูล (Data Balancing)

การปรับสมดุลข้อมูลด้วยการลดข้อมูลกลุ่มตัวอย่างในคลาสหลัก เป็นอีกหนึ่งวิธีที่นิยมนำมาใช้ในการแก้ปัญหาข้อมูลไม่สมดุล เนื่องจากสามารถลดทรัพยากรและเวลาที่ใช้ในการประมวลผลได้ ซึ่งการลดตัวอย่างข้อมูลในคลาสหลักนั้น ได้มีนักวิจัยนำเสนอผลงานไว้หลายวิธี เช่น วิธีลดข้อมูลด้วยกฎ Condensed Nearest Neighbor (CNN) [10] วิธี Tomek Link [11] วิธี Neighborhood Cleaning Rule (NCR) [12] การสุ่ม (Random Under Sampling : RUS) [13] เป็นต้น ในงานวิจัยนี้ใช้วิธีการลดข้อมูลโดยการสุ่มตัวอย่างข้อมูลด้วย

วิธีบูทสแทรป (Bootstrapping method) ซึ่งเป็นฟังก์ชันทางสถิติ โดยการแจกแจงจากการกระจายตัวของข้อมูล และทำการอนุมานเกี่ยวกับตัวอย่างจากข้อมูลที่มีการสุ่มตัวอย่างใหม่ โดยการหาค่าประมาณแบบจุดสามารถคำนวณได้จากสมการที่ (1)

(1)

$\frac{1}{B}$ให้ θ เป็นพารามิเตอร์ที่ต้องการประมาณในประชากร และให้ $\hat{\theta}_B$เป็นค่าประมาณของพารามิเตอร์ θ ด้วยวิธีบูทสแทรป ที่คำนวณจากข้อมูลตัวอย่างขนาด n โดยในการประมาณนั้นจะทำการคำนวณค่าประมาณของ θ จากชุดของตัวอย่าง ซึ่งจะได้ค่าประมาณคือ $\hat{\theta}_1^*$ เมื่อคำนวณครบทั้งหมดจำนวน B ครั้ง จะได้ค่าประมาณทั้งหมดของจำนวน B ตัว คือ $\hat{\theta}_1^*,\hat{\theta}_2^*,\ldots,\hat{\theta}_B^*$ จากนั้นนำมาสร้างฮิสโตแกรม (Histogram) โดยแต่ละตัวจะถูกกำหนดให้มีค่าความน่าจะเป็นเท่ากันคือ และการหาค่าประมาณแบบช่วงของพารามิเตอร์ θ ที่ระดับนัยสำคัญ α ดังสมการ (2)

$$P\left(\hat{\theta}_{BL} < \theta < \hat{\theta}_{BU}\right) = 1-\alpha \quad (2)$$

โดยค่า $\hat{\theta}_{BL}$ และ $\hat{\theta}_{BU}$ เกิดจากการคำนวณจากการแจกแจงค่าสถิติ $\hat{\theta}_1^*,\hat{\theta}_2^*,\ldots,\hat{\theta}_B^*$ ที่ได้ นำมาจัดเรียงลำดับค่าจากน้อยไปหาค่ามาก จากนั้นคำนวณหาค่าที่ตำแหน่งเปอร์เซ็นไทล์ที่ 100(α/2) กำหนดให้เป็น $\hat{\theta}_{BL}$ และหาค่าที่ตำแหน่งเปอร์เซ็นไทล์ที่ 100(1-α/2) กำหนดให้เป็น $\hat{\theta}_{BU}$ ดังนั้น ค่าความเชื่อมั่นของวิธีบุทสแทรป คือ $\left[\hat{\theta}_{BL},\hat{\theta}_{BU}\right]$ [14]

2.3 การคัดเลือกคุณลักษณะ (Feature Selection)

วิธีการคัดเลือกคุณลักษณะ โดยการคัดกรองจากค่าน้ำหนักด้วยค่ามาตรฐานอัตราเกน (Information Gain ratio : IGr) เป็นวิธีที่ถูกนำมาใช้แก้ปัญหามิติข้อมูลสูงในการทำเหมืองข้อความ [15] โดยการจัดอันดับคุณลักษณะด้วยการหาค่ามาตรฐานอัตราเกน ที่มีมากที่สุดในแต่ละมิติข้อมูล เพื่อเลือกเฉพาะคุณลักษณะเด่นของข้อมูล ซึ่งสามารถลดมิติข้อมูลได้โดยที่ส่งผลกระทบต่อคุณลักษณะของข้อมูลโดยรวมน้อยที่สุด วิธี IGr สามารถคำนวณได้ดังสมการ (3) - (4)

$$Entropy = -\sum_t p(j|t)\log_2 p(j|t) \quad (3)$$

โดยที่ $-\sum_t$ คือ ผลรวมความน่าจะเป็นในการเกิดของ j ที่เกิดขึ้นในคลาส t

$$Gain = Entropy(p) - \left[\sum_{i=1}^{k}\frac{n_i}{n}Entropy(i)\right] \quad (4)$$

โดยที่ $Entropy(p)$ คือค่าเอนโทรฟีของค่ารูท (Root)

$\sum_{i=1}^{k}\frac{n_i}{n}Entropy(i)$ คือ ค่าเอนโทรฟีของแต่ละโหนดย่อย

ในการวัดความน่าเชื่อถือของมิติของข้อมูล จะวัดจะค่า Gain Ratio โดยใช้ค่า SplitINFO ดังสมการ (5)

$$SplitINFO = -\sum_{i=1}^{k}\frac{n_i}{n}\log_2\frac{n_i}{n} \quad (5)$$

และคำนวณหาค่า GainRatio ได้ดังสมการ (6)

$$GainRatio = \frac{\Delta INFO}{SplitINFO} \quad (6)$$

2.4 การวัดประสิทธิภาพ

การวัดประสิทธิภาพขั้นตอนวิธีในการเรียนรู้ชุดข้อมูลไม่สมดุลนั้น จากการศึกษาพบว่า โดยส่วนใหญ่การประเมินประสิทธิภาพจะพิจารณาจากค่าความไว (Sensitivity) ค่าความจำเพาะ (Specificity) ค่า F-measure ค่า G-mean และ ค่า AUC (Area Under Curve) ซึ่งเป็นเครื่องมือวัดประสิทธิภาพที่นิยมนำมาใช้วัดการเรียนรู้ข้อมูลไม่สมดุล [1] [16] ดังสมการที่ (7) - (11)

$$TNR = \frac{TN}{TN+FP} \quad (7)$$

$$TPR = \frac{TP}{TP+FN} \quad (8)$$

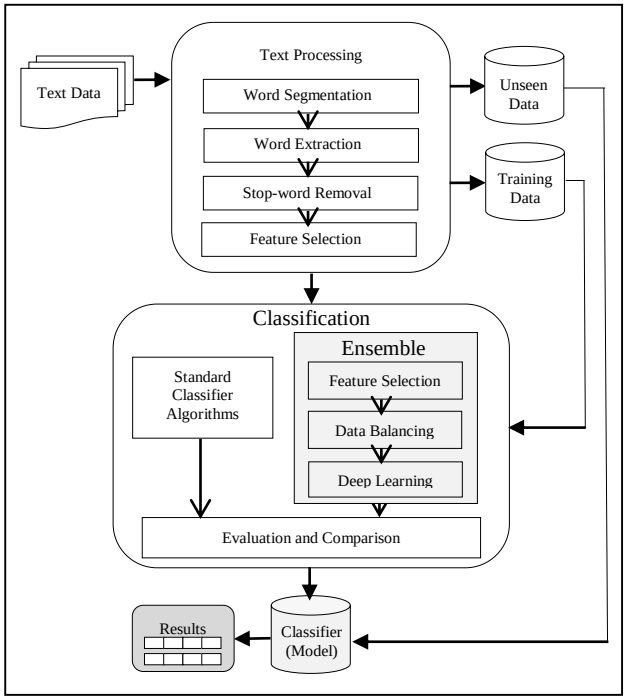
$$F-measure = 2\times\frac{precision\times recall}{precision+recall} \quad (9)$$

$$G-mean = \sqrt{TNR\times TPR} \quad (10)$$

$$AUC = \frac{TP\ rate + TN\ rate}{2} \quad (11)$$

3. วิธีดำเนินการวิจัย

งานวิจัยนี้มีวัตถุประสงค์เพื่อเพิ่มประสิทธิภาพการเรียนรู้เชิงลึกด้วยเทคนิคการคัดเลือกคุณลักษณะ และเทคนิคการปรับสมดุลข้อมูล เพื่อจำแนกประเภทข้อความที่ไม่สมดุล โดยกระบวนการทดลอง แสดงดังภาพที่ 2



ภาพที่ 2 : กระบวนการทดลอง

กระบวนการในการทดลองเริ่มด้วยการนำชุดข้อความมาจัดเตรียมข้อมูลก่อนนำไปประมวลผลด้วยเทคนิคการประมวลผลข้อความ (Text Processing) โดยจัดเตรียมชุดข้อมูลฝึกฝน (Training Data) และชุดข้อมูลทดสอบ (Unseen Data) จากนั้นทำการทดสอบขั้นตอนวิธีมาตรฐานในการจำแนกประเภท โดยนำมาเปรียบเทียบกับวิธีการที่นำเสนอ เพื่อสร้างโมเดลการจำแนกประเภท และนำไปใช้กับข้อมูลชุดทดสอบต่อไป

3.1 ชุดข้อมูลและการจัดเตรียมข้อมูลก่อนการประมวลผล

ชุดข้อมูลที่นำมาใช้ในการทดสอบในครั้งนี้ เป็นข้อความภาษาไทยที่นำมาจากเว็บไซต์ทวิตเตอร์ (www.twitter.com) แบ่งเป็นข้อมูลชุดฝึกฝน จำนวน 2,568 ข้อความ จำแนกกลุ่มข้อความออกเป็น 2 กลุ่ม (Binary Class) คือ กลุ่มข้อความที่เกี่ยวกับคอมพิวเตอร์และเทคโนโลยีสารสนเทศ (IT or Positive Class) และกลุ่มข้อความทั่วไป (Other or Negative Class) โดยจัดเตรียมข้อความให้อยู่ในรูปแบบที่มีโครงสร้างด้วยเทคนิคการประมวลผลข้อความ ซึ่งภายในมีกระบวนการหลัก คือ การตัดคำ (Word Segmentation) การกำจัดคำที่ไม่มีความหมาย (Stop-word Removal) การสกัดหน่วยคำ (Feature Extraction) และการเลือกคำคุณลักษณะ (Feature Selection) เมื่อผ่านกระบวนการประมวลผลข้อความแล้ว สามารถสกัดคุณลักษณะข้อมูล (Feature) ได้จำนวน 6,636 คุณลักษณะ ดังนั้นชุดข้อมูลที่ใช้ในงานวิจัยนี้จึงมีลักษณะเป็นชุดข้อมูลที่มีความไม่สมดุลสูง และมีมิติข้อมูลสูง

3.2 การจำแนกประเภทข้อความ

วิธีการแบบรวมกลุ่มที่นำเสนอ ประกอบไปด้วย 3 เทคนิค คือ 1) เทคนิคการเรียนรู้เชิงลึก 2) เทคนิคการคัดเลือกคุณลักษณะข้อมูล และ 3) เทคนิคกาปรับสมดุลข้อมูล งานวิจัยนี้จึงออกแบบเพื่อทดสอบการจำแนกประเภทไว้ 3 ระดับ คือ

1) ทดสอบประสิทธิภาพการจำแนกประเภทข้อมูล โดยเปรียบเทียบระหว่างขั้นตอนวิธีมาตรฐานที่ได้รับความนิยม จำนวน 5 ขั้นตอนวิธี เปรียบเทียบกับเทคนิคการเรียนรู้เชิงลึก ในการทดลองได้กำหนดพารามิเตอร์ของแต่ละขั้นตอนวิธี ดังนี้ Naïve Bayes กำหนด laplace_correction = true, Support Vector Machine (SVM) กำหนด kernel_type = polynomial Random Forest กำหนด criterion = gain_ratio และ voting_strategy = confidence vote, k-Nearest Neighbor (kNN) กำหนด kernel_type = dot, Neural Network (Multi Layer Perceptron) กำหนดให้ hidden layer โดย กำหนดชั้นซ่อน hd1 = 10, hd2 = 5 และกำหนด training cycles = 80 และ Deep Learning กำหนด activation = ExpRectifier และ epochs = 1.797 จากนั้นนำผลการจำแนกประสิทธิภาพของแต่ละขั้นตอนวิธีมาเปรียบเทียบประสิทธิภาพ

2) ทดสอบการคัดเลือกคุณลักษณะของข้อมูล โดยเปรียบเทียบระหว่างเทคนิคการเรียนรู้เชิงลึกที่ผ่านการฝึกฝนด้วยข้อมูลชุดปกติ (ข้อมูลไม่ผ่านการคัดเลือกคุณลักษณะ) และเทคนิคการเรียนรู้เชิงลึกที่ผ่านการฝึกฝนด้วยชุดข้อมูลที่มีการลดมิติข้อมูล โดยการคัดเลือกคุณลักษณะด้วยวิธีการคำนวณค่าน้ำหนัก IGr ซึ่งทดสอบคัดเลือกคุณลักษณะตามสัดส่วนจากน้อยไปมา คือ 10%, 20%, 30%, 40% และ 50% จากจำนวนคุณลักษณะทั้งหมด แล้วเปรียบเทียบเพื่อเลือกสัดส่วนที่ให้ประสิทธิภาพสูงสุด

3) ทดสอบการปรับสมดุลข้อมูล โดยเปรียบเทียบระหว่างเทคนิคการเรียนรู้เชิงลึกผ่านการฝึกฝนด้วยชุดข้อมูลที่มีการลดมิติข้อมูล เปรียบเทียบกับเทคนิคการเรียนรู้เชิงลึกผ่านการฝึกฝนด้วยชุดข้อมูลที่มีการลดมิติข้อมูลร่วมกับการปรับ

สมดุลข้อมูลด้วยเทคนิค SMOTE และ Random-under Sampling ด้วย Bootstrapping แล้วเปรียบเทียบประสิทธิภาพของแต่ละวิธี

3.3 การวัดประสิทธิภาพ

งานวิจัยนี้ใช้ 10 fold-cross validation ในการทดสอบประสิทธิภาพ โดยใช้เมตริกในการเปรียบเทียบประสิทธิภาพทั้งหมด 5 เมตริก คือ 1) ค่า Sensitivity) 2) ค่า Specificity 3) ค่า F-measure 4) ค่า G-mean และ 5) ค่า AUC

4. ผลการทดลอง

ผลการทดสอบเปรียบเทียบประสิทธิภาพในการจำแนกประเภทข้อความจำนวน 2,375 ข้อความ มีคุณลักษณะจำนวน 6,635 คุณลักษณะ ได้ผลการทดสอบดังตารางที่ 1

ตารางที่ 1 : เปรียบเทียบประสิทธิภาพการจำแนกประเภทข้อความของแต่ละขั้นตอนวิธี

Training data	Data total = 2,375 (Positive =377, Negative = 1,998)				
Measure matrix / Algorithms	TP rate	TN rate	F-measure	G-mean	AUC
Naïve-bayes	0.506	0.901	0.714	0.675	0.820
SVM	0.605	0.981	0.833	0.770	0.793
RandomForest	0.236	0.999	0.656	0.486	0.939
kNN	0.027	0.999	0.483	0.164	0.601
Neural Network	0.000	1.000	unknow	0.000	0.679
Deep Learning	0.706	0.947	0.714	0.818	0.929

ประสิทธิภาพในการเรียนรู้และจำแนกประเภทของเทคนิคการเรียรู้เชิงลึก สามารถเรียนรู้ข้อมูลที่ไม่สมดุลสูง และมีมิติข้อมูลสูงได้อย่างมีประสิทธิภาพ สามารถจัดการกับการเรียนรู้ที่มีการตัดสินใจที่เอนเอียงได้ดีกว่าขั้นตอนวิธีอื่น

การทดลองขั้นต่อไป คือการลดมิติข้อมูลโดยการคัดเลือกคุณลักษณะด้วยวิธีการคำนวณค่าน้ำหนักจากค่า IGr โดยเลือกคุณลักษณะตามสัดส่วนจากน้อยไปมาก คือ 10%, 20%, 30%, 40% และ 50% ได้ผลการทดลองดังแสดงในตารางที่ 2

จากการทดลองนำเทคนิคการคัดเลือกคุณลักษณะของข้อมูลที่มีมิติสูง พบว่า การเลือกคุณลักษณะข้อมูลในอัตราส่วน 20% สามารถลดจำนวนคุณลักษณะข้อมูล จาก 6,636 คุณลักษณะเหลือ 1,327 คุณลักษณะ ซึ่งลดลงมากถึง 5,309 คุณลักษณะ โดยที่ค่าประสิทธิภาพสูงขึ้น นั้นแสดงว่า การคัดเลือกคุณลักษณะด้วยวิธีการคำนวณค่าน้ำหนัก IGr ในอัตราส่วน 20% สามารถเพิ่มประสิทธิภาพให้กับการเรียนรู้ และช่วยลดความเอนเอียงในการตัดสินใจของเทคนิคการเรียนรู้เชิงลึกได้

ตารางที่ 2 : เปรียบเทียบประสิทธิภาพของเทคนิคการเรียนรู้เชิงลึกร่วมกับเทคนิคการคัดเลือกคุณลักษณะ

Trainning data	selection rate 10%=664 , 20%=1,327 , 30%=1,991 , 40%=2,654 , 50%=3,318				
Measure matrix / Selection rate	TP rate	TN rate	F-measure	G-mean	AUC
original features	0.706	0.947	0.714	0.818	0.929
featureSelection(10%)	0.698	0.958	0.726	0.817	0.898
featureSelection(20%)	0.735	0.954	0.740	0.837	0.923
featureSelection(30%)	0.700	0.926	0.672	0.805	0.896
featureSelection(30%)	0.671	0.876	0.574	0.767	0.896
FeatureSelection(50%)	0.748	0.838	0.580	0.792	0.923

และการทดลองขั้นสุดท้าย คือการนำเทคนิคการเรียนรู้เชิงลึกมาทำงานร่วมกับเทคนิคการปรับสมดุลข้อมูล และเทคนิคการคัดเลือกคุณลักษณะด้วย IGr ที่อัตราส่วน 20% ซึ่งในการปรับสมดุลข้อมูล ได้เปรียบเทียบระหว่างเทคนิค SMOTE โดยเพิ่มข้อมูลในอัตราส่วน 400% (อัตราส่วนที่เพิ่มเพื่อให้ใกล้เคียงกับจำนวนข้อมูลคลาสหลัก) และการสุ่มลดข้อมูลในคลาสหลักด้วยวิธีบูทสแทรป โดยลดในอัตราส่วน 80% (อัตราส่วนที่ลดเพื่อให้ใกล้เคียงกับจำนวนข้อมูลคลาสรอง) ผลการทดลองแสดงในตารางที่ 3

ตารางที่ 3 : เปรียบเทียบประสิทธิภาพของเทคนิคการเรียนรู้เชิงลึกร่วมกับการคัดเลือกคุณลักษณะและการปรับสมดุลข้อมูล

Trainning data	SMOTE : data = 3,466 instances (Positive=1,508 , Negative=1958) Bootstrapping data = 777 instances 1,327 (Positive=377 , Negative=400)				
Measure matrix / Ensemble	TP rate	TN rate	F-measure	G-mean	AUC
Deep Learning only	0.706	0.947	0.714	0.818	0.929
Deep Learning with Feature Selection	0.735	0.954	0.740	0.837	0.923
Deep Learning with Feature Selection and SMOTE (over-sampling)	0.897	0.648	0.762	0.762	0.822
Deep Learning with Feature Selection and Bootstrapping (under-sampling)	0.846	0.925	0.874	0.885	0.955

ผลการนำเทคนิคการปรับสมดุลข้อมูลมาใช้ร่วมกับการคัดเลือกคุณลักษณะของข้อมูล โดยเทคนิคการปรับสมดุลข้อมูลด้วยการลดข้อมูลในคลาสหลักด้วยวิธี Bootstrapping สามารถเพิ่มประสิทธิภาพให้เทคนิคการเรียนรู้เชิงลึกให้ดียิ่งขึ้น โดยได้ค่าความไวและความจำเพาะที่มีความใกล้เคียงกันมากขึ้น คือ

0.846 และ 0.925 นั่นคือความเอนเอียงในการจำแนกประเภทลดลง นอกจากนั้น ค่า F-measure เพิ่มขึ้นเท่ากับ 0.874 ค่า G-mean เพิ่มขึ้นเท่ากับ 0.885 และค่า AUC เพิ่มขึ้นเท่ากับ 0.955 ซึ่งเป็นค่าสูงที่สุดในการทดลอง ในทางตรงกันข้ามเทคนิคการปรับสมดุลข้อมูลโดยการเพิ่มข้อมูลในคลาสรองด้วย SMOTE กลับให้ประสิทธิภาพลดลง

5.สรุป อภิปรายผล และข้อเสนอแนะ

งานวิจัยนี้นำเสนอการการแก้ปัญหาข้อมูลไม่สมดุลสูงที่อยู่ในรูปแบบของข้อความ ซึ่งมีลักษณะเป็นข้อมูลที่มีมิติสูง โดยวิธีการแบบรวมกลุ่มด้วยเทคนิคการเรียนรู้เชิงลึกร่วมกับการคัดเลือกคุณลักษณะและการปรับสมดุลข้อมูล ผลการทดลองพบว่า วิธีที่นำเสนอสามารถจำแนกประเภทข้อความที่ไม่สมดุลสูงได้ดียิ่งขึ้น และลดความเอนเอียงในการตัดสินใจที่ให้ความสำคัญกับคลาสหลักให้ลดลงได้อย่างมีประสิทธิภาพ

ซึ่งจากการทดลองเทคนิคการปรับสมดุลข้อมูลด้วยการลดจำนวนข้อมูลในคลาสหลัก ด้วยวิธีบูทสแทรปสามารถทำการเรียนรู้และจำแนกประเภทข้อมูลที่ไม่สมดุลสูง และมีมิติข้อมูลสูงได้ดีกว่าเทคนิค SMOTE นอกจากนั้น ผลการทดลองยังแสดงให้เห็นว่า การใช้เทคนิคการคัดเลือกคุณลักษณะของข้อมูล นอกจากช่วยแก้ปัญหาข้อมูลที่มีมิติสูงแล้วนั้น ยังช่วยแก้ปัญหาข้อมูลไม่สมดุลสูงได้ในระดับที่น่าพอใจ

การศึกษาต่อไปในอนาคต ผู้วิจัยมีความมุ่งหวังในการศึกษาเพื่อปรับปรุงและพัฒนาวิธีการเรียนรู้ข้อมูลในลักษณะต่างๆ เช่น ข้อมูลที่มีความทับซ้อน ข้อมูลที่มีความคลาดเคลื่อน ข้อมูลเบาบาง เป็นต้น ซึ่งเป็นปัญหาที่มีระดับความยากและมีความท้าทายมากยิ่งขึ้น

เอกสารอ้างอิง

[1] H. He and E. A. Garcia, “Learning from Imbalanced Data,” *IEEE Trans. Knowl. Data Eng.*, vol. 21, no. 9, pp. 1263–1284, Sep. 2009.

[2] P. Vorraboot, S. Rasmequan, K. Chinnasarn, and C. Lursinsap, “Feature Selection for Imbalanced Datasets with Hybrid Approach,” presented at the 4th Conference on Knowledge and Smart Technology (KST), 2012, pp. 61–68.

[3] S. Satiman and T. Chintakovid, “Measuring Performance of Algorithms for Short Text Classification: A case study of Thai messages on Twitter,” in *NCCIT2011*, 2011, pp. 509–514.

[4] Z. Jiang, L. Li, D. Huang, and L. Jin, “Training word embeddings for deep learning in biomedical text mining tasks,” in *2015 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 2015, pp. 625–628.

[5] R. Ghosh, K. Ravi, and V. Ravi, “A novel deep learning architecture for sentiment classification,” in *2016 3rd International Conference on Recent Advances in Information Technology (RAIT)*, 2016, pp. 511–516.

[6] S. Guan, M. Chen, H. Y. Ha, S. C. Chen, M. L. Shyu, and C. Zhang, “Deep Learning with MCA-based Instance Selection and Bootstrapping for Imbalanced Data Classification,” in *2015 IEEE Conference on Collaboration and Internet Computing (CIC)*, 2015, pp. 288–295.

[7] J. Song, S. Qin, and P. Zhang, “Chinese text categorization based on deep belief networks,” in *2016 IEEE/ACIS 15th International Conference on Computer and Information Science (ICIS)*, 2016, pp. 1–5.

[8] I. Arel, D. C. Rose, and T. P. Karnowski, “Deep Machine Learning - A New Frontier in Artificial Intelligence Research [Research Frontier],” *IEEE Comput. Intell. Mag.*, vol. 5, no. 4, pp. 13–18, Nov. 2010.

[9] Y. Yan, M. Chen, M. L. Shyu, and S. C. Chen, “Deep Learning for Imbalanced Multimedia Data Classification,” in *2015 IEEE International Symposium on Multimedia (ISM)*, 2015, pp. 483–488.

[10] P. Hart, “The condensed nearest neighbor rule (Corresp.),” *IEEE Trans. Inf. Theory*, vol. 14, no. 3, pp. 515–516, May 1968.

[11] I. Tomek, “Two Modifications of CNN,” *IEEE Trans. Syst. Man Cybern.*, vol. SMC-6, no. 11, pp. 769–772, Nov. 1976.

[12] J. Laurikkala, “Improving Identification of Difficult Small Classes by Balancing Class Distribution,” *Springer AIME*, pp. 63–66, 2001.

[13] J. Prusa, T. M. Khoshgoftaar, D. J. Dittman, and A. Napolitano, “Using Random Undersampling to Alleviate Class Imbalance on Tweet Sentiment Data,” in *2015 IEEE International Conference on Information Reuse and Integration (IRI)*, 2015, pp. 197–202.

[14] K. Singh and M. Xie, “Bootstrap: A Statistical Method.” .

[15] C. Shang, M. Li, S. Feng, Q. Jiang, and J. Fan, “Feature selection via maximizing global information gain for text classification,” *Knowl.-Based Syst.*, vol. 54, pp. 298–309, Dec. 2013.

[16] S. M. Abd Elrahman and A. Abraham, “A Review of Class Imbalance Problem,” *J. Netw. Innov. Comput.*, vol. 1, pp. 332–340, 2013.

การรู้จำชื่อเฉพาะเกี่ยวกับภัยธรรมชาติจากข้อความทวิตเตอร์
NER Detection from Thai Tweets Related to Natural Disaster

มาลีรัตน์ โสดานิล (Maleerat Sodanil)[1] และยุทธนา ลังกาตุง (Yuttana Lungkatoong)[2]

[1]คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

[2]สำนักงานสรรพากรภาค 3

[1]maleerat.s@it.kmutnb.ac.th, [2]yuttana.la@gmail.com

บทคัดย่อ

ปัจจุบันสื่อสังคมออนไลน์ ถูกใช้งานอย่างแพร่หลาย ไม่ว่าจะเป็นการเผยแพร่เรื่องราวส่วนตัว ความรู้ การขายสินค้า และอื่น ๆ อีกจำนวนมาก รวมไปถึงเมื่อเกิดเหตุภัยพิบัติต่าง ๆ ขึ้น สื่อสังคมออนไลน์มักเป็นช่องทางแรกในการเผยแพร่ข้อมูลข่าวสารเหตุการณ์ภัยพิบัติที่เกิดขึ้น ซึ่งทวิตเตอร์ เป็นหนึ่งในสื่อสังคมออนไลน์อันดับต้น ๆ ที่มีปริมาณการเผยแพร่ข้อมูลข่าวสารจำนวนมหาศาล ซึ่งข้อมูลข่าวสารเหล่านี้ ประกอบไปด้วยข้อมูลที่มีประโยชน์ต่อการนำไปใช้งานเพื่อบริหารจัดการและการวางแผนการให้ความช่วยเหลือแก่ผู้ประสบภัยพิบัติ และอีกส่วนหนึ่งเป็นข้อมูลที่ไม่เกี่ยวข้องต่อการนำไปใช้งาน ทั้งนี้โดยธรรมชาติของข้อมูลข่าวสารที่ถูกเผยแพร่ทางสื่อสังคมออนไลน์นั้น มักเป็นข้อมูลแบบไม่มีโครงสร้าง ดังนั้นการสกัดข้อมูลเฉพาะที่เป็นประโยชน์ ต่อการนำไปใช้งานจึงเป็นการยาก ในงานวิจัยนี้ได้นำเสนอการรู้จำชื่อเฉพาะเกี่ยวกับภัยพิบัติธรรมชาติ เพื่อสกัดข้อมูลข่าวสารจากทวิตเตอร์ โดยใช้ข้อความเกี่ยวกับภัยพิบัติธรรมชาติจากทวิตเตอร์ ตั้งแต่วันที่ 28 มกราคม 2558 ถึงวันที่ 14 กันยายน 2558 จำนวน 100,000 คำ และใช้ขั้นตอนวิธีคอนดิชั่นนัล แรนดอม ฟิลด์ ในการรู้จำชื่อเฉพาะ ซึ่งเมื่อพิจารณาผลการประสิทธิภาพของแบบจำลองการรู้จำชื่อเฉพาะแสดงให้เห็นว่าวิธีการดังกล่าวมีประสิทธิภาพเพียงพอต่อการนำไปใช้งานจริง เมื่อเปรียบเทียบกับแบบจำลองการรู้จำชื่อเฉพาะที่มีลักษณะใกล้เคียงกัน

คำสำคัญ: ภัยพิบัติ การเรียนรู้ของเครื่อง การประมวลผลภาษาธรรมชาติ การสกัดข้อมูลสารสนเทศ การรู้จำชื่อเฉพาะ

Abstract

Social Network is the well-known media to distribute any story such as private data, knowledge commerce including when disaster is occurred. Social media is the first channel to announce and disseminate information to public. These information are useful beneficial for disaster management in order to help and organize just in time or in advance. Normally, data type can be either structured or unstructured which is difficult to extract only useful information. This research aims to develop the method that can be detect and recognize the entity name which is suitable for using in disaster management in the future. Data used in this research was collected between January 28, 2016 and September 14, 2016 in total of 100,000 words. Also Conditional Random Fields (CRF) algorithm is applied for name entity recognition. The performance was evaluated and compared with similar algorithms, given such a good score which can be applied to use in the real world.

Keyword: Natural Disaster, Machine Learning, Natural Language Processing, Information Extraction, Named Entity Recognition.

1. บทนำ

ภัยพิบัติทางธรรมชาติเป็นเรื่องที่มีผู้สนใจศึกษาวิจัยเป็นจำนวนมาก อันเนื่องมาจากผลของการเกิดภัยพิบัติมักก่อให้เกิดความเสียหายอย่างร้ายแรงทั้งชีวิตและทรัพย์สินของมนุษย์ ซึ่งเมื่อเกิดเหตุภัยพิบัติทางธรรมชาติขึ้น ผู้คนมักจะใช้ช่องทางต่างๆ ในการติดต่อสื่อสารกัน เพื่อการแพร่กระจายข่าวสาร การค้นหา

ข้อมูล การเฝ้าระวัง ตลอดจนไปถึงการขอความช่วยเหลือ และช่องทางที่นิยมในการใช้ติดต่อสื่อสารเป็นอย่างมากในปัจจุบัน คือ สื่อสังคมออนไลน์ (Social Media) ซึ่งลักษณะของการเผยแพร่ข้อมูลข่าวสารผ่านทางสื่อสังคมออนไลน์นั้นมีความรวดเร็วกว่าวิธีเผยแพร่ข้อมูลข่าวสารแบบดั้งเดิมอีกทั้งผู้ใช้งานสามารถเป็นได้ทั้งผู้เผยแพร่ข่าวสารและผู้รับข่าวสารในขณะเดียวกัน ทวิตเตอร์ (Twitter) เป็นเว็บไซต์ที่เปิดให้ส่งข้อความสั้น ๆ ที่มีความยาวไม่เกิน 140 ตัวอักษร เป็นหนึ่งในสื่อสังคมออนไลน์ที่ถูกใช้เพื่อเผยแพร่ข้อมูลข่าวสารเกี่ยวกับภัยพิบัติเป็นจำนวนมาก ข้อมูลข่าวสารเหล่านี้สามารถนำมาใช้ประโยชน์ในด้านต่าง ๆ เกี่ยวกับการจัดการและวางแผนเกี่ยวกับการให้ความช่วยเหลือผู้ประสบภัยได้ แต่ปัญหาในการนำข้อมูลข่าวสารดังกล่าวมาใช้งานนั้นเกิดจากปริมาณข้อมูลจากที่ผู้ใช้งานทำการเผยแพร่ผ่านทวิตเตอร์ในแต่ละวันนั้นมีจำนวนมหาศาล

หลายงานวิจัยมีการนำข้อมูลจากทวิตเตอร์ มาใช้เป็นข้อมูลเพื่อสร้างและทดสอบแบบจำลอง ยกตัวอย่างเช่น Kongthon et al [1] ได้วิเคราะห์ปริมาณการเผยแพร่ข้อมูลเกี่ยวกับอุทกภัยช่วงเดือนตุลาคม ถึง ธันวาคม พ.ศ. 2554 เผยให้เห็นว่ามีการใช้งานทวิตเตอร์เพื่อเผยแพร่ข้อมูลเกี่ยวกับอุทกภัยเพิ่มขึ้นเป็นจำนวนมากในช่วงที่เกิดอุทกภัยใหญ่ในประเทศไทยปี พ.ศ. 2554 นอกจากนั้น ยุทธนาและมาลีรัตน์ [2] ได้ศึกษาการจำแนกหมวดหมู่ข้อความภาษาไทยเกี่ยวกับแผ่นดินไหวจากทวิตเตอร์ โดยแบ่งออกเป็นสี่หมวดหมู่ คือ ข้อความเกี่ยวกับสถานการณ์แผ่นดินไหว ข้อความเกี่ยวกับการเฝ้าระวังการเกิดแผ่นดินไหว ข้อความเกี่ยวกับการร้องขอและการให้ช่วยเหลือผู้ประสบภัย ข้อความอื่น ๆ ที่ไม่เกี่ยวข้อง ซึ่งในแต่ละข้อความนั้นนอกจากแยกเป็นหมวดหมู่แล้วยังมีข้อมูลอื่น ๆ ทั้งที่เป็นประโยชน์ประกอบกันอยู่ในหนึ่งข้อความ เช่น ประเภทของภัยพิบัติ หน่วยงาน บุคคล สถานที่ และข้อความอื่นๆ ที่ไม่เกี่ยวข้องหรือไม่เป็นประโยชน์ โดยในงานวิจัยนี้ศึกษาการคัดแยกข้อความ ด้วยวิธีการรู้จำชื่อเฉพาะ (Named Entity Recognition: NER) ในการสกัดชื่อเฉพาะจากข้อความทวิตเตอร์ ที่เป็นภาษาไทยและเกี่ยวข้องกับภัยพิบัติธรรมชาติ เพื่อประโยชน์ในการบริหารจัดการและการวางแผนเกี่ยวกับภัยพิบัติทางธรรมชาติต่อไป

2. วรรณกรรมที่เกี่ยวข้อง

2.1 สื่อสังคมออนไลน์กับเหตุการณ์ภัยพิบัติ

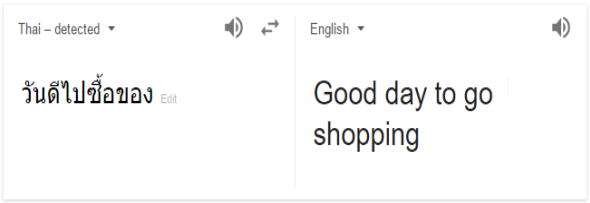
ปัจจุบันสื่อสังคมออนไลน์ได้รับความนิยมในการใช้งานสูง ไม่จำกัดเพียงแค่กลุ่มผู้ใช้ที่คุ้นเคยกับเทคโนโลยีอินเตอร์เน็ตเท่านั้น กลุ่มผู้ใช้บางกลุ่ม เช่น ผู้สูงอายุ เริ่มหันมาศึกษาใช้งานสื่อสังคมออนไลน์กันมากขึ้น ทั้งนี้ อาจเกิดจากอิทธิพลของการพัฒนาทางด้านเทคโนโลยีที่ทันสมัยมากขึ้นและสามารถเข้าถึงผู้ใช้ได้มากขึ้น ไม่ว่าจะเป็นเรื่องความสะดวก สบายในการใช้งาน หรือในด้านราคาที่ผู้ใช้งาน สามารถเข้าถึงได้มากขึ้น ซึ่งทวิตเตอร์เป็นหนึ่งในเว็บไซต์สื่อสังคมออนไลน์ ที่ผู้ใช้นิยมใช้งานเป็นอันดับต้นๆ ด้วยการเผยแพร่ข้อความขนาดไม่เกิน 140 ตัวอักษร ผู้ใช้งานจึงมักใช้ข้อความที่กระชับได้ใจความ ทำให้ผู้อ่านสามารถเข้าใจความหมายได้ง่าย

ในไม่กี่ปีที่ผ่านมาประเทศไทยประสบภัยพิบัติทางภัยธรรมชาติหลายครั้ง เช่น การเกิดอุทกภัยใหญ่ในประเทศไทยเมื่อปี 2554 ซึ่งในครั้งนั้นได้มีการวิจัยพบว่าปริมาณการเผยแพร่ข้อมูลเกี่ยวกับอุทกภัยเพิ่มขึ้นเป็นจำนวนมาก [1] แสดงให้เห็นว่าผู้ใช้งานในประเทศไทย ได้มีการนำทวิตเตอร์ มาใช้เป็นช่องทางในการเผยแพร่ข่าวสารภัยพิบัติเป็นจำนวนมาก ในต่างประเทศก็มีการใช้งานสื่อสังคมออนไลน์ในการเผยแพร่ข้อมูลข่าวสารภัยพิบัติทางธรรมชาติเช่นกัน ยกตัวอย่างเช่น การศึกษาข้อมูลจาก ไมโครบล็อก (Microblog) ในประเทศจีน ช่วงการเกิดแผ่นดินไหวที่เมืองยูซู (Yushu) ปี พ.ศ. 2553 [3] ซึ่งการเผยแพร่ข้อมูลข่าวสารมีจุดประสงค์ที่แตกต่างกันออกไป เช่น การเผยแพร่ข้อมูลการเกิดเหตุการณ์ภัยพิบัติ การเผยแพร่ข้อมูลเพื่อร้องขอหรือให้ความช่วยเหลือแก่ผู้ประสบภัยพิบัติ และการเฝ้าระวังติดตามการเกิดภัยพิบัติ ข้อมูลเหล่านี้ถูกส่งต่อและเผยแพร่ในสื่อสังคมออนไลน์อย่างรวดเร็ว ต่างจากการเผยแพร่ข้อมูลข่าวสารในอดีตที่ผู้เผยแพร่ข่าวสารส่วนใหญ่เป็นหน่วยงานทางด้านสื่อมวลชลเท่านั้น เช่น หนังสือพิมพ์ รายการวิทยุ รายการโทรทัศน์ ซึ่งผู้รับข่าวสารอาจไม่ได้รับข่าวสารได้ทันกับเหตุการณ์ภัยพิบัติที่เกิดขึ้น

2.2 Named Entity Recognition (NER)

Named Entity Recognition หรือ NER เป็นการระบุประเภทของชื่อเฉพาะที่อยู่ในข้อความ ยกตัวอย่างเช่น บุคคล สถานที่ องค์กร เป็นต้น ซึ่งวิธีการทำสามารถแบ่งออกเป็น 3 หมวดหมู่

ได้แก่ Rule-based, Machine Learning และ Hybrid Method [4] งานวิจัยส่วนใหญ่มุ่งเน้นในการระบุประเภทของคำจากข้อความ ที่มีรูปแบบ เช่น Wikipedia หนังสือพิมพ์ [4], [5] มีความ สำคัญที่นำไปใช้เกี่ยวกับกระบวนการประมวลผลภาษาทางธรรมชาติ (Natural Language Processing) หลากหลายในหลาย ๆ ด้าน เช่น การค้นคืนสารสนเทศ (Information Retrieval) การคัดแยกสารสนเทศ (Information Extraction) การแปลภาษาด้วยเครื่อง (Machine Translation) ซึ่งตัวอย่างที่ชัดเจนเห็นได้จากระบบการแปลภาษาด้วยเครื่อง หากเครื่องไม่สามารถแยกได้ว่าระหว่างชื่อเฉพาะและคำที่ต้องทำการแปลความหมายจะทำให้เกิดความผิดพลาดในการแปลได้ ยกตัวอย่างเช่น “วันดีไปซื้อของ” เครื่องอาจทำการแปลเป็น “Good day to go shopping” ดังภาพที่ 2



ภาพที่ 1: ตัวอย่างการแปลภาษาด้วยเครื่อง

มีงานวิจัยจำนวนมากที่ให้ความสนใจในกระบวนการทำการรู้จำชื่อเฉพาะซึ่งแต่ละภาษานั้นมีความยากง่ายแตกต่างกันไป หากพิจารณาแล้วภาษาอังกฤษมีความง่ายในการทำการรู้จำชื่อเฉพาะกว่าภาษาไทยหรือภาษาอื่น ๆ ที่มีลักษณะคล้ายกับภาษาไทย เช่น ภาษาจีน เนื่องจากภาษาอังกฤษมีการแบ่งคำด้วยช่องว่าง ทำให้ทราบขอบเขตของคำชัดเจน อีกทั้งในภาษาอังกฤษคำที่เป็นชื่อเฉพาะจะถูกเขียนขึ้นต้นด้วยตัวพิมพ์ใหญ่ ถึงแม้ว่าจะอยู่กลางประโยคก็ตามซึ่งแตกต่างกับภาษาไทยที่ไม่มีตัวบ่งบอกการตัดคำและชื่อเฉพาะ ทำให้มีนักวิจัยสนใจการเพิ่มประสิทธิภาพการรู้จำชื่อเฉพาะภาษาอื่นๆขึ้นมากมายนัชชาถิระสาโรช[6] ทำการวิจัยเปรียบเทียบประสิทธิภาพการรู้จำชื่อเฉพาะจากคลังข้อมูลที่ตัดข้อความแบบเป็นพยางค์และคลังข้อมูลที่ตัดข้อความแบบเป็นคำ โดยใช้คลังข้อมูลภาษาไทยจาก “BEST ”2009 ซึ่งใช้วิธีการรู้จำชื่อเฉพาะแบบ Conditional Random Field (CRF) ซึ่งเป็นการรู้จำชื่อเฉพาะแบบมีผู้สอน (Supervise Learning) กล่าวคือต้องมีการกำกับข้อมูลในคลังข้อมูลสำหรับเรียนรู้ก่อน ซึ่งในงานวิจัยดังกล่าวได้ทำการกับกำข้อมูลชื่อเฉพาะแบบธรรมดาร่วมกับการกำกับป้ายกำกับเพิ่มเติม BIOES ซึ่งเป็นการกำหนดขอบเขตเพิ่มเติม คือ B จุดเริ่มต้น, I ส่วนที่อยู่ในชื่อเฉพาะ, O ส่วนที่ไม่เกี่ยวข้อง, E จุดสิ้นสุด และ S เป็นป้ายกำกับที่บอกถึงคำที่เป็นอักษรตัวเดียว โดยผลการทดสอบประสิทธิภาพให้ผลใกล้เคียงกันทั้งการรู้จำชื่อเฉพาะที่ใช้คลังข้อมูลแบบตัดเป็นพยางค์และคลังข้อมูลแบบตัดเป็นคำ

งานวิจัยของ ยศวัธน์ แสนสิงห์ [7] ทำการทดสอบประสิทธิภาพในการรู้จำชื่อเฉพาะ เปรียบเทียบระหว่างการใช้กฎในกาหาขอบเขต ในการระบุประเภทของคำหลัก กับการใช้แบบจำลองฮิดเดนมาร์คอฟ (Hidden Markov Models-HMMs) และการใช้กฎในการหาขอบเขตร่วมกับ HMMs โดยใช้ข้อมูลจากเว็บไซต์ Thairath.co.th ตั้งแต่ปี 2552 และพิจารณารู้จำคำหลัก 3 ประเภท คือ ชื่อบุคคล ชื่องค์กร และสถานที่ ผลการวิจัยปรากฏว่าการใช้กฎในการหาขอบเขตในการระบุประเภทของคำหลัก ให้ประสิทธิภาพโดยรวมสูงสุด มีการนำ Named Entity Normalization (NEN) มาใช้ร่วมกับการทำการรู้จำชื่อเฉพาะในภาษาอังกฤษ เพื่อหารูปแบบของชื่อเฉพาะที่มีความหมายเดียวกันแต่เขียนไม่เหมือนกัน Liu และคนอื่น ๆ [4] ได้นำ NEN มาใช้ในการหารูปแบบบรรทัดฐานของชื่อเฉพาะโดยการเปรียบเทียบข้อความจาก Twitter หลาย ๆ ข้อความ ซึ่งในงานวิจัยได้ยกตัวอย่างชื่อเฉพาะ “Lady Gaga” และ “Lady Gaaaga” ต่างก็เป็นชื่อเฉพาะประเภทบุคคลและรูปแบบของการเขียนที่ใช้เป็นบรรทัดฐาน คือ “Lady Gaga” และจากผลการวิจัยแสดงให้เห็นว่าการนำ NEM มาใช้ร่วมกับการทำการรู้จำชื่อเฉพาะมีประสิทธิภาพสูงขึ้น Küçük และ Steinberger [8] ได้นำการหารูปแบบบรรทัดฐานของชื่อเฉพาะมาใช้เช่นกัน โดยใช้ในการวิจัยการรู้จำชื่อเฉพาะข้อความจาก Twitter ภาษาตุรกี ซึ่งมีการเขียนชื่อเฉพาะหลากหลายรูปแบบ ยกตัวอย่างเรื่องตัวพิมพ์เล็กพิมพ์ใหญ่ และการใช้เครื่องหมายการออกเสียง ซึ่งบางข้อความมีการเขียนกำกับแลบางข้อความไม่มีการเขียนกำกับ จึงได้ทำการทดลองเปรียบเทียบหาประสิทธิภาพการรู้จำชื่อเฉพาะโดยแบ่งของเป็นสองรูปแบบ การทดลองแรกใช้คลังข้อมูลแบบไม่มีการทำรูปแบบบรรทัดฐานชื่อเฉพาะและทำการทดลองเปรียบเทียบระหว่างการใช้ตัวพิมพ์ใหญ่และไม่ใช้ซึ่งผลการทดลองการไม่ใช้ตัวพิมพ์ใหญ่สามารถเพิ่มประสิทธิภาพการรู้จำชื่อ

เฉพาะประเภท ชื่อบุคคล สถานที่ และองค์กรเท่านั้น แต่เมื่อใช้กับชื่อเฉพาะ หลาย ๆ ประเภทประสิทธิภาพของการรู้จำชื่อเฉพาะลดลง ในการทดลองที่สองใช้คลังข้อมูลที่มีการทำรูปแบบบรรทัดฐานของชื่อเฉพาะ ผลการทดลองประสิทธิภาพของการรู้จำชื่อเฉพาะเพิ่มขึ้นเล็กน้อย Che และคนอื่น ๆ [9] ทำการศึกษาการรู้จำชื่อเฉพาะภาษาจีนโดยการใช้การรู้จำแบบสองภาษา ผู้วิจัยใช้วิธีเปรียบเทียบการจัดลำดับรูปประโยคระหว่างสองภาษา คือ ภาษาจีนและภาษาอังกฤษ ผลการวิจัยแสดงให้เห็นว่าการรู้จำชื่อเฉพาะโดยใช้สองภาษาร่วมกันโดยอาศัยการเปรียบเทียบการจัดเรียงรูปประโยคช่วยเพิ่มประสิทธิภาพในการรู้จำชื่อเฉพาะได้มากขึ้น

2.3 NER กับสื่อสังคมออนไลน์ (Twitter) และภัยพิบัติทางธรรมชาติ

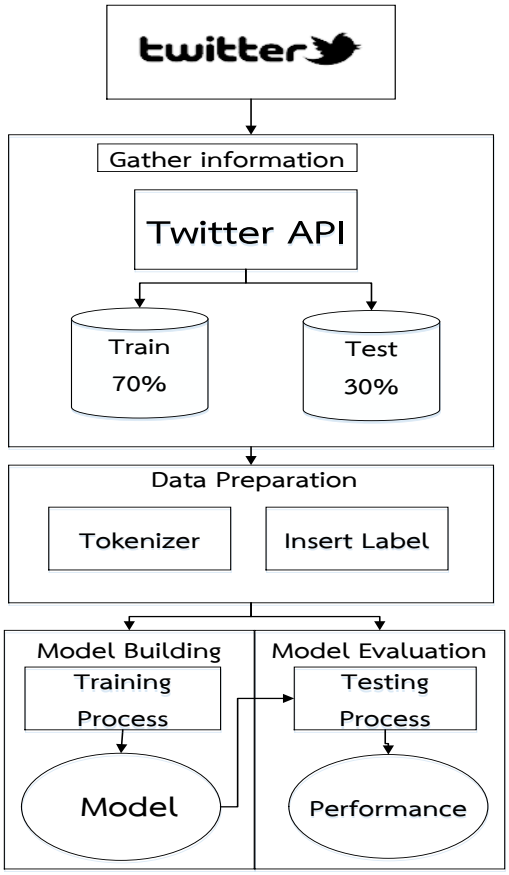
การใช้งานสื่อสังคมออนไลน์ถูกใช้เป็นช่องทางในการติดต่อสื่อสารและเผยแพร่ข้อมูลต่าง ๆ ซึ่งลักษณะของข้อมูลที่เผยแพร่มีหลากหลายรูปแบบ ทั้งเรื่องของการใช้ภาษาการสะกดตัวอักษร มีการเติมแต่งคำด้วยการเพิ่มตัวอักษรดัดแปลงไปจากการสะกดคำมาตรฐาน [8] ด้วยเหตุนี้งานวิจัยหลายชิ้นจึงได้นำ NEN มาใช้ [4] เพื่อหารูปแบบที่เป็นบรรทัดฐานของชื่อเฉพาะดังที่ได้กล่าวมาแล้ว Habib และ van Keulen [5] ได้นำเสนอ NEED4Tweet กระบวนการทำ Twitterbot เพื่อใช้ในการตอบข้อความบนเว็บไซต์ทวิตเตอร์ในส่วนที่ผู้ทดลองได้สร้างขึ้นเพื่อทำการวิจัยในกระบวนการทำงาน มีการเปรียบเทียบวิธีการในการรู้จำชื่อเฉพาะ ซึ่งได้นำการใช้การรู้จำชื่อเฉพาะด้วยวิธีการ CRF มาช่วยในกระบวนการสร้างชื่อเฉพาะ ทั้งนี้เพื่อใช้ในการสร้างข้อความสำหรับการตอบกลับ

นอกจากนี้ ผลจากการวิจัยยังแสดงให้เห็นถึงการใช้งานสื่อสังคมออนไลน์ในการเผยแพร่ข่าวสารจำนวนมาก ซึ่งมีจุดประสงค์ในการเผยแพร่ข่าวสารแตกต่างกันออกไปและทวิตเตอร์เป็นหนึ่งในสื่อสังคมออนไลน์ที่ถูกใช้งานเป็นจำนวนมาก ทำให้มีผู้สนใจนำข้อมูลข่าวสารจากทวิตเตอร์มาทำการวิจัยจำนวนมากเช่นกัน ข้อความจากทวิตเตอร์จะไม่มีลักษณะแบบแผน ลักษณะการเขียนต่างจากจากคลังข้อมูลที่ใช้ในการทำ NER ที่ผ่านมาในอดีตอีกทั้งภาษาไทยนั้นไม่มีการเว้นช่องว่างระหว่างคำ และข้อความทวิตเตอร์บางครั้งมีการเพิ่มตัวอักษรหรือดัดแปลงให้ต่างไปจากพจนานุกรมทำให้เกิดปัญหาในการตัดคำ ซึ่งงานวิจัยที่มีปัญหาเนื่องจากรูปแบบของภาษาเช่นเดียวกัน คือ การทำการรู้จำชื่อเฉพาะข้อความทวิตเตอร์ภาษาตุรกรี [8] พบปัญหาคล้ายกับการทำการรู้จำชื่อเฉพาะจากคลังข้อมูลทวิตเตอร์ภาษาไทย คือรูปแบบการสะกดคำที่ใช้ในทวิตเตอร์ มีความหลากหลายตามผู้ใช้งาน เช่น เรื่องการใช้เครื่องหมายเน้นเสียง บางข้อความมีการใช้ แต่ในบางข้อความไม่มีการใช้ ทำให้ชื่อเฉพาะที่ต้องทำการรู้จำมีความแตกต่างกันถึงแม้จะเป็นคำเดียวกัน การเพิ่มอักษรซ้ำเข้าไปในคำ ทำให้กลายเป็นคำที่ไม่มีความหมายและยากแก่การทำการรู้จำชื่อเฉพาะ อย่างไรก็ตามการพัฒนาประสิทธิภาพการรู้จำชื่อเฉพาะจากคลังข้อมูลทวิตเตอร์ภาษาไทยนั้นมีความจำเป็นเพื่อใช้เป็นเครื่องมือในการคัดแยกข้อมูลสารสนเทศที่มีประโยชน์ออกจากข้อความทวิตเตอร์เมื่อเกิดภัยพิบัติขึ้น

3. วิธีการดำเนินการวิจัย

3.1 การเก็บรวบรวมข้อมูล

สำหรับแหล่งข้อมูลที่ใช้ในการวิจัยนี้ใช้ข้อมูลจากทวิตเตอร์ ซึ่งโดยปกติแล้วทวิตเตอร์ไม่ไม่บริการข้อมูลย้อนหลังผ่านทาง API หลักที่เปิดให้บริการ แต่ทั้งนี้ทวิตเตอร์มีระบบค้นคืนสารสนเทศที่สามารถค้นหาข้อมูลย้อนหลังได้ซึ่งสามารถนำไปเป็นข้อมูลในงานวิจัยนี้ได้ ซึ่งการเก็บข้อมูลเพื่อใช้ทำคลังข้อมูลต้องทำการแยกข้อความออกจากกันเพื่อใช้เก็บลงสู่ฐานข้อมูล โดยใช้ Regular Expression ในการหารูปแบบของรหัส HTML ซึ่งในการเก็บลงสู่ฐานข้อมูลจะเก็บในลักษณะหนึ่งข้อความต่อหนึ่งระเบียน



ภาพที่ 2: กรอบแนวคิดการดำเนินการวิจัย

3.2 การจัดเตรียมข้อมูล

การจัดเตรียมข้อมูลเพื่อนำไปใช้สำหรับการสร้างแบบจำลองการรู้จำชื่อเฉพาะนั้น ข้อความจากทวิตเตอร์ต้องทำการตัดคำก่อนซึ่งในกระบวนการตัดคำเพื่อใช้เป็นฐานข้อมูลสำหรับสร้างแบบจำลองในงานวิจัยนี้ได้ใช้โปรแกรมตัดคำภาษาไทย (LexTo : Thai Lexeme Tokenizer)[1] โดยมีการดัดแปลงแก้ไขรหัสต้นฉบับให้โปรแกรมเพิ่มตัวอักษรพิเศษ (Pipe : "|") เพื่อใช้เป็นตัวอักษรอ้างอิงในการตัดคำ โดยในงานวิจัยครั้งนี้จะใช้ข้อมูลเกี่ยวกับภัยพิบัติแผ่นดินไหวในช่วงตั้งแต่วันที่ 1 เมษายน 2558 ถึงวันที่ 30 เมษายน 2558 และอุทกภัย ตั้งแต่วันที่ 28 มกราคม 2558 ถึงวันที่ 14 กันยายน 2558 จากข้อความบน Twitter ทั้งหมดจำนวน 100,000 คำ แบ่งโดยทำการแบ่งเป็นข้อมูลสำหรับเรียนรู้ร้อยละ 70 และข้อมูลทดสอบร้อยละ 30 งานวิจัยนี้ใช้โปรแกรม Stanford Named Entity Recognizer (NER) ในการสร้างแบบ จำลองการรู้จำชื่อเฉพาะ โดยมีการแบ่งกลุ่มของการรู้จำชื่อเฉพาะออกเป็น 6 กลุ่ม คือ ประเภทของภัยพิบัติ (Type) สถานที่ (Location) หน่วยงาน (Organization) บุคคล (Person) ระดับความเสียหายและผลกระทบ (Level) และคำอื่น ๆ ที่ไม่เกี่ยวข้อง

3.3 การสร้างแบบจำลองการจำแนกหมวดหมู่

เพื่อให้เป็นประโยชน์ต่อการบริหารจัดการและการวางแผนการช่วยเหลือผู้ประสบภัยธรรมชาติในประเทศไทย จึงดำเนินการจัดทำคลังข้อมูลเพื่อใช้ในการทดสอบจากข้อความทวิตเตอร์ภาษาไทยและทำการแบ่งประเภท ของชื่อเฉพาะออกเป็น 6 ประเภทคือ 1) คำที่เกี่ยวข้องกับชนิดของภัยพิบัติทางธรรมชาติ 2) คำที่เกี่ยวข้องกับบุคคล 3) คำที่เกี่ยวกับสถานที่ 4) คำที่เกี่ยวกับหน่วยงานหรือองค์กรต่างๆ 5) คำที่เกี่ยวกับระดับหรือผลกระทบที่เกิดจากภัยพิบัติ และ 6) คำอื่นที่ไม่เกี่ยวข้อง โดยใช้วิธี Conditional Random Field (CRF) ในการทำการรู้จำชื่อเฉพาะด้วยโปรแกรมจาก Stanford University

4. ผลการดำเนินงาน

ตารางที่ 1 ผลการประเมินความถูกต้องในการจำแนกหมวดหมู่ข้อความข่าวสาร โดยประสิทธิภาพด้านความแม่นยำ (P) ร้อยละ 91.58 ประสิทธิภาพด้านการค้นคืน (R) ร้อยละ 84.65 และ F-measure (F) ร้อยละ 87.98

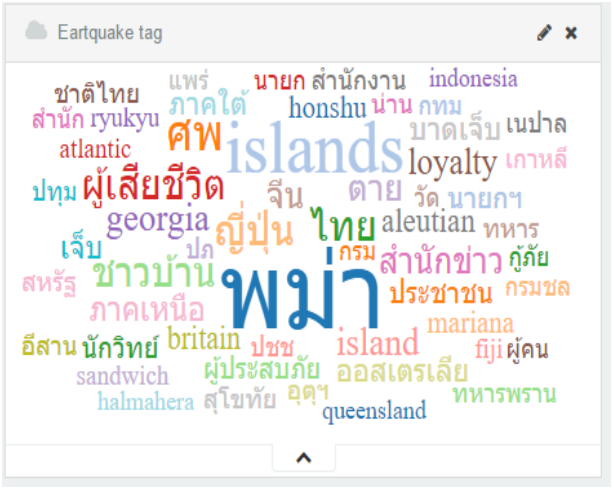
ตารางที่ 1: ประสิทธิภาพในการจำแนกหมวดหมู่

ชื่อเฉพาะ	P	R	F1
TYPE	0.9941	**0.9849**	**0.9895**
LOC	0.9384	0.8294	0.8805
ORG	0.8947	0.6623	0.7612
PER	0.9439	0.8211	0.8783
LEV	**1.0000**	0.8462	0.9167
OTR	0.8625	0.7995	0.8298
Total	**0.9158**	**0.8465**	**0.8798**

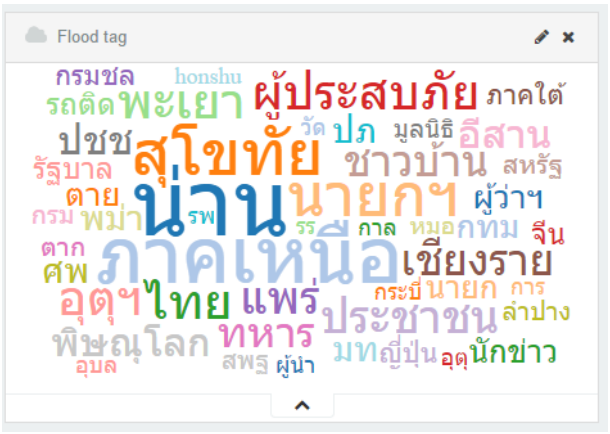
เมื่อนำแบบจำลองไปประยุกต์ใช้ในการสกัดชื่อเฉพาะออกจากข้อความที่ได้จาก Twitter ผ่านทาง Twitter API ผลที่ได้สามารถแยกคำที่เป็นชื่อเฉพาะในส่วนที่ต้องการออกมาได้เป็นอย่างดี และสามารถนำมาประยุกต์ใช้ในการแสดงความถี่ของคำ

[1] http://www.sansarn.com/lexto/

ที่เกิดขึ้นในช่วงเวลาปัจจุบันด้วย นอกจากนี้ยังได้นำเสนอการแสดงผลในรูปแบบ Tag Cloud ทำให้เห็นถึงสถานที่เกิดภัยธรรมชาติและกำลังมีการเผยแพร่ข้อมูลจากผู้ใช้งาน Twitter เรียงลำดับปริมาณการเผยแพร่ด้วยขนาดตัวอักษร ดังตัวอย่างจากภาพที่ 3 และ ภาพที่ 4



ภาพที่ 3: ตัวอย่างTag Cloud ขณะเกิดแผ่นดินไหวที่พม่า



ภาพที่ 4: Tag Cloud ขณะเกิดน้ำท่วมทางภาคเหนือของประไทย

5. สรุป

การสร้างแบบจำลองการรู้จำชื่อเฉพาะที่เกี่ยวกับภัยธรรมชาติ โดยใช้ข้อความจาก Twitter จำนวน 100,000 คำ ที่เกี่ยวข้องกับภัยธรรมชาติ และประเมินประสิทธิภาพของแบบจำลองโดยรวม ซึ่งผลของประสิทธิภาพด้านความแม่นยำร้อยละ 91.58 ประสิทธิภาพด้านการค้นคืนร้อยละ 84.65 และ F-measure ร้อยละ 87.98 แสดงให้เห็นว่าการสร้างแบบจำลองการรู้จำชื่อเฉพาะเกี่ยวกับภัยธรรมชาติจากข้อความทวิตเตอร์ มีประสิทธิภาพใกล้เคียงกับประสิทธิภาพของแบบจำลองการรู้จำชื่อเฉพาะที่เป็นภาษาไทยในด้านอื่น ๆ จึงเหมาะแก่การใช้เป็นแบบจำลองพื้นฐานสำหรับพัฒนาต่อเพื่อใช้ประโยชน์ในการคัดแยกสารสนเทศเกี่ยวกับภัยธรรมชาติจากข้อความทวิตเตอร์ภาษาไทยได้ต่อไป

กิตติกรรมประกาศ

งานวิจัยนี้ได้รับทุนสนับสนุนจากคณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

เอกสารอ้างอิง

[1] A. Kongthon, C. Haruechaiyasak, J. Pailai, and S. Kongyoung, "The role of Twitter during a natural disaster: Case study of 2011 Thai Flood", 2011.

[2] Lungkatoong, Y. & Sodanil, M., "Classification of open source earthquake disaster information in Thai language", 11th National Conference on Computing and Information Technology, .2015

[3] Z. Yanquan, Y. Lili, B. Van de Walle, and C. Han, "Classification of Microblogs for Support Emergency Responses: Case Study Yushu Earthquake in China", 2013, pp. 1553-1562.

[4] X. Liu, M. Zhou, F. Wei, Z. Fu, and X. Zhou, "Joint inference of named entity recognition and normalization for tweets", 2012, pp. 526-535.

[5] M. B. Habib, and M. van Keulen, "NEED4Tweet: a Twitterbot for tweets named entity extraction and disambiguation" 2015.

[6] นัชชา ธิระสาโรช "การรู้จำชื่อเฉพาะภาษาไทยการใช้แบบจำลองคอนดิชันนอลแรนด : อมฟิลด์ส" 2553.

[7] ยศวัธน์ แสนสิงห์ "การรู้จำชื่อเฉพาะภาษาไทยโดยใช้แบบจำลองฮิดเดนมาร์คอฟ" 2552.

[8] D. Küçük, and R. Steinberger, "Experiments to improve named entity recognition on Turkish tweets" *arXiv preprint arXiv:1410.8668*, 2014.

[9] W. Che, M. Wang, C. D. Manning, and T. Liu, "Named Entity Recognition with Bilingual Constraints", 2013, pp. 52-62

การพยากรณ์ประสิทธิภาพของผู้เรียนโดยใช้เทคนิควิธีการทำเหมืองข้อมูล
Predicting Student's Data for Performance by Data Mining

โชติกา เอี่ยมปริก (Chotika Iamprik)[1] และขวัญทัย สุดาเดช (Kwanruetai Sudadet)[2]
ภาควิชาวิทยาการคอมพิวเตอร์และสารสนเทศ คณะวิทยาศาสตร์ ศรีราชา มหาวิทยาลัยเกษตรศาสตร์ วิทยาเขตศรีราชา
[1]koysmilemelody11@gmail.com, [2]kwanruetaii2537@gmail.com

บทคัดย่อ

การวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาประสิทธิภาพของผู้เรียนโดยใช้เทคนิคการจำแนกข้อมูลหรือที่เรียกว่า classification ซึ่งเป็นหนึ่งในวิธีการทำเหมืองข้อมูล โดยใช้ปัจจัยการพยากรณ์คือ ผลการเรียนในวิชาบังคับตามหลักสูตรผลการพยากรณ์จะอยู่ในรูปแบบการจำแนกกลุ่มโดยมีกลุ่มดังนี้ 1.) ได้เกียรตินิยมอันดับที่ 1 2.) ได้เกียรตินิยมอันดับที่ 2 3.) สำเร็จการศึกษาภายใน 4 ปี 4.) สำเร็จการศึกษาภาภายใน 5-8 ปี และ 5.) ไม่สำเร็จการศึกษา

งานวิจัยชิ้นนี้ผู้วิจัยได้รวบรวมข้อมูล ตั้งแต่ปีการศึกษา 2553 – 2557 โดยแบ่งเป็นตามหลักสูตร(ปีที่ปรับปรุง 2550) หลักสูตรวิทยาศาสตร์บัณฑิต สาขาวิชาวิทยาการคอมพิวเตอร์ (ภาคปกติ/พิเศษ) จำนวนทั้งสิ้น 338 คน

จากผลการวิจัยสามารถสรุปได้ว่า การพยากรณ์ด้วยเทคนิคต้นไม้การตัดสินใจ(Decision Tree) โดยใช้อัลกอริธึม C4.5 (J48) มีประสิทธิภาพในการจำแนกข้อมูลโดยรวมได้ดีที่สุด โดยมีค่าความถูกต้องในการจำแนกกลุ่มของผู้เรียนที่ระดับสูงสุดเท่ากับ 35.80% และมีค่าความถูกต้องเพิ่มขึ้นเมื่อมีเกรดและเกรดเฉลี่ยของผู้เรียนตามลำดับแผนการเรียน ที่ระดับความถูกต้องในการจำแนกของผู้เรียน 85.80%

คำสำคัญ: การพยากรณ์ ประสิทธิภาพของผู้เรียน เทคนิควิธีการเหมืองข้อมูล การจำแนกประเภทข้อมูล ต้นไม้การตัดสินใจ

Abstract

This research object is to develop predicting learner's efficiency by using classification which is a kind of Data Mining Technique with predicting factors which are the results in required subject. The results are classified in groups following 1) 1st honor 2) 2nd honor 3) graduated with in 4 years 4) graduated within 5-8 years and 5) not graduated.

The data have been collected since education year 2010 until 2014 followed by a course (renovated in 2007). Bachelor of Science Computer Science (Regular / Special) A total of 338 people.

As the research results which are able to conclude that a decision tree prediction with algorithm C4.5 (J48) is the best to classification gross data, with the accuracy in identifying group of learner at the highest level 35.80% and accuracy increased with the grade and GPA. And the study by plan the level of accuracy in the classification of the study is 85.80%.

Keyword: Forecasting, Performance of the Students, Data Mining Techniques, Classification Data, Decision Tree.

1. บทนำ

การเก็บข้อมูลทางการศึกษา ถูกจัดเก็บจนมีปริมาณเพิ่มมากขึ้นทั้งในรูปแบบแฟ้มเอกสารและ รูปแบบเอกสารอิเล็กทรอนิกส์ที่ถูกจัดเก็บในรูปแบบของฐานข้อมูล หรือคลังข้อมูล [13] ดังนั้นงานวิจัยฉบับนี้ได้ทำการศึกษาผลการเรียนของผู้เรียนเพื่อนำมาแก้ไขประสิทธิภาพการเรียนรู้ของผู้เรียนให้มีเป้าหมายและจุดประสงค์ที่ชัดเจนที่จะทำให้ผู้เรียนมีผลการเรียนที่ดีขึ้นและตรงตามจุดประสงค์ของผู้เรียน งานวิจัยฉบับนี้ใช้ข้อมูลนิสิตคณะวิทยาศาสตร์ศรีราชา (Faculty of Science at Sriracha) โดยเลือกสาขาวิทยาการคอมพิวเตอร์(ภาคปกติ,ภาคพิเศษ) ในการวิเคราะห์ข้อมูล จัดการกับข้อมูล เพื่อทำการสร้างโมเดลในการหาประสิทธิภาพของผู้เรียนโดยการใช้เทคนิควิธีการทำ

เหมืองข้อมูลซึ่งได้ใช้การจำแนกกลุ่ม (Classification)ในการทำการพยากรณ์โดยมีหลักการพิจารณาจาก Target Class คือ 1.) ได้เกียรตินิยมอันดับที่ 1 2.) ได้เกียรตินิยมอันดับที่ 2 3.) สำเร็จการศึกษาภายใน 4 ปี 4.) สำเร็จการศึกษาภาภายใน 5-8 ปี และ 5.) ไม่สำเร็จการศึกษา โดยใช้ระเบียบการศึกษาและข้อบังคับของมหาวิทมหาวิทยาลัยเกษตรศาสตร์เป็นแนวทางการนำไปพยากรณ์ประสิทธิภาพการเรียนรู้ของผู้เรียนโดยใช้เทคนิควิธีการทำเหมืองข้อมูล (Data Mining) [7] เข้ามาช่วย โดยผู้เรียนสามารถพยากรณ์การทำนายประสิทธิภาพของตนเองได้ว่าตนเองอยู่ในประสิทธิภาพระดับใดของผลการพยากรณ์ เพื่อนำไปสู่การตัดสินใจให้นิสิตวางแผนการเรียนตามเป้าหมายของตนเองได้อย่างมีประสิทธิภาพยิ่งขึ้นโดย ใช้วิธีการทำเหมืองข้อมูล (Data Mining)

2. วรรณกรรมที่เกี่ยวข้อง

2.1 ประสิทธิภาพของผู้เรียน (The performance of the students)

ทุกระบบการศึกษาประกอบด้วยคุณภาพและความสามารถของผู้เรียน ได้รับการประเมินโดยให้เป็นเกรดในแต่ละวิชาเกี่ยวข้องกับการใช้เวลา ความพยายามที่นำมาใช้ เพื่อให้บรรลุจุดมุ่งหมายหรือวัตถุประสงค์อย่างใดอย่างหนึ่ง โดยเน้นไปที่ความสามารถความพยายามที่จะทำให้เกิดผลลัพธ์ที่มีประสิทธิภาพ การทดสอบประสิทธิภาพการศึกษาของผู้เรียนจะต้องทดสอบว่ามันคืออะไร ผู้เรียนสามารถรู้ประสิทธิภาพของตนเองได้ใช้เกณฑ์ในการพยากรณ์ประสิทธิภาพของผู้เรียนด้วยเกรดของนักเรียนในแต่ละหลักสูตรจะแทนได้ด้วยเกรดที่เป็นตัวอักษร[11]

2.2 การทำเหมืองข้อมูล (Data Mining)

การทำเหมืองข้อมูล (Data Mining) คือ [7] กระบวนการค้นหาสารสนเทศหรือข้อความรู้ที่อยู่ในฐานข้อมูลขนาดใหญ่ที่ซับซ้อนเพื่อนำข้อความรู้ที่ได้ไปใช้ประโยชน์ในการตัดสินใจ สารสนเทศที่ได้อาจนำมาสร้างการพยากรณ์หรือสร้างตัวแบบสำหรับการจำแนกกลุ่ม หรือแสดงความสัมพันธ์ระหว่างหน่วยต่างๆ หรือให้ข้อสรุปของสาระในฐานข้อมูล จะประกอบไปด้วยการนำกระบวนการทางสถิติและการเรียนรู้ผ่านระบบคอมพิวเตอร์ เพื่อสร้างตัวแบบ กฎเกณฑ์ รูปแบบ การพยากรณ์และข้อความรู้ จากฐานข้อมูลขนาดใหญ่ และเพื่อที่จะสามารถสกัดความรู้ที่อยู่ในข้อมูลขนาดใหญ่ได้

2.3 เทคนิคการจำแนกประเภทข้อมูล (Data Classification)

เทคนิคการจำแนกประเภทข้อมูล (Data Classification) [13] เป็นเทคนิคหนึ่งที่สำคัญของการสืบค้นความรู้บนฐานข้อมูลขนาดใหญ่ (Knowledge Discovery from very large Database: KDD) หรือ ดาต้าไมน์นิง (Data Mining) เทคนิคการจำแนกประเภทข้อมูลเป็นกระบวนการสร้างโมเดลจัดการข้อมูลให้อยู่ในกลุ่มที่กำหนดมาให้จากกลุ่มตัวอย่างข้อมูลที่เรียกว่าข้อมูลสอนระบบ (training data) ที่แต่ละแถวของข้อมูลประกอบด้วยฟิลด์หรือแอทริบิวท์จำนวนมาก แอทริบิวท์นี้อาจเป็นค่าต่อเนื่อง (continuous) หรือค่ากลุ่ม (categorical) โดยจะมีแอทริบิวท์แบ่ง (classifying attribute) ซึ่งเป็นตัวบ่งชี้คลาสของข้อมูล จุดประสงค์ของการจำแนกประเภทข้อมูลคือการสร้างโมเดลการแยกแอทริบิวท์หนึ่งโดยขึ้นกับแอทริบิวท์อื่น โมเดลที่ได้จากการจำแนกประเภทข้อมูลจะทำให้สามารถพิจารณาคลาสในข้อมูลที่ยังมิได้แบ่งกลุ่มในอนาคตได้

2.4 ต้นไม้ตัดสินใจ (Decision Tree)

Decision Tree เป็นการนำข้อมูลมาสร้างแบบจำลองการพยากรณ์ในรูปแบบการสร้างต้นไม้ และมีการทำงานแบบ Supervised Learning [8] คือการเรียนรู้ของโมเดลแบบมีครูสอนสามารถสร้างแบบจำลองการจัดหมวดหมู่ได้จากกลุ่มตัวอย่างข้อมูลที่กำหนดไว้ล่วงหน้า และพยากรณ์กลุ่มของรายการที่ยังไม่เคยนำมาจัดหมวดหมู่ได้ด้วยรูปแบบของ Tree โดยโครงสร้างประกอบไปด้วย Root Node Child และ Leaf Node อัลกอริทึมที่ใช้ในการสร้าง Decision Tree ได้แก่ ID3 Algorithm, C4.5 Algorithm, C5.0 Algorithm, CART Algorithm [9]

2.5 นาอีฟเบย์ (Naive Bayes)

ทฤษฎีของเบย์ (Bayes' Theorem) ถูกพัฒนาขึ้นโดย Thomas Bayes ใช้หลักการของความน่าจะเป็นแบบมีเงื่อนไขมาพัฒนาเป็นทฤษฎีบทดังกล่าวโดยมีรายละเอียดดังต่อไปนี้ ถ้าให้เอกภพสัมพัทธ์ U ประกอบด้วยเหตุการณ์ที่ไม่สามารถเกิดขึ้นได้พร้อมกัน จำนวน k เหตุการณ์คือ A1, A2, A3, …, Ak ดังรูปที่ 7.1 และให้ E เป็นเหตุการณ์หนึ่ง ในปริภูมิตัวอย่างที่เกิดจากการทดลองเดียวกันนี้และต้องเป็นส่วนหนึ่ง ของ Ai (i = 1,

2, 3, …, k) จะสามารถคำนวณความน่าจะเป็นแบบมีเงื่อนไขของเหตุการณ์หนึ่ง ใน Ai เมื่อเหตุการณ์ เกิดขึ้นแล้วได้สมการที่ (7.1) [6]

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

ภาพที่ 1 สมการที่(7.1)

2.6 ซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine (SVM)

เกิดจากการที่นำค่าของกลุ่มข้อมูลมาวางลงในฟีเจอร์สเปซ (Feature Space) จากนั้นจึงหาเส้นที่ใช้แบ่งข้อมูลทั้งสองออกจากกัน โดยจะสร้างเส้นแบ่ง (Hyper plane) ที่เป็นเส้นตรงขึ้นมา และเพื่อให้ทราบว่าเส้นตรงที่แบ่งสองกลุ่มออกจากกันเส้นตรงใดเป็นเส้นที่ดีที่สุด สำหรับรากฐานเดิมของ Support Vector Machine ถูกนำมาใช้กับข้อมูลที่เป็นเชิงเส้น แต่ในความเป็นจริงแล้วข้อมูลที่นำมาใช้ในระบบการสอนให้ระบบเรียนรู้ส่วนใหญ่มักเป็นข้อมูลแบบไม่เป็นเชิงเส้น ซึ่งสามารถแก้ปัญหาดังกล่าวด้วยการนำ Kernel Function มาใช้การจำแนกข้อมูลบนระนาบหลายมิติ จะใช้ส่วนการเลือกที่มีความเหมาะสมที่สุดเรียกว่า โครงสร้างในการคัดเลือก (feature selection) ซึ่งโครงสร้างในการคัดเลือกมาจากข้อมูลที่สอนให้ระบบเรียนรู้ [10]

2.7 งานวิจัยที่เกี่ยวข้อง

งานวิจัยที่นำเทคนิคการทำเหมืองข้อมูลมาช่วยสร้างโมเดลให้กับชุดข้อมูล โดยอาจใช้เพียงเทคนิคเดียว หรือมากกว่า 1 เทคนิคร่วมกันในการจำแนกข้อมูลที่อาศัยเทคนิคต้นไม้ตัดสินใจ [1], [2], [3] หรือการจำแนกข้อมูลแบบซัพพอร์ตเวกเตอร์เมชชีน [10] ซึ่งมีการเปรียบเทียบประสิทธิภาพการทำงานร่วมกันกับการใช้เพียงเทคนิคเดียว และจะเห็นได้ว่าการจำแนกข้อมูลนั้นหากใช้เพียงเทคนิคที่เหมาะสมกับข้อมูลจะส่งผลให้ผู้วิจัยสามารถทำงานได้มีประสิทธิภาพมากยิ่งขึ้น ดังนั้น เพื่อให้ทราบว่าเทคนิคใดเหมาะสมกับชุดข้อมูลที่นำมาใช้สร้างโมเดลมากที่สุด ผู้วิจัย [4], [5], [13] จึงต้องทำการเปรียบเทียบประสิทธิภาพระหว่างเทคนิคต่างๆกับชุดข้อมูลที่นำมาใช้ทำวิจัย ซึ่งค่าที่ผู้วิจัยนิยมนำมาใช้เปรียบเทียบประสิทธิภาพ คือ ค่าความถูกต้อง เป็นค่าที่ผู้วิจัยนำมาเปรียบเทียบเพื่อหาประสิทธิภาพความเหมาะสมของการเลือกใช้เทคนิคต่างๆ ในการนำมาสร้างโมเดล ได้ตรงกับความเหมาะสมของชุดข้อมูลมากที่สุดและผู้วิจัยอาจต้องใช้อัลกอริทึม [6], [8], [10], [11] ช่วยในการลดขนาดของข้อมูล (Feature Selection) ในการสร้างโมเดลในการหาคำตอบอย่างมีประสิทธิภาพ

3. วิธีการดำเนินการวิจัย

3.1 ศึกษาปัญหาและวิเคราะห์ข้อมูล

วิจัยฉบับนี้ผู้วิจัยได้รวบรวมข้อมูลการเรียนการสอนตามรายวิชา ในกลุ่มรายวิชาที่ลงทะเบียนเรียนตั้งแต่ผู้เรียนเริ่มทำการลงทะเบียนเรียน การให้คะแนนเกรดในแต่ละรายวิชา เก็บข้อมูลปีการศึกษา 2553 – 2557 โดยแบ่งเป็นตามหลักสูตร ใช้หลักสูตรปีที่ปรับปรุง 2550 วิทยาศาสตร์บัณฑิต สาขาวิชาวิทยาการคอมพิวเตอร์(ภาคปกติ/พิเศษ) มหาวิทยาลัยเกษตรศาสตร์ วิทยาเขตศรีราชา จำนวนทั้งสิ้น 338 คน จากการศึกษาพบว่า ผู้เรียนที่อยู่ในระหว่างการเรียนการสอนในแต่ละภาคเรียนการศึกษา แต่ละชั้นปีปัจจัยใดเป็นปัจจัยที่สำคัญในการวัดประสิทธิภาพของผู้เรียนที่เกิดขึ้นในระหว่างการศึกษาโดยผู้เรียนไม่สามารถวิเคราะห์ ประเมินประสิทธิภาพตนเองได้ ดังนั้นผู้วิจัยนำปัญหาที่เกิดขึ้นกับผู้เรียนมาแก้ไขด้วยวิธีที่ส่งผลต่อการพยากรณ์ประสิทธิภาพของผู้เรียนในอนาคต พิจารณาข้อมูลที่ได้จากการเรียนการสอน สามารถสรุปลักษณะข้อมูลที่ใช้ได้ข้อมูลที่ใช้วิเคราะห์งานวิจัยแบ่งออกเป็น 4 ส่วน ดังนี้คือ 1. ข้อมูลทั่วไปของผู้เรียน เป็นข้อมูลรหัสของผู้เรียน โดยวิจัยฉบับนี้ใช้ข้อมูลรหัสของผู้เรียนที่อยู่ในหลักสูตรการศึกษา 2. ข้อมูลที่ได้จากการเรียนของผู้เรียน โดยแบ่งตามระดับคะแนนตามเกรด (ดังตารางที่ 1) 3. คำนวณเกรดจากการเรียนของผู้เรียน โดยใช้ระเบียบการศึกษาและข้อบังคับบางประการของมหาวิทยาลัยเกษตรศาสตร์ 4. ระดับประสิทธิภาพเป็นดัชนีวัดประสิทธิภาพของแต่ละคลาสคำตอบ (Target Class)

ตารางที่ 1 ระดับคะแนนที่มีแต้ม

Parentage Grade	Letter Grade	Deseription
80-100	A	ดีเยี่ยม
75-79	B+	ดีมาก
70-69	B	ดี
65-69	C+	ค่อนข้างดี
60-64	C	พอใช้
55-59	D+	อ่อน
50-54	D	อ่อนมาก
0-49	F	ตก

3.2 การเตรียมข้อมูล (Data Preparation)

วิจัยฉบับนี้มุ่งเน้นการสร้างแบบจำลองเพื่อใช้สำหรับการพยากรณ์ประสิทธิภาพของผู้เรียน โดยใช้เทคนิควิธีการเหมืองข้อมูล 3 เทคนิค ได้แก่ ต้นไม้การตัดสินใจ (Decision Tree), นาอีฟเบย์ (Naïve Bays) , ซัพพอร์ตเวคเตอร์แมชชีน (Sequential minimal optimization) นำผลลัพธ์ค่าการพยากรณ์ประสิทธิภาพที่ได้จากการสร้างแบบจำลองด้วยเทคนิคต่างๆ เพื่อเปรียบเทียบประสิทธิภาพโดยรวมใช้ค่าความถูกต้องในการจำแนกกลุ่ม (Accuracy) เพื่อเป็นดัชนีวัดประสิทธิภาพของอัลกอริทึมและค่าเฉลี่ยโดยรวม (F-measure) เป็นดัชนีวัดประสิทธิภาพของผู้เรียนในแต่ละคลาสคำตอบ (Target Class) ได้จากแบบจำลองในการพยากรณ์ความถูกต้อง โดยผู้วิจัยทำการแบ่งตามแผนการเรียนโดยเลือกวิชาบังคับในแต่ละภาคเรียนการศึกษามีรายละเอียดแอทริบิวต์ดังตารางที่2

ตารางที่ 2 รายละเอียดชุดข้อมูล

ลำดับ	ชื่อแอทริบิวต์	ค่าตัวแปล	รายละเอียด	ประเภทข้อมูล
1	Student.ID	อักษร	รหัสประจำตัวนิสิต	ข้อมูลผู้เรียน
2 ถึง 37	Ex.SC.01371111	อักษร	รหัสวิชา	เกรดการเรียน
38	GPA	จำนวน	เกรดเฉลี่ย	เกรดการเรียนของผู้เรียนใช้กฎข้อบังคับของมหาวิทยาลัยเกษตรศาสตร์
39	Performance	ระดับ	5 ระดับ	ประสิทธิภาพ

3.3 การวิเคราะห์ข้อมูล (Data Analysis)

ผู้วิจัยแบ่งข้อมูลออกเป็น 8 ชุดข้อมูลประกอบไปด้วย (Feature) ซึ่งเป็นไปตามแผนการเรียนการสอน เมื่อพิจารณาตารางที่ 3 สามารถยกตัวอย่างได้ดังนี้ ชุดข้อมูลที่ 1 ประกอบไปด้วยลำดับตัวแปลที่ 1 และ39 (ข้อมูลผู้เรียน) ชุดข้อมูลที่ 2 ประกอบไปด้วยลำดับตัวแปลที่ 2 ถึง 8,38 และ39 (ปีการศึกษาภาคเรียนที่ 1 ภาคต้น) ชุดข้อมูลที่ 3 ประกอบไปด้วยลำดับตัวแปลที่ 2 ถึง 14,38 และ39 (ปีการศึกษาภาคเรียนที่ 1 ภาคปลาย) ชุดข้อมูลที่ 4 ประกอบไปด้วยลำดับตัวแปลที่ 2 ถึง 20,38 และ39 และ(ปีการศึกษาภาคเรียนที่ 2 ภาคต้น) ชุดข้อมูลที่ 5 ประกอบไปด้วยลำดับตัวแปลที่ 2 ถึง 27,38 และ39(ปีการศึกษาภาคเรียนที่ 2 ภาคปลาย) ชุดข้อมูลที่ 6 ประกอบไปด้วยลำดับตัวแปลที่ 2 ถึง 31,38 และ39(ปีการศึกษาภาคเรียนที่ 3 ภาคต้น) ชุดข้อมูลที่ 7 ประกอบไปด้วยลำดับตัวแปลที่ 2 ถึง 36,38 และ39 (ปีการศึกษาภาคเรียนที่ 3 ภาคปลาย) ชุดข้อมูลที่ 8 ประกอบไปด้วยลำดับตัวแปลที่ 2 ถึง 39 (ปีการศึกษาภาคเรียนที่ 4 ภาคต้น)

ผู้วิจัยใช้โปรแกรม WEKA 3.6 ในการสร้างแบบจำลองและการวิเคราะห์ข้อมูลโดยผู้วิจัยใช้ 10-fold Cross – validation คือ การแบ่งข้อมูลออกเป็น 10 ส่วน โดยที่แต่ละส่วนมีจำนวนเท่ากันหลังจากนั้นข้อมูลหนึ่งส่วนจะใช้เป็นตัวทดสอบประสิทธิภาพขอโมเดล ทำวนไปเช่นนี้ไปจนครบจำนวนที่แบ่งไว้ (ดร.เอกสิทธิ์,2014) นำไปใช้กับเทคนิค ต้นไม้การตัดสินใจ (Decision Tree), นาอีฟเบย์ (Naïve Bays) , ซัพพอร์ตเวคเตอร์แมชชีน (Sequential minimal optimization) จากนั้นนำข้อมูลที่วัดได้ไปสร้างแบบจำลอง เพื่อนำไปใช้สำหรับการพยากรณ์ประสิทธิภาพของผู้เรียนโดยใช้เทคนิควิธีการทำเหมืองข้อมูล

4.ผลการดำเนินงาน

จากการสร้างแบบจำลองและการวิเคราะห์ข้อมูลการพยากรณ์ประสิทธิภาพของผู้เรียนโดยใช้เทคนิควิธีการทำเหมืองข้อมูล ด้วยโปรแกรม WEKA 3.6 ตามวิธีการดำเนินการวิจัยและคำนวณค่าความถูกต้องของการจำแนกกลุ่ม (Accuracy) และค่าเฉลี่ยประสิทธิภาพโดยรวม (F-measure) จำนวน 8 ชุดข้อมูลตามขั้นตอนการวิจัยผลการวิเคราะห์สามารถแบ่งออกเป็น 2 ส่วนดังนี้

4.1 การเปรียบเทียบประสิทธิภาพของแบบจำลองที่ได้จากอัลกอริทึม

จากลำดับ 8 ชุดข้อมูลที่ผู้วิจัยสร้างขึ้นมีจำนวน 39 ลำดับแอทริบิวต์ แสดงให้เห็นถึงการพยากรณ์ประสิทธิภาพของ

ผู้เรียนในแต่ละหลักสูตรโดยแบ่งตามแผนการเรียนในแต่ละภาคการศึกษาตลอกการจบหลักสูตร ผู้วิจัยได้นำชุดข้อมูลทั้งหมดไปทดสอบกับอัลกอริทึมจำนวน 3 อัลกอริทึม ด้วยโปรแกรม WEKA 3.6 เพื่อเปรียบเทียบประสิทธิภาพของแบบจำลองที่ได้จากอัลกอริทึมโดยได้ผลการทดลองได้ดังต่อไปนี้

ตารางที่ 3 แสดงรายละเอียดการเปรียบเทียบประสิทธิภาพจากอัลกอริทึม

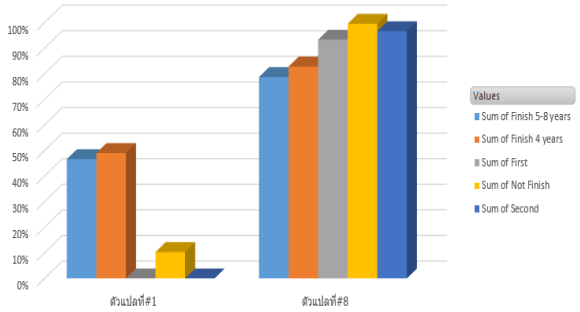
ชุดข้อมูล		ลำดับแอทริบิวต์	ค่าความถูกต้องของการจำแนกกลุ่ม (% Accuracy)			
ลำดับ	จำนวน		J48	Naïve Bays	SMO	Avg.
#1	2	1และ39 (ข้อมูลทั่วไปของผู้เรียน)	35.80	39.35	**39.35**	38.17
#2	9	2ถึง8+38+39 (ปีการศึกษาที่ 1 ภาคเรียนที่1)	**82.54**	81.95	71.01	78.50
#3	15	2ถึง14+38+39 (ปีการศึกษาที่ 1 ภาคเรียนที่2)	**79.88**	75.44	72.19	75.84
#4	21	2ถึง20+38+39 (ปีการศึกษาที่ 2 ภาคเรียนที่1)	**80.47**	76.04	74.56	77.02
#5	28	2ถึง27+38+39 (ปีการศึกษาที่ 2 ภาคเรียนที่2)	**81.07**	78.11	76.92	78.70
#6	32	2ถึง31+38+39 (ปีการศึกษาที่ 3 ภาคเรียนที่1)	**82.25**	80.18	78.99	80.47
#7	37	2ถึง36+38+39 (ปีการศึกษาที่ 3 ภาคเรียนที่2)	78.70	**80.47**	80.18	79.78
#8	38	2ถึง39 (ปีการศึกษาที่ 4 ภาคเรียนที่1)	**85.80**	80.77	83.14	83.23

จากตารางที่ 3 การเปรียบเทียบประสิทธิภาพอัลกอริทึมสามารถอธิบายได้ดังต่อไปนี้ ในลำดับชุดข้อมูลที่ #1 ข้อมูลทั่วไปของผู้เรียนมีค่าความถูกต้องในการจำแนกกลุ่ม (Accuracy) ประสิทธิภาพของผู้เรียนอยู่ในเกณฑ์ที่ต่ำ โดยมีค่าเฉลี่ยความถูกต้องของการจำแนกกลุ่มทั้ง 3 อัลกอริทึมที่ระดับ 39.35% ในลำดับชุดข้อมูลที่ #2 ถึง #8 มีค่าความถูกต้องของการจำแนกกลุ่มที่เพิ่มขึ้นตลอดจนจบกลักสูตรปีที่ปรับปรุง 2550 ได้แก่ อัลกอริทึม ต้นไม้การตัดสินใจ (Decision Tree) อยู่ที่ 85.80%

ตารางที่ 4 การเปรียบเทียบประสิทธิภาพของคลาสคำตอบแต่ละคลาสคำตอบ

ลำดับ	Finish 5-8 years	Finish 4 years	First	Not Finish	Second
#1	46.6%NB	49%J48	0%J48	10.3%J48	0%J48
#2	73.7%NB	76.4%NB	93.3%J48	99.5%J48	96.8%NB
#3	69.4%NB	72.3%NB	93.3%J48	99.5%J48	96.6%J48
#4	70%SMO	73.4%NB	93.3%J48	99.5%J48	96.6%J48
#5	72.4%SMO	73.8%SMO	93.3%J48	99.5%J48	96.6%J48
#6	74.2%NB	75.8%J48	93.3%J48	99.5%J48	96.6%J48
#7	75.7%NB	76.3%SMO	93.3%J48	99.5%J48	96.6%J48
#8	78.7%SMO	82.8%SMO	93.3%J48	99.5%J48	96.6%J48

เมื่อเปรียบเทียบประสิทธิภาพของคลาสคำตอบแต่ละคลาสคำตอบโดยหาค่าเฉลี่ยประสิทธิภาพโดยรวม (F-measure) ดังตารางที่ 4 แสดงให้เห็นว่า ชุดข้อมูลลำดับที่ #1 โดยมีค่าประสิทธิภาพแยกแต่ละคลาสคำตอบสูงสุดอยู่ที่ระดับ 49% สำหรับประสิทธิภาพของผู้เรียนในลำดับ Finish 4 year โดยอัลกอริทึม J48 ลำดับต่อมาได้แก่ ลำดับ Finish 5-8 year โดยอัลกอริทึม NB ทีระดับ 46.6% และ ลำดับ Not Finish โดยอัลกอริทึม J48 ทีระดับ 10.3% และเมื่อพิจารณาเกรดที่ได้ของผู้เรียนเมื่อผู้เรียนเรียนครบตามหลักสูตรการศึกษาในวิชาบังคับตลอดระยะเวลาการศึกษา ในชุดลำดับที่ (#2 ถึง #8) พบว่าอัลกอริทึมในการพยากรณ์ประสิทธิภาพของผู้เรียนที่มีประสิทธิภาพสูงสุดที่ระดับ 99.5% สำหรับประสิทธิภาพของผู้เรียนในลำดับ Not Finish โดยใช้อัลกอริทึม J48 เป็นอัลกอริทึมที่มีประสิทธิภาพการพยากรณ์สูงที่สุด ดังตัวอย่างภาพที่ 1



ภาพที่ 1: แสดงประสิทธิภาพของผู้เรียนทั้ง 5 Target Class โดยการจำแนกสีเป็นแผนภูมิแท่ง

5.สรุป

งานวิจัยฉบับนี้ได้ทำการทดสอบประสิทธิภาพของอัลกอริทึมที่ใช้ในการสร้างแบบจำลองเพื่อในการพยากรณ์ประสิทธิภาพของผู้เรียนโดยใช้เทคนิควิธีการทำเหมืองข้อมูลสำหรับการจำแนกประเภทข้อมูลและการทำนาย (Classification and Prediction) ผู้วิจัยเลือกใช้อัลกอริทึมในการทำเหมือนข้อมูลจำนวน 3 อัลกอริทึมได้แก่ ต้นไม้การตัดสินใจ (Decision Tree), นาอีฟเบย์ (Naïve Bays) , ซัพพอร์ตเวคเตอร์แมชชีน (Sequential minimal optimization) กับชุดข้อมูลผู้เรียนจำนวนทั้งสิ้น 338 คน ในขั้นตอนการหาประสิทธิภาพในแต่ละอัลกอริทึม ผู้วิจัยใช้ 10-fold Cross –

validation เพื่อใช้สำหรับวัดค่าประสิทธิภาพของแบบจำลองสำหรับค่าความถูกต้องของการจำแนกกลุ่ม (Accuracy) เป็นดัชนีชี้วัดประสิทธิภาพโดยรวมของอัลกอริทึม และใช้ค่าเฉลี่ยประสิทธิภาพโดยรวม (F-measure) เพื่อเป็นดัชนีชี้วัดประสิทธิภาพของแต่ละคลาสคำตอบ (Target Class) ที่ได้จากแบบจำลองการพยากรณ์ของผู้เรียน

ผลการวิจัยฉบับนี้พบว่า ต้นไม้การตัดสินใจ (Decision Tree) มีประสิทธิภาพการจำแนกข้อมูลโดยรวมสูงที่สุดโดยมีค่าความถูกต้องในการจำแนกกลุ่มผู้เรียน (Accuracy) สูงสุดที่ระดับ 85.80%

ผลการเปรียบเทียบประสิทธิภาพของคลาสคำตอบแต่ละคลาสคำตอบ (Target Class) โดยหาค่าเฉลี่ยประสิทธิภาพโดยรวม (F-measure) พบว่าการเรียนในแต่ละปีการศึกษา และแต่ละภาคเรียนการศึกษามีค่าเฉลี่ยของประสิทธิภาพโดยรวมที่แตกต่างกัน ดังนั้นเพื่อให้ได้ค่าการพยากรณ์ของผู้เรียนที่ดีจะต้องคำนึงถึงอัลกอริทึม เพื่อค้นหาอัลกอริทึมที่เหมาะสมในการพยากรณ์ของผู้เรียนโดยใช้เทคนิควิธีการทำเหมืองข้อมูล

เอกสารอ้างอิง

[1] Cuco, Carlos Herrero, et al. "An empirical analysis of the influence of classmates on the academic performance." Proceeding ITiCSE'13 Proceedings of the 18th ACM conference on Innovation and technology in computer science education. Association for Computing Machinery (ACM), 2013.

[2] Nasiri, Mahdi, and Behrouz Minaei. "Predicting GPA and academic dismissal in LMS using educational data mining: A case mining." E-Learning and E-Teaching (ICELET), 2012 Third International Conference on. IEEE, 2012.

[3]Yadin, Aharon. "Using unique assignments for reducing the bimodal grade distribution." ACM Inroads 4.1 (2013): 38-42.

[4] เสก สรรค์ วิลัย ลักษณ์, วิภา เจริญ ภัณฑารักษ์, และดวงดาว วิชาดากุล. "การใช้เทคนิคกาทำเหมืองข้อมูลเพื่อพยากรณ์ ผลการเรียนของนักเรียนโรงเรียนสาธิตแห่งมหาวิทยาลัย เกษตรศาสตร์ วิทยาเขตกำแพงแสน ศูนย์วิจัยและ พัฒนาการ ศึกษา." Veridian E-journal Science and Technology Silpakorn University 2.2 (2015): 1-17.

[5] ปรียารัตน์ นาคสุวรรณ์. "การทำนายผลสัมฤทธิ์ในการเรียน วิชาสถิติเบื้องต้นของนิสิตปริญญาตรี มหาวิทยาลัยบูรพา." วารสารวิทยาศาสตร์บูรพา Burapha Science Journal 17.1 (2015): 59-68.

[6] ดร.นพมาศ ปักเข็ม. การเรียนรู้แบบเบย์ (Bayesian Learning).[Online]: สืบค้นเมื่อวันที่ 4 กุมภาพันธ์ 2560 จาก http://mis.csit.sci.tsu.ac.th/noppamas/download/DataMining/DataMiningCh7V1.pdf

[7] สุชาดา กีระนันทน์. (2545). Data Mining. สัมมนาหนึ่งทศวรรษไอทีจุฬาฯ: การเพิ่มศักยภาพระบบการจัดการด้วยเทคโนโลยีสารสนเทศ. [Online]: สืบค้นเมื่อวันที่ 4 กุมภาพันธ์ 2560 จาก http://home.kku.ac.th/wichuda/Knowlage/6 DataMining/Datamining_Suchada.pdf

[8] เดชธรรม ศิริ และพยุง มีสัจ. "การจำแนกข้อมูลด้วยวิธีแบบร่วมกันตัดสินใจจากพื้นฐานขอ เทคนิคต้นไม้ตัดสินใจเทคนิคโครงข่ายประสาทเทียมและเทคนิคซัพพอร์ตเวกเตอร์แมชชีนร่วมกับการเลือกตัวแทนที่เหมาะสมด้วยขั้นตอนวิธีเชิงพันธุกรรม." Journal of King Mongkut's University of Technology North Bangkok (วารสารวิชาการ พระจอมเกล้า พระนครเหนือ) 21.2 (2011): 293-303.

[9] วิวัฒน์ ชินนาทศิริกุล.การทำเหมืองข้อมูลแบบจำลอง(Classification Decision Tree).[Online]: สืบค้นเมื่อวันที่ 4 กุมภาพันธ์ 2560 จาก http://digi.library.tu.ac.th/thesis/it/0966/title-biography.pdf

[10] จิรา แก้วสุวรรณ์."การตวจจับและการแก้ไขการวางตัวของภาพโดยใช้ซัพพอร์ตเวกเตอร์แมชชีน." Computer Technology King Mongkut's Institute of Technology North Bangkok. (2549):13-21.

[11] ปัจจัยที่มีผลต่อประสิทธิภาพของผู้เรียนผ่านเทคนิคการสอนแบบผู้เรียนมีส่วนร่วมกรณีศึกษา วิชาทฤษฎีสาขาสถาปัตยกรรมภายใน (Factors Affecting Student Efficiency through Cooperative / Collaborative Learning Technique: A Case Study of Interior Architectural Program) พรทิพย์ เรืองธรรม.

[12] ธนวัฒน์ ปัญญาวงค์. เทคนิคการจำแนกข้อมูล (Classification) [Online]: สืบค้นเมื่อวันที่ 4 กุมภาพันธ์ 2560 จาก https://staff.informatics.buu.ac.th/~komate/886464/%5B6%5D-Classification.pdf

[13] กฤตยา ทองผาสุก. "การเปรียบเทียบเทคนิคต้นไม้ตัดสินใจกฎนาอีฟเบย์ และเคเนียรเรสเนเบอร์เพื่อการจำแนกข้อมูล." The 7th National Conference on Computing and Information Technology (2011):30-35.

การพยากรณ์จำนวนผู้ป่วยโรคไข้เลือดออก ด้วยเทคนิคอนุกรมเวลา
Forecasting Number of Dengue Patients using Time Series Technique

ตรีรัตน์ เสริมทรัพย์ (Trirat Soemsap)[1] สุริฉาย มีหมู่ (Sureechay Meemoo)[2]
พัชรี อ่อนแก้ว (Patcharee Oankaew)[3] และพรวิมล คล่องสั่งสอน (Pornvimon Klongsangson)[4]
[1,2,3]สาขาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยราชภัฏอุดรธานี
[4]สาขาสถิติประยุกต์ คณะวิทยาศาสตร์ มหาวิทยาลัยราชภัฏอุดรธานี
[1]Treerat54@hotmail.com,[2] Jajakapom@gmail.com,[3]Noonjoy14@gmail.com, [4]Klongsangson@gmail.com

บทคัดย่อ

การพยากรณ์จำนวนผู้ป่วยโรคไข้เลือดออกที่ผ่านมาได้ใช้ข้อมูลสภาพภูมิอากาศมาพยากรณ์จำนวนผู้ป่วยในอนาคต แต่ในบางพื้นที่ไม่มีข้อมูลนี้เก็บไว้ประกอบกับโรคไข้เลือดออกเป็นโรคที่เกิดขึ้นตามฤดูกาล งานวิจัยนี้จึงใช้เทคนิคการวิเคราะห์อนุกรมเวลาแบบฉบับ ซึ่งประกอบด้วยการหาค่าแนวโน้ม ค่าการเปลี่ยนแปลงตามฤดูกาล และค่าการเปลี่ยนแปลงตามวัฏจักร ข้อมูลที่ใช้ประกอบด้วย ข้อมูลจำนวนผู้ป่วยโรคไข้เลือดออก สำหรับเรียนรู้ 10 ปี และใช้ 1 ปีสำหรับทดสอบ การหาค่าแนวโน้มใช้ 3 วิธี ได้แก่ วิธีกำลังสองน้อยที่สุด วิธีเฉลี่ยเคลื่อนที่ วิธีเฉลี่ยทีละครึ่ง การหาค่าดัชนีฤดูกาลใช้วิธีร้อยละของค่าแนวโน้ม และกำหนดให้ค่าการเปลี่ยนแปลงตามวัฏจักรมีค่าคงที่เป็น 0.1 ผลการทดลองพบว่าวิธีที่ใช้ค่าแนวโน้มเฉลี่ยเคลื่อนมีค่าความคลาดเคลื่อนของการพยากรณ์น้อยที่สุด RMSE = 2.54

คำสำคัญ : การพยากรณ์ โรคไข้เลือดออก อนุกรมเวลา

Abstract

Forecasting the number of patients with dengue in the past has used climate data to forecast the number of patients in the future. However, in some areas, there is not this data and also dengue fever is a disease caused by seasonal, so this research used time series analysis model which includes the estimation of the trend, the seasonal variation and the cyclical variation. The used experimental data for learning consisted of the number of dengue patients for 10 years and for testing 1 year. In this research we estimated the trend by using the least squares method, the moving average method and the semi-average method. The season index was estimated by using the percentage of the trend method and the cyclical variation is constant value which equal to 0.1. The results show that the moving average method has minimum error with RMSE = 2.54.

Keywords : Forecast, Dengue, Time Series.

1. บทนำ

โรคไข้เลือดออกเกิดจากเชื้อไวรัสที่มียุงลายเป็นพาหะนำโรค ซึ่งได้กลายเป็นปัญหาของด้านสาธารณสุขในหลายประเทศทั่วโลก รวมทั้งประเทศไทยด้วย เนื่องจากโรคไข้เลือดออกนั้นสามารถแพร่กระจายอย่างกว้างขวางและติดต่อได้อย่างรวดเร็ว ทำให้มีจำนวนผู้ป่วยเพิ่มมากยิ่งขึ้นโดยเฉพาะในชุมชนที่มีสภาพแวดล้อมที่เป็นแหล่งเพาะพันธ์ยุงลาย การพยากรณ์จำนวนผู้ป่วยโรคไข้เลือดออก จะช่วยให้ผู้ที่เกี่ยวข้องหาวิธีป้องกันการแพร่ระบาดของโรคได้ทันเวลา

เนื่องจากโรคไข้เลือดออกมียุงเป็นพาหะ ในการพยากรณ์จำนวนผู้ป่วยโรคไข้เลือดออกจึงนิยมใช้สภาพภูมิอากาศในการพยากรณ์ [1-3] แต่ประสบกับปัญหาที่ในบางพื้นที่ไม่ข้อมูลเหล่านั้นจัดเก็บไว้ ลักษณะของการเกิดโรคไข้เลือดออกจะคล้าย ๆ กันในทุกปี แต่จะมีรอบหรือวัฏจักรที่บางปีจะมีจำนวนผู้ป่วยสูง ในอำเภอเมือง จังหวัดอุดรธานี จะมีจำนวนผู้ป่วยสูงทุกสามปีโดยบางรอบสูงไม่มาก งานวิจัยนี้จึงใช้เทคนิคอนุกรมเวลาแบบฉบับที่ใช้จำนวนผู้ป่วยในอดีตมาพยากรณ์จำนวนผู้ป่วยใน

อนาคต ค่าที่นำมาใช้ในเทคนิคอนุกรมเวลาประกอบด้วย ค่าแนวโน้ม ค่าเปลี่ยนแปลงตามดูกาล ค่าเปลี่ยนแปลงตามวัฏจักร ค่าแนวโน้มที่ใช้ในการทดลองนี้ประมาณค่าได้จาก 3 วิธี ได้แก่ การหาค่าแนวโน้มใช้วิธีกำลังสองน้อยที่สุด วิธีเฉลี่ยเคลื่อนที่ วิธีเฉลี่ยทีละครึ่ง แล้วนำผลพยากรณ์ที่ได้เปรียบเทียบกัน

2. งานวิจัยที่เกี่ยวข้อง

Bharambe และ Kalbande [4] ได้สรุปงานวิจัยที่ผ่านมาเกี่ยวกับการพยากรณ์จำนวนผู้ป่วยโรคไข้เลือดออก โดยจำแนกเทคโนโลยีที่ใช้พยากรณ์เป็น เครือข่ายประสาทเทียม สื่อสังคมออนไลน์ เหมืองข้อมูล วิธีการผสมระหว่างเครือข่ายประสาทเทียมกับขั้นตอนวิธีพันธุกรรม และอนุกรมเวลา

ชะฎายุทธ์ ปัทเมฆ [5] พยากรณ์จำนวนผู้ป่วยโรคไข้เลือดออกเขตภาคตะวันออกของประเทศไทยปี 2556 – 2560 โดยใช้อนุกรมเวลาวิเคราะห์แบบปรับเรียบเอ็กซ์โปเนนเชียลพบว่า ถ้าขนาดและช่วงเวลาของการเกิดวัฏจักรไม่เท่ากัน การพยากรณ์จะไม่สามารถทำได้ด้วยความถูกต้องแน่นอน สำหรับรูปแบบการระบาดของโรคไข้เลือดออกของเขตภาคตะวันออกเป็นการสลับการระบาดสูงต่ำเป็นลักษณะเป็นปีเว้นปี

กิตติวิชญ์ ชวนินทวิสุทธิ์ และ สมจิตร อาจอินทร์ [6] พยากรณ์พื้นที่เสี่ยงต่อการเกิดโรคไข้เลือดออกโดยใช้เทคนิคเหมืองข้อมูล

ชาญชัยณรงค์ ทรงคาศรี [7] ได้ทำวิจัยเรื่องการพยากรณ์การเกิดโรคล่วงหน้า ใช้เทคนิคการทำให้เรียบยกกำลังสามของวินเตอร์

ตามที่ได้กล่าวมาแล้วการพยากรณ์จำนวนผู้ป่วยโรคไข้เลือดออกมีหลายเทคนิควิธี การพยากรณ์โดยใช้อนุกรมเวลาเหมาะกับที่พื้นที่ที่ขาดข้อมูลเกี่ยวกับสิ่งแวดล้อมเช่น สภาพอากาศ เหมาะกับโรคที่เกิดตามฤดูกาล และมีวัฏจักรของการเกิดโรค จึงนำเสนอการพยากรณ์โดยเทคนิคอนุกรมเวลาแบบฉบับ

3. เทคนิค ทฤษฎีที่ใช้ในการวิจัย

3.1 อนุกรมเวลาแบบฉบับ [8]

ข้อมูลหรือค่าสังเกตที่เปลี่ยนแปลงไปตามลำดับเวลาที่เกิดขึ้น ข้อมูลเหล่านี้ถูกเก็บรวบรวม ณ ช่วงเวลาต่าง ๆ ที่เท่า ๆ กัน หรือไม่ก็ได้ อนุกรมเวลาเป็นเรื่องเกี่ยวกับการเปลี่ยนแปลงค่าของตัวแปรในช่วงเวลาที่ผ่านไป ถ้าอนุกรมเวลาแสดงให้เห็นรูปแบบการเปลี่ยนแปลงในช่วงเวลาที่ผ่านมาในอดีต จะทำให้สามารถคาดการณ์ได้ว่าในอนาคตลักษณะการเปลี่ยนแปลงของข้อมูลควรอยู่ในรูปแบบใด ค่าพยากรณ์ (F) ในงานวิจัยนี้หาได้จากสมการ (1) [8]

$$F = T \times S \times C \qquad (1)$$

โดยที่ T เป็นค่าแนวโน้ม

S เป็นค่าเปลี่ยนแปลงตามฤดูกาล

C เป็นค่าเปลี่ยนแปลงตามวัฏจักร

ตารางที่ 1: ตัวอย่างการคำนวณหาค่า T โดยวิธีที่กำลังสองน้อยที่สุด

	Y	X	XY	X*X	T=a+bX	100*Y/T
1	4	-23	-92	529	6.251	63.99
2	6	-21	-126	441	7.617	78.771
3	10	-19	-190	361	8.983	111.321
4	8	-17	-136	289	10.349	77.302
5	22	-15	-330	225	11.715	187.793
6	32	-13	-416	169	13.081	244.63
7	22	-11	-242	121	14.447	152.281
8	31	-9	-279	81	15.813	196.041
9	24	-7	-168	49	17.179	139.705
10	7	-5	-35	25	18.545	37.746
11	10	-3	-30	9	19.911	50.223
12	10	-1	-10	1	21.277	46.999
13	5	1	5	1	22.643	22.082
14	0	3	0	9	24.009	0
15	4	5	20	25	25.375	15.764
16	5	7	35	49	26.741	18.698
17	26	9	234	81	28.107	92.504
18	32	11	352	121	29.473	108.574
19	67	13	871	169	30.839	217.257
20	55	15	825	225	32.205	170.781
21	52	17	884	289	33.571	154.896
22	36	19	684	361	34.937	103.043
23	35	21	735	441	36.303	96.411
24	24	23	552	529	37.669	63.713

ค่าแนวโน้มแสดงการเปลี่ยนแปลงของข้อมูลที่สนใจในระยะยาว การหาค่าแนวโน้มที่ใช้ทดลองนี้มี 3 วิธี

3.1.1 วิธีการหาค่าแนวโน้มกำลังสองน้อยที่สุด [8-9]

สำหรับข้อมูลอนุกรมเวลาข้อมูลที่ได้จะมีความสัมพันธ์กับเวลา

X เป็นเลขที่เวลาต่อเนื่องจากช่วงเรียนรู้

Y เป็นข้อมูลจริงในเลขที่ช่องเวลา X

n เป็นจำนวนเดือนที่ใช้เรียนรู้ในที่นี้คือ 120

จากตารางที่ 1 คำนวณหาค่า T ได้ดังนี้

$$T = a + bX \quad (2)$$

เมื่อ $$a = \frac{\sum Y}{n} \quad (3)$$

$$b = \frac{\sum xy}{\sum x 2} \quad (4)$$

ในการพยากรณ์การหาค่า T ใช้ a และ b จากช่วงเรียนรู้ X เป็นเลขที่เวลาต่อเนื่องจากช่วงเรียนรู้

3.1.2 วิธีการหาค่าแนวโน้มเฉลี่ยเคลื่อนที่

งานวิจัยนี้ใช้ค่าเฉลี่ยคลาดเคลื่อน 4 เดือน โดยนำจำนวนผู้ป่วย 4 เดือนติดกันมาหาค่าเฉลี่ย ดังตารางที่ 2 [8-9]

ตารางที่ 2: ตัวอย่างการคำนวณค่าวิธีเฉลี่ยเคลื่อนที่ 4 เดือน [10]

ปี	เดือน	จำนวนผู้ป่วย	ค่าเฉลี่ยเคลื่อนที่
2557	1	5	
2557	2	0	
2557	3	5	(5+0+5+1)/4 = 2.75
2557	4	1	(0+5+1+7)/4 = 3.25
2557	5	7	(5+1+7+28)/4=10.25
2557	6	28	(1+7+28+25)/4=15.25
2557	7	25	

3.1.3 วิธีการหาค่าแนวโน้มเฉลี่ยทีละครึ่ง

การหาเส้นแนวโน้ม วิธีนี้จะแบ่งข้อมูล เป็น 2 ส่วน แล้วคำนวณค่าเฉลี่ยของแต่ละส่วนแล้วให้เส้นแนวโน้มลากผ่าน

ตัวอย่าง การหาค่าเฉลี่ยทีละครึ่ง โดยจำนวนปีที่เรียนรู้เท่ากับ 4

n = จำนวนปีที่เรียนรู้ x จำนวนคาบต่อปี/2 = 4x3/2 = 6 (5)

$$k_1 = (n + 1) / 2 \quad , \quad k_2 = k_1 + n \quad (6)$$

$$y_1 = (\sum_{i=1}^{n} y)/n \qquad y_2 = (\sum_{i=n+1}^{2n} y)/n \quad (7)$$

$$m = (y_1 - y_2) / (k_1 - k_2) \quad (8)$$

$$b = (y_1 - m \times k_1) \quad (9)$$

จะได้ $$T = mX + b \quad (10)$$

ตารางที่ 3: ตัวอย่างการคำนวณค่าวิธีเฉลี่ยทีละครึ่ง [9-10]

ปี	คาบที่ (X)	จำนวนผู้ป่วย (Y)	k_1,k_2	y_1,y_2
2547	1	4		
	2	6		
	3	10	k_1=(6+1)/2=3.5	y_1=82/6=15.33
2548	4	8		
	5	22		
	6	32		
2549	7	22		
	8	31		
	9	24	**k_2=3.5+6=9.5**	y_2=104/6=17.33
2550	10	7		
	11	10		
	12	10		

3.2 ค่าเปลี่ยนแปลงตามฤดูกาล (S)

เป็นการวิเคราะห์ความผันแปรตามฤดูกาลที่เหมาะสมกับข้อมูลที่มีอิทธิพลของฤดูกาล งานวิจัยนี้ใช้วิธีหาค่าร้อยละของค่าแนวโน้ม มีขั้นตอนการหาดังนี้ ซึ่งค่าที่ได้จะเป็นดัชนีฤดูกาลของแต่ละเดือน ในเดือนเดียวกันของทุกปีจะมีค่านี้เท่ากัน [9]

1) นำ 100 มาคูณ กับจำนวนผู้ป่วย แล้วหารด้วยค่าแนวโน้มที่หาได้ เพื่อกำจัดค่าแนวโน้ม ดังสดมภ์ที่ 7 ของตารางที่ 1

ตารางที่ 4: ตัวอย่างการหาค่าเปลี่ยนแปลงตามฤดูกาล (ค่าแนวโน้มใช้วิธีที่ 2)

เดือน	2548	2549	2550	มัธยฐาน (1)	S=12x(1)/(2)
1		59.259	80.000	69.63	0.718
2		100.000	0.000	50	0.516
3	181.818	142.857	114.286	142.857	1.474
4	30.769	69.565	57.143	57.143	0.59
5	68.293	122.222	160.000	122.222	1.261
6	183.607	152.381	93.750	152.381	1.572
7	101.010	82.243	150.562	101.01	1.042
8	119.084	113.761	107.843	113.761	1.174
9	127.869	114.286	99.048	114.286	1.179
10	69.091	38.889	80.899	69.091	0.713
11	69.333	78.431	95.238	78.431	0.809
12	40.000	125.000	92.308	92.308	0.952
			รวม (2)	1163.12	

2)นำค่าที่ได้จากขั้นตอนที่หนึ่ง โดยข้อมูลแต่ละปีไปเรียงทีละเดือน เดือนเดียวกันอยู่แถวเดียวกัน แล้วหาค่ามัธยฐานแต่ละเดือน และปรับค่าให้ผลรวมเป็น 12 ดังตารางที่ 4 ซึ่งในทุกปีจะมีค่าเปลี่ยนแปลงตามฤดูกาลของเดือนหนึ่งเท่ากันทุกปี

3.3 ค่าเปลี่ยนแปลงตามวัฏจักร (C)

มีลักษณะคล้ายกับฤดูกาลเพียงแต่ว่าค่าเปลี่ยนแปลงตามฤดูกาลมีคาบหรือระยะเวลาสั้นกว่าค่าเปลี่ยนแปลงตามวัฏจักร แต่ในการทดลองนี้กำหนดให้เป็นค่าคงที่เท่ากับ 0.1 ทั้งนี้เพราะเมื่อครบวัฏจักร 3ปี บางรอบมีจำนวนผู้ป่วยสูงไม่มากเหมือนกับรอบอื่นๆ จึงใช้วิธีตกค้างไม่ได้ [5]

3.4 การคำนวณหาค่าพยากรณ์

การพยากรณ์จำนวนผู้ป่วยไข้เลือดออกคำนวณได้จากสมการที่ (1) ตัวอย่างการคำนวณมีดังตารางที่ 5 ตารางที่ 6 และตารางที่ 7 โดยวิธีที่ 1 หมายถึง วิธีกำลังสองน้อยที่สุด วิธีที่ 2 หมายถึง วิธีเฉลี่ยเคลื่อนที่ และวิธีที่ 3 คือ วิธีเฉลี่ยทีละครึ่ง [9], [11]

ตารางที่ 5: การหาผลพยากรณ์ของวิธีการหาค่าแนวโน้มวิธีที่ 1

ปี	เดือน	X	T	S	C	ผลพยากรณ์
2557	1	121	53.894	0.2972	0.1	1.602
2557	2	123	54.26	0.0890	0.1	0.483
2557	3	125	54.627	0.2635	0.1	1.439
2557	4	127	54.993	0.2691	0.1	1.480
2557	5	129	55.359	1.0401	0.1	5.758
2557	6	131	55.726	1.8717	0.1	10.430
2557	7	133	56.092	2.5093	0.1	14.075
2557	8	135	56.459	2.4313	0.1	13.727
2557	9	137	56.825	1.7826	0.1	10.129
2557	10	139	57.192	0.7392	0.1	4.228
2557	11	141	57.558	0.5259	0.1	3.027
2557	12	143	57.924	0.1812	0.1	1.050

3.5 ค่าความคลาดเคลื่อน [13]

ความคลาดเคลื่อน คือ ผลต่างระหว่างค่าที่พยากรณ์ได้กับค่าที่แท้จริง ถ้าค่าที่วัดได้ใกล้เคียงกับค่าจริงมากแสดงว่าการวัดนั้นมีความแม่นยำหรือความถูกต้องสูง ค่าความคลาดเคลื่อนที่ใช้ในงานวิจัยนี้คือ RMSE สาเหตุที่ไม่เลือกใช้ MAPE เนื่องจากการหา MAPE มีการหารด้วยจำนวนผู้ป่วยจริง ซึ่งจำนวนผู้ป่วยจริงในบางเดือนมีค่าเป็น 0

ตารางที่ 6: การหาผลพยากรณ์ของวิธีการหาค่าแนวโน้มวิธีที่ 2

ปี	เดือน	T	S	C	ผลพยากรณ์
2557	1	6	0.668	0.1	0.401
2557	2	5.25	0.901	0.1	0.473
2557	3	4.375	1.140	0.1	0.499
2557	4	9.5	0.714	0.1	0.678
2557	5	18.125	1.362	0.1	2.469
2557	6	34	1.364	0.1	4.639
2557	7	50.75	1.613	0.1	8.184
2557	8	54.125	1.281	0.1	6.931
2557	9	53	0.882	0.1	4.677
2557	10	38.875	0.742	0.1	2.884
2557	11	23.5	0.864	0.1	2.031
2557	12	11.375	0.468	0.1	0.533

ตารางที่ 7: การหาผลพยากรณ์ของวิธีการหาค่าแนวโน้มวิธีที่ 3

ปี	เดือน	X	T	S	C	ผลพยากรณ์
2557	1	121	53.352	0.297	0.1	1.587
2557	2	122	53.715	0.089	0.1	0.478
2557	3	123	54.078	0.263	0.1	1.424
2557	4	124	54.441	0.269	0.1	1.465
2557	5	125	54.804	1.039	0.1	5.696
2557	6	126	55.167	1.873	0.1	10.331
2557	7	127	55.530	2.510	0.1	13.939
2557	8	128	55.893	2.431	0.1	13.587
2557	9	129	56.256	1.782	0.1	10.026
2557	10	130	56.619	0.739	0.1	4.182
2557	11	131	56.982	0.526	0.1	2.997
2557	12	132	57.345	0.181	0.1	1.039

$$\text{MAPE} = \frac{100}{n}\sum_{t=1}^{n}\left|\frac{y_t - \hat{y}}{y_t}\right| \qquad (11)$$

$$\text{RMSE} = \sqrt{\sum_{t=1}^{n}\frac{(y_i - \hat{y})^2}{n}} \qquad (12)$$

n เป็นจำนวนเดือน t เป็นเดือนที่

y_t เป็นจำนวนผู้ป่วยจริงเดือนที่ t

$\hat{y}_t$ เป็นจำนวนผู้ป่วยที่พยากรณ์ได้เดือนที่ t

4. วิธีการดำเนินงาน

4.1 วิธีการ

วิธีการวิจัยการพยากรณ์จำนวนผู้ป่วยโรคไข้เลือดออกด้วยเทคนิคการวิเคราะห์อนุกรมเวลา ใช้ข้อมูลจำนวนผู้ป่วยโรคไข้เลือดออก อำเภอเมือง จังหวัดอุดรธานี ดังตารางที่ 8 โดยใช้ข้อมูล 10 ปี ย้อนหลังสำหรับเรียนรู้ ตั้งแต่ ปี พ.ศ. 2547–2556 และใช้ ปี พ.ศ. 2557 เป็นข้อมูลสำหรับทดสอบการพยากรณ์ โดยใช้เทคนิคอนุกรมเวลาแบบฉบับ ซึ่งต้องคำนวณหาค่าแนวโน้ม ค่าดัชนีฤดูกาล และค่าการเปลี่ยนแปลงตามวัฏจักร การหาค่าแนวโน้มเลือกใช้ 3 วิธี มาเปรียบเทียบผลการพยากรณ์ ได้แก่ วิธีกำลังสองน้อยที่สุด วิธีเฉลี่ยเคลื่อนที่ และวิธีเฉลี่ยทีละครึ่ง สำหรับค่าเปลี่ยนแปลงตามฤดูกาลและค่าเปลี่ยนแปลงตามวัฏจักร ใช้วิธีร้อยละของค่าแนวโน้มและวิธีกำหนดเป็นค่าคงที่ตามลำดับ [9], [10]

ตารางที่ 8: จำนวนผู้ป่วยโรคไข้เลือดออกรายเดือน ปี พ.ศ. 2547-2557

	47	48	49	50	51	52	53	54	55	56	57
1	5	4	5	9	8	2	18	1	3	19	2
2	0	6	0	3	1	9	17	1	0	13	6
3	5	10	4	1	2	10	15	0	6	18	0
4	1	8	5	8	2	8	32	0	5	24	2
5	7	22	26	29	18	35	79	7	21	51	1
6	28	32	32	59	27	41	117	26	55	181	5
7	25	22	67	86	49	87	234	15	82	299	5
8	39	31	55	89	17	65	217	23	81	199	5
9	39	24	52	50	11	36	74	17	41	77	9
10	19	7	36	29	18	26	26	4	32	14	1
11	13	10	35	11	8	25	10	6	45	8	0
12	4	10	24	3	2	8	3	0	8	1	0

4.2 เครื่องมือที่ใช้

เครื่องมือที่ใช้ในการทดลองประกอบด้วย

4.2.1 โปรแกรม Visual Studio 2010 (C#)

4.2.2 ระบบจัดการฐานข้อมูล Microsoft Access 2010

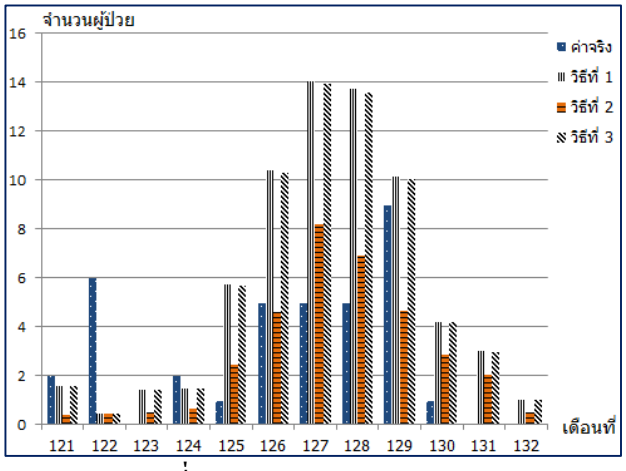
4.2.3 โน้ตบุ๊กคอมพิวเตอร์ core i5

5. ผลการทดลอง

จากตารางที่ 9 ตารางที่ 10 และภาพที่ 1 พบว่าวิธีที่ 2 ให้ผลการพยากรณ์ใกล้เคียงจำนวนผู้ป่วยจริงที่สุด ซึ่งค่าแนวโน้มของวิธีที่ 2 แตกต่างจากวิธีอื่นคือไม่เรียงเป็นเส้นตรง พิจารณาได้จากตารางที่ 5 ตารางที่ 6 และตารางที่ 7

ตารางที่ 9: เปรียบเทียบผลการพยากรณ์

ปี		เดือน	ค่าจริง	ผลพยากรณ์		
				วิธีที่ 1	วิธีที่ 2	วิธีที่ 3
2557		1	**2**	1.602	**0.401**	1.587
2557		2	**6**	0.483	**0.473**	0.478
2557		3	**0**	1.439	**0.499**	1.424
2557		4	**2**	1.48	**0.678**	1.465
2557		5	**1**	5.758	**2.469**	5.696
2557		6	**5**	10.43	**4.639**	10.331
2557		7	**5**	14.075	**8.184**	13.939
2557		8	**5**	13.727	**6.931**	13.587
2557		9	**9**	10.129	**4.677**	10.026
2557		10	**1**	4.228	**2.884**	4.182
2557		11	**0**	3.027	**2.031**	2.997
2557		12	**0**	1.05	**0.533**	1.039



ภาพที่ 1: กราฟเปรียบเทียบผลการพยากรณ์

ตารางที่ 10: ค่าความคลาดเคลื่อน RMSE

วิธีที่ 1	วิธีที่ 2	วิธีที่ 3
4.69	2.54	4.74

6. สรุปผล

การพยากรณ์จำนวนผู้ป่วยโรคไข้เลือดออกด้วยเทคนิคการวิเคราะห์อนุกรมเวลา โดยใช้อนุกรมเวลาแบบฉบับ และใช้ค่าแนวโน้ม 3 วิธี คือ วิธีการหาค่าแนวโน้มกำลังสองน้อยที่สุด วิธีการหาค่าแนวโน้มเฉลี่ยเคลื่อนที่ และวิธีการหาค่าแนวโน้มเฉลี่ยทีละครึ่ง เมื่อนำจำนวนผู้ป่วยโรคไข้เลือดออกมาวิเคราะห์เทคนิคอนุกรมเวลาแล้วพบว่า ทั้งสามวิธีพยากรณ์ช่วงเดือนกรกฎาคม สิงหาคม และกันยายนจะมีจำนวนผู้ป่วยสูงที่สุด ซึ่งสอดคล้องกับฤดูกาลของโรคไข้เลือดออก นอกจากนี้วิธีการหาค่าแนวโน้มเฉลี่ยเคลื่อนที่ซึ่งมีค่าแนวโน้มต่างจากอีกสองวิธี คือไม่เป็นเส้นตรง มีผลการพยากรณ์เปรียบเทียบใกล้เคียงกับจำนวนผู้ป่วยจริงมากที่สุด วิธีการหาค่าแนวโน้มกำลังสองน้อยที่สุดใกล้เคียงอันดับที่สอง และวิธีการหาค่าแนวโน้มเฉลี่ยทีละครึ่งเป็นลำดับสุดท้ายที่ใกล้เคียงค่าจริงน้อยที่สุด แต่ผลการทดลองวิธีที่หนึ่งและวิธีที่สามมีผลพยากรณ์ใกล้เคียงกันมากกราฟที่แสดงเกือบเป็นเส้นเดียวกัน วิธีการวิจัยการพยากรณ์จำนวนผู้ป่วยโรคไข้เลือดออกในอนาคตอาจจะเพิ่มจำนวนข้อมูล และพยากรณ์ด้วยวิธีอื่นแล้วนำผลลัพธ์มาเปรียบเทียบกันเพื่อให้ได้วิธีที่มีผลลัพธ์การพยากรณ์แม่นยำยิ่งขึ้น

7. กิตติกรรมประกาศ

ขอขอบคุณสำนักงานสาธารณสุขจังหวัดอุดรธานีที่อนุเคราะห์ข้อมูล และขอขอบคุณ NSC 2017 ที่ให้ทุนอุดหนุนการวิจัย

เอกสารอ้างอิง

[1] H. I. Datoc, R. Caparas, J. Caro, “Forecasting and Data Visualization of Dengue spread in the Philippine Visayas Island group”, *Information, Intelligence, Systems & Applications (IISA), 2016 7th International Conference*, Chalkidiki, Greece, 13-15 July 2016.

[2] H. M. Aburas, B. G. Cetiner, M. Sar, “Dengue confirmed-cases prediction: A neural network model”, *Expert Systems with Applications*, 37, 2010, pp. 4256–4260

[3] R. Lowe, T. C. Bailey, D. B. Stephenson, R. J. Graham, C. A.S. Coelho, M. S. Carvalho, C. Barcellos, “Spatio-temporal modelling of climate-sensitive disease risk: Towards an early warning system for dengue in Brazil”, *Computers & Geosciences*, 37, 2011, pp. 371 –381.

[4] A. A. Bharambe and D. R. Kalbande, “Techniques and Approaches for Disease Outbreak Prediction: A Survey,” *WIR'16 Proceedings of the ACM Symposium on Women in Research 2016*, Indore, India, March 2016, pp. 100-102

[5] ชะฎายุทธ์ ปัทมะ, “พยากรณ์จำนวนผู้ป่วยโรคไข้เลือดออกเขตภาคตะวันออกของประเทศไทยปี 2556-2560,” *วารสารวิชาการสาธารณสุข*, ปีที่ 25 ฉบับที่ 5 กันยายน – ตุลาคม 2557, หน้า 804-812.

[6] กิตติวิชญ์ ชวนินทวิสุทธิ์ และ สมจิตร อาจอินทร์, “ระบบพยากรณ์และควบคุมการระบาดโรคไข้เลือดออกโดยใช้เหมืองข้อมูลและภูมิสารสนเทศ,” การศึกษาอิสระปริญญาวิทยาศาสตรมหาบัณฑิต คณะวิทยาศาสตร์, มหาวิทยาลัยขอนแก่น, ขอนแก่น, 2554.

[7] ชาญชัยณรงค์ ทรงคาศรี, “รูปแบบการพยากรณ์โรคไข้เลือดออกในพื้นที่สำนักงานป้องกันควบคุมโรคที่ 6”, *วารสารสำนักงานป้องกันควบคุมโรคที่ 6*, ปีที่ 20 ฉบับที่ 1 ตุลาคม 2555 - มีนาคม 2556, ขอนแก่น, หน้า 65-81.

[8] R. M. Gor, *Forecasting Techniques*, Gnfc Info: Bodakdev.

[9] พรทิภา จรูญวรรณ, *อนุกรมเวลาและเลขดัชนี*, มหาวิทยาลัยราชภัฏอุดรธานี: อุดรธานี, 2542.

[10] ศิริลักษณ์ สุวรรณวงค์, *การวิเคราะห์อนุกรมเวลา*, สุวีริยาสาส์น: กรุงเทพ, 2535.

[11] วิชิต หล่อจีระชุณห์กุล, *เทคนิคการพยากรณ์*, พิมพ์ครั้งที่ 3, สถาบันบัณฑิตพัฒนบริหารศาสตร์: กรุงเทพ, 2548.

[12] สมเกียรติ เกตุเอี่ยม, *เทคนิคการพยากรณ์*, พิมพ์ครั้งที่ 2, มหาวิทยาลัยทักษิณ: สงขลา, 2548.

[13] E. Woschnagg, *Evaluating Forecast Accureacy*, University of Vienna: Vienna, 2004.

การวิเคราะห์ปัจจัยที่ส่งผลต่อบทวิจารณ์ของลูกค้าโดยใช้เหมืองความคิดเห็น
Analysis of Affecting Factors to Customer Reviews using Opinion Mining

ภัทราภรณ์ โชตินาเสียว (Pattaraporn Chotnasiaw)[1] วรารัตน์ สงฆ์แป้น (Wararat Songpan)[2]
สมจิตร อาจอินทร์ (Somjit Arch-int)[3] และสายยัญ สายยศ (Saiyan Saiyod)[4]
ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น
[1]pattarporn_c@kkumail.com, [2]wararat@kku.ac.th, [3]somjit@kku.ac.th, [4]saiyan@kku.ac.th

บทคัดย่อ

บทวิจารณ์ของลูกค้ามีแนวคิดที่หลากหลายและเป็นประโยชน์อย่างมากที่จะช่วยในการนำไปปรับปรุงสินค้าและการบริการของร้านค้าให้ดีขึ้น แต่เนื่องจากว่าบทวิจารณ์บางอันมีการวิจารณ์ไม่สอดคล้องกันกับค่าคะแนนดังกล่าวที่มีความคิดด้านบวกหรือลบต่อร้านค้า งานวิจัยนี้จึงเป็นการนำเสนอการวิเคราะห์ปัจจัยที่ส่งผลทำให้บทวิจารณ์เกิดผลตอบรับเชิงบวกหรือเชิงลบของลูกค้าที่เข้าไปใช้บริการ โดยใช้หลักการทำงานของเหมืองความคิดเห็น เพื่อสกัดหาคุณลักษณะที่ได้จากบทวิจารณ์มาใช้เป็นปัจจัยและนำไปวิเคราะห์ด้วยเทคนิคต้นไม้ตัดสินใจมาเป็นตัวจำแนกกลุ่มอีกทั้งได้เปรียบเทียบการคัดเลือกปัจจัยด้วย 3 วิธี ได้แก่ CfsSubsetEval, InfoGainAttribute และคัดเลือกปัจจัยจากคำที่มีความถี่ที่ปรากฏในบทวิจารณ์สูง นักวิจัยเก็บชุดข้อมูลได้นำมาจากเว็บบวงในดอทคอม กรณีศึกษาร้านกาแฟจำนวน 11,299 บทวิจารณ์ ได้สกัดคุณลักษณะจากบทวิจารณ์จำนวน 50 คุณลักษณะ ผลการทดลองพบว่าปัจจัยที่ส่งผลต่อผลตอบรับเชิงบวกและเชิงลบด้วยการเลือกจากคำที่มีความถี่ที่ปรากฏในบทวิจารณ์สูงและให้ผลดีที่สุดด้วยค่าความถูกต้อง 85.98%

คำสำคัญ: การวิเคราะห์ปัจจัย เหมืองความคิดเห็น
CfsSubsetEval InfoGainAttribute ต้นไม้ตัดสินใจ

Abstract

Customer reviews are important and useful to improve a product and service of service shop. However, the problem is some of customer reviews that are inconsistent with its review's score. Therefore, the customers are hard making decision to summarize the reviews to positive or negative. The purpose of this research finds out factor analysis method to help the reviews classification in positive or negative class using opinion mining. To finding sensible factors from the reviews, decision tree is used classifying and enhance to comparing the factors selection with CfsSubsetEval, InfoGainAttribute and high frequency appearance. The data are collected from www.wongnai.com in type of coffee shop, the total of reviews are 11,299 records and extract attributes of the highest 50 frequency factors. The result shows factors that make the most accuracy of effecting factor in positive and negative classify by high frequency appearance method in 85.98% of accuracy compared to other attribute selection methods.

Keyword: Factor Analysis, Opinion Mining,
CfsSubsetEval, InfoGainAttribute, Decision Tree.

1.บทนำ

ปัจจุบันมีข้อความแสดงความคิดเห็นของผู้คนอย่างมากมายบนอินเทอร์เน็ตซึ่งเป็นที่นิยมสำหรับผู้ที่ชอบใช้งาน ทั้งจากเว็บไซต์ บล็อก และเครือข่ายสังคมออนไลน์ที่อนุญาตให้คนสามารถเข้ามาแสดงข้อคิดเห็นและยังมีการให้ค่าคะแนนของสินค้าและการบริการด้วย ข้อความคิดเห็นเหล่านี้ช่วยเปิดพื้นที่ใหม่ให้กับธุรกิจต่างๆ และตลาดออนไลน์ เพราะข้อความต่างๆ มีผลทำให้ลูกค้าเลือกใช้หรือจะไม่ใช้สินค้าและการบริการได้

จากตัวอย่างความคิดเห็น “อาหารร้านนี้นิ้งทานมาสองรอบแล้ว แต่ยังไม่ประทับใจเหมือนเดิมเลย แอบเสียใจ เห็นเพื่อนๆ

ชมกันค่อนข้างเยอะ อีกอย่างรู้สึกว่าค่าอาหารแพงมาก และยังมีเซอวิชชาร์จอีก คุณภาพอาหารดีและอยากให้ปรับรสชาติค่ะจะดีกว่านี้ ส่วนเค้กเค้าก็อร่อยเหมือนเคยค่ะ" ลูกค้าให้คะแนนระดับ 3 ซึ่งความคิดเห็นอยู่ในด้านที่ไม่ดีเป็นส่วนใหญ่ แต่ค่าคะแนนที่ให้ถือว่าดีและทานได้ทำให้เกิดความลำเอียงในการให้คะแนนได้ และยังสร้างความสับสนให้กับลูกค้าและผู้ให้บริการที่เข้ามาสำรวจความเห็นเหล่านี้ซึ่งมีงานวิจัย[1] ได้กล่าวถึงปัญหานี้เช่นกัน ทำให้ตัดสินใจยากในการพิจารณาบทวิจารณ์ ดังนั้นจึงได้ทำการวิเคราะห์การแบ่งกลุ่มคำด้านบวกและลบเท่าๆ กัน อย่างไรก็ตามการเลือกปัจจัยเพื่อการวิเคราะห์บทวิจารณ์เป็นแบบไม่มีโครงสร้าง (Unstructured Text) ดังนั้นจึงควรหาวิธีการเลือกปัจจัยที่มีความยืดหยุ่นมากขึ้น

จากปัญหาดังกล่าว จึงมีแนวคิดในการสร้างประโยชน์ให้กับผู้ให้บริการจากบทวิจารณ์ของลูกค้ามีต่อร้านค้าอย่างแท้จริง โดยการทำเหมืองความคิดเห็น เพื่อสกัดหาสิ่งที่ต้องการศึกษา ผู้วิจัยจึงได้นำบทวิจารณ์จากเว็บวงในดอทคอม โดยการรวบรวมข้อมูลจากความคิดเห็นมาวิเคราะห์เพื่อหาปัจจัยที่ก่อให้เกิดความคิดเห็นผลตอบรับเชิงบวกและเชิงลบ โดยต้องการหาความสัมพันธ์ที่ได้จากคุณลักษณะ (Attribute) ของข้อมูลด้วยขั้นตอนวิธีต้นไม้ตัดสินใจที่เหมาะสำหรับการใช้วิเคราะห์ข้อมูลทุกรูปแบบ จากงานวิจัย[2] ได้นำเสนอเทคนิคการคัดเลือกปัจจัยที่นำมาใช้กับต้นไม้ตัดสินใจ ดังนั้นในการคัดเลือกปัจจัยของงานวิจัยนี้จึงใช้วิธี CfsSubsetEval, InfoGainAttribute และคัดเลือกจากคำที่มีความถี่ที่ปรากฏในบทวิจารณ์สูง ทั้ง 3 แบบเปรียบเทียบกัน เพื่อนำมาใช้ในการแสดงผลต่อความคิดเห็นที่ได้ผลลัพธ์เชิงบวกและลบ โดยพิจารณาภาพรวมจากประเภทของร้านค้า ซึ่งในงานวิจัยนี้ได้ยกตัวอย่างกลุ่มร้านกาแฟมาใช้เป็นกรณีศึกษาเพื่อที่จะนำไปปรับปรุงคุณภาพและการบริการของร้านกาแฟได้

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 เหมืองความคิดเห็น (Opinion Mining)

การทำเหมืองความคิดเห็น เป็นการจัดการกับความคิดของการนำเสนอความคิดเห็น ความรู้สึก และการกระทำในข้อความที่เกี่ยวข้องด้วยการประมวลผลภาษาธรรมชาติ (NLP) การวิเคราะห์ข้อความและภาษาศาสตร์[3] มีจุดมุ่งหมายเพื่อทำความเข้าใจการแสดงความคิดเห็นจากหัวข้อที่กำหนด อาจสะท้อนให้เห็นถึงการประเมินผลและระบุอารมณ์ในช่วงเวลาของผู้เขียน ณ เวลานั้น ซึ่งประเภทของการประมวลผลภาษาธรรมชาติสำหรับการติดตามอารมณ์ของลูกค้าที่เกี่ยวกับผลิตภัณฑ์ หรือเรียกว่า การวิเคราะห์ความรู้สึก (Sentiment Analysis) เป็นการพิจารณาข้อความในชุดของเอกสารที่มีความคิดเห็นหรือทัศนคติที่เป็นเชิงบวก เชิงลบ และเชิงกลาง นอกจากนี้ยังมีประโยชน์หลายอย่างที่สามารถช่วยให้นักการตลาดประเมินความสำเร็จของแคมเปญโฆษณา หรือการเปิดตัวผลิตภัณฑ์ใหม่ของสินค้าที่ระบุความนิยมของกลุ่มลูกค้าที่มีต่อผลิตภัณฑ์ ในการถ่ายทอดถึงอารมณ์ ความรู้สึก ของลูกค้าถือเป็นข้อมูลเชิงจิตวิทยาที่สามารถนำมาใช้วิเคราะห์หาพฤติกรรมของผู้บริโภคและพฤติกรรมทางสังคมเพื่อปรับปรุงสินค้าและบริการให้ตรงตามความต้องการลูกค้าต่อไป[4]

2.2 ต้นไม้ตัดสินใจด้วยขั้นตอนวิธี C4.5

เป็นขั้นตอนวิธีที่ถูกใช้มากที่สุดแบบหนึ่งในการเรียนรู้ของเครื่องโดยใช้คุณสมบัติของข้อมูลในการแยกแยะข้อมูลออกเป็นกลุ่มต่างๆ ประกอบด้วยโหนด (Node) แต่ละโหนดจะมีคุณลักษณะเป็นตัวทดสอบ กิ่งของต้นไม้ (Branch) คือค่าที่เป็นไปได้ของคุณลักษณะที่ถูกเลือกทดสอบ และใบ (Leaf) จะเป็นส่วนที่อยู่ล่างสุดได้ผลลัพธ์ของกลุ่มจำแนก (Class)[5] ซึ่งในการวัดความซับซ้อนของต้นไม้สินใจอาจจะถูกแบ่งในแง่ของขนาดต้นไม้ ทำให้ขั้นตอนวิธี C4.5 เป็นที่นิยมและมีวิธีการตัดเล็มกิ่งไม้ที่ไม่ให้มีขนาดใหญ่จนเกินไปและลดความไม่ถูกต้องในการแบ่งกลุ่มอีกด้วยงานวิจัย[6] ได้นำขั้นตอนวิธี C4.5 ซึ่งสามารถสร้างการหาค่าเอนโทรปี (Entropy) และค่าเกน (Gain) ตามสมการดังต่อไปนี้

$$I(s_1, s_2, \dots, s_n) = -\sum_{i=1}^{n} \frac{si}{s} \log_2 \frac{si}{s} \quad (1)$$

โดย S คือ จำนวนข้อมูลทั้งหมด ซึ่งมีจำนวนข้อมูล s_i ของ$class_i$ และ n คือ จำนวนคลาสทั้งหมด

จากนั้นหาค่าเอนโทปี (Entropy) เป็นค่าของผลรวมแต่ละแอทริบิวต์ A ที่สามารถแยกประเภทข้อมูลแต่ละคลาสของแต่ละค่าของแอทริบิวต์ m จำนวน ตามสมการดังต่อไปนี้

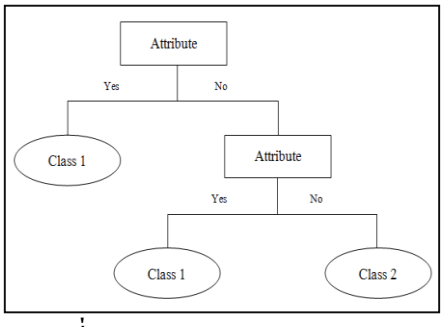
$$E(A) = \sum_{j=1}^{m} \frac{S_{1j}+\cdots+S_{nj}}{S} I(S_{1j}, \ldots, S_{nj}) \quad (2)$$

ดังนั้นค่าเกน (Gain) ของแอทริบิวต์ที่ถูกเลือกหาได้จากการแยกประเภทข้อมูลสารสนเทศทั้งหมดของแอทริบิวต์นั้นจึงคำนวณตามสมการดังต่อไปนี้

$$Gain(A) = I(s_1, s_2, \ldots, s_n) - E(A) \quad (3)$$

2.3 Information Gain Selection (InfoGain)

เป็นค่าเพื่อใช้ในการคัดเลือกคุณลักษณะด้วยการคำนวณค่า Gain สำหรับแต่ละคุณลักษณะ ถ้าคุณลักษณะใดมีค่า Gainสูงสุดจะถูกเลือกให้เป็นกลุ่มย่อยที่มีอำนาจต่อการจำแนกข้อมูล ในเครื่องมือ Weka จะเรียกว่า InfoGainAttribute



ภาพที่ 1: แสดงโครงสร้างต้นไม้ตัดสินใจ

2.4 Correlation based Feature Selection (CFS)

หรืออีกชื่อหนึ่ง CfsSubsetEval เป็นการหากลุ่มของคุณลักษณะที่ถูกประเมินค่า โดยพิจารณาดูจากกลุ่มของคุณลักษณะที่ถูกคัดเลือกสำหรับการจำแนกประเภทของข้อมูลกับข้อมูลที่เกี่ยวข้องกัน ซึ่งการจัดการกลุ่มย่อยของคุณลักษณะจะมีความสัมพันธ์สูงกับกลุ่มที่จำแนก และไม่สัมพันธ์กับกลุ่มจำแนกตัวอื่นๆ ส่วนคุณลักษณะที่ไม่เกี่ยวข้องจะถูกละทิ้ง เมื่อรวมกันแล้วให้ค่าผลทำนายเป็นกลุ่ม กลุ่มที่ให้ค่าสูงที่สุดจะต้องการกลุ่มของคุณลักษณะของข้อมูลที่มีค่าน้อยที่สุด

2.5 การสุ่มเพิ่มจำนวนตัวอย่างกลุ่มน้อย (Synthetic Minority Over-sampling Technique – SMOTE)

เป็นเทคนิคที่ช่วยแก้ปัญหาของการจำแนกข้อมูลที่ไม่สมดุล ซึ่งเป็นข้อมูลจำนวนตัวอย่างต่างกันมาก เมื่อนำข้อมูลที่แตกต่างกันไปทำการจำแนก ผลลัพธ์ที่ได้ทำให้การจำแนกอยู่ในกลุ่มเสียงข้างมาก ดังนั้นจึงแก้ปัญหาโดยสุ่มเพิ่มจำนวนข้อมูลในกลุ่มเสียงข้างน้อย (Minority Class) ให้มีปริมาณข้อมูลใกล้เคียงหรือเท่ากับข้อมูลในเสียงข้างมาก (Majority Class) เมื่อนำไปจำแนกประเภท จะช่วยให้ผลลัพธ์ในการสังเคราะห์กลุ่มข้างน้อยดีขึ้น[7]

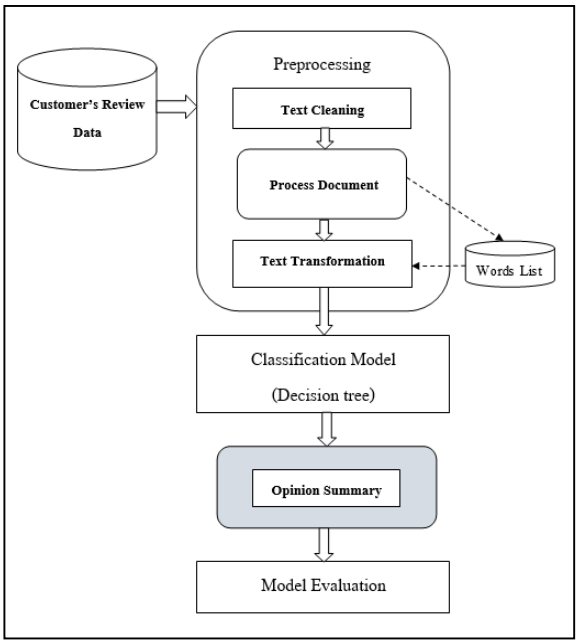
2.6 งานวิจัยที่เกี่ยวข้อง

เนื่องจากข้อมูลความคิดเห็นของลูกค้าจำนวนมากทำให้มีงานวิจัยทางด้านการทำเหมืองความคิดเห็นที่ได้วิเคราะห์จากทัศนคติ ความรู้สึก นึกคิด ของลูกค้ามีออกมาหลายงาน เช่น การวิเคราะห์ความพึงพอใจของลูกค้าและการค้นพบข้อมูลที่เป็นประโยชน์ที่สามารถสนับสนุนการตัดสินใจ โดยใช้ความคิดเห็นเกี่ยวกับโทรศัพท์มือถือ มีการแบ่งประเภทลักษณะการใช้งาน ฮาร์ดแวร์ ซอฟต์แวร์ ราคา และบริการ ที่แสดงด้านบวกและด้านลบของแต่ละหมวดหมู่[8] งานวิจัย[9] ได้สร้างระบบวิเคราะห์คำว่าคำใดแสดงทัศนคติต่อคุณลักษณะของร้านอาหารจาก 5 ประเภท คือ การบริการ รสชาติ ราคา สิ่งอำนวย ความสะดวก และบรรยากาศ ซึ่งแสดงความคิดเห็นที่มีข้อมูลทั้งด้านบวกและด้านลบในแต่ละคุณลักษณะของร้านอาหารแต่ละร้านจากเว็บบวงในดอทคอม อีกทั้งยังมีงานวิจัย[10] ที่เน้นการวิเคราะห์อารมณ์ความคิด ทัศนคติ และความรู้สึกของลูกค้าโดยใช้ถุงของคำ (Bag of Word) เพื่อดึงคุณลักษณะจากความคิดเห็นของลูกค้า นอกจากนี้งานวิจัย[1] ยังได้เลือกคุณลักษณะมาแยกคำดีและไม่ดีก่อนนำเข้าสู่การจำแนกข้อมูลมาใช้วิเคราะห์ความพึงพอใจเช่นกัน

ซึ่งงานวิจัยที่ทำในบทความนี้มีความแตกต่างจากงานวิจัยที่กล่าวมาโดยงานวิจัยดังกล่าวไม่ได้มีการเจาะจงหรือระบุถึงวิธีการเลือกคุณลักษณะที่ส่งผลต่อด้านบวกและด้านลบ ทำให้หลังการจำแนกข้อความที่ใช้ในการนำเสนอไม่ทราบถึงปัจจัยต่างๆ เพื่อนำไปปรับปรุงสินค้าและบริการ และการวิเคราะห์ปัจจัยที่ส่งผลต่อด้านบวกและด้านลบ ผู้วิจัยจึงมีแนวคิดในการวิเคราะห์การเลือกปัจจัยด้วยต้นไม้ตัดสินและใช้วิธี CfsSubsetEval, InfoGainAttribute และคัดเลือกจากคำที่มีความถี่ที่ปรากฏในบทวิจารณ์สูง โดยเปรียบเทียบผลการคัดเลือกปัจจัยที่ดีที่สุดก่อนนำไปประมวลผลการจำแนกข้อความต่อไป

3. วิธีการดำเนินการวิจัย

ในงานวิจัยนี้ การวิเคราะห์เหมืองความคิดเห็น มีขั้นตอนการทำงานของงานวิจัย ประกอบด้วยการทำงาน 5 ส่วน ดังนี้



ภาพที่ 2: กรอบงานวิจัยเหมืองความคิดเห็น

3.1 ขั้นตอนการเก็บรวบรวมข้อมูล

ผู้วิจัยได้เขียนโปรแกรมด้วยภาษา C# ในการดึงข้อมูลที่ได้จากความคิดเห็นของผู้ใช้งานในเว็บวงในดอทคอม โดยเลือกหมวดหมู่ร้านกาแฟ ที่อยู่ในพื้นที่กรุงเทพจำนวน 11,299 บทวิจารณ์ นำมารวบรวมไว้และส่งออกเป็นเอกสารประเภท Excel แล้วสร้างฟิลด์ที่ประกอบด้วย ชื่อผู้ใช้งาน (User), ค่าคะแนน (Rating), หัวข้อเรื่อง (Summary), ข้อความวิจารณ์ (Description), เวลา (Time) และชื่อร้านอาหาร (ShopName) ซึ่งฟิลด์ที่เลือกนำไปใช้สร้างแบบจำลอง คือ ข้อความวิจารณ์และค่าคะแนน

3.2 ขั้นตอนการเตรียมข้อมูล (Preprocessing)

นำข้อมูลที่ได้เข้าสู่การคัดกรองข้อมูล (Text Cleaning) เป็นการจัดการลบข้อมูลที่ไม่สมบูรณ์ ข้อมูลที่ขาดหายไป ข้อมูลที่เกินมาที่มีผลต่อการนำไปใช้งาน ต่อมาเป็นขั้นตอนกระบวนการจัดการเอกสาร (Process Document) นำข้อความวิจารณ์ไปแยกคำด้วย LexTo จะได้คำถูกตัดที่มีความหมายหลังจากนั้นนำเข้าเครื่องมือ RapidMiner เพื่อหาความถี่จากคำทั้งหมดที่ตัดมาได้ ซึ่งจำนวนคำที่ตัดได้ทั้งหมด 25,023 คำ แล้วค่อยเลือกคำที่มีความหมายจากคำที่ได้ เช่น “รสชาติ” ไปหาคำที่มีความหมายต่อท้ายเพิ่มในข้อความบทวิจารณ์ ได้ออกมาเป็นคำว่า “รสชาติดี” เป็นต้น แล้วนำมาหาคำที่ความหมายเหมือนกันแต่เขียนคำที่ไม่ถูกต้องซึ่งเกิดจากการพิมพ์ภาษาของแต่ละคนไม่เหมือนกันมีทั้งพิมพ์ถูกและพิมพ์ผิด จากนั้นจัดกลุ่มคำให้อยู่ในหมวดเดียวกัน เช่น คำว่า “รสชาต” และ “รสชาด” มาจัดเป็นคำเดียวกันได้เป็น “รสชาติ” และนับจำนวนความถี่รวมเข้าด้วยกัน จนได้ Words List ที่เป็นภาษาไทย ดังแสดงในตารางที่ 2 และเรียงคำด้วยความถี่ที่เกิดขึ้นจากมากไปน้อย จากนั้นนำ Words List และข้อมูลที่เตรียมไว้ ไปเข้าโปรแกรมที่เขียนด้วยภาษา C# จะเห็นว่า Words List คำไทยในตารางที่ 2 มีการเรียงความถี่มากที่สุดไปน้อยที่สุดด้วย 50 คุณลักษณะลงในตารางที่ 1 โดยโปรแกรมจะตรวจสอบคำที่ค้นพบใน Words List มีคำใดที่ตรงกับในข้อความ ถ้าพบคำในบทวิจารณ์จะแสดงเป็น “true” แต่ถ้าไม่พบคำจะแสดงเป็น “false” และกำหนดให้ค่าคะแนนระหว่าง 3-5 จัดกลุ่มความคิดเห็นเป็น “Positive” ค่าคะแนนระหว่าง 0-2 จัดกลุ่มความคิดเห็นเป็น “Negative” ดังตารางที่ 1

ตารางที่ 1: รูปแบบข้อมูลในบทวิจารณ์ที่ตรวจสอบคำที่ถูกค้นพบ

บทวิจารณ์	รสชาติดี	บรรยากาศดี	ประทับใจ	…	ถูกปาก	คะแนน	ความคิดเห็น (Opinion)
บทวิจารณ์ที่ 1	true	true	true	…	false	4	Positive
บทวิจารณ์ที่ 2	false	true	true	…	false	3	Positive
บทวิจารณ์ที่ 3	true	true	false	…	true	4	Positive
…	…	…	…	…	…	…	…
บทวิจารณ์ที่ n	false	false	false	…	true	1	Negative

ขั้นตอนแปลงข้อมูล (Text Transformation) เป็นการแปลงข้อมูลนำเข้าจากภาษาไทยเป็นภาษาอังกฤษ เพื่อเตรียมนำเข้าสู่ขั้นตอนการจำแนกข้อความเนื่องจากโปรแกรม Weka ไม่รองรับภาษาไทย จึงต้องแปลงข้อมูลนำเข้าเพื่อให้ง่ายต่อการนำไปประมวลผลในการจำแนกข้อความ ดังตารางที่ 2

ตารางที่ 2: ข้อมูลการนำเข้าจำนวน 50 คุณลักษณะ

คุณลักษณะ	ความถี่	คุณลักษณะ	ความถี่
1) รสชาติดี (Good Taste)	5,500	12) ร้านเล็ก (Small Café)	573
2) บรรยากาศดี (Good Atmosphere)	1,978	13) ราคาไม่แพง (Low Price)	547
3) ประทับใจ (Impressing)	1,706	14) นั่งชิล (Chill)	536
4) ตกแต่งร้านสวย-น่ารัก (Good Interior)	1,324	15) ร้านน่านั่ง (Nice Sit)	483
5) บริการช้า (Slow)	1,204	16) สะดวก (Convenient)	482
6) ราคาแพง (High Price)	891	17) ถูกใจ (Congenial)	455
7) ลูกค้าเยอะ (A lot of people)	800	18) รอคิว (Wait Queue)	427
8) ร้านอยู่ใกล้ (Near)	793	19) ร้านอยู่ไกล (Far Off)	421
9) บริการดี (Good Service)	764	20) ที่นั่งสบาย (Comfortable Sitting)	394
10) ไม่ประทับใจ (Unimpressing)	730	21) มีเมนูหลากหลาย (Various Stuff)	335
11) มีที่จอดรถ (Parking Area)	575	22) เงียบสงบ (Peaceful)	305

ตารางที่ 2: ข้อมูลการนำเข้าจำนวน 50 คุณลักษณะ (ต่อ)

คุณลักษณะ	ความถี่	คุณลักษณะ	ความถี่
23) คุ้มค่า (Worth)	301	37) กาแฟอร่อย (Delicious Coffee)	213
24) อาหารอร่อย (Delicious Food)	294	38) สมราคา (Reasonable Price)	210
25) ไม่อร่อย (Not Delicious)	289	39) มีเอกลักษณ์ (Unique)	184
26) ไม่เลี่ยน (Not to greasy)	280	40) ไม่มีที่จอด (No Parking Area)	172
27) น่ากิน-น่าทาน (Look Good Stuff)	268	41) โปรโมชั่น (Promotion)	166
28) ขนมอร่อย (Delicious Sweets)	267	42) อาหาร-บริการแย่ (Terribly)	148
29) สะอาด (Clean)	246	43) ลูกค้าไม่เยอะ (Less People)	146
30) เมนูแนะนำ (Recommended Menu)	242	44) ติดริมถนน (On the roadside)	141
31) เค้กอร่อย (Delicious Cake)	241	45) เมนูใหม่ (New Menu)	128
32) น่าสนใจ (Interesting)	220	46) พนักงานน่ารัก (Cute Staff)	124
33) กาแฟหอม (Good Smell Coffee)	219	47) อึดอัด (Awkward)	124
34) ยิ้มแย้ม (Smily)	216	48) ไม่มีรสชาติ (Seat is a lot)	122
35) บรรยากาศอบอุ่น-เป็นกันเอง (Warm Atmosphere)	216	49) เค้กนุ่ม (Soft Cake)	122
36) มีปลั๊กไฟ (Wall Plug)	214	50) ถูกปาก (Palatable)	116

3.3 ขั้นตอนการคัดเลือกคุณลักษณะของข้อมูล

เป็นวิธีที่ใช้เพื่อลดคุณลักษณะของข้อมูลและกำหนดลำดับความสำคัญของคุณลักษณะของข้อมูลใหม่เพื่อเพิ่มประสิทธิภาพในการจำแนกกลุ่มและได้ปัจจัยที่มีอำนาจในการจำแนกผลลัพธ์ที่ส่งผลในกลุ่มผลตอบรับเชิงบวกและลบมากที่สุดในการทดลองนี้ได้ใช้การคัดเลือกคุณลักษณะของข้อมูล 3 วิธี ได้แก่ CFS, InfoGain, และผู้วิจัยคัดเลือกจากคำที่มีความถี่ที่ปรากฏ โดยค่าที่ใช้ปรับพารามิเตอร์ในขั้นตอนวิธี C4.5 ได้แก่ ConfidenceFactor (CF) เท่ากับ 0.25, Numfold (Nf) เท่ากับ 3, unpruned เท่ากับ F โดยมีการตัดแต่งกิ่ง และค่าที่เปลี่ยนแปลงในแต่ละรอบที่ทดสอบคือ minNumObj (mNO) ปรับตั้งแต่ 2-10 ดังตารางที่ 4

3.4 ขั้นตอนการจำแนกข้อความ (Classification Model)

จากข้อมูลทั้งหมด 11,299 บทวิจารณ์ ได้แบ่งการจำแนกข้อความในรอบแรกเพื่อตรวจสอบความถูกต้อง พบว่า ข้อมูลไม่สมดุลกัน คือ จำนวนบทวิจารณ์ที่ตอบรับเป็น Positive จำนวน 10,767 บทวิจารณ์ และ Negative จำนวน 532 บทวิจารณ์ นำเข้าแบบจำลองได้ค่าความถูกต้องคิดเป็น 95.22% ได้ค่า (TP Rate=Negative) เท่ากับศูนย์ แต่ (TP Rate=Positive) เท่ากับหนึ่ง ทำให้ทราบว่าข้อมูลกระจุกตัวแต่ในกลุ่ม Positive มากกว่า เนื่องจากวิจัยนี้สนใจทั้งสองกลุ่ม จึงทำการสุ่มเพิ่มตัวอย่างกลุ่มน้อย (SMOTE) เป็นขั้นตอนวิธีการแก้ปัญหาข้อมูลที่ไม่สมดุลกัน โดยค่าที่ใช้ปรับพารามิเตอร์ของ SMOTE คือ percentage เท่ากับ 1,900 และ nearestNeighbor เท่ากับ 10 จนได้กลุ่มตัวอย่างความคิดเห็นเพิ่มเป็น Positive จำนวน 10,767 บทวิจารณ์ และNegative จำนวน 10,640 บทวิจารณ์ (ซึ่งจำนวนของกลุ่ม Negative ที่ได้ออกมาถูกปรับให้ใกล้เคียงกับกลุ่ม Positive มากที่สุด) รวมเป็น 21,407 บทวิจารณ์ ด้วยจำนวนทั้งหมด 50 คุณลักษณะ นำข้อมูลชุดข้อมูลการฝึกสอนเข้าแบบจำลองโดยใช้เทคนิคต้นไม้ตัดสินใจ C4.5 เพื่อวิเคราะห์หาปัจจัยที่มีผลต่อการทำให้เกิดผลตอบรับเชิงบวกและผลตอบรับเชิงลบ ได้ทำการแบ่งข้อมูลออกเป็นชุดข้อมูลการฝึกสอน (Training Set) และชุดข้อมูลทดสอบ (Testing Set) โดยใช้วิธี K-Fold Cross Validation โดยกำหนดค่า K เท่ากับ 5 ทำการทดสอบการจำแนกข้อความเพื่อให้แบบจำลองของต้นไม้ตัดสินใจที่ดีที่สุดออกมา

3.5 ขั้นตอนการทดสอบประสิทธิภาพ(Model Evaluation)

นำปัจจัยที่ได้จากการคัดเลือกคุณลักษณะของข้อมูลไปทดสอบเพื่อวัดประสิทธิภาพด้วยชุดข้อมูลทดสอบ แล้วทำการเปรียบเทียบวิธีเลือกคุณลักษณะข้อมูลพิจารณาด้วยการประเมินค่าความถูกต้อง (Accuracy) จากสมการที่ 4 ว่ามีความถูกต้องมากน้อยเพียงใด โดยสรุปข้อมูลที่มีผลตอบรับเชิงบวกและเชิงลบ

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+FP+FN+TN}} \quad (4)$$

ตารางที่ 3: ผลที่ใช้จำแนกบทวิจารณ์

		Predicted Class	
	กลุ่มจำแนก	ด้านบวก	ด้านลบ
Actual	ด้านบวก	True Positive (TP)	False Negative (FN)
	ด้านลบ	False Positive (FP)	True Negative (TN)

4. ผลการดำเนินงาน

ในการคัดเลือกปัจจัยก่อนการจำแนกข้อความ ทำการทดสอบหาปัจจัยโดยวิธี CFS ได้คัดเลือกมาทั้งหมด 35 ปัจจัย และ InfoGain ได้คัดเลือกมาทั้งหมด 38 ปัจจัย ส่วนความถี่ของคำที่ปรากฏได้คัดเลือกมาทั้งหมด 50 ปัจจัย นำมาเข้าแบบจำลองเพื่อทำการทดสอบจึงได้ผลการสร้างแบบจำลองด้วยต้นไม้ตัดสินใจที่ผ่านการคัดเลือกคุณลักษณะ และเปรียบเทียบค่าความถูกต้องในแต่ละกลุ่ม เป็นผลให้ปัจจัยที่ส่งผลต่อผลตอบรับเชิงบวกและลบ ซึ่งการปรับค่าพารามิเตอร์

CF = 0.25, Nf = 3, unpruned = F และ mNO = 3 ให้ค่าความถูกต้องดีที่สุด ด้วยวิธีคัดเลือกจากคำที่มีความถี่ที่ปรากฏในบทวิจารณ์สูง ผลที่ได้ในตารางที่ 4

ตารางที่ 4: การคัดเลือกข้อมูลและเปรียบเทียบค่า Accuracy

No.	ปรับค่า	Number of leave / Size of tree			Accuracy (%)		
	CF = 0.25, Nf = 3, unpruned = F mNO	CFS	InfoGain	ความถี่ของคำ	CFS	InfoGain	ความถี่ของคำ
1	2	88/175	93/185	116/231	84.21	84.62	85.92
2	3	81/161	89/177	97/193	84.22	84.47	85.98
3	4	78/155	85/169	89/177	84.13	84.36	85.84
4	5	72/143	75/149	85/169	84.03	84.27	85.75
5	6	70/139	68/135	80/159	83.93	84.16	85.59
6	7	62/123	62/123	73/145	83.91	84.14	85.53
7	8	59/117	58/115	69/137	83.89	84.11	85.49
8	9	58/115	57/113	67/133	83.76	84.02	85.36
9	10	53/105	51/101	85/169	83.73	84.00	85.75

ตารางที่ 5: ค่าความถูกต้องที่ดีที่สุดของแต่ละวิธีการคัดเลือกข้อมูล

	CFS	InfoGain	ความถี่ของคำ
Accuracy	84.22	84.62	85.98
TP Rate(class=Positive)	0.731	0.739	0.766
TN Rate(class=Negative)	0.955	0.953	0.955

จากตารางที่ 5 แสดงให้เห็นว่าแต่ละเทคนิคของการคัดเลือกคุณลักษณะของข้อมูลมีการให้ลำดับความสำคัญของคุณลักษณะที่แตกต่างกันออกไป ซึ่งกลุ่มผลตอบรับเชิงบวก (TP Rate=Positive) และเชิงลบ (TN Rate=Negative) มีค่าความถูกต้องสูงสุดคือ 0.766 และ 0.955 ตามลำดับและค่าความถูกต้อง (Accuracy) รวมทั้งสองกลุ่มคือ 85.98% โดยการคัดเลือกคุณลักษณะปัจจัยที่ได้จากความถี่ที่ปรากฏสูงจากต้นไม้ตัดสินใจที่ดีที่สุดคือ Low Price, Parking Area, Comfortable Sitting, Various Stuff, Delicious Food, Convenient, Delicious Coffee, Not to greasy, Far Off, Peaceful, Delicious Cake, Good Smell Coffee, Congenial, Delicious Sweets, Recommended Menu, Promotion, Warm Atmosphere, Worth, Small Cafe, A lot of people, Near, Terribly, Not Delicious และ High Price เป็นผลให้

ตารางที่ 6: ตัวอย่างกฎต้นไม้ตัดสินใจด้านผลตอบรับเชิงบวก

กฎ	กฎต้นไม้ตัดสินใจด้านผลตอบรับเชิงบวก	ผลลัพธ์
1	If Low Price = True	Positive
2	If Low Price = False And Parking Area = True	Positive
3	If Low Price = False And Parking Area = False And Comfortable Sitting = True	Positive
4	If Low Price = False And Parking Area = False And Comfortable Sitting = False And Various Stuff = True	Positive

ตารางที่ 7: ตัวอย่างกฎต้นไม้ตัดสินใจด้านผลตอบรับเชิงลบ

กฎ	กฎต้นไม้ตัดสินใจด้านผลตอบรับเชิงลบ	ผลลัพธ์
1	If Small Café = True And A lot of people = True	Negative
2	If Near = True And Terribly = True	Negative
3	If Not Delicious = True And Slow = True And High Price =True	Negative

ได้กฎการตัดสินใจที่ดีที่สุด 21 กฎ ซึ่งแสดงตัวอย่างของกฎที่ส่งผลต่อการวิจารณ์ของลูกค้า ดังตารางที่ 6 และตารางที่ 7

จากการวิเคราะห์กฎจากต้นไม้ตัดสินใจดังกล่าวได้มีลำดับคุณลักษณะปัจจัยที่ส่งผลต่อการวิจารณ์เชิงบวกและเชิงลบตามลำดับของระดับของต้นไม้ตัดสินใจ ซึ่งแสดงผลตามตารางที่ 8 ดังต่อไปนี้

ตารางที่ 8: ลำดับคุณลักษณะปัจจัยที่ส่งผลต่อการวิจารณ์จากต้นไม้ตัดสินใจที่ดีที่สุด

ลำดับที่	ปัจจัยที่ส่งผลเชิงบวก	ปัจจัยที่ส่งผลเชิงลบ
1.	ราคาถูก (Low Price)	ร้านเล็ก (Small Café)
2.	มีที่จอดรถ (Parking Area)	ลูกค้าเยอะ (A lot of people)
3.	ที่นั่งสบาย (Comfortable Sitting)	ร้านอยู่ใกล้ (Near)
4.	มีเมนูหลากหลาย (Various Stuff)	อาหาร-บริการแย่ (Terribly)
5.	อาหารอร่อย (Delicious Food)	ไม่อร่อย (Not Delicious)
6.	สะดวก (Convenient)	ราคาแพง (High Price)
7.	กาแฟอร่อย (Delicious Coffee)	
8.	ไม่เลี่ยน (Not to greasy)	
9.	ร้านอยู่ไกล (Far Off)	
10.	เงียบสงบ (Peaceful)	
11.	เค้กอร่อย (Delicious Cake)	
12.	กาแฟหอม (Good Smell Coffee)	
13.	ถูกใจ (Congenial)	
14.	ขนมอร่อย (Delicious Sweets)	
15.	เมนูแนะนำ (Recommended Menu)	
16.	โปรโมชั่น (Promotion)	
17.	บรรยากาศอบอุ่น-เป็นกันเอง (Warm Atmosphere)	
18.	คุ้มค่า (Worth)	

หลังจากจำแนกข้อความต้นไม้ตัดสินใจ พบว่า มีบางข้อความที่ยังจำแนกผิดพลาด เมื่อนำข้อความมาวิเคราะห์จากกฎที่ได้ ตัวอย่างในตารางที่ 9 บทวิจารณ์ที่สาม ให้เงื่อนไขตรงกับกฎที่ 1 “If Low Price = True” แต่ให้ผลลัพธ์เป็น Negative เมื่อเทียบผลลัพธ์จากกฎจริง สาเหตุที่ยังมีการจำแนกผิดพลาด เพราะว่าปัจจัยเพียงตัวเดียวที่ใช้เป็นเงื่อนไขไม่สามารถบอกข้อความทั้งหมด ทำให้คำตอบที่ได้อาจจะไม่ถูกต้องทั้งหมด เพราะมีหลายตัวแปรที่ส่งผลต่อผลลัพธ์เชิงบวกและลบ ซึ่งตัวอย่างที่นำเสนอถือว่ายังมีความผิดพลาดเกิดขึ้นเป็นส่วนน้อย เมื่อคิดจากบทวิจารณ์ทั้งหมดที่ให้ผลลัพธ์ถูกต้องตรงกับกฎที่ได้เป็นส่วนใหญ่

ตารางที่ 9: ตัวอย่างบทวิจารณ์เมื่อเปรียบเทียบจากกฎที่ได้

บทวิจารณ์	คะแนน	ความคิดเห็น	ผลจากกฎ	
			กฎ	ผลลัพธ์
เหมาะสำหรับสมัยนี้ที่คนออกไปเล่นเกมย่านอนุเสาวรีย์ และต้องการร้านนั่งพักพร้อมชาร์ตแบตไปในตัว! หรือคนที่ต้องการทำงานกลุ่มพร้อมชาร์ตโน้ตบุค (ต้องลองสังเกตุ ไม่ได้มีปลั๊กทุกที่นั่ง) ที่สำคัญอาหาร**ราคาไม่แพง** แพนเค้กรสชาติไม่เลว มีส่วนลดค่อนข้างบ่อย แต่ถ้านั่งแช่นาน อาจจะโดนสายตาของพนักงานมองมาได้	4	Positive	1	Positive
ร้านนี้เป็นเหมือนร้านที่นั่งปั่นงานและนั่งอ่านหนังสือของทั้งเด็กมัธยมและเด็กมหาลัย ร้านที่โต๊ะเยอะ แต่ที่ไม่ค่อยว่าง ข้อดีของที่นี่คงเป็นเปิด 24 ชั่วโมง ให้บรรยากาศเหมือนห้องสมุด อาหารและเครื่องดื่มเค้า**ไม่ค่อยอร่อย** และ**ราคาค่อนข้างสูง** แต่เค้าก็คงไม่กำไรมากนัก เพราะคนส่วากมานั่งยาวๆและใช้ไฟยาวๆเครื่องดื่มได้**ช้า** แต่ร้านนี้เหมาะกับอ่านหนังสือและปั่นงานยามดึกจริงๆ	2	Negative	3	Negative
ร้านกาแฟสไตล์แปลก ไม่มี wifi ให้ใช้ ที่นั่งไม่ค่อยสบายเท่าไหร่ โต๊ะเล็กนั่งทำงานก็ลำบาก แต่พนักงานดูแลเอาใจใส่ดีพอสมควร กาแฟรสชาติธรรมดา ราคากลางๆ**ไม่แพง**ไม่ถูก ชนิดของอาหารและเครื่องดื่มค่อนข้างน้อย ที่ติดมากคือเก้าอี้ที่นั่งไม่ค่อยสบาย ถ้าใครอยากลองของแปลก ก็แนะนำครับ	2	Negative	1	Positive
มีเค้กและเครื่องดื่มที่หลากหลาย แป้งเค้กที่นี่จะนุ่มมาก ชีสพายจะหนามากก บานอฟฟี่ก็อร่อย คาราเมลเข้มมาก อีกอย่างที่น่ากินคือเค้กทูโทน เนื้อนุ่มมาก แต่เสียดายที่**ร้านเล็ก**ไปนิดเวลา**คนมาเยอะ** จะไม่มีที่นั่งเลย และเค้กจะหมดเร็ว เพราะอร่อย	4	Positive	1	Negative

5. สรุปผล

จากผลการวิจัยสรุปได้ว่าการคัดเลือกปัจจัยโดยใช้ความถี่ที่ปรากฏสูงนั้นจะส่งผลให้การจำแนกข้อความบทวิจารณ์ที่เป็นทั้งด้านบวกและด้านลบนั้นได้ดีที่สุดและสามารถจำแนกปัจจัยทางด้านบวกและด้านลบได้ดีกว่าวิธีการอื่น แสดงให้เห็นว่าการสกัดปัจจัยหรือแอทริบิวต์แบบที่ใช้ความถี่สูงและปรากฏจริงในเอกสารจะสามารถจำแนกได้ดีและรองรับกับข้อความที่เป็นแบบไม่มีโครงสร้างและการเกิดภาษาใหม่ๆ ที่เกิดขึ้นในอนาคต

งานวิจัยในอนาคตผู้วิจัยจะมีการนำเสนอการคัดเลือกปัจจัยแบบยืดหยุ่นและมีการใช้การสื่อความหมายของคำเข้ามาช่วยในการจำแนกข้อความได้ดียิ่งขึ้น

เอกสารอ้างอิง

[1] พัชรนิกานต์ พงษ์ธนู, วรารัตน์ รุ่งวรวุฒิ, งามนิจ อาจอินทร์, และ สมจิตร อาจอินทร์. “วิเคราะห์ความพึงพอใจของลูกค้าจากข้อความคำแนะนำโดยการทำเหมืองความคิดเห็น” *The 2012 4th International Conference on Knowledge and Smart Technologies (KST)*. หน้า 53-60. มหาวิทยาลัยบูรพา, 2555.

[2] C. Jeeranan., “Analysis on Factors Affecting Normal-Grade Student Dismissal Using Decision Tree”. *SNRU Journal of Science and Technology*., Vol. 8, No. 2 (2016).

[3] B. Pang and L. Lee, "Opinion Mining and Sentiment Analysis," *Foundations and Trends in Information Retrieval* Vol. 2, Nos. 1-2 (2008), pp. 1-135.

[4] A. Kongthon, N. Angkawattanawit, C. Sangkeettrakarn, P. Palingoon and C. Haruechaiyasak, "Using an opinion mining approach to exploit Web content in order to improve customer relationship management," *PICMET 2010 Technology Management for Global Economic Growth*, Phuket, 2010, pp. 1-6.

[5] กิตติภักดี วัฒนะกุล. “คัมภีร์ระบบสนับสนุนการตัดสินใจและระบบผู้เชี่ยวชาญ”, กรุงเทพฯ : เคทีพีคอมพ์แอนด์คอนซัลท์, 2551.

[6] J. Ross Quinlan. C4.5: Programs for Machine Learning. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1993.

[7] N. V. Chawla K. W. Bowyer L. O. Hall W. P. Kegelmeyer "SMOTE: synthetic minority over-sampling technique," *Journal of Artificial Intelligence Research vol.* 16 Jan. 2002 pp. 321-357.

[8] Z. Li, L. Liu and C. Li, "Analysis of customer satisfaction from Chinese reviews using opinion mining," *6th IEEE International Conference on Software Engineering and Service Science (ICSESS)*, Beijing, 2015, pp. 95-99.

[9] สุทธิวรรณ จรรยาพัฒนากุล และ มาลีรัตน์โสดานิล, “ระบบวิเคราะห์ข้อความจากเว็บบอร์ดร้านอาหารโดยใช้เหมืองข้อมูลความคิดเห็น” *การประชุมทางวิชาการระดับชาติ ด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศ ครั้งที่ 9*. กรุงเทพฯ, 2556.

[10] Chumwatana.T, "Using sentiment analysis technique for analyzing Thai customer satisfaction from social media," *5th International Conference on Computing and Informatics (ICOCI) 2015*, 11-13 August 2015, Istanbul, Turkey.

การวิเคราะห์การทำนายการลาออกกลางคันของนักศึกษา ระดับประกาศนียบัตรวิชาชีพชั้นสูง โดยใช้เทคนิคเหมืองข้อมูล

Predictive Analytic for Student Dropout in High Vocational Certificate Using Data Mining Technique

ศักรินทร์ ภูพันนา (Sakkarin Phupanna)[1] ปวีณา วันชัย (Paweena Wanchai)[2]
วรารัตน์ สงฆ์แป้น (Wararat Songpan)[3] และนันท์นภัส เบญจมาศ (Nunnapus Benjamas)[4]
ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น
[1]Sphupanna@kkumail.com, [2]Wpaweena@kku.ac.th, [3]Wararat@kku.ac.th, [4]Nunnapus@kku.ac.th

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพในการจำแนกข้อมูลของอัลกอริทึม C4.5, Naïve Bayesian Learning และ Multilayer Perception เพื่อทำนายการลาออกกลางคันของนักศึกษาและวิเคราะห์หาปัจจัยที่เกี่ยวข้อง โดยใช้ข้อมูลพื้นฐานของนักศึกษาวิทยาลัยเทคโนโลยีภาคตะวันออกเฉียงเหนือ จังหวัดขอนแก่น ระหว่างปีการศึกษา 2556-2558 จำนวน 2,270 ชุดข้อมูล ใช้เทคนิคการทำ SMOTE แก้ไขปัญหาการไม่สมดุลกันของชุดข้อมูล โดยปรับค่า nearest Neighbors = 5 ทดสอบผลลัพธ์ด้วย 10-fold cross-validation และวัดประสิทธิภาพด้วยค่า Accuracy ผลการวิเคราะห์พบว่า Multilayer Perception ให้ค่าความแม่นยำสูงสุด 94.45% มากกว่า C4.5 และ Naïve Bayesian Learning ที่ร้อยละ 93.74% และ 83.74% ตามลำดับ และมีปัจจัยที่เกี่ยวข้องสูงสุด 5 อันดับ ได้แก่ ประเภทโรงเรียนเดิม ประเภทหลักสูตร เกรดเฉลี่ยก่อนเข้าศึกษา การกู้ยืมกองทุนเพื่อการศึกษาและสาขาวิชา

คำสำคัญ: การลาออกกลางคัน การเรียนรู้เบย์อย่างง่าย ต้นไม้ตัดสินใจ โครงข่ายประสาทเทียม

Abstract

The aim of this paper is to compare the performance of different classification techniques included C4.5, Naïve Bayesian Learning and Multilayer Perception algorithms. The algorithms were used to predict and analyze factors influencing student's decision to drop out. The data set was collected from Northeastern Technological College, Khon Kaen Province during the 2013-2015 academic years. The total number of sample data was 2,270 records. The SMOTE technique was also used for solving imbalanced problems of the data set. The nearest Neighbors was adjusted as 5. The cross-validation with 10 folds was used to evaluate the prediction accuracy. The results show that the performance of Multilayer Perception is 94.45% which higher than the C4.5 and Naïve Bayesian Learning, with 93.74% and 83.74% accordingly. The findings also indicate that student's decision to drop out significantly influenced by type of school, type of course, first grade, student loan and major.

Keyword: Student drop out, Naïve Bayesian Learning, Decision Tree, Multilayer Perception.

1. บทนำ

การจัดการการศึกษาด้านวิชาชีพ เป็นปัจจัยสำคัญในการพัฒนาบุคคลให้สอดคล้องกับตลาดแรงงาน เน้นการแก้ไขปัญหา สร้างองค์ความรู้ในอาชีพ มีคุณภาพ และมีเจตคติที่ดี

แต่ในสถานศึกษาหลายแห่งพบว่า ปัญหาการลาออกกลางคันของนักศึกษาเป็นสิ่งที่ต้องหาแนวทางในการแก้ไข จากการสำรวจและติดตามการลาออกกลางคันของผู้เรียนในระบบปกติ ปีการศึกษา 2552 – 2554 จำนวนสถานศึกษากลุ่มตัวอย่าง 353 แห่ง คิดเป็นร้อยละ 85 ของสถานศึกษา 416 แห่ง มีจำนวนนักศึกษาที่ลาออกกลางคัน คิดเป็นร้อยละ 27.52 ของภาพรวมทั้งหมด [1] โดยส่วนใหญ่เป็นนักศึกษาระดับประกาศนียบัตรวิชาชีพชั้นสูง (ปวส.) จากฐานข้อมูลวิทยาลัยเทคโนโลยีภาคตะวันออกเฉียงเหนือ ในปีการศึกษา 2556 – 2558 พบว่ามีจำนวนนักศึกษา ปวส. ทั้งหมด 3,146 คน จำนวนนักศึกษาที่ลาออกกลางคัน 326 คน คิดเป็นร้อยละ 10.36 ของภาพรวม [2]

ในปัจจุบันการทำเหมืองข้อมูลทางการศึกษา (Education Data Mining) เป็นที่นิยมอย่างมากโดยเฉพาะการจำแนกประเภทข้อมูล (Classification) เพื่อหาความสัมพันธ์ของข้อมูลและเปรียบเทียบประสิทธิภาพของอัลกอริทึม ดังงานวิจัยของ Pimpa Cheewaprakobkit ที่เปรียบเทียบ C4.5 กับ Neural Network เพื่อหาปัจจัยที่ส่งผลต่อระดับผลการเรียน ผลปรากฏว่า C4.5 ให้ค่าประสิทธิภาพสูงสุด [3] แตกต่างกับงานวิจัยของ Parneet Kaura และคณะ ที่เปรียบเทียบประสิทธิภาพในการจำแนกข้อมูลของอัลกอริทึม Multilayer Perception, Naïve Bayes, SMO, J48 และ REPTree ผลปรากฏว่า Multilayer Perception ให้ค่าประสิทธิภาพสูงสุด [4] และแตกต่างกับงานวิจัยของ S. B. Kotsiantis และคณะที่เปรียบเทียบ Naïve Bayes, 3-NN, Ripper, C4.5 และWinnow ผลปรากฏว่า Naïve Bayes ให้ค่าประสิทธิภาพสูงสุด [5] จากการทบทวนงานวิจัยที่เกี่ยวข้อง แม้จะมีการเปรียบเทียบอัลกอริทึมจากงานวิจัยต่างๆ ก็ไม่สามารถระบุได้ว่าอัลกอริทึมใดที่จำแนกประเภทข้อมูลได้ดีที่สุด

จากเหตุผลดังกล่าว งานวิจัยฉบับนี้จึงเสนอการวิเคราะห์การทำนายการลาออกกลางคันของนักศึกษาโดยใช้เทคนิคการทำเหมืองข้อมูลด้วยเทคนิควิธีต้นไม้ตัดสินใจ C4.5, เทคนิควิธีการเรียนรู้เบย์อย่างง่าย (Naïve Bayesian Learning) และเทคนิควิธีโครงข่ายประสาทเทียมแบบ Multilayer Perception (MLP) เพื่อเปรียบเทียบประสิทธิภาพของโมเดลและหาปัจจัยที่เกี่ยวข้อง เพื่อใช้เป็นแนวทางในการป้องกันและแก้ไขปัญหาการลาออกกลางคันของนักศึกษาในแต่ละปีการศึกษาต่อไป

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 อัลกอริทึม C4.5

เป็นอัลกอริทึมที่ถูกพัฒนามาจาก ID3 ใช้สร้างต้นไม้ตัดสินใจ (Decision Tree) โดยจะคัดเลือกคุณลักษณะ (Attribute) ที่มีความสัมพันธ์กันมากที่สุดมาไว้เป็นโหนดบนสุดเรียกว่า root node และหาคุณลักษณะที่มีความสัมพันธ์ถัดไปเรื่อยๆ ใช้ตัววัดเรียกว่า Information Gain เช่น ค่าของกลุ่มข้อมูล คือค่าคาดคะเนที่กลุ่มตัวอย่างต้องใช้จำนวนบิตในการแยกกลุ่มข้อมูล [6] นิยามดังสมการ

$$I(p,n) = -\frac{p}{p+n}\log_2\left(\frac{p}{p+n}\right) - \frac{n}{p+n}\log_2\left(\frac{n}{p+n}\right) \qquad (1)$$

ค่าคาดคะเนของข้อมูล (Entropy)

$$E(A) = \sum_{I=0}^{V} \frac{p_1 + n_1}{p+n} I(p_1, n_1) \qquad (2)$$

ค่าเกนข้อมูล (Data Gain) ดังสมการ

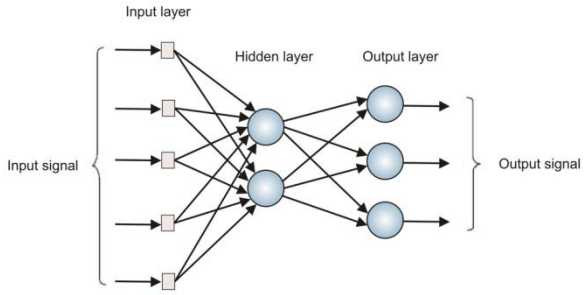
$$Gain(A) = I(p,n) - E(A) \qquad (3)$$

ซึ่งอัลกอริทึม C4.5 เหมาะสำหรับใช้วิเคราะห์ข้อมูลได้ทุกรูปแบบ และให้ประสิทธิภาพในการจำแนกข้อมูลได้ดี [12]

2.2 โครงข่ายประสาทเทียมแบบเพอร์เซ็ปตรอนหลายชั้น

เป็นโครงข่ายประสาทเทียมแบบหลายชั้น โหนดในแต่ละชั้นมีค่าน้ำหนักไม่เท่ากัน เหมาะกับงานที่มีความซับซ้อน โดยฝึกฝนแบบมีผู้สอน (Supervise) และใช้ขั้นตอนการส่งค่าย้อนกลับ (Backpropagation) ซึ่งประกอบด้วย 1.การส่งผ่านไปข้างหน้า (Forward Pass) คือ ข้อมูลจะเข้าโครงข่ายประสาทเทียมในชั้นข้อมูลเข้าและจะส่งผ่านไปอีกชั้นจนถึงชั้นข้อมูลออก และ 2.การส่งผ่านย้อนกลับ (Backward Pass) คือ ค่าน้ำหนักการเชื่อมต่อจะถูกปรับให้สอดคล้องกับกฎการแก้ไขข้อผิดพลาด คือผลต่างของผลตอบที่แท้จริงกับผลตอบเป้าหมายเกิดเป็นสัญญาณผิดพลาด ซึ่งสัญญาณที่ผิดพลาดนี้จะถูกส่งย้อนกลับเข้าโครงข่ายประสาทเทียมในทิศทางตรงกันข้ามกับการเชื่อมต่อ และ

ค่าน้ำหนักของการเชื่อมต่อจะถูกปรับจนกระทั่งผลตอบที่แท้จริงเข้าใกล้ผลตอบเป้าหมาย [13]



ภาพที่ 1: แสดงโครงข่าย MLP

2.3 การเรียนรู้เบย์อย่างง่าย

เป็นเทคนิควิธีที่ใช้อาศัยหลักความน่าจะเป็นเข้ามาช่วยในการคำตอบของประเภทตัวอย่างใหม่ซึ่งอยู่บนพื้นฐานของ Bayes Theorem และสมมติฐานที่ให้การเกิดของเหตุการณ์ต่างๆ ในการจัดกลุ่มข้อมูลที่เป็นอิสระต่อกัน (Independence) [11] คำนวณได้จากสมการ

$$P(h|D) = [P(D|h) \times P(h)] / P(D) \quad (4)$$

D คือ *ข้อมูลที่นำมาคำนวณแจกแจงความน่าจะเป็น posteriori probability ของเหตุการณ์ h*

P(h) คือ *ความน่าจะเป็นที่จะเกิดเหตุการณ์ H*

P(D คือ *ความน่าจะเป็นก่อนหน้าจากชุดข้อมูล D*

P(h|D) คือ *ความน่าจะเป็นของ h เมื่อรู้ D*

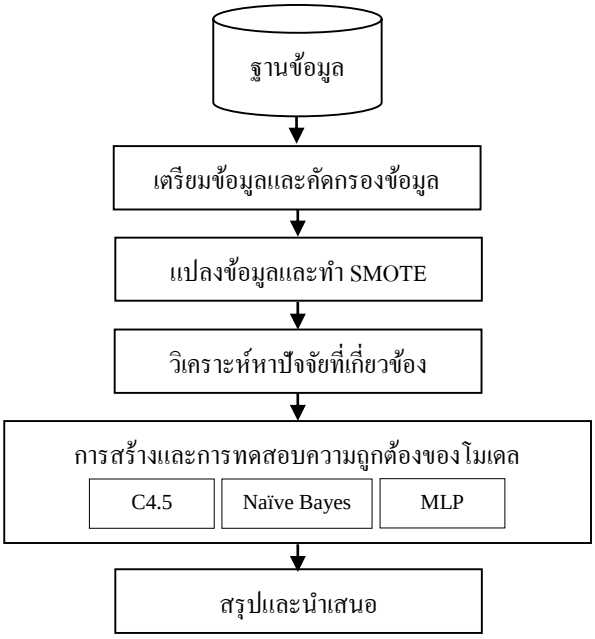
P(D|h) คือ *ความน่าจะเป็นของ D เมื่อรู้ h*

2.4 วิธีการสุ่มเพิ่มตัวอย่างกลุ่มน้อย

Synthetic Minority Over-sampling Technique (SMOTE) เป็นเทคนิควิธีที่ช่วยแก้ไขปัญหาการจำแนกข้อมูลที่ไม่สมดุลกันในแต่ละคลาสที่แตกต่างกันมาก เมื่อจำแนกข้อมูลจะทำให้ผลลัพธ์ที่ได้จากแบบจำลองเกิดการเรียนรู้และจำแนกข้อมูลไปในข้อมูลกลุ่มมากเท่านั้น มีหลักการทำงานโดยการสุ่มเพิ่มตัวอย่างขึ้นใหม่เฉพาะกลุ่มน้อยที่ใกล้เคียงตามจำนวน nearest Neighbors และจำนวนข้อมูลที่ต้องการให้สังเคราะห์เพิ่ม โดยกำหนดเป็นเปอร์เซ็นต์ (%) เพื่อให้เกิดการเรียนรู้และจำแนกข้อมูลในกลุ่มน้อยได้ดีขึ้น [8]

3. วิธีการดำเนินการวิจัย

ในงานวิจัยนี้ได้เสนอวิธีการทำเหมืองข้อมูลโดยใช้โปรแกรม Weka 3.6 ด้วยอัลกอริทึม C4.5, MLP และ Naïve Bayes เพื่อเปรียบเทียบประสิทธิภาพโมเดลและหาปัจจัยที่เกี่ยวข้องในการลาออกกลางคันของนักศึกษา



ภาพที่ 3: แสดงขั้นตอนการดำเนินงาน

3.1 การเตรียมข้อมูลและการคัดกรองข้อมูล

รวบรวมข้อมูลพื้นฐานของนักศึกษา จากฐานข้อมูลวิทยาลัยเทคโนโลยีภาคตะวันออกเฉียงเหนือ จังหวัดขอนแก่น ตั้งแต่ปีการศึกษา 2556-2558 คัดกรองเฉพาะข้อมูลที่มีความสมบูรณ์ทั้งหมด 1,178 ชุดข้อมูล มีแอทริบิวต์ที่เกี่ยวข้อง ดังตารางที่ 1

ตารางที่ 1: รายละเอียดของแอทริบิวต์ที่เกี่ยวข้อง

แอทริบิวต์	รายละเอียด	
Gender	เพศ	ชาย, หญิง
Age	อายุ	18-19, 20-21, 22-23, 24 ปีขึ้นไป
Major	สาขาวิชา	เทคโนโลยีคอมพิวเตอร์, โยธา, อิเล็กทรอนิกส์, เทคนิคยานยนต์, ไฟฟ้ากำลัง, เทคนิคการผลิต, การบัญชี, คอมพิวเตอร์ธุรกิจ, การตลาด
Course	หลักสูตร	ช่างอุตสาหกรรม, บริหารธุรกิจ

ตารางที่ 1: รายละเอียดของแอทริบิวต์ที่เกี่ยวข้อง (ต่อ)

TofC	ประเภทหลักสูตร	ปกติ, พิเศษ
MofO	อาชีพมารดา	เกษตรกรรม, แม่บ้าน, ลูกจ้างหรือรับจ้างทั่วไป, ค้าขายหรือธุรกิจส่วนตัว,ข้าราชการหรือพนักงานรัฐหรือรัฐวิสาหกิจ
FofO	อาชีพบิดา	เกษตรกรรม, พระสงฆ์, ลูกจ้างหรือรับจ้างทั่วไป, ค้าขายหรือธุรกิจส่วนตัว, ข้าราชการหรือพนักงานรัฐหรือรัฐวิสาหกิจ
TofS	ประเภทโรงเรียนเดิม	โรงเรียน, วิทยาลัย, กศน.
FG	เกรดเฉลี่ยก่อนเข้าศึกษา	< 2.00, 2.00-3.00, >3.00
SSN	จำนวนพี่น้องที่ศึกษา	1 คน, 2 คน, 3 คนขึ้นไป
SL	การกู้ยืมกองทุนเพื่อการศึกษ	กู้ยืม, ไม่กู้ยืม
Drop	สถานะ	ลาออก, ไม่ลาออก

3.2 การแปลงข้อมูลและการสุ่มเพิ่มตัวอย่างกลุ่มน้อย

เนื่องจากข้อมูลที่นำมาวิเคราะห์ในครั้งนี้ มีรูปแบบในการจัดเก็บทั้งตัวเลขและตัวอักษร จึงต้องแปลงให้อยู่ในรูปแบบเดียวกัน เช่น ประเภทหลักสูตร ปกติ เป็น TC1 พิเศษ เป็น TC2 หรือ ประเภทโรงเรียนเดิม โรงเรียน เป็น Type1 วิทยาลัย เป็น Type2 กศน. เป็น Type3 เป็นต้น

จากชุดข้อมูลทดสอบ สามารถแบ่งเป็นคลาส Yes (ลาออก) 52 ชุดข้อมูล และคลาส No (ไม่ลาออก) 1,126 ชุดข้อมูล ทำให้เกิดปัญหาในการจำแนกข้อมูลที่ไม่สมดุลกันในแต่ละคลาส ผู้วิจัยได้เลือกใช้เทคนิควิธี SMOTE กำหนดให้ค่า nearest Neighbors = 5 และ Percentage = 2100 เพื่อสุ่มเพิ่มคลาส Yes จาก 52 เป็น 1,144 ชุดข้อมูล เพราะเป็นคลาสที่สนใจและสุ่มเพิ่มจำนวนชุดข้อมูลได้ใกล้เคียงกับคลาส NO มากที่สุด ทำให้มีข้อมูลทั้งสิ้น 2,270 ชุดข้อมูล

3.3 วิเคราะห์หาปัจจัยที่เกี่ยวข้อง

การวิเคราะห์หาปัจจัยที่เกี่ยวข้องนั้น ใช้เทคนิคการจัดลำดับความสัมพันธ์กับตัวแปรกับกลุ่มข้อมูลในโปรแกรม Weka 3.6 ด้วยวิธี Info Gain Attribute, GainRatio AttributeEval และ OneRAttributeEval โดยพิจารณาจากค่า average merit ที่มากกว่าศูนย์ ทดสอบผลลัพธ์ด้วยวิธี K-fold cross-validation กำหนดค่า K = 10 ผลจากการวิเคราะห์คือปัจจัยทุกตัวมีความสัมพันธ์กับตัวแปรกลุ่มข้อมูลทั้งหมด จากนั้นนำมาจัดลำดับความสำคัญและคำนวณหาค่าเฉลี่ยแต่ละปัจจัย

3.4 การสร้างและการทดสอบความถูกต้องของโมเดล

นำชุดข้อมูลที่เตรียมไว้ทดสอบกับอัลกอริทึม C4.5 ปรับค่าพารามิเตอร์ คือ ค่าความเชื่อมั่น 0.25 ตัวอย่างขั้นต่ำของโหนดใบ 4 การตัดแต่งกิ่งต้นไม้ True ส่วนอัลกอริทึม MLP ปรับค่าพารามิเตอร์ คือ อัตราการเรียนรู้ 0.3 ตัวอย่างขั้นต่ำของโหนดใบ 0.6 รอบการเรียนรู้ 500 โหนดซ่อน 8 [11] และอัลกอริทึม Naïve Bayes ปรับค่าพารามิเตอร์ทุกตัวเป็น False กำหนดการทดสอบผลลัพธ์ด้วยวิธี K-fold cross-validation และวิธี Percentage split เพื่อเรียนรู้และทดสอบประสิทธิภาพค่าความแม่นยำ (Accuracy) ของโมเดลที่ดีที่สุด

4. ผลการดำเนินงาน

จากการทดสอบและวิเคราะห์หาปัจจัยที่ส่งผลกระทบในการลาออกกลางคันของนักศึกษา เพื่อดูความสัมพันธ์ตัวแปรกับกลุ่มข้อมูล ได้ผลลัพธ์ดังตารางที่ 2

ตารางที่ 2: ตารางแสดงผลลัพธ์การจัดลำดับความสัมพันธ์กับตัวแปรกับกลุ่มข้อมูล

No.	Info Gain Attribute	GainRatio AttributeEval	OneR AttributeEval
1	TofS	TofC	TofS
2	FG	TofS	FG
3	TofC	SL	TofC
4	Major	FG	Major
5	SL	Age	SL
6	Age	SSN	SSN
7	SSN	Major	Age
8	FofO	FofO	FofO
9	MofO	Course	Course
10	Course	Gender	Gender
11	Gender	MofO	MofO

จากตารางที่ 2 แสดงผลลัพธ์ลำดับความสำคัญแต่ละอัลกอริทึมที่แตกต่างกัน ดังนั้นการคำนวณหาค่าเฉลี่ยของปัจจัยที่ส่งผลกระทบสูงสุด 5 อันดับ จึงกำหนดให้ลำดับความสำคัญ

ของ Info Gain Attribute แทนด้วยตัวแปร I ลำดับความสำคัญของ GainRatioAttributeEval แทนด้วยตัวแปร G และลำดับความสำคัญของ OneRAttributeEval แทนด้วยตัวแปร O ดังสมการ

$$\text{ลำดับความสำคัญเฉลี่ย} = \frac{I + G + O}{\text{จำนวนอัลกอริทึมทั้งหมด}} \quad (5)$$

ตารางที่ 3: ตารางแสดงการจัดลำดับความสำคัญเฉลี่ย

ตัวแปร	ลำดับความสำคัญ			ลำดับความสำคัญเฉลี่ย	ลำดับความสำคัญ
	I	G	O		
Gender	11	10	10	10.33	10
Age	6	5	7	6.00	6
Major	**4**	**7**	**4**	**5.00**	**5**
Course	10	9	9	9.33	9
TofC	**3**	**1**	**3**	**2.33**	**2**
MofO	9	11	11	10.33	10
FofO	8	8	8	8.00	8
TofS	**1**	**2**	**1**	**1.33**	**1**
FG	**2**	**4**	**2**	**2.67**	**3**
SSN	7	6	6	6.33	7
SL	**5**	**3**	**5**	**4.33**	**4**

จากตารางที่ 3 แสดงการจัดลำดับความสำคัญเฉลี่ยของตัวแปร โดยตัวแปรที่มีค่าเฉลี่ยน้อยที่สุดคือตัวแปรที่ส่งผลกระทบในการลาออกกลางคันของนักศึกษามากที่สุด ใน 5 อันดับแรก ได้แก่ TofS, TofC, FG, SL และ Major

จากนั้นทดสอบกับโมเดลทั้ง 3 แบบโดยใช้ปัจจัยทั้งหมดเป็นชุดข้อมูลที่ 1 และใช้ปัจจัยที่เกี่ยวข้อง 5 อันดับ เป็นชุดข้อมูลที่ 2 ทดสอบผลลัพธ์ด้วยวิธี K-fold cross-validation ซึ่งจะแบ่งข้อมูลออกเป็นส่วนๆ เรียกว่า Fold และนำบางส่วนจากชุดข้อมูลมาทดสอบผลลัพธ์ของแบบจำลอง กำหนดให้ค่า K = 10 และวิธี Percentage split เป็นการแบ่งข้อมูลเพื่อเรียนรู้และทดสอบตามการกำหนดเปอร์เซ็นต์ โดยกำหนดเป็น 80 จากนั้นเปรียบเทียบค่าความแม่นยำ [9] ของโมเดลจากสมการ

$$\text{Accuracy} = \frac{TP + TN}{TP+FP+FN+TN} \quad (6)$$

TP คือ ค่าที่ทำนายถูกเชิงบวก *TN* คือ ค่าที่ทำนายถูกเชิงลบ

FP คือ ค่าที่ทำนายผิดเชิงบวก *FN* คือ ค่าที่ทำนายผิดเชิงลบ

ตารางที่ 4 : ผลการเปรียบเทียบประสิทธิภาพค่าความแม่นยำของโมเดล

อัลกอริทึม	Accuracy			
	ชุดข้อมูลที่ 1		ชุดข้อมูลที่ 2	
	K=10	Split=80%	K=10	Split=80%
C4.5	93.74	94.27	85.55	87.00
MLP	**94.45**	93.39	85.51	86.56
Naïve Bayes	83.74	84.36	83.52	84.14

จากตารางที่ 4 พบว่าชุดข้อมูลที่ 1 ที่ใช้ปัจจัยทั้งหมดเป็นชุดทดสอบ ด้วย 10-fold cross-validation ของอัลกอริทึม MLP นั้นให้ค่าความแม่นยำสูงที่สุดร้อยละ 94.45 มากกว่าอัลกอริทึม C4.5 และอัลกอริทึม Naïve Bayes ที่ร้อยละ 93.74 และ 83.74 ตามลำดับ

5. สรุป

จากการวิเคราะห์หาปัจจัยที่เกี่ยวข้องในการลาออกกลางคันของนักศึกษาและเปรียบเทียบประสิทธิภาพของอัลกอริทึมในงานวิจัยนี้ พบว่า อัลกอริทึมโครงข่ายประสาทเทียม Multilayer Perception ให้ค่าความแม่นยำสูงที่สุดร้อยละ 94.45 ซึ่งมากกว่าอัลกอริทึมต้นไม้ตัดสินใจ C4.5 และ Naïve Bayes ที่ร้อยละ 93.74 และ 83.74 ตามลำดับ เมื่อเปรียบเทียบค่าความแม่นยำของอัลกอริทึม MLP กับงานวิจัยของ Parneet Kaura และคณะ ที่ให้ค่าความแม่นยำสูงสุดที่ร้อยละ 75 [4] จะเห็นได้ว่างานวิจัยนี้ให้ค่าความแม่นยำที่สูงกว่าถึงร้อยละ 19.45 ซึ่งเป็นผลมาจากการจัดเตรียมข้อมูลที่มีกระบวนการรวมไปถึงการทำ SMOTE ที่แก้ไขปัญหาการจำแนกข้อมูลที่ไม่สมดุลกัน และมีปัจจัยที่เกี่ยวข้อง 5 อันดับแรก คือ TofS (ประเภทโรงเรียนเดิม), TofC (ประเภทหลักสูตร), FG (เกรดเฉลี่ยก่อนเข้าศึกษา), SL (การกู้ยืมกองทุนเพื่อการศึกษา) และ Major (สาขา) ซึ่งปัจจัยประเภทโรงเรียนเดิม สอดคล้องกับงานวิจัยของ ภัทร์พงศ์ พงศ์ภัทรกานต์ [10] ปัจจัยการกู้ยืมกองทุนเพื่อการศึกษา สอดคล้องกับงานวิจัยของ Jeeranan Chareonrat [7] และปัจจัยสาขาวิชา สอดคล้องกับงานวิจัยของ เยาวภา ภารสำเร็จ และคณะ [11] จากผลการ

วิเคราะห์ปัจจัยที่ได้จากชุดข้อมูลที่นำมาทำนายนั้นสามารถสรุปได้ว่า นักศึกษาที่จบการศึกษามาจากโรงเรียนมัธยม (ม.6) ไม่ได้กู้ยืมกองทุนเพื่อการศึกษา มีเกรดที่ต่ำกว่า 2.00 และเข้าเรียนในสาขาวิชาเทคนิคยานยนต์, ไฟฟ้ากำลังและเทคนิคการผลิตส่วนใหญ่จะลาออกกลางคัน เพราะเหตุนี้ ผู้บริหารหรืออาจารย์ที่ปรึกษาจะต้องเข้าใจสภาพครอบครัวของนักศึกษาและให้ความช่วยเหลือในเรื่องทางการเงิน โดยอาจจะจัดพิจารณาการผ่อนผันการชำระค่าเทอม พิจารณาให้กู้ยืมกองทุนเพื่อการศึกษา พิจารณาให้ทุนการศึกษาสำหรับผู้ที่เรียนดีแต่ยากจนหรือทุนกิจกรรม รวมไปถึงปัญหาความรู้พื้นฐานในแต่ละสาขาวิชา ดังนั้นอาจารย์ประจำสาขาอาจจะจัดสอนปรับพื้นฐานให้นักศึกษามีความรู้พอที่จะเรียนเนื้อหาหลักตอนเปิดเทอม เพื่อลดความเสี่ยงในการลาออกกลางคันของนักศึกษา

อย่างไรก็ตาม ถ้าต้องการวิเคราะห์ข้อมูลที่ละเอียดหรือให้มีประสิทธิภาพมากขึ้น ควรมีการเก็บรวบรวมข้อมูลปัจจัยตัวอื่นๆ ด้วย เช่น ข้อมูลพฤติกรรมของนักเรียน ข้อมูลสภาพการเรียน-การสอน โครงสร้างของหลักสูตรและสภาพแวดล้อมในการเรียน ข้อมูลเหล่านี้จะต้องจัดเก็บเพิ่มเติมโดยอาจจะใช้เครื่องมืออื่น เช่น แบบสอบถาม การสัมภาษณ์ เป็นต้น

เอกสารอ้างอิง

[1] สำนักติดตามและประเมินผลการอาชีวศึกษา, "*คู่มือโครงการการลดปัญหาการออกกลางคัน*," สำนักงานคณะกรรมการการอาชีวศึกษากระทรวงศึกษาธิการ, 2556.

[2] ฝ่ายแผนและนโยบาย, "*รายงานประเมินคุณภาพภายในสถานศึกษาเอกชน ประเภทอาชีวศึกษา*," วิทยาลัยเทคโนโลยีภาคตะวันออกเฉียงเหนือ, 2559.

[3] Pimpa Cheewaprakobkit, "Study of Factors Analysis Affecting Academic Achievement of Undergraduate Students in International Program," *Proceedings of the International MultiConference of Engineers and Computer Scientists (IMECS)*, Vol. 1, pp. 233-236, 2013.

[4] Parneet Kaura, Manpreet Singhb and Gurpreet Singh Josanc, "Classification and Prediction Based Data Mining Algorithms to Predict Slow Learners in Education Sector," *3rd International Conference on Recent Trends in Computing (ICRTC)*, Vol. 57, pp. 500-508, 2015.

[5] S. Kotsiantis, C. Pierrakeas, and P. Pintelas, "Efficiency of Machine Learning Techniques in Predicting Students' Performance In Distance Learning Systems," *Recent Advances In Mechanics And Related Fields University Of Patras, in Honour of Professor Constantine L. Goudas*, pp. 297-305, 2003.

[6] J. R. Quinlan "Induction of Decision Trees" *Machine Learning*, vol. 1, pp. 81-106, 1986.

[7] Jeeranan Chareonrat, "Analysis on Factors Affecting Normal-Grade Student Dismissal Using Decision Tree," *SNRU Journal of Science and Technology (SNRUJST)*, Vol. 8, No. 2, pp. 256-267, 2016.

[8] สุรพงษ์ เชี่ยวสกุลวัฒนา และสุกรี สินธุภิญโญ, "การจำแนกต้นไม้ตัดสินใจสำหรับชุดข้อมูลไม่สมดุลโดยใช้น้ำหนักต่างกันบนข้อมูลสังเคราะห์," *The Tenth National Conference on Computing and Information Technology (NCCIT)*, pp. 1-6, 2014.

[9] ชินพัฒน์ แก้วชินพร, "การจำแนกประเภทข้อมูลด้วยเทคนิคต้นไม้ตัดสินใจและการจัดกลุ่ม," คณะวิทยาศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง, 2553.

[10] ภัทร์พงศ์ พงศ์ภัทรกานต์, "การวิเคราะห์ปัจจัยที่ส่งผลต่อการพ้นสภาพของนักศึกษา ระดับปริญญาตรี โดยใช้คอมมิตตีแมชชีน," *The 6TH National Conference on Computing and Information Technology (NCCIT)*, pp. 491-496, 2010.

[11] เยาวภา ภารสำเร็จ, จิรัฏฐา ภูบุญอบ และวิรัฐ พงษ์ศิริ, "การเปรียบเทียบอัลกอริทึมเหมืองข้อมูลเพื่อวิเคราะห์ปัจจัยที่ส่งผลต่อระดับผลการเรียนของนักศึกษา," *วารสารวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยมหาสารคาม การประชุมทางวิชาการ "มหาวิทยาลัยมหาสารคามวิจัย ครั้งที่ 9: การวิจัยสู่ประชาคมอาเซียน 2"*, ปีที่ 9 ฉบับที่ 0 กันยายน พ.ศ. 2556 หน้า 281-289

[12] นงลักษณ์ พรมทอง, สุรพรรษ์ เพ็ญจำรัส และศักดิ์ชาย ตั้งประเสริฐ, "การเปรียบเทียบประสิทธิภาพการจำแนกข้อมูลด้วยเทคนิคการทำเหมืองข้อมูลระหว่างเคเนียเรสเนเบอร์ กฎพื้นฐาน และต้นไม้ตัดสินใจ," *The Eighth National Conference on Computing and Information Technology (NCCIT)*, pp. 540-546, 2012.

[13] จุฑารัตน์ ตั้งกิตติวัฒน์ และนลินภัสร์ ปรวัฒน์ปรียกร, "โมเดลการวิเคราะห์โรคในสุกร โดยใช้โครงข่ายประสาทเทียม," *The Tenth National Conference on Computing and Information Technology (NCCIT)*, pp. 26-31, 2014.

การค้นหายาสำหรับการรักษามะเร็งแบบมุ่งเป้าโดยการค้นหาในแนวลึกตามความเป็นศูนย์กลางบนเส้นทางการส่งสัญญาณของเซลล์

Drug Searching for Targeted Cancer Therapy by Centrality Depth First Search on Cell Signal Transduction Pathway

มัสยา สำเนียงงาม (Massaya Samnienggam)[1] และธนัสนี เพียรตระกูล (Tanasanee Phienthrakul)[2]

ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยมหิดล

[1]f.mas_tsu@hotmail.com, [2]tanasanee.phi@mahidol.ac.th

บทคัดย่อ

ในการรักษามะเร็งแบบมุ่งเป้า จำเป็นจะต้องทราบยีนที่ผิดปกติเพื่อเลือกใช้ยาได้อย่างเหมาะสม ในกรณีที่ไม่พบยาที่สามารถใช้กับยีนเป้าหมาย หรือผู้ป่วยไม่สามารถใช้ยาที่มีอยู่ได้ ต้องทำการค้นหายาอื่นทดแทน โดยอาศัยการวิเคราะห์เส้นทางการส่งสัญญาณของเซลล์เพื่อพิจารณายีนอื่นที่เกี่ยวข้อง งานวิจัยนี้จึงได้นำเสนอเทคนิคในการสำรวจเส้นทางการส่งสัญญาณของเซลล์เพื่อให้ได้ยาที่เหมาะสมต่อการรักษามะเร็งแบบมุ่งเป้าโดยใช้วิธีการค้นหาในแนวลึกตามความเป็นจุดศูนย์กลางของยีน ซึ่งได้ประยุกต์เทคนิคในการสำรวจกราฟเข้ากับการวัดความเป็นศูนย์กลาง เพื่อเพิ่มประสิทธิภาพในการค้นหายา ยาที่ค้นหาได้จะถูกนำมาให้คะแนนและประเมินผล ผลการวิจัยที่ได้แสดงให้เห็นว่า วิธีการค้นหาในแนวลึกตามความเป็นศูนย์กลางโดยวัดจากการคั่นกลางของยีนให้ชุดยาที่เหมาะสมโดยใช้เวลาในการวิเคราะห์เพียงเล็กน้อย

คำสำคัญ: การเลือกยา การรักษามะเร็งแบบมุ่งเป้า เส้นทางการส่งสัญญาณ ความเป็นศูนย์กลาง

Abstract

In targeted cancer therapy, the target gene must be known in order to select the suitable drugs. If there is no drug for the target gene or the patient cannot take the existing drugs, then other drugs must be searched by analyzing signal transduction pathway. This research proposes a technique to traverse in signal transduction pathway. This technique is used to find the suitable drugs for targeted cancer therapy using centrality depth first search (CDFS), which combines graph traversal algorithm and centrality measurement to increase the efficiency of drug discovery. Drug results will be scored and evaluated. The experimental results show that Betweenness Centrality on CDFS yields a set of suitable drugs within a reasonable time.

Keyword: Drug Selection, Targeted Cancer Therapy, Signal Transduction Pathway, Centrality.

1. บทนำ

มะเร็งเป็นหนึ่งในสิบอันดับของสาเหตุการเสียชีวิตในมนุษย์ และมีแนวโน้มจะเพิ่มขึ้นเรื่อยๆ ยิ่งกว่านั้น ผู้ป่วยมะเร็งจำนวนมากกว่าครึ่งอาศัยอยู่ในประเทศที่กำลังพัฒนา [1] ปัจจัยที่อาจเป็นสาเหตุของมะเร็ง ประกอบด้วย รูปแบบการดำเนินชีวิต สภาพแวดล้อม อาหาร สารเคมี และ ดีเอ็นเอ การรักษามะเร็งหลักๆ มี 3 วิธี ได้แก่ การผ่าตัด การฉายรังสี และ เคมีบำบัด ซึ่งการรักษาด้วยเคมีบำบัดเป็นวิธีการรักษาที่มีประสิทธิภาพ แต่ผู้ป่วยบางรายไม่สามารถทนต่อผลข้างเคียงที่เกิดขึ้นได้ ดังนั้นจึงมีวิธีการรักษาที่ส่งผลต่อเซลล์ปกติน้อยกว่าวิธีเคมีบำบัดเกิดขึ้น เรียกว่า การรักษามะเร็งแบบมุ่งเป้า ซึ่งเป็นวิธีการรักษาโดยให้ยาที่ส่งผลเฉพาะเจาะจงต่อเซลล์มะเร็ง เพื่อไปยับยั้งกระบวนการส่งสัญญาณ ที่เป็นสาเหตุของการแบ่งตัวในเซลล์มะเร็ง [2]

การรักษามะเร็งแบบมุ่งเป้ามักใช้เป็นวิธีการรักษาร่วมกับการรักษาวิธีอื่นๆ เพื่อเพิ่มประสิทธิภาพในการรักษาให้ดียิ่งขึ้น อย่างไรก็ตาม ในกรณีที่ไม่มียาที่ส่งผลต่อยีนเป้าหมาย เส้นทาง

การส่งสัญญาณ (Pathway) จะถูกนำมาวิเคราะห์หายีนที่เกี่ยวข้องกับยีนเป้าหมาย เพื่อหายาที่ช่วยในการรักษา

มีหลายงานวิจัยที่ศึกษาเกี่ยวกับเส้นทางการส่งสัญญาณของเซลล์และใช้เทคนิคในการสำรวจกราฟเข้ามาช่วยวิเคราะห์เส้นทางการส่งสัญญาณ เช่น Meng-Hsiun Tsai และคณะ [3] ได้ประยุกต์ใช้ยีนในการจำแนกเส้นทางเดินด้วยวิธีการ Dijkstra ร่วมกับทฤษฎีของเบย์ งานวิจัยดังกล่าวสามารถวิเคราะห์เส้นทางความสัมพันธ์ของเซลล์มะเร็งที่ซับซ้อนได้ แต่ยังคงต้องการการตรวจสอบจากผู้เชี่ยวชาญ เพื่อทำให้โมเดลการทำนายนี้มีความน่าเชื่อถือมากยิ่งขึ้น

Maya El Dayeh และ Nichael Hahsler [4] ได้ใช้วิธีการเดินแบบสุ่ม (Random walk) บนกราฟ เพื่อแก้ไขปัญหาความสมบูรณ์ของเส้นทางเดินกับเครือข่ายปฏิสัมพันธ์ระหว่างโปรตีน พวกเขาใช้เครือข่ายปฏิสัมพันธ์ระหว่างโปรตีนของยีสต์และเส้นทางเดินของยีสต์จำนวน 2 เส้นทาง จากนั้นจึงเปรียบเทียบความถูกต้องของวิธีที่นำเสนอกับงานอื่นๆ ที่ใช้วิธีการเดินแบบสุ่มเช่นเดียวกัน ผลที่ได้แสดงให้เห็นว่า วิธีที่พวกเขานำเสนอได้ความถูกต้องที่คล้ายกันหรือดีกว่างานอื่นที่นำมาเปรียบเทียบ

Deanna Petrochilos และ Neil Abernethy [5] ได้ทำการสำรวจความสัมพันธ์ระหว่างเครือข่ายของยีนและความเกี่ยวข้องทางมะเร็งวิทยา จากนั้นประเมินการทำนายของตัววัดความเป็นศูนย์กลางและสัมประสิทธิ์การจัดกลุ่มโดยใช้โมเดลเชิงเส้น และประยุกต์วิธีการเดินแบบสุ่มในการค้นหากลุ่มยีนที่เกี่ยวกับมะเร็ง ผลที่ได้แสดงให้เห็นว่า ยีนมะเร็งในเครือข่ายเมตาบอลิกและเครือข่ายการส่งสัญญาณมีความแตกต่างกันอย่างมีนัยสำคัญ ซึ่งทำให้สามารถจัดลำดับความสำคัญของยีนที่ทำให้เกิดโรค หรือ ทำนายสถานะของโรคได้

งานวิจัยนี้จึงมีแนวคิดที่จะพัฒนาการค้นหาและเลือกยาสำหรับวิธีการรักษามะเร็งแบบมุ่งเป้า โดยพิจารณาจากเส้นทางการส่งสัญญาณของเซลล์ โดยใช้เทคนิคสำรวจกราฟมาช่วยวิเคราะห์เส้นทางการส่งสัญญาณของเซลล์ แต่เนื่องจากงานวิจัยที่เกี่ยวข้องตามที่ได้ศึกษามานั้น ส่วนใหญ่เลือกใช้วิธีการสำรวจเส้นทางการส่งสัญญาณแบบสุ่ม ซึ่งใช้เวลานานกว่าจะเจอเป้าหมายที่ต้องการ หรือ วิธีการสำรวจแบบอื่นนั้นไม่สามารถใช้กับงานวิจัยนี้ได้ เช่น วิธีการ Dijkstra ซึ่งต้องใช้น้ำหนักของขอบมาพิจารณา แต่งานวิจัยนี้ยังไม่มีข้อมูลของน้ำหนักขอบ ดังนั้น งานวิจัยนี้จึงได้ประยุกต์วิธีการค้นหาในแนวลึกร่วมกับการวัดความเป็นศูนย์กลาง เกิดเป็นเทคนิคที่ชื่อว่า การค้นหาในแนวลึกตามความเป็นศูนย์กลาง ซึ่งการวัดความเป็นศูนย์กลางจะช่วยแนะนำเส้นทางที่คาดว่าจะพบยาที่เกี่ยวข้องกับยีนเป้าหมายได้เร็วขึ้น และงานวิจัยนี้ยังได้ออกแบบระบบการคิดคะแนน เพื่อใช้เป็นตัววัดผล โดยการเปรียบเทียบคะแนนยาที่ได้จากแต่ละรูปแบบของการวัดความเป็นศูนย์กลาง ได้แก่ ความเป็นศูนย์กลางโดยวัดจากระดับ ความเป็นศูนย์กลางโดยวัดจากความใกล้ชิด และ ความเป็นศูนย์กลางโดยวัดจากการคั่นกลาง และ การเปรียบเทียบคะแนนยาที่ได้จากแต่ละเทคนิคการสำรวจเส้นทางการส่งสัญญาณของเซลล์ ได้แก่ การค้นหาในแนวลึก การค้นหาในแนวกว้าง และ การค้นหาในแนวลึกตามความเป็นศูนย์กลาง

การค้นหายาสำหรับงานวิจัยนี้ได้นำเส้นทางการส่งสัญญาณของมะเร็งจากแหล่งข้อมูลที่มีการกล่าวอ้างถึงในหลายงานวิจัย อย่าง KEGG [6] ซึ่งข้อมูลความสัมพันธ์ระหว่างยีนในเส้นทางการส่งสัญญาณเดียวกันที่ได้จาก KEGG จะถูกนำมาวิเคราะห์หาเป้าหมายของยาในกรณีที่ไม่มียาที่ส่งผลต่อยีนเป้าหมายโดยตรง ในงานวิจัยนี้เลือกใช้ข้อมูลยาที่มีความน่าเชื่อถือจาก ClinicalTrials.gov [7] และ Drugs@FDA [8] เพื่อเป็นข้อมูลประกอบการพิจารณาในการเลือกยาที่เหมาะสม

2. การวัดความเป็นศูนย์กลางสำหรับการสำรวจเส้นทาง

งานวิจัยนี้ใช้การวัดความเป็นศูนย์กลางเพื่อแนะนำเส้นทางเดินในการสำรวจเส้นทางการส่งสัญญาณ ซึ่งประกอบด้วยการวัดดังต่อไปนี้

2.1 ความเป็นศูนย์กลางโดยวัดจากระดับ

ความเป็นศูนย์กลางโดยวัดจากระดับ (Degree centrality) คือ จำนวนของการเชื่อมโดยตรงที่เกิดขึ้นกับจุด หรือจำนวนของจุดที่ใกล้เคียงที่สุด [9] ซึ่งสามารถคำนวณได้ดังสมการที่ 1

$$C_D(i) = k_i = \sum_j A_{ij} \qquad (1)$$

โดยที่ $C_D(i)$ คือ ค่าความเป็นจุดศูนย์กลางของจุด i, k_i คือ จำนวนของเส้นที่เชื่อมถึงจุด i, และ A_{ij} จะเท่ากับ 1 ถ้า i และ j เชื่อมโยงกันและเท่ากับ 0 ถ้าเป็นกรณีอื่นๆ

2.2 ความเป็นศูนย์กลางโดยวัดจากความใกล้ชิด

ความเป็นศูนย์กลางโดยวัดจากความใกล้ชิด (Closeness centrality) คือ ผลรวมของระยะทางระหว่างจุดที่พิจารณากับจุดอื่นทุกจุดภายในกราฟ [10] ซึ่งสามารถคำนวณได้ดังสมการที่ 2

$$C_C(i) = \frac{1}{\sum_j d(i,j)} \quad (2)$$

โดยที่ $C_C(i)$ คือ ค่าความเป็นจุดศูนย์กลางของจุด i และ $d(i,j)$ คือ ระยะทางระหว่างจุด i และจุด j

2.3 ความเป็นศูนย์กลางโดยวัดจากการคั่นกลาง

ความเป็นศูนย์กลางโดยวัดจากการคั่นกลาง (Betweenness centrality) คือ จำนวนครั้งที่จุดที่พิจาณาทำตัวเป็นสะพานเชื่อมระหว่างเส้นทางที่สั้นที่สุด [11] ซึ่งสามารถคำนวณได้ดังสมการที่ 3

$$C_B(i) = \sum_{s \neq t \neq i} \frac{\sigma_{st}(i)}{\sigma_{st}} \quad (3)$$

โดยที่ $C_B(i)$ คือ ค่าความเป็นจุดศูนย์กลางของจุด $i, \sigma_{st}(i)$ คือ จำนวนของเส้นทางที่สั้นที่สุดระหว่างจุด s และจุด t ที่ผ่านจุด i, σ_{st} คือ จำนวนทั้งหมดของเส้นทางที่สั้นที่สุดระหว่างจุด s และจุด t

3. การค้นหาในแนวลึกตามความเป็นศูนย์กลางของยีน (Centrality Depth First Search: CDFS)

การค้นหาในแนวลึกตามความเป็นจุดศูนย์กลางของยีนหรือ CDFS เป็นวิธีที่ประยุกต์การค้นหาแนวลึก (Depth First Search) ซึ่งเป็นการสำรวจลึกลงไปในกราฟจนกระทั่งพบจุดสิ้นสุดโดยจะเริ่มต้นจากโหนดต้นทาง ร่วมกับการวัดความเป็นศูนย์กลาง เพื่อช่วยแนะนำทิศทางในการสำรวจเส้นทางการส่งสัญญาณ โดยไม่ย้อนกลับไปที่โหนดพ่อแม่ ซึ่งจะช่วยลดระยะเวลาในการค้นหาในกรณีที่เส้นทางการส่งสัญญาณมีขนาดใหญ่มาก ขั้นตอนการทำงานของเทคนิคนี้มีดังต่อไปนี้

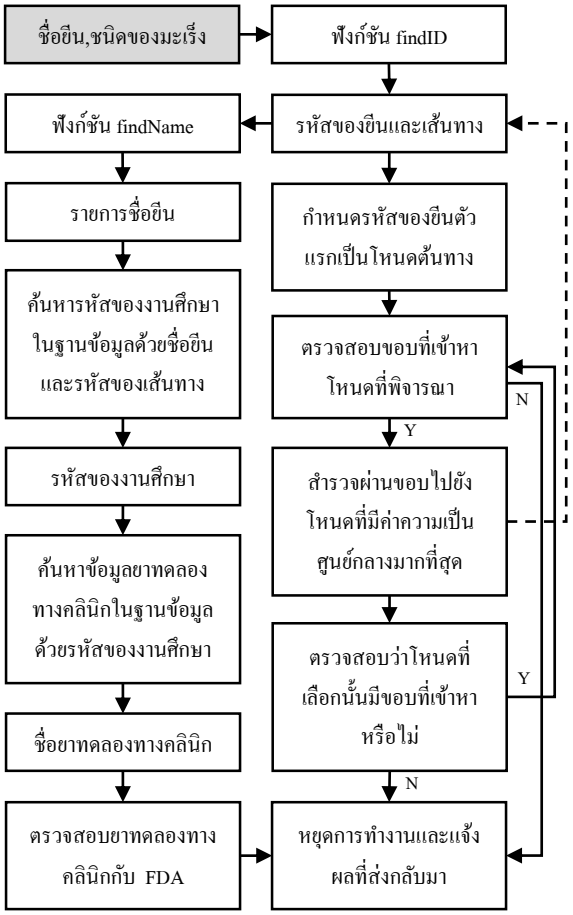
(1) เลือกโหนดต้นทาง

(2) ตรวจสอบว่ามีเส้นหรือความสัมพันธ์ในทิศที่เข้าหาโหนดที่พิจารณาอยู่หรือไม่ ถ้าไม่พบเส้นที่เข้าหาโหนดที่พิจารณาให้ยกเลิกการทำงาน

(3) ถ้าพบเส้นที่เข้าหาโหนดที่พิจารณาให้สำรวจผ่านเส้นไปยังโหนดที่มีค่าความเป็นศูนย์กลางมากที่สุดในบรรดาโหนดที่ค้นพบในเลเวลที่พิจารณา

(4) ตรวจสอบว่าโหนดที่เลือกเดินไปนั้นมีเส้นที่เข้าหาหรือไม่ ถ้ามีให้กลับไปทำซ้ำขั้นตอนที่ 3 ถ้าไม่มีให้หยุดการทำงาน

CDFS จะถูกประยุกต์เข้ากับระบบให้คะแนนยา ซึ่งสามารถอธิบายได้ดังภาพที่ 1

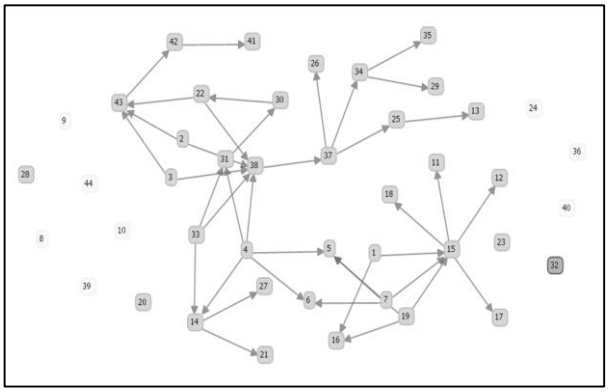


ภาพที่ 1: ระบบการเลือกยา

ระบบเริ่มด้วยผู้ใช้งานใส่ชื่อยีนและชนิดของมะเร็ง จากนั้นชื่อยีนและชนิดของมะเร็งจะถูกนำไปจับคู่กับข้อมูลในฐานข้อมูลเพื่อหารหัสของยีนและเส้นทางเดิน จากนั้นระบบจะนำรหัสเหล่านั้นไปค้นหาชื่ออื่นของยีน ซึ่งจะได้รายการชื่อยีนออกมา ชื่อยีนจากรายการจะถูกนำมาค้นหารหัสของงานศึกษาทางคลินิกในฐานข้อมูล ซึ่งจะใช้ค้นหาข้อมูลยาทดลองทางคลินิกในฐานข้อมูลอีกครั้ง

จากนั้นระบบจะทำการตรวจสอบยาทดลองทางคลินิกเหล่านี้กับข้อมูลจาก FDA รหัสของยีนตัวแรกที่ผู้ใช้งานใส่เข้ามา

ในระบบจะถูกกำหนดให้เป็นโหนดเริ่มต้น ตัวอย่างของเส้นทางการส่งสัญญาณแสดงดังภาพที่ 2



ภาพที่ 2: เส้นทางการส่งสัญญาณของมะเร็งเม็ดเลือดขาว

4. ระบบให้คะแนนยา (Scoring Drug System)

การให้คะแนนยา เป็นวิธีประเมินยาและจัดลำดับยาตามคะแนนเพื่อแนะนำยาที่เหมาะสมแก่การรักษามะเร็งแบบมุ่งเป้า ข้อมูลที่นำมาใช้คำนวณคะแนนประกอบด้วย 2 ส่วน ได้แก่ ข้อมูลยา (clinicaltials.gov และ Drugs@FDA) และข้อมูลเส้นทางการส่งสัญญาณของมะเร็งจาก KEGG ซึ่งคะแนนแบ่งออกเป็น 5 ส่วน ได้แก่ Clinical Trials Score, Study Phase Score, Study Status Score, FDA Score, และ Hop Score

1) **Clinical Trials Score** มีค่าเป็น 1 ถ้าชื่อยืนถูกค้นพบในข้อมูลของ clinicaltrials.gov และ ถ้าไม่พบจะมีค่าเป็น 0

2) **Study Phase Score** มีค่าตามระยะการทดลองในงานศึกษาตามข้อมูลจาก clinicaltrials.gov ในกรณีที่ระบบพบงานศึกษาที่ศึกษายืนตัวเดียวกัน ระบบจะเลือกงานศึกษาที่อยู่ในระยะทดลองมากที่สุด ระยะในการทดลองประกอบด้วย 4 ระยะ ได้แก่ (1) ระยะตรวจคัดกรอง (2) กำหนดขั้นตอนการทดสอบ (3) การทดสอบขั้นสุดท้าย และ (4) การอนุมัติงานศึกษา

3) **Study Status Score** เป็นการให้คะแนนตามสถานะของงานศึกษาตามข้อมูลจาก clinicaltrials.gov การให้คะแนนจะกำหนดตามความหมายและความสำคัญของสถานะของงานศึกษานั้นๆ ในกรณีที่มีงานศึกษาที่ทำการศึกษายาตัวเดียวกัน ระบบจะเลือกสถานะจากงานศึกษาที่มีระยะการทดลองมากที่สุดและวันที่เสร็จสมบูรณ์ล่าสุด

สถานะที่เป็นไปได้ประกอบด้วย “ยังไม่ได้รับการคัดเลือก” หรือ “ยกเลิก” หรือ “ถอน” คะแนนจะเท่ากับ 0 ถ้าสถานะของงานศึกษา คือ “กำลังคัดเลือก” หรือ “ลงทะเบียน” คะแนนจะเท่ากับ 1 ถ้าสถานะคือ “ใช้งานอยู่ หยุดการคัดเลือก” คะแนนจะเท่ากับ 2 ถ้าสถานะคือ “เสร็จสมบูรณ์” คะแนนจะเท่ากับ 3 และถ้าสถานะของงานศึกษา คือ “ถูกระงับ” คะแนนจะเท่ากับ 0.5 ซึ่งสถานะนี้สามารถถูกเปลี่ยนให้เป็น “ถอน” หรือ “กำลังคัดเลือก” ได้ในภายหลัง

4) **FDA Score** จะเท่ากับ 1 ถ้า ชื่อยาถูกค้นพบในฐานข้อมูลของ FDA และถ้าไม่พบจะเท่ากับ 0

5) **Hop Score** กำหนดให้เป็นสัดส่วนผกผันของระยะทางจากโหนดต้นทางไปสู่โหนดที่พิจารณาบวกหนึ่ง ถ้าชื่อยาถูกพบจากโหนดมากว่าหนึ่งโหนดในเส้นทางการส่งสัญญาณเดียวกัน ระบบจะเลือกคะแนนในส่วนนี้จากโหนดที่มีจำนวนก้าว (hop) มากที่สุด

คะแนนแต่ละส่วนจะถูกทำให้ค่าอยู่ระหว่าง 0 ถึง 1 โดย Study Phase Score และ Study Status Score จะถูกหารด้วยคะแนนที่มากที่สุดของแต่ละส่วนคะแนน จากนั้นคะแนนแต่ละส่วนจะถูกคูณด้วยค่าน้ำหนักที่งานวิจัยนี้ได้กำหนดเป็นค่าเริ่มต้นตามความสำคัญของคะแนนแต่ละส่วนไว้ดังสมการที่ 4

$$\frac{(C\times w_C+P\times w_P+S\times w_S+F\times w_F+H\times w_H)}{Full\ score}\times 100 \qquad (4)$$

โดยที่ C คือ Clinical Trials score

P คือ Study Phase score

S คือ Study Status score

F คือ FDA score

H คือ Hop score และ

w_C, w_P, w_S, w_F และ w_H คือ ค่าน้ำหนัก ซึ่งในงานวิจัยนี้ทดสอบได้ค่าที่เหมาะสมเป็น 1, 1, 1, 4, และ 7 ตามลำดับ ทำให้ $Full\ score$ มีค่าเท่ากับ 14

5. การเปรียบเทียบผลการวัดความเป็นศูนย์กลาง

ในการทดสอบเพื่อเลือกวิธีการวัดความเป็นศูนย์กลางของ CDFS ได้ทำการทดลองเปรียบเทียบวิธีวัดความเป็นศูนย์กลาง 3 แบบ คือ ความเป็นศูนย์กลางโดยวัดจากระดับ ความเป็นศูนย์กลาง

โดยวัดจากความใกล้ชิด และ ความเป็นศูนย์กลางโดยวัดจากการคั่นกลาง

งานวิจัยนี้ได้นำยีนจำนวน 3,871 ตัว จาก14 เส้นทางการส่งสัญญาณที่ถูกนำมาใช้เป็นตัวทดสอบสำหรับการหายาที่เกี่ยวข้อง ทำการคำนวณคะแนนของยาโดยใช้วิธีการที่นำเสนอในหัวข้อที่ 4 และคำนวณคะแนนเฉลี่ยของยาสิบอันดับแรก ผลที่ได้แสดงดังตารางที่ 1

ตารางที่ 1: คะแนนเฉลี่ยของยาตามการวัดความเป็นจุดศูนย์กลาง

Pathway ID	Degree		Closeness		Betweenness	
	Score (%)	*Time(sec.)*	*Score (%)*	*Time(sec.)*	*Score (%)*	*Time(sec.)*
hsa05210	**33.728**	*35.039*	**33.728**	37.879	**33.728**	35.759
hsa05211	**43.083**	13.118	43.053	*12.227*	43.053	12.695
hsa05212	**57.219**	33.847	57.024	33.549	57.150	*33.469*
hsa05213	**45.461**	*65.150*	**45.461**	65.370	**45.461**	66.168
hsa05214	**52.129**	58.562	**52.129**	56.019	**52.129**	*55.663*
hsa05215	**50.200**	32.196	**50.200**	*31.821*	**50.200**	33.794
hsa05216	**42.255**	*12.504*	**42.255**	12.700	**42.255**	12.610
hsa05217	**4.737**	0.638	**4.737**	*0.601*	**4.737**	*0.601*
hsa05218	**56.630**	20.753	**56.630**	*20.692*	**56.630**	20.872
hsa05219	**37.769**	*9.581*	**37.769**	9.740	**37.769**	9.686
hsa05220	**38.607**	5.652	**38.607**	5.609	**38.607**	*5.469*
hsa05221	42.748	7.016	42.747	*6.768*	**47.904**	13.291
hsa05222	35.242	4.061	35.242	*4.034*	**37.706**	4.152
hsa05224	**52.675**	15.086	**52.675**	*14.779*	**52.675**	15.037
Total (%)	44.761	23.723	44.729	***23.560***	**45.202**	24.014

จากตารางที่ 1 แสดงให้เห็นว่า เส้นทางการส่งสัญญาณหลายเส้นทางตามการวัดความเป็นศูนย์กลางมีคะแนนยาเฉลี่ยเท่ากัน ซึ่งบ่งบอกได้ว่าลำดับยาของเส้นทางการส่งสัญญาณเหล่านั้นเหมือนกัน แต่คะแนนเฉลี่ยของยาตามการวัดความเป็นศูนย์กลางโดยวัดจากระดับ และ วัดจากความใกล้ชิด มีค่าน้อยกว่าคะแนนเฉลี่ยของยาตามการวัดความเป็นศูนย์กลางโดยวัดจากการคั่นกลาง ฉะนั้นจึงสามารถสรุปได้ว่า การวัดความเป็นศูนย์กลางโดยวัดจากการคั่นกลางเหมาะสมสำหรับ CDFS มากที่สุดตามคะแนนเฉลี่ยของยาสิบอันดับแรก

อย่างไรก็ตาม การวัดความเป็นศูนย์กลางโดยวัดจากการคั่นกลางใช้เวลาในการวิเคราะห์มากที่สุด แต่ก็ไม่ต่างจากการวัดความเป็นจุดศูนย์กลางแบบอื่นมากนัก งานวิจัยนี้จึงเลือกใช้การวัดความเป็นศูนย์กลางโดยวัดจากการคั่นกลางบน CDFS ในการเปรียบเทียบผลคะแนนยาตามวิธีการสำรวจกราฟ

6. การเปรียบเทียบผลตามวิธีการสำรวจกราฟ

นอกจากการเปรียบเทียบผลคะแนนของยาตามการวัดความเป็นจุดศูนย์กลางแล้ว งานวิจัยนี้ยังได้ทำการเปรียบเทียบผลคะแนนของยาตามวิธีการสำรวจกราฟด้วยวิธีการสำรวจกราฟที่ถูกนำมาเปรียบเทียบประกอบด้วย การค้นหาในแนวลึก (DFS) การค้นหาในแนวกว้าง (BFS) และ การค้นหาในแนวลึกตามความเป็นจุดศูนย์กลางของยีน (CDFS) งานวิจัยนี้ได้เลือกยาสิบอันดับแรกจากระบบการให้คะแนนยาและคำนวณเช่นเดียวกันกับการเปรียบเทียบแบบแรก ผลลัพธ์ที่ได้แสดงดังตารางที่ 2

ตารางที่ 2: คะแนนเฉลี่ยของยาตามเทคนิคสำรวจกราฟ

Pathway ID	DFS		BFS		CDFS	
	Score (%)	*Time (sec.)*	*Score (%)*	*Time (sec.)*	*Score (%)*	*Time (sec.)*
hsa05210	32.771	29.565	33.487	*19.477*	**33.728**	35.759
hsa05211	35.426	12.870	37.050	13.220	**43.053**	*12.695*
hsa05212	56.286	*30.131*	55.904	33.390	**57.150**	33.469
hsa05213	**45.461**	*62.926*	45.444	69.799	**45.461**	66.168
hsa05214	47.456	*55.608*	47.775	58.159	**52.129**	55.663
hsa05215	48.891	68.967	49.784	70.563	**50.200**	*33.794*
hsa05216	39.047	47.236	41.345	52.838	**42.255**	*12.610*
hsa05217	**4.737**	2.072	**4.737**	1.620	**4.737**	*0.601*
hsa05218	**56.759**	84.551	56.708	87.913	56.630	*20.872*
hsa05219	33.981	32.412	35.867	49.345	**37.769**	*9.686*
hsa05220	38.198	20.371	38.414	19.123	**38.607**	*5.469*
hsa05221	46.335	69.386	46.502	78.636	**47.904**	*13.291*
hsa05222	36.625	14.0549	36.525	13.477	**37.706**	*4.152*
hsa05224	47.829	98.599	47.898	106.774	**52.675**	*15.037*
Total (%)	42.538	50.076	43.011	53.353	**45.202**	***24.014***

จากตารางที่ 2 จะเห็นว่า CDFS มีคะแนนเฉลี่ยของยาสูงที่สุด และ ยังใช้เวลาในการวิเคราะห์น้อยกว่าเทคนิคตัวอื่น

ซึ่งสามารถสรุปได้ว่า CDFS เหมาะสมสำหรับสนับสนุนการเลือกยามากที่สุดด้วยคะแนนเฉลี่ยของยาสิบอันดับแรกมากที่สุด

7. สรุป

งานวิจัยนี้ได้นำเสนอวิธีการค้นหาในแนวลึกตามความเป็นศูนย์กลางของยีน หรือ CDFS เพื่อสำรวจเส้นทางการส่งสัญญาณของมะเร็งและแนะนำยาที่เหมาะสมสำหรับการรักษามะเร็งแบบมุ่งเป้าด้วยการวิเคราะห์เส้นทางการส่งสัญญาณของเซลล์ นอกจากนั้นยังได้นำเสนอวิธีการให้คะแนนยาเพื่อจัดอันดับตามความน่าเชื่อถือของยา และการส่งผลกระทบกับยีนตัวอื่นๆ ซึ่งสามารถนำไปใช้เรียงลำดับยาที่เป็นผลลัพธ์จากการสืบค้นเพื่อแนะนำยาที่เกี่ยวข้อง

การประเมินผลแบ่งออกเป็น 2 ส่วน ได้แก่ การเปรียบเทียบผลคะแนนของยาตามการวัดความเป็นศูนย์กลาง ได้แก่ ความเป็นศูนย์กลางโดยวัดจากระดับ ความเป็นศูนย์กลางโดยวัดจากความใกล้ชิด และความเป็นศูนย์กลางโดยวัดจากการคั่นกลาง เพื่อหาการวัดความเป็นศูนย์กลางที่เหมาะสมสำหรับ CDFS ผลลัพธ์แสดงให้เห็นว่า การวัดจากการคั่นกลางเหมาะสมกับ CDFS มากที่สุดด้วยคะแนนยาเฉลี่ยสิบอันดับแรกที่มากที่สุด

การเปรียบเทียบผลคะแนนของยาตามเทคนิคสำรวจกราฟ ได้แก่ การค้นหาในแนวลึก (DFS) การค้นหาในแนวกว้าง (BFS) และ การค้นหาในแนวลึกตามความเป็นศูนย์กลางของยีน (CDFS) เพื่อหาวิธีการที่เหมาะสมสำหรับการแนะนำยา ผลลัพธ์แสดงให้เห็นว่า CDFS เหมาะสำหรับสนับสนุนการเลือกยามากที่สุดด้วยคะแนนเฉลี่ยของยาสิบอันดับแรกมากที่สุดและเวลาเฉลี่ยที่ใช้ประมวลผลน้อยที่สุด แต่การตัดสินใจสุดท้ายว่าจะเลือกยาตัวใดมาใช้ในการรักษายังขึ้นอยู่กับการตัดสินใจของแพทย์ เทคนิคนี้เป็นเพียงส่วนช่วยสนับสนุนการตัดสินใจ โดยเสนอตัวเลือกยาที่เพิ่มมากขึ้นให้ผู้ทำการตัดสินใจได้พิจารณา

เอกสารอ้างอิง

[1] World Health Organization: WHO. [Cited 2015 March 20]. Available from: http://www.who.int/

[2] Narongsak Kiatikajornthada, “Targeted Therapy”, *Health Spot*, Bumrungrad International, August 36, 2013. [Cited 2015 March 21]. Available from: https://www.bumrungrad.com/healthspot/August-2013/targeted-therapy

[3] Meng-Hsiun Tsai, et al, “Genetic Regulatory Networks Established by Shortest Path Algorithm and Conditional Probabilistic for Ovarian Carcinoma Microarray Data”, *Proceeding of the Ninth International Conference on Machine Learning and Cybernetics*, Qingdao, pp. 32-36, July 11-14, 2010.

[4] Maya El Dayeh and Michael Hahsler, “Biological Pathway Completion using Network Motifs and Random Walks on Graphs”, *Proceeding of the 2012 IEEE*, pp. 229-236, 2012.

[5] Deanna Petrochilos and Neil Abernethy, “Assessing Network Characteristics of Cancer Associated Genes in Metabolic and Signaling Networks”, *Proceeding of the 2012 IEEE*, pp. 290-297, 2012.

[6] Kyoto Encyclopedia of Genes and Genomes: Pathway. [Cited 2016 August 27]. Available from: http://www.genome.jp/kegg/pathway.html

[7] ClinicalTrials.gov. [Cited 2016 August 27]. Avalible from: https://clinicaltrials.gov

[8] FDA@Drugs. [Cited 2016 August 27]. Avalible from: https://www.accessdata.fda.gov

[9] Linton C. Freeman, “Centrality in Social Networks Conceptual Clarification”, *Social Network*, 1(1978/79)215-239.

[10] Alex Bavelas, “Communication patterns in task-oriented group”, *J. Acoust. Soc. Am*, 22(6):725-730, 1950.

[11] Linton C. Freeman, “A set of measures of centrality based upon betweenness”, *Sociometry*, 40:35-41, 1997.

การแนะนำลูกค้ากลุ่มเป้าหมายบนทวิตเตอร์สำหรับสินค้าผ้าทอพื้นเมืองภาคใต้
Potential Customer Recommendation on Twitter for Local Woven Fabric Products

นิชากรณ์ พันธ์คง (N. Pankong)[1] เสาวลักษณ์ อร่ามพงศานุวัต (S. Arampongsanuwat)[2]
และนริศรา นาคเมธี (N. Nakmaetee)[3]
[1]สาขาคอมพิวเตอร์และเทคโนโลยีสารสนเทศ คณะวิทยาศาสตร์ มหาวิทยาลัยทักษิณ
[2]ภาควิชาคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยศิลปากร
[3]สาขาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยเทคโนโลยีราชมงคลพระนคร
[1]nichakorn.tsu@hotmail.com, [2]arampongsanuwat_s@su.ac.th, [3]narisara.n@rmutp.ac.th

บทคัดย่อ

บทความฉบับนี้มีวัตถุประสงค์เพื่อค้นหาความสนใจของลูกค้ากลุ่มเป้าหมาย โดยใช้ประโยชน์จากข้อความบนทวิตเตอร์ เพื่อเปรียบเทียบเทคนิคการจำแนกประเภทข้อความที่เหมาะสมกับผลิตภัณฑ์ผ้าทอพื้นเมืองภาคใต้ โดยการสร้างแบบจำลองความสนใจของผู้ใช้ และการแนะนำผลิตภัณฑ์ผ้าพื้นเมืองภาคใต้ ซึ่งประกอบด้วย 3 ขั้นตอนคือ 1) เก็บรวบรวมข้อมูลจากข้อความบนทวิตเตอร์ 2) การจำแนกประเภทสินค้าด้วยเหมืองข้อความ 3) อนุมานความสนใจสินค้าของผู้ใช้ โดยเทคนิคข้อมูลหลากหลายมิติ (Multi Dimensions Data) ด้วยการคำนวณค่าน้ำหนักความสนใจของผู้ใช้ในภาพรวม (User Interest Weights) ผนวกกับการคำนวณค่าน้ำหนักความสนใจของผู้ใช้เฉพาะบุคคล (User Interest Ranking) ผลการทดลองพบว่า เทคนิคที่มีประสิทธิภาพสูงสุดในการจำแนกประเภทข้อความคือเทคนิคเพื่อนบ้านใกล้เคียง ที่ระดับค่าอัตราการรู้จำ 90% เมื่ออนุมานความสนใจสินค้าจากผู้ใช้ด้วยเทคนิคค่าเฉลี่ยความสนใจเนื้อหา พบว่า ประสิทธิภาพการอนุมานสินค้าและการนำแนะนำสินค้ามีค่าความถูกต้องที่ 65%

คำสำคัญ: การอนุมานความสนใจของผู้ใช้ โมเดลการแนะนำทวิตเตอร์

Abstract

The objective of this paper is to discover the potential customer interest extraction by taking advantage of tweets on Twitter to compare the text classification techniques appropriate to the Thai southern woven fabric that are suitable for developing user interest recommendation. In order to develop the recommendation model consists of three steps: 1) Collect retweet messages on Twitter by using Twitter API and extract keyword by longest matching based on dictionary based and Thai fabric dictionary. 2) Product classification by text mining. 3) Inference the preference product of user with Multi Dimensions Data technique, which is a combination of User Interest Weights (UIW), and User K-Nearest Neighbor algorithm, yielded the F-measure of 90%. Moreover, the inference algorithm and the recommended product model give the accuracy of 65%.

Keyword: Inferring User Interests, Recommendation Model, Twitter.

1. บทนำ

การเปลี่ยนแปลงที่สำคัญในยุคไทยแลนด์ 4.0 คือการพัฒนาเศรษฐกิจดิจิทัล โดยธุรกิจนำสื่อสังคมออนไลน์มาประยุกต์ใช้เพื่อการวางแผนยุทธศาสตร์ด้านการตลาดและการสนับสนุนการตัดสินใจ ที่สามารถรองรับการทำงานด้านการบริหารลูกค้าสัมพันธ์และการกำหนดลูกค้ากลุ่มเป้าหมายของแต่ละผลิตภัณฑ์ ทำให้เกิดการบริการข้อมูลเกี่ยวกับผลิตภัณฑ์ อันจะเป็นผลดีต่อธุรกิจรายย่อย (SME) ที่สามารถใช้งานสื่อสังคมออนไลน์เพื่อวางกลยุทธ์การแข่งขันด้านการตลาด และการเพิ่มช่องทางการค้าให้กับธุรกิจรายย่อย โดยนำรูปแบบการปฏิสัมพันธ์บนสื่อสังคมออนไลน์ที่สามารถรองรับขั้นตอนการ

ตัดสินใจซื้อสินค้า คือ การพิจารณาคุณสมบัติสินค้า การแสดงความสนใจสินค้าจากผู้ใช้ การซื้อสินค้าและบริการหลังการขาย เป็นต้น [1] ,[2], [3]

อย่างไรก็ตามแม้ประเทศไทยจะมีการใช้งานสื่อสังคมออนไลน์อย่างแพร่หลาย แต่มีการประยุกต์เพื่อสร้างกลยุทธ์แข่งขันให้กับผลิตภัณฑ์ชุมชนยังไม่เพียงพอโดยเฉพาะอุตสาหกรรมสิ่งทอและเครื่องนุ่งห่มขนาดเล็ก มักขาดความเข้าใจด้านการตลาดและไม่สามารถสกัดข้อมูลด้านการตลาดออกมาเป็นกลยุทธ์ ส่งผลให้ขาดการวางตำแหน่งทางการแข่งขันของธุรกิจ (Strategic Positioning) ที่ชัดเจน เน้นผลิตสิ่งที่ชำนาญโดยไม่ให้ความสำคัญการศึกษาลูกค้ากลุ่มเป้าหมายและการรักษาฐานลูกค้าปัจจุบัน ส่งผลให้ผลิตสินค้าขาดอัตลักษณ์ที่ชัดเจน

ดังนั้น การวิเคราะห์ลูกค้ากลุ่มเป้าหมาย โดยการสื่อสารเพื่อเพิ่มโอกาสในการสร้างแบรนด์และมูลค่าเพิ่มให้กับสินค้านับว่าเป็นเรื่องที่สำคัญที่ผู้ประกอบการสามารถเข้าถึงข้อมูลเชิงลึกด้านความต้องการสินค้าของลูกค้า โดยเฉพาะข้อมูลที่สำคัญคือ ข้อมูลทางการตลาด ความเข้าใจเรื่องพฤติกรรมผู้บริโภค การวางตำแหน่งของสินค้า และช่องทางการจัดจำหน่ายที่เหมาะสมกับกลุ่มเป้าหมาย เป็นต้น [4]

ผู้วิจัยจึงมีแนวคิดในการพัฒนาโมเดลระบบแนะนำข้อมูลความสนใจผลิตภัณฑ์ผ้าทอพื้นเมืองภาคใต้เฉพาะบุคคล เพื่อช่วยแนะนำข้อมูลผลิตภัณฑ์ให้ตรงตามความสนใจของกลุ่มลูกค้าอย่างแท้จริง และนำเสนอแนวคิดในการพัฒนาเทคนิคการค้นหาความสนใจเฉพาะบุคคลเพื่อค้นหาลูกค้ากลุ่มเป้าหมาย โดยใช้ประโยชน์จากการรีทวีต (Retweet) ข้อความบนทวิตเตอร์ ซึ่งเป็นพฤติกรรมบนสื่อสังคมออนไลน์ในการแบ่งปันข้อความจากพฤติกรรมการแสดงความสนใจต่อสินค้าในแบบโดยตรง (Explicit Relationship)

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ปัจจุบันการทอผ้ายังคงเป็นมรดกทางวัฒนธรรม โดยเฉพาะการทอผ้าในภาคใต้ที่มีชื่อเสียงในเรื่องของการทอผ้ายก ทั้งประเภทผ้ายกฝ้าย ยกไหม ยกดิ้นเงินดิ้นทอง สำหรับลวดลายยังคงอนุรักษ์ลวดลายดั้งเดิมไว้ เช่น ลายราชวัตร ลายดอกพิกุล ลายลูกแก้ว ลายก้านแย่ง และลายรูปสัตว์ต่างๆ แหล่งทอผ้ายกที่สำคัญในภาคใต้ คือ จังหวัดนครศรีธรรมราช สุราษฎร์ธานี สงขลา และตรัง การทอผ้าในภาคใต้ของไทยได้ใช้ภูมิปัญญาในการสืบทอดการทอลวดลายผ้าดั้งเดิมและมีการประดิษฐ์คิดค้นลวดลายขึ้นมาใหม่จนเป็นเอกลักษณ์เฉพาะของผ้าพื้นเมืองภาคใต้ [5]

2.1 โมเดลการวิเคราะห์ความสนใจ (User Interests Model)

เทคโนโลยีของระบบแนะนำมีส่วนสำคัญต่อการพัฒนานวัตกรรมของเว็บสังคมออนไลน์ ซึ่งมักมีกระบวน การค้นพบความสนใจที่หลากหลายของผู้ใช้และทำให้เกิดกลุ่มผู้ใช้หลากหลายล้านคน โดย Berkovsky และคณะ [6] ทำการวิจัยเกี่ยวกับการตรวจสอบลักษณะเฉพาะของบุคคลเพื่อแสดงผลลัพธ์ความสนใจของผู้ใช้แต่ละคน โดยสามารถคัดกรองตามความนิยมสินค้าและบริการสามารถแบ่งเป็น 2 ประเภทคือ โมเดลการรวมกลุ่ม (Aggregation Model) คือ การรวมกลุ่มผู้ใช้จากค่าความคล้ายคลึงกันของประวัติผู้ใช้ โดยการใช้ค่าน้ำหนักความพึงพอใจจากผู้ใช้ และโมเดลการทำนายกลุ่ม (Prediction Aggregation Model) คือการวิเคราะห์ค่าความคล้ายคลึงระหว่างประวัติของผู้ใช้ โดยแบ่งเป็น 3 ประเภทดังนี้

1) วิธีการการแนะนำโดยเนื้อหา (Content Based) เป็นเทคนิคการให้คำแนะนำที่ใช้พื้นฐานจากข้อมูลเนื้อหาแนะนำความสนใจของผู้ใช้ โดยตัดสินใจจากคุณลักษณะของข้อมูลที่สามารถบ่งชี้ได้ว่าผู้ใช้มีความชอบเนื้อหา เทคนิคที่นิยมใช้คือ วิธีความคล้ายคลึงโคซายน์ (Cosine Similarity) เป็นวิธีการหาค่าความคล้ายคลึงกันจากค่าความต่างของมุมของข้อมูล 2 กลุ่มที่เกิดขึ้นบนพื้นที่เวกเตอร์ (Vector Space) ซึ่งค่าความคล้ายคลึงกันแบบโคซายน์นี้จะมีค่าอยู่ระหว่าง 0 - 1 เท่านั้น

2) การแนะนำแบบแบบพึ่งพาผู้ใช้ร่วม (Collaborative Filtering) คือเทคนิคการใช้ข้อมูลพฤติกรรมของผู้ใช้และผู้เกี่ยวข้อง เพื่อใช้แนะนำการกระทำที่น่าสนใจโดยทำนายสิ่งของที่มีประโยชน์ให้แก่ผู้ใช้ซึ่งจะขึ้นอยู่กับสิ่งของที่ผู้ใช้คนอื่นให้คะแนนไว้ก่อนหน้า วิธีการพึ่งพาผู้ใช้ร่วมมี 2 แบบ คือ วิธีการกรองข้อมูลแบบพึ่งพาผู้ใช้ร่วม (User Based) และการกรองข้อมูลแบบวัตถุร่วม (Item-Based)

3) การแนะนำแบบผสมผสาน (Hybrid Recommender) วิธีผสมผสานโดยการรวมการแนะนำแบบพิจารณาเนื้อหา และแบบการกรองแบบพึ่งพาร่วมกันเข้าด้วยกัน ด้วยเทคนิคการผสมผสานแบบร่วมและแยกกัน (Combining and Separate) เทคนิคการ

ผสมผสานแบบเพิ่ม (Adding Content Based) การนำผลลัพธ์ของการแนะนำแบบพิจารณาเนื้อหามาเพิ่มในวิธีการแบบการกรองร่วมกัน และเทคนิคการผสมผสานแบบเพิ่ม (Adding Collaborative) การนำผลลัพธ์ของวิธีการแบบการกรองร่วมกันมาเพิ่มในการแนะนำแบบพิจารณาเนื้อหา [7], [8], [10]

2.2 เหมืองข้อความ

การวิเคราะห์เหมืองข้อความในงานวิจัยนี้ จำแนกกระบวนการเตรียมข้อมูลออกเป็น 3 กระบวน ดังนี้ การเลือกคุณลักษณะ (Feature Selection) การกลั่นกรองข้อความ (Text Cleaning) การแทนข้อความ (Text Representation) เพื่อการจำแนกประเภทเนื้อหา (Text Categorization) วิธีการเรียนรู้ของเครื่อง (Machine Learning Approach) กลไกการเรียนรู้ของเครื่องสามารถสร้างกฎและประเมินทางเลือกในการตัดสินใจตามกฎแต่ละข้อเป็นเครื่องมือตัวหนึ่งที่จะทำให้ได้มาซึ่งองค์ความรู้และยังสามารถสร้างนวัตกรรมใหม่ขึ้นมาได้ [9] วิธีการเรียนรู้ของเครื่องประกอบด้วยวิธีการต่าง ๆ เช่น วิธีเพื่อนบ้านที่ใกล้เคียงที่สุด (K-Nearest Neighbor: KNN) คือวิธีการจำแนกประเภทข้อมูล โดยการหาความใกล้เคียงของข้อมูลอินพุตที่ต้องการทำนาย กับข้อมูลของสิ่งที่มีอยู่ก่อนแล้ว และนำมาเรียงลำดับตัวที่ใกล้ที่สุด และเลือกใกล้ที่สุด โดยจำนวนตัวที่ใกล้ที่สุดที่จะเลือกให้แทนจำนวน k ตัว ซึ่งกำหนดให้ เป็นค่าพารามิเตอร์จำนวนเพื่อนบ้านที่สนใจ เพื่อการตรวจสอบเงื่อนไขที่เหมือนกันหรือใกล้เคียงกันมากที่สุด ถ้า k=3 หมายถึงจะพิจารณาข้อมูลจำนวน 3 ชุดที่อยู่ใกล้เคียงจุดที่ต้องการพิจารณา เพื่อคำนวณระยะห่าง (Distance) ของข้อมูลที่ต้องการพิจารณากับกลุ่มข้อมูลตัวอย่าง เทคนิควิธีการวัดระยะห่างสามารถทำได้หลายวิธี เช่น Euclidian Distance, Manhattan Distance เป็นต้น โดยจัดเรียงลำดับของระยะห่างและเลือกพิจารณาชุดข้อมูลที่ใกล้จุดที่ต้องการพิจารณาตามจำนวน k ที่กำหนดไว้ พิจารณาข้อมูลจำนวน k ชุด และพิจารณาประเภทข้อมูลใดที่ใกล้จุดที่พิจารณาเป็นจำนวนมากที่สุด [11]

2.3 งานวิจัยที่เกี่ยวข้อง

Shih และคณะ [12] ทำการวิจัยเกี่ยวกับการแนะนำสไตล์เสื้อผ้าให้กับลูกค้าด้วยเทคนิคการกรองแบบพึ่งพาร่วมกันจากข้อความบนทวิตเตอร์ โดยเก็บรวบรวมข้อมูลเกี่ยวกับคุณสมบัติผ้าจำนวน 60 ประเภท สไตล์ของเสื้อผ้า 41 สไตล์ และคำสำคัญที่ใช้แสดงถึงความรู้สึกกับสินค้าที่เกี่ยวข้องกันกันจำนวน 28 คำสำคัญ ผลการวิจัยพบว่าความพึงพอใจและความรู้สึกของลูกค้าที่มีต่อสินค้าเป็นเรื่องที่มีความสำคัญ อันจะส่งผลต่อการการพยากรณ์ความสนใจสินค้าโดยเฉพาะสินค้ากลุ่มผ้าและเครื่องแต่งกาย ซึ่งมีตลาดทางด้านพาณิชย์อิเล็กทรอนิกส์ที่มีขนาดใหญ่มาก ดังนั้นการพัฒนาระบบแนะนำเพื่อช่วยให้นักออกแบบเสื้อผ้าสามารถผลิตสินค้าตามความต้องการของลูกค้าได้ถือเป็นเรื่องสำคัญ โดยพบว่า ความยาวของกระโปรงส่งผลต่อความพึงพอใจของลูกค้าผู้หญิงมากที่สุด

นอกจากนั้น เทคนิคการกรองแบบพึ่งพาร่วมกัน Seo, J.-H., และ คณะ[13] นำมาใช้พัฒนาระบบแนะนำเพื่อการออกแบบที่สามารถเชื่อมต่อกับความรู้สึกและการออกแบบผลิตภัณฑ์ เพื่อเป็นสื่อกลางในการติดต่อสื่อสารระหว่างนักออกแบบและร้านค้าในธุรกิจการออกแบบ โดยระบบสามารถวางแผนผลิตภัณฑ์ของผู้ขายและเชื่อมต่อกับโรงงานอุตสาหกรรมสิ่งทอ เพื่อคำนวณค่าความคล้ายคลึงระหว่างผู้ใช้โดยใช้เทคนิคการหาค่าเพื่อนบ้านที่ใกล้เคียงที่สุด เพื่อวิเคราะห์ผู้ใช้ที่มีพฤติกรรมใกล้เคียงที่สุด โดยหาค่าน้ำหนักระหว่างผู้ใช้และค่าสัมประสิทธิ์สหสัมพันธ์เพียร์สัน เพื่อวิเคราะห์หาผู้ใช้ที่มีความคล้ายคลึงกันในการออกแบบเสื้อผ้า

3. วิธีการดำเนินการวิจัย

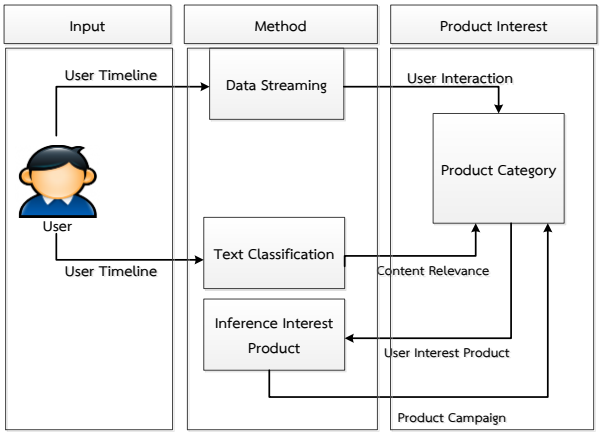
3.1 การพัฒนาแบบจำลองการแนะนำลูกค้ากลุ่มเป้าหมาย

งานวิจัยนี้เสนอขั้นตอนการพัฒนาแบบจำลองการแนะนำจากเนื้อหาของผู้ใช้บนทวิตเตอร์ซึ่งประกอบด้วย 3 ขั้นตอนคือ

3.1.1 เก็บรวบรวมข้อมูลจากผู้ประกอบการที่ใช้งานทวิตเตอร์โดยการโพสต์หรือแบ่งปันข้อมูลเกี่ยวกับสินค้าผ้าพื้นเมืองภาคใต้ที่ กลุ่มเป้าหมายทั้งหมดเป็นร้านค้าออนไลน์ที่ขายสินค้าผ้าพื้นเมืองภาคใต้ จำนวน 30 ร้าน โดยแต่ละร้านมีผู้ติดตามอย่างน้อย 400 คน เพื่อดำเนินการเก็บข้อมูลของผู้ใช้โดยพัฒนาโปรแกรมประยุกต์สำหรับเก็บรวบรวมข้อมูลจากกระดานสถานะโดยใช้ทวิตเตอร์เอพีไอในการเชื่อมต่อกับข้อมูลผู้ใช้

3.1.2 วิธีการวิเคราะห์ความสนใจของผู้ใช้จากความสัมพันธ์ แบบทางตรง จากการแบ่งปันข้อความไปยังผู้อื่น (Retweet) การโพสต์ข้อความ (Tweet) และการถูกกล่าวถึง

หรือนำไปใส่ในข้อความ Tweets ของผู้อื่น โดยการสกัดคำสำคัญจากข้อความบนทวิตเตอร์ ด้วยเทคนิคการเลือกคำที่ยาวที่สุด (Longest Matching) โดยอ้างอิงจากพจนานุกรม และทดลองเปรียบเทียบประสิทธิภาพการจำแนกประเภทเนื้อหาจาก 4 เทคนิค คือ วิธีโครงสร้างต้นไม้ วิธีการข่ายงานเบย์ วิธีเพื่อนบ้านใกล้เคียง และวิธีการซัพพอร์ตเวกเตอร์แมชชีน เพื่อนำผลที่ได้ไปทำการวิเคราะห์กลุ่มสินค้าที่ผู้ใช้สนใจ (Content Interest Product) จากผลการจำแนกประเภทเนื้อหาในแต่ละข้อความของผู้ใช้แต่ละคน โดยสรุปขั้นตอนวิธีการวิเคราะห์ความสนใจสินค้าของลูกค้าบนสื่อสังคมออนไลน์โดยใช้เนื้อหาดังภาพที่ 1



ภาพที่ 1: ขั้นตอนวิธีการวิเคราะห์ความสนใจสินค้า

จากภาพที่ 1 เมื่อเก็บรวบรวมข้อมูลผู้ใช้ทวิตเตอร์ ทำการพิจารณาประเภทสินค้าจากเนื้อหาที่ผู้ใช้โพสต์ เพื่อจัดทำโมเดลชุดการเรียนรู้ และเปรียบเทียบประสิทธิภาพจาก 4 เทคนิค คือ วิธีเพื่อนบ้านใกล้เคียงที่สุด วิธีโครงสร้างต้นไม้ วิธีการข่ายงานเบย์ และวิธีการซัพพอร์ตเวกเตอร์แมชชีน เพื่อประเมินประสิทธิภาพโดยใช้ค่าความแม่นยำ ค่าความระลึกถึง และค่าอัตราการรู้จำ

3.1.3 การอนุมานความสนใจสินค้า (User Inference) โดยใช้เทคนิคข้อมูลหลากหลายมิติ เป็นการผนวกความสนใจของผู้ใช้ที่มีต่อสินค้า (User Product Rating: UPR) และการแบ่งปันเนื้อหาบนของผู้ใช้ที่รีทวีตเกี่ยวกับสินค้า (User Content Rating: UCR) ดังตัวอย่างข้อมูลในตารางที่ 1 และ 2 ตามลำดับ ซึ่งผลจากการผนวกข้อมูลจาก UPR และ UCR เกิดเป็นเมทริกซ์ความสัมพันธ์จากความสนใจสินค้าร่วมกับเนื้อหาที่ผู้ใช้รีทวีต (User Product Preference: UPP) ดังตารางที่ 3

ตารางที่ 1: เมทริกซ์ความสนใจสินค้าของผู้ใช้ (UPR)

Users	Products		
	p1	p2	p3
u1	5	2	1
u2	1	4	4
u3	1	0	5

จากตารางที่ 1 เมื่อกำหนดให้ระบบแนะนำมีผู้ใช้จำนวน 3 คน (U = u1,u2,u3) และมีสินค้าจำนวน 3 รายการ (P = p1, p2, p3) โดยกำหนดให้ความพึงพอใจต่อสินค้า p1 – p3 อยู่ในระดับ 0 - 5 คะแนน

ตารางที่ 2: เมทริกซ์เนื้อหาที่ผู้ใช้รีทวีตเกี่ยวกับสินค้า

Products	Content			
	urt1	*urt2*	*urt3*	*urt4*
p1	1	0	0	0
P2	1	1	1	0
P3	1	0	0	1

จากตารางที่ 2 เป็นรายละเอียดเกี่ยวกับเนื้อหาที่ผู้ใช้รีทวีตเกี่ยวกับสินค้าแต่ละรายการ เมื่อกำหนดให้มีผู้ใช้รีทวีตเกี่ยวกับสินค้าจำนวน 4 คน ($U = \{urt1, urt2, urt3, urt4\}$) โดยแต่ละข้อความสามารถอธิบายคุณลักษณะที่จำแนกเนื้อหาตามประเภทสินค้า ในรูปแบบข้อมูล Binary จากการสกัดคำสำคัญตามพจนานุกรมผ้าไทย โดยแต่ละข้อความอาจกล่าวถึงสินค้าได้หลายประเภท

จากตัวอย่างในตารางที่ 3 เป็นการผนวกความสนใจสินค้าของผู้ใช้และความสนใจสินค้าแยกตามเนื้อหาที่รีทวีต ดังนั้นเมื่อสินค้ารายการที่ p1 มีการรีทวีตจากผู้ใช้ คนที่ urt1 และ urt2 โดยกำหนดให้การแทนค่า Binary ดังนี้ 1= มีคำสำคัญที่บ่งบอกคุณลักษณะ และ 0= ไม่มีคำสำคัญที่สามารถบ่งบอกคุณลักษณะความสนใจสินค้าจากผู้ใช้ที่รีทวีตข้อความ

ตารางที่ 3: เมทริกซ์ความสนใจจากสินค้าและข้อความ (UPP)

Users	Content					
	Products	*urt1*	*urt2*	*urt3*	*urt4*	Total
u1	p1	1	0	0	0	1
	p2	1	1	1	0	3
	p3	1	0	0	1	2
u2	p1	1	0	0	0	1
	p3	1	0	0	1	2
u3	p2	1	1	1	0	3
	p3	1	0	0	1	2

- การคำนวณค่าน้ำหนักความสนใจของผู้ใช้ (User Interest Weights) เป็นขั้นตอนการวิเคราะห์ค่าความพึงพอใจจากภาพรวมของผู้ใช้ที่มีต่อสินค้าทั้ง 9 ประเภทข้อความ โดยคำนวณจากความถี่ในการทวีตสินค้าแต่ละประเภทและทำการปรับค่าระดับคะแนนให้อยู่ระว่าง 0 – 1 คะแนน ดังสมการที่ 1

$$UIW_{x,c} = \left(\frac{1}{\text{MAX}_{c\in C_i}}\right)\sum_{i\in p_x} r_{i,c} \qquad (1)$$

เมื่อกำหนดให้ p_x แทนเซตของสินค้าที่ผู้ใช้พึงพอใจ และ $r_{i,a}$ แทนค่าข้อความที่เกี่ยวกับประเภทสินค้าที่ i โดยที่ 1 แทนข้อความที่เกี่ยวกับประเภทสินค้า และ 0 แทนข้อความที่ไม่เกี่ยวกับประเภทสินค้า

โดยที่ทำการปรับค่าถ่วงน้ำหนักคะแนนจะให้ C_i แทนเซตของข้อความรีทวีตของสินค้าประเภทที่ i โดยมีตัวอย่างการคำนวณค่าน้ำหนักความสนใจของผู้ใช้คนที่ 1 ดังนี้

เมื่อ Max = 3

$UIW_{u1,rt1} = 3/3 = 1$ $UIW_{u1,rt2} = 1/3 = 0.33$

$UIW_{u1,rt3} = 1/3 = 0.33$ $UIW_{u1,rt4} = 1/3 = 0.33$

ตารางที่ 4: ค่าน้ำหนักความสนใจของผู้ใช้ (UIW)

Users	Content			
	urt1	*urt2*	*urt3*	*urt4*
u1	1	0.33	0.33	0.33
u2	1	0	0	0.5
u3	1	0.5	0.5	0.5

จากตารางที่ 4 สามารถอนุมานความสนใจของผู้ใช้คนที่ u1 จากน้ำหนักความสนใจ โดยเขียนในรูปของ vector ดังนี้ UIW_{u1} = (1, 0.33, 0.33, 0.33) เพราะฉะนั้นผู้ใช้คนที่ 1 คือ u1 แสดงความสนใจจากสินค้าข้อความรีทวีตที่ 1 คือ urt1 ที่ระดับคะแนน 1 มากที่สุด

- การคำนวณค่าน้ำหนักความสนใจของผู้ใช้เฉพาะบุคคล (User Interest Ranking) โดยวิเคราะห์เนื้อหาที่เกี่ยวข้องตามประเภทสินค้าตามความพึงพอใจของผู้ใช้ x ที่มีต่อสินค้า และปรับค่าความชอบจากช่วงคะแนน 1-5 จากตารางที่ 1 ให้เป็น 0-1 ตามระดับความสนใจ ด้วยการหารด้วยค่าของประเภทสินค้าที่มีค่าสูงสุด ดังสมการที่ 2

$$I_{x,i,c} = \frac{r_{i,c}.uiw_c}{\text{MAX}_{c\in C_i}\left(r_{i,c}.uiw\right)} \qquad (2)$$

เมื่อกำหนดให้ $I_{x,i,c}$ แทนระดับความสนใจของผู้ใช้ x จากข้อความที่ c ของสินค้าประเภท i จากนั้นนำค่าระดับความสนใจของสมการที่ 2 มาคูณกับค่าน้ำหนักความสนใจของผู้ใช้โดยการแทนค่า Binary ความสนใจ ที่มีค่า 0 หรือ 1 ด้วยค่า uiw_c และคำนวณค่าความสนใจเฉพาะบุคคล ดังสมการที่ 3

$$IRP_{x,i,c} = I_{x,i,c} \cdot ui_i \qquad (3)$$

เมื่อกำหนดให้ $IRP_{x,i,c}$ คือค่าความสนใจที่ผู้ใช้ x จากการสกัดคุณลักษณะข้อความที่ c ของสินค้าประเภทที่ i และ ui_i แทนค่าความชอบโดยรวมที่ผู้ใช้ x มีต่อสินค้าประเภทที่ i

ดังตัวอย่าง เมื่อผู้ใช้คนที่ u1 เลือกสินค้ารายการที่ p2 เมื่อทำการแทนค่า Binary ด้วยค่าน้ำหนักความสนใจรวม $UIW_{x,c}$ ของผู้ใช้คนที่ u1 โดยสามารถคำนวณค่าความสนใจของผู้ใช้คนที่ u1 แยกตามประเภทสินค้าและเนื้อหาที่รีทวีตได้ดังนี้

$$RP_{u1,p1,urt1} = 1x1 = 1$$
$$RP_{u1,p2,urt1} = 1x3 = 3$$
$$RP_{u1,p3,urt1} = 1x2 = 2$$

เพราะฉะนั้น สามารถอนุมานความสนใจสินค้าของผู้ใช้ u1 จากข้อความรีทวีตที่ urt1 โดยเขียนในรูปของ vector ดังนี้ $RP_{u1,urt1}$ = (1, 3, 2) ดังนั้นจากข้อความที่รีทวีตพบว่า ผู้ใช้มีโอกาสที่จะซื้อสินค้ารายการที่ p2 มากที่สุด รองลงมาเป็นรายการ p3 และ p1 ตามลำดับ

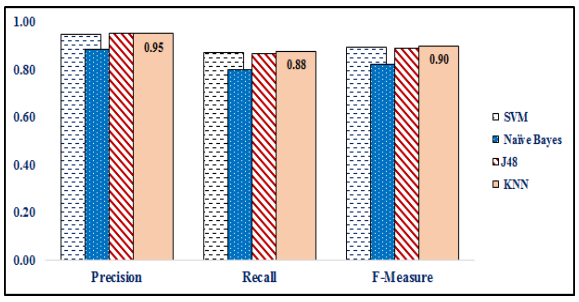
3.1.4 การคำนวณค่าความคล้ายคลึงระหว่างผู้ใช้ ด้วยเทคนิคที่นิยมใช้คือวิธีความคล้ายคลึงโคซายน์ (Cosine Similarity) ดังสมการที่ 4 โดยพิจารณาคะแนนสูงสุด 3 อันดับ

$$sim\left(U_{i,}\ I_{j}\right) = \cos\left(U_{i,}\ I_{j}\right) = \frac{\sum_{i=1}^{k} u_{i,l}\, i_{j,l}}{\sqrt{\sum_{i=1}^{k} u_{i,l}^{2}}\sqrt{\sum_{i=1}^{k} i_{i,l}^{2}}} \qquad (4)$$

4. ผลการดำเนินงาน

จากการเก็บรวบรวมข้อมูลกลุ่มเป้าหมายจำนวน 30,000 ข้อความ นำมาสกัดคำสำคัญ ผลการวิจัยพบว่า จำนวนคำสำคัญของพจนานุกรมผ้าไทยและสามารถสกัดประเภทสินค้าทั้งหมด 9 ประเภท ประกอบด้วย ผ้าบาติก ผ้าทอพิมพ์ลาย ผ้าทอเกาะยอ ผ้าขาวม้า ผ้าถุง ผ้าคลุมไหล่ เสื้อเชิ้ต ผ้าไหม และชุดไทย

ผลการวิเคราะห์ความสนใจของผู้ใช้ ด้วยเทคนิคเหมืองข้อความ เพื่อจำแนกประเภทสินค้าที่ผู้ใช้สนใจจากเนื้อหา ดังภาพที่ 2 พบว่าประสิทธิภาพค่าอัตราการรู้จำของเทคนิคเพื่อนบ้านใกล้เคียงที่ ให้ประสิทธิภาพการจำแนกประเภทเนื้อหาสูงที่สุด ที่ระดับ 90% ค่าความแม่นยำที่ระดับ 95% และค่าระลึกถึงที่ระดับ 88% โดยการแบ่งข้อมูลออกเป็น 10 ส่วน โดยที่แต่ละส่วนมีจำนวนข้อมูลเท่ากัน จากนั้นข้อมูลหนึ่งส่วนจะใช้เป็นตัวทดสอบประสิทธิภาพของโมเดลทำวนไปเช่นนี้จนครบจำนวนที่แบ่งไว้ ดังภาพที่ 2



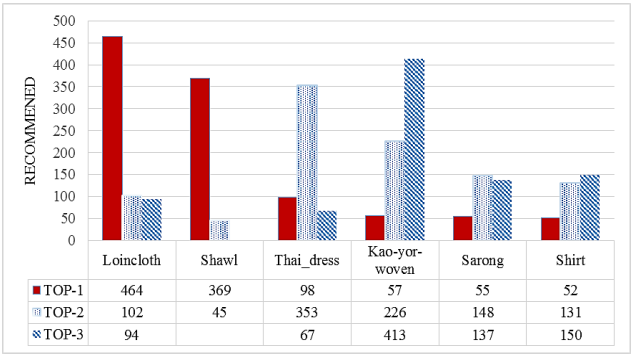
ภาพที่ 2: การเปรียบเทียบประสิทธิภาพการจำแนกประเภทเนื้อหา

เมื่อเปรียบเทียบค่าอัตราการรู้จำ พบว่าประสิทธิภาพเฉลี่ยโดยรวมดีที่สุดด้วยวิธีการเพื่อนบ้านใกล้เคียงพบว่า สินค้าประเภท ผ้าขาวม้า ผ้าลายดอก ผ้าทอเกาะยอ ผ้าถุง และเสื้อเชิ้ต ที่ระดับ 100% นอกจากนั้นสินค้าที่มีประสิทธิภาพอัตราการรู้จำต่ำสุดคือ สินค้าประเภทผ้าไหม ดังตารางที่ 5

ตารางที่ 5: การเปรียบเทียบค่าอัตราการรู้จำแยกตามเทคนิคการจำแนกเนื้อหาและประเภทสินค้า

ประเภทสินค้า	เทคนิคการจำแนกเนื้อหา			
	SVM	Naïve Bayes	J48	KNN
ผ้าขาวม้า	0.528	0.486	0.501	1.000
ผ้าลายดอก	1.000	1.000	1.000	1.000
ผ้าทอเกาะยอ	0.914	0.689	0.920	1.000
ผ้าบาติก	0.963	0.866	0.961	0.999
ผ้าถุง	0.907	0.819	0.912	1.000
ผ้าไหม	0.858	0.762	0.795	0.965
ผ้าคลุมไหล่	0.880	0.879	0.880	0.976
ชุดไทย	0.982	0.945	0.977	0.993
เสื้อเชิ้ต	0.806	0.818	0.818	1.000

ในขณะที่การอนุมานความสนใจสินค้าด้วยเทคนิคข้อมูลหลากหลายมิติของผู้ใช้บนทวิตเตอร์ ได้ทำการทดลอง จากชุดการเรียนรู้ของโมเดลจำนวน 30,000 ข้อความ และชุดทดสอบจำนวน 10,000 ผลการทดลองพบว่า การอนุมานความสนใจเพื่อการแนะนำสินค้าที่ตรงกับความสนใจของผู้ใช้มากที่สุดอันดับ 1 คือ ผ้าขาวม้า เป็นผลิตภัณฑ์ที่ได้รับการแนะนำให้กับผู้ใช้สูงสุดจำนวน 464 คน คิดเป็นร้อยละ 46 รองลงมาเป็นผ้าคลุมไหล่ แนะนำให้กับผู้ใช้จำนวน 369 คน คิดเป็นร้อยละ 37 และชุดไทย แนะนำให้กับผู้ใช้จำนวน 98 คน คิดเป็นร้อยละ 9.8 จากผู้ใช้จำนวน 1,000 คน ตามลำดับ ดังภาพที่ 3 ทั้งนี้การประเมินผลโมเดลการแนะนำสามารถวัดค่าความถูกต้องจากการแนะนำสินค้าได้ที่ระดับความถูกต้องร้อยละ 65



ภาพที่ 3: การเปรียบเทียบการแนะนำสินค้าให้กับผู้ใช้

5. สรุป

ผลการจำแนกประเภทเนื้อหาบนทวิตเตอร์ พบว่า ข้อมูลอยู่ในสภาวะของความแม่นยำสูง (High Precision) และค่าความระลึกถึงสูง (High Recall) อันเกิดจากสาเหตุคำสำคัญของประเภทผ้าทอพื้นเมือง ร้านค้ามีการใช้กลุ่มคำสำคัญของสินค้าแต่ละประเภทใกล้เคียงกันไม่มีการกระจายการใช้งานคำสำคัญในสินค้าที่มีความแตกต่างกัน ส่งผลให้ประสิทธิภาพการจำแนกเนื้อหาแต่ละประเภทสินค้าสูง โดยประเภทสินค้าที่มีประสิทธิภาพอัตราการรู้จำต่ำสุด พบว่าเป็นกลุ่มผ้าไหม อันมีสาเหตุเกิดจาก การโพสต์เนื้อหาเกี่ยวกับสินค้ามักมีเนื้อหาที่โพสต์มีข้อความอื่น ๆ ปะปนจำนวนมาก แต่มีคำสำคัญที่อยู่ในกลุ่มคำที่นำมาประมวลผลมีน้อย จนส่งผลให้ประสิทธิภาพต่ำสุด

ผลการอนุมานความสนใจของสินค้าจากโมเดลการแนะนำสินค้า พบว่า ความต้องการของผู้บริโภคบนตลาดออนไลน์มีสินค้าที่ได้รับความสนใจสูงสุดคือ ผ้าขาวม้า ผ้าคลุมไหล่ และชุดไทย จากผลการวิเคราะห์ดังกล่าว สามารถนำไปใช้เป็นแนวทางในการพัฒนาผลิตภัณฑ์ที่มีการแปรรูปภาพผ้าทอพื้นเมืองภาคใต้ และภาคอื่น ๆ เพื่อเพิ่มส่วนแบ่งในตลาดผู้บริโภค และขยายช่องทางการผลิตสินค้าที่หลากหลายมากขึ้นตามความต้องการของผู้ใช้

6. ข้อเสนอแนะ

แนวทางการพัฒนางานวิจัยเพิ่มเติม ควรมีการวิเคราะห์สินค้าที่มีโอกาสขายร่วมกัน เพื่อพัฒนาแผนการตลาดรองรับการขยายตัวสินค้าที่มีความเกี่ยวเนื่องกับผ้าทอพื้นเมืองในอนาคตต่อไป รวมทั้งการวิเคราะห์คุณสมบัติสินค้าที่มีผลต่อการตัดสินใจเลือกซื้อสินค้า เช่น ลายผ้า สี วัสดุที่ใช้ในการทอ และประวัติความเป็นมาของแหล่งผลิต เป็นต้น

เอกสารอ้างอิง

[1] Chen, J., et al, “Short and Tweet: Experiments on Recommending Content from Information Streams,” *In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, New York, NY, USA: ACM, pp. 1185–1194, 2010.

[2] Banerjee, N., and Chakraborty, D, “User interests in social media sites: an exploration with micro-blogs,*” In Proceeding of the 18th ACM conference on Information and knowledge management - CIKM ’09*, pp. 1823–1826, 2009.

[3] Hassan, S., Nadzim, S. Z. A. and Shiratuddin, N. “Strategic Use of Social Media for Small Business Based on the AIDA Model,” *Procedia - Social and Behavioral Sciences*, pp. 262–269, 2015.

[4] การดี เลียวไพโรจน์, ภูมิพร ธรรมสถิตย์เดช . “*แผนปฏิบัติการส่งเสริมวิสาหกิจขนาดกลางและขนาดย่อมรายสาขาอุตสาหกรรมสิ่งทอและเครื่องนุ่งห่ม*.” สำนักงานส่งเสริมวิสาหกิจขนาดกลางและขนาดย่อม, หน้า 19–127, 2555.

[5] พัชรินทร์ เทพสุวรรณ, “วัฒนธรรมผ้า : ผ้าทอพื้นเมืองภาคใต้ สืบค้นจาก https://www.culture.pn.psu.ac.th/... /0742544001407291764.doc

[6] Berkovsky, et al,“Group-based Recipe Recommendations: Analysis of Data Ag-gregation Strategies.” *In Proceedings of the 4th ACM Conference on Recommender Systems*, pp. 111-118, 2010.

[7] M. Singh, “Social Network Based Recommender System.” *IEEE*, pp 5-6, 2013.

[8] Zafarani, R., et al, “Recommendation in Social Media. ” *In Social Media Mining: An Introduction.* Cambridge: Cambridge University Press, pp. 245-270, 2014.

[9] Rajaraman, A., and Ullman, J. D, “Mining of Massive Datasets,” *Lecture Notes for Stanford CS345A Web Mining*, Vol, 67. pp. 1-513, 2011.

[10] Ma, H., Jia, M., Zhang, D., and Lin, X, “Combining tag correlation and user social relation for microblog recommendation,” Information *Sciences*, pp. 325–337, 2017.

[11] Kanavos, A., et al, “Large Scale Implementations for Twitter Sentiment Classification,” *Algorithms*, Vol. 10, no. 1, pp.33, 2017.

[12] Shih, B.Y., Chen, C.Y. and Chen, Z.-S, “An Empirical Study of an Internet Marketing Strategy for Search Engine Optimization,” H*uman Factors and Ergonomics in Manufacturing*, Vol. 16, no. 1, pp. 61–81, 2006.

[13] Seo J.H. and Lee, K. P, “Development of Website Design Personalization Service Using Design Recommender System,*” In Proceedings of the 6th Asian Design Conference*, pp. 100–110, 2003.

การจำแนกน้ำหอมด้วยจมูกอิเล็กทรอนิกส์
Perfume Classification using Electronic Nose

พนิดา หล่อวงศ์ตระกูล (Panida Lorwongtrakool)[1] เฉลิมชัย หล่อวงศ์ตระกูล (Chaloemchai Lowongtrakool)[2]
และนิธินันท์ มาตา (Nitinan Mata)[3]
[1,2]สาขาวิชาเทคโนโลยีสารสนเทศและการสื่อสาร คณะวิทยาศาสตร์และศิลปศาสตร์ มหาวิทยาลัยเทคโนโลยีราชมงคลอีสาน
[3]สาขาวิชาเทคโนโลยีสารสนเทศ คณะวิทยาศาสตร์ มหาวิทยาลัยราชภัฏบุรีรัมย์
[1]panidajlo@gmail.com, [2]zdysys@gmail.com, [3]mata0711@gmail.com

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาจมูกอิเล็กทรอนิกส์ (e-nose) สำหรับการจำแนกน้ำหอมชนิดต่างๆ โดยใช้อาเรย์เซ็นเซอร์วัดการตอบสนองของก๊าซชนิดต่างๆ จากสารประกอบของน้ำหอมแต่ละกลิ่น ในการทดลองใช้ตัวอย่างน้ำหอม จากน้ำหอมแท้และน้ำหอมเลียนแบบอย่างละ 3 ยี่ห้อ ประกอบด้วยชุดข้อมูลที่มีคุณลักษณะ 3 รูปแบบ ได้แก่ 1) ชุดข้อมูลที่ประกอบด้วยทุกคุณลักษณะ 2) ชุดข้อมูลที่ผ่านการลดมิติโดยการสกัดคุณลักษณะ ด้วย PCA และ 3) ชุดข้อมูลที่ผ่านการลดมิติโดยใช้การคัดเลือกคุณลักษณะด้วย CfsSubsetEval, Correlation และ Gain Ratio ทำการจำแนกด้วยวิธี Multilayer Perceptron, RBF, SVM, J48, และ Bayesnet ทำการสอนและทดสอบแบบ 10 Folds Cross Validation เปรียบเทียบผลการทดลองโดยค่าความถูกต้อง ผลการทดลองแสดงให้เห็นว่า เมื่อใช้ชุดข้อมูลที่ประกอบด้วยทุกคุณลักษณะ ร่วมกับวิธี SVM Kernel Type Polynomial สามารถจำแนกน้ำหอมชนิดต่างๆ ได้ค่าความถูกต้องที่ 96.49% เห็นได้ว่าจมูกอิเล็กทรอนิกส์ มีความสามารถที่จะรู้จำกลิ่นน้ำหอมชนิดต่างๆได้ และสามารถนำจมูกอิเล็กทรอนิกส์ไปประยุกต์ใช้ในการจำแนกน้ำหอม การตรวจสอบน้ำหอมได้อย่างมีประสิทธิภาพ

คำสำคัญ: จมูกอิเล็กทรอนิกส์ การจำแนก กลิ่น น้ำหอม การลดมิติ

Abstract

The aim of this research was to develop electronic nose (e-nose) for classifying different perfumes that have different qualities by using array sensor to measure response of odor compound. In the experiment, samples of 3 brands from inspire perfume and counterfeit were used. The 3 datasets of the experiment composed of 1) all features, 2) feature extraction with PCA, and 3) feature selection with CfsSubsetEval, Correlation and Gain Ratio. Classification was done by algorithm such as Multilayer Perceptron, RBF, SVM, J48 and Bayesnet. Training and Testing were done by 10 folds cross validation and compared by accuracy. The result showed that SVM with all features have accuracy at 96.49%. Therefore, e-nose was able to recognize different types of odor and can be applied to classify or detect perfume efficiently.

Keyword: Electronic Nose, Classification, Odor, Perfume, Dimensionality Reduction.

1. บทนำ

อุตสาหกรรมต่างๆ ให้ความสำคัญในการตรวจสอบสิ่งปลอมปนไม่ว่าจะเป็นในวัตถุดิบ อาหาร เครื่องดื่ม จึงมีนักวิจัยพยายามคิดค้นวิธีการตรวจสอบ ซึ่งหนึ่งในนั้นคือการตรวจสอบด้วยกลิ่น โดยใช้จมูกอิเล็กทรอนิกส์ ที่ปัจจุบันถูกนำไปใช้อย่างแพร่หลาย เช่น การตรวจสอบการปนเปื้อนของเชื้อจุลินทรีย์ หรือสารเคมีที่อาจเกิดในขั้นตอนการผลิต การตรวจความสดของเนื้อหมู [1] ตรวจสอบเชื้อราในสตรอเบอรี่ [2] ตรวจสอบการปลอมปนของกาแฟ [3] แม้กระทั่งการวินิจฉัยโรคด้านการแพทย์ [4] [5] ซึ่งต่างก็ใช้จมูกอิเล็กทรอนิกส์เข้ามาตรวจสอบ

สำหรับอุตสาหกรรมที่ให้ความสำคัญต่อกลิ่นมากที่สุดเป็นอันดับต้น ๆ คือ อุตสาหกรรมน้ำหอม ซึ่งกำลังประสบปัญหาการปลอมแปลงผลิตภัณฑ์ น้ำหอมบางยี่ห้อผลิตจากวัตถุดิบที่มีคุณภาพและอาจมีราคาแพง ในขณะที่น้ำหอมเลียนแบบใช้วัตถุดิบที่ไม่ค่อยมีคุณภาพ แต่จำหน่ายในราคาถูกจึงทำให้อุตสาหกรรมน้ำหอมเลียนแบบทั่วโลกเพิ่มขึ้นอย่างรวดเร็ว และพยายามที่จะเลียนแบบน้ำหอมแท้ที่มีคุณภาพและราคาสูง การปลอมแปลงน้ำหอม ก่อให้เกิดความเสียหายในทางเศรษฐกิจ [6] และยังส่งผลต่อผู้บริโภคเนื่องจากน้ำหอมเลียนแบบ ใช้วัตถุดิบที่มีคุณภาพต่ำ และสารเคมีที่เป็นอันตรายเมื่อผู้บริโภคใช้เป็นเวลานาน นำไปสู่ปัญหาสุขภาพได้ น้ำหอมเลียนแบบเหล่านี้มักจะจำหน่ายในราคาที่ถูกกว่าและทำกล่องบรรจุภัณฑ์ได้อย่างใกล้เคียง ซึ่งยากในการตรวจสอบ

จมูกมนุษย์สามารถจำแนกกลิ่น เพื่อตรวจสอบคุณภาพของอาหาร ตรวจสอบน้ำหอม ตรวจสอบคุณภาพของสิ่งแวดล้อม ตรวจสอบการรั่วไหลของสารเคมี หรือแม้กระทั่งอาวุธหรือยาเสพติด แต่จมูกมนุษย์มีข้อจำกัดในการใช้งาน ที่สามารถดมกลิ่นได้บางกลิ่น ประสิทธิภาพการดมกลิ่นขึ้นอยู่กับความพร้อมของสุขภาพร่างกาย และมนุษย์เกิดความเหนื่อยล้าได้ ที่สำคัญคือ จมูกมนุษย์ไม่เหมาะสำหรับดมกลิ่นสารพิษต่าง ๆ ที่เป็นอันตราย อีกทั้งยังวัดประเมินได้ค่อนข้างยาก [7] จมูกอิเล็กทรอนิกส์ ได้จำลองและเลียนแบบการดมกลิ่นของมนุษย์ และสามารถนำไปประยุกต์ใช้ในอุตสาหกรรมได้ ผู้วิจัยจึงเล็งเห็นความสำคัญ ในการพัฒนาจมูกอิเล็กทรอนิกส์เพื่อจำแนกกลิ่นน้ำหอมแท้และน้ำหอมเลียนแบบ โดยในการพัฒนานี้จะประกอบด้วยส่วนของฮาร์ดแวร์ซึ่งใช้อาเรย์เซ็นเซอร์ในการรับกลิ่น และส่วนของการวิเคราะห์ ซึ่งใช้เทคนิคเหมืองข้อมูลในการจำแนกกลิ่น

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 จมูกอิเล็กทรอนิกส์

จมูกอิเล็กทรอนิกส์ (Electronic Nose: e-nose) เป็นเครื่องมือที่ประกอบด้วยอาร์เรย์เซ็นเซอร์ทางเคมี ที่มีความไวตอบสนองต่อกลิ่นและมีระบบการรู้จำที่สามารถจำกลิ่นที่มีความง่าย และซับซ้อนได้ [8] หลักการทำงานพื้นฐานของจมูกอิเล็กทรอนิกส์ การทำงานเริ่มจากรับสัญญาณด้วยอาเรย์เซ็นเซอร์ เข้ามาในรูปแบบสัญญาณอะนาล็อก ทำการแปลงสัญญาณอะนาล็อกเป็นดิจิตอล และเข้าสู่การประมวลผลสัญญาณ โดยอาศัยการรู้จำในการวิเคราะห์รูปแบบของกลิ่น (Fingerprint) เพื่อระบุว่าเป็นกลิ่นใด

เซ็นเซอร์ที่ใช้ในการทดลองเป็นแบบ Metal Oxides ซึ่ง แต่ละชนิดจะมีความไวต่อก๊าซแตกต่างกันดังนี้ MQ-2 เป็นเซ็นเซอร์ที่ไวต่อก๊าซไวไฟและควัน สามารถตรวจจับความเข้มข้นของก๊าซที่ติดไฟในอากาศได้ ไวต่อก๊าซมีเทน ควันไฟ และควันบุหรี่ เซ็นเซอร์ MQ-3 ไวต่อแอลกอฮอล์ (Alcohol) และเบนซิน (Benzene) เซ็นเซอร์ MQ4 ไวต่อก๊าซมีเทน (Methane) หรือ CNG เซ็นเซอร์ MQ-5 LPG ไวต่อโพรเพน (Propane) เซ็นเซอร์ MQ-7 สำหรับตรวจจับก๊าซคาร์บอนมอนอกไซด์ (Carbon Monoxide) เซ็นเซอร์ MQ-6 ไวต่อแอลพีจี (LPG) และบิวเธน (Butane) MQ-8 ไวต่อไฮโดรเจน (Hydrogen) เซ็นเซอร์ MQ-135 สำหรับตรวจวัดคุณภาพของอากาศ ซึ่งใช้การตรวจจับก๊าซที่หลากหลายในอากาศ ประกอบด้วยแอมโมเนีย (Ammonia: NH3) ไนโตรเจนออกไซด์ (Nitrogen Oxide), แอลกอฮอล์ เบนซิน ควัน และคาร์บอนไดออกไซด์ (Carbon Dioxide) ซึ่งเหมาะกับการใช้ในออฟฟิศหรือโรงงาน [9] [10]

2.2 การลดมิติ

การลดมิติ (Dimensionality Reduction) เป็นสิ่งที่ควรคำนึงถึงในขั้นตอนการเตรียมข้อมูล เพื่อลดปัญหามิติของข้อมูล (Curse of Dimension) ทำได้โดยนำสัญญาณที่ได้จากเซ็นเซอร์มาสกัดหรือเลือกคุณลักษณะ เพื่อความถูกต้องในการจำแนกหรือจัดกลุ่มในขั้นตอนสุดท้าย อีกทั้งลดการใช้ทรัพยากรต่างๆ เช่น ระยะเวลาในการประมวลผล พื้นที่ในหน่วยความจำ รวมถึงการกำจัดมิติที่ไม่เกี่ยวข้องหรือข้อมูลรบกวนอีกด้วย โดยทั่วไปการลดมิติข้อมูลมี 2 แนวทาง ได้แก่ การสกัดคุณลักษณะและการคัดเลือกคุณลักษณะ

การสกัดคุณลักษณะ (Feature Extraction) มีวัตถุประสงค์ เพื่อให้ได้คุณลักษณะ (Feature) จากการเปลี่ยน (Transformation) ด้วยสมการทางคณิตศาสตร์ ซึ่งการเปลี่ยนนี้จะลดมิติของอินพุตและเก็บข้อมูลที่เกี่ยวข้องกับเป้าหมายไว้ทั้งหมด ไม่จำเป็นต้องทำการเลือกคุณลักษณะเพื่อเป็นตัวแทน การสกัดคุณลักษณะมีหลายวิธีการวิธีที่ใช้กันแพร่หลายได้แก่

วิธี PCA (Principal Component Analysis) และ LDA (Linear Discriminant Analysis), FFT (Fast Fourier Transform, DWT (Discrete Wavelet Transform) ซึ่งเป็นวิธีการแบบ Direct Use of Subspace Projection Methods [9] ข้อดีของ PCA คือ การลดมิติ และหลีกเลี่ยงการ Over Fitting โดยหลักการของ PCA จะใช้มิติที่มีความแปรปรวนและแปลงให้อยู่ใน 2-3 มิติ เพื่อประมวลผลโดยไม่สนใจข้อมูลการจำแนก

การคัดเลือกคุณลักษณะ (Feature Selection) เป็นการคัดเลือกคุณลักษณะที่มีความเกี่ยวข้องกับคลาสเป้าหมาย เพื่อนำไปจำแนกหรือจัดกลุ่มในขั้นตอนสุดท้าย โดยเลือกกลุ่มของคุณลักษณะ (Subset of Features) จากข้อมูลต้นฉบับ (Original Data) ช่วยลดความซ้ำซ้อน (Data Reduction) หรือคุณลักษณะที่ไม่เกี่ยวข้อง [10-12] อีกทั้งยังส่งผลต่อความถูกต้องในการพยากรณ์และการจัดกลุ่มอีกด้วย [13]

2.3 การรู้จำรูปแบบ

เทคโนโลยีที่สำคัญอีกประการ คือ เทคโนโลยีการรู้จำรูปแบบ (Pattern Recognition) ที่สามารถนำข้อมูลสัญญาณจากเซ็นเซอร์เข้ามา และตัดสินใจว่าเป็นกลิ่นอะไร จากการศึกษาพบว่า มีการวิเคราะห์องค์ประกอบ โดยใช้วิธีทางสถิติ PCA และ LDA ในการวิเคราะห์ ทั้งนี้ข้อมูลของกลิ่นแต่ละชนิดย่อมมีลักษณะของข้อมูลแตกต่างกัน อย่างไรก็ตามในงานวิจัย [14] พบว่ามีความถูกต้องในระดับหนึ่งเท่านั้น เนื่องจากข้อมูลที่ได้มีลักษณะแบบไม่เป็นเชิงเส้น (Non-linear) จึงนำ Machine Learning มาใช้ เพื่อเพิ่มความถูกต้องในการจำแนกหรือจัดกลุ่ม และสามารถจัดการกับข้อมูลที่ไม่เป็นเชิงเส้นด้วยอัลกอริทึมต่าง ๆ เช่น Fuzzy ARTMAP, Neural Network, SVM (Support Vector Machine) และ Clustering Algorithms เป็นต้น

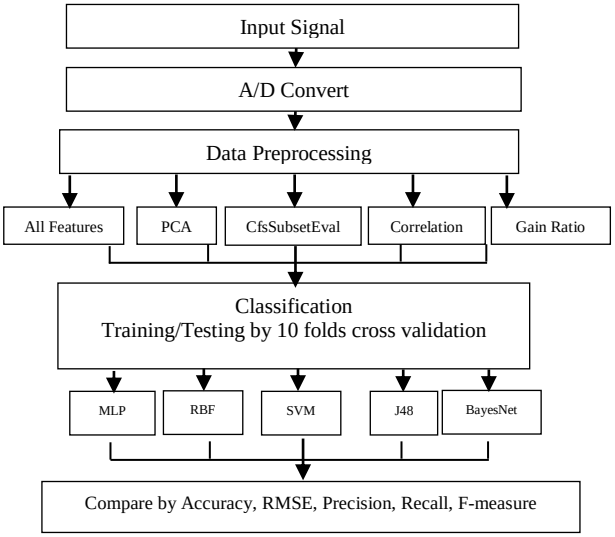
2.4 งานวิจัยที่เกี่ยวข้อง

มีการพัฒนาจมูกอิเล็กทรอนิกส์ โดยใช้เซ็นเซอร์แบบ Metal Oxide Semiconductor (MOS) ทั้งหมด 4 หัว ได้แก่ TGS2620,TGS2602,TGS2201A และ TGS2201B ซึ่งเป็นสิ่งประดิษฐ์ที่มีขนาดเล็ก ต้นทุนต่ำ และสามารถตรวจจับกลิ่นตัวอย่างได้หลากหลายอย่างมีประสิทธิภาพ โดยการตรวจจับก๊าซหลักๆ ได้แก่ เบนซิน โทลูอีน คาร์บอนมอนอกไซด์ และแอมโมเนีย งานวิจัยดังกล่าวเป็นการเลือกพิจารณาหัวเซ็นเซอร์ที่เหมาะสมกับก๊าซในแต่ละรูปแบบ ในการจำแนกรูปแบบของกลิ่นที่ได้ อัลกอริทึมที่ใช้ในการประมวลผลวิเคราะห์รูปแบบของกลิ่นได้แก่ 1) Kernel Principle Component Analysis (KPCA) ซึ่งเป็นการดึงเอาค่าคุณลักษณะของข้อมูลออกมาในรูปแบบ Hybrid Feature Extraction โดยอยู่บนพื้นฐานของ Kernel Transformation และ PCA 2) Liner Discriminant Analysis (LDA) ซึ่งเป็นเทคนิคการจำแนก โดยมีชุดข้อมูลสำหรับฝึกสอน และข้อมูลสำหรับทดสอบเปรียบเทียบกัน ผลที่ได้จากงานวิจัยนี้คือ กลิ่นแต่ละรูปแบบไม่จำเป็นต้องใช้หัวเซ็นเซอร์ครบทุกตัว แต่อาจจะใช้เพียงเซ็นเซอร์แค่บางตัวก็สามารถแยกแยะกลิ่นในแต่ละประเภทได้ [15]

การจำแนกและตรวจสอบน้ำหอมด้วยระบบดมกลิ่นอิเล็กทรอนิกส์ เป็นการจำแนกกลิ่นของน้ำหอมแท้และน้ำหอมเลียนแบบโดยใช้เครื่องมือสำเร็จรูปและซอฟต์แวร์ Nose Pattern Editor ใช้ในการวิเคราะห์ โดยประกอบด้วย 2 อัลกอริทึมในการวิเคราะห์ 1) Classical Feature Selection (R/R0) โดยที่ R คือค่าความต้านทานการระเหยของสารประกอบในน้ำหอมที่ปล่อยออกมา และ R_0 คือค่าความต้านทานเริ่มต้นของเซ็นเซอร์ในอากาศ 2) PCA คือการวิเคราะห์องค์ประกอบจากข้อมูลหลายตัวแปรโดยพิจารณาความแปรปรวนของข้อมูล และทำการลดมิติลง ผลที่ได้จากการทดลองแสดงให้เห็นว่าค่าของเอทิลแอลกอฮอล์ในน้ำหอมมีผลกับการจำแนกประเภทของน้ำหอม [6]

3. วิธีการดำเนินการวิจัย

จมูกอิเล็กทรอนิกส์ที่พัฒนาขึ้นประกอบด้วยแผงวงจรอิเล็กทรอนิกส์ ชุดอาเรย์เซ็นเซอร์ (MQ-2, MQ-3, MQ-4, MQ-7, MQ-8, MQ-135) ขวดแก้วความจุ 850 ml สำหรับวางตัวอย่างน้ำหอม แผงวงจรอิเล็กทรอนิกส์จะเชื่อมต่อกับคอมพิวเตอร์โน๊ตบุคผ่านพอร์ต RS232 ด้วยโปรแกรม LabVIEW สัญญาณอะนาล็อกที่ได้จากการวัดการตอบสนองของก๊าซชนิดต่างๆ จากสารประกอบของน้ำหอมแต่ละกลิ่นจะถูกแปลงเป็นดิจิตอล บันทึกค่าเป็นตัวเลข และส่งออกเป็นไฟล์ชนิด .csv สำหรับในขั้นการเตรียมข้อมูล (Data Preprocessing) จะกำจัดข้อมูลรบกวน (Noise) และ จัดข้อมูลให้อยู่ในรูปแบบที่เหมาะสม เพื่อนำไปวิเคราะห์ จำแนกและประเมินผล โดยขั้นตอนการทดลองแสดงดังภาพที่ 1



ภาพที่ 1: ขั้นตอนการทดลอง

3.1 การเก็บรวบรวมข้อมูลจากตัวอย่าง

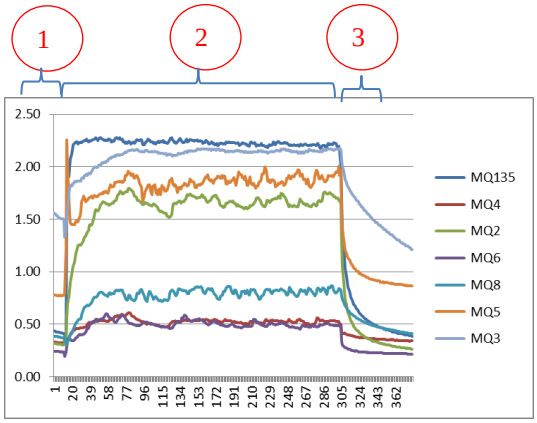
ตัวอย่างน้ำหอมที่นำมาใช้ในการทดสอบประกอบด้วยน้ำหอมแท้และน้ำหอมเลียนแบบจำนวน 3 ยี่ห้อ ได้แก่ Chanel Chance Eau Tendre Green, Chloe', Britney Spears เก็บข้อมูลจากตัวอย่างโดยฉีดลงบนแผ่นสำลี และเก็บในขวดแก้วที่เตรียมไว้ ใช้อาเรย์เซ็นเซอร์รับค่าการตอบสนองที่มีต่อก๊าซชนิดต่าง ๆ โดยขั้นตอนการรวบรวมข้อมูลจากเซ็นเซอร์แบ่งออกเป็น 3 ช่วง ได้แก่

1) ก่อนวัดการตอบสนอง

2) ระหว่างเซ็นเซอร์ตอบสนองต่อก๊าซของตัวอย่างน้ำหอม

3) หลังจากเอาเซ็นเซอร์ออกจากตัวอย่างน้ำหอม

โดยกำหนดเวลาในแต่ละขั้นตอนเป็น 20 วินาที 300 วินาที 60 วินาที ตามลำดับ และเริ่มวัดค่าสำหรับตัวอย่างถัดไปจนครบทุกตัวอย่าง ดังนั้นแต่ละตัวอย่างจะใช้เวลาประมาณ 420 วินาที หรือ 7 นาที ภาพสัญญาณที่ได้จากเซ็นเซอร์ดังภาพที่ 2



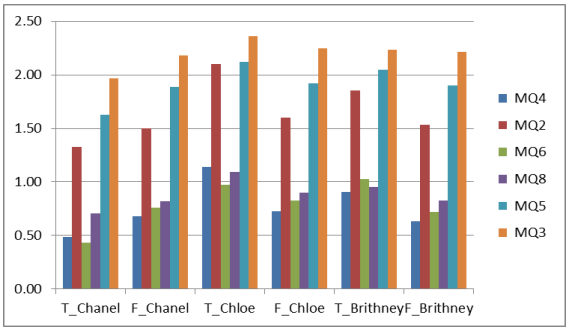
ภาพที่ 2: สัญญาณที่ได้จากการตอบสนองของเซ็นเซอร์

3.2 การเตรียมข้อมูล

เตรียมข้อมูลสัญญาณที่ได้จากกลิ่นน้ำหอมทั้ง 6 ชนิด กำจัดข้อมูลรบกวนออก โดยข้อมูลของน้ำหอมที่ได้จาก 7 เซ็นเซอร์จะประกอบด้วย 7 แอตทริบิวต์ จำนวน 380 เรคอร์ด จากนั้นรวมข้อมูลของน้ำหอมทั้ง 6 ชนิด เข้าด้วยกัน ดังนั้นชุดข้อมูลจะมีขนาด 7 แอตทริบิวต์ จำนวน 2,280 เรคอร์ด เพื่อใช้เป็นข้อมูลสำหรับสอนและทดสอบ

3.3 การวิเคราะห์สัญญาณเซ็นเซอร์

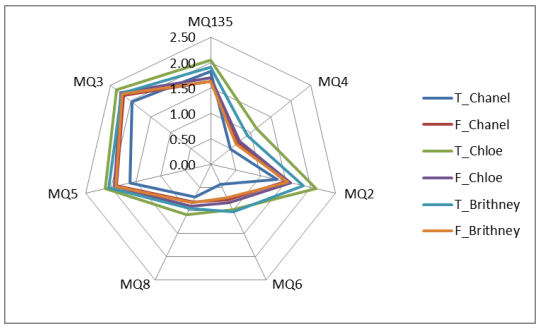
ทำการวิเคราะห์ข้อมูลเบื้องต้น โดยใช้ค่าเฉลี่ยของการตอบสนองของแต่ละเซ็นเซอร์ ที่มีต่อสารประกอบของน้ำหอมแต่ละกลิ่นทั้ง 6 ชนิด ดังภาพที่ 3



ภาพที่ 3: การตอบสนองของเซ็นเซอร์ต่อน้ำหอมทั้ง 6 ชนิด

การวิเคราะห์ลักษณะของข้อมูล เพื่อพิจารณาในเบื้องต้นว่าข้อมูลมีความเป็นเชิงเส้นหรือไม่เป็นเชิงเส้น มากน้อยเพียงใด ด้วยการพล็อตแผนภาพเรดาร์ โดยใช้ค่าเฉลี่ยที่ได้จากการตอบสนองของแต่ละเซ็นเซอร์ที่มีต่อสารประกอบของน้ำหอมแต่ละชนิด ดังภาพที่ 4 ทั้งนี้รูปแบบของการตอบสนองของเซ็นเซอร์ ที่มีการรวมเซ็นเซอร์แต่ละประเภทเข้าด้วยกัน

ชุดเซ็นเซอร์ที่เหมาะสมที่ใช้ในอาร์เรย์ ควรสร้างสัญญาณที่ไม่ขึ้นต่อกัน ซึ่งในเบื้องต้นพบว่า ข้อมูลมีลักษณะที่ไม่เป็นเชิงเส้น มีรูปแบบที่คล้ายกัน มีการซ้อนทับของจุดต่าง ๆ ยากต่อการจำแนก ดังนั้นจึงต้องใช้วิธีการแบบไม่เป็นเชิงเส้นในการจำแนก เพื่อให้การรู้จำมีประสิทธิภาพ



ภาพที่ 4: ค่าเฉลี่ยการตอบสนองของแต่ละเซ็นเซอร์

3.4 ขั้นตอนการทดลอง

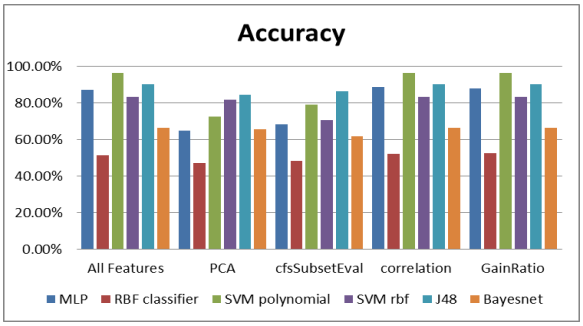
งานวิจัยนี้มีขั้นตอนการทดลอง โดยเริ่มจากการทำความสะอาดข้อมูล กำจัดข้อมูลรบกวน และทำการทดลองโดยชุดข้อมูลที่มีคุณลักษณะ 3 รูปแบบ ได้แก่ 1) ชุดข้อมูลที่ประกอบด้วยทุกคุณลักษณะ 2) ชุดข้อมูลที่ผ่านการลดมิติด้วย PCA และ 3) ชุดข้อมูลที่ผ่านการเลือกคุณลักษณะด้วยวิธี CfsSubsetEval, Correlation และ Gain Ratio ทำการจำแนกด้วยวิธี Multilayer Perceptron, RBF, SVM, J48, และ BayesNet ทำการสอนและทดสอบแบบ 10 Folds Cross Validation จากนั้นนำผลที่ได้มาเปรียบเทียบประสิทธิภาพ

4. ผลการทดลอง

ผลการทดลองจำแนกกลิ่นน้ำหอมแท้และน้ำหอมเลียนแบบอย่างละ 3 ชนิด ได้แก่ Channel , Chloe', Britney Spears โดยใช้เทคนิคเหมืองข้อมูล ด้วยวิธีการต่างๆ ให้ผลความถูกต้องดังตารางที่ 1

ตารางที่ 1: ผลการทดลอง

Method	Correct	Incorrect	RMSE
MLP	87.32%	12.68%	0.1785
RBF classifier	51.54%	48.46%	0.3133
SVM (Polynomial)	**96.49%**	3.51%	0.1081
SVM (RBF)	83.46%	16.54%	0.2348
J48	90.35%	9.65%	0.1684
Bayesnet	66.36%	33.64%	0.2939
MLP + PCA	64.74%	35.26%	0.2744
RBF + PCA	47.19%	52.81%	0.3245
SVM (Polynomial) + PCA	72.63%	27.37%	0.302
SVM (RBF)+ PCA	81.80%	18.20%	0.2463
J48 + PCA	84.34%	15.66%	0.2
BayesNet + PCA	65.53%	34.47%	0.2743
MLP + CfsSubsetEval	68.51%	31.49%	0.2484
RBF + CfsSubsetEval	48.29%	51.71%	0.3103
SVM (Polynomial)+ CfsSubsetEval	78.99%	21.01%	0.2646
SVM (RBF)+ CfsSubsetEval	70.57%	29.43%	0.3132
J48 + CfsSubsetEval	86.32%	13.68%	0.1933
BayesNet + CfsSubsetEval	61.84%	38.16%	0.2859
MLP + Correlation	88.60%	11.40%	0.1679
RBF + Correlation	52.24%	47.76%	0.3126
SVM Polynomial + Correlation	**96.49%**	3.51%	0.1081
SVM RBF + Correlation	83.46%	16.54%	0.2348
J48 + Correlation	90.44%	9.56%	0.1679
BayesNet + Correlation	66.36%	33.64%	0.2939
MLP + Gain Ratio	87.76%	12.24%	0.1806
RBF + Gain Ratio	52.59%	47.41%	0.3107
SVM (Polynomial) + Gain Ratio	**96.49%**	3.51%	0.1081
SVM (RBF) + Gain Ratio	83.46%	16.54%	0.2348
J48 + Gain Ratio	90.44%	9.56%	0.1681
BayesNet + Gain Ratio	66.36%	33.64%	0.2939



ภาพที่ 5: กราฟเปรียบเทียบผลความถูกต้อง

จากภาพที่ 5 เมื่อเปรียบเทียบผลที่ได้จากการทดลอง พบว่าโดยส่วนใหญ่วิธีที่ให้ค่าความถูกต้องสูงคือ SVM (Kernel Polynomial) และเมื่อนำการลดมิติข้อมูลด้วย PCA มาใช้ร่วมทำให้ค่าความถูกต้องลดลงทุกวิธี ซึ่ง PCA อาจไม่เหมาะที่จะนำมาใช้กับชุดข้อมูลนี้ และเมื่อนำการเลือกคุณลักษณะด้วย CfsSubsetEval ซึ่งเป็นการเลือกชุดของคุณลักษณะที่ดีที่สุด คือ MQ4, MQ2, MQ6, MQ3 มาร่วมกับการจำแนกด้วยวิธีการต่างๆนั้น ทำให้ค่าความถูกต้องสูงขึ้นกว่าวิธี PCA และเมื่อใช้ Correlation ซึ่งเป็นการหาค่าสหสัมพันธ์ของคุณลักษณะและจัดเรียงจากคุณลักษณะที่มีค่าสหสัมพันธ์จากมากไปน้อย คือ MQ4, MQ6, MQ8, MQ2, MQ3, MQ5, และ MQ135 พบว่าเมื่อนำ Correlation ร่วมกับ SVM จะได้ค่าความถูกต้องที่ 96.49% เท่ากันกับวิธี Gain Ratio ที่ได้ค่าความถูกต้องที่ 96.49% เช่นกัน โดยการเรียงลำดับตามค่า Gain คือ MQ2, MQ4, MQ6, MQ3, MQ5, MQ135, และ MQ8

5. สรุปและอภิปรายผล

การจำแนกน้ำหอมโดยใช้อาเรย์เซ็นเซอร์ ซึ่งประกอบด้วย 7 เซ็นเซอร์ เพื่อตรวจจับก๊าซของสารประกอบของกลิ่นน้ำหอม พิจารณาในเบื้องต้นพบว่าข้อมูลมีลักษณะแบบไม่เป็นเชิงเส้น ดังนั้นเพื่อความถูกต้องมากยิ่งขึ้น ควรเลือกเทคนิคที่เหมาะกับลักษณะของข้อมูลแบบไม่เป็นเชิงเส้น จากการทดลองสรุปได้ว่าวิธีการ SVM Kernel Type Polynomial ให้ค่าความถูกต้องค่อนข้างสูง และเมื่อใช้วิธี SVM ควบคู่กับการลดมิติของข้อมูลด้วยวิธี PCA ส่งผลให้ความถูกต้องของการจำแนกลดลง เนื่องจาก PCA พิจารณาแค่ค่าความแปรปรวน โดยไม่สนใจข้อมูลการจำแนก ในขณะเดียวกัน การคัดเลือกคุณลักษณะที่เกี่ยวข้องก่อนการนำไปประมวลผล จะส่งผลให้ความถูกต้องเพิ่มมากขึ้นในระดับหนึ่ง จากผลการทดลอง จะเห็นได้ว่าควรใช้เซ็นเซอร์ทุกตัว เนื่องจากเซ็นเซอร์ทั้งหมดส่งผลต่อการจำแนก ทำให้ค่าความถูกต้องมากที่สุด และเพื่อให้โมเดลสามารถใช้งานได้อย่างมีประสิทธิภาพอย่างต่อเนื่อง ควรพัฒนาอัลกอริทึมให้สามารถเรียนรู้ข้อมูลใหม่ ๆ ที่จะเข้ามาได้อย่างชาญฉลาด

เอกสารอ้างอิง

[1] Huang, L., et al., "Nondestructive measurement of total volatile basic nitrogen (TVB-N) in pork meat by integrating near infrared spectroscopy, computer vision and electronic nose techniques". *Food Chemistry*, 2014. 145(0): p. 228-236.

[2] Pan, L., et al., "Early detection and classification of pathogenic fungal disease in post-harvest strawberry fruit by electronic nose and gas chromatography–mass spectrometry." *Food Research International*, 2014. 62(0): p. 162-168.

[3] Brudzewski, K., S. Osowski, and A. Dwulit, "Recognition of Coffee Using Differential Electronic Nose". *Instrumentation and Measurement, IEEE Transactions on*, 2012. 61(6): p. 1803-1810.

[4] Westenbrink, E., et al., "Development and application of a new electronic nose instrument for the detection of colorectal cancer." *Biosensors and Bioelectronics*, 2015. 67(0): p. 733-738.

[5] Schmekel, B., F. Winquist, and A. Vikström, "Analysis of breath samples for lung cancer survival." *Analytica Chimica Acta*, 2014. 840(0): p. 82-86.

[6] Cano, M., et al., "Rapid discrimination and counterfeit detection of perfumes by an electronic olfactory system." *Sensors and Actuators B: Chemical*, 2011. 156(1): p. 319-324.

[7] Loutfi, A., et al., "Electronic noses for food quality: A review." *Journal of Food Engineering*, 2015. 144(0): p. 103-111.

[8] Gardner, J.W. and P.N. Bartlett, "A brief history of electronic noses." *Sensors and Actuators B: Chemical*, 1994. 18(1–3): p. 210-211.

[9] Marco, S. and A. Gutierrez-Galvez, "Signal and Data Processing for Machine Olfaction and Chemical Sensing: A Review." *Sensors Journal, IEEE*, 2012. 12(11): p. 3189-3214.

[10] Huang, Y., et al. "A two-phase feature selection method using both filter and wrapper." in *Systems, Man, and Cybernetics, 1999*. IEEE SMC '99 Conference Proceedings. 1999 IEEE International Conference on. 1999.

[11] Guangrong, L., et al. "A novel unsupervised feature selection method for bioinformatics data sets through feature clustering." in *Granular Computing*, 2008. GrC 2008. IEEE International Conference on. 2008.

[12] Adiguzel, Y. and H. Kulah, "Breath sensors for lung cancer diagnosis." *Biosensors and Bioelectronics*, 2015. 65(0): p. 121-138.

[13] Bharti, K.K. and P.K. Singh, "Hybrid dimension reduction by integrating feature selection with feature extraction method for text clustering." *Expert Systems with Applications*, 2015. 42(6): p. 3105-3114.

[14] Haddi, Z., et al., "E-Nose and e-Tongue combination for improved recognition of fruit juice samples." *Food Chemistry*, 2014. 150(0): p. 246-253.

[15] Zhang, L., F. Tian, and G. Pei, "A novel sensor selection using pattern recognition in electronic nose." *Measurement*, 2014. 54(0): p. 31-39.

การจัดกลุ่มและวิเคราะห์ตามแนวทฤษฎีพหุปัญญาด้วยเทคนิคดาต้าไมนิ่ง
Cluster and Analysis Based on Multiple Intelligences Theory Using Data Mining

ณรงค์เดช รัตนานนท์เสถียร (Narongdech Rattananonsathien)[1] , อภิชาติ บัวตูม (Apichat Buatoom)[2]
และมนต์สวรรค์ พลอยมุกดา (Monsawan Ploymookda)[3]
[1]สาขาวิชาคอมพิวเตอร์ศึกษา คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏกาญจนบุรี
[2]สาขาวิชาเทคโนโลยีการศึกษาและคอมพิวเตอร์ศึกษา คณะครุศาสตร์ มหาวิทยาลัยราชภัฏกาญจนบุรี
[3]สาขาวิชาเทคโนโลยีสารสนเทศ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏกาญจนบุรี
narongdech@kru.ac.th, apichabuatoomt@gmail.com, monsawan@kru.ac.th

บทคัดย่อ

การวิจัยนี้มีจุดประสงค์เพื่อจัดกลุ่มนักศึกษาและวิเคราะห์คุณลักษณะของนักศึกษาแต่ละกลุ่มตามแนวทฤษฎีพหุปัญญาด้วยเทคนิคดาต้าไมนิ่ง โดยกลุ่มตัวอย่างคือนักศึกษาสาขาเทคโนโลยีการศึกษาและคอมพิวเตอร์ศึกษา ชั้นปีที่ 1-4 จำนวน 205 คน คณะครุศาสตร์ มหาวิทยาลัยราชภัฏกาญจนบุรี เครื่องมือที่ใช้ในการวิจัยได้แก่ แบบสอบถามด้านพหุปัญญาสำหรับทดสอบความความถนัดพหุปัญญาของ โฮเวิร์ด การ์ดเนอร์ และเทคนิคการจัดกลุ่มวิธี SimpleK-means แบบ Use Training Set และวิเคราะห์คุณลักษณะด้วยดาต้าไมนิ่ง ผลวิจัยพบว่าจำนวนกลุ่มของนักศึกษาที่จัดตามความถนัด ตามทฤษฎีพหุปัญญาเท่ากับ 2 กลุ่ม โดยกลุ่มที่ 1 ร้อยละ 29 และกลุ่มที่ 2 ร้อยละ 71

คำสำคัญ : ทฤษฎีพหุปัญญา การจัดกลุ่ม การเรียนรู้ คุณลักษณะตามแนวพหุปัญญา ดาต้าไมนิ่ง

Abstract

The present study aims at clustering and analyzing characteristics of students based on multiple intelligence domains using data mining technique. A group of 205 first-year to fourth-year students majoring in Information Technology and Computer Education, Faculty of Education, Kanchanaburi Rajabhat University was selected as the sample of the study. Multiple Intelligences Test based on Howard Gardner's MI Model assessing student aptitudes and simple K-means clustering technique Use Training Set were adopted as research instruments, and the student characteristics were analyzed using data mining technique. According to the results, there were 2 groups of students categorized by student aptitudes and multiple intelligences theory. There were 2 groups of participants as follows: first group is 29% and second group is 71%.

Keyword: Multiple Intelligences Theory, Learning Cluster, Characteristics of Multiple Intelligences, data mining.

1.บทนำ

การศึกษายุค Thailand 4.0 เป็นยุคในการเตรียมความพร้อมความรู้ให้กับคน เพื่อให้เป็นผู้ที่รักจะเรียน เน้นการสร้างคนให้มีทักษะในศตวรรษที่ 21 โดยเน้นทักษะในการคิดวิเคราะห์เป็นหลัก ดังนั้นการจัดการเรียนการสอนตามแนวปฏิรูปการเรียนรู้ที่เน้นผู้เรียนเป็นสำคัญ สามารถนำการเรียนรู้ตามทฤษฎีพหุปัญญาที่ได้รับความนิยม ซึ่งเป็นทฤษฎีที่มีแนวคิดของการจัดกิจกรรมการเรียนรู้สนองความแตกต่างระหว่างบุคคล เป็นการเรียนรู้จัดการเรียนรู้ได้ตรงตามความสามารถของผู้เรียนเพื่อเพิ่มทักษะในการคิดวิเคราะห์การเรียนรู้อย่างมีประสิทธิภาพ ตลอดจนส่งเสริมศักยภาพที่มีในตัวนักศึกษา [1]

เหตุผลดังกล่าวข้างต้นทำให้ผู้วิจัยมีความตระหนักถึงความจำเป็นต่อการส่งเสริมความสามารถการเรียนรู้ให้เกิดกับผู้เรียน

ได้อย่างอย่างเต็มประสิทธิภาพ ผู้วิจัยจึงได้นำแนวการจัดการเรียนรู้ตามแนวทฤษฎีพหุปัญญาและเทคนิคดาต้าไมนิ่งซึ่งเป็นวิธีที่ได้รับความนิยม โดยวิเคราะห์ข้อมูลของนักศึกษาด้วยเทคนิคดาต้าไมนิ่ง [1] ตามแนวทฤษฎีพหุปัญญา เพื่อนำผลการวิจัยครั้งนี้ไปวิเคราะห์และสร้างกิจกรรมส่งเสริมการเรียนรู้โดยยึดศักยภาพ ความถนัดและสติปัญญาให้ได้รับการพัฒนาทุกๆ ด้านไปพร้อมๆ กัน เพื่อพัฒนาการเรียนรู้ของนักศึกษาให้มีประสิทธิภาพ และผลการเรียนรู้ที่ดียิ่งขึ้นต่อไป

2. วัตถุประสงค์

2.1 เพื่อวิเคราะห์คุณลักษณะของนักศึกษาแต่ละกลุ่มตามแนวทฤษฎีพหุปัญญาด้วยเทคนิคดาต้าไมนิ่ง

2.2 เพื่อจัดกลุ่มนักศึกษาตามความถนัดของนักศึกษาตามแนวทฤษฎีพหุปัญญาด้วยเทคนิคดาต้าไมนิ่ง

3. ขอบเขตการวิจัย

3.1 ทักษะพหุปัญญามี 9 ด้านมีการจัดกลุ่มทางพหุปัญญา (Multiple Intelligence Domains) [3,4] ประกอบด้วย 3 กลุ่มหลัก

กลุ่มที่ 1 วิเคราะห์ (Analytic) ประกอบด้วย ด้านที่ 6 สติปัญญาด้านตรรกะและคณิตศาสตร์ (Logical Intelligence), ด้านที่ 7 สติปัญญาด้านดนตรี/จังหวะ (Musical Intelligence), ด้านที่ 8 สติปัญญาด้านความเข้าใจธรรมชาติ (Naturalistic Intelligence)

กลุ่มที่ 2 กลุ่มพินิจพิจารณา (Introspective) ประกอบด้วย ด้านที่ 2 สติปัญญาด้านจิตวิญญาณและการดำรงอยู่ของชีวิต (Existential intelligence), ด้านที่ 4 สติปัญญาด้านตนเองหรือการเข้าใจตนเอง (Intrapersonal Intelligence), ด้านที่ 9 สติปัญญาด้านมิติสัมพันธ์(Spatial Intelligence)

กลุ่มที่ 3 กลุ่มปฏิสัมพันธ์ (Interactive) ประกอบด้วย ด้านที่ 1 การเคลื่อนไหวปัญญาด้านร่างกายและเคลื่อนไหว (Kinesthetic Intelligence), ด้านที่ 3 สติปัญญาด้านมนุษยสัมพันธ์ (Interpersonal Intelligence), ด้านที่ 5 ปัญญาด้านภาษา (Linguistic Intelligence)

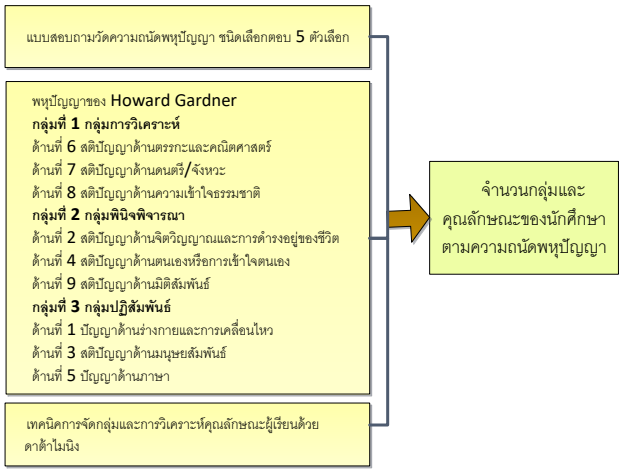
3.2 แบบสอบถามด้านพหุปัญญาสำหรับทดสอบความความถนัดพหุปัญญาของ โฮเวิร์ด การ์ดเนอร์ [5] ใช้เป็นเครื่องมือการวิจัย โดยแบบสอบถามแบ่งเป็น 2 ส่วนคือ ส่วนข้อมูลทั่วไปและส่วนแบบสอบถาม ชนิดเลือกตอบ 5 ตัวเลือก จำนวน 90 ข้อ แบ่งระดับความคิดเห็นเพื่อวัดความสามารถทางพหุปัญญา 5 ระดับ ได้แก่ ระดับ 5, 4, 3, 2 และ 1 หมายถึงนักศึกษามีระดับความคิดเห็นอยู่ในระดับ มากที่สุด, มาก, ปานกลาง, น้อย และน้อยที่สุด ตามลำดับ

3.3 กลุ่มตัวอย่างที่ใช้ในการวิจัย คือ นักศึกษาสาขาเทคโนโลยีการศึกษาและคอมพิวเตอร์ศึกษาชั้นปีที่ 1-4 จำนวน 205 คน คณะครุศาสตร์ มหาวิทยาลัยราชภัฏกาญจนบุรี

3.4 โปรแกรม Orange Canvas เวอร์ชัน 2.7 และโปรแกรม Weka เวอร์ชัน 3.8 เป็นโปรแกรมแบบ OpenSource ที่ใช้งานง่ายและได้รับความนิยมในการนำมาใช้ทางด้านดาต้าไมนิ่ง (data mining) สำหรับงานวิจัยนี้ได้นำโปรแกรมดังกล่าวมาวิเคราะห์คุณลักษณะและจัดกลุ่มตามความถนัดของนักศึกษาตามแนวพหุปัญญา

4. วิธีดำเนินการวิจัย

4.1 กรอบแนวคิดในการวิจัย ผู้วิจัยได้ศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องกับทฤษฎีพหุปัญญา ดาต้าไมนิ่ง แล้วนำมาวิเคราะห์และสังเคราะห์เพื่อกำหนดกรอบแนวคิดในการวิจัย แสดงได้ดังภาพที่ 1



ภาพที่ 1 กรอบแนวคิดในการวิจัย

4.2 ขั้นตอนการดำเนินงานวิจัย ดังนี้

4.2.1 นักศึกษาทดสอบด้วยแบบสอบถามตามทฤษฎีพหุปัญญา ชนิดเลือกตอบ 5 ตัวเลือก จำนวน 90 ข้อ

4.2.2 จัดเตรียมข้อมูล (pre-processing) สำหรับประมวลผลด้วยเทคนิคดาต้าไมนิ่ง

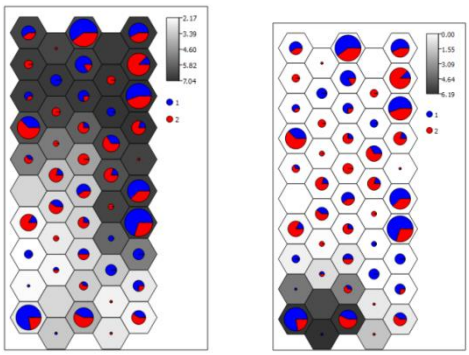
4.2.3 นำข้อมูลที่ผ่านการจัดเตรียมข้อมูลมาทดสอบเทคนิคดาต้าไมนิ่ง เพื่อจัดกลุ่ม)Clustering (ได้ผลลัพธ์เป็นจำนวนกลุ่มนักศึกษาที่มีความถนัดแต่ละด้าน

4.2.4 สรุปผลลัพธ์และวิเคราะห์ข้อมูล เพื่อนำผลของการวิจัยครั้งนี้ไปสร้างรูปแบบการจัดการเรียนรู้ตามหลักพหุปัญญาให้สอดคล้องกับความถนัดในด้านต่างๆ โดยการสร้างเนื้อหากิจกรรม การเรียนรู้ให้กับนักศึกษา เพื่อจะได้เป็นการพัฒนาการเรียนรู้ ของนักเรียน นักศึกษาให้มีประสิทธิภาพ และผลการเรียนรู้ที่ดียิ่งขึ้นต่อไป

5. ผลการวิจัย

ผลการวิจัยสรุปได้ 3 ข้อดังนี้ 1) การวิเคราะห์คุณลักษณะทั่วไปของนักศึกษาที่นำมาศึกษาในงานวิจัยนี้, 2) การวิเคราะห์จำนวนกลุ่มและคุณลักษณะของนักศึกษาตามความถนัดพหุปัญญาเน้นตามกระบวนการใน 3 กลุ่ม 9 ด้าน [6] และ 3) การวิเคราะห์จำนวนกลุ่มและคุณลักษณะของนักศึกษาตามความถนัดพหุปัญญา เพื่อศึกษาว่ากลุ่มตัวอย่างที่สนใจมีความถนัดพหุปัญญาอย่างไร ดังรายละเอียดต่อไปนี้

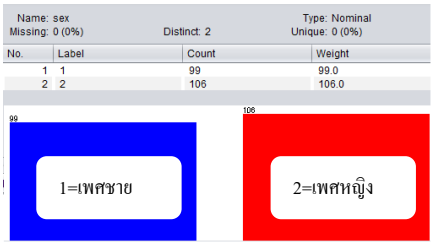
5.1 การวิเคราะห์คุณลักษณะทั่วไปของนักศึกษา ด้วยโปรแกรม Orange Canvas ด้วย Self-Organizing Map (SOM) และโปรแกรม Weka ดังนี้



(ก) รายวิชาที่ทำคะแนนได้สูงอยู่เสมอ (ข) ลักษณะการเรียนรู้ที่ชอบ

ภาพที่ 2 คุณลักษณะด้านการเรียนของนักศึกษา

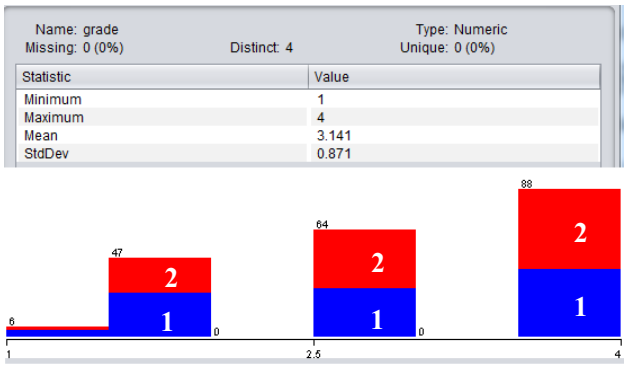
จากภาพที่ 2 (ก) พบว่าสีพื้นเข้มส่วนใหญ่นักศึกษาเป็นเพศหญิงเรียนรายวิชาคอมพิวเตอร์ พลศึกษา สุขศึกษาเป็นวิชาที่ทำคะแนนได้สูงอยู่เสมอ ส่วนสีพื้นสีขาวพบว่านักศึกษาเป็นเพศชายส่วนใหญ่เรียนรายวิชาภาษา เช่น ไทย อังกฤษ ที่ทำคะแนนได้สูงอยู่เสมอ จากภาพที่ 2 (ข) นักศึกษาส่วนใหญ่เป็นเพศหญิงมีลักษณะการเรียนที่ชอบเรียนรู้หลายๆ รูปแบบผสมผสานกัน



(ก) จำนวนเพศนักศึกษา

ภาพที่ 3 ความสัมพันธ์ระหว่างเพศกับระดับผลการศึกษา

จากภาพที่ 3 กำหนดให้หมายเลข 1 คือเพศชาย และหมายเลข 2 คือเพศหญิง และจากภาพที่ 3(ก) พบว่านักศึกษาส่วนใหญ่เป็นเพศหญิง



(ข) จำนวนเพศในแต่ละช่วงระดับผลการศึกษา

ภาพที่ 3 (ต่อ) ความสัมพันธ์ระหว่างเพศกับระดับผลการศึกษา

จากภาพที่ 3 (ข) พบว่าภาพรวมนักศึกษามีระดับการศึกษาเฉลี่ยอยู่ที่ 3.141 และค่าเบี่ยงเบนมาตรฐาน 0.871 และนักศึกษาส่วนใหญ่มีระดับผลการศึกษาที่ระดับ 4 ซึ่งเกรดเฉลี่ยของนักศึกษามีผลต่อกิจกรรมการเรียนรู้ตามหลักพหุปัญญาที่ระดับนัยสำคัญ 0.05 [7]

5.2 การวิเคราะห์จำนวนกลุ่มและคุณลักษณะของนักศึกษาตามความถนัดพหุปัญญาเน้นตามกระบวนการใน 3 กลุ่ม 9 ด้าน

การวิเคราะห์คุณลักษณะนักศึกษาในข้อนี้ใช้เทคนิคการจัดกลุ่ม (Cluster) วิธี SimpleK-means แบบ Use Training Set

ด้วยโปรแกรม Weka ผลการวิเคราะห์ตามพหุปัญญาแบ่งตามกลุ่มปัญญา 3 กลุ่ม สามารถนำไปสร้างรูปแบบการจัดการเรียนรู้สร้างเนื้อหา กิจกรรม การเรียนรู้ให้กับนักศึกษาเพื่อพัฒนาการเรียนรู้ ที่เน้นกระบวนการตามกลุ่มพหุปัญญาที่แตกต่างกัน หากมีจุดประสงค์พัฒนาศักยภาพนักศึกษาในด้านที่ต่างกัน สรุปได้ดังนี้

5.2.1 คุณลักษณะนักศึกษากลุ่มที่ 1 การวิเคราะห์

นักศึกษากลุ่มนี้มีคุณลักษณะทั่วไปคือ ชอบการเรียนรู้จากการทัศนศึกษาจากแหล่งความรู้ภายนอก ชอบขอคำแนะนำจากผู้อื่น มีความถนัดในการเขียนและการลงมือปฏิบัติและชอบทำงานร่วมกับผู้อื่น โดยนักศึกษามีความถนัดตามพหุปัญญาแบ่งได้ 2 กลุ่มคือ

กลุ่มที่ 1 คิดเป็นร้อยละ 38 มีคุณลักษณะที่สำคัญดังนี้ด้านที่ 6: เป็นนักศึกษาที่มีเหตุผลในการพูดคุยกับผู้อื่นเสมอ, ด้านที่ 7: ชอบเรียนวิชาดนตรีและใช้เวลาว่างในการเล่นดนตรีฟังเพลง ฮัมเพลงและเป็นนักศึกษาที่สามารถแยกแยะจังหวะเพลงได้ดี ด้านที่ 8: เป็นนักศึกษาที่มีความเข้าใจธรรมชาติสูง เช่น ชอบอ่านหนังสือ นิตยสาร เกี่ยวกับการท่องเที่ยวตามธรรมชาติ และชอบเที่ยวตามธรรมชาติ

กลุ่มที่ 2 คิดเป็นร้อยละ 62 มีคุณลักษณะที่สำคัญดังนี้ด้านที่ 6: เป็นนักศึกษาที่มีเหตุผลในการพูดคุยกับผู้อื่นเสมอ, ด้านที่ 7: เป็นนักศึกษาชอบฟังเพลง ชอบฮัมเพลงทำให้เกิดความสุข, ด้านที่ 8: เป็นนักศึกษาที่ชอบไปเที่ยวตามธรรมชาติ

5.2.2 คุณลักษณะนักศึกษากลุ่มที่ 2 พินิจพิจารณา

นักศึกษากลุ่มนี้มีคุณลักษณะทั่วไปคือ ชอบเรียนวิชาดนตรี เรียนรู้จากแหล่งความรู้ภายนอกชอบขีดเขียนหรือแต่งเติมรูปภาพเป็นผู้ที่สังเกตการเปลี่ยนแปลงดินฟ้าอากาศ ชอบเลี้ยงสัตว์และชอบไปสวนสัตว์โดยนักศึกษามีความถนัดตามพหุปัญญาแบ่งได้ 2 กลุ่มคือ

กลุ่มที่ 1 คิดเป็นร้อยละ 41 มีคุณลักษณะที่สำคัญดังนี้ด้านที่ 2: เป็นผู้ที่รับรู้ถึงการเกิด แก่ เจ็บ ตายว่าเป็นเรื่องธรรมดาโดยเข้าใจหลักของศาสนาที่ตนนับถือและสามารถนำปรัชญามาใช้ในการดำเนินชีวิตได้ ด้านที่ 4: เป็นผู้ที่ชอบอิสระและมีความเป็นตัวของตัวเอง และเข้าใจตนเองว่ามีความเก่งหรือไม่เก่งในด้านใด ชอบมีงานอดิเรกที่ทำให้ตนเองอยู่ตามลำพัง ด้านที่ 9: เป็นผู้ที่ชอบเกี่ยวกับงานด้านศิลปะได้แก่ การวาดรูปและขีดเขียน ออกแบบสิ่งของ

กลุ่มที่ 2 คิดเป็นร้อยละ 59 มีคุณลักษณะที่สำคัญดังนี้ด้านที่ 2: เป็นผู้ที่รับรู้ถึงการเกิด แก่ เจ็บ ตายว่าเป็นเรื่องธรรมดาและสามารถนำปรัชญามาใช้ในการดำเนินชีวิตได้ ด้านที่ 4: เป็นผู้ที่ชอบอิสระและมีความเป็นตัวของตัวเอง ด้านที่ 9: เป็นผู้ที่ชอบเกี่ยวกับงานด้านศิลปะเช่นเดียวกัน

5.2.3 คุณลักษณะนักศึกษากลุ่มที่ 3 ปฏิสัมพันธ์

นักศึกษากลุ่มนี้มีคุณลักษณะทั่วไปคือมีความถนัดด้านการเขียนมากที่สุด ทำให้ชอบขีดเขียนและตกแต่งรูปภาพ ชอบทัศนศึกษาจากแหล่งเรียนรู้ภายนอก วิชาดนตรีเป็นวิชาที่ทำคะแนนได้สูงสุดเสมอและเป็นผู้ที่ชอบช่วยเหลือผู้อื่นโดยนักศึกษามีความถนัดตามพหุปัญญาแบ่งได้ 2 กลุ่มคือ

กลุ่มที่ 1 คิดเป็นร้อยละ42 มีคุณลักษณะที่สำคัญดังนี้ด้านที่ 1: การเคลื่อนไหวร่างกายชอบการขี่จักรยาน ด้านที่ 3: ชอบทำงานและไปเที่ยวเป็นกลุ่ม ด้านที่ 5: เป็นผู้ที่เขียนอธิบายสิ่งต่างๆ ได้ดี

กลุ่มที่ 2 คิดเป็นร้อยละ 58 มีคุณลักษณะที่สำคัญดังนี้ด้านที่ 1: เป็นผู้ที่ชอบหาอะไรทำยามว่าง เช่น การเต้นประกอบเพลง การขี่จักรยาน ด้านที่ 3: มีคุณลักษณะเช่นเดียวกันกับในกลุ่มที่ 1 ด้านที่ 5:เป็นนักศึกษาที่เรียนรู้ทางภาษาได้ดี เช่น การแต่งเนื้อเรื่องจากภาพ การหาความหมายของคำศัพท์ การชอบเรียนวิชาภาษาอังกฤษ เป็นต้น

จากข้อมูลข้างต้นเพื่อวิเคราะห์คุณลักษณะของนักศึกษาทั้งหมดโดยเน้นคุณลักษณะแต่ละกลุ่ม ตามความถนัดและความชอบเรียนรู้ของนักศึกษา ได้แก่ กลุ่มการวิเคราะห์,กลุ่มพินิจพิจารณาและกลุ่มปฏิสัมพันธ์ สามารถสรุปได้ดังตารางที่ 1 เพื่อผู้สอนสามารถเน้นการพัฒนาศักยภาพตามกลุ่ม

ตารางที่ 1 สรุปคุณลักษณะนักศึกษาในแต่ละกลุ่ม

กลุ่มคุณลักษณะ	กลุ่ม	ร้อยละ	ด้านคุณลักษณะที่สำคัญ
การวิเคราะห์	1	38	6-8
	2	62	
พินิจพิจารณา	1	41	2, 4, 9
	2	59	
ปฏิสัมพันธ์	1	42	1, 3, 5
	2	58	

จากตารางพบว่านักศึกษาสาขาเทคโนโลยีการศึกษาและคอมพิวเตอร์ศึกษาชั้นปีที่ 1-4 มีคุณลักษณะกลุ่มการวิเคราะห์สูงสุด คือร้อยละ 62 และคุณลักษณะในกลุ่มอื่นนักศึกษามีความชอบใกล้เคียงกัน

ดังนั้นจากผลการวิเคราะห์ในตารางที่ 1 พบว่าผู้สอนสามารถพัฒนาศักยภาพนักศึกษาโดยเน้นกลุ่มตามคุณลักษณะจะมีกิจกรรมที่ต้องพัฒนาตามด้านคุณลักษณะที่สำคัญ กล่าวคือ หากผู้สอนต้องการพัฒนาศักยภาพผู้เรียนตามกลุ่มที่ 1 การวิเคราะห์ ควรเสริมกิจกรรมที่มีการใช้เหตุผลในการทำงานและมีกิจกรรมที่เรียนรู้นอกห้องเรียน เป็นต้น แต่หากผู้สอนต้องการพัฒนาศักยภาพผู้เรียนตามกลุ่มที่ 2 พินิจพิจารณา ควรเสริมกิจกรรมจากแหล่งเรียนรู้ภายนอก แต่หากผู้สอนต้องการพัฒนาศักยภาพตามกลุ่มที่ 3 ปฏิสัมพันธ์ ควรมีการเสริมกิจกรรมที่เน้นการทำงานเป็นกลุ่มและเป็นกิจกรรมที่เน้นการอธิบายด้วยวิธีการเขียน เป็นต้น

5.3 การวิเคราะห์จำนวนกลุ่มและคุณลักษณะของนักศึกษาตามความถนัดพหุปัญญา

นักศึกษามีคุณลักษณะทั่วไปคือ วิชาที่ทำคะแนนได้สูงสุดอยู่เสมอคือวิชาดนตรีนักศึกษามีความถนัดการเขียนมากกว่าวิธีอื่น เช่น การอ่าน การฟังกิจกรรมที่ชอบมากคือกิจกรรมการเขียนหรือตกแต่งภาพลักษณะการเรียนที่ชอบมากที่สุดคือการทัศนศึกษาจากแหล่งความรู้ภายนอกชอบขอคำแนะนำจากผู้อื่น มากกว่าวิธีอื่น เช่น ชอบเกี่ยวกับตัวเลข และชอบเลี้ยงสัตว์หรือไปสวนสัตว์ มากกว่าวิธีอื่น เช่น การใช้สังคมออนไลน์ การเขียนนวนิยาย โดยนักศึกษามีความถนัดตามพหุปัญญา แบ่งได้ 2 กลุ่ม คือ

กลุ่มที่ 1 คิดเป็นร้อยละ 29 โดยนักศึกษามีระดับความคิดเห็นอยู่ในระดับมากขึ้นไปมีคุณลักษณะดังนี้

กลุ่มพินิจพิจารณาด้านที่ 2: สติปัญญาด้านจิตวิญญาณและการดำรงอยู่ของชีวิต ที่นักศึกษาสามารถนำหลักปรัชญาไปใช้ในการดำเนินชีวิตได้ และด้านที่ 4: สติปัญญาด้านตนเองหรือการเข้าใจตนเอง ที่นักศึกษารับรู้ถึงการเกิด แก่ เจ็บ ตาย ว่าเป็นเรื่องธรรมดา มีความเข้าใจตนเองว่าชอบอิสระและเป็นตัวของตัวเองกลุ่ม 3 กลุ่มปฏิสัมพันธ์ด้านที่ 3: สติปัญญาด้านมนุษยสัมพันธ์นักศึกษาชอบเที่ยว พูดคุยเป็นกลุ่ม

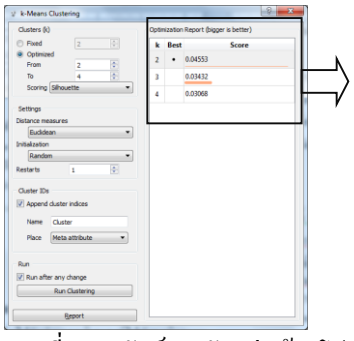
กลุ่มที่ 2 คิดเป็นร้อยละ 71 โดยนักศึกษามีระดับความคิดเห็นอยู่ในระดับมากขึ้นไปมีคุณลักษณะดังนี้

กลุ่มพินิจพิจารณา ด้านที่ 2: สติปัญญาด้านจิตวิญญาณและการดำรงอยู่ของชีวิต ที่นักศึกษาเข้าใจสัจธรรมชีวิตและเข้าใจหลักของศาสนาที่ตนนับถือ ด้วยการรับรู้ถึงการเกิด แก่ เจ็บ ตายว่าเป็นเรื่องธรรมดาโดยนำหลักปรัชญาไปในใช้ในการดำเนินชีวิตได้ และด้านที่ 4: สติปัญญาด้านตนเองหรือการเข้าใจตนเอง ที่นักศึกษาเป็นคนชอบอิสระ มีความเข้าใจตนเองว่าเก่งด้านใดและไม่เก่งด้านใด กลุ่ม 3 กลุ่มปฏิสัมพันธ์ ด้านที่ 1: ปัญญาด้านร่างกายและการเคลื่อนไหว นักศึกษาเรียนรู้ในการขี่จักรยานหรือเล่นสเก็ตเป็นเรื่องง่าย และสามารถทำงานฝีมือได้ดี และด้านที่ 3:สติปัญญาด้านมนุษยสัมพันธ์นักศึกษาชอบทำงาน เล่นเกม เที่ยวเป็นกลุ่ม สามารถเข้ากับผู้อื่นได้ดี

ตารางที่ 2 ผลลัพธ์การจัดกลุ่มในแต่ละวิธี

การจัดกลุ่มวิธี	จำนวนกลุ่ม	sum of squared errors	Time (seconds)
SimpleKMeans	2 กลุ่ม (29%, 71%)	10600.567	0.03
MakeDensityBasedClusterer	2 กลุ่ม (36%, 64%)	10600.567	0.06

ตารางที่ 2 ใช้โปรแกรม weka สำหรับประมวลผลโดยกำหนดใช้โหมด Cluster คือ Use training set และแสดงผลลัพธ์การจัดกลุ่มข้อมูลนักศึกษาด้วยวิธี SimpleKMeans และวิธี MakeDensityBasedClusterer พบว่าทั้งสองวิธีให้ผลลัพธ์การจัดกลุ่มเท่ากันคือ 2 กลุ่มแต่วิธี SimpleKMeans ใช้เวลาในการประมวลผลเร็วกว่า ดังนั้นงานวิจัยนี้ได้เลือกการแบ่งกลุ่มด้วยวิธี SimpleKMeans และได้ทดสอบการกลุ่มด้วยโปรแกรม Orange Canvas ด้วยการจัดกลุ่มวิธี k-means



ภาพที่ 4 ผลลัพธ์การจัดกลุ่มด้วยโปรแกรม Orange Canvas

จากภาพที่ 4 การจัดกลุ่มวิธี k-means พบว่าจำนวนกลุ่มที่เหมาะสมที่สุดคือ 2 กลุ่มด้วยคะแนน 0.04553

6. สรุปอภิปรายผล

ด้วยงานวิจัยต่อไปจักนำผลการวิเคราะห์ไปพัฒนากิจกรรมตามความถนัดและความชอบของนักศึกษาโดยเลือกนักศึกษาสาขาวิชาเทคโนโลยีการศึกษาและคอมพิวเตอร์ศึกษาทุกชั้นปี คณะครุศาสตร์ มหาวิทยาลัยราชภัฏกาญจนบุรีเป็นกลุ่มตัวอย่าง จากนั้นนำไปทดสอบและประเมินผลการเรียนรู้ของนักศึกษาก่อนและหลังมีการพัฒนากิจกรรมดังกล่าว ดังนั้นผู้วิจัยได้เลือกใช้ดาต้าไมนิ่งสำหรับการวิเคราะห์คุณลักษณะของนักศึกษาตามความถนัดและความชอบ เนื่องจากเป็นวิธีที่ได้รับความนิยมและเป็นวิธีที่ไม่ซับซ้อน

โดยผลการวิเคราะห์คุณลักษณะของนักศึกษาสามารถสรุปผลได้ดังนี้ 1) การวิเคราะห์คุณลักษณะทั่วไปของนักศึกษาพบว่านักศึกษาส่วนใหญ่เป็นเพศหญิง และมีระดับการศึกษาเฉลี่ย 3.141 2) การวิเคราะห์จำนวนกลุ่มและคุณลักษณะของนักศึกษาตามความถนัดพหุปัญญาเน้นตามกระบวนการ พบว่าสามารถแบ่งกลุ่มพหุปัญญานักศึกษาได้ 2 กลุ่ม และ 3) การวิเคราะห์จำนวนกลุ่มและคุณลักษณะของนักศึกษาตามความถนัดพหุปัญญาพบว่าแบ่งกลุ่มพหุปัญญาได้ 2 กลุ่ม มีความถนัดในการพินิจพิจารณาและการปฏิสัมพันธ์ โดยกลุ่มที่ 1 ร้อยละ 29 มีระดับความคิดเห็นอยู่ในระดับมากขึ้นไป โดยกิจกรรมที่ควรเป็นกิจกรรมตามหลักปรัชญาการดำเนินชีวิต การทำงานเป็นกลุ่มและกลุ่มที่ 2 ร้อยละ 71 มีระดับความคิดเห็นอยู่ในระดับมากขึ้นไป โดยกิจกรรมควรเป็นกิจกรรมเกี่ยวกับการดำรงชีวิตโดยอาศัยหลักของศาสนาที่ตนนับถือ, เป็นกิจกรรมที่ให้ความอิสระ เน้นการเคลื่อนไหวทางร่างกาย และทำงานเป็นกลุ่ม ดังนั้นจากผลวิเคราะห์ข้างต้นสามารถนำไปสร้างรูปแบบการจัดการเรียนรู้ตามหลักพหุปัญญาให้สอดคล้องกับความถนัดในด้านต่างๆ โดยมีเนื้อหา กิจกรรม การเรียนรู้ให้สอดคล้องกับความถนัดของนักศึกษาต่อไป

เอกสารอ้างอิง

[1] สุมาลี สิกเสน, “การวิเคราะห์ความสามารถทางการเรียนรู้ตามแนวทางพหุปัญญา ของโฮเวิร์ด การ์ดเนอร์ ของนักศึกษากลุ่มโปรแกรมวิชาคอมพิวเตอร์ มหาวิทยาลัยราชภัฏนครปฐม”, *The Eighth National Conference on Computing and Information Technology 2012,* มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระครเหนือ.

[2] เอกสิทธิ์ พัชรวงศ์สักดา, "การวิเคราะห์ข้อมูลด้วยเทคนิคดาต้า ไมน์นิง เบื้องต้น",บริษัท เอเซีย ดิจิตอลการพิมพ์ จำกัด. กรุงเทพ, 2557.

[3] Walter Mckenzie. “Multiple Intelligence Domain”.Avalible Online at http://surfaquarium.com/MI/mi_domains.htm

[4] Walter Mckenzie. “Multiple Intelligence and Instructional Technology”.The Association for Supervision and Curriculum Development conference,2005.

[5] Armstrong T. 1994. **Multiple intelligence** (Online).AvailableURL: http://www.thomasarmstrong.com/multiple_intelligences.htm

[6] Karim Hajhashemi and Wong Bee Eng. “A Validation Study of the Persian Version of McKenzie’s Multiple Intelligences Inventory to Measure Profiles of Pre-University stuent” vol. 18, no. 2, pp. 343-355, 2010.

[7] ปราโมช ตงฉิน, “การวิเคราะห์กิจกรรมการเรียนรู้ ตามแนวทางพหุปัญญาของ นักศึกษาชั้นปีที่ มหาวิทยาลัยราชภัฏเพชรบุรี 1” *The Eighth National Conference on Computing and Information Technology 2012*, มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระครเหนือ.

การประเมินระยะห่างระหว่างยานพาหนะด้วยเซ็นเซอร์บนโทรศัพท์มือถือ
Distance Estimation Through Mobile Sensing

พุทธิพงษ์ เล็กขาว (Puttipong Leakkaw)[1] และสุขสันต์ พาณิชพาพิบูล (Sooksan Panichpapiboom)[2]
คณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
[1]59606002@kmitl.ac.th, [2]sooksan@alumni.cmu.edu

บทคัดย่อ

ระยะห่างระหว่างยานพาหนะบนท้องถนนถือว่าเป็นหนึ่งในข้อมูลการจราจรที่สามารถสะท้อนลักษณะสภาพการจราจรได้เป็นอย่างดี อีกทั้งยังสามารถนำไปประยุกต์ใช้ในการประเมินหาความหนาแน่นของการจราจร ซึ่งนับว่าเป็นหนึ่งในข้อมูลการจราจรที่มีความสำคัญ โดยในปัจจุบันการหาระยะห่างระหว่างยานพาหนะจะใช้ข้อมูลจากวีดีโอหรือรูปภาพที่เก็บจากกล้องวงจรปิดที่ติดตั้งไว้ตามถนน ซึ่งมีความยุ่งยากในการติดตั้ง และขาดความยืดหยุ่นในการใช้เก็บรวบรวมข้อมูลการจราจร อย่างไรก็ตามความก้าวหน้าทางด้านเทคโนโลยีบนโทรศัพท์มือถือ ทำให้การประมวลผลรูปภาพและวีดีโอผ่านสมาร์ทโฟนมีความเป็นไปได้มากยิ่งขึ้น ในบทความนี้เราจะนำเสนอวิธีการในการประเมินระยะห่างระหว่างยานพาหนะจากภาพถ่ายมุมมองด้านหน้าด้วยสมาร์ทโฟน ซึ่งใช้พลังงานและทรัพยากรในการประมวลผลน้อยกว่าการใช้ข้อมูลจากวีดีโอ และมีค่าความผิดพลาดเฉลี่ยประมาณ 1 เมตร ซึ่งนับว่ามีความแม่นยำเพียงพอต่อการนำไปใช้เป็นข้อมูลการจราจรได้

คำสำคัญ: ระยะห่างระหว่างยานพาหนะ เซ็นเซอร์เคลื่อนที่ ข้อมูลการจราจร

Abstract

Distance between vehicles is one of traffic information which can indicate the traffic condition. Obviously, on a congested road, the average distance will be shorter than the free-flow traffic condition. Moreover, it can be used to calculate the traffic density which is an important traffic information. Currently, the distance between vehicles can be calculated from fixed sensors data such as traffic surveillance camera. However, using fixed sensors is not effective in terms of cost, coverage, and convenience. A mobile sensing approach, which vehicles act as mobile sensors and collect the traffic information as they travel, is more appealing. In this paper, we introduce a method for estimating the distance between vehicles with a perspective image captured by a smartphone. It uses lower power and resources than video processing. The accuracy of proposed method is acceptable for a traffic application.

Keyword: Distance Estimation, Mobile Sensing, Traffic Sensing

1. บทนำ

การประเมินระยะห่างระหว่างยานพาหนะนั้นสามารถทำให้เราทราบลักษณะของการจราจรในเบื้องต้นได้ ยกตัวอย่างเช่น ถ้าหากระยะห่างระหว่างยานพาหนะมีมาก แสดงว่าการจราจรมีความคล่องตัว และในทางตรงกันข้ามหากระยะห่างระหว่างยานพาหนะมีน้อย แสดงว่าการจราจรติดขัด อีกทั้งหากเราสามารถเก็บข้อมูลระยะห่างระหว่างยานพาหนะนี้ได้จากยานพาหนะหลายๆคันบนท้องถนน เราจะสามารถนำไปใช้ในการหาความหนาแน่นของการจราจร ซึ่งเป็นข้อมูลการจราจรที่สำคัญอีกชนิดหนึ่งได้อีกด้วย

โดยทั่วไปในการหาระยะห่างระหว่างยานพาหนะนั้นสามารถทำได้โดยการใช้ข้อมูลจากกล้องวีดีโอที่ติดตั้งไว้ตามถนน ซึ่งยุ่งยากในการติดตั้ง และมีราคาสูง อีกทั้งยังขาดความยืดหยุ่นในการจะนำไปใช้เป็นเซ็นเซอร์สำหรับเก็บข้อมูลการจราจรให้ครอบคลุมทุกเส้นทาง ด้วยเหตุนี้ในปัจจุบันจึงมีความพยายามใช้เทคโนโลยีที่สามารถหาได้ง่าย และราคาถูกกว่ามาใช้ในการหาระยะห่างระหว่างยานพาหนะ

ในปัจจุบันสมาร์ทโฟนเข้ามามีบทบาทสำคัญมากขึ้นในงานด้านการจราจร เนื่องจากสมาร์ทโฟนในปัจจุบันมาพร้อมกับเซ็นเซอร์ที่หลากหลาย และมีผู้ใช้งานเป็นจำนวนมาก ดังนั้นเราจึงสามารถเปลี่ยนยานพาหนะไม่ว่าจะเก่าหรือใหม่ให้เป็นเซ็นเซอร์เคลื่อนที่สำหรับเก็บข้อมูลการจราจรได้โดยใช้สมาร์ทโฟน โดยเฉพาะอย่างยิ่งในปัจจุบันที่สมาร์ทโฟนมีการพัฒนาระบบประมวลผล และกล้องถ่ายภาพให้มีประสิทธิภาพมากขึ้น ทำให้เริ่มมีการนำสมาร์ทโฟนมาใช้ในการประมวลผลรูปภาพและวีดีโอ ซึ่งนับว่าเป็นประโยชน์อย่างยิ่งกับงานด้านการจราจร โดยในปัจจุบันเริ่มมีการนำกล้องของสมาร์ทโฟนมาใช้ในการตรวจหาสิ่งกีดขวาง หรือคำนวณระยะห่างระหว่างวัตถุ ซึ่งส่วนใหญ่จะมุ่งเน้นไปที่การประมวลผลข้อมูลจากวีดีโอ ยกตัวอย่างเช่น iOnRoad [1] ที่สามารถประเมินระยะห่างระหว่างยานพาหนะได้เช่นกัน แต่อย่างไรก็ตามยังไม่มีรายงานความแม่นยำของวิธีการดังกล่าว อีกทั้งการประมวลผลวีดีโอบนสมาร์ทโฟนนั้น จะต้องใช้ทรัพยากรในการประมวลผลสูง และบริโภคพลังงานในการทำงานมาก ซึ่งนับว่าเป็นข้อจำกัดที่สำคัญของสมาร์ทโฟน งานวิจัยนี้จึงมีแนวคิดที่จะใช้การประเมินหาระยะห่างระหว่างยานพาหนะจากภาพถ่าย แทนการใช้วีดีโอ

2. งานวิจัยที่เกี่ยวข้อง

ในงานวิจัยด้านระบบการจราจรอัจฉริยะในปัจจุบัน มีการนำสมาร์ทโฟนมาทำหน้าที่เป็นเซ็นเซอร์เคลื่อนที่สำหรับเก็บข้อมูลการจราจรอย่างแพร่หลาย ยกตัวอย่างเช่นในงานวิจัยที่ [2] มีการประยุกต์ใช้เซ็นเซอร์วัดความเร่งบนสมาร์ทโฟนในการประเมินหาความเร็วเฉลี่ยของยานพาหนะ และในงานวิจัยที่ [3] ไมโครโฟนบนโทรศัพท์ถูกใช้ในการประเมินสภาพการจราจรจากเสียงแตรบนท้องถนนในประเทศอินเดีย เนื่องจากผู้ขับขี่ในประเทศอินเดียจะนิยมบีบแตรหากการจราจรหนาแน่น อย่างไรก็ตามในความก้าวหน้าของเทคโนโลยีบนโทรศัพท์มือถือ ทำให้การประมวลผลรูปภาพ และวีดีโอจากกล้องของสมาร์ทโฟน เริ่มถูกนำมาใช้ในงานด้านการจราจรมากขึ้น

การใช้ภาพหรือวีดีโอในงานด้านการจราจรส่วนใหญ่จะมุ่งเน้นไปที่การตรวจหาคุณสมบัติของสภาพแวดล้อมบนท้องถนน เช่นงานวิจัยที่ [4], [5] มีการใช้กล้องในการตรวจจับเส้นแบ่งช่องการจราจร รวมไปถึงวิธีการตรวจหาวัตถุต่างๆที่อยู่ด้านหน้า เช่นยานพาหนะ ในงานวิจัยที่ [6], [7] หรือ สิ่งกีดขวางต่างๆในงานวิจัยที่ [8], [9] โดยมีบางส่วนของงานที่กล่าวมาสามารถประมาณระยะห่างระหว่างกล้องกับวัตถุที่อยู่ด้านหน้าได้ แต่อย่างไรก็ตามงานส่วนใหญ่ใช้การประมวลผลด้วยคอมพิวเตอร์เป็นหลัก ซึ่งมีความยุ่งยากในการติดตั้งอุปกรณ์ จึงไม่สะดวกมากนักต่อการนำไปใช้เป็นเซ็นเซอร์สำหรับเก็บข้อมูลการจราจรระหว่างการเดินทางในชีวิตประจำวัน

อย่างไรก็ตามในปัจจุบันเริ่มมีการศึกษาการใช้งานสมาร์ทโฟนในการประเมินหาระยะห่างระหว่างยานพาหนะมากขึ้น ยกตัวอย่างเช่นในงานที่ [10] เป็นการใช้กล้องจากสมาร์ทโฟนในการตรวจหาสิ่งกีดขวางที่อยู่ด้านหน้า และประเมินระยะห่างจากข้อมูลวีดีโอ แต่ความแม่นยำของวิธีการดังกล่าวยังไม่ได้ถูกประเมินไว้แต่อย่างใด อีกทั้งการประมวลผลวีดีโอนั้นจะใช้ทรัพยากรมาก และมีอัตราการบริโภคพลังงานค่อนข้างสูง จึงเป็นอุปสรรค์สำคัญของสมาร์ทโฟนรุ่นเก่าที่ประสิทธิภาพการทำงานยังไม่สูงมากนัก

3. วิธีการทดลอง

ในบทความนี้เราจะนำเสนอวิธีการประเมินระยะห่างระหว่างยานพาหนะของเรากับยานพาหนะที่อยู่ด้านหน้าโดยใช้ภาพถ่ายจากสมาร์ทโฟนที่ติดตั้งไว้ ซึ่งการใช้ภาพถ่ายในการหาระยะระหว่างวัตถุนั้นจะใช้ทรัพยากร และพลังงานน้อยกว่าการใช้วีดีโอ แต่อย่างไรก็ตามการการหาระยะห่างดังกล่าวก็ทำได้ยากกว่าด้วยเช่นกัน เนื่องจากการใช้ภาพเพียงภาพเดียวในการหาระยะระหว่างวัตถุนั้น จะไม่มีภาพวัตถุ หรือสภาพแวดล้อมก่อนหน้าใดๆให้อ้างอิงเพื่อใช้ประมาณตำแหน่งของวัตถุนั้นๆเลย โดยในส่วนนี้ของบทความจะอธิบายวิธีการที่ใช้ในการประเมินระยะห่างระหว่างยานพาหนะ ภายใต้ทรัพยากรที่มีจำกัดของสมาร์ทโฟน

3.1 ตั้งค่าอุปกรณ์

ในการทดลองนี้เราใช้โทรศัพท์ iPhone 5s สำหรับเก็บรวบรวมข้อมูลรูปภาพ ซึ่งสมาร์ทโฟนดังกล่าวจะถูกติดตั้งไว้บริเวณด้านหลังกระจกหน้ารถในแนวนอน และพัฒนา

โปรแกรมประยุกต์สำหรับถ่ายภาพ และบันทึกองศาความเอียงของกล้องที่ติดตั้งไว้ ด้วยแอ็กเซเรโรมิเตอร์ (Accelerometer) โดยตั้งค่าให้เมื่อติดตั้งสมาร์ทโฟนไว้ในแนวตั้งฉากกับพื้นถนน ค่าความเอียงจะอยู่ที่ 0° หากเงยกล้องขึ้นค่าองศาของความเอียงจะเป็นลบ และหากคว่ำกล้องลงค่าองศาของความเอียงจะเป็นบวก ค่าความเอียงนี้มีความสำคัญในการคำนวณหาระยะห่างระหว่างยานพาหนะด้านหน้าเป็นอย่างยิ่ง เนื่องจากถ้าหากติดตั้งสมาร์ทโฟนด้วยความเอียงที่แตกต่างกัน ยานพาหนะที่ปรากฏในภาพก็จะอยู่ในตำแหน่งที่แตกต่างกันไปด้วย ส่งผลให้ยากที่จะหาระยะห่างที่แท้จริง ดังนั้นเราจึงจำเป็นต้องมีวิธีที่จะชดเชยผลกระทบดังกล่าว

3.2 เตรียมรูปภาพ

ในขั้นตอนการเตรียมรูปภาพ จะเป็นการชดเชยผลกระทบของความเอียงที่แตกต่างกันในการติดตั้งสมาร์ทโฟน ด้วยการปรับตำแหน่งของภาพให้เนื้อหาภายในภาพอยู่ในตำแหน่งที่ใกล้เคียงกัน ด้วยวิธีปรับจุดยึดภาพ ($Anchor\ Point$) ดังนั้นเราจึงทำการทดลองหาความสัมพันธ์ของจุดยึดภาพ กับ ความเอียงในการติดตั้งสมาร์ทโฟนทุก 1° ตั่งแต่ $-10°$ ไปจนถึง 10° โดยพบว่าความสัมพันธ์ระหว่างข้อมูลดังกล่าวเป็นแบบเชิงเส้น และมีสมการที่แทนความสัมพันธ์ระหว่างข้อมูลทั้งสองนี้คือ

$$Anchor\ Point = (-0.0271 \times \theta) + 1$$

โดยที่ $Anchor\ Point$ คือจุดยึดของภาพที่จะทำให้ตำแหน่งของภาพเปลี่ยนไป จนเนื้อหาภายในภาพอยู่ในตำแหน่งที่ใกล้เคียงกัน ดังนั้นวิธีการเตรียมรูปภาพนี้ทำให้เราสามารถติดตั้งสมาร์ทโฟนด้วยความเอียงเท่าไรก็ได้ โดยที่ไม่ส่งผลกระทบต่อการนำไปคำนวณหาระยะห่างระหว่างยานพาหนะ

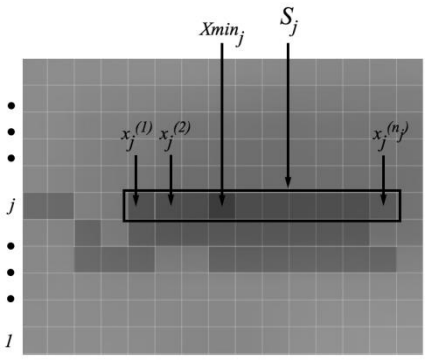
3.3 การเปลี่ยนแปลงรูปภาพ

ในการหาระยะห่างระหว่างยานพาหนะเราจะต้องทำการเปลี่ยนแปลงรูปภาพให้อยู่ในมุมมองของนก (Inverse Perspective Mapping) เพื่อให้ระยะของวัตถุในภาพมีความสอดคล้องกับระยะในโลกจริง ซึ่งตัวอย่างของการเปลี่ยนแปลงภาพ แสดงในภาพที่ 1 โดยในการทดลองนี้เราพัฒนาโปรแกรมประยุกต์บนระบบปฏิบัติการ iOS ด้วยภาษา Objective-C ซึ่งมีฟังก์ชันที่สามารถแปลงรูปภาพให้อยู่ในมุมมองของนกได้คือ



ก. ยานพาหนะด้านหน้าในมุมมองปกติ ข. ยานพาหนะด้านหน้าหลังจากทำ Inverse Perspective Mapping

ภาพที่ 1: ภาพมุมมองของนกที่แปลงจากภาพมุมมองด้านหน้า



ภาพที่ 2: ตัวอย่างคุณสมบัติของเงาที่เราต้องการค้นหาในภาพ

CATransformation3D ที่อยู่ในรูปแบบของเมตริก 4 มิติ โดยในตำแหน่งแถวที่ 2 คอลัมน์ที่ 4 ของเมตริกดังกล่าว คือค่าที่สามารถแปลงภาพให้อยู่ในมุมมองของนกได้ ซึ่งค่าที่ใช้จะอยู่ในช่วง 0-1 ขึ้นอยู่กับขนาดของภาพตั้งต้น โดยขนาดภาพตั้งต้นที่ใช้ในการทดลองนี้คือ 123×185 พิเซล ซึ่งมีค่าที่เหมาะสมเท่ากับ 0.0172

3.4 ประเมินระยะห่างระหว่างยานพาหนะ

หลังจากที่แปลงรูปภาพด้านหน้าให้เป็นภาพมุมมองของนกเรียบร้อยแล้ว เราจะต้องหาตำแหน่งของยานพาหนะที่อยู่ในภาพ และแปลงตำแหน่งดังกล่าวให้เป็นระยะในโลกจริง ดังนั้นเราจะแบ่งขั้นตอนนี้ออกเป็น 2 ส่วนย่อยด้วยกัน

3.4.1 การตรวจหายานพาหนะ

ข้อสังเกตสำคัญที่เราพบในภาพคือใต้ตัวรถจะมีเงาเสมอ ดังนั้นเราจะหาตำแหน่งของยานพาหนะในภาพได้ด้วยการค้นหาเงา ซึ่งเป็นคุณสมบัติที่ชัดเจน และมีอยู่กับยานพาหนะทุกคัน โดยเราจะสแกนหาแถบสีเข้ม (S) ของแต่ละแถวจากล่างขึ้นบน ซึ่งนิยามของ S ที่เราต้องการหา คือตำแหน่งที่มีค่าความเข้มน้อยที่สุดในแถว ($Xmin$) พร้อมตำแหน่งข้างเคียง (x) ที่มีค่าความเข้มต่างจาก $Xmin$ ไม่เกิน 0.2 ที่อยู่ต่อเนื่องกันเป็น

จำนวนมากที่สุดในแถว และสมาชิกทั้งหมดใน S จะต้องมีค่าความเข้มน้อยกว่า 0.55 โดยแสดงตัวอย่างในภาพที่ 2 ซึ่งค่าความเข้มของสมาชิกใน S แต่ละตำแหน่งสามารถหาได้จากสมการ

$$x = \sqrt{r^2 + g^2 + b^2} \qquad (1)$$

ค่าความเข้ม จะอยู่ในช่วง $0 - \sqrt{3}$, r, g และ b คือค่าสี RGB ของตำแหน่งนั้นๆ ซึ่งค่าของแต่ละสีอยู่ในช่วง 0-1 เช่นกัน

แผนภาพการทำงานของอัลกอริทึมตรวจหายานพาหนะแสดงในภาพที่ 3 โดยเราเก็บข้อมูลของ S ในแต่ละแถว ประกอบด้วย n แทนความยาวของ S, d แทนค่าเฉลี่ยความเข้มของทุกจุดที่เป็นสมาชิกใน S ตามสมการ

$$d_j = \frac{\sum_{i=1}^{n_j} x_j^{(i)}}{n_j} \qquad (2)$$

d_j คือค่าเฉลี่ยความเข้มของ S ในแถวที่ j, x คือค่าความเข้มของแต่ละตำแหน่งที่เป็นสมาชิกของ S ในแถวที่ j, md แทนค่ามัธยฐานของ d จากทุกแถวก่อนหน้า จากสมการ

$$md_j = Median(d_{j1}, \cdots, d_{j-1}) \qquad (3)$$

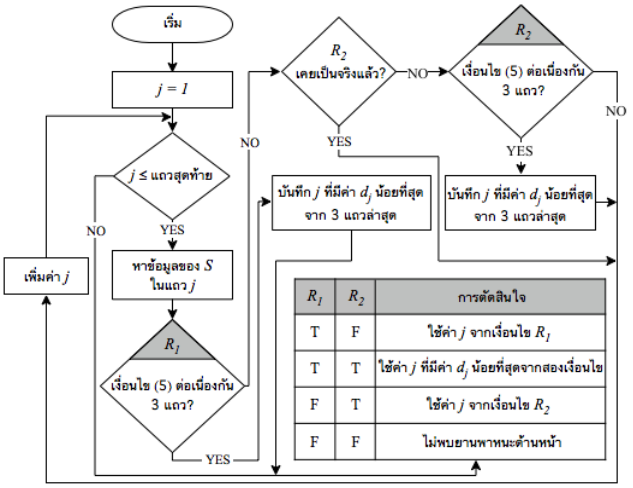
และ mn แทนค่ามัธยฐานความยาวของ S จากทุกแถวก่อนหน้า จากสมการ

$$mn_j = Median(n_{j1}, \cdots, n_{j-1}) \qquad (4)$$

โดยแถวที่เป็นเงาของยานพาหนะ จะต้องมีคุณสมบัติตรงตามเงื่อนไขที่กำหนดอย่างน้อยหนึ่งเงื่อนไข ต่อเนื่องกัน 3 แถว โดยเงื่อนไขแรกคือ

$$n_j \geq \sigma \; AND \; d_j < \frac{md_j}{1.25} \qquad (5)$$

σ คือค่าความยาวขั้นต่ำของ S ที่จะสามารถแยกเงาของยานพาหนะออกจากเงาของวัตถุอื่นๆได้ ซึ่งขึ้นอยู่กับความถี่ในการเลือกตำแหน่งในแถว โดยเงื่อนไขนี้สามารถทำงานได้อย่างมีประสิทธิภาพกับภาพที่มีความสว่างน้อยถึงปานกลาง แต่ในภาพที่มีความสว่างมาก ค่า n ของแต่ละแถวจะน้อยกว่า σ เสมอ



ภาพที่ 3: แผนภาพแสดงการทำงานของอัลกอริทึมตรวจหายานพาหนะ

นั่นทำให้ไม่สามารถหาเงาของยานพาหนะในภาพที่สว่างมากได้ จึงมีการเพิ่มเงื่อนไขที่ 2 สำหรับภาพที่มีความสว่างมากดังนี้

$$n_j \geq mn_j + \tau \qquad (6)$$

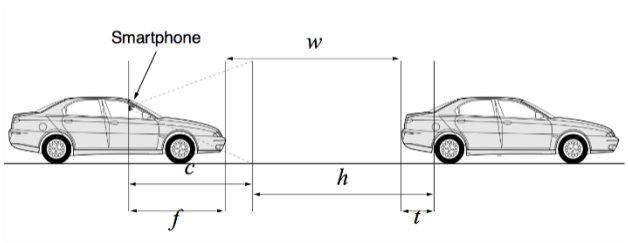
โดย τ คือค่าความยาวขั้นต่ำของ S ที่เรากำหนดไว้สำหรับภาพที่สว่างมาก ซึ่งค่า τ ขึ้นอยู่กับความถี่ของการเลือกตำแหน่งในแถว

ในขั้นตอนสุดท้าย เราจะตัดสินใจเลือกแถวที่เป็นเงาของยานพาหนะในภาพจากตารางการตัดสินใจในภาพที่ 3 โดยค่า j ที่ได้จะเป็นระยะห่างระหว่างจุดที่กล้องมองเห็นกับเงาบริเวณล้อหลังของยานพาหนะด้านหน้าในภาพ(ในหน่วยพิกเซล)

3.4.2 ประเมินระยะห่างระหว่างยานพาหนะ

หลังจากได้ระยะห่างระหว่างจุดที่กล้องมองเห็นกับเงาบริเวณล้อหลังของยานพาหนะในภาพ เราจะต้องแปลงระยะห่างดังกล่าวให้เป็นระยะห่างในโลกจริง ด้วยการหาอัตราส่วนของจำนวนพิกเซลในภาพต่อระยะ 1 เมตรโดยใช้ภาพที่มีการทำสัญลักษณ์ไว้ทุกๆ 1 เมตรบนถนน มาแปลงให้อยู่ในมุมมองของนก และช่องห่างระหว่างสัญลักษณ์ในภาพจะใช้อ้างอิงเท่ากับ 1 เมตรในโลกจริง ซึ่งขึ้นอยู่กับขนาดภาพตั้งต้นและความสูงของสมาร์ทโฟนเหนือพื้นถนน โดยอัตราส่วนของจำนวนพิกเซลต่อเมตรที่ใช้ในการทดลองนี้คือ 16.5 พิกเซล

อย่างไรก็ตามระยะห่างระหว่างยานพาหนะคือระยะระหว่างกันชนหน้าถึงกันชนท้ายของยานพาหนะที่อยู่ด้านหน้า ดังนั้น



ภาพที่ 10: ปัจจัยในการคำนวณหาช่องห่างระหว่างยานพาหนะ

ปัจจัยต่างๆในการหาระยะห่างดังกล่าวแสดงในภาพที่ 4 โดยที่ w คือระยะห่างระหว่างยานพาหนะ, h คือระยะห่างระหว่างจุดที่กล้องมองเห็นถึงบริเวณล้อหลังของยานพาหนะด้านหน้า, c คือระยะจุดบอดของกล้อง, f คือระยะระหว่างสมาร์ทโฟนถึงกันชนหน้ารถ และ t คือระยะระหว่างกันชนท้ายกับล้อหลังของรถ ดังนั้นระยะห่างระหว่างยานพาหนะสามารถคำนวณได้จากสมการ

$$w = \big(h + (c - f)\big) - t \qquad (7)$$

4. ตั้งค่าตัวแปรและประเมินประสิทธิภาพ

4.1 ตั้งค่าตัวแปร

ค่าของตัวแปรที่ถูกกำหนดไว้ในแต่ละส่วนของการทำงานในการทดลองนี้แสดงได้ตามตารางที่ 1 โดยการตรวจหายานพาหนะ เราจะทำการสแกนทุกๆ 5 แถว จากล่างขึ้นบน เนื่องจากหากแสกนด้วยความถี่ของแถวที่น้อยเกินไป จะส่งผลให้อัตราความคลาดเคลื่อนของตำแหน่งยานพาหนะที่ตรวจจับได้ผิดพลาดไปมาก หรือหากสแกนด้วยความถี่ของแถวที่มากเกินไป จะส่งผลให้ระยะเวลาในการประมวลผลสูงขึ้นมากเช่นเดียวกัน โดยตัวแปรที่สำคัญในขั้นตอนการตรวจหายานพาหนะ ประกอบด้วย σ และ τ ซึ่งค่าทั้งสองขึ้นอยู่กับความถี่ในการเลือกตำแหน่งในแถว โดยเราทำการทดลองด้วยความถี่ดังกล่าวที่แตกต่างกัน ประกอบด้วย 1, 5 และ 10 ตามลำดับ โดยที่ σ และ τ ของแต่ละความถี่ในการเลือกตำแหน่งในแถว แสดงในตารางที่ 2 และในส่วนของการประเมินระยะห่าง ค่า t ที่ใช้ คือค่าความยาวโดยประมาณของกันชนท้าย ค่า f ที่ใช้ เป็นค่าระยะห่างโดยประมาณระหว่างสมาร์ทโฟนกับกันชนหน้ารถ และค่า c ขึ้นอยู่กับขนาดของภาพตั้งต้น และความสูงของสมาร์ทโฟนเหนือพื้นถนน

ตารางที่ 1: ค่าคงที่ของตัวแปรในแต่ละส่วนการทำงาน

ส่วนการทำงาน	ตัวแปร	ค่าที่ใช้
ตรวจหารถ	σ	ขึ้นอยู่กับความถี่ในการเลือกตำแหน่งในแถว
	τ	
ประเมินระยะห่างระหว่างยานพาหนะ	t	0.5
	f	1.6
	c	3.64

ตารางที่ 2: ค่าของตัวแปรในแต่ละความถี่ของการเลือกตำแหน่งในแถว

ความถี่ในการเลือกตำแหน่ง (พิกเซล)	ค่าตัวแปรในเงื่อนไข	
	σ	τ
1	55	20
5	8	3
10	5	2

ตารางที่ 3: ค่าเฉลี่ยความผิดพลาดและระยะเวลาในการประมวลผลในแต่ละความถี่ของการเลือกตำแหน่งในแถว

ระยะจริง (เมตร)	ค่าเฉลี่ยความผิดพลาดของแต่ละความถี่ (เมตร)			เวลาประมวลผลเฉลี่ยของแต่ละความถี่ (วินาที)		
	1	5	10	1	5	10
4	0.96	0.46	0.92	3.3	0.4	0.2
6	0.55	0.53	0.46	5.5	0.9	0.4
8	0.48	0.64	0.56	4.8	0.9	0.4
10	0.59	0.83	0.66	6.3	1.1	0.6
12	1.02	1.02	1.19	8.6	1.5	0.8
14	1.33	1.44	1.36	9.7	1.8	0.9
16	1.49	1.57	1.50	11.3	2.1	1.1
18	2.15	2.73	2.72	13.5	2.6	1.3

ซึ่งในการทดลองนี้ติดตั้งสมาร์ทโฟนที่ความสูงประมาณ 1.15 เมตร เหนือพื้นถนน ดังนั้นค่า c ที่ใช้ คือระยะจุดบอดของกล้องสำหรับระดับความสูงดังกล่าว

4.2 ข้อมูลที่ใช้ในการทดลอง

ในการทดลองนี้เราเก็บตัวอย่างภาพทั้งหมด 480 ภาพ ในสภาพแวดล้อมที่แตกต่างกันทั้ง แสง,เงา, พื้นผิวถนน รวมไปถึงระยะห่างระหว่างยานพาหนะ ซึ่งประกอบไปด้วยระยะ 4, 6, 8, 10, 12, 14, 16 และ18 เมตร โดยบันทึกภาพในขณะที่รถหยุดนิ่งบนถนนจริง เนื่องจากการหาระยะห่างที่แท้จริงในขณะ

ยานพาหนะเคลื่อนที่เพื่อนำมาใช้อ้างอิงนั้นเป็นไปได้ยาก จากนั้นเราจะแบ่งข้อมูลทั้งหมดออกเป็น 2 ชุด ชุดแรก 160 ภาพ สำหรับใช้ในการปรับค่าของตัวแปรให้มีประสิทธิภาพสูงสุดในการตรวจหายานพาหนะในภาพและอีก 320 ภาพ สำหรับทดสอบความแม่นยำของวิธีการที่นำเสนอ

4.3 ผลการทดลอง และความแม่นยำ

เราประเมินหาระยะห่างระหว่างยานพาหนะด้วยการใช้ความถี่ของการเลือกตำแหน่งในแถวที่แตกต่างกัน ซึ่งความแม่นยำจะวัดจากค่าความผิดพลาดของระยะห่างที่ประเมินได้กับค่าระยะห่างจริงที่วัดและบันทึกไว้ของแต่ละภาพ ซึ่งความแม่นยำและระยะเวลาในการประมวลผลของแต่ละความถี่ในการเลือกตำแหน่งแสดงในตารางที่ 3 ซึ่งสังเกตได้ว่า ช่วงระยะห่างระหว่างยานพาหนะที่ต่ำกว่า 12 เมตร มีความแม่นยำค่อนข้างสูง โดยมีความผิดพลาดเฉลี่ยไม่ถึง 1 เมตร และเพิ่มขึ้นเล็กน้อยในระยะห่างที่มากขึ้น สูงสุด 2.67 เมตรที่ระยะห่าง 18 เมตร และส่วนของระยะเวลาในการประมวลผลนั้น หากสแกนทุกตำแหน่งในแถวจะใช้เวลาในการประมวลผลสูงมาก และลดลงเมื่อความห่างของความถี่มากขึ้น ซึ่งหากเทียบกับความแม่นยำแล้ว การสแกนทีละ 10 ตำแหน่งมีความแม่นยำแตกต่างจากการสแกนทุกตำแหน่งในแถวเพียงเล็กน้อยเท่านั้น ด้วยระยะเวลาในการทำงานที่แตกต่างกันเกือบ 10 เท่าตัว

5. สรุป

ในบทความวิจัยนี้ เราได้นำเสนอวิธีการประเมินหาระยะห่างระหว่างยานพาหนะ ด้วยภาพถ่ายมุมมองด้านหน้าจากกล้องของสมาร์ทโฟนที่ติดตั้งไว้ด้านหลังกระจกหน้ารถ ซึ่งเห็นได้ว่าสมาร์ทโฟนสามารถใช้เป็นเซ็นเซอร์เคลื่อนที่สำหรับเก็บข้อมูลระยะห่างระหว่างยานพาหนะในขณะเดินทาง ด้วยความแม่นยำที่สามารถนำไปใช้เป็นข้อมูลการจราจรได้ โดยมีความคลาดเคลื่อนเฉลี่ยเพียง 10% เท่านั้น และเนื่องจากวิธีการที่นำเสนอนี้ใช้รูปภาพในการประมวลผล จึงใช้ทรัพยากรในการประมวลผลน้อยกว่าการประมวลผลวีดีโอ จึงสามารถใช้ในสมาร์ทโฟนที่มีประสิทธิภาพไม่สูงนักได้ และโดยเฉพาะข้อมูลระยะห่างระหว่างยานพาหนะนี้ หากได้มาจากยานพาหนะหลายๆคันในเส้นทางเดียวกัน จะสามารถนำไปใช้ในการประเมินหาความหนาแน่นของการจราจร ซึ่งนับว่าเป็นข้อมูลการจราจรที่สำคัญในเส้นทางนั้นๆได้อีกด้วย

เอกสารอ้างอิง

[1] “iOnRoad Augmented Driving Pro”, *available: http://www.ionroad.com/*.

[2] S. Panichpapiboon and P. Leakkaw, “Traffic Sensing Through Accelerometers,” *IEEE Transactions on Vehicular Technology*, vol. 65, no. 5, pp. 3559 – 3567, 2016.

[3] R. Banerjee and A. Sinha, “Trigsense: Accelerometer triggered audio sensing for traffic condition monitoring” *Proceedings of the IEEE Int. Conf. Intell. Sensors, Sensor Netw., and Info. Processing*, Apr 2013.

[4] J. Wang, T. Mei, B. Kong, and H. Wei, “An approach of lane detection based on inverse perspective mapping” *Proceedings of the IEEE Intl. Conf. Intelligent Transportation Systems (ITSC)*, Oct 8-11, 2014.

[5] M. Aly, “Real time detection of lane markers in urban streets” *Proceedings of the IEEE Intelligent Vehicles Symposium (IV)*, Jun 4-6, 2008.

[6] S. Tuohy, D. O’Cualain, E. Jones, and M. Glavin, “Distance determination for an automobile environment using inverse perspective mapping in opencv” *Proceedings of the IET Irish Signals and Systems Conference*, Jun 23-24, 2010.

[7] J. Chang and C. Cho, “Vision-based front vehicle detection and its distance estimation” *Proceedings of the IEEE International Conference on Systems Man and Cybernetics*, Oct 8-11, 2006.

[8] R. Itu and R. Danescu, “An efficient obstacle awareness application for android mobile devices” *Proceedings of the IEEE Intelligent Computer Communication and Processing (ICCP)*, Sep 4-6, 2014.

[9] M. Rezaei, M. Terauchi, and R. Klette, “Robust vehicle detection and distance estimation under challenging lighting conditions,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 5 pp. 2723-2743, 2015.

[10] R. G. Danescu, R. Itu, and A. Petrovai, “Sensing the driving environment with smart mobile devices” *Proceedings of the IEEE Intelligent Computer Communication and Processing (ICCP)*, Sep 3-5, 2015.

ระบบติดตามรถประจำทางแบบเรียลไทม์บนโทรศัพท์เคลื่อนที่
A System of Real-time Bus Tracking on Mobile Phone

สุทัศน์ บรรจง (Suthat Banchong)[1] บัญชา จุมพล (Buncha Joompol)[2] และจันทิมา พลพินิจ (Jantima Polpinij)[3]
ห้องปฏิบัติการวิจัยเชิงปัญญา (Intellect Laboratory)
ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาการสารสนเทศ มหาวิทยาลัยมหาสารคาม
[1]suthat.ban@msu.ac.th, [2]buncha.chu@msu.ac.th,[3] jantima.p@msu.ac.th

บทคัดย่อ

งานนี้นำเสนอการพัฒนาระบบต้นแบบของการติดตามรถประจำทางแบบเรียลไทม์บนโทรศัพท์เคลื่อนที่บนระบบปฏิบัติการแอนดรอยด์ที่มีต้นทุนต่ำ แนวคิดเบื้องหลังการพัฒนาระบบนี้คืออินเทอร์เน็ตในทุกสรรพสิ่ง ที่เป็นการเชื่อมต่ออุปกรณ์ ต่างๆ(นอกเหนือจากการใช้งานอุปกรณ์ทั่วไป เช่นคอมพิวเตอร์และสมาร์ทโฟน) ไปยังอินเทอร์เน็ต เป้าหมายของระบบต้นแบบนี้คือการติดตามรถประจำทางแบบเรียลไทม์ และการตรวจสอบความหนาแน่นของผู้ใช้บริการบนรถโดยสาร ซึ่งแอพพลิเคชั่นนี้น่าจะช่วยผู้ใช้งานสามารถวางการเดินทางได้

คำสำคัญ: การติดตามรถประจำทางแบบเรียลไทม์ แอนดรอยด์ โทรศัพท์เคลื่อนที่ อาร์ดุยโน่ อินเทอร์เน็ตในทุกสรรพสิ่ง

Abstract

This work presents a prototype development of real-time bus tracking on Android mobile phone with low cost. The concept behind this prototype is Internet of Things (*IoT*) *that refers to the connection of devices* (*other than typical fare such as computers and smartphones*) *to the Internet. The objective of this prototype system is to track bus in real-time, and monitor the number of passengers on the bus. As the result, this application may help users to plan for their travel in a city or university.*

Keyword: Real-time bus tracking, Mobile phone, Android, Arduino, Internet of Things

1. บทนำ

GPS Tracking [1], [2] เป็นการใช้ GPS ในการติดตามตำแหน่งของบางสิ่งบางอย่างในปัจจุบัน เช่น รถ, เด็ก, คนชรา, หรือแม้แต่สัตว์เลี้ยง โดยอาจจะมีฟังก์ชันเสริมเข้าไปในตัว GPS Tracking เพื่อให้มันสามารถบอกสิ่งอื่นได้อีกนอกจาก พิกัด Latitude กับ Longitude เท่านั้น ปัจจุบัน GPS Tracking นิยมใช้ในการติดตามรถยนต์มากที่สุด โดยได้มีการเพิ่มเซนเซอร์เข้ามาเพื่อหาข้อมูลอื่นๆ ในรถที่ต้องการทราบ เช่น ความเร็วของรถที่วิ่งในขณะนั้น หรือแม้แต่ทิศทางในการวิ่ง และอื่นๆ อีกมากมาย GPS Tracking ในการติดตามรถยนต์ที่มี GPS เป็นตัวช่วยในการรู้ตำแหน่งของรถยนต์ที่ต้องการติดตามนั้น มีได้หลายรูปแบบทั้งแบบติดตามย้อนหลัง หรือ Off-line และแบบติดตามแบบเรียลไทม์ (Real-Time) [1], [2], [3], [4]

ปัจจุบันนี้การติดตามรถประจำทางแบบเรียลไทม์ (Real Time Bus Tracking) [3], [4] มีการใช้กันอย่างแพร่หลายไปทั่วโลก โดยเฉพาะการใช้เพื่อการติดตามรถโดยสารประจำทางระหว่างจังหวัดและรถขนส่งสินค้า [3], [4], [5], [6] เพื่อใช้ตรวจสอบพฤติกรรมการขับขี่ทุกรูปแบบ และป้องกันปัญหาที่มีสาเหตุมาจากผู้ขับขี่ เช่น การทุจริตน้ำมัน ขับนอกเส้นทาง นำรถใช้ส่วนตัว หรือการลักขโมยสินค้า และผู้ใช้ยังสามารถเรียกดูรายงานย้อนหลัง โดยสามารถตรวจสอบการใช้น้ำมัน เส้นทางการวิ่ง และความเร็ว เป็นต้น [4], [5]

อย่างไรก็ตาม สำหรับรถโดยสารที่ใช้เพื่อการเดินทางระยะสั้น เช่น ในตัวจังหวัด หรือระหว่างอำเภอและจังหวัด หรือในมหาวิทยาลัย ส่วนใหญ่จะยังไม่มีการใช้งาน GPS Tracking เนื่องจากอาจจะไม่คุ้มค่าในเรื่องการลงทุน ทำให้ผู้ใช้บริการรถโดยสารเหล่านั้นไม่สามารถทราบถึงสถานะของรถโดยสารที่

ต้องการติดตาม โดยเฉพาะข้อมูลที่สำคัญ เช่น ตำแหน่งของรถโดยสารในปัจจุบัน เป็นต้น ทำให้ผู้ใช้บริการรถโดยสารเหล่านั้นยากที่จะวางแผนการเดินทางได้อย่างเหมาะสม

จากเหตุผลข้างต้น ทำให้เกิดแนวคิดสำหรับงานวิจัยฉบับนี้ ซึ่งเป็นการนำเสนอการพัฒนาระบบต้นแบบการติดตามรถประจำทางแบบเรียลไทม์บนโทรศัพท์เคลื่อนที่ผ่าน Web-based GPS Tracking [3], [4] สำหรับการติดตามสถานะของรถโดยสารที่ใช้เพื่อการเดินทางระยะสั้น โดยมีเครื่องมือหลักที่ใช้ในการพัฒนาระบบคือ Arduino และเซนเซอร์ GPS ในการตรวจจับตำแหน่ง สำหรับผู้ใช้งานระบบฯ เพื่อการตรวจสอบตำแหน่งหรือสถานะของรถประจำทางเพียงเข้าไปยังเว็บไซต์หรือการใช้งานเว็บไซต์ผ่านแอพพลิเคชั่นบนโทรศัพท์เคลื่อนที่ที่เชื่อมต่อไปยังอินเทอร์เน็ต ผู้ใช้บริการก็สามารถรับรู้ถึงตำแหน่งของรถโดยสารเป้าหมายที่ต้องการใช้บริการ ตลอดจนความหนาแน่นของผู้โดยสารบนรถประจำทางเป้าหมาย [7] ซึ่งเครื่องมือที่คณะผู้จัดทำนำเสนอน่าจะช่วยให้ผู้ใช้บริการรถโดยสารระยะสั้นสามารถวางแผนการเดินทางของตนเองได้อย่างเหมาะสมมากยิ่งขึ้น

2. แนวคิด เทคนิค และเครื่องมือที่เกี่ยวข้อง

2.1 อินเทอร์เน็ตในทุกสรรพสิ่ง (Internet of Things: IoT)

IoT [8] คือ เทคโนโลยีอินเทอร์เน็ตที่เชื่อมต่ออุปกรณ์และเครื่องมือต่างๆ เข้าด้วยกัน เช่น โทรศัพท์มือถือ รถยนต์ โทรทัศน์ เป็นต้น โดยเครื่องมือต่างๆ จะสามารถเชื่อมโยงและสื่อสารกันได้โดยผ่านระบบอินเตอร์เน็ตนั่นเอง ซึ่งเราสามารถนำขีดความสามารถของ IoT มาช่วยในการบริหารจัดการและการพัฒนาสิ่งใหม่เพื่อตอบสนองต่อการทำงาน โดยใช้ข้อมูลต่างๆ ที่ถูกรวบรวมผ่านเทคโนโลยีเซนเซอร์ (Sensor Technology) แล้วส่งผ่านเครือข่ายอินเทอร์เน็ตเพื่อมาเก็บไว้ในฐานข้อมูลกลาง โดย “Thing” ในที่นี้ก็คือสิ่งที่เราสนใจควบคุม (Control) หรือตรวจสอบ (Monitor) ที่ถูกตรวจจับโดยเซนเซอร์ และเซนเซอร์จะส่งสัญญาณไปสู่เซิร์ฟเวอร์ผ่านระบบเครือข่าย ซึ่งอาจจะเป็นเครือข่ายภายใน (Intranet) หรือ เครือข่ายภายนอก (Extranet) ที่เป็นระบบปิด หรือ เครือข่ายอินเทอร์เน็ตที่เป็นระบบเปิดจากนั้นผู้ใช้ระบบจะสามารถตรวจสอบสถานะของ “Thing” ได้ โดยที่ไม่ต้องเสียเวลาไปนั่งตรวจสอบหรือจัดการกับ “Thing” เหล่านั้น



ภาพที่ 1: กรอบแนวคิดของอินเตอร์เน็ตในทุกสรรพสิ่ง

2.2 อาร์ดุยโน่ (Arduino)

อาร์ดุยโน่ [9] เป็นบอร์ดไมโครคอนโทรลเลอร์ตระกูล AVR ที่มีการพัฒนาแบบ Open Source บอร์ด Arduino ถูกออกแบบมาให้ใช้งานได้ง่าย ดังนั้นจึงเหมาะสำหรับผู้เริ่มต้นศึกษา ทั้งนี้ผู้ใช้งานยังสามารถดัดแปลง เพิ่มเติม และพัฒนาต่อยอดได้ทั้งตัวบอร์ดหรือโปรแกรม บอร์ดอาร์ดุยโน่ยังง่ายในการต่ออุปกรณ์เสริมต่างๆ คือผู้ใช้งานสามารถต่อวงจรอิเล็กทรอนิกส์จากภายนอกแล้วเชื่อมต่อเข้ามาที่ขา I/O ของบอร์ด หรือเพื่อความสะดวกสามารถเลือกต่อกับบอร์ดเสริม (Arduino Shield) ประเภทต่างๆ เช่น XBee Shield, Music Shield, และ Relay Shield เป็นต้น

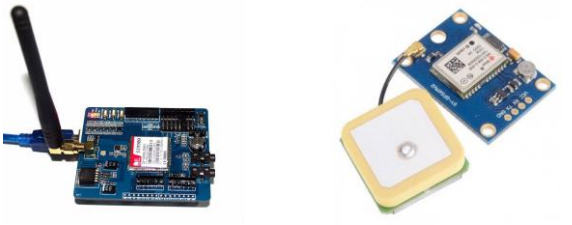
ในงานวิจัยนี้ใช้อาร์ดุยโน่ Uno R3 บอร์ดไมโครคอนโทรลเลอร์แบบโอเพนซอร์ซออกแบบมาให้ใช้งานได้ง่าย ใช้ชิพ ATmega328 รันที่ความถี่ 16 MHz หน่วยความจำแฟลช 32 KB แรม 2 KB บอร์ดใช้ไฟเลี้ยง 7 ถึง 12 V มีระดับแรงดันไฟฟ้าในการทำงานและขาสัญญาณอยู่ที่ 5 V (TTL) มีอินพุตแบบดิจิตอลหรือเอาท์พุต 14 ขา (เป็น PWM ได้ 6 ขา) มีอินพุตแบบอนาล็อก 6 ขา Serial UART 1 ชุด I2C 1 ชุด SPI 1 ชุด เขียนโปรแกรมบนซอฟต์แวร์ อาร์ดุยโน่ IDE และโปรแกรมผ่านพอร์ตยูเอสบี (USB)



ภาพที่ 2: เมนบอร์ดอาร์ดุยโน่ Uno R3

2.3 Arduino sim900 Module

Arduino sim900 Module [10] มีความสามารถเหมือนโทรศัพท์ ที่สามารถโทรเข้า/โทรออก รับสาย/วางสาย ส่ง SMS ส่งคำสั่ง ussd เชื่อมต่ออินเตอร์เน็ต หรือทำงานต่างๆ ได้เหมือนโทรศัพท์มือถือ ซึ่งในงานนี้จะใช้ Arduino sim900 Module ในการเชื่อมต่อกับอินเตอร์เน็ต เพื่อที่จะให้ GPS ทำงานได้และสามารถส่งค่าไปยังเซิร์ฟเวอร์ (Server) ได้



(a) Arduino sim900 Module (b) Arduino GPS Module

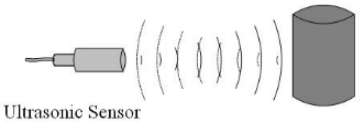
ภาพที่ 3: Arduino sim900 Module และ Arduino GPS Module

2.4 Arduino GPS Module

GPS (Global Positioning System) [11] เป็นระบบบอกพิกัดบนโลกที่อาศัยตำแหน่งของดาวเทียม GPS เป็นพิกัดอ้างอิง โดยเครื่องรับ GPS สามารถคำนวณตำแหน่งบนพื้นโลกได้จากการวัดระยะทางระหว่างเครื่องรับและดาวเทียม GPS ซึ่งต้องมีอย่างน้อย 4 ดวง เพื่อที่จะสามารถนำเอาความสัมพันธ์ไปสร้างเป็นสมการ 4 สมการซึ่งจำเป็นสำหรับการคำนวณหาตัวแปร 4 ตัว (Xr, Yr, Zr และ cdt) [11]

2.5 เซนเซอร์อัลตราโซนิก (Ultrasonic Sensor)

อัลตราโซนิก [12] เป็นเซนเซอร์ที่ทำงานโดยอาศัยคลื่นเสียงที่มีความถี่สูงกว่า 20 กิโลเฮิร์ต (kHz) ซึ่งเป็นคลื่นในย่านที่มนุษย์ไม่สามารถได้ยินเสียง เซนเซอร์ชนิดอัลตราโซนิกทำงานโดยอาศัยการกระจาย หรือการเคลื่อนที่ของคลื่นเสียงไปกระทบกับพื้นผิวของตัวกลาง ซึ่งอาจเป็นของแข็งหรือของเหลว บางส่วนของคลื่นเสียงจะแทรกผ่านเข้าไปในตัวกลางนั้น และส่วนใหญ่ของคลื่นความถี่สูงนี้จะสะท้อนกลับเรียกว่า “Echo” โดยช่วงเวลาของการสะท้อนกลับของคลื่นเสียงเป็นสัดส่วนโดยตรงกับระยะห่างระหว่างวัตถุกับเซนเซอร์



ภาพที่ 4: แสดงการทำงานของคลื่นอัลตราโซนิก

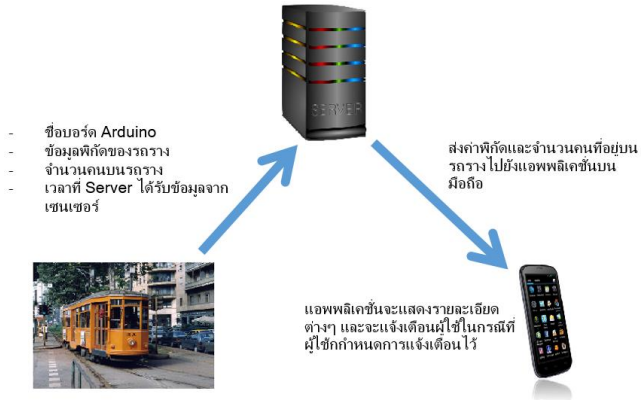
อย่างไรก็ตาม ควรเว้นระยะห่างของเซนเซอร์และวัตถุที่ต้องการตรวจสอบให้เหมาะสม เพื่อไม่ให้คลื่นเสียงที่ส่งออกไปมีการรบกวนกัน เพื่อประสิทธิภาพและความแม่นยำในการตรวจจับของตัวเซนเซอร์ นอกจากนี้การวางสิ่งของที่จะทำการตรวจจับ ควรจัดวางสิ่งของที่จะทำการตรวจจับให้มีระนาบที่จะสามารถสะท้อนคลื่นเสียงที่ส่งไปกลับมาได้ ตามคุณสมบัติของคลื่นเสียงที่ว่า มุมตกกระทบเท่ากับมุมสะท้อน เพื่อให้ได้ประสิทธิภาพในการตรวจจับดีที่สุด

2.6 Google Maps APIs

Google Maps [13] คือบริการของ Google ที่ให้บริการเทคโนโลยีด้านแผนที่ประสิทธิภาพสูง ใช้งานง่าย และให้ข้อมูลของธุรกิจในท้องถิ่น ได้แก่ ที่ตั้งของธุรกิจรายละเอียดการติดต่อ และเส้นทางการขับขี่ โดยบริการแผนที่นี้เริ่มต้นให้บริการตั้งแต่กลางปี ค.ศ. 2005 เป็นบริการฟรี จัดให้แก่ผู้ใช้ทั่วโลกส่วนประกอบที่สำคัญที่ดึงดูดผู้ใช้งานเป็นอย่างมาก คือแผนที่และภาพถ่ายดาวเทียมคุณภาพดี ซึ่งครอบคลุมพื้นผิวโลกในมาตราส่วนต่าง ๆ ตามความเหมาะสม

3. การออกแบบและพัฒนาระบบ

กรอบการดำเนินงานของการพัฒนาระบบตรวจสอบตำแหน่งการเดินรถประจำทางบนระบบปฏิบัติการแอนดรอยด์ ได้ดังต่อไปนี้



ภาพที่ 5: กรอบการดำเนินงานของระบบ

ขั้นตอนที่ 1: *รวบรวมข้อมูล* (*Data Collection*)

ในขั้นตอนนี้ จะเป็นการสังเกต เวลาที่รถประจำทางแต่ละคันที่วิ่งให้บริการ เวลาโดยเฉลี่ยในแต่ละรอบการโดยสาร จำนวนผู้ใช้บริการรถประจำทาง พฤติกรรมการขึ้น - ลง รถ

ประจำทางของผู้ใช้บริการรถประจำทาง จุดที่รถประจำทางจอดรับผู้โดยสาร และเส้นทางการเดินรถประจำทางในมหาวิยาลัยมหาสารคาม

ขั้นตอนที่ 2: *การออกแบบ (Design)* - จะประกอบด้วย 5 ขั้นตอนได้แก่

(1) ออกแบบการทำงานของระบบเซ็นเซอร์ตรวจจับจำนวนผู้โดยสารที่ขึ้น - ลง รถประจำทางด้วย Ultrasonic sensor และส่งค่าจำนวนของความหนาแน่นของผู้บริการบนรถประจำทางในเวลานั้น ๆ ไปยังฐานข้อมูลเพื่อแสดงแก่ผู้ใช้

(2) ออกแบบระบบการส่ง – รับค่า GPS จากตำแหน่งของรถประจำทางไปยัง Server และส่งต่อไปยังผู้ใช้แอพพลิเคชั่น

(3) ออกแบบฐานข้อมูล (Database Design) เพื่อเก็บข้อมูลของจำนวนความหนาแน่นของผู้ใช้บริการ และตำแหน่งของรถประจำทางในเวลานั้น ๆ เพื่อส่งไปยังผู้ใช้

(4) ออกแบบและโปรแกรมควบคุมอุปกรณ์คอนโทรลเลอร์ เซ็นเซอร์ และโมดูลต่าง ๆ ให้เชื่อมต่อกันและทำงานได้ตามที่ต้องการ

(5) ออกแบบโปรแกรมและส่วนของการติดต่อผู้ใช้งาน (User Interface) วางแผนฟังก์ชันการทำงานของแอพพลิเคชั่น ให้เชื่อมต่อกับ Server ได้ และให้มีความสะดวกต่อการใช้งาน

ขั้นตอนที่ 3: *การพัฒนาระบบ (Implementation)* - จะแบ่งออกเป็น 3 ส่วน ได้แก่

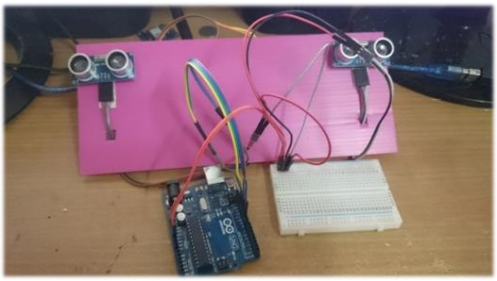
(1) การพัฒนาระบบเพื่อควบคุมการทำงานของอุปกรณ์ Arduino โดยใช้โปรแกรม Arduino IDE ซึ่งอยู่บนพื้นฐานการเขียนโปรแกรมของภาษาซี

(2) การจำลอง Server เพื่อการจัดเก็บข้อมูลจำนวนความหนาแน่นของผู้ใช้บริการและพิกัดตำแหน่งของรถประจำทาง เพื่อส่งค่าไปยังแอพพลิเคชั่นบนมือถือของผู้ใช้ได้

(3) ส่วนของขั้นตอนการพัฒนาแอพพลิเคชั่นบนระบบแอนดรอยด์ โดยใช้โปรแกรม Android-studio

3.1 อาร์ดุยโน่และเซนเซอร์อัตราโซนิก

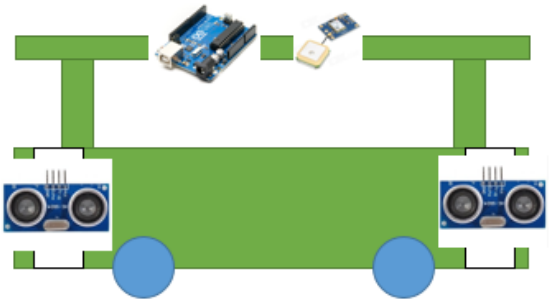
งานวิจัยนี้ออกแบบให้รถโดยสารมีประตูทางขึ้น-ลงจำนวน 2 ประตู โดย 1 ประตู จะมีเซนเซอร์อัตราโซนิกจำนวน 2 ตัว โดยจะวางตำแหน่งของเซนเซอร์อัตราโซนิกตัวที่ 1 จะอยู่ข้างนอก ตัวที่ 2 จะอยู่ถัดจากตัวที่ 1 เข้าไปข้างใน ห่างกันประมาณ 20 เซนติเมตร ซึ่งมีหลักการคือ เมื่อเซ็นเซอร์ตัวที่ ทำงาน (Active) ซึ่งก็หมายความว่ามีผู้โดยสารเดินผ่านเซนเซอร์ตัวที่ 1 ไปแล้ว จากนั้นถ้าเซนเซอร์ตัวที่ 2 ทำงานตามในช่วงเวลาที่กำหนดไว้ในโปรแกรม ระบบก็จะนับเป็นผู้โดยสารเดินขึ้นมาบนรถ และในทางกลับกับ เมื่อเซ็นเซอร์ตัวที่ 2 ทำงานแล้วถ้าเซนเซอร์ตัวที่ 1 ทำงานตามในช่วงเวลาที่กำหนดไว้ในโปรแกรม ระบบก็จะนับเป็นผู้โดยสารเดินลงจากรถ การวางตำแหน่งของเซนเซอร์ดูได้จากภาพที่ 6



ภาพที่ 6: การต่อเชื่อมเซนเซอร์อัลตราโซนิกและบอร์ดอาร์ดุยโน่

จากภาพที่ 6 เนื่องจากเซนเซอร์อัตราโซนิกแต่ละตัวต้องการไฟเลี้ยง เพื่อให้ใช้งานได้ที่ 5.0V ซึ่งตัวบอร์ด อาร์ดุยโน่ UNO R3 มีพอร์ทจ่ายไฟที่ 5.0V เพียงพอร์ทเดียว ดังนั้น จึงจำเป็นต้องมีการใช้ Breadboard เข้ามาช่วยในการกระจายสัญญาณไฟ

สำหรับการวางตำแหน่งของแผงวงจรดังกล่าวบนรถโดยสาร จะวางไว้ด้านบนประตูที่ห่างจากผู้โดยสารที่จะเดินขึ้นลงประมาณ 50 เซนติเมตร ซึ่งแผนผังการการติดตั้งอุปกรณ์คอนโทรลเลอร์บนรถโดยสารสามารถแสดงได้ดังภาพที่ 7



ภาพที่ 7: ตำแหน่งการติดตั้งอุปกรณ์คอนโทรเลอร์บนรถโดยสาร

4. ผลการทดสอบระบบ

ในการทดสอบระบบติดตามรถประจำทางแบบเรียลไทม์บนโทรศัพท์เคลื่อนที่ จะทำการทดสอบใน 2 ส่วน ซึ่งเป็นองค์ประกอบหลักของระบบ นั่นคือ การทดสอบการตรวจวัดความหนาแน่นของผู้ใช้บริการบนรถประจำทางด้วยเซนเซอร์อัลตราโซนิก และการทดสอบการประมาณเวลาของการมาถึงของรถประจำทางจากตำแหน่งของผู้ใช้แอปพลิเคชันและตำแหน่งของรถประจำทาง

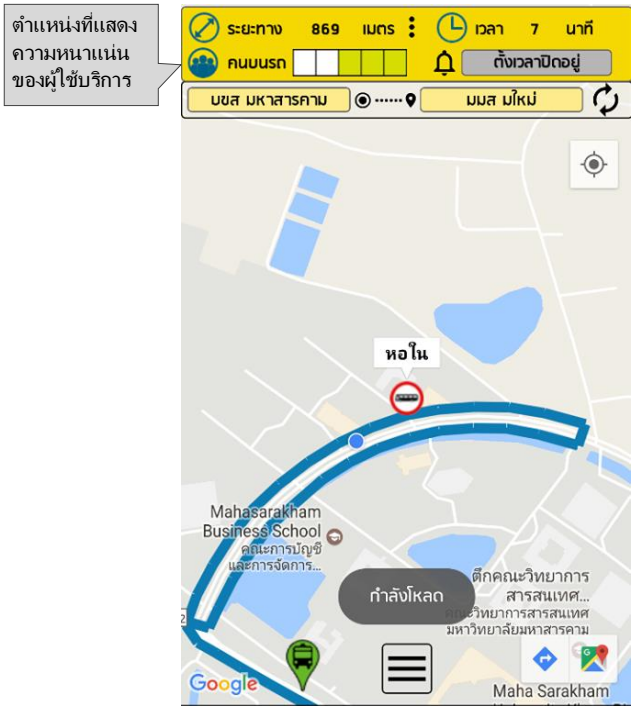
4.1 การทดสอบการตรวจวัดความหนาแน่นของผู้ใช้บริการบนรถประจำทางด้วยเซนเซอร์อัลตราโซนิก

ในส่วนของการใช้เซนเซอร์อัลตราโซนิกในการตรวจวัดความหนาแน่นของผู้ใช้บริการบนรถประจำทาง โดยในการแสดงความหนาแน่นของผู้ใช้บริการบนรถประจำทางนั้น จะแสดงเป็นแถบสี โดยความหมายของแถบสีสามารถแสดงได้ดังตารางที่ 1

ตารางที่ 1: ความหมายของแถบสี

แถบสี	ความหมาย	จำนวนผู้ใช้บริการโดยประมาณการ
ขาว	ว่าง	ไม่มีผู้ใช้บริการ
เขียวอ่อน	น้อยมาก	ประมาณ 1 - 10 คน
เขียวเข้ม	น้อย	ประมาณ 11 - 20 คน
เหลือง	ปานกลาง	ประมาณ 21 - 30 คน
ส้ม	หนาแน่น	ประมาณ 31 - 40 คน
แดง	หนาแน่นมาก	มากกว่า 40 คน

อย่างไรก็ตาม แม้การตรวจจับความหนาแน่นของผู้ใช้บริการบนรถประจำทางด้วยเซนเซอร์อัลตราโซนิกจะสามารถแสดงผลได้ค่อนข้างน่าพอใจ เพราะจากการทดสอบการตรวจจับความหนาแน่นของผู้ใช้บริการบนรถประจำทางด้วยเซนเซอร์อัลตราโซนิกจำนวน 50 ครั้ง พบว่าให้ผลลัพธ์ถูกต้อง (Accuracy) คิดเป็นร้อยละ 80 โดยข้อผิดพลาดที่เกิดขึ้นนั้น เกิดขึ้นในกรณีที่ผู้โดยสารขึ้นรถประจำทางมาพร้อมๆ กัน 2 คน ทำให้การตรวจจับเกิดข้อผิดพลาด เนื่องจากลักษณะการสะท้อนกลับของเสียงขึ้นอยู่กับมุมตกกระทบที่ทำให้เสียงกระจายไปในทิศทางต่างๆ จึงอาจจะไม่เหมาะกับวัตถุที่มีลักษณะเป็นก้อนที่ไม่สม่ำเสมอ ซึ่งรูปร่างของคนก็เป็นเช่นนั้น ดังนั้นผลที่ได้จากการสะท้อนกลับของคลื่นอัลตราโซนิกที่ใช้กับวัตถุลักษณะนี้จะมีความเที่ยงตรงลดลง [12]



ภาพที่ 8: การแสดงความหนาแน่นของผู้ใช้บริการบนรถประจำทางของแอปพลิเคชันที่ใช้บนโทรศัพท์เคลื่อนที่

4.2 การทดสอบการประมาณเวลาของการมาถึงของรถประจำทางจากตำแหน่งของผู้ใช้แอปพลิเคชันและตำแหน่งของรถประจำทาง

ในการทดสอบการประมาณเวลาของการมาถึงของรถประจำทางจากตำแหน่งของผู้ใช้แอปพลิเคชันและตำแหน่งของรถประจำทาง เราได้ทดสอบระบบด้วยผู้ช่วยที่แตกต่างกันทั้งสิ้น 30 คน ที่มีจุดรอรถที่แตกต่างกัน ซึ่งในการประมาณเวลาจะอาศัยตำแหน่งของรถโดยสารที่ส่งผ่านเซนเซอร์ GPS ขณะเดียวกันก็มีการส่งตำแหน่งของผู้ใช้ผ่าน GPS บนโทรศัพท์เคลื่อนที่ เพื่อใช้ในการประมาณระยะเวลาของการมาถึงของรถโดยสาร

จากการทดสอบพบว่าการประมาณระยะเวลาของการมาถึงของรถโดยสารสำหรับจำนวนผู้ทดสอบทั้ง 30 คน ให้ความถูกต้องร้อยละ 75 สาเหตุของการประมาณที่ผิดพลาด เนื่องจากในระหว่างตำแหน่งของรถโดยสารและตำแหน่งของผู้ใช้งานอาจจะมีจุดรอรถโดยสารหลายๆ จุด และในบางจุดอาจจะมี

ผู้โดยสารรอใช้บริการหลายคน ทำให้ระยะในการให้บริการในแต่ละจุดเพิ่มมากขึ้น ซึ่งนำไปสู่ความผิดพลาดที่อาจจะเกิดขึ้นได้ นอกจากนี้ยังรวมไปถึงการสูญหายของสัญญาณ GPS ในบางช่วงเวลาอีกด้วย

5. สรุป

งานวิจัยนี้ได้นำเสนอการออกแบบและพัฒนาระบบต้นแบบราคาประหยัดของระบบติดตามรถประจำทางแบบเรียลไทม์บนโทรศัพท์เคลื่อนที่แบบ Web-based GPS Tracking สำหรับการติดตามสถานะของรถโดยสารที่ใช้เพื่อการเดินทางระยะสั้น ภายใต้แนวคิดอินเทอร์เน็ตในทุกสรรพสิ่ง (Internet of Things) สำหรับตรวจสอบตำแหน่งของรถประจำทางผ่านโทรศัพท์เคลื่อนที่ระบบปฏิบัติการแอนดรอยด์ (Android) เพื่อให้ผู้รอใช้บริการรถประจำทางสามารถตรวจสอบตำแหน่งของรถประจำทางแบบเรียลไทม์ (Real-time) สำหรับคาดการณ์เวลาของการมาถึง ณ ตำแหน่งของการรอคอย ตลอดจนความหนาแน่นของผู้โดยสารบนรถประจำทางเป้าหมาย เพื่อประกอบการวางแผนการเดินทางของตนเองได้อย่างเหมาะสมมากยิ่งขึ้น ซึ่งเครื่องมือหลักที่ใช้ในการพัฒนาระบบจะเป็นการประยุกต์ใช้อุปกรณ์ไมโครคอนโทรลเลอร์ Arduino และเซนเซอร์ GPS ในการตรวจจับตำแหน่ง และส่งข้อมูลผ่านเว็บแอพพลิเคชั่นในการเก็บข้อมูลตำแหน่งของรถประจำทาง โดยผู้ที่ใช้โทรศัพท์เคลื่อนที่สามารถตรวจสอบตำแหน่งของรถประจำทางผ่านแอพพลิเคชั่นบนมือถือของตนเองได้

เอกสารอ้างอิง

[1] R.Ramani, S.Valarmathy, N.S. Vanitha, S.Selvaraju, and M.Thiruppathi, “Vehicle Tracking and Locking System Based on GSM and GPS”, *International Journal of Intelligent Systems and Applications*, Vol. 9, pp. 86-93, 2013.

[2] P. Verma and J.S Bhatia, “Design and Development of GPS-GSM based Tracking System with Google Map based Monitoring”, *International Journal of Computer Science, Engineering and Applications* (*IJCSEA*), Vol.3, No.3, 2013.

[3] S. Gharge, M. Chhaya, G. Chheda, J. Deshpande, and N. Gajra, “Real Time Bus Monitoring System Using GPS”, *International Journal of Engineering Science and Technology*, Vol.2, No. 3, 2012.

[4] M. Kumbhar, M. Survase, P. Mastud, A.Salunke, and S. Sirdeshpande, “Real Time Web Based Bus Tracking System”, *International Research Journal of Engineering and Technology* (*IRJET*), Vol. 3, Issue 2, 2016.

[5] V. Pandey, A. Jha, A. Tiwari, and V. Paarcha, “Real Time Bus Position and Time Monitoring System”, *International Journal of Science Technology & Engineering*, Vol. 1, Issue 10, 2015.

[6] S. Chandurkar, S. Mugade, S. Sinha, M. Misal, and P. Borekar, “Implementation of Real Time Bus Monitoring and Passenger Information System”, *International Journal of Scientific and Research Publications*, Vol. 3, Issue 5, 2013.

[7] K. Ganesh, M. Thrivikraman, J. Kuri, H. Dagale, G. Sudhakar, and S. Sanyal, , “Implementation of a Real Time Passenger Information System”, *Int. J. Engg. Sc. & Mgmt.*, Vol. 2, Issue 2, 2012.

[8] J.A. Stankovic, “Research Directions for the Internet of Things”, *IEEE Internet of Things Journal*, pp 3-9, 2014.

[9] D.K. Rath, “Arduino Based: Smart Light Control System”, *International Journal of Engineering Research and General Science*, Vol. 4, Issue 2, 2016.

[10] M.S. Zaghloul, “GSM-GPRS Arduino Shield (GS-001) with SIM 900 chip module in wireless data transmission system for data acquisition and control of power induction furnace”, *International Journal of Scientific & Engineering Research*, Vol. 5, Issue 4, 2014.

[11] C. Joseph, A.D. Ayyappan, A.R. Aswini, and B.D. Bharathy, “GPS/GSM Based Bus Tracking System (BTS)”, *International Journal of Scientific & Engineering Research*, Vol. 4, Issue 12, pp. 176-179, 2014.

[12] B. Mustapha, A. Zayegh, and R.K. Begg, “Ultrasonic and Infrared Sensors Performance in a Wireless Obstacle Detection System”, *The 1st International Conference on Artificial Intelligence, Modelling & Simulation*, 2013.

[13] M. Konarski and W. Zabierowski, “Using Google Maps API along with technology .NET”, *The International Conference on Modern Problems of Radio Engineering, Telecommunications and Computer Science* (*TCSET*), 2010.

เกตเวย์เพื่อการเชื่อมโยงเครือข่ายเฉพาะกิจสมาร์ทโฟนเข้ากับเครือข่ายอินเทอร์เน็ต
Gateway Application for Smartphone Ad hoc Network and Internet

วัชรธร หลายคุณากร (W. Laikhunakhorn)[1] ธวัชชัย โกวิชสกุล (T. Kovicsakul)[2]
ธนานพ ทองถาวร (T. Thongthavorn)[3] และสุเมธ ประภาวัต (S. Prabhavat)[4]
คณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้า เจ้าคุณทหารลาดกระบัง กรุงเทพฯ
[1]Watcharathorn.l@gmail.com, [2]Mos_12_5@hotmail.com, [3]Thananop@it.kmitl.ac.th, [4]Sumet@it.kmitl.ac.th

บทคัดย่อ

ปัจจุบันการติดต่อสื่อสารมีบทบาทสำคัญต่อการดำเนินชีวิต ซึ่งเครือข่ายอินเทอร์เน็ตเป็นช่องทางหนึ่งที่ใช้เพื่อแลกเปลี่ยนข้อมูลข่าวสารได้ ทั้งในสถานการณ์ปกติและสถานการณ์ฉุกเฉิน แต่เมื่อเกิดเหตุการณ์ที่โครงข่ายสื่อสารพื้นฐานได้รับความเสียหาย เช่น ในสถานการณ์ภัยพิบัติ ผู้ประสบเหตุในพื้นที่ดังกล่าว จะไม่สามารถติดต่อสื่อสารและเข้าถึงเครือข่ายอินเตอร์เน็ตได้ ดังนั้น การมีเครือข่ายสื่อสารที่ไม่ต้องใช้โครงสร้างพื้นฐาน เป็นช่องทางสำรองสำหรับการติดต่อสื่อสารและเข้าถึงเครือข่ายอินเทอร์เน็ตจึงเป็นสิ่งที่จำเป็น บทความนี้จึงนำเสนอแนวทางการพัฒนาแอปพลิเคชันสำหรับการติดต่อสื่อสารภายในเครือข่ายสมาร์ตโฟน และระหว่างเครือข่ายสมาร์ตโฟนกับเครือข่ายอินเทอร์เน็ต โดยประยุกต์ใช้เทคโนโลยีเครือข่ายเฉพาะกิจเคลื่อนที่และเทคโนโลยีเกตเวย์เชื่อมโยงระหว่างกัน

คำสำคัญ: อินเทอร์เน็ตเกตเวย์ สมาร์ตโฟนแอปพลิเคชัน เครือข่ายเฉพาะกิจสมาร์ตโฟน

Abstract

Communication plays an important role in today's life as it is one of fundamentals of living. Internet is a source of information for people in everyday life and crisis situations. However, the communication and internet access services can be disrupted due to failures in network infrastructure, for example in disaster areas. This paper proposes an application of Mobile ad-hoc network (MANET) requiring no network infrastructure.

We develop a mobile application providing communication service and Internet connection in smartphone ad-hoc networks based on MANET and integrated Internet-MANET.

Keyword: Internet Gateway, Smart Phone Application, Ad-hoc Network

1. บทนำ

ปัจจุบันเทคโนโลยีเพื่อการสื่อสารมีความก้าวไกลและมีบทบาทสำคัญในชีวิตเราเป็นอย่างมาก แต่เทคโนโลยีเหล่านี้ส่วนใหญ่ล้วนแต่อาศัยโครงสร้างพื้นฐาน (Infrastructure) หากโครงสร้างพื้นฐานเกิดขัดข้องจะส่งผลต่อการติดต่อสื่อสารในวงกว้าง คณะผู้จัดทำจึงเสนอระบบเครือข่ายเฉพาะกิจสมาร์ทโฟน (Mobile Ad-hoc Network: MANET) เพื่อเป็นเครือข่ายสำรองเมื่อโครงสร้างพื้นฐานไม่สามารถใช้งานได้

MANET [1] เป็นเครือข่ายเฉพาะกิจที่ไม่อาศัยโครงสร้างพื้นฐานในการทำงาน ซึ่งทำให้เทคโนโลยีดังกล่าวมีความเหมาะสมที่จะนำไปใช้งานเป็นช่องทางการสื่อสารสำรองในสถานการณ์ที่โครงสร้างพื้นฐานไม่สามารถใช้งานได้ ยกตัวอย่างเช่นเหตุการณ์ภัยพิบัติเป็นต้น อย่างไรก็ตามช่องทางแลกเปลี่ยนข้อมูลข่าวสารบนเครือข่ายอินเทอร์เน็ตนั้นยังคงมีประสิทธิภาพมากกว่าในด้านการเรียกใช้งานบริการต่าง ๆ ซึ่งเทคโนโลยีในปัจจุบันยังไม่สนับสนุนการเชื่อมต่อระหว่างเครือข่ายเฉพาะกิจกับเครือข่ายอินเทอร์เน็ต

คณะผู้จัดทำจึงพัฒนากลไกเชื่อมโยงเครือข่ายเฉพาะกิจสมาร์ตโฟนเข้ากับเครือข่ายอินเทอร์เน็ตบนอุปกรณ์สมาร์ตโฟน

2. ทฤษฎีหรือแนวคิดที่เกี่ยวข้อง

2.1 เทคโนโลยีเครือข่ายเฉพาะกิจเคลื่อนที่ (Mobile Ad-hoc Network: MANET)

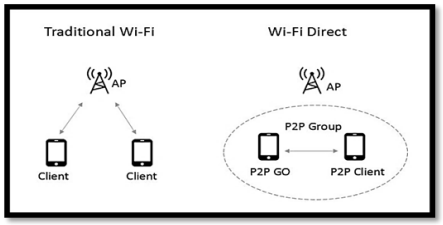
ในปัจจุบันเทคโนโลยีที่ถูกนำมาใช้เพื่อสนับสนุนการทำงานบนเครือข่ายเฉพาะกิจเคลื่อนที่บนสมาร์ตโฟนได้แก่

2.1.1 Wi-Fi Direct (Wi-Fi Peer-to-Peer)

Wi-Fi Direct [2] ทำงานคล้าย Wi-Fi ในแบบดั้งเดิม เป็นเทคโนโลยีที่มีในอุปกรณ์สมาร์ตโฟนแทบทุกเครื่องและ

สามารถนำมาประยุกต์ใช้ในการสร้างเครือข่ายเฉพาะกิจสมาร์ตโฟน

ลักษณะการทำงานคือการทำให้สมาร์ตโฟนเชื่อมต่อกันแบบโหนดต่อโหนด (Peer-to-Peer) ที่สำคัญคือเทคโนโลยี Wi-Fi Direct ไม่จำเป็นต้องเชื่อมต่อผ่าน Access point (AP) ทางกายภาพแต่จะใช้ซอฟต์แวร์ AP ในการเชื่อมต่อแทน โดยการเชื่อมต่อจะถูกจัดเป็นกลุ่มๆ ในแต่ละกลุ่มจะมีอุปกรณ์ที่ทำหน้าที่เป็น Group Owner (GO) ทำหน้าที่เป็น AP และยังเป็นเครื่องที่คอยแจกจ่าย IP ให้กับ P2P Client



ภาพที่ 2.1: ภาพเปรียบเทียบการทำงานระหว่าง Wi-Fi แบบดั้งเดิมและ Wi-Fi Direct

2.1.2 กลไกการค้นหาเส้นทางของเครือข่ายเฉพาะกิจเคลื่อนที่

โพรโทคอลการค้นหาเส้นทางของเครือข่ายเฉพาะกิจเคลื่อนที่สามารถแบ่งกลุ่มตามลักษณะการทำงานของโพรโทคอลได้เป็น 2 กลุ่มคือ

1) โพรโทคอลค้นหาเส้นทางแบบโพรแอคทีฟ

เป็นการเก็บข้อมูลเส้นทางการส่งทั้งหมดไว้ในตารางเส้นทาง และมีการปรับปรุงตารางตามเวลาที่กำหนด ทำให้สามารถส่งข้อมูลไปยังโหนดปลายทางได้ทันทีเมื่อต้องการ ตัวอย่างเช่น DSDV [3] เป็นต้น

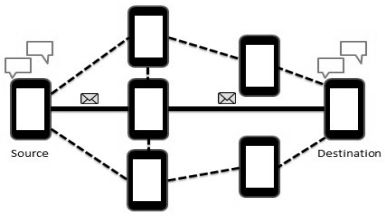
2) โพรโทคอลค้นหาเส้นทางแบบรีแอคทีฟ

เป็นกลไกที่จะทำงานก็ต่อเมื่อมีความต้องการจะส่งข้อมูลเท่านั้น ไม่มีการเก็บตารางเส้นทางไว้ล่วงหน้า แต่จะค้นหาเส้นทางทันทีเมื่อต้องการส่งข้อความจากโหนดต้นทางไปยังโหนดปลายทาง ตัวอย่างเช่น DSR [4] และ AODV [5]

2.2 โครงงานการศึกษาประสิทธิภาพกลไกค้นหาเกตเวย์บนเครือข่ายเฉพาะกิจเคลื่อนที่

จากการศึกษาข้อมูลจากงานวิจัยเครือข่ายติดต่อสื่อสารเฉพาะกิจบนสมาร์ตโฟนจากปริญญานิพนธ์คณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง ประจำปีการศึกษา 2558 [6] เป็นงานวิจัยที่ศึกษาทฤษฎีพื้นฐาน และประสิทธิภาพทางกายภาพของเครือข่ายเฉพาะกิจเคลื่อนที่บนสมาร์ตโฟน และพัฒนาออกมาในรูปแบบแอปพลิเคชันโปรแกรมแชทบนระบบปฏิบัติการแอนดรอยด์

การทำงานของแอปพลิเคชันคือ เมื่อมีสมาร์ตโฟนเปิดใช้งานแอปพลิเคชันอย่างน้อย 2 เครื่อง แอปพลิเคชันสามารถรับและส่งข้อความภายในเครือข่ายเฉพาะกิจสมาร์ตโฟนได้ภายในรัศมีที่กำหนด ยังไม่สามารถส่งข้อความออกไปยังเครือข่ายอื่นๆได้ดังภาพที่ 2.2



ภาพที่ 2.2: แสดงการทำงานของแอปพลิเคชันจากงานวิจัยที่ศึกษา

2.3 กลไกการค้นหาเกตเวย์ (Gateway Discovery)

การเชื่อมโยงระหว่างเครือข่ายเฉพาะกิจกับเครือข่ายอินเทอร์เน็ตต้องอาศัยกลไกเกตเวย์เป็นตัวกลางระหว่างสองเครือข่ายนี้ โดยปกติแล้วเครือข่ายเฉพาะกิจนั้นจะไม่มีกลไกค้นหาโหนดเกตเวย์ แต่ในปัจจุบันมีงานวิจัยที่ได้ศึกษาและพัฒนากลไกค้นหาเกตเวย์โดยจำแนกได้เป็น 2 โพรโทคอลย่อยๆ ได้ดังต่อไปนี้

2.3.1 การค้นหาเกตเวย์แบบรีแอคทีฟ

การค้นหาเกตเวย์แบบรีแอคทีฟ [7][8] เป็นการค้นหาเกตเวย์จากโหนดต้นทางโดยโหนดต้นทางจะสร้างข้อความ Route Request (RREQ) และกระจายไปยังโหนดข้างเคียงทั้งหมด เมื่อโหนดข้างเคียงได้รับ แต่ไม่มีโหนดใดเป็นเกตเวย์ ก็จะไม่ตอบกลับข้อความไปยังโหนดต้นทาง และทำการกระจาย

ข้อความ RREQ ไปยังโหนดข้างเคียงลักษณะเดิม จนเมื่อพบโหนดที่เป็นเกตเวย์ เกตเวย์จะตอบกลับด้วยข้อความ Route Reply

2.3.2 การค้นหาเกตเวย์แบบโพรแอคทีฟ

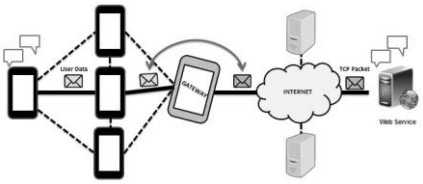
การค้นหาเกตเวย์แบบโพรแอคทีฟ [7] เป็นการค้นหาเกตเวย์โดย จะเริ่มที่โหนดเกตเวย์ทำการกระจายข้อความ Gateway Advertisement (GWADV) เป็นช่วง ๆ ให้กับโหนดที่อยู่ข้างเคียง เมื่อโหนดข้างเคียงได้รับข้อความ GWADV ก็จะทำการกระจายต่อๆ ไป ในข้อความ GWADV นั้นจะระบุเส้นทางไปยังโหนดเกตเวย์ โดยแต่ละโหนดข้างเคียงต้องตัดสินในการรับข้อมูล หากโหนดข้างเคียงยังไม่มีข้อมูลเส้นทางไปยังเกตเวย์ก็จะรับข้อมูล และเก็บบันทึกลงในตารางเส้นทางของโหนดนั้นๆ

3. แนวคิดและการออกแบบแอปพลิเคชัน

จากงานวิจัยก่อนหน้า [3] ผู้วิจัยจึงมีแนวคิดที่จะพัฒนาเครือข่ายเฉพาะกิจเคลื่อนที่ให้สามารถส่งข้อความออกไปยังเครือข่ายอื่นๆ เช่นเครือข่ายอินเทอร์เน็ต โดยอาศัยกลไกเกตเวย์เป็นตัวกลางระหว่างสองเครือข่ายโดยมีรายละเอียดดังนี้

3.1 การออกแบบระบบการเชื่อมโยงเครือข่ายเฉพาะกิจสมาร์ทโฟนกับเครือข่ายอินเตอร์เน็ต

หลักการคือ เมื่อโหนดต้นทางภายในเครือข่ายเฉพาะกิจต้องการส่งข้อความไปยังโหนดปลายทาง โหนดต้นทางจะไม่ทราบว่าปลายทางนั้นอยู่ภายในเครือข่ายเฉพาะกิจหรืออยู่นอกเครือข่ายเฉพาะกิจ โหนดต้นทางจะทำการส่งข้อความไปยังโหนดปลายทางเสมือนโหนดปลายทางอยู่ในเครือข่ายเฉพาะกิจ แต่หากไม่มีการตอบกลับมา (Timeout) ภายในเวลา 30 วินาที โหนดต้นทางจะเข้าใจว่าโหนดปลายทางอยู่นอกเครือข่ายเฉพาะกิจ และส่งข้อความดังกล่าวไปหาเกตเวย์โดยใช้กลไกค้นหาเกตเวย์ (Gateway Discovery) ในการค้นหาเส้นทางที่สามารถใช้งานได้ เกตเวย์ที่เชื่อมต่ออินเทอร์เน็ตอยู่จะส่งต่อข้อความดังกล่าวไปยังโหนดปลายทางที่อยู่บนเครือข่ายอินเตอร์เน็ตดังภาพที่3.1



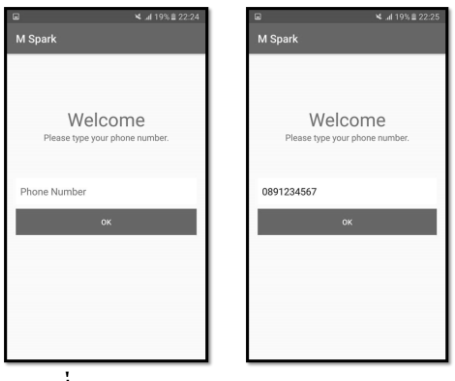
ภาพที่ 3.1: แสดงการทำงานแอปพลิเคชันของแอปพลิเคชันใหม่

ในการออกแบบแอปพลิเคชันผู้วิจัยใช้กลไกค้นหาเส้นทาง AODV เนื่องจากงานวิจัยก่อนหน้า[5] กล่าวว่า AODV เป็นโพรโทคอลรีแอคทีฟที่จะทำงานก็ต่อเมื่อโหนดต้นทางต้องการส่งข้อความไปยังโหนดปลายทางเท่านั้น เหตุดังกล่าวทำให้เกิดโอเวอร์เฮดต่ำและสิ้นเปลืองแบนด์วิดท์น้อยกว่าเมื่อเทียบกับโพรโทคอลอื่น

จากการศึกษางานวิจัยการศึกษาประสิทธิภาพกลไกค้นหาเกตเวย์บนเครือข่ายเฉพาะกิจเคลื่อนที่ [9] ซึ่งเปรียบเทียบประสิทธิภาพระหว่างกลไกค้นหาเกตเวย์แบบรีแอคทีฟและโพรแอคทีฟ ทั้ง 2 แบบมีข้อดีและข้อเสียแตกต่างกันทำให้ผู้วิจัยออกแบบแอปพลิเคชันให้ใช้งานได้ทั้งแบบรีแอคทีฟและโปรแอคทีฟตามความเหมาะสมโดยสามารถตั้งค่ากลไกค้นหาเกตเวย์ของเครื่องที่ทำหน้าที่เป็นเกตเวย์ได้จากปุ่ม Setting ภายในแอปพลิเคชัน

3.2 การออกแบบส่วนติดต่อผู้ใช้

เมื่อเปิดแอปพลิเคชันนี้ขั้นตอนแรกเป็นการเข้าสู่ระบบโดยผู้ใช้จะต้องกรอกเบอร์โทรศัพท์เพื่อใช้งานเป็น Username ดังตัวอย่างภาพที่ 3.2



ภาพที่ 3.2: แสดงหน้าเข้าสู่ระบบของแอปพลิเคชัน

เมื่อเข้าสู่ระบบเรียบร้อยแล้วจะเป็นโปรแกรมแชตลักษณะดังภาพที่ 3.3

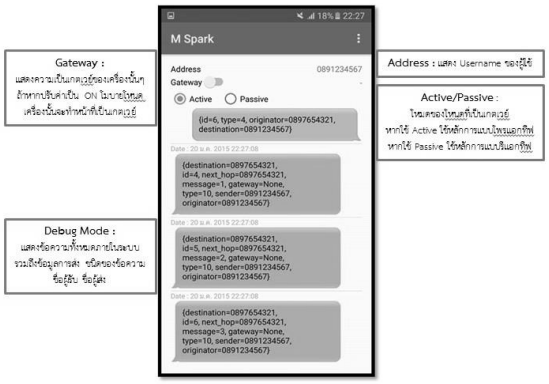


ภาพที่ 3.3: แสดงการรับ-ส่งข้อความในโปรแกรมแชต

หากต้องการเข้าไปตั้งค่าให้โปรแกรมแชต ให้กดปุ่มที่อยู่มุมบนขวาดังภาพที่ 3.4 และภาพที่ 3.5



ภาพที่ 3.4: แสดงปุ่มตั้งค่า (Settings) ในโปรแกรมแชต



ภาพที่ 3.5: แสดงการทำงานภายในของการตั้งค่าแชต

- Gateway แสดงความเป็นเกตเวย์ของเครื่องนั้น ถ้าเปิดโมบายโหนดนั้นจะทำหน้าที่เป็นเครื่องเกตเวย์

- Active/Passive เป็นรูปแบบของเครื่องที่เป็นเกตเวย์ (กลไกค้นหาเกตเวย์)

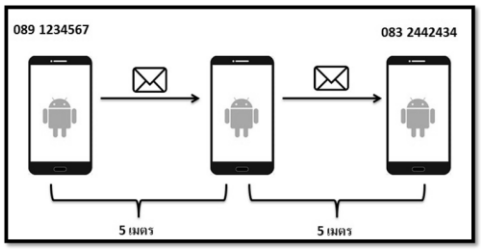
- Active ให้เกตเวย์ทำงานแบบโพรแอกทีฟ คือเกตเวย์จะประกาศ Gateway Advertisement ทำให้โมบายโหนดอื่นๆ ทราบว่าตนเป็นเกตเวย์

- Passive ให้เกตเวย์ทำงานแบบรีแอกทีฟคือเกตเวย์จะรอการส่งข้อความ Gateway Request เข้ามาจากโมบายโหนดอื่นๆ และตอบกลับด้วย Gateway Reply เพื่อยืนยันเส้นทาง

4. การประเมินและวิเคราะห์การทำงาน

การทดสอบใช้งานแอปพลิเคชันจะแบ่งออกเป็น 2 ส่วนคือ การรับ-ส่งข้อความภายในเครือข่ายเฉพาะกิจ และการรับ-ส่งข้อความโดยเชื่อมต่อไปยังเครือข่ายอินเตอร์เน็ตผ่านเกตเวย์

1. การรับและส่งข้อความภายในเครือข่ายเฉพาะกิจสมาร์ตโฟน ทดสอบโดยใช้สมาร์ตโฟน 3 เครื่องโดยใช้เบอร์โทรศัพท์มือถือเป็น Username ทดสอบโดยทำการส่งข้อความจากโมบายโหนดซ้ายสุด (089-1234567) ไปยังเครื่องขวาสุด (083-2442434) โดยอาศัยกลไกมัลติฮอปให้โมบายโหนดเครื่องตรงกลางเป็นทางผ่านส่งต่อข้อความไปยังโมบายโหนดขวาสุด



ภาพที่ 4.1: แสดงการจำลองการทดสอบการรับส่งข้อความ

ผลการทดสอบ

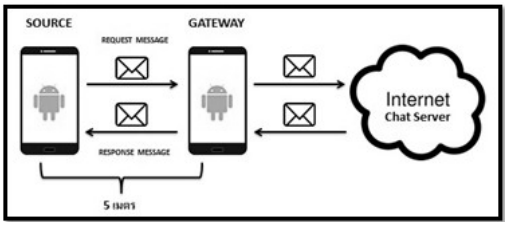
ตารางที่ 1: แสดงตัวเลขผลการทดสอบการส่งข้อความภายในเครือข่ายเฉพาะกิจสมาร์ตโฟน

	ระยะห่าง 1 เมตร/เครื่อง	ระยะห่าง 5 เมตร/เครื่อง	ระยะห่าง 10 เมตร/เครื่อง
ส่งข้อความครั้งแรก	23 s.	25 s.	30 s.
ส่งข้อความครั้งที่สอง	10 s.	11 s.	11 s.
ส่งข้อความครั้งที่สาม	8 s.	9 s.	7 s.
ส่งข้อความครั้งที่สี่	5 s.	6 s.	10 s.

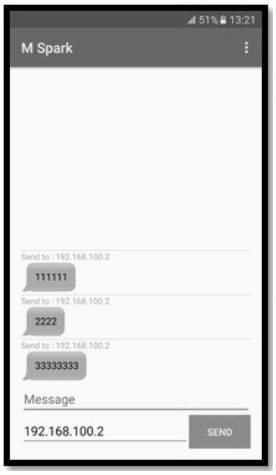
สรุปได้ว่า การรับส่งข้อความภายในเครือข่ายเฉพาะกิจส่วนมาก จะกินเวลาในการส่งข้อความครั้งแรกประมาณ 20-30 วินาที เนื่องจากแต่ละโหนดต้องทำการค้นหาเส้นทางใหม่ ยังไม่มีบันทึกไว้ ทำให้ใช้เวลานาน ในการรับส่งข้อความครั้งต่อๆ ไป จะใช้เวลาประมาณ 5-10 วินาที ซึ่งน้อยกว่าครั้งแรกเพราะมีเส้นทางบันทึกไว้ในตารางเก็บเส้นทางเรียบร้อยแล้ว

2. การเชื่อมต่อไปยังเครือข่ายอินเทอร์เน็ตผ่านเกตเวย์ ทดสอบโดยใช้สมาร์ตโฟน 2 เครื่อง และ แชตเซิร์ฟเวอร์ 1 เครื่อง ซึ่งแชตเซิร์ฟเวอร์เป็นเซิร์ฟเวอร์เฉพาะที่ถูกตั้งขึ้นเพื่อรอรับข้อความจากเครือข่ายเฉพาะกิจผ่านเครือข่ายอินเทอร์เน็ต

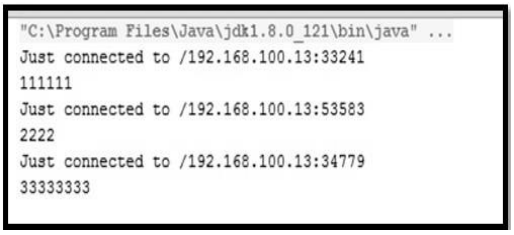
สมาร์ตโฟน 2 เครื่องจะมีเครื่องหนึ่งเป็นโหนดต้นทางที่อยู่ในเครือข่ายเฉพาะกิจและอีกเครื่องหนึ่งเป็นโหนดเกตเวย์ที่เชื่อมต่ออินเทอร์เน็ตอยู่ โดยโหนดต้นทางจะส่งข้อความไปยังแชตเซิร์ฟเวอร์ที่ออนไลน์อยู่บนเครือข่ายอินเทอร์เน็ตโดยอาศัยกลไกเกตเวย์ในการเชื่อมต่อ หากส่งข้อความไปถึงแชตเซิร์ฟเวอร์ได้ ข้อความจะแสดงบนหน้าจอของแชตเซิร์ฟเวอร์ โดยการทดสอบสามารถดูได้จากภาพที่ 4.2, 4.3 และ 4.4



ภาพที่ 4.2: แสดงการจำลองการทดสอบเชื่อมต่อไปยังแชตเซิร์ฟเวอร์บนเครือข่ายอินเทอร์เน็ต



ภาพที่ 4.3: แสดงการส่งข้อความจากโหนดต้นทางไปหาแชตเซิร์ฟเวอร์ 192.168.100.2



ภาพที่ 4.4: แสดงหน้าจอของแชตเซิร์ฟเวอร์เมื่อได้รับข้อความจากโหนดต้นทาง

ผลการทดสอบ

ตารางที่ 2: แสดงตัวเลขผลการทดสอบการส่งข้อความไปยังแชตเซิร์ฟเวอร์บนเครือข่ายอินเทอร์เน็ต

	ระยะห่าง 1 เมตร/เครื่อง	ระยะห่าง 5 เมตร/เครื่อง	ระยะห่าง 10 เมตร/เครื่อง
ส่งข้อความครั้งแรก	68 s.	60 s.	85 s.
ส่งข้อความครั้งที่สอง	62 s.	141 s.	116 s.
ส่งข้อความครั้งที่สาม	110 s.	62 s.	178 s.
ส่งข้อความครั้งที่สี่	101 s.	102 s.	107 s.

สรุปได้ว่า การส่งข้อความไปยังแชตเซิร์ฟเวอร์โดยมีเกตเวย์เป็นตัวเชื่อมระหว่างเครือข่ายเฉพาะกิจ และเครือข่ายอินเทอร์เน็ต สามารถทำได้ ไม่เกิดข้อความสูญหาย แต่พบปัญหาด้านเวลาในการส่ง จากตารางจะเห็นได้ว่าใช้เวลาส่งมากกว่าการส่งข้อความภายในเครือข่ายเฉพาะกิจ เนื่องจากต้อง

รอ 30 วินาทีจน Timeout จึงจะใช้กลไกค้นหาเกตเวย์ และมีความไม่แน่นอนของเวลาในการส่ง โดยมีสาเหตุหลักมาจากความแปรปรวนสะสมของเวลาที่ใช้ในการค้นหาเกตเวย์และเวลาที่ใช้ในการเชื่อมต่อกับแชตเซิร์ฟเวอร์ผ่านเครือข่ายอินเทอร์เน็ต

5. บทสรุป

แอปพลิเคชันนี้พัฒนาขึ้นเพื่อสร้างเครือข่ายเฉพาะกิจให้ใช้งานบนสมาร์ทโฟน และเป็นเครือข่ายที่ใช้ติดต่อสื่อสารแลกเปลี่ยนข้อมูลได้ โดยไม่อาศัยโครงสร้างพื้นฐานเหมาะที่จะใช้เป็นเครือข่ายสำรองเมื่อเกิดเหตุการณ์ฉุกเฉิน โดยแอปพลิเคชันนี้สามารถติดตั้งในสมาร์ทโฟนบนระบบปฏิบัติการแอนดรอยด์ตั้งแต่เวอร์ชั่น 4.1 (Jelly Bean) เป็นต้นไป

หลังจากการทดสอบประสิทธิภาพของเครือข่ายพบว่า การรับ-ส่งข้อความภายในเครือข่ายสามารถทำได้ เมื่อมีเครื่องที่เปิดใช้งานแอปพลิเคชันอยู่ภายในรัศมีของเครือข่ายซึ่งจากการทดสอบมีระยะรัศมีสูงสุดประมาณ 70 เมตร และใช้เวลาในการรับ-ส่งข้อความในครั้งแรกประมาณ 20-35 วินาที และข้อความต่อๆไปประมาณ 5-12 วินาที เหตุที่ครั้งแรกใช้เวลานานเพราะเครื่องต้นทางและเครื่องปลายทางยังไม่มีการบันทึกเส้นทางเก็บไว้ในตารางเก็บเส้นทาง

การส่งข้อความจากเครือข่ายเฉพาะกิจไปยังเครือข่ายอินเทอร์เน็ตผ่านกลไกเกตเวย์สามารถทำได้ แต่เวลาในการส่งข้อความยังไม่เสถียรจากการทดสอบจะอยู่ในช่วง 1-3 นาที สาเหตุจากความแปรปรวนสะสมของเวลาที่ใช้ในการค้นหาเกตเวย์และเวลาที่ใช้ในการเชื่อมต่อกับแชตเซิร์ฟเวอร์ผ่านเครือข่ายอินเทอร์เน็ตและปัจจัยภายนอกอื่นๆเช่น ความเร็วของอินเทอร์เน็ต, สัญญาณรบกวน ที่เครือข่ายเฉพาะกิจไม่สามารถควบคุมได้ เป็นต้น

กิตกรรมประกาศ

งานวิจัยนี้ได้รับการสนับสนุนทุนวิจัยงบประมาณแผ่นดิน ปีงบประมาณ 2559 จากสถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง ผ่านการพิจารณาจากสำนักงานคณะกรรมการวิจัยแห่งชาติ (วช.)

เอกสารอ้างอิง

[1] Sunil taneja and Ashwani kush. “A survey of routing protocols in mobile ad hoc networks”, *Proc. International journal of innovation management and technology*, vol. 1, no.3, august, 2010.

[2] Jimmy Tieu and Sihan Ye. “Wi-Fi Direct Services” Master’s Thesis Department of Electrical and Information Technology, Faculty of Engineering, LTH, Lund University. June 201

[3] วาริส จันอิ, สกุณา เจริญปัญญาศักดิ์, “การวิเคราะห์เชิงเปรียบเทียบในทอพอโลยีและโปรโทรคอลการค้นหาเส้นทางสำหรับเครือข่ายไร้สายแบบเฉพาะกิจ”, ภาควิชาวิศวกรรมคอมพิวเตอร์, คณะวิศวกรรมศาสตร์, มหาวิทยาลัยสงขลานครินทร์

[4] D.B. Johnson and D.A. Maltz, “Dynamic Source Routing in Ad Hoc Wireless Networks.” Computer Science Department Carnegie Mellon University. 1996

[5] C.E. Perkins, E. M. Belding-Royer, and S. R. Das, “Ad hoc On-Demand Distance Vector (AODV) Routing for Internet Connectivity,” *Radio Science Conference*, 2007.NRSC 2007.National, Cairo, 2007.

[6] กฤชนิพันธ์ เดิมธรณินทร์, อรุณกมล ธรรมโกฎิ์ และ สุเมธ ประภาวัต, “เครือข่ายติดต่อสื่อสารเฉพาะกิจบนสมาร์ทโฟน” การประชุมวิชาการระดับชาติด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศ ครั้งที่ 7, 2015.

[7] H. El-Moshrify, M. A. Mangoud and M. Rizk, “Gateway Discovery in Ad hoc On-Demand Distance Vector (AODV) Routing for Internet Connectivity,*" Radio Science Conference*, 2007. NRSC 2007. National, Cairo, 2007, pp. 1-8.

[8] Matthias Rosenschon. “Internet Gateway Discovery for Mobile Ad Hoc Networks” Submitted for the Degree of Doctor of PhilosophySchool of Engineering and Mathematical Sciences, CITY University London. 2 October 2007

[9] ธเนศ พุ่มไพจิตร, วรรณิศา กำมะหยี่, “การศึกษาประสิทธิภาพกลไกค้นหาเกตเวย์บนเครือข่ายเฉพาะกิจ” การประชุมวิชาการระดับชาติด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศ ครั้งที่ 8, 2016.

กระบวนการเลือกเส้นทางของทีซีพีแบบหลายเส้นทาง ในเครือข่ายที่กำหนดโดยซอฟต์แวร์

Path Selection Algorithm for Multipath TCP in Software-Defined Networking

จิรศักดิ์ จุลวัจน์ (Jirasak Jullawat) เกียรติคุณ กาวิละ (Kiattikun Kawila)
*กุลิสร์ ณ นคร (Kulit Na Nakorn) และกุลธิดา โรจน์วิบูลย์ชัย (Kultida Rojviboonchai)**
ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
{Jirasak.ju, Kiattikun.ka}@student.chula.ac.th, {Kulit.n, Kultida.r}@chula.ac.th

บทคัดย่อ

ความสามารถของ Multipath TCP (MPTCP) ที่เพิ่ม Throughput และเป็น Load balance ซึ่งคงทนต่อความเสียหาย (Fault Tolerance) จากเส้นทางเชื่อมต่อที่มีปัญหา แต่ด้วยเครือข่ายที่มีความซับซ้อนในปัจจุบัน ทำให้การที่ส่งข้อมูลโดยใช้ MPTCP เกิดการใช้เส้นทางร่วมกัน แต่หากเป็นการส่งข้อมูลภายใต้เครือข่ายที่กำหนดโดยซอฟต์แวร์ที่ทราบถึงลักษณะของข้อมูลที่ถูกส่งภายในระบบ และสามารถเขียนโปรแกรมควบคุมการทำงานได้ทั้งระบบนั้น จะช่วยเลือกเส้นทางที่เหมาะสมได้ งานวิจัยนี้จึงเสนอกระบวนการระบุ Subflow ที่ถูกต้องแม่นยำและกระบวนการเลือกเส้นทางที่ใช้ข้อมูลแบนด์วิดท์ที่พร้อมใช้งานแบบเรียลไทม์และค่าความหน่วง เพื่อประมาณค่า Throughput ในการเลือกคู่เส้นทางที่จะได้ค่า Throughput สูงสุดจากเส้นทางทั้งหมดในเครือข่าย ผลลัพธ์จากการวัดประสิทธิภาพในการทดลองนี้พบว่าสามารถส่งข้อมูลที่ได้ค่า Throughput ที่สูงกว่าการเลือกเส้นทางแบบสุ่ม 76 เปอร์เซ็นต์

คำสำคัญ: ทีซีพีแบบหลายเส้นทาง โอเพนโฟลว เครือข่ายที่กำหนดโดยซอฟต์แวร์ แบนด์วิดท์ ค่าหน่วง

Abstract

The features of Multipath TCP (MPTCP) are throughput improvement, load balancing and fault tolerance from used link is down. However, the current network is complicated. Therefore, data transmission will share the same path. Then, if we use MPTCP in Software-Defined Networking which knows the characteristics of the packet and also allow for programming to control the path selection. These processes may use for select the suitable paths. This study proposed the accurate subflow identification algorithm and path selection algorithm which use real-time traffic and delay for estimating throughput. The results of performance evaluation show the throughput was 76% improved than the random path selection.

Keyword: Multipath-TCP, OpenFlow, Software-Defined Networking, Bandwidth, Delay

1. บทนำ

ในปัจจุบันนี้ คอมพิวเตอร์ มือถือ แท็บเล็ต รวมไปถึงเครื่องเซิร์ฟเวอร์มักจะมีส่วนต่อขยาย (Network Interface) จำนวน 2 ส่วนต่อขยายหรือมากกว่านั้น อุปกรณ์เครือข่ายและโครงสร้างพื้นฐานของระบบเครือข่ายสามารถรองรับการส่งข้อมูลพร้อมกันหลายเส้นทางจากอุปกรณ์หนึ่งไปยังอุปกรณ์หนึ่งได้ การส่งข้อมูลโดยมากจะใช้โพรโทคอลที่มีชื่อว่า Transmission Control Protocol (TCP) [1] แต่อย่างไรก็ตาม TCP สามารถทำได้เพียงแค่ส่งข้อมูลในเส้นทางเดียวเท่านั้น ดังนั้น TCP ในปัจจุบันจึงไม่ได้ใช้ประโยชน์จากการที่มีเส้นทางการส่งข้อมูลที่มากขึ้น และด้วยเหตุนี้จึงมีผู้พัฒนา Multipath TCP ขึ้น

Multipath TCP (MPTCP) [2] คือ ส่วนต่อขยาย (Extension) ของ TCP ปกติ ที่จัดการควบคุมการส่งข้อมูลในหลายเส้นทาง ใน MPTCP หนึ่งการเชื่อมต่อจะสามารถสร้าง Subflow ย่อยได้ และแต่ละ Subflow สามารถใช้ในการส่ง

*Corresponding author

ข้อมูลผ่านเส้นทางหรือส่วนต่อขยายที่ต่างกันได้ จะเห็นว่า MPTCP สามารถใช้ประโยชน์ได้มากกว่า TCP หากการเชื่อมต่อของ TCP ถูกขัดขวางและสูญเสียการเชื่อมต่อ เมื่อต้องการเชื่อมต่ออีกครั้ง การเชื่อมต่อแบบ TCP ต้องการทำกระบวน Three-way Handshake ยิ่งไปกว่านั้น หากการเชื่อมถูกขัดขวางในขณะที่กำลังส่งข้อมูลที่มีขนาดใหญ่ ภายหลังการเชื่อมต่ออีกครั้งข้อมูลนี้อาจจะต้องเริ่มต้นส่งใหม่อีกครั้งหนึ่ง ดังนั้น MPTCP สามารถช่วยแก้ปัญหานี้ได้ โดยใช้เส้นทางอื่นที่ยังสามารถใช้งานได้และยังเชื่อมต่ออยู่ ภายหลังหากเส้นทางที่ถูกขัดขวางได้รับการแก้ไขแล้ว จะสามารถส่งข้อมูลในเส้นทางนี้ได้อีกครั้ง ในปัจจุบัน MPTCP ยังไม่ได้ถูกใช้อย่างแพร่หลาย แต่อย่างไรก็ตามระบบปฏิบัติการ iOS ได้มีการใช้ MPTCP และยินยอมให้ iPhone และ iPad สร้างการเชื่อมต่อสำรองผ่านเครือข่ายสัญญาณโทรศัพท์ [3]

ในทางทฤษฎีแล้ว MPTCP มีประสิทธิภาพที่ดี แต่เนื่องจากความซับซ้อนของระบบเครือข่ายและการเลือกเส้นทางแบบสั้นที่สุด (Shortest Path) ทำให้การเชื่อมต่อจากเครื่องต้นทางไปยังเครื่องปลายทางด้วยโพรโทคอล MPTCP อาจจะมีการใช้เส้นทางร่วมกันและเกิดคอขวด (Bottleneck) ในเครือข่ายขึ้นได้ แต่หากใช้ MPTCP ภายใต้เครือข่ายที่กำหนดโดยซอฟต์แวร์ ซึ่งทราบคุณลักษณ์แพ็คเก็ตของ MPTCP และยอมให้สามารถเขียนโปรแกรมเพื่อใช้ในการเลือกเส้นทาง น่าจะสามารถใช้เลือกเส้นทางที่เหมาะสมให้กับ MPTCP ได้

งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนากระบวนการเลือกเส้นทางของ MPTCP ภายใต้เครือข่ายที่กำหนดโดยซอฟต์แวร์หรือ Software-Defined Networking (SDN) ซึ่งเป็นเครือข่ายแห่งอนาคต [4] โดยใช้อุปกรณ์ควบคุมหรือ Controller เป็นส่วนควบคุมและทำงานผ่านโพรโทคอล OpenFlow [5] กระบวนการทำงานของ Controller จะศึกษาคุณลักษณ์ของแพ็คเกต MPTCP Header และจะจำแนกแต่ละการเชื่อมต่อของ MPTCP นอกจากนี้แล้วในการเลือกเส้นทางของแต่ละ Subflow จะใช้ข้อมูลจากค่าแบนด์วิดท์ที่พร้อมใช้งานแบบเรียลไทม์ (Real-time Available Bandwidth) และความหน่วง (Delay) ระหว่างต้นทางและปลายทางสำหรับการเลือกเส้นทางที่ดีที่สุด งานวิจัยนี้จะศึกษาบน Mininet Network Emulator [6] และ Floodlight Controller [7] ทำการทดลองบนโครงสร้างเครือข่ายที่มีการส่งข้อมูลแบบ UDP [8] ผสมด้วย (Background UDP Traffic) และได้ทำการวัดผลจากค่า Throughput โดยใช้โปรแกรม Iperf [9]

2. งานวิจัยที่เกี่ยวข้อง

หลังจากเครือข่าย SDN เริ่มได้รับความสนใจ และ Paasch และคณะ [10] ได้พัฒนา MPTCP ใน Linux Kernel ขึ้น หลังจากนั้น Pol และคณะ [11] ได้ใช้เครือข่าย SDN ในการจัดการส่งข้อมูล MPTCP โดยทำการกำหนดค่า VLAN สำหรับแต่ละส่วนต่อขยายของทั้งเครื่องต้นทางและเครื่องปลายทางก่อนที่จะส่งข้อมูล จำนวนของ VLAN ขึ้นกับจำนวน Subflow ของ MPTCP ซึ่ง Controller สามารถระบุ Subflow ได้จากหมายเลข VLAN ที่ถูกกำหนดไว้ใน Header ของแพ็คเก็ต และทำการกำหนดเส้นทางจากเส้นทางที่มีความสามารถส่งข้อมูลได้เยอะที่สุดตลอดเส้นทางระหว่างเครื่องต้นทางและเครื่องปลายทาง ซึ่งช่วยในการเพิ่ม Throughput อย่างไรก็ตาม ข้อจำกัดของกระบวนการนี้คือต้องทำการกำหนด VLAN และเส้นทางระหว่างเครื่องต้นทางและเครื่องปลายทางก่อนจึงจะสามารถใช้งานได้

ดังนั้นเพื่อลดความยุ่งยากหรือซับซ้อนในการตั้งค่า VLAN การศึกษาของ Sandri และคณะ [12] และ Zannettou และคณะ [13] ได้ศึกษาลักษณะของแพ็คเกต MPTCP Header และพบว่าสามารถแบ่งแยก Subflow แต่ละ Subflow ได้ ช่วยให้ Controller เลือกจัดการกับแพ็คเกตแต่ละ Subflow ซึ่งสามารถเพิ่ม Throughput ให้ข้อมูลที่ส่งอันเนื่องมาจากการเลือกส่งแพ็คเกตไปคนละเส้นทาง อย่างไรก็ตามงานวิจัยทั้งสองนี้ ไม่สามารถจำแนก Subflow จากหลายการเชื่อมต่อที่มาจากต้นกำเนิดเดียวกัน และส่งในเวลาใกล้เคียงกันได้ อีกทั้งกระบวนการเลือกเส้นทางยังเป็นแบบการเลือกเส้นทางที่สั้นที่สุดเท่านั้น ซึ่งอาจจะมีเส้นทางอื่นที่มีประสิทธิภาพในการส่งที่ดีกว่า

3. วิธีการดำเนินงานวิจัย

เพื่อให้ Controller สามารถควบคุมการส่งข้อมูลที่เป็น MPTCP ผ่านเครือข่าย SDN ได้อย่างถูกต้องนั้น ในการศึกษานี้จะแบ่งการทำงานออกเป็น 2 ส่วนย่อย ส่วนแรกระบุว่าแพ็คเก็ตที่เข้ามาเป็น Subflow ที่เท่าไร ส่วนย่อยที่สองคือ เลือก

เส้นทางที่เหมาะสมเพื่อส่งแพ็คเก็ตนั้น โดยงานวิจัยนี้จะนำเสนอกระบวนการทำงานของ Controller เป็น 2 ส่วน ดังนี้

3.1 กระบวนการระบุ Subflow ใน SDN

ในส่วนนี้จะอธิบายกระบวนการระบุ Subflow ใน SDN โดยจะเริ่มจากการอธิบายการทำงานของ MPTCP หลังจากนั้นเราจะอธิบายว่า MPTCP จะสามารถใช้ในเครือข่าย SDN ได้อย่างไร

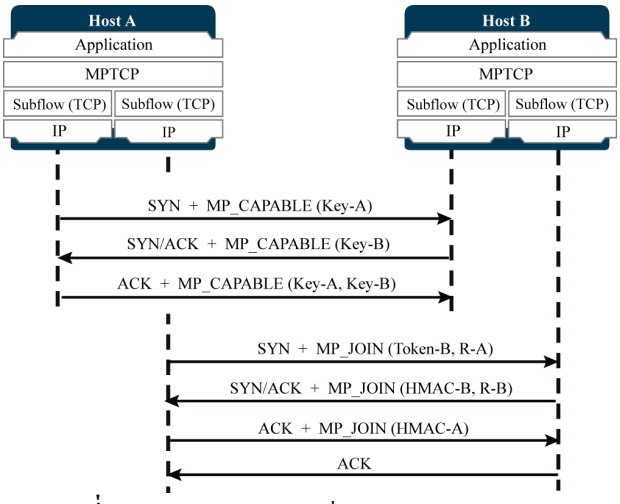
ตามภาพที่ 1 เมื่อเริ่มการเชื่อมต่อ MPTCP จะใช้กระบวนการ Three-way Handshake เหมือน TCP ปกติ ทั้งกระบวนการเริ่มต้น Flow หลัก Subflow รอง โดยที่ส่วนของ TCP Option จะมีการเพิ่มส่วนของ MP_CAPABLE ใน Flow หลัก MP_JOIN ใน Subflow อื่น ๆ ซึ่งหนึ่งในพารามิเตอร์ที่จำเป็นคือ Token ซึ่ง Token คือ Cryptographic Hash ของค่า Key ที่ถูกสร้างขึ้นและส่งไปให้เครื่องอีกฝั่งทราบ โดยงานวิจัยนี้เลือกใช้ค่า Token ที่เกิดจาก Key ของฝั่ง MPTCP Server (Host B) จุดเด่นของ Token คือในหนึ่งการเชื่อมต่อค่า Token จะเป็นค่าเดียวกันเสมอ ดังนั้นเราจึงใช้ประโยชน์ตรงนี้ เพื่อมาใช้ในการแบ่งแยก Subflow ที่อยู่ในการเชื่อมต่อ MPTCP เดียวกันได้

รายละเอียดของการคำนวณค่า Token ซึ่งได้กำหนดไว้ใน [2] ว่า Token จะถูกสร้างโดย SHA-1 [14] อัลกอริทึม และเลือกใช้แค่ 32 bits บนของผลลัพธ์ โดยสามารถเขียนได้ดังสมการที่ (1)

$$Token_{Sender} = Upper_{32}(SHA\text{-}1(Key_{Sender})) \qquad (1)$$

อัลกอริทึมที่เราเสนอนั้นใช้แนวคิดของ OpenFlow ที่แยกส่วนการส่งต่อแพ็คเก็ต (Data Plane) และส่วนการตัดสินใจ (Control Plane) ซึ่งเกิดขึ้นใน OpenFlow Switch ตัวเดียวกัน โดย Control Plane นี้จะถูกควบคุมด้วย Controller เมื่อมี Flow ของข้อมูลใหม่เข้ามาที่ OpenFlow Switch จากนั้น Switch จะส่งข้อมูลนั้น (PACKET_IN) ไปหา Controller หลังจากนั้น Controller จะจัดการกับ PACKET_IN โดยจะตรวจสอบข้อมูล Header ของแพ็คเก็ต

ความคิดพื้นฐานนี้ใช้จำแนกแต่ละ Subflow ของ MPTCP ซึ่ง MPTCP จะประกอบไปด้วยการเชื่อมต่อ TCP ปกติหลายการเชื่อมต่อ การศึกษาชิ้นนี้ได้สร้าง Mapping Table สำหรับเก็บข้อมูล โดยในแต่ละแถวคือแต่ละการเชื่อมต่อ TCP ค่า



ภาพที่ 1: แสดงโครงสร้างข้อมูลที่ส่งผ่านโปรโตคอล MPTCP

Token จะเป็นตัวเลขที่ถูกใส่ไว้ในแต่ละแถวและจะถูกใช้เพื่อจำแนกการเชื่อมต่อ TCP ที่มาจากการเชื่อมต่อ MPTCP เดียวกัน นอกจากนี้แต่ละแถวยังมีค่าตัวแปรที่ชื่อว่า Subflow ซึ่งไว้เก็บลำดับของแต่ละ Subflow

ภาพที่ 2 จะแสดงกระบวนการทำงานของอัลกอริทึมในแต่ละขั้นตอน และรูปที่ 3 แสดงตัวอย่างเมื่อมีการส่งข้อมูล MPTCP จำนวน 2 การเชื่อมต่อมาในเวลาเดียวกัน แต่ละการเชื่อมต่อมี 2 Subflows ในแผนภาพส่วนบนแสดงทิศทางของแพ็คเก็ตแต่ละขั้นตอน และตารางส่วนล่างแสดงถึงการเปลี่ยนแปลงของ Mapping Table ในแต่ละขั้นตอน การทำงานเริ่มจากขั้นตอนที่ 1 เมื่อแพ็คเก็ตแรกที่มาจากเครื่องต้นทาง (MPTCP Client) ไปยังเครื่องปลายทาง (MPTCP Server) เมื่อ Controller ได้รับแพ็คเก็ตใหม่ และตรวจพบว่าเป็น MPTCP (บรรทัดที่ 2) และพบว่าเป็น MP_CAPABLE (บรรทัดที่ 12) และแพ็คเก็ตนี้เป็น TCP SYN (บรรทัดที่ 13) Controller จะค้นหาเส้นทางให้แพ็คเก็ตนี้ (บรรทัดที่ 14) ไว้สำหรับเพิ่มไปใน Flow Table ของ OpenFlow switch หลังจากนั้น Controller จะใส่ต้นทางปลายทางและตั้งค่า Flow เป็น Subflow ที่ 1 ใน Mapping Table (บรรทัดที่ 15) แต่อย่างไรก็ตามในขั้นตอนนี้จะยังไม่สามารถระบุตัวเลข Token เนื่องจากยังไม่มีข้อมูลของ Token ในแพ็คเก็ตที่เข้ามา โดยจะกลับมาใส่ในขั้นตอนที่ 2 ในส่วนของขั้นตอนที่ 2 3 และ 4 นั้นจะมีส่วนที่คล้ายคลึงกับขั้นตอนที่ 1 ดังแสดงตามภาพที่ 2 และภาพที่ 3 แต่จะต่างกันตรงกระบวนการระบุค่า Token ซึ่งขั้นตอนที่ 2 จะได้ค่า Token โดยการคำนวณจากสมการที่ (1) (บรรทัดที่ 18) ขั้นตอนที่ 3 จะสามารถได้ค่า Token มาโดยตรง

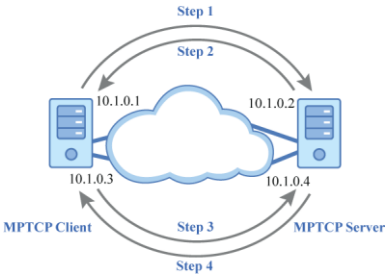
```
Input: Flow entry
Output: OpenFlow rules
1:  p ← Packet header received by the controller
2:  if p is MPTCP packet then
3:      ip_src ← Extract source IP from p
4:      ip_dest ← Extract destination IP from p
5:      port_src ← Extract source port from p
6:      port_dest ← Extract destination port from p
7:      match ← Create match from IP and Port
8:      if match in PathCache then
9:          path ← get path from PathCache
10:     else
11:         mptcp_type ← Extract MPTCP type from TCP option in p
12:         if mptcp_type == MP_CAPABLE then
13:             if SYN packet then
14:                 path ← Get path from Path Selection Algorithm
15:                 Store match and path in PathCache
16:             else if SYN-ACK packet then
17:                 senderKey ← Extract Sender key from p
18:                 token ← Hash a senderKey
19:                 Map token to SYN packet path in PathCache
20:                 path ← Get revert of path from PathCache with IP
                        and Port information
21:                 Store match and path in path PathCache
22:             end if
23:         end if
24:         if mptcp_type == MP_JOIN then
25:             if SYN packet then
26:                 token ← Extract Receiver Token from TCP option
27:                 path ← Get another path
28:                 Store a new match and path to PathCache
29:             else if SYN-ACK packet then
30:                 token ← Lock up from PathCache with IP and Port
                      information
31:                 path ← Get revert of path from PathCache with IP
                      and Port information
32:                 Store match and path in PathCache
33:             end if
34:         end if
35:     end if
36:     Add rules to switches with path information.
37: end if
```

ภาพที่ 2: แสดงโค้ดเทียมของอัลกอริทึมการระบุ Subflow

จากแพ็คเก็ต ขั้นตอนที่ 4 ต้องใช้ค่า Token จากขั้นตอนที่ 3 โดยจะค้นหาแถวของขั้นตอนที่ 3 โดยหาค่า Source ที่เท่ากับ Destination ของแพ็คเก็ตนี้ (บรรทัดที่ 30)

3.2 ขั้นตอนการเลือกเส้นที่ใช้ส่งข้อมูล

เนื่องจากเครื่องผู้ส่งที่เป็นผู้กำหนดจำนวนของ Subflow ไม่ทราบจำนวน Subflow ที่ควรสร้าง และการส่งจำนวน Subflow ที่มากเกินไปจะเป็น Overhead ให้กับเครือข่ายโดยไม่เกิดผลประโยชน์ ดังนั้นงานวิจัยนี้จึงเสนอการเลือกเส้นทางที่ใช้สำหรับ 2 Subflows อันเนื่องมาจากเป็นกรณีที่สามารถเกิดขึ้นได้เยอะที่สุดในปัจจุบัน ดังสังเกตจากคอมพิวเตอร์ มือถือ แท็บเล็ต รวมไปถึงเครื่องเซิร์ฟเวอร์มักจะมีส่วนต่อขยายอย่างต่ำจำนวน 2 ส่วนต่อขยาย และเครือข่ายที่นิยมทำ Redundant เพื่อเพิ่มเสถียรภาพ



Step 1 :

Source	Destination	Token	Subflow
10.1.0.1:55604	10.1.0.2:5001		1
10.1.0.1:34211	10.1.0.2:5001		1

Step 2 :

Source	Destination	Token	Subflow
10.1.0.1:55604	10.1.0.2:5001	1234	1
10.1.0.1:34211	10.1.0.2:5001	5678	1
10.1.0.2:5001	10.1.0.1:55604	1234	1 (reply)
10.1.0.2:5001	10.1.0.1:34211	5678	1 (reply)

Step 3 :

Source	Destination	Token	Subflow
10.1.0.1:55604	10.1.0.2:5001	1234	1
10.1.0.1:34211	10.1.0.2:5001	5678	1
10.1.0.2:5001	10.1.0.1:55604	1234	1 (reply)
10.1.0.2:5001	10.1.0.1:34211	5678	1 (reply)
10.1.0.3:50812	10.1.0.4:5001	1234	2
10.1.0.3:37222	10.1.0.4:5001	5678	2

Step 4 :

Source	Destination	Token	Subflow
10.1.0.1:55604	10.1.0.2:5001	1234	1
10.1.0.1:34211	10.1.0.2:5001	5678	1
10.1.0.2:5001	10.1.0.1:55604	1234	1 (reply)
10.1.0.2:5001	10.1.0.1:34211	5678	1 (reply)
10.1.0.3:50812	10.1.0.4:5001	1234	2
10.1.0.3:37222	10.1.0.4:5001	5678	2
10.1.0.4:5001	10.1.0.3:50812	1234	2 (reply)
10.1.0.4:5001	10.1.0.3:37222	5678	2 (reply)

ภาพที่ 3: แสดงตัวอย่างการเริ่มต้นส่งข้อมูล

การเลือกเส้นทางควรจะได้เส้นทางที่ส่งข้อมูลแล้วจะได้ค่า Throughput ที่สูง ดังนั้น Throughput จึงเป็นค่าสำคัญที่เราต้องการทราบเพื่อเปรียบเทียบในการเลือกเส้นทางแต่ละรูปแบบที่สามารถส่งข้อมูลได้ ข้อมูลที่ Controller หาได้และมีผลกับ Throughput ในการส่งข้อมูลของ MPTCP คือแบนด์วิดท์ และค่าความหน่วงของแต่ละเส้นทาง ซึ่งค่าแบนด์วิดท์ที่เราจะใช้คือค่าแบนด์วิดท์ทั้งหมดลบออกจากแบนด์วิดท์ที่ถูกใช้อยู่ในขณะนั้น ส่วนค่าความหน่วงจะมีผลทำให้ การทำ Three-Way Handshake ทำได้ช้าและยังมีผลกับการส่งข้อมูลแบบหลายเส้นทางเพราะค่าความหน่วงของ 2 เส้นทางที่ต่างกันจะทำให้เกิดปัญญา Head-of line blocking และจะส่งผลกับ Throughput ด้วย แต่ค่าความสำคัญของตัวแปรเหล่านี้ที่ไม่เท่ากันและขึ้นกับลำดับของเส้นทางที่เริ่มส่งด้วย จึงทำให้ประมาณค่า Throughput จากข้อมูลเหล่านี้ได้ยาก

งานวิจัยชิ้นนี้จึงเสนอการหาค่าความสัมพันธ์ของแบนด์วิดท์เส้นทางที่ 1 แบนด์วิดท์เส้นทางที่ 2 ค่าความหน่วงเส้นทางที่ 1 และค่าความหน่วงเส้นทางที่ 2 เพื่อประมาณหาค่า

Throughput โดยใช้สมการที่ได้จากกระบวนการทำ Multiple Linear Regression การหาข้อมูลเริ่มต้นเพื่อใช้ในกระบวนการทำ Multiple Linear Regression เริ่มต้นจากกำหนดค่าข้างต้นประจำแต่เส้นทางตามตารางที่ 1 จากนั้นทดลองส่งข้อมูลแบบ MPTCP ไปทุกรูปแบบการจับคู่เส้นทาง เพื่อทดสอบหาค่า Throughput ซึ่งการทดลองนี้ในแต่ละรูปแบบการทดลองจะทำการทดลอง 3 ครั้งและหาค่าเฉลี่ย ซึ่งผลการทดลองจะแสดงตามรูปที่ 4 (ก) และนำผลมาสร้างสมการที่ (2) ซึ่งได้จากการคำนวณสมการ Multiple Linear Regression โดยใช้ภาษา R ผ่านโปรแกรม RStudio เพื่อประมาณหาค่า Throughput โดยกำหนดให้ BW1 คือแบนด์วิดท์เส้นทางที่ 1 BW2 คือแบนด์วิดท์เส้นทางที่ 2 D1 คือค่าความหน่วงเส้นทางที่ 1 D2 คือค่าความหน่วงเส้นทางที่ 2 และ TP คือค่า Throughput

$$TP = 0.021327 + (0.857795 \times BW1) + (0.811956 \times BW2) - (0.004140 \times D1) - (0.003819 \times D2) \quad (2)$$

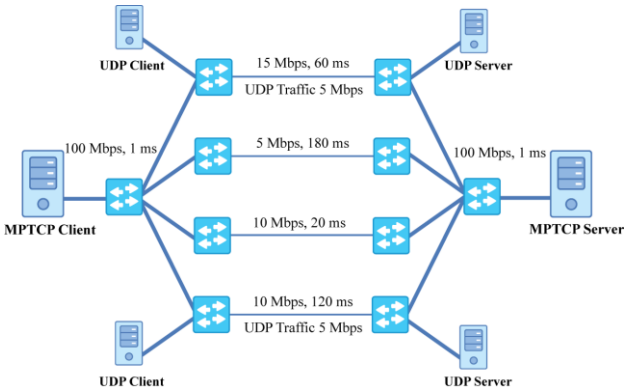
ภาพที่ 4 (ข) แสดงกราฟผลของ Throughput ที่ได้จากการแทนค่าสมการที่ (2) พบว่าเมื่อเปรียบเทียบกราฟในภาพที่ 4 (ก) และ (ข) แล้วมีลักษณะที่ใกล้เคียงกัน งานวิจัยนี้จึงนำสมการนี้ไปใช้ใน Controller เพื่อการประมาณค่า Throughput ของแต่ละคู่เส้นทางที่เป็นไปได้ และทำการเลือกเส้นทางที่มีการประมาณค่า Throughput ได้สูงสุด

4. ผลการดำเนินงาน

งานวิจัยนี้ได้ใช้ Mininet Network Emulator 2.2.1 [6], Ubuntu 14.04.4 LTS, MPTCP: Stable release v0.91.2 [10]

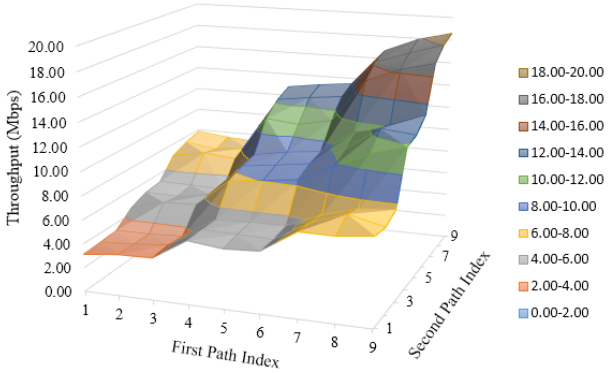
ตารางที่ 1: คุณสมบัติแต่ละเส้นทาง จากต้นทางถึงปลายทาง

Path Index	Bandwidth (Mbps)	Delay (ms)
1	2	200
2	2	100
3	2	20
4	5	200
5	5	100
6	5	20
7	10	200
8	10	100
9	10	20

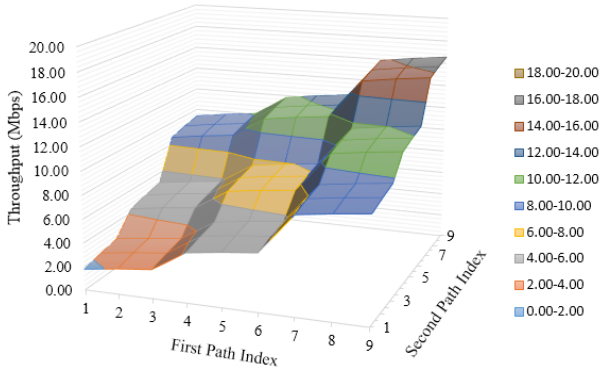


ภาพที่ 5: โครงสร้างเครือข่ายที่ใช้ในการทดลอง

และ Iperf 2.0.8 [9] ทำการทดลองบนโครงสร้างเครือข่ายตามรูปภาพที่ 5 การทดลองสามารถใช้ 1 ส่วนต่อขยายของเครื่องส่งและเครื่องรับได้นั้นอันเนื่องด้วย MPTCP เวอร์ชั่น 0.90 ขึ้นไปสามารถกำหนดให้สามารถส่งได้มากกว่า 1 Subflow ใน 1 คู่ของ IP Address โดยการทดลองจะออกแบบเป็นสองรูปแบบ คือ รูปแบบแรก Controller ทำการเลือกเส้นทางแบบสุ่ม (Random) อันเนื่องมาจากค่าเริ่มต้นของ Controller ที่เลือกเส้นทางแบบเส้นที่สุดนั้น ในกรณีมีเส้นทางที่สั้นที่สุดเท่ากันมากกว่า 1 เส้นทาง โดยที่แต่ละเส้นทางมีความสามารถในการ

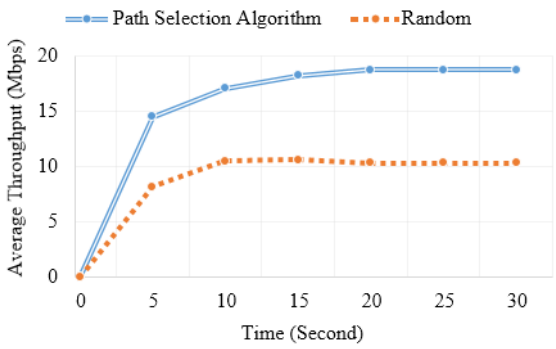


(ก) ผล Throughput จากการทดลอง



(ข) ผล Throughput จากการแทนค่าสมการ

ภาพที่ 4: แสดงค่า Throughput ในแต่ละคู่เส้นทาง



ภาพที่ 6: แสดงค่า Throughput โดยเฉลี่ยจากการทดลอง

ส่งที่ต่างกัน จะทำเลือกเส้นทางใดทางหนึ่งมาใช้ เปรียบเสมือนกับการสุ่มนั้นเอง โดยไม่ได้พิจารณาแพ็คเกต MPTCP Header หรือข้อมูลอื่นแต่อย่างใด และแบบที่สองคือ Controller ทำกระบวนการระบุ Subflow ตามหัวข้อ 3.1 และกระบวนการเลือกเส้นทางตามหัวข้อ 3.2

การทดลองจะตั้งค่าแบนด์วิดท์ ค่าความหน่วง และตลอดการทดลองจะมีข้อมูล UDP ผสมตามเส้นทางที่ระบุในภาพที่ 5 โดยแต่ละรูปแบบจะทดลอง 5 ครั้ง โปรแกรม Iperf จะแสดงผลของ Throughput ทุก ๆ 5 วินาที จากนั้นหาค่าเฉลี่ยของผลที่ได้ สามารถแสดงในภาพที่ 6

5. สรุป

งานวิจัยนี้แสดงให้เห็นว่าสามารถนำ SDN เข้ามาช่วยควบคุมการส่งข้อมูลที่ใช้โพรโทคอล MPTCP ผลการทดลองพบว่ากระบวนการเลือกเส้นทางที่ได้นำเสนอไปนี้สามารถทำให้ค่าเฉลี่ยของ Throughput ดีขึ้น 76 เปอร์เซ็นต์ เมื่อเปรียบเทียบกับการส่งแบบสุ่มเส้นทาง อันเนื่องมาจากเป็นการเลือกเส้นทางให้แต่ละ Subflow ที่ไม่ใช้เส้นทางเดียวกันอันทำให้เกิดคอขวด และกระบวนการนี้เป็นการเลือกเส้นทางที่คำนวณค่า Throughput ไว้ก่อนที่จะส่งข้อมูล โดยอัลกอริทึมที่นำเสนอนี้สามารถระบุแต่ Subflow แยกออกกันได้อย่างถูกต้องจากข้อมูลในแพ็คเก็ต การพิจารณาเลือกเส้นทางจากการพิจารณาค่าตัวแปรที่มีผลกับค่า Throughput ของการส่งข้อมูลแบบ MPTCP ซึ่งค่าความสำคัญของแต่ละตัวแปรถูกคำนวณจากกระบวนการทำ Multiple Linear Regression

เอกสารอ้างอิง

[1] J. Postel. (1981). *TRANSMISSION CONTROL PROTOCOL*. Available: https://tools.ietf.org/html/rfc793

[2] A. Ford, Raiciu, C., Handley, M., and O. Bonaventure. (2013). *TCP Extensions for Multipath Operation with Multiple Addresses*. Available: https://tools.ietf.org/html/rfc6824

[3] (2016). *Use Multipath TCP to create backup connections for iOS*. Available: https://support.apple.com/en-hk/HT201373

[4] B. A. A. Nunes, M. Mendonca, X. N. Nguyen, K. Obraczka, and T. Turletti, "A Survey of Software-Defined Networking: Past, Present, and Future of Programmable Networks," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 3, pp. 1617-1634, 2014.

[5] N. McKeown *et al.*, "OpenFlow: enabling innovation in campus networks," *SIGCOMM Comput. Commun. Rev.*, vol. 38, no. 2, pp. 69-74, 2008.

[6] B. Lantz and B. Heller. *Mininet*. Available: http://mininet.org

[7] R. Izard. *Floodlight OpenFlow Controller*. Available: http://www.projectfloodlight.org

[8] J. Postel. (1980). *User Datagram Protocol*. Available: https://tools.ietf.org/html/rfc768

[9] A. Tirumala, F. Qin, J. Dugan, J. Ferguson, and K. Gibbs. (2010). *Iperf*. Available: https://sourceforge.net/projects/iperf/

[10] *MultiPath TCP - Linux Kernel implementation*. Available: http://multipath-tcp.org/

[11] R. v. d. Pol *et al.*, "Multipathing with MPTCP and OpenFlow," in *2012 SC Companion: High Performance Computing, Networking Storage and Analysis*, 2012, pp. 1617-1624.

[12] M. Sandri, A. Silva, L. A. Rocha, and F. L. Verdi, "On the Benefits of Using Multipath TCP and Openflow in Shared Bottlenecks," in *2015 IEEE 29th International Conference on Advanced Information Networking and Applications*, 2015, pp. 9-16.

[13] S. Zannettou, M. Sirivianos, and F. Papadopoulos, "Exploiting path diversity in datacenters using MPTCP-aware SDN," in *Proceedings - IEEE Symposium on Computers and Communications*, 2016, vol. 2016-August, pp. 539-546.

[14] D. Eastlake 3rd and P. Jones. (2001). *US Secure Hash Algorithm 1 (SHA1)*. Available: https://tools.ietf.org/html/rfc3174

อุปกรณ์เสริมสำหรับรถเข็นคนพิการโดยใช้เทคโนโลยีเครือข่ายแห่งสรรพสิ่ง

Accessories for wheelchair by using Internet of Things technology

ณัฐกิตติ์ วัชรเธียรสกุล(Nuttakit Vatcharatiansakul) [1] *สิทธิกร เจริญทรัพย์(Sithikorn Charoensub)* [1]
และปานวิทย์ ธุวะนุติ (Panwit Tuwanut)[1]
[1]*คณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง*
ggguuyy@gmail.com,koshiz1994@gmail.com,panwit@it.kmtil.ac.th

บทคัดย่อ

เนื่องจากในปัจจุบันประเทศไทยได้ก้าวเข้าสู่สังคมผู้สูงอายุ ทำให้ผู้สูงอายุเป็นจำนวนมากมีความต้องการใช้รถเข็นในการเคลื่อนที่ไปตามที่ต่างๆ และในปัจจุบันรถเข็นที่มีฟังก์ชั่นการทำงานแบบพิเศษต่างๆ ได้ ก็มีราคาแพง และต้องนำเข้าจากต่างประเทศ ดังนั้นในงานวิจัยนี้จึงเป็นการนำเสนอการประดิษฐ์อุปกรณ์เสริมสำหรับรถเข็นให้มีความอัจฉริยะ มีความทันสมัยสะดวกต่อการใช้งาน มีความปลอดภัยต่อผู้สูงอายุ และราคาย่อมเยา โดยอุปกรณ์เสริมที่นำเสนอนี้จะมี ฟังก์ชั่นการทำงานหลักๆ ประกอบด้วย ส่วนของการเฝ้าระวังสุขภาพ (Health Monitoring) ส่งข้อมูลสัญญาณชีพไปยังโรงพยาบาลหรือสถานพยาบาลที่ได้กำหนดไว้ ส่วนตรวจจับการเคลื่อนที่ของรถเข็นให้เคลื่อนที่ด้วยความปลอดภัย ส่วนระบุตำแหน่งของรถเข็น และส่วนควบคุมเบาะไม่ให้เกิดแผลกดทับเมื่อนั่งบนรถเข็นเป็นระยะเวลานาน การทำงานทั้งหมดของรถเข็นนี้ทำงานอยู่บนแนวคิดของเครือข่ายแห่งสรรพสิ่งสามารถสั่งการ ควบคุม และเฝ้าระวังผ่านแอพพลิเคชั่นบนมือถือและโปรแกรมประยุกต์บนเว็บ

คำสำคัญ: เฝ้าระวังสุขภาพ เครือข่ายแห่งสรรพสิ่ง

Abstract

These days Thailand has entering into aging society are leading lots of elderly people need to use wheelchair for moving to another place. Moreover, wheelchair which has special functions are expensive and also importing from other countries. Therefore, this research is presented about the invention of smart wheelchair which are modern, easy to use, safety for elderly people and affordable prices. The smart wheelchair has offered main functions including Health Monitoring - Sending pulse to a hospital or clinic as defined, controlling the movement of wheelchair safely, defining the position of the wheelchair and controlling the seat to prevent the pressure sores when using the wheelchair for a long time. All of these are based on Internet of Things which can be controlled and monitored via mobile web application and web application.

Keyword: Health Monitoring, IoT, MQTT

1. บทนำ

เนื่องจากในปัจจุบันประเทศไทยได้ก้าวเข้าสู่สังคมผู้สูงอายุ มีจำนวนผู้สูงอายุเป็นจำนวนมาก จนกลายเป็นประเด็นยุทธศาสตร์การวิจัยของชาติด้านผู้สูงอายุและสังคมสูงอายุ และปัญหาหนึ่งที่พบคือ ผู้สูงอายุที่มีอายุมากๆ หรือ ผู้ป่วยแบบพักรักษาอยู่ที่บ้าน จำเป็นจะต้องนั่งรถเข็นไปตามที่ต่างๆ และในปัจจุบันพบว่า รถเข็นที่มีอยู่ตามท้องตลาดก็เป็นรถเข็นแบบปกติ ไม่มีฟังก์ชั่นพิเศษใดๆ ในขณะที่รถเข็นที่มีฟังก์ชั่นพิเศษต่างๆ ก็จะมีราคาสูงขึ้นไปตามลำดับ และต้องนำเข้าจากต่างประเทศ ทำให้ผู้สูงอายุส่วนใหญ่ไม่มีโอกาสได้ใช้งาน

ดังนั้นจึงเกิดแนวคิดในการพัฒนาอุปกรณ์เสริมสำหรับรถเข็นให้มีความอัจฉริยะ มีความทันสมัย มีความสะดวกสบายต่อผู้สูงอายุ และมีราคาย่อมเยาว์ โดยอุปกรณ์เสริมที่นำเสนอในงานวิจัยนี้ จะประกอบด้วย อุปกรณ์ตรวจเฝ้าระวังสุขภาพแบบพกพา (Portable health monitoring) โดยสามารถวัดสัญญาณชีพที่จำเป็น ได้แก่ ความดันโลหิต สัญญาณคลื่นไฟฟ้าหัวใจ อุณหภูมิร่างกาย สัญญาณลมหายใจ เมื่อมีความผิดปกติเกิดขึ้นกับผู้สูงอายุ ก็จะส่งสัญญาณดังกล่าวไปยังโรงพยาบาล เพื่อให้แพทย์หรือพยาบาลได้สังเกตความผิดปกติ ทำให้แพทย์สามารถวินิจฉัยในเบื้องต้นได้ สามารถให้คำแนะนำการรักษาในขั้นต้นแก่ผู้สูงอายุได้ และรถเข็นดังกล่าวยังคำนึงถึงการเกิดแผลกดทับเมื่อนั่งรถเข็นเป็นเวลานาน จึงมี

ระบบปั๊มลมอัตโนมัติเพื่อเป่าเข้ามายังเบาะแบบพิเศษที่ช่วยลดการเกิดแผลกดทับเมื่อนั่งเป็นเวลานาน นอกจากนี้แล้ว อีกฟังก์ชั่นหนึ่งของรถเข็นที่นำเสนอคือ มีระบบตรวจจับการล้มหรือ การเคลื่อนที่ที่ผิดปกติ เช่น เคลื่อนที่ไวเกินไป ซึ่งอาจจะเกิดจากการเลื่อนลงพื้นที่ลาดชันเป็นต้น ช่วยลดอุบัติเหตุที่อาจเกิดขึ้นกับผู้สูงอายุได้

อุปกรณ์เสริมสำหรับรถเข็นคนพิการที่นำเสนอนี้ มีราคาย่อมเยาว์ ทำงานบนเทคโนโลยีเครือข่ายแห่งสรรพสิ่งและส่งข้อมูลสัญญาณชีพที่มีความถูกต้องเป็นไปตามมาตรฐานสากลได้แบบเรียลไทม์ทำให้ผู้รับข้อมูลเช่น ญาติพี่น้อง หมอ สามารถตอบสนองต่อความผิดปกติได้ทันที และผู้ป่วยมีความสะดวกสบายในการนั่งไม่เป็นแผลกดทับ

2. เอกสารและงานวิจัยที่เกี่ยวข้อง

จากการสืบค้นงานวิจัยที่ผ่านมา ที่เกี่ยวข้องกับรถเข็น (Wheel Chair) ในฐานข้อมูลของ IEEE พบว่างานวิจัยส่วนใหญ่มุ่งเน้นไปที่การควบคุมทิศทางการเคลื่อนที่ของรถเข็นโดยใช้อวัยวะต่างๆ อาทิเช่น การควบคุมด้วยดวงตา เช่นงานวิจัยของ Umar Mohammad และคณะ [1] ได้นำเสนอการควบคุมทิศทางการเคลื่อนที่ของรถเข็นด้วยการเคลื่อนที่ของดวงตา หรือในงานวิจัยของ Manoj Challagundla และคณะ [2] ได้นำเสนอการควบคุมด้วยการกระพริบตา ต่อมางานวิจัยของ Shamim Kaiser และคณะ [3] ได้นำเสนอการควบคุมโดยการใช้กล้ามเนื้อแขนโดยติดเซ็นเซอร์เพื่อวัดสัญญาณ EMG ที่บริเวณแขน ต่อมาในงานวิจัยของ Angelo Zerbetto Neto และคณะ [4] ได้นำเสนอต้นแบบการควบคุมโดยใช้การเอียงของลำคอเช่น ก้ม เงย เอียงซ้าย ขวา เพื่อควบคุมทิศทางของรถเข็น นอกจากนี้แล้วก็ยังมีการนำเสนอการควบคุมโดยใช้คลื่นสมอง อาทิเช่นงานวิจัยของ R. S. Naveen [5] และ S-. Y. Cho [6] สรุปเป็นตารางดังนี้

ตารางที่ 1. Research summary

Organ	Research
Eyes	[1],[2]
Bicep and Tricep	[3]
Neck mussle	[4]
Brain wave	[5],[6]

ต่อมา Shraddha Uddhav Khadilkar [7] ได้นำเสนอการใช้ Smart Phone มาช่วยในการควบคุม โดยการควบคุมผ่าน mobile application หรือการใช้ gyroscope sensor ในการเอียงโทรศัพท์ไปในทิศต่างๆ จากนั้นก็จะส่งข้อมูลผ่าน Bluetooth ไปควบคุมล้อให้เคลื่อนที่ตามที่ต้องการ ต่อมางานวิจัยของ S A Akash และคณะ [8] ได้นำเสนอ การใช้แนวคิดของ Internet of Things มาใช้ร่วมกับรถเข็น โดยมีฟังก์ชั่นหลักๆ ได้แก่ Control Module โดยใช้ joy stick หรือเสียงพูด หรือ Accelerometer ในการควบคุม หรือควบคุมผ่าน Web ฟังก์ชั่นต่อมาคือ Perception module ได้แก่ ส่วนตรวจจับสิ่งกีดขวาง ส่วนตรวจจับความลาดเอียง และฟังก์ชั่นสุดท้ายคือ Utility ได้แก่ ส่วนตรวจจับการเปลี่ยนถุงปัสสาวะ การเตือนว่าแบตเตอรี่ใกล้หมด เป็นต้น

จากงานวิจัย [8] ทำให้คณะวิจัยมีความสนใจที่มุ่งเน้นไปที่การพัฒนาอุปกรณ์เสริมสำหรับรถเข็น Wheel Chair สำหรับผู้สูงอายุ ที่มีความทันสมัย สะดวกต่อการใช้งาน มีความปลอดภัยต่อผู้สูงอายุ และราคาย่อมเยาว์ โดยมีฟังก์ชั่นการทำงานหลักประกอบด้วย ส่วนของการเฝ้าระวังสุขภาพ)Health Monitoring) ส่งข้อมูลสัญญาณชีพไปยังโรงพยาบาลหรือสถานพยาบาลที่ได้กำหนดไว้ ส่วนของการควบคุมการเคลื่อนที่ของรถเข็นให้เคลื่อนที่ด้วยความปลอดภัย ส่วนระบุตำแหน่งของรถเข็น และส่วนควบคุมเบาะไม่ให้เกิดแผลกดทับเมื่อนั่งบนรถเข็นเป็นระยะเวลานาน การทำงานทั้งหมดของรถเข็นนี้ทำงานอยู่บนแนวคิดของ Internet of Things สามารถสั่งการควบคุม และเฝ้าระวังผ่าน Mobile application

3. แนวคิด ทฤษฎี ที่เกี่ยวข้อง

3.1 เทคโนโลยีเครือข่ายของสรรพสิ่ง (internet of things)

คำว่า Internet of Things มีการให้คำจำกัดความของหลายความหมาย จนกระทั่งล่าสุด สถาบันวิศวกรรมไฟฟ้าและวิศวกรรมอิเล็กทรอนิกส์นานาชาติ) IEEEได้ออกนิยามของ (คำว่าInternet of Things (IoT) [9] “An IoT is a network that connects uniquely identifiable Things to the Internet. The Things have sensing/actuation and potential programmability capabilities. Through the exploitation of unique identification and sensing, information about the Thing can be collected and the state of the Thing can be changed from anywhere,

anytime, by anything.” แปลเป็นไทยในมุมมองของผู้วิจัย “Internet of Things” คือ โครงข่ายที่เชื่อมต่อสรรพสิ่งใดๆ ซึ่งมีลักษณะที่ระบุเอกลักษณ์ได้ เข้ากับเครือข่ายอินเตอร์เน็ตผ่านโพรโทคอลการสื่อสารทั้งแบบใช้สายและไร้สาย โดยที่ สรรพสิ่งใดๆ จะมีความสามารถในการรับรู้ ตอบสนองและมีศักยภาพในการโปรแกรมหรือสั่งการการทำงานได้จากประโยชน์ของความสามารถในการระบุเอกลักษณ์และความสามารถในการรับรู้นี้สารสนเทศบนสรรพสิ่งใดๆจะสามารถถูกรวบรวม จัดเก็บและสถานภาพของสรรพสิ่งใดๆก็สามารถที่จะเปลี่ยนแปลงได้ตลอดเวลาจากสถานที่ใดๆเวลาใดๆโดยสิ่งใดๆ” จากนิยามดังกล่าว ทำให้สรุปได้ว่า Internet of Things หรือ เครือข่ายแห่งสรรพสิ่งจะมีลักษณะดังนี้

- เป็นการเชื่อมต่อกันระหว่างสรรพสิ่งใดๆ
- เชื่อมต่อสรรพสิ่งใดๆ เข้ากับโครงข่ายอินเตอร์เน็ต
- สรรพสิ่งใดๆ มีลักษณะที่ระบุเอกลักษณ์ได้
- สามารถพบได้ทั่วไป
- มีความสามารถในการ รับรู้ และตอบสนอง ได้
- เป็นระบบฝังตัวแบบอัจฉริยะ
- มีความสามารถในการสื่อสารทำงานร่วมกันได้
- สามารถตั้งค่าด้วยตนเองได้
- สามารถโปรแกรมหรือสั่งการการทำงานได้

โดยเทคโนโลยีที่ทำให้เครือข่ายแห่งสรรพสิ่งเกิดขึ้นได้จริงและสร้างผลกระทบในวงกว้างได้ แบ่งออกเป็นสามกลุ่มได้แก่ เทคโนโลยีที่ช่วยให้สรรพสิ่งรับรู้ข้อมูลในบริบทที่เกี่ยวข้อง เช่น เซ็นเซอร์ เทคโนโลยีที่ช่วยให้สรรพสิ่งมีความสามารถในการสื่อสาร เช่น ระบบสมองกลฝังตัว รวมถึงการสื่อสารแบบไร้สายที่ใช้พลังงานต่ำ อาทิ Zigbee, 6LowPAN, Low-power Bluetooth และเทคโนโลยีที่ช่วยให้สรรพสิ่งประมวลผลข้อมูล ในบริบทของตน เช่น เทคโนโลยีการประมวลผลแบบคลาวด์ และเทคโนโลยีการวิเคราะห์ข้อมูลขนาดใหญ่ หรือ Big Data Analytics

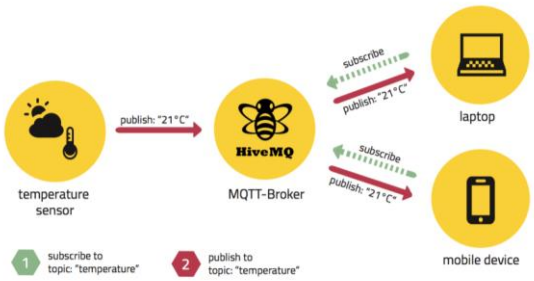
3.2 โปรโตคอล MQTT

ในการสื่อสารข้อมูลระหว่างอุปกรณ์ใดๆนั้น จำเป็นจะต้องใช้โปรโตคอลในการสื่อสาร โดยเฉพาะใน เครือข่ายแห่งสรรพสิ่งนั้น ปัจจุบันมีโปรโตคอลที่ได้รับความนิยมอยู่สองชนิดคือ CoAP Protocol และ MQTT Protocol โดยในงานวิจัยนี้จะมุ่งเน้นการประยุกต์ใช้ MQTT Protocol [10] MQTT Protocol เริ่มต้นคิดค้นโดย Andy Stanford Clark จากบริษัท IBM และ Arlen Nipper จากบริษัท Eurotech ในปี คศ. 1999เป็นการคิดค้นเพื่อแก้ปัญหาในการสื่อสารสารสนเทศในท่อส่งน้ำมัน ข้อมูลในการสื่อสาร เช่น อุณหภูมิ ความดัน และสัญญาณควบคุมวาวล์ท่อน้ำมัน การสื่อสารข้อมูลนั้นเป็นข้อมูลตลอดแนวท่อน้ำมันยาว กิโลเมตร ซึ่งข้อมูลทั้งหมดต้อง 17 การคิดค่าใช้จ่ายตามแบนด์วิดท์ที่เช่าใช้ส่งผ่านดาวเทียม ซึ่งมีงาน ปัญหาที่เกิดขึ้นนั้นทำให้มีค่าใช้จ่ายใช้ในการสื่อสารข้อมูลสูง ดังนั้น เพื่อลดค่าใช้จ่ายดังกล่าวจึงได้พัฒนาโปรโตคอลแบบ Message Queuing Telemetry Transport (MQTT) เพื่อลดค่าใช้จ่ายดังกล่าว โดยที่โปรโตคอล MQTT นี้ มีลักษณะคือ

1. Open เป็นมาตรฐานเปิดสามารถนำไปพัฒนาได้โดยไม่เสียค่าใช้จ่าย
2. Lightweight เป็นโปรโตคอลขนาดเล็ก มีขนาด Header แค่ 2 bytes
3. Reliable เป็นโปรโตคอลที่มีความน่าเชื่อถือ สามารถทำ Quality of Services (QoS) ได้
4. Simple เป็นโปรโตคอลที่ง่ายต่อการใช้งาน ใช้รูปแบบการทำงานแบบ Publish/Subscribe

การทำงานของโปรโตคอล MQTT ประกอบไปด้วยตัวกลางที่เรียกว่า Broker อธิบายขั้นตอนการทำงานคร่าวๆ คือ

1. เมื่อ Sensor ทำการอ่านค่าที่วัดได้ จะทำการ Publish ข้อมูลดังกล่าว ไปยัง Topic ที่กำหนด บน MQTT Broker
2. MQTT Broker จะทำการ Push ข้อมูลที่ได้ ไปยัง Client หรืออุปกรณ์ใดๆ ที่ทำการ Subscribe Topic นั้นๆ แสดงดังภาพที่ 1



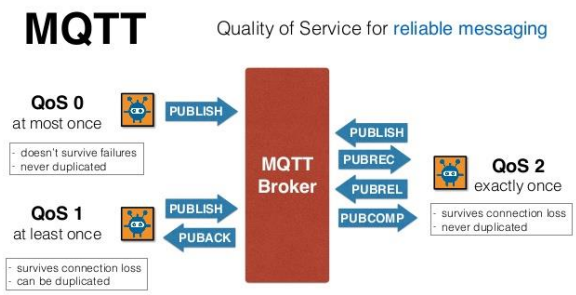
ภาพที่ 1 Publish and Subscribe method for MQTT Protocol

นอกจากนี้คุณสมบัติเด่นอีกประการของ MQTT Protocol คือ ความสามารถในการทำ QoS โดยแบ่งได้ 3 ระดับคือ

QoS 0 มีคุณสมบัติคือ At most once คือ ส่งไปแน่ๆ 1 ครั้ง

QoS 1 มีคุณสมบัติคือ At least once คือ ได้รับแน่ๆ 1 ครั้ง

QoS 2 มีคุณสมบัติคือ exactly once คือ ได้รับอย่างถูกต้องแน่ๆ 1 ครั้ง



ภาพที่ 2 QoS of MQTT

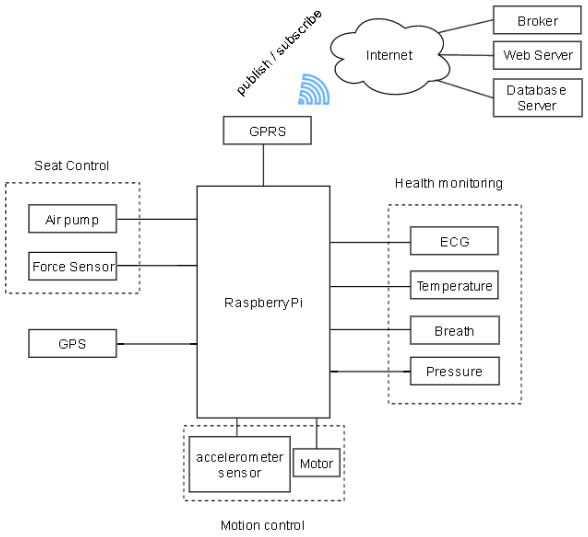
3.3 แนวคิดของงานวิจัย

พัฒนาอุปกรณ์เสริมสำหรับรถเข็นที่นำเทคโนโลยีมาประยุกต์ใช้กับอุปกรณ์ที่นำมาติดกับรถเข็นในลักษณะของเครือข่ายแห่งสรรพสิ่ง (Internet of Things) ที่อุปกรณ์แต่ละตัวสามารถติดต่อสื่อสารกันได้ (Sensing) โดยนำอุปกรณ์ Raspberry pi ซึ่งเปรียบเสมือนเครื่องคอมพิวเตอร์ขนาดเล็กมาต่อเข้ากับอุปกรณ์รับค่าต่างๆ (Sensor) และตัวRaspberry Pi ยังสามารถเชื่อมต่อกับอินเตอร์เน็ตโดยใช้ MQTT Protocol เป็นสื่อกลางสำหรับการรับส่งข้อมูล และตอบสนองความต้องการของผู้สูงอายุ สามารถเฝ้าระวังและส่งข้อมูลสัญญาณชีพต่างๆ ได้ สามารถควบคุมการเคลื่อนที่ของรถ สามารถควบคุมเบาะเพื่อลดแผลกดทับ และส่งตำแหน่งที่อยู่ผ่านระบบจีพีเอสส่งไปยังเว็บแอพพลิเคชั่นแบบเรียลไทม์ในส่วนของโปรแกรมประยุกต์บนเว็บและฐานข้อมูลนั้น จะมีการทำการพัฒนาด้วย Node.js ร่วมกับ HTML5 CSS3 และ Javascript ซึ่งจะทำการเชื่อมต่อกับฐานข้อมูลแบบ NoSQL ด้วย MongoDB ในส่วนของ Platform ที่สามารถใช้งานได้บนหน้าโปรแกรมประยุกต์บนเว็บและโปรแกรมประยุกต์บนเว็บที่ทำงานบนมือถือ เช่น Internet Explorer, Google Chrome, Mozilla Firefox, Safari(iOS)

4. การออกแบบระบบ

4.1 Block Diagram ของระบบ

ในส่วนของอุปกรณ์เสริมที่นำเสนอมี Block Diagram แสดงดังภาพที่ 3 ประกอบด้วย 3 ส่วนหลักๆ คือ Health Monitoring , ส่วนของ Seat Control และส่วนของ Motion Control



ภาพที่ 3 Block diagram

4.2 การวิเคราะห์ความต้องการของระบบ

จากการศึกษา รวบรวมข้อมูล และวิเคราะห์ความต้องการ หรือ ประสิทธิภาพของอุปกรณ์ต่างๆที่มีอยู่ในปัจจุบัน สามารถนำมาสรุปความต้องการของระบบที่ต้องการจะออกแบบได้ดังนี้

ความต้องการที่เป็นหน้าที่หลักของระบบ

- ผู้ดูแลและแพทย์สามารถตรวจสอบค่าสัญญาณ ECG, Temperature, Breath rate, Status แบบเรียลไทม์ได้
- สามารถวัดค่า ECG, Temperature, Breath rate ได้แบบเรียลไทม์
- สามารถช่วยป้องกันแผลกดทับได้โดยมีเบาะเป่าลมสำหรับผู้ป่วย
- แพทย์สามารถแจ้งเตือนมายังผู้ดูแลเมื่อเกิดความผิดปกติที่เกิดขึ้นกับผู้สูงอายุ และผู้พิการ
- ผู้ดูแลสามารถตรวจสอบตำแหน่งของผู้สูงอายุ และผู้พิการที่นั่งอยู่บนรถเข็นได้

- ผู้ดูแลสามารถตรวจสอบความเร็วของอุปกรณ์ที่ติดตั้งบนรถเข็นได้
- แพทย์สามารถตรวจสอบค่าย้อนหลังในลักษณะของ Log ได้
- มีหน้าเว็บเพจสำหรับแสดงค่าต่างๆของผู้สูงอายุ และผู้พิการได้
- สามารถหยุดรถเข็นได้อัตโนมัติเมื่อไม่มีผู้เข็น

5. ผลการทดลอง

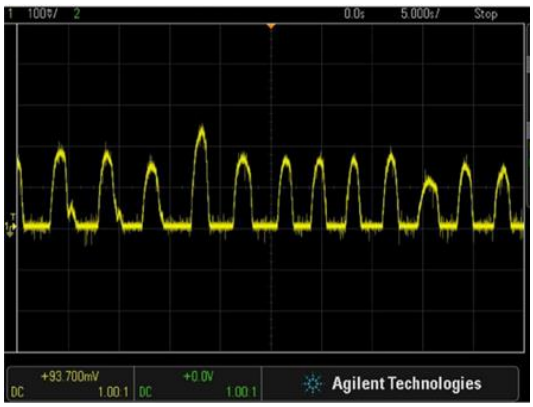
5.1 ผลการทดลองส่วนของ Health Monitoring

ทำการติดเซ็นเซอร์ต่างๆ และทำการวัดผลที่ได้จากเซ็นเซอร์ต่างๆ แสดงดังภาพที่ 4-9

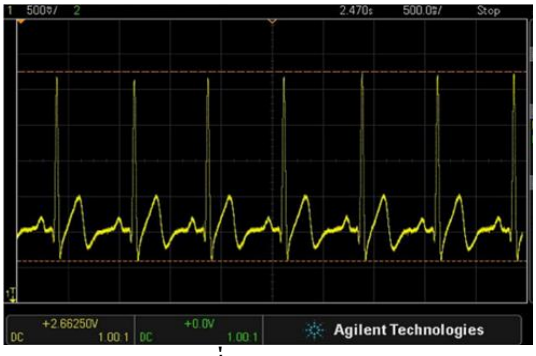
โดยเซนเซอร์ตรวจวัดระดับลมหายใจจะสามารถวัดค่าระดับลมหายใจของผู้ใช้งานได้ ECG Sensor จะใช้ในการวัดสัญญาณชีพจรของผู้ใช้งานในระหว่างที่นั่งบนรถเข็น และได้ทำการติดตั้งเซนเซอร์วัดความดันโลหิตและอุณหภูมิร่างกายที่สามารถวัดได้ในขณะใช้งาน



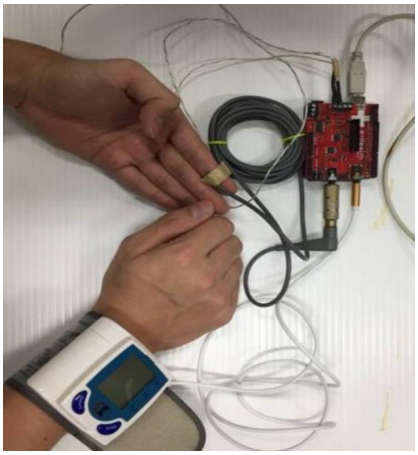
ภาพที่ 4. Breath Sensor



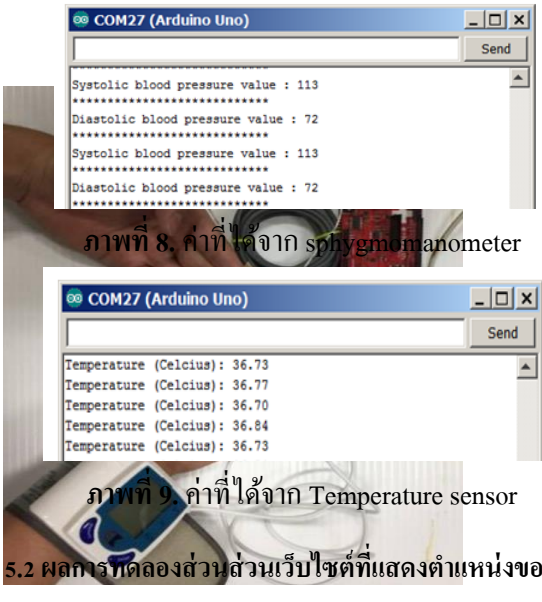
ภาพที่ 5. Breath Signal



ภาพที่ 6. ECG Signal



ภาพที่ 7. sphygmomanometer และ Temperature sensor



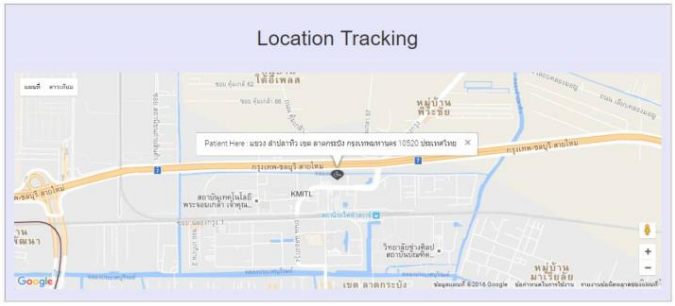
ภาพที่ 8. ค่าที่ได้จาก sphygmomanometer

ภาพที่ 9. ค่าที่ได้จาก Temperature sensor

5.2 ผลการทดลองส่วนส่วนเว็บไซต์ที่แสดงตำแหน่งของรถเข็น

ใช้ GPS module ส่งค่าตำแหน่งปัจจุบันไปแสดงผลบนเว็บไซต์ แสดงผลตำแหน่งดังภาพที่ 10

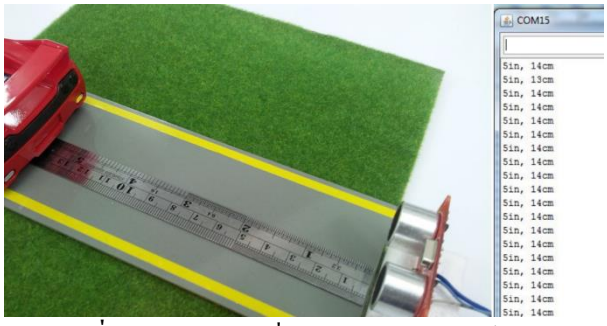
Figure 8. Values from sphygmomanometer



ภาพที่ 10. Location by GPS display on Web Application

5.3 ผลการทดลองส่วนของการตรวจจับสิ่งกีดขวาง

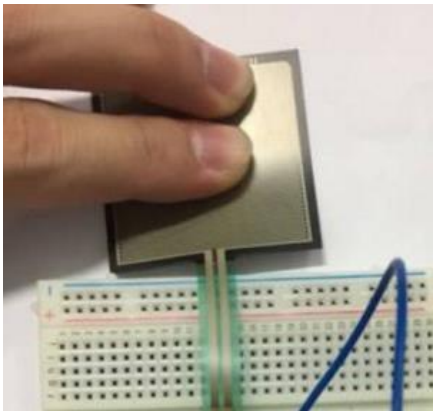
ใช้เซนเซอร์อัลตราโซนิคในการตรวจจับสิ่งกีดขวางที่อยู่ข้างหน้ารถเข็น แสดงระยะที่วัดได้นำมาใช้ในการหยุดรถเข็นแบบอัตโนมัติ แสดงผลดังภาพที่ 11

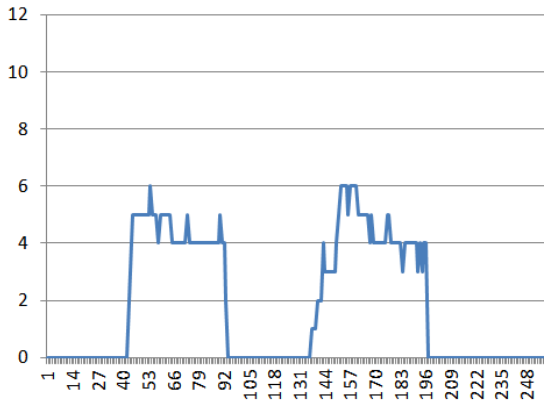


ภาพที่ 11. การวัดระยะสิ่งกีดขวางของอุปกรณ์โดยใช้เซนเซอร์อัลตราโซนิค

5.4 ผลการทดลองส่วนของการควบคุมเบาะ

ใช้เซนเซอร์วัดแรงกดในการวัดแรงกดของผู้ใช้รถเข็นและจับเวลาที่นั่งเพื่อไปควบคุมการปรับลมของเบาะป้องกันแผลกดทับ แสดงผลดังภาพที่ 12





ภาพที่ 12. กราฟแสดงผลของเซนเซอร์ที่ติดกับเบาะป้องกันแผลกดทับ

5.5 ผลการทดลองส่วนเว็บไซต์ที่แสดงผลข้อมูลสัญญาณชีพ

ระบบสารสนเทศที่พัฒนาบนโปรแกรมประยุกต์บนเว็บเพื่อแจ้งข้อมูลผู้ป่วยให้แพทย์ หรือ ผู้ดูแลทราบถึงข้อมูลสุขภาพผู้ป่วยผ่านทาง Websocket แสดงผลดังภาพที่ 13



ภาพที่ 13. Real-Time System of Health monitoring

6. สรุปผล

ในบทความวิจัยนี้เป็นการประยุกต์ใช้เทคโนโลยีการเชื่อมต่อของสรรพสิ่ง ในการพัฒนาอุปกรณ์เสริมเพื่อเชื่อมต่อกับรถเข็นที่มีอยู่ใช้งานทั่วไป ให้มีความอัจฉริยะ ทันสมัย สะดวกต่อการใช้งาน มีความปลอดภัยต่อผู้สูงอายุ และราคาย่อมเยาว์ โดยมี ฟังก์ชั่นการทำงานประกอบด้วย ส่วนของการเฝ้าระวังสุขภาพ (Health Monitoring) ส่วนตรวจจับการเคลื่อนที่ของรถเข็น ส่วนระบุตำแหน่งของรถเข็น และส่วนควบคุมเบาะ และส่งค่าต่างๆ ไปแสดงผลบน Web Application แบบเรียลไทม์ ผ่าน Web Socket ทำให้บุคคลต่างๆ ที่เกี่ยวข้องสามารถดูค่าต่างๆ ได้พร้อมๆ กัน สามารถตอบสนองต่อความผิดปกติของผู้ใช้งานรถเข็นได้ทันที

เอกสารอ้างอิง

[1] Umar Mohammad; Mohammad Anas , “Design of a low cost DIY moving wheel chair using ATmega1284P based on retina movement for the persons disabled with quadriplegia”, 2015 Annual IEEE India Conference (INDICON)

[2] Manoj Challagundla; K. Yogeshwar Reddy; N. Harsha Vardhan,“Automatic motion control of powered wheel chair by the movements of eye blink”, Advanced Communication Control and Computing Technologies (ICACCCT), 2014 International Conference on

[3] Shamim Kaiser; Zamshed I. Chowdhury; Shamim Mamun; Amir Hussain; Mufti Mahmud,“Solar Powered Wheel Chair for Physically Challenged People Using Surface EMG Signal”, Computational Intelligence, 2015 IEEE Symposium Series on

[4] Angelo Zerbetto Neto; Alexandre Mesquita; Marilda Machado Spindola; Marcos Magnani,“Prototype of a Wheelchair Controlled by Cervical Movements”, Bioinformatics and Bioengineering (BIBE), 2014 IEEE International Conference on

[5] R. S. Naveen; Anitha Julian,“Brain computing interface for wheel chair control”, Computing, Communications and Networking Technologies (ICCCNT),2013 Fourth International Conference on

[6] S-. Y. Cho; A. P. Winod; K. W. E. Cheng,“Towards a Brain-Computer Interface based control for next generation electric wheelchairs”, Power Electronics Systems and Applications, 2009. PESA 2009. 3rd International Conference on

[7] Shraddha Uddhav Khadilkar; Narendra Wagdarikar ,“Android phone controlled voice, gesture and touch screen operated smart wheelchair”,Pervasive Computing (ICPC), 2015 Int Conference on

[8] S A Akash; Akshay Menon; Arpit Gupta; Md Waheeb Wakeel; M N Praveen; P. Meena, "A novel strategy for controlling the movement of a smart wheelchair using internet of things",Global Humanitarian Technology Conference - South A sia Satellite (GHTC-SAS), 2014 IEEE

[9 Roberto Minerva,Abyi Biru, Domenico Rotondi, “Towards a definition of the Internet of Things (IoT)”, IEEE, 2015.

[`10] Dave Locke. MQ telemetry transport (mqtt) v3.1 protocol specification. IBM developerWorks Technical Library, available at http://www.ibm.com/developerworks/library/ws-mqtt/, 2016

การขยายขนาดการให้บริการแบบอัตโนมัติในระบบด็อกเกอร์
An Auto-scaling Mechanism in Docker System

ธีรภัทร์ ขุนเพ็ชร์(Teerapat Khunpeth)[1] และสุชา สมานชาติ(Sucha Smanchat)[2]
[1]ภาควิชาการสื่อสารข้อมูลและเครือข่าย [2]ภาควิชาเทคโนโลยีสารสนเทศ
คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
[1]s5707011858053@email.kmutnb.ac.th, [2]sucha.s@it.kmutnb.ac.th

บทคัดย่อ

ปัจจุบันมีเทคโนโลยีการทำเครื่องคอมพิวเตอร์เสมือนหลากหลายรูปแบบเพื่อให้มีการใช้งานเครื่องคอมพิวเตอร์ที่มีสมรรถนะสูงอย่างคุ้มค่าให้มากที่สุดและเทคโนโลยีลินุกซ์คอนเทนเนอร์(Linux Container: LXC) หรือด็อกเกอร์(Docker) กำลังได้รับความนิยมในการใช้งานเนื่องจากสามารถนำไปประยุกต์ใช้งานในรูปแบบการให้บริการต่าง ๆ ที่ไม่ยุ่งยาก เช่น การให้บริการเว็บไซต์ เป็นต้น อย่างไรก็ตามหากมีการให้บริการเพื่อตอบสนองกับผู้ใช้บริการเป็นจำนวนมากจนทำให้ทรัพยากรในการให้บริการไม่เพียงพอต่อการใช้งานจึงมีความจำเป็นที่จะต้องขยายขนาดการให้บริการ ผู้วิจัยจึงได้ทำพัฒนาสถาปัตยกรรมการขยายขนาดการให้บริการแบบอัตโนมัติในระบบด็อกเกอร์ขึ้นให้มีความพร้อมในการให้บริการตลอดเวลาโดยใช้เทคโนโลยีด็อกเกอร์

งานวิจัยนี้ได้ดำเนินการพัฒนาสถาปัตยกรรมการขยายขนาดการให้บริการอัตโนมัติในระบบด็อกเกอร์ที่ประกอบด้วยส่วนการเฝ้าระวัง ส่วนการตัดสินใจ และส่วนการขยายขนาดการให้บริการโดยใช้ค่าของการใช้งานหน่วยประมวลผลมาเป็นค่าที่ใช้ในการตัดสินใจว่าจะขยายขนาดหรือลดขนาดและมีการทดสอบการทำงานและทดสอบความพร้อมในการให้บริการของกลไกที่ใช้ในงานวิจัยนี้ โดยกลไกสามารถลดและขยายขนาดการให้บริการได้ตามที่ผู้ใช้งานตั้งค่าไว้โดยมีข้อผิดพลาดจากการทำงานเล็กน้อย

คำสำคัญ: การขยายขนาดการให้บริการแบบอัตโนมัติ ด็อกเกอร์ ลินุกซ์คอนเทนเนอร์

Abstract

Currently, there are many options for virtualization of computer systems in order to maximize the utilization of physical computing hardware. Among these technologies, Docker has recently emerged as a leading Linux container implementation, which has gained its popularity for its easiness to be applied to various service applications such as website services. However, as more users access and use a Docker system, the computing resources initially provided for the Docker system may not be able handle such high workload causing the services to be irresponsive. Thus, there is a need for an auto-scaling mechanism to automatically allocate computing resources for the Docker system. This research proposes an architecture of an auto-scaling mechanism to provide high availability in the Docker system.

The proposed auto-scaling architecture is composed of a monitoring component, a scaling decision maker, and a scaling executor. The scaling decision is based on CPU utilization metric in order to scale-out and scale-in the computing resources. The evaluation result shows that the proposed architecture can satisfactorily performs auto-scaling for the Docker system according to the user-defined CPU utilization threshold.

Keyword: Auto-scaling Mechanism, Docker, Linux Container.

1. บทนำ

ปัจจุบันเทคโนโลยีการประมวลผลบนอินเทอร์เน็ตหรือคลาวด์ (Cloud) เป็นที่ได้รับความนิยมเป็นอย่างมาก เนื่องจากเครื่องคอมพิวเตอร์มีสมรรถนะในการประมวลผลที่สูงจึงมีผู้ให้บริการจำนวนมากเช่นกัน ด้วยเหตุผลนี้จึงเป็นข้อดีต่อกับผู้ที่ต้องการใช้งานเนื่องจากผู้ใช้งานสามารถเลือกการใช้งานได้ตาม

ต้องการ เช่น สามารถเลือกหน่วยความจำหลัก (Random Access Memory : RAM) หน่วยความจำสำรอง (Hard Disk Drive) และเครือข่ายคอมพิวเตอร์ (Network) เป็นต้น ให้มีความเหมาะสมในการใช้งาน และด้วยเทคโนโลยีที่คอมพิวเตอร์มีประสิทธิภาพสูงร่วมกับความเร็วของการใช้งานอินเทอร์เน็ตจึงทำให้ผู้ใช้งานสามารถที่จะจัดการสิ่งต่าง ๆ เช่น เลือกทรัพยากรเป็นต้น ได้ด้วยตัวเองในรูปแบบที่ผู้ให้บริการจัดเตรียมไว้ให้

การให้บริการต่างๆ ผ่านระบบคลาวด์นั้นจำเป็นที่จะต้องรักษาข้อกำหนดการให้บริการระหว่างผู้ให้บริการและผู้ใช้บริการในระบบคอมพิวเตอร์ หรือ Service Level Agreements (SLA) ให้มีคุณภาพที่พร้อมให้บริการอยู่ตลอดเวลา ปัจจัยที่จะทำให้ระบบไม่สามารถให้บริการได้ เช่น ทรัพยากรเครื่องคอมพิวเตอร์มีไม่เพียงพอต่อการให้บริการ หรือ แอพพลิเคชั่นมีการทำงานอันผิดพลาดจากการทำงานของตัวมันเอง เป็นต้น สิ่งหนึ่งที่เราสามารถรักษา SLA ให้มีความพร้อมใช้งานตลอดเวลานั้น คือการดูแลและจัดการทรัพยากรของระบบคอมพิวเตอร์ทั้งหมดให้มีการจัดสรรได้ตามความเหมาะสมต่อความต้องการใช้งาน

การเพิ่มจำนวนเครื่องคอมพิวเตอร์หรือเพิ่มทรัพยากรของเครื่องคอมพิวเตอร์ที่ให้บริการ นับว่าเป็นอีกวิธีการหนึ่งที่จะเพิ่มประสิทธิภาพให้กับระบบให้มีความสามารถรองรับกับจำนวนการร้องขอการใช้บริการทรัพยากรจากผู้ใช้บริการ การขยายทรัพยากร (Scaling) เพื่อรองรับการทำงานนั้นมีอยู่ด้วยกัน 2 รูปแบบ คือ แบบการขยายตามแนวนอน เป็นการเพิ่มปริมาณเครื่องคอมพิวเตอร์และอีกรูปแบบ คือ การขยายตามแนวตั้งซึ่งเป็นการเพิ่มประสิทธิภาพให้กับเครื่องคอมพิวเตอร์ เช่น การเพิ่มหน่วยความจำหลัก เพิ่มหน่วยการประมวลผล เป็นต้น โดยทั่วไปแล้วระบบที่มีกลไกการขยายความสามารถในการรอบรับการใช้งาน (Scalability) จะประกอบด้วย 3 ส่วนหลัก ๆ ได้แก่ ส่วนการเก็บข้อมูลทรัพยากรต่างๆ ของการใช้งาน (Agent) ส่วนของการเฝ้าระวังการใช้ทรัพยากรของเครื่องคอมพิวเตอร์ (Monitoring) และส่วนของการตัดสินใจในการขยายระบบ (Decision Maker) [1] กระบวนการนี้จะเป็นตัวขับเคลื่อนให้มีการตัดสินใจในการขยายการรองรับการใช้งาน

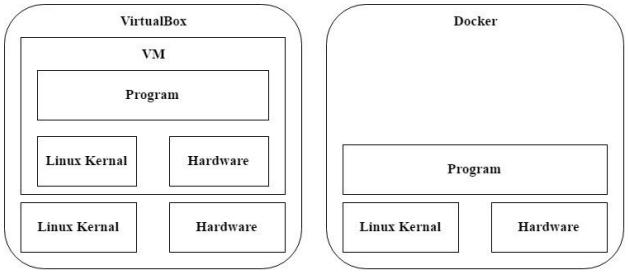
ดังนั้นด้วยความต้องการใช้งานบริการต่างๆ ผ่านทางอินเทอร์เน็ตที่มีมากขึ้นในยุคปัจจุบัน ทำให้ในบางครั้งการจัดการทรัพยากรต่างๆ ในรูปแบบการใช้งาน VM แบบเดิมอาจจะใช้เวลาค่อนข้างมากในการขยายการรองรับการให้บริการให้ตอบสนองความต้องการต่อการใช้งานได้ทันเวลา และเทคโนโลยีที่สะดวกและรวดเร็วต่อการสร้างคอมพิวเตอร์เสมือนด้วย LXC นั้นสามารถที่จะนำมาทดแทนการเพิ่มขยายในรูปแบบเดิมให้ตอบสนองต่อความต้องการของผู้ใช้งานและผู้ให้บริการ ดังนั้นผู้วิจัยจึงมีแนวคิดที่จะพัฒนากลไกการเพิ่มขยายนี้ให้มีรูปแบบการตรวจจับสถานะที่ต้องการทรัพยากรเพิ่มเติมจนถึงการสั่งงานให้มีการเพิ่มขยายการบริการแบบอัตโนมัติ เพื่อที่จะเป็นการลดค่าใช้จ่ายต่างๆ ที่จะเกิดขึ้นในส่วนของผู้ให้บริการได้และเพื่อให้มีการจัดสรรใช้งานทรัพยากรต่างๆ ของเครื่องคอมพิวเตอร์โดยเกิดประโยชน์สูงสุด

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 ด็อกเกอร์ (Docker)

ด็อกเกอร์ (Docker) คือ ลินุกซ์คอนเทนเนอร์ชนิดหนึ่งซึ่งถูกพัฒนาเป็นโปรเจกต์โอเพ่นซอร์ส ใช้หลักการเดียวกันกับ Linux Container โดยมีจุดประสงค์ที่จะช่วยให้นักพัฒนาสามารถสร้างแอพพลิเคชั่น นำแอพพลิเคชั่นไปติดตั้งและสั่งให้แอพพลิเคชั่นทำงานได้ง่ายขึ้น โดยที่คอนเทนเนอร์ที่เกิดขึ้นมาแต่ละตัวนั้นจะแยกการทำงานออกจากกันอย่างอิสระไม่ว่าจะเป็นโปรเซสการทำงานและระบบเครือข่ายโดยใช้ฮาร์ดแวร์ตัวเดียวกัน โดยด็อกเกอร์กำลังได้รับความนิยมในการพัฒนาและการนำไปใช้งานในหลาย ๆ ด้าน ด้วยเหตุผลที่ว่ามีความยืดหยุ่นและมีความเร็วในการทำงานที่ค่อนข้างสูง [2]

ด็อกเกอร์มีการทำงานในสถาปัตยกรรมแบบแม่ข่ายและลูกข่าย (Client-Server) โดยจะมีด็อกเกอร์ไคลแอนต์ (Docker client) ทำหน้าที่เชื่อมต่อกับระบบประมวลผลด็อกเกอร์ (Docker engine) เพื่อที่จะสั่งให้แอพพลิเคชั่นต่างๆ ทำงานได้ ตัวระบบประมวลผลด็อกเกอร์นั้นสามารถรองรับการสั่งงานผ่าน RESTful APIs จึงทำให้ง่ายต่อการสั่งงานทางไกลหรือพัฒนาแอพพลิเคชั่นต่างๆ มาเรียกใช้งาน ด็อกเกอร์จะทำงานได้ต้องมีส่วนประกอบต่างๆ ดังต่อไปนี้ Docker host, Docker client, Docker images, Docker registries, Docker Containers และ Dockerfile ซึ่งด็อกเกอร์มีคุณสมบัติที่แตกต่างกับ Virtual Machine ต้องที่มีขนาดเล็กกว่า ทำให้ใช้ทรัพยากรที่น้อยกว่าอีกด้วย แสดงดังภาพที่ 1



ภาพที่ 1: การเปรียบเทียบระหว่าง VM กับ Docker [2]

2.2 Kubernetes

Kubernetes เป็นระบบที่บริษัท Google พัฒนาขึ้นมาเพื่อใช้สำหรับจัดการเกี่ยวกับลินุกซ์คอนเทนเนอร์ที่อยู่ในระบบสภาพแวดล้อมแบบกลุ่ม (Cluster) โดยหลักแล้วจะถูกนำมาใช้ในขั้นตอนส่วนของการนำเอาลินุกซ์คอนเทนเนอร์ไปใช้งาน (Deployment) ซึ่ง Kubernetes จะจัดการเรื่องระบบเครือข่ายและจัดภาระการทำงานให้กับลินุกซ์คอนเทนเนอร์ [3] ผู้ใช้งานสามารถสั่งงานผ่านเครื่องสั่งงานด้วยโปรแกรม kubectl และการให้บริการต่าง ๆ จะให้บริการบนเครื่องสั่งงาน ซึ่งรายละเอียดของ Kubernetes นั้นสามารถทำงานได้ด้วยส่วนประกอบดังต่อไปนี้ Master server, Worker Server และหน่วยการทำงานของ Kubernetes [4]

2.2.1 เครื่องสั่งงาน (Master server)

เป็นหน่วยในการควบคุมการทำงานของ Kubernetes แบบกลุ่ม ซึ่งผู้ที่จะใช้งานจะสั่งงานต่าง ๆ ผ่านทางเครื่องสั่งงานนี้ไปยังเครื่องที่ทำหน้าที่รับงาน (Worker server) โดยเครื่องสั่งงานจะมีโปรแกรมที่คอยตอบโต้กับระบบ Kubernetes ดังนี้

2.2.1.1 Etcd

Etcd จะเหมือนกับส่วน Etcd ของ CoreOS เป็นการให้บริการการเก็บข้อมูลแบบkey-value ซึ่งแต่ละเครื่อง Server จะใช้การให้บริการนี้เป็นตัวกลางในการสื่อสารกันภายในกลุ่ม เช่น สถานการณ์แสดงผลต่าง ๆ ของเครื่อง Server ภายในกลุ่ม โดยการใช้งาน Etcd นั้นจะมีการเก็บข้อมูลในรูปแบบของ HTTP/JSON เมื่อต้องการเรียกดูข้อมูลก็จะเรียกข้อมูลของ key ที่ต้องการและ Etcd จะตอบกลับค่าที่ต้องการกลับมาให้

2.2.1.2 API Service Server

ส่วนของ API Service Server ถือว่าเป็นส่วนสำคัญของเครื่องสั่งและจัดการงานต่าง ๆ ของเครื่องภายในกลุ่มโดยจะอ่านค่าสถานะต่าง ๆ ขึ้นมาจากส่วน Etcd และการใช้งานการบริการ API Service Server นั้นสามารถใช้การเรียก RESTful ได้ ถือว่าเป็นข้อดีที่ทำให้นักพัฒนาสามารถที่จะพัฒนาเครื่องมือขึ้นมาติดต่อกับ API Service Server ได้เองโดยผ่าน RESTful และในงานวิจัยนี้ก็ได้พัฒนาโปรแกรมขึ้นมาด้วยการใช้ RESTful ติดต่อกับ API Service Server

2.2.1.3 Controller Manager Server

ส่วนการให้บริการในส่วนนี้จะใช้ในการจัดการงานเกี่ยวกับการสร้างคอนเทนเนอร์ โดยจะอ่านค่าสถานะต่าง ๆ ของคอนเทนเนอร์ขึ้นมาจากส่วน Etcd และส่วนนี้ก็จะมีหน้าที่ในการจัดการลดหรือขยายการให้บริการด้วยเช่นกัน

2.2.1.4 Scheduler Server

เป็นส่วนในการจัดการบริหารงานในแต่ละเครื่อง Server ภายในกลุ่ม ส่วนนี้จะตรวจสอบไม่ให้มีการใช้งานทรัพยากรเกินกว่าทรัพยากรว่างที่มีอยู่ เมื่อมีคอนเทนเนอร์เกิดขึ้นมาใหม่ส่วนนี้จะจัดการให้ว่าเครื่องใดจะเป็นเครื่องที่รองรับคอนเทนเนอร์ตัวที่เกิดขึ้นมาใหม่นี้

2.2.2 เครื่องรับงาน (Worker Server)

เครื่องรับงานจะทำหน้าที่รอรับงานจากเครื่องสั่งงานโดยจะเชื่อมต่อผ่านระบบเครือข่าย บนเครื่องรับงานนี้จะมีโปรแกรมด็อกเกอร์ทำงานอยู่ เมื่อมีการสร้างงานขึ้นมาใหม่จากเครื่องสั่งงานแล้วเครื่องรับงานจะดำเนินการสร้างด็อกเกอร์คอนเทนเนอร์ขึ้นตามที่ได้รับงานมา ประกอบด้วยส่วนการทำงานดังต่อไปนี้

2.2.2.1 Kubelet Service

Kubelet จะเป็นส่วนที่ติดต่อรับข้อมูลกับเครื่องสั่งงานว่าเครื่องสั่งงานนั้นสั่งให้สร้างคอนเนอร์แบบใดมาและจะคอยรับคำสั่งต่าง ๆ จากเครื่องสั่งงาน

2.2.2.2 Proxy Service

จะทำหน้าที่ส่งการร้องขอการให้บริการจากด็อกเกอร์โฮสต์เข้าไปยังด็อกเกอร์คอนเทนเนอร์และส่งผลลัพท์กลับคืน ซึ่งส่วนของ Proxy service นี้สามารถที่จะจัดการการให้บริการในรูปแบบการกระจายงานได้ (Load Balancing)

2.2.3 หน่วยการทำงานของ Kubernetes

ในการทำงานของแอพพลิเคชั่นคอนเทนเนอร์ที่ทำงานอยู่ในสภาพแวดล้อม Kubernetes จะประกอบด้วยหน่วยการทำงานดังต่อไปนี้

2.2.3.1 Pods

Pods คือชื่อเรียกของลินุกซ์คอนเทนเนอร์ที่มีแอพพลิเคชั่นทำงานอยู่ ในหนึ่ง Pods อาจจะมีลินุกซ์คอนเทนเนอร์อยู่เป็นจำนวนหลายตัวก็ได้แต่โดยทั่วไปแล้วจะนิยมให้มีแค่หนึ่งตัวในการรองรับการให้บริการของแอพพลิเคชั่น และในแอพพลิเคชั่นก็อาจจะมีจำนวนมากกว่าหนึ่ง Pods ขึ้นอยู่กับความสามารถในการรองรับการให้บริการของบริการนั้น ซึ่งในงานวิจัยนี้จะเน้นไปในการลดและขยายการให้บริการแบบอัตโนมัติด้วยการเพิ่มและลดจำนวน Pods ที่จะเข้ามารองรับการให้บริการ

2.2.3.2 Services

Services คือชื่อเรียกส่วนการเชื่อมต่อกับลินุกซ์คอนเทนเนอร์ โดยส่วนนี้จะนำส่งการร้องขอใช้บริการต่าง ๆ ไปให้ยังลินุกซ์คอนเทนเนอร์ที่อยู่ใน Pods อาจจะกล่าวอีกอย่างหนึ่งว่า Services คือส่วนที่เป็นอินเตอร์เฟสไว้เชื่อมต่อระหว่างผู้ใช้บริการกับลินุกซ์คอนเทนเนอร์ที่ให้บริการ

2.2.3.3 Replication Controllers

Replication Controllers เป็นส่วนที่คอยจัดการ Pods ในเรื่องของการขยายขนาดการให้บริการในแนวนอน (Horizontal) โดยปกติเมื่อมีการสร้าง Pods ขึ้นแล้วส่วนของ Replication Controller จะถูกเปิดการใช้งานขึ้นมาพร้อมกัน หน้าที่อีกหนึ่งอย่างของ Replication Controllers คือการจัดการในเรื่องรุ่นของแอพพลิเคชั่น (Version) ที่ทำงานในคอนเทนเนอร์ เช่นหากเดิมมี Pods ที่ให้บริการแอพพลิเคชั่นในรุ่นที่หนึ่งและหากผู้ที่ให้บริการต้องการนำแอพพลิเคชั่นในรุ่นที่สองใช้งานแทนที่แอพพลิเคชั่นรุ่นที่หนึ่งทำงานอยู่ ส่วนของ Replication Controllers จะนำแอพพลิเคชั่นรุ่นที่สองขึ้นมาทำงานและเมื่อตรวจสอบแล้วว่าแอพพลิเคชั่นรุ่นสองทำงานได้แล้ว Replication Controllers ก็จะดำเนินการปิดแอพพลิเคชั่นรุ่นหนึ่งลงโดยอัตโนมัติ

2.3 ระบบฐานข้อมูลแบบอนุกรมเวลา (Time Series Databases)

ฐานข้อมูลที่มีความสัมพันธ์กับอันตรภาคเวลา โดยทั่วไปแล้วนิยมใช้กับการเก็บข้อมูลที่ต้องมีช่วงเวลามาเป็นตัวชี้วัดระดับ สำหรับงานวิจัยนี้ได้เลือกใช้ระบบจัดการฐานข้อมูลแบบอนุกรมเวลาที่มีชื่อว่า InfluxDB ซึ่งเป็นโปรแกรมโอเพ่นซอร์สที่มีความสามารถจัดการฐานข้อมูลแบบอนุกรมเวลา โดยสามารถจัดเก็บข้อมูลได้หลากหลายและมีความยืดหยุ่นในการเรียกดูหรือสืบค้นข้อมูลขึ้นมาแสดงผล การใช้งานโปรแกรมจัดการฐานข้อมูลอนุกรมเวลา InfluxDB สามารถใช้คำสั่ง SQL เหมือนกับการใช้งานโปรแกรมจัดการฐานข้อมูลเชิงสัมพันธ์ทั่วไปในการเรียกดูข้อมูลขึ้นมาแสดงผล และยังสามารถให้ผู้ใช้งานติดต่อกับฐานข้อมูลผ่านทาง HTTP RESTful ได้อย่างสะดวก [5]

2.4 การขยายขนาดการให้บริการ

2.4.1 การขยายการรองรับการใช้งานในยุคก่อนและหลังการใช้ Cloud

การขยายการรองรับการใช้งาน ได้มีการศึกษาและดำเนินการมาอย่างยาวนาน ปัจจัยที่มีผลต่อการพิจารณาการขยายรองรับการใช้งานนั่นคือค่าการทำงานของหน่วยประมวลผลและความสามารถของการใช้งานเครือข่ายข้อมูล (Bandwidth) เพื่อที่จะรองรับการทำงานของแอพพลิเคชั่นต่างๆ ได้อย่างต่อเนื่องและมีประสิทธิภาพ หากปัจจัยที่ได้กล่าวมานั้นไม่เพียงพอต่อการใช้งานก็จะดำเนินการจัดซื้อจัดจ้างอุปกรณ์เข้ามาเสริมใหม่เพื่อรักษาความพร้อมให้บริการอยู่เสมอ ต่อมาในยุคของการเทคโนโลยีเครื่องคอมพิวเตอร์เสมือนซึ่งได้มีการใช้งานเทคโนโลยีคลาวด์เข้าด้วยนั้นยังคงมีการวัดประสิทธิภาพของหน่วยประมวลผลเช่นเดิม แต่จะมีการกำหนดช่วงของค่าที่เหมาะสมในการดำเนินการขยายความสามารถ คือ มีการตั้งค่าเส้นขีดบน (Upper Scale) และเส้นขีดล่าง (Lower Scale) [6] หากค่าของหน่วยประมวลผลเกินเส้นขีดบนเมื่อไรก็จะทำการขยายเครื่องคอมพิวเตอร์เพื่อรองรับการใช้บริการและหากค่าของหน่วยประมวลผลน้อยกว่าเส้นขีดล่างเมื่อไรก็จะทำการลดจำนวนของเครื่องคอมพิวเตอร์ลงเพื่อรักษาค่าใช้จ่ายต่าง ๆ ที่ได้เกิดขึ้นจากผู้ให้บริการคลาวด์

2.4.2 อัลกอริทึมในการขยายเครื่องคอมพิวเตอร์แบบอัตโนมัติ

อัลกอริทึมการขยายความสามารถของเครื่องคอมพิวเตอร์ขณะทำงานอยู่ [7] ได้นำตัวแปรในการตัดสินใจขยายความสามารถ เช่น ระยะเวลาของ SLA ลักษณะงาน และค่าประสิทธิภาพของหน่วยประมวลผลมาพิจารณา เป็นต้น เนื่องจากงานวิจัยดังกล่าวเกี่ยวข้องกับการขยายความสามารถ

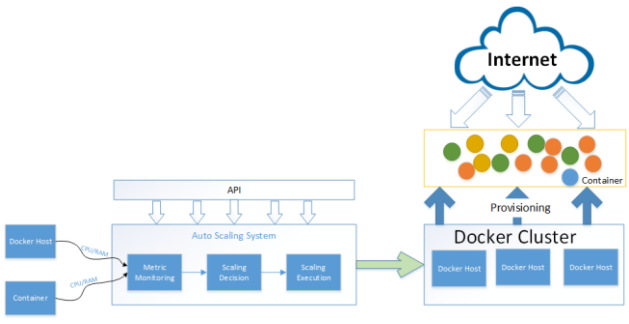
ของเครื่องคอมพิวเตอร์ตามลักษณะงานที่เข้ามาเป็นลำดับขั้น (Workflow) จึงต้องคำนึงถึงลักษณะงานและจัดลำดับความสำคัญของงานในการขยายความสามารถ ซึ่งอัลกอริทึมดังกล่าวได้ทดสอบใช้งานบน Amazon AWS EC2 พบว่าสามารถทำงานได้ตามที่ผู้วิจัยได้กำหนด สามารถที่จะเพิ่มและลดขนาดความสามารถของเครื่องคอมพิวเตอร์ที่ใช้ประมวลผลได้ตามระยะเวลาที่ต้องการ อันเนื่องมาจากข้อตกลง SLA แล้วค่าใช้จ่ายที่ได้เกิดขึ้นบนการใช้บริการคลาวด์

3. วิธีการดำเนินงาน

การออกแบบและพัฒนาระบบการขยายขนาดการให้บริการแบบอัตโนมัติในระบบด็อกเกอร์สามารถที่จะใช้งานระบบให้ทำงานร่วมกับการขยายขนาดของด็อกเกอร์คอนเทนเนอร์จำนวนหลายตัวที่อยู่ในระบบด็อกเกอร์โฮสต์เดียวกัน ผู้วิจัยได้แบ่งขั้นตอนในการดำเนินงานวิจัยออกเป็น 6 ขั้นตอน ดังต่อไปนี้ คือ

1) ออกแบบโครงสร้างของระบบด็อกเกอร์ แสดงดังภาพที่ 2 ซึ่งประกอบด้วย

- Agent คือ ส่วนการเก็บข้อมูลทรัพยากรใช้งาน Metric
- Monitoring คือ ส่วนของการจัดเก็บค่าสถิติต่างๆ จากตัว Agent โดยออกแบบให้เก็บเป็นฐานข้อมูลแบบอนุกรมเวลา (Time Series Database)
- Scaling Decision คือ ส่วนของการตัดสินใจในการทำงานว่าควรจะดำเนินการขยายเพื่อรองรับการให้บริการหรือไม่ตามความเหมาะสมจากค่าที่ได้กำหนดไว้
- Scaling Execution คือ ส่วนของการดำเนินการสั่งงานให้มีการขยายเพื่อรองรับการให้บริการและมีการจัดสรรทรัพยากรได้อย่างเหมาะสม



ภาพที่ 2: โครงสร้างระบบรองรับการขยายตัวแบบอัตโนมัติในระบบด็อกเกอร์

2) ออกแบบส่วนในการเก็บข้อมูลการใช้งานทรัพยากรและการเฝ้าสังเกตข้อมูลการใช้งาน

ในส่วนของการเก็บข้อมูลทรัพยากรการใช้งานจะเป็นระบบที่มีชื่อว่า Heapster ซึ่งระบบ Heapster คือ ส่วนในการจัดการกับค่าของทรัพยากรต่างๆ โดยจะอ่านค่าจาก API ของด็อกเกอร์โฮสต์นั้นว่าแต่ละตัวมีการใช้งานทรัพยากรเป็นอย่างไร โดยดำเนินการอ่านค่าข้อมูลทรัพยากรการใช้งานทุก ๆ 1 นาที เมื่ออ่านค่าได้สำเร็จจะส่งค่าออกไปเก็บยังฐานข้อมูล ในที่นี้ได้เลือกใช้โปรแกรมจัดการระบบฐานข้อมูลแบบอนุกรมเวลา

ในส่วนของการเฝ้าสังเกตข้อมูลการใช้งานนั้นได้ออกแบบให้มีโปรแกรมที่ทำงานเป็นรอบเวลาโดยจะดึงข้อมูลเพื่อตรวจสอบทุก ๆ 1 นาที โดยจะทำการดึงข้อมูลออกจากฐานข้อมูลแบบอนุกรมเวลาที่ใช้ InfluxDB เป็นโปรแกรมจัดการฐานข้อมูล ซึ่งส่วนของการเก็บข้อมูลจะเขียนข้อมูลการใช้งานทรัพยากรลงในฐานข้อมูลนี้ โดยจะต้องพัฒนาส่วนในการดึงข้อมูลมาประมวลผลว่าสมควรที่จะลดหรือขยายขนาดการให้บริการในช่วงเวลานั้นหรือไม่

3) ออกแบบส่วนการดำเนินการขยายหรือลดขนาดการให้บริการ เมื่อรับค่าข้อมูลจากการเฝ้าสังเกตแล้ว นำมาทำการคำนวณหาค่าของ CPU ที่ใช้งานในช่วงเวลานั้น ดังสมการที่ 3-1

$$cpu_i = \frac{cpu_time_{(i,t)} - cpu_time_{(i,t-1)}}{60 \times 10^9} \times core_i \qquad (3\text{-}1)$$

โดยที่ cpu_i คือ ค่าของการใช้งาน CPU ของ pod_i

$core_i$ คือ จำนวน CPU ที่ pod_i ใช้งานมีหน่วยเป็นมิลลิคอร์(millicore)

$cpu_time_{(i,t)}$ คือ ค่าของการใช้งาน CPU ณ เวลา t หน่วยเป็นไมโครวินาที(microsecond)

$cpu_time_{(i,t-1)}$ คือ ค่าของการใช้งาน CPU ณ เวลา t-1 หน่วยเป็นไมโครวินาที(microsecond)

จากนั้นทำการคำนวณหาค่า CPU เฉลี่ยของทุก Pods ที่ให้บริการในช่วงเวลานั้น เพื่อนำไปในการจำนวน Pods ดังสมการที่ 3-2

$$cpu_{avg} = \frac{\sum_{i=1}^{n_{current}} cpu_i}{n_{current}} \qquad (3\text{-}2)$$

โดยที่ cpu_{avg} คือ ค่าเฉลี่ยของการใช้งานหน่วยประมวลผลของ Pods นั้น

$n_{currect}$ คือ จำนวน Pods ที่ให้บริการอยู่ในปัจจุบัน

cpu_i คือ ค่าของการใช้งาน CPU ของแต่ละ Pods

หลังจากนั้นนำค่าที่ได้ไปคำนวณหาจำนวน Pods ดังสมการที่ 3-3

$$n_{new} = \left\lceil \frac{cpu_{avg}}{cpu_{target}} \right\rceil \times n_{current} - n_{current} \quad (3\text{-}3)$$

โดยที่ n_{new} คือจำนวนของ Pods ใหม่ที่จะดำเนินการขยายหรือลดขนาด

$n_{currect}$ คือ จำนวนของ Pods ในปัจจุบันที่ให้บริการ

cpu_{avg} คือ ค่าเฉลี่ยของการใช้งานหน่วยประมวลผลของ Pods นั้น

cpu_{target} คือค่าของหน่วยประมวลผลที่ผู้ใช้งานได้ใส่ค่าเข้ามาว่าจะต้องอยู่ในเกณฑ์ที่เท่าไร

4) ทดสอบระบบ จะเป็นการทดสอบการทำงานของระบบโดยการจำลองการใช้งานการบริการเพื่อสังเกตพฤติกรรมในการขยายขนาดการให้บริการโดยอัตโนมัติโดยพิจารณาว่าระบบสามารถทำงานได้ถูกต้องตามที่ได้วางแผนการทดสอบได้หรือไม่ โดยทำการจำลองเหตุการณ์ออกเป็น 3 เหตุการณ์ ดังนี้ เหตุการณ์ที่ 1 ไม่มีการใช้บริการ จำนวนผู้ใช้เท่ากับ 0 เหตุการณ์ที่ 2 มีการใช้บริการ จำนวนผู้ใช้บริการเท่ากับ 1600 หน่วงเวลาการใช้บริการ 5 นาที และเหตุการณ์ที่ 3 หยุดการใช้บริการ จำนวนผู้ใช้บริการเท่ากับ 0

5) ประเมินประสิทธิภาพ จะเป็นการประเมินประสิทธิภาพของการรองรับการให้บริการด้วยการขยายขนาดการให้บริการได้สัมพันธ์กับปริมาณการใช้บริการ โดยจำลองผู้ใช้งานเริ่มต้นที่จำนวน 800 ผู้ใช้งานต่อวินาที ทำการเรียกใช้บริการเว็บไซต์เป็นช่วงเวลาต่อเนื่อง และพิจารณาว่าการเรียกใช้งานเว็บไซต์นั้นสามารถแสดงผลได้อย่างถูกต้องไม่มีข้อผิดพลาด

4. ผลการดำเนินงาน

4.1 ผลการทดสอบส่วนการดำเนินการขยายขนาดการให้บริการ

ทดสอบการหาค่าจำนวนการขยายขนาดการให้บริการโดยอาศัยค่าที่ได้จากการคำนวณการใช้งานหน่วยประมวลผลของ Pods การทดสอบในกระบวนการนี้ได้ตั้งค่าเอาไว้ว่าเมื่อมีการใช้งานหน่วยประมวลผลของด็อกเกอร์คอนเทนเนอร์และ Pods มากกว่าร้อยละ 30 ให้ดำเนินการสั่งการขยายขนาดการให้บริการ แต่ในกระบวนการทดสอบนี้ติดปัญหาที่เครื่องคอมพิวเตอร์ที่ได้นำมาทำเป็นระบบคลัสเตอร์นั้นมีทรัพยากรไม่เพียงพอต่อการสร้างด็อกเกอร์คอนเทนเนอร์ที่มีจำนวนมากกว่า 40 คอนเทนเนอร์ จึงจำเป็นต้องจำกัดการขยายขนาดที่มากที่สุด คือ 40 คอนเทนเนอร์เพื่อให้โปรแกรมยังคงสามารถทำงานต่อได้

ตารางที่ 1: การทดสอบการทำงานการปรับขนาดการให้บริการ

นาที	CPU	จำนวน	เหตุการณ์	จำนวนผู้ใช้บริการ
0	0	1	ไม่มีการใช้บริการ	0
1	100	4	จำลองการใช้บริการ	1600
2	ช่วงหน่วงเวลาการขยายขนาด	4		1600
3		4		1600
4		4		1600
5	42.256	6		1600
6	ช่วงหน่วงเวลาการขยายขนาด	6		1600
7		6		1600
8		6		1600
9	29.098	6	จำลองการใช้บริการ	1600
10	29.039	6		1600
11	27.163	6	หยุดการจำลองการใช้บริการ	0
12	0	1		0
13	ช่วงหน่วงเวลาการลดขนาด	1		0
14		1		0
15		1		0
16		1		0

ตารางที่ 1 แสดงผลการทำงานเมื่อกำหนดให้มีการขยายขนาดการให้บริการอัตโนมัติเมื่อกำหนดให้มีการให้บริการเว็บไซต์ด้วยโปรแกรม Nginx บนด็อกเกอร์คอนเทนเนอร์หรือ Pods ที่มีจำนวนเท่ากับ 1 และเมื่อมีการใช้งานหน่วยประมวลผลมากกว่าร้อยละ 40 ให้ทำการขยายขนาดการให้บริการอัตโนมัติและเมื่อการใช้งานหน่วยประมวลผลน้อยกว่าร้อยละ 30 ให้มีการลดขนาดการให้บริการโดยเป็นไปตามค่าที่ได้จากการคำนวณ

จากการทดสอบในนาทีที่ 1 ได้มีการจำลองการเรียกใช้บริการจำนวนมากด้วยการจำลองให้มีผู้ใช้งานจำนวน 1600 หน่วยและทำการเรียกการใช้งานต่อเนื่องเป็นเวลา 5 นาที ผลลัพธ์ คือ ระบบได้มีการสั่งให้เพิ่มจำนวนของด็อกเกอร์คอนเทนเนอร์ขึ้น

เป็น 4 คอนเทนเนอร์และหลังจากนั้นจะมีช่วงเวลาในการหน่วงเวลาเพื่อรอการสร้างด็อกเกอร์คอนเทนเนอร์และความพร้อมในการเก็บข้อมูลการใช้งานทรัพยากร ในนาทีที่ 10 ได้มีการตรวจสอบใหม่อีกครั้งพบว่ามีการใช้งานทั้ง 4 ด็อกเกอร์คอนเทนเนอร์มีการใช้งานเฉลี่ยแล้วร้อยละ 42.256 จำเป็นต้องขยายขนาดการให้บริการเพิ่มขึ้นไปเป็น 6 คอนเทนเนอร์ ในนาทีที่ 12 ได้มีการตรวจสอบการใช้งานทรัพยากรอีกครั้งพบว่ามีการใช้งานหน่วยประมวลของด็อกเกอร์คอนเทนเนอร์ทั้ง 6 คอนเทนเนอร์เฉลี่ยอยู่ที่ร้อยละ 0 ประกอบมีการสิ้นสุดการเรียกใช้บริการจึงทำให้ไม่มีการใช้งานหน่วยประมวลผล เมื่อระบบตรวจสอบแล้วได้ผลว่าลดขนาดจำนวนการให้บริการลงเหลือ 1 คอนเทนเนอร์ และมีช่วงหน่วงเวลาการลดขนาดอยู่ที่ 4 นาที และหลังจากนาทีที่ 16 เป็นต้นไปนั้นไม่มีการใช้งานจึงคงจำนวนด็อกเกอร์คอนเทนเนอร์อยู่ที่ 1 คอนเทนเนอร์เช่นเดิม

4.2 ทดสอบความพร้อมของการให้บริการ

การทดสอบความสามารถในการให้บริการเว็บไซต์ด้วยโปรแกรม Nginx บนด็อกเกอร์คอนเทนเนอร์เป็นการทดสอบที่มีจุดประสงค์คือระบบการให้บริการต่าง ๆ ภายในระบบคลัสเตอร์เดียวกันจะต้องมีความพร้อมให้บริการตลอดเวลาและมีความราบรื่นในการเพิ่มหรือขยายขนาดการให้บริการซึ่งจะต้องเกิดข้อผิดพลาดให้น้อยที่สุดมี การทดสอบใช้เครื่องมือในการจำลองผู้ใช้งานด้วยการใช้โปรแกรม Apache Jmeter

การทดสอบได้จำลองผู้ใช้งานให้เพิ่มขึ้นและลดลงตามช่วงเวลาเพื่อทดสอบความพร้อมในการให้บริการของด็อกเกอร์คอนเทนเนอร์โดยกำหนดให้มีการขยายขนาดการให้บริการเมื่อมีระดับการใช้งานหน่วยประมวลผลมากกว่าร้อยละ 30 ของการใช้งาน ซึ่งการให้บริการนี้จะให้บริการเว็บไซต์และทดสอบด้วยการใช้โปรแกรมจำลองผู้ใช้งานเริ่มต้นที่จำนวน 800 ผู้ใช้งานต่อวินาทีโดยให้เรียกใช้บริการเว็บไซต์เป็นช่วงเวลาต่อเนื่องไป โดยจะมีการวัดผลการให้บริการคือ การตอบสนองสำเร็จหมายความว่าการเรียกใช้งานเว็บไซต์สามารถแสดงผลได้อย่างถูกต้องไม่มีข้อผิดพลาด (มีการแสดงผลการตอบสนองของรหัส HTTP เป็น 200 หมายความว่าสำเร็จ) เกิดขึ้นและการตอบสนองล้มเหลวหมายความว่าการเรียกใช้งานเว็บไซต์นั้นให้การแสดงผลที่ผิดพลาด เช่น การบริการมีความไม่พร้อมให้บริการ (มีการแสดงผลการตอบสนองของรหัส HTTP 503 หมายความว่าไม่พร้อมให้บริการ) มีผลที่ได้ดังต่อไปนี้

ตารางที่ 2: ผลการทดสอบการให้บริการ

เวลา (นาที)	จำนวนผู้ใช้งาน	ค่าเฉลี่ยการใช้ CPU	จำนวน Pods ปัจจุบัน	จำนวน Pods ที่ต้องขยาย/ลดขนาด (Target)	สำเร็จ (%)
1	800	0	1	1	100
2	800	43.727	1	2	100
5	800	55.653	2	4	100
8	1200	36.264	4	5	100
11	1200	28.54	5	5	100
12	1200	24.1	5	5	100
13	100	17.28	5	3	100
17	100	7.078	3	1	99.99
21	100	20.25	1	1	100
22	100	19.48	1	1	100
23	100	17.603	1	1	100
24	1500	100	1	4	100
27	1500	42.518	4	6	100
30	1500	32.339	6	7	100
33	1	0.27	7	1	100
37	1	0.27	1	1	100

จากตารางที่ 2 ได้แสดงผลของการทดสอบความพร้อมของการให้บริการในการใช้กลไกการขยายขนาดการให้บริการอัตโนมัติในระบบด็อกเกอร์โดยแสดงผลตามเวลา (นาที) จำนวนผู้ใช้งานจะเป็นค่าที่ถูกจำลองขึ้นมาจากโปรแกรมจำลองผู้ใช้งาน การใช้งานหน่วยประมวลผลจะเป็นค่าเฉลี่ยของหน่วยประมวลผลของด็อกเกอร์คอนเทนเนอร์ที่ทำงานในการให้บริการนั้น จำนวน Pods ปัจจุบัน คือ จำนวนของด็อกเกอร์คอนเทนเนอร์ที่ทำงานในการให้บริการอยู่ในช่วงเวลานั้น จำนวน Pods ที่ต้องขยายหรือลดขนาดคือจำนวนที่ได้จากการตัดสินใจในการขยายหรือลดขนาดของด็อกเกอร์คอนเทนเนอร์ ค่าสำเร็จคืออัตราส่วนของจำนวนการเรียกใช้บริการและได้ผลสำเร็จต่อจำนวนการเรียกใช้บริการทั้งหมด ค่าล้มเหลว คืออัตราส่วนของจำนวนการเรียกใช้บริการและได้ผลล้มเหลวต่อ

จำนวนการเรียกใช้บริการทั้งหมดค่าที่ได้จากตารางที่ 2 สามารถอธิบายได้ดังต่อไปนี้ ในช่วงที่มีการเพิ่มขนาด

การให้บริการต่อผู้ใช้บริการที่เพิ่มขึ้นเพื่อให้สามารถรองรับการทำงานได้มีประสิทธิภาพในนาทีที่ 2, 5, 8 และ 24 จะพบว่ามีค่าเฉลี่ยของการใช้งานหน่วยประมวลผลที่เพิ่มขึ้นและมีการขยายขนาดของด็อกเกอร์คอนเทนเนอร์เพิ่มขึ้นตามจำนวนที่ได้จากการคำนวณและเมื่อทดสอบความสามารถในการให้บริการพบว่ามีอัตราสำเร็จร้อยละ 100 นั่นหมายความว่าเมื่อขยายขนาดการให้บริการจะไม่มีผลกระทบต่อการให้บริการที่ได้มีการให้บริการอยู่ก่อนหน้านี้และในช่วงที่มีการลดขนาดการให้บริการในนาทีที่ 13, 17 และ 33 ในนาทีที่ 17 จะพบว่าอัตราความพร้อมในการให้บริการมีค่าอยู่ที่ร้อยละ 99.99 สาเหตุที่เป็นเช่นนี้เพราะในการสร้างการให้บริการ (Service) ใน Kubernetes ได้มีการกำหนดให้มีลักษณะเป็นการกระจายงาน (Load balancing)ให้กับด็อกเกอร์คอนเทนเนอร์ที่อยู่ภายใต้ Replication Controller และในการลดจำนวนของด็อกเกอร์คอนเทนเนอร์ภายในการให้บริการนี้มีโอกาสที่จะพบช่วงที่กำลังลบด็อกเกอร์คอนเทนเนอร์ออกไปจึงทำให้ตัวกระจายงานนั้นจ่ายงานให้กับด็อกเกอร์คอนเทนเนอร์ที่กำลังโดนลบแต่เมื่อจ่ายงานให้แล้วไม่มีการให้บริการจึงได้ตอบสนองเป็นไม่พร้อมให้บริการ (รหัส HTTP คือ 503) ดังนั้นย่อมมีโอกาสที่จะเกิดเหตุการณ์เช่นนี้จึงทำให้การตอบสองความพร้อมในการให้บริการนั้นมีความผิดพลาดอยู่ร้อยละ 0.01 ในช่วงเวลาดังกล่าว ในการทดสอบความพร้อมในการให้บริการเมื่อใช้กลไกการขยายขนาดการให้บริการอัตโนมัติในระบบด็อกเกอร์นั้นมีความสามารถที่จะให้บริการได้โดยมีความผิดพลาดเฉพาะระหว่างการลดขนาดของด็อกเกอร์คอนเทนเนอร์เพียงเล็กน้อยเท่านั้น

5. สรุป

งานวิจัยนี้ผู้วิจัยได้ตรวจจับการใช้งานทรัพยากรหน่วยประมวลผลของด็อกเกอร์คอนเทนเนอร์และนำค่าการใช้งานหน่วยประมวลผลมาเข้าสู่ระบบการเฝ้าระวังจากด็อกเกอร์คอนเทนเนอร์หลาย ๆ ตัวในระบบการทำงานแบบกลุ่มและตรวจสอบการใช้งานหากมีการใช้งานทรัพยากรมากกว่าค่าที่ผู้ใช้ได้ระบุในระบบแล้วส่วนของการขยายขนาดการให้บริการก็จะดำเนินการต่อไปโดยอาศัยค่าที่ได้จากการเฝ้าระวังมาคำนวณเพื่อตัดสินใจว่าจะขยายหรือลดขนาดการให้บริการเป็นจำนวนเท่าไรในระบบด็อกเกอร์

ผลการทดสอบการทำงานของระบบพร้อมกับทดสอบความพร้อมของการให้บริการเมื่อมีการขยายหรือลดขนาดแบบอัตโนมัติผลที่ได้ คือ ระบบมีการลดและขยายขนาดเป็นไปตามการใช้งานหน่วยประมวลผลและยังมีความพร้อมในการให้บริการของส่วนการให้บริการภายในด็อกเกอร์โดยที่มีความผิดพลาดเพียงเล็กน้อยเท่านั้น และในส่วนของกระบวนการเฝ้าสังเกตข้อมูลที่เพิ่มขึ้นมานั้น การใช้งาน CPU จะขึ้นอยู่กับจำนวน Pods หากมีการใช้งานของ Pods มากก็ส่งผลทำให้การใช้งาน CPU มากขึ้นเช่นกัน สำหรับงานวิจัยนี้มีการใช้งานจำนวน Pods ไม่มาก ทำให้การใช้งาน CPU น้อยและไม่มีนัยสำคัญ

เอกสารอ้างอิง

[1] Nikravesh, A.Y., S.A. Ajila, and L. Chung-Horng. Measuring Prediction Sensitivity of a Cloud Auto-scaling System. in Computer Software and Applications Conference Workshops (COMPSACW), 2014 IEEE 38th International. 2014.

[2] Joy, A.M. Performance comparison between Linux containers and virtual machines. in Computer Engineering and Applications (ICACEA), 2015 International Conference on Advances in. 2015.

[3] Kubernetes. What is Kubernetes? Available online at https://kubernetes.io/docs/concepts/overview/what-is-kubernetes/

[4] Key Concepts of Kubernetes. Available online at http://blog.arungupta.me/key-concepts-kubernetes/

[5] influxdata. Available online at https://www.influxdata.com/

[6] Han, R., et al., Lightweight Resource Scaling for Cloud Applications, in Proceedings of the 2012 12th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (ccgrid 2012). 2012, IEEE Computer Society. p. 644-651.

[7] Heinze, T., et al., Auto-scaling techniques for elastic data stream processing, in Proceedings of the 8th ACM International Conference on Distributed Event-Based Systems. 2014, ACM: Mumbai, India. p. 318-321.

สถาปัตยกรรมสำหรับระบบบันทึกข้อมูลบนคลาวด์แบบหลายผู้ให้บริการ
An architecture for data storage system on multiple cloud service providers

มนัสชัย ปรีชาหาญ (Manutchai Preechahan)[1] และสุชา สมานชาติ(Sucha Smanchat)[2]
ภาควิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
[1]s5807011858150@email.kmuntb.ac.th, [2]sucha.smanchat@acm.org

บทคัดย่อ

การใช้ระบบคลาวด์ในการจัดเก็บไฟล์บนอินเตอร์เน็ตเป็นอีกทางเลือกหนึ่งที่มีประสิทธิภาพและน่าเชื่อถือ โดยผู้ให้บริการมีพื้นที่สำหรับจัดเก็บข้อมูลโดยไม่มีค่าใช้จ่ายได้ปริมาณหนึ่ง เมื่อมีความต้องการพื้นที่จัดเก็บไฟล์เพิ่มขึ้น การใช้งานหลายผู้ให้บริการเป็นวิธีหนึ่งที่เพิ่มพื้นที่การจัดเก็บข้อมูลได้ แต่การจัดเก็บไฟล์โดยใช้ผู้ให้บริการหลายแห่งอาจสร้างความสับสนในการใช้งานได้

งานวิจัยนี้จึงนำเสนอสถาปัตยกรรมสำหรับระบบบันทึกข้อมูลบนคลาวด์แบบหลายผู้ให้บริการ (Google Drive, Dropbox และ Amazon S3) โดยทำงานบนระบบ Docker โดยที่ขนาดไฟล์ที่อัพโหลดและพื้นที่คงเหลือของแต่ละผู้ให้บริการเป็นค่าในการตัดสินใจแบ่งไฟล์และเลือกผู้ให้บริการ ผู้วิจัยได้พัฒนาระบบต้นแบบขึ้นเพื่อทดสอบสถาปัตยกรรมนี้ และจากการทดสอบตามสถานการณ์ที่กำหนด ระบบที่ได้พัฒนาตามสถาปัตยกรรมที่นำเสนอสามารถทำงานได้อย่างถูกต้องโดยไม่เกิดความสูญเสียของไฟล์ ช่วยให้ผู้ใช้สามารถใช้งานพื้นที่เก็บไฟล์บนผู้ให้บริการคลาวด์หลายแห่งได้โดยควบคุมเพียงที่เดียว

คำสำคัญ: คลาวด์ จัดเก็บไฟล์ สถาปัตยกรรม ด็อกเกอร์

Abstract

Nowadays, people use the computer system to their keep datas. Nonetheless, there are many risks such as virus and ransomware. Therefore, making backup copies of data on on the Internet is one of the best way to avoid such problems. Usually, cloud storage providers offer a limited free storage space. As the need for the storage grows, using free storage from multiple cloud providers is one of the options. However, managing files stored on multiple cloud providers can cause confusing making it harder to use.

This research proposes an architecture for data storage system on multiple cloud service providers (Google Drive, Dropbox and Amazon S3) hosted on Docker platform. The decision of splitting file and selecting cloud provider for storing file fragments are based on the size of the whole file and the available free space on cloud providers. For the evaluation, a prototype system is implemented realizing the proposed architecture. From the evaluation of different scenarios, the prototype system performs correctly without losing file integrity. This propose architecture thus help to manage files stored on multiple cloud providers from a single point.

Keyword: Cloud, Storage, Architecture, Docker

1. บทนำ

ปัจจุบันการจัดเก็บไฟล์บนอินเตอร์เน็ต ได้มีผู้ให้บริการเป็นจำนวนมาก เช่น Google Drive, Dropbox และ Amazon S3 เป็นต้น การใช้บริการจัดเก็บไฟล์บนอินเตอร์เน็ตจึงได้รับความนิยม โดยผู้ให้บริการแต่ละรายจะมีพื้นที่จัดเก็บไฟล์ประมาณหนึ่งให้ โดยไม่มีค่าบริการ หากผู้ใช้ต้องการเพิ่มพื้นที่ต้องชำระค่าบริการเป็นรายเดือน หากผู้ใช้ไม่ต้องการชำระค่าบริการและต้องการพื้นที่การเก็บไฟล์เพิ่มขึ้น อาจทำได้โดยจะเป็นการใช้บริการจัดเก็บไฟล์บนอินเตอร์เน็ตจากผู้ให้บริการหลายรายเพื่อให้ได้พื้นที่จัดเก็บไฟล์เพิ่มมากขึ้น โดยแต่ละผู้ให้บริการจะ

มีหน้าจอการใช้งานที่แตกต่างกัน และขณะใช้งานร่วมกันจะต้องเปลี่ยนหน้าต่างการใช้งานไปมา ทำให้ผู้ใช้งานเกิดความสับสน

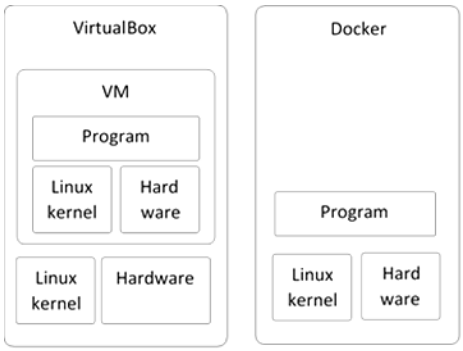
จากปัญหาดังกล่าว ทำให้เกิดแนวคิดพัฒนาวิธีการหรือระบบที่ผู้ใช้สามารถจัดการพื้นที่จัดเก็บไฟล์จากแต่ละผู้ให้บริการ ผ่านระบบที่สร้างขึ้น เพื่ออำนวยความสะดวกและง่ายต่อการจัดการข้อมูลการจัดเก็บไฟล์บนระบบ

งานวิจัยนี้จึงมีจุดประสงค์เพื่อนำเสนอสถาปัตยกรรมสำหรับจัดเก็บไฟล์บนอินเตอร์เน็ตโดยใช้บริการจัดเก็บไฟล์จากหลายผู้ให้บริการ

2. เทคโนโลยีและงานวิจัยที่เกี่ยวข้อง

2.1 ด็อกเกอร์ (Docker)

Docker [1]เป็นโปรแกรมที่ใช้ทำ Virtualization มีคุณสมบัติแตกต่างกับ Virtual Machine (เช่น VirtualBox) ตรงที่มีขนาดเล็กกว่าเพราะ Docker ทำ Virtualization บนระดับ Kernel ทำให้มีขนาดที่เล็กกว่า Virtual Machine



ภาพที่ 1: ความแตกต่างของ Docker และ VirtualBox

จากภาพที่ 1 Docker เป็น Software ที่ใช้ทำงานบนระบบปฏิบัติการแตกต่างกับ Virtualbox ที่ต้องจำลองทรัพยากรระบบ เช่น หน่วยประมวลผล หน่วยความจำ การเชื่อมต่อเครือข่าย เป็นต้น ทำให้ระบบที่ใช้งาน Virtualbox ต้องใช้ทรัพยากรของระบบมากกว่า Docker

Docker มีหลักการออกแบบการทำงานคือ “1 Container 1 Task” Container หมายถึง ระบบที่ได้จำลองสภาพแวดล้อมในการทำงาน เพื่อใช้ในการประมวลผลเพียงอย่างเดียว และทำการติดต่อกันผ่าน RESTFul Web Service ภายในเพื่อให้การออกแบบระบบมีความยืดหยุ่น สามารถเพิ่มการประมวลผลในแต่ละส่วนได้

2.2 Node.js

Node.js [2] เป็นสภาพแวดล้อม (Environment) ที่ใช้ในการพัฒนาแอปพลิเคชันโดยใช้ภาษา Javascript และ google JavaScript engine V8 เป็นคอมไพเลอร์ google JavaScript engine V8 มีความแตกต่างจากคอมคอมไพเลอร์อื่นที่ปรับกระบวนการทำงานจากทำงานทีละคำสั่ง (ทำงานแรกเสร็จแล้วถึงทำงานบรรทัดถัดไป) เป็นนำงานที่ต้องการประมวลผลมาเข้าลำดับคิวเวลาและทำงานตามคิวเวลาแทน ทำให้งานที่ทำอยู่อาจไม่เสร็จแต่คอมไพเลอร์ทำงานถัดไปทันที ดังนั้นการแก้ปัญหาดังกล่าวต้องใช้การเขียนฟังก์ชัน Callback

Callback คือฟังก์ชันรอการทำงานเสร็จ เมื่องานที่ทำอยู่เสร็จจะส่งค่าให้ฟังก์ชันอื่นนำผลลัพธ์ที่ได้ไปทำงานต่อ ยกตัวอย่างเช่น ฟังก์ชันการอัพโหลดไฟล์ของระบบที่พัฒนาเมื่อมีคำสั่งในคิวอัพโหลดไฟล์มา ระบบจะอัพโหลดไฟล์หลายไฟล์ในเวลาเดียวกัน และเมื่อแต่ละคำสั่งอัพโหลดทำงานเสร็จจะ Callback ให้อัพเดทข้อมูลในฐานข้อมูล เป็นต้น

จากตัวอย่างจะเห็นถึงประโยชน์ของการใช้ Node.js ทำให้ระบบที่พัฒนาไม่ต้องรอคิวการทำงานทีละส่วน

2.3 RESTFul Web service

RESTFul Web Service (Representational State Transfer) [3] คือ รูปแบบการรับส่งข้อมูลเป็น Method โดยการใช้งานจะอ้างอิงตามคำสั่งของ Method ของ HTTP Protocol

ตารางที่ 1:ตัวอย่าง Method ของ RESTFul Web Service

Method	ลักษณะการทำงาน	รูปแบบ URL
GET	แสดงข้อมูลทั้งหมด	http://goo.co/item/
	แสดงข้อมูลเฉพาะรายการ	http://goo.co/item/12
POST	เพิ่มข้อมูล	http://goo.co/item/
PUT	แก้ไขข้อมูล	http://goo.co/item /1
DELETE	ลบข้อมูล	http://goo.co/item /2

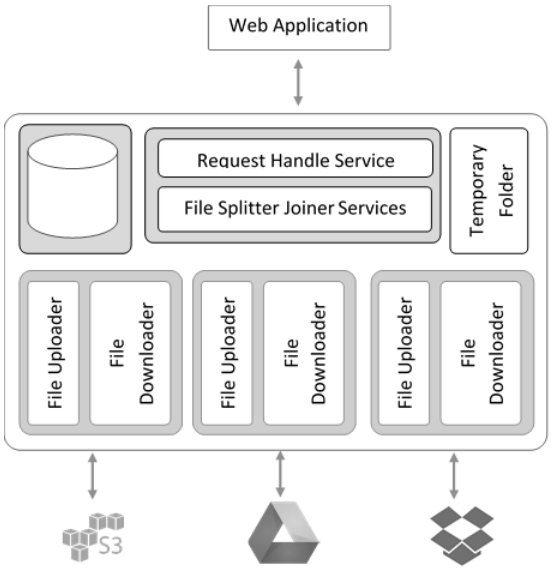
จากตารางที่ 1 การเขียนติดต่อกันผ่านทาง RESTFul Web Service มีความชัดเจนในการติดต่อกันระหว่างกันโดยผ่าน HTTP Protocol เช่นต้องการอัพโหลดไฟล์จะใช้คำสั่ง PUT ผ่านทาง URL http://localhost/upload เป็นต้น

RESTFul Web Service สามารถส่งข้อมูลได้หลากหลายรูปแบบ เช่น JSON HTML และ XML รวมไปถึงไฟล์รูปภาพ ไฟล์ ซึ่งภายในระบบที่พัฒนา การติดต่อกันทั้งหมดที่เกิดขึ้นภายในสถาปัตยกรรมจะเป็นในรูปแบบ RESTFul ทั้งหมด

3. วิธีการดำเนินการวิจัย

3.1 ภาพรวมและโครงสร้างของสถาปัตยกรรม

ภาพรวมของระบบประกอบด้วย 2 ส่วนหลักคือ 1) เว็บแอปพลิเคชัน ใช้ในการติดต่อกับผู้ใช้งาน สร้างขึ้นเพื่อใช้ในการทดสอบระบบ 2) ส่วนประมวลผลหลักจะทำหน้าที่รับคำสั่งจาก เว็บแอปพลิเคชัน และทำงานตามคำสั่งที่ได้รับจากผู้ใช้ มีหน้าที่แบ่งไฟล์ เลือก Cloud Storage Services เก็บบันทึกข้อมูลของผู้ใช้ โดยมีโครงสร้างของสถาปัตยกรรมดังนี้



ภาพที่ 2: สถาปัตยกรรมระบบ

จากภาพที่ 2 โครงสร้างสถาปัตยกรรมของระบบแบ่งการทำงานออกเป็น 3 ส่วน โดยในแต่ละส่วนมี Container เป็นของตัวเอง โดยมีส่วนประกอบดังนี้

File Management Module เป็นส่วนที่ใช้ในการติดต่อกับผู้ใช้งาน โดยมีส่วนประกอบย่อย แบ่งเป็น 2 ส่วนได้แก่ Request Handle Service จัดการการติดต่อกับผู้ใช้งานผ่าน RESTFul Web Service และส่วน File Splitter Joiner Services ส่วนการแบ่งไฟล์/ต่อไฟล์ ตัดสินใจเลือก Cloud Storage Services ในการบันทึก

Temporary Folder ที่เก็บไฟล์ชั่วคราวเมื่อผู้ใช้อัพโหลดไฟล์จะถูกบันทึกไว้ที่แฟ้มนี้ก่อนเพื่อรออัพโหลด และใช้สำหรับการแบ่งและรวมไฟล์ เมื่อระบบใช้งานไฟล์แล้วจะลบออกจากระบบ

Database เป็นส่วนเก็บข้อมูลของไฟล์ที่บันทึกใน Temporary Folder ที่อยู่ของไฟล์บน Cloud Storage Services token ของ user แต่ละคน จำนวนพื้นที่คงเหลือ

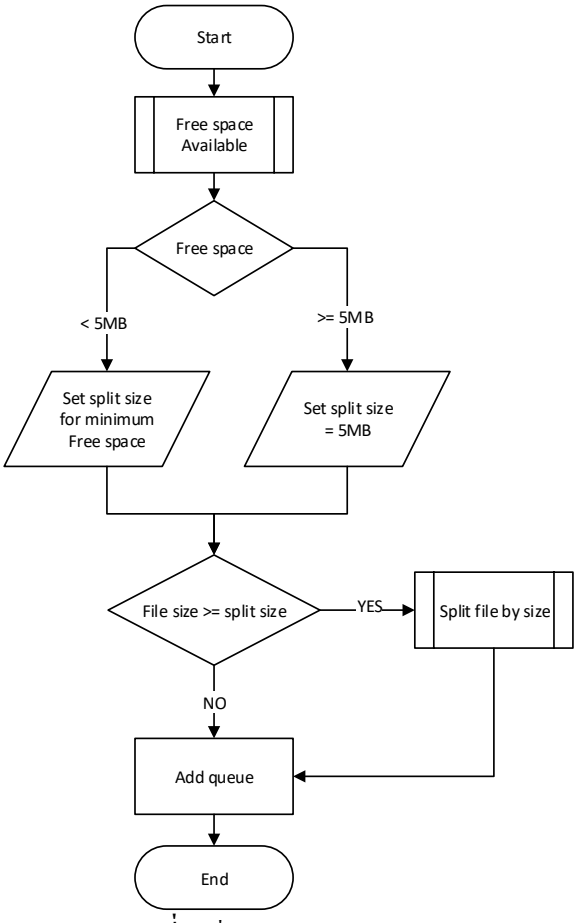
Files Upload/Download Provider แยกตาม Cloud Storage Services มีหน้าที่ Upload และ Download ไฟล์ และบันทึกการข้อมูลไปยังฐานข้อมูลโดยการอ่านไฟล์จะเข้าถึงผ่าน Temporary Folder

3.2 การแบ่งไฟล์และเลือก Cloud Storage Services

เมื่อผู้ใช้งานอัพโหลดไฟล์ไปยัง Temporary Folder สำเร็จ ระบบจะบันทึกข้อมูลของไฟล์ที่ได้รับลงฐานข้อมูล และอ่านขนาดของไฟล์ โดยขนาดไฟล์ที่ใช้ในการแบ่งไฟล์อัพโหลดคือ 5 MB เนื่องจากเป็นขนาดสูงสุดที่สามารถอัพโหลดแบบ simple upload ของ Google drive API [4] งานวิจัยนี้เลือกการอัพโหลดแบบ simple upload เนื่องจากไม่จำเป็นต้องมีข้อมูลอธิบายไฟล์ (metadata) ระบบที่พัฒนาเมื่อแบ่งไฟล์จะมีผลคือไม่สามารถระบุ metadata ได้ การแบ่งไฟล์

วิธีการแบ่งไฟล์มีกระบวนการดังนี้ นำเข้าไฟล์โดยจัดให้อยู่ในรูปแบบ Byte Array และแบ่งตามขนาดที่กำหนด การแบ่งไฟล์ในลักษณะนี้ จะไม่ทำให้ข้อมูลในไฟล์เสียหาย เนื่องจากไม่ได้อ่านข้อมูลโดยตรง แต่มีข้อจำกัดที่เมื่อรวมไฟล์แล้วจำเป็นต้องรวมให้ถูกลำดับของไฟล์ ไม่สามารถรวมไฟล์ข้ามลำดับได้

การแบ่งไฟล์และเลือก Cloud Storage Services จะมีเงื่อนไขการตัดสินใจดังนี้



ภาพที่ 3 เงื่อนไขการตัดสินใจแบ่งไฟล์

จากภาพที่ 3 เงื่อนไขการตัดสินใจแบ่งไฟล์ เมื่อได้รับไฟล์ 1) ตรวจสอบพื้นที่คงเหลือในระบบ 2) ถ้าในระบบเหลือน้อยกว่า 5 MB ให้ตั้งค่า spilt size เป็นค่าที่น้อยสุดของ Cloud Storage Services ที่มี 3) ถ้าในระบบเหลือมากกว่า 5 MB ให้ตั้งค่า split size เท่ากับ 5 MB 4) ถ้าขนาดไฟล์ที่ได้รับน้อยกว่า ให้เข้าคิวอัพโหลด 5) ถ้าขนาดไฟล์ที่ได้รับมากกว่า ให้เข้าฟังก์ชันแบ่งไฟล์ตามขนาด split size และเข้าคิวอัพโหลด

หลังจากที่เข้าคิวอัพโหลดแล้วใช้ node-schedule [5]ให้ส่งคำสั่งไปยัง File Splitter Joiner Services มีลักษณะการทำงานเป็น schedule กำหนดเวลาการทำงานไว้ทุกๆ 1 วินาทีเมื่อครบเวลาที่กำหนด schedule ตรวจสอบคิวการอัพโหลด หากมีข้อมูลของคิว schedule เลือกคิวอัพโหลดแรกสุดมาทำงาน เนื่องจากคำสั่งการแบ่งไฟล์ จะแบ่งไฟล์ตามขนาดที่กำหนดก่อน ไฟล์สุดท้ายจะเป็นไฟล์เศษเหลือจากตัดไฟล์ เมื่อบันทึกไฟล์ลงคิวจึงจำเป็นต้องใช้ไฟล์แรกสุดของคิวก่อนเสมอ File Splitter Joiner Services อ่านพื้นที่คงเหลือของผู้ใช้จากฐานข้อมูล เลือก Cloud Storage ที่มีขนาดพื้นที่เหลือน้อยสุดก่อน เนื่องจากจะใช้พื้นที่หมดก่อนและสั่งอัพโหลด เมื่อส่งคำสั่งแล้วเปลี่ยนแปลงข้อมูลพื้นที่คงเหลือของผู้ใช้บนฐานข้อมูล

3.3 การรวมบริการ

จากภาพที่ 2 ระบบที่ได้ออกแบบการทำงานเป็นส่วนย่อยทั้งหมด 4 ส่วน แต่ละส่วนจะมี container แยกตามส่วน และใช้ Docker Compose [6] ทำหน้าที่รวมบริการทั้งหมดเข้าด้วยกัน โดย Docker Compose เป็นชุดคำสั่งที่ใช้สำหรับใช้ทำงานร่วมกันของ Docker Container และในแต่ละ Container สามารถติดต่อกันได้ผ่าน link การใช้งาน docker compose มีสคริปต์ไฟล์ ชื่อ docker-compose.yml และใช้คำสั่ง docker compose ในการทำงาน

4. ผลการดำเนินงานวิจัย

การทดสอบจะใช้สถาณการณ์จำลอง 3 รูปแบบ เพื่อให้ครอบคลุมกับการทำงานของระบบที่พัฒนา

รูปแบบที่ 1 อัพโหลดไฟล์ขนาดไม่เกิน 5 MB โดยกำหนดพื้นที่คงเหลือของ Cloud Storage Services ทุกบริการ 20MB มีผลลัพธ์ดังนี้ docker-compose ทำงานตามได้ตามขั้นตอน เมื่อผู้ใช้อัพโหลดไฟล์ ไฟล์จะถูกเปลี่ยนชื่อเพื่อป้องกันกรณีผู้ใช้อัพโหลดไฟล์ชื่อเดียวกัน จากนั้น File Splitter Joiner Services จะเลือก Cloud Storage Services ที่เหลือน้อยสุด แต่สถานการณ์กำหนดให้มีขนาดเท่ากัน ทำให้เลือกตาม id ที่อยู่อันดับแรกแทน โดยไม่มีการเปลี่ยนชนิดไฟล์ ระบบที่ได้พัฒนาสามารถทำงานตามภาพที่ 4

```
backend    | file size:1190986
backend    | http://dropbox:4000/up/705/1
backend    | Successfully uploaded data to Dropbox
file-1493815092395.docx
```

ภาพที่ 4 การอัพโหลดไฟล์ขนาดไม่เกิน 5 MB

รูปแบบทดสอบที่ 2 โดยให้อัพโหลดไฟล์ขนาด 18 MB โดยกำหนดให้พื้นที่ของ Cloud Storage Services ดังนี้ Google Drive 20 MB Dropbox 10 MB และ S3 5 MB จากภาพที่ 3 ระบบที่พัฒนามีการทำงานดังนี้ 1) แบ่งไฟล์ออกเป็น 4 ไฟล์ แต่ละไฟล์มีขนาด 5 MB และไฟล์สุดท้าย 3 MB 2) อัพ

โหลดไฟล์ตามพื้นที่คงเหลือ ได้ดังนี้ อัพโหลดไฟล์แรกที่ S3 อัพโหลดไฟล์ 2 และ3 ที่ Dropbox และไฟล์สุดท้ายที่ Google Drive ผลการอัพโหลดผ่านระบบที่พัฒนา (มีเงื่อนไขการเลือก) ระบบได้ทำงานตามที่ได้ที่สร้างเงื่อนไขไว้ และการทำงานสามารถทำงานแบบ Asynchronous ได้จริง โดยการคำสั่งการอัพโหลดไฟล์ จะสั่งให้อัพโหลดไปยัง S3 ก่อนเป็นงานแรก แต่งานที่เสร็จก่อนคือ Dropbox ดังภาพที่ 5

```
backend    | filesize:18611095

backend    | http://s3:4001/up/714/3
backend    | http://dropbox:4000/up/715/1
backend    | http://dropbox:4000/up/716/1
backend    | http://drive:4002/up/717/2

backend    | Successfully uploaded data to Dropbox
file-1493829369289.zip.sf-part3

backend    | Successfully uploaded data to Dropbox
file-1493829369289.zip.sf-part2

backend    | Successfully uploaded data to Drive
0BxBJahT8Z3mYWFBqYmpObktIcGs

backend    | Successfully uploaded data to S3
```

ภาพที่ 5 ผลการอัพโหลดไฟล์ขนาดเกิน 5 MB

รูปแบบทดสอบที่ 3 โดยมีเงื่อนไข ใช้พื้นที่โดยมีข้อจำกัดของพื้นที่จัดเก็บไฟล์ดังนี้ ให้พื้นที่จัดเก็บของ Google Drive คงเหลือ 1 MB พื้นที่จัดเก็บของ Dropbox คงเหลือ 0.5 MB และให้พื้นที่คงเหลือ S3 คงเหลือ 3 MB ไฟล์มีขนาด 3.1 MB ระบบแบ่งไฟล์ตามเงื่อนไขได้ดังนี้ เมื่อขนาดพื้นที่คงเหลือน้อยสุดน้อยกว่า 5 MB ให้แบ่งไฟล์ตามขนาดที่น้อยที่สุดของพื้นคงเหลือ ระบบจะแบ่งไฟล์ออกเป็น 7 ไฟล์ ไฟล์ที่ 1-6 ไฟล์ละ 0.5 MB ไฟล์สุดท้ายมีขนาด 0.1 MB ระบบสามารถแบ่งไฟล์และเลือกผู้ให้บริการได้ตามเงื่อนไข โดยใช้พื้นที่คงเหลือของ Dropbox ในไฟล์ที่ 1 ในส่วนของไฟล์ที่ 2 และ 3 จะใช้พื้นที่คงเหลือของ Google Drive ไฟล์ที่ 4 -7 จะใช้พื้นที่คงเหลือของ S3 ดังภาพที่ 6

```
backend    |  filesize:3195928

backend    |  http://dropbox:4000/up/736/1, usage 524288
remaining 0
backend    |  http://drive:4002/up/737/2, usage 524288
remaining 524288
backend    |  http://drive:4002/up/738/2, usage 524288
remaining 0
backend    |  http://s3:4001/up/739/3, usage 524288
remaining 2621440
backend    |  http://s3:4001/up/740/3, usage 524288
remaining 2097152
backend    |  http://s3:4001/up/741/3, usage 524288
remaining 1572864
backend    |  http://s3:4001/up/742/3, usage 50200
remaining 1522664
```

ภาพที่ 6 ผลการอัพโหลดไฟล์ขนาดไม่เกิน 5 MB แบบมีเงื่อนไข

5. สรุปผลงานวิจัยและข้อเสนอแนะ

ผู้วิจัยได้ออกแบบและพัฒนาสถาปัตยกรรมสำหรับการแบ่งและกระจายไฟล์โดยอัตโนมัติ ซึ่งประกอบด้วย File Management Moduleใช้ในการติดต่อกับผู้ใช้งานผ่าน RESTFul Web Service ตลอดจนวิธีการแบ่งขนาดไฟล์และการเลือก Cloud Storage Services ในการอัพโหลด พร้อมทั้งออกแบบฐานข้อมูลสำหรับการพัฒนาส่วน File Upload/Download Provider สำหรับการอัพโหลดและดาวน์โหลดไฟล์ไปยังผู้ให้บริการทั้ง 3 ผู้ให้บริการ (Google Drive , Dropbox , Amason S3)

จากการทดสอบทั้งสองส่วนได้แก่ การทดสอบการใช้ docker-compose ในการทำงาน สามารถสร้าง Container และทำงานประสานกันตามสถาปัตยกรรมได้ถูกต้อง และการทดสอบการอัพโหลดไฟล์ โดยแบ่งเป็น การอัพโหลดไฟล์ขนาดเล็กที่ไม่ต้องแบ่งไฟล์ (ขนาดไม่เกิน 5 MB) และการอัพโหลดไฟล์ขนาดใหญ่แบบมีเงื่อนไขของพื้นที่คงเหลือของ Cloud Storage Services สามารถทำงานได้ถูกต้อง ผลการทดสอบแสดงให้เห็นว่าสถาปัตยกรรมที่นำเสนอในงานวิจัยนี้สามารถใช้งานได้จริง ทำให้ผู้ใช้สามารถใช้งานพื้นที่เก็บไฟล์บนผู้ให้บริการคลาวด์หลายแห่งได้โดยควบคุมผ่านทางระบบที่สร้างตามสถาปัตยกรรมนี้เพียงที่เดียว ซึ่งจะช่วยลดความยุ่งยากและความสับสนในการจัดการเก็บไฟล์บนผู้ให้บริการหลายๆ แห่ง

เอกสารอ้างอิง

[1] Wang, C. *"Containers 101: Linux containers and Docker explained."* 2016; Available from: http://www.infoworld.com/article/3072929/linux/containers-101-linux-containers-and-docker-explained.html.

[2] Shaikh, S. *"Learn the Concept of Callback in Node.JS."* 2017; Available from: https://www.eduonix.com/blog/web-programming-tutorials/learn-concept-callback-node-js/.

[3] Oracle. *"The Java EE 6 Tutorial."* 2013 [26/04/2017]; Available from: http://docs.oracle.com/javaee/6/tutorial/doc/gijqy.html.

[4] Google Inc.,. *"Uploading Files."* 2017; Available from: https://developers.google.com/drive/v3/web/manage-uploads.

[5] Baxley, R. *"node-schedule."* 2017;Available from: https://github.com/node-schedule/node-schedule.

[6] Inc., D. *"Overview of Docker Compose"*. 2017; Available from: https://docs.docker.com/compose/overview/.

การประยุกต์ใช้เทคนิคการจัดเก็บและค้นคืน เพื่อการตรวจสอบการบังคับใช้นโยบายความมั่นคงด้านสารสนเทศ
Applying Information Storage and Retrieval Technique for Auditing Information Technology Security Policy Enforcement

รัฐศิริ เทสสิริ (Ruttasiri Tessiri)[1] และ นครทิพย์ พร้อมพูล (Nakornthip Prompoon)[2]
[1]ภาควิชาวิศวกรรมคอมพิวเตอร์, จุฬาลงกรณ์มหาวิทยาลัย
[2]ภาควิชาวิศวกรรมคอมพิวเตอร์, จุฬาลงกรณ์มหาวิทยาลัย
Ruttasiri.T@Student.chula.ac.th, Nakornthip.S@chula.ac.th

บทคัดย่อ

งานวิจัยนี้นำเสนอเครื่องมือวัดที่สนับสนุนกระบวนการตรวจสอบ ด้วยการหาความคล้ายคลึงของข้อความจากนโยบายความมั่นคงด้านสารสนเทศและข้อกำหนดความต้องการของระบบด้านเทคโนโลยีสารสนเทศเป็นกรณีศึกษา โดยใช้เทคนิคการจัดเก็บและค้นคืนสารสนเทศด้วยแบบจำลองปริภูมิเวคเตอร์ และการให้น้ำหนักของคำจากความถี่ของคำที่ปรากฏในเอกสารและจำนวนเอกสารทั้งหมดที่คำนั้นปรากฏ ผลการทดลองด้วยข้อมูลตัวอย่างจากระบบงานจริง 3 ระบบ คำสอบถามที่สร้างจากนโยบายความมั่นคงจำนวน 30 รายการสอบถาม จากจำนวน 400 นโยบายด้านความมั่นคง ได้ค่าเฉลี่ยของค่าระลึกและค่าความแม่นยำเป็น 0.44 และ 0.47 ตามลำดับจากค่ากำหนด 0.5 เครื่องมือที่สร้างขึ้นนี้มีฟังก์ชันหลักสนับสนุนการตรวจสอบ ที่ครอบคลุมการจัดเก็บ ค้นคืน และแสดงผลของรายการความต้องการที่สอดคล้องกับนโยบายความมั่งคงแต่ละรายการที่กำหนดไว้ได้

คำสำคัญ: การตรวจสอบสารสนเทศ การวัดค่าความคล้ายคลึง และ นโยบายความมั่นคง

Abstract

This research presents a tool that supports an auditing process for textual similarity measurement between two documents using security policy and system requirements specification of information technology system as a case study. Information storage and retrieval technique based on Vector Space Model was used to implement the proposed tool. Term frequency and inverse document frequency are used to assign weight of each term in each document. The experimental result from the three real projects and thirty queries generated from four hundred security policies shows that average precision and recall is 0.44 and 0.47 respectively for 0.5 threshold. The proposed tool has the main functions for an auditing process that covers the storage, retrieval and display a list of system requirements specification that contains the predefined security policy.

Keyword: Information Technology Audits, Similarity Measurement and Security Policy

1. บทนำ

ปัจจุบันในองค์กรต่างๆ ได้นำเทคโนสารสนเทศมาใช้ในสนับสนุนการดำเนินการขององค์กร เพื่อให้สามารถบรรลุพันธกิจหลัก นโยบาย และกฎระเบียบได้ การตรวจสอบสารสนเทศเป็นสิ่งที่จำเป็นเพื่อเข้ามาควบคุมกระบวนการทำงานว่าเป็นไปตามสิ่งที่ได้กำหนดไว้หรือไม่ [1] โดยต้องมีการกำหนดวัตถุประสงค์ของการตรวจสอบ และนำไปสู่กิจกรรมเพื่อให้เกิดมาตรฐานในการประเมิน ปรับปรุงและควบคุมให้เป็นไปตามแนวทางการบริหารความเสี่ยง [2] ขององค์กร

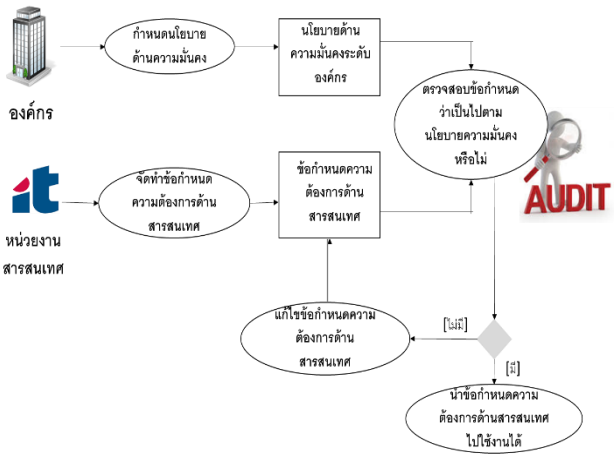
นโยบายด้านความมั่งคงเป็นสิ่งที่มีความสำคัญอย่างยิ่ง และ

จำเป็นต้องนำไปปฏิบัติโดยต้องกำหนดเป็นส่วนหนึ่งของข้อกำหนดความต้องการของระบบสารสนเทศที่จะพัฒนาขึ้น เนื่องจากจะเป็นข้อมูลนำเข้าของกระบวนการออกแบบและการพัฒนาระบบ ที่จะช่วยให้ระบบที่สร้างขึ้นนั้นมีความมั่นคง และมีความไว้วางใจในการใช้งานได้

นโยบายด้านความมั่นคงประกอบไปด้วย การรักษาความลับ บูรณภาพ และสภาพพร้อมใช้งานของระบบ รวมไปถึงระบุตัวตน ภาวะส่วนตัว การพิสูจน์ตัวตน และการอนุญาตเข้าถึงระบบ เป็นต้น [3] ซึ่งองค์กรต่างๆ ต้องกำหนดรายละเอียดในแต่ละด้านตามบริบทและความรับผิดชอบของระบบที่พัฒนา

แนวทางการตรวจสอบจำแนกเป็น 2 แบบ [1] คือ 1) การตรวจสอบโดยผู้ตรวจสอบภายใน เป็นการตรวจสอบเบื้องต้นที่อาจจะคำนึงถึงต้นทุนขององค์กร ซึ่งอาจส่งผลด้านความน่าเชื่อถือ 2) การตรวจสอบโดยใช้ผู้ตรวจสอบภายนอก ที่หน่วยงานจัดจ้าง โดยทั้งสองรูปแบบมีวัตถุประสงค์เดียวกันคือเพื่อตรวจสอบให้แน่ใจว่าโครงการได้ดำเนินการ เป็นไปตามวัตถุประสงค์ กฎ ระเบียบ และข้อบังคับด้านความมั่นคงของการพัฒนาซอฟต์แวร์ ซึ่งกระบวนการตรวจสอบผู้รับการตรวจสอบจะจัดเตรียมข้อมูลให้แก่ผู้ตรวจสอบตามคำร้องขอ ส่วนใหญ่คือผลผลิตจากโครงการ ได้แก่ ข้อกำหนดความต้องการของระบบ (System Requirements Specification) ข้อกำหนดการออกแบบ (Design Specification) และ รหัสต้นฉบับ (Source Code) เป็นต้น ผู้ตรวจสอบจำเป็นต้องมีความรู้และเข้าใจวัตถุประสงค์ กฎ ระเบียบ และข้อบังคับของการพัฒนาซอฟต์แวร์[4]ขององค์กร เพื่อนำมาเปรียบเทียบและรายงานผลให้แก่ผู้บริหารเพื่อนำไปวิเคราะห์เพื่อการปรับปรุงและกำหนดแนวทางการดำเนินการต่อไป โดยมีภาพรวมของการทำงานดังภาพที่ 1

ดังนั้นผู้วิจัยจึงนำเทคนิคการจัดเก็บและค้นคืนเพื่อการตรวจสอบการบังคับใช้นโยบายความมั่นคงด้านสารสนเทศมาประยุกต์ใช้และสร้างเครื่องมือเพื่อวัดความคล้ายคลึงของเอกสารอย่างอัตโนมัติ โดยใช้นโยบายความมั่นคงด้านสารสนเทศและข้อกำหนดความต้องการของระบบด้านเทคโนโลยีสารสนเทศเป็นกรณีศึกษา เพื่อลดปัญหาต่างๆ ที่



ภาพที่ 1: ภาพรวมกระบวนการทำงานตรวจสอบสารสนเทศระหว่างนโยบายความมั่นคงและข้อกำหนดสารสนเทศ

กล่าวมา และช่วยลดเวลาในการค้นหาประเด็นที่ต้องการตรวจสอบได้อย่างรวดเร็ว

2. วรรณกรรมที่เกี่ยวข้อง

ผู้วิจัยแยกงานวิจัยออกเป็น 3 ประเภท โดยกลุ่มงานวิจัยแรกคืองานวิจัยที่กล่าวถึง การตรวจสอบสารสนเทศโดยให้เครื่องมือเข้ามาประยุกต์และแก้ไขปัญหา กลุ่มงานวิจัยที่เกี่ยวข้องกับกระบวนการค้นคืนสารสนเทศและแสดงผล และงานวิจัยที่การถึงการยกตัวอย่างงานวิจัยที่เกี่ยวข้องกับนโยบายความมั่นคง

งานวิจัยที่กล่าวถึงการตรวจสอบสารสนเทศ โดยงานวิจัยนี้กล่าวถึงการตรวจสอบว่าเป็นแนวทางสำคัญสำหรับองค์กร[1] เพื่อที่จะรับประกันว่าองค์กรมีการควบคุมภายในที่ดี และเป็นไปตามข้อกำหนดความต้องการ และได้กล่าวถึงค่าใช่จ่ายที่สูงของการตรวจสอบ และความซับซ้อนของการตรวจสอบ การวิจัยนี้ใช้ตัวอย่างองค์กรจากการทบทวน

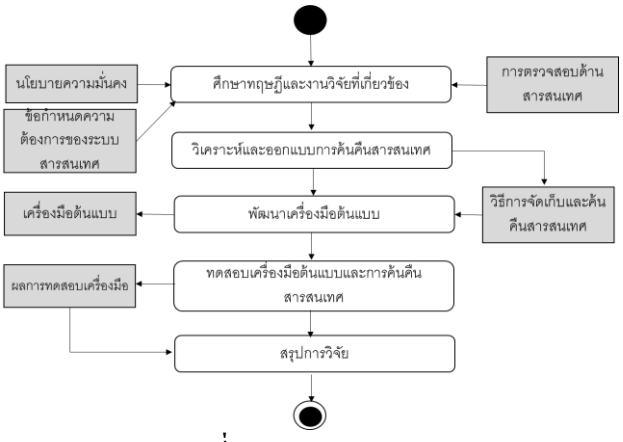
งานวิจัยกลุ่มต่อมาจะกล่าวถึงการค้นคืนสารเทศโดยนำมาประยุกต์ใช้ในงานต่างๆ งานวิจัยที่ผู้วิจัยได้สร้างเครื่องมือบนพื้นฐานไวยากรณ์ความมั่นคงที่สร้างขึ้น ให้ผู้ใช้สามารถป้อนข้อมูลนำเข้าที่จำเป็นต่อการสร้างความต้องการความมั่นคง และสามารถค้นคืนผลลัพธ์ที่ได้จากเครื่องมือคือความต้องการที่มีองค์ประกอบด้านความมั่นคงในรูปประโยค ภาษาอังกฤษ [5]โดยผู้วิจัยได้นำกระบวนการจัดเก็บและการค้นคืนสารสนเทศ ภาษาอังกฤษมาใช้เนื่องด้วยกรณีตัวอย่างที่ผู้วิจัยนำมาทดลองเป็นบริษัทระหว่างประเทศ ซึ่งเอกสารทั้งหมดเป็นภาษาอังกฤษ ส่วนอีกงานวิจัยได้กล่าวถึง หลักการจัดเก็บและ

ค้นคืนสารสนเทศของแบบจำลองปริภูมิเวคเตอร์ เพื่อช่วยสนับสนุนการค้นคืนบริบทความมั่นคงจากส่วนของเอกสารมาตรฐานความมั่นคงสารสนเทศนั้น ได้ตรงตามความต้องการของผู้ใช้ [6]

อีกหนึ่งงานวิจัยได้นำเสนอพื้นฐานของสารสนเทศด้านการตรวจสอบความปลอดภัย โดยกล่าวถึงกระบวนการตรวจสอบสารสนเทศด้านการตรวจสอบความปลอดภัย [7] เนื่องด้วยข้อมูลเป็นสิ่งที่สำคัญมากมากต่อองค์กร ข้อมูลมีความละเอียดอ่อน ซึ่งบุคคลอื่นไม่ควรถูกอนุญาตให้เข้าถึงข้อมูลได้ โดยเฉพาะอย่างยิ่งคู่แข่งซึ่ง ผู้วิจัยนำองค์ความรู้ด้านการตรวจสอบสารสนเทศด้านความมั่นคงมาปรับใช้กับโครงงานวิจัย ดังศึกษาจากกรณีตัวอย่าง [8]

3. วิธีการดำเนินการวิจัย

ผู้วิจัยนำเสนอเครื่องมือวัดความคล้ายคลึงของเอกสารโดยใช้นโยบายความมั่นคงด้านสารสนเทศและข้อกำหนดความต้องการของระบบด้านเทคโนโลยีสารสนเทศเป็นกรณีศึกษาเพื่อสนับสนุนและแก้ไขปัญหาค่าใช้จ่ายที่มีราคาสูงรวมถึงการลดระยะเวลาการทำงาน โดยใช้เครื่องมือที่ผู้วิจัยพัฒนาขึ้นมาเองโดยมีหลักการทำงานและกระบวนการดังต่อไปนี้



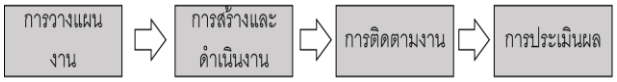
ภาพที่ 2: กระบวนการดำเนินงานวิจัย

จากภาพที่ 2 กระบวนการดำเนินงานวิจัยมีขั้นตอนดังต่อไปนี้

1) ศึกษาทฤษฎีต่างๆที่เกี่ยวข้องกับงานวิจัย

ผู้วิจัยศึกษาทฤษฎี การตรวจสอบสารสนเทศ นโยบายความมั่นคงสารสนเทศ และ การค้นคืนสารสนเทศ โดยมีรายละเอียดของทฤษฎีต่างๆ ดังต่อไปนี้

1.1) การตรวจสอบสารสนเทศ ผู้วิจัยศึกษาเพื่อนำกระบวนการดังกล่าวมาปรับใช้กับข้อกำหนดความต้องการการเพื่อพัฒนาเครื่องมือเพื่อสนับสนุนการทำงานของกระบวนการตรวจสอบที่เรียกว่าโคบิต (COBIT: Control Objective for Information and Related Technology) [2] มีขั้นตอนดังภาพที่ 3



ภาพที่ 3: กระบวนการตรวจสอบ

จากภาพที่ 3 กระบวนการตรวจสอบ แบ่งกระบวนการตรวจสอบออกเป็น การวางแผนงาน การสร้างและดำเนินการ การจัดส่ง การติดตาม และการประเมินผล

1.2) นโยบายความมั่นคงสารสนเทศ [1] มีวัตถุประสงค์เพื่อกำหนดทิศทางและให้การสนับสนุนการดำเนินการด้านความมั่นคงสำหรับสารสนเทศภายในองค์กร เพื่อให้สอดคล้องกับข้อกำหนดทางธุรกิจ กฎหมาย และระเบียบปฏิบัติ ผู้บริหารระดับสูงต้องจัดทำเอกสารนโยบายความมั่งคง (Information Security Policy Document) ขององค์กรอย่างเป็นลายลักษณ์อักษร

1.3) การจัดเก็บและค้นคืนสารสนเทศ เป็นศาสตร์ที่ช่วยในการสนับสนุนให้มีการจัดเก็บและค้นคืน สารสนเทศแบบอัตโนมัติโดยจะขึ้นอยู่กับเนื้อหาและสาระสำคัญที่ปรากฏในเอกสารเหล่านั้น เนื่องจากสารสนเทศมีปริมาณมากและหลากหลายประเภท[5],[6] ทำให้ยากต่อการสืบค้น ฉะนั้นการจัดเก็บและค้นคืนสารสนเทศจึงเข้ามามีบทบาทสำคัญ

ผู้วิจัยใช้ Porter Algorithm [5],[6] ในการหารากศัพท์โดยมีขั้นตอนการทำงาน ดังตารางที่ 1

ตารางที่ 1 ตัวอย่างขั้นตอนการทำงานของ Porter Algorithm

Step	Suffix	Replacement	Example
1a	sses	ss	caresses → caress
1b	ed	NULL	plastered → plaster
1b1	at	Ate	conflated→ conflate
1c	y	I	happy → happi
2	ational	Ate	relational → relate
3	ful	NULL	hopeful → hope
4	al	NULL	revival → reviv

1.4) การทำดรรชนีอัตโนมัติ เป็นการระบุความสำคัญของคำในเซตดรรชนี เพื่อบ่งบอกถึงความสำคัญของคำจากการปรากฏขึ้นของคำภายในเอกสารใดๆ ในชุดข้อมูล (Collection) จากการหาค่าน้ำหนักของคำสำหรับ งานวิจัยนี้ได้เลือกการกำหนดค่าน้ำหนักของคำตามแนวคิดของความถี่ของคำที่ปรากฏในเอกสารและน้ำหนักของคำในแต่ละเอกสาร [5],[6] ซึ่งประกอบด้วย 2 กระบวนการคือ การนับความถี่ของคำ (Term Frequency: TF) เป็นการนับจำนวนครั้งของคำนั้นๆ ที่ปรากฏ ในเอกสารโดยสามารถคำนวณค่าความถี่ของคำได้ จากสมการ1

$$tf_{i,j} = 1 + \log_2 f_{i,j} \quad \text{.....(1)}$$

กำหนดให้

$f_{i,j}$ คือ ความถี่ของเทอม j ในเอกสารที่ i

การหาความถี่ของเอกสารแบบผกผัน (Inverse Document Frequency: IDF) คือการคำนวณหาสัดส่วนของเทอมใดๆที่ปรากฏในเอกสารทั้งหมดภายในชุดข้อมูล (Collection) ด้วยสมการ 2

$$idf_i = \log_2 \frac{N}{n_i} \quad \text{.....(2)}$$

กำหนดให้

N คือ จำนวนเอกสารทั้งหมดในชุดข้อมูล (Collection)

n_i คือ จำนวนเอกสารที่ปรากฏเทอม i

ดังนั้นสามารถคำนวณค่าน้ำหนักของเทอมหรือความสำคัญของเทอมได้จากสมการ 3

$$w_{i,j} = \left[1 + \log_2 f_{i,j}\right] \times \log_2 \frac{N}{n_i} \quad \text{.....(3)}$$

เมื่อ $w_{i,j}$ คือค่าน้ำหนักของเทอม i ในเอกสารที่ j

การเปรียบเทียบความคล้าย ระหว่างข้อคำถามกับเอกสารสามารถเปรียบเทียบความคล้ายได้จากค่าสัมประสิทธิ์โคซายน์ (Cosine Coefficient) [6] จากแบบ จำลองเชิงปริภูมิ (Vector Space Model) [9] ดังสมการ 4

$$sim(d_j, q) = \frac{\sum_{i=1}^{t} w_{i,j} \times w_{i,q}}{\sqrt{\sum_{i=1}^{t} w_{i,j}^2} \times \sqrt{\sum_{j=1}^{t} w_{i,q}^2}} \quad \text{.....(4)}$$

กำหนดให้

$w_{i,j}$ คือ ค่าน้ำหนักของเทอม i ในเอกสาร j

$w_{i,q}$ คือ ค่าน้ำหนักของเทอม i ในข้อคำถาม q

$w_{i,j}^2$ คือ ค่าน้ำหนักของเทอม i ใดๆทั้งหมด ในเอกสาร j

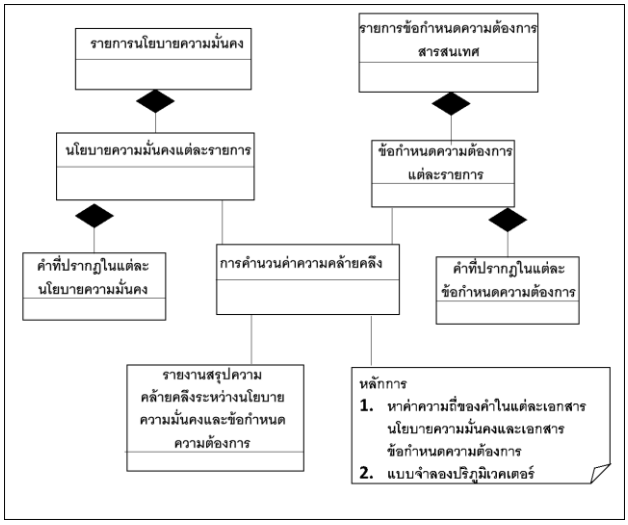
$w_{i,q}^2$ คือ ค่าน้ำหนักของเทอม i ใดๆทั้งหมด ในข้อคำถาม

ค่าความคล้ายระหว่างเอกสารและข้อคำถามอยู่ระหว่าง 0 ถึง 1 หากค่าความคล้ายมีค่าเข้า ใกล้ 1 หมายความว่า เอกสารและข้อคำถามมีความคล้ายกันมา

2) วิเคราะห์และออกแบบการค้นคืนสารสนเทศ

วิเคราะห์และออกแบบการค้นคืนสารสนเทศระหว่างเอกสารของเอกสารนโยบายความมั่นคงและข้อกำหนดความต้องการ ดังภาพที่ 4

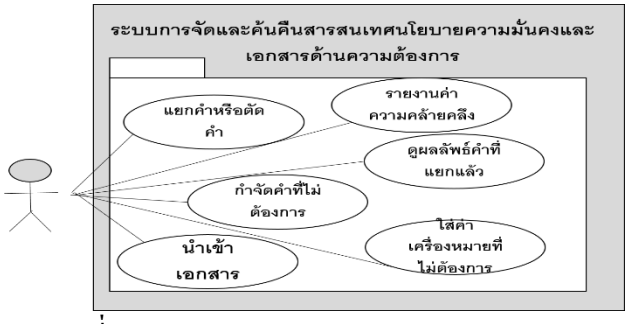
ภาพที่ 4: การหาค่าความคล้ายคลึงสารสนเทศของเอกสารของนโยบายความมั่นคงและข้อกำหนดความต้องการของระบบ



จากภาพที่ 4 นโยบายความมั่นคงจะประกอบไปด้วยรายการของนโยบายความมั่นคงและแต่ละรายการจะประกอบด้วยคำที่ปรากฏในนโยบายความมั่นคง และรายการข้อกำหนดความต้องการ จะประกอบไปด้วยข้อกำหนดความต้องการแต่ละรายการ และภายในแต่ละรายการจะประกอบด้วยคำที่ปรากฏในข้อกำหนดความต้องการ และหาค่าความคล้ายคลึงโดยใช้ แบบจำลองเชิงปริภูมิ [5],[6]

3) วิเคราะห์และออกแบบการค้นคืนสารสนเทศ

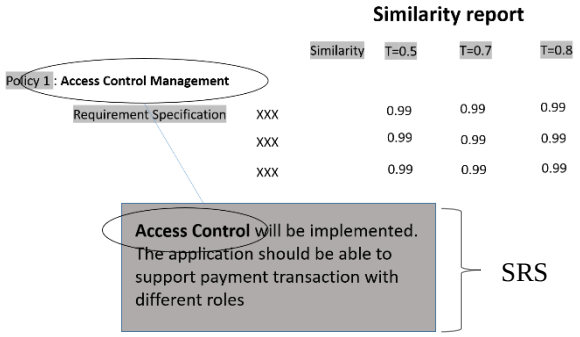
ความต้องการเชิงหน้าที่ของระบบการจัดเก็บและค้นคืนสารสนเทศของเอกสารนโยบายความมั่นคงและเอกสารข้อกำหนดความต้องการ อธิบายได้ด้วยแผนภาพยูสเคสของระบบดังรูปต่อไปนี้



ภาพที่ 5: ยูสเคสของระบบการจัดการและค้นคืนสารสนเทศนโยบายด้านความมั่นคงและเอกสารด้านความต้องการ

4) พัฒนาเครื่องมือต้นแบบ

จากความต้องการที่ได้ระบุไว้ ผู้วิจัยได้ทำการพัฒนาเครื่องมือต้นแบบเพื่อสนับสนุนฟังก์ชันในงานเพื่อนำมาประยุกต์ใช้กับการทำงานจริง



ภาพที่ 6: ตัวอย่างรายงานค่าความคล้ายคลึงและเอกสารข้อกำหนดความต้องการของระบบ

จากภาพที่ 6 ระบบหาค่าความคล้ายคลึงระหว่างเอกสารข้อกำหนดความต้องการของระบบและนโยบายความมั่นคงโดยเก็บข้อมูลลงฐานข้อมูล และสามารถระบุได้ว่าเอกสารความต้องการเล่มไหนที่มีความเกี่ยวข้องกับนโยบายด้านความมั่นคง ถ้าค่ามีค่ามากแสดงว่าเอกสารดังกล่าวเกี่ยวข้องกับนโยบายด้านความมั่นคงข้อนั้นนั่นเอง

4. ผลการดำเนินการวิจัย

ผู้วิจัยได้ดำเนินการใช้ระบบเพื่อให้ได้ผลลัพธ์เพื่อนำมาตรวจสอบและประเมินผลของระบบ

4.1 ผลลัพธ์ของงานวิจัย

ระบบสามารถคำนวนค่าความคล้ายคลึงระหว่างเอกสารนโยบายความมั่นคงและเอกสารข้อกำหนดความต้องการของระบบ และออกรายงานเพื่อให้ผู้ตรวจสอบสารสนเทศเพื่อสามารถนำไปใช้งานได้

4.2 การประเมินผลงานวิจัย

ผู้วิจัยประเมินคุณภาพของเครื่องมือต้นแบบในงานวิจัยนี้จะอ้างอิงจากมาตรฐาน ไอเอสโอ 9126 [9] ซึ่งในงานวิจัยนี้จะใช้ทั้งสิ้น 2 ปัจจัย โดยระบุมาตรวัดของแต่ละปัจจัย ดังตารางที่ 2

ตารางที่ 2 มิติปัจจัยการประเมินคุณภาพและมาตรวัดที่ใช้

หน้าที่การทำงาน (Functionality)	ความเหมาะสม (Suitability)	ความครบถ้วนของฟังก์ชันการทำงานของเครื่องมือต้นแบบ
	ความถูกต้อง (Accuracy)	ความถูกต้องของฟังก์ชันการทำงานเครื่องมือต้นแบบ

โดยจะทำการทดสอบฟังก์ชันงานต่างๆ โดยมีรายละเอียดดังตารางที่ 3

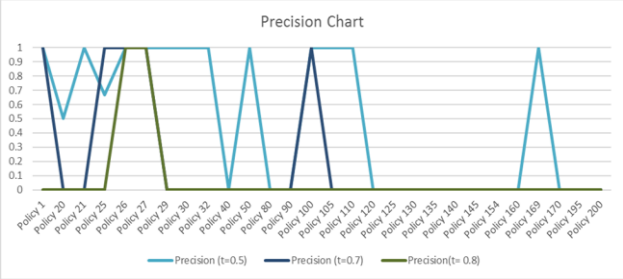
ตารางที่ 3 รายละเอียดข้อมูลที่ใช้ในการทดลอง

รายละเอียด	รายการ
นโยบายความมั่นคง	400
เอกสารข้อกำหนดความต้องการ	3
รายการสอบถาม (Query)	30
ค่ากำหนด	0.5, 0.7, 0.8

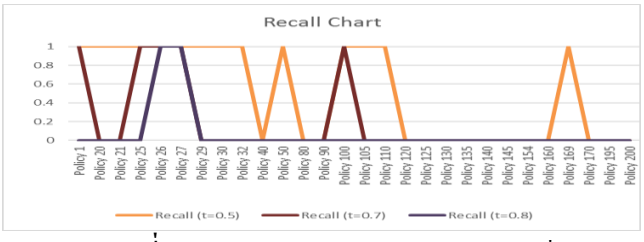
ผู้วิจัยใช้ค่าความแม่นยำ (Precision) และ ค่าระลึก (Recall)ในการประเมินประสิทธิผลการค้นคืน [10] โดยได้ผลของการประเมินค่าความแม่นยำและค่าระลึกดังภาพที่ 7 และ 8

จากภาพที่ 7 ค่าความแม่นยำที่ได้โดยเฉลี่ย ของค่ากำหนด 0.5, 0.7 และ 0.8 เป็น 0.43, 0.17 และ 0.07 ตามลำดับ

จากภาพที่ 8 ค่าระลึกที่ได้โดยเฉลี่ย ของค่ากำหนด 0.5, 0.7 และ 0.8 เป็น 0.47, 0.17 และ 0.07 ตามลำดับ



ภาพที่ 7: แผนภาพค่าความแม่นยำของแต่ละนโยบายความมั่นคง



ภาพที่ 8: แผนภาพค่าระลึกของแต่ละนโยบายความมั่นคง

5. สรุปผล

ผลการวิจัยค่าความคล้ายคลึงระหว่างนโยบายความมั่นคงและเอกสารข้อกำหนดความต้องการ แสดงว่านโยบายความมั่นคงที่ถูกกล่าวถึงในเอกสารข้อกำหนดความต้องการมักจะเป็นนโยบายทั่วไปที่จดจำได้มักปรากฏเป็นนโยบายที่อยู่ในลำดับต้นๆของเอกสาร เช่น การจัดการรายชื่อผู้ใช้ระบบและรหัสผ่าน (User and Password Management) จะถูกกล่าวถึงในทุกๆ เอกสารข้อกำหนดความต้องการ ส่วนนโยบายความมั่นคงที่เฉพาะเจาะจงมักจะไม่ถูกกล่าวถึงซึ่งปรากฏในลำดับท้ายๆของเอกสาร เช่น การเข้ารหัสลับ (Cryptography) ซึ่งไม่ถูกระบุลงเอกสารข้อกำหนดความต้องการใดๆเลย แสดงให้เห็นว่าผู้พัฒนาระบบควรได้รับการอบรมให้มีความเข้าใจในข้อกำหนดนโยบายความมั่นคงอย่างถ่องแท้ เพื่อให้เข้าใจและสามารถนำนโยบายความมั่นคงไปใช้ในการจัดทำโครงการสารสนเทศได้อย่างมีประสิทธิภาพ

ทั้งนี้หากมีจำนวนเอกสารของโครงการสารสนเทศมากขึ้นมาใช้ในการทดลอง จะทำให้เครื่องมือที่สร้างมีความน่าเชื่อถือมากยิ่งขึ้น อีกทั้งการเพิ่มความสามารถในการสืบค้นโดยใช้ข้อคำถามที่เป็นวลี หรือประโยคได้ จะช่วยให้ได้ผลของการสืบค้นที่มีความแม่นยำมากยิ่งขึ้น

เอกสารอ้างอิง

[1] Tiago Rosario, Ruben Pereira and Miguel Mira da Silva, “Formalization of the IT Audit Management Process “,*IEEE, Enterprise Distributed Object Computing Conference Workshops (EDOCW), 2012 IEEE 16th International*

[2] Robert Moeller, “*IT Audit, Control and Security*”, John Wiley & Son IC.,2010.

[3] Mattord., M.E.W.a.H.J., “*Principla of Information Security 4th Edition*”, Course Technology ,2003

[4] M. Schumacher, E. Fernandez-Buglioni, D.Hybertson, F. Buschmann and P. Sommerlad, “Security Patterns: Integrating Security and Systems Engineering”, John Wiley & Sons, 2005.

[5] Kawin Supaporn, Nakornthip Prompoon and Thongchai Rojkansadan, “Enterprise Assets Security Requirements Construction from ESRMG Gramma base on Security Patterns”, *IEEE, Software Engineering Conference, 2007. APSEC 2007. 14th Asia-Pacific.*

[6] Poramaporn Ruamjinda and Nakornthip Prompoon, “Framework for information security standards storage and retrieval using security pattern”, *IEEE, Software Engineering and Service Science (ICSESS), 2013 4th IEEE International Conference on*.

[7] Edward C. Lo and Mike Marchand, “Security Audit: A Case Study”, *IEEE, Electrical and Computer Engineering, 2004. Canadian Conference on.*

[8] https://www.beartai.com/news/it-thai-news/115217,18-05-2017

[9] Baeza-Yates, R.A., Ribeiro-Neto, B., “*Modern Informatiuon Retrieval”*,Addison-Sesleyt Longman Publishing Co., Inc., 1999

[10] Xusheng Xiao, Amit Paradkar, Surech Thummalapenta and Tao Xie, “Automated Extraction of Security Policies from Natural Language Software Document*”, IEEE Transactions on Software Engineering* (Volume: PP, Issue: 99)

ผลกระทบของอุณหภูมิต่อการควบคุมพลังงานการส่งข้อมูลสำหรับเครือข่ายเซ็นเซอร์ไร้สาย
Temperature Effect on Transmission Power Control for Wireless Sensor Networks

วิลาวรรณ รักผกาวงศ์ (Wilawan Rukpakavong)[1] และ สิริกันยา นิลพานิช (Sirikunya Nilpanich)[2]
สาขาวิชาวิทยาการคอมพิวเตอร์คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยธรรมศาสตร์
[1]wilawan@cs.tu.ac.th, [2] skn@cs.tu.ac.th

บทคัดย่อ

เครือข่ายเซ็นเซอร์ไร้สายประกอบด้วย เซ็นเซอร์และหน่วยประมวลผล, หน่วยเก็บข้อมูล และหน่วยสื่อสารข้อมูล โดยที่แต่ละโหนดจะมีขนาดเล็ก และใช้อุปกรณ์ไร้สายพลังงานต่ำ เนื่องจากการมีความจุพลังงานแบบจำกัด การควบคุมพลังงานในการส่งข้อมูลจึงเป็นวิธีหนึ่งที่ใช้ในการยืดอายุการทำงานของโหนด บทความวิจัยนี้จะเสนอกลไกเสริมในการควบคุมพลังงานการส่งข้อมูลให้มีประสิทธิภาพ โดยนำค่าของอุณหภูมิซึ่งเป็นปัจจัยหนึ่งที่มีผลกระทบต่อความเข้มของสัญญาณ เพื่อนำมาใช้ในการคำนวนหาค่าพลังงานที่เหมาะสม สำหรับส่งข้อมูล ความเป็นไปได้ของงานวิจัยนี้ได้ทดสอบกับอุปกรณ์จริงโดยใช้ค่าการใช้พลังงานและอัตราการหายของข้อมูล เป็นตัวตรวจวัดผลการปฏิบัติการ

คำสำคัญ: อุณหภูมิ เครือข่ายเซ็นเซอร์ไร้สาย การควบคุมพลังงานการส่งข้อมูล

Abstract

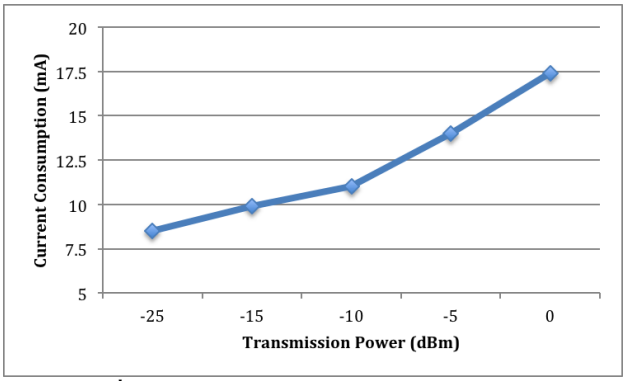
The Wireless Sensor Networks consist of sensors and actuators with processing, storage and communication capabilities. Each node is a small, low-power wireless device. Due to the limited energy capacity, Transmission Power Control (TPC) is one alternative to minimise transmission energy consumption in order to extend lifetime of a node. This paper extends an energy-efficient transmission power control mechanism by utilising temperature, which is one of the key factors that impact the Received Signal Strength (RSS) to calculate optimal transmission power. The feasibility of the proposed schemes is validated by testbed experiments using performance metrics, such as, energy consumption and packet loss rate.

Keyword: Temperature, Wireless Sensor Networks, Transmission Power Control.

1. บทนำ

เนื่องจากการมีพลังงานที่จำกัด การใช้พลังงานเป็นปัจจัยที่สำคัญสำหรับอุปกรณ์เซ็นเซอร์ไร้สาย การใช้พลังงานของอุปกรณ์ตามมาตรฐาน LAN ไร้สาย IEEE 802.11 ไม่สามารถนำมาใช้กับอุปกรณ์ที่ใช้แบตเตอรี่ เพื่อให้มีอายุการใช้งานที่ยาวนานได้ ทำให้เครือข่ายเซ็นเซอร์ไร้สายจะใช้อุปกรณ์ที่อิงตามมาตรฐานของ IEEE 802.15.4 ซึ่งได้มีการปรับปรุงเพื่อการสื่อสารที่มีประสิทธิภาพและประหยัดพลังงานโดยมีอัตราส่งข้อมูลต่ำ การควบคุมพลังงานส่งข้อมูล (Transmission Power Control-TPC) เป็นวิธีการที่กำหนดในตัวรับส่งสัญญาณวิทยุแบบ IEEE 802.15.4 เพื่อให้โหนดสามารถเปลี่ยนระดับพลังงานการส่งซึ่งส่งผลให้เกิดการใช้พลังงานจากการส่งข้อมูลหลายรูปแบบ ตัวอย่างเช่น CC2420 [1] สามารถเปลี่ยนกำลังรับส่งข้อมูลในระหว่างทำงานได้ ความสัมพันธ์ระหว่างกำลังรับส่งข้อมูลกับการใช้กระแสไฟของตัวรับส่งสัญญาณวิทยุตามข้อมูลทางเทคนิคอุปกรณ์ CC2420 [1] แสดงในภาพที่ 1 การใช้พลังงานจะลดลงประมาณ 51% ถ้าทุกโหนดส่งแพ็คเก็ตทั้งหมดในค่าระดับพลังงานต่ำสุด ซึ่งจะส่งผลให้อายุการใช้งานของโหนดยาวนานขึ้น นอกจากนี้ TPC ยังมีข้อดีอื่นอีกเช่น [2, 3, 4] สามารถลดขนาดของพื้นที่ชนกันของข้อมูล (Collision Domain) และสามารถนำไปสู่การลดการใช้พลังงานของโหนดข้างเคียงได้ อย่างไรก็ตามในการลดพลังงานการส่งข้อมูลนั้น อาจจะเพิ่มอัตราการสูญหายของแพ็คเก็ตเนื่องจากความแรงของสัญญาณไร้สายที่ลดลงและไม่คงที่ ในบางงาน

วิจัย [5, 6, 7, 8] รายงานว่าอุณหภูมิเป็นหนึ่งในปัจจัยที่มีอิทธิพลต่อความแรงของสัญญาณ ดังนั้นจึงเป็นเรื่องท้าทายที่จะค้นหาค่าพลังการรับส่งที่ดีที่สุด เพื่อช่วยตอบสนองความต้องการด้านความน่าเชื่อถือในการสื่อสารได้โดยคำนึงถึงผลกระทบของอุณหภูมิที่มีต่อความแรงของสัญญาณในการรับข้อมูล



ภาพที่ 1: ค่าการใช้พลังงานกับระดับพลังงานที่ใช้ส่งข้อมูล

2. วรรณกรรมที่เกี่ยวข้อง

Lin และคณะ [5] สังเกตว่าการเปลี่ยนแบบรายวันของความแรงสัญญาณการรับ (Received Signal Strength-RSS) อยู่ที่ประมาณ 6 เดซิเบลซึ่งอาจเกิดจากอุณหภูมิ หรือ/และ ความชื้นและปัจจัยอื่นๆ Bannister และคณะ [6] ศึกษาความสัมพันธ์ระหว่าง RSS และอุณหภูมิในทะเลทรายโซโนรานทางตะวันตกเฉียงใต้ของสหรัฐอเมริกา ซึ่งอุณหภูมิในช่วงฤดูร้อนอาจเปลี่ยนแปลงจาก 25°C ถึง 65°C พวกเขาพบว่าค่า RSS มีแนวโน้มลดลงเมื่ออุณหภูมิเพิ่มขึ้น ส่งผลต่อคุณภาพการส่งข้อมูลระหว่างโหนด พวกเขาเสนอสมการเพื่อชดเชยความแรงของสัญญาณที่ลดลงดังนี้:

$$P_r - T_L(T) \geq P_s \qquad (1)$$

โดยที่ P_r คือพลังงานที่วัดได้ที่เครื่องฝั่งผู้รับ, P_s คือค่าขั้นต่ำของความแรงสัญญาณที่สามารถรับได้ และ T_L (T) เป็นผลกระทบจากอุณหภูมิที่ T ° C ซึ่งคำนวณได้จาก:

$$T_L(T) = 0.1996 * (T - 25) \qquad (2)$$

โดยที่ T คืออุณหภูมิระหว่าง 25°C ถึง 65°C ต่อมา Lee และ Chung [7] ขยายสมการโดยกำหนด T_L (T) เป็น $RSSI_{loss}$ การศึกษาของพวกเขามุ่งเน้นไปที่ตัวรับส่งสัญญาณวิทยุ CC2420 เท่านั้นและเสนอสมการในการปรับกำลังรับส่งเป็น:

$$P_{level} = ((P_{out} + 40) / 12)^{2.91} \qquad (3)$$

โดยค่า P_{level} เป็นระดับพลังงานที่ส่ง และ P_{out} พลังงานขาออก (dBm) ของ CC2420 Boano และคณะ [8] ศึกษาผลกระทบของอุณหภูมิบนแพลตฟอร์ม CC2420 และ CC2520 โดยแสดงให้เห็นว่าความสัมพันธ์ระหว่างอุณหภูมิกับ RSS สำหรับแพลตฟอร์มต่างๆ สามารถแทนด้วยค่าได้ในรูปแบบฟังก์ชั่นเส้นตรง โดยมีการใช้ค่าคงที่ในสมการที่ต่างกันจากการศึกษาผลงานทั้งหมดก่อนหน้านี้ มุ่งเน้นเฉพาะผลกระทบของอุณหภูมิต่อโหนดผู้ส่ง แต่บทความนี้จะนำเสนอรูปแบบการปรับระดับกำลังรับส่ง ซึ่งครอบคลุมผลกระทบต่ออุณหภูมิของทั้งผู้ส่งและผู้รับสัญญาณ

3. วิธีดำเนินการวิจัย

ในงานก่อนหน้านี้ [9] พลังงานขั้นต่ำที่จำเป็นสำหรับการส่ง (P_{min}) สามารถคำนวณได้ดังนี้

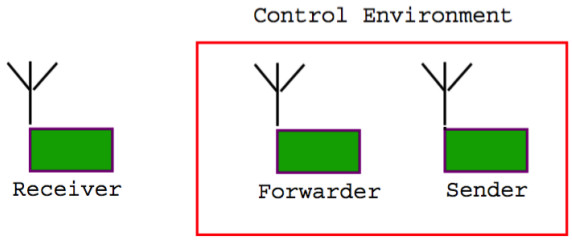
$$P_{min} = RSS_{min} - RSS_{max} + P_{max} \qquad (4)$$

โดยที่ RSS_{min} เป็นขั้นต่ำของความแรงสัญญาณในการรับข้อมูล และ P_{max} เป็นพลังงานสูงสุดที่สามารถส่งได้ โดยตัวอย่างเช่นสำหรับ CC2420 ค่าเหล่านี้มีค่าเท่ากับ -92 dBm และ 0 dBm ตามลำดับ [1] RSS_{max} เป็นค่าที่ได้รับจากการที่โหนดผู้ส่งใช้ค่าพลังงานสูงสุดในการส่ง ซึ่งสามารถวัดได้ระหว่างขั้นตอนการค้นหาเราเตอร์หรือเกตเวย์ที่อยู่ใกล้เคียง

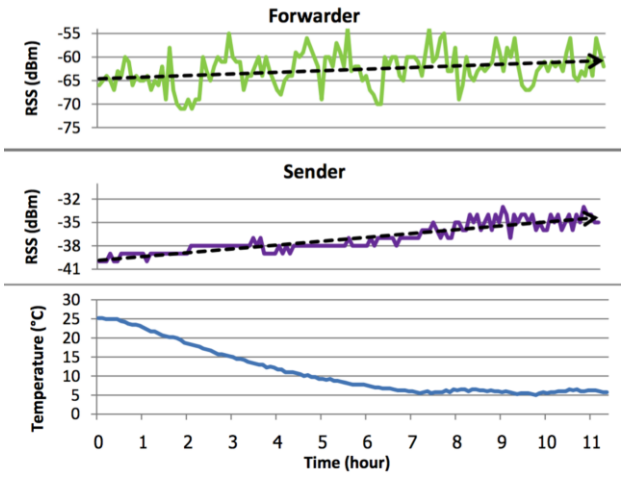
3.1 ผลกระทบจากอุณหภูมิ

การทดลองเพื่อหาค่าผลกระทบจากอุณหภูมินี้ดำเนินการในอาคารโดยใช้อุปกรณ์ N740 NanoSensor [10] โหนดเหล่านี้ใช้ตัวรับสัญญาณวิทยุ Chipcon 2420 ซึ่งทำงานที่คลื่นความถี่ 2.4 GHz และใช้ระบบปฏิบัติการ Contiki โดยแอปพลิเคชันที่ใช้สำหรับการทดสอบประกอบด้วย 3 โหนดคือ ผู้ส่ง (Sender) ผู้ส่งต่อ (Forwarder) และผู้รับ (Receiver) สำหรับ โหนดผู้ส่งและผู้ส่งต่อจะอยู่ในสภาพแวดล้อมที่

ควบคุม ในขณะที่โหนดผู้รับจะอยู่ในสภาพห้องปกติที่อุณหภูมิคงที่ตามภาพที่ 2 โหนดทั้งหมดมีเซ็นเซอร์วัดอุณหภูมิ โดยเริ่มจากวางผู้ส่งและผู้ส่งต่อ (ห่อด้วยถุงพลาสติก) ไว้ในตู้เย็นที่ไม่ได้เสียบปลั๊ก เมื่อเสียบปลั๊กตู้เย็นอุณหภูมิของตู้เย็นจะลดลงจาก 25°C ถึง 5°C ทุกๆ 5 นาที ผู้ส่งจะอ่านค่า อุณหภูมิ 10 ครั้งเพื่อส่งไปยังผู้ส่งต่อ จากนั้นแพ็คเก็ตเหล่านี้จะถูกส่งต่อไปยังผู้รับโดยแนบค่า RSS ของผู้ส่งที่วัดได้ที่เครื่องผู้ส่งต่อมาพร้อมกับข้อมูล หลังจากนั้นผู้รับจะนำข้อมูล RSS ของแพ็คเก็ตที่ได้รับจากผู้ส่งต่อรวมทั้งค่า RSS ที่วัดได้จากสัญญาณจากสัญญาณของเครื่องผู้ส่งต่อ เพื่อรวบรวมสถิติ



ภาพที่ 2: การทดลองผลกระทบจากอุณหภูมิ



ภาพที่ 3: การทดลองผลกระทบจากอุณหภูมิ

ผลของค่า RSS ที่เก็บรวบรวมได้จากการลดอุณหภูมิของผู้ส่งและผู้ส่งต่อแสดงในภาพที่ 3 จากกราฟสามารถสังเกตได้ว่าเมื่อค่าอุณหภูมิเปลี่ยนแปลง มีผลต่อความแรงของสัญญาณของผู้ส่งต่อ เมื่ออุณหภูมิลดลงจาก 25°C ถึง 5°C ค่า RSS ของผู้ส่งต่อจะเพิ่มขึ้นประมาณ 4 dBm ซึ่งหมายความว่า RSS จะเพิ่มขึ้นประมาณ 0.2 dBm เมื่ออุณหภูมิลดลง 1°C ซึ่งแสดงถึงผลกระทบของอุณหภูมิในกรณีที่เครื่องผู้ส่งมีการเปลี่ยนของค่าอุณหภูมิแต่เครื่องผู้รับอุณหภูมิคงที่ (อุณหภูมิห้องประมาณ 25°C) นอกจากนี้ยังพบว่าเมื่อทั้งเครื่องผู้ส่งและเครื่องผู้รับอยู่ในอุณหภูมิที่ลดลง 1°C ค่าความแรงสัญญาณที่รับได้จะเพิ่มขึ้น 7 dBm (หรือประมาณ 2 เท่า) จากผลการทดลองนี้แสดงให้เห็นถึงแนวโน้มเช่นเดียวกับผลลัพธ์ที่ได้จากการทดลองของ Bannister [6], Boano [8] และ Luomala [11]

3.2 การออกแบบการชดเชยพลังงาน

จากสมการ (4) อุณหภูมิเริ่มต้นของทั้งผู้ส่งและตัวรับสัญญาณจะต้องถูกบันทึก (เรียกว่า T_s และ T_r) ผลกระทบของอุณหภูมิสำหรับการปรับกำลังส่งคำนวนได้จาก

$$PT_{min}(T_{ns},T_{nr}) = PT_{min} + \tau((T_{ns} - T_s) + (T_{nr} - T_r)) \quad (5)$$

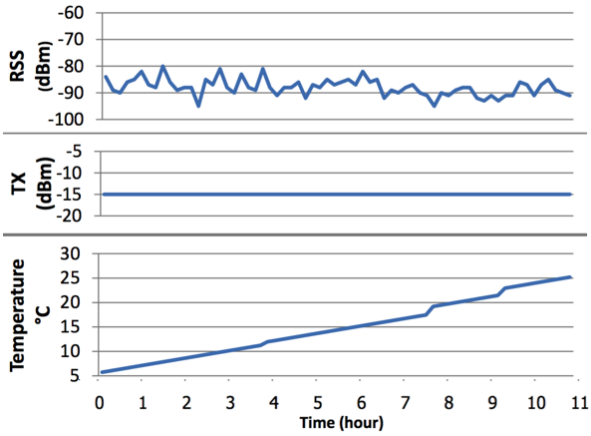
โดยที่ T_{ns} และ T_{nr} เป็นค่าอุณหภูมิใหม่ของผู้ส่งและตัวรับสัญญาณ ค่า τ คือค่าเอฟเฟ็กต์อุณหภูมิตามแต่ละแพลตฟอร์มวิทยุเช่นประมาณ 0.2 dBm สำหรับ CC2420 ซึ่งหาได้จากการทดลองผลกระทบอุณหภูมิในหัวข้อก่อนหน้า

4. ผลการทดลอง

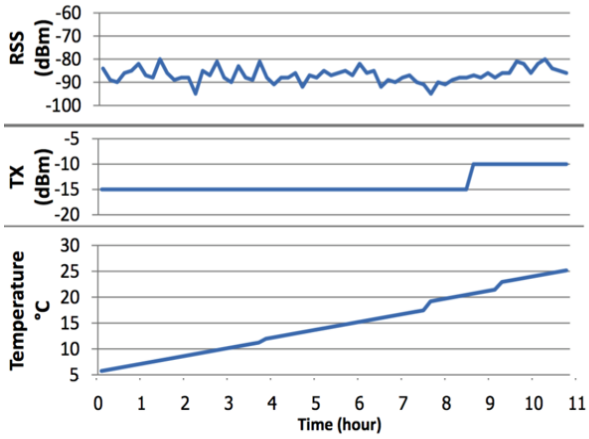
เพื่อตรวจสอบสมการข้างต้น จะมีวัดผลและเปรียบเทียบระหว่างสองเทคนิคคือ Static ตามสมการ (4) และสมการที่ใส่ผลต่ออุณหภูมิ คือ Temp. Effect ตามสมการ (5) โดยจะมีสองโหนดเรียกว่าผู้ส่งและตัวรับ จะอยู่ในสภาพแวดล้อมควบคุมโดยมีอุณหภูมิเพิ่มขึ้นจากเริ่มต้นประมาณ 5°C จนกระทั่งค่าอุณหภูมิ 25°C ค่าเริ่มต้น RSS_{max} จากผู้ส่งไปยังเครื่องรับคือ -70 dBm โดยค่าการส่งข้อมูลด้วยระดับพลังงานสูงสุด (P_{max}) คือ 0dBm) ค่าสัญญาณขั้นต่ำที่รับสัญญาณ (RSS_{min}) คือ -92 dBm ดังนั้นตามสมการ (4) ระดับพลังงานในการส่ง (TX) ที่เหมาะสมของวิทยุแบบ CC2420 สำหรับการส่งคือ -15 dBm จากผลการส่งแบบ Static จะให้ค่าสิ้นเปลืองพลังงานต่ำดัง แสดงในตารางที่ 1 อย่างไรก็ตามเมื่ออุณหภูมิเพิ่มขึ้นตามภาพที่ 4 และ 5 เทคนิคแบบ static จะไม่เปลี่ยนกำลังในการส่งข้อมูล ดังนั้นอัตราข้อมูลสูญหายจะมากขึ้น ในทางตรงกันข้ามเทคนิค Temp. Effect จะให้ผลการใช้พลังงานจะสูงกว่าเล็กน้อย แต่มีอัตราข้อมูลสูญหายที่ต่ำกว่า

ตารางที่ 1: ผลการทดลองจากการเพิ่มค่าอุณหภูมิ

Technique	Packet Loss (%)	Energy (J)
Static	6.06	444.0
Temp. Effect	3.03	444.5



ภาพที่ 4: อุณหภูมิที่เพิ่มขึ้นกับการทำงานของเทคนิค Static



ภาพที่ 5: อุณหภูมิที่เพิ่มขึ้นกับการทำงานของเทคนิค Temp. Effect

5. สรุป

ในบทความนี้ได้มีการตรวจสอบผลกระทบของอุณหภูมิ ต่อ RSS นอกจากนี้ยังนำเสนอสมการที่ชดเชยผลของอุณหภูมิด้วยการทดสอบ เพื่อประเมินประสิทธิภาพของกลไกที่เสนอเทียบกับสมการก่อนหน้านี้ โดยผลการทดลองแสดงให้เห็นว่าเทคนิคที่นำเสนอใหม่มีประสิทธิภาพโดยรวมที่ดีกว่าเทคนิคที่มีอยู่เช่น ทำให้อัตราการสูญหายแพ็คเก็ตต่ำลง ในขณะที่ใช้พลังงานเพิ่มขึ้นเล็กน้อย ซึ่งการสูญหายของแพ็ตเก็ตหมายถึงการส่งข้อมูลที่ใช้พลังงานโดยสูญเปล่า ดังนั้นการปรับเปลี่ยนการใช้พลังงานควรต้องคำนึงถึงผลกระทบของอุณหภูมิ ต่อสัญญาณในการรับส่งข้อมูลประกอบด้วย

งานวิจัยต่อไปควรจะพิจารณาถึงปัจจัยอื่นที่มีผลกระทบต่อสัญญาณของข้อมูลไร้สาย เช่น ความชื้น หรือ สัญญาณรบกวน เป็นต้น

เอกสารอ้างอิง

[1] 2.4 GHz IEEE 802.15.4 / ZigBee-ready rf tranceiver. CC2420 Data Sheet (Rev. B), March 2007.

[2] J. Gomez , “A case for variable-range transmission power control in wireless multihop networks” *INFOCOM 04*; 425–1436, 2004.

[3] J. Philip Monks, “Transmission power control for enhancing the performance of wireless packet data networks” *Technical report* 2001.

[4] E S. Jung and N H. Vaidya, *A power control mac protocol for ad hoc networks*, 2002.

[5] S. Lin, J. Zhang, G. Zhou, L. Gu, T. He and J A. Stankovic, “ATPC: Adaptive transmission power control for wireless sensor networks,” *Proceedings of the Fourth International Conference on Embedded Networked Sensor Systems (SenSys)*, 2006.

[6] K. Bannister, G. Giorgetti, K. Eep and S. Gupta “Wireless sensor networking for hot applications: Effects of temperature on signal strength, data collection 32. and localization” *Proceedings of the ACM 15th Workshop on Embedded Networked Sensors* 2008.

[7] J. Lee and K. Chung “An efficient transmission power control scheme for temperature variation in wireless sensor networks,” *Sensors*, Vol. 3, pp. 3078–3093, 2003.

[8] C. Boano, J. Brown, Z. He, U. Roedig, and T. Voigt “Low-Power Radio Communication in Industrial Outdoor Deployments: The Impact of Weather Conditions and ATEX-compliance” *Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*, pp. 159–176, 2009.

[9] W. Rukpakavong, I. Phillips and L. Guan,“ Neighbour discovery for transmit power adjustment in ieee 802.15.4 using RSSI” *Proceedings of New Technologies, Mobility and Security (NTMS), 2011 4th IFIP International Conference*, 2011.

[10] Sensinode Ltd “Sensinode ltd. joins texas instruments low-power rf developer network, offers ipv6 wireless network solutions based on 6lowpan standard technology”, 2008.

[11] J. Luomala and I. Hakala “Effects of Temperature and Humidity on Radio Signal Strength in Outdoor Wireless Sensor Networks” *Proceedings of the Federated Conference on Computer Science and Information Systems* 2015.

ระบบป้องกันการโจรกรรมรถจักรยานยนต์โดยใช้เทคนิคการรู้จำใบหน้าผ่านคลาวด์ภายใต้แนวคิดอินเตอร์เน็ตเพื่อทุกสิ่ง

Anti-Theft Motorcycle System using Face Recognition via Cloud Computing under Concept on Internet of Things

พชรอร เพชรสมุทร (Patcharaon Petchsamutr)[1] และ มหศักดิ์ เกตุฉ่ำ (Mahasak Ketcham)[1]
ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
[1]s5807021857521@email.kmutnb.ac.th, [2]mahasak.k@it.kmutnb.ac.th

บทคัดย่อ

ปัญหาพิเศษนี้มีวัตถุประสงค์เพื่อพัฒนาระบบป้องกันการโจรกรรมรถจักรยานยนต์โดยใช้เทคนิคการรู้จำใบหน้าผ่านคลาวด์ภายใต้แนวคิดอินเทอร์เน็ตเพื่อทุกสิ่ง สำหรับใช้ในการป้องกันการโจรกรรมรถจักรยานยนต์ ในปัญหาพิเศษนี้ได้ใช้เทคนิค การตรวจจับใบหน้า (Face Detection) และการรู้จำใบหน้า (Face Recognition) โดยการประมวลผลผ่านบอร์ด Raspberry Pi โดยวัดผลจากการนำภาพใบหน้าบุคคลที่รับจากกล้องมาทำการเปรียบเทียบภาพใบหน้าบุคคลที่เจ้าของรถจักรยานยนต์เมื่อตรวจพบว่าภาพใบหน้าบุคคลที่เป็นเจ้าของรถจักรยานยนต์และภาพใบหน้าบุคคลที่รับจากกล้องนั้นไม่ตรงกัน ระบบจะทำการแจ้งเตือนบนแอปพลิเคชันผ่านระบบเครือข่ายอินเทอร์เน็ต 3G โดยใช้อุปกรณ์ 3G Shield และ GPS Module เพื่อติดตามรถจักรยานยนต์ โดยส่งเป็นภาพใบหน้าบุคคลที่รับจากกล้อง และบอกตำแหน่งที่ตั้งของรถจักรยานยนต์ ณ ปัจจุบันทันที

จากการทดสอบระบบ อุปกรณ์และแอปพลิเคชันสามารถทำการแจ้งเตือนบอกตำแหน่งของรถจักรยานยนต์และภาพที่รับเข้ามาจากกล้องได้อย่างแม่นยำ เพื่อช่วยเพิ่มประสิทธิภาพในการป้องกันการโจรกรรมจักรยานยนต์ และสามารถนำไปใช้ได้จริง

คำสำคัญ: รถจักรยานยนต์ การรู้จำใบหน้า การประมวลผลแบบกลุ่มเมฆ อินเทอร์เน็ตเพื่อทุกสิ่ง

Abstract

The Special problem aimed at to develop an Anti-Thief Motorcycle System Using Image Recognition via Cloud Computing under Concept on Internet of Things for use in the prevention of theft motorcycle. The special problem is used face detection and face recognition techniques are processed by the raspberry pi board by measuring the image received from the camera to compare the original owners of the motorcycle when it detects that the original image and the image received from the camera that does not match. The system will be notified of the application via the 3G network using 3G Shield and GPS Module to track of the motorcycle. By sending a face image receiving by a camera and tell the location of the current a motorcycle immediately.

The results of the system are equipment and application to alert position of motorcycle and incoming image from the camera accurately for increase the effectiveness of the theft motorcycle and the system can use in the current.

Keyword: Motorcycle, Face Recognition, Cloud Computing, Internet of Thing.

1. บทนำ

ประเทศไทยมีผู้ใช้รถจักรยานยนต์เป็นจำนวนมาก โดยจากข้อมูลสถิติรถจักรยานยนต์ ที่จดทะเบียนในประเทศไทยกับทางกรมการขนส่งทางบก ณ วันที่ 31 ธันวาคม 2558 [1] มีจำนวนทั้งสิ้น 35,546,520 คัน โดยแบ่งเป็น รถจักรยานยนต์ส่วนบุคคล (Motorcycle) จำนวน 20,308,201 คัน และ รถจักรยานยนต์สาธารณะ (Public Motorcycle) จำนวน 189,362 คัน รวม ทั้ง 2 ประเภท 20,407,563 คัน จากสถิติข้างต้นรถจักรยานยนต์ ถือ

เป็นพาหนะที่มีผู้ใช้งานจำนวนมาก เนื่องจากการใช้รถยนต์ที่มากขึ้น มีการจราจรแออัด รถจักรยานยนต์สามารถช่วยให้ผู้คนเดินทางได้สะดวกที่สุด ยานพาหนะมีขนาดเล็ก สามารถใช้เส้นทางลัดต่างๆเพื่อไปให้ถึงจุดหมายได้ตามเวลาที่นัดหมาย และมีการใช้งานอย่างแพร่หลายในทุกๆสายอาชีพ

การโจรกรรม หมายถึง การกระทำที่นำเอาทรัพย์สินของผู้อื่นโดยเจตนา เพื่อยึดทรัพย์เหล่านั้นมาเป็นของตน ซึ่งทรัพย์ไม่ได้หมายถึงเฉพาะเงินทอง อาจหมายถึงสิ่งอื่นที่มีค่า เช่น ข้อมูล ความลับ ทรัพย์สินทางปัญญา [2] ซึ่งปัจจุบันการโจรกรรมรถจักรยานยนต์มีหลากหลายรูปแบบ การงัดเบ้ากุญแจ การเหลาคอรถ การใช้กุญแจปลอม การต่อสายตรง รวมถึงการยกรถไป ซึ่งทุกการกระทำสร้างความเสียหายและความเดือดร้อนให้กับเจ้าของรถอีกด้วย เนื่องจากเป็นยานพาหนะที่ใช้ในชีวิตประจำวัน สถิติจากสำนักงานตำรวจแห่งชาติ ว่าด้วยเรื่องการโจรกรรมรถจักรยานยนต์ ข้อมูล ณ ปี 2544-2553 [3] ได้ระบุถึงจำนวนการรับแจ้งความการโจรกรรมที่มากขึ้น แต่การติดตามส่งคืนเจ้าของนั้นน้อยลง



ภาพที่ 1 : การใช้เหล็กงัดรูกุญแจรถจักรยานยนต์ (ซ้าย) และ การต่อสายตรงรถจักรยานยนต์ (ขวา)

Zhigang Liu, Anqi Zhang และ Shaojun Li [4] ระบบนี้ควบคุมโดยโมดูล RFID เพื่อเปิดและปิด เมื่อรถถูกขโมยจะมีการเซ็นเซอร์การสั่นสะเทือนและเซ็นเซอร์อินฟราเรดแบบ Pyroelectric และโมดูล GSM จะส่งข้อมูลตำแหน่งที่ได้รับจากโมดูล GPS ไปยังโทรศัพท์มือถือของเจ้าของ ดังนั้นเจ้าของสามารถตรวจสอบตำแหน่งของยานพาหนะด้วยโทรศัพท์มือถือระบบปฏิบัติการแอปพลิเคชันแอนดรอยด์ ซอฟต์แวร์ที่พัฒนาขึ้นโดยทีมงาน ระบบนี้ใช้โทรศัพท์มือถือระบบปฏิบัติการแอปพลิเคชันแอนดรอยด์ เป็นโทรศัพท์เคลื่อนที่ซึ่งสะดวกและยืดหยุ่นกว่าผลิตภัณฑ์ชนิดอื่นๆ เนื่องจากเจ้าของสามารถตรวจสอบและติดตามตำแหน่งของรถยนต์ได้ทันทีด้วยแอปพลิเคชัน เมื่อรถยนต์ถูกโจรกรรม

เพื่อแก้ไขปัญหาดังกล่าว ผู้วิจัยจึงมีแนวคิดที่จะนำเทคนิคการรู้จำใบหน้ามาใช้ร่วมกับระบบป้องกันการสูญหายมาใช้ในรถจักรยานยนต์ โดยใช้เทคนิค SIFT (Scale Invariant Feature Transform) มาช่วยในการเปรียบเทียบความคล้ายคลึงระหว่างภาพใบหน้าบุคคลที่รับเข้าจากกล้องและภาพใบหน้าบุคคลที่เป็นเจ้าของรถจักรยานยนต์บนฐานข้อมูลคลาวด์ เพื่อช่วยในการป้องกันโจรกรรมรถจักรยานยนต์ โดยการจับภาพใบหน้าบุคคลที่กระทำการโจรกรรมรถจักรยานยนต์ เพื่อเป็นหลักฐานในการจับกุมผู้กระทำการโจรกรรม ด้วยอัลกอริทึมที่ใช้สามารถระบุตัวตนของผู้ต้องสงสัย ที่กระทำการโจรกรรม ณ ขณะนั้น ส่งผลให้สะดวกต่อการระบุตัวตนผู้ต้องสงสัยในการก่อเหตุการโจรกรรมขึ้น

2. หลักการและเทคนิคที่ใช้ในการวิจัย

2.1 ขั้นตอนการรู้จำใบหน้า (Face Recognition Algorithm)

ขั้นตอนในการรู้จำใบหน้าในปัจจุบันสามารถ จำแนกตามวิธีการได้ดังนี้

2.1.1 Knowledge-based methods วิธีการนี้ใช้กฎเกณฑ์พื้นฐานความรู้ของมนุษย์ของส่วนประกอบสำคัญบนใบหน้า อุปสรรคที่สำคัญของวิธีการนี้ คือ การหากฎเกณฑ์ที่จำเพาะเจาะจงซึ่งใช้ในการจำแนกหน้าของมนุษย์นั้นทำได้ยาก เพราะถ้าหากระบุรายละเอียดมากเกินไปจะมีผลทำให้การตรวจจับใบหน้าทำได้ยาก หรือ ถ้าภาพใบหน้านั้นมีรายละเอียดน้อยเกินไป จะทำให้การตรวจจับใบหน้าเกิดข้อผิดพลาดได้ง่าย และนอกเหนือจากนั้น สำหรับการตรวจจับใบหน้าที่มีตำแหน่งทิศทางนอกเหนือไปจากใบหน้า จะทำการตรวจจับทำได้ยากและมีความซับซ้อนมากขึ้น

2.1.2 Feature invariant approaches วิธีการนี้ใช้สำหรับการค้นหาลักษณะเด่น (Feature) ที่แตกต่างกันของแต่ละใบหน้าในการตรวจจับสมมติฐานของวิธีการดังกล่าวตั้งอยู่บนพื้นฐานของการสังเกตของมนุษย์ซึ่งสามารถที่จะจดจำวัตถุใดๆ ได้ไม่ว่าจะมีการเปลี่ยนแปลงตำแหน่งการจัดวาง หรือสภาพแสงที่แตกต่างกันเป็นต้น ซึ่งจะต้องมีคุณลักษณะบางอย่างที่ไม่ขึ้นอยู่กับผลของการเปลี่ยนแปลงดังกล่าวในการ ตรวจหาลักษณะเด่นบนใบหน้า

ปัญหาหนึ่งของวิธีการนี้คือ คุณลักษณะบนใบหน้าสามารถถูกทำลายได้โดยง่ายด้วยสภาพแสง สัญญาณ รบกวน และการถูกบดบัง อีกทั้งขอบของจุดเด่นบนใบหน้าอาจจะตรวจพบได้ยากในบางกลุ่มใบหน้า รวมถึงสภาพแสงเงาก็ยังมีผลกระทบโดยตรงกับการตรวจจับขอบของวัตถุ

2.1.3 Template matching method เป็นการเปรียบเทียบภาพที่ต้องการค้นหากับโครงสร้างแบบจำลองของใบหน้ามาตรฐาน โดยเก็บข้อมูลความสัมพันธ์อย่างอิสระของส่วนต่างๆ บนใบหน้าได้แก่โครงสร้างใบหน้า ตา จมูก และปาก ในท่าหน้าตรงตัวอย่างเช่น Sakai, Nagao and Fujibayashi, (1969) เสนอโครงสร้างแบบจำลองแผ่นแบบย่อย (subtemplates model) โดยใช้ตัวกรองโซเบล (sobel filter) หาเส้นขอบ เพื่อหาตำแหน่งความน่าจะเป็นของส่วนย่อยต่างๆบนใบหน้าที่สามารถเข้ากันได้ดีที่สุดกับแบบจำลองแผ่นแบบย่อยให้เป็นตำแหน่งของใบหน้าที่ต้องการต่อมา Tsukamoto, Lee, and Tsuji (1994) เสนอแบบจำลองคุณภาพสำหรับรูปแบบหน้า (qualitative model for face pattern : QMF) โดยใช้พารามิเตอร์ของแสงสว่างและเส้นขอบใบหน้าเป็นแบบจำลองของใบหน้า และเทคนิคที่นิยมใช้อีกอันหนึ่งคือการวิเคราะห์องค์ประกอบหลัก (principal component analysis: PCA) โดย Jung, Lee C.W., and Lee Y.C (2002) ศึกษาการค้นหาตำแหน่งที่แน่นอนของใบหน้าโดยใช้หลักการพื้นฐานทางด้านเรขาคณิต วิเคราะห์องค์ประกอบของใบหน้าให้อยู่ในรูปของสมการทางคณิตศาสตร์

2.1.4 Appearance-based methods วิธีการนี้จะใช้โมเดลที่ได้มาจากการเรียนรู้ของกลุ่มของรูปภาพตัวอย่าง ทั้งที่เป็นใบหน้า และไม่ใช่ใบหน้าโดยทั่วไปแล้ววิธีการนี้จะอยู่บนพื้นฐานของเทคนิคการวิเคราะห์ทางสถิติและการเรียนรู้ของเครื่อง (Machine Learning) ซึ่งใช้ในการหาคุณลักษณะความสัมพันธ์ของรูปใบหน้า และรูปที่ไม่ใช่ใบหน้า คุณลักษณะที่ได้รับการเรียนรู้จะถูกจัดในรูปแบบของแบบจำลองการกระจาย (Distribution Models) หรือ ฟังก์ชันการจำแนก (Discriminant Functions) ในขณะเดียวกัน การลดมิติ (Dimension Reduction) ก็ถูกนำมาใช้เพื่อ เพิ่มประสิทธิภาพในการคำนวณ และการตรวจจับ ใบหน้าอีกด้วย

2.1.5 วิธีหน้าลักษณะเฉพาะ (Eigenface Methods) เป็นการวิเคราะห์องค์ประกอบพื้นฐานของใบหน้า ด้วยกระบวนการทางสถิติของใบหน้าที่หลากหลาย โดยหน้าลักษณะเฉพาะคือเซตของเวกเตอร์ลักษณะเฉพาะ (Eigen vector) ที่สามารถหาได้จากเมทริกซ์ความแปรปรวน ร่วมเกี่ยว (covariance matrix) สร้างเป็นแบบจำลองของใบหน้าที่รวมเอาลักษณะเด่นต่างๆของภาพใบหน้าตัวอย่างมาร่วมกันเพื่อหาค่าเฉพาะเจาะจงขององค์ประกอบบนใบหน้า ซึ่งในแต่ละบุคคลจะเป็นค่าเฉพาะของบุคคลนั้นๆ

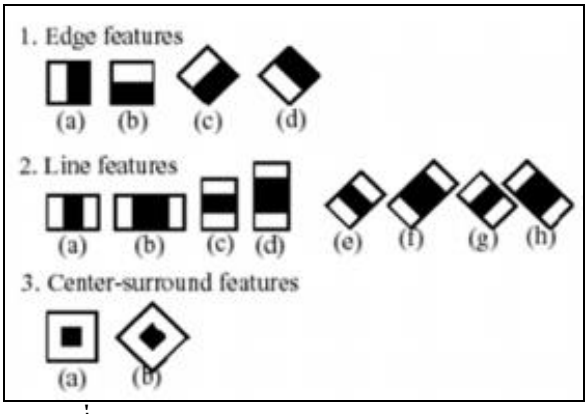
2.2 SIFT (Scale Invariant Feature Transform)

นำมาใช้ในการวิเคราะห์เพื่อหาจุดเด่นของรูปภาพที่ได้รับเข้ามา โดยไม่ขึ้นกับขนาดหรือทิศทางของวัตถุในภาพ ซึ่งลักษณะเด่นที่ได้จากวัตถุลักษณะเดียวกันจะมีลักษณะที่คล้ายคลึงกัน จึงมักนามาใช้เกี่ยวกับการจับคู่วัตถุจากภาพสองภาพที่มีวัตถุในภาพเหมือนกันที่ได้รับความนิยมในการศึกษาเพื่อพัฒนาปรับปรุงประสิทธิภาพในการรู้จำหรือจาแนกวัตถุจากภาพ โดยมีขั้นตอนหลักๆ ได้แก่ การตรวจหาคุณลักษณะเด่น (Feature detection) และการจับคู่ภาพ (Key point matching) ที่มีประสิทธิภาพในการรู้จำได้รวดเร็วโดยทดสอบความทนทานต่อการเปลี่ยนแปลงที่เกิดขึ้นต่างๆ จึงนับได้ว่าเป็นวิธีการพื้นฐานในการพัฒนาปรับปรุงประสิทธิภาพในขั้นตอนการตรวจหาคุณลักษณะเด่นของวิธีอื่น ๆ ได้แก่ PCA-SIFT, GSIFT, ASIFT และ SURF เป็นต้น การคำนวณเพื่อตรวจหาจุดสนใจ (Key point) ในภาพที่มีการเปลี่ยนแปลงเชิง 2 มิติของระดับความเข้มแสงรอบๆ จุดสนใจนั้น อัลกอริทึมนี้ได้รับความนิยมโดยมีผู้นำมาศึกษาและพัฒนาปรับปรุงประสิทธิภาพอย่างเนื่อง โดยเฉพาะการพัฒนาในหัวข้อของการรู้จำวัตถุจากภาพ (Object-based recognition)โดยมีขั้นตอนการทำงานหลัก ๆ 3 ขั้นตอน ได้แก่ การหาและกาหนดจุดสนใจ (Key point detection) การอธิบายจุดสนใจ (Key point description) และการจับคู่ภาพ (Matching)

2.3 Haar Like-Feature

Viola และ Jones ได้นำเสนอวิธีการตรวจจับและตีความวัตถุภายในภาพด้วยหลักการของ Haar Waralet สำหรับสร้างภาพสี่เหลี่ยมที่เรียกว่า Feature โดยภาพนี้แสดงถึงผลต่างระหว่างพื้นที่ส่วนสีขาว และส่วนที่เป็นสีดำ Feature นี้สามารถ

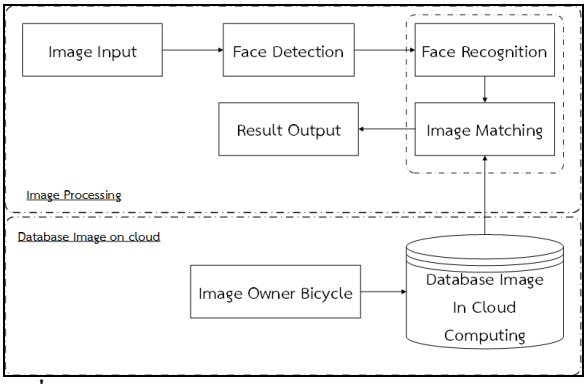
เปลี่ยนแปลงขนาดและตำแหน่งได้ เพื่อใช้สำหรับการตรวจจับลักษณะบนรูปแบบต่างๆเช่น เส้นตรง, วงกลม เป็นต้น ดังแสดงภาพที่ 2



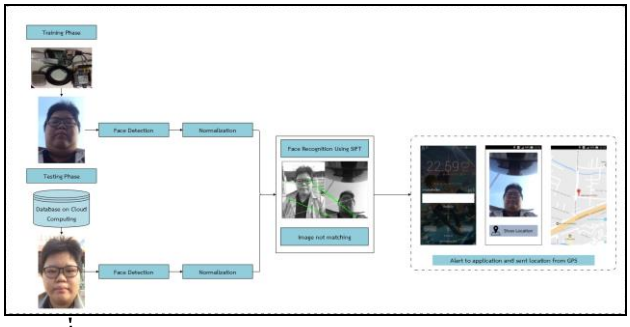
ภาพที่ 2 : รูปแบบของ Feature สำหรับการตรวจจับแบบต่างๆ

3. วิธีการดำเนินการวิจัย

การศึกษาและรวบรวมข้อมูลมุ่งเน้นศึกษาและรวบรวมพฤติกรรมการโจรกรรมและการป้องกันการโจรกรรมรถจักรยานยนต์ เพื่อนำมาพัฒนาการจัดทำปัญหาพิเศษนี้ ซึ่งพบว่าเทคโนโลยีปัจจุบันได้มีการนำรูปแบบ เช่น การแจ้งเตือนไปบนโทรศัพท์มือถือผ่านคลาวด์เพื่อให้ทราบว่ารถยนต์ที่จอดไว้มีการถูกโจรกรรมหรือไม่ การพัฒนาระบบรู้จำใบหน้าใช้ยืนยันตนเพื่อช่วยให้จับผู้ก่อเหตุได้รวดเร็วมากยิ่งขึ้น ดังนั้นผู้จัดทำปัญหาพิเศษจึงได้นำแนวคิดนี้มาช่วยในการพัฒนาปัญหาพิเศษนี้โดยสามารถนำมาเป็นองค์ความรู้ใหม่ๆและนำมาช่วยป้องกันการโจรกรรมได้ระดับหนึ่ง ซึ่งจะทำการ ซึ่งจะนำระบบที่พัฒนาขึ้นไปติดตั้งไว้ในรถจักรยานยนต์ก่อนนำไปทดสอบเพื่อให้ผลการทดสอบใกล้เคียงกับเหตุการณ์การโจรกรรม เพื่อให้ระบบสามารถทำงานได้อัตโนมัติและนำผลที่ได้มาดำเนินการให้เกิดประโยชน์สูงสุด โดยมีแนวคิดในการตรวจจับ 5 ขั้นตอน ดังแสดงในภาพที่ 3 และภาพรวมการทำงานของระบบดังแสดงในภาพที่ 4 และอธิบายขั้นตอนการทำงานของระบบดังต่อไปนี้



ภาพที่ 3 : ขั้นตอนกระบวนการของระบบป้องกันการโจรกรรมรถจักรยานยนต์โดยใช้เทคนิคการรู้จำภาพผ่านคลาวด์ภายใต้แนวคิดอินเทอร์เน็ตเพื่อทุกสิ่ง



ภาพที่ 4 : ภาพรวมของระบบป้องกันการโจรกรรมรถจักรยานยนต์โดยใช้เทคนิคการรู้จำภาพผ่านคลาวด์ภายใต้แนวคิดอินเทอร์เน็ตเพื่อทุกสิ่ง

3.1 กระบวนการในส่วนของการรับภาพ (Image Input)

กล้องจะทำการรับภาพใบหน้าบุคคลจากสถานที่ซึ่งเป็นการทำงานแบบ Real-Time โดยภาพที่ได้จะเป็นรูปภาพบุคคลที่เป็นอาจจะเป็นเจ้าของรถจักรยานยนต์หรือไม่ใช่เจ้าของรถจักรยานยนต์ ลักษณะในการติดตั้งกล้อง จะติดตั้งที่บริเวณใกล้เคียงกับหน้าปัดไมล์ของรถจักรยานยนต์ โดยภาพที่ได้จะเป็นทั้งภาพที่มีแสงสว่างมากและแสงสว่างน้อย โดยแสดงดังภาพที่ 5 ข้อมูลอีกส่วนหนึ่งคือ ภาพบุคคลที่เป็นเจ้าของรถจักรยานยนต์ซึ่งเก็บฐานข้อมูลรูปภาพไว้บนคลาวด์ โดยแสดงดังภาพที่ 6



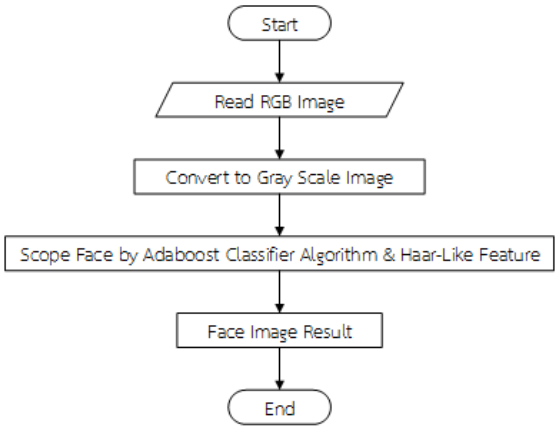
ภาพที่ 5 : แสดงการจับภาพใบหน้าบุคคลจากรถจักรยานยนต์ที่เป็นทั้งภาพที่มีแสงสว่างมากและแสงสว่างน้อย



ภาพที่ 6 : ภาพบุคคลที่เป็นเจ้าของรถจักรยานยนต์

3.2 กระบวนการตรวจจับใบหน้า (Face Detection)

ในขั้นตอนนี้ กล้องจะทำการจับภาพใบหน้าของบุคคลกระทำการโจรกรรมรถจักรยานยนต์ ณ ปัจจุบัน และทำการหาขอบเขตของบริเวณที่คาดว่าจะเป็นใบหน้า โดยแสดงขั้นตอนการตรวจจับใบหน้าที่ได้จากกล้องดังภาพที่ 7 นี้



ภาพที่ 7 : ขั้นตอนการตรวจจับภาพส่วนที่เป็นใบหน้า

จากแผนภาพการประมวลผลภาพด้านบนจะได้ผลลัพธ์เป็นใบหน้าของบุคคล จากรูปภาพจากกล้องที่ติดอยู่ที่หน้าปัดไมล์รถจักรยานยนต์ ซึ่งสามารถอธิบายอัลกอริทึมที่ใช้ได้ดังต่อไปนี้

1) Read RGB Image: นำเข้าภาพจากกล้องติดอยู่ที่หน้าปัดไมล์รถจักรยานยนต์ ซึ่งเป็นภาพ RGB color model

2) Convert to Gray Scale: แปลงภาพที่รับเข้าจาก RGB color model เป็น gray scale color model โดยใช้สมการในแปลงภาพจาก RGB เป็น Gray scale แสดงดังภาพที่ 8



ภาพที่ 8 : แปลงภาพที่รับเข้าจาก RGB เป็น Gray Scale

$$Gray = 0.299 \times R + 0.587 \times G + 0.114 \times B$$

3) Scope Face: หาตำแหน่งขอบเขตของใบหน้าโดยการใช้ Adaboost classifier Algorithm ซึ่งใช้พื้นฐานการตรวจจับใบหน้าของ Haar-like features ในการจำแนกส่วนที่เป็นใบหน้าออกจากภาพรวมทั้งหมด

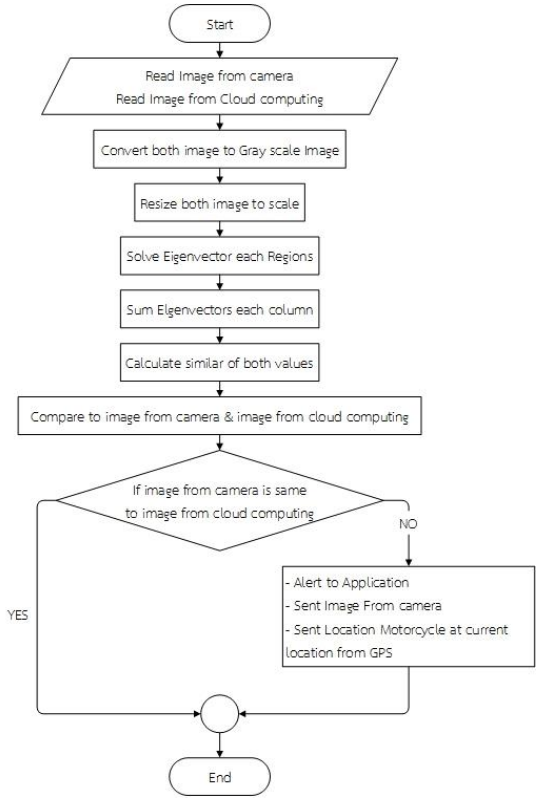
4) Face Image Result: ผลลัพธ์ที่ได้จะเป็นรูปภาพที่มีขอบเขตบริเวณหน้า แสดงดังภาพที่ 9



ภาพที่ 9 : รูปภาพที่มีขอบเขตบริเวณใบหน้า

3.3 กระบวนการรู้จำใบหน้า (Face Recognition) และกระบวนการเปรียบเทียบภาพใบหน้า (Image Matching)

ในขั้นตอนนี้ระบบจะทำการรู้จำใบหน้า และทำการเปรียบเทียบภาพใบหน้ากับฐานข้อมูล โดยใช้เทคนิคการจับคู่แบบ SIFT (Scale Invariant Feature Transform) เมื่อได้ข้อมูลรูปภาพใบหน้าที่ได้จากกล้องนั้นมาเปรียบเทียบกับภาพในฐานข้อมูล โดยมีกระบวนการดังภาพที่ 10 นี้



ภาพที่ 10 : ขั้นตอนการรู้จำใบหน้า และ การเปรียบเทียบใบหน้า

จากแผนภาพการประมวลผลภาพด้านบนจะได้ผลลัพธ์การรู้จำใบหน้าของบุคคล จากรูปภาพจากกล้องที่ติดอยู่ที่หน้าปัดไมล์รถจักรยานยนต์และภาพเจ้าของรถจักรยานยนต์

จากในฐานข้อมูลบนคลาวด์ ซึ่งสามารถอธิบายอัลกอริทึมที่ใช้ได้ดังต่อไปนี้

1) Read Image from camera & Image from cloud computing: นำเข้าภาพจากกล้องที่ติดอยู่ที่หน้าปัดไมล์รถจักรยานยนต์ ซึ่งเป็นภาพ RGB color model และภาพจากฐานข้อมูลบนคลาวด์

2) Convert both image to gray scale: แปลงภาพที่รับเข้าจากกล้อง และ ภาพจากฐานข้อมูลบนคลาวด์ เป็น gray scale

3) Resize both image to scale: ปรับขนาดของภาพทั้งสองให้มีขนาด 100x100 เท่ากัน เพื่อง่ายต่อการประมวลผล ดังภาพที่ 11



ภาพที่ 11 : ปรับขนาดภาพใบหน้าทั้งสองภาพ

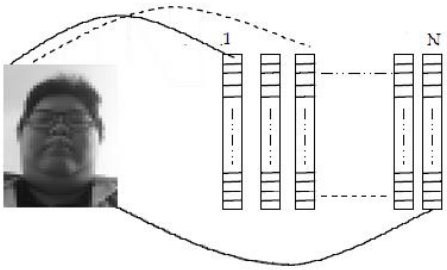
4) Solve Eigenvector each regions: คำนวณหาค่า eigenvectors ของแต่ละพื้นที่ (region) ดังภาพที่ 12 เพื่อหาความแตกต่างของแต่ละมิติ คือ



ภาพที่ 12 : คำนวณหาค่า eigenvectors

$$g_i = f_i - f_\mu \, , i = 1, 2, 3, \ldots, N_p$$

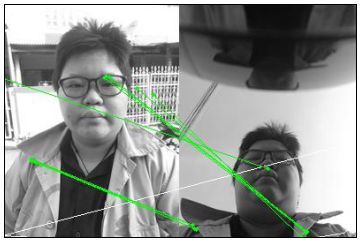
5) Sum Eigenvectors each column: หาค่าผลรวมของค่าความต่างของมิติในแต่ละคอลัมน์ $c = \sum_{n=1}^{N_p} G$ จะได้ชุดของค่าผลรวมของค่าความต่างของมิติในแต่ละคอลัมน์ $c_n \, , n = 1, 2, 3, \ldots, N$ ดังภาพที่ 13



ภาพที่ 13 : หาค่าผลรวมของค่าความต่างของมิติ

6) Calculate similar of both values: การนำค่า C_n ของทั้งภาพทั้งสองมาเปรียบเทียบถึงความใกล้เคียงกันของค่า C_n ดังกล่าว เพื่อหาผลลัพธ์ความใกล้เคียงกันของภาพใบหน้าบุคคลที่รับจากกล้อง และ ภาพใบหน้าบุคคลที่เป็นเจ้าของรถจักรยานยนต์ในฐานข้อมูลคลาวด์

7) Compare Image from camera & Image from cloud computing: ทำการเปรียบเทียบความคล้ายคลึงของภาพใบหน้าทั้งสองภาพโดยใช้เทคนิคการจับคู่แบบ SIFT (Scale Invariant Feature Transform) ดังภาพที่ 14

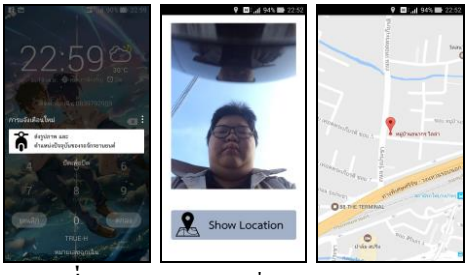


ภาพที่ 14 : การเปรียบเทียบรูปภาพทั้งสองโดยใช้เทคนิคการจับคู่แบบ SIFT (Scale Invariant Feature Transform)

โดยระบบมีเงื่อนไขว่าถ้าภาพที่รับจากกล้องติดที่หน้าปัดไมล์รถจักรยานยนต์ใกล้เคียงกับภาพเจ้าของรถจักรยานยนต์บนฐานข้อมูล จะไม่ทำการแจ้งเตือนใดๆ แต่ถ้าไม่มีความใกล้เคียงจะทำการแจ้งเตือนไปยังแอปพลิเคชันโดยระบบจะทำการส่งภาพบุคคลที่รับเข้าจากกล้องและแจ้งตำแหน่งปัจจุบันของรถจักรยานยนต์

3.4 การแสดงผลลัพธ์ (Result Output)

การแสดงผลลัพธ์ของการประมวลผล ระบบจะแสดงผลลัพธ์ไปยัง Application บนหน้าจอโทรศัพท์สมาร์ทโฟนบนระบบปฏิบัติการแอนดรอยด์ ในลักษณะการแจ้งเตือน (Alarm) โดยระบบจะแสดงผลเป็นโดยการส่งรูปภาพบุคคลที่กระทำการโจรกรรมรถจักรยานยนต์ และทำการแจ้งพิกัด ตำแหน่งที่อยู่ปัจจุบันของรถจักรยานยนต์ ดังแสดงตัวอย่างของโปรแกรมดังภาพที่ 15 ซึ่งสามารถช่วยให้ทำการติดตามหารถจักรยานยนต์ตามตำแหน่งของ GPS ได้



ภาพที่ 15 : ตัวอย่างระบบที่แสดงในแอปพลิเคชัน

4. ผลการดำเนินการวิจัย

ระบบป้องกันการโจรกรรมรถจักรยานยนต์ นี้ได้ทำการทดสอบโดยใช้บุคคลที่ไม่ใช่เจ้าของรถจักรยานยนต์จำนวน 5 คน ทดสอบกับระบบจำนวน 20 ครั้ง โดยใช้ทั้งภาพใบหน้าบุคคลที่เป็นทั้งเวลากลางวันและกลางคืนมาทำการเปรียบเทียบความคล้ายคลึงของภาพใบหน้าบุคคลที่เป็นเจ้าของรถจักรยานยนต์

GPS Module ใช้เวลาโดยเฉลี่ย 45-90 วินาทีในการติดตามตำแหน่งรถจักรยานยนต์ ในสภาพอากาศปกติ ทุกการติดตามตำแหน่งรถจักรยานยนต์ใช้เวลาโดยเฉลี่ย 150 วินาที จากเวลาของการแจ้งเตือนบนแอพลิเคชัน แต่ในสภาพอากาศที่แปรปรวนอาจใช้ระยะเวลาในการติดตามนานกว่า ในการแจ้งเตือนผ่านแอปพลิเคชันขึ้นอยู่กับความเร็วของเครือข่ายอินเทอร์เน็ต 3G และพื้นที่ที่ใช้งานเครือข่ายอินเทอร์เน็ต 3G ดังตารางที่ 1

ตารางที่ 1 : ระบบป้องกันการโจรกรรมรถจักรยานยนต์

การทดสอบระบบ	ระยะเวลาเฉลี่ยในการประมวลผล	ผลการทดสอบ
1. การรู้จำใบหน้า	45 วินาที	ร้อยละ 75
2. ความเร็วในการเข้าถึงข้อมูล	90 วินาที	ร้อยละ 60
3.ระยะเวลาในการแจ้งเตือน	90 วินาที	ร้อยละ 65
4. ตำแหน่งของรถจักรยานยนต์	60 วินาที	ร้อยละ 75

ผลการทดสอบระบบยังคงมีความคลาดเคลื่อนของการรู้จำใบหน้าที่รับเข้ามาจากกล้องและภาพใบหน้าบุคคลที่เป็นเจ้าของรถจักรยานยนต์ ในทางด้านสภาพอากาศที่มีความแปรปรวนทำให้การติดตามรถจักรยานยนต์ยากขึ้น และเครือข่ายอินเทอร์เน็ต 3G ขึ้นอยู่กับพื้นที่ที่ใช้งาน

5. สรุปผลการวิจัย

ระบบป้องกันการโจรกรรมรถจักรยานยนต์โดยใช้เทคนิคการรู้จำใบหน้าผ่านคลาวด์ภายใต้แนวคิดอินเทอร์เน็ตเพื่อทุกสิ่งใช้สำหรับเพื่อป้องกันการโจรกรรมรถจักรยานยนต์ โดยใช้เทคนิคการรู้จำใบหน้า (Face Recognition) จะมีการทดสอบโดยการนำเข้าข้อมูลภาพใบหน้าบุคคลที่รับเข้ามาจากกล้อง ซึ่งใช้กล้องและภาพใบหน้าบุคคลที่เป็นเจ้าของรถจักรยานยนต์จากฐานข้อมูลบนคลาวด์ หลังจากนั้นระบบจะทำการเปรียบเทียบภาพใบหน้าบุคคลทั้งสอง ว่ามีความคล้ายคลึงกันหรือไม่ โดยใช้เทคนิค SIFT (Scale Invariant Feature Transform) ถ้าทำการเปรียบเทียบภาพใบหน้าบุคคลทั้งสองภาพแล้วว่ามีความคล้ายคลึงกัน จะไม่ทำการส่งแจ้งเตือนใดๆ แต่ถ้าทำการเปรียบเทียบภาพใบหน้าบุคคลทั้งสองภาพแล้วว่าไม่มีความคล้ายคลึงกัน ระบบจะทำการแจ้งเตือนไปยังแอพลิเคชันส่งข้อมูลผ่านอินเทอร์เน็ตโดยใช้ 3G Shield เป็นตัวเชื่อมต่ออินเทอร์เน็ต โดยทำการส่งภาพใบหน้าบุคคลที่รับเข้าจากกล้องและทำการแจ้งพิกัดตำแหน่งของรถจักรยานยนต์ ณ ปัจจุบันโดยใช้ GPS Module ช่วยในการติดตามรถจักรยานยนต์

แนวทางการพัฒนาระบบป้องกันโจรกรรมรถจักรยานยนต์ต่อในอนาคต โดยการเพิ่ม IR Sensor เข้ามาตรวจจับการสั่นสะเทือนของการเคลื่อนที่รถจักรยานยนต์ ก่อนที่จะทำการรู้จำใบหน้าของบุคคล เพื่อสามารถลดภัยคุกคามและความสูญเสียจากการโจรกรรมรถจักรยานยนต์ได้

เอกสารอ้างอิง

[1] กรมการขนส่งทางบก. (2558). [ออนไลน์]. สถิติรถจักรยานยนต์จดทะเบียน. [สืบค้นวันที่ 20 กันยายน 2559].จาก https://data.go.th/DatasetDetail.aspx?id=19d8d378-6dfd-4a51-a16a-0dd5db308e14

[2] สำนักงานราชบัณฑิตยสภา. (2553). [ออนไลน์].โจรกรรม. [สืบค้นวันที่ 21 กันยายน 2559]. จาก http://www.royin.go.th/?knowledges=โจรกรรม-๕-เมษายน-๒๕๕๓

[3] LittleLee. (7 มีนาคม 2554). [ออนไลน์].สถิติรถจักรยานยนต์ถูกโจรกรรม. [สืบค้นวันที่ 24 กันยายน 2559].จาก http://www.oknation.net/blog/LittleLee/2011/03/07/entry-1

[4] Zhigang Liu, Anqi Zhang and Shaojun Li, “Vehicle Anti-theft Tracking System Based on Internet of things”, IEEE 2013

[5] Wei Yuan,“Automatic Tracking System of Vehicles Based on GPS and GSM,” Shanxi Electronic Technology ,vol.4, no.4, pp.28-29, April. 2011.

[6] HuaqunGuo,et al, “An Automotive Security System for Anti-Theft.”, 2009

[7] Shihab A. Hameed,et al, “Car Monitoring,Alerting and Tracking Model.”, 2010

การประเมินผลประสิทธิภาพการเข้ารหัสลับแบบ เออีเอส
AES: Performance Evaluation of Symmetric Encryption

ลักษนันท์ พลอยวัฒนาวงศ์ (Luxsanan Ploywattanawong)[1] และศิรปัฐช์ บุญครอง (Sirapat Boonkrong)[2]
[1]ภาควิชาเทคโนโลยีสารสนเทศ [2] ภาควิชาการสื่อสารข้อมูลและเครือข่าย
คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
[1] luxsanun29@gmail.com, [2] sirapat.b@it.kmutnb.ac.th

บทคัดย่อ

เทคโนโลยีรักษาความมั่นคงปลอดภัยของสารสนเทศเป็นเรื่องที่ท้าท้ายในการใช้งานอินเตอร์เน็ตและเครือข่าย ปัจจุบันการเข้ารหัสลับเป็นการป้องกันการบุกรุกประเภทหนึ่งที่ได้รับความนิยมในรับส่งข้อมูลที่เป็นความลับ ซึ่งมีนักวิชาการคิดค้นอัลกอริธึมหลายประเภทเพื่อใช้ในการเข้ารหัสลับ งานวิจัยนี้ได้ทำการศึกษาวิธีการเข้ารหัสแบบสมมาตร โดยการทดลองวิเคราะห์และเปรียบเทียบวิธีการเข้ารหัสลับของ AES ในด้านความเร็ว ระยะเวลา ขนาดของไฟล์การเข้าและถอดรหัส ถึงความแตกต่างกันของแต่ละรูปแบบและขนาดของกุญแจที่แตกต่างกันในการเข้ารหัสและถอดรหัสของข้อมูล ซึ่งพบว่าผลการเข้ารหัสแสดงให้เห็นว่า AES โหมด cfb ของกุญแจทั้ง 3 ขนาดมีประสิทธิภาพในด้านความเร็วที่ดีกว่า และใช้เวลาการเข้ารหัสน้อยที่สุด เมื่อเปรียบเทียบกับการเข้ารหัส AES โหมดอื่นๆ ที่ทดสอบร่วมกัน หากเทียบกันแล้ว cfb1 ใช้เวลาในการเข้ารหัสและถอดรหัสมากที่สุด แต่โดยรวม AES ทุกโหมดทำงานใช้เวลาในการเข้ารหัสและถอดรหัสต่างกันเพียงเล็กน้อย

คำสำคัญ: เออีเอส การเข้ารหัสลับ การเปรียบเทียบ เออีเอส

Abstract

Information technology security is the challenge of the Internet and network. Currently, the encryption that has gained popularity in the transmission of cryptography. The researchers created an algorithm for use in more types of encryption. This paper studies symmetric encryption. An analyzing and comparison mode of operation of encryption AES in speed, timing and size of the file encode and decode. To know the different styles and sizes of each of the different keys for encryption and decryption of data. It was found, the simulation results show that the AES mode cfb of the three key performance measures in terms the better of speed and minimal of time in encoding. In comparison to other tested AES encryption mode together. The Cfb1 takes to encode and decode the most compared to other modes. Therefore, all AES modes was worked encryption and decryption takes on a little different.

Keyword: AES, Encryption, Comparison AES.

1. บทนำ

นวัตกรรมและเทคโนโลยีการสื่อสารในปัจจุบันมีความก้าวหน้าอย่างต่อเนื่อง ส่งผลให้เกิดรูปแบบการสื่อสารบนช่องทางที่หลากหลาย นอกจากนั้นยังได้มีความพยายามที่จะพัฒนาระบบและอุปกรณ์ต่างๆ ให้สามารถเชื่อมต่อกันได้หลากหลายแม้ขณะกำลังเคลื่อนที่ ส่งผลดีด้านความสะดวกรวดเร็วในการติดต่อสื่อสารของมนุษย์ แต่การสื่อสารบนเครือข่ายนั้นยังคงไม่ปลอดภัยและอาจทำให้เกิดภัยคุกคามในการโจรกรรมข้อมูลความลับ ส่งผลให้เกิดความเสียหายต่อบุคคลและองค์กร จากจุดเริ่มต้นของความก้าวหน้าและปัญหาที่เกิดขึ้น ทำให้เกิดความคิดในการที่จะรักษาความลับในการรับส่งข้อมูล เราเรียกว่า เทคโนโลยีการเข้ารหัสลับ

มนุษย์รู้จักการเข้ารหัสมาตั้งแต่สมัยโรมัน นักวิชาการต่างๆ ได้ศึกษาวิธีการเข้ารหัสจากในอดีตมาเป็นพื้นฐานแนวคิดการรักษาความปลอดภัยของข้อมูลอิเล็กทรอนิกส์และพัฒนาเรื่อยมาจนถึงปัจจุบัน และกลายเป็นกระบวนการทางคณิตศาสตร์ในการเข้ารหัสลับ (Cryptographic Algorithms) โดยการสร้างสิ่งที่อยู่ในรูปตัวอักษร อักขระ ตัวเลข หรือสัญลักษณ์ใดๆ ขึ้นมา โดยเรียกสิ่งนั้นว่า “กุญแจ” (Key) และใช้

“กุญแจ” (Key) เป็นกลไกสำคัญในการ “เข้ารหัส” (Encryption) และ “ถอดรหัส” (Decryption)

การเข้ารหัสข้อมูลเป็นเทคโนโลยีที่ถูกนำขึ้นมาประกอบการตัดสินใจด้านการวางแผนระบบการจัดการสารสนเทศบนเครือข่าย ซึ่งองค์กรต่างให้ความสำคัญอย่างมากเพื่อรักษาความลับของชั้นข้อมูล ความต้องการของเทคโนโลยีการเข้ารหัสข้อมูล เช่น สามารถระบุตัวตนของผู้เข้าถึงข้อมูล (Authenticity) สามารถรักษาความลับ (Confidentiality) สามารถรักษาความถูกต้องและสมบูรณ์ของข้อมูล (Integrity) สามารถป้องกันการปฏิเสธความรับผิดชอบของการเข้าถึงข้อมูล (Non-repudiation)

จากการศึกษาพบว่า องค์กรส่วนใหญ่มีการจัดการข้อมูลในรูปแบบของอิเล็กทรอนิกส์ ซึ่งข้อมูลต่างๆ มีชั้นความลับที่แตกต่างกันตามตำแหน่งงานของแต่ละบุคคล และข้อมูลต่างๆ เหล่านั้นมีความสำคัญต่อความอยู่รอดทางธุรกิจขององค์กร จึงจำเป็นต้องมีการนำเทคโนโลยีการเข้ารหัสลับเข้ามาช่วยในการจัดการรักษาความลับในการรับส่งข้อมูลระหว่างบุคคลขึ้น ผู้วิจัยจึงได้ทำการศึกษาวิธีการเข้ารหัสแบบสมมาตร โดยการทดลอง วิเคราะห์และเปรียบเทียบวิธีการเข้ารหัสลับ (Mode of operation) ของ เออีเอส โดยนำมาเปรียบเทียบกันเพื่อให้เห็นถึงความแตกต่างกันของการใช้งานแต่ละรูปแบบและเลือกใช้งานได้อย่างถูกต้องเหมาะสมโดยเปรียบเทียบในด้านความเร็ว เวลา ขนาดของไฟล์การเข้าและถอดรหัสของการดำเนินการในแต่ละอัลกอริธึม เพื่อหาอัลกอริธึมที่มีรูปแบบการดำเนินงานที่มีประสิทธิภาพมากกว่า

2. วรรณกรรมที่เกี่ยวข้อง

เทคโนโลยีการเข้ารหัสลับหรือวิทยาการการเข้ารหัสลับเป็นวิธีการที่จะช่วยแปลงข้อความให้เป็นรหัสที่บุคคลอื่นไม่สามารถอ่านเข้าใจได้ หากบุคคลเหล่านั้นไม่มีกุญแจถอดรหัสข้อความด้วยขั้นตอนวิธีการ (Algorithm) การเข้ารหัสจึงเป็นกลไกรักษาความมั่นคงปลอดภัยที่นิยมใช้อย่างแพร่หลายในปัจจุบัน คำว่า “การเข้ารหัส” [1] เกี่ยวข้อกับศาสตร์และศิลป์ของการศึกษาด้านรหัสลับ (Cryptology)

2.1 การเข้ารหัสแบบ AES

AES (Advanced encryption standard) [2] เป็นวิทยาการเข้ารหัสที่เป็นที่นิยมในปัจจุบัน จากการออกแบบโดยชาวเบลเยี่ยมชื่อ Joan Daemen และ Vincent Rijmen ในปี ค.ศ. 2001 ด้วยวิธีการคำนวณที่ซับซ้อนตามลำดับขั้นตอน โดยจัดอยู่ในกลุ่มการเข้ารหัสแบบสมมาตร (Symmetric Key Cryptosystem) เป็นการเข้ารหัสอย่างง่าย โดยมีความเร็วในการทำงานที่ดี ใช้หน่วยความจำในการทำงานน้อย มีมาตรฐานที่กำหนดให้ Key มีขนาด 128, 192 และ 256 บิต ข้อมูลที่จะนำมาเข้ารหัสมีขนาด 128 บิต บนเมตริกซ์ 4 x 4 และมีรอบในการเข้ารหัสเป็นมาตรฐาน ดังนี้

1) ขนาด 128 bit keys เข้ารหัส 10 รอบ
2) ขนาด 192 bit keys เข้ารหัส 12 รอบ
3) ขนาด 256 bit keys เข้ารหัส 14 รอบ

โดยแต่ละรอบจะทำการเข้ารหัส 4 แบบ เพื่อให้เกิดความปลอดภัยสูง อัลกอริธึมจะทำการเข้ารหัส SubBytes, ShiftRows, MixColumns, และ XorRoundKey ตามลำดับ ซึ่งการถอดรหัสสามารถทำได้โดยกระบวนการย้อนกลับของการเข้ารหัสเพื่อให้ได้ค่า Plaintext ออกมาสมบูรณ์

สำหรับอัลกอริธึมที่นิยมใช้ ได้แก่ Rijindael Algorithm ซึ่งเป็นมาตรฐานที่มีประสิทธิภาพและมีความปลอดภัยสูง และยังสามารถนำไปใช้กับอุปกรณ์เคลื่อนที่ได้อย่างเช่น กำหนดเพียงให้เลื่อนพยัญชนะออกไปอีก 1 ตำแหน่ง “ดอกหอม” หากเลื่อนตำแหน่งพยัญชนะไป 1 ตัว จะได้การเข้ารหัสเป็น “ตฮขพฮย” ดังนั้นเมื่อผู้ส่งใช้กุญแจในการเข้ารหัสแล้ว ต้องส่งมอบกุญแจนั้นให้กับผู้รับอีกฝ่าย ซึ่งการเข้ารหัสแบบดังกล่าวต้องใช้กุญแจดอกเดียวกันในการถอดรหัส และต้องมีการเก็บรายละเอียดกุญแจเป็นความลับ เพื่อความปลอดภัยของข้อมูลอิเล็กทรอนิกส์

2.2 Cipher block modes operation

1) Cipher block chaining (CBC) มีการทำงานโดยข้อมูลที่จะป้อนเข้าสู่อัลกอริธึมการเข้ารหัสนั้น จะเกิดจากผลลัพธ์ของการทำ XOR ระหว่าง Plaintext block ปัจจุบันกับ Ciphertext block จากการทำงานก่อนหน้า โดยใช้คีย์เดียวกันในทุกการเข้ารหัส

ซึ่งจะเป็นผลให้ข้อมูลถูกเปลี่ยนก่อนที่จะเข้าสู่กระบวนการเข้ารหัส จากการเปลี่ยนนี้จะเกิดขึ้นในลักษณะแบบลูกโซ่

2) Cipher feedback (CFB) มีการทำงานคล้ายกับ CBC โดยค่า Ciphertext block ก่อนการนำไปประมวลผลการเข้ารหัสแล้วนำมา XOR กับ Plaintext block ปัจจุบัน

3) Electronic codebook (ECB) มีการแบ่งข้อมูลออกเป็นกลุ่ม ตามอัลกอริธึมที่เลือกใช้ เช่น 128 บิต สำหรับ AES หรือ 64 บิต สำหรับ DES จากนั้นเข้ารหัสด้วยกุญแจที่กำหนด สำหรับวิธีการ ECB นี้ให้ความปลอดภัยในระดับต่ำ เนื่องจากใช้กุญแจเดิมในการเข้ารหัสแต่ละชุดข้อมูล

4) Output feedback (OFB) มีการทำงานคล้ายกับ CFB แต่ข้อมูลที่ส่งต่อไปยังการเข้ารหัสชุดถัดไปคือรหัสที่มีการเข้ารหัสแล้วเพื่อให้มีการทำการเข้ารหัสซ้ำ

2.3 การทบทวนวรรณกรรม และสารสนเทศที่เกี่ยวข้อง

จากการศึกษาและทบทวนวรรณกรรมที่เกี่ยวข้องกับการเข้ารหัสลับแบบ AES ในการสร้างมุมมองและแนวคิดพบว่า การวิจัยด้านการออกแบบและบูรณาการ AES เพื่อออกแบบสถาปัตยกรรมการเข้ารหัสและถอดรหัสให้มีความรวดเร็ว เหมาะกับการใช้งานบนอุปกรณ์ต่างๆ [3] งานวิจัยเกี่ยวกับการประเมินประสิทธิภาพ AES ด้วยโปรแกรมที่ต่างจากงานวิจัยชิ้นนี้ [4] การศึกษาการเปรียบเทียบประสิทธิภาพของอัลกอริทึมการเข้ารหัสข้อมูล [5] นอกจากนั้นบางงานวิจัยศึกษาวิเคราะห์ประสิทธิภาพของอัลกอริธึม และนำมาเปรียบเทียบการเข้ารหัสลับแบบ DES, AES และ Blowfish [6] เปรียบเทียบประสิทธิภาพในการส่งของ AES บนพื้นฐานการประมวลผลที่แตกต่างกันโดยใช้ 32-bit CPUs, 64-bit CPUs, cheap 8-bit, smart-card CPUs, และฮาร์ดแวร์เฉพาะในการทดสอบ. [7] การบูรณาการเรื่องการรักษาความปลอดภัยของการเข้ารหัสลับแบบสมมาตร [8], [9] จะเห็นว่ามีงานวิจัยถูกพัฒนาขึ้นอย่างมากในด้านของการศึกษา วิเคราะห์ และเปรียบเทียบบนข้อจำกัดทางทรัพยากรที่แตกต่างกัน อาจได้ผลลัพธ์ในการทดสอบที่เหมือนหรือต่างกันได้

สำหรับการวิจัยนี้ผู้วิจัยมีความสนใจที่จะศึกษาวิเคราะห์เปรียบเทียบวิธีการเข้ารหัสลับแบบสมมาตรโดยเจาะจงทำการทดสอบกับแบบ AES โดยเริ่มจากการวางแผน เลือกเครื่องมือติดตั้ง จำลองเปรียบเทียบผล บนทรัพยากรที่มีอยู่ภายในองค์กร เพื่อนำผลที่ได้มาประยุกต์ใช้กับงานได้อย่างเหมาะสม

3. วิธีการดำเนินการวิจัย

งานวิจัยด้านความปลอดภัยในปัจจุบัน ได้มีการทดสอบบนเครือข่ายโครงสร้างพื้นฐานและอุปกรณ์สื่อสารที่แตกต่างกัน [10] สำหรับงานวิจัยนี้ได้ใช้การจำลองบนเครื่องแม่ข่ายภายในมหาวิทยาลัยเทคโนโลยีราชมงคลสุวรรณภูมิ ศูนย์สุพรรณบุรี สาขาเทคโนโลยีสารสนเทศ คณะวิทยาศาสตร์และเทคโนโลยี ใช้ VM Ware ในการจำลองและประมวลผลการทดลองบนเครื่องแม่ข่าย HP DL380 Gen7 ทำงานอยู่บน CPU 1 core RAM 512 MB และทดสอบด้วย Openssl โปรแกรมเปิดแหล่งจำลองเครือข่าย PuTTY ระบบปฏิบัติการ Linux Redhat CentOS 7.3 แสดงผลผ่านเครื่องโน๊ตบุกส่วนบุคคล การทดลองจะดำเนินการซ้ำ 15 ครั้ง เพื่อคำนวณหาค่าเฉลี่ยของช่วงเวลาในการประมวลผลการเข้ารหัสและถอดรหัสข้อมูลด้วยวิธีการเข้ารหัสแบบ AES ที่ใช้โหมดต่างกัน ดังตารางที่ 1

3.1 วางแผนการทดลอง

เลือกทดลอง Cryptographic Algorithms แบบ AES ซึ่งทำการเปรียบเทียบ และสร้างไฟล์บนโปรแกรม Openssl ด้วยการพิมพ์ข้อมูลชนิดอักขระทั้งหมดจนได้ไฟล์ขนาด 20214 ไบต์ จำนวน 1 ไฟล์และตั้งชื่อว่า cherry.txt ซึ่งเป็นไฟล์ Plaintext สำหรับนำไปใช้กับการเข้ารหัส จากนั้นทำการเข้ารหัส AES ทั้ง 18 แบบ ไปที่ไฟล์ orange1.txt ถึง orange18.txt และถอดรหัสไปที่ไฟล์ Melon1.txt ถึง Melon 18.txt ซึ่งมีการแยกไฟล์แสดงผลการเข้ารหัสและถอดรหัสออกจากกัน เพื่อให้เห็นค่าความเร็ว ความแตกต่างของขนาดของไฟล์ก่อนการเข้ารหัสและถอดรหัส เพื่อให้ตรงกับปัจจัยที่ศึกษาคือ ความเร็วการเข้ารหัส/ถอดรหัส และขนาดของข้อมูลที่ถูก Encrypt ของแต่ละ Key size ในแต่ละโหมด

ตารางที่ 1: การออกแบบและผลลัพธ์เวลาของ Algorithm ในทดลอง

Algorithm	Mode	Key Size (bits)	Block Size (bits)	Encryption Time (s)	Decryption Time (s)
AES	cbc	128	128	0.006	0.007
AES	cfb	128	128	0.003	0.005
AES	cfb1	128	128	0.020	0.019
AES	cfb8	128	128	0.007	0.005
AES	ecb	128	128	0.007	0.006
AES	ofb	128	128	0.006	0.004
AES	cbc	192	128	0.006	0.006
AES	cfb	192	128	0.004	0.006
AES	cfb1	192	128	0.021	0.022
AES	cfb8	192	128	0.006	0.006
AES	ecb	192	128	0.004	0.005
AES	ofb	192	128	0.006	0.004
AES	cbc	256	128	0.005	0.006
AES	cfb	256	128	0.004	0.006
AES	cfb1	256	128	0.026	0.027
AES	cfb8	256	128	0.006	0.006
AES	ecb	256	128	0.004	0.005
AES	ofb	256	128	0.005	0.004

3.2 ขั้นตอนการทดสอบ

เริ่มการทดสอบ 1) โดยทำการสร้างไฟล์ .txt ผ่านคำสั่งบนโปรแกรม Openssl บนเครื่องแม่ข่าย เพื่อใช้ในการเข้ารหัสและใช้ในแสดงผลการเข้ารหัส 2) ทำการเข้ารหัสจากไฟล์ที่สร้าง ในการทดลองนี้ให้ใช้ไฟล์ที่ชื่อ cherry.txt เป็นไฟล์ในการเข้ารหัสหรือเป็น plaintext ตามตารางการวางแผนทั้ง 18 แบบ 3) ทำการเข้ารหัสทั้ง 18 แบบ ที่มีความแตกต่างกันที่ Key size, Mode แต่ขนาดของไฟล์ plaintext เป็นไฟล์เดียวกันขนาดเท่ากัน 4) ทำการตรวจสอบผลการเข้ารหัสจากไฟล์แสดงผลการเข้ารหัส ในการทดลองนี้ให้แสดงผลในชื่อไฟล์ orange1… 18 5) ทำการถอดรหัสจากไฟล์ orange1… 18 และแสดงผลในชื่อไฟล์ melon1.txt และทำการบันทึกผลการทดลอง โดยผู้วิจัยจะทำการทดสอบเป็นจำนวนทั้งสิ้น 15 รอบ 6) นำผลการทดสอบไปคำนวณหาค่าเฉลี่ยรวมของช่วงเวลาในการประมวลผลการเข้ารหัสข้อมูลและถอดรหัสข้อมูลด้วยโปรแกรม Excel ดังสมการ

$$\text{mean } x = (w_1x_1 + w_2x_2 + w_3x_3 + \ldots + w_nx_n)/n$$

กำหนดให้ w_nx_n คือ เวลาที่ใช้ในการประมวลการเข้า/ถอดรหัส

n คือ จำนวนรอบของการเข้า/ถอดรหัส

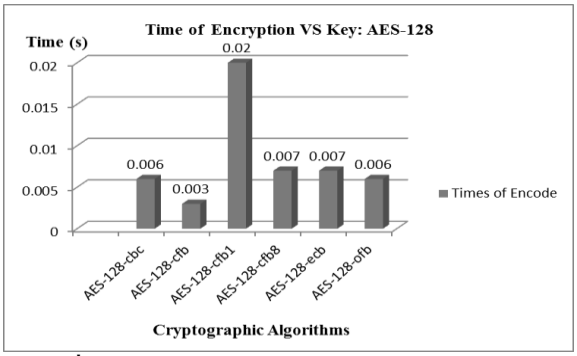
x คือ ผลรวมของเวลาในการเข้า/ถอดรหัส

4. ผลการดำเนินงาน

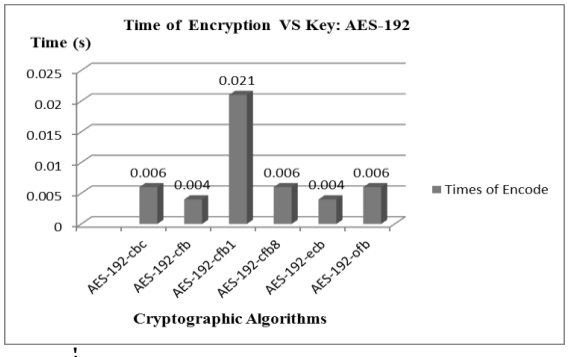
ผลการดำเนินงานที่เกิดขึ้นจากการใช้ Openssl แสดงให้เห็นความเร็วการโหลดข้อมูลและขนาดไฟล์ที่แตกต่างกัน จากการใช้งานขั้นตอนและวิธีการเข้ารหัสลับของแต่ละโหมด โดยแบ่งผลการเปรียบเทียบประสิทธิภาพเป็น 2 ด้าน

4.1 ผลการเปรียบเทียบประสิทธิภาพความเร็วในการเข้ารหัส

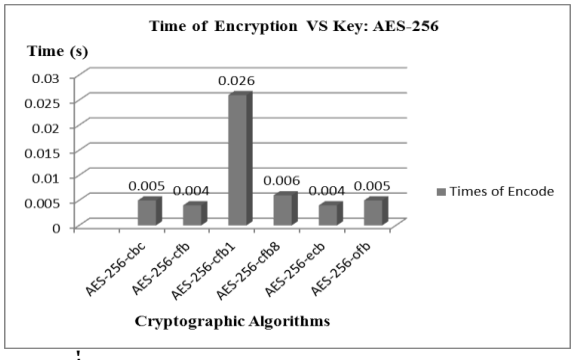
โปรแกรม Openssl ได้ทำการเข้ารหัสของอัลกอริธึม AES โดยใช้การเข้ารหัสขนาด 128, 192, และ 256 ซึ่งมีโหมดที่แตกต่างกันทั้ง 18 แบบ ผลการทดลองแสดงให้เห็นว่าแบบ AES-128 bit Mode cfb ใช้ความเร็วในการเข้ารหัสน้อยที่สุด 0.003s และแบบ AES-128 bit Mode cfb1 ใช้ความเร็วในการเข้ารหัสมากที่สุด 0.026s และในโหมดอื่นๆนั้นใช้ความเร็วไม่ต่างกันโดยใช้ความเร็วประมาณ 0.004s – 0.007s ดังภาพที่ 1-3



ภาพที่ 1: การเปรียบเทียบประสิทธิภาพความเร็วของ AES-128 ในการเข้ารหัสที่โหมดต่างกัน



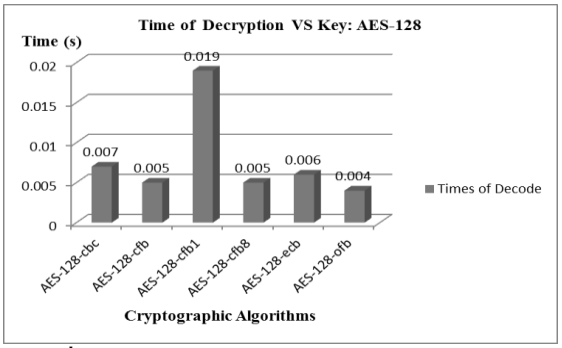
ภาพที่ 2: การเปรียบเทียบประสิทธิภาพความเร็วของ AES-192 ในการเข้ารหัสที่โหมดต่างกัน



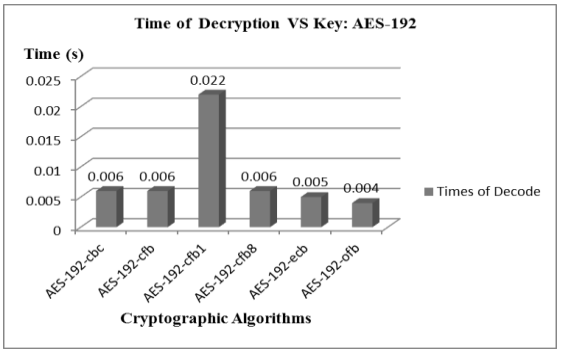
ภาพที่ 3: การเปรียบเทียบประสิทธิภาพความเร็วของ AES-256 ในการเข้ารหัสที่โหมดต่างกัน

4.2 ผลการเปรียบเทียบประสิทธิภาพความเร็วในการถอดรหัส

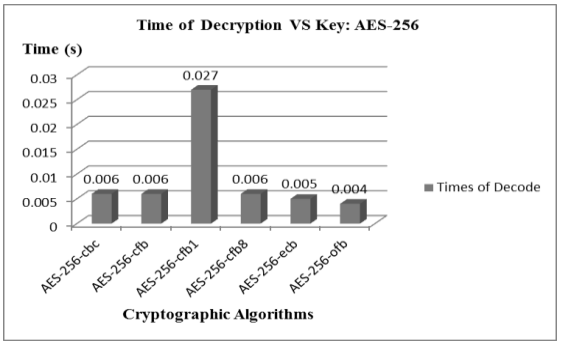
การถอดรหัสของอัลกอริธึม AES ทั้ง 18 แบบ แสดงให้เห็นว่า แบบ AES-128, 192, และ 256 bit Mode obf ใช้ความเร็วในการถอดรหัสน้อยที่สุด 0.004s และแบบ AES-256 bit Mode cfb1 ใช้ความเร็วในการถอดรหัสมากที่สุด 0.027s และในโหมดอื่นๆ นั้น ใช้ความเร็วไม่ต่างกัน โดยใช้ความเร็วประมาณ 0.005s – 0.007s ดังภาพที่ 4-6



ภาพที่ 4: การเปรียบเทียบประสิทธิภาพความเร็วของ AES-128 ในการถอดรหัสที่โหมดต่างกัน



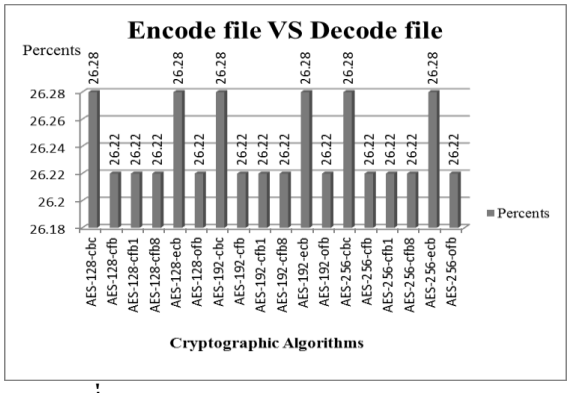
ภาพที่ 5: การเปรียบเทียบประสิทธิภาพความเร็วของ AES-192 ในการถอดรหัสที่โหมดต่างกัน



ภาพที่ 6: การเปรียบเทียบประสิทธิภาพความเร็วของ AES-256 ในการถอดรหัสที่โหมดต่างกัน

งานวิจัยนี้ให้ความสนใจในการเปรียบเทียบขนาดไฟล์ก่อนการเข้ารหัสและหลังการถอดรหัสย้อนกลับ โดยศึกษาถึงขนาดไฟล์ Plaintext, Ciphertext เพื่อสามารถนำไปวางแผนการบริหารจัดการเครือข่ายได้อย่างเหมาะสม

การเข้ารหัสของอัลกอริธึมทั้ง 18 แบบ แสดงให้เห็นถึงขนาดไฟล์ที่ทำการเข้ารหัสลับและขนาดไฟล์การถอดรหัสลับ พบว่า ไฟล์ที่ทำการ Encode และ Decode ทั้ง 18 แบบ มีขนาดไฟล์ของอัลกอริธึมแต่ละโหมดมีขนาดไม่ต่างกันมาก ขนาดไฟล์เข้ารหัสจะมีขนาดใหญ่กว่าขนาดไฟล์ที่ถูกถอดรหัส จะสังเกตได้จากสัดส่วนค่าร้อยละความต่างกันของแต่โหมด ห่างกันอยู่ที่ร้อยละ 26.22 – 26.28 ดังภาพที่ 7



ภาพที่ 7: การเปรียบเทียบขนาดไฟล์ก่อนการเข้ารหัสและหลังการถอดรหัส

5. สรุปผลการวิจัยและแนวทางในการพัฒนาต่อ

ผลการจำลองแสดงให้เห็นว่า AES โหมด cfb ของกุญแจทั้ง 3 ขนาดมีประสิทธิภาพในด้านความเร็วที่ดีกว่า ซึ่งใช้เวลาการเข้ารหัสน้อยที่สุด เมื่อเปรียบเทียบกับการเข้ารหัส AES โหมดอื่นๆ ที่ทดสอบร่วมกัน หากเทียบกันแล้ว cfb1 ใช้เวลาในการเข้ารหัสและถอดรหัสมากที่สุด แต่โดยรวม AES ทุกโหมดทำงานด้วยเวลาในการเข้ารหัสและถอดรหัสต่างกันเพียงเล็กน้อย นอกจากนั้นยังพบว่า หากเปรียบเทียบขนาดไฟล์ของการเข้ารหัสลับทั้ง 18 แบบ มีขนาดไฟล์ของอัลกอริธึมแต่ละโหมดมีขนาดไม่ต่างกันมาก แต่จะสังเกตได้ว่าขนาดไฟล์เข้ารหัสจะมีขนาดใหญ่กว่าขนาดไฟล์ถอดรหัส

สำหรับงานวิจัยนี้สามารถนำไปใช้ในการจำลองเปรียบเทียบกับการเข้ารหัสลับรูปแบบอื่นๆ อาจโดยการเปรียบเทียบบนเครื่องประมวลผลที่แตกต่างกัน จำนวนไฟล์ Plaintext ที่มีขนาดต่างกัน และทดสอบการเข้ารหัสด้วยอัลกอริธึมแบบต่างๆ เพื่อเปรียบเทียบความปลอดภัยสูงสุด และสามารถนำไปประยุกต์ใช้งานกับการเข้ารหัสลับได้อย่างเหมาะสม

6. กิตติกรรมประกาศ

การทำงานวิจัยครั้งนี้ประสบความสำเร็จลุล่วงด้วยดี เพราะความกรุณาของรองศาสตราจารย์ ดร.ศิรปัฐช์ บุญครอง ที่กรุณาให้คำแนะนำปรึกษาในการออกแบบงานวิจัย จนทำให้งานวิจัยครั้งนี้มีความสมบูรณ์

เอกสารอ้างอิง

[1] Stallings, W. (2006). *Cryptography and network security: principles and practices*. Pearson Education India.

[2] Daemen, J., & Rijmen, V. (2013). *The design of Rijndael: AES-the advanced encryption standard*. Springer Science & Business Media.

[3] Lu, C. C., & Tseng, S. Y. (2002). Integrated design of AES (Advanced Encryption Standard) encrypter and decrypter. In *Application-Specific Systems, Architectures and Processors, 2002. Proceedings. The IEEE International Conference on* (pp. 277-285). IEEE.

[4] Elbirt, A. J., Yip, W., Chetwynd, B., & Paar, C. (2000, April). An FPGA-based performance evaluation of the AES block cipher candidate algorithm finalists. In *proceedings of the 3rd AES Candidate Conference*.

[5] Nadeem, A., & Javed, M. Y. (2005, August). A performance comparison of data encryption algorithms. In *Information and communication technologies, 2005. ICICT 2005. First international conference on* (pp. 84-89). IEEE.

[6] Thakur, J., & Kumar, N. (2011). DES, AES and Blowfish: Symmetric key cryptography algorithms simulation based performance analysis. *International journal of emerging technology and advanced engineering*, *1*(2), 6-12.

[7] Schneier, B., Kelsey, J., Whiting, D., Wagner, D., Hall, C., & Ferguson, N. (1999). Performance comparison of the AES submissions.

[8] Fujisaki, E., & Okamoto, T. (1999, August). Secure integration of asymmetric and symmetric encryption schemes. In *Annual International Cryptology Conference* (pp. 537-554). Springer Berlin Heidelberg.

[9] Bellare, M., Desai, A., Jokipii, E., & Rogaway, P. (1997, October). A concrete security treatment of symmetric encryption. In *Foundations of Computer Science, 1997. Proceedings., 38th Annual Symposium on* (pp. 394-403). IEEE.

[10] Ahmad, R., & Ismail, W. (2017). Performance Comparison of the Improved Power-Throughput AES and Blowfish Algorithms on FPGA. In *9th International Conference on Robotic, Vision, Signal Processing and Power Applications* (pp. 19-25). Springer Singapore.

การเปรียบเทียบประสิทธิภาพการแยกตัวประกอบของตัวเลขจำนวนเต็ม โดยใช้อัลกอริทึม Fermat และ Pollard Rho

Comparison of Integer Factorization using Fermat and Pollard Rho Algorithm

ปริญญา นาโท (Parinya Natho)[1] และศิรปัฐช์ บุญครอง (Sirapat Boonkrong)[2]
[1]ภาควิชาเทคโนโลยีสารสนเทศ [2] ภาควิชาการสื่อสารข้อมูลและเครือข่าย
คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
[1]p.natho@hotmail.com, [2]sirapat.b@it.kmutnb.ac.th

บทคัดย่อ

การใช้งาน Cryptographic RSA เป็นการเข้ารหัสแบบอสมมาตรวิธีหนึ่งที่ได้รับความนิยมใช้งานในการจัดการด้านการรักษาความปลอดภัยของแอพพลิเคชั่น เช่น กระบวนการยืนยันตัวตน การสร้างลายเซ็นดิจิทัล โดยการเข้ารหัสแบบ RSA จะประกอบด้วยกุญแจส่วนตัวและกุญแจสาธารณะ ซึ่งได้จากการคำนวณจากตัวเลขสองจำนวนที่มีขนาดใหญ่ ปัญหาด้านความปลอดภัยของ RSA จะขึ้นอยู่กับความยากในการแยกตัวประกอบของตัวเลขที่มีขนาดใหญ่ ซึ่งในงานวิจัยฉบับนี้ได้ทำการเปรียบเทียบประสิทธิภาพการแยกตัวประกอบของตัวเลขจำนวนเต็มโดยใช้อัลกอริทึม Fermat และอัลกอริทึม Pollard Rho ทำการทดสอบด้วยตัวเลขที่เป็นจำนวนเต็มที่ได้จากการสุ่มของโปรแกรมขึ้นมาและตัวเลขจำนวนคี่ที่กำหนดตามรูปแบบ ผลการทดลองพบว่าตัวเลขที่เป็นจำนวนเต็มที่ได้จากการสุ่มขึ้นมาอัลกอริทึม Pollard Rho มีความเร็วในการแยกตัวประกอบมากกว่าอัลกอริทึม Fermat ส่วนตัวเลขที่เป็นจำนวนคี่พบว่าอัลกอริทึม Fermat จะมีความเร็วในการแยกตัวประกอบมากกว่าอัลกอริทึม Pollard Rho

คำสำคัญ: กุญแจอสมมาตร การแยกตัวประกอบ RSA

Abstract

Cryptographic RSA are one of the most popular methods used in Security applications. They are used in authentication, digital signatures. RSA cryptography consists of a private key and public key. This is calculated of two large prime numbers. The RSA security problem issues on the difficulty of factorization. This research, propose a comparison performance of integer factorization between Fermat algorithm and Pollard Rho algorithm. This will be experimenting with a prime random number and that comes up with some custom odd numbers. The experiment results show that Pollard Rho algorithm can factor the modulus very rapidly when experimenting with random numbers, but odd number Fermat algorithm can factor the modulus very rapidly.

Keyword: Asymmetric Cryptography, Factorization, RSA.

1. บทนำ

ในปัจจุบันความก้าวหน้าทางด้านเทคโนโลยีการสื่อสาร การรับส่ง และการจัดเก็บข้อมูลในรูปแบบอิเล็กทรอนิกส์ถือว่ามีความสำคัญเป็นอย่างมาก เมื่อข้อมูลส่งจากผู้ส่งไปยังผู้รับจะต้องมีความปลอดภัยแนวความคิดหลักเกี่ยวกับความปลอดภัยของข้อมูลคือ การรักษาความลับของข้อมูล (Confidentiality) ความถูกต้องสมบูรณ์ของข้อมูล (Integrity) และความพร้อมใช้งาน (Availability) [1] โดยแนวความคิดเหล่านี้ส่วนใหญ่จะใช้วิธีการรักษาความปลอดภัยด้วยกระบวนการเข้ารหัส (Encryption) เป็นการแปลงข้อความหรือข้อมูลอิเล็กทรอนิกส์ในรูปที่อ่านได้ (Plaintext) ให้เป็นอีกรูปแบบหนึ่งที่เปลี่ยนแปลงไปจากเดิมซึ่งไม่สามารถอ่านได้ (Ciphertext) และถ้าหากต้องการแปลงข้อความหรือข้อมูลอิเล็กทรอนิกส์ที่อยู่ในรูปแบบที่อ่านไม่ได้ (Ciphertext) ให้

กลับมาอยู่ในรูปแบบที่สามารถอ่านได้ (Plaintext) จะเรียกว่ากระบวนการถอดรหัส (Decryption) [1][2] การเข้ารหัสและการถอดรหัสไม่เป็นเพียงแค่การรักษาความปลอดภัยของข้อมูลแต่ยังสามารถใช้สำหรับตรวจสอบการยืนยันตัวบุคคลในการเข้าใช้งาน สามารถแบ่งประเภทของการเข้ารหัสได้ 2 ประเภท ได้แก่ การเข้าแบบสมมาตร (Symmetric Cryptography) และการเข้ารหัสแบบอสมมาตร (Asymmetric Cryptography)

สำหรับการเข้ารหัสแบบอสมมาตร (Asymmetric Cryptography) จะมีกุญแจคู่ (Key Pairs) อยู่โดยกุญแจคู่นี้จะถูกสร้างขึ้นมาพร้อมกัน โดยกุญแจตัวที่หนึ่งหรือกุญแจส่วนตัว (Private Key) จะเอาไว้ใช้สำหรับถอดรหัสข้อมูล ส่วนกุญแจอีกตัวหนึ่งจะเอาไว้ใช้สำหรับเข้ารหัสข้อมูล หรือที่เรียกว่ากุญแจสาธารณะ (Public Key) [2-6] อัลกอริทึม RSA เป็นตัวอย่างอัลกอริทึมหนึ่งในการเข้ารหัสแบบอสมมาตร ซึ่งในอัลกอริทึม RSA กุญแจสาธารณะและกุญแจส่วนตัวที่สร้างขึ้นโดยจะมีความเกี่ยวข้องในเชิงฟังก์ชันจำนวนเฉพาะขนาดใหญ่คู่หนึ่ง ทำให้ความปลอดภัยของตัวอัลกอริทึมขึ้นอยู่กับความยากในการแยกตัวประกอบของตัวเลขที่มีขนาดใหญ่ โดยเฉพาะผลคูณของตัวเลขจำนวนเฉพาะขนาดใหญ่

ในงานวิจัยฉบับนี้ได้ทำการเปรียบเทียบประสิทธิภาพการแยกตัวประกอบของตัวเลขจำนวนเต็ม โดยใช้อัลกอริทึม Fermat และ Pollard Rho โดยตัวชี้วัดที่จะนำมาวิเคราะห์ประสิทธิภาพของการทำงานจะประกอบด้วยความสามารถของอัลกอริทึมที่สามารถแยกตัวประกอบของตัวเลขที่ได้จากการสุ่มของโปรแกรมและตัวเลขที่ได้จากการกำหนดซึ่งจะเป็นตัวเลขจำนวนคี่ที่ซ้ำกันจำนวน 5 หลักด้านหลังจากตัวเลขทั้งหมด 8 หลักจำนวน 5 ชุดข้อมูล และตัวเลขจำนวนคี่ที่ซ้ำกัน 4 หลักหน้าและซ้ำกัน 2 หลักถัดไปและซ้ำกัน 2 หลักท้ายสุดอีกจำนวน 4 ชุดข้อมูล งานวิจัยนี้ได้จัดลำดับหัวข้อตามลำดับคือ ส่วนที่ 2 อธิบายถึงทฤษฎีและงานวิจัยที่เกี่ยวข้อง ส่วนที่ 3 วิธีดำเนินการวิจัย ส่วนที่ 4 ผลการดำเนินการวิจัย และส่วนที่ 5 สรุปผลการวิจัย

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 การเข้ารหัสแบบ RSA

การเข้ารหัสแบบ RSA [7] เป็นวิธีการเข้ารหัสแบบหนึ่งของการเข้ารหัสแบบอสมมาตรถูกนำเสนอครั้งแรกในปี 1978 โดย Ron Rivest, Adi Shamir และ Leonard Adleman ที่ MIT โดยที่ RSA นั้นได้มาจากนามสกุลของทั้งสามคน การเข้ารหัสแบบ RSA ถูกนำมาใช้งานอย่างแพร่หลายในการทำธุรกรรมอิเล็กทรอนิกส์ เช่นลายเซ็นดิจิตอล โดยการเข้ารหัสจะต้องมีกุญแจส่วนตัว (Private Key) และกุญแจสาธารณะ (Public Key) ซึ่งสร้างจากตัวเลขที่สุ่มขึ้นมา และนำมาผ่านอัลกอริทึม RSA โดยกุญแจสาธารณะจะเป็นตัวที่สามารถเปิดเผยและใช้ร่วมกันได้ ส่วนกุญแจส่วนตัวจะมีอยู่เฉพาะผู้ที่รับข้อมูลและผู้ส่งข้อมูลเท่านั้น

กระบวนการสร้างกุญแจส่วนตัวและกุญแจสาธารณะขึ้นมานั้น ในตอนแรกสุด ผู้ใช้งานจะต้องเลือกตัวเลขจำนวนเฉพาะแบบสุ่ม p และ q ที่มีค่ามาก ๆ ขึ้นมา 2 ค่า จากนั้นทำการคำนวณหาผลคูณ $N = pq$ จากนั้นผู้ใช้งานจะเลือกตัวเลขจำนวนเต็มแบบสุ่มขึ้นมาอีกหนึ่งค่า เพื่อใช้เป็นกุญแจเข้ารหัสลับ e โดยที่ e และ $(p-1)(q-1)$ ต้องเป็นจำนวนเฉพาะสัมพันธ์ร่วมกัน (Relatively Prime) หลังจากนั้นคำนวณหาค่ากุญแจถอดรหัส d จากสมการ

$$ed = 1 \bmod (p-1)(q-1)$$
$$\text{โดยที่} \quad d = e^{-1} \bmod (p-1)(q-1) \qquad (1)$$

ในการเข้ารหัสข้อความต้นฉบับสามารถใช้สมการได้ดังนี้

$$C = M^e \bmod N \qquad (2)$$

สำหรับการถอดรหัสข้อความสามารถใช้สมการได้ดังนี้

$$M = C^d \bmod N \qquad (3)$$

2.2 อัลกอริทึม Fermat

อัลกอริทึม Fermat [8] เป็นอัลกอริทึมในการแยกตัวประกอบของตัวเลขจำนวนเต็มโดยใช้หลักการที่ว่าตัวเลขจำนวนเต็ม (N) ซึ่ง (N) ควรเป็นตัวเลขจำนวนคี่ แต่กรณีถ้าเป็นจำนวนคู่สามารถแยกตัวประกอบ 2 จำนวนออกไป จากนั้นหาตัวประกอบโดยหาค่า a และ b ที่เป็นจำนวนเต็มน้อยที่สุดที่ทำให้ $a - b \le \sqrt{N} \le a + b$ โดยที่ $a - b = N / (a + b)$ และถ้าแยกตัวประกอบได้ $N = N * 1$ จะได้ว่า N เป็นตัวเลขจำนวนเฉพาะ โดยทำการทดสอบทีละเงื่อนไขและบวกค่า a เพิ่มขึ้นไปทีละหนึ่ง

จนพบคำตอบที่ได้ตัวประกอบเข้าสู่รากของ N มากที่สุด เขียนเป็นรหัสเทียมได้ดังนี้

```
FermatFactor(N): // N ควรเป็นตัวเลขจำนวนคี่
    a ⟵ ceil(sqrt(N))
    b2 ⟵ a * a – N
    while b2 isn't a square:
        a ⟵ a + 1
        b2 ⟵ a * a – N
    endwhile
return a – sqrt(b2) // หรือ a + sqrt(b2)
```

จากคุณสมบัติดังกล่าวของอัลกอริทึม Fermat ประสิทธิภาพในการแยกตัวประกอบกรณีที่แย่ที่สุดคือเมื่อ N เป็นจำนวนเฉพาะ ซึ่งจะทำให้ประสิทธิภาพการทำงานเป็น O(N)

2.3 อัลกอริทึม Pollard Rho

อัลกอริทึม Pollard Rho [9] เป็นอัลกอริทึมในการแยกตัวประกอบของตัวเลขที่มีค่ามาก โดยอาศัยคุณสมบัติของการหารเพื่อให้หาตัวประกอบของตัวเลขจำนวนนั้นได้เร็วโดยถ้าหากต้องการหาตัวเลข 2 ตัว x และ y ที่มีคุณสมบัติหารด้วยตัวเลข p แล้วมีเศษเหลือเท่ากัน หากให้ p เป็นตัวประกอบหนึ่งของ n เมื่อ n คือตัวเลขที่ต้องการหาตัวประกอบ จะได้ว่า $p \leq GCD(x - y, n) \leq n$ เนื่องจาก p หาร x -y และ n ลงตัว เขียนเป็นรหัสเทียมได้ดังนี้

```
x⟵2, y ⟵ 2; d⟵ 1
While d = 1:
    x ⟵ f(x) // f(x) = (x^2 + 1) mod n
    y ⟵ f(f(y))
    d ⟵ GCD(|x -y|, n)
if d = n, return failure.
Else, return d.
```

อัลกอริทึม Pollard Rho จะใช้ฟังก์ชันมอดูโล n สำหรับการสร้างลำดับตัวเลขสุ่มเทียมโดยลำดับทั้งสองจะมีความเร็วในการเลื่อนลำดับต่างกัน 2 เท่า กล่าวคือ ลำดับหนึ่งจะเลื่อนไป 2 ขั้น ในขณะที่อีกลำดับจะเลื่อนไปเพียง 1 ขั้นเท่านั้น หากให้ x และ y แทนตัวเลขในลำดับทั้ง 2 ในขณะนั้น การหาหารร่วมมาก (GCD) ของ $|x - y|$ และ n โดยถ้าหาก GCD มีค่าเท่ากับ n จะทำให้เกิดข้อผิดพลาด เนื่องจาก GCD จะมีค่าเท่ากับ n ได้เมื่อ $x = y$ นั่นเอง แต่ถ้าหาก GCD มีค่าเท่ากับ 1 แสดงว่าหาตัวประกอบยังไม่พบ จะกลับไปเลือก x และ y จากลำดับตัวเลขสุ่มเทียมตัวต่อไป และสุดท้ายถ้าหาก GCD เท่ากับค่าอื่นแสดงว่าค่าที่ได้นั้นคือตัวประกอบของ n นั่นเอง

2.4 งานวิจัยที่เกี่ยวข้อง

กระบวนการแยกตัวประกอบของตัวเลขจำนวนเต็มนั้นไม่สามารถชี้วัดได้ว่าอัลกอริทึมแบบไหนสามารถแยกตัวประกอบของตัวเลขได้ทุกตัว เพราะว่ารูปแบบของตัวเลขนั้นมีรูปแบบของตัวเลขหลายรูปแบบรวมถึงจำนวนหลักของตัวเลขจำนวนเต็ม ซึ่งยังเป็นปัญหาในการแยกตัวประกอบของตัวเลขจำนวนเต็มอีกประการหนึ่งเช่นกัน โดยในปี 2015 มีนักวิจัยได้ทำการทดลองแยกตัวประกอบของตัวเลขจำนวนเต็มด้วยอัลกอริทึม Dixon's อัลกอริทึม Fermat และอัลกอริทึม Pollard Rho โดยตัวเลขจำนวนเต็มที่ใช้ทดสอบได้พัฒนาด้วยโปรแกรมภาษาจาวาในการสุ่มตัวเลขที่เป็นจำนวนเต็มทั่วไปจำนวนสูงสุด 10 หลัก ผลการทดสอบพบว่าอัลกอริทึม Pollard Rho สามารถประมวลผลการแยกตัวประกอบได้เร็วกว่าเมื่อเทียบกับอัลกอริทึมอีกสองตัว [10] และเมื่อจำนวนหลักเพิ่มขึ้นทำให้ตัวเลขมีขนาดใหญ่ตามขึ้นไปด้วยซึ่งอัลกอริทึม Pollard Rho สามารถนำมาพัฒนาโปรแกรมแยกตัวประกอบสำหรับตัวเลขจำนวนเต็มที่มีขนาดใหญ่ประสิทธิภาพได้ดี [11]

3. วิธีดำเนินการวิจัย

งานวิจัยฉบับนี้มีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพการแยกตัวประกอบของตัวเลขจำนวนเต็ม โดยการเปรียบเทียบความเร็วในการประมวลผลการแยกตัวประกอบของอัลกอริทึม Fermat กับอัลกอริทึม Pollard Rho เพื่อทดสอบหาว่าอัลกอริทึมใดเหมาะสมกับรูปแบบตัวเลขชนิดใดที่ใช้ในการแยกตัวประกอบ สำหรับเครื่องมือที่ผู้วิจัยใช้ทดสอบผู้วิจัยเลือกใช้ภาษา JAVAในการพัฒนาโปรแกรมและ JDK Version 1.8 เพื่อใช้ทดสอบซึ่งทำงานบนเครื่องคอมพิวเตอร์ Laptop ที่ติดตั้งบนระบบปฏิบัติการ Ubuntu 16.04 64 bit ความเร็ว CPU Core i3-3227U 1.90 GHz หน่วยความจำ 4 GB โดยใช้ตัวเลขที่เป็นตัวเลขจำนวนเต็มทดสอบกับอัลกอริทึมทั้งสองเพื่อหาระยะเวลาความเร็วในการประมวลผลการแยกตัวประกอบสำเร็จ

ขั้นตอนการทดสอบผู้วิจัยได้กำหนดตัวเลขจำนวนเต็มในรูปแบบทั่วไปโดยใช้การสุ่มจากโปรแกรมที่พัฒนาขึ้นขนาด

จำนวนสูงสุด 10 หลักโดยทำการทดสอบจำนวน 10 ครั้ง และกำหนดรูปแบบตัวเลขจำนวนเต็มคี่โดยมีรูปแบบการกำหนดตัวเลข [12] ดังตารางที่ 1

ตารางที่ 1: แสดงรูปแบบตัวเลขจำนวนเต็มคี่ที่ใช้ทดสอบ

ลำดับที่	รูปแบบตัวเลข
1	13311111
2	17733333
3	19155555
4	13377777
5	33399999
6	11117733
7	33331155
8	55553355
9	77773311

ในการเปรียบเทียบระยะเวลาด้านความเร็วในการประมวลผลแยกตัวประกอบผู้วิจัยได้กำหนดรูปแบบคำสั่งการแสดงเวลาการประมวลผลในส่วนของฟังก์ชันหลักดังภาพที่ 1

```
long startTime = System.currentTimeMillis();
... your program ....
long endTime = System.currentTimeMillis();
long totalTime = endTime - startTime;
System.out.println("Total time=" +totalTime);
```

ภาพที่ 1: คำสั่งแสดงเวลาการประมวลผล

4. ผลการดำเนินการวิจัย

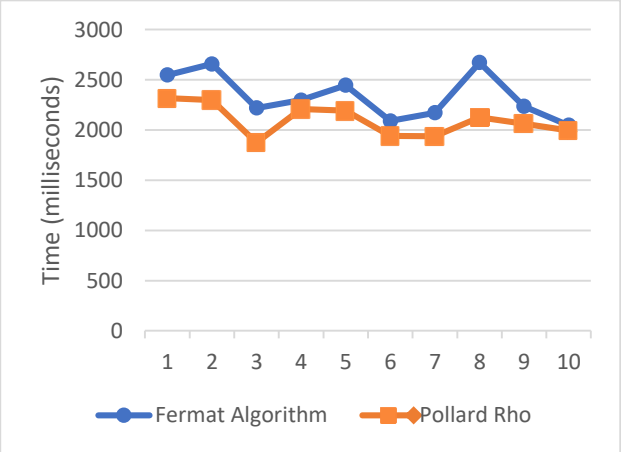
งานวิจัยนี้ได้ทำการเปรียบเทียบประสิทธิภาพการแยกตัวประกอบของตัวเลขจำนวนเต็มโดยใช้อัลกอริทึมของ Fermat กับอัลกอริทึมของ Pollard Rho ผู้วิจัยทำการทดสอบโดยแบ่งข้อมูลออกเป็น 2 ส่วนคือ ตัวเลขที่ได้จากโปรแกรมทำการสุ่มและตัวเลขจำนวนคี่ที่ได้จากการกำหนดขึ้นมาตามรูปแบบที่กำหนด จากนั้นทำการเปรียบเทียบประสิทธิภาพด้านความเร็วในการแยกตัวประกอบของตัวเลขจำนวนเต็มของแต่ละอัลกอริทึม

4.1 ผลการทดสอบด้วยตัวเลขที่สุ่มโดยโปรแกรม

ทำการทดสอบโดยการใช้โปรแกรมสุ่มเลือกตัวเลขที่เป็นตัวเลขจำนวนเต็มขนาดจำนวน 10 หลักโดยทำการทดสอบทั้งหมด 10 รอบ จากนั้นทำการหาเวลาเฉลี่ยรวมทั้งหมด โดยผลที่ได้จากการทดสอบแสดงดังตารางที่ 2 และเมื่อทำการเปรียบเทียบในรูปแบบกราฟสามารถแสดงได้ดังภาพที่ 2

ตารางที่ 2: แสดงผลการเปรียบเทียบตัวเลขที่สุ่ม

ครั้งที่	เวลาในการประมวลผล (มิลลิวินาที)	
	Fermat Algorithm	Pollard Rho
1	2548	2316
2	2658	2297
3	2219	1875
4	2297	2209
5	2446	2192
6	2091	1942
7	2172	1938
8	2672	2125
9	2235	2063
10	2047	1994



ภาพที่ 2 : แสดงการเปรียบเทียบตัวเลขที่ได้จากการสุ่ม

จากตารางและภาพที่แสดงพบว่ากรณีที่เป็นตัวเลขจำนวนเต็มที่ได้จากการสุ่มของโปรแกรมประสิทธิภาพการแยกตัวประกอบของอัลกอริทึม Pollard Rho สามารถประมวลผลได้เร็วกว่าอัลกอริทึม Fermat

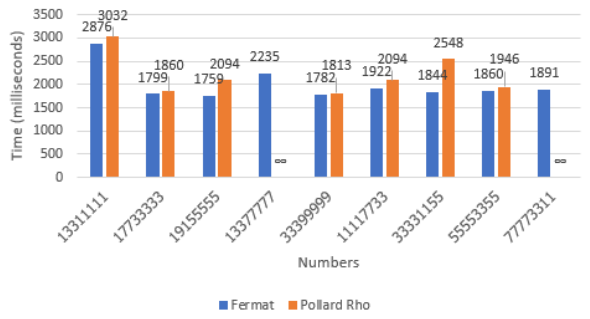
4.2 ผลการทดสอบด้วยตัวเลขที่กำหนดเอง

ทำการทดสอบโดยกำหนดรูปแบบตัวเลขจำนวนคี่ที่ซ้ำกันจำนวน 5 หลักด้านหลังจากตัวเลขทั้งหมด 8 หลักจำนวน 5 ชุดข้อมูล และตัวเลขจำนวนคี่ที่ซ้ำกัน 4 หลักหน้าและซ้ำกัน 2 หลักถัดไปและซ้ำกัน 2 หลักท้ายสุดอีกจำนวน 4 ชุดข้อมูล ผลจากการทดสอบแสดงได้ดังตารางที่ 3 และเมื่อทำการเปรียบเทียบในรูปแบบกราฟสามารถแสดงผลได้ดังภาพที่ 3

ตารางที่ 3: แสดงการเปรียบเทียบตัวเลขที่กำหนดเอง

ตัวเลขที่ทดสอบ	เวลาในการประมวลผล (มิลลิวินาที)	
	Fermat Algorithm	Pollard Rho
13311111	2876	3032
17733333	1799	1860
19155555	1759	2094
13377777	2235	*
33399999	1782	1813
11117733	1922	2094
33331155	1844	2548
55553355	1860	1946
77773311	1891	*

* ไม่สามารถแยกตัวประกอบได้



ภาพที่ 3 : แสดงการเปรียบเทียบตัวเลขจำนวนคี่ที่กำหนดเอง

จากตารางกรณีที่กำหนดตัวเลขที่เป็นจำนวนคี่ตามรูปแบบที่กำหนด พบว่า อัลกอริทึม Fermat จะมีประสิทธิภาพด้านความเร็วในการแยกตัวประกอบมากกว่าอัลกอริทึม Pollard Rho และมีตัวเลขจำนวนคี่ตามรูปแบบที่กำหนดบางตัวที่อัลกอริทึม Pollard Rho ไม่สามารถแยกตัวประกอบได้

5. สรุปผลการดำเนินงาน

งานวิจัยนี้ได้ทำการเปรียบเทียบประสิทธิภาพการแยกตัวประกอบของตัวเลขจำนวนเต็มโดยเปรียบเทียบอัลกอริทึม Fermat กับอัลกอริทึม Pollard Rho ผลจากการทดลองในสถานการณ์ที่เป็นตัวเลขจำนวนเต็มรูปแบบทั่วไปหรือตัวเลขที่ได้จากการสุ่มของโปรแกรม พบว่า อัลกอริทึม Pollard Rho จะมีประสิทธิภาพความเร็วในการแยกตัวประกอบมากกว่าอัลกอริทึม Fermat ในขณะที่สถานการณ์ที่กำหนดรูปแบบตัวเลขจำนวนคี่ที่ซ้ำกัน พบว่าอัลกอริทึม Fermat สามารถประมวลผลการแยกตัวประกอบได้เร็วกว่าอัลกอริทึม Pollard Rho ซึ่งแสดงให้เห็นว่ากรณีที่ตัวเลขเป็นจำนวนเต็มคี่นั้นอัลกอริทึม Fermat จะมีความเหมาะสมในการแยกตัวประกอบมากที่สุดซึ่งเป็นไปตามทฤษฎีและสอดคล้องกับงานวิจัยที่เกี่ยวข้อง อย่างไรก็ตามในอนาคตสำหรับการทดสอบควรมีการเพิ่มขนาดของตัวเลขจำนวนเต็มที่นำมาทดลองอีกทั้งยังต้องทดสอบด้วยอัลกอริทึมในการแยกตัวประกอบในรูปแบบอื่นอีกด้วย

เอกสารอ้างอิง

[1] M. Abadi, "Security Protocols and Specifications", in Foundations of Software Science and Computation Structures, Springer-Verlag Berlin Germany. Vol. 1578 pp.1-13 (1999).

[2] W. Stallings, "Cryptography and Network Security 4 Ed". Prentice Hall. pp. 58-309(2005).

[3] D. M. Bressoud, "Factorizations and primality testing", in: Springer-Verlag, pp. 21-22, New York , 1989.

[4] P. Lehman, R. Sherman, "Factoring Large Integers", in: Mathematics of Computation, pp. 637-646, 1974.

[5] J. McKee, "Speeding Fermat's factoring method", in: Mathematics of Computation, pp. 1729-1737, 1999.

[6] O. Oystein, "Euler's Factorization Method", in: Number Theory and Its History, pp. 59-64, 1976.

[7] R.L.Rivest, A.Shamir and L.Adleman "A method for obtaining digital signatures and Public-Key Cryptosystems", Communications of the ACM 21 (1978), 120-126.

[8] P. Lehman, R. Sherman, "Factoring Large Integers", in: Mathematics of Computation, vol. 28, pp. 637-646, 1974.

[9] J. M. Pollard, "A Monte Carlo method for factorization", in: BIT Numerical Mathematics, Vol. 15, pp. 331–334, 1975.

[10] D. Miller, V. Huber and M. Hagaman, C. Morrision, "Comparing the heuristic running times and time complexities of integer factorization algorithms", in Supercompting Challenge, Team 14, 2015.

[11] C. L. Duta, L. Gheorghe and N. Tapus. "Framework for Evaluation and Comparison of Integer Factorization Algorithms". SAI Computing Conference. 2016.

[12] A. Murthy, "Sequences A061507", in The On-line Encyclopedia of Integer Sequences.

ระบบไอโอทีฮับสำหรับเครือข่ายเนมดาต้า
An IoT Hub System for Named Data Networking

ธงชัย เจือจันทร์ (Thongchai Chuachan)[1] ภัทรภัฏ ไหวดี (Pattarapat Waidee)[2]
และศิริมงคล มีศิริ (Sirimongkon Meesiri)[3]
สาขาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏสุรินทร์
[1]Thongchai.c@srru.ac.th , [2]January20@windowslive.com, [3]Sirimongkon@srru.ac.th

บทคัดย่อ

อินเทอร์เน็ตในทุกสิ่งเป็นหนึ่งในเครือข่ายที่ถูกใช้กับอุปกรณ์ที่มีประสิทธิภาพหลากหลาย และต้องอาศัยโบรคเกอร์ในการเชื่อมต่อระหว่าง อุปกรณ์ไอโอทีกับผู้ใช้ผ่านอินเทอร์เน็ตถึงแม้จะมีการเชื่อมต่อทางกายภาพในระดับท้องถิ่น งานวิจัยนี้จึงปรับปรุงโดยใช้เครือข่ายเอ็นดีเอ็นเป็นเครื่องมือ โดยการพัฒนาระบบไอโอทีฮับช่วยให้การเชื่อมต่ออุปกรณ์ไอโอทีได้ทั้งผ่านอินเทอร์เน็ต และเครือข่ายท้องถิ่นแบบอัตโนมัติ แล้วทดสอบด้วยระบบเทสเบด ผลการทดลองแสดงถึงประสิทธิภาพและและประสิทธิผลของระบบเครือข่ายอินเทอร์เน็ตในทุกสิ่งดีขึ้น

คำสำคัญ: เครือข่ายเอ็นดีเอ็น เนมดาต้า ไอโอที เอ็มคิวทีที โบรคเกอร์ ไอโอทีฮับของเครือข่ายเอ็นดีเอ็น

Abstract

Internet of Things (IoT) is crucial in modern networking, and has been used by very heterogeneous devices and systems. Furthermore, IoT is limited to recognize and to swap between local networks and IoT broker. In this paper, we have enhanced an IoT accessibility by prototyping an NDN IoT-Hub system. Our NDN IoT-Hub system has been tested in an NDN Testbed. Experimental results have shown that the NDN IoT-Hub system improves IoT networks in several aspects.

Keyword: NDN Networks, Named Data Networking, IoT, MQTT Broker, NDN IoT-Hub.

1. บทนำ

Internet of Thing (IoT) [1] เป็นแนวทางการออกแบบการเชื่อมต่อสื่อสารสำหรับอุปกรณ์อินเตอร์เน็ตสำหรับทุกสิ่งอย่าง ซึ่งปัจจุบันยังคงทำงานอยู่บนสถาปัตยกรรมเครือข่าย Transmission Control Protocol / Internet Protocol (TCP/IP) [2], [3] และมีปัญหาหลายด้านดังที่ Zhang และคณะ [4] ได้นำเสนอไว้ เช่น ความมั่นคง การส่งแพ็กเก็ตทางเดียว (Stateless) การจัดการความคับคั่ง (Congestion) การเร้าข้อมูลจากระยะไกล (Remote Area) และการกระจายข้อมูล เป็นต้น

หนึ่งในโพรโทคอลที่ออกแบบมาใช้กับ IoT คือ MQTT [5] ที่ได้ออกแบบให้แพ็กเก็ตมีขนาดเล็ก เชื่อมต่อระหว่างอินเทอร์เน็ตและอุปกรณ์ IoT ด้วยวิธี Publish และ Subscribe โดย MQTT ยังคงใช้ระบบ Centralize Server ทำหน้าที่เป็น Broker ส่งต่อคำสั่งระหว่างผู้ใช้และอุปกรณ์ IoT และยังคงมีปัญหาหลายด้าน เช่น ต้องอาศัย Broker แม้การส่งคำสั่งไปยังอุปกรณ์จะอยู่ในเครือข่ายท้องถิ่น (Local Area Network : LAN) ในบางสถานการณ์ระบบ Domain Name System (DNS) ให้บริการได้ยากจากปัญหาการใช้เครือข่ายไร้รูปแบบ (Infrastructure Less) การส่งแพ็กเกตแบบ Broadcast ทำให้อัตราสิ้นเปลืองแบตเตอร์รีมีสูงขึ้น และถูกจู่โจมได้ง่าย เป็นต้น ดังนั้นในเครือข่าย IoT จึงยังมีหลายปัจจัยที่เป็นเรื่องท้าทายสำหรับการออกแบบเครือข่ายสำหรับอุปกรณ์ IoT

Named Data Networking (NDN) [6], [7] เป็นหนึ่งในสถาปัตยกรรมเครือข่าย ที่มีแนวคิดที่จะออกแบบและพัฒนาการเชื่อมต่อเครือข่ายในระบบการสื่อสารยุคใหม่ มุ่งเน้นให้ผู้ใช้สามารถรับข้อมูลได้อย่างรวดเร็ว (Best Effort) การเข้าถึงแหล่งข้อมูลจะใช้ชื่อของวัตถุ (Named Data Object: NDO) เป็นสิ่งบ่งชี้ข้อมูลของผู้ให้บริการ (Producer) เข้าถึงได้ง่ายจาก

การใช้ NDO ในงานวิจัยนี้จึงออกแบบโพรโทคอลสำหรับการเชื่อมต่อของอุปกรณ์ IoT โดยใช้เครือข่าย NDN ออกแบบรูปแบบ NDO ให้เหมาะกับการใช้กับอุปกรณ์ IoT และพัฒนาระบบ NDN IoT-Hub สำหรับการส่งและรับคำสั่งของอุปกรณ์ IoT จากเครือข่าย WAN แล้วเพิ่มกลไกการสลับใช้เครือข่าย LAN อัตโนมัติหากผู้ใช้กับอุปกรณ์ IoT อยู่ใน LAN เดียวกันได้

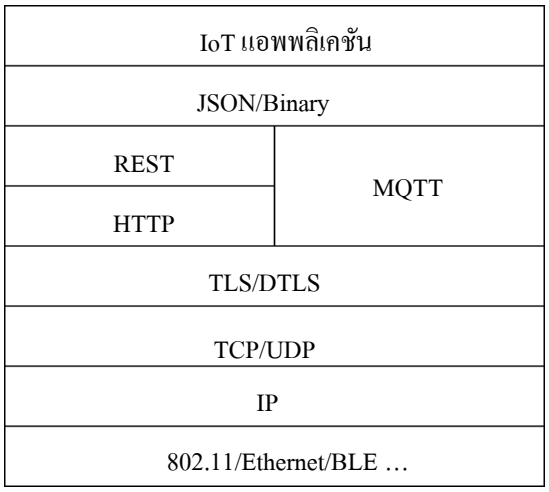
ต้นแบบของงานวิจัยนี้พัฒนาต่อยอดจากซอฟต์แวร์เร้าเตอร์ Named Forwarding Daemon (NFD) [8] ของเครือข่าย NDN และทดสอบบนระบบ NDN Bone เทียบกับการใช้ Broker ของโพรโทคอล MQTT และการเร้าแพ็กเก็ตเมื่ออุปกรณ์ IoT และผู้ใช้อยู่ใน LAN เดียวกัน ผลจากการทดสอบระบบพบว่างานวิจัยนี้ช่วยให้การเข้าถึงอุปกรณ์ IoT มีประสิทธิภาพการเชื่อมต่อที่สูงขึ้น และลดการใช้ทรัพยากรณ์ของอุปกร IoT ได้

2. วรรณกรรมและงานวิจัยที่เกี่ยวข้อง

2.1 อินเทอร์เน็ตในทุกสิ่ง

Internet of Thing (IoT) [1] เป็นวิธีการเชื่อมต่อสื่อสารสำหรับอุปกรณ์เข้าสู่เครือข่ายซึ่งปัจจุบันมีการออกแบบโพรโทคอลหลายแบบดังแสดงในภาพที่ 1 ที่มี IoT แอพพลิเคชันทำงานอยู่ในชั้นบนสุดของระบบเครือข่าย และปัจจุบันนิยมใช้ JSON และ Binary เป็นภาษาที่แอพพลิเคชันใช้ ในขณะที่ชั้นรองลงมาจะเป็นชั้นของ MQTT และ HTTP ทำงานร่วมกับ REST ซึ่งจะคอยเป็นส่วนเชื่อมระหว่างแอพพลิเคชันและชั้น ทรานสปอร์ต (Transport Layer) ส่วนชั้นที่อยู่ต่ำกว่า MQTT และ HTTP เป็นส่วนของการเข้ารหัสข้อมูลอย่าง Transport Layer Security (TLS) และ Datagram Transport Layer Security (DTLS) โดยในชั้นนี้จะสามารถถูกควบคุมให้มีการเข้ารหัสหรือไม่เข้ารหัสได้จากการคอนฟิกของผู้พัฒนาแอพพลิเคชันและผู้คอนฟิกเครือข่าย จากนั้นจะเป็นโครงสร้างของเครือข่าย TCP/IP ปกติ โดยประกอบด้วยโพรโทคอล TCP และ UDP เป็นส่วนของ ทรานสปอร์ต และ IP ใช้ในการขนส่งระหว่างต้นทางและปลายทาง จากนั้นจะเป็นโพรโทคอลที่ทำงานในเครือข่าย LAN เช่น 802.11 Ethernet และ BLE เป็นต้น

จากภาพที่ 1 จะเห็นว่า IoT มีชั้นของเครือข่าย (Network Stack) จึงทำให้การขนส่งข้อมูลอาจเกิดความล่าช้าจากปัจจัยหลายประการ เช่น การประมวลผล ขนส่ง และความหนาแน่นของแพ็กเก็ตในเครือข่าย เป็นต้น นอกจากนี้ IoT ที่อาศัยการขนส่งข้อมูลด้วยโพรโทคอล TCP/IP ยังไม่อาจสลับการรับและส่งคำสั่งบนเครือข่าย WAN และ LAN โดยอัตโนมัติ ถึงแม้ผู้ใช้จะอยู่ใน LAN เดียวกันกับอุปกรณ์ IoT ยากต่อการจัดการอาจเกิดปัญหาการถูกจู่โจม และมีความล่าช้าได้ งานวิจัยนี้จึงออกแบบระบบ IoT โดยอาศัยโพรโทคอลของ NDN มาปรับปรุงให้ระบบเครือข่ายของอุปกรณ์ IoT ทำงานได้อย่างมีประสิทธิภาพมากขึ้น



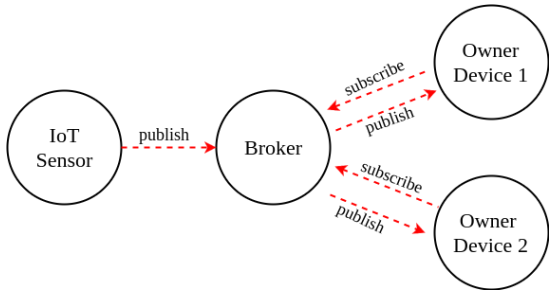
ภาพที่ 1: สถาปัตยกรรมเครือข่ายของ IoT

2.2 MQTT

Message Queue Telemetry Transport (MQTT) [5] เป็นโพรโทคอลที่ทำหน้าที่แบบ Publish และ Subscribe ขนาดเล็กระหว่างสองปลายทาง โดยออกแบบด้วยแนวคิดที่จะจัดการข้อมูลปริมาณสูงและใช้แบนด์วิดท์ต่ำ ทำงานได้บนโพรโทคอล TCP/IP ซึ่ง MQTT มีขั้นสถาปัตยกรรมดังภาพที่ 2 ประกอบด้วยอุปกรณ์ IoT ที่เชื่อมต่อกับ Broker และเมื่ออุปกรณ์ เช่น Sensor แพ็กเก็ตจะถูก Publish ไปยัง Broker เครื่องลูกข่ายที่ Subscribe ไปที่ Broker ก็จะได้รับแพ็กเพ็กเก็ตผ่านทาง Broker แทนการเชื่อมต่อไปยัง Sensor โดยตรง ดังนั้นอุปกรณ์อย่าง Owner Device 1 และ 2 ต้องเชื่อมต่อกับ Broker ตลอดเวลาเพื่อรับข้อมูล

สำหรับวิธีการระบุอุปกรณ์ MQTT ใช้การกำหนดชื่อเป็นลำดับชั้น (Hierarchy) ซึ่งจะช่วยให้ Broker สามารถแยกความเป็นเจ้าของ ของอุปกรณ์ IoT ได้ แต่ MQTT ต้องอาศัย Broker สูง อาจเกิดความล้มเหลว ณ จุดเดียว (Single Point of Failure) และ MQTT อาศัยเพียง IP ในการเชื่อมต่อ ซึ่งมีปัญหา

หลายด้าน เช่น การใช้กับเครือข่ายที่ไม่มีโครงสร้าง (Infrastructure Less) ได้ยาก เกิดดีเลย์ ทำงานแบบ Stateless และการปรับใช้จริงทำได้ยาก



ภาพที่ 2: รูปแบบการเชื่อมต่อของการใช้โพรโทคอล MQTT

2.3 Named Data Networking

Named Data Networking (NDN) [6], [7] เป็นหนึ่งในสถาปัตยกรรมเครือข่ายที่เน้นด้านการกระจายเนื้อความ (Content Distribution) โดยการออกแบบชื่อของ Content ให้เป็นแบบลำดับชั้นและมีคำนำหน้า (Prefix) เช่น /ndn/th/ac/srru แล้วเพิ่มลำดับชั้นเพื่อระบุชื่อไฟล์ /ndn/th/ac/srru/comsci/index.html/1 เป็นต้น การเร้าแพ็กเก็ตในเครือข่าย NDN จึงใช้ Prefix เก็บไว้ในตาราง Forwarding Information Base (FIB) เพื่อใช้เป็นเครื่องมือในการค้นหาเส้น นอกจากนี้ NDN ยังสามารถสำรอง (Cache) แพ็กเก็ตไว้ในอุปกรณ์เร้าเตอร์ของ NDN เพื่อใช้ซ้ำได้

การใช้ Prefix จึงเป็นส่วนสำคัญที่จะช่วยให้การจัดการอุปกรณ์ IoT ทำได้ง่าย เพราะแต่ละอุปกรณ์ IoT จะมี Prefix เป็นของตนเอง และสามารถเร้าแพ็กเก็ตจากเครือข่าย WAN แบบ Multipath สามารถรับรู้การเชื่อมต่อแบบ LAN ได้อย่างอัตโนมัติ แต่ปัจจุบันเครือข่าย NDN ยังคงมีปัญหาการจัดการตาราง FIB สำหรับอุปกรณ์ IoT แบบอัตโนมัติ งานวิจัยนี้จึงต้องการพัฒนา NDN IoT-Hub สำหรับการจัดการ Prefix เพื่อเชื่อมต่อกับอุปกรณ์ IoT ในเครือข่าย WAN แบบอัตโนมัติได้

2.5 งานวิจัยก่อนหน้า

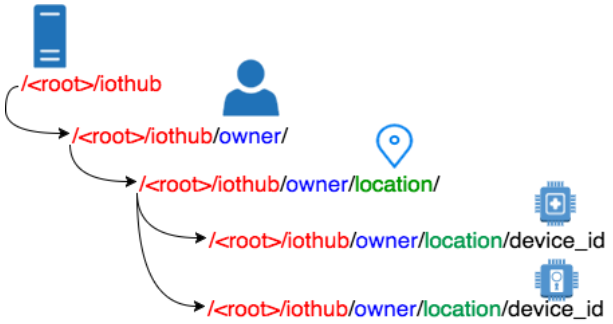
ได้มีความพยายามออกแบบและพัฒนาระบบเครือข่ายสำหรับอุปกรณ์ IoT ไว้หลายด้าน [9] หนึ่งในแนวทางการพัฒนาที่สำคัญ คือ ทำให้ IoT ทำงานได้บนสถาปัตยกรรมเครือข่าย TCP/IP เดิม ดังเช่น [10], [11] ที่ปรับปรุงรูปแบบการเชื่อมต่อของเกตเวย์ (Gateway) ให้มีความสามารถใจการติดต่อกับอุปกรณ์ IoT ได้อย่างอัตโนมัติ และการพัฒนาเกตย์เวย์ให้สนับสนุนโพรโทคอลของ IoT ที่หลากหลาย แต่ระบบเหล่านี้ก็ยังคงต้องอาศัย Broker และยังคงขาดกลไกการแยกแยะวิธีเร้าแพ็กเก็ตสำหรับ IoT ในเครือข่าย LAN และ WAN แบบอัตโนมัติได้

ในงานวิจัยบางส่วนออกแบบวิธีการกำหนดชื่อที่เป็นเอกลักษณ์ (ID) และการควบคุมการเข้าถึงข้อมูลของอุปกรณ์ IoT ดังที่ Fremantle และคณะ [12] ได้นำเสนอไว้ ซึ่งมีความยุ่งยากซับซ้อนต่อวิธีการเข้ารหัสและจัดการ ID ของอุปกรณ์ IoT จำนวนมหาศาล ในขณะที่การนำสถาปัตยกรรมเครือข่าย NDN มาใช้จะช่วยลดความยุ่งยากในการจัดการ ID ของอุปกรณ์ และสามารถนำชื่อ ID ของอุปกรณ์มาสร้างเป็น Prefix และเร้าแพ็กเก็ตได้ งานวิจัยนี้จึงออกแบบแนวทางการสร้าง Prefix และนำไปใช้เป็น ID แล้วนำไปสู่การเร้าแพ็กเก็ต ลดความซับซ้อน เร้าแพ็กเก็ตได้ง่าย และรับรู้รูปแบบการเชื่อมต่อเครือข่ายแบบ LAN และ WAN ได้อย่างอัตโนมัติ

3. การออกแบบและพัฒนาระบบระบบ

3.2 การออกแบบ Prefix สำหรับอุปกรณ์ IoT

การออกแบบ Prefix สำหรับอุปกรณ์ IoT ของงานวิจัยนี้จะเพิ่มส่วนขยายดังแสดงในภาพที่ 3 โดย IoT Hub Server (IoT-HS) จะเป็น Prefix <root> สำหรับจัดการอุปกรณ์ IoT เช่น /ndn/th/ac/srru/iothub จากนั้นระบบจะตามด้วยเจ้าของ (Owner) อุปกรณ์ ซึ่งแต่ละ Owner อาจมีอุปกรณ์อยู่หลายตำแหน่ง จึงต้องระบุตำแหน่ง (Location) และใช้กฎควบคุมการเพิ่มตำแหน่งของผู้ใช้ IoT และสุดท้ายจะเป็นชื่อของอุปกรณ์ IoT ที่เชื่อมเข้าสู่ระบบ IoT-HS



ภาพที่ 3: การกำหนดชื่อ Prefix สำหรับอุปกรณ์ IoT

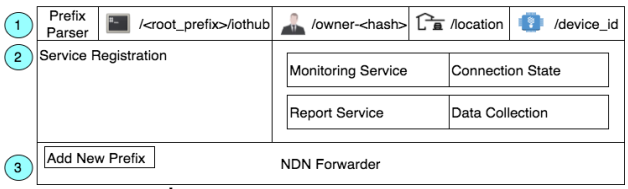
3.1 การออกแบบระบบ NDN IoT-Hub

การเพิ่ม Prefix เข้าสู่ระบบ NDN IoT-Hub ใหม่จากงานวิจัยนี้ ใช้ 3 ขั้นตอนดังแสดงในภาพที่ 4 ที่ประกอบด้วย Prefix Parser, Service Registration และการเพิ่ม Prefix ให้กับ NDN Forwarder ดังนี้

1. Prefix Parser (*PP*) ใช้สำหรับการตรวจสอบรูปแบบของ Prefix ที่มีการร้องขอ ซึ่ง Prefix จะต้องมีส่วนประกอบ 4 ส่วน คือ 1) Root Prefix เช่น /ndn/th/ac/srru และตามด้วย iothub 2) เจ้าของอุปกรณ์ IoT ที่จะใช้งานระบบ IoT-HS ซึ่งจะตามด้วยการ Hash ข้อมูลด้วยอัลกอริธึม SHA1 เพื่อป้องกันการชนกัน (Collision) ของชื่อ Prefix 3) ตำแหน่ง (Location) ของอุปกรณ์ IoT ใช้สำหรับระบุตำแหน่ง เช่น โรงรถ เป็นต้น และ 4) Device ID ใช้สำหรับบ่งชี้ตัวอุปกรณ์ เช่น หลอดไฟ ประตูโรงรถ เป็นต้น

2. Service Registration (*SR*) เป็นส่วนของการตรวจสอบในขั้นตอน Monitoring Service และ Report Service ของอุปกรณ์หลังจากขั้นตอน *PP* โดยระบบจะคอยตรวจสอบสถานะอุปกรณ์และใช้ความสามารถของ NDN สำรองข้อมูลไว้ในช่วงเวลา T_1 ถึง T_2 ตามแต่ละอุปกรณ์ IoT ดังนั้นข้อมูลดังกล่าวจึงถูกกระจายและถูกใช้ซ้ำได้ และลดภาระการร้องขอข้อมูลไปยังอุปกรณ์ IoT ได้

3. New Prefix (*NP*) เป็นส่วนเชื่อมต่อระหว่างกลไก *PP* และ *SR* ให้เข้าสู่ระบบการส่งต่อ (Forward) แพ็กเก็ต ในเครือข่าย NDN



ภาพที่ 4: โครงสร้างของระบบ NDN IoT-Hub

3.2 การพัฒนาระบบ

การพัฒนาระบบแบ่งออกเป็น 3 ส่วน คือ NDN IoT-Hub (NDN-IH), IoT Remote Controller (IoT-RC) และ IoT-Controller (IoT-Con) โดยทั้ง 3 ส่วนจะใช้โปรแกรม Named Forwarding Daemon (NFD) เป็นเครื่องมือการส่งต่อแพ็กเก็ต และมีการทำงานร่วมกันดังต่อไปนี้

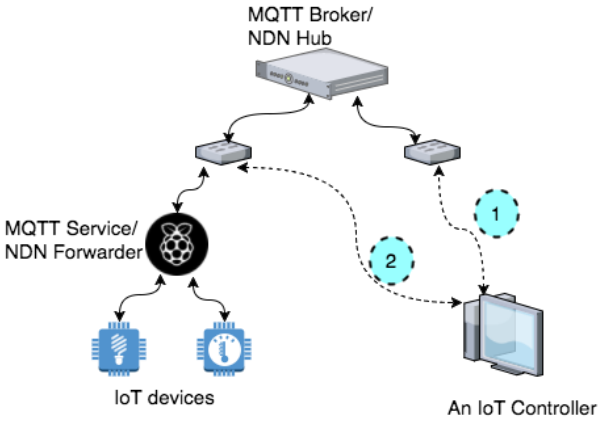
1. ส่วนของ NDN-IH รอรับการเชื่อมต่อเข้าสู่เครือข่าย NDN ของอุปกรณ์ IoT โดยใช้ไลบรารี NDN-JS คอยตรวจสอบ Prefix ที่มีค่าเริ่มต้นด้วย /<root_prefix>/iothub ก่อนที่จะเข้าสู่กลไก *PP*, *SR* และ *NP* ซึ่ง NDN-IH จะสนับสนุนการเชื่อมต่อเครือข่าย NDN ที่วางซ้อน (Overlay) อยู่บนโพรโทคอล TCP และ UDP เนื่องจาก NDN-IH ยังคงต้องอาศัยโพรโทคอล TCP/IP ซึ่งถูกใช้เป็นมาตรฐานการสื่อสารในปัจจุบัน

2. IoT-RC เป็นหน้าต่างโปรแกรมที่จะเชื่อมเข้าสู่เครือข่าย NDN ที่สนับสนุนการสื่อสารกับ NDN-IH และ IoT-Con ได้โดยอัตโนมัติ ซึ่งเป็นการดึงคุณสมบัติของเครือข่าย NDN ทำงานได้กับเครือข่าย Ethernet และ TCP/IP โดย IoT-RC พัฒนาโดยใช้ Node.js และเชื่อมต่อกับ NFD ด้วยไลบรารี NDN-JS

3. IoT-Con ใช้ Raspbery-PI 3 Model B ติดตั้งโปรแกรม NFD เพื่อใช้เครือข่าย NDN และใช้พอร์ต GPIO เป็นตัวควบคุมอุปกรณ์อิเล็กทรอนิกส์ โดยในงานวิจัยนี้ทำการควบคุมหลอด Light-emitting diode (LED) จำนวน 5 หลอด

3.3 การทดลองและประเมินผล

การทดลองของงานวิจัยนี้ใช้วิธีทดสอบระบบบนเครือข่าย Testbed ดังแสดงในภาพที่ 5 โดยได้ทำการประเมินโดยเทียบประสิทธิภาพการใช้เครือข่าย NDN สำหรับระบบ NDN IoT Hub และการใช้โพรโทคอล MQTT ที่มีโครงสร้างการเชื่อมต่อเครือข่ายแบบเดียวกัน แล้วประเมินค่า Latency, Load และ Path Selection โดยในงานวิจัยนี้ได้แยกทดลองสองรอบ รอบละ 30 ครั้งแบ่งเป็น 2 กรณี คือ 1) กรณีที่เชื่อมต่อแบบ WAN ในเส้นหมายเลข 1 ซึ่งทั้งระบบ NDN IoT Hub และ MQTT ต้องอาศัยศูนย์กลางในการควบคุม และ 2) การเชื่อมต่อแบบ LAN ตามหมายเลข 2 ซึ่งการทดลองแต่ละครั้งมีการคอนฟิกค่าเบื้องต้น คือ IP, DNS และ Gateway สำหรับการใช้ MQTT และไม่มีการกำหนดค่าเบื้องต้นทางเครือข่าย เพราะ NDN มีคุณสมบัติการเชื่อมต่อโดยใช้ชื่อ Prefix ของอุปกรณ์แล้ว



ภาพที่ 5: แผนผังการทดลอง

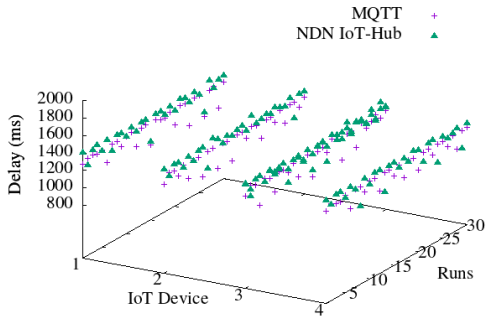
ตารางที่ 1: การกำหนดค่าพารามิเตอร์ในการทดลอง

Parameters	Settings
Runs	30
Protocols	NDN, MQTT
Strategy	NDN, Broadcast Strategy
โพรโทคอลที่เชื่อมกับ NDN-IH และ MQTT Broker	TCP
Connection	LAN, WAN
Link Speed	100 Mb/s

ตารางที่ 1 แสดงค่าพารามิเตอร์ที่ถูกกำหนดสำหรับการทดลองในงานวิจัยนี้ โดยการทดลองจะทำซ้ำทั้งหมด 30 ครั้ง เพื่อใช้ประเมินความเชื่อมั่นทางสถิติที่ 95% ซึ่งการทดลองนี้ใช้โพรโทคอลสองประเภท คือ NDN และ MQTT โดย NDN จะมีส่วนขยาย IoT-HS จากงานวิจัยนี้ แล้วทำการเปรียบเทียบกับโพรโทคอลสำหรับ IoT ที่ถูกใช้อย่างกว้างขวางอย่าง MQTT แต่ยังคงอาศัยเพียงสถาปัตยกรรมเครือข่าย TCP/IP อยู่ และทั้งสองโพรโทคอลจะเชื่อมต่อกับส่วนกลาง นั่นคือ MQTT Broker และ NDN IoT-Hub ซึ่งจะเชื่อมต่อโดยใช้โพรโทคอล TCP แล้วสร้างการเชื่อมต่อทั้งแบบ LAN และการจำลองเครือข่าย WAN สำหรับงานวิจัย เพื่อสังเกตพฤติกรรมการสื่อสารของทั้งสองโพรโทคอลสำหรับ IoT

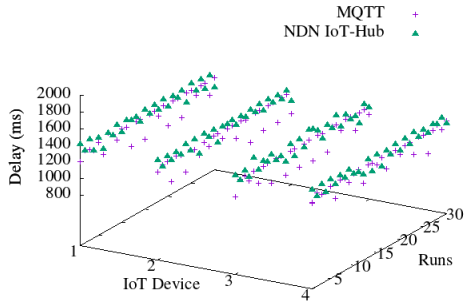
4. ผลการทดลอง

4.1 ผลการประเมินดีเลย์



ภาพที่ 6: ดีเลย์ระหว่าง MQTT และ NDN IoT-Hub ที่เชื่อมต่อบนเครือข่าย LAN

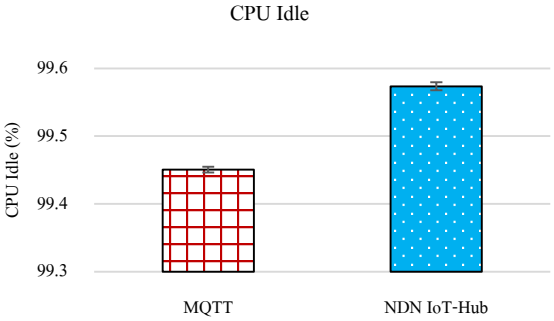
การตรวจสอบค่าดีเลย์เมื่อมีการเชื่อมต่อระหว่างผู้ใช้และอุปกรณ์ IoT บนเครือข่าย WAN แบบจำลองดังแผนผังการทดลองในภาพที่ 5 เส้นหมายเลข 1 จะเห็นว่าการใช้โพรโทคอล MQTT และการใช้ระบบ NDN IoT-Hub มีความแตกต่างกันเพียงเล็กน้อยดังแสดงในภาพที่ 6 ซึ่งมีค่าดีเลย์อยู่ระหว่าง 1257.4 ± 11.7 และ 1331.8±11.0 ตามลำดับ โดยการค่าดีเลย์ที่สูงขึ้นของการใช้ระบบ NDN IoT-Hub มีสูงกว่าเนื่องจาก NDN ออกแบบให้มีการเข้ารหัส จึงทำให้แพ็กเก็ตมีขนาดใหญ่กว่าและเป็นที่มาของค่าดีเลย์ที่สูงขึ้นเล็กน้อย



ภาพที่ 7: ดีเลย์ระหว่าง MQTT และ NDN IoT-Hub ที่เชื่อมต่อบนเครือข่าย WAN

สำหรับการเชื่อมต่อแบบ LAN ดังภาพที่ 7 จะเห็นว่าการใช้โพรโทคอล MQTT และระบบ NDN IoT-Hub ดีเลย์มีค่าใกล้เคียงกันอยู่ที่ 1246.8 ± 13.3 ms และ 1328.8±8.7 ms ตามลำดับ แต่การใช้ระบบ NDN IoT-Hub จะไม่ต้องอาศัยการเชื่อมต่อไปที่ส่วนกลาง ทำให้ช่วยลดปัญหาความพร้อมใช้งาน (Availability) ได้

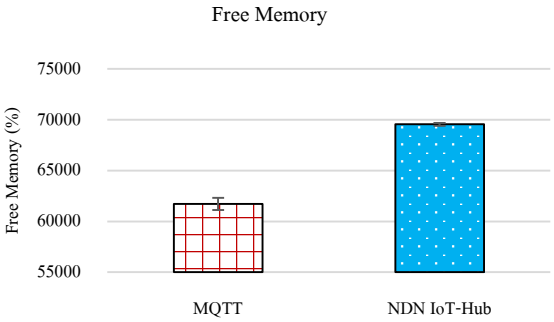
4.3 ผลการประเมินด้านภาระ (Load)



ภาพที่ 8: อัตรา CPU ว่าง

ภาระของแต่ละส่วนในระบบของการใช้โพรโทคอล MQTT และ NDN IoT-Hub แสดงในภาพที่ 8 และภาพที่ 9 แสดงถึง Load ที่ถูกใช้และมีอัตราร้อยละ CPU ว่าง ของการประมวลผลในอุปกรณ์ IoT ซึ่งจะเห็นว่าการประมวลผลของโปรแกรม NDN IoT-Hub มีสัดส่วนการใช้ทรัพยากร CPU ต่ำกว่าการใช้งานโพรโทคอล MQTT แสดงถึงประสิทธิภาพการจัดการ (Handle) แพ็กเก็ตที่มีประสิทธิภาพมากกว่าได้

สำหรับภาพที่ 9 เป็นสัดส่วน Memory ว่าง เมื่อมีการใช้งานอุปกรณ์ IoT โดยใช้โมดูลของโพรโทคอล MQTT และระบบ NDN IoT-Hub ซึ่งจะเห็นว่าการใช้ทรัพยากร Memory ของอุปกรณ์ IoT ต่ำกว่าหากใช้ซอฟต์แวร์ของระบบ NDN IoT-Hub ซึ่งมีเหลืออยู่ถึง 69579±153.4 MB ในขณะที่การใช้โปรแกรมที่ใช้โพรโทคอล MQTT อย่างง่ายจะเหลือ Memory เพียง 62060.5±563.0 MB



ภาพที่ 9: สัดส่วน Memory ว่าง

4.2 วิเคราะห์การอาศัยส่วนกลางเป็นปัจจัยเชื่อมต่อสื่อสารกับอุปกรณ์ IoT

การเชื่อมต่อโดยอาศัยส่วนกลางเพียงปัจจัยเดียว เช่น โพรโทคอล MQTT จะต้องใช้ Broker และ ระบบ NDN IoT-Hub ต้องอาศัย NDN-IH เป็นเครื่องมือเชื่อมต่อระหว่างอุปกรณ์ IoT และผู้ใช้ ซึ่งระบบส่วนกลางนี้เสี่ยงต่อการถูกจู่โจมนำไปสู่ปัญหา Single point of failure ได้ และต้องอาศัยเสถียรภาพของระบบส่วนกลางและเครือข่ายอย่างมาก ถึงแม้ผู้ใช้และอุปกรณ์ IoT จะอยู่ใน LAN เดียวกัน เช่น ในบ้าน เป็นต้น จากการทดลองของงานวิจัยนี้พบว่า การไม่พร้อมใช้งานของ Broker จากการใช้โพรโทคอล MQTT ระบบจะล้มเหลวร้อยละ 100 จากการเชื่อมต่อจากเครือข่าย LAN เดียวกับอุปกรณ์ IoT เพียงเพราะเชื่อมเข้าสู่อินเทอร์เน็ตไม่ได้ ในขณะที่การใช้ระบบ NDN IoT-Hub จะยังคงเชื่อมต่อกับอุปกรณ์ IoT ได้ร้อยละ 100 ถึงแม้ NDN-IH จะไม่พร้อมใช้และการเชื่อมต่อสู่อินเทอร์เน็ตล้มเหลว

5. สรุป

IoT ถูกใช้กับอุปกรณ์ที่มีความหลากหลาย และปัจจุบันการสื่อสารระหว่างอุปกรณ์และผู้ใช้ยังคงอาศัย Broker เป็นส่วนกลาง แม้จะเป็นการสื่อสารแบบเครือข่ายท้องถิ่น งานวิจัยนี้จึงนำความสามารถในการเร้าแพ็กเก็ตด้วยชื่อ Prefix ของเครือข่าย NDN มาปรับปรุงวิธีการเชื่อมต่อสื่อสารในเครือข่าย IoT และพัฒนาระบบ IoT Hub สำหรับเครือข่าย NDN เพื่อเป็นจุดเชื่อมต่อกับอุปกรณ์ IoT ผ่านอินเทอร์เน็ต และสามารถสลับการใช้ในเครือข่ายท้องถิ่นอย่างอัตโนมัติ จากการวัดและประเมินผลด้านดีเลย ประสิทธิภาพการทำงาน และการเร้าแพ็กเก็ตระหว่างเครือข่าย WAN และ LAN แสดงถึงประสิทธิภาพและประสิทธิผลของระบบ NDN IoT-Hub และการใช้เครือข่าย NDN สนับสนุนการเชื่อมต่อเครือข่ายของอุปกรณ์ IoT ได้

เอกสารอ้างอิง

[1] H. Tschofenig, J. Arkko, D. Thaler, and D. McPherson, “Architectural Considerations in Smart Object Networking,” IETF, RFC 7452, March 2015.

[2] “Transmission Control Protocol,” Defense Advanced Research Projects Agency, 4676 Admiralty Way, Marina del Rey, California 90291, RFC 793, September 1981.

[3] “Internet Protocol,” IETF, 4676 Admiralty Way, Marina del Rey, California 90291, RFC 791, September 1981.

[4] W. Shang, Y. Yu, R. Droms, and L. Zhang, “Challenges in IoT Networking via TCP/IP Architecture,” NDN Technical Report NDN-0038, February 2016.

[5] “Message Queuing Telemetry Transport,” ISO, IEC 20922:2016, June 2016.

[6] V. Jacobson, D. Smetters, J. Thornton, M. Plass, N. Briggs, and R. Braynard, “Networking Named Content,” 55 [1]: 117–124, January 2012.

[7] L. Zhang, A. Afanasyev, J. Burke, V. Jacobson, kc claffy, P. Crowley, C. Papadopoulos, L. Wang, and B. Zhang, “Named Data Networking,” 44 [3]: 66–73, July 2014.

[8] NFD Team, “NFD Developer’s Guide,” NDN Technical Report NDN-0021, February 2015.

[9] A. Al-Fuqaha, M. Guizani, and M. Mohammadi, “Internet of Things: A Survey on Enabling Technologies, Protocols, and Applications,” 17 [4]: 2347–2376, June 2015.

[10] P. Desai, P. Anantharam, and A. Sheth, “Semantic Gateway as a Service Architecture for IoT Interoperability,” in Proceedings of IEEE International Conference on Mobile Services, New York, NY, USA, 2 July 2015, pp. 313–319.

[11] S.-M. Kim, H.-S. Choi, and W.-S. Rhee, “IoT home gateway for auto-configuration and management of MQTT devices,” in Proceedings of IEEE Conference on Wireless Sensors, Melaka, Malaysia, 24 August 2015, pp. 12–17.

[12] P. Fremantle, B. Aziz, J. Kopecký, and P. Scott, “Federated Identity and Access Management for the Internet of Things,” in Proceedings of International Workshop on Secure Internet of Things, Wroclaw, Poland, 10 September 2014, pp. 10–17.

การรู้จำอารมณ์ของเสียงพูดโดยการใช้ความแปรปรวนของบริบทของเสียงพูด
Emotional Speech Recognition Using Context-Dependent Variation

ณิธยาน์ สกลพิสิษฐ์ (Nithaya Sakolpisit)[1] สิปปวิชญ์ แซ่หวาง (Sippawit Saewang)[2]
และกันต์พงษ์ วรรัตน์ปัญญา (Kuntpong Woraratpanya)[3]
คณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
[1]tkbink@hotmail.com, [2]ripple_chen@hotmail.co.th, [3]kuntpong@it.kmitl.ac.th

บทคัดย่อ

ปัจจุบันการรู้จำอารมณ์จากเสียงพูดได้กลายเป็นส่วนหนึ่งของการติดต่อสื่อสารระหว่างมนุษย์กับคอมพิวเตอร์ โดยมุ่งหวังให้เครื่องคอมพิวเตอร์สามารถสื่อสารได้ใกล้เคียงกับมนุษย์มากที่สุด แต่อย่างไรก็ตามเสียงพูดของมนุษย์โดยทั่วไปมักจะมีสภาพแวดล้อมรบกวนผสมอยู่กับเนื้อเสียงจริง จึงกลายเป็นสิ่งที่ยากต่อการรู้จำของคอมพิวเตอร์ ดังนั้นในงานวิจัยนี้จึงได้นำเสนอวิธีการแยกเนื้อเสียงพูดจริงออกจากเสียงที่เกิดจากสภาพแวดล้อมรบกวนโดยใช้ความแปรปรวนของบริบทของเสียงพูด หลังจากนั้นนำเนื้อเสียงพูดจริงมาสกัดคุณลักษณะด้วยวิธีการสกัดความถี่ฟอร์แมนต์ จากผลการทดลองเมื่อทดสอบกับคลังข้อมูล อีโมล่า (EMOLA) พบว่าวิธีการที่นำเสนอมีประสิทธิภาพในการรู้จำอารมณ์จากเสียงพูดมากกว่าวิธีการที่นำมาเปรียบเทียบ ซึ่งเป็นผลมาจากการสกัดคุณลักษณะเด่นจากเสียงพูดที่ไม่มีสภาพแวดล้อมรบกวน

คำสำคัญ: การรู้จำอารมณ์ของเสียงพูด คุณลักษณะเด่นของฟอร์แมนต์ การแยกเสียงพูด ความแปรปรวนของบริบทของเสียงพูด สภาพแวดล้อมรบกวน

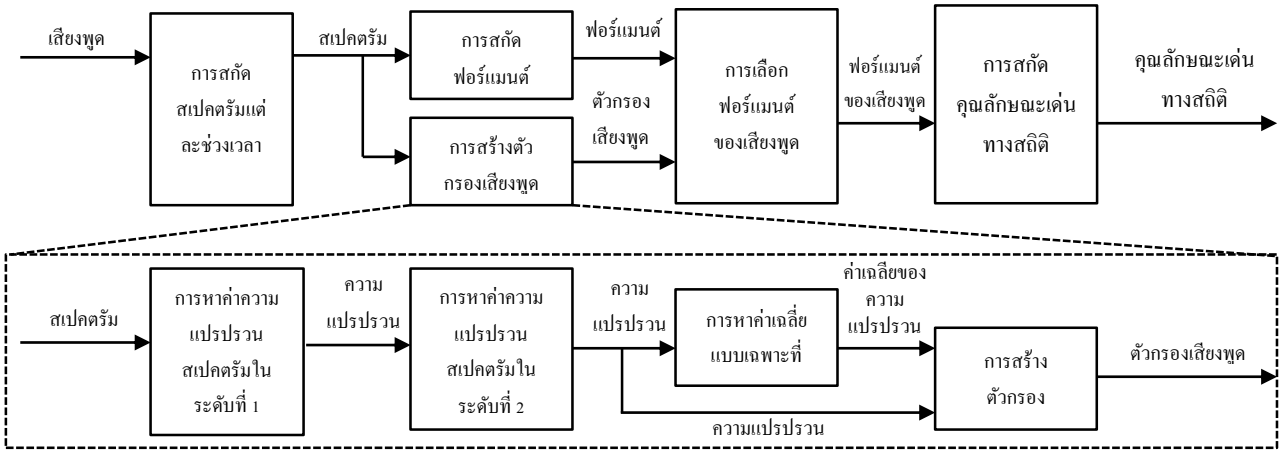
Abstract

Currently, emotional speech recognition is becoming an important part of human-computer interaction (HCI). It aims to make the man-machine interaction more natural and friendly. However, its efficiency is always reduced in noisy environment. This becomes a grand challenge for improving the emotional speech recognition. Therefore, this paper proposes context-dependent variation for extracting emotional speech in the noisy environment. The proposed method applies the local variance of speech spectrum to serve as features of speech extraction. This method can extract emotional speech based on formant frequency cleanly. Furthermore, when the proposed method is tested with EMOLA emotional speech corpus, the experimental result shows that the recognition rate significantly increases when compared to the one of baseline methods. This achievement comes from the effectiveness of extracting the emotional speech in the noisy environment.

Keyword: Emotional Speech Recognition, Formant Feature, Speech Extraction, Context-Dependent Variation, Noisy Environment

1. บทนำ

การสื่อสารด้วยเสียงพูดเป็นกระบวนการที่ง่ายสำหรับมนุษย์ต่อการรับรู้อารมณ์ของผู้พูด แต่เป็นกระบวนการที่ยากต่อการรับรู้ของคอมพิวเตอร์ เนื่องจากในความเป็นจริง เสียงพูดไม่ได้มีเฉพาะเนื้อเสียงเพียงอย่างเดียว แต่เป็นการผสมกันระหว่างเนื้อเสียงจริง และเสียงที่เกิดจากสภาพแวดล้อมรบกวน เช่น เสียงลม เสียงจากคนรอบข้าง เป็นต้น ส่งผลให้การรับรู้และการตีความหมายอารมณ์ของผู้พูดของคอมพิวเตอร์มีความผิดพลาดได้ง่าย ดังนั้น จึงมีงานวิจัยที่พยายามพัฒนาการรู้จำอารมณ์จากเสียงพูด โดยการหาวิธีการสกัดคุณลักษณะเด่นของสัญญาณเสียงพูด งานวิจัยโดยทั่วไปนิยมใช้การสกัดคุณลักษณะเด่นของเสียงด้วย 3 วิธีการ ได้แก่ การหาค่าสัมประสิทธิ์ความถี่สเกลเมล (Mel-Frequency Cepstral Coefficient: MFCC) เป็นวิธีการสกัดคุณลักษณะเด่นมาจากการรับรู้จากการฟังของมนุษย์ สัจพจน์ (Prosodic) เป็นวิธีการสกัดคุณลักษณะเด่นมาจากองค์ประกอบพื้นฐานของเสียงของมนุษย์ และการสกัดความถี่



ภาพที่ 1: กรอบแนวความคิดของวิธีการใช้ความแปรปรวนของบริบทของเสียงพูด

ฟอร์แมนต์ (Formant) เป็นวิธีการที่สามารถสกัดคุณลักษณะเด่นมาจากการเปล่งเสียงของมนุษย์ [1] เนื่องจากฟอร์แมนต์สามารถอธิบายลักษณะจำเพาะ (Characteristic) ของอารมณ์จากเปล่งเสียงของมนุษย์ได้ดี ดังนั้น งานวิจัยนี้จึงเลือกใช้การสกัดคุณลักษณะเด่นด้วยวิธีการการสกัดความถี่ฟอร์แมนต์

อย่างไรก็ตาม เสียงพูดไม่ได้มีเฉพาะเนื้อเสียงเพียงอย่างเดียว แต่เป็นการผสมกันระหว่างเนื้อเสียงจริงและเสียงที่เกิดจากสภาพแวดล้อมรบกวน เพื่อช่วยเพิ่มประสิทธิภาพให้ฟอร์แมนต์สามารถสกัดคุณลักษณะเด่นได้ดีมากขึ้น ในงานวิจัยนี้ได้นำเสนอวิธีการแยกเนื้อเสียงพูดจริงออกจากเสียงที่เกิดจากสภาพแวดล้อมรบกวนโดยใช้ความแปรปรวนของบริบทของเสียงพูด หลังจากนั้นนำเนื้อเสียงพูดจริงมาสกัดคุณลักษณะด้วยวิธีการสกัดความถี่ฟอร์แมนต์ ดังภาพที่ 1 และวัดประสิทธิภาพของคุณลักษณะเด่นด้วยวิธีการหาระยะทางแบบยุคลิด (Euclidean distance) การรู้จำอารมณ์จากเสียงพูดภาษาไทยในงานวิจัยนี้ใช้ฐานข้อมูลภาษาไทยที่เป็นมาตรฐาน อีโมล่า [2] จากศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ (NECTEC)

งานวิจัยฉบับนี้แบ่งออกเป็น 5 หัวข้อ ได้แก่ บทนำ ทฤษฎีและงานวิจัยที่เกี่ยวข้องกับการรู้จำอารมณ์ วิธีการดำเนินการวิจัย วิธีการทดสอบ และผลการทดลอง จากนั้นจึงสรุปผลและอภิปรายในหัวข้อสุดท้าย

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 การสกัดความถี่ฟอร์แมนต์

การสกัดความถี่ฟอร์แมนต์ คือ วิธีการหาเซปตรัมที่มีขนาด (Magnitude) สูงที่สุด ประกอบด้วย 3 ขั้นตอน คือ การสกัดสเปคตรัม การสกัดเซปสตรัม และการเลือกจุดสูงสุดของเซปสตรัม

ขั้นตอนที่ 1 สกัดสเปคตรัมจากเสียงพูดโดยใช้การแปลงฟูเรียร์จากสัญญาณที่ถูกแบ่งกรอบแล้ว ดังสมการที่ (1)

$$X_u^{(i)} = \sum_{n=0}^{N_u-1} x_n^{(i)} \cdot e^{-2\pi i\left(\frac{un}{N_u}\right)} \qquad (1)$$

โดยที่ u, $X_u^{(i)}$, n, $x_n^{(i)}$ และ N_u คือ อันดับของความถี่ของสเปคตรัม, สเปคตรัมความถี่อันดับที่ u ของสัญญาณในกรอบเวลาที่ i, ลำดับของตัวอย่างสัญญาณ, สัญญาณในกรอบเวลาที่ i และจำนวนของความถี่ของสเปคตรัม ตามลำดับ

ผลการสกัดสเปคตรัมจากสัญญาณเสียงพูดจากภาพที่ 2(ก) จะถูกแสดงด้วยสเปคโตแกรมดังภาพที่ 2(ข)

ขั้นตอนที่ 2 สกัดเซปสตรัม ($C_v^{(i)}$) โดยใช้ฟังก์ชันลอกการิทึมร่วมกับการแปลงฟูเรียร์ย้อนกลับ (Inverse Fourier transform) เพื่อทำให้ข้อมูลอยู่บนโดเมนคิวเฟรนซี (Quefrency domain) ดังสมการที่ (2)

$$C_v^{(i)} = \sum_{u=1}^{N_u} \log\left(X_u^{(i)}\right) \cdot \cos\left(\frac{\pi}{N_u}\left(uv + \frac{u}{2}\right)\right) \qquad (2)$$

โดยที่ v และ $C_v^{(i)}$ คือ อันดับคิวเฟรนซีของเซปสตรัม และเซปตรัมลำดับที่ v ตามลำดับ

ขั้นตอนที่ 3 เลือกฟอร์แมนต์ ($F^{(i)}$) จากจุดสูงสุดของเซปสตรัม ดังสมการที่ (3)

$$F^{(i)} = \left\{C_v^{(i)} \middle| C_{v-1}^{(i)} < C_v^{(i)} \text{ and } C_v^{(i)} > C_{v+1}^{(i)}\right\} \qquad (3)$$

โดยที่ $F^{(i)}$ คือ ฟอร์แมนต์ของกรอบเวลาที่ i

จุดสูงสุดของเซปสตรัมแต่ละจุด คือ ฟอร์แมนต์ของแต่ละลำดับ ซึ่งความถี่ฟอร์แมนต์จะถูกคำนวณจากอันดับคิวเฟรนซี แล้วจะถูกนำมาใช้เป็นคุณลักษณะเด่นที่เรียกว่า “ความถี่ฟอร์แมนต์”

2.2 งานวิจัยที่เกี่ยวข้อง

งานวิจัยที่เกี่ยวข้องกับการรู้จำอารมณ์ของเสียงพูด มีดังนี้

D. Ververidis และคณะ [3] ได้รวบรวมและสรุปงานวิจัยที่เกี่ยวข้องกับการรู้จำอารมณ์จากเสียงพูด ใน 3 ประเด็น ได้แก่ (1) ความหลากหลายทางด้านภาษา เช่น ภาษาอังกฤษ ภาษาเยอรมัน ภาษาญี่ปุ่น และภาษาจีน เป็นต้น (2) ความหลากหลายทางด้านคุณลักษณะเด่นที่ใช้ เช่น Pitch, Formant, MFCC, Teager-energy, Vocal tract cross-section area และ (3) ความหลากหลายทางด้านวิธีการจำแนกอารมณ์ เช่น Hidden Markov model (HMM), Artificial Neural Network (ANN), Linear discriminant analysis (LDA), *k*-nearest neighbor, และ Support vector machine (SVM)

Y. Yoshitomi และคณะ [4] นำเสนอวิธีการรู้จำอารมณ์จากเสียงพูดร่วมกับใบหน้า โดยสกัดคุณลักษณะเด่นของเสียงจาก Pitch, Energy และ Derivative จากนั้นทำการรู้จำโดยใช้ HMM ส่วนการรู้จำจากใบหน้า จะสกัดคุณลักษณะโดยใช้ Discrete cosine transform (DCT) จากนั้นทำการรู้จำโดยใช้นิวรอนเน็ตเวิร์ก (Neural Network: NN)

ต่อมา Y. Wang และคณะ [1] นำเสนอวิธีการรู้จำอารมณ์จากเสียงพูดร่วมกับใบหน้า โดยสกัดคุณลักษณะเด่นของเสียงจาก Prosodic, MFCC และ Formant ส่วนการสกัดคุณลักษณะจากใบหน้าใช้ Gabor wavelet จากนั้นจะทำการรวมคุณลักษณะเด่นจากทั้งสองชุด โดยใช้วิธีการ Stepwise และทำการรู้จำโดยใช้ LDA

P. Sukkummek และคณะ [5] นำเสนอวิธีการเลือกคุณลักษณะเด่นสำหรับการรู้จำอารมณ์จากเสียงพูด โดยใช้วิธีการ Minimum redundancy maximum relevance (mRMR) ซึ่งวิธีการนี้จะใช้คุณลักษณะเด่นมาจากหลายวิธีการ เช่น Linear Prediction Filter Coefficient (LPC), MFCC, Prosodic, Energy, Zero-crossing rate และวัดประสิทธิภาพการเลือกคุณลักษณะเด่นโดยใช้คลังข้อมูลภาษาอังกฤษและภาษาไทย

M. A. Kammoun และคณะ [6] นำเสนอวิธีการสกัดความถี่ฟอร์แมนต์จากสมูทสเปคตรัม (Smooth spectrum) โดยสเปคตรัมจะถูกทำให้ราบเรียบด้วยวิธีการ Temporal windowing จากนั้นนำสมูทสเปคตรัมไปสกัดคุณลักษณะด้วยวิธีการสกัดฟอร์แมนต์

3. วิธีการดำเนินงานวิจัย

จากงานวิจัยที่เกี่ยวข้องกับการสกัดความถี่ฟอร์แมนต์เป็นวิธีสกัดคุณลักษณะเด่นที่มีประสิทธิภาพ สามารถอธิบายลักษณะจำเพาะของอารมณ์จากการเปล่งเสียงของมนุษย์ได้ดี แต่ในทางปฏิบัติการรู้จำอารมณ์จากเสียงพูดมีเสียงรบกวนจากสภาพแวดล้อม ซึ่งมีผลต่อประสิทธิภาพของการรู้จำอารมณ์ของเสียงพูดเมื่อใช้ความถี่ฟอร์แมนต์ ดังนั้น งานวิจัยนี้จึงนำเสนอการใช้ความแปรปรวนของบริบทของเสียงพูดสำหรับสกัดอารมณ์ของเสียงพูดออกจากสภาพแวดล้อมรบกวน ขั้นตอนของการรู้จำอารมณ์จากเสียงพูดประกอบด้วย 2 กระบวนการหลัก คือ การสกัดคุณลักษณะเด่นของอารมณ์ของเสียงพูด โดยใช้ความแปรปรวนของบริบทของเสียงพูด และการจำแนกอารมณ์ของเสียงพูดโดยใช้ระยะทางแบบยุคลิด

กระบวนการแรก คือ การสกัดคุณลักษณะเด่นของเสียงพูดโดยใช้ความแปรปรวนของบริบทของเสียงพูดเพื่อสกัดอารมณ์ของเสียงพูดจากสภาพแวดล้อมรบกวน ซึ่งมีวิธีการทำงาน ดังนี้

ขั้นตอนที่ 1 แบ่งสัญญาณ (x) ออกเป็นกรอบสัญญาณ ($x^{(i)}$) โดยแต่ละกรอบสัญญาณจะมีขนาดและคาบที่ทับซ้อนกัน (Overlapping)

ขั้นตอนที่ 2 สกัดสเปคตรัมของแต่ละกรอบสัญญาณโดยใช้การแปลงฟูริเยร์ดังสมการที่ (1) ซึ่งได้ผลดังแสดงในภาพที่ 2(ข)

ขั้นตอนที่ 3 สกัดฟอร์แมนต์ ($F^{(i)}$) ดังสมการที่ (2) และ (3)

ขั้นตอนที่ 4 สร้างตัวกรอง (H_i) จากวิธีการสร้างตัวกรองเสียงพูด สามารถทำได้ดังนี้

ขั้นตอนที่ 4.1 คำนวณหาค่าความแปรปรวนสเปคตรัมในระดับที่ 1 (s_i) ในแต่ละกรอบสัญญาณดังสมการที่ (4) ซึ่งได้ผลดังแสดงในภาพที่ 2(ค)

$$s_i = \frac{1}{N_u}\sum (X_u^{(i)} - \bar{X}^{(i)})^2 \quad (4)$$

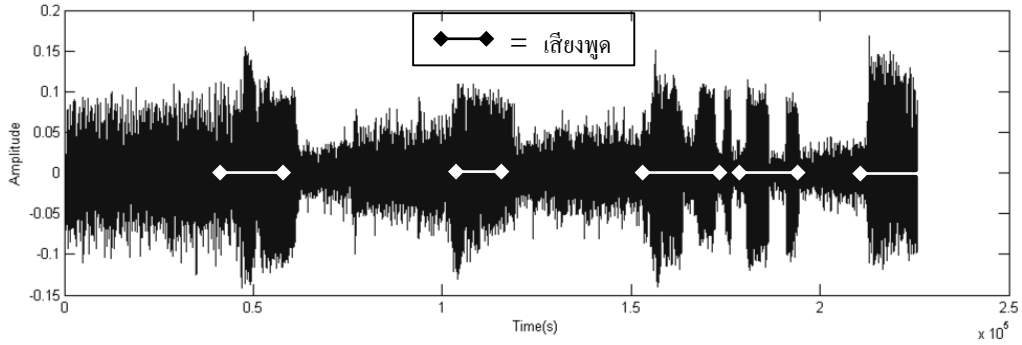
โดยที่ $\bar{X}^{(i)}$ คือค่าเฉลี่ยของสเปคตรัมของแต่ละกรอบสัญญาณ

ขั้นตอนที่ 4.2 คำนวณหาค่าความแปรปรวนสเปคตรัมในระดับที่ 2 (s'_i) ในแต่ละกรอบสัญญาณดังสมการที่ (5) ซึ่งได้ผลดังแสดงในภาพที่ 2(ง)

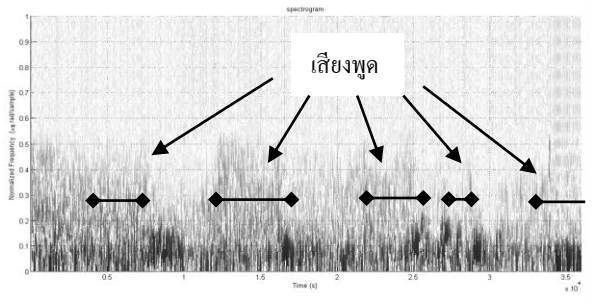
$$s'_i = \frac{1}{W}\sum_{k=i}^{i+W} (s_k - \bar{s}_i)^2 \quad (5)$$

โดยที่ W คือ ขนาดของกรอบความแปรปรวน

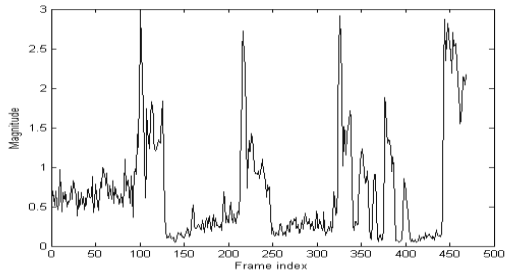
ขั้นตอนที่ 4.3 คำนวณหาค่าเฉลี่ยแบบเฉพาะ (m_i) ของความแปรปรวนสเปคตรัมในระดับที่ 2 (s'_i) ในแต่ละกรอบ



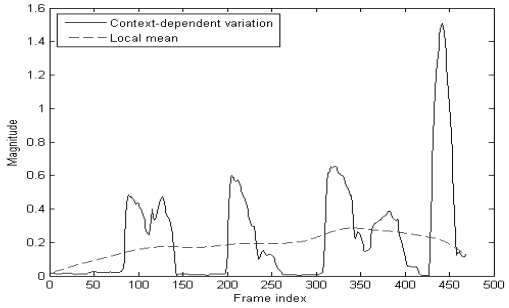
(ก) ตัวอย่างสัญญาณเสียงจากฐานข้อมูลอีโมล่า (EMOLA) มีลักษณะภายใต้สภาพแวดแวดล้อมประกอบด้วย เสียงลม และเสียงเครื่องจักร



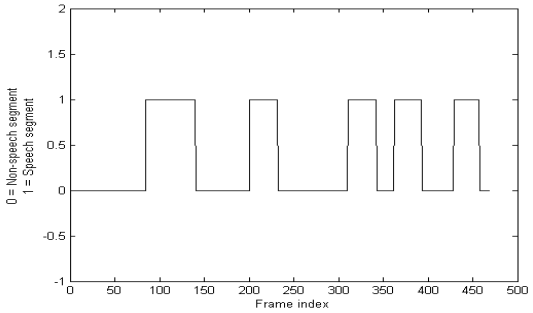
(ข) สเปคตรัมของเสียงพูดที่ถูกแบ่งกรอบสัญญาณแล้ว



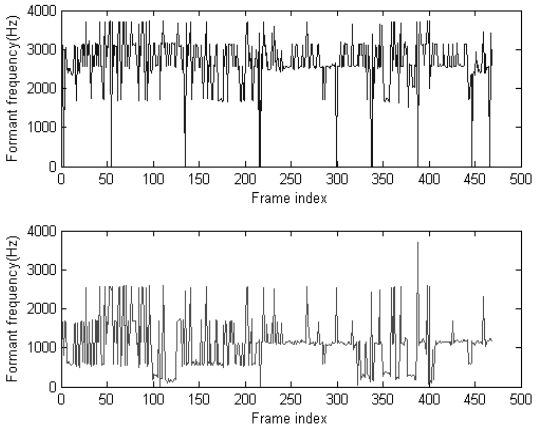
(ค) ความแปรปรวนของสเปคตรัมในแต่ละกรอบสัญญาณ



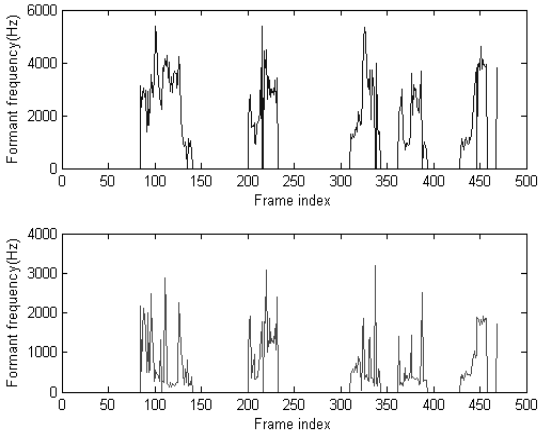
(ง) ความแปรปรวนของบริบทของคำพูด และ การค่าเฉลี่ยเฉพาะที่



(จ) ช่วงคำพูดของแต่ละกรอบสัญญาณ



(ฉ) ความถี่ฟอร์แมนต์ลำดับที่ 1 (ล่าง) และ 3 (บน) ของเสียงพูดที่ยังไม่ผ่านตัวกรอง



(ช) ความถี่ฟอร์แมนต์ลำดับที่ 1 (ล่าง) และ 3 (บน) ของเสียงพูดที่ผ่านตัวกรอง

ภาพที่ 2: ผลลัพธ์ของแต่ละขั้นตอนในการสกัดคุณลักษณะเด่นของเสียงพูดด้วยวิธีการใช้ความแปรปรวนของบริบทของเสียงพูด

ตารางที่ 1: จำนวนตัวอย่างที่ใช้ฝึกฝนและทดสอบ

อารมณ์	จำนวนตัวอย่าง ฝึกฝน	จำนวนตัวอย่าง ทดสอบ
โกรธ	472	471
มีความสุข	545	544
เศร้า	172	171
ปกติ	226	226
รวม	1,415	1,412

สัญญาณดังสมการที่ (6) ซึ่งได้ผลดังแสดงในภาพที่ 2(ง)

$$m_i = \frac{1}{2w+1} \sum_{l=i-w}^{l+w} s_i' \qquad (6)$$

โดยที่ w คือ ขนาดของกรอบความแปรปรวนของบริบทคำพูด

ขั้นตอนที่ 4.4 สร้างตัวกรอง (H_i) จากค่าเฉลี่ยแบบเฉพาะที่ (m_i) ดังสมการที่ (7) ซึ่งได้ผลดังแสดงในภาพที่ 2(จ)

$$H_i = \begin{cases} 1, & s_i' > m_i \\ 0, & s_i' \le m_i \end{cases} \qquad (7)$$

ขั้นตอนที่ 5 นำฟอร์แมนต์ของเสียงพูด (F') จากภาพที่ 2(ฉ) มาผ่านการกรองช่วงสัญญาณดังสมการที่ (8) ซึ่งได้ผลดังแสดงในภาพที่ 2(ช)

$$F' = \{F^{(i)} \mid H_i > 0\} \qquad (8)$$

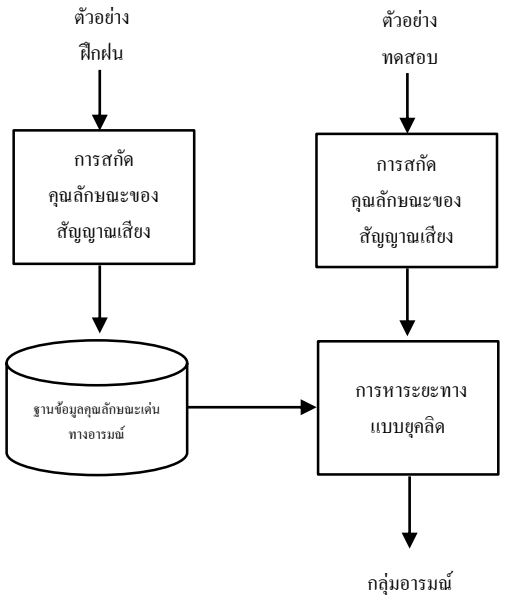
ขั้นตอนที่ 6 คำนวณความถี่ฟอร์แมนต์ และสกัดคุณลักษณะเด่นทางสถิติ ได้แก่ ค่าเฉลี่ย ค่าความแปรปรวน ค่าสูงสุด ค่าต่ำสุด และค่ามัธยฐาน ดังงานวิจัยที่ [1]

หลังจากที่ทำการสกัดคุณลักษณะเด่นด้วยวิธีการข้างต้น กระบวนการต่อไปจะทำการแยกคุณลักษณะเด่นชุดฝึกฝนที่ได้มาทำการสร้างแบบจำลองรู้จำอารมณ์ด้วยวิธีการหาระยะทางแบบยุคลิด และใช้ชุดทดสอบเพื่อวัดประสิทธิภาพของวิธีการที่สร้างขึ้นดังภาพที่ 3 โดยรายละเอียดการจำแนกอารมณ์จะกล่าวในขั้นตอนการทดลอง

4. การทดลอง

การทดลองนี้มีจุดประสงค์เพื่อวัดประสิทธิภาพของการรู้จำอารมณ์จากเสียงพูดโดยใช้ความแปรปรวนของบริบทของอารมณ์ของเสียงพูดจากสภาพแวดล้อมรบกวน

การทดลองนี้ใช้คลังข้อมูลเสียงภาษาไทยจาก Emotional speech corpus from Lakorn (EMOLA) [2] จากศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ (NECTEC) ประกอบด้วยข้อมูลเสียงพูด ซึ่งถูกรวบรวมจากละครโทรทัศน์ มี 4 อารมณ์ ได้แก่ โกรธ มีความสุข เศร้า และปกติ โดยอัตราส่วนของข้อมูล



ภาพที่ 3: กรอบแนวความคิดของการรู้จำอารมณ์

ชุดฝึกฝน (Training) ต่อข้อมูลชุดทดสอบ (Testing) คือ 50:50 สามารถแจกแจงได้ดังตารางที่ 1 งานวิจัยนี้ได้ทำการทดลองทั้งหมด 3 วิธีการ ได้แก่ (1) วิธีการสกัดความถี่ฟอร์แมนต์ (2) วิธีการสกัดความถี่สกัดฟอร์แมนต์ร่วมกับการแบ่งช่วงพยางค์ [7] (3) วิธีการสกัดความถี่ฟอร์แมนต์ร่วมกับการใช้ความแปรปรวนของบริบทของเสียงพูด และการกำหนดพารามิเตอร์ที่ใช้ในการทดลองมีดังนี้ กรอบเวลาของสัญญาณเสียงแต่ละกรอบเวลามีความยาว 30 มิลลิวินาที มีช่วงคาบเกี่ยว 20 มิลลิวินาที ขนาดของกรอบความแปรปรวนแบบท้องถิ่น (W) เท่ากับ 31 และขนาดของกรอบของค่าขีดแบ่งแบบท้องถิ่น (Local threshold: w) เท่ากับ 180 และจำนวนของฟอร์แมนต์ ใช้ 3 อันดับต่อเฟรม

ผลการทดลองปรากฏดังตารางที่ 2 พบว่าวิธีการสกัดความถี่ฟอร์แมนต์ร่วมกับการใช้ความแปรปรวนของบริบทของมีประสิทธิภาพในรู้จำอารมณ์จากเสียงพูดของอารมณ์โกรธ 45.71 % อัตราการรู้จำของอารมณ์มีความสุข 48.16 % อัตราการรู้จำของอารมณ์เศร้า 22.81 % อัตราการรู้จำของอารมณ์ปกติ 26.55 % และอัตราการรู้จำของอารมณ์ โดยเฉลี่ย 40.80 % ซึ่งสูงกว่าวิธีการสกัดความถี่ฟอร์แมนต์ และวิธีการสกัดความถี่ฟอร์แมนต์ร่วมกับการแบ่งช่วงพยางค์ในทุกๆ อารมณ์ จากตารางที่ 3 เนื่องจากวิธีการที่นำเสนอคุณลักษณะเด่นจะถูกสกัดมาจากเนื้อเสียงแท้จริงซึ่งถูกแยกจากสภาพแวดล้อมรบกวนด้วยวิธีความแปรปรวนของบริบทของเสียงพูด อย่างไรก็ตามฐานข้อมูลที่นำมาใช้ในการทดสอบนี้

ตารางที่ 2: การเปรียบเทียบผลการรู้จำอารมณ์จากเสียงพูดของวิธีการนำเสนอกับวิธีการสกัดความถี่ฟอร์แมนต์ และวิธีร่วมกับการแบ่งช่วงพยางค์

วิธีการ	ค่าความถูกต้องของการรู้จำอารมณ์ (%)				
	โกรธ	มีความสุข	เศร้า	ปกติ	ค่าเฉลี่ย
สกัดความถี่ฟอร์แมนต์ [1]	33.69	47.98	21.64	24.34	36.25
สกัดความถี่ฟอร์แมนต์ร่วมกับการแบ่งช่วงพยางค์ [7]	34.55	38.97	14.62	18.14	31.20
วิธีการที่นำเสนอ	**45.71**	**48.16**	**22.81**	**26.55**	**40.80**

ตารางที่ 3: การเปรียบเทียบข้อดีและข้อเสียระหว่างการใช้ความแปรปรวนของบริบทของเสียงพูด และไม่ใช้ความแปรปรวนของบริบทของเสียงพูด

	การใช้ความแปรปรวนของบริบทของเสียงพูด	การไม่ใช้ความแปรปรวนของบริบทของเสียงพูด
ข้อดี	เสียงรบกวนถูกกำจัด ทำให้การสกัดคุณลักษณะมีความเจาะจงมากยิ่งขึ้น ความแม่นยำในการรู้จำเพิ่มขึ้น	มีความรวดเร็วในการสกัดคุณลักษณะ และเนื้อเสียงมีความครบถ้วน
ข้อเสีย	เนื้อเสียงบางส่วนอาจถูกกำจัดออกไปกับเสียงรบกวน	เสียงรบกวนยังคงอยู่ ทำให้คุณลักษณะที่ได้มีความเจาะจงค่อนข้างน้อย ความแม่นยำในการรู้จำลดลง

มีความท้าทายอยู่หลายประการ เช่น เสียงต่างเพศ เสียงต่างวัย เสียงต่างสำเนียง เสียงดนตรีพื้นหลัง เสียงรอบข้างเป็นต้น ซึ่งส่งผลให้ยากต่อการจำแนกทำให้ค่าประสิทธิภาพในการรู้จำอารมณ์จากเสียงพูดไม่สูงมากนัก

5. สรุป

ในงานวิจัยนี้ได้นำเสนอวิธีการแยกเนื้อเสียงพูดจริงออกจากเสียงที่เกิดจากสภาพแวดล้อมรบกวน โดยใช้ความแปรปรวนของบริบทของเสียงพูด หลังจากนั้นนำเนื้อเสียงพูดจริงมาสกัดคุณลักษณะด้วยวิธีการสกัดความถี่ฟอร์แมนต์ และวัดประสิทธิภาพของคุณลักษณะเด่นด้วยวิธีการหาระยะทางแบบยุคลิด (Euclidean distance) การรู้จำอารมณ์จากเสียงพูดภาษาไทย โดยในงานวิจัยนี้ใช้ฐานข้อมูลภาษาไทยที่เป็นมาตรฐาน อีโมล่า จากศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ (NECTEC) จากผลการทดลอง พบว่าวิธีการสกัดความถี่ฟอร์แมนต์ร่วมกับการใช้ความแปรปรวนของบริบทของมีประสิทธิภาพในรู้จำอารมณ์จากเสียงพูด โดยอัตราการรู้จำของอารมณ์โดยรวม 40.80 % ซึ่งมีประสิทธิภาพสูงกว่าวิธีการสกัดความถี่ฟอร์แมนต์ และวิธีการสกัดความถี่ฟอร์แมนต์ร่วมกับการแบ่งช่วงพยางค์ในทุก ๆ เนื่องจากวิธีการที่นำเสนอ คุณลักษณะเด่นจะถูกสกัดมาจากเนื้อเสียงแท้จริงซึ่งถูกแยกจากสภาพแวดล้อมรบกวนด้วยวิธีความแปรปรวนของบริบทของเสียงพูด

เอกสารอ้างอิง

[1] Y. Wang and L. Guan, "Recognizing human emotional state from audiovisual signals," *IEEE Transactions on Multimedia*, vol. 10, no. 5, pp. 936–946, 2008.

[2] S. Kasuriya, T. Teeramunkong, and C. Wutiwiwatchai, "Developing a Thai emotional speech corpus," *International Conference Oriental held jointly with 2013 Conference on Asian Spoken Language Research and Evaluation*, pp. 1–5, 2013.

[3] D. Ververidis and C. Kotropoulos, "Emotional speech recognition: resources, features, and methods," *Speech Communication*, vol. 48, no. 9, pp. 1162–1181, 2006.

[4] Y. Yoshitomi, S.-I. Kim, T. Kawano, and T. Kilazoe, "Effect of sensor fusion for recognition of emotional states using voice, face image and thermal image of face," *Proceedings 9th IEEE International Workshop on Robot and Human Interactive Communication*, pp. 178–183, 2000.

[5] P. Sukhummek, S. Kasuriya, T. Theeramunkong, C. Wutiwiwatchai, and H. Kunieda, "Feature selection experiments on emotional speech classification," *12th International Conference on Electrical Engineering/-Electronics, Computer, Telecommunications and Information Technology*, pp. 1–4, 2015.

[6] M. A. Kammoun, D. Gargouri, M. Frikha and A. Ben Hamida, "Cepstral method evaluation in speech formant frequencies estimation," *Industrial Technology, 2004. IEEE. 2004 IEEE International Conference on*, Vol. 3, pp. 1612-1616, 2004.

[7] N. Chanamool, B. Unhalekajit, and M. Phothisonothai, "Development of computer software for pathological analyzing of Thai language speakers," *KMUTT Research and Development Journal*, vol. 33, no. 4, pp. 319–328, 2010.

การรู้จำอารมณ์ของเสียงพูดโดยใช้วิธีจำแนกแบบสองระดับ
Emotional Speech Recognition Using Two-level Classification

สิปปวิชญ์ แซ่หวาง (Sippawit Saewang)[1] ณิธยาน์ สกลพิสิษฐ์ (Nithaya Sakolpisit)[2]
และกันต์พงษ์ วรรัตน์ปัญญา (Kuntpong Woraratpanya)[3]
[1]คณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
[1]ripple_chen@hotmail.co.th, [2]tkbink@hotmail.com, [3]kuntpong@it.kmitl.ac.th

บทคัดย่อ

การรู้จำอารมณ์ของเสียงพูดมีการพัฒนาอย่างต่อเนื่องเพื่อประยุกต์ใช้งานอย่างหลากหลาย เช่น การประเมินสุขภาพจิตโดยผ่านระบบโทรศัพท์หรือเครื่องมือติดต่อสื่อสาร และการตรวจจับอารมณ์ของผู้เรียนของระบบการเรียนผ่านระบบเครือข่าย โดยทั่วไปคุณลักษณะเด่นของเสียงพูดจะถูกตรวจจับโดยใช้สัมประสิทธิ์เซปสตรัมบนสเกลเมลซึ่งใช้สำหรับการรู้จำเสียงพูด อย่างไรก็ตามเมื่อใช้คุณลักษณะเด่นสัมประสิทธิ์เซปสตรัมบนสเกลเมลมาจำแนกอารมณ์ของเสียงพูดมักจะทำให้เกิดการจำแนกกลุ่มไม่ถูกต้อง ดังนั้นงานวิจัยนี้จึงนำเสนอวิธีจำแนกแบบสองระดับสำหรับการรู้จำอารมณ์ของเสียงพูด โดยในระดับที่ 1 ใช้การแบ่งกลุ่มแบบเคมีนและวิธีการหาระยะทางแบบยุคลิด การแบ่งเพื่อแยกคุณลักษณะเด่นระหว่างกลุ่มของอารมณ์ หลังจากนั้นกลุ่มของอารมณ์ที่คลุมเครือจะถูกจำแนกโดยใช้คุณลักษณะเด่นภายในกลุ่มในระดับที่ 2 ด้วยวิธีการหาระยะทางแบบยุคลิด โดยฐานข้อมูลที่ใช้ คือ Ryerson Emotion Database (RED) ซึ่งเป็นเสียงพูดภาษาอังกฤษประกอบด้วย 6 อารมณ์ ได้แก่อารมณ์โกรธ อารมณ์รังเกียจ อารมณ์กลัว อารมณ์มีความสุข อารมณ์เศร้า และอารมณ์ประหลาดใจ จากผลการทดลองพบว่าวิธีการที่นำเสนอให้ประสิทธิภาพในการจำแนกดีกว่าวิธีการอื่น

คำสำคัญ: การรู้จำอารมณ์ของเสียงพูด การจำแนกแบบสองระดับ อัลกอริทึมการแยกกลุ่มแบบเคมีน คุณลักษณะเด่นภายในกลุ่ม และคุณลักษณะเด่นภายนอกกลุ่ม

Abstract

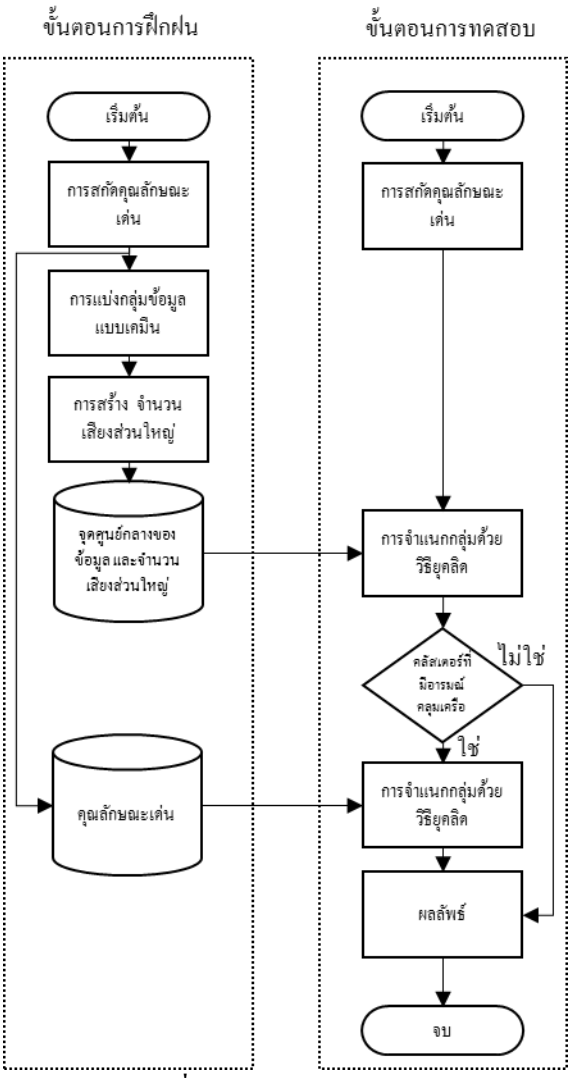
Emotional speech recognition has been developed for various applications such as mental health assessment via phones or other voice communications and emotional learner detection in E-learning. In general, speech feature is detected by means of Mel Frequency Cepstral Coefficient (MFCC) widely used in speech recognition. However, when the MFCC feature is used for emotional speech recognition, it is mostly misclassified. Therefore, this paper proposes a two-level classification for the emotional speech recognition. In this way, k-means clustering algorithm is used to classify inter-cluster features in the first level so that the clear emotions are completely separated. Then the ambiguous emotions are classified by intra-cluster features in the second level. In this stage, the simple Euclidean Distance is used for classification. Based on Ryerson Emotion Database (RED) in the English language including six emotions—anger, disgust, fear, happy, sadness, and surprise, the experimental results show that the proposed method outperforms the conventional methods, including k-Nearest Neighbors (k-NN), Decision Tree, Linear Discriminant Analysis (LDA) and Support Vector Machine (SVM).

Keyword: emotional speech recognition, two-level classification, *k*-means clustering algorithm, inter-cluster features, intra-cluster feature.

1. บทนำ

การรู้จำอารมณ์จากเสียงพูด คือ การที่คอมพิวเตอร์สามารถรับรู้เสียงพูดของมนุษย์และแปลความหมายได้ว่าในขณะที่พูดผู้พูดมีอารมณ์อย่างไร เช่น มีอารมณ์โกรธ มีความสุข เศร้า ประหลาดใจ หรืออารมณ์ปกติ เป็นต้น ปัจจุบันการรู้จำอารมณ์จากเสียงพูดได้มีการนำมาประยุกต์ใช้ในหลาย ๆ ด้าน เช่น ทางด้านการแพทย์ ได้มีการนำมาใช้ในการวิเคราะห์สุขภาพจิตทั้งแบบการสื่อสารโดยตรงและแบบออนไลน์ เพื่อช่วยตัดสินใจว่าควรจะให้คำปรึกษาอย่างไรจึงจะเหมาะสมกับอารมณ์ของผู้ป่วยในขณะนั้น ด้านการศึกษาแบบออนไลน์ถูกนำมาใช้เพื่อช่วยในการตัดสินใจเกี่ยวกับการออกแบบสื่อการเรียนการสอนให้เหมาะสมกับผู้เรียน เป็นต้น

ด้วยความหลากหลายของงานที่นำไปประยุกต์ใช้ ทำให้มีผู้วิจัยพยายามพัฒนาวิธีการรู้จำอารมณ์จากเสียงพูดเพื่อให้มีความแม่นยำและเหมาะสมกับการประยุกต์ใช้งานสำหรับแต่ละงานมากขึ้น D. Ververidis และ C. Koutropulos [1] ได้สำรวจงานวิจัยเกี่ยวกับการรู้จำอารมณ์จากเสียงพูด ได้แก่ (1) ความหลากหลายของภาษาที่ใช้ในการรู้จำอารมณ์ เช่น ภาษาอังกฤษ ภาษาญี่ปุ่น ภาษาจีน เป็นต้น (2) คุณลักษณะเด่นที่เหมาะสมต่อการรู้จำอารมณ์ของมนุษย์ เช่น ระดับของเสียงพูด (Pitch), สัมประสิทธิ์เซปสตรัมบนสเกลเมล (Mel-Frequency Cepstral Coefficient: MFCC), ฟอร์แมนต์ (Formant) และ (3) วิธีการจำแนกอารมณ์ เช่น Support Vector Machine (SVM), Neural Network (NN), Linear Discriminant Analysis (LDA), *k*-Nearest Neighbours (*k*-NN) และการรู้จำอารมณ์จากเสียงพูดร่วมกับการรู้จำใบหน้าถูกนำเสนอโดย Y. Yoshitomi และคณะ [2] และ Y. Wang และคณะ [3] ในงานวิจัยนี้ใช้วิธีการสกัดคุณลักษณะเด่นของเสียงจาก Pitch, Energy และ Derivative และรู้จำอารมณ์จากเสียงพูดด้วยโครงข่ายประสาทเทียม (Neural Network: NN) การสกัดคุณลักษณะเด่นของเสียงจาก Prosodic, MFCC และFormant แล้วทำการรู้จำด้วย Linear discriminant analysis (LDA) ที่ได้นำเสนอไว้ใน [4] P. Sukkummek และคณะ [5] ให้ความสำคัญในการเลือกคุณลักษณะเด่นจากคุณลักษณะเด่นที่ผ่านกระบวนการสกัดแล้วเพื่อให้ได้คุณลักษณะเด่นที่เหมาะในการรู้จำอารมณ์มากขึ้น



ภาพที่ 1: แผนภาพขั้นตอนการทำงาน

โดยใช้วิธีการ minimum Redundancy Maximum Relevance (mRMR) ส่วน M. A. Kammoun และคณะ [6] สกัดคุณลักษณะเด่นด้วยฟอร์แมนต์อยู่บนสเปคตรัมที่ผ่านการสมูทแล้ว ซึ่งในการสมูทจะใช้วิธี Temporal windowing

จากงานวิจัยข้างต้นจะเห็นว่าในการรู้จำอารมณ์จากเสียงพูดนั้น จะมุ่งเน้นในการสกัดคุณลักษณะเด่นโดยทั่วไปคุณลักษณะเด่นที่นิยมใช้กันอย่างกว้างขวางในงานรู้จำอารมณ์จากเสียงพูด คือ MFCC อย่างไรก็ตามคุณลักษณะเด่น MFCC สำหรับการรู้จำอารมณ์ของเสียงพูดค่อนข้างมีความคล้ายคลึงกัน ซึ่งส่งผลให้การจำแนกอารมณ์มีความผิดพลาดได้ง่าย ดังนั้นงานวิจัยนี้จึงนำเสนอวิธีจำแนกคุณลักษณะเด่น MFCC แบบสองระดับ โดยคุณลักษณะเด่นของข้อมูลชุดฝึกฝน (Training)

จะถูกแบ่งกลุ่มในระดับที่ 1 ด้วยวิธีการแบ่งกลุ่มแบบเคมีน (*k*-means) เพื่อแบ่งกลุ่มคุณลักษณะเด่นของอารมณ์ออกเป็นคลัสเตอร์ แล้วนำไปใช้จำแนกคุณลักษณะเด่นของข้อมูลชุดฝึกฝนและข้อมูลชุด ทดสอบ (Testing) ด้วยวิธีการหาระยะทางแบบยุคลิด (Euclidean distance) แต่หากพบว่าผลลัพธ์ในแต่ละคลัสเตอร์มีมากกว่า 1 อารมณ์ คุณลักษณะเด่นจะถูกนำมาจำแนกอีกครั้งหนึ่งในระดับที่ 2 เฉพาะกลุ่มของอารมณ์ที่เกี่ยวข้อง การรู้จำอารมณ์จากเสียงพูดภาษาอังกฤษในงานวิจัยนี้ใช้ฐานข้อมูล Ryerson Emotion Database (RED) งานวิจัยฉบับนี้แบ่งออกเป็น 5 หัวข้อ ได้แก่ หัวข้อบทนำ หัวข้อทฤษฎีที่เกี่ยวข้อง หัวข้อวิธีดำเนินการวิจัย หัวข้อการทดลองและผลการทดลอง และหัวข้อสรุปผล

2. ทฤษฎีที่เกี่ยวข้อง

2.1 การสกัดคุณลักษณะจากเสียง

พื้นฐานของการรู้จำอารมณ์จากเสียง ประกอบด้วยปัจจัยสำคัญสองประการ คือ การสกัดคุณลักษณะเด่น และการจำแนกคุณลักษณะเด่นเพื่อจำแนกอารมณ์ของเสียงพูด จากศึกษาค้นคว้าพบว่าคุณลักษณะเด่นที่นิยมใช้กันอย่างกว้างขวาง คือ Mel-Frequency Cepstral Coefficient (MFCC) [7] ถูกใช้อย่างแพร่หลายในงานวิจัยด้านการรู้จำเสียงพูด คุณลักษณะเด่นนี้มีการเลียนแบบคุณสมบัติของเสียงพูดจำเพาะให้ใกล้เคียงกับการได้ยินของมนุษย์ โดยนำกระบวนการ Cepstral Analysis มาใช้ ซึ่ง MFCC นั้นได้ใช้ตัวกรองที่เรียกว่า Mel-filter bank เพื่อทำการกรองค่าแมกนิจูด (Magnitude) บนความถี่จากต่ำไปสูง เพื่อทำให้ค่าสัมประสิทธิบนความถี่สูงน้อยลงและทำให้สัมประสิทธิบนความถี่ต่ำมากขึ้น ซึ่งจะได้ค่าสัมประสิทธิมาทั้งหมด 13 ค่าจากทั้งสัญญาณ ซึ่งในงานวิจัยนี้ได้เน้นไปทางการสกัดอารมณ์ จึงสนใจแนวโน้มหรือภาพรวมของเสียง จึงคำนวณหาค่าเฉลี่ย (Mean) ค่ามัธยฐาน (Median) ค่าส่วนเบี่ยงเบน-มาตรฐาน (Standard Deviation) ค่าสูงสุด (Max) และค่าต่ำสุด (Min) จากทั้งสัญญาณ ซึ่งจะทำให้ได้ จำนวนคุณลักษณะเด่นทั้งหมด 65 ชนิด [3]

ตารางที่ 1: การกำหนดชื่อย่อของแต่ละอารมณ์ของเสียงพูด

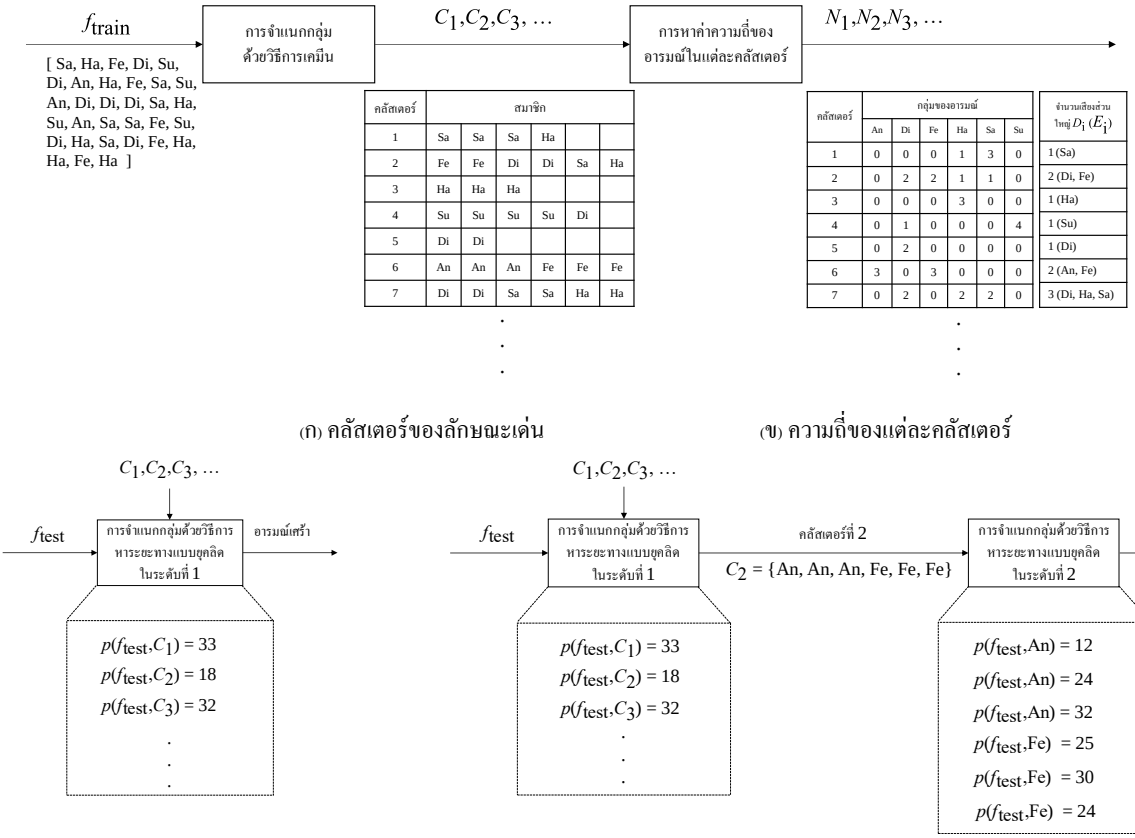
อารมณ์	ตัวย่อ
อารมณ์โกรธ	An
อารมณ์รังเกียจ	Di
อารมณ์กลัว	Fe
อารมณ์ความสุข	Ha
อารมณ์เศร้า	Sa
อารมณ์ประหลาดใจ	Su

2.2 การจำแนกอารมณ์

วิธีการในการจำแนกอารมณ์มีหลายวิธีการ เช่น *k*-Nearest Neighbour (*k*-NN) [8] ซึ่งเป็นวิธีการที่ง่ายที่สุด โดยการเปรียบเทียบค่าของข้อมูลที่อยู่ในกลุ่มข้อมูลอื่นกับข้อมูลที่สนใจว่ามีค่าใกล้เคียงกับกลุ่มข้อมูลใดมากที่สุด Linear Discriminant Analysis (LDA) [9] มีการจัดการแยกคุณลักษณะเด่นของข้อมูล ออกเป็นกลุ่ม ๆ แล้วทำให้ข้อมูลจากกลุ่มเดียวกันเข้าใกล้กันมากขึ้น และข้อมูลจากกลุ่มต่างกันจะอยู่ห่างกันมากที่สุด Support Vector Machine (SVM) [9] วิธีการนี้จะให้อินพุทที่ใช้ฝึกเป็นเวคเตอร์ในสเปซ *N* มิติ จากนั้นทำการสร้างไฮเปอร์เพลน (Hyperplane) เพื่อแยกกลุ่มของเวคเตอร์นำเข้า Decision tree [10] เป็นแผนผังที่แสดงความสัมพันธ์ระหว่างเงื่อนไขและลำดับของความสัมพันธ์ ตั้งแต่เงื่อนไขแรกจนถึงเงื่อนไขสุดท้าย ซึ่งช่วยในการตัดสินใจในการจำแนกกลุ่ม *k*-means clustering [10] เป็นการแบ่งกลุ่มโดยสุ่มจุดศูนย์กลางของข้อมูล (Centroid) ตามจำนวนกลุ่ม (*k*) ที่สนใจ จากนั้นจะแบ่งกลุ่มของข้อมูลตามจำนวนของจุดศูนย์กลางโดยที่มีผลรวมของค่าความแตกต่างของระยะทางของแต่ละกลุ่มที่น้อยที่สุด เป็นต้น

3. วิธีการดำเนินการวิจัย

กระบวนการรู้จำอารมณ์จากเสียงพูดโดยทั่วไปจะประกอบด้วยสองกระบวนการหลัก คือ การสกัดคุณลักษณะเด่น และการจำแนกคุณลักษณะเด่นของอารมณ์ คุณลักษณะเด่นที่ใช้ในงานวิจัยนี้ คือ MFCC อย่างไรก็ตาม



ภาพที่ 2: ตัวอย่างการทำงานของการรู้จำอารมณ์ของเสียงพูดโดยใช้วิธีจำแนกแบบสองระดับ

คุณลักษณะเด่น MFCC สำหรับการรู้จำอารมณ์ของเสียงพูดค่อนข้างมีความคล้ายคลึงกันค่อนข้างมาก ส่งผลให้การจำแนกอารมณ์มีความผิดพลาดได้ง่าย ดังนั้นจึงนำเสนอวิธีจำแนกแบบสองระดับสำหรับการรู้จำอารมณ์ของเสียงพูด ในงานวิจัยนี้จะแบ่งข้อมูลในฐานข้อมูลออกเป็นข้อมูลชุดฝึกฝน (Training) และข้อมูลชุดทดสอบ (Testing) ซึ่งคุณลักษณะเด่นของข้อมูลชุดฝึกฝน จะถูกแบ่งกลุ่มในระดับที่ 1 ด้วยวิธีการแบ่งกลุ่มแบบเคมีน (*k*-means) เพื่อแบ่งกลุ่มคุณลักษณะเด่นของอารมณ์ออกเป็นคลัสเตอร์ (Cluster) จากนั้นจะจำแนกคุณลักษณะเด่นระหว่างคุณลักษณะเด่นของข้อมูลชุดฝึกฝนและข้อมูลชุดสอบ ด้วยวิธีการหาระยะทางแบบยุคลิด (Euclidean distance) แต่หากพบว่าอารมณ์เด่นในแต่ละคลัสเตอร์มีมากกว่า 1 อารมณ์ คุณลักษณะเด่นจะถูกนำมาจำแนกอีกครั้งหนึ่งในการจำแนกระดับที่ 2 เฉพาะกลุ่มของอารมณ์ที่เกี่ยวข้อง ดังภาพที่ 1 วิธีจำแนกแบบสองระดับสำหรับการรู้จำอารมณ์ของเสียงพูด มีวิธีการทำงานดังนี้

ขั้นตอนที่ 1 สกัดคุณลักษณะเด่น (x_j) แบบ MFCC [3]

ขั้นตอนที่ 2 ทำการแบ่งกลุ่มคุณลักษณะเด่นด้วยวิธีการแบ่งกลุ่มแบบเคมีน ซึ่งแต่ละกลุ่มจะถูกเรียกว่าคลัสเตอร์ ผลลัพธ์จากขั้นตอนนี้ คือ คลัสเตอร์ของลำดับที่ i (C_i) และจุดศูนย์กลางของคลัสเตอร์ (M_i) ดังภาพที่ 2(ก)

ขั้นตอนที่ 3 หาเซตของอารมณ์ที่มีความถี่มากที่สุดหรือเรียกว่าอารมณ์เด่น E ของคลัสเตอร์ที่ i ซึ่งหาได้จากสมการที่ (1)

$$\square_{\square} = \arg\max_{\square} (\square_{\square,\square}) \quad (1)$$

โดยที่ $n_{i,e}$ คือ จำนวนของอารมณ์ e ในคลัสเตอร์ที่ i ซึ่งสามารถนับได้จากสมการที่ (2)

$$n_{i,e} = \text{count}(\{ \square_{\square} \mid \square_{\square} \in C_i \text{ and } \square_{\square} \in e \}) \quad (2)$$

ตารางที่ 2: ค่าความถูกต้องของการรู้จำอารมณ์ของเสียงพูด

วิธีการจำแนก	ค่าความถูกต้องของการรู้จำ (เปอร์เซ็นต์)						
	อารมณ์ โกรธ	อารมณ์ รังเกียจ	อารมณ์ กลัว	อารมณ์ ความสุข	อารมณ์ เศร้า	อารมณ์ ประหลาดใจ	ค่าเฉลี่ย
k-NN, $k = 1$	63.64	38.46	41.67	72.73	80.00	25.00	52.17
k-NN, $k = 3$	63.64	46.15	33.33	72.73	70.00	25.00	50.72
k-NN, $k = 5$	54.55	46.15	50.00	72.73	90.00	16.67	53.62
k-NN, $k = 7$	45.45	61.54	50.00	72.73	100.00	16.67	56.52
k-NN, $k = 9$	54.55	38.46	50.00	54.55	80.00	16.67	47.83
Decision Tree	36.36	46.15	25.00	45.45	30.00	0.00	30.43
LDA	63.64	15.38	8.33	27.27	20.00	0.00	21.74
SVM	63.64	53.85	25.00	27.27	60.00	33.33	43.48
TLC	**63.64**	53.85	41.67	**73.73**	**100.00**	**41.67**	**60.87**

โดยที่ *e* คือ อารมณ์ของคุณลักษณะเด่นดังตารางที่ 1

ขั้นตอนที่ 4 นับจำนวนอารมณ์ที่มีความถี่สูงสุด (D_i) ในแต่ละคลัสเตอร์ดังสมการที่ (3)

$$D_i = \text{count}(E_i) \tag{3}$$

ขั้นตอนที่ 5 กำหนดคุณลักษณะของคลัสเตอร์ โดยคลัสเตอร์ที่มีจำนวนอารมณ์ที่มีความถี่สูงสุด เพียงอารมณ์เดียว คลัสเตอร์นั้นจะถูกกำหนดเป็นคลัสเตอร์ที่มีอารมณ์ชัดเจน (C^1) ในขณะที่คลัสเตอร์ที่มีจำนวนอารมณ์ความถี่สูงสุดที่เท่ากันมากกว่า 1 อารมณ์ คลัสเตอร์นั้นจะถูกกำหนดเป็นคลัสเตอร์ที่มีอารมณ์คลุมเครือ (C^2) ดังแสดงในสมการที่ (4) และ (5) ตามลำดับ

$$C^1 = \{C_i \mid D_i \neq 1\} \tag{4}$$

$$C^2 = \{C_i \mid D_i > 1\} \tag{5}$$

ขั้นตอนที่ 6 จำแนกคุณลักษณะของอารมณ์ที่จะทดสอบ f_{test} ด้วยวิธีการจำแนกสองระดับ โดยเริ่มจากระดับที่หนึ่งนำ f_{test} มาเปรียบเทียบกับค่าเฉลี่ยของคุณลักษณะเด่นแต่ละคลัสเตอร์ ดังภาพที่ 2(ค) ถ้า f_{test} ใกล้กับค่าเฉลี่ยของ C^1 ในคลัสเตอร์ใดจะจำแนกว่าเป็นอารมณ์ของเสียงนั้น ในขณะที่ถ้า f_{test} ใกล้กับค่าเฉลี่ยของ C^2 คุณลักษณะเด่นจะถูกนำไปจำแนกด้วยวิธีการการวัดระยะห่างยุคลิดเฉพาะกลุ่มของอารมณ์ที่เกี่ยวข้อง ดังภาพที่ 2(ง) วิธีการจำแนกกลุ่มสองระดับสามารถอธิบายได้ดังสมการที่ (6)

$$\Box(f_{test}, M) = \begin{cases} e, & p(f_{test}, M) \Box C^1 \\ p(f_{test}, X_{C_i}), & p(f_{test}, M) \Box C^2 \end{cases} \tag{6}$$

โดยที่ X_{c_i} คือ คุณลักษณะเด่นของคลัสเตอร์ C ในลำดับที่ i และ $p(f_{test}, M)$ คือ ผลลัพธ์จากการจำแนกกลุ่มระหว่างคุณลักษณะเด่นของชุดทดสอบกับจุดศูนย์กลางของแต่ละคลัสเตอร์ ขณะที่ $p(f_{test}, X_{c_i})$ คือ ผลลัพธ์จากการจำแนกกลุ่มในระดับที่สองระหว่างคุณลักษณะเด่นกับคุณลักษณะเด่นเฉพาะอารมณ์ที่ชัดเจนในคลัสเตอร์ที่ i โดยอาศัยการหาระยะทางแบบยุคลิด

4. การทดลอง

การทดสอบประสิทธิภาพของการรู้จำอารมณ์ของเสียงพูดได้ใช้ฐานข้อมูลของ Ryerson Emotion Database (RED) [3] ประกอบด้วย 6 อารมณ์ ได้แก่ โกรธ (Anger) รังเกียจ (Disgust) กลัว (Fear) มีความสุข (Happy) เศร้า (Sadness) และประหลาดใจ (Surprise) ในการทดลองนี้เลือกใช้เฉพาะเสียงพูดภาษาอังกฤษเท่านั้น ข้อมูลที่ได้จะถูกแบ่งออกเป็นชุดข้อมูลฝึกฝน และชุดข้อมูลทดสอบโดยมีอัตราส่วนเป็น 50:50 โดยข้อมูลเสียงทั้งหมดจะถูกสกัดคุณลักษณะเด่นด้วย MFCC จากนั้นจึงทำการจำแนกกลุ่มของอารมณ์โดยใช้วิธีการจำแนกแบบสองระดับ (Two-level Classification: TLC) เปรียบเทียบกับวิธีการจำแนกแบบอื่น ได้แก่ วิธี *k*-NN โดยทดลองกำหนดค่า $k = \{1, 3, 5, 7, 9\}$, Decision Tree แบ่งโหนดโดยใช้ Gini's diversity index, ในวิธี LDA และ SVM ใช้ฟังก์ชั่นเส้นตรง (Linear kernel) ดังตารางที่ 2 ในการ

ทดลองทั้งหมดนี้ได้ถูกคำนวณโดยใช้โปรแกรม MATLAB ทำการทดลอง [11]

ผลการทดลองถูกแสดงไว้ในตารางที่ 2 พบว่าวิธี TLC มีความถูกต้องในการรู้จำอารมณ์จากเสียงพูดโดยเฉลี่ยเท่ากับ 60.87% ซึ่งสูงกว่าวิธีการจำแนกแบบอื่น ๆ เมื่อแยกพิจารณาในแต่ละอารมณ์พบว่าค่าความถูกต้องของการรู้จำด้วยวิธี TLC ของอารมณ์โกรธ อารมณ์มีความสุข อารมณ์เศร้า และอารมณ์ประหลาดใจ สูงกว่าหรือเท่ากับวิธีการจำแนกแบบอื่นที่นำมาเปรียบเทียบ เนื่องจากวิธีที่นำเสนอสามารถจำแนกคุณลักษณะเด่นของ MFCC ที่เป็นกลุ่มอารมณ์คลุมเครือได้

5. สรุปผล

งานวิจัยนี้นำเสนอวิธีจำแนกแบบสองระดับสำหรับการรู้จำอารมณ์ของเสียงพูด โดยในระดับที่ 1 ใช้การแบ่งกลุ่มแบบเคมีน เพื่อแยกคุณลักษณะเด่นของ MFCC ระหว่างกลุ่มของอารมณ์ ดังนั้นกลุ่มของอารมณ์ที่ชัดเจน จะสามารถแบ่งแยกได้อย่างถูกต้อง หลังจากนั้นกลุ่มของอารมณ์ที่คลุมเครือจะถูกจำแนกโดยใช้คุณลักษณะเด่นของ MFCC ภายในกลุ่มในระดับที่ 2 จากผลการทดลองพบว่าวิธีการที่นำเสนอให้ประสิทธิภาพในการจำแนกดีกว่าวิธีการที่นำมาเปรียบเทียบ

อย่างไรก็ตาม วิธีจำแนกแบบสองระดับจะเป็นการอิงถึงความถี่ของแต่ละกลุ่มข้อมูลของแต่ละคลัสเตอร์นั้น ซึ่งอาจจะมีความไม่เท่าเทียมกันของกลุ่มข้อมูลที่นำมาฝึกฝน ทำให้การจำแนกกลุ่มอาจจะมีความคลาดเคลื่อนได้

เอกสารอ้างอิง

[1] D. Ververidis and C. Kotropoulos, “Emotional speech recognition: resources, features, and methods,” *Speech Communication*, vol. 48, no. 9, pp. 1162–1181, 2006.

[2] Y. Yoshitomi, S.-I. Kim, T. Kawano, and T. Kilazoe, “Effect of sensor fusion for recognition of emotional states using voice, face image and thermal image of face,” *Proceedings 9th IEEE International Workshop on Robot and Human Interactive Communication. IEEE RO-MAN 2000* (*Cat. No.00TH8499*), pp. 178–183, 2000.

[3] Y. Wang, L. Guan, “Recognizing human emotion from audiovisual signals,” *IEEE Transactions on Multimedia*, Vol. 10, No. 5, pp. 936-946, 2008.

[4] Y. Wang, L. Guan and A.N. Venetsanopoulos, “Kernel based fusion with application to audiovisual emotion recognition,” *IEEE Trans. on Multimedia*, vol. 14, no. 3, pp. 597-607, 2012.

[5] P. Sukhummek, S. Kasuriya, T. Theeramunkong, C. Wutiwiwatchai, and H. Kunieda, “Feature selection experiments on emotional speech classification,” *12th International Conference on Electrical Engineering/-Electronics, Computer, Telecommunications and Information Technology* (*ECTI-CON*), pp. 1–4, 2015.

[6] M. A. Kammoun, D. Gargouri, M. Frikha and A. Ben Hamida, “Cepstral method evaluation in speech formant frequencies estimation,” *Industrial Technology, 2004. IEEE ICIT '04. 2004 IEEE International Conference on*, Vol. 3, pp. 1612-1616, 2004.

[7] A. Sengupta and R. Ghosh, “Modelling a voice activated speaker identification system using MFCC-Pitch-Formant Vector,” *Journal of The Institution of Engineers* (*India*) *Series BSP* (*March–May 2012*) *93*(*1*), pp. 51–56, 2012.

[8] V. L. Kaundanya, A. Patil and A. Panat, “Performance of k-NN classifier for emotion detection using EEG signals,” *2015 International Conference on Communications and Signal Processing* (*ICCSP*), Melmaruvathur, pp. 1160-1164, 2015.

[9] Y. Xie and T. Zhang, “A fault diagnosis approach using SVM with data dimension reduction by PCA and LDA method,” *2015 Chinese Automation Congress* (*CAC*), Wuhan, pp. 869-874, 2015.

[10] C. woo Kim, S. H. Ahn and T. Yoon, “Comparison of flavivirus using datamining-Apriori, K-means, and decision tree algorithm,” *2017 19th International Conference on Advanced Communication Technology* (*ICACT*), Bongpyeong, pp. 454-457, 2017.

[11] Introduction to Computer Programming With MATLAB Department of Phonetics and Linguistics, Univ. *College Landon* [*Online*]. *Available :www.phon.ucl.ac.uk/courses/spsci/matlab/, 2004*.

รหัสผ่านรูปแบบตารางตัวเลขสุ่ม
Random Number Gird Password

วิริยะ ลาวัลย์ปรีดา (Wiriya Lawanpreda)[1] และณัฐธนนท์ หงส์วริทธิ์ธร (Nuttanont Hongwarittorrn)[2]
ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยธรรมศาสตร์
[1]wlawanpreda@gmail.com, [2]nth@cs.tu.ac.th

บทคัดย่อ

ปัญหาที่น่าสนใจของงานวิจัยด้านการพิสูจน์ตัวตนคงไม่พ้นประสิทธิภาพทั้งในแง่การจำรหัสผ่านได้และความปลอดภัยของการพิสูจน์ตัวตน โดยปกติการพิสูจน์ตัวตนที่ประสิทธิภาพด้านการใช้งานจะทำให้ประสิทธิภาพด้านความปลอดภัยลดลง ในทางกลับกันการเพิ่มความปลอดภัยจะทำให้การใช้งานของผู้ใช้ยุ่งยากเช่นกัน สำหรับงานวิจัยนี้จำได้ออกแบบระบบที่ช่วยให้ผู้ใช้สามารถจำรหัสผ่านได้ง่ายทั้งคงความปลอดภัยที่สูงโดยการผสมรวมข้อดีของทั้งรหัสผ่านแบบรูปภาพและการกรอกรหัสผ่านอักษร ในส่วนของการลงทะเบียนผู้ใช้ต้องเลือกรูปภาพจากรูปภาพจำนวนหนึ่งแล้วกำหนดรหัสผ่าน โดยคาดว่ารูปภาพจะช่วยเตือนถึงรหัสผ่านเมื่อผู้ใช้เห็นรูปภาพ ในส่วนของการพิสูจน์ตัวตนผู้ใช้จะแปลงรหัสผ่านทีละตัวอักษรเป็นตัวเลข ตัวเลขจะแสดงไว้คู่ตัวอักษรอยู่บนภาพ เช่น สมมติให้ผู้ใช้ต้องการกรอกรหัสผ่าน book ผู้ใช้ต้องกรอก 7111139 หาก 7 11 และ 39 แทนด้วย b o และ k ตามลำดับ จากการศึกษานำร่องจากผู้เข้าร่วมการทดลอง 23 คน ได้ผลลัพธ์ในระดับที่ดี โดยระยะเวลาที่ใช้ในกระบวนการพิสูจน์ตัวตนโดยเฉลี่ยคือ 34.84 วินาที ซึ่งเมื่อคิดเวลาเฉลี่ยการใส่รหัสผ่านตั้งแต่ตัวอักษรตัวแรกจนครบทั้งหมดจะเป็น 23.91 วินาที ซึ่งคาดว่าจะสามารถลดระยะเวลาลง หากผู้ใช้งานคุ้นเคยกับวิธีการแล้ว

คำสำคัญ: การพิสูจน์ตัวตน รหัสผ่านแบบตาราง
ความปลอดภัย

Abstract

The main research problem of authentication systems is a new designed password which is easy to remember and hard to be attacked. Research has attempted to create new passwords which improve more usability and security. Generally, when the password is easy to remember for users, it will be easy to be attacked. Therefore, the challenge of research in this area to create a new password which can achieve higher usability and security. Graphical passwords have been studied and believed that they would be good alternatives that users could recall their passwords more easily than text passwords and could be more secured to be hacked also. The current research aimed to propose a new design password which was expected to be easier for users to recall the passwords and to be more difficult to be attacked. The proposed password was a kind of hybrid of letter and picture password. In the password registration phase, users would need to use a picture from a list, and input words as their passwords. In the login phase, users would need to input numeric letters randomly appeared on each character. For example, a user input "book" as his password. He would have to input 7111139 if 7, 11 and 39 were appeared on b, o, and k respectively. The pilot test with 23 participants produced a positive result. The average authentication phase took 34.84sec, and the time users input their passwords was 23.91sec. The participants were satisfied with the new proposed password even though they may not be familiar with. The usability measure should be improved when users get used to this password.

Keyword: Authentication, Gird Password, Security

1. บทนำ

การพิสูจน์ตัวตนโดยทั่วไปคือรหัสผ่านอักษรโดยการกรอกรหัสผ่านลงกล่องข้อความ ก่อนที่จะสามารถพิสูจน์ตัวตนจำเป็นต้องให้ผู้ใช้ตกลงกับระบบกำหนดรหัสผ่านที่ใช้ ไม่ว่าจะ

เป็นที่ผู้ใช้เป็นผู้สร้างขึ้นเองหรือระบบกำหนดให้ก็ตาม ข้อดีของวิธีการพิสูจน์ตัวตนรูปแบบนี้คือผู้ใช้สามารถเรียนรู้การใช้งานได้อย่างเร็ว แต่ประเด็นปัญหาที่ยังพบอยู่ไม่ว่าจะเป็นจากทางด้านมีผู้ไม่หวังดีโจมตี เช่น การแอบมองจากด้านหลัง การโจมตีจากการดักรหัสผ่านจากคีย์บอร์ด หรือปัญหาที่เกิดจากตัวผู้ใช้เอง [1] เช่น การลืมรหัสผ่าน การตั้งรหัสผ่านที่มีความปลอดภัยต่ำ ดังนั้นระบบจึงเพิ่มความปลอดภัยด้วยกำหนดกฎการตั้งรหัสผ่าน และไม่ควรการใช้รหัสผ่านที่มีความหมายหรือข้อมูลส่วนตัว แต่ผู้ใช้คนหนึ่งจำเป็นต้องใช้ระบบสารสนเทศหลายระบบมากหากผู้ใช้จำเป็นต้องตั้งรหัสผ่านที่สลับซับซ้อนเพื่อความปลอดภัยในระบบที่ใช้อยู่ทุกระบบจึงเป็นเรื่องยากที่จะให้ผู้ใช้จำรหัสผ่านได้ ผู้ใช้ส่วนใหญ่จึงมักใช้รหัสผ่านที่เป็นชุดเดียวกันหรือคล้ายคลึงกันในทุกระบบที่ใช้การพิสูจน์ตัวตน ซึ่งจะทำให้ความปลอดภัยของการพิสูจน์ตัวตนลดลง เพิ่มความเสี่ยงที่จะเกิดผลเสียกับระบบที่ใช้งานอยู่

การพิสูจน์ตัวตนโดยใช้รูปภาพเป็นทางเลือกที่ดีเนื่องจากผลการศึกษาทางจิตวิทยา พบว่าสมองของคนเราสามารถจำข้อมูลที่เป็นรูปภาพได้ดีกว่าข้อมูลที่เป็นข้อความ เพราะรูปภาพเป็นสิ่งที่น่าดึงดูดมากกว่าข้อความที่เป็นแค่ตัวอักษร รูปภาพสามารถสื่อแทนความหมายของผู้ใช้แต่ละคน คาดว่าการนำรูปภาพมาเป็นรหัสผ่านจะช่วยให้จดจำรหัสผ่านได้ดีกว่าการจดจำตัวอักษรหลายตัวและยิ่งมากขึ้นในกรณีการจดจำกลุ่มของตัวอักษรที่ไม่มีความหมาย รหัสผ่านรูปภาพได้ถูกพัฒนามาอย่างต่อเนื่องจึงมีรูปแบบที่แตกต่างและหลากหลาย ไม่ว่าจะเป็นการกดตำแหน่งในรูปภาพให้เหมือนตอนลงทะเบียน ลากรูปภาพไปยังช่องคำตอบ ซึ่งส่วนใหญ่จะมีประสิทธิภาพด้านการใช้งาน ผู้ใช้งานสามารถพิสูจน์ตัวตนโดยง่าย และง่ายต่อการจำได้ของรหัสผ่าน แต่ยังมีจุดอ่อนด้านความปลอดภัย อาทิ การโจมตีโดยการแอบมองจากด้านหลัง ซึ่งหากสามารถหาวิธีการพิสูจน์ตัวตนที่มีความสามารถที่ดีทั้งในด้านประสิทธิภาพการใช้งานช่วยให้ผู้ใช้งานสามารถพิสูจน์ตัวตนโดยง่าย ผู้ใช้จดจำรหัสผ่านได้ง่ายขึ้นและในด้านความปลอดภัยที่ระบบคอมพิวเตอร์จะมีระดับความปลอดภัยที่สูงขึ้น ผู้ใช้ย่อมรู้สึกปลอดภัยและมีความสุขใจในการใช้งานระบบคอมพิวเตอร์มากขึ้น

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ความหมายของการพิสูจน์ตัวตนคือการกระทำหรือกระบวนการเพื่อพิสูจน์ ตรวจสอบ ค้นหาว่าเป็นบุคคลที่ต้องการพิสูจน์ตัวตนหรือไม่ ซึ่งจะได้ผลลัพธ์ของกระบวนการคือใช่หรือไม่ใช่

Khetani และคณะ [2] นำเสนอการแบ่งหมวดหมู่ของกระบวนการพิสูจน์ตัวตน โดยลักษณะหนึ่งที่น่าสนใจคือ Knowledge Base โดยวิธีการพิสูจน์ตัวตนรูปแบบนี้คือ การพิสูจน์ตัวตนที่ใช้ข้อมูลเฉพาะหรือความรู้เฉพาะเพื่อแยกตัวบุคคลหมายถึงไม่ว่าจะเป็นบุคคลใดหากมีข้อมูลรหัสผ่านจะสามารถพิสูจน์ตัวตนได้สำเร็จ



ภาพที่ 1: ตัวอย่างการพิสูจน์ตัวตนของ Towhidi [3]

Towhidi [3] ได้นำเสนอวิธีการพิสูจน์ตัวตนโดยใช้การกรอกรหัสผ่านใส่กล่องข้อความแต่รหัสผ่านที่กรอกจะได้มาจากการเลือกรูปภาพชุดหนึ่งที่ตั้งไว้เมื่อตอนลงทะเบียนเพื่อแก้ปัญหาการแอบมองจากด้านข้าง โดยรูปภาพที่เลือกใช้เป็นรูปภาพใบหน้าบุคคลที่คาดว่าจะสามารถจำได้ดีกว่ารูปภาพทั่วไป ในการลงทะเบียนผู้เข้าร่วมการทดลองต้องเลือกรูปจำนวนหนึ่งเพื่อตั้งเป็นรหัสผ่านตามภาพที่ 1 ซึ่งอักษรใต้รูปภาพแต่ละรูปจะใช้วิธีสุ่มออกมาในแต่ละครั้งที่พิสูจน์ตัวตน ยังมี Tuscano [4] ได้พิสูจน์ตัวตนลักษณะนี้โดยได้ผลการศึกษาว่าถึงแม้ว่ารูปภาพที่แสดงจะถูกเบลอ ผู้ใช้สามารถพิสูจน์ตัวตนได้สำเร็จ ทั้งยังเพิ่มประสิทธิภาพความปลอดภัยในด้านการสุ่มรหัสผ่านจากผู้ไม่หวังดีที่รู้จักกับผู้ใช้ จะเห็นได้ว่าการพิสูจน์ตัวตนโดยใช้อักษรที่อยู่ข้างใต้รูปภาพที่เลือกไว้ในการลงทะเบียนนำอักษรไปกรอกลงกล่องข้อความเพื่อพิสูจน์ตัวตน โดยผลลัพธ์ที่ได้น่าสนใจไม่ว่าเรื่องประสิทธิภาพด้านความปลอดภัยให้ที่สูงขึ้น แต่ยังมีจุดอ่อนตรงผู้ไม่หวังดีสามารถคาด

เดารหัสผ่านได้โดยการสุ่มเพราะมีงานวิจัยได้วิธีว่าการเลือกรหัสผ่านรูปภาพที่เป็นใบหน้าบุคคลจะสามารถคาดเดาได้จากลักษณะของผู้ใช้งานหรือการเลือกรูปภาพที่มีจุดเด่นสำคัญช่วยให้จำง่าย เช่น การเลือกรูปภาพใบหน้าบุคคลที่สวมแว่น การเลือกใบหน้าบุคคลที่มีชนชาติเดียวกับผู้ใช้ เป็นต้น และที่น่าสนใจอีกจุดคือการพิสูจน์ตัวตนของงานวิจัยนี้จำเป็นต้องจำรูปภาพใบหน้าบุคคลถึงสี่ภาพเป็นอย่างน้อยจะมีสิทธิที่ผู้ใช้งานไม่สามารถจำรูปภาพได้เมื่อระยะเวลาผ่านไป หรือจำเป็นต้องใช้งานระบบเช่นนี้เป็นจำนวนมากเป็นการสร้างภาระในการจำของผู้ใช้งานเพิ่มขึ้น

Alese และคณะ [5] นำเสนอวิธีการพิสูจน์ตัวตนรูปแบบหนึ่ง โดยแนวคิดสำคัญมาจาก Fake Pointers คือข้อตกลงที่กำหนดไว้ให้รู้แค่ระหว่างตัวระบบกับผู้ใช้งาน ขั้นตอนการลงทะเบียนให้ผู้ใช้งานให้ผู้ใช้กำหนดรูปสัญลักษณ์แต่ละรูปกับหมายเลขระหว่าง 1 - 9 และจำนวนรหัสผ่านระหว่าง 3 - 6 ตัว สำหรับการพิสูจน์ตัวตนตามภาพที่ 2 ระบบจะทำการแสดงสัญลักษณ์ตามจำนวนรหัสผ่านที่เลือกไว้ ผู้ใช้ต้องกรอกตัวเลขที่กำกับต่อสัญลักษณ์โดยต้องเรียงตามลำดับ จะเห็นได้ว่าระบบจะมีความปลอดภัยไม่สามารถคาดเดารหัสผ่านแต่ละรอบได้ เนื่องจากการแสดงของสัญลักษณ์เป็นการสุ่ม และยังสามารถเปลี่ยนรูปแบบสัญลักษณ์เป็นรูปแบบที่ผู้ใช้คาดว่าจะเหมาะกับผู้ใช้



ภาพที่ 2: ตัวอย่างการพิสูจน์ตัวตนของ Alese [5]

Purushothaman [6] ออกแบบการพิสูจน์ตัวนรูปแบบหนึ่งเพื่อเน้นพัฒนาประสิทธิภาพความปลอดภัย โดยการลงทะเบียนด้วยการกรอกคำโดยใช้อักษร a - z และ 1 - 9 โดยไม่จำกัดตัวอักษร วิธีการพิสูจน์ตัวตนระบบจะแสดงตารางอักษรกับตัวเลขโดยแต่ละตัวจะแยกกันอยู่ในช่องซึ่งจะมีสีพื้นหลังที่แตกต่างกัน ผู้ใช้จะต้องทำการแปลงรหัสผ่านที่ตั้งไว้ทีละตัวอักษรเป็นอักษรตัวแรกของสีพื้นหลังใต้ตัวอักษร

จากงานวิจัยที่ผ่านมาจะได้เห็นว่าทั้งหมดได้ให้ความสำคัญกับความปลอดภัย หรือเปลี่ยนจากรูปแบบการจำรูปภาพที่ผู้ใช้เลือกไว้เมื่อตอนลงทะเบียนเป็นวิธีการรูปแบบอื่นไม่ว่าจะเนื่องจากปัญหาเรื่องที่ต้องใช้พื้นที่ในการแสดงผลรูปภาพที่คู่กับตัวอักษรต้องใช้พื้นที่ แต่อาจเป็นการเพิ่มความยุ่งยากต่อการใช้งาน สุดท้ายเห็นได้ว่าแต่ละวิธียังคงไม่สามารถสร้างความเหมาะสมได้ระหว่างประสิทธิภาพในด้านความปลอดภัยและประสิทธิภาพในด้านการใช้งาน ทางผู้วิจัยเห็นถึงความสำคัญของการพิสูจน์ตัวตนจึงต้องการหาวิธีการที่สามารถทำให้ผู้ใช้งานที่ง่ายต่อการพิสูจน์ตัวตนและต้องช่วยให้ผู้ใช้ลงภาระในการจำรหัสผ่านที่จำเป็นต้องใช้งานในระบบจำนวนมาก ทั้งยังมีประสิทธิภาพทางด้านความปลอดภัย

3. ระเบียบวิธีการศึกษา

3.1 การออกแบบการทดลอง

ส่วนใหญ่ระบบที่ออกแบบเป็นรหัสผ่านรูปภาพ รหัสผ่านที่ผู้ใช้งานต้องจำจะเป็นรูปภาพ ไม่ว่าจะเป็นตำแหน่งของรูปภาพที่เลือกไว้หรือชุดรูปภาพหนึ่งชุดแต่ต่างจากรหัสผ่านแบบกล่องข้อความที่ผู้ใช้จำเป็นต้องจำรหัสผ่านเป็นชุดอักษรและอักขระหนึ่งชุด ซึ่งจากงานวิจัยที่ผ่านมาจะเห็นได้ว่าการจำรูปภาพมีประสิทธิภาพที่ดีกว่าชุดอักษรและอักขระแต่ในงานวิจัยนี้สนใจว่าหากเราใช้รูปภาพเพื่อเป็นตัวช่วยเตือนความจำไปยังรหัสผ่านตัวอักษรจะช่วยลดภาระในการจำรหัสผ่าน

จากการสำรวจเบื้องต้นมีผู้เข้าร่วมการทดลอง 20 คน เพื่อทดสอบว่า ผู้ใช้เมื่อเห็นภาพเดิมที่ตนเองเคยมาก่อนแล้วจะทำให้นึกถึงคำเดิมที่หรือไม่ ทำการสำรวจโดยให้ผู้เข้าร่วมการทดลองเลือกรูปภาพ 1 รูปจากรูปภาพทั้งหมด 4 รูป จากนั้นให้บอกคำที่มีความหมายสำหรับตนเองกับรูปภาพที่เลือก ทิ้งระยะเวลาจากนั้น 1 สัปดาห์ นำรูปภาพที่ผู้เข้าร่วมการทดลองเลือกไว้ไปให้ดูจะสามารถจำได้ถึงคำที่เคยระบุไว้หรือไม่ ซึ่งผลลัพธ์ที่ได้ทุกคนสามารถนึกถึงคำที่เคยระบุไว้ได้ มีประเด็นที่น่าสนใจคือในรูปภาพที่เหมือนกันคำที่ผู้เข้าร่วมการทดลองแต่ละคนเลือกจะแตกต่างกันไปตามประสบการณ์ของผู้เข้าร่วมการทดลอง จากแนวคิดดังกล่าวทางผู้วิจัยได้ออกแบบรูปแบบ

การพิสูจน์ตัวตนที่แตกต่างเพื่อช่วยทั้งลดภาระในการจำรหัสผ่านและยังคงมีประสิทธิภาพในด้านความปลอดภัย

3.2 เครื่องมือที่ใช้

ในงานวิจัยได้ใช้ภาษา HTML และ Javascript ซึ่งจะถูกแสดงผลผ่านบราวเซอร์ ทางด้านเซิฟเวอร์ใช้ Firebase Hosting โดยการประมวลผลทั้งหมดจะใช้ภาษา Javascript และในส่วนของฐานข้อมูลใช้ Firebase Realtime Database ที่เป็น NoSQL

3.3 การเก็บรวบรวมข้อมูล

งานวิจัยนี้เป็นการศึกษาออกแบบระบบการพิสูจน์ตัวตน โดยขณะผู้เข้าร่วมการทดลองทำการพิสูจน์ตัว ระบบจะมีการเก็บข้อมูลการใช้งาน และแบบสอบถามเก็บข้อมูลความพึงพอใจของผู้เข้าร่วมการทดลอง

3.4 วิธีการทดลอง

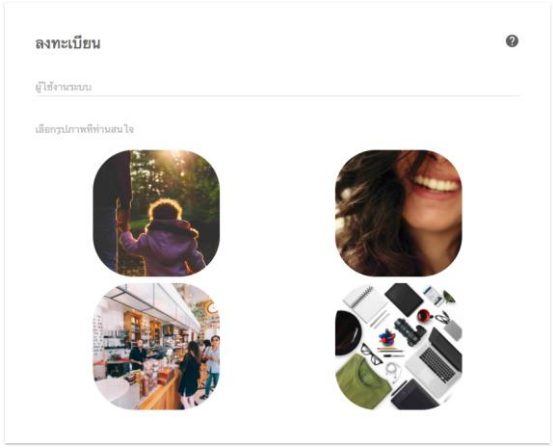
ให้ผู้เข้าร่วมการทดลองเข้าใช้งานระบบผ่านบราวเซอร์ โดยการทดลองถูกแบ่งออกเป็น 5 ขั้นตอน คือ 1. ส่วนหน้าเริ่มต้นการทดลองจะมีรายละเอียดเพื่อให้ผู้เข้าร่วมการทดลองทราบถึงวัตถุประสงค์และจะมีคู่มืออธิบายการใช้งานระบบเบื้องต้น 2. ส่วนการลงทะเบียน ผู้เข้าร่วมการทดลองจะทำการสร้างชื่อผู้ใช้งาน เลือกรูปภาพและรหัสผ่าน 3. ผู้เข้าร่วมการทดลองจะทดลองทำการพิสูจน์ตัวตนเป็นการทำให้คุ้นเคยกับวิธีการและคุ้นกับการแปลงรหัสผ่านจากตัวอักษรเป็นตัวเลข 4. ผู้เข้าร่วมการทดลองจะทำการพิสูจน์ตัวตน ในครั้งนี้ระบบจะทำการจับเวลา 5. เมื่อผู้เข้าร่วมการทดลองทำการพิสูจน์ตัวตนสำเร็จจะมีแบบสอบถามเกี่ยวกับความพึงพอใจของผู้ใช้งาน

3.4.1 ขั้นตอนการลงทะเบียน

เริ่มที่ผู้ใช้งานกรอกชื่อผู้ใช้งานในกล่องข้อความ ถัดมาเลือกรูปภาพ 1 รูป ตามภาพที่ 3 เมื่อเลือกรูปภาพสำเร็จจะมีกล่องข้อความสำหรับกรอกรหัสผ่าน โดยรหัสผ่านจะต้องเป็นตัวอักษรภาษาอังกฤษเท่านั้น ทางผู้วิจัยคาดหวังว่ารูปภาพที่เลือกจะสอดคล้องหรือมีความเชื่อมโยงกับรหัสผ่าน เมื่อผู้ใช้รูปภาพจะสามารถนึกถึงรหัสผ่านได้ภายหลัง

3.4.2 ขั้นตอนการพิสูจน์ตัวตน

เริ่มที่ผู้ใช้กรอกชื่อผู้ใช้งานที่เคยลงทะเบียนไว้ แล้วกดที่ปุ่มด้านข้างเพื่อดึงข้อมูลมาแสดงผล จากนั้น ระบบจะดึงรูปภาพที่ได้เลือกไว้ มาแสดงเป็นพื้นหลังโดยคาดหวังช่วยเตือนความทรงจำสำหรับรหัสผ่านของผู้ใช้ คู่กับตัวอักษรและตัวเลข 1 ชุด โดยใน 1 ชุดนี้จะเป็น ตัวอักษรภาษาอังกฤษ a - z เรียงตามลำดับ และมีตัวเลขแสดงคู่กับตัวอักษรแต่ละตัวตามภาพที่ 4 ตัวเลขดังกล่าวจะถูกสุ่มขึ้นใหม่ทุกครั้งที่ดึงข้อมูลรูปภาพมาแสดงผล ซึ่งผู้ใช้งานต้องทำการแปลงรหัสผ่านที่ตั้งไว้เป็นตัวเลขโดยค้นหาตัวอักษรที่เป็นรหัสผ่านแล้วนำตัวเลขข้างใต้ตัวอักษรมาใช้กรอกเป็นรหัสผ่านลงในกล่องข้อความทีละตัวจนครบ



ภาพที่ 3: หน้าลงทะเบียน

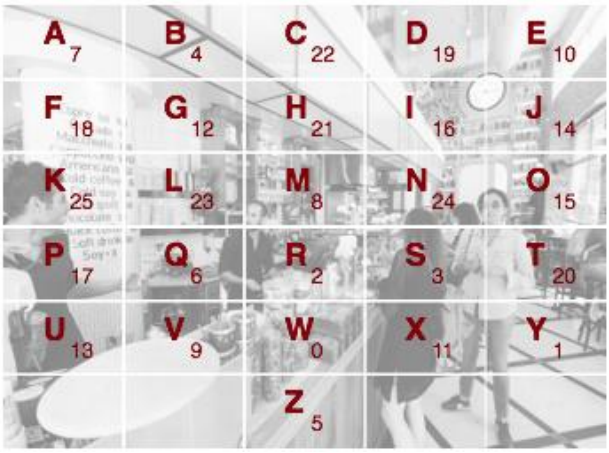


ภาพที่ 4: หน้าการพิสูจน์ตัวตน

3.4.3 ตัวอย่างการพิสูจน์ตัวตน

สมมติให้ผู้ใช้งานทำส่วนการลงทะเบียนสำเร็จแล้ว โดยใช้ชื่อผู้ใช้งานคือ Scenario ได้เลือกรูปภาพ และรหัสผ่านเป็นคำว่า iberry ในขั้นตอนการพิสูจน์ตัวตน ระบบจะแสดงรูปภาพที่ผู้ใช้เลือกไว้ในขั้นตอนลงทะเบียน ซึ่งภายในรูปจะมีอักษร a - z และแต่ละตัวอักษรจะมีตัวเลขกำกับอยู่ดังภาพที่ 5 ผู้ใช้จะแปลง

รหัสผ่านทีละตัวอักษรเป็นตัวเลขข้างใต้ตัวอักษรจนครบ สมมติรหัสผ่านเป็นคำว่า iberry ผู้ใช้จะกรอกตัวเลขเป็น 16410221 ในช่องรหัสผ่าน



ภาพที่ 5: ตัวอย่างรูปภาพและตัวอักษรคู่ตัวเลขที่ระบบสร้างขึ้น

4. ผลการทดลอง

จากผู้เข้าร่วมการทดลองทั้งหมด 23 คน ได้ผลการทดลองดังต่อไปซึ่งจะแบ่งออกเป็น 3 รูปแบบ ได้แก่ ผลการทดลองของการพิสูจน์ตัวตนของผู้เข้าร่วมการทดลอง ผลการประเมินความพึงพอใจของผู้เข้าร่วมการทดลอง และการวิเคราะห์ประสิทธิภาพทางด้านความปลอดภัย

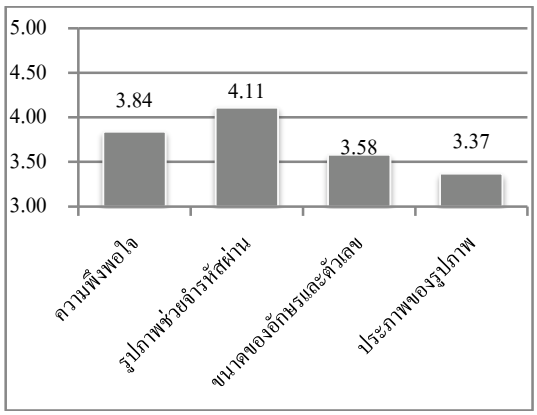
4.1 ผลการทดลองของการพิสูจน์ตัวตนของผู้เข้าร่วมการทดลอง

ผู้เข้าร่วมการทดลองทุกคนสามารถพิสูจน์ตัวตนสำเร็จ โดยสามารถพิสูจน์ตัวตนสำเร็จในครั้งแรก 91.30% และ 8.70% สามารถยืนยันตัวสำเร็จในครั้งที่ 2 จากข้อสังเกตที่ไม่สามารถพิสูจน์ตัวตนสำเร็จในครั้งแรกเนื่องจากไม่คุ้นเคยในการแปลงรหัสผ่าน

ค่าเฉลี่ยจำนวนตัวอักษรที่ใช้ตั้งเป็นรหัสผ่านคือ 6.04 ตัวอักษร เวลาเฉลี่ยที่ผู้เข้าร่วมการทดลอง(ตั้งแต่ระบบแสดงรูปภาพ) คือ 34.84 วินาที และเมื่อคิดแค่เวลาเฉลี่ยการใส่รหัสผ่านตั้งแต่ตัวอักษรตัวแรกจนครบทั้งหมด คือ 23.91 วินาที ซึ่งคาดว่าระยะเวลาจะลดลงเมื่อผู้ใช้งานคุ้นเคยกับวิธีการพิสูจน์ตัวตนรูปแบบนี้มากขึ้น โดยทั้งระยะเวลาและความปลอดภัยยังคงแปรผันตามจำนวนของการตั้งรหัสผ่าน เช่นเดียวกับวิธีการพิสูจน์ตัวตนรูปแบบอื่นอยู่

4.2 ผลการประเมินความพึงพอใจของผู้เข้าร่วมการทดลอง

ผลการวิเคาะห์ข้อมูลจากแบบสอบถามมีกว่า 76.70% ของผู้เข้าร่วมการทดลองที่คาดว่าจะสามารถจำรหัสผ่านได้ เมื่อเห็นรูปภาพที่เลือกไว้อีกครั้ง แต่มีแค่ผู้เข้าร่วมการทดลอง 58.69% ที่ตั้งรหัสผ่านมีความเชื่อมโยงกับรูปภาพที่เลือกไว้



ภาพที่ 10: สรุปคะแนนแบบสอบถามของผู้เข้าร่วมการทดลอง

จากภาพที่ 10 เห็นได้ว่าคะแนนเฉลี่ยความพึงพอใจของผู้ใช้งานอยู่ในระดับที่ดี โดยส่วนความพึงพอใจต่อวิธีการใช้งาน เนื่องจากเป็นวิธีการใหม่ต้องใช้เวลาเพื่อเพิ่มความคุ้นเคยด้านความสามารถในการจำรหัสผ่านได้ผู้ทดสอบส่วนใหญ่ก็คาดว่าเมื่อเห็นรูปภาพจะสามารถระลึกถึงรหัสผ่านได้ถือเป็นการช่วยลดภาระในการจำรหัสผ่านของผู้ใช้ ซึ่งในส่วนขนาดตัวอักษรและตัวเลขและประเภทของรูปภาพที่อยู่ในระดับคะแนนที่น้อยกว่า เพราะทางผู้วิจัยพยายามออกแบบระบบที่สามารถใช้งานได้ในทุกอุปกรณ์ จึงกำหนดตัวอักษรและตัวเลขมาในขนาดที่พอมองเห็นได้ไม่ได้ใหญ่ชัดเจนมากนัก และเนื่องจากเป็นการศึกษานำร่องจึงไม่ได้มีรูปภาพให้เลือกมากนัก

4.3 การวิเคราะห์ประสิทธิภาพด้านความปลอดภัย

ในการวิเคราะห์ประสิทธิภาพด้านความปลอดภัยนี้จะวิเคราะห์แยกตามเทคนิคที่นิยมใช้ในการโจมตีการพิสูจน์ตัวตน

Brute Force เป็นการสุ่มรหัสผ่านไล่ไปเรื่อยๆจนกว่าจะโจมตีสำเร็จ โดยวิธีการทั่วไปโอกาสจะเพิ่มขึ้นเสมอเนื่องจากรหัสผ่านจะเป็นชุดเดิมเสมอ แต่สำหรับวิธีการพิสูจน์ตัวตนของงานวิจัยนี้ที่รหัสผ่านจะเปลี่ยนไปในทุกครั้งจะเป็นเรื่องยากขึ้นสำหรับการโจมตี โอกาสในการโจมตีกรณีสุ่มตัวเลข จากอักษร 1 ตัวจะมีโอกาสถูกแปลงเป็นตัวเลขระหว่าง 0 - 99 แสดง

ว่าการตั้งรหัสผ่าน n ตัว ความน่าจะเป็นที่สุ่มถูกอยู่ที่ 100^n จะเห็นได้ว่าจำนวนตัวอักษรของรหัสผ่านยังส่งเพิ่มระดับความน่าจะเป็นในการคาดเดาหรือระดับความปลอดภัยอย่างชัดเจน

Keystroke เป็นการโจมตีที่จะทำการดักข้อมูลการกดคีย์บอร์ดซึ่งจะไม่สามารถใช้การกับวิธีการพิสูจน์ตัวตนของงานวิจัยนี้ได้ เนื่องจากรหัสผ่านที่ต้องกรอกจะไม่เหมือนเดิมในทุกๆรอบเพราะผู้ใช้จะแปลงรหัสผ่านเป็นตัวเลขที่ถูกสุ่มขึ้นมา

Shoulder Surfing วิธีการแอบมองจากด้านข้าง เมื่อเทียบเฉพาะการแอบมองการพิมพ์ข้อความ 1 ชุด กับตัวเลข 1 ชุด หากไม่ใช่คำที่คุ้นเคย ผู้พิมพ์จะพิมพ์ข้อความได้ช้าอาจทำให้ผู้ไม่หวังดีแอบมองได้ง่ายขึ้น แต่ต่างกับการพิมพ์ตัวเลขที่มีบริเวณแป้นพิมพ์แคบทั้งสามารถพิมพ์ได้รวดเร็วและป้องกันการแอบมาง่ายกว่า โดยสำหรับรหัสผ่านรูปแบบนี้จำเป็นต้องแอบมองทั้งหน้าจอผู้ใช้และการกดคีย์บอร์ดเนื่องจากรหัสผ่านที่ผู้ใช้กรอกไม่สามารถใช้ได้ในครั้งถัดไป บุคคลไม่หวังดีต้องแปลงชุดตัวเลขที่ได้กลับไปตัวอักษร แต่ด้วยเนื่องจากการสุ่มตัวเลขระหว่าง 0 - 99 มีโอกาสตัวเลขสองหลักจะต้องแปลงเป็นอักษรตัวเดียวหรือสองตัวซึ่งเพิ่มความยุ่งยากต่อผู้บุคคลไม่หวังดีอย่างมาก

Dictionary และ Guessing ทั้งสองวิธีนี้จะไม่สามารถโจมตีรหัสผ่านรูปแบบนี้ได้โดยตรงเนื่องจากต้องการวิธีการแปลงคำเป็นตัวเลขก่อนแต่มีโอกาสที่บุคคลที่ไม่หวังดีเดาสุ่มคำที่เป็นรหัสผ่านได้ โดยเฉพาะบุคคลที่ไม่หวังดีที่คุ้นเคยกับผู้ใช้เป็นผู้โจมตีเอง

โดยรวมแล้ววิธีการพิสูจน์ตัวตนของงานวิจัยนี้ได้พัฒนาในด้านความปลอดภัยขึ้นในหลายๆทาง ยิ่งวิธีการโจมตีที่ใช้คอมพิวเตอร์เป็นตัวคำนวณหารหัสผ่านจะมีความปลอดภัยมากขึ้นเมื่อเทียบกับวิธีอื่น

5. สรุปผลการศึกษา

งานวิจัยนี้เป็นการศึกษานำร่องเพื่อศึกษาออกแบบการพิสูจน์ตัวตนโดยวิธีการพิสูจน์ตัวตนซึ่งมีการใช้รูปภาพและการพิมพ์ใช้ตัวตัวเลขซึ่งเป็นรหัสปรากฏตัวอักษรตามรหัสผ่านของผู้ใช้ เป็นการปรับปรุงเสนอแนวการสร้างรหัสที่เพิ่มประสิทธิภาพทั้งแง่การใช้งานและความปลอดภัย และจากการทดลองในเบื้องต้นพบว่า ผู้ใช้ส่วนใหญ่พอใจในการใช้งานการพิสูจน์ตัวตนรูปแบบนี้ถึงแม้อาจจะยังไม่ค่อยคุ้นเคยเนื่องจากจะต้องแปลงรหัสผ่านเป็นตัวเลขเพื่อกรอกเป็นข้อมูลเพื่อการพิสูจน์ตัวตน และจากการวิเคราะห์ด้านความปลอดภัย อาจจะคาดหวังว่าได้ว่ารหัสผ่านรูปแบบนี้จะมีระดับความปลอดภัยเพิ่มขึ้น แต่ยังคงต้องทำการทดลองเพิ่มเติมทั้งในจำนวนผู้เข้าร่วมการทดลองที่มากขึ้นและการทดลองที่กำหนดให้มีผู้ไม่หวังดีแอบมองการกรอกรหัสผ่าน ตลอดจนการทดลองที่ทิ้งระยะเวลาเพื่อทดสอบว่าผู้ใช้จะยังจดจำรหัสผ่านได้มากน้อยแค่ไหนเมื่อมีรูปภาพเป็นสิ่งช่วยเตือนความจำ (Retention Over Time)

เอกสารอ้างอิง

[1] A. Adams and M. A. Sasse, “Users are not the enemy”, *Communications of the ACM*, vol. 42, no. 12, pp. 40-46, 1999.

[2] V. Khetani , J. Nicholas , A. Bongirwar and A. Yeole, “Securing Web Accounts Using Graphical Password Authentication through Watermarking”, *International Journal of Computer Trends and Technology*, vol. 9, no. 6, pp. 269-274, 2015.

[3] F. Towhidi, “An Enhancement on Passface Graphical Password Authentication”, *Journal of Computer Science and Technology*, vol. 26, no. 6, pp. 988–999, 2013.

[4] G. Tuscano, “Graphical password authentication using Pass faces”, *Journal of Engineering Research and Applications*, vol. 5, no. 3, pp. 60-64, 2015.

[5] B. K. Alese, A. A. Omojowo, A. T. Adesuyi, A F. Thompson, O. S. Adewale, and F. O. Osuolale, “An Enhanced Graphical Password Technique Using Fake Pointers”, *Proceedings of Informing Science & IT Education Conference*, pp. 79-89, 2015.

[6] G. R. Purushothaman and K. Ashwini, “A Novel Two Step Random Colored Grid Process: Graphical Password Authentication System”, *International Journal of Computer Networks and Communications Security*, vol. 4, no. 2, pp. 52-55, 2016.

การใช้การเรียนรู้แบบเชิงลึกสำหรับการรู้จำตัวอักษรพิมพ์ไทย
Using Deep Learning for Printed Thai Characters Recognition

ปองคุณ ปานเกิด (Pongkun Pankerd)[1] ตวงสิทธิ์ ธิราบุตร (Tuangsit Tirabutr)[2] ทศพล ทำดวง (Thotsapon Tamduang)[3]
และสุนีย์ พงษ์พินิจภิญโญ (Sunee Pongpinigpinyo)[4]*
ภาควิชาคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยศิลปากร วิทยาเขตพระราชวังสนามจันทร์
[1]Pankerd_P@su.ac.th, [2]Tirabutr_T@su.ac.th, [3]Tamduang_T@su.ac.th, [4]Pongpinigpinyo_S@su.ac.th

บทคัดย่อ

การเรียนรู้แบบเชิงลึกเป็นการเรียนรู้ของเครื่องที่มีการพัฒนาล่าสุดที่ให้ผลลัพธ์ความถูกต้องแม่นยำมากที่สุดในด้านการรู้จำภาพ การรู้จำตัวอักษรเป็นประเภทหนึ่งของการรู้จำภาพเนื่องจากตัวอักษรที่เป็นอินพุทถูกแปลงด้วยเครื่องแสกนเนอร์หรือกล้องถ่ายรูป ผลของการแปลงจะได้เป็นรูปภาพตัวอักษร งานวิจัยนี้นำสนอโครงข่ายประสาทเทียมแบบคอนโวลูชันซึ่งเป็นการเรียนรู้ของเครื่องประเภทหนึ่งสำหรับใช้ในการรู้จำตัวอักษรพิมพ์ไทยที่ประกอบด้วย 44 ตัวอักษร และรูปแบบอื่นๆ ที่ไม่ใช่ตัวอักษร 21 รูปแบบ โครงข่ายประสาทเทียมแบบคอนโวลูชันสามารถเรียนรู้ชุดคุณลักษณะเฉพาะของตัวอักษรที่เป็นรูปภาพโดยอัตโนมัติ TensorFlow ถูกนำมาประยุกต์ใช้สำหรับการเรียนรู้คุณลักษณะ ผลการทดลองแสดงให้เห็นว่าโครงข่ายประสาทเทียมแบบคอนโวลูชันประสบความสำเร็จที่สามารถจำแนกตัวอักษรพิมพ์ไทยได้อย่างถูกต้อง 99.84%

คำสำคัญ: การเรียนรู้แบบเชิงลึก, การรู้จำตัวอักษรพิมพ์ไทย, โครงข่ายประสาทเทียมแบบคอนโวลูชัน, เทนเซอร์โฟลว

Abstract

Deep learning is the latest development of machine learning which has achieved the most accuracy results in the field of image recognition. Character recognition is a type of image recognition since input characters are digitized by an optical scanner or camera. The resulting of digitize is the character images. This paper presents convolutional neural networks (CNN) as the type of machine learning are used for recognizing printed Thai characters which has 44 characters and 21 symbols which are non-characters. CNN can automatically learn a unique set of character image features. TensorFlow is applied for learning features. The experimental results show that CNN are able to achieve printed Thai character which has recognition rate of 99.84%.

Keyword: deep learning, printed Thai character recognition, convolutional neural networks, TensorFlow

1. บทนำ

การรู้จำภาพตัวอักษรพิมพ์ภาษาไทย (printed Thai character recognition)ได้มีการศึกษาวิจัยอย่างต่อเนื่องเพื่อปรับปรุงประสิทธิภาพการรู้จำภาพตัวอักษรให้มีความแม่นยำมากที่สุด โดยมีวัตถุประสงค์เพื่อนำไปประยุกต์ใช้ให้เกิดความถูกต้องมากที่สุดของการแปลงข้อมูลเอกสารกระดาษเข้าสู่ระบบคอมพิวเตอร์ด้วยการจับภาพ เอกสารต่างๆ ผ่านเครื่องสแกนเนอร์หรือกล้องดิจิทัล ซึ่งภาพที่ได้จะมีการจัดเก็บไว้ในรูปไฟล์ BMP JPG หรือ PDF เมื่อนำไฟล์ภาพที่ได้เข้าสู่ระบบโปรแกรมจะทำการแปลงไฟล์ภาพนั้นๆ เป็นข้อความที่ผู้ใช้สามารถแก้ไขได้เช่นเดียวกับไฟล์เอกสารในระบบ คอมพิวเตอร์ทั่วไป ทำให้การแปลงเอกสารจากต้นฉบับกระดาษเป็นข้อมูลดิจิทัลทำได้อย่างมี ประสิทธิภาพและรวดเร็วมากยิ่งขึ้น เป็นการลดขั้นตอนการพิมพ์เอกสารใหม่

งานวิจัยรู้จำภาพตัวอักษรพิมพ์ภาษาไทยใน [1] ได้นำเสนอโดยใช้เทคนิคด้านการวิเคราะห์ตัวประกอบสำคัญและโครงข่ายประสาทเทียม โดยนำข้อมูลภาพตัวอักษรที่ได้จากการพิมพ์ด้วยเครื่องพิมพ์เลเซอร์ที่ความละเอียด 600 จุดต่อนิ้ว และนำเอกสารให้ถูกอ่านโดยผ่านเครื่องสแกนเนอร์ที่มีความละเอียด 200 จุด

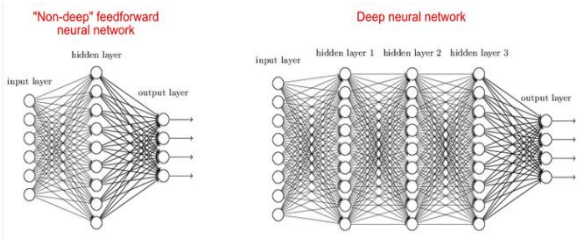
ต่อนิ้วจำนวน 3,264 ตัวอักษร ซึ่งประกอบด้วยตัวอักษรแบบ AngsanaUPC, BrowaliaUPC, CordiaUPC, DilleniaUPC, EucrosiaUPC และ FreesiaUPC แต่ละแบบประกอบด้วยตัวอักษรขนาด 16, 18, 20, 22, 24, 26, 28 และ 36 จุดเป็นตัวอักษรต้นแบบ งานวิจัย [2] ใช้โปรเจ็กชั่นฟีเจอร์และฟัซซี่นิวรอลในการรู้จำอักษรตัวพิมพ์ภาษาไทย ตัวอักษรที่ปรับให้มีขนาดมาตรฐานจะอยู่ภายในกรอบ 64*64 จุดภาพ จากนั้นจึงทำการโปรเจ็คชั่นภาพตัวอักษรตามแนวแกนตั้งและแกนนอน หาค่ามากที่สุดของจำนวนจุดภาพแต่ละช่วง นำค่าที่ได้ในแต่ละช่วงนั้นไปจัดกลุ่มอยู่ในช่วงต่างๆ ของฟัซซี่เซตด้วยวิธีการของฟัซซี่ลอจิก โดยการหาความสัมพันธ์ของค่าสูงสุดที่ได้ในแต่ละช่วงของตัวอักษรต้นแบบ เพื่อสร้างเป็นกฎของฟัซซี่ในการรู้จำตัวอักษร แล้วนำเอาโครงข่ายเซลล์นิวรอลมาเรียนรู้กฎการรู้จำแบบฟัซซี่อักษรตัวพิมพ์ภาษาไทยเพื่อการรู้จำ ซึ่งแตกต่างจาก [3] ที่ใช้วิธีการต่อขยายรหัส, การใช้คะแนนนิยมสูงสุด และใช้วิธีการรู้จำทดสอบกับตัวอักษร นอกจากนี้ใน [4] ได้นำเสนอการเพิ่มประสิทธิภาพของระบบรู้จำตัวอักษรภาษาไทยโดยการซ่อมแซมตัวอักษรไม่สมบูรณ์โดยวิเคราะห์จุดเชื่อมต่อของตัวอักษรจากโครงร่างของตัวอักษร ซึ่งสังเคราะห์ด้วยวิธี k-Segment Principal Curve และขั้นตอนที่สี่ ปรับปรุงคุณภาพของตัวอักษรภายหลังการซ่อมแซม โดยวิธี Morphological Analysis เพื่อปรับแต่งขอบของตัวอักษรและเติมช่องว่างภายในตัวอักษรเพื่อให้คุณภาพของตัวอักษรมีความสมบูรณ์มากขึ้น ในงานวิจัย [5] นำเสนอวิธีการเรียนรู้ตัวอักษรไทยที่ไม่มีหัว เนื่องจากคุณลักษณะของอักษรไทยจะมีองค์ประกอบต่างๆ สำหรับตัวอักษรที่แตกต่างกัน เช่น หัวกลม, เส้นซิกแซก, เส้นโค้ง, โดยใช้วิธีโครงข่ายประสาทเทียม และสหสัมพันธ์แบบไข้ว (Hierarchical Cross Correlation) ซึ่งทดลองกับจำนวนตัวอักษร 12 ตัวอักษรที่ไม่มีหัวกลม งานวิจัย [6] นำเสนออัลกอริทึม Heuristic Incremental Integer Programming ที่ประยุกต์จากแนวความคิด Integer Programming และใช้คลังข้อมูลพจนานุกรมภาษาไทยเพิ่มประสิทธิภาพของวิธีการเรียนรู้ตัวอักษรไทยที่ไม่ชัดเจน ตัวอักษรขาดหายที่ปรากฏบนไฟล์เอกสารทางประวัติศาสตร์ไทยและเอกสารจากการส่งโทรสาร โดยสร้างเป็นตัวแบบเซ็ตพาร์ทิชัน (set-partitioning) แล้วใช้เทคนิคการค้นหาพาร์ทิชันชุดที่ดีที่สุดของตัวอักษรที่มีความชัดเจนมากที่สุดสร้างเป็นตัวอักษรไทยขึ้นใหม่ที่มีความถูกต้องมากกว่า 93%

ในงานวิจัยนี้นำเสนอการรู้จำภาพตัวอักษรพิมพ์ภาษาไทยรูปสระและรูปวรรณยุกต์ภาษาไทย โดยใช้โครงข่ายประสาทเทียมเชิงลึก (Deep Learning Neural Network) ที่เป็นแบบการเรียนรู้เชิงลึก (deep learning) ซึ่งเป็นรูปแบบการเรียนรู้ของเครื่องที่เป็นที่ยอมรับในงานวิจัยที่แสดงให้เห็นว่ามีประสิทธิภาพมากที่สุดในการแก้รู้จำภาพซึ่งดีกว่าโครงข่ายประสาทเทียม ((Artificial Neural Networks)

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.2 โครงข่ายประสาทเทียมแบบคอนโวลูชัน (Convolutional Neural Networks: CNN) [7]

เป็นรูปแบบหนึ่งของการเรียนรู้ของเครื่อง (machine learning) ที่จำลองรูปแบบประเภทโครงข่ายประสาทเทียม (Artificial Neural Networks: ANN) ให้เป็นรูปแบบการเรียนรู้เชิงลึก (deep learning) ที่เป็นลักษณะโครงข่ายประสาทเทียมเชิงลึก (Deep Learning Neural Network) โดยมีวัตถุประสงค์ที่จะสร้างเครื่องมือซึ่งมีความสามารถในการ เรียนรู้การจดจำรูปแบบ (Pattern Recognition) และการสร้างความรู้ใหม่ (Knowledge Extraction) เช่นเดียวกับความสามารถที่มีในสมองมนุษย์ โครงข่ายประสาทเทียมเชิงลึกมีการเรียนรู้วิธีการแทนข้อมูลอย่างมีประสิทธิภาพ เนื่องจากการเรียนรู้เชิงลึกมีโครงข่าย hidden layer หลายชั้นที่มากกว่าใน ANN แสดงในภาพที่ 1 ซึ่งเป็นการแบ่งปันค่าถ่วงน้ำหนักหลายๆ ตำแหน่งในภาพนำเข้าทำให้การคำนวณค่าแกรเดียนต์ที่หลายตำแหน่งได้ค่าใหญ่พอให้ทำเกิดการปรับค่าถ่วงน้ำหนักได้อย่างเหมาะสม



ภาพที่ 1: จำนวน hidden layer ของ ANN และ CNN [8]

2.2 TensorFlow [9]

เป็น open source library ที่ google พัฒนาขึ้นด้วยภาษาไพทอน (Python) เป็นไลบรารีสำหรับใช้พัฒนาการเรียนรู้ของเครื่อง เพื่อตอบสนองความต้องการของระบบที่นำมาใช้ในการ

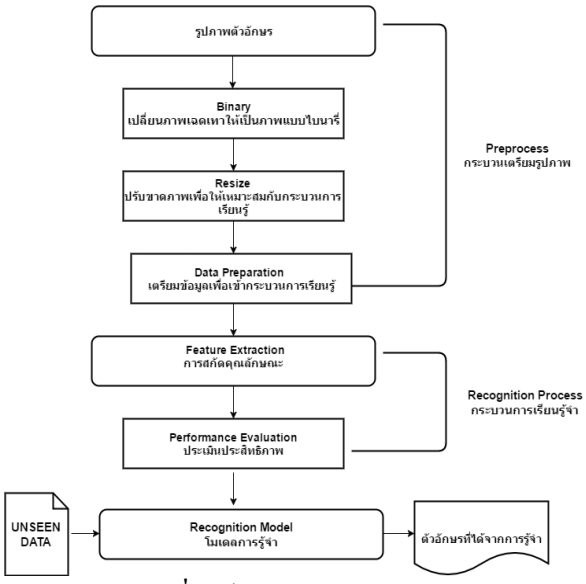
สร้างและ train โครงข่ายประสาทเทียมเชิงลึกเพื่อหารูปแบบและสร้างความรู้ TensorFlow แสดงการคำนวณโดยใช้ Data Flow Grpah ซึ่งแต่ละโหนดของ TensorFlow แทนการคำนวณทางคณิตศาสตร์ ในขณะที่เส้นเชื่อมโยง (edge) แทนข้อมูลอาเรย์แบบหลายมิติ (tensors) ซึ่งทั้งสองอย่างนี้จะทำงานเชื่อมระหว่างกัน

2.3 Optical character recognition (OCR) [10]

การแปลงไฟล์ภาพเอกสารประเภทต่างๆ เช่นเอกสารที่ได้จากการสแกนไฟล์ภาพ (ไฟล์ประเภท PDF, JPG หรือ BMP) หรือภาพที่ถ่ายโดยกล้องดิจิทัล ให้เป็นไฟล์ข้อความโดยอัตโนมัติซึ่งไฟล์ข้อความนั้นสามารถแก้ไขและสามารถค้นหาได้ หลักการประมวลผลของ OCR เป็นไปตามลักษณะของภาษาที่จะนำมาแปลง เนื่องจากแต่ละภาษามีรูปแบบที่แตกต่างกัน เทคโนโลยีการรู้จำตัวอักษรภาษาไทยมีการพัฒนาให้สามารถรองรับการรู้จำตัวอักษรภาษาไทยที่มีความแตกต่างของขนาดและรูปแบบตัวอักษร โดยเฉพาะอย่างยิ่งตัวอักษรภาษาไทยที่มีความคล้ายคลึงกันมากเช่น ด-ค-ต, ก-ถ-ภ, ผ-ฝ-พ-ฟ โครงสร้างของระบบ OCR ประกอบไปด้วยขั้นตอนการทำงานหลัก 3 ขั้นตอนได้แก่ 1) ขบวนการประมวลผลขั้นต้น (Pre-Processing) เป็นขั้นตอนในการปรับแต่งและจัดเตรียมข้อมูลให้เหมาะสมกับขั้นตอนการรู้จำ ขั้นตอนนี้ประกอบด้วยการกรองข้อมูลแทรกซ้อน (Noise Filtering), การปรับแต่งข้อมูล (Normalization), การตัดแบ่งพื้นที่ใช้งาน (Cropping), การสกัดลักษณะสำคัญ (Feature Extraction) 2) เป็นส่วนที่พิจารณารูปตัวอักษรที่ส่งเข้าไปเป็นรหัสตัวอักษรอะไรมีงานวิจัยที่นำเสนอวิธีการหลากหลายเพื่อนำมาใช้ให้ได้ผลการทำงานที่ดีที่สุด เมื่อภาพส่งเข้าสู่ระบบรู้จำ ระบบจะแปลงภาพเป็นข้อความตัวอักษรเพื่อส่งผลลัพธ์ออกสู่กระบวนการประมวลผลขั้นปลาย 3) ขบวนการประมวลผลขั้นปลาย (Post-Processing) เป็นส่วนการตรวจสอบและแก้ไขข้อความเพื่อดูความถูกต้อง ของการสะกดคำและไวยากรณ์ภาษา ก่อนที่เนื้อหาดิจิทัลต่างๆ จะนำไปใช้ประโยชน์ต่อไป

3. วิธีการดำเนินการวิจัย

งานวิจัยนี้มีวิธีในการดำเนินงานวิจัยหลักๆทั้งสิ้น 3 ขั้นตอน ซึ่งประกอบไปด้วย 1) การเตรียมรูปภาพตัวอักษรพิมพ์ไทยเพื่อใช้ในการรู้จำ 2) กระบวนการการรู้จำ และ 3) กระบวนการทดสอบ ขั้นตอนการดำเนินงานวิจัยแสดงในดังภาพที่ 1



ภาพที่ 1: ขั้นตอนวิธีการดำเนินงาน

3.1 การเก็บรวบรวมข้อมูล และการเตรียมข้อมูล

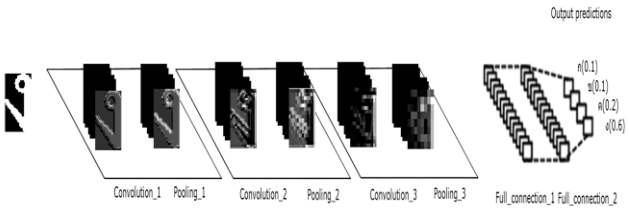
คลังข้อมูลภาพตัวอักษรพิมพ์ไทยสำหรับการรู้จำตัวอักษรภาษาไทยได้นำมาจากศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ (NECTEC) [11] ที่มีจำนวน 48,880 ภาพ และนามสกุลไฟล์ภาพเป็น .bmp โดยแบ่งเป็นตัวอักษรพยัญชนะไทย 44 รูปแบบ และรูปแบบสระบางส่วน, รูปวรรณยุกต์ ทั้งหมด 21 รูปแบบ (ทั้งหมด 65 class) ในงานวิจัยนี้ใช้ภาพรูปตัวอักษร 18 รูปภาพของแต่ละพยัญชนะ, รูปสระบางส่วนและรูปวรรณยุกต์ ดังแสดงในภาพที่ 2 ดังนั้นข้อมูลภาพตัวอักษรพิมพ์ไทยที่ใช้ในงานวิจัยมีทั้งหมด 1,170 ภาพ (ได้มาจาก 18*65) เลือกรูปแบบตัวอักษร 8 ฟอนต์จากคลังข้อมูลภาพที่มีทั้งหมด 18 ฟอนต์ ได้แก่ Angsana, Browalia, Cordia, Dillenia, Eucrosia, Freesia, Jasmine และ Kodchiang ขนาดฟอนต์ 16 ตัวอักษรพิมพ์ไทยเป็นแบบอักษรปกติไม่หนาและไม่ตัวเอียง แต่ละภาพมีขนาด 31x51 pixels เพื่อเป็นการเตรียมรูปตัวอักษรให้มีหลายขนาดสำหรับใช้ในการรู้จำ จึงได้ทำการปรับขนาดรูปภาพตัวอักษร 3 ขนาด คือ 28x28, 32x32 และ 64x64 pixels



ภาพที่ 2: ตัวอย่างของรูปตัวอักษร, รูปสระ, รูปวรรณยุกต์

การพัฒนากระบวนการการรู้จำในงานวิจัยนี้ใช้ TensorFlow ซึ่งเป็น library ของ google ที่รองรับการทำงานโครงข่ายประสาทเทียมแบบคอนโวลูชั่น การประมวลผลทำงานบนระบบปฏิบัติการ WINDOWS 7 รุ่น 64 bit ซึ่งมีหน่วยประมวลผล Intel(R) Core(TM) i7-4500U CPU 1.80GHz หน่วยความจำ 8GB และ หน่วยประมวลผลภาพ NVIDIA GeForce 840M ข้อมูลรูปภาพที่ใช้ในการเรียนรู้จำได้ใช้ตามขนาดของข้อมูลรูปภาพที่เป็น 3 ขนาด 28x28, 32x32 และ 64x64 pixels ตามลำดับ

ข้อมูลรูปภาพที่นำมาใช้ในการรู้จำถูกแบ่งเป็น 2 ส่วน ส่วนแรกมีจำนวนข้อมูล80 % เพื่อใช้สำหรับการฝึกสอน (training) ส่วนที่สองมีจำนวนข้อมูล20 % เพื่อใช้สำหรับการทดสอบ เพื่อวัดค่าความแม่นยำในการเรียนรู้จำ และเข้าสู่กระบวนการการสกัดคุณลักษณะ (feature extraction) โดยใช้กระบวนการ (Conv, ReLU and Maxpool) ใน TensorFlow เพื่อสกัดและแยกคุณลักษณะอย่างอัตโนมัติของรูปภาพตัวอักษรที่อยู่ในชั้น hidden layers ของโครงข่ายประสาทเทียมแบบคอนโวลูชั่น โดยที่คุณลักษณะที่ถูกแยกจะถูกแบ่งไปตามตัวพยัญชนะที่กำหนด (class) ตัวอย่างถูกแสดงดังภาพที่ 3



ภาพที่ 3: ตัวอย่างการสกัดคุณลักษณะจากรูปภาพ

การทดลองของงานวิจัยนี้เลือกใช้โมเดลสำหรับการรู้จำ 2 แบบ ได้แก่ แบบที่ 1คือ 2 Conv, 2 ReLU and 2 maxpool และแบบที่ 2 คือ 3 Conv, 3 ReLU and 3 maxpool ทดลองกับรูปภาพอักษรทั้ง 3 ขนาดที่ได้กล่าวไปแล้ว ดังนั้นรูปแบบการทดลองทั้งหมดได้ 5 รูปแบบการทดลองดังนี้

Model A ใช้ภาพขนาด 28*28 pixels, (2 Conv, 2 ReLU, 2 maxpool) มี 16 และ 32 hidden layer nodes

Model B ใช้ภาพขนาด 32*32 pixels, (2 Conv, 2 ReLU, 2 maxpool) มี 16 และ 32 hidden layer nodes

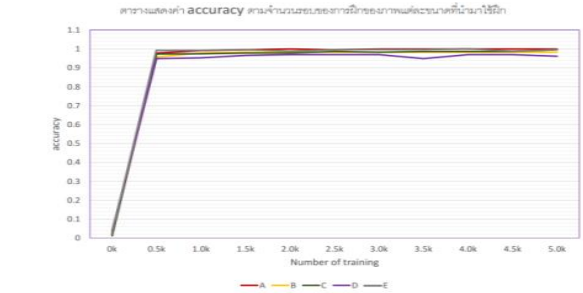
Model C ใช้ภาพขนาด 32*32 pixels, (3 Conv, 3 ReLU, 3 maxpool) มี 16, 32 และ 64 hidden layer nodes

Model D ใช้ภาพขนาด 64*64 pixels, (2 Conv, 2 ReLU, 2 maxpool) มี 16 และ 32 hidden layer nodes

Model E ใช้ภาพขนาด 64*64 pixels (3 Conv, 3 ReLU, 3 maxpool) มี 16, 32 และ 64 hidden layer nodes

Model A มีการปรับขนาดของกระบวนการ maxpool 2 ครั้ง ทำให้ผลลัพธ์ของภาพที่ได้มีขนาดเหลือเป็น ของภาพ 4/1 ดังนั้นเมื่อนำภาพขนาด 28*28 pixels ผ่านกระบวนการ maxpool 2 ครั้ง เหลือภาพขนาด 7*7pixels จึงไม่สามารถทำกระบวนการ maxpool ครั้งที่3 ได้ ภาพขนาด 28*28 pixels จึงผ่านกระบวนการ 2 Conv, 2 ReLU และ 2 maxpool เท่านั้น

ในการทดลองใช้รอบในการฝึกสอนทั้งหมด 5,รอบ 000 หลังจากทำการเรียนรู้จำแล้วจึงทำการบันทึกโมเดลเพื่อใช้ในการทดสอบกับภาพข้อมูลใหม่ (unseen data) ผลค่าความถูกต้องของการของกระบวนการเรียนรู้จำ แสดงดังภาพที่ 4



	0k	0.5k	1.0k	1.5k	2.0k	2.5k	3.0k	3.5k	4.0k	4.5k	5.0k
A	0.0128	0.9786	0.9914	0.9957	1	0.9957	1	1	1	1	1
B	0.0256	0.9572	0.9786	0.9829	0.9829	0.9829	0.9829	0.9829	0.9829	0.9829	0.9829
C	0.0128	0.9735	0.9743	0.9786	0.9786	0.9871	0.9829	0.9871	0.9871	0.9871	0.9957
D	0.0384	0.9487	0.9529	0.9658	0.97	0.97	0.97	0.9487	0.97	0.97	0.9615
E	0.0213	0.9914	0.9914	0.9957	0.9871	0.9957	0.9957	0.9957	1	0.9871	0.9957

ภาพที่ 4: กราฟและตารางที่แสดงค่าความแม่นยำของภาพแต่ละขนาดที่นำมาฝึกซึ่งแสดงผลตามจำนวนรอบของการฝึก

จากภาพที่ 4 แสดงให้เห็นว่า Model A ให้ค่าความแม่นยำของการรู้จำของมากที่สุด 98.29% โดยหากเปรียบเทียบการใช้เวลาในการประมวลผลเมื่อแบ่งตามการ *convolution* ออกเป็น 2 Conv และ 3 Conv ส่วน Model A, B และD ที่เป็น 2 Conv ใช้เวลาในการประมวลผลการรู้จำ 9, 11, 44 นาที ตามลำดับ Model C, และE ที่เป็น 3 Conv ใช้เวลาในการประมวลผลการรู้จำ 15, 63 นาที ตามลำดับ

3.3 การทดสอบกับภาพข้อมูลใหม่

เพื่อทดสอบประสิทธิภาพ model ของการรู้จำ จึงได้ทำการทดสอบกับภาพข้อมูลใหม่ (unseen data) ที่เป็นภาพข้อความที่เตรียมขึ้นมา และภาพตัวอักษรพิมพ์ไทย สำหรับภาพข้อมูลที่เป็นข้อความเช่นข้อความ “ก็สามารถทำให้ต้นข้าวเกิดโรค” ได้จากการแปลงไฟล์เอกสารเป็นไฟล์ภาพที่มีขนาด 1943*1289 pixels นามสกุลไฟล์ภาพเป็น .jpg โดยมีขั้นตอนดังนี้

3.3.1 ทำ threshold เพื่อทำให้ภาพสีเฉดเทาเปลี่ยนเป็นภาพ Binary ที่มีค่าสีเพียง 2 ค่า คือ 0 และ 1

3.3.2 การแยกตัวอักษร (segment) หลังจากภาพเอกสารได้ผ่านการทำ threshold แล้วทำให้ตัวอักษรที่สนใจนั้นมีค่าสีเป็นสีดำทั้งหมด มาทำการแยกตัวภาพอักษร, รูปสระ และรูปวรรณยุกต์ ออกจากกันโดยใช้ Flood-Fill Algorithm [12] เนื่องจากการแยกตัวอักษรอาจจะไม่ถูกต้องทั้งหมด จึงมีการตรวจสอบและแก้ไขให้การแยกตัวอักษรนั้นถูกต้อง

3.3.3 ปรับขนาดภาพ เนื่องจากภาพที่ผ่านขั้นตอนต่างๆ ทำให้ภาพที่ได้มีขนาดไม่เท่ากัน จึงต้องทำการปรับขนาดของรูปภาพให้เท่ากัน เพื่อใช้ในขั้นตอนการทดสอบโมเดลสำหรับการรู้จำ

สำหรับภาพตัวอักษรพิมพ์ไทยที่นำมาทดสอบได้จากคลังข้อมูลรูปภาพจาก NECTEC ที่เป็นคนละชุดกับชุดข้อมูลที่ใช้ในการเรียนรู้ จำนวน 20ภาพของแต่ละพยัญชนะ, รูปสระและรูปวรรณยุกต์ ดังนั้นข้อมูลภาพทั้งหมด 1,350 รูปภาพ (20*65 รูปภาพ) และนำภาพมาปรับขนาดภาพให้เหมาะสม เพื่อนำไปใช้ในการรู้จำภาพตัวอักษรใหม่

.4 ผลการดำเนินงานวิจัย

จากการนำ Model A, B, C, D และ E มาทำการทดสอบภาพกับข้อมูลใหม่ (unseen data) ที่เป็นข้อความ และภาพตัวอักษรได้ผลการทดลองแสดงในตารางที่ 1 และตารางที่ 2 ตามลำดับ

ตารางที่ 1: ข้อมูลภาพข้อความที่มาใช้ในการทดสอบกับ model ต่างๆ

Models	Text Image
A	ก็สามารถทำให้ต้นข้าวเกิดโรค
B	ก็สามารถทำให้ต้นข้าว**ว์กิด**โรค
C	ก็สามารถทำให้ต้นข้าวเกิดโรค
D	ก็สามารถ**ท่**ให้ต้นข้าวเกิดโรค
E	ก็สามารถ**ท่ร**ห้ให้ต้นข้าวเกิดโรค

จากตารางที่ 1 Model A และ Model C สามารถรู้จำตัวอักษรทุกตัวอักษรของข้อความถูกต้องทุกตัวอักษร ส่วน Model B รู้จำรูปผิดไป 1 ตัวเป็นรูปวรรณยุกต์ไม้เอก (ว่) ซึ่งที่ถูกต้องคือ “เ” Model D รู้จำรูปสระผิดไป 1 ตัวเป็นรูปวรรณยุกต์ไม้เอก (ท่) ซึ่งที่ถูกต้องคือรูปสระ “ำ” Model E รู้จำรูปสระผิดไป 2 ตัวเป็นรูปวรรณยุกต์ (ท่) และ “ร” ซึ่งที่ถูกต้องคือรูปสระ “ำ” และรูป ”ใ” ตามลำดับ

ตารางที่ 2: ผลการทดสอบประสิทธิภาพการรู้จำของภาพตัวอักษรที่ใช้ทดสอบกับ model ต่างๆ

Models	A	B	C	D	E
ค่าความถูกต้อง	**0.9984**	0.9969	0.9969	0.9884	0.9976

จากตาราง 2 ที่แสดงประสิทธิภาพการรู้จำของ model ต่างๆ ที่ทำภาพตัวอักษรพิมพ์ไทย จะเห็นได้ว่า Model A ให้ค่าประสิทธิภาพการรู้จำมากที่สุดที่มีค่าความแม่นยำ 99.84% เมื่อใช้ทดสอบกับภาพตัวอักษรพิมพ์ไทยที่เป็นรูปภาพใหม่ (unseen data)

จากผลการทดลองที่แสดงให้เห็นว่า Model A ให้ค่าประสิทธิภาพการรู้จำมากที่สุด จึงได้ทำการทดลองเปรียบเทียบประสิทธิภาพโดยใช้โปรแกรม Tesseract-OCR [13] ซึ่งเป็นโครงข่ายประสาทเทียมแบบ Long Short-Term Memory (LSTM) จัดเป็นโครงข่ายประเภท Recurrent Neural Network (RNN) โดยใช้ข้อมูลที่เป็น unseen data ชุดเดียวกับที่ใช้ทดสอบกับ Model Aใน ซึ่งผลการทดลองเปรียบเทียบประสิทธิภาพความถูกต้องระหว่าง Tesseract-OCR และ Model A แสดงดังตารางที่ 3

ตารางที่ 2: ผลการทดลองเปรียบเทียบประสิทธิภาพความถูกต้องระหว่าง Tesseract-OCR และ Model A

	Model A	Tesseract-OCR
ค่าความถูกต้อง	**0.9984**	0.7668

5. สรุปและอภิปราย

จากผลการทดลองที่เป็นการเรียนรู้การรู้จำภาพตัวอักษรพิมพ์ไทย และเปรียบเทียบประสิทธิภาพการเรียนรู้จำที่ทดสอบกับข้อความรูปภาพ และภาพที่เป็น unseen data แสดงให้เห็นว่า Model A ที่ใช้ขนาดภาพ 28x28 pixels เมื่อเทียบกับ model อื่นๆ นั้นให้ประสิทธิภาพในการรู้จำตัวอักษรได้ถูกต้องมากที่สุด

99.84% โดยการรู้จำผิดพลาดไป 2 ตัวอักษรจาก 65 ตัวอักษร คือ จาก “ซ” รู้จำเป็น “ช” ทั้ง 2 ตัว

สำหรับภาพตัวอักษรขนาด 64x64 pixels เหมาะสมกับการใช้โครงข่ายประสาทเทียมแบบคอนโวลูชั่นที่มี 3 Conv, 3 ReLU, และ 3 maxpool มากที่สุดโดยให้ประสิทธิภาพในการรู้จำตัวอักษรได้ถูกต้อง 99.76% ได้รู้จำผิดพลาดไป 3 ตัวอักษรจาก 65 ตัวอักษร คือ ‘ซ’ รู้จำเป็น ‘ฆ’ 2ตัว และ ‘เ’ ทำนายเป็น ‘ท’ 1 ตัว ดังนั้นเมื่อเปรียบเทียบกับ Model A จะเห็นได้ว่าภาพของตัวอักษรพิมพ์ไทยขนาด 28x28 pixels เหมาะสมที่จะนำมาใช้กับการรู้จำภาพตัวอักษรพิพม์ไทยโดยใช้โครงข่ายประสาทเทียมแบบคอนโวลูชั่นสำหรับการรู้จำที่ใช้กระบวนการที่เป็น 2 Conv, 2 ReLU และ 2 Maxpool ใน TensorFlow ที่ให้ความถูกต้องของการรู้จำตัวอักษร 99.84%

สำหรับการเลือกใช้ฟังก์ชัน Conv, ReLU และMaxpool ซึ่งอยู่ในเทคโนโลยี Convolution Neural Network (CNN) ในกระบวนการการสกัดคุณลักษณะโดยที่ CNN จะทำการคัดเลือกคุณลักษณะเอง จนได้คุณลักษณะที่มีความสัมพันธ์กับผลลัพธ์มากที่สุด ซึ่งแตกต่างจาก Machine learning แบบเดิม ที่จะต้องให้มนุษย์เป็นคนเลือกว่าจะใช้คุณลักษณะใดในการเรียนรู้ และถ้าเลือกใช้คุณลักษณะที่ไม่เกี่ยวข้องกับผลลัพธ์ก็จะทำให้โปรแกรมเกิดความผิดพลาดในการประมวลผล

ผลของการทดลองดังกล่าวสามารถนำไปประยุกต์ใช้กับเนื้อหาข้อความภาษาไทยที่เป็นรูปภาพโดยใช้ในการรู้จำตัวอักษรแบบการเรียนรู้เชิงลึก โครงข่ายประสาทเทียมแบบคอนโวลูชั่น ที่ให้ประสิทธิภาพความถูกต้องในการรู้จำตัวอักษรมากที่สุดแปลงเป็นข้อความที่แก้ไขได้ แล้วนำข้อความมาแปลงเป็นอักษรเบรลล์สำหรับผู้ใช้ที่เป็นผู้พิการทางสายตาต่อไป

เอกสารอ้างอิง

[1] ธเนศ ศรีวิรุฬห์ชัย, “การรู้จำตัวอักษรพิมพ์ภาษาไทยโดยใช้เทคนิคด้านการวิเคราะห์ตัวประกอบสำคัญและนิวรอลเน็ตเวิร์ก”, *วิทยานิพนธ์ปริญญามหาบัณฑิต*, ภาควิชาวิศวกรรมคอมพิวเตอร์, คณะวิศวกรรมศาสตร์, จุฬาลงกรณ์มหาวิทยาลัย, 2541.

[2] อุกฤษณ์ มารังค์, ชม กิ้มปาน, และ พิศิษฏ์ โภคารัตน์กุล, “ระบบฟัซซี่นิวรอลในการรู้จำอักษรตัวพิมพ์ภาษาไทย”, *การประชุมทางวิชาการของมหาวิทยาลัยเกษตรศาสตร์ ครั้งที่ 38 สาขาวิศวกรรมศาสตร์และสาขาอุตสาหกรรมเกษตร*, 1-4 กุมภาพันธ์ 2543.

[3] วราภรณ์, อภิรัตนตระกูล, “การรู้จำตัวอักษรพิมพ์ภาษาไทยโดยวิธีรันแรงท์โค้ดร่วมกับฮีสโตแกรมและการวิเคราะห์บิทแมพ”, *วิทยานิพนธ์ปริญญามหาบัณฑิต*, สาขาเทคโนโลยีสารสนเทศ, มหาวิทยาลัยศรีปทุม, 2544.

[4] วรพจน์ พีระวิทย์, “การเพิ่มประสิทธิภาพของระบบรู้จำตัวอักษรภาษาไทยโดยการซ่อมแซมตัวอักษรไม่สมบูรณ์”, *วิทยานิพนธ์มหาบัณฑิต*, วิศวกรรมคอมพิวเตอร์, วิศวกรรมศาสตร์, มหาวิทยาลัยเกษตรศาสตร์, 2547.

[5] A. Thammano and P. Duangphasuk, “Printed Thai Character Recognition Using Hierarchical Cross-Correlation ARTMAP”, *Proceedings of the 17th IEEE International Conference on Tools with Artificial Intelligence*, December 17, 2013.

[6] S. Chaivatna and S. Tangwongsan, "An Optimal Approach towards Recognizing Broken Thai Characters in OCR Systems", Proceedings of 2012 International Conference on Digital Image Computing Techniques and Applications (DICTA), Dec 3-5, 2012.

[7] L. Deng and D. Yu, “Deep Learning: Methods and Applications”, *Foundations and Trends in Signal Processing*, vol. **7,** issue 3–4, 2014.

[8] Deep Neural Network [Online] Available from: *https://stats.stackexchange.com/questions/182734/what-is-the-difference-between-a-neural-network-and-a-deep-neural-network*, access date: February 15, 2017.

[9] TensorFlow [Online] Available from: *https://www.tensorflow.org/*, access date: March 1, 2017.

[10] Thai OCR [Online] Available from: *http://thaiocr.phaisarn.com/ocr.html*, access date: February 10, 2017.

[11] คลังข้อมูลภาพอักษร [Online] Available from: http://www.nectec.or.th/corpus/index.php?league=bd, access date: November 5, 2016.

[12] Flood-Fill Algorithm [Online] Available from: *http://stackoverflow.com/questions/2783204/flood-fill-using-a-stack*, access date: January 7, 2017.

[13] R. Smith, “An Overview of the Tesseract OCR Engine”, *Proceedings of the 9th International Conference on Document Analysis and Recognition*, 2007, pp. 629–633.

การแก้ไขความคลาดเคลื่อนของสีด้วยการกรองไกด์อิมเมจ
Rectifying Chromatic Aberration using Guided Image Filtering

พิชามญชุ์ สินลาลับ (P.Sinlalab) [1] ปรวัฒน์ วิสูตรศักดิ์ (P.Visutsak) [2] และสรร รัตนสัญญา (S. Ratanasanya) [3]
ภาควิชาวิทยาการคอมพิวเตอร์และสารสนเทศ คณะวิทยาศาสตร์ประยุกต์ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
[1]pichamon.s@nrru.ac.th, [2]porawat.v@sci.kmutnb.ac.th, [3]san.ratanasanya@gmail.com

บทคัดย่อ

ความคลาดเคลื่อนของสี (Chromatic Aberration) เกิดจากการกระเจิงของแสงสีต่างๆ ที่มีความยาวคลื่นต่างกัน ซึ่งเป็นส่วนประกอบของแสงสว่างที่มาตกกระทบเลนส์ (Lens) คุณภาพต่ำ งานวิจัยที่ผ่านมายังไม่สามารถแก้ไขเฉพาะบริเวณที่เกิดความคลาดเคลื่อนของสีได้อย่างเหมาะสม เป็นเพียงการลดทอนความคลาดเคลื่อนของสีด้วยการปรับเป็นสีเทา บทความนี้นำเสนอการแก้ไขความคลาดเคลื่อนของสี (Chromatic Aberration) บนพื้นฐานของการกรองไกด์อิมเมจ (Guided Image Filtering) โดยใช้ประโยชน์จากคุณสมบัติที่น่าสนใจของการกรองไกด์อิมเมจ จากผลการทดลองในเบื้องต้น ให้ผลลัพธ์เป็นที่น่าพอใจ เนื่องจากสามารถแก้ไขความคลาดเคลื่อนของสีได้ใกล้เคียงกับสีธรรมชาติมากขึ้นเมื่อเทียบกับวิธีการอื่น ๆ ก่อนหน้า

คำสำคัญ: ตัวกรองไกด์อิมเมจ,ความคลาดเคลื่อนของสี, การแก้ไขความคลาดเคลื่อนของสี

Abstract

Chromatic Aberration (CA) occurs by the dispersion of rays of colors, each color has different wavelength, consisting in the ray of light passing through low quality lens. The existing research in the literatures were not be able to rectify the CA although the CA was suppressed by modifying it toward color of gray. This paper presents rectification of CA based on Guided Image Filtering. By exploiting the interesting properties of guided image filtering. The preliminary experiment showed an impressive results in a way that the CA was rectified toward the real color of particular objects as close as possible compared to the existing techniques.

Keyword: Guided Image Filter, Chromatic Aberration, Correction of Chromatic Aberration

1. บทนำ

ปัจจุบันภาพถ่ายมีบทบาทเกี่ยวข้องกับชีวิตประจำวันของคนเราเพิ่มมากขึ้น เรามักจะใช้ภาพถ่ายเป็นสื่อกลางในการถ่ายทอดเรื่องราวต่าง ๆ คุณภาพของภาพถ่ายจึงเป็นเรื่องที่สำคัญ หากมีสิ่งผิดปกติบนภาพถ่าย ย่อมจะส่งผลต่อคุณภาพของภาพถ่าย หากภาพนั้นใช้ในเชิงธุรกิจ อาจส่งผลต่อความน่าเชื่อถือและมูลค่าของงานได้ สิ่งผิดปกติที่มักจะพบในภาพถ่ายคือความคลาดเคลื่อนหรือความผิดเพี้ยนของสี สังเกตได้จากขอบของวัตถุในภาพจะเป็นสีม่วงหรือสีอื่นๆ และไม่มีความคมชัด ซึ่งเกิดจากแสงของแต่ละสี เดินทางผ่านเลนส์ (Lens) และ ไม่ได้ตกกระทบไปยังจุดเดียวกัน [1] ความคลาดเคลื่อนของสีจะเกิดในบริเวณที่มีแสงสว่างมาก จึงมักเกิดกับเลนส์ที่มีการเปิดรูรับแสงกว้างในขณะที่ทำการถ่ายภาพ และมักเกิดกับเลนส์ที่มีทางยาวโฟกัส (Focal Length) มาก ๆ เนื่องจากแสงมีการเดินทางระยะทางมากทำให้เกิดการกระจายตัวของแสงมากขึ้นตามไปด้วย [2]

จากการที่ได้ศึกษางานวิจัยที่มีการแก้ไขปัญหาความคลาดเคลื่อนของสีโดยวิธีทางตรรกะ(Logical) หรือการใช้คอมพิวเตอร์ในการแก้ไขจุดที่เกิดความคลาดเคลื่อนของสี นั้นพบว่างานวิจัยที่ผ่านมา เช่น [3], [4], [5], [6], [7] ยังไม่สามารถแก้ไขปัญหาได้อย่างสมบูรณ์ เป็นเพียงการลดทอนความคลาดเคลื่อนของสี โดยลดทอนให้เข้าใกล้สีเทาหรือสีดำเท่านั้น ดังนั้นงานวิจัยนี้จึงนำเสนอวิธีการแก้ไขความคลาดเคลื่อนของสีโดยการใช้การกรองไกด์อิมเมจ เพื่อปรับ

ค่าสีเฉพาะบริเวณที่เกิดความคลาดเคลื่อนของสี ด้วยการใช้ประโยชน์จากคุณสมบัติที่น่าสนใจของการกรองไกด์อิมเมจมาทำการปรับค่าสีที่คลาดเคลื่อนให้มีความใกล้เคียงกับสีตามธรรมชาติมากที่สุด ซึ่งผู้วิจัยคาดหวังว่าวิธีการนี้จะสามารถแก้ไขความคลาดเคลื่อนของสีได้สมบูรณ์มากยิ่งขึ้น

2. งานวิจัยที่เกี่ยวข้อง

ปัญหาความคลาดเคลื่อนของสีเริ่มได้รับความสนใจเมื่อไม่นานมานี้ งานวิจัยที่เกี่ยวข้องกับความคลาดเคลื่อนของสีกับจึงยังมีจำนวนไม่เยอะมาก ซึ่งยุทธศักดิ์และคณะ [8] ได้แบ่งงานวิจัยที่เกี่ยวข้องออกเป็น 2 กลุ่ม โดยงานวิจัยในกลุ่มแรก จะทำการตรวจสอบภาพถ่ายทีละพิกเซล (Pixel) เพื่อหาบริเวณที่เกิดความคลาดเคลื่อนของสี และทำการปรับให้ค่าสีแดง สีเขียว และสีน้ำเงินในบริเวณดังกล่าวให้มีความใกล้เคียงกัน ส่งผลให้บริเวณนั้นกลายเป็นสีเทา หรือสีดำ [3], [4], [5], [6] ซึ่งการตรวจสอบในระดับพิกเซลนั้นมีจุดบกพร่องคือใช้เวลาในการประมวลผลนาน นอกจากนั้นวิธีการเหล่านี้ยังไม่สามารถแยกแยะบริเวณที่เกิดความคลาดเคลื่อนของสีออกจากบริเวณที่เป็นสีธรรมชาติได้อีกด้วย จึงทำให้ไม่สามารถลดทอนความคลาดเคลื่อนของสีได้อย่างถูกต้อง ต่อมา ณัฐชัย และคณะ [9] ได้เสนอการนำค่าคอนทราสต์เฉพาะที่ (Local Contrast) มาใช้ในการตรวจจับบริเวณที่เกิดความคลาดเคลื่อนของสีแบบอัตโนมัติ และใช้ทิศทางของเกรเดียนต์ (Gradient Direction) ของแสงมาช่วยแยกแยะระหว่างสีที่เกิดขึ้นจริงตามธรรมชาติและสีที่เกิดจากความคลาดเคลื่อนของสีได้อย่างถูกต้องชัดเจนมากขึ้นอย่างมีนัยสำคัญ อย่างไรก็ตามสำหรับการแก้ไขความผิดเพี้ยนของสีที่เกิดขึ้นจากความคลาดเคลื่อนของสีนั้น ยังคงใช้วิธีการลดทอนค่าสีจนเข้าใกล้สีเทา หรือสีดำอยู่นั่นเอง

งานวิจัยในกลุ่มที่สองใช้ ตัวกรอง (Filter) มาช่วยให้การประมวลผลเร็วขึ้น โดย Jung และคณะ [7] ได้เสนอการแก้ไขความคลาดเคลื่อนของสีโดยแก้บริเวณขอบภาพทั้งหมดด้วยวิธีการกรองไกด์อิมเมจ (Guided Image Filtering) ซึ่งวิธีการนี้เป็นการผสานภาพ 2 ภาพเข้าด้วยกัน โดยนำโครงสร้างหลักจาก ภาพ I หรือ ภาพไกด์ (Guidance Image) มาเป็นตัวชี้นำ เพื่อถ่ายทอดโครงสร้างหลักของภาพนั้นไปยังภาพผลลัพธ์ q_i (Output Image) ดังสมการที่ (1)

$$q_i = \bar{a}_i I + \bar{b}_i \tag{1}$$

ซึ่ง $\bar{a}_i$, $\bar{b}_i$เป็นค่าเฉลี่ยสัมประสิทธิ์เชิงเส้นและ i คือเลขลำดับที่ หรือ ดัชนี (Index) ของพิกเซลภายในภาพ ดังนั้นภาพไกด์จึงเป็นภาพที่มีรายละเอียดต่างๆ ที่ต้องการ เช่น วัตถุต่างๆในภาพสมบูรณ์ที่สุด และที่ภาพผลลัพธ์จะปรากฏโครงสร้างหลักของภาพไกด์ ซึ่งได้แก่ ขอบของวัตถุต่างๆในภาพนั้น สำหรับรายละเอียด และความคมชัดอื่นๆของภาพและที่ภาพผลลัพธ์จะปรากฏโครงสร้างหลักของภาพไกด์ ซึ่งได้แก่ ขอบของวัตถุต่างๆในภาพนั้น ผลลัพธ์นั้น จะได้จากภาพที่ทำหน้าที่เป็นตัวกรองนำเข้า (Filtering Input Image) หรือ p ซึ่งจะทำการถ่ายโอนรายละเอียดของสีและความคมชัดจากภาพ p ไปยัง q ผ่านทางค่าเฉลี่ยสัมประสิทธิ์เชิงเส้น $\bar{a}_i$ และ $\bar{b}_i$ เนื่องจาก $\bar{a}_i$ และ $\bar{b}_i$ เป็นความสัมพันธ์ระหว่างภาพ I และ p ดังแสดงในสมการที่ (2),(3) และ (4)

$$a_i = \frac{\frac{1}{\omega}\sum_{i\in\omega_k} I_i p_i - \mu_k \bar{p}_k}{\sigma_k^2 + \varepsilon} \tag{2}$$

$$b_i = \bar{p}_k - a_k \mu_k \tag{3}$$

$$\bar{p}_k = \sum_{i\in\omega_k} p_i \tag{4}$$

โดยที่ μ_k เป็นค่าเฉลี่ย(Mean) σ_k^2 เป็นค่าความแปรปรวน (Variance) ของรูปภาพ a_i เป็นสัมประสิทธิ์เชิงเส้นจากการหาความแปรปรวนร่วมเกี่ยวระหว่างภาพสองภาพ b_i เป็นสัมประสิทธิ์เชิงเส้นที่แสดงถึงรายละเอียดของสีจากภาพที่เป็นตัวกรองเกิดจากการ นำค่าเฉลี่ยของภาพที่เป็นตัวกรองมาลบด้วย a_k ซึ่งเป็นโครงสร้างของภาพไกด์และ $\bar{p}_k$ เป็นค่าเฉลี่ยของภาพที่เป็นตัวกรอง

วิธีการของ Jung ใช้ภาพที่ชั้นสีเขียวเป็นภาพไกด์ จึงเปรียบเสมือนเป็นการปรับค่าของภาพที่ชั้นสีแดงและชั้นสีน้ำเงิน ให้มีความใกล้เคียงกับค่าของภาพที่ชั้นสีเขียวให้มากที่สุด มีผลให้ความคลาดเคลื่อนของสีบริเวณขอบภาพลดหายไป เพราะค่าของทั้ง 3 สีมีความใกล้เคียงกัน แต่วิธีการกรองไกด์อิมเมจสามารถทำได้เพียงปรับปรุงคุณภาพของภาพให้ดีขึ้น

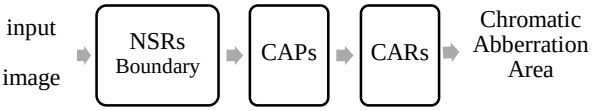
เท่านั้น ทำให้ไม่สามารถแก้ไขเฉพาะบริเวณที่เกิดความคลาดเคลื่อนของสีได้ นอกจากนี้ก็ยังไม่สามารถแยกบริเวณส่วนที่เกิดความคลาดเคลื่อนของสีออกจากบริเวณอื่นได้

จากการศึกษาคุณสมบัติของภาพไกด์และภาพที่เป็นตัวกรอง ผู้วิจัยมีแนวคิดว่าโครงสร้างของภาพและรายละเอียดของสีนั้นสามารถนำมาใช้เพื่อปรับค่าความคลาดเคลื่อนของสีให้มีความใกล้เคียงกับสีจริงของภาพนั้นๆ ซึ่งในงานวิจัยที่ผ่านมา พบว่ายังไม่มีงานใดสามารถแก้ไขความคลาดเคลื่อนของสีได้ เป็นเพียงแค่การลดทอนให้เข้าใกล้สีดำหรือสีเทา เพื่อให้มองไม่เห็นความคลาดเคลื่อนเท่านั้น

3. วิธีการที่นำเสนอ

3.1 การแก้ไขความคลาดเคลื่อนของสีด้วยการกรองไกด์อิมเมจ

งานวิจัยนี้เริ่มจากการแยกส่วนที่เกิดความคลาดเคลื่อนของสีออกจากบริเวณอื่นก่อน ซึ่งจะทำให้สามารถแก้ไขเฉพาะบริเวณที่เกิดความคลาดเคลื่อนของสีได้ โดยใช้ขั้นตอนการดำเนินงานการตรวจจับความคลาดเคลื่อนของสีอัตโนมัติของณัฐชัยและคณะ[9] ซึ่งมีรายละเอียดดังภาพที่ 1



ภาพที่ 1: ขั้นตอนการดำเนินงานของการตรวจจับความคลาดเคลื่อนสีแบบอัตโนมัติ

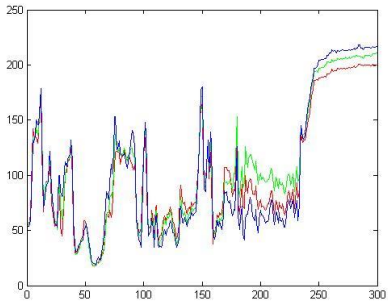
ขั้นตอนแรกเริ่มจากการคำนวณหาบริเวณที่เป็นจุดสว่างในภาพที่เรียกว่าการหา Near saturated regions (NSRs) หลังจากนั้นทำการค้นหาจุดที่มีความเป็นไปได้ที่จะเกิดความคลาดเคลื่อนของสีเรียกบริเวณดังกล่าวว่า Chromatic Aberration Points (CAPs) โดยค้นหาจากบริเวณเส้นขอบของรูปภาพทั้ง 4 แกน และขยายบริเวณที่ค้นหาออกเพื่อให้ครอบคลุมบริเวณที่เกิดความคลาดเคลื่อนของสีทั้งหมด หรือ Chromatic Aberration Regions (CARs) ด้วยการใช้ทิศทางของเกรเดียนต์ของแสง (Gradient Direction) และ Dilation ช่วยในการระบุว่าบริเวณดังกล่าวเป็นบริเวณที่เกิดความคลาดเคลื่อนของสีขึ้นจริง เมื่อได้บริเวณที่เกิดความคลาดเคลื่อนของสีแล้ว จะเห็นได้ว่าในบริเวณดังกล่าว ค่าความเข้มแสง (Intensity) ของสีแดง และ / หรือ สีน้ำเงิน มีความแตกต่างจากค่าความเข้มแสงของสีเขียวอย่างเห็นได้ชัดดังแสดงในภาพที่ 2 ซึ่งจะเห็นได้ว่ามีความความเคลื่อนของสีตามเส้นขอบของวัตถุ ซึ่งสีที่คลาดเคลื่อนนั้น เกิดจากสีแดงและสีน้ำเงิน จึงทำให้ดูเป็นสีม่วง



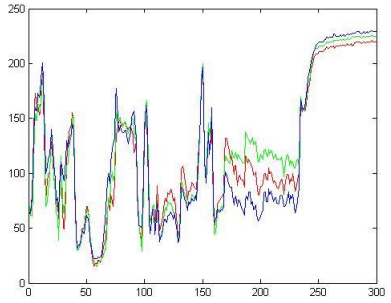
(a) (b)

ภาพที่ 2: ภาพแสดงการเปรียบเทียบภาพต้นฉบับ(a)และภาพบริเวณที่เกิดความคลาดเคลื่อนของสี(b)

หากมองภาคตัดขวางของภาพที่ 2 เพื่อดูความเข้มของสีทั้ง 3 สี จะปรากฏดังแสดงในภาพที่ 3



(a)



(b)

ภาพที่ 3: กราฟแสดงค่าความเข้มของแสงสีแดง เขียวและน้ำเงิน จากภาพต้นฉบับ(a)ในภาพที่2และ(b)ในบริเวณที่เกิดความคลาดเคลื่อนของสีจากภาพที่ 2

จากภาพที่ 3 จะเห็นได้ว่า เส้นกราฟค่าความเข้มของแสงทั้ง 3 สีในภาพที่ 3(a) มีความใกล้เคียงกัน ส่วนในภาพที่ 3(b) จะมีบางช่วงที่ทั้ง 3 สีมีความแตกต่างกันค่อนข้างมาก ซึ่งบริเวณดังกล่าวก็คือความคลาดเคลื่อนของสีนั่นเอง

เมื่อทำการพิจารณาสมการที่ (3) จะพบว่า สมการนี้อธิบายถึงความแตกต่างโดยเฉลี่ยระหว่างภาพไกด์ และส่วนที่เหมือนกันระหว่างภาพต้นฉบับและภาพไกด์ซึ่งก็คือโครงสร้างของภาพและสี เพราะฉะนั้นหากทำการถ่ายโอนโครงสร้างของภาพ และสีไปยังภาพผลลัพธ์ด้วยวิธีการกรองไกด์อิมเมจโดยตรง ผลลัพธ์ที่ได้คือความคลาดเคลื่อนของสีก็จะถูกถ่ายโอนไปยังภาพผลลัพธ์ด้วย ถึงแม้ว่าการที่นำภาพของชั้นสีเขียวมาเป็นภาพไกด์จะช่วยทำให้รายละเอียดของสีในภาพที่ชั้นสีแดง และสีน้ำเงินมีความใกล้เคียงกับรายละเอียดของสีภาพที่ชั้นสีเขียวมากขึ้นแล้วก็ตาม แต่ก็ยังไม่สามารถทำให้สีของภาพนั้นเข้าใกล้สีจริงตามธรรมชาติได้ เป็นเพียงการทำให้เข้าใกล้สีเทา หรือสีดำ ซึ่งสามารถลดทอนความคลาดเคลื่อนของสีได้ในระดับหนึ่ง

ดังนั้นหากทำการปรับลดความแตกต่างของสีในบริเวณที่เกิดความคลาดเคลื่อนของสีได้ ก็จะสามารถลดทอน หรือแก้ไขความคลาดเคลื่อนของสีได้อย่างเป็นธรรมชาติมากขึ้น หากเปรียบเทียบในมุมมองของพลังงานความเข้มของแสงแล้ว นั่นคือการปรับค่าพลังงานของสีแดงและสีน้ำเงินให้เหมาะสม โดยอ้างอิงจากพลังงานของสีเขียว เพราะฉะนั้นแนวคิดเบื้องต้นในการแก้ไขความคลาดเคลื่อนของสีจึงเป็นการนำค่าพลังงานเฉลี่ย (Root Mean Square Power) มาประยุกต์ใช้ โดยสามารถคำนวณได้จากสมการที่ (5)

$$X_i = \frac{X_i}{\sqrt{2}}, X_i = \{R_i, B_i \mid i \in CARs\} \quad (5)$$

เมื่อ X_i แทนค่าของสีในชั้นสีแดงและสีน้ำเงิน ที่ตำแหน่ง i ซึ่งเป็นตำแหน่งที่เกิดความคลาดเคลื่อนของสี เพื่อให้พลังงานความคลาดเคลื่อนของสีในชั้นสีแดงและสีน้ำเงินถูกแทนที่ด้วยค่าเฉลี่ยของพลังงาน ซึ่งมีแนวโน้มที่ความคลาดเคลื่อนสีของในบริเวณนั้นๆ จะลดลงอย่างเป็นธรรมชาติ

3.2 วิธีการวัดผลและเปรียบเทียบประสิทธิภาพ

ผู้วิจัยจะทำการวัดผลประสิทธิภาพของการทดลอง โดยใช้ Ratting Scale [10] โดยให้บุคคลทั่วไปพิจารณาภาพที่แก้ไขความคลาดเคลื่อนของสีแล้วจากวิธีการของ Jung และคณะ[7] ซึ่งเป็นการแก้บริเวณขอบภาพทั้งหมดด้วยวิธีการกรองไกด์อิมเมจ โดยต่อไปนี้จะเรียกว่าวิธีของ Jung และวิธีการที่นำเสนอ แล้วให้คะแนนความเหมือนกับภาพต้นฉบับที่ไม่มีความคลาดเคลื่อนของสี โดยที่คะแนนอยู่ในช่วง 1-5 เมื่อ 1 หมายถึง ไม่เหมือน และ 5 หมายถึงเหมือนมาก แล้วนำคะแนนเหล่านั้นจากการประเมินของบุคคลทั่วไปมาหาค่าเฉลี่ยเลขคณิต(Arithmetic Mean) ในการทดลองนี้มีจำนวน 5 คน ภาพที่ใช้ในการทดลองมีทั้งหมด 50 ภาพ เป็นภาพที่มีขนาด 150 x 100 พิกเซล เครื่องคอมพิวเตอร์ที่ใช้ในการประมวลผลภาพเพื่อทำการทดลองมีความเร็ว 2.50 GHz มีหน่วยความจำหลัก (RAM) 4 GB ใช้ระบบปฏิบัติการวินโดวส์ 8 แบบ 64 บิท (Bits)

โดยทั่วไปในการวัดคุณภาพของรูปภาพจะนิยมใช้เครื่องมือทางคณิตศาสตร์ เช่น Peak Signal to Noise Ratio (PSNR), Signal-to-Noise Ratio (SNR) หรือ Structural Similarity Index (SSIM) เป็นตัวชี้วัด แต่เนื่องจาก PSNR เป็นการตรวจสอบความแตกต่างของพิกเซลเท่านั้นจึงทำให้ไม่สามารถแยกแยะความต่างทางโครงสร้างของภาพและสีรวมทั้งข้อผิดพลาดอื่นๆ ได้ ทำให้ภาพที่มีค่า PSNR เท่ากัน อาจมีคุณภาพต่างกันเมื่อดูด้วยตาได้ สำหรับ SSIM นั้นได้รับการพัฒนาและนำเสนอโดย Wang และคณะ [12] เป็นการวัดคุณภาพของภาพจาก 3 ปัจจัย ได้แก่ ความสว่าง คอนทราสต์ และโครงสร้างของภาพ เพื่อกำจัดข้อด้อยของ PSNR และ SNR แต่ทว่าวิธีการนี้ยังมีข้อจำกัดคือค่าที่คำนวณได้อาจมีความแตกต่างกันมากในภาพที่มีคุณภาพเท่าหรือใกล้เคียงกันขึ้นอยู่กับพารามิเตอร์ที่ใช้ในการคำนวณ จึงทำให้มีความสัมพันธ์แบบไม่เชิงเส้น (Non-Linear Relationship) กับระบบการมองเห็นของมนุษย์ (Human Visual System) [13] ส่งผลให้ค่าที่วัดได้อาจไม่ตอบสนองต่อคุณภาพได้เทียบเท่ากับการมองเห็นของมนุษย์ ซึ่งข้อเสียนี้จะเห็นได้ชัดเมื่อทำการเปรียบเทียบภาพที่มีการเบลอ (Blurred Images) [14] ซึ่งภาพที่มีความคลาดเคลื่อนสีเป็นภาพที่มีลักษณะดังกล่าว ดังนั้นในงานวิจัยนี้จึงใช้วิธีการวัดประสิทธิภาพโดยใช้การประเมินจากบุคคลทั่วไป

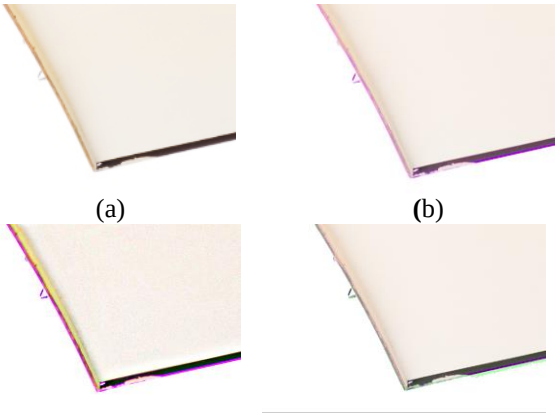
4. ผลการทดลอง

ในงานวิจัยครั้งนี้ ผู้วิจัยจะทำการเปรียบเทียบผลการทดลองกับวิธีของ Jung โดยมีรายละเอียด ดังนี้

4.1 การแก้ไขความคลาดเคลื่อนของสี

ผลการทดลองการแก้ไขความคลาดเคลื่อนของสีโดยใช้ค่า

RMS เปรียบเทียบผลการทดลองกับวิธีของ Jung ดังแสดงในภาพที่ 4, 5 และ 6

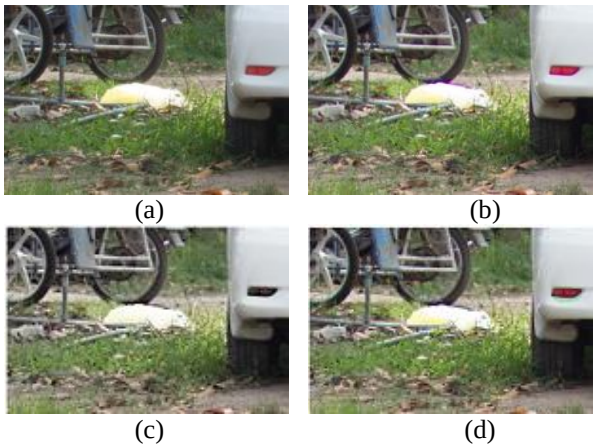


(a) (b)

(c) (d)

ภาพที่ 4: ผลการแก้ไขความคลาดเคลื่อนของสี ภาพต้นฉบับที่ไม่มีความคลาดเคลื่อนของสี (a)และ ภาพต้นฉบับที่มีความคลาดเคลื่อนของสี (b) ภาพที่แก้ไขโดยวิธีของJung (c) วิธีที่นำเสนอ (d)

จากภาพที่ 4 ภาพถ่ายต้นฉบับดังที่แสดงในภาพที่ 4 (a) ไม่มีความคลาดเคลื่อนของสี ภาพที่ 4 (b) มีความคลาดเคลื่อนของสีที่บริเวณขอบของวัตถุ จากการแก้ไขความคลาดเคลื่อนของสี จะเห็นว่าวิธีของ Jung ในภาพที่ 4 (c)นั้น ทำให้สีของบริเวณรอบข้างเปลี่ยนไปด้วย ส่วนการแก้ไขความคลาดเคลื่อนของสีด้วยวิธีที่นำเสนอในภาพที่ 4 (d) นั้นแก้ไขเฉพาะขอบของวัตถุ มีความใกล้เคียงกับภาพต้นฉบับในภาพที่ 4 (a)

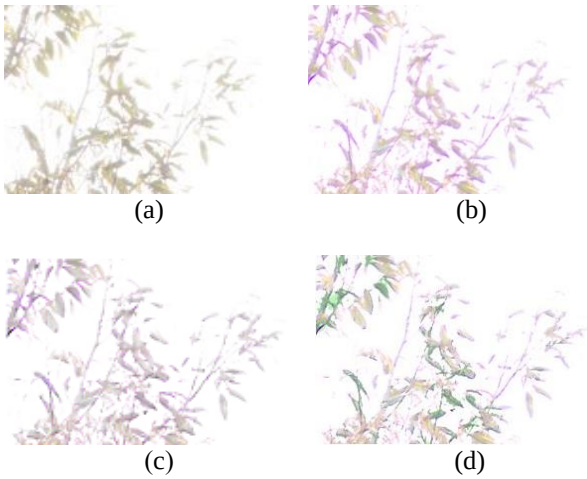


(a) (b)

(c) (d)

ภาพที่ 5: ผลการแก้ไขความคลาดเคลื่อนของสี ภาพต้นฉบับที่ไม่มีความคลาดเคลื่อนของสี (a) และ ภาพต้นฉบับที่มีความคลาดเคลื่อนของสี (b) ภาพที่แก้ไขโดยวิธีของ Jung (c) วิธีที่นำเสนอ (d)

จากภาพที่ 5 ภาพถ่ายต้นฉบับดังที่แสดงในภาพที่ 5 (a) ไม่มีความคลาดเคลื่อนของสี ภาพที่ 5 (b) จะเห็นว่าเกิดความคลาดเคลื่อนสีที่บริเวณกรอบไฟของท้ายรถยนต์และวัตถุที่อยู่กลางภาพ เมื่อใช้วิธีของ Jung ในภาพที่ 5 (c) นั้น บริเวณขอบวัตถุกลางภาพจะถูกแก้ไขให้เป็นดำและบริเวณกรอบไฟท้ายรถยนต์จะมีสีที่เปลี่ยนไปด้วย เมื่อใช้วิธีแก้ไขความคลาดเคลื่อนของสีด้วยวิธีที่นำเสนอในภาพที่ 5 (d) จะเห็นว่าบริเวณขอบวัตถุกลางภาพจะไม่ใช่สีดำและบริเวณกรอบไฟท้ายรถยนต์ยังเป็นสีแดงเพียงแต่ลดความเข้มลงไปเล็กน้อย



(a) (b)

(c) (d)

ภาพที่ 6: ผลการแก้ไขความคลาดเคลื่อนของสี ภาพต้นฉบับที่ไม่มีความคลาดเคลื่อนของสี (a)และ ภาพต้นฉบับที่มีความคลาดเคลื่อนของสี (b) ภาพที่แก้ไขโดยวิธีของ Jung (c) วิธีที่นำเสนอ (d)

จากภาพที่ 6 ภาพถ่ายต้นฉบับดังที่แสดงในภาพที่ 5 (a) ไม่มีความคลาดเคลื่อนของสี ภาพถ่ายต้นฉบับในภาพที่ 6 (b) มีความคลาดเคลื่อนของสีเกิดขึ้นบริเวณใบไม้ค่อนข้างมาก จากการแก้ไขความคลาดเคลื่อนของของสีด้วยวิธีของ Jung ในภาพที่ 6 (c) นั้น ทำให้สีของใบไม้รอบข้าง เปลี่ยนไปด้วย ส่วนการแก้ไขความคลาดเคลื่อนของสีด้วยวิธีที่นำเสนอในภาพที่ 6 (d) นั้นบริเวณใบไม้รอบข้างไม่ได้ถูกแก้ไขไปด้วย

ผู้วิจัยได้นำภาพที่แก้ไขความคลาดเคลื่อนของสีแล้วจากวิธีของ Jung และวิธีการที่นำเสนอ ให้บุคคลทั่วไปพิจารณาให้คะแนนความเหมือนกับภาพต้นฉบับที่ไม่มีความคลาดเคลื่อนของสีโดยมีผลคะแนนดังตารางที่ 1

ตารางที่ 1: ผลการประเมินคุณภาพของรูปภาพจากบุคคลทั่วไป

วิธีการวิจัย	คะแนนการพิจารณาของบุคคลทั่วไปคนที่					คะแนนเฉลี่ย
	1	2	3	4	5	
วิธีของ Jung	4	3	3	4	3	3.6
วิธีที่นำเสนอ	5	5	4	4	4	4.4

จากตารางที่ 1 จะเห็นว่าบุคคลทั่วไปจำนวน 5 คน ได้ให้คะแนนวิธีการที่นำเสนอเฉลี่ย 4.4 มากกว่าวิธีของ Jung ซึ่งได้คะแนนเฉลี่ย 3.6 แสดงให้เห็นว่าวิธีการที่นำเสนอนั้นสามารถแก้ไขความคลาดเคลื่อนของสีได้ดีกว่าซึ่งหากสังเกตจากผลลัพธ์รูปภาพที่นำมาเปรียบเทียบกัน จะพบว่าวิธีการที่นำสามารถแก้ไขเฉพาะบริเวณที่เกิดความคลาดเคลื่อนของสีได้ โดยไม่มีผลกระทบกับบริเวณใกล้เคียง

5. สรุปผลและข้อเสนอแนะ

งานวิจัยนี้นำเสนอวิธีการแก้ไขความคลาดเคลื่อนของสีด้วยการกรองไกด์อิมเมจโดยนำคุณลักษณะของโครงสร้างของภาพและสีมาพิจารณาเพื่อใช้ในการแก้ไขความคลาดเคลื่อนของสีที่เกิดขึ้นในภาพ โดยมีสมมติฐานว่าความคลาดเคลื่อนสีเป็นพลังงานของความแตกต่างของสี ดังนั้นจึงประยุกต์หลักการของ RMS Power ในการแก้ไขความคลาดเคลื่อนสี และจากผลการทดลองเบื้องต้น (Preliminary Experiment) เมื่อให้บุคคลทั่วไปทำการประเมินผลโดยการให้คะแนนความเหมือนกับภาพต้นฉบับที่ไม่มีความคลาดเคลื่อนของสี พบว่าวิธีที่นำเสนอได้รับคะแนนเฉลี่ยจากบุคคลทั่วไปว่ามีประสิทธิภาพในการแก้ไขความคลาดเคลื่อนสีเป็น 4.4 ซึ่งมากกว่าวิธีของ Jung ที่ได้คะแนนเฉลี่ย 3.6

ถึงแม้ว่าวิธีการที่นำเสนอนั้นจะได้ผลที่ดีกว่าวิธีการที่ผ่านมาแต่จากการศึกษาการกรองไกด์อิมเมจผู้วิจัยพบว่าคุณสมบัติโครงสร้างของรูปภาพและโครงสร้างรายละเอียดของสีนั้นมีความน่าสนใจในเชิงสถิติและสามารถนำมาวิเคราะห์เพื่อหาแนวทางที่เหมาะสมในการลดทอนค่าของสีในบริเวณที่เกิดความคลาดเคลื่อนสีให้ดีขึ้นได้ ซึ่งเป็นวิธีหนึ่งที่จะนำไปพัฒนาและปรับปรุงการแก้ไขความคลาดเคลื่อนสีเพิ่มเติมต่อไป

เอกสารอ้างอิง

[1] T. Luhmann, H .Hastedt and W.Tecklenburg, "Modelling of Chromatic Aberration For High Precision Photogrammetry," *IAPRS Volume XXXVI*, Part 5, Dresden 25-27, 2006.

[2] E. Hecht, *Optics. 3rd*, Addison-Wesley Publishing Company: 150-155, 1998.

[3] S.W.Chung, B.K. Kim and W. J.Song, "Removing Chromatic aberration by digital image processing," *Optical Engineering*, vol. 46, no. 6, pp. 067002-1–067002-10, 2010.

[4] S. B. Kang, "Automatic removal of purple fringing from Images," *U.S. Patent 7577292*, Washington, DC: U.S. Patent and Trademark Office, 2009.

[5] H. J. Ju, D. K. Lee and R. H. Park." Chromatic Aberration Removal in Narrow Color Regions of Digital Images," *16th IEEE International Conference on Consumer Electronics*, pp. 31-32, 2012.

[6] B. K. Kim and R. H. Park, "Automatic detection and correction of purple fringing using the gradient information and desaturation," *2008 16th European Signal Processing Conference*, Lausanne, pp. 1-5, 2008.

[7] C.D.Jung, D.W.Jang, H.S.Kim and R. H. Park, "Color Fringe Correction Using Guided Image Filtering" *18th IEEE International Conference on Consumer Electronics*, pp. 1-2, 2014.

[8] ยุทธศักดิ์ จินาพร, สุวัจชัย กมลสันติโรจน์ และ สรร รัตนสัญญา, "การศึกษาวิธีการค้นหาและแก้ไขความผิดเพี้ยนของสีในภาพดิจิตอล," *การประชุมวิชาการระดับประเทศด้านเทคโนโลยีสารสนเทศ ครั้งที่ 8 (NCIT 2016)*, หน้า 528-533, 2559.

[9] ณัฐชัย งามเสงี่ยม, สุวัจชัย กมลสันติโรจน์ และ สรร รัตนสัญญา, "การตรวจจับความคลาดเคลื่อนสีอัตโนมัติโดยใช้คุณสมบัติเฉพาะที่," *วารสารสถิติประยุกต์และเทคโนโลยีสารสนเทศ* ปีที่ 1 ฉบับที่ 2 กรกฎาคม-ธันวาคม พ.ศ. 2559 หน้า 8-23.

[10] ชัชวาลย์ เรืองประพันธ์, "*สถิติพื้นฐาน*" ภาพวิชาสถิติ คณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น, 2556.

[11] K. He, J. Sun, and X. Tang, "Guided image filtering," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 35, no. 6, pp. 1397–1409, 2013.

[12] Z.Wang"Image quality assessment: from error visibility to structural imilarity," *IEEE Transactions on Image Processing* ,vol 13, Issue 4, 2004.

[13] K.Okarma, "Quality Assessment of images with Multiple Distortions using Combined Metrics," *ELEKTRONIKA IR ELEKTROTECHNIKA*, ISSN 1392 1215, VOL. 20, NO. 6, 2014.

[14] Z.Wang , L.Li , S.Wu , Yi Xia , Z. Wan and C. Cai, "A New Image Quality Assesement Algorithm based on SSIM and Multiple Regressions" *International Journal of Signal Processing, Image Processing and Pattern Recognition*, Vol.8, No.11,2015.

การปรับปรุงไกด์อิมเมจฟิลเตอร์ในการค้นหาความคลาดเคลื่อนสีในภาพดิจิทัล
Improvement on Guided Image Filter for Chromatic Aberration Detection

ยุทธศักดิ์ จินาพร (Y.Jinaporn)[1] ปรวัฒน์ วิสูตรศักดิ์ (P.Visutsak)[2] และสรร รัตนสัญญา (S. Ratanasanya)[3]
ภาควิชาวิทยาการคอมพิวเตอร์และสารสนเทศ คณะวิทยาศาสตร์ประยุกต์ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
[1]yuttasak.ji@rd.go.th, [2]porawat.v@sci.kmutnb.ac.th, [3]san.ratanasanya@gmail.com

บทคัดย่อ

ความคลาดเคลื่อนสีที่เกิดขึ้นกับภาพถ่ายดิจิทัล (Digital Photo) เป็นหนึ่งในสาเหตุหลักที่ทำให้คุณภาพของภาพลดลง บทความนี้นำเสนอการนำปรับปรุง การกรองไกด์อิมเมจ (Guided Image Filtering) เพื่อให้สามารถตรวจจับและแยกความแตกต่างระหว่างความคลาดเคลื่อนสี และสีที่เกิดขึ้นจริงตามธรรมชาติได้อย่างมีประสิทธิภาพ โดยใช้ลักษณะของความคลาดเคลื่อนสีและคุณสมบัติของการกรองไกด์อิมเมจมาช่วยตรวจสอบบริเวณที่เกิดความคลาดเคลื่อนสีขึ้นจริง ผลการทดลองแสดงให้เห็นว่าวิธีการที่นำเสนอสามารถแยกบริเวณที่เกิดความคลาดเคลื่อนสี ออกจากบริเวณที่มีสีจริงในธรรมชาติได้อย่างถูกต้องถึงร้อยละ 95.29 โดยเฉลี่ย

คำสำคัญ: ความคลาดเคลื่อนสี, การกรองไกด์อิมเมจ, ตัวกรองไกด์, การค้นหาความคลาดเคลื่อนสี, ขอบสี

Abstract

Chromatic Aberration (CA) in digital images is one of the main cause to degrade the quality of those images. This paper presents the improvement of Guided Image Filtering in order for it to be able to detect and distinguish CA and non-CA area effectively. The proposed method exploits characteristics of CA and properties of Guided Image Filtering together to determine whether it is the CA area or not. The experimental results show that the proposed method can ascertain CA areas in the images with the accuracy of 95.29% on the average.

Keyword: Chromatic Aberration, Guided Image Filtering, Guided Filter, Chromatic Aberration Detection, Color Fringe.

1. บทนำ

เนื่องจากเทคโนโลยีการถ่ายภาพ แต่งภาพ ที่ทันสมัย และกล้องดิจิทัล (Digital Camera) ทำให้ผู้คนนิยมถ่ายภาพมากขึ้น เพื่อตอบสนองต่อความต้องการของผู้บริโภค ผู้ผลิตจึงลดต้นทุนการผลิต โดยใช้ เลนส์ (Lens) คุณภาพต่ำลง เพื่อให้ราคาของกล้องดิจิทัลถูกลง ส่งผลให้เกิดความคลาดเคลื่อนสีได้มากขึ้น ทำให้คุณภาพของภาพถ่ายลดลงด้วย เนื่องจาก เมื่อแสงขาว ซึ่งประกอบไปด้วยแสงสีต่างๆ ที่มีความยาวคลื่น (Wavelength) แตกต่างกัน เดินทางผ่านเลนส์คุณภาพต่ำ จะเกิดการกระจายของแสงได้ง่าย ทำให้แสงแต่ละสีตกกระทบลงบนเซ็นเซอร์รับภาพ (Image Sensor) ไม่ตรงจุดโฟกัส (Focal Point) จึงเกิดความคลาดเคลื่อนสีขึ้นในภาพถ่ายดิจิทัล โดยจะเกิดขึ้นบริเวณขอบของวัตถุ [1]

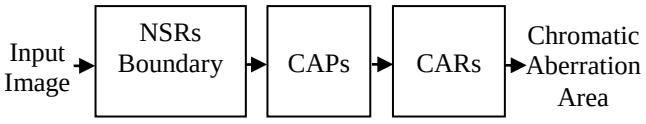
ปัจจุบันการแก้ปัญหาความคลาดเคลื่อนสีได้แบ่งออกเป็น 2 แนวทาง คือ การใช้เลนส์อโปโครแมท (Apochromat Lens) ที่มีการกระจายแสงต่ำ ซึ่งเป็นเลนส์ชนิดพิเศษที่ช่วยลดการกระจายแสง แต่ตัวเลนส์มีความเปราะบางและราคาค่อนข้างสูง [2] และการใช้โปรแกรมประยุกต์ในการค้นหาและแก้ไขบริเวณที่เกิดความคลาดเคลื่อนสี แต่อย่างไรก็ตาม โปรแกรมประยุกต์ต่าง ๆ ที่มีในปัจจุบัน ยังไม่สามารถแก้ไขปัญหานี้ได้โดยอัตโนมัติ จึงจำเป็นต้องให้ผู้ใช้ ช่วยในการกำหนดบริเวณที่ต้องการค้นหา หรือแก้ไข ถึงแม้ว่าแนวทางนี้จะมีค่าใช้จ่ายน้อยและมีความยืดหยุ่นสูง แต่ใช้ระยะเวลาในการแก้ไขภาพค่อนข้างนาน และคุณภาพของผลลัพธ์ที่ได้ขึ้นกับความชำนาญ

ส่วนบุคคล จึงมีผู้วิจัยหลายกลุ่มสนใจที่จะแก้ไขปัญหานี้ได้โดยอัตโนมัติ ซึ่งยุทธศักดิ์ และคณะ [3] ได้แบ่งงานวิจัยที่เกี่ยวกับความคลาดเคลื่อนสีที่ผ่านมาออกเป็น 2 กลุ่ม กลุ่มที่หนึ่ง มีการค้นหาและแก้ไขความคลาดเคลื่อนสี [4], [5], [6], [7], [8] ซึ่งงานวิจัยในกลุ่มนี้ใช้เวลาค่อนข้างนาน เนื่องจากมีการค้นหาในทุกพิกเซล (Pixel) แล้วจึงทำการแก้ไขความคลาดเคลื่อนสีจึงทำให้ต้องใช้เวลาในการประมวลผลช้าที่สุด หรือบิ๊กโอ (Big O) เป็น $O(n^2)$ โดยที่ n คือจำนวนพิกเซล แต่มีเพียงวิธีการที่นำเสนอโดย ณัฐชัย และคณะ [8] ที่สามารถแยกแยะความคลาดเคลื่อนสีออกจากสีที่เกิดขึ้นตามธรรมชาติได้ จากการประยุกต์ใช้ทิศทางของเกรเดียนต์ ทำให้ได้ผลลัพธ์ที่ถูกต้องมากขึ้นอย่างมีนัยสำคัญ ส่วนกลุ่มที่สอง ไม่มีการค้นหา แต่มีการแก้ไขความคลาดเคลื่อนสี โดยใช้การกรองไกด์อิมเมจ (Guided Image Filtering) [9] ในการแก้ไขความคลาดเคลื่อนสี ซึ่งสามารถแก้ไขได้อย่างรวดเร็ว แต่ทว่าวิธีนี้ก็ยังไม่สามารถแยกแยะความแตกต่างของสีที่เกิดขึ้นตามธรรมชาติและสีที่เกิดจากความคลาดเคลื่อนสีได้

ดังนั้นงานวิจัยนี้จึงนำเสนอวิธีการปรับปรุงการกรองไกด์อิมเมจ โดยนำคุณลักษณะของบริเวณที่เกิดความคลาดเคลื่อนสีและคุณสมบัติของตัวกรองไกด์อิมเมจมาใช้ในการแยกแยะความแตกต่างของสีที่เกิดจากความคลาดเคลื่อนสีและสีที่เกิดขึ้นตามธรรมชาติ

2. วรรณกรรมที่เกี่ยวข้อง

ณัฐชัย และคณะ [8] ได้ทำการพัฒนากรอบการทำงาน (Framework) ของการตรวจจับความคลาดเคลื่อนสีแบบอัตโนมัติจากการศึกษาจุดเด่นของงานวิจัยก่อนหน้า [4], [5], [6], [7] ดังแสดงในภาพที่ 1



ภาพที่ 1: กรอบการทำงานของการตรวจจับความคลาดเคลื่อนสีแบบอัตโนมัติของณัฐชัย และคณะ [8]

จากภาพที่ 1 กรอบการทำงานนี้ เริ่มจากการหาบริเวณที่มีค่าความสว่างภายในรูปภาพสูง หรือ Near Saturated Regions (NSRs) เนื่องจากความคลาดเคลื่อนสีโดยมากเกิดที่ขอบของวัตถุที่อยู่ใกล้กับบริเวณที่มีแสงสว่างมาก จากนั้นทำการหาบริเวณที่น่าจะเกิดความคลาดเคลื่อนสี หรือ Candidate Regions (CRs) โดยดูจากค่าของเส้นขอบและคอนทราสต์ (Contrast) ในบริเวณโดยรอบ NSRs เมื่อได้ CRs แล้ว จึงนำบริเวณดังกล่าวมาทำการตรวจสอบว่าเป็นบริเวณที่เกิดความคลาดเคลื่อนสีจริงหรือไม่ โดยการใช้ทิศทางของเกรเดียนต์ (Gradient) ของสี และเพื่อให้ได้บริเวณที่เกิดความคลาดเคลื่อนสีที่ถูกต้องมากขึ้น จึงใช้วิธีการ Dilation ช่วยในการตรวจสอบด้วย แต่อย่างไรก็ตามวิธีการนี้ต้องตรวจสอบทีละพิกเซล จึงใช้เวลาในการประมวลผลนาน การนำตัวกรอง (Filter) มาประยุกต์ใช้จึงอาจสามารถแก้ปัญหาเรื่องความเร็วในการประมวลผลได้

ในปีค.ศ. 2014 Jung และคณะ [9] ได้นำเสนอวิธีที่เรียกว่าการกรองไกด์อิมเมจ (Guided Image Filtering) ซึ่งเป็นวิธีการปรับปรุงคุณภาพของภาพโดยการหาความสัมพันธ์ระหว่างภาพ 2 ภาพในรูปแบบเชิงเส้น จึงใช้เวลาในการประมวลผลเป็น O(n) และนำเอาลักษณะเด่นของภาพแต่ละภาพออกมาผสานกัน โดยที่ทั้งสองภาพนั้นอาจจะเป็นภาพเดียวกัน หรือเป็นภาพที่มีความแตกต่างกันก็ได้ ซึ่งวิธีการกรองไกด์อิมเมจจะถ่ายโอนโครงสร้างของภาพ I หรือภาพไกด์ (Guidance Image) ไปยัง q หรือ ภาพผลลัพธ์ (Output Image) ดังสมการที่ 1

$$q_i = (\overline{a_i} I_i + \overline{b_i}) \quad (1)$$

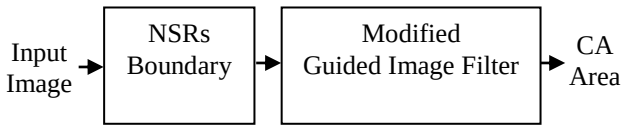
โดยที่ $\overline{a_i}$ และ $\overline{b_i}$ เป็นค่าเฉลี่ยสัมประสิทธิ์เชิงเส้น และ i เป็นจำนวนพิกเซลภายในพื้นที่การคำนวณ [10]

ด้วยการใช้ภาพที่ชั้นสีเขียวเป็นภาพไกด์ จะทำให้ค่าของภาพที่ชั้นสีแดงและชั้นสีน้ำเงินเข้าใกล้ค่าของชั้นสีเขียวมากที่สุด เปรียบเสมือนแสงทั้ง 3 สี ตกลงตรงจุดรับภาพในตำแหน่งเดียวกัน ซึ่งจะทำให้ไม่เกิดความคลาดเคลื่อนสีที่บริเวณขอบภาพ เพราะค่าสีเขียว ค่าสีแดง และค่าสีน้ำเงินตรงบริเวณที่ไม่มีความคลาดเคลื่อนสีจะมีค่าที่ใกล้เคียงกัน แต่เนื่องจากคุณสมบัติโดยตรงของการกรองไกด์อิมเมจเป็นเพียงวิธีที่ใช้ปรับปรุงคุณภาพของภาพให้มีความสมบรูณ์มากขึ้น จึงยังไม่

สามารถใช้ค้นหาบริเวณที่เกิดความคลาดเคลื่อนสี ทำให้ไม่สามารถแก้ไขเฉพาะบริเวณที่เกิดความคลาดเคลื่อนสีได้ และยังไม่สามารถแยกแยะความแตกต่างของสีที่เกิดขึ้นตามธรรมชาติและสีที่เกิดจากความคลาดเคลื่อนสีได้อีกด้วย

3. วิธีการดำเนินการวิจัย

จากการศึกษางานวิจัยที่เกี่ยวข้อง ผู้วิจัยจึงมีแนวคิดที่จะปรับปรุงวิธีการค้นหาบริเวณที่เกิดความคลาดเคลื่อนสีโดยใช้การกรองไกด์อิมเมจเป็นพื้นฐาน เพื่อเพิ่มความรวดเร็วในการประมวลผล อีกทั้งยังได้พัฒนาการค้นหาและแยกแยะบริเวณที่เกิดความคลาดเคลื่อนสีสำหรับการกรองไกด์อิมเมจ เพื่อให้มีประสิทธิภาพ และความถูกต้องที่มากขึ้น จึงได้ทำการออกแบบกรอบการทำงานดังแสดงในภาพที่ 2



ภาพที่ 2: กรอบการทำงานของการค้นหาบริเวณที่เกิดความคลาดเคลื่อนสีที่นำเสนอ

จากภาพที่ 2 กรอบการทำงานนี้ นำการหา NSRs มารวมกับการกรองไกด์อิมเมจที่ผู้วิจัยได้ปรับปรุงให้สามารถแยกแยะความคลาดเคลื่อนสีได้ ซึ่งรายละเอียดของขั้นตอนต่าง ๆ มีดังนี้

3.1 NSRs Boundary

เนื่องจากบริเวณที่เกิดความคลาดเคลื่อนสีจะมีค่าความสว่างมากกว่าร้อยละ 90 ขึ้นไป [5] จึงทำการหาบริเวณที่เกิดแสงสว่างภายในภาพ โดยการเปลี่ยนระบบสีจากเดิมที่เป็น RGB ให้เป็นระบบสี YCbCr จากนั้นทำการหาบริเวณที่เกิดแสงสว่างโดยกำหนดค่าขีดแบ่ง (T) เป็น 0.8 ซึ่งเท่ากับร้อยละ 90 ของค่าความสว่าง (Y) ดังสมการที่ 2

$$NSR = \{Y_{(i,j)} \in I \mid Y_{(i,j)} > T\} \quad (2)$$

เมื่อ NSR คือ บริเวณที่เกิดแสงสว่างภายในภาพ Y หมายถึงค่าของแสงภายในภาพในระบบสี YCbCr ที่พิกัด (i,j) ของภาพ I

3.2 การค้นหาความคลาดเคลื่อนสีด้วยตัวกรองไกด์อิมเมจ

ผู้วิจัยนำเสนอการปรับปรุงการกรองไกด์อิมเมจเพื่อให้สามารถค้นหาบริเวณที่เกิดความคลาดเคลื่อนสีได้ โดยทำการปรับปรุงอัลกอริทึม (Algorithm) ในส่วนของการไกด์ภาพ โดยแทนค่าตัวแปร I ให้อยู่ในรูปของค่าความแตกต่างกันของสีดังสมการที่ 3 และ 4 ตามลำดับ

$$I_i = R_i - G_i \quad (3)$$

$$\overline{I_i} = \overline{R_i} - \overline{G_i} \quad (4)$$

โดยที่ R_i คือค่าของสีแดง G_i คือค่าของสีเขียว $\overline{R_i}$ คือค่าเฉลี่ยของสีแดง และ $\overline{G_i}$ คือค่าเฉลี่ยของสีเขียว เนื่องด้วยคุณลักษณะเฉพาะของบริเวณที่เกิดความคลาดเคลื่อนสี ค่าของสีแดง และสีน้ำเงินจะแตกต่างจากค่าของสีเขียวอย่างเห็นได้ชัด และเนื่องจากระบบการมองเห็นของมนุษย์ (Human Visual System) มีความอ่อนไหว (Sensitive) ต่อสีแดงมากที่สุด [11] ดังนั้นหากใช้ค่าความต่างของสีแดงกับสีอื่น ๆ แทนตัวภาพไกด์แล้ว ด้วยคุณลักษณะดังกล่าวก็จะทำให้ เห็นความแตกต่างของบริเวณที่เกิดความคลาดเคลื่อนสีกับบริเวณที่เป็นสีจริงในธรรมชาติได้อย่างชัดเจนมากขึ้น และทำให้บริเวณดังกล่าวแยกออกจากกันได้อย่างชัดเจน จากสมการที่ 1 ตัวแปร $\overline{b_i}$ ซึ่งเป็นส่วนที่ให้โครงสร้างของสีภายในภาพนั้นตามที่นำเสนอใน [10] สามารถคำนวณได้จาก

$$\overline{b_i} = \overline{p_i} - a_i \overline{I_i} \quad (5)$$

โดยที่ $\overline{p_i}$ คือค่าเฉลี่ยตัวกรองภาพ a_i คือสัมประสิทธิ์เชิงเส้น $\overline{I_i}$ คือค่าเฉลี่ยภาพไกด์ ซึ่งเมื่อนำค่าของตัวแปร $\overline{b_i}$ มาแทนลงในสมการที่ 1 พบว่า

$$q_i = (\overline{a_i} I_i + \overline{p_i} - a_i \overline{I_i}) \quad (6)$$

เมื่อนำค่าความแตกต่างของสีมาเป็นตัวไกด์ดังนำเสนอในสมการที่ 3 และ 4 มาแทนค่าลงในสมการที่ 6 และทำการจัดรูปสมการใหม่ จะได้สมการที่แสดงความสัมพันธ์ระหว่างค่าความคลาดเคลื่อนสีและค่าความแตกต่างของสีดังแสดงในสมการที่ 7

$$q_i = (\overline{a_i} R_i + b_i) - (\overline{a_i} G_i - a_i \overline{G_i}) \quad (7)$$

โดยที่ q_i คือบริเวณที่เกิดความคลาดเคลื่อนสีที่เกิดจากค่าความแตกต่างของสีแดงกับสีเขียว b_i คือสัมประสิทธิ์เชิงเส้น และทำเช่นเดียวกันกับชั้นของสีแดงและสีน้ำเงิน

3.3 วิธีการวัดผลและเปรียบเทียบประสิทธิภาพ

ในการทดลอง ผู้วิจัยใช้ภาพทั้งหมด 50 ภาพ โดยแบ่งภาพออกเป็น 2 กลุ่ม กลุ่มละ 25 ภาพคือ กลุ่มภาพที่เกิดความคลาดเคลื่อนสีที่ไม่มีพื้นหลัง และกลุ่มภาพที่เกิดความคลาดเคลื่อนสีที่มีพื้นหลัง โดยขนาดของภาพที่ใช้ในการทดลองคือ 370 x 247 พิกเซล และทำการเปรียบเทียบประสิทธิภาพของวิธีที่นำเสนอ กับวิธีการประยุกต์ใช้ทิศทางของเกรเดียนต์ที่นำเสนอใน [8] ซึ่งเป็นวิธีการที่มีประสิทธิภาพและให้ค่าความถูกต้องในการค้นหาความคลาดเคลื่อนสีมากกว่างานวิจัยอื่น ๆ ที่ผ่านมา โดยวัดค่าความถูกต้อง (Accuracy) และค่าความผิดพลาด (False Positive Rate) ด้วยวิธี Confusion Matrix ซึ่งเป็นวิธีการเปรียบเทียบระหว่างผลลัพธ์ของแต่ละวิธีการกับการค้นหาความคลาดเคลื่อนสีด้วยมนุษย์ที่เป็นช่างภาพอิสระจำนวน 3 ท่าน ดังสมการที่ 8 และ 9 ตามลำดับ

$$\text{Accuracy} = \left(\frac{TP + TN}{TP + TN + FP + FN}\right) * 100 \quad (8)$$

$$\text{FPR} = \left(\frac{FP}{TN + FP}\right) * 100 \quad (9)$$

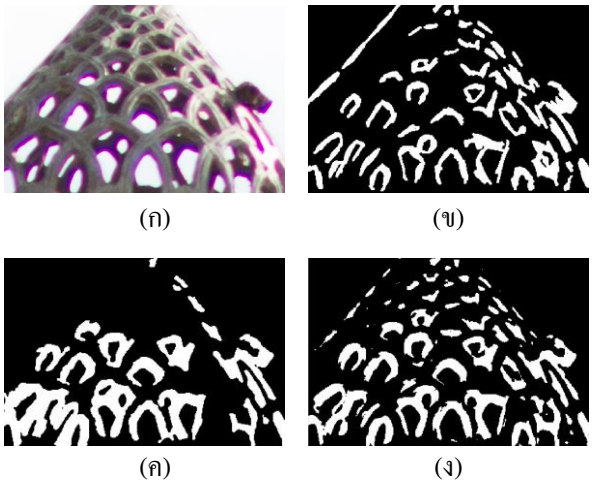
นอกจากนั้นยังทำการเปรียบเทียบเวลาในการประมวลผลอีกด้วย ในการทดลองนี้ใช้เครื่องคอมพิวเตอร์ที่มีความเร็วอยู่ที่ 2.80 GHz มีหน่วยความจำหลัก (RAM) อยู่ที่ 4 GB ใช้ระบบปฏิบัติการวินโดวส์ 7 แบบ 64 bits (Windows 7-64bits) และภาษาโปรแกรมที่ใช้คือ MATLAB เวอร์ชั่น 2013a

4. ผลการทดลอง

เพื่อให้เห็นประสิทธิภาพของการค้นหาบริเวณที่เกิดความคลาดเคลื่อนสี ผู้วิจัยจึงได้แบ่งผลการทดลองออกเป็น 3 ส่วน ดังต่อไปนี้

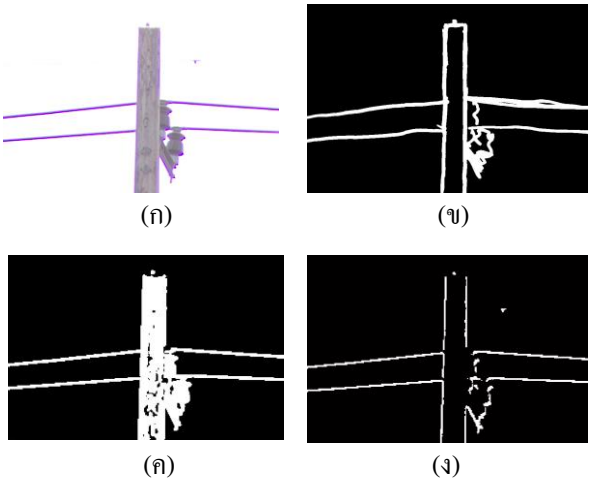
4.1 ผลการทดลองกลุ่มภาพที่ไม่มีพื้นหลัง

ภาพที่ 3 และ 4 แสดงตัวอย่างของผลการทดลองจากภาพทั้งหมดในกลุ่ม



(ก) (ข) (ค) (ง)

ภาพที่ 3: ผลการทดลองค้นหาความคลาดเคลื่อนสี (ก) ภาพต้นฉบับ (ข) ผู้เชี่ยวชาญ (ค) วิธีการใช้ทิศทางของเกรเดียนต์ (ง) วิธีการที่นำเสนอ



(ก) (ข) (ค) (ง)

ภาพที่ 4: ผลการทดลองค้นหาความคลาดเคลื่อนสี (ก) ภาพต้นฉบับ (ข) ผู้เชี่ยวชาญ (ค) วิธีการใช้ทิศทางของเกรเดียนต์ (ง) วิธีการที่นำเสนอ

จากภาพที่ 3 และ 4 แสดงให้เห็นว่าวิธีการที่นำเสนอสามารถค้นหาบริเวณที่เกิดความคลาดเคลื่อนสีได้เพิ่มมากขึ้นและใกล้เคียงกับการค้นหาด้วยผู้เชี่ยวชาญ ซึ่งมากกว่าวิธีการใช้ทิศทางของเกรเดียนต์ เนื่องจากวิธีการดังกล่าวใช้ทิศทางของเกรเดียนต์ในการคัดแยกบริเวณที่เกิดความคลาดเคลื่อนสีกับสีจริงในธรรมชาติตรงบริเวณขอบภาพ ซึ่งในบริเวณนี้จะมีค่าเกรเดียนต์ใกล้เคียงกับค่าความสว่างของแสง จึงส่งผลให้ไม่สามารถคัดแยกบริเวณที่เกิดความคลาดเคลื่อนสีในบริเวณดังกล่าวได้ ซึ่งต่างจากวิธีการที่นำเสนอที่ใช้ค่าความแตกต่าง

ของสีเป็นตัวไกด์ ความสว่างของแสงจึงไม่ส่งผลต่อการคัดแยกความคลาดเคลื่อนสีออกจากสีที่เกิดขึ้นตามธรรมชาติ และเมื่อทดลองกับภาพทั้งหมดและนำไปวัดผลกับผู้เชี่ยวชาญ ค่าความถูกต้องและค่าความผิดพลาดเฉลี่ยที่ได้แสดงในตารางที่ 1

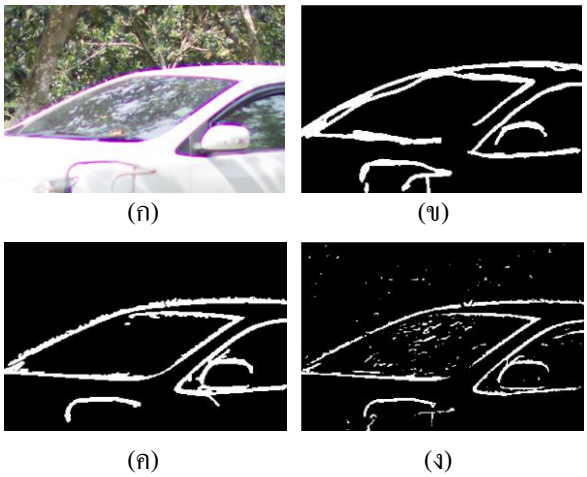
ตารางที่ 1: ผลการวัดค่าความถูกต้องและค่าความผิดพลาดเฉลี่ยของกลุ่มภาพที่ไม่มีพื้นหลัง

วิธีการวิจัย	ค่าความถูกต้องเฉลี่ย (ร้อยละ)	ค่าความผิดพลาดเฉลี่ย (ร้อยละ)
วิธีการใช้ทิศทางของเกรเดียนต์	93.45	5.98
วิธีที่การนำเสนอ	95.29	1.99

จากตารางที่ 1 แสดงให้เห็นว่าวิธีการที่นำเสนอมีค่าความถูกต้องเฉลี่ยอยู่ที่ร้อยละ 95.29 และมีค่าความผิดพลาดเฉลี่ยอยู่ที่ร้อยละ 1.99 ส่วนวิธีการใช้ทิศทางของเกรเดียนต์ มีค่าความถูกต้องเฉลี่ยอยู่ที่ร้อยละ 93.45 และค่าความผิดพลาดเฉลี่ยอยู่ที่ร้อยละ 5.98 จากค่าที่ได้แสดงให้เห็นว่าวิธีการที่นำเสนอสามารถค้นหาบริเวณที่เกิดความคลาดเคลื่อนสีได้เพิ่มมากขึ้นกว่าวิธีการใช้ทิศทางของเกรเดียนต์ ซึ่งสอดคล้องกับภาพที่ได้จากการทดลอง

4.2 ผลการทดลองกลุ่มภาพที่มีพื้นหลัง

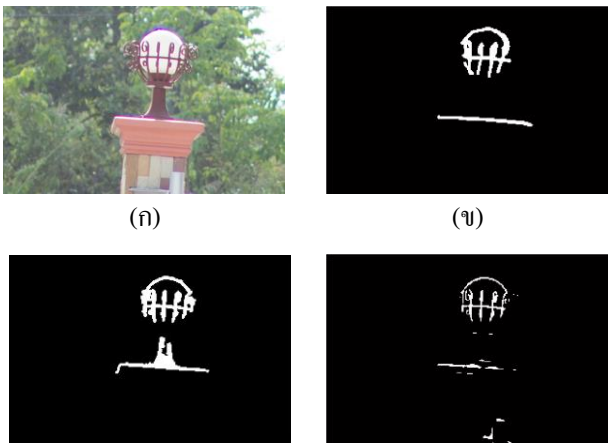
ภาพที่ 5 และ 6 แสดงตัวอย่างของผลการทดลองจากภาพทั้งหมดในกลุ่ม



(ก) (ข)

(ค) (ง)

ภาพที่ 5: ผลการทดลองค้นหาความคลาดเคลื่อนสี (ก) ภาพต้นฉบับ (ข) ผู้เชี่ยวชาญ (ค) วิธีการใช้ทิศทางของเกรเดียนต์ (ง) วิธีการที่นำเสนอ



(ก) (ข)

(ค) (ง)

ภาพที่ 6: ผลการทดลองค้นหาความคลาดเคลื่อนสี (ก) ภาพต้นฉบับ (ข) ผู้เชี่ยวชาญ (ค) วิธีการใช้ทิศทางของเกรเดียนต์ (ง) วิธีการที่นำเสนอ

จากภาพที่ 5 และ 6 แสดงให้เห็นว่าวิธีการที่นำเสนอค้นหาบริเวณที่เกิดความคลาดเคลื่อนสีได้ใกล้เคียงกับผู้เชี่ยวชาญ และเมื่อทดลองกับภาพทั้งหมดในกลุ่มนี้ค่าเฉลี่ยที่ได้แสดงในตารางที่ 2

ตารางที่ 2: ผลการวัดค่าความถูกต้องและค่าความผิดพลาดเฉลี่ยของกลุ่มภาพที่มีพื้นหลัง

วิธีการวิจัย	ค่าความถูกต้องเฉลี่ย (ร้อยละ)	ค่าความผิดพลาดเฉลี่ย (ร้อยละ)
วิธีการใช้ทิศทางของเกรเดียนต์	91.93	5.76
วิธีการที่นำเสนอ	94.46	2.34

จากตารางที่ 2 แสดงให้เห็นว่าวิธีการที่นำเสนอมีค่าความถูกต้องเฉลี่ยอยู่ที่ร้อยละ 94.46 และมีค่าความผิดพลาดเฉลี่ยอยู่ที่ร้อยละ 2.34 ส่วนวิธีการใช้ทิศทางของเกรเดียนต์ มีค่าความถูกต้องเฉลี่ยอยู่ที่ร้อยละ 91.93 และค่าความผิดพลาดเฉลี่ยอยู่ที่ร้อยละ 5.76 จากค่าที่ได้แสดงให้เห็นว่าวิธีการที่นำเสนอสามารถค้นหาบริเวณที่เกิดความคลาดเคลื่อนสีได้เพิ่มมากขึ้นกว่าวิธีการใช้ทิศทางของเกรเดียนต์ ซึ่งสอดคล้องกับภาพที่ได้จากการทดลอง

4.3 เวลาในการประมวลผล

จากการทดลองกับกลุ่มภาพทั้งหมด จึงทำให้ได้ผลของเวลาเฉลี่ยในการประมวลผล ซึ่งได้แสดงในตารางที่ 3

ตารางที่ 3: เวลาเฉลี่ยในการประมวลผล

วิธีการวิจัย	ค่าเฉลี่ยของเวลาในการประมวลผล (วินาที)
วิธีการใช้ทิศทางของเกรเดียนต์	1564.64
วิธีการที่นำเสนอ	0.85

จากตารางที่ 3 แสดงให้เห็นว่า วิธีการที่นำเสนอนั้นมีค่าเฉลี่ยของเวลาในการประมวลผลอยู่ที่ 0.85 วินาที และวิธีการใช้ทิศทางของเกรเดียนต์ มีค่าเฉลี่ยอยู่ที่ 1564.64 วินาที จากค่าเฉลี่ยของเวลาที่ได้จึงทำให้ทราบว่าวิธีการที่นำเสนอมีความรวดเร็วในการประมวลผลมากกว่าวิธีการใช้ทิศทางของเกรเดียนต์อย่างเห็นได้ชัด สาเหตุที่ทำให้วิธีการดังกล่าวมีความล่าช้าในการประมวลผลเนื่องจากมีการค้นหาบริเวณที่เกิดความคลาดเคลื่อนสีในลักษณะของการค้นหาทีละพิกเซลจึงทำให้เกิดความล่าช้าในการประมวลผลด้วยเวลา $O(n^2)$ ส่วนวิธีการที่นำเสนอจะประมวลผลในลักษณะเชิงเส้นและเวลาที่ใช้ขึ้นอยู่กับขนาดของภาพซึ่งก็คือ $O(n)$ จึงมีความรวดเร็วกว่ามาก

5. สรุปผลการวิจัย

งานวิจัยนี้นำเสนอวิธีการปรับปรุงการกรองไกด์อิมเมจเพื่อให้สามารถค้นหาบริเวณที่เกิดความคลาดเคลื่อนสีได้อย่างมีประสิทธิภาพ และรวดเร็วยิ่งขึ้น โดยการไกด์ภาพด้วยคุณลักษณะเฉพาะของบริเวณที่เกิดความคลาดเคลื่อนสีที่อยู่ในรูปของค่าความแตกต่างระหว่างสีแดงกับสีอื่น ๆ ซึ่งผลที่ได้ทำให้บริเวณที่เกิดความคลาดเคลื่อนสีถูกแยกออกจากบริเวณอื่นที่เป็นสีที่เกิดขึ้นตามธรรมชาติได้โดยอัตโนมัติ นอกจากนี้ยังได้นำ NSR มาช่วยพิจารณาหาโอกาสการเกิดความคลาดเคลื่อนสีในภาพด้วย จึงทำให้ได้ค่าความถูกต้องเฉลี่ยสูงสุดถึงร้อยละ 95.29 และมีค่าความผิดพลาดเฉลี่ยต่ำสุดเพียงร้อยละ 1.99 นอกจากนี้ยังมีค่าเฉลี่ยของเวลาที่ใช้ในการประมวลผลอยู่ที่ 0.85 วินาที ซึ่งน้อยกว่าวิธีการอื่น ๆ ที่ผ่านมา จากการวิจัยจึงสรุปได้ว่างานวิจัยนี้มีประสิทธิภาพในการค้นหาบริเวณที่เกิดความคลาดเคลื่อนสี และสามารถแยกแยะสีที่เกิดขึ้นตามธรรมชาติกับสีที่เกิดความคลาดเคลื่อนสีโดยอัตโนมัติได้ และมีจุดเด่นในเรื่องของความเร็วในการประมวลผล แต่งานวิจัยนี้ยังไม่ได้ทดลองกับภาพที่ผ่านกระบวนการประมวลผลภาพด้วยโปรแกรมแต่งภาพมาก่อนเช่น การเพิ่มความเข้มของสีและลดส่วนสว่างภายในภาพ ซึ่งเป็นปัญหาที่น่าสนใจในการนำไปพัฒนาต่อในอนาคต

เอกสารอ้างอิง

[1] W. T. Welford, *Aberrations of Optical Systems*, Taylor and Francis Group: New York, 1986.

[2] Rudolf Kingslake, *Lens Design Fundamentals*, Academic Press: New York, 1978.

[3] ยุทธศักดิ์ จินาพร, สุวัจชัย กมลสันติโรจน์ และ สรร รัตนสัญญา, "การศึกษาวิธีการค้นหาและแก้ไขความผิดเพี้ยนของสีในภาพดิจิตอล," *การประชุมวิชาการระดับประเทศด้านเทคโนโลยีสารสนเทศ ครั้งที่ 8 (NCIT 2016)*, หน้า 528-533, 2559.

[4] S. B. Kang, "Automatic removal of purple fringing from images," *U.S. Patent 7577292*, Washington, DC: U.S. Patent and Trademark Office, 2009.

[5] B. K. Kim and R. H. Park, "Automatic detection and correction of purple fringing using the gradient information and desaturation," *2008 16th European Signal Processing Conference*, Lausanne, pp. 1-5, 2008.

[6] S. W. Chung, B. K. Kim, and W. J. Song, "Removing chromatic aberration by digital image processing," *Optical Engineering*, vol. 46 no. 6. pp. 067002-1–067002-10, 2010.

[7] H. J. Ju, D. K. Lee, and R. H. Park, "Chromatic Aberration Removal in Narrow Color Regions of Digital Images," *16th IEEE International Conference on Consumer Electronics 2012 (ICCE 2012)*, pp. 31-32, 2012.

[8] ณัฐชัย งามเสงี่ยม, สุวัจชัย กมลสันติโรจน์ และ สรร รัตนสัญญา, "การตรวจจับความคลาดเคลื่อนสีอัตโนมัติโดยใช้คุณสมบัติเฉพาะที่," *วารสารสถิติประยุกต์และเทคโนโลยีสารสนเทศ* ปีที่ 1 ฉบับที่ 2 กรกฎาคม-ธันวาคม พ.ศ. 2559 หน้า 8-23.

[9] C. D. Jung, D. W. Jang, H. S. Kim, and R. H. Park, "Color Fringe Correction Using Guided Image Filtering," *18th IEEE International Symposium on Consumer Electronics (ISCE 2014)*. pp. 1-2, 2014.

[10] K. He, J. Sun, and X. Tang, "Guided image filtering," *IEEE Trans*. Pattern Analysis and Machine Intelligence, vol. 35, no. 6, pp. 1397–1409, 2013.

[11] Rafael C. Gonzalez and Richard E. Woods, *Digital Image Processing*, Pearson: New Jersey, 2007.

การนับจำนวนประชากรภายในอาคารโดยใช้วีดิทัศน์เป็นฐาน
A Video-Based People Counting System in indoor spaces

มาโนช ประชา (Manoch Pracha)[1] และศิริชัย เตรียมล้ำเลิศ (Sirichai Triamlumlerd)[2]
[1,2]ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี
manoch.p@en.rmutt.ac.th, sirichai.t@en.rmutt.ac.th

บทคัดย่อ

ปัจจุบันเรื่องเวลาในการให้บริการมีความสำคัญต่อความรู้สึกของผู้คนที่มาใช้บริการ ไม่ว่าจะเป็นจุดให้ข้อมูลลูกค้าหรือจุดจ่ายเงินตามสถานที่ให้บริการลูกค้าหรือห้างสรรพสินค้าต่างๆ ในกรณีที่มีผู้ใช้บริการเป็นจำนวนมากจำเป็นต้องเปิดจุดให้บริการเพิ่ม เพื่อช่วยลดเวลาในการรอ สำหรับหน่วยงานทางการศึกษาเองก็มีจุดให้บริการกับนักศึกษาอยู่หลายบริเวณเช่น จุดบริการแบบครบวงจรของคณะ มหาวิทยาลัย จุดบริการแบบครบวงจรของคณะต่างๆ จุดธุรการของสาขาต่างๆ หรือแม้แต่ห้องเรียน จากปัญหาดังกล่าว ได้เกิดแนวความคิดในการทำระบบวัดความหนาแน่นของผู้ใช้บริการ ระบบตรวจนับจำนวนประชากรภายในอาคาร เพื่อนำมาช่วยแก้ไขปัญหา และยังนำไปใช้วัดความหนาแน่นของจำนวนนักศึกษาในห้องเรียนได้ด้วย ระบบที่จัดทำนี้จะใช้วีดิทัศน์ในการเก็บภาพเพื่อตรวจสอบความเคลื่อนไหวและนับจำนวนประชากรมนุษย์ในบริเวณนั้น ด้วยการนำภาพที่ได้ในแต่ละช่วงเวลามาตรวจจับโครงร่างศีรษะของมนุษย์และทิศทางการเคลื่อนที่ เพื่อใช้ในการนับจำนวนประชากร ระบบทำงานได้เป็นที่น่าพอใจ โดยมีประสิทธิภาพเท่ากับ 91.25% ในบริเวณที่มีจำนวนคนอยู่ไม่มาก สามารถนำไปประยุกต์ใช้งานจริงได้

คำสำคัญ: ระบบนับคน ตรวจจับคน ติดตามคน

Abstract

At competitive period of time, waiting time is one of the crucial problems that directly affects to the individual feel of users. Many services attempt to reduce or even eliminate the wasted time as best as possible. For instance, in case that there are many customers waiting in a queue for: paying supply bills at a counter service; buying things in a department store; getting stamped on their passports at immigration bureau. In those waiting situations, such service providers require to gain their performance of services, for example, by opening a new counter to serve the waiting customers. The problem of waiting for services also occurs in universities, e.g., long waiting queue in one stop service office. Accordingly, this research proposes a design and implementation of a people counting system to automatically detect and count people in a particular indoor area where people usually wait for a service. The system has been designed based on image processing techniques for detecting, tracking, and counting heads of people. The preliminary experiment result shows 91.25% of accuracy where the number of people less than 10 which can be practically used in any indoor area.

Keyword: People Counting System, Human Detection, People Tracking.

1. บทนำ

สถิติของจำนวนประชากรถือเป็นตัวชี้วัดหนึ่งที่สำคัญที่สามารถนำมาใช้ประกอบในการตัดสินใจในการบริหารจัดการกิจกรรมต่างๆ เช่น ความหนาแน่นของประชากรในพื้นที่หนึ่ง ณ. ช่วงเวลาใดๆ ในห้างสรรพสินค้า บอกให้รู้ว่าพื้นที่ดังกล่าวในช่วงเวลานั้นมีโอกาสที่ลูกค้าจะจับจ่ายใช้สอย และรับรู้ข่าวสารโฆษณาจากห้างสรรพสินค้าสูงเป็นพิเศษ จำนวนผู้โดยสารบนเครื่องบินที่สอดคล้องกับจำนวนผู้โดยสารที่ทำ

การลงทะเบียนเที่ยวบินแล้ว บอกให้รู้ว่าไม่มีผู้โดยสารตกหล่นเที่ยวบินก่อนทำการบิน จำนวนผู้เข้าฟังการบรรยายในห้องประชุม บ่งบอกให้เห็นถึงความน่าสนใจหรือความสำคัญของการบรรยายนั้น

ถึงแม้ว่าระบบการนับจำนวนประชากรจะมีหลากหลายรูปแบบ เช่น ระบบนับจำนวนประชากรด้วยกลไกหรือเซ็นเซอร์ที่ติดตั้งบริเวณทาง เข้า-ออก ระบบนับประชากรโดยอาศัยรังสีคลื่นใต้แดง ระบบดังกล่าวมีข้อด้อยและมีข้อจำกัดบางประการ อาทิเช่น มีค่าใช้จ่ายสูง ไม่สะดวกในการติดตั้งและขนย้าย ไม่ยืดหยุ่นกับหลายสภาพพื้นที่

ดังนั้น ระบบนับจำนวนประชากรภายในอาคารโดยใช้วิดิทัศน์เป็นฐานจึงจำเป็นต้องถูกพัฒนาขึ้น เพื่อใช้นับจำนวนประชากรในบริเวณใดๆ ภายในอาคาร โดยเป็นระบบที่ไม่ยึดติดกับจำนวนและตำแหน่งทาง เข้า-ออก ของอาคารหรือห้องภายในอาคาร และยังคงความสะดวกในการติดตั้งระบบ

2. วรรณกรรมที่เกี่ยวข้อง

ระบบตรวจจับและติดตามบุคคลจากกล้องวิดีโอเริ่มมีบทบาทต่อการใช้ชีวิตประจำวันของมนุษย์มากขึ้นระบบดังกล่าวถูกวิจัยและพัฒนาขึ้นเพื่ออำนวยความสะดวกและสนับสนุนกิจกรรมมนุษย์ในแง่มุมต่าง ๆ เช่นการตรวจจับและนับจำนวนคนบนทางเท้า [1], [2] การตรวจจับและนับจำนวนผู้โดยสารที่ยืนรอ ณ จุดรับส่งผู้โดยสารรถประจำทาง [3] การตรวจจับและนับจำนวนคนในบริเวณที่สนใจภายในอาคาร [4], [5] หุ่นยนต์ตรวจจับและติดตามมนุษย์เพื่อช่วยเหลือและพัฒนาคุณภาพชีวิตสำหรับผู้สูงวัย [6] ฯลฯ

ประสิทธิภาพของระบบตรวจจับและติดตามบุคคลได้รับผลกระทบโดยตรงจากสภาพแวดล้อมที่ถูกนำไปประยุกต์ใช้งาน ดังนั้นงานวิจัยเกี่ยวกับระบบดังกล่าวจึงถูกจำแนกได้เป็น 2 ประเภทตามสภาพแวดล้อมคือ ภายในอาคารและภายนอกอาคาร งานภายในอาคารตัวอย่างเช่น [7], [8] งานประเภทนี้มีอุปสรรคที่เกิดขึ้นได้แก่ บุคคลเป้าหมายถูกบดบังด้วยโครงสร้างของตัวอาคาร ขนาดพื้นที่จำกัด งานภายนอกอาคารตัวอย่างเช่น [1], [2], [3] และอุปสรรคที่มักเกิดขึ้นกับงานภายนอกอาคารได้แก่ สภาวะแสงที่เปลี่ยนแปลงในระหว่างวัน ซึ่งส่งผลให้ข้อมูลภาพของบุคคลที่สนใจไม่คงที่ อย่างไรก็ตามอุปสรรคที่เกิดขึ้นร่วมกันของงานตรวจจับและติดตามบุคคลทั้งภายในและภายนอกอาคารได้แก่ บุคคลที่สนใจถูกบดบังโดยวัตถุใด ๆ แสงเงาที่เกิดขึ้นจากวัตถุรวมถึงตัวบุคคลเอง เสื้อผ้าที่บุคคลสวมใส่มีทรงและสีที่หลากหลาย ท่าทางและกิจกรรมของมนุษย์ที่มีความหลากหลายและเปลี่ยนแปลงอยู่เสมอเช่น การกอดกัน การจับมือกัน การลุก การนั่ง การเดิน ฯลฯ

สภาวะอุปสรรคหนึ่งที่เกิดขึ้นเสมอ และเป็นเป้าหมายหนึ่งที่นักวิจัยหลายกลุ่มให้ความสนใจและพยายามจะเอาชนะในงานตรวจจับและติดตามบุคคลทั้งภายในและภายนอกอาคารคือสภาวะบุคคลที่สนใจถูกบดบังโดยวัตถุใด ๆ อุปสรรคนี้เกิดขึ้นเป็นปกติและเป็นข้อจำกัดของระบบที่ใช้กล้องเดี่ยว [4], [6], [9] ในการจับภาพบุคคล จึงทำให้มีนักวิจัยหลายกลุ่ม [5][7][10] พยายามเอาชนะอุปสรรคดังกล่าว โดยอาศัยกล้องหลายตัวจากหลายตำแหน่งในการรับภาพและนำภาพเหล่านั้นมาประมวลผลร่วมกัน ถึงแม้ว่าการประมวลผลภาพจากกล้องหลายตัวจะแก้ปัญหาการบดบังโดยวัตถุใด ๆ ได้ดีในระดับหนึ่ง แต่ต้องแลกกับการประมวลผลที่ซับซ้อน ใช้เวลานาน มีค่าใช้จ่ายที่สูงกว่าการประมวลผลภาพจากกล้องเดี่ยว

AI Amain Hosain และ Anil Saha [11] นำเสนอวิธีการติดตามคนจากภาพวิดีโอโดยใช้หลักการ Color Histogram และ Correlogram ในขั้นตอนการเปลี่ยนแปลงรูปร่างของภาพได้ใช้หลักการของ Morphological Image Closing และใช้การเปลี่ยนเฟรมของภาพ (Similarity Measure) ผลของการวิจัยนี้สามารถใช้ได้ดีกับสภาวะแวดล้อมในอาคารแต่ความสามารถของระบบจะลดลงเมื่อคนที่อยู่ในพื้นที่ทดลองแต่งกายด้วยสีของเสื้อผ้าที่คล้ายกัน

Luca Del Pizzo และ คณะ [12] นำเสนอวิธีการใหม่สำหรับการนับประชากรจากกล้องที่ติดตั้งไว้ด้านบนทั้งในส่วนของในอาคารและนอกอาคาร ผลลัพธ์ที่ได้จากการวิจัยแสดงให้เห็นว่า การนับจำนวนประชากรเมื่อกระจายตัวกันมีความถูกต้อง 98% ทั้ง RGB และ Depth Sensors ทั้งนี้การตรวจนับประชากรที่เกาะกลุ่มกันมีความถูกต้องมากใกล้เคียงกับการ

ตรวจนับแบบกระจายตัวเมื่อใช้ Depth Sensors แต่จะมีความถูกต้องลดลงเมื่อใช้ RGB

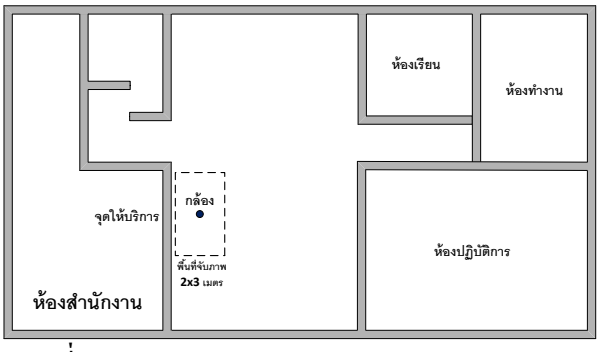
Jau-Woei Perng และ คณะ [13] นำเสนอระบบนับจำนวนผู้โดยสารในรถประจำทางผ่านกล้องซึ่งติดตั้งอยู่ด้านบน โดยเสนอวิธีการแก้ปัญหาจากแสงสว่างเมื่อมีการเปิดประตูรถประจำทางและหาวัตถุที่เคลื่อนไหวโดยใช้เทคนิค Morphological ในการกรองสิ่งรบกวน (Noise) และการเติมสีลงไปในวัตถุ ส่วนการนับจำนวนคนจะนับเมื่อมีวัตถุเคลื่อนผ่านเส้นที่กำหนด โดยระบบที่นำเสนอให้ความแม่นยำที่ 87%

Pongsakon Bamrungthai และ Suteera Puengsawad [14] ได้พัฒนาระบบนับจำนวนคนภายในอาคารแบบอัตโนมัติโดยใช้กล้องที่ติดต่อกับระบบอินเทอร์เน็ตติดตั้งอยู่ในตำแหน่งของประตูทางเข้า โดยการตรวจนับภาพบุคคลจะใช้หลักการ (Subtraction) การเปลี่ยนแปลงรูปภาพ (Morphological) และการตรวจจับกลุ่มของจุดภาพที่มีสีขาว-ดำ (Blob Detection) ระบบยังสามารถตรวจนับจำนวนคนที่ผ่านเข้าและออกประตูได้ด้วย

Bin Li และ คณะ [15] นำเสนอวิธีการนับจำนวนคนโดยจับภาพหัวคนและการติดตามความเคลื่อนไหวของคนที่ผ่านเข้ามาในพื้นที่โดยกล้องที่ติดตั้งอยู่ด้านบน หลักการที่ใช้ในระบบนี้ ใช้การสกัดภาพเบื้องหน้า (Foreground Extraction) ในขั้นตอนแรกและใช้ตัวจำแนก LBP Feature Base Adaboost เพื่อตรวจจับหัวคน ต่อมาใช้อัลกอริทึม Meanshift ในการติดตามคน ระบบสามารถตรวจนับจำนวนคนโดยมีค่าความถูกต้องแม่นยำที่ 96%

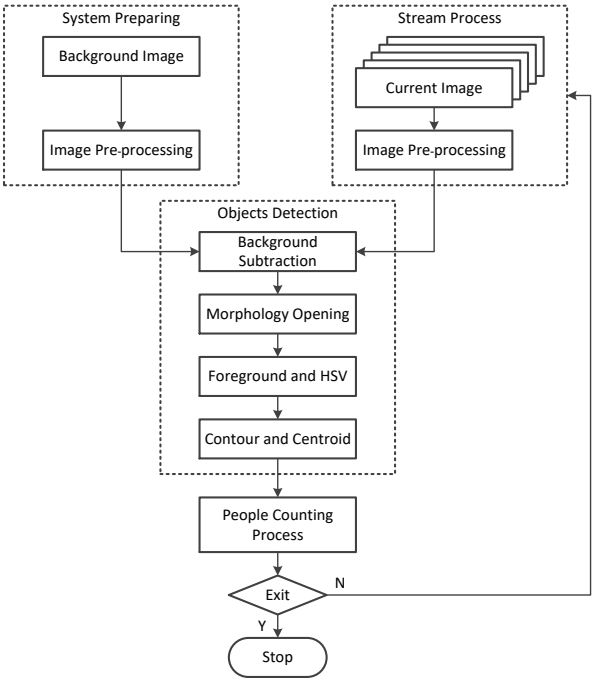
3. วิธีการดำเนินการวิจัย

ระบบการนับจำนวนประชากรภายในอาคารได้จำลองระบบและทดสอบภายในอาคารภาควิชาวิศวกรรมคอมพิวเตอร์ โดยทำการติดตั้งกล้องวีดิทัศน์ไว้ที่ระดับเพดานของอาคารความสูง 4 เมตร ในมุมมองด้านบน (Top View) ตรงจุดให้บริการของหน่วยงาน โดยกรอบสี่เหลี่ยมคือมุมมองภาพสำหรับการตรวจนับ ดังภาพที่ 1



ภาพที่ 1: ผังอาคารและการวางตำแหน่งกล้องสำหรับการตรวจนับ

โดยระบบที่พัฒนาขึ้นมีกระบวนการทำงานอยู่ 4 ส่วนหลัก ดังนี้ การเตรียมข้อมูลภาพพื้นหลัง การจับภาพปัจจุบัน การตรวจสอบการเคลื่อนไหวของวัตถุ และการนับจำนวนประชากร ดังภาพที่ 2



ภาพที่ 2: กระบวนการทำงานของระบบตรวจนับ

หลักการทำงานของระบบเริ่มจากการเก็บภาพพื้นหลัง (Background Image : BG) เป็นเฟรมตั้งต้น และทำการอ่านภาพจากกล้องวีดิทัศน์จากเฟรมปัจจุบัน (Current Image : CR) เพื่อลดปัญหาของเรื่องค่าสีภายใต้ภาพที่มีความแตกต่างกัน ระบบทำการปรับเปลี่ยนสีของภาพ (Image Pre-processing) เป็นสีเทา (Gray Scale) เพื่อให้เหมาะกับการวิเคราะห์หา

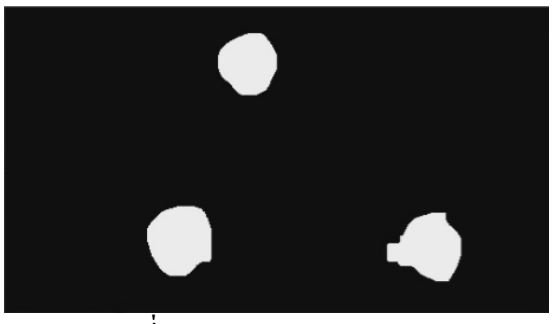
จำนวนประชากร และนำภาพทั้งคู่ลบกัน (Background Subtraction) จะได้ส่วนที่แตกต่างกันของทั้ง 2 ภาพ ดังภาพที่ 3 เก็บไว้ในตาราง Mask

$$Mask(x, y) = |CR(x, y) - BG(x, y)| \quad (1)$$

นำข้อมูล Mask เข้ากระบวนการเปลี่ยนแปลงรูปร่าง ด้วยหลักการ Morphological Processing หลักการนี้จะทำการขยาย (Dilation) และการกร่อนขนาด (Erosion) โดยมี 2 วิธีคือ Opening และ Closing โดยระบบเลือกใช้แบบ Opening คือการทำ Erosion ก่อน จากนั้นจึงทำ Dilation กระบวนการนี้ทำให้ได้ภาพที่เรียบเนียนและการกรองสิ่งรบกวน (Noise) ออกจาก Mask ดังภาพที่ 4



ภาพที่ 3: ผลของ Background Subtraction



ภาพที่ 4: ผลของ Morphological Opening

นำข้อมูล Mask แปลงกลับข้อมูลภาพบริเวณที่เคลื่อนไหวเป็น RGB โดยการนำข้อมูลใน Mask ทำการ Bitwise AND กับ Current Image จะได้ข้อมูล Foreground แบบ RGB กลับมา และเปลี่ยนพื้นหลังเป็นสีขาว

$$Foreground(x, y) = \begin{cases} CR(x, y), Mask(x, y) \,\&\, CR(x, y) > 0 \\ White\ Pixel, Otherwise \end{cases} \quad (2)$$

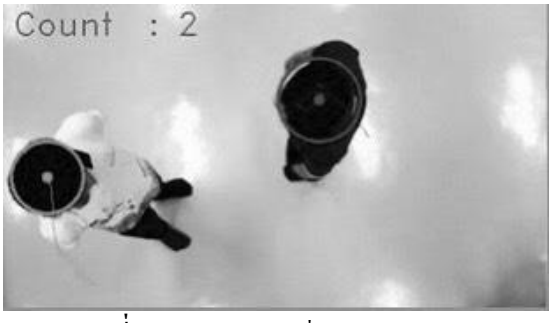
นำข้อมูล Foreground ปรับโทนของสีด้วยหลักการ HSV โดยที่ Hue คือสีของภาพ Saturation คือ ปริมาณความอิ่มตัวของสี และ Value คือค่าปริมาณความสว่างของภาพ จะได้ภาพที่มีสีอยู่ในโทนเดียวกัน แล้วเข้ากระบวนการ Morphological แบบ Opening อีกครั้งเพื่อปรับรูปร่าง ก่อนเข้ากระบวนการนับจำนวนประชากร

การนับจำนวนประชากร (People Counting) เป็นการนำข้อมูลใน Foreground มาตรวจสอบ Contour ของแต่ละจุดมีขนาดและรัศมีที่ใกล้เคียงกับศีรษะมนุษย์ และอยู่ในเฟรมต่อเนื่องมา 5 เฟรมหรือไม่ โดยปกติมนุษย์จะมีการเคลื่อนไหวตลอดเวลาที่ยืน จำนวน 5 เฟรมเพียงพอในการเริ่มในการตรวจจับ ถ้าเข้าเงื่อนไขจะทำการนับเป็น 1 คน พร้อมทั้งวาดจุดกึ่งกลางศีรษะ (Centroid) และลากเส้นระหว่างจุดกึ่งกลางศีรษะ พร้อมทั้งวาดวงกลมล้อมศีรษะ และแสดงตัวเลขที่นับได้

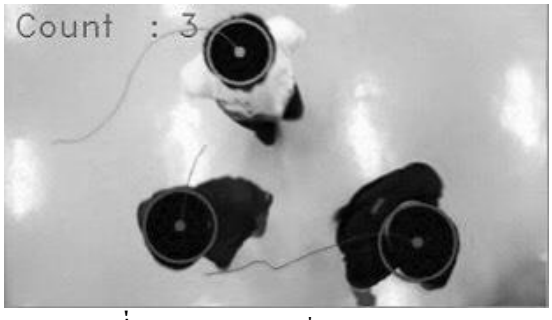
4. ผลการดำเนินงาน

ระบบได้มีการทดสอบการทำงานภายในอาคารภาควิชาวิศวกรรมคอมพิวเตอร์ โดยติดกล้องบริเวณเพดานของชั้น กล้องที่ใช้มีคุณภาพดังนี้ ความละเอียด (Resolution) 2.0M pixel และอัตราเฟรม (Frame rate) 30 fps ทำให้ได้ข้อมูลภาพพื้นหลังที่ครอบคลุมพื้นที่จริง 2x3 เมตร

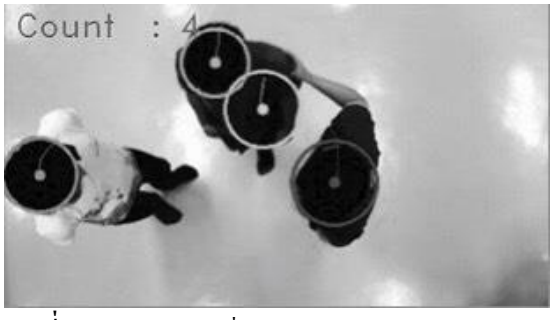
ผลจากการทดลองด้วยหลักการที่ออกแบบ เมื่อมีคนเดินผ่านหน้ากล้อง ระบบจะทำการตรวจนับพร้อมแสดงตัวเลขที่นับได้ และวาดวงกลมตรงส่วนศีรษะของคนที่ยืนอยู่ภายใต้ภาพ ดังตัวอย่างในภาพที่ 5 ที่มีจำนวน 2 คน และ ภาพที่ 6 ที่มีจำนวน 3 คนอยู่ภายใต้ภาพ



ภาพที่ 5: ผลการตรวจนับเมื่อมีคนเดินผ่าน 2 คน



ภาพที่ 6: ผลการตรวจนับเมื่อมีคนเดินผ่าน 3 คน



ภาพที่ 7: ผลการตรวจนับที่เกิดความผิดพลาดจากโครงสร้างคล้าย

ผลที่ได้ทำการทดสอบระบบได้แสดงให้เห็นในตารางที่ 1 โดยในการทดสอบนั้นได้มีการให้นักศึกษาเดินผ่านหน้ากล้องจำนวน 1 คน 2 คน 3 คน และ 4 คน พร้อมกัน เป็นจำนวน 20 รอบ ทั้งนี้บริเวณที่ใช้ในทดสอบอยู่ในพื้นที่ของอาคารและควบคุมสิ่งแวดล้อมของภาพได้

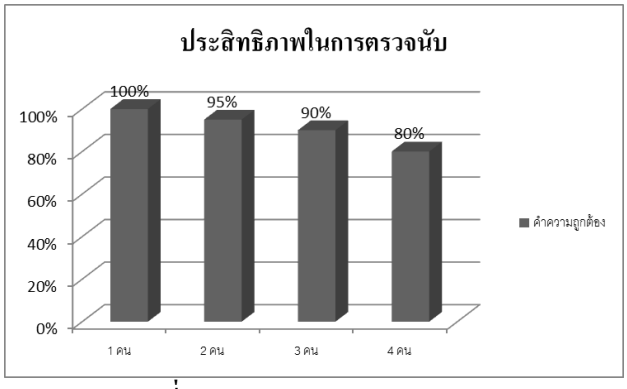
ตารางที่ 1: ผลการทดสอบระบบการนับจำนวนคน

จำนวนคนในการทำลอง (คน)	จำนวนการเดินผ่านระบบ (ครั้ง)	ผลการนับที่ถูกต้อง	ความถูกต้อง (%)
1	20	20	100%
2	20	19	95%
3	20	18	90%
4	20	16	80%
		ประสิทธิภาพของระบบ	91.25%

จากผลการทดสอบระบบจะเห็นได้ว่ามีความผิดพลาดเกิดขึ้นในบางครั้ง เนื่องด้วยหลักการที่ได้ใช้ในการการนับคนนั้นใช้ส่วนหัวในการนับ ซึ่งมีลักษณะกลมและมีสีดำ แต่เมื่อคนที่เดินผ่านมาในระบบใส่ชุดสีดำ หรือการยืนที่ทับซ้อนกัน ก็จะทำให้เกิดการผิดพลาดในการตรวจสอบ และอีกทั้งตัวกล้องที่ได้นำมาใช้มีคุณภาพตามท้องตลาดเมื่อมีการเดินผ่านอย่างรวดเร็วระบบก็ไม่สามารถตรวจนับได้ ซึ่งแสดงผลความผิดพลาดที่เกิดดังภาพที่ 7 และภาพที่ 8 โดยประสิทธิภาพของระบบแสดงได้ดังกราฟ ภาพที่ 9



ภาพที่ 8: ผลการตรวจนับที่เกิดความผิดพลาดจากการสวมเสื้อดำ



ภาพที่ 9: ประสิทธิภาพในการตรวจนับบุคคล

5. สรุป

ผลที่ได้จากการทำสอบระบบจะมีประสิทธิภาพมากเมื่อมีจำนวนคนน้อยที่เดินผ่านหน้ากล้องของระบบ โดยประสิทธิภาพที่ได้จากการทดสอบเมื่อมีคนเดินผ่านระบบ 1 คน อยู่ที่ 100% 2 คน อยู่ที่ 95% 3 คน อยู่ที่ 90% และ 4 คนอยู่ที่ 80% จะเห็นได้ว่าประสิทธิภาพจะต่ำลงเมื่อคนเริ่มมากขึ้น เนื่องจากเมื่อมีจำนวนคนมากขึ้นจะเกิดโอกาสการทับซ้อนของคนมากขึ้นในภาพ ทำให้ประสิทธิภาพในการนับให้ครบทุกคน

ในภาพต่ำลง และสีของเสื้อผ้าที่สวมใส่ก็มีผลต่อประสิทธิภาพ จากการทดสอบระบบได้ค่าประสิทธิภาพรวมในการทดลอง 1-4 คนของระบบอยู่ที่ 91.25%

ระบบที่ได้ออกแบบทำการติดตั้งภายในอาคาร โดยส่วนใหญ่แล้วอาคารสำนักงานมีความสูงประมาณ 4 เมตร เมื่อเก็บภาพคนยืนภายใต้กล้องที่ศีรษะมนุษย์สูงจากพื้นอีก 1.60 เมตร ทำให้เหลือพื้นที่รองรับปริมาณของประชากรภายใต้ภาพได้ไม่มากพอถือเป็นข้อจำกัดของงาน แนวทางในการพัฒนา ถ้าเปลี่ยนกล้องเป็นแบบ 360 องศา จะทำให้มุมมองของภาพที่มองเห็นได้รอบด้าน จะรองรับการตรวจนับประชากรได้มากขึ้นได้

เอกสารอ้างอิง

[1] K. Chen, Z. Zhang, “Pedestrian Counting With Back-Propagated Information and Target Drift Remedy”, *IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS: SYSTEMS*, VOL. 47, NO. 4, pp. 639-647, APRIL 2017.

[2] M. Stubenschrott, T. Matyus, and C. Kogler, “Real-time estimation of pedestrian inflow rates from saturated sensor counting data in a complex metro station”, *Intelligent Transportation Systems (ITSC), 2015 IEEE 18th International Conference on. IEEE*, pp. 1958-1963, 2015.

[3] G. Garcia-Bunster, M. Torres-Torriti, “A density-based approach for effective pedestrian counting at bus stops”, *Proceedings of the 2009 IEEE International Conference on Systems, Man, and Cybernetics*, pp. 3434-3439, 2009.

[4] Q. Ye, “A Robust Method for Counting People in Complex Indoor Spaces”, *2nd International Conference on Education Technology and Computer (ICETC)*, Vol. 2, pp.450-454.

[5] R. Mohedano, C. R. del-Blanco, et al., “Robust 3D People Tracking and Positioning System in A Semi-Overlapped Multi-Camera Environment”, *15th IEEE International Conference on Image Processing (ICIP 2008)*, pp. 2656-2659.

[6] M. Volkhardt, C. Weinrich, and H. M. Gross, “People tracking on a mobile companion robot”, *Systems, Man, and Cybernetics (SMC), 2013 IEEE International Conference on IEEE*, pp. 4354-4359, 2013.

[7] K. Qian, X. Ma, X. dai, et al., “A multi-camera approach to tracking and localization of people with coexisting robots”, *Proceedings of the 7th World Congress on Intelligent Control and Automation*, pp. 5162-5167, 2008.

[8] S. Velipasalar, Y. L. Tian, and A. Hampapur “Automatic counting of interacting people by using a single uncalibrated camera”, *Multimedia and Expo, 2006 IEEE International Conference on IEEE*, pp. 1265-1268, 2006.

[9] B. Hdioud, A. Ezzahout, Y. Hadi, et al., “A real-time people tracking system based on trajectory estimation using single field of camera view”, *Computer Applications Technology (ICCAT), 2013 International Conference on IEEE*, 2013.

[10] T. Yang, Y. Zhang, X. Tong, et al., “A new hybrid synthetic aperture imaging model for tracking and seeing people through occlusion”, *IEEE Transactions on Circuits and Systems for Video Technology*, pp. 1461-1475, 2013.

[11] A. A. Hosain, A. Saha, “An appearance based human motion tracker”, *Informatics, Electronics & Vision (ICIEV), 2013 International Conference on IEEE*, 2013.

[12] L. D. Pizzo, P. Foggia, A. Greco, et al. “A versatile and effective method for counting people on either RGB or depth overhead cameras”, *Multimedia & Expo Workshops (ICMEW), 2015 IEEE International Conference on IEEE*, 2015.

[13] J. W. Perng, T. Y. Wang, Y. W. Hsu, et al. “The design and implementation of a vision-based people counting system in buses”, *System Science and Engineering (ICSSE), 2016 International Conference on IEEE*, 2016.

[14] P. Bamrungthai, S. Puengsawad, “Robust people counting using a region-based approach for a monocular vision system”, *Science and Technology (TICST), 2015 International Conference on IEEE*, pp. 308-312, 2015.

[15] Li, Bin, et al. "A people counting method based on head detection and tracking." *Smart Computing (SMARTCOMP), 2014 International Conference on. IEEE*, 2014.

การเปรียบเทียบประสิทธิภาพการค้นคืนและการจำแนกหมวดหมู่ภาพด้วยคุณลักษณะรูปร่าง

A Comparison of Image Retrieval and Classification Based on Shape Feature

สุกัญชลิกา บุญมาธรรม (Sukanchalika Boonmatham)[1] สุปราณี ห้อมา (Supranee Horma)[2] เสาวนีย์ ปรัชญาเกรียงไกร (Soawanee Prachayagringkai)[3] พรศักดิ์ ปรีเลขา (Pornsak preelakha)[4] เจษฎา บุญสิริ (Jedsada Boonsiri)[5] วรรณนภา โพธิ์ผลิ (Wannapa Phopli)[6] และมาลีรัตน์ โสดานิล (Maleerat Sodanil)[7]

ภาควิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

[1]Sukanchalika.boo@mail.pbru.ac.th, [2]Supranee.h@email.kmutnb.ac.th, [3]Soawanee.p@email.kmutnb.ac.th, [4]Pornsakpre@pim.ac.th,[5]s5907011956011@email.kmutnb.ac.th, [6]Wannapa.p@email.kmutnb.ac.th, [7]Maleerat.s@it.kmutnb.ac.th

บทคัดย่อ

งานวิจัยนี้นำเสนอการค้นคืนภาพด้วยการหาความคล้ายคลึงด้วยเทคนิคการสกัดรูปร่างโดยการสกัดขอบภาพ ซึ่งใช้ชุดข้อมูลภาพกลุ่มสัตว์ แบ่งออกเป็น 3 ประเภท ได้แก่ ภาพโค ม้า และนกกระจอก จำนวนทั้งสิ้น 57 ภาพ จากเว็บไซต์ที่ให้บริการคลังภาพ และใช้การค้นคืนด้วยวิธีการวัดระยะห่างแบบยูคลิเดียนและวัดความคล้ายคลึงแบบโคซายน์ ผลการวัดประสิทธิภาพการค้นคืนภาพด้านความแม่นยำพบว่าวิธีแบบโคซายน์ให้ค่าความแม่นยำสูงกว่าวิธีแบบยูคลิเดียน และการจำแนกหมวดหมู่ชุดข้อมูลภาพกลุ่มสัตว์ ด้วยเทคนิคซัพพอร์ตเวกเตอร์แมชชีน ผลการทดลองพบว่า ให้ค่าความถูกต้องในการจำแนกหมวดหมู่สูงกว่าเทคนิค เคเนียเรสเนเบอร์ โดยมีค่าความถูกต้องร้อยละ 80.70

คำสำคัญ: การค้นคืนเชิงเนื้อหาของภาพ การจำแนกข้อมูลภาพ คุณลักษณะรูปร่าง การวัดค่าความคล้ายคลึง

Abstract

This research presents an image retrieval with similarity by utilizing a boundary extraction. We use 57 images of animal group from the gallery-service website, which are categorized into 3 types: a cow, a horse and a sparrow, as a data set. The similarity between images is measured by "Cosine Similarity" and "Euclidean Distance" respectively. A result of the precision assessment reveals that "Cosine Similarity" give the higher precision value than "Euclidean Distance". And A result of the accuracy assessment in animal group image classification between "Support Vector Machine" and "k-nearest neighbor" reveals that "Support Vector Machine" provide higher accuracy value than "k-nearest neighbor" with an accurate rate at 80.70%.

Key words: CBIR, Image Classification, Shapes Feature, Similarity Measure

1. บทนำ

ในสังคมยุคสารสนเทศและสังคมออนไลน์ ข้อมูลมัลติมีเดีย (Multimedia Information) มีแนวโน้มเพิ่มขึ้นอย่างรวดเร็ว เช่น รูปภาพ วิดีโอ และเสียง โดยเฉพาะอย่างยิ่งข้อมูลรูปภาพที่ไม่เพียงแต่มีจำนวนเพิ่มขึ้น แต่ยังมีความหลากหลายทั้งในรูปแบบของภาพ ชนิดของภาพ ขนาดของภาพ และอื่น ๆ ซึ่งภาพเหล่านั้นสามารถแบ่งออกได้เป็นภาพมนุษย์ เช่น ภาพใบหน้าคน และภาพที่ไม่ใช่มนุษย์ เช่น ภาพสัตว์ ภาพสถานที่ และภาพทิวทัศน์

การค้นคืนภาพเป็นสิ่งที่ดึงดูดความสนใจของนักวิจัย [1] หลายวิธีได้ถูกนำเสนอขึ้นมาเพื่อใช้ในการค้นคืนภาพ ซึ่งวิธีการเหล่านี้สามารถแบ่งออกเป็น 2 รูปแบบ คือ การค้นคืนภาพโดยใช้พื้นฐานทางตัวอักษร(Text-based Image Retrieval) เป็นวิธีการค้นคืนภาพด้วยคำสำคัญโดยในภาพแต่ละภาพจะมี

การให้คำสำคัญเพื่ออธิบายภาพและนำมาเปรียบเทียบหาค่าคำสำคัญที่ตรงกับข้อความที่สอบถาม และ การค้นคืนภาพโดยใช้พื้นฐานทางเนื้อหาของภาพ (Content-based Information Retrieval-CBIR) เป็นวิธีการค้นคืนภาพสอบถาม (Query Image) ด้วยการค้นหาภาพในฐานข้อมูลที่มีลักษณะเหมือนหรือคล้ายกับภาพสอบถาม และพิจารณาความคล้ายคลึงกันของภาพจากค่าสี (Color) พื้นผิว (Texture) หรือรูปร่าง (Shape) ซึ่งวิธีการต่าง ๆ เหล่านี้ถูกพัฒนาเพื่อให้การค้นคืนภาพมีความถูกต้องและมีประสิทธิภาพสูงสุด รูปร่างเป็นคุณลักษณะที่สำคัญของรูปภาพ [2] เนื่องจากรูปร่างมีข้อมูลภาพที่น่าสนใจที่สุดสำหรับการรับรู้ของมนุษย์ [3] และมีลักษณะที่บ่งชี้ถึงจุดเด่นของวัตถุใด ๆ ในภาพจึงสามารถใช้สำหรับการตรวจจับวัตถุได้ [4] รวมถึงมีประโยชน์สำหรับระบบค้นคืนสารสนเทศ เช่น ระบบค้นคืนรูปภาพที่มีลักษณะเฉพาะ

งานวิจัยนี้ให้ความสนใจเกี่ยวกับ การค้นคืนภาพโดยใช้พื้นฐานทางเนื้อหาของภาพ จากความคล้ายคลึงกันของภาพรูปร่าง จึงนำเสนอเทคนิคการสกัดคุณลักษณะรูปร่าง (Shape Feature Extraction) ซึ่งเป็นคุณลักษณะสำคัญที่ช่วยให้สามารถจดจำวัตถุ [5] สำหรับการค้นคืนภาพ โดยการสกัดคุณลักษณะรูปร่างนี้จะสามารถอธิบายภาพที่เหมือนหรือคล้ายคลึงกัน [6] เพื่อการค้นคืนที่มีความแม่นยำได้ ด้วยเทคนิค Boundary-based Shape Representation ซึ่งเป็นเทคนิคในการอธิบายภาพด้วยเส้นแสดงรูปร่าง จากนั้นนำภาพมาทดสอบการค้นคืนด้วยวิธีการวัดความคล้าย (Similarity) จาก 2 วิธีการ คือ Euclidean Distance และ Cosine Similarity เพื่อเปรียบเทียบค่าความแม่นยำในการค้นคืนภาพที่ได้ และการจำแนกหมวดหมู่ของภาพ (Classification) จาก 2 เทคนิค คือ Support Vector Machine (SVM) และ k-Nearest Neighbor (k-NN) เพื่อเปรียบเทียบประสิทธิภาพการจำแนกหมวดหมู่ของภาพ

2. วรรณกรรมที่เกี่ยวข้อง

2.1 คุณลักษณะรูปร่าง (Shapes Feature)

รูปร่าง (Shape) เป็นคุณลักษณะสำคัญที่จะช่วยให้สามารถจดจำวัตถุได้ ตัวบ่งชี้รูปร่างถูกนำมาใช้ประโยชน์มากที่สุดในการรับรู้และการจำแนกประเภทของวัตถุ สำหรับการจำแนกด้วยตนเอง และการจำแนกด้วย Computer Applications [5] วิธีการ Boundary-based Shape Representation เป็นเทคนิคในการอธิบายภาพโดยใช้เส้นแสดงรูปร่างของภาพที่สนใจ เช่น เส้นรอบวงของวงกลมและวงรี เป็นต้น [7], [8], [9] โดยวิธีการสกัดคุณลักษณะรูปร่าง มี 2 รูปแบบ คือ Boundary-based เป็นวิธีการสกัดข้อมูลรูปร่างบนเส้นขอบเขตของภาพ และ Region-based เป็นวิธีการสกัดข้อมูลรูปร่างจากบริเวณของภาพ โดยวิธีการที่กล่าวมาเป็นวิธีการสกัดข้อมูลรูปร่างจากพิกเซลทั้งหมดที่มีอยู่ภายในภาพที่สนใจ มีหลายวิธีที่สามารถทำงานได้ด้วยการสกัดคุณลักษณะแบบ Boundary-based เช่น วิธีตัวอธิบายฟูเรียร์ (Fourier Descriptor) วิธีตัวอธิบายเวฟเล็ต (Wavelet Descriptor) วิธีอธิบายพื้นที่ความโค้ง (Curvature Scale Space Descriptors) เป็นต้น [10]

ในบทความนี้เลือกใช้เทคนิค Boundary-based ในการทำ Shape Representation [11], [12] เทคนิคนี้อาศัยหลักการจากแนวคิดที่ว่ามนุษย์ที่ใช้เส้นแสดงรูปร่างเป็นหลักในการแยกแยะรูปร่าง โดยเทคนิค Boundary-based นั้น แบ่งออกเป็นสองกลุ่ม ได้แก่ Continuous Approach (Global Approach) และ Discrete Approach (Structural Approach)

Continuous Approach (Global Approach) [13], [14] เป็นกลุ่มเทคนิคที่อธิบายรูปร่างในภาพรวมทั้งหมด ซึ่งการอธิบายนั้นสามารถทำได้หลายวิธี เช่น การใช้คุณลักษณะของรูปร่าง การกำหนดจุดบนเส้นแสดงรูปร่าง การใช้ Shape Signature [15], [16], [17] เป็นต้น กลุ่มเทคนิคนี้มีข้อดีคือ มีกระบวนการจับคู่ที่ตรงไปตรงมา ง่ายต่อการนำไปใช้แต่ก็มีจุดอ่อนที่สำคัญคือ มีโอกาสผิดพลาดเนื่องมาจากลักษณะภาพที่มีความคลุมเครือต่างๆ สูง เนื่องจากใช้ข้อมูลส่วนน้อยของรูปร่างเท่านั้น และไม่สามารถใช้กับรูปร่างที่ไม่มีเส้นแสดงรูปร่างที่ชัดเจน ด้วยข้อจำกัดดังกล่าว เทคนิคในกลุ่ม Continuous Approach จึงมักถูกใช้ควบคู่กับเทคนิคประเภทอื่นเป็นหลัก ตัวอย่างเทคนิคในกลุ่มนี้ เช่น Fourier Descriptors เป็นต้น

Discrete approach (Structural Approach) แตกต่างจากกลุ่มเทคนิค Continuous Approach ตรงที่กลุ่มเทคนิคนี้จะทำการแบ่งรูปร่างออกเป็นส่วนย่อย ๆ เช่น สี่เหลี่ยม วงกลมหรือเส้นโค้ง โดยเรียกส่วนย่อยเหล่านี้ว่า Primitive แล้วใช้การจัดการโครงสร้างความสัมพันธ์ระหว่าง Primitive ในการอธิบายภาพ แต่กลุ่มเทคนิคนี้มีข้อเสียคือ ไม่มีความเสถียร

ในการอธิบายภาพ เนื่องจากไม่มีข้อกำหนดตายตัวว่าแต่ละรูปร่างต้องมี Primitive เท่าใด นอกจากนี้เทคนิคในกลุ่มนี้มีกระบวนการที่ยุ่งยากซับซ้อนเกินกว่าจะนำไปใช้ได้จริง รวมถึงไม่สามารถอธิบายรูปร่างในภาพรวม ส่งผลให้ผลลัพธ์การทำการจับคู่เกิดข้อผิดพลาดได้ง่าย เช่น จับคู่ยอดวิหารให้คล้ายกับกล้วย เป็นต้น

2.2 การวัดค่าความคล้ายคลึง (Similarity Measure)

ระบบการค้นคืนภาพมีวิธีการวัดค่าความคล้ายคลึงอยู่หลายวิธี ซึ่งในงานนี้ใช้วิธีการวัดค่าความคล้ายคลึงในการค้นคืนภาพ ดังนี้

2.2.1 Euclidean Distance [18] หรือเรียกอีกอย่างหนึ่งว่า L2 distance เป็นวิธีการที่ถูกนำมาใช้เพื่อวัดระยะห่างระหว่างจุด 2 จุด (a, b) เพื่อค้นหาความคล้ายคลึงกันของข้อมูล โดยใช้วิธีการคำนวณจากสมการ (1)

$$EU(a,b) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2} \qquad (1)$$

2.2.2 Cosine Similarity [19] เป็นวิธีการหาความคล้ายคลึงกันจากค่าความต่างของมุมของข้อมูล 2 อันที่เกิดขึ้นบนพื้นที่เวกเตอร์ (Vector Space) ซึ่งความคล้ายคลึงกันแบบโคซายน์นี้จะมีค่าอยู่ระหว่าง 0-1 เท่านั้น โดยใช้วิธีการคำนวณจากสมการ (2)

$$sim(d,q) = \frac{\sum_{i=1}^{n} d_i \times q}{(\sum_{i=1}^{n} d_i^2 \times \sum_{i=1}^{n} q^2)^{\frac{1}{2}}} \qquad (2)$$

โดยที่ d คือ เอกสารต่างๆ ในคอลเล็คชั่น (Collection) และ q คือ Query ที่ผู้ใช้เพิ่มเข้าไป

2.3 การจำแนกข้อมูลภาพ (Image Classification)

การจำแนกข้อมูลภาพเป็นการสร้างโมเดลหรือตัวจำแนกข้อมูล (Classifier) เพื่อทำนายหมวดหมู่ของข้อมูล (Class) งานวิจัยนี้เลือกใช้การจำแนกข้อมูล 2 รูปแบบ ดังนี้

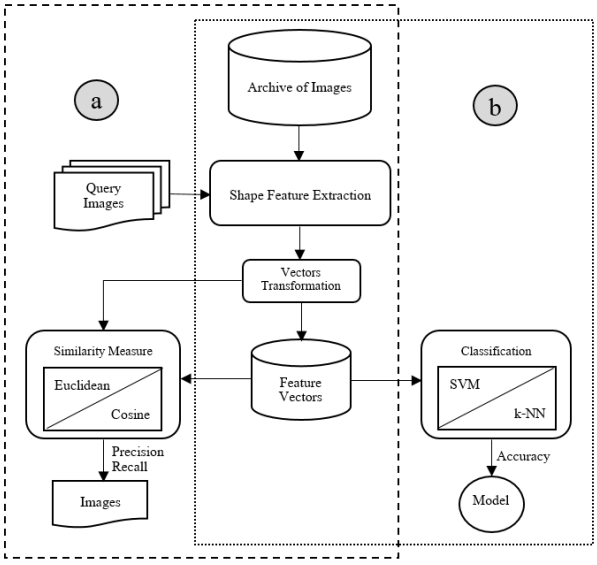
2.3.1 ซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine) คือ วิธีการเรียนรู้แบบมีผู้สอน ใช้จำแนกรูปแบบหรือกลุ่มของข้อมูล โดยอาศัยระนาบมาใช้ในการแบ่งเขตของข้อมูล [20] เมื่อมีชุดข้อมูลสำหรับฝึกมาให้และแต่ละข้อมูลจะถูกแบ่งไปในแต่ละประเภท ซัพพอร์ตเวกเตอร์แมชชีนจะสร้างแบบจำลองที่สามารถพยากรณ์ได้ว่าข้อมูลนี้จะจัดอยู่ในประเภทใด

2.3.2 เคเนียเรสเนเบอร์ (k-Nearest Neighbor) เป็นหนึ่งในวิธีการที่นิยมใช้กับปัญหาในลักษณะการจำแนกหมวดหมู่ [21], [22] โดยใช้วิธีการคำนวณเพื่อวัดระยะทางระหว่างข้อมูลเพื่อพิจารณาว่าข้อมูลนั้นเข้าใกล้กับจุดใดมากที่สุด โดยจะคำนวณเพื่อวัดระยะห่างระหว่างข้อมูลตามกลุ่มที่กำหนด หรือกำหนดค่า k ถ้าข้อมูลเข้าใกล้จุด k ใดนั้น หมายความว่าข้อมูลนั้นจะอยู่คลาสหรือกลุ่มนั้น ซึ่งการกำหนดค่า k เป็นจำนวนคี่ เพื่อให้สามารถตัดสินใจได้ว่าจะอยู่กลุ่มใดเมื่อการคำนวณระยะทางมีค่าเท่ากัน โดยเทคนิคในการวัดระยะที่นิยมนำมาใช้ด้วยวิธีการหาระยะทางแบบ Euclidean Distance โดยการวัดระยะทางว่าอยู่ใกล้กับจุดใดมากที่สุดเพื่อทำนายคลาสของข้อมูล

3. วิธีการดำเนินการวิจัย

งานวิจัยนี้นำเสนอการค้นคืนภาพสัตว์ และการจำแนกหมวดหมู่ของข้อมูลภาพ โดยมีขั้นตอนการดำเนินการ ดังนี้

จากภาพที่ 1 แสดงขั้นตอนการดำเนินการวิจัยการค้นคืนภาพสัตว์ และการจำแนกหมวดหมู่ของข้อมูลภาพ ซึ่งมีขั้นตอนดังนี้

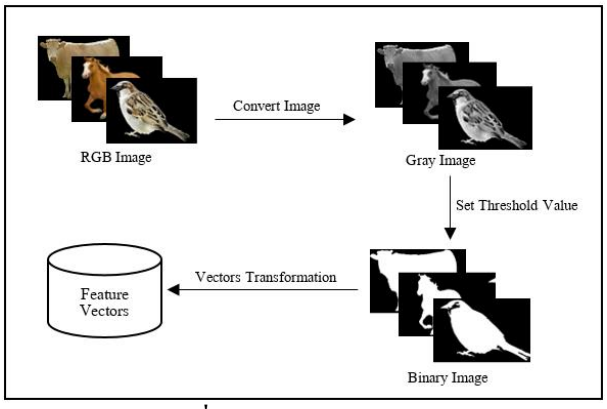


ภาพที่ 1: ขั้นตอนการดำเนินการวิจัย

3.1 การสกัดข้อมูลภาพ

ชุดข้อมูลที่นำมาใช้ในการสกัดข้อมูลภาพ ประกอบไปด้วยภาพสัตว์จำนวนทั้งสิ้น 57 ภาพ แบ่งเป็นภาพโค (Cows) จำนวน 19 ภาพ ภาพม้า (Houses) จำนวน 20 ภาพ และภาพนกกระจอก (Sparrows) จำนวน 18 ภาพ จากเว็บไซต์ http://pngimg.com [23] เป็นชนิดภาพสี (RGB) แบบพื้นหลังโปร่งใส (Background Transparent) ขนาด 420x300 พิกเซล

นำภาพเข้าสู่กระบวนการสกัดคุณลักษณะรูปร่าง (Shape Feature Extraction) โดยการแปลงภาพสีให้เป็นภาพสีเทา (Gray Image) และแปลงภาพสีเทาให้เป็นภาพขาวดำ (Binary Image) ที่ได้กำหนดค่า Threshold ไว้ตามค่าสีของภาพสีเทา



ภาพที่ 2: ขั้นตอนการสกัดข้อมูลภาพ

จากนั้นจึงนำภาพขาวดำที่ได้เข้าสู่กระบวนการสกัดขอบภาพ (Boundary Feature) จนกระทั่งได้ค่าเมทริกซ์ขนาด 420x300 จากนั้นนำไปแปลงเป็นค่าตามคุณลักษณะแบบเวกเตอร์ (Vector Transformation) จะได้เวกเตอร์ที่มีขนาด 126,000 แอททริบิวต์ต่อภาพ นำไปเก็บไว้ในคลังเวกเตอร์คุณลักษณะ (Feature Vectors) เพื่อเข้าสู่กระบวนการค้นคืนและการจำแนกหมวดหมู่ภาพต่อไป

3.2 การค้นคืนภาพ (Image Retrieval)

การค้นคืนภาพเป็นการหาความคล้ายคลึง (Similarity) ของภาพสอบถามกับภาพในคลังภาพ ด้วยวิธีการ Euclidean Distance เปรียบเทียบกับวิธีการ Cosine Similarity เพื่อหาความคล้ายคลึงกันมากที่สุด ระหว่างภาพสอบถามกับภาพในคลัง ภาพ จากนั้นวัดค่าความแม่นยำ (Precision) และค่าความระลึก (Recall) ของการค้นคืนภาพสัตว์ กระบวนการค้นคืนภาพแสดงดังภาพที่ 2 (a)

3.3 การจำแนกหมวดหมู่ภาพ (Image Classification)

กระบวนในการจำแนกหมวดหมู่ของภาพ จะใช้ข้อมูลที่อยู่คลังเวกเตอร์คุณลักษณะ (Feature Vectors) ที่ได้จากการสกัดข้อมูลภาพในกระบวนการขั้นตอนที่ 3.1 นำมาใช้ในการจำแนกหมวดหมู่โดยแบ่งจัดแบ่งเป็น 3 คลาส ตามจำนวนภาพ โค ม้า และนกกระจอกเทศ ซึ่งงานวิจัยนี้ใช้เทคนิค SVM เปรียบเทียบกับเทคนิค k-NN เพื่อหาประสิทธิภาพการจำแนกหมวดหมู่ด้วยวิธีการทดสอบแบบ 10-fold Cross-Validation โดยแบ่งชุดข้อมูลออกเป็น 10 ชุดย่อย (fold) ชุดข้อมูล 9 ชุด เป็นชุดสำหรับฝึก (Training set) ส่วนที่เหลือ 1 ชุด เป็นชุดข้อมูลสำหรับทดสอบ (Test set) จากนั้นวัดค่าความถูกต้อง (Accuracy) ของการจำแนกหมวดหมู่ภาพสัตว์ กระบวนการจำแนกหมวดหมู่ภาพแสดงดังภาพที่ 2 (b)

4. ผลการทดลอง

4.1 ผลการค้นคืนภาพด้วยการวัดค่าความคล้ายคลึง

จากการทดลองค้นคืนภาพโดยการหาความคล้ายคลึงด้วยวิธีการ Euclidean Distance เปรียบเทียบกับวิธีการ Cosine Similarity จากภาพสอบถามคนละชุดกับคลังภาพ พบว่า การค้นคืนภาพด้วยการหาความคล้ายคลึงวิธีการ Cosine Similarity ให้ค่าความแม่นยำในการค้นคืนภาพในกลุ่มสัตว์ชนิดเดียวกัน มากกว่าวิธีการ Euclidean Distance ในการค้นคืนภาพโค ม้า และนกกระจอก โดยวิธีการ Cosine Similarity ให้ค่าความแม่นยำและค่าความระลึกในการค้นคืนภาพ เท่ากัน คือ 0.737, 0.800 และ 1.000 ตามลำดับ และวิธีการ Euclidean Distance ให้ค่าความแม่นยำและค่าความระลึกในการค้นคืนภาพเท่ากัน คือ 0.474, 0.300 และ 0.889 ตามลำดับ แสดงดังตารางที่ 1

ตารางที่ 1: ประสิทธิภาพการค้นคืนภาพด้วยการวัดค่าความคล้ายคลึง

Similarity Measure	Class	Precision	Recall
Cosine Similarity	Cows	0.737	0.737
	Houses	0.800	0.800
	Sparrows	1.000	1.000
Euclidean Distance	Cows	0.474	0.474
	Houses	0.300	0.300
	Sparrows	0.889	0.889

ในการค้นคืนจะทำการค้นคืนภาพที่มีค่าความคล้ายคลึงกันมากที่สุดอันดับแรก โดยผลการค้นคืนภาพสัตว์ที่ทำการวัดค่า

ความคล้ายคลึงที่ได้จากการคำนวณ แสดงดังตารางที่ 2

ตารางที่ 2: ตัวอย่างผลการค้นคืนภาพสัตว์

Query Image	First Similarity Retrieved Image
	
	
	

4.2 ผลการหาประสิทธิภาพการจำแนกหมวดหมู่

จากการทดลองจำแนกหมวดหมู่ภาพ ด้วยเทคนิค SVM เปรียบเทียบกับเทคนิค k-NN เพื่อหาประสิทธิภาพการจำแนกหมวดหมู่ภาพ พบว่า การจำแนกหมวดหมู่ภาพด้วยเทคนิค SVM ให้ค่าความถูกต้องในการจำแนกเท่ากับ 80.70 ซึ่งได้ค่าความถูกต้องสูงกว่าเทคนิค k-NN ซึ่งได้ค่าความถูกต้องในการจำแนกเท่ากับ 73.68 แสดงดังตารางที่ 3

ตารางที่ 3: ผลการวัดประสิทธิภาพการจำแนกหมวดหมู่

Classification Technique	Class	Accuracy (%)
SVM	Cows	80.70
	Houses	
	Sparrows	
k-NN	Cows	73.68
	Houses	
	Sparrows	

5. สรุป

งานวิจัยนี้นำเสนอการค้นคืนภาพด้วยการหาความคล้ายคลึงด้วยเทคนิคการสกัดภาพ Boundary-based ในการทำ Shape Feature Extraction โดยใช้ชุดข้อมูลภาพกลุ่มสัตว์ จำนวน 57 ภาพ ผลการวัดประสิทธิภาพความแม่นยำและค่าระลึกในการค้นคืนภาพด้วยวิธี Cosine Similarity ให้ค่าความแม่นยำในการค้นคืนภาพในกลุ่มสัตว์ชนิดเดียวกันสูงกว่าวิธี Euclidean Distance และการจำแนกหมวดหมู่ชุดข้อมูลภาพกลุ่มสัตว์ พบว่า ผลการวัดประสิทธิภาพด้วยเทคนิค SVM ให้ค่าความถูกต้องในการจำแนกหมวดหมู่ภาพสูงกว่าเทคนิค k-NN จากการทดลอง พบว่า การสกัดคุณลักษณะด้วยเทคนิคการสกัดภาพ Boundary-based ทำให้เกิดแอททริบิวต์จำนวนมาก ซึ่งส่งผลต่อกระบวนการประมวลผล คือทำให้ใช้ระยะเวลาในการประมวลผลเป็นเวลานาน ดังนั้น ในกระบวนการสกัดขอบรูปร่างของภาพครั้งต่อไปจึงควรพิจารณาหาวิธีการลดจำนวนแอททริบิวต์ หรืออาจเลือกใช้คุณลักษณะอื่นที่น่าสนใจ ที่จะส่งผลให้การทำงานดังกล่าวมีประสิทธิภาพเพิ่มขึ้น และอาจส่งผลให้มีความแม่นยำในการค้นคืนข้อมูลภาพเพิ่มมากขึ้น

เอกสารอ้างอิง

[1] N. S. Mansoori, M. Nejati, P. Razzaghi and S. Samavi, "Bag of visual words approach for image retrieval using color information," *2013 21st Iranian Conference on Electrical Engineering (ICEE)*, Mashhad, 2013, pp. 1-6.

[2] ไกรศักดิ์ เกษร. “การค้นคืนสารสนเทศมัลติมีเดีย.” [Online]. Available: http://student.nu.ac.th/kraisakk/ir/IR/6.Chapter04.pdf [Accessed: 20-Mar-2017].

[3] P. Desai, J. Pujari, S. Parvatikar, “Image Retrieval Using Shape Feature: A Study”. In: Das V.V., Thankachan N. (eds) Computational Intelligence and Information Technology. *Communications in Computer and Information Science*, vol 250. Springer, Berlin, Heidelberg.

[4] C. S. Gode and A. S. Khobragade, “Object detection using color clue and shape feature,” *in Wireless Communications, Signal Processing and Networking (WiSPNET), International Conference on*, 2016, pp. 464–468.

[5] N. V. Revollo, C. A. Delrieux and R. González-José, "Set of bilateral and radial symmetry shape descriptor based on contour information," in *IET Computer Vision*, vol. 11, no. 3, pp. 226-236, 4 2017.

[6] J. C. Figueiredo, F. G. M. Neto and I. C. de Paula, "Contour-based feature extraction for image classification and retrieval," *2016 35th International Conference of the Chilean Computer Science Society (SCCC)*, Valparaíso, 2016, pp. 1-7.

[7] Z. Dengsheng and L. Guojun, “Study and evaluation of

different Fourier methods for image retrieval," *Image and Vision Computing*, vol. 23, no. 1, pp. 33–49, Jan. 2005.

[8] Z. Dengsheng and L. Guojun, "Shape-based image retrieval using generic Fourier descriptor," *Signal Processing: Image Communication*, vol. 17, no. 10, pp. 825–848, Nov. 2002.

[9] Z. Dengsheng and L. Guojun, "Review of shape representation and description techniques," *Pattern Recognition*, vol. 37, no. 1, pp. 1–19, Jan. 2004.

[10] G. Zhang, Z. M. Ma, Q. Tong, Y. He and T. Zhao, "Shape Feature Extraction Using Fourier Descriptors with Brightness in Content-Based Medical Image Retrieval," *2008 International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, Harbin, 2008, pp. 71-74.

[11] K. Schindler and D. Suter, "Object detection by global contour shape," *Pattern Recognition*, vol. 41, no. 14, pp. 3736–3748, Dec. 2008.

[12] Z. Dengsheng and L. Guojun, "A comparative study of curvature scale space and Fourier descriptors for shape-based image retrieval," *Journal of Visual Communication and Image Representation*, vol. 14, no.1, pp. 39–57, Mar. 2003.

[13] R. B. Yadav, N. K. Nishchal, A. K. Gupta, and V. K. Rastogi, "Retrieval and classification of shape-based objects using Fourier, generic Fourier, and wavelet-Fourier descriptors technique: A comparative study," *Optics and Lasers in Engineering*, vol. 45, no. 6, pp. 695–708, Jun. 2007.

[14] E. Sokic and S. Konjicija, "Phase preserving Fourier descriptor for shape-based image retrieval," Signal Processing: Image Communication, vol. 40, pp. 82–96, Jan. 2016.

[15] S. Xin, P. Lei, and W. Xiao-Jun, "Multi-scale contour flexibility shape signature for Fourier descriptor," *Journal of Visual Communication and Image Representation*, vol. 26, pp. 161–167, Jan. 2015.

[16] A. El-ghazal, O. Basir, and S. Belkasim, "Farthest point distance: A new shape signature for Fourier descriptors," *Signal Processing: Image Communication*, vol. 24, no. 7, pp. 572–586, Aug. 2009.

[17] X. Bai, X. Yang, and L. J. Latecki, "Detection and recognition of contour parts based on shape similarity," *Pattern Recognition*, vol. 41, no. 7, pp. 2189–2199, Jul. 2008.

[18] M.D. Malkauthekar, "*ANALYSIS OF EUCLIDEAN DISTANCE AND MANHATTAN DISTANCE MEASURE IN FACE RECOGNITION* ," in *Third International Conference on Computational Intelligence and Information Technology (CIIT 2013)*, 2013, pp. 503-507.

[19] ไกรศักดิ์ เกษร. *การค้นหาข้อมูลเชิงความหมาย: แนวคิดใหม่ของโปรแกรมการค้นหา (Search Engine) และแนวทางการพัฒนาในอนาคต*. วารสารวไลยอลงกรณ์ ปริทัศน์ ปีที่ 1, ฉบับที่ 2, กรกฎาคม-ธันวาคม 2554.

[20] "Support Vector Machines (SVM)." [Online]. Available: http://www.statsoft.com/Textbook/Support-Vector-Machines. [Accessed: 20-Mar-2017].

[21] S. Begum, D. Chakraborty, and R. Sarkar, "Data Classification Using Feature Selection and kNN Machine Learning Approach," in *2015 International Conference on Computational Intelligence and Communication Networks (CICN)*, 2015, pp. 811–814.

[22] İ. G. Tekdemir and B. Alboyacı, "Improvement of power swing detection performance of a distance relay by using k-NN algorithm," in *2015 9th International Conference on Electrical and Electronics Engineering (ELECO)*, 2015, pp. 541–545.

[23] "PNG images with transparent background." [Online]. Available: http://pngimg.com/. [Accessed: 29-Mar-2017].

การจำแนกชนิดของพลอยโดยใช้คุณลักษณะของสี
Automatic Gem Classification using Color-based Features

ปัทมพร เก่งการรบ (Pattamaporn Kengkarnrob)[1] และเสาวลักษณ์ วรรธนาภา (Saowaluk Watanapa)[2]
สาขาวิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยธรรมศาสตร์
[1]preawploy@gmail.com, [2]scw@cs.tu.ac.th

บทคัดย่อ

งานวิจัยนี้นำเสนอวิธีการจำแนกพลอยจากภาพถ่าย โดยวิเคราะห์จากคุณลักษณะของสี ได้แก่ RGB HSV และ YCbCr ซึ่งเป็นประโยชน์อย่างมากสำหรับผู้ซื้อที่ขาดความรู้และความชำนาญในการจำแนกชนิดของพลอยในการเลือกซื้อเบื้องต้น ด้วยเหตุนี้งานวิจัยนี้จึงได้พัฒนาระบบสำหรับการจำแนกพลอยด้วยคุณลักษณะของสีแบบอัตโนมัติและทดสอบประสิทธิภาพการจำแนกพลอยโดยใช้เทคนิคการเรียนรู้ของเครื่องจักร (Machine Learning) ได้แก่ Mimimum Distance และ K-Nearest Neighbor ซึ่งผลการทดลองพบว่าการจำแนกชนิดโดยวิธี Minimum Distance มีความถูกต้อง 87.83% และการจำแนกโดยใช้ข้อมูลใกล้ที่สุด K ตัว โดยให้ K = 1, 3, 5 และ 7 ซึ่งมีความถูกต้อง 96.33 96.17% 95.17% และ 94.83% ตามลำดับ

คำสำคัญ: การจำแนกประเภทอัญมณี คุณลักษณะของสี

Abstract

This paper proposes an approach to the classification of gems, using color-based features including RGB HSV and YCbCr. Such a categorization can be highly beneficial to numerous customers in Thai gem and jewelry markets for their preliminary decision. In this research, we also implement an automated gem classification tool adopting a couple of machine learning techniques, namely Minimum Distance and K-nearest Neighbor. With regard to the experimental results, the percentage of classification accuracy via minimum distance is 87.83% accurate, while the figures of the K-nearest neighbor designated at 1, 3, 5, and 7 account for 96.33% 96.17% 95.17% and 94.83% respectively.

Keyword: Classification Gems, Color Models

1. บทนำ

ปัจจุบันพลอยเป็นที่นิยมอย่างมากทั้งในประเทศไทยและต่างประเทศ เนื่องจากมีความสวยงามและเพื่อการนำโชค ทำให้พลอยเป็นสินค้าส่งออกเป็นอันดับสี่ และนำเข้าเป็นอันดับห้า ซึ่งกำลังขยายตัวมากในตลาดอัญมณีของไทย [1]

อย่างไรก็ตามพลอยเป็นอัญมณีที่เกิดจากแร่ธาตุหลายตระกูลและมีสีหลากหลายสี บางชนิดมีความคล้ายคลึงกันแต่ราคาแตกต่างกันมาก เม็ดหนึ่งมีราคาหลักร้อยส่วนอีกเม็ดอาจราคาหลักหมื่น ทำให้ผู้จำหน่ายพลอยบางรายใช้ช่องว่างดังกล่าวในการหลอกลวงผู้บริโภคที่ขาดความชำนาญในเรื่องดังกล่าว

จากปัญหาดังกล่าว งานวิจัยนี้จึงนำเสนอวิธีการจำแนกชนิดของพลอยที่เป็นที่นิยมในประเทศไทย โดยประยุกต์วิธีการเธรดโชลด์ [2],[3],[4] และมอร์โฟโลยี [5] ที่สำคัญงานวิจัยนี้ใช้คุณลักษณะของสี ได้แก่ RGB HSV และ YCbCr [6] ในการปรับปรุงภาพและประยุกต์การเรียนรู้ของเครื่องจักร (Machine Learning) ประกอบด้วย Minimum distance และ K-Nearest Neighbor [7] ในการจำแนกชนิดของพลอย

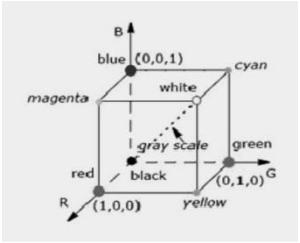
2. วรรณกรรมที่เกี่ยวข้อง

ในส่วนนี้นำเสนอเกี่ยวกับ ทฤษฎีที่เกี่ยวข้อง ได้แก่ แบบจำลองสี (Color Models) การแยกส่วนของภาพ (Image Segmentation) การจำแนกประเภทข้อมูลภาพ (Image Classification) และงานวิจัยที่เกี่ยวข้อง

2.1 แบบจำลองสี (Color Models)

(1) แบบจำลองสี RGB

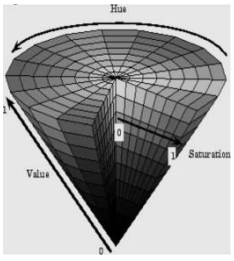
เป็นแบบจำลองที่ประกอบด้วยแม่สีหลัก 3 สี ได้แก่ สีแดง (Red) สีเขียว (Green) และสีน้ำเงิน (Blue) เมื่อนำแบบจำลองนี้มาระบุบนพิกัดคาร์ทีเซียน (Cartesian coordinate system) ดังภาพที่ 2.1 จะเห็นได้อย่างชัดเจนว่าแบบจำลองดังกล่าวถูกแสดงด้วยลูกบาศก์ (Cube) ที่มีจุดกำเนิด (0,0,0) คือค่าสีดำ และ (1,1,1) คือค่าสีขาว สำหรับสีเทา(Gray Scale) คือเส้นทแยงมุมจากจุดสีดำมาที่ตำแหน่งจุดสีขาว[6] โดยแบบจำลองนี้นิยมใช้ในงานด้านโทรทัศน์และหน้าจอคอมพิวเตอร์



ภาพที่ 2.1 : แบบจำลองสี RGB

(2) แบบจำลองสี HSV

เป็นแบบจำลองที่อยู่บนพื้นฐานการมองเห็นของมนุษย์ อีกทั้งยังใช้กันอย่างแพร่หลายในวงการคอมพิวเตอร์วิสัยทัศน์สำหรับขั้นตอนการแบ่งส่วนภาพ[6] โดยระบบสี HSV ประกอบด้วยค่าสี (Hue : H) ค่าความอิ่มตัวของสี (Saturation : S) และค่าความสว่างของสี (Value : V) มีประโยชน์สำหรับการประมวลผลภาพเป็นอย่างมาก เนื่องจากมีความผันแปรต่อแสงน้อย ทำให้เหมาะสมในงานถ่ายภาพอัญมณีที่ต้องถ่ายจากแหล่งแสงที่ควบคุมได้ยาก เช่น ความวาวและความสะท้อนแสงของอัญมณี[8] ภาพที่ 2.2 แสดงแบบจำลองดังกล่าว



ภาพที่ 2.2 : แบบจำลองสี HSV

(3) แบบจำลองสี YCbCr

แบบจำลองนี้ใช้สำหรับงานด้านวิดีโอและถูกกำหนดไว้ในมาตรฐาน ITU-R BT.601-5 และ ITU-R BT.709-5 ของสหภาพโทรคมนาคมระหว่างประเทศ(International Telecommunication Union) ซึ่งเป็นที่นิยมในสัญญาณทีวียุโรป [6] ซึ่งองค์ประกอบของแบบจำลองสี YCbCr นี้ ค่า Y หมายถึงค่าที่ใช้เก็บข้อมูลค่าความเข้มของแสง (Luminance) ส่วนค่า Cb และ Cr หมายถึง ค่าที่ใช้เก็บข้อมูลของแสงสีน้ำเงินและสีแดงตามลำดับ

2.2 การแยกส่วนของภาพ (Image Segmentation)

การแยกส่วนของภาพ เป็นเทคนิคหนึ่งของการประมวลผลภาพ โดยมีวัตถุประสงค์ในการแยกวัตถุ (Object) ที่สนใจออกจากพื้นหลัง (Background) ของภาพ โดยวิธีที่นิยมใช้กันอย่างแพร่หลายมี 2 วิธี [12] ได้แก่ การหาขอบภาพ (Edge Detection) เทคนิคการทำเธรดโชลด์ (Threshold) ในการวิจัยนี้เลือกการทำเธรดโชลด์และมอร์โฟโลจี ในการปรับปรุงคุณภาพภาพ

(1) การทำเธรดโชลด์ (Threshold)

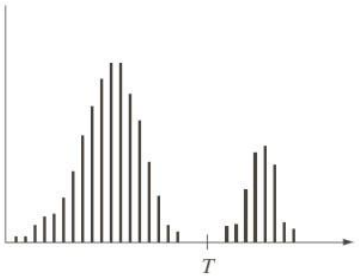
เป็นเทคนิคที่ไม่ซับซ้อนในการแบ่งส่วนภาพ ซึ่งเป็นการแยกวัตถุ (Object) ออกจากพื้นหลัง (Background) โดยพิจารณาจากฮีสโตแกรม (Histogram) ของภาพ โดยการนำภาพสีมาแปลงให้เป็นภาพสีเทาและแปลงภาพเป็นไบนารี (Binary) ที่มีค่าระดับความเข้มเทาเพียงสองระดับ ได้แก่ ขาวและดำ[2],[3]

โดยทั่วไปการกำหนดค่าเธรดโชลด์ (T) คือเลือกค่าเฉลี่ย แต่หากพิกเซลของวัตถุมีความสว่างมากกว่าพื้นหลังสม่ำเสมอและไม่มีสัญญาณรบกวน (Noise) ค่าเธรดโชลด์ที่เป็นค่าเฉลี่ยหรือค่ากลางจะทำงานได้ดี[4] โดยการทำเธรดโชลด์สามารถแบ่งเป็นการกำหนดค่าเธรดโชลด์หนึ่งระดับ (Single Threshold) และเธรดโชลด์หลายระดับ (Multiple thresholding) ซึ่งงานวิจัยนี้เลือกวิธีการกำหนดค่าเธรดโชลด์ด้วยวิธี Otsu ซึ่งเป็นวิธีที่นิยมแพร่หลายในงานอุตสาหกรรม [4]

การกำหนดค่าเธรดโชลด์หนึ่งระดับ (Single Threshold) เป็นการกำหนดค่าความเข้มให้มีค่าที่สามารถแยกความแตกต่างของวัตถุและพื้นหลังได้อย่างชัดเจน การวิเคราะห์เลือกค่าเธรดโชลด์ (T) จะวิเคราะห์จากภาพฮีสโตแกรมดังภาพที่ 2.3 เมื่อได้รับค่าดังกล่าวแล้ว จะทำการแปลงภาพตั้งต้นเป็นภาพไบนารี่ที่มีเพียงค่า 0 และ 1 โดยหากค่าสีในพิกเซลไหนมีค่าน้อยกว่า T

จะกลายเป็น 0 และ 1 ในกรณีที่มีสีในพิกเซลมีค่ามากกว่า T ตามสมการที่ 2.1 [13]

$$g(x,y) = 1, if\ f(x,y) > T \quad (2.1)$$
$$0, if\ f(x,y) \leq T$$



ภาพที่ 2.3 : แสดงฮีสโตแกรมของภาพเธรดโชลด์หนึ่งระดับ

(2) การแบ่งส่วนภาพด้วยวิธีการมอร์โฟโลจี (Morphological Image Segmentation)

การแบ่งส่วนภาพด้วยวิธีการมอร์โฟโลจีเป็นวิธีการปรับปรุงการแบ่งส่วนภาพที่มีประสิทธิภาพกว่าแบบเธรดโชลด์ เนื่องจากบางครั้งภาพต้นฉบับอาจมีการสะท้อนของแสงและเงาของวัตถุ หรือวัตถุมีส่วนที่ทับซ้อนกัน ทำให้การแบ่งส่วนภาพโดยเธรดโชลด์อาจได้ผลลัพธ์ที่น้อยกว่า โดยกระบวนการมอร์โฟโลจีมีหลายกหลายวิธี เช่น Erosion, Dilation, Opening, Closing, Thinning และ Thickening เป็นต้น โดยงานวิจัยนี้เลือกวิธีการทำ Opening [5] ซึ่งสามารถอธิบายได้ดังนี้

การทำ Openning [5] คือการทำ Erosion เพื่อให้ภาพมีขนาดเล็กลงและทำการกำจัดจุดเล็กของภาพที่ไม่ต้องการ และตามด้วยการทำ Dilation ซึ่งเป็นการนำผลภาพจากขั้นตอนก่อนหน้ามาขยาย ซึ่งผลลัพธ์ที่ได้จากวิธี Opening ภาพจะถูกเปิดกว้างมากขึ้น โดยไม่เสียรูปทรง สมการที่ 2.2 โดยที่ I คือภาพต้นฉบับ

$$opening(I) = (dilation(erosion(I)) \quad (2.2)$$

2.3 การจำแนกประเภทข้อมูลภาพ (Image Classification)

การเรียนรู้ของเครื่องจักร (Machine Learning) เป็นเทคนิคที่ใช้สำหรับคอมพิวเตอร์เพื่อสนับสนุนการเรียนรู้ โดยจะมีเทคนิคหลากหลายวิธี ได้แก่ การจำแนกประเภทโดยวิธีวัดระยะทางน้อยที่สุด (Minimum Distance) การจำแนกประเภทโดยใช้ข้อมูลใกล้ที่สุด K ตัว (K-Nearest Neighbors) และซัพพอร์ตเวคเตอร์แมชชีน (Support Vector Machines) เป็นต้น งานวิจัยนี้สนใจเทคนิคการจำแนกประเภทโดยวิธีวัดระยะทางน้อยที่สุด (Minimum Distance) และการจำแนกประเภทโดยใช้ข้อมูลใกล้ที่สุด K ตัว (K-Nearest Neighbors) เนื่องจากการศึกษาพบว่า เป็นวิธีที่รวดเร็วต่อการฝึกระบบ มีประสิทธิภาพสูงในกรณีที่ข้อมูลมีขนาดใหญ่และมีสัญญาณรบกวน ซึ่งหากข้อมูลการตัดสินใจมีความซับซ้อนวิธีนี้สามารถสร้างโมเดลที่มีประสิทธิภาพ[7]

การจำแนกประเภทโดยวิธีวัดระยะทางน้อยที่สุด (Minimum Distance) โดยวิธีการนี้จะมีโมเดลแต่ละคลาสเป็นเวคเตอร์ของค่าเฉลี่ยกลางของเซตของเวคเตอร์ฟีเจอร์ในแต่ละคลาส โดยจะทำการตัดสินว่าวัตถุที่กำลังพิจารณา ซึ่งมีเวคเตอร์ฟีเจอร์ x เป็นคลาสใดจากโมเดลของคลาสที่อยู่ใกล้เคียงกับตัวมันมากที่สุด สามารถเขียนสมการดังสมการที่ 2.3 [5]

$$m_j = \frac{1}{n_j}\sum_{i=1}^{n_j} x_i^j \quad (2.3)$$

การจำแนกประเภทโดยใช้ข้อมูลใกล้ที่สุด K ตัว (K-Nearest Neighbors : KNN) โดยวิธีการใช้ข้อมูลใกล้ที่สุด K ตัว ซึ่งพิจารณาตัวที่ใกล้ที่สุดโดยการโหวตของเวคเตอร์ฟีเจอร์ ถ้า K ตัวนั้นเป็นเวคเตอร์ฟีเจอร์ของคลาสใดมากที่สุด วัตถุทดสอบก็จะถูกจำแนกว่าเป็นคลาสนั้น ซึ่งจะสามารถประยุกต์ใช้มาตรวัดความแตกต่าง เช่น การวัดระยะทางที่ใกล้ที่สุดแบบยูคลิเดียน (Euclidean distance) ดังสมการที่ 2.4[5]

$$dist\ (X_1, X_2) = \sqrt{\sum_{i=1}^{n}(X_{1i} - X_{2i})^2} \quad (2.4)$$

2.4 งานวิจัยที่เกี่ยวข้อง

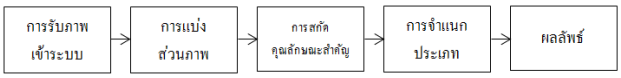
จากการศึกษางานวิจัยที่ผ่านมาพบว่า งานวิจัยด้านการประมวลผลและการจำแนกข้อมูลมีหลากหลาย แต่ยังไม่มีงานวิจัยใดที่เกี่ยวข้องโดยตรงในการจำแนกชนิดของพลอยจากภาพถ่าย ดังนั้นผู้วิจัยได้ศึกษางานวิจัยที่ใกล้เคียงจากทั่วโลก เช่น งานวิจัยหนึ่งที่ทำการวิจัยเกี่ยวกับระบบตรวจสอบสีเพชรจากภาพถ่าย โดยใช้อัลกอริทึมโซเบล (Sobel) ในการตรวจจับขอบของเพชร และใช้ค่า Saturation จากโมเดล HSV ในการแบ่งระดับสีเพชร จากนั้นจำแนกสีของเพชรโดยใช้วิธีวัดค่า

ระยะทางที่ใกล้ที่สุด (Euclidean distance) [8] โดยงานวิจัยนี้ผู้วิจัยสนใจวิธีการสกัดคุณลักษณะสำคัญของภาพและการจำแนกประเภท และงานวิจัยต่อมาได้ศึกษางานวิจัยอีกงานพบว่ามีงานวิจัยเกี่ยวกับการจำแนกชนิดอัญมณีที่มีข้อบกพร่อง [9] โดยใช้การจำแนกชนิดของข้อมูลโดยอัลกอริทึม decision trees และ decision thresholds สำหรับการฝึกระบบและใช้เทคนิคทางการเรียนรู้ของเครื่องจักร (Machine Learning) หลายเทคนิคเปรียบเทียบกัน โดยผลการทดลองพบว่าการจำแนกชนิดโดยใช้ข้อมูลใกล้ที่สุด (KNN) มีความถูกต้องสูงและมีการประมวลผลของข้อมูลที่รวดเร็ว ต่างจากการจำแนกชนิดโดยซัพพอร์ตเวคเตอร์แมชชีน (SVM) ซึ่งเป็นอัลกอริทึมที่มีความซับซ้อน ทำให้มีการประมวลผลที่ช้า งานวิจัยต่อมาที่ผู้วิจัยสนใจเกี่ยวกับระบบสี ได้แก่ งานวิจัยเกี่ยวกับการพัฒนาระบบการสื่อสารเรื่องสี เพื่อประเมินคุณภาพของอัญมณีในการออกใบรับรองคุณภาพอัญมณีของสถาบันวิจัยพัฒนาอัญมณีและเครื่องประดับแห่งชาติ [10] ซึ่งเป็นงานวิจัยเกี่ยวกับเรื่องสีเป็นหลัก โดยอาศัยระบบสีของมันเซลล์ (Munsell) ซึ่งจะบอกถึงลักษณะสีสัน(Hue) ความสว่างของสี(Value) และความอิ่มตัวของสี (Chroma) โดยจะแปลงค่ารหัสสีที่ได้เป็นชื่อสีสากลในระบบสี ISCC-NBS ซึ่งเป็นรหัสและชื่อสีสากลในการสื่อสารเรื่องสีของอัญมณี หลังจากนั้นผู้วิจัยได้ทำงานศึกษางานวิจัยที่ทำการจำแนกดอกไม้ ได้แก่ งานวิจัยที่จำแนกดอกไม้จาก 25 ชนิดทั่วโลก โดยใช้อัลกอริทึม KNN ในการจำแนกชนิดของดอกไม้[11] ซึ่งผลลัพธ์มีความถูกต้อง 98.88% โดยมีงานวิจัยอื่นที่นำเสนอวิธีที่แตกต่างกันออกไป เช่น Nearest Neighbor Support Vector Machine Multiple Kernel และ Probabilistic Neural Network พบว่าวิธีการจำแนกชนิดโดยใช้ข้อมูลใกล้ที่สุด K ตัว ยังคงเป็นวิธีที่ได้ผลลัพธ์ดี ในงานวิจัยที่จำแนกชนิดโดยสนใจคุณลักษณะด้านสีเป็นหลัก

3. วิธีการดำเนินการวิจัย

ภาพที่ 3.1 แสดงภาพรวมการดำเนินงานวิจัยนี้ โดยมี 5 ขั้นตอน ได้แก่ 1. การรับภาพเข้าระบบ (Image Acquistion) 2. การแบ่งส่วนภาพ (Segmentation) 3. การสกัดคุณลักษณะ (Feature Extraction) 4. การจำแนกประเภท (Classification) และ 5. ผลลัพธ์

ขั้นตอนที่ 1 การรับภาพเข้าระบบ (Image Acquistion) ในขั้นตอนการเตรียมภาพ ผู้วิจัยใช้กล้องถ่ายรูปดิจิตอลความละเอียด 14.2 ล้านพิกเซล ซึ่งถ่ายจากด้านหน้าของพลอยทั้งหมด 12 ชนิดชนิดละ 50ภาพ รวมทั้งหมด 600 ภาพ โดยภาพเป็นไฟล์นามสกุล JPEG และพลอยถ่ายภายในตู้แสงที่อุณหภูมิ 5,000 เคลวิน ซึ่งเป็นแสงที่มีความใกล้เคียงแสงธรรมชาติในตอนกลางวัน [10] และถ่ายภาพพลอยจากด้านบนของตู้แสงที่มีระยะห่างระหว่างกล้องกับตู้แสง 6 นิ้ว[14] การจัดวางกล้องเพื่อถ่ายพลอยและภาพตัวอย่างพลอยแต่ละชนิดแสดงอยู่ภาพที่ 3.2 และ 3.3 ตามลำดับ



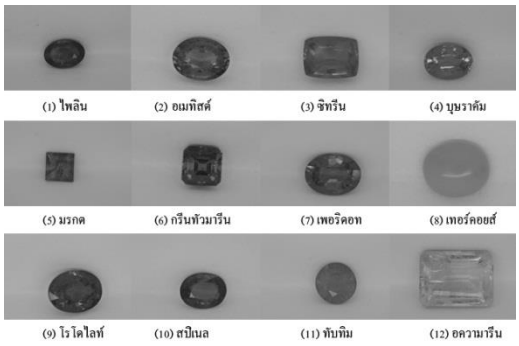
ภาพที่ 3.1 : ขั้นตอนการดำเนินงาน



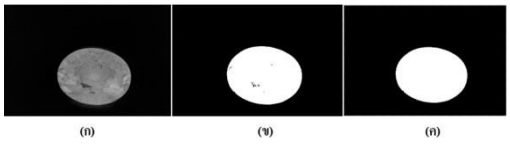
ภาพที่ 3.2 : การจัดวางกล้องเพื่อถ่ายภาพพลอย

ขั้นตอนที่ 2 การแบ่งส่วนภาพ (Image Segmentation) โดยวิเคราะห์จากค่า Saturation ที่แปลงจาก RGB เป็น HSV เนื่องจากค่า Saturation จากสเปซสี HSV แสดงความอิ่มตัวของสี ซึ่งเม็ดพลอยแต่ละชนิดจะมีค่า Saturation สูงกว่าพื้นหลังอย่างชัดเจน จึงใช้ค่านี้ในการแบ่งส่วนภาพเม็ดพลอยออกจากพื้นหลังได้อย่างมีประสิทธิภาพ โดยการแบ่งส่วนภาพด้วยภาพระดับสีเทาจากสเปซสี RGB นั้น ยังขาดประสิทธิภาพ เนื่องจากเม็ดพลอยที่ใช้ในการทดลองมีหลากหลายสี ทำให้ยากต่อการแบ่งส่วนภาพ จากนั้นทำเธรดโชลด์ด้วยวิธี Otsu[15] ซึ่งปริมาณความเข้มของแสงตรงเม็ดพลอยกับพื้นหลังจะแตกต่างกันชัดเจน จึงเลือกใช้ค่าเธรดโชลด์หนึ่งระดับ โดยในแต่ละภาพจะได้ค่าเธรดโชลด์อัตโนมัติ ดังภาพที่ 2.3 หลังจากนั้นทำการ

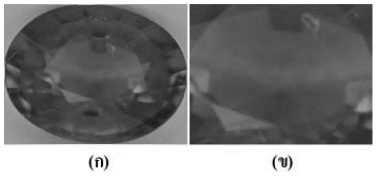
ปรับปรุงภาพด้วยวิธี Morphology Opening โดยใช้มาสก์(mask) ขนาด 3 x 3 พิกเซล เพื่อกำจัดรายละเอียดขนาดเล็กของภาพ แสดงในภาพที่ 3.4 จากนั้นกำหนด Bounding Box เพื่อสร้าง Region ที่เราสนใจ (Region of Interest : ROI) โดยมีเฉพาะสีของพลอย แล้วทำการกำหนดกรอบสี่เหลี่ยมจัตุรัส จากจุดศูนย์กลางของภาพ โดยการวัดระยะความกว้างครึ่งหนึ่ง จากแกน X และวัดระยะทางความสูงครึ่งหนึ่งของแกน Y ซึ่งผลลัพธ์จะแสดงดังภาพที่ 3.5



ภาพที่ 3.3 : ตัวอย่างภาพถ่ายพลอยแต่ละชนิด



ภาพที่ 3.4 : การแปลงภาพ RGB เป็น HSV (ก) ภาพไบนารี่ (ข) และการทำ Morphology Opening (ค)



ภาพที่ 3.5 : ส่วนของวัตถุที่เราสนใจ (ก) และการกำหนดกรอบสี่เหลี่ยมจัตุรัส (ข)

ขั้นตอนที่ 3 การสกัดคุณลักษณะ (Feature Extraction) เก็บค่าเฉลี่ยของแต่ละคลาสทั้งหมด 12 คลาส ทั้งเก้าคุณลักษณะที่ได้จากโมเดลสีแบบ RGB HSV และ YCbCr แสดงดังภาพที่ 3.6 โดยเหตุผลที่ใช้ทั้งเก้าคุณลักษณะเพราะเมื่อนำค่าสีจากพลอยที่มีสีใกล้เคียงกัน เช่น ซิทรินกับบุษราคัม ซึ่งมีสีเหลืองส้มเหมือนกัน, มรกตกับกรีนทัวมาลีนซึ่งมีสีเขียว, และทับทิมสปิเนลกับโรโดไลท์ซึ่งมีสีแดง มาพล็อตกราฟย่อย 2 มิติ เช่น R-G, G-B, R-B, H-S, H-V, Y-Cb, และ Y-Cr เพื่อวิเคราะห์หาการกระจายตัวของค่าสี จะเห็นได้ว่าทั้งเก้าคุณลักษณะนี้มีความสำคัญในการจำแนกชนิดพลอยดังกล่าว

	ค่าเฉลี่ยของเก้าคุณลักษณะ								
ชนิดของพลอย	R	G	B	H	S	V	Y	Cb	Cr
อเมทิสต์	95.26	72.76	123.47	74.84	41.65	48.43	89.23	146.89	134.23
อะความารีน	115.76	137.86	142.81	52.49	19.46	56.19	129.21	133.43	117.94
บลูแซฟไฟร์	39.50	51.22	89.28	60.49	53.39	35.09	60.71	146.41	120.11
ซิทริน	160.70	78.03	4.00	7.79	97.49	63.02	96.99	83.19	169.60
มรกต	31.33	103.01	65.71	41.26	71.07	40.40	82.41	122.27	99.19
กรีนทัวมาลีน	65.98	100.35	72.38	36.01	41.70	39.38	90.62	120.87	114.88
เพอริดอท	105.80	118.53	23.60	18.98	80.65	46.49	105.24	88.12	129.22
โรโดไรท์	85.95	48.29	44.07	17.81	48.83	33.71	66.73	120.63	144.86
ทับทิม	124.37	45.84	72.73	92.47	64.15	48.77	78.16	128.18	160.54
สปิเนล	96.60	49.19	40.19	3.35	57.43	37.88	69.54	117.09	149.48
เทอร์คอยส์	71.09	147.17	152.77	50.73	53.79	60.48	123.40	141.67	94.18
บุษราคัม	155.06	108.52	17.79	11.05	88.73	60.81	112.27	81.21	154.92

ภาพที่ 3.6 : ค่าเฉลี่ยในแต่ละคลาส ทั้งเก้าคุณลักษณะ

ขั้นตอนที่ 4 การจำแนกชนิดของพลอย (Classification) ในขั้นตอนนี้ใช้เทคนิคการวัดระยะทางที่ใกล้ที่สุดแบบยูคลิเดียน โดยใช้เทคนิคทางการเรียนรู้สองเทคนิคเปรียบเทียบกัน ได้แก่ Minimum Distance และ K-Nearest Neighbor โดยบันทึกผลการทดลองในรูปแบบ Confusion Matrix

4. ผลการดำเนินงาน

ในการจำแนกชนิดของพลอย ผู้วิจัยได้พัฒนาโปรแกรมสำหรับจำแนก โดยประยุกต์ตามขั้นตอนวิธีที่กล่าวในหัวข้อที่ 3 ซึ่งโปรแกรมดังกล่าวถูกพัฒนาโดยเครื่องมือ MATLAB

ภาพที่ 3.7 แสดงผลการทดลองวิธีจำแนกชนิดด้วย Minimum Distance มีผลความถูกต้อง 87.83% และอีกวิธีหนึ่งโดยวิธีการจำแนกชนิดแบบ KNN โดยกำหนดให้ K=1 มีความถูกต้องมากที่สุดคิดเป็น 96.33% แสดงดังภาพที่ 3.8 และให้ K = 3, 5, และ 7 มีความถูกต้อง 96.17% 95.17% และ 94.83% ตามลำดับ โดยเมื่อเปรียบเทียบ 2 วิธีนี้ วิธีการจำแนก

ชนิด โดยกำหนดให้ K=1 ระบบมีความถูกต้องเพิ่มขึ้นจากวิธีเก่าคิดเป็น 8.5%

	Confusion Matrix for Minimum Distance												
		Predicted											
		AMT	AQM	BSP	CIT	EMR	GTM	PER	RHO	RUB	SPN	TQS	YSP	รวม	Accuracy
Actual class	AMT	42	8											50	84%
	AQM		50											50	100%
	BSP	1		46					3					50	92%
	CIT				49								1	50	98%
	EMR					43	7							50	86%
	GTM		8			7	28		7					50	56%
	PER							50						50	100%
	RHO								40		10			50	80%
	RUB									50				50	100%
	SPN								15		35			50	70%
	TQS											50		50	100%
	YSP				6								44	50	88%
	Average Accuracy														87.83%

ภาพที่ 3.7 ผลการจำแนกประเภทโดย Minimum Distance

	Confusion Matrix for K Nearest Neighbor by K = 1												
		Predicted											
		AMT	AQM	BSP	CIT	EMR	GTM	PER	RHO	RUB	SPN	TQS	YSP	รวม	Accuracy
Actual class	AMT	50												50	100%
	AQM		50											50	100%
	BSP			50										50	100%
	CIT				50									50	100%
	EMR					49	1							50	98%
	GTM					1	49							50	98%
	PER							50						50	100%
	RHO								41		9			50	82%
	RUB									50				50	100%
	SPN								11		39			50	78%
	TQS											50		50	100%
	YSP												50	50	100%
	Average Accuracy														96.33%

ภาพที่ 3.8 ผลการจำแนกประเภทโดยให้ K = 1

5. สรุป

งานวิจัยนี้ทำการศึกษาคุณลักษณะสำคัญ (Features) ที่เหมาะสม ในการจำแนกชนิดพลอย โดยอาศัยคุณลักษณะของสีและแสงเป็นหลัก ซึ่งผลการทดลองโดยนำค่าของคุณลักษณะข้างต้นมาใช้ในการจำแนกชนิดพลอยพบว่าผลการทดลองสามารถให้ค่าความถูกต้อง (Accuracy) ไม่ต่ำกว่า 80 เปอร์เซ็นต์ และให้ค่าความถูกต้องสูงสุดถึง 96.33%

ในขั้นตอนต่อไป ผู้วิจัยจะทำการจำแนกชนิดของพลอยโดยเจาะจงทำเฉพาะพลอยที่มีสีที่ใกล้เคียงกันมากที่สุด ได้แก่ ซิทรินกับบุษราคัม มรกตกับกรีนทัวมาลีน และโรโดไลท์ทับทิมกับสปิเนล โดยจะนำภาพพลอยที่ถ่ายด้วยกล้องมือถือและถ่ายในสภาพแวดล้อมที่เปิด มาทำการทดลองเพื่อให้เป็นประโยชน์กับผู้บริโภคมากขึ้น

เอกสารอ้างอิง

[1] ศูนย์ข้อมูลอัญมณีและเครื่องประดับ สถาบันวิจัยและพัฒนาอัญมณีและเครื่องประดับแห่งชาติ (องค์การมหาชน) “สถานการณ์การนำเข้าส่งออกสินค้าอัญมณีและเครื่องประดับไทย” พ.ศ. 2558

[2] R.Yogamangalam, and B.Karthikeyan, “Segmentation techniques comparison in image processing,” IJET,2013.

[3] L. Haitao, and L.Shengpu, “An Algorithm and Implementation for Image Segmentation,” *International journal of signal processing, image processing and pattern recognition*, vol. 9, no. 3, pp. 125-132, 2016.

[4] บุญธรรม ภัทราจารุกุล,*การประมวลผลภาพดิจิตอลเบื้องต้น*, : ซีเอ็ดยูเคชั่น, ไทย, 2556

[5] ปกรณ์ แก้วตระกูลพงษ์,*การประยุกต์อัลกอริทึมแมชีนวิชันโดยใช้ NI Vision*, : จุฬาลงกรณ์มหาวิทยาลัย, ไทย, 2559

[6] A.Ibraheem, M.Hasan, Z.Khan, and K.Mishra, “Understanding color models: a review” ARPN Journal of Science and Technology, vol. 2, NO. 3, 2012.

[7] N.Bhatia, and Vandana, “Survey of nearest neighbor techniques,” *International Journal of Computer Science and Information Security*, vol. 8, no. 2, 2010.

[8] อรทัย โรจนเจริญวงศ์, “ระบบตรวจสอบสีเพชรจากภาพถ่าย,” *International Conference on Knowledge and Smart Technologies (KST2010)*,2553

[9] P.Hurtik, M.Burda, and I.Perfilieva, “An image recognition approach to classification of jewelry stone defects,” IEEE,2013.

[10] ศักดา ศิริพันธุ์,”โครงการพัฒนาระบบการสื่อสารเรื่องสี เพื่อการประเมินคุณภาพของอัญมณี ตามแผนปรับโครงสร้างอุตสาหกรรม ระยะที่ 2,” 2546

[11] D.Guru, Y.Sharath, and S.Manjunath, “Texture features and knn in classification of flower images,” RTIPPR,2010

[12] สมเกียรติ อุดมหรรษากุล,*การประมวลผลภาพดิจิตอลเบื้องต้น*, : ท้อป, ไทย, 2554

[13] Stefano Ferrari,“ Image segmentation,” academic year. 2011–2012.

[14] สุมาลี เทพโสพรรณ,วิเคราะห์อัญมณี, : สถาบันอัญมณีวิทย์, ไทย, 2551.

[15] Nobuyuki Otsu, “A threshold selection method from gray-level histograms,” *IEEE Transactions on Systems Man, and Cybernetics*, vol. 9, no. 1, pp. 62-66, 1979.

การเปรียบเทียบประสิทธิภาพอัลกอริทึมในการค้นคืนภาพด้วยสีและพื้นผิว
Comparison of Algorithm's Performance Based on Color and Texture in Image Retrieval

ประยุทธ สุระเสนา (Prayoot Surasena)[1] ศุภเทพ สติมั่น (Supathep Satiman)[2] เอกลักษณ์ ฉิมจารย์ (Ekkalak Chimjan)[3]
สุทัศน์ กำมณี (Sutat Gammanee)[4] เชาวลิต สมบูรณ์พัฒนากิจ (Chaovalit Somboonpattanakit)[5]
อิทธิกร ธรรมจันทึก (Ittikon Thammachantuek)[6] และมาลีรัตน์ โสดานิล (Maleerat Sodanil)[7]
[1]สาขาวิชาคอมพิวเตอร์และเทคโนโลยีสารสนเทศ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏนครสวรรค์
[2]สาขาวิชาคอมพิวเตอร์ธุรกิจ คณะบริหารธุรกิจและการจัดการ มหาวิทยาลัยราชภัฏอุบลราชธานี
[3]สาขาวิชาเทคโนโลยีสารสนเทศและการสื่อสาร คณะวิทยาศาสตร์และศิลปศาสตร์ มหาวิทยาลัยเทคโนโลยีราชมงคลอีสาน
[4]สาขาวิชาเทคโนโลยีสารสนเทศ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏกาญจนบุรี
[5-7]ภาควิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
[1]Prayoot@nsru.ac.th, [2]Supathep.s@ubru.ac.th, [3]Ekkalak.ch@rmuti.ac.th, [4]Sutat@kru.ac.th
[5]s5907011956126@email.kmutnb.ac.th, [6]s5907011956029@email.kmutnb.ac.th and [7]Maleerat.s@it.kmutnb.ac.th

บทคัดย่อ

การค้นคืนภาพแบบอิงเนื้อหา (CBIR) เป็นวิธีการหนึ่งที่ได้รับความนิยมในการพัฒนาระบบค้นคืนภาพ ลักษณะเด่นที่นักวิจัยมักนำมาใช้ในการศึกษาเพื่อพัฒนาอัลกอริทึมในการค้นคืนภาพคือ สีและ พื้นผิวของภาพ เนื่องจากสามารถนำมาใช้ค้นคืนภาพได้อย่างมีประสิทธิภาพ จึงเกิดอัลกอริทึมในการค้นคืนภาพด้วยสีและพื้นผิวขึ้นจำนวนมาก แต่พบว่า การศึกษาเปรียบเทียบประสิทธิภาพการทำงานของอัลกอริทึมที่มีอยู่ยังไม่แพร่หลายมากนัก บทความนี้นำเสนอผลการเปรียบเทียบประสิทธิภาพการทำงานของอัลกอริทึมที่ได้รับความนิยมอย่าง Color and Edge Directivity Descriptor (CEDD) Fuzzy Color and Texture Histogram (FCTH) และอัลกอริทึม Joint Composite Descriptor (JCD) ผลการทดลองพบว่า JCD มีความแม่นยำในการค้นคืนสูงที่สุดโดยมีค่าความถูกต้องอยู่ที่ร้อยละ 69

คำสำคัญ: การค้นคืนภาพ การค้นคืนภาพด้วยสี การค้นคืนภาพด้วยพื้นผิว อัลกอริทึมการค้นคืนภาพ ประสิทธิภาพการค้นคืนภาพ

Abstract

An image retrieval with the use of CBIR is a well-known mechanism to generate image retrieval system. According to computing researchers, one of the main purposes of image retrieval is the development of algorithm in retrieving pictures. Using colors and image surfaces, these two applicable mechanisms are effective implementation in image retrieving with the aim to increase a greater number of algorithm potentiality. However, there is a limited number of study with regard to the use of colors and image surfaces in assessing algorithm potentiality. This article will present a comparative study of different models of algorithm potentiality, for example, CEDD, FCTH and JCD. Based on the experiment, it has found that JCD model illustrates the most accurate system in bringing back images with 69% statistical measurement.

Keyword: Image, Image Retrieval, Image Retrieval using Color, Color and Texture, Image Retrieval using Texture.

1. บทนำ

ในปัจจุบันความก้าวหน้าทางเทคโนโลยีสารสนเทศ ได้เข้ามามีบทบาทสำคัญในชีวิตประจำวันของเรามากขึ้นโดยเฉพาะเทคโนโลยีในระบบดิจิทัล มีการเผยแพร่ข้อมูลข่าวสาร การแบ่งปันข้อมูล [1] ทั้งในรูปแบบของรูปภาพ วิดีโอ กราฟิก และรูปแบบอื่นๆ ผ่านทางเครือข่ายการสื่อสารและสื่อสังคมออนไลน์ ทำให้มีการใช้ประโยชน์จากการประมวลผลภาพในด้านต่าง ๆ ทั้งการสำรวจและการวิเคราะห์ข้อมูล เช่น การวิเคราะห์ข้อมูลทางการแพทย์ การสำรวจข้อมูลทรัพยากรธรรมชาติ การพยากรณ์อากาศจากการประมวลภาพถ่ายดาวเทียม และการวิเคราะห์ข้อมูลจากคลังข้อมูลภาพเป็นต้น

การค้นคืนภาพแบบอิงเนื้อหา (Content Based Image Retrieval ; CBIR) [1-2] เป็นเทคนิคการค้นคืนภาพโดยใช้เนื้อหาของภาพในการสืบค้นข้อมูล และเป็นเทคนิคที่ได้รับความนิยมในการจัดทำระบบการค้นคืนภาพ เนื้อหาของภาพที่มักจะนำมาใช้ในเทคนิคนี้ได้แก่ สี (Color) รูปร่าง (Shape) และพื้นผิว (Texture) ของภาพ การศึกษางานวิจัยในช่วงระยะเวลาที่ผ่านมาพบว่า อัลกอริทึม Color and Edge Directivity Descriptor (CEDD), Fuzzy Color and Texture Histogram (FCTH) และ Joint Composite Descriptor (JCD) เป็นอัลกอริทึมที่มีประสิทธิภาพและได้รับความนิยมในการค้นคืนภาพด้วยสีและลักษณะของพื้นผิว [4], [5], [6], [15], [16], [17] แต่การศึกษาในลักษณะที่เปรียบเทียบประสิทธิภาพการทำงานของทั้ง 3 อัลกอริทึมนั้นยังพบไม่มากนัก งานวิจัยฉบับนี้จึงมีวัตถุประสงค์เพื่อศึกษาการทำงาน และนำเสนอการเปรียบเทียบประสิทธิภาพของทั้งสามอัลกอริทึม เมื่อนำมาทดสอบกับชุดข้อมูลที่มีความยากและท้าทายยิ่งขึ้น

2. งานวิจัยที่เกี่ยวข้อง

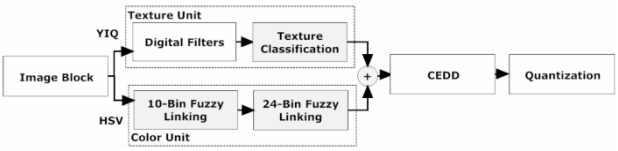
2.1 Texture Similarity

Texture Similarity หรือการค้นคืนภาพด้วยคุณลักษณะพื้นผิวเป็นคุณลักษณะที่สำคัญในการค้นคืนภาพชนิดหนึ่งนอกจากการค้นคืนภาพด้วยคุณลักษณะของสี (color feature) คุณลักษณะด้านรูปทรง (shape feature) การค้นคืนด้วยลักษณะพื้นผิวคือการค้นคืนด้วยลักษณะเด่นที่มีรูปแบบซ้ำกันในบริเวณหนึ่งซึ่งได้มีการใช้วิธีการค้นคืนจากการเทียบเคียงมุมมองของมนุษย์โดยในงานวิจัยในช่วงแรกได้แบ่งพื้นผิวออกเป็น 6 ลักษณะ [5] การค้นคืนด้วยพื้นผิวในช่วงดังกล่าวมีการใช้แนวคิดการค้นหาความคล้ายแบบ point-by-point โดยใช้ค่า mean squared error (MSE) และ peak signal-to-noise ratio (PSNR) [6] จากนั้นจึงได้มีการพัฒนางานวิจัย [7],[8] ขึ้นเป็นการค้นคืนด้วยพื้นผิวนิยมที่มีลักษณะเป็นสีเทา (Grey Scale) และเปรียบเทียบในลักษณะของพิกเซลในบริเวณใกล้เคียง ในงานวิจัย [9] มีการนำเสนอวิธีการ structural texture similarity (STSIMs) [10],[11] ซึ่งแตกต่างจากงานวิจัยในคุณลักษณะของการค้นคืนด้วยพื้นผิวในงานที่ผ่านมา โดยพยายามทำการประเมินผลแบบ point-by-point ให้เป็นแบบโครงสร้างโดยการใช้หลักสถิติดูความสัมพันธ์เป็นพื้นที่หลังจากนั้นจึงได้พัฒนาเพิ่มจนเป็น STSIM-2 และ STSIM-M [9],[12] ซึ่งมีการประยุกต์ใช้ค่า mean และค่า correlation มาหารูปแบบที่มีความซับซ้อนสูงจนถึงปัจจุบันได้นำเอาแนวคิดเรื่องการค้นคืนด้วยสีมาประยุกต์รวมเข้ากับการค้นคืนแบบพื้นผิวโดยตั้งอยู่บนแนวคิดว่ามนุษย์ไม่สามารถรับรู้สีเป็นจำนวนมากพร้อมกันได้ [13],[14]

Compact Composite Descriptors (CCD) เป็นกลุ่มของ low lever features ที่ใช้เก็บรายละเอียดของข้อมูลประเภทมัลติมีเดียซึ่งใช้ได้ทั้งรูปภาพและวีดีโอ เป็น features extraction ที่มีจุดเด่นในเรื่องของระยะเวลาที่ใช้ในการประมวลผล ถูกออกแบบให้มีการเก็บ index ที่มีขนาดเล็กเพื่อลดเวลาในการทำงาน โดย Compact Composite Descriptors แบ่งออกเป็น ประเภท 2 คือ กลุ่มแรกจะใช้ข้อมูลจากสี (color) และพื้นผิว (texture) มาเป็นปัจจัยในการพิจารณาคัดแยกเนื้อหาของภาพ (visual content) ที่เป็นภาพสี เช่น อัลกอริทึม Color and Edge Directivity Descriptor (CEDD) และ Fuzzy Color and Texture Histogram (FCTH) ส่วนอีกกลุ่มหนึ่งจะใช้ข้อมูลจากระดับความสว่าง (brightness) และพื้นผิว (texture) มาเป็นปัจจัยในการพิจารณาคัดแยกเนื้อหาภาพ (visual content) ที่เป็นภาพในระดับสีเทา (grayscale images) เช่น ภาพที่ใช้ทางการแพทย์รังสีวิทยา เป็นต้น ซึ่งในงานวิจัยนี้ให้ความสนใจกับภาพสีจึงทำการศึกษาโครงสร้างของตัว Descriptor ของอัลกอริทึม CEDD และ FCTH

2.2 Color and Edge Directivity Descriptor (CEDD)

CEDD เป็นอัลกอริทึมที่มีการประยุกต์ใช้คุณลักษณะการค้นคืนของพื้นผิวผสมกับการค้นคืนด้วยสีที่ได้รับความนิยมและมีประสิทธิภาพ [15] หลักการทำงาน แสดงดังภาพที่ 1



ภาพที่ 1: การทำงานของ CEDD

หลักการทำงานของ CEDD คือใช้คุณลักษณะในระดับ Low-Level โดยมีการใช้ค่าฮิสโตแกรมของค่าสีและค่าพื้นผิวประมวลผลโดย Fuzzy-Linking Histogram งานวิจัยดังกล่าวได้แปลงค่าสีให้อยู่ในโหมด Hue Saturation Value (HSV) และจัดให้อยู่ในรูปแบบ 10 เฉดสี แต่ละเฉดสีแทนด้วยค่าสีทั้งหมด 10 สีในส่วนของค่าพื้นผิวได้ใช้มาตรฐาน MPGE-7 ใช้ค่าฮิสโตแกรม 5 หลักในการจำแนกค่าพื้นผิวโดยใช้แนวคิดของ Image Block ซึ่งแบ่งออกเป็น 4 Sub Blcok เมื่อหาค่าผลลัพธ์ออกมาจะได้ค่าที่มีความเกี่ยวเนื่องกันในพื้นที่ใกล้เคียงกัน 6 พื้นที่ โดยคำนวณหาระยะห่างระหว่างภาพของอัลกอริทึม CEDD ดังสมการที่ (1)

$$T_{ij} = t(x_i, x_j) = \frac{x_i^T x_j}{x_i^T x_i + x_j^T x_j - x_i^T x_j} \quad (1)$$

โดย x^T คือ ค่าทรานสโพสของเวกเตอร์ x

i คือ ค่า Sub Block ในแนวตั้งจากเส้นทะแยงมุม 45 องศา

j คือ ค่า Sub Block ในแนวนอน จากเส้นทะแยงมุม 135 องศา

คุณลักษณะของ Sub Block เป็นเส้นแบบไม่มีทิศทาง แทนด้วย $f_v(k), f_h(k), f_{d\text{-}45}(k), f_{d\text{-}135}(k)$ และ $f_{nd}(k)$ โดยที่ k = 0,…,3 แทนพื้นที่ของ Sub Block โดยขนาดของ Edge แทนด้วย $m_v(i,j), m_h(i,j), m_{d\text{-}45}(i,j), m_{d\text{-}135}(i,j)$, และ $m_{nd}(i,j)$ สำหรับ (i,j) ของ Sub Block คำนวณได้จากสมการ (2) - (6)

$$m_v(i,j) = \left| \sum_{k=0}^{3} a_k(i,j) \times f_v(k) \right| \quad (2)$$

$$m_h(i,j) = \left| \sum_{k=0}^{3} a_k(i,j) \times f_h(k) \right| \quad (3)$$

$$m_{d-45}(i,j) = \left| \sum_{k=0}^{3} a_k(i,j) \times f_{d-45}(k) \right| \quad (4)$$

$$m_{d-135}(i,j) = \left| \sum_{k=0}^{3} a_k(i,j) \times f_{d-135}(k) \right| \quad (5)$$

$$m_{nd}(i,j) = \left| \sum_{k=0}^{3} a_k(i,j) \times f_{nd}(k) \right| \quad (6)$$

โดยคำนวณค่า max ได้จากสมการ (7) และปรับค่า m ของ Image Block ทั้ง 6 ด้วยสมการ (8)

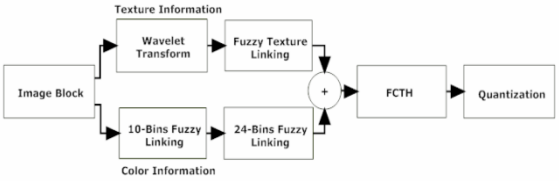
$$\max = MAX(m_v, m_h, m_{d-45}, m_{d-135}, m_{nd}) \quad (7)$$

$$m_v' = \frac{m_v}{\max}, m_h' = \frac{m_h}{\max}, m_{d-45}' = \frac{m_{d-45}}{\max}, m_{d-135}' = \frac{m_{d-135}}{\max}, m_{nd}' = \frac{m_{nd}}{\max} \quad (8)$$

ค่าผลลัพธ์ที่ได้จะเป็นสารสนเทศของพื้นผิวจากแต่ Image Block ที่เป็นค่าฮิสโตรแกรมของ 6 พื้นที่ ซึ่งจะสอดคล้องกับ Region ของภาพ ซึ่งอัลกอริทึมจะทำการตรวจสอบว่าค่า max สูงกว่าค่าที่กำหนด (Threshold) หรือไม่ ซึ่งค่า Threshold จะเป็นตัวจำแนกกลุ่ม Image Block ว่าควรเป็น Texture Block หรือ Non Texture Block

2.3 Fuzzy color and Texture Histogram (FCTH)

ในกลุ่มอัลกอริทึมของการจำแนกสีและพื้นผิวยังมีการเสนออัลกอริทึมอื่นมาใช้ในการค้นคืนได้แก่ Fuzzy color and Texture histogram (FCTH) [16] มีกระบวนการทำงานดังภาพที่ 2



ภาพที่ 2: การทำงานของ FCTH

จากภาพแสดงการทำงานของอัลกอริทึมโดยนำตรรกะคลุมเครือ (Fuzzy Logic) 3 ระบบ มาใช้ในการแบ่งสีและพื้นผิว โดยจะทำการแปลงค่าของสีในระบบ HSV และแบ่งสีออกเป็น 10 เฉดสี โดยแบ่งเป็น ช่อง H ใช้พื้นที่เท่ากับ 8 ช่อง S ใช้พื้นที่เท่ากับ 2 และช่อง V ใช้พื้นที่เท่ากับ 3 ในส่วนของพื้นผิวจะใช้ High Frequency จาก Haar Wavelet Transform ใน Fuzzy เพื่อให้อยู่ในรูปของการหาความสัมพันธ์ของ 8 พื้นที่

จากจากศึกษาพบว่า ผลการทดสอบในหลายๆชุดข้อมูล โดยทั่วไปพบว่าอัลกอริทึม FCTH มีค่าความแม่นยำสูงกว่าอัลกอริทึม CEDD แต่ก็มีในบางชุดข้อมูลที่กลับพบว่า CEDD มีผลลัพธ์ที่ดีกว่า จึงมีแนวคิดในการนำ descriptors ทั้งสองตัวนี้มาทำงานร่วมกันโดยยังอยู่บนหลักการทำงานเดิมที่ใช้ข้อมูลจากสีและพื้นผิว แต่จะมีการนำค่าผลลัพธ์ที่ออกจาก 2fuzzy system มาทำการ join กันก่อนนำไปใช้ในการระบุค่าสีและพื้นผิวที่อยู่ในแต่ละ sub-regions

2.4 Joint Composite Descriptor (JCD)

JCD เป็นอัลกอริทึมที่นำ CEDD และ FCTH มาปรับให้มีประสิทธิภาพมากยิ่งขึ้น ซึ่งยังอยู่ในกลุ่มของการใช้สีและพื้นผิวในการค้นคืน [17] อัลกอริทึมดังกล่าวได้มีการรวมความสามารถของอัลกอริทึม CEDD ที่ใช้ความสามารถการค้นคืนโดยใช้พื้นผิวและสีประกอบกัน และอัลกอริทึม FCTH ที่ใช้การค้นคืนด้วยเทคนิค Fuzzy เข้าด้วยกัน อัลกอริทึม JCD ได้ปรับปรุงบางส่วนจากทั้งสองอัลกอริทึมโดยแบ่งพื้นที่ผิวเป็น 7 พื้นที่และพื้นที่ย่อยอีก 24 พื้นที่ แบ่งสีในการคำนวณออกเป็น 24 เฉดสี ใช้ Fuzzy โดยใช้ข้อมูลจาก MPEG-7 ตัวอย่างการคำนวณผลลัพธ์ของอัลกอริทึม JCD ที่มีความสัมพันธ์กับค่าที่ได้จากการคำนวณของ CEDD และ FCTH ดังสมการที่ (9) – (11)

$$JCD(j)_0^i = \frac{FCTH(j)_o^i + FCTH(j)_4^i + CEDD(j)_o^i}{2} \quad (9)$$

$$JCD(j)_1^i = \frac{FCTH(j)_1^i + FCTH(j)_5^i + CEDD(j)_2^i}{2} \quad (10)$$

$$JCD(j)_2^i = CEDD(j)_4^i \quad (11)$$

โดยที่ j คือด้วยภาพใดๆ และ i คือสัญลักษณ์แทน 24 เฉดสี ($i \in [0, 23]$)

3. การทดสอบประสิทธิภาพอัลกอริทึม

3.1 ชุดข้อมูลในการทดลอง

งานวิจัยนี้ต้องการทดสอบประสิทธิภาพของเทคนิคที่ใช้ในการค้นคืนภาพ โดยใช้การพิจารณาจากคุณลักษณะสำคัญของภาพ 2 คุณลักษณะ ได้แก่ สี (Color) และ พื้นผิว (Texture) ในการทดลองจะใช้ชุดข้อมูลภาพ Recognition Benchmark Images [18] จาก Center for Visualization & Virtual Environments ซึ่งเป็นข้อมูลภาพถ่ายจำนวน 10,200 ภาพ ประกอบไปด้วย ภาพจำนวน 2,550 ชุด เป็นชุดภาพที่ใช้มีขนาด 640x480 พิกเซล และมีนามสกุลของภาพเป็น .jpg โดยแต่ละชุดมี 4 ภาพ ประกอบด้วย ภาพคำถาม 1 ภาพ และ ภาพเฉลย 3 ภาพโดยเป็นภาพเกี่ยวกับเครื่องใช้ต่างๆ ภายในบ้าน อาคาร สำนักงาน ดอกไม้ ต้นไม้ เครื่องใช้ไฟฟ้า วิวธรรมชาติ และอื่นๆ ซึ่งแต่ละภาพนั้นจะถูกถ่ายทั้งหมด 4ครั้ง ในมุมที่แตกต่างกัน ดังแสดงในภาพที่ 3

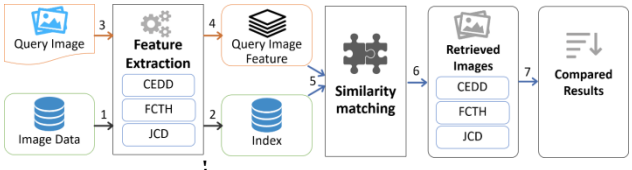


ภาพที่ 3: ตัวอย่างภาพจากชุดข้อมูล Recognition Benchmark Images

ซึ่งในชุดข้อมูลมีภาพที่มีระดับความยากง่ายหลายระดับ เช่นในกลุ่มของภาพที่เป็นปกซีดี (CD-covers) จะคัดแยกได้ง่ายกว่ากลุ่มของภาพดอกไม้ หรือในด้านความคมชัดของภาพมีภาพถ่ายบางส่วนถูกถ่ายในพื้นที่แสงสว่างน้อยจึงได้ภาพที่มีความคมชัดน้อยกว่าภาพที่ถ่ายกลางแจ้ง ซึ่งอาจส่งผลต่ออัลกอริทึมที่ใช้ในการค้นคืนภาพได้ ชุดข้อมูลภาพนี้จึงเป็นชุดข้อมูลที่มีความน่าสนใจและมีความท้าทาย ผู้วิจัยจึงได้เลือกมาใช้ในการทดลองในครั้งนี้

3.2 การออกแบบการทดลอง

กระบวนการทดลอง ได้ทำการทดสอบและเปรียบเทียบประสิทธิภาพการค้นคืนภาพด้วยสีและพื้นผิวด้วย 3 อัลกอริทึม คือ CEDD, FCTH และ JCD โดยกระบวนการทดลอง แสดงดังภาพที่ 4



ภาพที่ 4: กรอบแนวคิดในการทดสอบ

กระบวนการทดลองเริ่มด้วยการนำข้อมูลชุดทอสอบไปผ่านกระบวนการ feature extraction ด้วย อัลกอริทึมคือ 3

CEDD, FCTH และ JCD นำ feature ที่ได้จากแต่ละอัลกอริทึมไปสร้างเป็น index จากนั้นจะนำภาพที่เป็นภาพคำถามจาก กลุ่มไปเข้ากระบวน 2550 feature extraction เพื่อสร้างเป็น query image feature เสร็จแล้วนำไปหาความคล้ายคลึงกันของภาพที่เก็บไว้ใน index จะได้ภาพค้นคืนกลับมา กลุ่มคือ ภาพจากอัลกอริทึม 3 CEDD, FCTH และ JCD นำผลลัพธ์จากทั้ง อัลกอริทึม มาเปรียบเทียบกันว่าผลที่ 3 ได้กลับมาตรงกับคำตอบหรือไม่ โดยใช้สมการ

3.3 เครื่องมือที่ใช้ในการทดสอบประสิทธิภาพ

การทดสอบประสิทธิภาพกับชุดข้อมูลภาพ Recognition Benchmark Images ด้วยโปรแกรม Lucene Image Retrieval (LIRE) ซึ่งเป็นเครื่องมือที่พัฒนาด้วยชุดคำสั่งภาษา JAVA ในการทำ text indexing และ searching ที่นิยมนำไปพัฒนา search engine และ recommendation systems โดยในงานนี้นำมาใช้เป็นเครื่องมือในการทดลองการค้นคืนภาพ โดยใช้ฟังก์ชันในการทำ feature extraction การค้นหา และสร้าง index ด้วยพื้นฐานของลูซีน

3.4 การทดสอบประสิทธิภาพของระบบ

ทดสอบการค้นคืนภาพในแต่ละอัลกอริทึม จะวัดประสิทธิภาพ โดยใช้ค่าความแม่นยำ (Precision) ดังสมาการ (12) – (13)

$$Precision\ (k) = \frac{x}{x+y} \times 100\% \quad (12)$$

$$Average\ Precision\ = \frac{x \sum_{k=1}^{n} (Precision\ (k))}{n} \quad (13)$$

โดย X คือ จำนวนภาพที่ตรงกับความต้องการที่ค้นคืนได้

Y คือ จำนวนภาพที่ไม่ตรงกับความต้องการที่ค้นคืนได้

n คือ จำนวนภาพคำถาม (Query Image) ทั้งหมด

k คือ ค่า precision ของภาพทดสอบแต่ละภาพ

4. ผลการทดลอง

การทดลองใช้ข้อมูลภาพทดสอบ จำนวน 2, 550ชุด จากชุดข้อมูลภาพ Recognition Benchmark Images ซึ่งแต่ละภาพคำถาม นั้นจะประกอบไปด้วย 4 ภาพต่อ 1 ชุด โดยดำเนินการหาค่า similarity และวัดค่า hits score ของแต่ละวิธีการค้นคืนด้วยโปรแกรม LIRE ผลการวัดประสิทธิภาพด้านความแม่นยำจากภาพคำถาม พบว่าอัลกอริทึม JCD มีค่าเฉลี่ยความแม่นยำสูงที่สุดในการค้นคืนภาพด้วยการสกัดคุณลักษณะด้านสีและพื้นผิว โดยผลการทดลองและตัวอย่างผลลัพธ์ แสดงในตารางที่ 1 -2

ตารางที่ 1 : ผลการวัดความแม่นยำในการค้นคืนภาพ

อัลกอริทึม	Average Precision
Color and Edge Directivity Descriptor (CEDD)	0.65
Fuzzy Color and Texture Histogram (FCTH)	0.67
Joint Composite Descriptor (JCD)	0.69

ตารางที่ 2 : ตัวอย่างเปรียบเทียบผลลัพธ์การค้นคืนภาพ

อัลกอริทึม	รูปค้น	ผลลัพธ์
CEDD		
FCTH		
JCD		

5.สรุป

งานวิจัยนี้นำเสนอการเปรียบเทียบอัลกอริทึมที่ใช้ในการค้นคืนภาพ โดยใช้เทคนิคการสกัดคุณลักษณะจากสีและพื้นผิวของภาพ และผลการวัดประสิทธิภาพด้านความแม่นยำของอัลกอริทึมที่ใช้ คือ Color and Edge Directivity Descriptor (CEDD), Fuzzy Color and Texture Histogram (FCTH) และ Joint Composite Descriptor (JCD) โดยนำไปทดสอบกับชุดข้อมูลภาพจำนวน 10,200 ภาพ และใช้ภาพทั้งหมด 2550 ภาพ เป็นภาพคำถาม (Query Image) ผลการทดลองพบว่าอัลกอริทึม Joint Composite Descriptor (JCD) มีความแม่นยำในการค้นคืนสูงที่สุด เนื่องจากเป็นอัลกอริทึมที่รวมเอาค่าผลลัพธ์ที่ได้เกิดจากสอง fuzzy system มาทำการ join กัน ซึ่งทั้ง CEDD และ FCTH มีการแบ่งพื้นที่ของภาพ (texture areas) ไม่เหมือนกันทำให้เกิดการสร้าง indexing ที่แตกต่างกันด้วย เมื่อนำมาปรับค่าเข้าสู่ส่วนกลางจึงทำให้สามารถเพิ่มความแม่นยำในการค้นคืนให้แก่ระบบค้นคืนภาพได้ ส่วนในด้าน

เวลาที่ใช้ในการประมวลผลนั้นมีความแตกต่างกับ CEDD และ FCTH เพียงเล็กน้อยเท่านั้น โดยพบว่า JCD มีค่าความถูกต้องสูงที่สุดอยู่ที่ร้อยละ 69 อย่างไรก็ตามงานวิจัยด้านนี้ยังมีเทคนิคการสกัดคุณลักษณะอื่นที่น่าสนใจ เช่น รูปร่าง (Shape) หรือใช้หลายคุณลักษณะประกอบกัน เพื่อหาอัลกอริทึมที่มีประสิทธิภาพในการค้นคืนที่เหมาะสมกับลักษณะของชุดข้อมูลแบบต่าง ๆ ได้ หรือนำอัลกอริทึมเหล่านั้นไปพัฒนาต่อยอดเพื่อให้ได้ผลการค้นคืนที่มีความแม่นยำมากขึ้น

เอกสารอ้างอิง

[1] ประภาพร กุลลิ้มรัตน์ชัย ,“Image Retrieval Fundamentals" *EAU Heritage Journal 40* ปีที่ 2556 .ค.ธ-.ค.ก 2 ฉบับที่ 7.

[2] กิติ์สุชาต พสุภา, ภาณวี ฉัตรค้ำจุนเจริญและ โชติรส วุฒิเลิศเดชา, “การทำนายอารมณ์ของมนุษย์ที่มีต่อรูปภาพโดยใช้คุณลักษณะของภาพและเครื่องมือติดตามการมองเห็น” *JOURNAL OF INFORMATION SCIENCE AND TECHNOLOGY* ปีที่ 5 2 ฉบับที่,.2558

[3] สุกัญญา สุขะปุณพันธ์กันต์พงษ์ วรรัตน์ปัญญาและ สุพจน์ นิตย์ , ,สุวัฒน์“การสร้างดัชนีภาพและค้นคืนภาพโดยอาศัยพื้นฐานของโครงสร้างต้นไม้*” PSU PIC 2014*.

[4] A.K. Alhassan and A. Alfaki, “Color and Texture Fusion-Based Method for Content-Based Image Retrieva” *2017 International Conference on Communication, Control, Computing and Electronics Engineering (ICCCCEE)*

[5] H.Tamura, S. Mori, T.Yamawaki, “Textural Features Corresponding to Visual Perception*” IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS* 1978.

[6] Eckert and A. P. Bradley, “Perceptual quality metrics applied to still image compression,” *Signal Process*, Vol. 70, no. 3, pp. 177–200, .1998

[7] Z. Wang, A. C. Bovik, H. R. Sheikh, andE. P. Simoncelli, “Image quality assessment:From error visibility to structural similarity,” *IEEE Trans Image Process*, Vol. 13, no. 4, pp. 600–612, 2004.

[8] Z. Wang and E. P. Simoncelli, ‘‘Translation insensitive image similarity in complex Table 2 Pearson’s r for Different Analysis Methods Pappas et al.: Image Analysis: Focus on Texture Similarity*” in Int. Proc. IEEE Conf. Acoust. Speech Signal Process*.,Vol. 101, no. 9, pp. 573-576, 2005.

[9] J. Zujovic, T. N. Pappas, and D. L. Neuhoff, ‘‘Structural texture similarity metrics for image analysis and retrieval,’’ *IEEE Trans.Image Process*., Vol. 22, no. 7, pp. 2545–2558, .2013

[10] R. M. Haralick, K. Shanmugam, and I. Dinstein, “Textural features for image classification,” *IEEE Trans. Syst. Man, Cybern*., Vol. 3, no. 6, pp. 610–621, .1973

[11] T. Mita, T. Kaneko, B. Stenger, and O. Hori, “Discriminative feature co-occurrence selection for object detection,” *IEEE Trans. Pattern Anal. Mach. Intell*., vol. 30, no. 7, pp. 1257–1269, .2008

[12] J. Zujovic, ‘‘Perceptual texture similarity metrics,’’ Ph.D. dissertation, Electr., Eng., Comput. Sci. Dept., Northwestern Univ., Evanston, IL, USA, .2011

[13] J. Chen, T. N. Pappas, A. Mojsilovic, and B. E. Rogowitz, ‘‘Adaptive perceptual color-texture image segmentation,’’ *IEEE Trans. Image Process*., Vol. 14, no. 10, pp. 1524–1536, .2005

[14] A. Mojsilovic´, J. Kovac˘evic´, J. Hu,R. J. Safranek, and S. K. Ganapathy, ‘‘Matching and retrieval based on the vocabulary and grammar of color patterns,’’ *IEEE Trans. Image Process*., Vol. 1, no. 1, pp. 38–54, .2000

[15] S. Chatzichristofis and Y. Boutalis, “Cedd: Color and edge directivity descriptor: A compact descriptor for image indexing and retrieval,” Vol. 5008. Springer, 2008, p. 312.

[16] Chatzichristofis, S.A. and Boutalis, Y.S., ., “Fcth: Fuzzy color and texture histogram-a low level feature for accurate image retrieval*,” in Proceedings of the 9th International Workshop on Image Analysis for Multimedia Interactive Services*, WIAMIS, 2008, pp. 191–196.

[17] A. Savvas Chatzichristofis,N. Papamarkos and Boutalis Yiannis, , “AUTOMATIC IMAGE ANNOTATION AND RETRIEVAL USING THE JOINT COMPOSITE DESCRIPTOR,” *Informatics (PCI), 2010 14th Panhellenic Conference* on, Sept. 2010.

[18] D. Nistér and H. Stewénius. “Scalable recognition with a vocabulary tree, “ In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Volume 2, pp. 2161-2168, 2006.

การจำแนกภาพใบสมุนไพรไทยด้วยคุณลักษณะพื้นผิวเฉพาะส่วน
Classification of Thai Herbal Leaf Images Based on Local Texture Feature

พุทธศักดิ์ สมชัย (Phuttasak Somchai)[1] และสมปอง เวฬุวนาธร (Sompong Valuvanathon)[2]
ภาควิชาคณิตศาสตร์ สถิติและคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยอุบลราชธานี
[1]pozedon_nig@hotmail.com, [2]sompong.v@ubu.ac.th

บทคัดย่อ

บทความวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาขั้นตอนวิธีการในการจำแนกภาพใบสมุนไพรไทยด้วยคุณลักษณะพื้นผิวเฉพาะส่วน โดยใช้วิธีเกรเลเวลโค-ออคเคอร์ เรนซ์ แมทริก (GLCM) ในการสกัดคุณลักษณะ ใช้การจำแนกด้วยวิธีซัพพอร์ตเวกเตอร์แมชชีน (SVM) ร่วมกับเรเดียลเบสิสฟังก์ชันเคอร์เนล (RBF) ด้วยการกำหนดค่าพารามิเตอร์ซีและแกมม่าด้วยมือ แล้วทดสอบประสิทธิภาพของโมเดลด้วยวิธีไขว้กันแบบหลายเท่า (K-fold cross validation) โดยทำการทดลองกับชุดข้อมูลรูปถ่ายภาพใบสมุนไพรไทยที่จัดทำขึ้น จากใบไม้ที่มีลักษณะเส้นลายใบแบบตาข่ายขนนก 10 ชนิดๆ ละ 10 ภาพ โดยทำข้อมูลเป็น 2 ชุดคือ ชุดด้านหน้าและหลังของใบ จำนวน 200 ภาพ และถ่ายโดยใช้แสงส่องที่ใบไม้โดยตรงจำนวน 200 ภาพ ผลการทดลองแสดงให้เห็นว่าอัตราการจำแนกใบสมุนไพรไทยมีความถูกต้องสูงสุดร้อยละ 80.00

คำสำคัญ: สมุนไพรไทย เกรเลเวลโค-ออคเคอร์เรนซ์ แมทริก เรเดียลเบสิสฟังก์ชันเคอร์เนล ซัพพอร์ตเวกเตอร์แมชชีน

Abstract

This paper presents method for Thai herbal leaf images classification based on local texture feature. The Grey Level Co-occurrence Matrix (GLCM) is used to feature extraction. The Support Vector Machine with Radial Basis Function (RBF) kernel are used to classification, and manual adjust parameter C and Gamma. Then, test model performance by K-fold cross validation. In the experiment, we make Thai herbal leaf data set. There are 10 species, take a photo both front and rear side of each leaf 200 images and take a photo by using light through the leaf 200 images. The result of experiment show that the classification accuracy of the images with light through the leaf higher than the images without light through the leaf. The classification accuracy rates up to 80.00%.

Keyword: Thai herbal, GLCM, RBF, SVM

1. บทนำ

สมุนไพรไทย นับเป็นพืชที่มีสรรพคุณในการบรรเทาอาการเจ็บ ป่วย หรือไข้ได้ ในทางแพทย์แผนไทย สมุนไพร [1] ตามพจนานุกรมฉบับราชบัณฑิตยสถาน พ.ศ. 2525 หมายถึง พืชที่ใช้ ทำเป็นเครื่องยา สมุนไพรกำเนิดมาจากธรรมชาติ และมีความหมายต่อชีวิตมนุษย์โดยเฉพาะในทางสุขภาพ อันหมายถึงทั้งการส่งเสริมสุขภาพและการรักษาโรค ความหมายของยาสมุนไพรในพระราชบัญญัติยา พ.ศ. 2510 ได้ระบุว่ายาสมุนไพร หมายความว่า ยาที่ได้จากพฤกษาชาติ สัตว์หรือแร่ธาตุ ซึ่งมิได้ผสมปรุงหรือแปรสภาพ เช่น พืชก็ยังเป็นส่วนของราก ลำต้น ใบ ดอก ผล ฯลฯ ซึ่งมิได้ผ่านขั้นตอนการแปรรูปใดๆ แต่ในทางการค้า สมุนไพรมักจะถูกดัดแปลงในรูปแบบต่างๆ เช่น ถูกหั่นให้เป็นชิ้นเล็กลง บดเป็นผงละเอียด หรืออัดเป็นแท่งแต่ในความรู้สึกของคนทั่วไปเมื่อกล่าวถึงสมุนไพร มักนึกถึงเฉพาะต้นไม้ที่นำมาใช้เป็นยาเท่านั้น สมุนไพรนับว่าเป็นวัตถุดิบและภูมิปัญญาที่มีควรค่าแก่การรักษา ความรู้เกี่ยวกับสมุนไพรของชาวบ้านที่สืบทอดกันมา รวมทั้งความรู้ที่เกิดขึ้นจากการประยุกต์ใช้หรือการสกัดสารที่มีฤทธิ์ทางยา ที่ได้มาจากสมุนไพร ไม่ว่าจะเป็น จากลำต้น ราก ใบ ผล หรือดอกก็ตาม เพื่อนำมาใช้ในการรักษา หรือบรรเทาอาการป่วยของผู้ป่วยโรคต่างๆ ได้ แม้ว่าสมุนไพรจะมีประโยชน์ที่จะช่วยในการเยียวยารักษาได้ แต่ก็เป็นการยากที่จะทำการจำแนก

พืชสมุนไพรได้ถูกต้อง เพราะพืชสมุนไพรบางชนิดนั้นหายากหรือบางชนิดอาจมีความคล้ายคลึงกับใบไม้จากพืชพันธ์ชนิดอื่น ทำให้ไม่สามารถจำแนก และนำสมุนไพรมาใช้ได้ทันทีโดยที่ไม่ได้มีการตรวจสอบจากผู้เชี่ยวชาญ ดังนั้นผู้วิจัยจึงมีแนวคิดที่จะพัฒนาขั้นตอนวิธีการขึ้นมาโดยใช้คุณลักษณะเกรเลเวลโคออคเคอร์เรนซ์แมทริก และใช้ LibSVM ร่วมกับ RBFkernel ในการจำแนกใบพืชสมุนไพร โดยใช้ K-fold Crass validation ในการวัดประสิทธิภาพ โดยให้มีความใกล้เคียงกับการจำแนกโดยผู้เชี่ยวชาญด้านสมุนไพรมากที่สุด

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 ทฤษฎีที่เกี่ยวข้อง

2.1.1 การจำแนกคุณลักษณะของใบพืชสมุนไพร [2]

การจำแนกคุณลักษณะของใบพืชสมุนไพรจำเป็นต้องมีวิธีการในการค้าหา หรือรูปแบบเพื่อให้คอมพิวเตอร์เข้าใจและแสดงผลของคุณลักษณะออกมา โดยการพิจารณาจากรูปร่างลักษณะของใบในการจำแนก แบ่งออกเป็น 4 ลักษณะ คือ

1) ลักษณะของใบ
2) ลักษณะของปลายใบ
3) ลักษณะของขอบใบ
4) ลักษณะของลายใบ ประกอบด้วย 2 ลักษณะ [3] ดังนี้

ก. เส้นใบ ให้สังเกตเส้นกลางใบซึ่งต่อเป็นเนื้อเดียวกับก้านใบ จากเส้นกลางใบแยกออกเป็นเส้นใบ ซึ่งจะแยกแขนงออกไปอีกเป็นเส้นแขนงใบ

ข. การเรียงของเส้นใบ

- เส้นใบแบบขนานพบในพืชใบเลี้ยงเดี่ยวแบ่งเป็นเรียงตามยาวของใบและเส้นใบขนานกันตามขวางของใบ
- เส้นใบแบบตาข่ายพบในพืชใบเลี้ยงคู่ ซึ่งมี 2 แบบ คือแบบตาข่ายขนนก(ดังภาพที่ 1)และตาข่ายแบบรูปมือ(ดังภาพที่ 2)



ภาพที่ 1 เส้นลายใบแบบตาข่ายขนนก



ภาพที่ 2 : เส้นลายใบตาข่ายแบบรูปมือ

2.1.2 Grey Level Co-occurrence Matrix (GLCM) [4]

การวัดพื้นผิวของวัตถุแบบ GLCM ได้กำเนิดขึ้นจากพื้นผิวเทียของภาพที่เกิดขึ้น ถูกนำเสนอโดย Haralick ในราวปี 1970 นักวิเคราะห์ภาพจำนวนมาก พวกเขามีจุดมุ่งหมายที่สร้างแนวคิดหรือหลักการที่จะใช้ในซอฟต์แวร์ เพื่อที่จะเพิ่มประสิทธิภาพในการจำแนก ผลงานในขั้นต้นนั้นจำเป็นต้องมีการบีบอัดความรู้ทางด้านคณิตศาสตร์เพื่อกระบวนการสร้างความรู้ความเข้าใจที่ยากนี้ สำหรับนักเรียนและนักวิเคราะห์ชั้นแนวหน้า การสอนเกี่ยวกับ GLCM นี้ถูกพัฒนาขึ้นเพื่อช่วยให้กลุ่มคนเหล่านี้ให้เข้าใจ และถูกนำไปใช้กันอย่างกว้างขวางตั้งแต่ปี 1999 เป็นต้นมา

GLCM [5] มีการคำนวณทางสถิติหลายอย่าง โดย สถิติเหล่านี้ให้ข้อมูลเกี่ยวกับพื้นผิวของภาพ ซึ่งมีดังนี้

- Contrast : เป็นการวัดค่าพื้นที่รูปแบบต่างๆ ใน Gray-Level Co-occurrence Matrix

$$\sum_{i,j} |i-j|^2 \, p(i,j) \qquad (1)$$

เมื่อ i คือ ค่าของเมทริกซ์ในแนวนอน
j คือ ค่าของเมทริกซ์ในแนวตั้ง
p(i,j) คือชุดของค่าสถิติลำดับที่สอง ของการเปลี่ยนเปลงระดับสีเทาระหว่าง I และ j

- Correlation : เป็นการวัดค่าการรวมกันของความน่าจะเป็นที่จะเกิดขึ้นของคู่พิกเซลที่ระบุไว้

$$\sum_{i,j} \frac{(i-\mu j)(j-\mu j)p(i,j)}{\sigma_i \sigma_j} \qquad (2)$$

เมื่อ i คือ ค่าของเมทริกซ์ในแนวนอน
j คือ ค่าของเมทริกซ์ในแนวตั้ง
p(i,j) คือ ชุดของค่าสถิติลำดับที่สอง ของการเปลี่ยนเปลงระดับสีเทาระหว่าง i และ j
μ คือ ค่าเฉลี่ยของ p
σ คือ ค่าความแปรปรวนของความเข้มของพิกเซลมีความสัมพันธ์ที่มีส่วนทำให้เกิดเป็น GLCM

- Energy : คือค่าผลรวมขององค์ประกอบยกกำลังสองใน GLCM หรือเรียกอีกอย่างหนึ่งว่า Angular Second Moment

$$\sum_{i,j} p(i,j)^2 \quad (3)$$

เมื่อ i คือ ค่าของเมทริกซ์ในแนวนอน
j คือ ค่าของเมทริกซ์ในแนวตั้ง
p(i,j) คือ ชุดของค่าสถิติลำดับที่สอง ของการเปลี่ยนเปลงระดับสีเทาระหว่าง i และ j

- Homogeneity : เป็นการวัดค่าความใกล้เคียงกันของการกระจายองค์ประกอบใน GLCM สำหรับ GLCM เชิงเส้นทแยงมุม

$$\sum_{i,j} \frac{p(i,j)}{1+|i-j|} \quad (4)$$

เมื่อ i คือ ค่าของเมทริกซ์ในแนวนอน
j คือ ค่าของเมทริกซ์ในแนวตั้ง
p(i,j) คือ ชุดของค่าสถิติลำดับที่สอง ของการเปลี่ยนเปลงระดับสีเทาระหว่าง i และ j

2.1.3 Radial Basis Function (RBF) [6]

RBF หรืออาจจะหมายถึง Multivariable Function (อาจจะเรียกว่า Multivariable) เป็นการรวมเอาฟังก์ชั่นกฎพื้นฐานเชิงเส้น รวมเป็นฟังก์ชั่นเดี่ยว (Radial Basis Function) ซึ่งเป็นที่นิยมและสามารถนำมาใช้ในงานที่มีมากกว่าหนึ่งมิติ มักจะถูกนำไปใช้ในฟังก์ชั่นการประมาณการของข้อมูล ซึ่งจะรู้จักเฉพาะในจำนวนจำกัดของจุด (จำนวนที่ยากต่อการประเมินผล) เพื่อการประเมินค่าของฟังก์ชั่นการประมาณการให้สามารถถูกนำไปใช้ได้อย่างมีประสิทธิภาพ

Radial basis functions เป็นวิธีหนึ่งที่มีประสิทธิภาพและถูกนำมาใช้อยู่บ่อยๆ การนำไปใช้งานเพิ่มเติม รวมถึงเขตข้อมูลที่สำคัญของเครือข่ายประสาทและทฤษฎีการเรียนรู้ เริ่มแรกนั้น Radial basis functions ก็คือ Radially Symmetric Functions ที่ถูกปรับโดยคะแนนในพื้นที่ของ Euclidean หลายมิติ และหลังจากนั้นก็ถูกรวมเป็นเส้นตรงกลายมาเป็นพื้นที่ประมาณการของข้อมูลที่พึ่งพาอาศัย ข้อมูลที่พึ่งพาอาศัยกันนี้ทำให้ช่องว่างที่เกิดขึ้นเพื่อให้เหมาะสำหรับการให้การประมาณกลุ่มที่มีขนาดใหญ่ของฟังก์ชั่นที่กำหนด นอกจากนี้ยังเปิดประตูไปสู่การมีอยู่และผลลัพธ์ที่เป็นเอกลักษณ์ สำหรับ การแก้ไขข้อมูลที่กระจัดกระจายโดย Radial basis functions ในการตั้งค่าทั่วไป (โดยเฉพาะอย่างยิ่งในหลายๆ มิติ)

2.1.4 Support Vector Machine [7]

Support Vector Machine หรือ SVM เป็นอัลกอริทึมที่สามารถนำมาช่วยแก้ปัญหาการจำแนกข้อมูล ใช้ในการวิเคราะห์ข้อมูลและจำแนกข้อมูล โดยอาศัยหลักการของการหาสัมประสิทธิ์ของสมการเพื่อสร้างเส้นแบ่งแยกกลุ่มข้อมูลที่ถูกป้อนเข้าสู่กระบวนการสอนให้ระบบเรียนรู้ โดยเน้นไปยังเส้นแบ่งแยกแยะกลุ่มข้อมูลได้ดีที่สุด

แนวความคิดของ Support Vector Machine เกิดจากการที่นำค่าของกลุ่มข้อมูลมาวางลงในฟีเจอร์สเปซ (Feature Space) จากนั้นจึงหาเส้นที่ใช้แบ่งข้อมูลทั้งสองออกจากกันโดยจะสร้างเส้นแบ่ง (Hyperplane) ที่เป็นเส้นตรงขึ้นมา และเพื่อให้ทราบว่าเส้นตรงที่แบ่งสองกลุ่มออกจากกันนั้น เส้นตรงใดเป็นเส้นที่ดีที่สุด

สำหรับรากฐานเดิมของ SVM ถูกนำมาใช้กับข้อมูลที่เป็นเชิงเส้น แต่ในความเป็นจริงแล้วข้อมูลที่นำมาใช้ในระบบการสอนให้ระบบเรียนรู้ส่วนใหญ่มักเป็นข้อมูลแบบไม่เป็นเชิงเส้น ซึ่งสามารถแก้ปัญหาดังกล่าวด้วยการนำ Kernel Function มาใช้การจำแนกข้อมูลบนระนาบหลายมิติ จะใช้ส่วนการเลือกที่มีความเหมาะสมที่สุดเรียกว่า โครงสร้างในการคัดเลือก (feature selection) ซึ่งโครงสร้างในการคัดเลือกมาจากข้อมูลที่สอนให้ระบบเรียนรู้ จำนวนเซตของโครงสร้างที่ใช้อธิบายในกรณีหนึ่ง เรียกว่า เวกเตอร์ (vector) ดังนั้นจุดมุ่งหมายของตัวแบบ SVM คือ แบ่งแยกกลุ่มของเวกเตอร์ในกรณีนี้ด้วยหนึ่งกลุ่มของตัวแปรเป้าหมายที่อยู่ข้างหนึ่งของระนาบ และกรณีของกลุ่มอื่นที่อยู่ทางระนาบต่างกัน ซึ่งเวกเตอร์ที่อยู่บนระนาบหลายมิติทั้งหมดเรียกว่า ซัพพอร์ตเวกเตอร์ (Support Vectors)

2.2 ทฤษฎีที่เกี่ยวข้อง

2.2.1 อนุสรณ์ อุ่นท้าว และสมปอง เวฬุวนาธร [8] ได้นำเสนอการรู้จำท่ามือภาษาไทยโดยใช้การจำลองแบบอาณาจักรมดร่วมกับซัพพอร์ตเวกเตอร์แมชชีน แบบเรเดียลฟังก์ชันร่วมกับค่าความถี่ของทิศทางตามค่าเกรเดียนท์ (Histograms of oriented gradients: HOG) และการวิเคราะห์

องค์ประกอบหลัก (Principal component analysis: PCA) โดยนำมาประยุกต์ใช้ร่วมกับแบบจำลองสร้างส่วนเพิ่มผ่านกระบวนการตรวจสอบประสิทธิภาพการคาดการณ์ของโมเดลด้วยวิธีไขว้กันแบบหลายเท่า โดยทดลองกับชุดข้อมูลรูปถ่ายท่ามือภาษามือไทยที่จัดทำขึ้น จำนวน 1,420 ภาพ ซึ่งประกอบด้วยท่ามือจากผู้แสดงจำนวน 10 คน คนละ 142 ท่า ผลการทดลองทำให้ประสิทธิภาพของการรู้จำท่าภาษามือที่มีความถูกต้องในการรู้จำสูงที่สุด 96.3%

2.2.2 Cecilia Di Ruberto และ Lorenzo Putzu [9] ได้นำเสนอวิธีการรู้จำใบไม้ที่รวดเร็ว ในการจำแนกโดยใช้ SVM และคุณลักษณะเวกเตอร์ที่มีมิติของภาพสูง ผู้วิจัยได้ทำการศึกษาวิธีการในการสกัดคุณลักษณะจากเอกสารงานวิจัยต่างๆที่นำเสนอรูปแบบและวิธีการในการสกัดคุณลักษณะที่รวดเร็วและมีความแม่นยำเมื่อนำมาทำการจำแนก โดยวิธีที่ผู้วิจัยนำเสนอ คือ Shape Description, Invariant Moment, Border Description, Color Description (Centroid Radii) และ Texture Feature แล้วนำคุณลักษณะที่ได้มาทำการประมวลผลร่วมกัน ซึ่งทำให้ได้ผลลัพธ์ที่มีความแม่นยำในการรู้จำสูง เมื่อเปรียบเทียบกับผลงานวิจัยชิ้นอื่นๆ ที่นำมาเปรียบเทียบ ผลที่ได้คือ ความแม่นยำของการรู้จำอยู่ 99% เมื่อทดสอบการรู้จำ Linear SVM และ 98.8% บน Radiant Basis Function กับ Support Vector Machine

2.2.3 Xinhong Zhang และ Fan Zhang [10] ได้นำเสนอการสกัดคุณลักษณะของภาพใบยาสูบในงานประชุมด้านการประมวลผลภาพและสัญญาณในปี 2008 โดยนำเอาเทคนิคต่างๆมาใช้ในการแก้ไขปัญหาของการสกัดคุณลักษณะและวิเคราะห์ใบยาสูบ เช่น คุณลักษณะของสี (Color Features) รูปร่าง (Shape Features) ขนาดและลายพื้นผิว (Texture Features) โดยใช้ GLCM (Grey-Level scale Co-occurrence Matrix) ในการการสกัดคุณลักษณะลายพื้นผิว คำนวณด้วยการใช้ข้อมูลทางสถิติของ GLCM ดังนี้ Texture Energy, Texture Entropy และค่า Contrast ซึ่งผู้วิจัยได้สรุปข้อผลการวิจัยว่าสามารถนำเอาข้อมูลที่ได้จากการสกัดคุณลักษณะไปพัฒนาเครื่องจักรอัตโนมัติให้สามารถจำแนกใบยาสูบได้

2.2.4 รักษ์วริน วรรณศิลป์ [2] ได้นำเสนอการตรวจหาคุณลักษณะเด่นด้วยวิธี SIFT (Scale Invariant Feature Transform) สำหรับใบสมุนไพรพื้นบ้าน การวัดประสิทธิภาพของการตรวจหาคุณลักษณะเด่น โดยการทดลองใช้ ชุดข้อมูลทดสอบเป็นรูปภาพของใบสมุนไพรทั้งหมด 3 ชนิด ได้แก่ ใบมะรุม ใบสะระแหน่ และใบบัวบก การทดลองและวัดประสิทธิภาพด้วยการหาค่า F-measure ผลการทดลองพบว่าชุดทดสอบที่สามารถตรวจหาคุณลักษณะได้ดีที่สุดคือ ชุดทดสอบที่ 3 (ใบบัวบก) มีค่า F-measure เท่ากับ 60.0 % และค่าเฉลี่ยของชุดทดสอบทั้งหมด คือ ค่า F-measure ค่าความแม่นยำ (Precision: P) และ ค่าการระลึก (Recall: R) เท่ากับ 67.0%, 69.0% และ 63.0% ตามลำดับ

3. วิธีการดำเนินการวิจัย

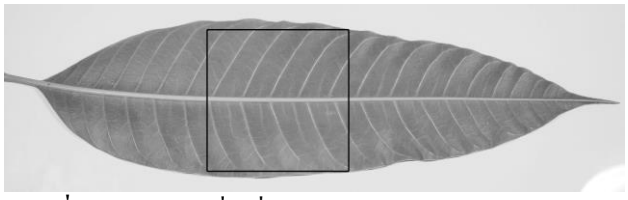
3.1 เลือกใบพืชสมุนไพรที่มีลักษณะจำเพาะในการรู้จำ คือเป็นพืชที่มีลายใบแบบร่างแห จำนวน 10 ชนิดจากพืชสมุนไพรพื้นบ้านที่มีในท้องถิ่นภาคตะวันออกเฉียงเหนือ

ตารางที่ 2: ข้อมูล dataset และจำนวนของข้อมูล

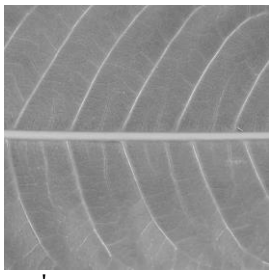
ลำดับ	ชื่อพืชสมุนไพร	ลำดับ	ชื่อพืชสมุนไพร
1	ใบขนุน	6	ใบมะรุม
2	ใบมะกรูด	7	ใบย่านาง
3	ใบโหระพา	8	ใบลำไย
4	ใบกระโดน	9	ใบหูเสือ
5	ใบมะม่วง	10	ใบมะปราง

3.2 นำใบพืชสมุนไพรที่คัดเลือก โดยถ่ายทั้งด้านหน้าและด้านหลังของใบวางบนกระดาษสีขาวด้วยกล้องดิจิตอล รวมทั้งการถ่ายโดยใช้แสงส่องที่ใบไม้โดยตรง

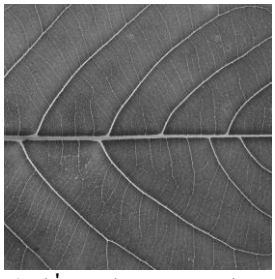
3.3 ทำการตัดเอาบางส่วนของใบไม้โดยการตัดเอาบางส่วนของภาพใบไม้โดยใช้พื้นที่สี่เหลี่ยมจัตุรัสที่กว้างที่สุดแต่ไม่เกินขนาดของความกว้างของใบไม้จากช่วงกลางของใบ แล้วทำการปรับขนาดของภาพให้มีขนาด 1000x1000 พิกเซล จากนั้นทำการแปลงไฟล์ภาพจากภาพสี RBG มาเป็นภาพโทนสีเทา (Grey Scale)



ภาพที่ 2 : ภาพใบไม้ที่เปลี่ยนเป็นโทนสีเทา และทำการตัดภาพบางส่วนของใบไม้



ภาพที่ 3 : ภาพใบไม้ที่ได้ทำการตัดภาพบางส่วนของใบไม้และปรับขนาดของภาพเป็น 1000x1000 พิกเซล



ภาพที่ 4 : ภาพใบไม้ที่ถ่ายด้วยวิธีการใช้แสงส่องผ่านโดยตัดภาพบางส่วนและปรับขนาดของภาพเป็น 1000x1000 พิกเซล

3.4 จากนั้นนำภาพที่ได้ในขั้นตอนที่ 3.3 มาผ่านกระบวนการ สกัดคุณลักษณะของภาพโดยใช้วิธี GLCM Feature (Gray Level Co-occurrence Matrix) ซึ่งจะได้ขนาดของคุณลักษณะเท่ากับ 4 ค่า คือค่า Contrast, Correlation, Energy, Homogeneity

3.5 นำข้อมูลที่ได้จากกระบวนการที่ 3.4 มาสร้างเป็น dataset 2 ชุดดังนี้ 1) ข้อมูลใบไม้ที่ถ่ายทั้งด้านหน้าและด้านหลัง (GLCM_FR) จำนวน 200 ข้อมูล 2) ข้อมูลใบไม้ที่ถ่ายโดยใช้แสงส่องผ่านใบไม้ (GLCM_Light) จำนวน 200 ข้อมูล

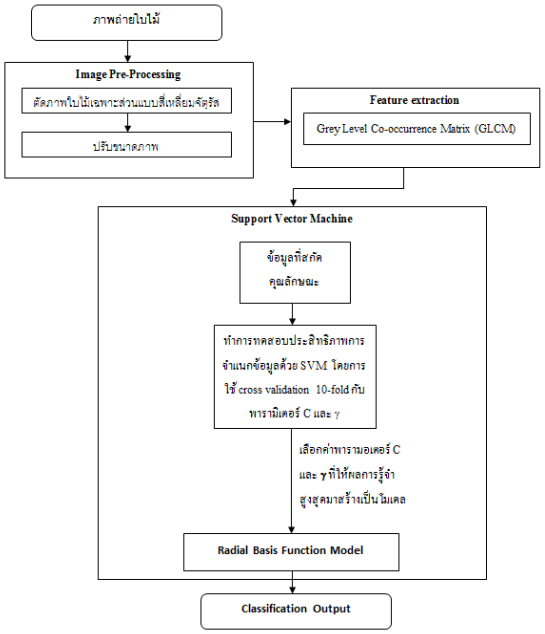
ตารางที่ 2: ข้อมูล dataset และจำนวนของข้อมูล

ชื่อชุดข้อมูล	คำอธิบาย	จำนวนข้อมูล
1. GLCM_FR	ข้อมูลใบไม้ที่ถ่ายทั้งด้านหน้าและด้านหลัง	200
2. GLCM_Light	ข้อมูลใบไม้ที่ถ่ายโดยใช้แสงส่องผ่านใบไม้	200

3.6 นำข้อมูลที่ได้จากกระบวนการที่ 3.5 มาผ่านกระบวนการจำแนกโดยใช้โปรแกรม LibSVM (windows) โดยใช้วิธีในการจำแนก โดยใช้ SVM ร่วมกับ RBFKernel โดยกำหนดค่าพารามิเตอร์ C และ ค่าพารามิเตอร์ gamma ด้วยมือ แล้วทำการวัดประสิทธิภาพของโมเดลด้วย K-fold Cross-Validation

3.7 หาค่า C และ gamma ที่ให้ประสิทธิภาพการจำแนกของโมเดลสูงสุดมาเป็นโมเดลในการทดสอบ

3.8 เปรียบเทียบและอภิปรายผลลัพธ์ที่ได้จากการประมวลผล



ภาพที่ 5 : ภาพขั้นตอนการดำเนินงานของงานวิจัย

4. การทดลอง

จากการทดลองผู้วิจัยได้ทำการทดลองโดยการใช้โปรแกรม LibSVM ร่วมกับ RBFKernel โดยการปรับค่าพารามิเตอร์ C ระหว่าง 1 - 100 และพารามิเตอร์ gamma

ระหว่าง 0.01 - 5 ซึ่งผลการทดลองคือ การจำแนกของชุดข้อมูล GLCM_FR มีประสิทธิภาพการจำแนก สูงสุดที่ร้อยละ 54.50 และผลการจำแนกของชุดข้อมูล GLCM_Light มีประสิทธิภาพการจำแนกสูงสุดที่ร้อยละ 80.00

ตารางที่ 3 : ผลการทำการทดลองด้วย K-fold Cross-Validation

ชื่อชุดข้อมูล	Accuracy Rate
1. GLCM_FR	54.50%
2. GLCM_Light	80.00%

5. สรุปผลการทดลองและอภิปรายผล

จากการทดลองสรุปได้ว่า การจำแนกของชุดข้อมูล GLCM_Light นั้นมีความแม่นยำสูงสุดเพราะ เมื่อถ่ายภาพโดยการส่องแสงผ่านใบไม้จะทำให้สามารถเห็นลายใบชัดเจนยิ่งขึ้น เมื่อเทียบกับการถ่ายภาพใบไม้ที่วางบนแผ่นกระดาษสีขาวเป็นฉากหลัง และเมื่อเปรียบเทียบกับชุดข้อมูล GLCM _FR พบว่ามีความชัดของลายใบน้อยกว่ามากไม่สามารถมองเห็นลายใบได้ชัดเจน

การเก็บรวบใบไม่สมุนไพรเพื่อนำมาสร้างเป็นชุดข้อมูลนั้นจำเป็นใช้การคัดเลือกใบไม้ที่มีความสมบูรณ์และเจริญเติบโตเต็มที่ รวมถึงโรคของใบที่ทำให้เกิดจุดด่างหรือรอยบนใบไม้เองก็มีผลต่อการสกัดคุณลักษณะเช่นกัน รวมไปถึงการใช้กล้องดิจิตอลในการถ่ายภาพ ต้องใช้องค์ประกอบของแสงที่เหมาะสม การตั้งค่าต่างๆของกล้องเพื่อให้ได้ภาพที่คมชัดที่สุดและสิ่งที่จำเป็นอีกวิธีหนึ่งคือการเตรียมข้อมูลของภาพที่เหมาะสมเพื่อที่จะสกัดคุณลักษณะของภาพให้ได้มีประสิทธิภาพที่สุด

เอกสารอ้างอิง

[1] วิกิพีเดีย สารานุกรมเสรี. th.wikipedia.org. ม.ป.ป. https://th.wikipedia.org/wiki/สมุนไพร (10 สิงหาคม 2558 ที่เข้าถึง).

[2] รักษ์วริน วรรณศิลป์ และกฤตชัย บุญสิวนนท์. “การตรวจหาคุณลักษณะเด่นด้วยวิธี SIFT สำหรับสมุนไพรพื้นบ้าน” *The Tenth National Conference on Computing and Information Technology - NCCIT2014*, 2557: 553-558.

[3] ครูนันทนา สำเภา. http://www.nana-bio.com/. ม.ป.ป. http://www.nana-bio.com/elearning/plant%20organ/ leaf.htm (15 ตุลาคม 2558 ที่เข้าถึง).

[4] Mryka Hall-Beyer. www.fp.ucalgary.ca. ม.ป.ป. http://www.fp.ucalgary.ca/mhallbey /the_glcm.htm (10 ตุลาคม 2558 ที่เข้าถึง).

[5] The MathWorks, Inc. www.mathworks.com. ม.ป.ป. https://uk.mathworks.com/ help/images/gray-level-co-occurrence-matrix-glcm.html (10 เมษายน 2559 ที่เข้าถึง).

[6] Martin Buhmann. http://www.scholarpedia.org. 2010.http://www.scholarpedia.org/ article/Radial_basis_function (9 ตุลาคม 2558 ที่เข้าถึง).

[7] tsukasaz. *http://www.mindphp.com/.* 5 มิถุนายน 2555. http://www.mindphp.com/ forums/viewtopic.php? t=12298 (10 ตุลาคม 2558 ที่เข้าถึง).

[8] อนุสรณ์ อุ่นท้าว และสมปอง เวฬุวนาธร. “การรู้จำท่ามือภาษามือไทยโดยใช้การจำลองแบบอาณาจักรมดร่วมกับซัพพอร์ตเวกเตอร์แมชชีน” The Eleventh National Conference on Computing and Information Technology : NCCIT2015, 2015: 137-142.

[9] Cecilia Di Ruberto and Lorenzo Putzu. “A Fast Leaf Recognition Algorithm based on SVM Classifier and High Dimensional Feature Vector.” 2014 International Conference on Computer Vision Theory and Applications (VISAPP), 2557.

[10] Xinhong Zhang and Fan Zhang. “Images Features Extraction of Tobacco Leaves.” 2008 Congress on Image and Signal Processing, 2551: 773-776.

แอพพลิเคชันวิเคราะห์โรคและแนะนำข้อมูลสมุนไพรไทยด้วยเทคนิคอิงกฎเกณฑ์
Rule Based Diagnostic Application for Thai Herbal Medicines

ชยาพร แก่นสาร์ (Chayaporn Kaensar)
ภาควิชาคณิตศาสตร์ สถิติ และคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยอุบลราชธานี
chayaporn.k@ubu.ac.th

บทคัดย่อ

ปัจจุบันแนวคิดเรื่องการดูแลสุขภาพได้รับความนิยมและถูกกำหนดให้เป็นเรื่องสำคัญอันดับต้นๆของผู้คนยุคนี้ เช่น การรับประทานอาหารคลีน การออกกำลังกาย การใช้สารสกัดจากธรรมชาติ หรือแม้แต่การใช้สมุนไพรรักษาโรค โดยเฉพาะในทางการแพทย์ สมุนไพรหลายชนิดได้รับการตรวจสอบและรับรองว่าสามารถรักษาหรือบำบัดอาการป่วยหรือโรคหลายชนิดได้ ซึ่งนับเป็นทางเลือกในการรักษาอาการของผู้ป่วยได้อีกทางหนึ่ง ดังนั้นแอพพลิเคชันวิเคราะห์โรคและแนะนำข้อมูลสมุนไพรด้วยเทคนิคอิงกฎเกณฑ์ จึงได้ถูกพัฒนาขึ้นเพื่อช่วยวิเคราะห์โรคหรืออาการป่วยโดยสร้างกฎความสัมพันธ์จากองค์ความรู้ในด้านข้อมูลโรคจำนวน 35 โรคเพื่อนำมาวิเคราะห์และแนะนำสมุนไพรในฐานข้อมูลจำนวน 80 ชนิดแก่ผู้ใช้ โดยกฎเกณฑ์ต่างๆ ได้ถูกตรวจสอบและประเมินความถูกต้องจากผู้เชี่ยวชาญทางการแพทย์และเภสัชกรจำนวน 3 คน ซึ่งระบบนี้ได้ถูกพัฒนาด้วยภาษาจาวาและทำงานบนระบบปฏิบัติการแอนดรอยด์ นอกจากนี้ระบบยังมีส่วนจัดการข้อมูลสุขภาพของผู้ใช้เพื่อนำมาวิเคราะห์ร่วมกับเทคนิคข้างต้นในการนำเสนอข้อมูลให้ตรงกับข้อมูลสุขภาพของผู้ใช้นั้นๆมากขึ้น โดยในการทดสอบระบบนั้น ได้แบ่งออกเป็น 2 ส่วนคือ ส่วนแรกเป็นการทดสอบประสิทธิภาพกฎการอ้างอิงด้วยวิธี Confusion Matrix ซึ่งได้ผลความถูกต้องเท่ากับ 93.33% และส่วนที่สองเป็นการทดสอบโดยวัดระบบความพึงพอใจจากผู้เชี่ยวชาญทางการแพทย์ จำนวน 3 คน และผู้ป่วยรวมถึงผู้ใช้งานทั่วไปจำนวน 30 คน ซึ่งพบว่ามีระดับความพึงพอใจโดยเฉลี่ยเท่ากับ 3.93 และ 3.99 (SD = 0.23 และ 0.37) ตามลำดับ

คำสำคัญ: เทคนิคอิงกฎเกณฑ์ การวิเคราะห์โรคภัย สมุนไพรไทย การทดสอบเส้นทางแอนดรอย์แอพพลิเคชัน

Abstract

In recent years, we have found that number of people have more concerns with their health and therefore have increased their expectations in consumption, exercise, using natural materials or even herbal medicines. Furthermore, the use of herbal medicines and alternative methods of treatment are widely perceived standard of care and safe. Thus, we proposed the Rule Based Diagnostic Application for Thai Herbal Medicines. This system can be used for Android Platforms. We collected and defined the set of conditions that describe symptoms and their possible causes about 35 diseases and 80 kinds of herbal information for analyzing and recommending Thai herbal medicines for users which are examined and tested by three medical experts including Doctors and Pharmacist. Furthermore, the system also allow users to update their personal health record that help system analyzed and provide information according to User Health Information. For experimental result, we divided experimental into two sets: First, we validated the rule based technique using Confusion Matrix and we found that the correctly classified instances is 93.33%. Second, we conducted experiment measuring by medical experts, patients and user satisfaction based on Likert Type Scale and we found that the medical export continued to record satisfaction mean score were 3.93 (S.D=0.23) and users rated the result score were 3.99 (S.D=0.37) respectively.

Keyword: Rule Based System, Diagnostic Application, Thai Herbal Medicines, Android Application

1. บทนำ

ปัจจุบันแนวคิดเรื่องการดูแลสุขภาพได้รับความนิยมและถูกกำหนดให้เป็นเรื่องสำคัญอันดับต้นๆของผู้คนยุคนี้ ดังจะเห็นได้จากการที่คนส่วนใหญ่หันมาใส่ใจดูแลสุขภาพตนเองในหลายๆ ด้านกันมากขึ้น เริ่มตั้งแต่การปรุงหรือการรับประทานอาหารคลีน การเลือกทานผักผลไม้ปลอดสารพิษ การใช้สิ่งของประจำวันที่ต้องสกัดจากธรรมชาติ หรือแม้แต่การใช้สมุนไพรเป็นทางเลือกในการรักษาโรคและอาการเจ็บป่วย

โดยเฉพาะด้านการแพทย์ พบว่านอกจากการใช้ยาเคมีรักษาอาการทั่วไปแล้ว ยังสามารถใช้สมุนไพรหรือยาที่สกัดจากธรรมชาติมาเป็นทางเลือกหนึ่งในการรักษาหรือบำบัดอาการป่วยได้ ซึ่งยาหรือสมุนไพรดังกล่าวนั้นได้รับการตรวจสอบและรับรองจากกระทรวงสาธารณสุขว่าสามารถรักษาอาการเจ็บป่วยได้ ซึ่งปัจจุบันกระทรวงสาธารณสุขได้รับรองคุณภาพและมาตรฐานของสมุนไพรไว้ทั้งสิ้น 80 ชนิด [1], [2], [3] โดยได้จัดเก็บและรวบรวมข้อมูลสมุนไพรไว้ในหลายรูปแบบเช่น ในรูปแบบเว็บหรือแอพพลิเคชันบนอุปกรณ์พกพา แต่ยังมีข้อจำกัดคือเป็นเพียงแหล่งรวบรวมข้อมูลสำหรับสืบค้น ไม่มีการนำเทคนิคการวิเคราะห์อาการเจ็บป่วยเข้ามาใช้ร่วมด้วย หรือหากมีการใช้เทคนิคช่วยตัดสินใจ ก็มีหลายแอพพลิเคชันที่ยังขาดการนำประวัติผู้ใช้เข้ามาแนะนำข้อมูลร่วมด้วย

ดังนั้นงานวิจัยนี้จึงต้องการนำเสนอระบบวิเคราะห์โรคและแนะนำข้อมูลสมุนไพรไทยด้วยเทคนิคอิงกฎเกณฑ์ ซึ่งได้นำแนวคิดการสร้างกฎความสัมพันธ์ (Rule Based System) มาใช้ในการสร้างองค์ความรู้และเชื่อมโยงกฎในเรื่องโรคและอาการเจ็บป่วย รวมถึงสมุนไพรที่รักษาอาการในโรคต่างๆ ได้ เพื่อนำมาใช้วิเคราะห์และเปรียบเทียบข้อมูลก่อนแนะนำและแสดงผลข้อมูลสมุนไพรที่เป็นประโยชน์แก่ผู้ใช้ต่อไป โดยระบบได้ถูกออกแบบและพัฒนาบนแอนดรอยด์แอพพลิเคชัน ซึ่งประกอบด้วยการทำงาน 3 ส่วนคือ ส่วนแรกคือการวิเคราะห์โรค ซึ่งผู้ใช้สามารถเลือกอาการของโรค จากนั้นระบบจะวิเคราะห์ว่าเป็นโรคอะไรและแนะนำการรักษาโดยสมุนไพรใด ส่วนที่สองคือการค้นหาข้อมูลสมุนไพร ซึ่งส่วนนี้จะประกอบด้วยส่วนการสืบค้นและแสดงรายละเอียดของสมุนไพรนั้น เช่น ชื่อทางวิทยาศาสตร์ ชื่อท้องถิ่น แหล่งที่พบ สรรพคุณ วิธีการใช้ หรือข้อควรระวัง เป็นต้น และส่วนที่สามคือการกำหนดค่าข้อมูลผู้ใช้ ซึ่งจะช่วยให้ผู้ใช้กำหนดข้อมูลสุขภาพส่วนตัวได้ เช่น ข้อมูลน้ำหนัก ส่วนสูง การแพ้ยา หรือโรคประจำตัว ซึ่งจะเป็นข้อมูลที่ระบบจะนำไปใช้วิเคราะห์ร่วมกับกฎความสัมพันธ์ก่อนแนะนำและแสดงผลข้อมูลต่อไป

2. วัตถุประสงค์งานวิจัย

2.1 เพื่อพัฒนาแอพพลิเคชันที่ช่วยวิเคราะห์อาการของโรคต่างๆ และแนะนำสมุนไพรด้วยเทคนิคอิงกฎเกณฑ์แก่ผู้ใช้

2.2 เพื่อศึกษาความพึงพอใจของการใช้งานระบบแนะนำข้อมูลสมุนไพรไทยด้วยเทคนิคอิงกฎเกณฑ์

3. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ในการดำเนินงานวิจัยนั้น ได้มีการศึกษาและนำทฤษฎีที่เกี่ยวข้องมาประยุกต์ใช้ดังนี้คือ

3.1 เทคนิคอิงกฎเกณฑ์ เป็นเทคนิคที่ใช้กฎมาช่วยในการทำนายและตัดสินใจข้อมูล โดยหลักการคือจะนำองค์ความรู้เฉพาะด้านของผู้เชี่ยวชาญมาสร้างเป็นกฎและเงื่อนไขต่างๆ เพื่อให้คอมพิวเตอร์สามารถสร้างผลลัพธ์ได้คล้ายกับสิ่งที่ผู้เชี่ยวชาญได้กำหนดไว้ ซึ่งข้อดีคือเป็นเทคนิคที่ง่าย และเหมาะกับปัญหาที่มีกฎหรือเงื่อนไขที่ชัดเจน แต่ข้อจำกัดคือกฎเกณฑ์ดังกล่าวไม่สามารถเปลี่ยนแปลงเงื่อนไขได้บ่อย ดังนั้นจึงเหมาะกับองค์ความรู้ที่มีเงื่อนไขตายตัวและเป็นข้อเท็จจริง ซึ่งจะใช้ในการจำลองความรู้หรือประสบการณ์ของผู้เชี่ยวชาญโดยการให้เหตุผลหรือการตัดสินใจต่อสถานการณ์ต่างๆ ของบุคคลนั้น ซึ่งสามารถเขียนได้ในรูปแบบกฎเงื่อนไข คือ เงื่อนไข IF (ถ้า) <เงื่อนไข> THEN (แล้ว) <ผลที่ตามมา> [4], [5]

3.2 งานวิจัยที่เกี่ยวข้อง ในการวิเคราะห์อาการเพื่อระบุโรคใดๆ นั้น นับว่าเป็นปัญหาที่มีงานวิจัยศึกษาและรองรับอยู่เป็นจำนวนมาก โดยกลไกหนึ่งที่นิยมนำไปประยุกต์ใช้คือระบบผู้เชี่ยวชาญ (Expert System) ที่ช่วยสร้างกฎและเงื่อนไขความสัมพันธ์ขององค์ความรู้ ซึ่งถูกกำหนดโดยผู้เชี่ยวชาญในสาขานั้นๆ [5] ดังจะเห็นได้จากในปี 1970 ได้มีการนำเอาระบบชาญฉลาด (Intelligent System) เข้ามาช่วยประมวลผลด้านการแพทย์ นั่นคือมีการพัฒนาระบบ MYCIN โดยมหาวิทยาลัยสแตนฟอร์ด สหรัฐอเมริกา ภายใต้ชื่อโครงการ Heuristic Programming Project [6] ซึ่งระบบดังกล่าวถูกพัฒนาขึ้นเพื่อช่วยวิเคราะห์ยาปฏิชีวนะ (Antibiotic) ที่ต่อต้าน

จุลินทรีย์ (Microbes) ในร่างกาย โดยในยุคนั้นระบบ MYCIN ได้ถูกใช้และมีการอ้างอิงไปยังงานวิจัยอื่นเพื่อต่อยอดผลผลิตเป็นจำนวนมาก โดยได้สร้างกฎความสัมพันธ์จากความรู้ทางการแพทย์ไว้มากกว่า 500 เงื่อนไข แต่ข้อจำกัดคือความซับซ้อนของเงื่อนไขและการแก้ไขเปลี่ยนแปลงข้อมูลนั้นทำได้ยาก ดังนั้นต่อมาจึงมีการนำเสนอผลงานหรือแนวคิดเพื่อแก้ปัญหา รวมทั้งประยุกต์ใช้ระบบความเชี่ยวชาญอื่นๆ เพื่อนำมาปรับใช้ในงานทางการแพทย์ ยกตัวอย่างเช่น

ระบบ NES [7] เป็นการใช้หลักการของระบบผู้เชี่ยวชาญเข้ามาช่วยสร้างกฎและความสัมพันธ์ของอาการต่างๆ เพื่อให้ระบบช่วยวิเคราะห์หรือวินิจฉัยโรคที่เกี่ยวข้องกับโรคทางสมอง (Neurological Disorder) โดยใช้เทคนิคอิงกฎเกณฑ์และต้นไม้ตัดสินใจ (Decision Tree) เข้ามาช่วยสร้างโมเดล โดยพัฒนาบนเว็บด้วยภาษา ASP.NET

L.V. Critien et al. [8] เป็นการใช้เทคนิคอิงกฎเกณฑ์เข้ามาช่วยวิเคราะห์และออกแบบอาการต่างๆ ที่เกี่ยวข้องกับโรคหัวใจล้มเหลว เพื่อระบุแนวทางและวินิจฉัยความเสี่ยงของการเกิดโรคในเบื้องต้น โดยได้พัฒนากฎและความสัมพันธ์ต่างๆจากผู้เชี่ยวชาญทางการแพทย์ ซึ่งนำเสนอข้อมูลในรูปแบบเว็บ

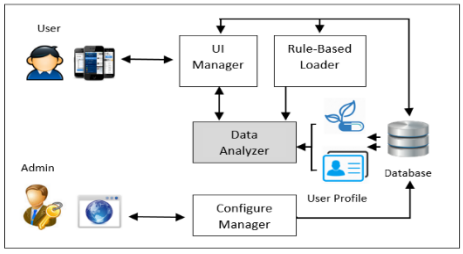
S.Phutthachan et al. [9] นำเทคนิคออนโทโลจี (Ontology) เข้ามาช่วยสร้างเครือข่ายเชิงความหมายสำหรับยา ซึ่งมีการนำชื่อยา อาการป่วย สรรพคุณ และขนาดที่ใช้มาสร้างความสัมพันธ์ระหว่างคลาสที่เกิดขึ้น ซึ่งผลลัพธ์จะได้ระบบสืบค้นสารสนเทศด้านยาที่ผู้ใช้สามารถค้นหา และให้รายละเอียดข้อมูลส่วนอื่นๆ ที่เกี่ยวข้องได้ แต่อย่างไรก็ตามการค้นหาข้อมูลยังมีข้อจำกัดคือความถูกต้องของข้อมูลจะขึ้นอยู่กับเครือข่ายออนโทโลจีที่ผู้พัฒนาได้สร้างความสัมพันธ์ขึ้น และเนื่องจากเป็นระบบออฟไลน์ ทำให้การแก้ไขข้อมูลใดๆ อาจต้องกำหนดรูปแบบความสัมพันธ์ใหม่

S.Bhat and K.Aishwarya [10] นำเสนอระบบแนะนำยาใหม่ให้กับผู้ใช้ซึ่งมีสรรพคุณใกล้เคียงหรือคล้ายกับยาที่เคยใช้ในปัจจุบันเพื่อเป็นทางเลือกอื่นให้ผู้ใช้ กรณีผู้ใช้นั้นอาจมีอาการแพ้ยาหรือข้อจำกัดอื่น เช่น ราคา กลิ่น รสชาติ ยี่ห้อ หรือแหล่งจัดจำหน่าย เป็นต้น ซึ่งได้ใช้เทคนิคการแนะนำข้อมูลแบบ Item-based Recommendation เข้ามาประยุกต์เพื่อตรวจสอบหาความคล้ายกันของยา โดยมีการพิจารณาข้อมูลของผู้ใช้นั้นๆ ร่วมด้วย เช่น ข้อมูลทั่วไป ที่อยู่ ยี่ห้อ หรือมาตรฐานยาที่สนใจ ยาที่ใช้เป็นประจำ เพื่อนำข้อมูลทั้งสองส่วนมาจำแนกและแนะนำข้อมูลยาใหม่แก่ผู้ใช้ต่อไป

อย่างไรก็ดี พบว่าในการสร้างกฎหรือองค์ความรู้ที่มีอาการของโรคที่ชัดเจน นิยมใช้การอิงกฎเกณฑ์เข้ามาช่วยสร้างระบบการตัดสินใจและแนะนำข้อมูลยา นอกจากนี้ยังพบว่ามีงานวิจัยบางส่วนที่ไม่ได้นำข้อมูลผู้ใช้ไปเป็นปัจจัยหนึ่งในการวิเคราะห์เพื่อคัดกรองหรือแสดงข้อมูลที่เป็นประโยชน์ร่วมด้วย ดังนั้นเพื่อให้การแนะนำข้อมูลสมุนไพรไทยตามอาการหรือโรคที่เกี่ยวข้องมีความสอดคล้องกับข้อมูลสุขภาพของผู้ใช้นั้นๆมากขึ้น งานวิจัยนี้จึงนำข้อมูลผู้ใช้มาเป็นปัจจัยหนึ่งในการพิจารณาร่วมกับการสร้างกฎความสัมพันธ์เพื่อสร้างระบบแนะนำข้อมูลสมุนไพรไทยที่เป็นประโยชน์และสอดคล้องกับอาการเจ็บป่วยหรือข้อมูลสุขภาพของผู้ใช้นั้นๆ มากขึ้น

4. โครงสร้างระบบ

ในการทำงานของระบบนั้นจะเริ่มจากผู้ใช้ล็อกอินเข้าสู่ระบบและระบบจะแสดงหน้าต่างสอบถามข้อมูลสุขภาพพื้นฐานจากผู้ใช้ เช่น อายุ น้ำหนัก ส่วนสูง โรคประจำตัวหรือประวัติการแพ้ยาเป็นต้น จากนั้นจะบันทึกข้อมูลผู้ใช้ ซึ่งส่วนนี้ผู้ใช้สามารถแก้ไขหรือเปลี่ยนแปลงได้ในภายหลัง โดยเมื่อเข้าสู่หน้าหลักแล้ว ส่วนจัดการแสดงผล (User Interface Manager) จะแสดงส่วนการทำงานในฟังก์ชันต่างๆ ในหน้าหลัก เช่น เมื่อผู้ใช้เลือกใช้งานฟังก์ชันการวิเคราะห์อาการเพื่อตรวจสอบโรค ระบบจะโหลดกฎความสัมพันธ์จากส่วน Rule Based Loader ที่ได้กำหนดไว้และส่งไปวิเคราะห์พร้อมตรวจสอบกับข้อมูลสมุนไพรและข้อมูลผู้ใช้ในส่วนวิเคราะห์ข้อมูล (Data Analyzer) เพื่อประมวลผลและคัดกรองข้อมูลตามกฎและเงื่อนไขต่างๆ ก่อนส่งผลข้อมูลสมุนไพรกลับไปยังส่วน UI Manager เพื่อแสดงผลต่อไป นอกจากนี้ระบบยังมีส่วนจัดการข้อมูลสมุนไพรใน Configure Manager ที่ให้ผู้ดูแลระบบ (Admin) สามารถเพิ่มหรือแก้ไขข้อมูลสมุนไพร หรือข้อมูลอาการเจ็บป่วยให้มีความทันสมัยมากขึ้นผ่านเว็บจัดการของระบบ ดังภาพที่ 1



ภาพที่ 1: โครงสร้างระบบ

5. การดำเนินการวิจัย

ในการดำเนินการวิจัย ประกอบด้วยขั้นตอนต่างๆ ดังนี้

5.1 การรวบรวมข้อมูลและการสร้างกฎโดยผู้เชี่ยวชาญ

ข้อมูลที่นำมาวิเคราะห์จะเก็บรวบรวมมาจาก 2 ส่วน โดยส่วนแรกคือส่วนอาการและโรคที่เจ็บป่วย จำนวน 35 โรคที่พบในชีวิตประจำวัน (3 กลุ่มอาการหลัก) [11] ดังแสดงในตาราง 1. โดยข้อมูลแต่ละกลุ่มประกอบด้วยรายละเอียดและอาการของโรคที่พบบ่อย นอกจากนี้ยังมีการรวบรวมข้อมูลส่วนอื่น เช่น ความรุนแรงของอาการ (Scale) และเงื่อนไขหรือปัจจัยเสี่ยงที่ ทำให้เกิดโรคด้วย (Condition) เป็นต้น

ตารางที่ 1: จำนวนข้อมูลโรคในกลุ่มโรคต่างๆ ภายในระบบ

กลุ่มโรค	จำนวน	ตัวอย่างโรคที่พบบ่อย
ระบบทางเดินอาหาร	15	โรคกระเพาะอาหาร โรคถุงน้ำดีอักเสบ โรคลำไส้อักเสบ โรคไส้ติ่งอักเสบ
ผิวหนัง	9	โรคอีสุกอีใส โรคเริม กลาก เกลื้อน โรคฝี หรือโรคผื่นจากการแพ้
ระบบทางเดินหายใจ	11	ไข้หวัด ไข้หวัดใหญ่ วัณโรค ต่อม ทอนซิลอักเสบ หลอดลมอักเสบ

จากนั้นนำข้อมูลหรือบริบทที่เกี่ยวข้องจากการรวบรวมข้อมูลจากเอกสารทางการแพทย์ เช่น โรค อาการ หรือปัจจัยเสี่ยงที่ทำให้เกิดโรค [11], [12], [13] มาสร้างกฎความสัมพันธ์ ซึ่งถูกตรวจสอบโดยผู้เชี่ยวชาญทางการแพทย์ ดังตารางที่ 2

ตารางที่ 2: ตัวอย่างข้อมูลกลุ่มโรคระบบทางเดินอาหาร

กลุ่มโรค	เงื่อนไขการเลือกอาการ	ความรุนแรง	ผลการวิเคราะห์โรค
ระบบทางเดินอาหาร	ปวดจุกแน่นที่ท้องส่วนบน	เรื้อรังหรือไม่ระบุ	โรคกระเพาะอาหาร
	ปวดใต้ชายโครงซ้าย		
	ปวด จุก ท้องก่อนหรือหลังทานอาหาร		
	แสบแน่นหน้าอก		
	ปวดท้อง บริเวณชายโครงด้านขวา	ไม่ระบุหรือเฉียบพลัน	ถุงน้ำดีอักเสบ
	คลื่นไส้ อาเจียน		
	มีไข้สูง		
	ตัวซีด		
	ปวดท้องบริเวณลิ้นปี่จนถึงด้านหลัง	ไม่ระบุ	ตับอ่อนอักเสบ
	คลื่นไส้ อาเจียน		
	ปวดแน่นบริเวณหน้าท้อง	ฉับพลัน	ช่องท้องอักเสบ
	มีไข้สูง		
	หัวใจเต้นเร็ว หายใจเร็ว		
	ปวดแน่นบริเวณหน้าท้อง	ไม่ระบุ	อาการท้องอืด ท้องเฟ้อ แน่นจุกเสียด

จากตารางข้างต้น สามารถอธิบายตัวอย่างความสัมพันธ์ได้คือ ถ้าเลือกกลุ่มโรคทางเดินอาหาร และมีอาการปวดแน่นบริเวณท้องส่วนบน มีภาวะอาเจียน รวมถึงมีไข้ และตัวซีดร่วมด้วย แสดงว่ามีโอกาสเป็นโรคถุงน้ำดีอักเสบ ในขณะที่หากมีเพียงอาการปวดท้องส่วนบนหรือบริเวณใต้ชายโครงซ้าย แต่ไม่มีภาวะอาเจียน หรือมีไข้ หรือปวดแบบเฉียบพลัน อาจมีแนวโน้มเป็นโรคกระเพาะอาหาร เป็นต้น

ในขณะที่ส่วนที่สองเป็นการแทนกฎของอาการและสมุนไพรที่ใช้รักษา ซึ่งประกอบด้วยข้อมูลอาการหรือโรค ชื่อสมุนไพร (Herb Name) คุณลักษณะของยา (Herbal Nature) รสชาด (Herbal Favor) สรรพคุณ (Efficacy) เป็นต้น [14] ดังแสดงตัวอย่างกฎความสัมพันธ์ดังตารางที่ 3

ตารางที่ 3: ตัวอย่างข้อมูลยาสมุนไพร

ยาจากสมุนไพร	โรคและอาการที่ใช้ยาสมุนไพรรักษา
ยาขมิ้นชัน	อาการท้องอืด ท้องเฟ้อ แน่นจุกเสียด
ยาขิง	อาการท้องอืด ท้องเฟ้อ แน่นจุกเสียด อาการคลื่นไส้ อาเจียน
ยาชุมเห็ดเทศ	อาการท้องผูก
ยามะขามแขก	อาการท้องผูก
ยากล้วย	โรคกระเพาะอาหาร, อาการท้องเสีย
ยาฟ้าทะลายโจร	อาการท้องเสีย อาการไอ ระคายคอจากเสมหะ อาการไข้

โดยนอกจากสมุนไพรจะมีประโยชน์และสรรพคุณในการรักษาโรคได้แล้ว ในทางการแพทย์ยังมีข้อควรระวังในการใช้ยาดังกล่าวด้วย เช่น หากผู้ใช้กำลังตั้งครรภ์หรือให้นมบุตร ควรหลีกเลี่ยงการใช้ฟ้าทะลายโจร เพราะอาจเกิดอาการช็อกขั้นรุนแรงได้ [1], [2] ดังสรุปกฎได้ในตารางที่ 4.

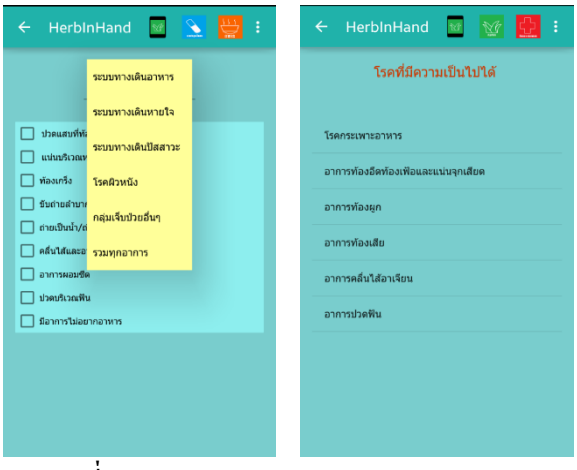
ตารางที่ 4: ตัวอย่างสมุนไพรในกลุ่มที่ควรระวัง

ชื่อโรคและอาการ	สมุนไพรที่ควรหลีกเลี่ยง
โรคนิ่วในถุงน้ำดี	ขมิ้น ขิง
กำลังตั้งครรภ์และให้นมบุตร	ขมิ้น ชุมเห็ดเทศ มะขามแขก ฟ้าทะลายโจร
มีภาวะทางเดินอาหารอุดตัน	ชุมเห็ดเทศ มะขามแขก
มีอาการท้องผูก	กล้วย
ผู้มีอาการแพ้ฟ้าทะลายโจร	ฟ้าทะลายโจร
มีประวัติเป็นโรคหัวใจรูห์มาติค	ฟ้าทะลายโจร
มีประวัติเป็นโรคมะเร็งผิวหนัง	บัวบก
แพ้ว่านหางจระเข้	ว่านหางจระเข้
ไดมีการทำงานบกพร่อง	กระเจี๊ยบแดง
มีประวัติเป็นโรคไตอักเสบ	ฟ้าทะลายโจร

5.2 การพัฒนาระบบ

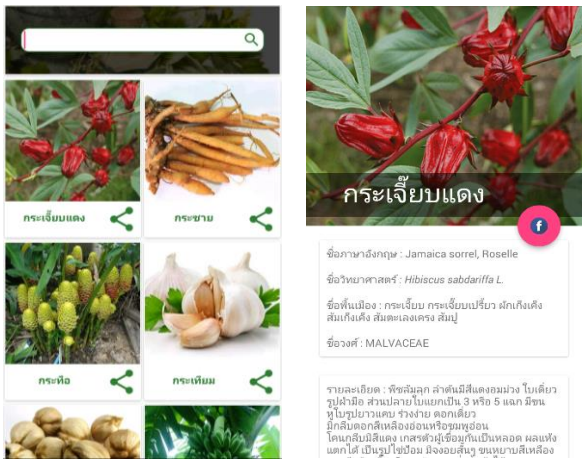
ในขั้นตอนนี้จะนำกฎและเงื่อนไขที่ได้จากขั้นตอนแรกมาประมวลผลและพัฒนาในรูปแบบแอนดรอยด์แอพพลิเคชัน ซึ่งประกอบด้วยฟังก์ชันการทำงาน 3 ส่วนคือ

5.2.1 ส่วนการวิเคราะห์โรค เป็นส่วนการทำงานหลักที่นำเทคนิคอิงกฎเกณฑ์มาใช้งาน โดยประกอบด้วยการเลือกเมนูวิเคราะห์โรคและกลุ่มอาการเพื่อจำแนกข้อมูล จากนั้นระบบจะประมวลผลและใช้เทคนิคอิงกฎเกณฑ์มาช่วยตัดสินใจ ก่อนแสดงรายชื่อโรคที่เกี่ยวข้องต่อไป ดังภาพที่ 2



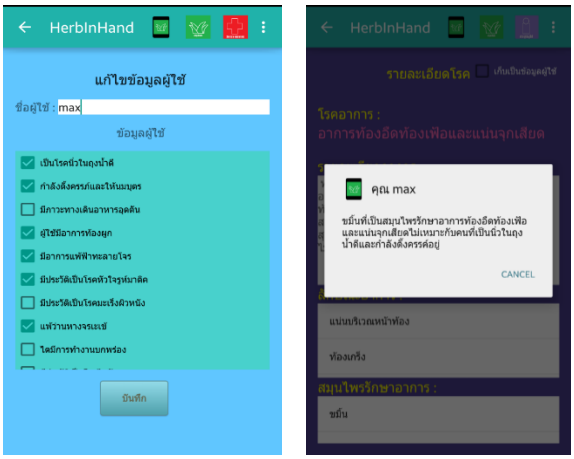
ภาพที่ 2: ตัวอย่างหน้าจอส่วนการวิเคราะห์อาการและโรค

5.2.2 ส่วนค้นหาข้อมูลโรคและสมุนไพร โดยผู้ใช้สามารถค้นหาข้อมูลสมุนไพร รวมทั้งรายละเอียดต่างๆ เช่น ชื่อทางวิทยาศาสตร์ ส่วนที่นำไปสกัด ขนาดและวิธีใช้ รวมถึงสรรพคุณได้ ดังแสดงในภาพที่ 3



ภาพที่ 3: ตัวอย่างหน้าจอส่วนการค้นหาและรายละเอียดข้อมูลสมุนไพร

5.2.3 ส่วนจัดการข้อมูลผู้ใช้ โดยระบบจะมีการนำข้อมูลสุขภาพและเงื่อนไขต่างๆ ของผู้ใช้แต่ละรายไปวิเคราะห์ร่วมด้วย กล่าวคือหากผู้ใช้มีโรคประจำตัว หรืออาการแพ้ยาใดๆ ระบบจะแนะนำหรือนำเสนอข้อมูลสมุนไพรที่อาจมีสรรพคุณช่วยบำบัดโรคนั้นในขณะที่มีการแสดงผลข้อมูล และในทางตรงกันข้ามระบบจะหลีกเลี่ยงการแนะนำสมุนไพรที่อาจเสี่ยงต่อโรคประจำตัวหรือการแพ้ยาของผู้ใช้นั้นๆ โดยจะแสดงผลผ่านการแจ้งเตือนหน้าจอ เช่น หากผู้ใช้มีประวัติเป็นนิ่วในถุงน้ำดี หรืออยู่ในภาวะตั้งครรภ์ ต้องหลีกเลี่ยงการใช้ยาขมิ้นชันเนื่องจากอาจส่งผลต่อการสะสมของนิ่ว หรือการบีบตัวของมดลูกได้ ซึ่งกรณีนี้ระบบจะแจ้งเตือนข้อมูลในการใช้สมุนไพรดังกล่าวแก่ผู้ใช้ ดังภาพที่ 4



ภาพที่ 4: ตัวอย่างหน้าจอส่วนการค้นหาและรายละเอียดข้อมูลสมุนไพร

6. การทดสอบระบบและผลการวิจัย

ผู้วิจัยได้แบ่งการทดสอบระบบออกเป็น 2 ส่วนคือ

6.1 การทดสอบประสิทธิภาพของกฎการอ้างอิง โดยผู้วิจัยได้เก็บข้อมูลเพิ่มเติมจากประวัติการป่วย พร้อมโรคของผู้ป่วยจำนวน 30 กรณี และนำมาแทนค่าข้อมูล ก่อนทดสอบด้วยวิธี Confusion Matrix [5] ซึ่งพบว่ากฎเกณฑ์ที่ใช้ในระบบ (Predicted Value) ให้ค่าที่ตรงกับข้อมูลและอาการของผู้ป่วยจริง (Actual Value) จำนวน 28 กรณี หรือคิดเป็นร้อยละ 93.33

6.2 การทดสอบระดับความพึงพอใจของผู้ใช้ โดยส่วนแรกเป็นการประเมินความพึงพอใจจากผู้เชี่ยวชาญทางการแพทย์ 1 คนและเภสัชกรด้านสมุนไพร 2 คน ซึ่งพบว่ามีระดับความพึงพอใจโดยเฉลี่ยเท่ากับ 3.93 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.23 ในขณะที่การประเมินผลในส่วนที่สองเป็นการวัดระดับความพึงพอใจจากผู้ใช้งานจำนวน 30 คน ซึ่งประกอบด้วยผู้ป่วยที่มาใช้บริการในศูนย์สุขภาพจำนวน 10 คน และผู้ใช้งานทั่วไปจำนวน 20 คน ซึ่งจากตารางที่ 5 เป็นการสรุปผลการวัดในด้านต่างๆ เช่น ระบบสามารถวิเคราะห์อาการป่วยได้

สอดคล้องกับโรคที่เป็น รวมทั้งมีการแนะนำสมุนไพรที่มีประโยชน์หรือมีสรรพคุณต่ออาการป่วยนั้น นอกจากนี้ยังวัดผลในด้านของความสะดวกในการใช้งานและการเข้าถึงข้อมูลของผู้ใช้ด้วย ซึ่งพบว่ามีค่าเฉลี่ยเท่ากับ 3.99 และมีส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.37 ตามลำดับ

ตารางที่ 5: ผลประเมินความพึงพอใจของผู้ใช้ที่มีต่อระบบงาน

ด้านการประเมิน	$\bar{x}$	S.D	ระดับ
ข้อมูลที่แนะนำมีความสอดคล้องต่อผู้ใช้งาน	3.90	0.92	ปานกลาง
ข้อมูลที่แนะนำเป็นประโยชน์ต่อผู้ใช้งาน	4.07	0.64	มาก
มีรูปแบบการนำเสนอข้อมูลที่ง่ายและชัดเจน	4.07	0.78	มาก
การนำเสนอข้อมูลมีความครบถ้วน	4.00	0.79	มาก
ระบบน่าสนใจและน่าใช้งาน	3.93	0.69	ปานกลาง
ระดับความพึงพอใจโดยเฉลี่ย	3.99		
ส่วนเบี่ยงเบนมาตรฐาน	0.37		

7. สรุปผลและการพัฒนาในอนาคต

แอพพลิเคชันวิเคราะห์โรคและแนะนำข้อมูลสมุนไพรไทยด้วยเทคนิคอิงกฎเกณฑ์ เป็นระบบวิเคราะห์โรคซึ่งใช้เทคนิคอิงกฎเกณฑ์เข้ามาช่วยในการสร้างกฎและเงื่อนไขแทนองค์ความรู้จากผู้เชี่ยวชาญ พร้อมวิเคราะห์โรคจากอาการเจ็บป่วย จำนวน 35 โรค ซึ่งจะนำมาวิเคราะห์เพื่อแนะนำยาสมุนไพรในฐานข้อมูลจำนวน 80 ชนิดแก่ผู้ใช้ โดยกฎเกณฑ์ต่างๆ จะถูกตรวจสอบและประเมินความถูกต้องจากผู้เชี่ยวชาญทางการแพทย์ ซึ่งระบบนี้ถูกพัฒนาด้วยภาษาจาวาและทำงานบนระบบปฏิบัติการแอนดรอยด์ นอกจากนี้ระบบยังมีส่วนจัดการข้อมูลผู้ใช้ เพื่อนำข้อมูลนั้นมาวิเคราะห์ร่วมกับเทคนิคข้างต้นก่อนนำมาคัดกรองและนำเสนอข้อมูลให้ตรงกับข้อมูลสุขภาพของผู้ใช้นั้นๆ มากขึ้น

ในส่วนการทดสอบระบบนั้น ได้แบ่งออกเป็น 2 ส่วนคือ ส่วนแรกเป็นการทดสอบประสิทธิภาพของกฎการอ้างอิงด้วยวิธี Confusion Matrix ซึ่งพบว่าได้ค่าความถูกต้องเท่ากับร้อยละ 93.33 ในขณะที่ส่วนที่สองได้ประเมินระดับความพึงพอใจโดยผู้เชี่ยวชาญทางการแพทย์ จำนวน 3 คน และผู้ป่วยรวมถึงผู้ใช้งานทั่วไปจำนวน 30 คน ซึ่งมีค่าเฉลี่ยเท่ากับ 3.93 และ 3.99 (SD = 0.23 และ 0.37) ตามลำดับ

โดยในการพัฒนาระบบงานในอนาคตนั้น จะมีการเพิ่มขอบเขตของอาการและโรคให้ครอบคลุมตามที่กำหนดไว้ในทางการแพทย์มากขึ้น นอกจากนี้จะมีการเพิ่มประสิทธิภาพในส่วนการวิเคราะห์โรค โดยเพิ่มเทคนิคการแนะนำข้อมูลประเภทอื่นเข้ามาทำงานร่วมกันในรูปแบบ Hybrid Recommendation เพื่อเพิ่มความถูกต้องของการวิเคราะห์โรคและการแนะนำยาสมุนไพรให้มากขึ้น

เอกสารอ้างอิง

[1] อาสาฬา เชาวน์เจริญ และคณะ, คู่มือเภสัชกรรมสมุนไพร เล่มที่1: สมุนไพรไม่ใช่ยาขม, กรุ่งเทพฯ: สภาเภสัชกรรม, 2558

[2] ชยันต์ พิเชียรสุนทร และวิเชียร จีรวงส์, คู่มือเภสัชกรรมแผนไทย เล่ม 2: เครื่องยาพฤกษวัตถุ, กรุ่งเทพฯ: ซีเอ็ดยูเคชั่น จำกัด, 2556.

[3] ศุภชัย ติยวรนันท์ และคณะ, คู่มือเภสัชกรรมแผนไทย เล่ม 6: เภสัชกรรม, กรุ่งเทพฯ: ซีเอ็ดยูเคชั่น จำกัด, 2556.

[4] Suh and Sang C, Practical Applications of Data Mining, Jones & Bartlett Learning, 2012

[5] H. Jiawei, Data Mining: Concepts And Techniques, Amsterdam : Elsevier/Morgan Kaufmann, 2012

[6] T. Ellis et al., The Stanford Heuristic Programming Project: Goals & Activities, AI Magazine Vol1. No.1, pp.25-30, 1980.

[7] Al-Hajji. A., Rule-Based expert system for diagnosis and symptom of Neurological Disorders. The 10th International Conference of Intelligent Computing Methodologies, pp.135-143, 2012

[8] L.V. Critien et al., A Rule Based System for Diagnosing and Treating Chronic Heart Failure, Workshop on Computational Intelligence (UKCI), London, 2014

[9] S.Phutthachan et al., A Framework for mapping thai drugs using a pharmaceutical ontology entension of SNOMED CT, The 11th International Joint Conference on Computer Science and Software Engineering, pp. 313-318, 2014

[10] S.Bhat and K.Aishwarya, Item-Based Hybrid Recommender System For Newly Marketed Pharmaceutical Drugs, International Conference on Advances in Computing, Communications and Informatics (ICACCI), pp.2107-2111, 2013.

[11] ปิติกานต์ บูรณาภาพ, 21 โรคเฝ้าระวังเพื่อสุขภาพในชีวิตประจำวัน, สำนักพิมพ์อินทรีย์, กรุงเทพ, 2558.

[12] M.A. Papadakis and S.J. McPhee, Current Medical Diagnosis & Treatment, McGraw-Hill Education, New York, 2013.

[13] ปราณี ทู้ไพเราะ, คู่มือการตรวจวินิจฉัยโรค, สำนักพิมพ์เอ็นพีเพลส, กรุงเทพฯ, 2558

[14] B.G. Katzung and A.J. Trevor, Basic and Clinical Pharmacology, Mcgraw-Hill Education, New York, 2015.

การพัฒนาระบบวินิจฉัยโรคปาล์มน้ำมันโดยใช้เทคนิคต้นไม้ตัดสินใจ
Development of Oil Palm Disease Diagnosis System by using Decision Tree

วงกต ศรีอุไร (Wongkot Sriurai)[1] พงษ์เพชร ถาพร(Phongphet Thaporn)[2] และวัชรินทร์ นามปัญญา (Watcharin Nampanya)[3]

[1,2,3]ภาควิชาคณิตศาสตร์ สถิติ และคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยอุบลราชธานี

[1]wongkot.s@ubu.ac.th, [2]phongphet.th.54@ubu.ac.th, [3]watcharin.na.54@ubu.ac.th

บทคัดย่อ

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อสร้างโมเดลในการวินิจฉัยโรคปาล์มน้ำมันโดยใช้เทคนิคต้นไม้ตัดสินใจและพัฒนาระบบวินิจฉัยโรคปาล์มน้ำมัน ข้อมูลที่ใช้ในการวิเคราะห์โรคเป็นข้อมูลโรคปาล์มน้ำมันระหว่างปี พ.ศ. 2552 ถึง พ.ศ. 2556 ซึ่งข้อมูลเหล่านี้จะถูกนำมาสร้างโมเดลโดยใช้โปรแกรม Weka ระบบนี้พัฒนาขึ้นในรูปแบบของเว็บแอปพลิเคชันโดยใช้ภาษาพีเอชพีร่วมกับฐานข้อมูลมายเอสคิวแอล ผลการพัฒนาและทดสอบระบบ พบว่าระบบสามารถวินิจฉัยโรคปาล์มน้ำมันได้ และระบบสามารถแนะนำวิธีการป้องกันกำจัดโดยใช้สารเคมี สำหรับการทดสอบความถูกต้องของโมเดลโดยใช้โปรแกรม Weka พบว่าโมเดลการวินิจฉัยโรคปาล์มน้ำมันมีค่าความถูกต้องเท่ากับ 94.27% ผลการประเมินความพึงพอใจจากผู้ใช้ทั่วไปจำนวน 30 คน พบว่ามีค่าเฉลี่ยของความพึงพอใจเท่ากับ 4.34 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.63 แสดงให้เห็นว่าผู้ใช้มีความพึงพอใจต่อระบบที่พัฒนาขึ้นนี้อยู่ในระดับมาก

คำสำคัญ: ระบบวินิจฉัย โรคปาล์มน้ำมัน ต้นไม้ตัดสินใจ

Abstract

The objectives of this research are to create a model for oil palm disease analysis using decision tree and to develop a diagnosis system of oil palm disease. The data used in this analysis is oil palm disease data, from 2009 to 2013. The research has created a model by using Weka software. The system was developed in form of a web application by using PHP language with MySQL database. The development and testing results showed that the system was able to diagnose for oil palm disease and recommend chemicals for disease preservation. The accuracy of the proposed model is 94.27%, tested by using Weka. The satisfaction of the system is evaluated by 30 general users using questionnaires. The experimental result gives the mean 4.34 with the standard deviation 0.63. It showed that the users' satisfaction towards this system was in the high level.

Keyword: Diagnosis System, Oil Palm Disease, Decision Tree

1. บทนำ

เนื่องจากปาล์มน้ำมันเป็นพืชที่ให้ผลผลิตน้ำมันปาล์มดิบต่อหน่วยพื้นที่ปลูกสูงกว่าพืชน้ำมันชนิดอื่นทุกชนิด การปลูกปาล์มน้ำมันสามารถให้ผลผลิตสูงถึงประมาณ 520 กิโลกรัมต่อไร่ ซึ่งให้ผลผลิตสูงกว่าถั่วเหลืองถึง 10 เท่า [1] ปัจจุบันปาล์มน้ำมันถูกใช้ในอุตสาหกรรมพลังงานไบโอดีเซล 38 เปอร์เซ็นต์ น้ำมันพืชและอุตสาหกรรมอาหาร 51 เปอร์เซ็นต์ และอุตสาหกรรมที่ไม่ใช่อาหาร เช่น สบู่ 11 เปอร์เซ็นต์ จากข้อดีดังกล่าวทำให้ผู้อุปโภคบริโภคมีความต้องการใช้น้ำมันปาล์มเพิ่มสูงขึ้นทั้งในด้านอาหารและพลังงาน ทำให้มีเกษตรกรสนใจหันมาปลูกพืชปาล์มน้ำมันกันมากขึ้น แต่เกษตรกรผู้ปลูกปาล์มน้ำมันยังต้องประสบปัญหาต่างๆ ที่ทำให้ผลผลิตลดลง เช่น โรคปาล์มน้ำมันที่สร้างความเสียหายต่อการเพาะปลูก ซึ่งจะเป็นผลกระทบให้ผลผลิตปาล์มน้ำมันลดลง ในปัจจุบันมีเว็บไซต์ที่ให้ความรู้ด้านการปลูกปาล์มน้ำมันแก่เกษตรกรผู้ปลูกปาล์ม หรือผู้ที่สนใจ แต่เว็บไซต์

ดังกล่าวมีลักษณะให้ข้อมูลเพียงอย่างเดียว และจำนวนข้อมูลที่รวบรวมไว้มีจำนวนมากและหลากหลาย ยากต่อการศึกษาและค้นหาข้อมูล กว่าจะได้ข้อมูลที่ต้องการต้องใช้เวลานาน

ดังนั้นคณะผู้วิจัยจึงเล็งเห็นความสำคัญของการพัฒนาระบบวินิจฉัยโรคปาล์มน้ำมันโดยใช้เทคนิคต้นไม้ตัดสินใจ โดยเทคนิคนี้จะช่วยวิเคราะห์หาอาการผิดปกติของปาล์มน้ำมัน ว่าเป็นโรคอะไร โดยพิจารณาจากลักษณะอาการของปาล์มที่เกิดโรค สาเหตุที่ทำให้เกิดโรค ระยะการเจริญเติบโตของปาล์มน้ำมัน บริเวณที่เกิดโรค การแพร่ระบาด สภาพแวดล้อมที่ทำให้เกิดโรค และมีการแนะนำวิธีการป้องกันกำจัดและการรักษา ระบบดังกล่าวจะช่วยให้ผู้ใช้มีความสะดวกในการค้นหาข้อมูลและการวิเคราะห์โรคที่เกิดในปาล์มน้ำมัน

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 ทฤษฎีที่เกี่ยวข้อง

2.1.1 ความรู้พื้นฐานเกี่ยวกับโรคที่เกิดกับปาล์มน้ำมัน

โรคที่เกิดกับปาล์มน้ำมัน เป็นอาการผิดปกติที่เกิดขึ้นกับปาล์มน้ำมัน ซึ่งส่งผลต่อปริมาณการให้ผลผลิตปาล์มที่ลดลง โดยอาการผิดปกติดังกล่าว เกิดจากสาเหตุที่แตกต่างกันและมีวิธีการป้องกันและการกำจัดที่แตกต่างกัน [2] ตัวอย่างของโรคที่เกิดกับปาล์มน้ำมันมีดังนี้

-โรคใบไหม้ ลักษณะอาการ พบอาการของโรคครั้งแรกบนยอดที่ยังไม่คลี่หรือบนใบที่เริ่มคลี่สองใบแรกในระยะแรกเกิดจุดเล็กๆ ลักษณะโปร่งใสกระจายอยู่ทั่วไปบนใบ ต่อมาแผลมีการพัฒนาเห็นชัดเจนขึ้น เมื่อแผลเจริญเต็มที่มีรูปร่างกลมสีน้ำตาลดำบุ๋มตรงกลาง ขอบแผลนูนมีลักษณะเป็นมัน แผลมีวงสีเหลืองล้อมรอบ

-โรคผลเน่า ลักษณะอาการ เปลือกนอกของผลจะนุ่มมีสีดำ โดยจะเริ่มจากโคนหรือปลายผลเข้ามา โดยมากจะเกิดกับผลที่สุกแล้ว

2.1.2 ต้นไม้ตัดสินใจ (Decision Tree) [3] เป็นการจำแนกประเภทข้อมูลซึ่งมีลักษณะเหมือนโครงสร้างต้นไม้ภายในต้นไม้ประกอบไปด้วยโหนด (Node) ซึ่งแต่ละโหนดจะแสดงถึงการตัดสินใจบนข้อมูลของคุณสมบัติต่างๆ กิ่งของต้นไม้แสดงถึงค่าหรือผลลัพธ์ที่ได้จากการทดสอบ และใบเป็นสิ่งที่อยู่ล่างสุดของต้นไม้ตัดสินใจจะแสดงถึงกลุ่มของข้อมูล (Class) โดยเทคนิคที่ใช้ในการสร้างต้นไม้ตัดสินใจมีหลายวิธี เช่น

C4.5 [3], [4] เป็นอัลกอริทึมที่ใช้สร้างต้นไม้ตัดสินใจที่พัฒนาโดย Ross Quinlan มีพื้นฐานเดียวกับอัลกอริทึม ID3 โดยขั้นตอนวิธีที่ใช้สร้างต้นไม้ตัดสินใจสำหรับแบ่งกลุ่มข้อมูลจะใช้หลักการของ Information gain เช่น เดียวกับ ID3 แต่จะมีส่วนเพิ่มเติมจาก ID3 เข้ามาโดยอัลกอริทึม C4.5 ได้เพิ่มการใช้ค่ามาตรฐานอัตราส่วนเกน (Gain Ratio criterion) ในการตัดสินใจเลือกคุณลักษณะที่จะใช้เป็นโหนดรากหรือโหนดระดับต่างๆ บนต้นไม้ ดังสมการที่ 1

$$SplitInfo(A) = -\sum_{j=1}^{k} \frac{|D_j|}{|D|} x \log_2 \left(\frac{|D_j|}{|D|} \right) \qquad (1)$$

โดยที่ k เป็นระดับค่าการกระจายของข้อมูล

|D| เป็นจำนวนข้อมูลในฐานข้อมูล D

|Dj| เป็นจำนวนข้อมูลในฐานข้อมูล D ที่มีค่า j เป็นระดับการกระจายของข้อมูล A

ค่าสารสนเทศของการแบ่งแยก (Split information) นี้จะแสดงถึงระดับการกระจายของข้อมูล เมื่อนำค่านี้ไปหารค่าเกนจะได้ค่ามาตรฐานอัตราส่วนเกน สามารถเขียนในรูปสมการที่ 2 ดังนี้

$$GainRatio(D) = \frac{Gain(A)}{SplitInfo(A)} \qquad (2)$$

ค่ามาตรฐานส่วนเกนช่วยแก้ไขความเอนเอียงของค่าเกนได้ โดยทำให้ค่ามาตรฐานอัตราส่วนในการแบ่งด้วยคุณลักษณะที่มีการกระจายสูงถูกปรับลดลง

Random Forest [5] เป็นเทคนิคหนึ่งที่สร้างโมเดลจากการสุ่มเลือกชุด Training Data และสุ่มเลือกคุณลักษณะออกมาหลาย ๆ ชุด จากนั้นจะนำมาสร้าง Decision Tree หลาย ต้น แต่ละต้นจะให้คำตอบออกมา และในขั้นตอนสุดท้ายจะนำคำตอบของต้นไม้แต่ละต้นมารวมกัน เพื่อดูว่าค่าไหนเหมาะสมที่สุด

2.2 งานวิจัยที่เกี่ยวข้อง

ชื่นนภา บุญตาเพศและวงกต ศรีอุไร (2555) [6] ได้พัฒนาระบบวินิจฉัยโรคพริกโดยใช้เทคนิคการจำแนกข้อมูล เพื่อใช้วิเคราะห์ความผิดปกติของพริก โดยอาศัยข้อมูลจากลักษณะ

อาการ พื้นที่และบริเวณเกิดโรค ระยะการเติบโต เชื้อที่ทำให้เกิดโรค วิเคราะห์ว่าพริกเป็นโรคอะไร และแนะนำสารเคมีที่จำเป็นต้องใช้ในการรักษา ซึ่งสร้างโมเดลจากโปรแกรม Weka โดยใช้เทคนิคต้นไม้ตัดสินใจ (Decision Tree) และพัฒนาเว็บไซต์เพื่อใช้ในการวิเคราะห์โรค ผลของการทดสอบโมเดลวินิจฉัยโรคพริกพบว่ามีค่าความถูกต้องเท่ากับ 98.90%

พัชราภรณ์ ราชประดิษฐ์ และ จันทนา จันทราพรชัย (2556) [7] ได้พัฒนาระบบผู้เชี่ยวชาญเพื่อวินิจฉัยโรคข้าว โดยใช้ CLIPS ข้อมูลที่ใช้ศึกษาเป็นโรคข้าวที่เกิดจากเชื้อโรค 3 กลุ่ม ได้แก่ เชื้อแบคทีเรีย เชื้อรา เชื้อไวรัส จำนวน 16 โรค ซึ่งข้อมูลเหล่านี้ถูกนำไปสร้างกฎ ซึ่งจำแนกตามลักษณะอาการของโรคโดยใช้อัลกอริทึม C4.5 ในโปรแกรม Weka ทั้งนี้มีการตรวจสอบความถูกต้องของกฎโดยใช้นักวิชาการเกษตร 10 คน เพื่อยืนยันความถูกต้องของแต่ละกฎ ผลการตรวจสอบพบว่ากฎมีความถูกต้อง 94.5 %

B.Venkatalakshmi และ M.V Shivsankar (2014) [8] ได้ศึกษาการนำเทคนิคเหมืองข้อมูลมาใช้ในการวินิจฉัยโรคหัวใจ โดยมีการนำเทคนิค Naïve Bayes และ Decision Tree มาใช้ในการสร้างโมเดลเพื่อวินิจฉัยโรคหัวใจ ผลการทดสอบประสิทธิภาพของโมเดล พบว่าเทคนิค Decision Tree ให้ค่าความถูกต้อง (Accuracy) สูงกว่า เทคนิค Naïve Bayes

จากงานวิจัยข้างต้นสรุปได้ว่า งานวิจัยที่เกี่ยวกับการวินิจฉัยโรคสามารถประยุกต์ใช้เทคนิคสำหรับนำมาวิเคราะห์โรคได้หลากหลายเทคนิค เช่น ต้นไม้ตัดสินใจ หรือระบบผู้เชี่ยวชาญ ดังนั้นในงานวิจัยนี้จึงประยุกต์ใช้เทคนิคต้นไม้ตัดสินใจสำหรับการวิเคราะห์โรคที่เกิดกับปาล์มน้ำมัน

3. วิธีดำเนินการวิจัย

3.1 ศึกษาทฤษฎี งานวิจัยที่เกี่ยวข้องและรวบรวมข้อมูลที่เกี่ยวข้องกับโรคปาล์มน้ำมัน

ข้อมูลที่นำมาวิเคราะห์เป็นข้อมูลโรคที่เกิดในปาล์มน้ำมันที่ได้รวบรวมมาจากหนังสือทั้งหมด 3 เล่มได้แก่ หนังสือปาล์มน้ำมันการปรับปรุงการขยายพันธุ์ การปลูกและการจัดการสวน [9] หนังสือคู่มือการปลูก-แปรรูปเชิงการค้าปาล์มน้ำมัน [10] และหนังสือคู่มือการปลูกปาล์มน้ำมัน [11] ซึ่งข้อมูลที่นำมาวิเคราะห์เป็นข้อมูลที่อยู่ระหว่างปีพ.ศ. 2552-2556

3.2 สร้างโมเดลโดยใช้โปรแกรม Weka 3.8.1 หลังจากรวบรวมข้อมูลในหัวข้อที่ 3.1 แล้ว ดำเนินการเตรียมข้อมูลเพื่อนำมาใช้สร้างโมเดล ซึ่งมีรายละเอียดดังตารางที่ 1

ตารางที่ 1: ข้อมูลที่ใช้วิเคราะห์โรคที่เกิดในปาล์มน้ำมัน

ข้อมูล	ค่าข้อมูล
1. การแพร่ระบาด	- โรคที่ไม่สามารถแพร่ระบาดได้ - โรคที่สามารถแพร่ระบาดได้
2. สาเหตุที่ทำให้เกิดโรค	- เชื้อสาเหตุมาจากเชื้อรา - เชื้อสาเหตุมาจากเชื้อเห็ด - เชื้อสาเหตุมาจากเชื้อแบคทีเรีย - การเข้าทำลายของแมลง - การถ่ายทอดทางพันธุกรรม -ไม่ทราบสาเหตุที่แน่ชัด
3. บริเวณที่เกิดโรค	- เมล็ด - ราก - ลำต้น - ก้านใบ - ใบ - ใบยอด - ยอด - ผล (บริเวณที่เกิดโรคบางโรคอาจพบ 2 ที่ เช่น ยอดและปลายใบ)
4.ระยะการเจริญเติบโตของปาล์มน้ำมัน	- ระยะเมล็ด - ระยะต้นกล้า - ระยะก่อนให้ผลผลิต
5.สภาพแวดล้อมที่ทำให้เกิดโรค	- สภาพแวดล้อมที่เหมาะสมต่อการเกิดเชื้อรา - สภาพแวดล้อมหลังจากการย้ายต้นกล้า - สภาพแวดล้อมที่มีอากาศแห้งแล้งหรือมีความชื้นน้อย - สภาพดินเป็นที่ลุ่มและมีน้ำขัง - สภาพอากาศชื้นในฤดูฝน - สภาพที่มีการปลูกปาล์มแบบ

	ปลูกชิดติดกัน - สภาพดินที่เคยปลูกยางพาราและมีการระบาดของโรค -ไม่ทราบสภาพอากาศที่แน่ชัด
6.ลักษณะอาการของปาล์มที่เกิดโรค	- ยอดอ่อนเน่าตาย - รากมีลักษณะกรอบ เนื้อเยื่อภายในแห้งเป็นผง - ใบมีแผลและมีจุดเล็กๆ มากมายเรียงตัวกันบนแผล - ลำต้นหักพับที่บริเวณโคนต้น - ผลปาล์มมีอาการเน่าแห้ง (เนื่องจากลักษณะอาการมีจำนวนมาก จึงยกตัวอย่างเพียงบางส่วน)
7.โรคปาล์มน้ำมัน จำนวน 17 โรค	(1) โรคบราวน์เยิม (2) โรคแอนแทรคโนส (3) โรคใบไหม้ (4) โรคใบจุด (5) โรคบลาส (6) โรคทางใบบิด (7) โรคใบจุดสาหร่าย (8) โรคลำต้นเน่า (9) โรคลำต้นส่วนบนเน่า (10) โรคผลร่วง (11) โรคทะลายเน่า (12) โรคยอดเน่า (13) โรคก้านใบเน่า (14) โรคผลเน่า (15) โรคเหี่ยว (16) โรคตาเน่า-ใบเล็ก (17) โรคที่เกิดจากเชื้อ Schizophyllum commune

ข้อมูลในตารางที่ 1 จะถูกแปลงให้เป็นไฟล์ .CSV เพื่อนำไปสร้างโมเดลในโปรแกรม Weka 3.8.1 โดยเลือกวิธีการทดสอบประสิทธิภาพด้วยวิธี 10- fold Cross validation โดยสร้างโมเดลจาก 2 เทคนิค ได้แก่ Random Forest และ C4.5 จากนั้นเปรียบเทียบประสิทธิภาพว่าเทคนิคไหนให้ประสิทธิภาพสูงกว่า เมื่อเปรียบเทียบแล้วนำโมเดลที่มีประสิทธิภาพที่สูงกว่าไปประยุกต์ใช้ในการพัฒนาระบบต่อไป

การวัดประสิทธิภาพของโมเดลใช้สมการหาค่าความถูกต้อง (Accuracy) [4] ของโมเดลจากสมการที่ 3

$$Accuracy = \frac{TP + TN}{(TP + FP + FN + TN)} \quad (3)$$

จากสมการข้างต้นอธิบายค่าของตัวแปรได้ดังนี้ ค่า True Positive (TP) ค่า True Negative (TN) ค่า False Positive (FP) ค่า False Negative (FN)

3.3 วิเคราะห์และออกแบบระบบ หลังจากสร้างโมเดลสำหรับวิเคราะห์โรคที่เกิดในปาล์มน้ำมันแล้ว คณะผู้วิจัยได้วิเคราะห์ความต้องการของระบบ และออกแบบระบบ โดยแบ่งกลุ่มผู้ใช้งานออกเป็น 3 กลุ่ม แต่ละกลุ่มสามารถใช้งานระบบได้ ดังนี้

3.3.1 ผู้ใช้งานทั่วไปสามารถสมัครสมาชิก สามารถกรอกข้อมูลเพื่อวินิจฉัยโรคปาล์มน้ำมัน สามารถดูผลการวินิจฉัยโรคปาล์มน้ำมันและสามารถดูวิธีป้องกันกำจัดโรคปาล์มน้ำมัน

3.3.2 สมาชิกสามารถกรอกข้อมูลเพื่อวินิจฉัยโรคปาล์มน้ำมัน สามารถดูผลการวินิจฉัยและวิธีการป้องกันโรคปาล์มน้ำมัน สามารถบันทึกผลการวินิจฉัยโรคปาล์มน้ำมัน และสามารถค้นหาวิธีป้องกันกำจัดโรคปาล์มน้ำมันโดยใช้ชื่อโรคในการค้นหา

3.3.3 ผู้ดูแลระบบสามารถแก้ไขข้อมูลของโรคปาล์มน้ำมัน สามารถจัดการข้อมูลการป้องกันกำจัดโรคปาล์มน้ำมันสามารถจัดการข้อมูลสารเคมีที่ใช้ในการรักษาโรคปาล์มน้ำมัน สามารถเพิ่ม ลบ ข้อมูลโรคปาล์มน้ำมัน สามารถเพิ่ม ลบ แก้ไขข้อมูลสมาชิก และสามารถดู/ลบผลการวินิจฉัยของสมาชิก

3.4 พัฒนาระบบ คณะผู้วิจัยได้พัฒนาระบบวินิจฉัยโรคปาล์มน้ำมันโดยใช้เทคนิคต้นไม้ตัดสินใจด้วยภาษาพีเอชพี (PHP) ร่วมกับฐานข้อมูลมายเอสคิวแอล (MySQL) และมีการใช้เทคโนโลยี Bootstrap มาช่วยในการพัฒนาระบบเพื่อช่วยให้สามารถพัฒนาระบบได้ง่ายและรวดเร็วมากยิ่งขึ้น

3.5 ทดสอบระบบ ใช้วิธีทดสอบระบบแบบ Black Box Testing เป็นการทดสอบฟังก์ชันต่างๆที่มีในระบบว่าระบบสามารถทำงานได้ตรงตามความต้องการของผู้ใช้หรือไม่ จากนั้นนำระบบนี้ไปให้กลุ่มตัวอย่างทดลองใช้งานจำนวน 30 คน และประเมินความพึงพอใจในการใช้ระบบ

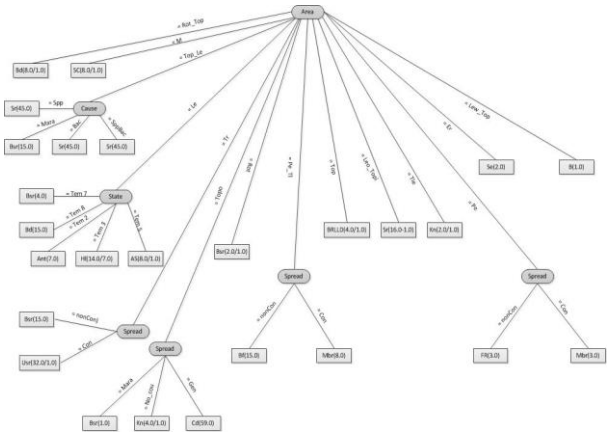
4. ผลการดำเนินการวิจัย

4.1 ผลที่ได้จากการสร้างโมเดลด้วยโปรแกรม Weka 3.8.1

ตารางที่ 2: ผลการวัดประสิทธิภาพของโมเดล

เทคนิค	Accuracy
Random Forest	90.86%
C4.5	94.27%

ผลการวัดประสิทธิภาพของโมเดลในตารางที่ 2 พบว่าเทคนิค C4.5 มีค่าความถูกต้องของโมเดลสูงกว่าเทคนิค Random Forest ดังนั้นงานวิจัยนี้จึงเลือกใช้เทคนิค C4.5 ในการสร้างโมเดลสำหรับวินิจฉัยโรคที่เกิดในปาล์มน้ำมัน (ดังภาพที่ 1) และนำโมเดลไปประยุกต์ใช้ในการพัฒนาระบบระบบวินิจฉัยโรคปาล์มน้ำมัน



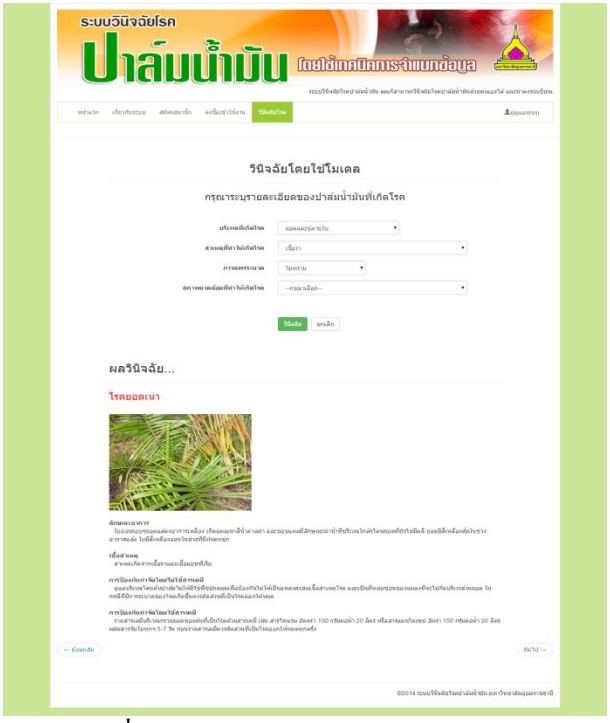
ภาพที่ 1: แผนภาพ treeโมเดลการวินิจฉัยโรคปาล์มน้ำมัน

4.2 ผลที่ได้จากการพัฒนาระบบ

ระบบวินิจฉัยโรคปาล์มน้ำมันโดยใช้เทคนิคต้นไม้ตัดสินใจ ประกอบด้วยส่วนต่างๆ ดังนี้



ภาพที่ 2: หน้าหลักระบบวินิจฉัยโรคปาล์มน้ำมัน



ภาพที่ 3: หน้าผลวินิจฉัยโรคปาล์มน้ำมันโดยใช้โมเดลต้นไม้ตัดสินใจ

4.3 ผลการทดสอบระบบ

คณะผู้วิจัยได้ทดสอบระบบวินิจฉัยโรคปาล์มน้ำมันโดยใช้วิธีการทดสอบแบบ Black box testing ผลการทดสอบพบว่าระบบสามารถทำงานได้ตามวัตถุประสงค์ที่กำหนดไว้

4.4 ผลการประเมินความพึงพอใจของการใช้ระบบ

กำหนดระดับความพึงพอใจ [12] ของการใช้ระบบวินิจฉัยโรคปาล์มน้ำมัน ไว้ดังนี้

4.51 – 5.00 หมายถึง มากที่สุด

3.51 – 4.50 หมายถึง มาก

2.51 – 3.50 หมายถึง ปานกลาง

1.51 – 2.50 หมายถึง น้อย

1.00 – 1.50 หมายถึง น้อยที่สุด

คณะผู้วิจัยได้ศึกษาความพึงพอใจของผู้ใช้ระบบวินิจฉัยโรคปาล์มน้ำมัน จำนวน 30 คน ได้ผลดังนี้

ตารางที่ 3: ผลประเมินความพึงพอใจของผู้ใช้ระบบวินิจฉัยโรคปาล์มน้ำมัน

รายการประเมิน	$\overline{X}$	S.D.	การแปลผล
1. ด้านการตรงตามความต้องการของผู้ใช้	4.27	0.61	มาก
2. ด้านการทำงานได้ตามฟังก์ชันงานของระบบ	4.41	0.67	มาก
3. ด้านความง่ายต่อการใช้งานระบบ	4.33	0.64	มาก
4. ด้านการรักษาความปลอดภัยของข้อมูลในระบบ	4.35	0.61	มาก
ค่าเฉลี่ย	4.34	0.63	มาก

จากตารางที่ 3 สรุปได้ว่าความพึงพอใจของผู้ใช้โดยภาพรวมมีความพึงพอใจอยู่ในระดับมาก ($\overline{X} = 4.34$, $S.D. = 0.63$) โดยผลการประเมินด้านการทำงานได้ตามฟังก์ชันงานของระบบมีค่ามากที่สุด ($\overline{X} = 4.41$, $S.D. = 0.67$)

5. สรุปผลการวิจัย

บทความนี้ได้ดำเนินการสร้างโมเดลสำหรับวิเคราะห์โรคปาล์มน้ำมันโดยใช้เทคนิคต้นไม้ตัดสินใจ (Decision Tree) โปรแกรมที่ใช้สร้างโมเดล คือ โปรแกรม Weka 3.8.1 ผลของการทดสอบประสิทธิภาพของโมเดลโดยใช้เทคนิค C4.5 พบว่าโมเดลวินิจฉัยปาล์มน้ำมันมีค่าความถูกต้องเท่ากับ 94.27% ระบบนี้พัฒนาด้วยภาษาพีเอชพี (PHP) ร่วมกับฐานข้อมูลมายเอสคิวแอล (MySQL) หลังจากที่พัฒนาระบบเรียบร้อยแล้วได้นำระบบไปทดสอบด้านการใช้งาน และทดสอบความพึงพอใจจากผู้ใช้ทั่วไปจำนวน 30 คน ผลการประเมินความพึงพอใจจากผู้ใช้ทั่วไปมีค่าเฉลี่ยเท่ากับ 4.34 และมีค่าเบี่ยงแบนมาตรฐานเฉลี่ยเท่ากับ 0.63 สามารถสรุปได้ว่าระบบที่พัฒนาขึ้นมีค่าความพึงพอใจอยู่ในระดับมาก

เอกสารอ้างอิง

[1] อาริยา จูดคง, “ปาล์มน้ำมัน วิกฤตหรือโอกาส” เข้าถึงได้จาก: http://www.doa.go.th/pibai/pibai/n14/v_8-sep/rai.html , 2555.

[2] รักษ์ พฤกษชาติ, “คู่มือการปลูก-แปรรูปเชิงการค้าปาล์มน้ำมัน”, กรุงเทพมหานคร : นีออน์, 2552.

[3] Ian H. Witten and E. Frank, “Data Mining: Practical Machine Learning Tools And Techniques 3rd” , 2011.

[4] Tan et.al, “Introduction to Data mining”, 2014.

[5] Weka, “Available: http://www.cs.waikato.ac.nz/ml/index.html”, 2016.

[6] ชื่นนภา บุญตาเพศ และวงกต ศรีอุไร “ระบบวินิจฉัยโรคพริกโดยใช้เทคนิคการจำแนกข้อมูล”, The Eighth National Conference on Computing and Information Technology, 2555, หน้า 496-501.

[7] พัชราภรณ์ ราชประดิษฐ์ และจันทนา จันทราพรชัย, “ระบบผู้เชี่ยวชาญเพื่อวินิจฉัยโรคข้าว”, Veridian E-Journal, SU Vol.6 No. 1 January – April, 2556.

[8] B.Venkatalakshmi and M.V Shivsankar, “Heart Disease Diagnosis Using Predictive Data mining”, International Conference on Innovations in Engineering and Technology (ICIET’14), 2014, pp. 1873-1877.

[9] สถานวิจัยพืชกรรมปาล์มน้ำมัน คณะทรัพยากรธรรมชาตและสำนักวิจัยและพัฒนา มหาวิทยาลัยสงขลานครินทร์“ปาล์มน้ำมัน การปรับปรุงขยายพันธุ์ การปลูกและกาจัดการสวน”, สงขลา : มหาวิทยาลัยสงขลานครินทร์, 2553.

[10] รักษ์ พฤกษชาติ, “คู่มือการปลูก-แปรรูปเชิงการค้าปาล์มน้ำมัน”, กรุงเทพมหานคร : นีออน์, 2552.

[11] บุษบา สัตประเสริฐ, “คู่มือการปลูกปาล์มน้ำมัน”, กรุงเทพมหานคร : ศูนย์ส่งเสริมและฝึกอบรม

[12] ราตรี นันทสุคนธ์, “การวิจัยในชั้นเรียนและการวิจัยพัฒนาการเรียนการสอน”, กรุงเทพมหานคร, 2554, หน้า 133 – 136.

ระบบตรวจสอบสินค้าในพิพิธภัณฑ์ผ้าด้วยสมาร์ทโฟนโดยประยุกต์ใช้เทคโนโลยีคิวอาร์โค้ด

System Verify Product Details in Textile Museum by Smartphone using QR code Technology

ณัฐวดี หงษ์บุญมี(Nattavadee Hongboonmee)[1] และสิทธิพล ปราจันทร์(Sittipon Prajan)[2]*

ภาควิชาวิทยาการคอมพิวเตอร์และเทคโนโลยีสารสนเทศ คณะวิทยาศาสตร์ มหาวิทยาลัยนเรศวร

[1]nattavadeeho@nu.ac.th , [2]Sittiponp56@nu.ac.th

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาระบบตรวจสอบสินค้าภายในพิพิธภัณฑ์ผ้าโดยใช้เทคโนโลยีคิวอาร์โค้ดเป็นเครื่องมือในการตรวจสอบข้อมูลสินค้าและช่วยตัดสินใจเลือกซื้อสินค้าภายในพิพิธภัณฑ์ผ้า โดยพัฒนาระบบเป็นเว็บแอปพลิเคชัน ใช้ภาษา PHP, JavaScript และ Bootstrap framework ในการพัฒนาและเชื่อมโยงกับฐานข้อมูลด้วย MySQL เพื่ออำนวยความสะดวกให้กับเจ้าหน้าที่สามารถใช้ระบบสร้างคิวอาร์โค้ด ส่วนการตรวจสอบข้อมูลสินค้าผู้ใช้สามารถสแกนคิวอาร์โค้ดด้วยสมาร์ทโฟน ระบบได้รับการประเมินประสิทธิภาพโดยผู้เชี่ยวชาญจำนวน 4 คน และประเมินความพึงพอใจโดยผู้ใช้ทั่วไปจำนวน 30 คน ผลการประเมินจากผู้เชี่ยวชาญพบว่าระบบมีประสิทธิภาพอยู่ในระดับดี มีค่าเฉลี่ย 4.02 ค่าส่วนเบี่ยงเบนมาตรฐาน 0.55 และผลการประเมินความพึงพอใจจากผู้ใช้ทั่วไปอยู่ในระดับดี มีค่าเฉลี่ยเท่ากับ 3.96 ค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับ0. 4 1สามารถสรุปได้ว่าระบบที่พัฒนาขึ้นนี้ มีประสิทธิภาพสามารถนำไปใช้ในการตรวจสอบข้อมูลสินค้าในพิพิธภัณฑ์ผ้าได้

คำสำคัญ: ระบบตรวจสอบสินค้า คิวอาร์โค้ด สมาร์ทโฟน พิพิธภัณฑ์ผ้า

Abstract

This research aims to develop system verify product details within the textile museum using QR code technology as a tool to check the product information and to purchase goods within the textile museum. By developing web application using PHP, JavaScript and Bootstrap framework to develop a system and connect databases with MySQL for manage database. For convenience, the officer can use the system to create a QR code. The monitoring information is for users to scan a QR code with a Smartphone. System performance has been evaluated by experts of 4 people and the satisfaction of general user 3 0 people. The results of the expert evaluation showed that the efficiency of the system was at a good level, with an average of 4.02, a standard deviation of 0.55. And the results of the evaluation of user satisfaction are good. With an average of 3.96, standard deviation of 0.41. It can be concluded that this developed system can effectively be used for system verify product details in the textile museum.

Keyword: Verify Product Details System, QR code, Smartphone, Textile Museum

1. บทนำ

ปัจจุบันการอนุรักษ์และสืบสานภูมิปัญญาท้องถิ่นด้านผ้าทอนับวันจะเลือนหายไปตามกาลเวลา เพื่อเป็นการอนุรักษ์ไว้ซึ่งภูมิปัญญาชาวบ้าน ดังนั้นมหาวิทยาลัยนเรศวรร่วมกับโครงการร้านจิตรลดา ซึ่งเป็นโครงการเนื่องมาจากพระราชดำริในสมเด็จพระนางเจ้าฯ พระบรมราชินีนาถในรัชกาลที่ ร่วม 9สนับสนุนงบประมาณเกิดเป็นพิพิธภัณฑ์ผ้าภายในมหาวิทยาลัยนเรศวร [1] เพื่อให้เป็นแหล่งเรียนรู้และรวบรวมผลิตภัณฑ์จากผ้าประเภทต่างๆ จัดแสดงผ้าฝ้ายและผลิตภัณฑ์จากโครงการ โดยได้จัดแสดงฉลองพระองค์ในสมเด็จพระนางเจ้า ฯ พระบรมราชินีนาถที่ทรงพระกรุณาโปรดเกล้า ฯ พระราชทานแก่มหาวิทยาลัยนเรศวร นอกจากนี้ยังมีการจัดแสดงผ้าและวิถีชีวิตของกลุ่มชาติพันธุ์ต่างๆ ที่มีการอพยพเข้ามาและมีอยู่เดิมในเขตภาคเหนือ ผ้าโบราณที่มีการนำมาใช้ในพิธีกรรมต่างๆ

นิทรรศการวิถีชีวิตของกลุ่มชาติพันธุ์ใน 9 จังหวัดภาคเหนือตอนล่าง รวมถึงห้องจัดแสดงพิพิธภัณฑ์ผ้าจิตรลดาและการรวบรวมผลิตภัณฑ์งานฝีมือจากแหล่งต่างๆ ภายในประเทศมาจำหน่ายภายในพิพิธภัณฑ์

เนื่องจากความก้าวหน้าของเทคโนโลยีสารสนเทศได้เข้ามามีบทบาทในชีวิตประจำวันมาอย่างต่อเนื่อง โดยเฉพาะเว็บเทคโนโลยีที่ช่วยในการจัดการข้อมูลผ่านเครือข่ายอินเทอร์เน็ตก่อให้เกิดความสะดวกในด้านต่างๆ และเนื่องจากพิพิธภัณฑ์ผ้าซึ่งเป็นสถานที่ที่มีการจัดแสดงผ้าไทย ผ้าทอท้องถิ่นหลายชนิด นอกจากนี้ยังจำหน่ายสินค้าของที่ระลึกและสินค้าอื่นๆ อีกมากมาย แต่จำนวนเจ้าหน้าที่ในการให้บริการยังมีจำนวนน้อยไม่เพียงพอในการให้รายละเอียดข้อมูลเกี่ยวกับผ้าและสินค้าต่างๆ ภายในพิพิธภัณฑ์ ซึ่งต้องสอบถามรายละเอียดข้อมูลจากเจ้าหน้าที่เท่านั้น ทำให้เกิดปัญหาเวลามีผู้เข้าชมพิพิธภัณฑ์จำนวนมากในเวลาพร้อมๆกัน ดังนั้นถ้าสามารถนำเทคโนโลยีมาประยุกต์ใช้เพื่อให้สามารถตรวจสอบข้อมูลเกี่ยวกับผ้าและสินค้าได้ด้วยตนเองประกอบการตัดสินใจในการเลือกซื้อสินค้าภายในพิพิธภัณฑ์ผ้า ก็จะช่วยอำนวยความสะดวกและเพิ่มประสิทธิภาพให้กับการเข้าชมพิพิธภัณฑ์มากขึ้น

ดังนั้นงานวิจัยนี้จึงนำเสนอระบบตรวจสอบข้อมูลสินค้าในพิพิธภัณฑ์ผ้าด้วยสมาร์ทโฟนโดยใช้คิวอาร์โค้ด เนื่องจากสมาร์ทโฟนเป็นอุปกรณ์อิเล็กทรอนิกส์เคลื่อนที่ ช่วยอำนวยความสะดวกในด้านต่างๆ ทำให้การดำเนินชีวิตเป็นไปได้โดยง่าย โดยระบบจะอำนวยความสะดวกให้เจ้าหน้าที่สามารถสร้างคิวอาร์โค้ดได้สะดวกรวดเร็ว เพื่อนำไปติดที่สินค้า และอำนวยความสะดวกให้กับผู้ที่มาใช้บริการพิพิธภัณฑ์ผ้าสามารถสแกนคิวอาร์โค้ดด้วยตนเองเพื่อตรวจสอบข้อมูลสินค้าและช่วยตัดสินใจในการเลือกซื้อสินค้าอีกด้วย

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ในการดำเนินงานและพัฒนาระบบ มีการศึกษาและนำทฤษฎีงานวิจัยที่เกี่ยวข้องมาประยุกต์ใช้ดังนี้คือ

2.1 แนวคิดเกี่ยวกับ Quick Response Code

Quick Response Code หรือ QR code คือบาร์โค้ดสองมิติที่มีต้นกำเนิดมาจากประเทศญี่ปุ่น [2] โดยบริษัท Denso-Wave มุ่งเน้นให้สามารถอ่านได้อย่างรวดเร็ว โดยนิยมใช้กับโทรศัพท์มือถือ รุ่นที่มีกล้องถ่ายภาพ คิวอาร์โค้ดถูกพัฒนาให้สามารถบรรจุข้อมูลได้ทั้งแนวตั้งและแนวนอนทำให้สามารถบรรจุข้อมูลได้ประมาณ 4,000 ตัวอักษรหรือ 200 เท่าของบาร์โค้ด 1 มิติ [3] นอกจากนี้บาร์โค้ด 2 มิติสามารถถอดรหัสได้แม้ภาพบาร์โค้ดบางส่วนมีความเสียหาย อุปกรณ์ที่ใช้อ่านและถอดรหัสบาร์โค้ด 2 มิติ มีตั้งแต่เครื่องอ่านแบบเลเซอร์เหมือนกับบาร์โค้ด 1 มิติ จนถึงโทรศัพท์มือถือแบบมีกล้องถ่ายรูปในตัว ซึ่งติดตั้งโปรแกรมถอดรหัสไว้ภายในเครื่อง ส่วนลักษณะของบาร์โค้ด 2 มิติ มีมากมายตามชนิดของบาร์โค้ดเช่น วงกลม สี่เหลี่ยมจัตุรัสหรือสี่เหลี่ยมผืนผ้า



ภาพที่ 1: บาร์โค้ด 2 มิติแบบ QR code

2.2 แนวคิดเกี่ยวกับ Bootstrap Framework

Bootstrap คือชุดคำสั่งที่ประกอบด้วยภาษา CSS HTML และ JavaScript เป็นชุดคำสั่งที่ถูกพัฒนาขึ้นมาเพื่อกำหนดกรอบหรือรูปแบบการพัฒนาเว็บไซต์ในส่วนของการปฏิสัมพันธ์กับผู้ใช้งานหรือเรียกว่า Front-end Framework ใช้สำหรับพัฒนาเว็บไซต์ส่วนการแสดงผล ซึ่งแตกต่างจากภาษาประเภท Server Side Script เช่น ภาษา PHP Python โดย Bootstrap กำหนดให้นักพัฒนาสามารถนำไปใช้งานพัฒนาเว็บไซต์ได้แบบโอเพนซอร์สในชื่อว่า Bootstrap Framework [4]

2.3 แนวคิดเกี่ยวกับ Responsive Web

Responsive Web เป็นเทคนิคการเขียนเว็บเพื่อปรับรูปแบบการแสดงผลให้เหมาะสมกับอุปกรณ์พกพา [5] เช่น การปรับขนาดตัวอักษรเพื่อรองรับการสัมผัสด้วยนิ้วมือ การสร้าง Responsive Web Design ใช้เทคนิคหลายอย่างร่วมกันไม่ว่าจะเป็น Fluid Grid, Flexible Images และ CSS3 Media Queries การทำเว็บไซต์แบบ Responsive มีประโยชน์คือการทำ Responsive เพียงแค่เว็บไซต์เดียวก็รองรับทุกอุปกรณ์ ช่วยให้เซิฟเวอร์ไม่ต้องทำงานหนักและประหยัดเวลาในการจัดทำเว็บไซต์

2.4 งานวิจัยที่เกี่ยวข้อง

โป่งลางและประเสริฐศักดิ์ [6] ได้ทำการศึกษาเรื่องระบบการตรวจสอบการเข้าเรียนของนิสิตโดยใช้ QR-CODE: กรณีศึกษาคณะศิลปศาสตร์และวิทยาศาสตร์ สาขาวิทยาการคอมพิวเตอร์ โดยประยุกต์ใช้รหัสคิวอาร์โค้ดร่วมกับข้อมูลการลงทะเบียนของนักศึกษา เพื่อพัฒนาระบบสำหรับตรวจสอบรายชื่อนิสิตที่เข้าเรียน ผลจากการทำงานพบว่าระบบการตรวจสอบรายชื่อบนอุปกรณ์เคลื่อนที่ผ่านรหัสคิวอาร์โค้ดสามารถช่วยให้การทำงานโดยรวมสะดวกและรวดเร็วมากยิ่งขึ้น

จุฑารัตน์และณัฐวี [7] ได้ทำการวิจัยเรื่องระบบจัดการครุภัณฑ์ด้วยเทคโนโลยีคิวอาร์โค้ดบนแอนดรอยด์โฟน เป็นการพัฒนาระบบจัดการครุภัณฑ์โดยใช้คิวอาร์โค้ดเป็นเครื่องมือในการจัดการและตรวจสอบข้อมูลครุภัณฑ์ คณะวิทยาศาสตร์และเทคโนโลยี สถาบันเทคโนโลยีปทุมวัน โดยพัฒนาระบบเป็นแอปพลิเคชันบนแอนดรอยด์ด้วยภาษาจาวาและใช้มายเอสคิวแอลเป็นระบบจัดการฐานข้อมูล ผลการประเมินพบว่าระบบมีประสิทธิภาพอยู่ในระดับดีที่สุด คิดเป็นร้อยละ 99 สามารถนำไปประยุกต์ใช้ในการจัดการครุภัณฑ์ได้เป็นอย่างดี

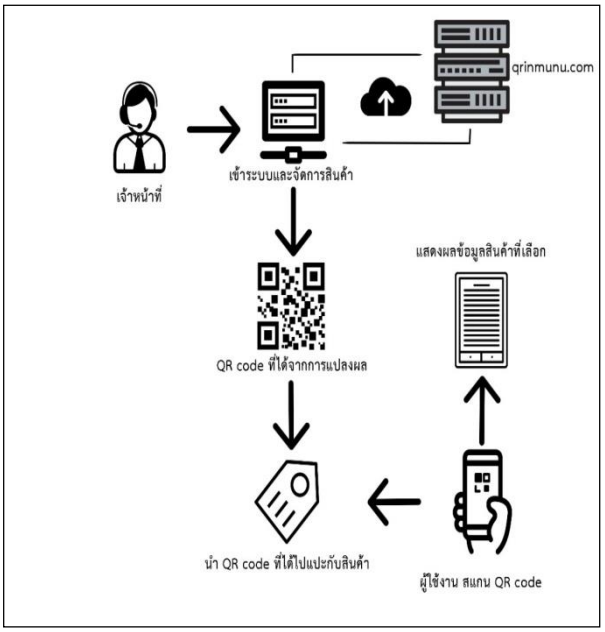
นฤเทพและคณะ [8] ได้วิจัยเรื่องการประยุกต์ใช้คิวอาร์โค้ดในการบันทึกข้อมูลการเข้าร่วมกิจกรรมเพื่อสำเร็จการศึกษา งานวิจัยนี้กำหนดให้มีการบันทึกข้อมูลเกี่ยวกับนักศึกษาลงในคิวอาร์โค้ด แล้วให้ผู้มีส่วนเกี่ยวข้องบันทึกการเข้าร่วมกิจกรรมผ่านบริการซิงก์และฝากข้อมูลในรูปแบบออนไลน์ จากนั้นจะทำการเชื่อมโยงข้อมูลที่ได้สู่ระบบของสำนักทะเบียน ซึ่งจะเป็นผู้ตรวจสอบข้อมูลการเข้าร่วมกิจกรรมของนักศึกษา จากผลการวิจัยพบว่า ผลการประเมินระบบจากผู้ใช้อยู่ในระดับดี สามารถช่วยให้การบันทึกการเข้าร่วมกิจกรรมของนักศึกษามีความถูกต้องชัดเจนและสามารถตรวจสอบได้

3. วิธีการดำเนินการ

3.1 การศึกษาเครื่องมือและวิเคราะห์ปัญหาที่เกี่ยวข้อง

เครื่องมือที่ใช้ในงานวิจัยนี้ คือ ภาษา PHP ภาษา JavaScript Bootstrap Framework และโปรแกรม MySQL เครื่องมือสำหรับการประเมินประสิทธิภาพระบบใช้แบบสอบถามผู้เชี่ยวชาญและแบบสอบความพึงพอใจผู้ใช้ทั่วไป

การรวมรวมข้อมูลเกี่ยวกับสินค้าในพิพิธภัณฑ์ผ้าได้จากการสอบถามเจ้าหน้าที่ที่เกี่ยวข้อง สามารถแบ่งข้อมูลสินค้าในพิพิธภัณฑ์ผ้าออกเป็น 3 ส่วนคือ ผลิตภัณฑ์ผ้า สินค้าโอทอป และของที่ระลึก กำหนดขอบเขตการทำงานของระบบได้ดังนี้ 1) เจ้าหน้าที่พิพิธภัณฑ์ผ้าสามารถสร้างคิวอาร์โค้ดได้สะดวก 2) ผู้ใช้ทั่วไปหรือลูกค้าสามารถใช้สมาร์ทโฟนสแกนข้อมูลคิวอาร์โค้ดเพื่อตรวจสอบข้อมูลสินค้าได้ 3) ข้อมูลสินค้าต้องสามารถปรับปรุงให้ทันสมัยอยู่เสมอ โดยสามารถแสดงขั้นตอนการทำงานของระบบได้ ดังภาพที่ 2



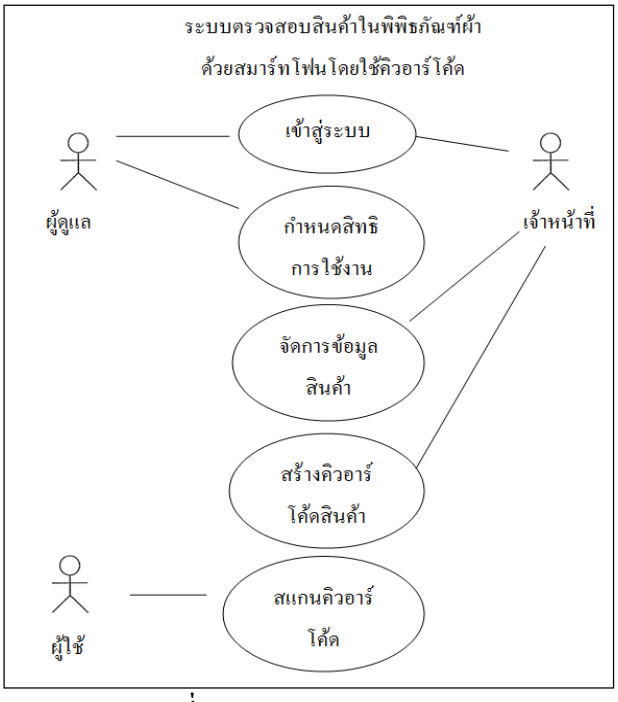
ภาพที่ 2: ขั้นตอนการทำงานระบบตรวจสอบสินค้าด้วยคิวอาร์โค้ด

จากภาพที่ 2 ขั้นตอนการทำงานเริ่มจากเจ้าหน้าที่พิพิธภัณฑ์ผ้าเข้าเว็บไซต์ระบบและดำเนินการจัดการเกี่ยวกับข้อมูลสินค้า สร้างคิวอาร์โค้ดข้อมูลสินค้าแต่ละชนิด จากนั้นนำคิวอาร์โค้ดไปแปะติดที่ตัวสินค้า เมื่อลูกค้าต้องการทราบข้อมูลสินค้า จะทำการสแกนคิวอาร์โค้ดด้วยสมาร์ทโฟน แล้วรายละเอียดข้อมูลสินค้าจะแสดงผลบนหน้าจอสมาร์ทโฟน

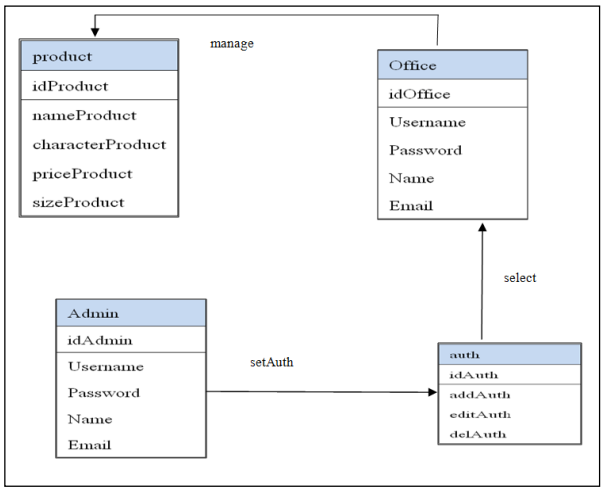
3.2 การวิเคราะห์และออกแบบระบบ

การวิเคราะห์และออกแบบระบบใช้แผนภาพ UML ประกอบด้วย Use Case Diagram ดังภาพที่ 3 โดยกำหนดสิทธิการใช้งานระบบเป็น 3 ส่วนคือ 1) ผู้ดูแลระบบ สามารถกำหนดสิทธิการเข้าระบบได้ 2) เจ้าหน้าที่ สามารถจัดการข้อมูลสินค้าและสร้างคิวอาร์โค้ดของสินค้า และ 3) ผู้ใช้ทั่วไป สามารถใช้

สมาร์ทโฟนสแกนคิวอาร์โค้ดเพื่อตรวจสอบรายละเอียดข้อมูลสินค้าได้



ภาพที่ 3: Use Case Diagram ของระบบ



ภาพที่ 4: Class Diagram ของระบบ

การออกแบบ Class Diagram ของระบบ ประกอบด้วยคลาสผู้ดูแลระบบ (Admin) คลาสกำหนดสิทธิ (auth) คลาสเจ้าหน้าที่ (Officer) และคลาสสินค้า (product) รายละเอียดดังภาพที่ 4

3.3 การพัฒนาระบบ

การพัฒนาระบบเลือกใช้การพัฒนาด้วยเว็บเทคโนโลยี เพื่อให้สามารถบันทึกข้อมูลและแก้ไขข้อมูลได้แบบออนไลน์ โดยเลือกใช้เครื่องมือในการพัฒนาระบบคือภาษา PHP ภาษา JavaScript และ Bootstrap framework เป็นเครื่องมือหลักในการพัฒนาและเชื่อมโยงกับฐานข้อมูลสินค้าด้วย MySQL สำหรับคิวอาร์โค้ดใช้ภาษา JavaScript เป็นภาษาหลักในการพัฒนา

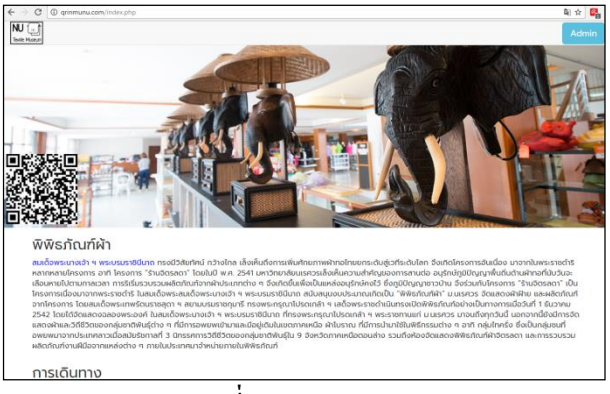
3.4 การทดสอบและประเมินประสิทธิภาพของระบบ

การทดสอบระบบเป็นขั้นตอนที่สำคัญคือการทดสอบระบบที่พัฒนาก่อนที่จะนำมาใช้งานจริง เพื่อลดข้อผิดพลาดที่จะเกิดขึ้น เมื่อพบข้อผิดพลาดจะดำเนินการแก้ไขปรับปรุงให้มีความสมบูรณ์ต่อไป หลังจากทดสอบระบบแล้วจึงนำระบบไปประเมินผลการใช้งานของระบบ โดยทดสอบประสิทธิภาพระบบกับผู้เชี่ยวชาญและผู้ใช้ทั่วไป

4. ผลการดำเนินการ

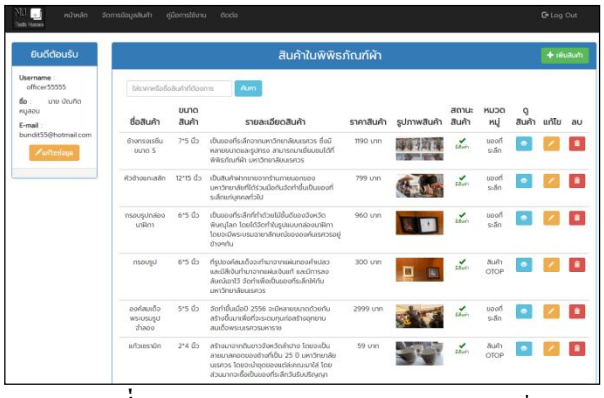
4.1 ผลการพัฒนาระบบ

การพัฒนาระบบตรวจสอบข้อมูลสินค้าในพิพิธภัณฑ์ผ้าด้วยสมาร์ทโฟนโดยใช้คิวอาร์โค้ด คณะผู้วิจัยได้ทำการพัฒนาระบบโดยนำเสนอในรูปแบบของ Web Responsive ซึ่งสามารถปรับการแสดงผลตามขนาดของหน้าจอได้ทั้งเดสก์ทอปและสมาร์ทโฟนประกอบไปด้วยรายละเอียดดังต่อไปนี้



ภาพที่ 5: หน้าจอหลักของระบบ

หน้าจอหลักของระบบ แสดงดังภาพที่ 5 เจ้าหน้าที่พิพิธภัณฑ์ต้องล็อกอินเพื่อเข้าไปยังหน้าจออื่นๆ ของระบบ เมื่อเข้าระบบแล้วเจ้าหน้าที่สามารถจัดการข้อมูลสินค้า เช่น เพิ่ม ลบและแก้ไขข้อมูลสินค้า ตัวอย่างหน้าจอแสดงดังภาพที่ 6



ภาพที่ 6: หน้าจอจัดการข้อมูลสินค้าสำหรับเจ้าหน้าที่



ภาพที่ 7: หน้าจอการสร้างคิวอาร์โค้ด

หน้าจอการสร้างคิวอาร์โค้ดของสินค้า เจ้าหน้าที่ใส่ข้อมูล URL ของสินค้าแล้วกดสร้างคิวอาร์โค้ด ระบบจะสร้างคิวอาร์โค้ดให้เจ้าหน้าที่นำไปแปะติดที่ตัวสินค้า ลูกค้าสามารถสแกนตรวจสอบข้อมูลสินค้าผ่านสมาร์ทโฟนได้ ดังภาพที่ 7 และภาพที่ 8



ภาพที่ 8: แสดงผลการสแกนคิวอาร์โค้ดในสมาร์ทโฟน

ส่วนของบุคคลทั่วไปต้องมีแอปพลิเคชันที่สามารถอ่านคิวอาร์โค้ดเพื่อใช้ในการตรวจสอบข้อมูลสินค้า เช่นแอปพลิเคชัน QR code reader หรือแอปพลิเคชัน Line เป็นต้น นอกจากนี้สมาร์ทโฟนต้องเชื่อมต่อกับเครือข่ายอินเทอร์เน็ต

4.2 ผลการประเมินระบบ

ผลการประเมินระบบใช้การประเมินคุณภาพระบบจากผู้เชี่ยวชาญและประเมินผลความพึงพอใจจากผู้ใช้ทั่วไป ผู้เชี่ยวชาญในงานวิจัยนี้ได้แก่ ผู้เชี่ยวชาญด้านการพัฒนาระบบสารสนเทศจำนวน 2 คน และเจ้าหน้าที่ของพิพิธภัณฑ์ผ้าจำนวน 2 คน รวมทั้งหมด 4 คน มีเกณฑ์ให้คะแนนแบบ 5 คะแนน (คะแนนเท่ากับ 1 คือต้องปรับปรุง – คะแนนเท่ากับ 5 คือดีมาก)

ตารางที่ 1: ผลการประเมินคุณภาพระบบจากผู้เชี่ยวชาญ

รายการประเมิน	ระดับคุณภาพ		
	$\bar{X}$	S.D.	การแปลผล
1. ด้านความถูกต้องในการทำงานของระบบ	4.05	0.55	ดี
2. ด้านความปลอดภัยของระบบ	4.11	0.47	ดี
3. ด้านความเร็วในการทำงานของระบบ	3.89	0.63	ดี
สรุปผลการประเมิน	**4.02**	**0.55**	**ดี**

ผลการประเมินคุณภาพจากผู้เชี่ยวชาญ พบว่าระบบมีคุณภาพการทำงานอยู่ในระดับดี มีค่าเฉลี่ยรวมเท่ากับ 4.02 ค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.55 ดังแสดงรายละเอียดในตารางที่ 1

ตารางที่ 2: ผลการประเมินความพึงพอใจของผู้ใช้

รายการประเมิน	ระดับความพึงพอใจ		
	$\bar{X}$	S.D.	การแปลผล
1. ด้านความสะดวกและง่ายต่อการใช้งานระบบ	3.89	0.20	ดี
2. ด้านความถูกต้องในการทำงานของระบบ	4.20	0.67	ดี
3. ด้านความปลอดภัยของระบบ	4.00	0.05	ดี
4. ด้านความสามารถระบบตรงตามความต้องการ	3.76	0.68	ดี
5. ความพึงพอใจในภาพรวมของระบบ	3.95	0.45	ดี
สรุปผลการประเมิน	**3.96**	**0.41**	**ดี**

การประเมินผลการดำเนินงานระบบใช้ประเมินจากความพึงพอใจของผู้ใช้ได้แก่ ผู้ใช้งานทั่วไปจำนวน 30 คน มีเกณฑ์ให้คะแนนแบบ 5 คะแนน ผลการประเมินพบว่าผู้ใช้งานทั่วไปมีความพึงพอใจอยู่ในระดับดี ค่าเฉลี่ยรวมเท่ากับ 3.96 ค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.41 แสดงรายละเอียดในตารางที่ 2

5. สรุป

การศึกษาวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาระบบตรวจสอบข้อมูลสินค้าในพิพิธภัณฑ์ผ้า มหาวิทยาลัยนเรศวรด้วยเทคโนโลยีคิวอาร์โค้ด เครื่องมือในการพัฒนาระบบใช้ภาษา PHP เป็นภาษาหลักในการพัฒนาใช้ Bootstrap Framework สำหรับการสร้าง Web Responsive ภาษา JavaScript สำหรับการสร้างคิวอาร์โค้ดและ MySQL เป็นระบบจัดการฐานข้อมูล

ผลการพัฒนาระบบแบ่งการทำงานเป็น 2 ส่วนคือสำหรับเจ้าหน้าที่ สามารถจัดการข้อมูลสินค้าและสร้างคิวอาร์โค้ดผ่านเว็บแอปพลิเคชันได้ ผู้ใช้บริการพิพิธภัณฑ์สามารถสแกนคิวอาร์โค้ดผ่านสมาร์ทโฟนเพื่อตรวจสอบข้อมูลสินค้าด้วยตนเองได้ ผลการประเมินระบบโดยผู้เชี่ยวชาญจำนวน 4 คน และประเมินความพึงพอใจโดยผู้ใช้งานทั่วไปจำนวน 30 คน พบว่าการประเมินคุณภาพระบบโดยผู้เชี่ยวชาญ มีค่าเฉลี่ยคุณภาพเท่ากับ 4.02 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.55 ซึ่งแสดงว่าระบบมีคุณภาพการใช้งานอยู่ในระดับดี และการประเมินความพึงพอใจของระบบโดยผู้ใช้ทั่วไป ค่าเฉลี่ยเท่ากับ 3.96 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.41 ซึ่งแสดงว่าผู้ใช้ทั่วไปมีความพึงพอใจต่อระบบอยู่ในระดับดี สามารถสรุปได้ว่าระบบนี้มีประสิทธิภาพสามารถนำไปใช้ในการตรวจสอบข้อมูลสินค้าในพิพิธภัณฑ์ผ้าได้

ข้อเสนอแนะเพิ่มเติมจากกลุ่มตัวอย่างที่ทดสอบระบบคือ 1) เพิ่มการใส่เสียงบรรยายข้อมูลสินค้าเพื่อให้ระบบมีความสมบูรณ์มากขึ้น 2) การแสดงผลในอินเทอร์เน็ตเอกซ์พลอเรอร์มีบางส่วนที่แสดงผลได้ไม่สมบูรณ์เนื่องจาก Bootstrap framework ไม่รองรับอินเทอร์เน็ตเอกซ์พลอเรอร์เวอร์ชั่นเก่า

เอกสารอ้างอิง

[1] สถานอารยธรรมศึกษา โขง-สาละวิน. พิพิธภัณฑ์ผ้า . (ออนไลน์). แหล่งที่มา : http://www.nuks.nu.ac.th/v3/?page_id=91. 20 พฤษภาคม 2559.

[2] ดวงกมล นาคะวัจนะ. “QR Code”, *วารสารประกาย* ปีที่ 8 ฉบับที่ 85 พ.ศ. 2554 หน้า 36.

[3] ThaiKidCom. บาร์โค้ด 2 มิติเทคโนโลยีแห่งอนาคต. (ออนไลน์). แหล่งที่มา : http://school.thaikidscom.com/index.php?option=com_content&view=article&id=152:2010-08-27-07-18-24&catid=38:techniques-for-using-various-applications&Itemid=116. 11 เมษายน 2559.

[4] จีราวุธ วารินทร์, *สร้างเว็บไซต์อีคอมเมิร์ซแบบ Responsive ด้วย PHP Bootstrap*, สำนักพิมพ์ซิมพลิฟาย : กรุงเทพฯ, 2558.

[5] SiamHTML. Responsive Web Design คืออะไร (ออนไลน์). แหล่งที่มา : http://www.siamhtml.com/responsive-web-design-คืออะไร/. 22 เมษายน 2559.

[6] โป่งล่าง เพ็ชรรุ่งและประเสริฐศักดิ์ อู่อรุณ, “ระบบตรวจสอบการเข้าเรียนของนิสิตโดยใช้ QR-CODE: กรณีศึกษาคณะศิลปศาสตร์และวิทยาศาสตร์ สาขาวิทยาการคอมพิวเตอร์” *The 3rd ASEAN Undergraduate Conference in Computing (AUCC2015),* Chachoengsao, Thailand, April 29-30, 2015.

[7] จุฑารัตน์ โถชัย และณัฐวี อุตกฤษฎ์, “ระบบจัดการครุภัณฑ์ด้วยเทคโนโลยีคิวอาร์โค้ดบนแอนดรอย์โฟน” *The 11th National Conference on Computing and Information Technology (NCCIT2015),* Bangkok, Thailand, July 2-3, 2015.

[8] นฤเทพ สุวรรณธาดา สมคิด แซ่หลีและสรเดช ครุฑจ้อน, “การประยุกต์ใช้คิวอาร์โค้ดในการบันทึกข้อมูลการเข้าร่วมกิจกรรมเพื่อสำเร็จการศึกษาโดยนำเข้าข้อมูลผ่านระบบเก็บข้อมูลออนไลน์” *The 9 th National Conference on Computing and Information Technology (NCCIT2013)* Bangkok, Thailand, May 9-10, 2013.

ระบบสารสนเทศเพื่อการจัดการสำหรับการออกแบบเพื่อมวลชนในประเทศไทย: ไทย-อยู่ดี
The Management Information System for Universal Design in Thailand: Thai-UDee

สุไรยา เจริญเระ (Suraiya Charornreh)[1] ฮาซือนี มะสลอง (Hasuenee Masalong)[2]
พรทิพย์ บุญเต็ม (Porntip Boontem)[3] และ ณัฐพงศ์ ทองเทพ (Nattapong Tongtep)[4]
สาขาวิชาเทคโนโลยีสารสนเทศและการสื่อสาร คณะเทคโนโลยีและสิ่งแวดล้อม มหาวิทยาลัยสงขลานครินทร์ วิทยาเขตภูเก็ต
[1]s.charornreh@gmail.com, [2]hasuenee@gmail.com, [3]pornthip785@gmail.com, [4] nattapong.t@phuket.psu.ac.th

บทคัดย่อ

งานวิจัยนี้เป็นการวิเคราะห์และออกแบบระบบสารสนเทศเพื่อการจัดการสำหรับการออกแบบเพื่อมวลชนในประเทศไทย หรือเรียกว่า "ไทย-อยู่ดี" โดยระบบดังกล่าว จะเป็นเว็บแอปพลิเคชันที่มุ่งเน้นอำนวยความสะดวกให้กับผู้บกพร่องทางร่างกาย ผู้ดูแลคนพิการ และเพื่อทุกคน ให้สามารถดำรงชีวิตได้อย่างเท่าเทียมกันในสังคม สามารถค้นหาสถานที่ต่าง ๆ ที่มีสิ่งอำนวยความสะดวกหรือบริการที่สอดคล้องกับสภาพแวดล้อมความต้องการของแต่ละบุคคลได้อย่างสะดวกรวดเร็ว โดยระบบจะแบ่งประเภทบุคคลออกเป็น 10 ประเภท ได้แก่ กลุ่มเพศทางเลือก คนตาบอด คนพิการที่ใช้วีลแชร์ คนพิการที่ไม่ใช้วีลแชร์ คนพร้อมสัตว์เลี้ยง คนมุสลิม เด็ก ผู้สูงอายุ พระภิกษุ และสตรีมีครรภ์ ผลลัพธ์ที่ได้จากการวิจัยนี้จะช่วยส่งเสริมการท่องเที่ยวของประเทศไทย และช่วยยกระดับคุณภาพการใช้ชีวิตของผู้พิการ และผู้สูงอายุ

คำสำคัญ: การออกแบบเพื่อมวลชน ระบบสารสนเทศ เว็บแอปพลิเคชัน

Abstract

This research work focuses on analysis, design and development of a web-based management information system for universal design in Thailand so called "Thai-UDee". The system helps people with disabilities, caregivers, and everyone live equally in the society by conveniently searching for facilities or services which are appropriate to people's needs. Ten types of people are LGBT, blind, people with disabilities who use wheelchair, people with disabilities who do not use wheelchair, people with pets, Muslim, children, elder, monk, and pregnant women. The expected outcome of this research can be used to promote tourism in Thailand and improve the quality of living for people with disabilities and the elderly.

Keywords: Universal Design, Information System, Web Application.

1. บทนำ

จากจำนวนคนพิการและผู้สูงอายุที่มีมากขึ้นเมื่อรวมกับประชากรเด็ก สตรีมีครรภ์ หรือผู้ทุพพลภาพอื่น ๆ แล้ว พบว่าในอนาคตประชากรเหล่านี้จะมีสัดส่วนรวมแล้วมากกว่าคนทั่วไป การขับเคลื่อนเรื่องสภาพแวดล้อมที่เหมาะสมกับคนพิการและผู้สูงอายุจึงเป็นประเด็นที่ทั่วโลกกำลังรณรงค์กันอย่างแพร่หลาย เพื่อให้คนพิการและผู้สูงอายุสามารถใช้ชีวิตได้อย่างเท่าเทียมโดยไม่รู้สึกว่าตนเองเป็นภาระของผู้อื่นในสังคม คณะผู้วิจัยได้เล็งเห็นถึงความสำคัญดังกล่าวนี้จึงได้นำแนวคิดการออกแบบเพื่อมวลชน (Universal Design: UD) มาวิเคราะห์และพัฒนาเป็นระบบสารสนเทศสำหรับประเทศไทย โดยระบบสารสนเทศจะประกอบด้วย สถานที่ต่าง ๆ ที่มีอุปกรณ์ สิ่งอำนวยความสะดวก หรือบริการที่สอดคล้องกับสภาพแวดล้อมความต้องการพิเศษของคนพิการและผู้สูงอายุในรูปแบบเว็บแอปพลิเคชัน เพื่อให้คนทุกกลุ่ม โดยเฉพาะคนพิการและผู้สูงอายุสามารถค้นหาสิ่งต่าง ๆ เหล่านี้ได้อย่างสะดวก รวดเร็ว และใช้งานได้ง่าย ตามแนวทางแห่งพระราชบัญญัติส่งเสริมและพัฒนาคุณภาพชีวิตคนพิการ พ.ศ. 2550 มาตรา 37 บัญญัติให้ออกกฎกระทรวงเพื่อกำหนดลักษณะหรือการจัดให้มีอุปกรณ์ สิ่งอำนวยความสะดวก หรือบริการในอาคาร สถานที่ หรือบริการสาธารณะอื่น ๆ เพื่อให้คนพิการสามารถเข้าถึงและใช้ประโยชน์ได้อย่างทั่วถึงและเป็นธรรม

2. วรรณกรรมและทฤษฎีที่เกี่ยวข้อง

2.1 Universal Design การออกแบบเพื่อมวลชน เป็นแนวความคิดสากลที่องค์การสหประชาชาติได้พยายามเผยแพร่และส่งเสริม จากแนวความคิดเดิมเพื่อให้คนพิการได้รับสิ่งอำนวยความสะดวกในการดำรงชีวิตในอาคารและสิ่งแวดล้อม ตามโครงการ Promotion of Non-Handicapping Physical Environment for Disabled Persons และได้มีการพัฒนาตามลำดับเป็น Accessible Design, Adaptable Design, Barrier Free Design ซึ่งในที่สุดก็เป็นที่ยอมรับโดยทั่วไป

หลักการของ Universal Design ที่เริ่มด้วยความเท่าเทียมกันในการใช้สอยของผู้ใช้ที่ต่างวัยและต่างความสามารถ (Equitable User) ปรับเปลี่ยนการใช้ได้ (Flexible User) ใช้ง่ายด้วยตนเอง (Simple and Intuitive) การสื่อความหมายเป็นที่เข้าใจ (Perceptible Information) มีระยะเผื่อไว้กันผิดพลาด (Tolerance for Error) เบาแรง (Low Physical Effort) และมีขนาดและที่ว่างเพื่อการเข้าถึงและใช้ได้ (Size and Space for Approach and Use) [1].

2.2 กำหนดลักษณะ หรือการจัดให้มีอุปกรณ์ สิ่งอำนวยความสะดวก หรือบริการในอาคาร สถานที่ หรือบริการสาธารณะอื่น ๆ เพื่อให้คนพิการสามารถเข้าถึงและใช้ประโยชน์ได้ พ.ศ. 2555 [2].

ข้อ 1 “สิ่งอำนวยความสะดวก” หมายความว่า เครื่องมือเพื่อการช่วยเหลือทั้งภายในและภายนอกอาคาร หรือสถานที่ โดยการสร้าง ติดตั้ง หรือดัดแปลงให้เครื่องมือดังกล่าวเป็นส่วนหนึ่งของอาคารหรือสถานที่ เพื่อให้คนพิการเข้าถึงและใช้ประโยชน์ได้บนพื้นฐานของความต้องการพิเศษของคนพิการแต่ละประเภท

“บริการ” หมายความว่า การปฏิบัติ การดูแล การให้ความช่วยเหลือ หรือการให้ความสะดวก เพื่อให้คนพิการเข้าถึงและใช้ประโยชน์ได้บนพื้นฐานของความต้องการพิเศษของคนพิการแต่ละประเภทโดยไม่เก็บค่าใช้จ่ายในการให้บริการนั้น และให้หมายความรวมถึงการสนับสนุนวัสดุ อุปกรณ์ หรือสิ่งอำนวยความสะดวกแก่คนพิการ ในการปฏิบัติกิจวัตรประจำวัน หรือการมีส่วนร่วมทางสังคมได้อย่างเท่าเทียมกับบุคคลทั่วไป

ข้อ 2 อุปกรณ์หรือสิ่งอำนวยความสะดวกที่จัดให้มีในอาคารหรือสถานที่ตามกฎกระทรวงนี้ ต้องมีสภาพมั่นคงแข็งแรงและปลอดภัยในการใช้งาน เพื่อให้คนพิการ สามารถเข้าถึง และใช้ประโยชน์ได้อย่างเหมาะสมและสอดคล้องกับความต้องการพิเศษของคนพิการแต่ละประเภท

ข้อ 3 การจัดให้มีอุปกรณ์ สิ่งอำนวยความสะดวก หรือบริการในอาคารหรือสถานที่เพื่อให้คนพิการสามารถเข้าถึงและใช้ประโยชน์ได้ ให้นำบทบัญญัติกฎหมายว่าด้วยการควบคุมอาคารมาใช้บังคับโดยอนุโลม

ข้อ 4 อาคารที่มีผู้ปฏิบัติงานเป็นคนพิการร่วมอยู่ด้วย ให้มีอุปกรณ์สิ่งอำนวยความสะดวก หรือบริการเพื่อให้คนพิการสามารถเข้าถึงและใช้ประโยชน์ได้อย่างเหมาะสม และสอดคล้องกับความต้องการพิเศษของคนพิการแต่ละประเภท

ข้อ 5 อาคารหรือสถานที่ของหน่วยงานของรัฐ องค์กรเอกชน หรือองค์กรอื่นใดให้มีอุปกรณ์สิ่งอำนวยความสะดวก หรือบริการเพื่อให้คนพิการสามารถเข้าถึงและใช้ประโยชน์ได้อย่างหนึ่งอย่างใด

3. วิธีการดำเนินงาน

3.1 การศึกษารวบรวมข้อมูลและวิเคราะห์ระบบ

ได้ศึกษาคู่มือการออกแบบสภาพแวดล้อมสำหรับคนพิการและคนทุกวัย โดยเก็บรวบรวมข้อมูลจากแหล่งต่าง ๆ นำมาสร้างเป็นฐานองค์ความรู้ เพื่อทำการวิเคราะห์ข้อมูล แยกประเภท และจัดหมวดหมู่ของข้อมูล

3.2 แยกประเภทและจัดหมวดหมู่ของข้อมูล

นำข้อมูลที่ได้มาวิเคราะห์และทำการจำแนกข้อมูลออกเป็น 2 ส่วน ดังนี้

- ส่วนของข้อมูลประเภทบุคคล จำแนกบุคคลออกเป็น 10 ประเภท ดังภาพที่ 1



ภาพที่ 1: สัญลักษณ์ประเภทบุคคล 10 ประเภท

- ส่วนของข้อมูลสิ่งอำนวยความสะดวก จำแนกหมวดหมู่ออกเป็น 25 หมวด 104 รายการ ดังภาพที่ 2



ภาพที่ 2: สัญลักษณ์สิ่งอำนวยความสะดวก 25 หมวดหมู่

3.3 การคำนวณหาร้อยละความเหมาะสมของสถานที่ต่อประเภทบุคคล

จากสิ่งอำนวยความสะดวก 104 รายการ นำมาเข้ากระบวนการวิเคราะห์องค์ความรู้ จำแนกได้สิ่งอำนวยความสะดวกที่เหมาะสมกับบุคคลทั้ง 10 ประเภท ดังนี้

ตารางที่ 1: จำนวนรายการสิ่งอำนวยความสะดวกจาก 104 รายการที่มีความเหมาะสมกับบุคคลทั้ง 10 ประเภท

จำแนกสิ่งอำนวยความสะดวกที่เหมาะสมต่อประเภทบุคคล	กลุ่มเพศทางเลือก	คนตาบอด	คนพร้อมอัศว์เลี้ยง	คนพิการที่ใช้วีลแชร์	คนพิการที่ไม่ใช้วีลแชร์	คนมุสลิม	เด็ก	ผู้สูงอายุ	พระภิกษุ	สตรีมีครรภ์
จำนวนสิ่งอำนวยความสะดวกใน 104 รายการ ที่เหมาะสมกับบุคคลทั้ง 10 ประเภท	55	75	56	94	80	68	56	65	57	58

ตารางที่ 2: ตัวอย่างตารางตรวจสอบความเหมาะสมของสิ่งอำนวยความสะดวกกับประเภทบุคคล

หมวดหมู่	สิ่งอำนวยความสะดวก	กลุ่มเพศทางเลือก	คนตาบอด	คนพร้อมอัศว์เลี้ยง	คนพิการที่ใช้วีลแชร์	คนพิการที่ไม่ใช้วีลแชร์	คนมุสลิม	เด็ก	ผู้สูงอายุ	พระภิกษุ	สตรีมีครรภ์	รวม
1	ที่จอดรถ											
	- มีสัญลักษณ์รูปคนพิการที่พื้น ขนาด 90x90 เซนติเมตร	0	0	0	1	0	0	0	0	0	0	1
	- มีป้ายสัญลักษณ์รูปคนพิการขนาดไม่น้อยกว่า 30x30 เซนติเมตร ติดสูงจากพื้น 2.00 เมตร	0	0	0	1	0	0	0	0	0	0	1
	- ช่องจอดรถ มีขนาด 2.40x6 เมตร มีที่ว่างด้านข้างรถไม่น้อยกว่า 1-1.40 เมตร	0	0	0	1	0	0	0	0	0	0	1
	- มีแสงสว่างเพียงพอ	0	0	0	1	0	0	0	0	0	0	1
2	ทางลาด											
	- อัตราส่วนความสูงต่อทางลาดอยู่ที่ 1:12	1	1	1	1	1	1	1	1	1	1	10
	- มีราวจับทั้ง 2 ข้าง	1	1	1	1	1	1	1	1	1	1	10
	- มีขอบกันตกสูงจากพื้น 5 เซนติเมตร	1	1	1	1	1	1	1	1	1	1	10
	- มีขนาดความกว้าง วัสดุพื้นผิวและความชันที่เหมาะสม	1	1	1	1	1	1	1	1	1	1	10
	- มีการปักมุมพื้นส่วนที่ต่างระดับกันไม่เกิน 45° องศา	1	1	1	1	1	1	1	1	1	1	10
	- มีพื้นผิววัสดุไม่ลื่น ไม่มีร่องและตะแกรงระบายน้ำ	1	1	1	1	1	1	1	1	1	1	10
	รวม	55	75	56	94	80	68	56	65	57	58	664

การคำนวณหาร้อยละความเหมาะสมของสถานที่แต่ละสถานที่ว่ามีความเหมาะสมกับกลุ่มบุคคลแต่ละประเภทร้อยละเท่าไร โดยใช้สูตรดังนี้

$$\frac{\text{สิ่งอำนวยความสะดวกที่สถานที่นั้นมี}}{\text{สิ่งอำนวยความสะดวกทั้งหมด}} \times 100$$

ตัวอย่าง การคำนวณร้อยละความเหมาะสมอ้างอิงจากสิ่งอำนวยความสะดวกของสถานที่ท่องเที่ยวแหลมพรหมเทพ จังหวัด ภูเก็ต สำหรับผู้สูงอายุ

$$F_p = \frac{\#ActualF_p}{\#ExpectedF_p} \times 100 = \frac{18}{65} \times 100 = 27.69$$

จากสูตร F (Facilities) คือ สิ่งอำนวยความสะดวกของแต่ละสถานที่ p (people) คือ ประเภทบุคคล

F_p คือ ร้อยละความเหมาะสมของสิ่งอำนวยความสะดวกแต่ละสถานที่ต่อประเภทบุคคล ในที่นี้คือร้อยละความเหมาะสมของแหลมพรหมเทพสำหรับผู้สูงอายุ

$\#ActualF_p$ คือ จำนวนรายการสิ่งอำนวยความสะดวกตามจริงของสถานที่ที่มีความเหมาะสมต่อประเภทบุคคล สำหรับผู้สูงอายุมีทั้งหมด 18 รายการ ดังตารางที่ 3

$\#ExpectedF_p$ คือ จำนวนรายการของสิ่งอำนวยความสะดวกที่ควรจะมีต่อประเภทบุคคล สำหรับผู้สูงอายุมีทั้งหมด 65 รายการ ดังตารางที่ 1

จากการคำนวณหาร้อยละความเหมาะสมของสถานที่ท่องเที่ยวแหลมพรหมเทพซึ่งมีสิ่งอำนวยความสะดวกที่เหมาะสมต่อประเภทบุคคล สำหรับผู้สูงอายุได้ผลลัพธ์ดังนี้ ความเหมาะสมของสถานที่นี้คิดเป็นร้อยละ 27.69 หรือ 27.7% ของสิ่งอำนวยความสะดวกทั้งหมดที่ประเภทบุคคลควรจะมี

ตารางที่ 3: ตารางร้อยละความเหมาะสมของแหลมพรหมเทพสำหรับบุคคลทั้ง 10 ประเภท

ตัวอย่างการคำนวณ	กลุ่มเพศทางเลือก	คนตาบอด	คนพร้อมอัศว์เลี้ยง	คนพิการที่ใช้วีลแชร์	คนพิการที่ไม่ใช้วีลแชร์	คนมุสลิม	เด็ก	ผู้สูงอายุ	พระภิกษุ	สตรีมีครรภ์	โดยรวม
จำนวนรายการของสิ่งอำนวยความสะดวกตามจริง	20	13	18	21	20	28	14	18	19	17	18.8
จำนวนรายการของสิ่งอำนวยความสะดวกที่ควรจะมี	55	75	56	94	80	68	56	65	57	58	104
ร้อยละความเหมาะสมของสิ่งอำนวยความสะดวก	36.4%	17.3%	32.1%	23.3%	25%	41.2%	25.0%	27.7%	33.3%	29.3%	18.1%

จากตารางการคำนวณนำมาแสดงให้อยู่ในรูปของแถบแสดงสถานะ (Progress Bar) เพื่อให้มองเห็นภาพรวมได้อย่างชัดเจนมากยิ่งขึ้น ดังภาพที่ 3



ภาพที่ 3: แถบแสดงสถานะ (Progress Bar) ร้อยละสิ่งอำนวยความสะดวกของแหลมพรหมเทพต่อประเภทบุคคล

3.4 เครื่องมือที่ใช้ในการพัฒนา

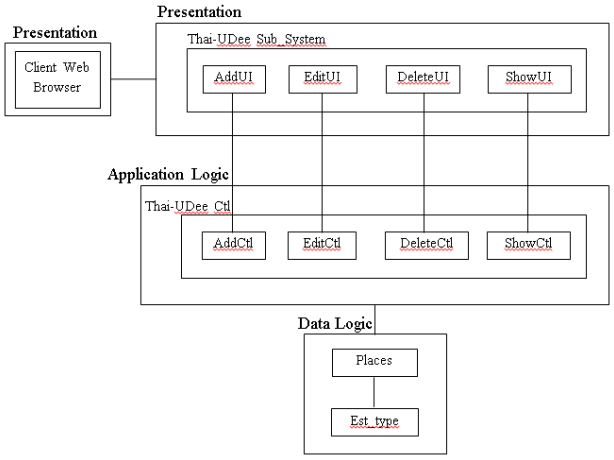
พัฒนาระบบโดยใช้ภาษา PHP, HTML, JavaScript และ CSS ในการเขียนเว็บแอปพลิเคชัน และใช้ XAMPP เป็นตัวจำลอง Server เพื่อทดสอบระบบ และใช้ฐานข้อมูล MySQL

Tools อื่น ๆ ที่นำมาช่วยในการพัฒนาระบบ ได้แก่ Laravel Framework ซึ่งเป็นตัวช่วยในการเขียนเว็บแอปพลิเคชันให้ง่ายมากขึ้น Bootstrap และ Illustrator นำมาช่วยในการพัฒนาเพื่อให้เว็บแอปพลิเคชันมีความสวยงาม และน่าสนใจมากยิ่งขึ้น

3.5 รายละเอียดการพัฒนา

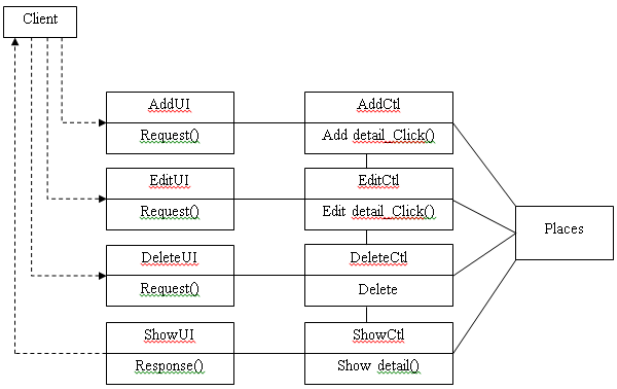
แผนภาพและสถาปัตยกรรมต่าง ๆ ที่ใช้สำหรับอธิบายขั้นตอนการทำงานของระบบสารสนเทศไทย-อยู่ดี

- การออกแบบระบบงานย่อย เป็นสถาปัตยกรรมในการออกแบบระบบงาน ฟังก์ชันการทำงาน และองค์ประกอบย่อยต่าง ๆ ภายในระบบ



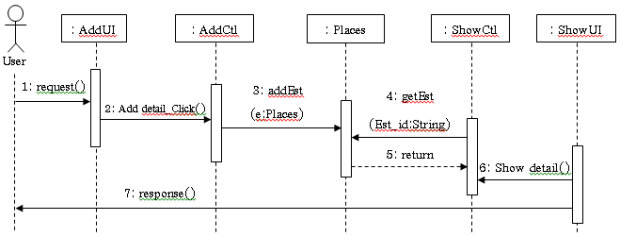
ภาพที่ 4: การออกแบบระบบงานย่อย

- องค์ประกอบภายในระบบงานย่อย ซึ่งจะมีฟังก์ชันหลัก ๆ คือ เพิ่มข้อมูล แก้ไขข้อมูล ลบข้อมูล และแสดงข้อมูล

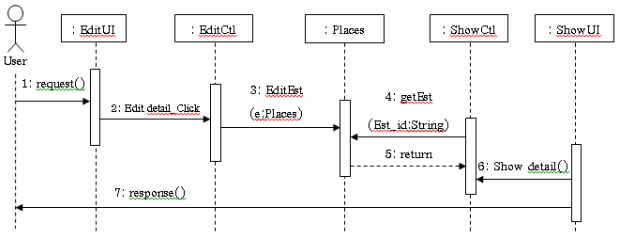


ภาพที่ 5: องค์ประกอบภายในระบบงานย่อย

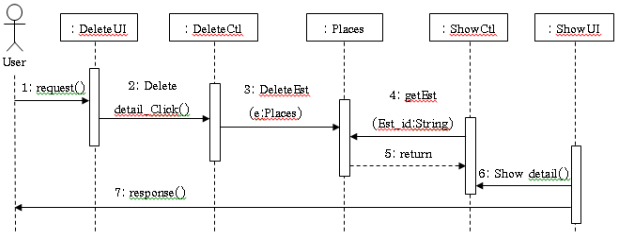
- แผนภาพ Sequence diagram ซึ่งจะแยกองค์ประกอบหลัก ๆ เป็น Interface: Add ดังภาพที่ 6 Interface: Edit ดังภาพที่ 7 Interface: Delete ดังภาพที่ 8 และ Interface: Show ดังภาพที่ 9



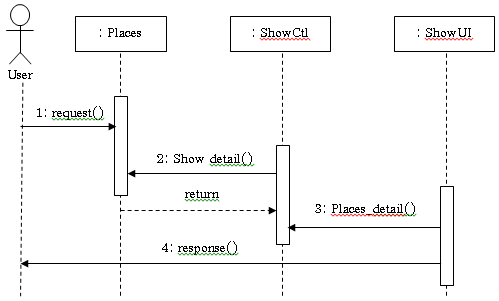
ภาพที่ 6: แผนภาพ Sequence diagram Interface: Add



ภาพที่ 7: แผนภาพ Sequence diagram Interface: Edit

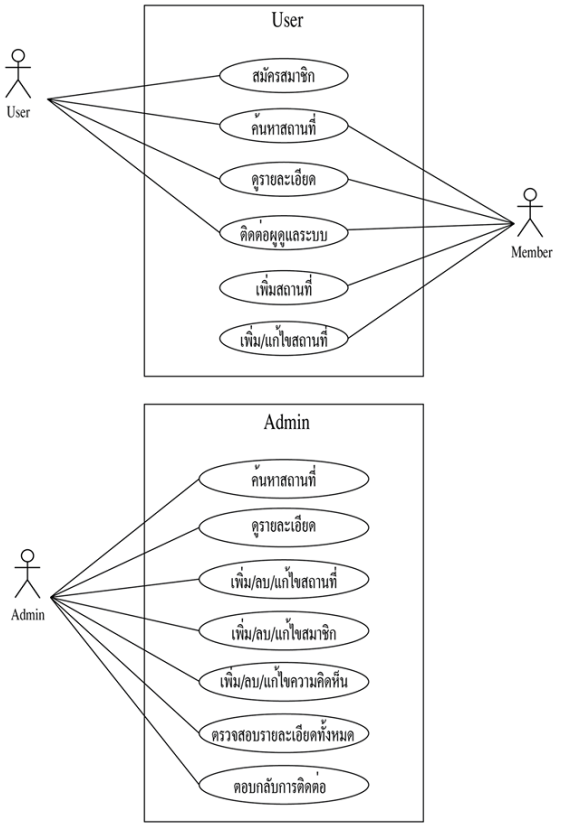


ภาพที่ 8: แผนภาพ Sequence diagram Interface: Delete



ภาพที่ 9: แผนภาพ Sequence diagram Interface: Show

- แผนภาพ Use case diagram อธิบายฟังก์ชันการทำงานต่าง ๆ ของระบบกับผู้ใช้ระบบ



ภาพที่ 10: แผนภาพ Use case diagram

4. ผลการดำเนินงาน

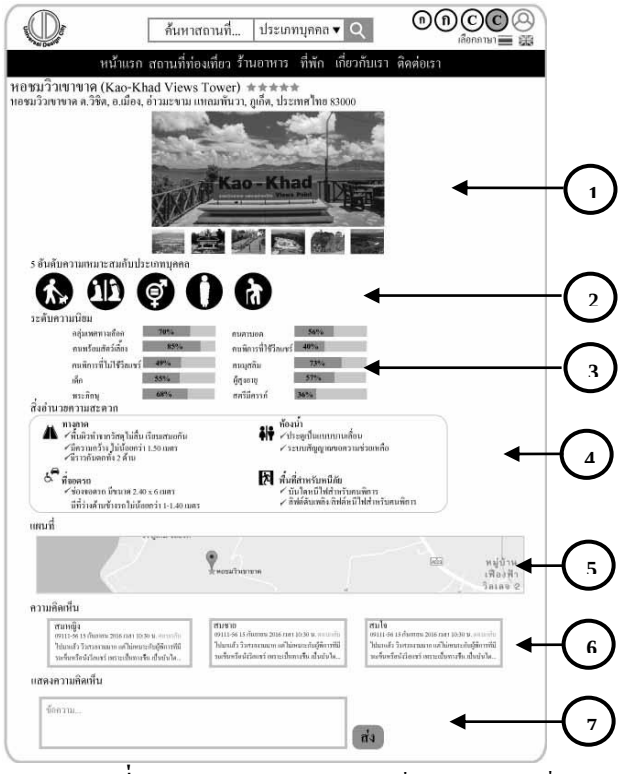
4.1 การทำงานของระบบ



ภาพที่ 11: หน้าจอแสดงหน้าแรกของเว็บไซต์

จากภาพที่ 11: อธิบายการใช้งานเว็บไซต์

1. ผู้ใช้สามารถค้นหาสถานที่ และเลือกประเภทบุคคลได้ตามความต้องการ
2. ผู้ใช้สามารถสมัครสมาชิกเพื่อเข้าสู่ระบบ สามารถย่อ/ขยายขนาดตัวอักษรได้ 2 ระดับ สามารถเปลี่ยนสีพื้นหลัง และรองรับการใช้งานภาษาไทย และภาษาอังกฤษ
3. แท็บเมนูที่เข้าถึงการใช้งานของหน้าเว็บต่าง ๆ
4. ไอคอนแสดงประเภทบุคคลเพื่อให้ผู้ใช้สามารถเข้าถึงข้อมูลได้อย่างสะดวกรวดเร็ว
5. รายการสถานที่ต่าง ๆ เพื่อให้ผู้ใช้สามารถเลือกดูข้อมูลรายละเอียดของสถานที่นั้น ๆ



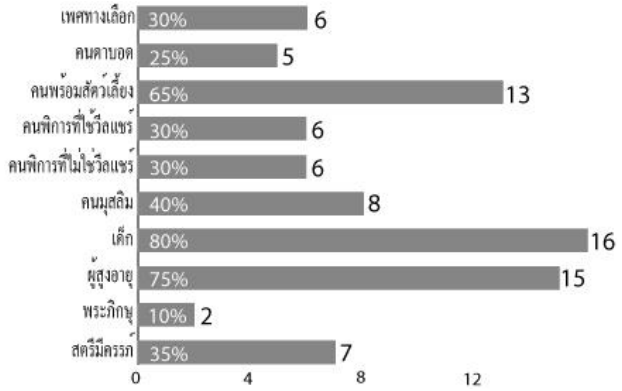
ภาพที่ 12: หน้าจอแสดงรายละเอียดเพิ่มเติมของสถานที่

จากภาพที่ 12: อธิบายการใช้งานเว็บไซต์

1. รูปภาพบรรยากาศของสถานที่ที่ผู้ใช้เลือก
2. ไอคอนแสดง 5 อันดับความเหมาะสมกับประเภทบุคคลต่อสถานที่ที่ผู้ใช้เลือก
3. ระดับความนิยมของประเภทบุคคลแต่ละประเภทที่มีต่อสถานที่นั้น ๆ
4. สิ่งอำนวยความสะดวกที่รองรับต่อประเภทบุคคลที่สถานที่นั้นมี
5. แผนที่แสดงตำแหน่งของสถานที่
6. ข้อมูลแสดงความคิดเห็นของสมาชิก
7. กล่องแสดงความคิดเห็น

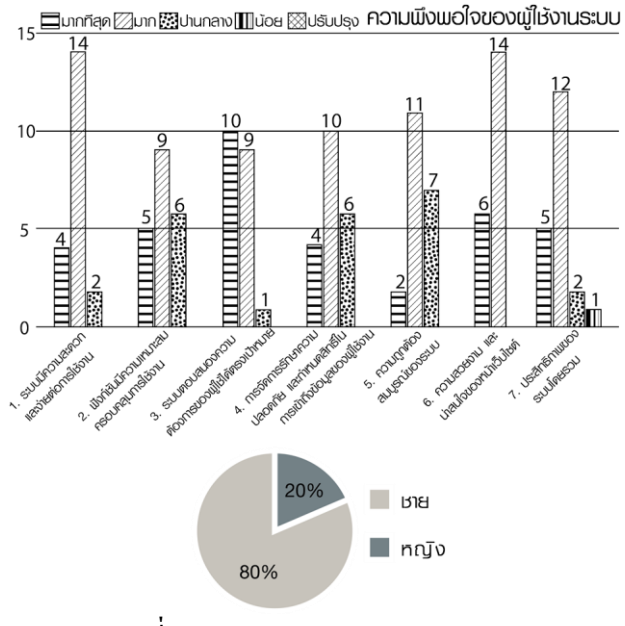
4.2 ผลการทดสอบระบบ

ในการทดสอบระบบผู้วิจัยไม่สามารถนำกลุ่มบุคคลทั้ง 10 ประเภทมาทำการทดสอบใช้งานระบบได้ทั้งหมด จึงได้เลือกกลุ่มบุคคลที่มีความเกี่ยวข้อง และไม่ได้เกี่ยวข้องโดยตรง ซึ่งอาจจะเป็นคนในครอบครัว เพื่อน หรือผู้ที่ดูแลให้ความช่วยเหลือกลุ่มบุคคลทั้ง 10 ประเภท จำนวน 20 คน ทำการทดสอบใช้งานระบบและประเมินผลความพึงพอใจ ได้ผลสรุปดังนี้



ภาพที่ 13: กราฟแสดงกลุ่มคนที่เกี่ยวข้องกับบุคคลทั้ง 10 ประเภท

จากกราฟแสดงกลุ่มคนที่มีความเกี่ยวข้องกับบุคคลทั้ง 10 ประเภท จำนวน 20 คน เมื่อทดสอบใช้งานระบบแล้ว ได้ผลสรุปว่า ใน 20 คนของผู้ที่มาทดสอบใช้งานระบบมีความเกี่ยวข้องกับกลุ่มบุคคลทั้ง 10 ประเภท ดังภาพที่ 13 ยกตัวอย่างเช่น พระภิกษุจำนวน 2 คน คิดเป็นร้อยละ 10% เพศทางเลือก คนพิการที่ใช้วีลแชร์ และคนพิการที่ไม่ใช้วีลแชร์จำนวน 6 คน คิดเป็นร้อยละ 30% เป็นต้น



ภาพที่ 14: กราฟแสดงความพึงพอใจของผู้ใช้งานระบบ

จากกราฟแสดงความพึงพอใจของผู้ใช้งานระบบ มีเกณฑ์การวัดทั้งหมด 5 ระดับด้วยกัน คือ พึงพอใจมากที่สุด = มาก = ปานกลาง = น้อย = และปรับปรุง = ตามลำดับ ดังภาพที่ 14 จากการทดสอบระบบได้ผลสรุปว่า ผู้ใช้ส่วนใหญ่มีความพึงพอใจอยู่ในระดับมาก ซึ่งก็คือผู้ใช้ส่วนใหญ่มีความพึงพอใจมากในการใช้งานระบบสารสนเทศไทย: อยู่-ดี

5. สรุปผลการดำเนินงาน

งานวิจัยนี้ได้พัฒนาระบบเพื่อเพิ่มคุณภาพชีวิตของคนในสังคมโดยเฉพาะคนพิการและผู้สูงอายุให้สามารถดำรงชีวิตได้อย่างเท่าเทียมกับกลุ่มบุคคลอื่น ซึ่งทางผู้วิจัยได้พัฒนาระบบขึ้นมาในรูปแบบของเว็บแอปพลิเคชันที่สามารถค้นหาสถานที่ต่าง ๆ และมีสิ่งอำนวยความสะดวกที่สอดคล้องกับความต้องการของบุคคลแต่ละประเภทเหล่านี้ได้อย่างสะดวก รวดเร็ว และใช้งานได้ง่าย

เนื่องจากข้อมูลของสถานที่ในระบบยังมีไม่มากพอ พร้อมกันนั้นระบบยังไม่ได้เป็นที่รู้จักของกลุ่มเป้าหมาย จึงมีความจำเป็นอย่างยิ่งสำหรับความต้องการข้อมูลสถานที่จากผู้ใช้งานระบบจริงเป็นส่วนสำคัญ โดยเฉพาะผู้ประกอบการต่าง ๆ หรือกลุ่มคนทุกประเภทที่สามารถเพิ่มข้อมูลสถานที่ให้กับระบบได้

เอกสารอ้างอิง

[1] รศ.ไตรรัตน์ จารุทัศน์, *คู่มือการออกแบบสภาพแวดล้อม สำหรับคนพิการ และคนทุกวัย*, หน่วยวิจัยสภาพแวดล้อมที่เหมาะสมกับผู้สูงอายุและคนพิการ คณะสถาปัตยกรรมศาสตร์ จุฬาลงกรณ์, กรุงเทพฯ, 2552

[2] รศ.ไตรรัตน์ จารุทัศน์, *คู่มือรายการอุปกรณ์สิ่งอำนวยความสะดวกหรือบริการที่สอดคล้องกับความต้องการพิเศษของคนพิการ*, หน่วยปฏิบัติการวิจัยสภาพแวดล้อมที่เหมาะสมกับผู้สูงอายุและคนพิการ คณะสถาปัตยกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย, กรุงเทพฯ, 2555

[3] SURANART NIAMCOME, Responsive Web Design (ออนไลน์), แหล่งที่มา: http://www.siamhtml.com/responsive web-design, 8 พฤษภาคม 2559

[4] วิกิพีเดีย สารานุกรมเสรี, *กลุ่มบุคคลที่มีความหลากหลายทางเพศ (ออนไลน์),* https://th.wikipedia.org/wiki/ กลุ่มบุคคลที่มีความหลากหลายทางเพศ, 7 มีนาคม 2558

[5] กภ.ปทิตตา มีหนุน, ความพิการ ความหมาย และประเภทเพื่อการขอรับสิทธิประโยชน์จากรัฐ (ออนไลน์), http://www.pt.mahidol.ac.th/knowledge/?p=70, 28 มีนาคม 2555

การคัดเลือกคู่ค้าโดยการบูรณาการวิธีการตัดสินใจแบบหลายเกณฑ์และวิธีการเขียนโปรแกรมแบบหลายเป้าหมาย กรณีศึกษา บริษัท เอ็นทีเอ็น เทรดดิ้ง-ประเทศไทย จำกัด

Supplier Selection Using Integrated Multi-Criteria Decision Making and MCGP : A Case Study of NTN Trading Thailand Company Limited

นิตญา โพธิ์ดา (N. Phoda)[1], สิรภัทร เชี่ยวชาญวัฒนา (S. Chiewchanwattana)[2] และคำรณ สุนัติ (K. Sunat)[3]

ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น

[1]nittaya.poda@gmail.com, [2]sunkra@kku.ac.th, [3]khamron_sunat@yahoo.com

บทคัดย่อ

งานวิจัยนี้นำเสนอการคัดเลือกคู่ค้าโดยการบูรณาการวิธีการตัดสินใจแบบหลายเกณฑ์ (Multi-Criteria Decision Making) ซึ่งใช้วิธีฟัซซี่ท็อปซิส (Fuzzy TOPSIS) สำหรับแก้ไขปัญหาการตัดสินใจในการเลือกที่มีความหลากหลาย โดยการนำข้อมูลเกณฑ์การเลือกคู่ค้ามาประเมินคู่ค้าที่ดีที่สุดและใช้วิธีการเขียนโปรแกรมแบบหลายเป้าหมาย (Multi-Choice Goal Programming : MCGP) เพื่อแก้ไขปัญหาการสั่งซื้อที่ไม่มีการกำหนดเป้าหมายทำให้สินค้ามีต้นทุนที่สูงในงานวิจัยนี้ใช้ข้อมูลเกณฑ์การประเมินคู่ค้า มูลค่ารวมการสั่งซื้อ จำนวนสั่งซื้อรายไตรมาส ราคาสินค้ารวมต้นทุนต่อชิ้นและข้อมูลเป้าหมายการสั่งซื้อ ผลการวิจัยพบว่าวิธีฟัซซี่ท็อปซิส ได้รับค่าสัมประสิทธิ์ความใกล้ (Closeness Coefficient : CC) ดังนี้ 1. คู่ค้ารายที่ 3 = 0.72 2. คู่ค้ารายที่ 2 = 0.58 3. คู่ค้ารายที่ 1= 0.55 นำค่า CC ที่ได้รับคำนวณร่วมกับวิธี MCGP ได้ค่าดังนี้ 1. จำนวนรวมการสั่งซื้อสูงสุดวิธี MCGP ได้ค่าปริมาณการสั่งซื้อที่น้อยกว่าวิธีเดิมจาก 200 ชิ้นเป็น 170 ชิ้น 2. ต้นทุนรวมการสั่งซื้อต่ำสุดวิธี MCGP ได้ค่ามูลค่าต้นทุนรวมการสั่งซื้อที่ต่ำกว่าวิธีเดิมจาก 3,056,790 บาท เป็น 2,520,000 บาท

คำสำคัญ: การตัดสินใจแบบหลายเกณฑ์ ฟัซซี่ท็อปซิส การเขียนโปรแกรมแบบหลายเป้าหมาย

Abstract

This research presented a Supplier Selection Using Integrated method using Fuzzy TOPSIS for resolving a variety of decision-making decisions and select the best partner from partner selection criteria and used a multi-choice goal programming (MCGP) technique to solve the issue of unspecified order targeting that results in higher cost. This research employed following data: partner selection criteria, total order amount, order quantity per quarter, product price with cost, and order target. The results show that Fuzzy TOPSIS provided closeness coefficient (CC) as follows: 1. Partner 3 = 0.72, 2. Partner 2 = 0.58, 3. Partner 1= 0.55. This CC value was then calculated with MCGP as follows: 1. For the maximum total order quantity, MCGP resulted in less total order quantity than the previous order method from 200 items to 170 items, 2. For the minimum total order cost, MCGP resulted in less total cost than the previous method from 3,056,790 baht to 2,520,000 baht.

Keyword: Multi-Criteria Decision Making, Fuzzy TOPSIS, Multi-Choice Goal Programming.

1. บทนำ

ปัจจุบันธุรกิจอุตสาหกรรมเครื่องมือแพทย์ในประเทศไทยมีการขยายตัวเพิ่มขึ้นอย่างต่อเนื่องที่มูลค่าทางการตลาดในประเทศเพิ่มขึ้นจาก 25,958 ล้านบาท เมื่อปี 2553 และคาดว่ามูลค่าจะเพิ่มขึ้นเป็น 38,000 ล้านบาทในปี 2558 [1] ประเทศไทยได้นำเข้าเครื่องมือและอุปกรณ์ทางการแพทย์ปีละประมาณ 40,000 ล้านบาท คิดเป็น 80% ของปริมาณที่ใช้จริง [2] จากการขยายตัวของธุรกิจที่เพิ่มขึ้นทำให้เกิดการแข่งขันเพิ่มมากขึ้น ดังนั้นการเลือกคู่ค้าที่ดีก็เป็นอีกหนึ่งวิธีที่จะช่วยให้องค์กรได้เปรียบคู่แข่งขัน ในการเลือกคู่ค้าที่เหมาะสมนั้นต้องพิจารณาเกณฑ์ต่างๆ มากมาย เกณฑ์ในการประเมินที่ใช้มีทั้งเกณฑ์เชิงปริมาณและเกณฑ์เชิงคุณภาพ Babak Daneshvar Rouyendegh และคณะได้นำเสนอเกณฑ์ที่ใช้ในการประเมินคู่ค้าทั้งหมด 17 เกณฑ์ [3] ซึ่งเกณฑ์ที่ใช้แตกต่างกันไปตามลักษณะของธุรกิจ

ในการดำเนินธุรกิจเราไม่สามารถเลือกคู่ค้าเพียงรายเดียวได้ เนื่องจากคู่ค้าไม่สามารถตอบสนองความต้องการของบริษัทได้ทุกเรื่อง ทำให้บริษัทต้องทำการพิจารณาเลือกซื้อสินค้ากับคู่ค้าหลายราย ซึ่งไม่ใช่เรื่องง่ายที่จะพิจารณาเลือกคู่ค้าแต่ละรายเมื่อมีการตัดสินใจร่วมกันหลายคน วิธีที่ใช้สำหรับการเลือกมีหลายวิธี เช่น analytic hierarchy process (AHP), analytic network process (ANP), fuzzy sets theory (FST), goal programming (GP) [3] เป็นต้น วิธีการดังกล่าวไม่สามารถครอบคลุมเป้าหมายในการสั่งซื้อ เช่น การหาค่าตัวเลขที่เหมาะสมในการสั่งซื้อ การหาต้นทุนรวมต่ำสุดในการสั่งซื้อ จากเป้าหมายที่มีความหลากหลายทำให้ไม่สามารถใช้วิธี Goal Programming model [4] แก้ไขได้ Chin-Nung Liao และคณะจึงได้เสนอวิธีไฮบริดโดยใช้วิธี Fuzzy TOPSIS มาช่วยในการคัดเลือกคู่ค้าที่มีเกณฑ์ในการเลือกที่หลากหลายและใช้วิธี MCGP [5] มาช่วยแก้ไขปัญหาการหาค่าตัวเลขที่เหมาะสมในการสั่งซื้อ การหาต้นทุนรวมต่ำสุด

งานวิจัยนี้ได้ศึกษาวิธีการคัดเลือกคู่ค้า กรณีศึกษา บริษัท เอ็นทีเอ็น เทรดดิ้ง - ประเทศไทย จำกัด ซึ่งปัญหาเดิมพบว่าในการสั่งซื้อสินค้ากับคู่ค้าไม่มีการกำหนดเป้าหมายและปริมาณการสั่งซื้อ ทำให้มีต้นทุนในการสั่งซื้อที่สูงและไม่มีการประเมินคุณลักษณะต่างๆ ของคู่ค้าที่ชัดเจน เช่น คุณภาพ การส่งสินค้าตรงเวลา เป็นต้น การสั่งซื้อเป็นเพียงการใช้ความรู้สึกและความคุ้นเคยกับคู่ค้าบางรายเท่านั้น ทำให้การสั่งซื้อสินค้ากับคู่ค้าแต่ละราย ได้ตัวเลขที่ยังไม่เหมาะสม ดังนั้นในงานวิจัยนี้จึงได้นำเสนอการคัดเลือกคู่ค้าโดยประยุกต์การบูรณาการวิธีการตัดสินใจแบบหลายเกณฑ์และวิธีการเขียนโปรแกรมแบบหลายเป้าหมาย กรณีศึกษา บริษัท เอ็นทีเอ็น เทรดดิ้ง-ประเทศไทย จำกัด จากผลการทดลองพบว่าการคัดเลือกคู่ค้าด้วยวิธี Fuzzy TOPSIS ให้ค่าสัมประสิทธิ์ในการจัดลำดับคู่ค้าได้เหมาะสม และการคำนวณด้วยวิธี MCGP ได้ปริมาณตัวเลขการสั่งซื้อและตัวเลขต้นทุนการสั่งซื้อที่เหมาะสมกว่าวิธีการเดิม

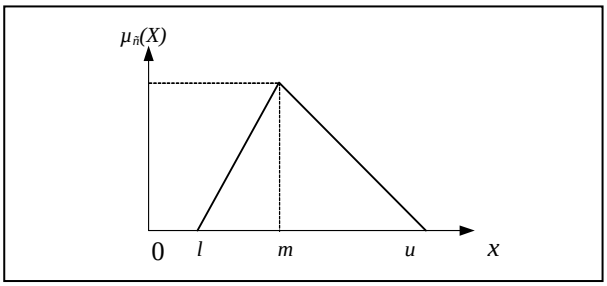
2. วรรณกรรมที่เกี่ยวข้อง

2.1 ฟัซซี่เซต (Fuzzy Set)

ฟัซซี่ (Fuzzy) เป็นทฤษฎีที่ถูกคิดค้นขึ้นโดยศาสตราจารย์ลอตฟี ซาเดห์ในปี ค.ศ. 1965 ฟัซซี่ถูกพัฒนาจากทฤษฎีเซตวิภัชนัยโดยเป็นการใช้เหตุผลแบบประมาณ[3] ซึ่งแตกต่างจากเซตแบบดั้งเดิมที่ใช้เหตุผลแบบเด็ดขาดฟัซซี่ถูกนำมาประยุกต์ใช้งานเพื่อจำลองการตัดสินใจของผู้เชี่ยวชาญต่อปัญหาที่ซับซ้อน สามารถนิยามได้ดังนี้[6]

1) $\tilde{a}$ เป็นฟัซซี่เซตในเอกภพสัมพัทธ์ (Universe of Discourse) ของ X ถูกแสดงถึงคุณลักษณะโดยฟังก์ชั่นความเป็นสมาชิก $\mu_{\tilde{a}}(x)$ แต่ละสมาชิก x ใน X เป็นจำนวนจริงที่อยู่ในช่วง 0 –1 เช่น 0.2, 0.5 เป็นต้น ค่าฟังก์ชั่น $\mu_{\tilde{a}}(x)$ เป็นระดับความเป็นสมาชิกของ x ใน $\tilde{a}$

2) ฟัซซี่มีหลายแบบในรูปแบบฟังก์ชั่นที่ได้รับความนิยมมากที่สุดรูปแบบหนึ่งคือฟังก์ชั่นฟัซซี่รูปสามเหลี่ยม (Triangular fuzzy number: TFN) เนื่องจากง่ายต่อการนำมาคำนวณซึ่งตัวเลขฟัซซี่ใน 1 เซต จะมี $\tilde{a}=(l,m,u)$ ที่เป็นระดับความเป็นสมาชิกซึ่งสามารถแสดงเป็นรูปภาพได้ดังภาพที่ 1



ภาพที่ 1: ตัวเลขฟัซซี่รูปสามเหลี่ยม (Triangular fuzzy number)

จากภาพที่ 1 สามารถแสดงความเป็นสมาชิกฟังก์ชั่นและเงื่อนไขได้ดังนี้ l ถึง m เป็นช่วงบนแกน x ที่ทำให้ฟังก์ชั่นเพิ่มขึ้น m ถึง u เป็นค่าในช่วงที่ฟังก์ชั่นลดลง $l \le m \le u$ สามารถเขียนได้ดังสมการต่อไปนี้

$$\mu_{\tilde{n}}(x)=\begin{cases}0 & \text{เมื่อ } x<l; x>u \\ \dfrac{x-l}{m-l} & \text{เมื่อ } l \le x \le m \\ \dfrac{u-x}{u-m} & \text{เมื่อ } m \le x \le u\end{cases} \quad (1)$$

2.2 ฟัซซี่ท็อปซิส (Fuzzy TOPSIS)

Fuzzy TOPSIS (Technique for order preference by similarity to an ideal solution) ได้รับการพัฒนาโดย Hwang & Yoon ในปี ค.ศ. 1981 วิธีการ TOPSIS ใช้สำหรับแก้ไขปัญหาการตัดสินใจในการเลือกที่มีความหลากหลายในการ

ตัดสินใจ ทางเลือกที่มีค่าสัมประสิทธิ์ความใกล้ชิดมากที่สุดแสดงให้เห็นถึงทางเลือกที่ดีที่สุดและเป็นค่าที่อยู่ใกล้ค่าอุดมคติเชิงบวก (Fuzzy Positive Ideal Solution: FPIS) และไกลค่าอุดมคติเชิงลบ (Fuzzy Negative Ideal Solution: FNIS) มากที่สุด ขั้นตอนของ Fuzzy TOPIS [3] มีรายละเอียดดังนี้

2.2.1 กำหนดทางเลือก กำหนดเกณฑ์ในการประเมินและจัดกลุ่มผู้เชี่ยวชาญในการตัดสินใจ

2.2.2 กำหนดตัวแปรภาษาที่เหมาะสมสำหรับการให้น้ำหนักความสำคัญของเกณฑ์ $\tilde{w}_j = (l_{ij}, m_{ij}, u_{ij})$ เมื่อ j คือเกณฑ์แต่ละเกณฑ์และกำหนดตัวแปรภาษาที่เหมาะสมสำหรับให้คะแนนทางเลือกที่กำหนดไว้ในแต่ละเกณฑ์ของแต่ละทางเลือก $(\tilde{x}_{ij})$ ตัวแปรภาษาที่กำหนดจะแทนค่าด้วยเลขฟัซซี (Triangular fuzzy number: TFN)

2.2.3 รวมค่าน้ำหนักของเกณฑ์และค่าคะแนนในแต่ละเกณฑ์ทั้งหมดที่ได้จากผู้เชี่ยวชาญทุกคนด้วยสมการที่ (2) และสมการที่ (3)

$$\tilde{x}_{ij} = \frac{1}{k}\left[\tilde{x}_{ij}^1 + \tilde{x}_{ij}^2 + ... + \tilde{x}_{ij}^k\right] \quad ; i = 1,2,...,\ m\ ;\ j = 1,2,\ ...,\ n \tag{2}$$

$$\tilde{w}_j = \frac{1}{k}\left[\tilde{w}_j^1 + \tilde{w}_j^2 + ... + \tilde{w}_j^k\right] \quad ; j = 1,2,\ ...,\ n \tag{3}$$

i คือ แถวแทนทางเลือก, j คือ หลักแทนเกณฑ์การประเมิน k คือ จำนวนผู้เชี่ยวชาญ, m คือ จำนวนคู่ค้า, n คือ จำนวนเกณฑ์

2.2.4 สร้างเมทริกซ์ตัดสินใจฟัซซี ด้วยสมการที่ (4)

$$\tilde{D} = \begin{array}{c} \\ A_1 \\ A_2 \\ \vdots \\ A_m \end{array} \begin{array}{c} \begin{array}{cccc} C_1 & C_2 & ... & C_n \end{array} \\ \begin{bmatrix} \tilde{x}_{11} & \tilde{x}_{12} & ... & \tilde{x}_{1n} \\ \tilde{x}_{21} & \tilde{x}_{22} & ... & \tilde{x}_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ \tilde{x}_{m1} & \tilde{x}_{m2} & ... & \tilde{x}_{mn} \end{bmatrix} \end{array} ;\ \tilde{w} = [\tilde{w}_1, \tilde{w}_2, ..., \tilde{w}_n] \tag{4}$$

$i = 1,2,...,m\ ;\ j = 1,2,...,n$

C_n คือ ค่าเกณฑ์การตัดสินใจที่ n, A_m คือ ค่าทางเลือกที่ m

2.2.5 ทำการนอร์มัลไลเมทริกซ์ (Normalize) การตัดสินใจฟัซซี่ดังสมการต่อไปนี้

$$\tilde{R} = [\tilde{r}_{ij}]_{mxn} \quad , i = 1,2,...,m; \quad j = 1,2,...,n \tag{5}$$

สมการข้างต้นสามารถคำนวณเป็นรายละเอียดดังต่อไปนี้

$$\tilde{r}_{ij} = \left(\frac{l_{ij}}{U_j^*}, \frac{m_{ij}}{U_j^*}, \frac{u_{ij}}{U_j^*}\right), \qquad U_j^* = \max\ u_{ij} \tag{6}$$

l_{ij} คือ ตัวเลขฟัซซี่ตัวที่ 1, m_{ij} คือ ตัวเลขตัวเลขฟัซซี่ตัวที่ 2 u_{ij} คือ ตัวเลขฟัซซี่ตัวที่ 3 ค่าที่มากที่สุดในแถวที่ i คอลัมน์ j

2.2.6 กำหนดค่าน้ำหนักกับเมทริกซ์การตัดสินใจของฟัซซี่ที่ผ่านการนอร์มัลไลจากข้อ 2.2.5 โดยใช้สมการที่ (7)

$$\tilde{V} = [\tilde{v}_{ij}]_{mxn} \quad , i = 1,2,...,m; \quad j = 1,2,...,n \tag{7}$$

เมื่อ

$$\tilde{v}_{ij} = \tilde{r}_{ij} \otimes \tilde{w}_j \quad , i = 1,2,...,m; \quad j = 1,2,...,n \tag{8}$$

โดย $\otimes$ คือการนำตัวเลขหลักและแถวในตารางเมทริกซ์มาคูณกันโดยตรง เช่น $(a_1, b_1, c_1) \otimes (a_2, b_2, c_2)$ หมายถึง $(a_1 \times a_2, b_1 \times b_2, c_1 \times c_2)$ และแต่ละ $\tilde{v}_{ij} = (a_{ij}, b_{ij}, c_{ij})$

2.2.7 กำหนดวิธีการแก้ปัญหาในอุดมคติเชิงบวกของฟัซซี่ (fuzzy positive-ideal solution: FPIS) แทนด้วย S^+ และวิธีการแก้ปัญหาอุดมคติเชิงลบของฟัซซี่ (fuzzy negative-ideal solution: FNIS) แทนด้วย S^- สามารถคำนวณโดยใช้สมการที่ (9) และ (10)

$$S^+ = (\tilde{v}_1^+, \tilde{v}_2^+, ..., \tilde{v}_n^+) \tag{9}$$

$$S^- = (\tilde{v}_1^-, \tilde{v}_2^-, ..., \tilde{v}_n^-) \tag{10}$$

โดย $\tilde{v}_j^+ = \max\{\tilde{v}_{ij}\}$ และ $\tilde{v}_j^- = \min\{\tilde{v}_{ij}\}$ ซึ่ง $\tilde{v}_j$ เป็นน้ำหนักปรกติของ TFNs $i = 1,2,...,m;\ j = 1,2,...,m;$

2.2.8 คำนวณระยะห่างแต่ละทางเลือกจากค่าในอุดมคติเชิงบวก $FNIS\ (d^+)$ และค่าในอุดมคติเชิงลบ $FNIS\ (d^-)$ ระยะห่างระหว่างตัวเลขฟัซซี่สามเหลี่ยม เช่น $A_1 = (l_1, m_1, u_1)$ และ $A_2 = (l_2, m_2, u_2)$ สามารถคำนวณได้ดังสมการที่ (11) และ (12)

$$d(A_1, A_2) = \sqrt{\frac{1}{3}\left[(l_1 - l_2)^2 + (m_1 + m_2)^2 + (u_1 - u_2)^2\right]}$$

$$d_i^+ = \Sigma_{j=1}^n\ d(\tilde{v}_{ij}, \tilde{v}_j^+) \quad , i = 1,2,...,m \tag{11}$$

$$d_i^- = \Sigma_{j=1}^n\ d(\tilde{v}_{ij}, \tilde{v}_j^-) \quad , i = 1,2,...,m \tag{12}$$

โดย $\tilde{v}_j^+ = \max\{\tilde{v}_{ij}\}$ และ $\tilde{v}_j^- = \min\{\tilde{v}_{ij}\}$ ซึ่ง $\tilde{v}_j$ เป็นน้ำหนักปรกติของ TFNs

2.2.9 คำนวณค่าสัมประสิทธิ์ความใกล้ชิด (Closeness Coefficient: CC_i) และจัดลำดับของทางเลือกตามค่าสัมประสิทธิ์ที่ได้หลังจากได้รับค่าระยะทาง d^+ และ d^- จะคำนวณค่าสัมประสิทธิ์ความใกล้ชิดของแต่ละทางเลือกโดยใช้สมการที่ (13)

$$CC_i = \frac{d_i^-}{d_i^+ + d_i^-}, i = 1,2,...,m \tag{13}$$

2.3 การเขียนโปรแกรมหลายเป้าหมาย (Multi-Choice Goal Programming)

Multi-Choice Goal Programming (MCGP) [4] เป็นวิธีการวิเคราะห์วางแผนเพื่อรับมือกับการตัดสินใจที่มีเกณฑ์ความต้องการที่หลากหลายซึ่งจะให้ผู้เชี่ยวชาญ (Decision Makers: DMs) ตั้ง ระดับความต้องการออกเป็นหลายส่วน

Fuzzy Goal Program Model (FGP) เป็นแนวคิดใหม่ที่เกี่ยวกับเรื่องขอบเขตบน $(G_{i,\max})$ และขอบเขตล่าง $(G_{i,\min})$ ของระดับความต้องการโดยที่ y_i เป็นค่าที่นำไปสู่เป้าหมายของ MCGP และมีค่าต่อเนื่องซึ่ง $(G_{i,\min}) \le y_i \le (G_{i,\max})$ ดังนั้น MCGP แบ่งเป็นสองประเภทดังต่อไปนี้ตามทางเลือกของการแก้ไข MCGP [3]:

แบบที่ 1: กรณี "ค่ามากดีที่สุด "(the more the better)

$$Max \sum_{x=1}^{n} \left[w_i \left(d_i^+ + d_i^- \right) + \alpha (e_i^+ + e_i^-) \right] \quad (14)$$

$$s.t.\ f_i(X) - d_i^+ + d_i^- = y_i, i = 1,2,...,n \quad (15)$$

$$y_i - e_i^+ + e_i^- = G_{i,\max}, i = 1,2,...,n \quad (16)$$

$$G_i,_{\min} \le y_i \le G_i,_{\max}$$

$$d_i^+, d_i^-, e_i^+, e_i^- \ge 0, i = 1,2,...,n$$

แบบที่ 2 กรณี “ค่าน้อยดีที่สุด” (the less the better)

$$Min \sum_{x=1}^{n} \left[w_i \left(d_i^+ + d_i^- \right) + \alpha (e_i^+ + e_i^-) \right] \quad (17)$$

$$s.t.\ f_i(X) - d_i^+ + d_i^- = y_i, i = 1,2,...,n \quad (18)$$

$$y_i - e_i^+ + e_i^- = G_{i,\min}, i = 1,2,...,n \quad (19)$$

$$G_i,_{\min} \le y_i \le G_i,_{\max}$$

$$d_i^+, d_i^-, e_i^+, e_i^- \ge 0, i = 1,2,...,n$$

$X \in F$ (F เป็นเซตที่เป็นไปได้ X คือค่าใดๆ ในระบบ) ซึ่ง d_i^+ และ d_i^- เป็นค่าเบี่ยงเบนในทางบวกและลบที่สอดคล้องกับเป้าหมายที่ i $|f_i(x) - y_i|$ ส่วน e_i^+ และ e_i^- เป็นมูลค่าตัวแปรตัดสินใจที่ต้องการหา ค่าเหล่านั้นมีความเบี่ยงเบนไปยังเป้าหมายที่มีหลายทางเลือกทางบวกและทางลบที่สอดคล้องกับ $|y_i - G_{i,\min}|$ กรณีค่าน้อยดีที่สุด และ α_i คือค่าน้ำหนักที่คูณกับ $|y_i - G_{i,\min}|$ หรือ $|y_i - G_{i,\max}|$ กรณี ค่ามากดีที่สุด และ α_i คือค่าน้ำหนักที่คูณกับ $|y_i - G_{i,\max}|$ ผู้เชี่ยวชาญเป็นผู้กำหนดค่า $G_{i,\min}$ และ $G_{i,\max}$

2.4 งานวิจัยที่เกี่ยวข้อง

ผู้จัดทำได้ศึกษางานวิจัยที่เกี่ยวข้องและนำวิธีการที่ได้ศึกษามาประยุกต์ใช้ในงานวิจัยนี้ดังตารางที่ 1

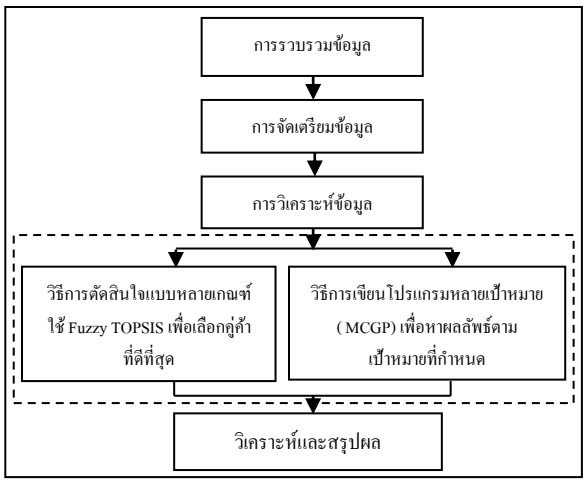
ตารางที่ 1: สรุปงานวิจัยที่เกี่ยวข้อง

ผู้เขียน	นำเสนอวิธี	ข้อดีบทความ
Babak Daneshvar Rouyendegh และคณะ [3]	Fuzzy TOPSIS และ Multi-Choice Goal Programming :MCGP	จัดลำดับคู่ค้าด้วยวิธี Fuzzy TOPSIS จาก CC และใช้วิธี MCGP หาตัวเลขตามเป้าหมายการสั่งซื้อ สามารถครอบคลุมเงื่อนไขในการสั่งซื้อสินค้าแต่ต้องอาศัยผู้ที่มีความชำนาญในการวิเคราะห์ข้อมูลกำหนดเป้าหมาย
O.Jadidi และคณะ [8]	TOPSIS และ fuzzy multi-objective model	งานวิจัยนี้นอกจากประเมินคู่ค้าตามความพึงพอใจของผู้เชี่ยวชาญแล้วยังมีการเขียนโปรแกรมเชิงเป้าหมายเพื่อหาปริมาณการสั่งซื้อที่เหมาะสม
Chin-Nung Liao และคณะ [5]	Fuzzy TOPSIS และ Multi-choice goal programming (MCGP)	มีการเลือกคู่ค้าตามความพึงพอใจของผู้เชี่ยวชาญและยังทำให้บริษัทสามารถจัดสรรปริมาณการสั่งซื้อได้ตรงตามเป้าหมายที่ต้องการ แต่ต้องอาศัยผู้เชี่ยวชาญในการประเมิน

จากการศึกษางานวิจัยที่เกี่ยวข้องพบว่าวิธี **Fuzzy TOPSIS** สามารถนำมาแก้ไขปัญหาการตัดสินใจเลือกที่มีความไม่แน่นอนและวิธี **MCGP** สามารถนำมาประยุกต์ใช้แก้ไขปัญหาการกำหนดเป้าหมายในการสั่งซื้อที่เหมาะสมได้ ดังนั้นผู้วิจัยจึงได้เลือกวิธีการดังกล่าวมาใช้ในงานวิจัยนี้

3. วิธีการดำเนินการวิจัย

งานวิจัยนี้ได้นำเสนอการคัดเลือกคู่ค้าโดยการบูรณาการวิธีการตัดสินใจแบบหลายเกณฑ์ด้วย Fuzzy TOPSIS และวิธี MCGP มีขั้นตอนงานดังภาพที่ 2



ภาพที่ 2: ขั้นตอนการดำเนินงานวิจัย

3.1 การรวบรวมข้อมูลและการจัดเตรียมข้อมูล

ในหัวข้อนี้เป็นการกำหนดตัวแปรภาษาคะแนนความสำคัญของเกณฑ์และคะแนนความพึงพอใจคู่ค้าแบ่งออกเป็น 7 ระดับ

แทนด้วยตัวเลขฟัซซี่ดังตารางที่ 2 และนำข้อมูลประวัติการสั่งซื้อในหนึ่งไตรมาตร จำนวน 36,000 รายการ มาจัดกลุ่มชนิดของสินค้าที่ซื้อบ่อยเลือกชนิดสินค้าที่ต้องการ คำนวณราคาสินค้ารวมต้นทุนต่อหน่วย กำหนดตัวเลขปริมาณการสั่งซื้อกับคู่ค้าแต่ละรายและมูลค่าต้นทุนรวมในการสั่งซื้อนำเสนอผู้บริหารเพื่อกำหนดนโยบายดังนี้ ในการสั่งซื้อสินค้าหนึ่งไตรมาตร จัดซื้อต้องนำเข้าสินค้าชนิดเดียวกันขั้นต่ำ 100 ชิ้น แต่ไม่เกิน 300 ชิ้น แบ่งการสั่งซื้อจากคู่ค้า ดังนี้ รายที่ 1 จำนวน 50 ชิ้น รายที่ 2 จำนวน 70 ชิ้น รายที่ 3 จำนวน 80 ชิ้น ต้นทุนรวมการสั่งซื้อต้องไม่น้อยกว่า 2,522,800 บาท แต่ไม่เกิน 4,452,000 บาท ราคาสินค้ารวมต้นทุน/หน่วย คู่ค้า 1 ~= 14,000 บาท คู่ค้า 2 ~= 15,500 บาท คู่ค้า 3 ~= 15,000 บาท

เกณฑ์ที่ใช้สำหรับประเมินคู่ค้า มีทั้งหมด 18 เกณฑ์ดังนี้ 1. ส่งสินค้าครบตามจำนวน 2. ส่งสินค้าตรงเวลา 3. ไม่มีข้อผิดพลาด 4. ราคาแข่งขันในท้องตลาดได้ 5. การรับประกัน 6. เงื่อนไขการชำระเงิน 7. เงื่อนไขการขนส่ง 8. ภาพลักษณ์องค์กร 9. มีความยืดหยุ่น 10. คุณภาพ 11. การสั่งซื้อขั้นต่ำ 12. กำลังการผลิต 13. มีความซื่อสัตย์ 14. การบริการ 15. สถานที่รับ- ส่งสินค้าที่มีความสะดวก 16. การติดต่อสื่อสารรวดเร็ว 17. มีแค็ตตาล็อคและสินค้าตัวอย่าง 18. มีใบอนุญาตินำเข้าสินค้าเครื่องมือแพทย์ แทนค่าเกณฑ์ด้วยตัวแปร C_i

ตารางที่ 2: แทนค่าตัวเลขฟัซซี่ (TFN) สำหรับตัวแปรภาษาที่ให้คะแนนความสำคัญกับเกณฑ์และคะแนนความพึงพอใจคู่ค้า

คะแนน	ความสำคัญของเกณฑ์		ความพึงพอใจในคู่ค้า	
	ความสำคัญ	TFN	พึงพอใจ	TFN
1	ต่ำมาก	(0, 0, 0.1)	แย่มาก	(0, 0, 1)
2	ต่ำ	(0, 0.1, 0.3)	แย่ปานกลาง	(0,1, 3)
3	ปานกลาง	(0.1, 0.3, 0.5)	แย่	(1, 3, 5)
4	ปานกลาง	(0.3, 0.5, 0.7)	ปานกลาง	(3, 5, 7)
5	สูงปานกลาง	(0.5, 0.7, 0.9)	ดีปานกลาง	(5, 7, 9)
6	สูง	(0.7, 0.9, 1.0)	ดี	(7, 9, 10)
7	สูงมาก	(0.9, 1.0, 1.0)	ดีมาก	(9, 10, 10)

เมื่อผู้เชี่ยวชาญทุกคนได้ประเมินคู่ค้าตามแบบประเมินแล้ว จะทำการแปลงข้อมูลแบบประเมินคู่ค้าเป็นตัวเลขฟัซซี่เพื่อนำไปคำนวณตามวิธีของ Fuzzy TOPSIS ดังตารางที่ 3

3.2 การวิเคราะห์ข้อมูล

ในขั้นตอนการวิเคราะห์ข้อมูลในส่วนของ Fuzzy TOPSIS ผู้จัดทำได้นำข้อมูลการประเมินที่ได้รับจากผู้เชี่ยวชาญไปคำนวณซึ่งจะแสดงในหัวข้อ 4.1 ในส่วนของวิธี MCGP ผู้จัดทำได้อธิบายตัวแปรและค่าพารามิเตอร์ที่ใช้ในหัวข้อ 3.2.1 และตารางที่ 4 จากนั้นกำหนดเป้าหมาย กำหนดเงื่อนไขการสั่งซื้อสินค้าตามที่บริษัทต้องการในหัวข้อ 3.2.2 สร้างสมการจากเป้าหมายที่กำหนด สำหรับเขียนโปรแกรมด้วยวิธี MCGP จะได้สมการที่ (20) – (24) รายละเอียดดังนี้

3.2.1 กำหนดตัวแปรตัดสินใจดังนี้

X_i คือ จำนวนการสั่งซื้อแต่ละบริษัท

Y_i คือ ตัวเลขจำนวนเต็ม ถ้ามีการสั่งซื้อให้ Y มีค่าเท่ากับ 1 ถ้าไม่มีการสั่งซื้อให้ Y มีค่าเท่ากับ 0

ตารางที่ 3: ตัวอย่างการแปลงข้อมูลแบบประเมินคู่ค้าเป็นตัวเลขฟัซซี่

ลำดับ	เกณฑ์ที่ใช้ในการพิจารณา	คะแนนระดับความสำคัญ	คะแนนความพึงพอใจ		
			คู่ค้ารายที่ 1	คู่ค้ารายที่ 2	คู่ค้ารายที่ 3
1.	C1	(0.9,1.0,1.0)	(9,10,10)	(9,10,10)	(9,10,10)
2.	C2	(0.7,0.9,1.0)	(5,7,9)	(7,9,10)	(9,10,10)
3.	C3	(0.9,1.0,1.0)	(9,10,10)	(9,10,10)	(9,10,10)
4.	C4	(0.5,0.7,0.9)	(3,5,7)	(5,7,9)	(9,10,10)
5.	C5	(0.5,0.7,0.9)	(5,7,9)	(5,7,9)	(5,7,9)
6.	C6	(0.5,0.7,0.9)	(3,5,7)	(9,10,10)	(3,5,7)
7.	C7	(0.9,1.0,1.0)	(9,10,10)	(9,10,10)	(9,10,10)
8.	C8	(0.7,0.9,1.0)	(9,10,10)	(7,9,10)	(5,7,9)
9.	C9	(0.5,0.7,0.9)	(9,10,10)	(9,10,10)	(0,1,3)
10.	C10	(0.9,1.0,1.0)	(9,10,10)	(9,10,10)	(9,10,10)
11.	C11	(0.9,1.0,1.0)	(9,10,10)	(9,10,10)	(9,10,10)
12.	C12	(0.7,0.9,1.0)	(5,7,9)	(7,9,10)	(9,10,10)
13.	C13	(0.5,0.7,0.9)	(5,7,9)	(5,7,9)	(5,7,9)
14.	C14	(0.7,0.9,1.0)	(9,10,10)	(7,9,10)	(7,9,10)
15.	C15	(0.9,1.0,1.0)	(9,10,10)	(9,10,10)	(7,9,10)
16.	C16	(0.5,0.7,0.9)	(9,10,10)	(9,10,10)	(0,1,3)
17.	C17	(0.9,1.0,1.0)	(9,10,10)	(9,10,10)	(9,10,10)
18.	C18	(0.9,1.0,1.0)	(9,10,10)	(9,10,10)	(9,10,10)

ตารางที่ 4: ค่าพารามิเตอร์ที่ใช้ในสมการในแต่ละเป้าหมาย

ค่าพารามิเตอร์	ความหมาย
F	จำนวนสินค้าที่ต้องการ
CC_i	ค่าสัมประสิทธิ์ของคู่ค้า
q_i	ราคาสินค้ารวมต้นทุนของคู่ค้า
d_j^+, d_j^-	ค่าเบี่ยงเบนค่าสูงสุดและต่ำสุดของแต่ละเป้าหมาย (Goal – j)
e_j^+, e_j^-	ค่าเบี่ยงเบนค่าสูงสุดและต่ำสุดของตัวเลขระดับความต้องการที่ผู้เชี่ยวชาญกำหนดในแต่ละเป้าหมาย ($y_t - G_{t\max/\min}$)
S_i	จำนวนกำลังการผลิตของคู่ค้าแต่ละราย
$G_{t\max/\min}$	ค่าตัวเลขสูงสุด-ต่ำสุด ที่ผู้เชี่ยวชาญกำหนดในแต่ละเป้าหมาย

3.2.2 กำหนดสมการตามเป้าหมายที่ต้องการดังนี้

สมการเป้าหมายตามวิธีของ MCGP

$$Min \sum_{j=1}^{k}(d_j^+ + d_j^- + e_j^+ + e_j^-) \quad k = |G| \tag{20}$$

G หมายถึง จำนวนเป้าหมายที่กำหนด

เป้าหมายที่ 1 จำนวนรวมการสั่งซื้อสูงสุด (Maximize total value of purchasing: TVP) สมการดังนี้

$$\sum_{i=1}^{n} CC_i.X_i + d_j^- - d_j^+ \geq G_{1\min} \tag{21}$$

เป้าหมายที่ 2 ต้นทุนรวมการสั่งซื้อต่ำสุด (Minimize total cost of purchasing: TCP) สมการดังนี้

$$\sum_{i=1}^{n} q_i.X_i + d_j^- - d_j^+ \geq G_{2\min} \text{ หรือ } \leq G_{2\max} \tag{22}$$

เงื่อนไขความต้องการ

$$\sum_{i=1}^{n} X_i + d_j^- - d_j^+ \leq F \tag{23}$$

$$X_i \leq S_i.Y_i \tag{24}$$

4. ผลการดำเนินการวิจัย

4.1 การตัดสินใจแบบหลายเกณฑ์ (Multi-Criteria Decision Making) ด้วยวิธี Fuzzy TOPSIS

ผลการทดลองได้จากการนำข้อมูลการประเมินคู่ค้าที่ได้รับจากผู้เชี่ยวชาญจำนวน 3 ท่าน ซึ่งให้คะแนนความสำคัญของเกณฑ์ที่ใช้แตกต่างกัน สามารถนำมาคำนวณตามวิธีการ Fuzzy TOPSIS ดังนี้

4.1.1 รวมค่าน้ำหนักของเกณฑ์และค่าคะแนนในแต่ละเกณฑ์ทั้งหมดที่ได้จากผู้เชี่ยวชาญทุกคนด้วยสมการที่ (2) และสมการที่ (3) สร้างเมทริกซ์ตัดสินใจฟัซซี่ ด้วยสมการที่ (4) จะได้ข้อมูลดังตารางที่ 5

4.1.2 ทำการนอร์มัลไลเมทริกซ์ (Normalize) การตัดสินใจฟัซซี่ด้วยสมการ (5)

4.1.3 กำหนดค่าน้ำหนักกับเมทริกซ์การตัดสินใจของฟัซซี่ที่ผ่านการนอร์มัลไลจากข้อ 4.1.2 โดยใช้สมการที่ (7)

4.1.4 กำหนดวิธีการแก้ปัญหาในอุดมคติเชิงบวกของฟัซซี่ (fuzzy positive-ideal solution: FPIS) แทนด้วย S^+ และวิธีการแก้ปัญหาอุดมคติเชิงลบของฟัซซี่ (fuzzy negative-ideal solution: FNIS) แทนด้วย S^- สามารถคำนวณโดยใช้สมการที่ (9) และ (10) ได้ผลลัพธ์ดังนี้

s^+=[(1.0,1.0,1.0),(1.0,1.0,1.0),(1.0,1.0,1.0),(0.97,0.97,0.97),(0.97,0.97,0.97),(0.97,0.97,0.97),(0.84,0.84,0.84),(1.0,1.0,1.0),(0.97,0.97,0.97),(0.71,0.71,0.71),(1.0,1.0,1.0),(1.0,1.0,1.0),(1.0,1.0,1.0),(0.97,0.97,0.97),(1.0,1.0,1.0),(1.0,1.0,1.0),(1.0,1.0,1.0),(1.0,1.0,1.0),(1.0,1.0,1.0)]

s^- = [(0.75, 0.75, 0.75),(0.64, 0.64, 0.64),(0.63, 0.63, 0.63),(0.39, 0.39, 0.39),(0.56, 0.56, 0.56),(0.39, 0.39, 0.39),(0.65, 0.65, 0.65),(0.58, 0.58, 0.58),(0.31, 0.31, 0.31),(0.71, 0.71, 0.71),(0.75, 0.75, 0.75),(0.58, 0.58, 0.58),(0.56, 0.56, 0.56),(0.64, 0.64, 0.64),(0.69, 0.69, 0.69),(0.63, 0.63, 0.63),(0.63, 0.63, 0.63),(0.81, 0.81, 0.81)]

4.1.5 คำนวณระยะห่างแต่ละทางเลือกจากค่าในอุดมคติเชิงบวก $FNIS\ (d^+)$ และค่าในอุดมคติเชิงลบ $FNIS\ (d^-)$ ระยะห่างระหว่างตัวเลขฟัซซี่สามเหลี่ยมด้วยสมการที่ (11) และ (12) จะได้ข้อมูลดังตารางที่ 6 และตารางที่ 7

ตารางที่ 5: เมทริกซ์ตัดสินใจฟัซซี่ (Fuzzy Decision Matrix)

เกณฑ์	คู่ค้ารายที่ 1	คู่ค้ารายที่ 2	คู่ค้ารายที่ 3	น้ำหนัก
C1	(0.9,1,1)	(8.33,9.67,10)	(8.33,9.67,10)	(9,10,10)
C2	(0.83,0.97,1)	(7.67,9,9.67)	(8.33,9.67,10)	(9,10,10)
C3	(0.9,1,1)	(7,8.67,9.67)	(7.67,9.33,10)	(7.67,9,9.67)
C4	(0.77,0.9,0.97)	(5,7,8.67)	(5.67,7.67,9.33)	(8.33,9.67,10)
C5	(0.77,0.9,0.97)	(7,8.67,9.67)	(7,8.67,9.67)	(7,8.67,9.67)
C6	(0.77,0.9,0.97)	(5,7,8.67)	(9,10,10)	(5,7,8.67)
C7	(0.9,1,1)	(7,8.67,9.67)	(7,8.67,9.67)	(7,8.67,9.67)
C8	(0.83,0.97,1)	(9,10,10)	(7.67,9.33,10)	(7,8.67,9.67)
C9	(0.77,0.9,0.97)	(8.33,9.67,10)	(7.67,9.33,10)	(4,5.67,7.33)
C10	(0.9,1,1)	(7.67,9,9.67)	(7.67,9,9.67)	(7.67,9,9.67)
C11	(0.9,1,1)	(8.33,9.67,10)	(8.33,9.67,10)	(8.33,9.67,10)
C12	(0.83,0.97,1)	(7,8.67,9.67)	(7.67,9.33,10)	(8.33,9.67,10)
C13	(0.77,0.9,0.97)	(7.67,9,9.67)	(7,8.67,9.67)	(7,8.67,9.67)
C14	(0.83,0.97,1)	(7.67,9.33,10)	(7.67,9.33,10)	(8.33,9.67,10)
C15	(0.9,1,1)	(8.33,9.67,10)	(8.33,9.67,10)	(7.67,9.33,10)
C16	(0.9,1,1)	(8.33,9.67,10)	(7,8.67,9.67)	(7.67,9.33,10)
C17	(0.9,1,1)	(7,8.67,9.67)	(7,9,10)	(7,8.67,9.67)
C18	(0.77,0.9,0.97)	(5,7,8.67)	(5.67,7.67,9.33)	(8.33,9.67,10)

4.1.6 คำนวณค่าสัมประสิทธิ์ความใกล้ชิด (Closeness Coefficient: CC_i) และจัดลำดับคู่ค้าตามค่าสัมประสิทธิ์ที่ได้

หลังจากได้รับค่าระยะทาง d^+ และ d^- ใช้สมการที่ (13) จะได้ข้อมูลดังตารางที่ 8

ตารางที่ 6: ระยะห่างระหว่าง FPIS S^+ และ FNIS S^-

	FPIS S^+			FNIS S^-		
เกณฑ์/คู่ค้า	คู่ค้ารายที่ 1	คู่ค้ารายที่ 2	คู่ค้ารายที่ 3	คู่ค้ารายที่ 1	คู่ค้ารายที่ 2	คู่ค้ารายที่ 3
C1	0.15	0.15	0.11	0.19	0.19	0.6
C2	0.22	0.18	0.15	0.23	0.27	0.60
C3	0.23	0.18	0.19	0.24	0.28	0.56
C4	0.40	0.35	0.2	0.29	0.35	0.62
C5	0.25	0.25	0.25	0.28	0.28	0.52
C6	0.40	0.17	0.40	0.29	0.48	0.45
C7	0.21	0.21	0.21	0.25	0.25	0.56
C8	0.15	0.21	0.26	0.35	0.31	0.53
C9	0.20	0.23	0.49	0.53	0.51	0.37
C10	0.17	0.17	0.17	0.21	0.21	0.56
C11	0.15	0.15	0.15	0.19	0.19	0.58
C12	0.26	0.21	0.18	0.27	0.31	0.60
C13	0.22	0.25	0.25	0.29	0.28	0.52
C14	0.21	0.21	0.18	0.26	0.26	0.58
C15	0.15	0.15	0.18	0.24	0.24	0.57
C16	0.15	0.23	0.18	0.30	0.24	0.58
C17	0.23	0.22	0.23	0.24	0.26	0.54
C18	0.4	0.35	0.2	0.29	0.35	0.62

ตารางที่ 7: ระยะห่างของแต่ละทางเลือกของ d^+ และ d^-

คู่ค้า/ระยะห่าง	d^+	d^-
คู่ค้ารายที่ 1	3.86	4.81
คู่ค้ารายที่ 2	3.63	5.07
คู่ค้ารายที่ 3	3.89	9.93

ตารางที่ 8: การจัดอันดับคู่ค้าตามค่าสัมประสิทธิ์ของความใกล้ชิด

คู่ค้ารายที่	ค่าสัมประสิทธิ์ความใกล้ชิด *(CC)*	จัดลำดับ
1	0.55	3
2	0.58	2
3	0.72	1

จากตารางที่ 8 วิธีของฟัซซี่ท็อปซิสสามารถช่วยในการตัดสินใจเลือกได้ดังนี้ คู่ค้ารายที่ 3 ได้ค่า CC = 0.72 จะถูกพิจารณาเลือกเป็นลำดับที่ 1 คู่ค้ารายที่ 2 ได้ค่า CC = 0.58 จะถูกพิจารณาเลือกเป็นลำดับที่ 2 และ คู่ค้ารายที่ 1 ได้ค่า CC = 0.55 จะถูกพิจารณาเลือกเป็นลำดับที่ 3

4.2 การเขียนโปรแกรมหลายเป้าหมาย (Multi-Choice Goal Programming: MCGP)

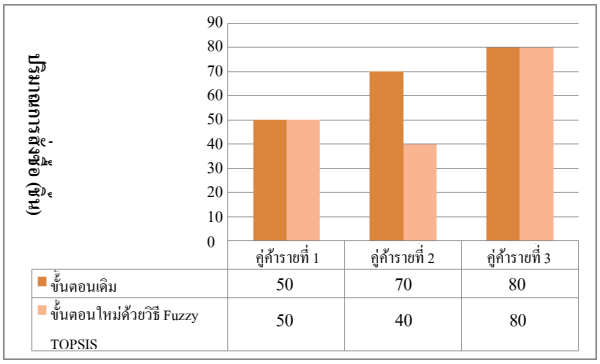
ในงานวิจัยนี้ผู้จัดได้นำค่า CC ที่ได้รับจากตารางที่ 8 มาคำนวณร่วมกับวิธี MCGP โดยใช้โปรแกรม Lindo เวอร์ชั่น 16 ในการแก้ไขปัญหาตามสมการที่ (20) – (24) ได้ผลลัพธ์ดังตารางที่ 9

ตารางที่ 9: ปริมาณการสั่งซื้อสินค้าที่ได้รับจากการคำนวณและตัวเลขต้นทุนรวมการสั่งซื้อ

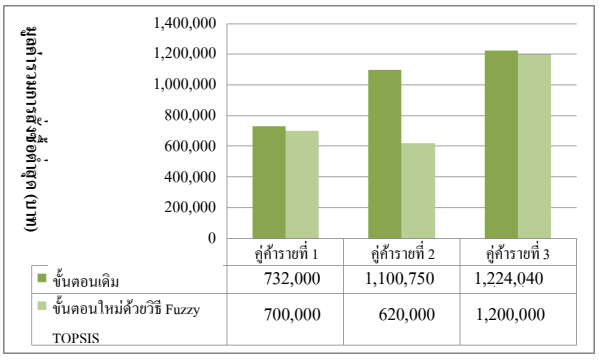
คู่ค้า	ตัวแปรตัดสินใจ	ปริมาณการสั่งซื้อ	ต้นทุนรวม
รายที่ 1	X1	50 ชิ้น	700,000 บาท
รายที่ 2	X2	40 ชิ้น	620,000 บาท
รายที่ 3	X3	80 ชิ้น	1,200,000 บาท

4.3 เปรียบเทียบผลลัพธ์ระหว่างวิธีการเดิมกับวิธี Fuzzy TOPSIS และ MCGP

ผู้จัดทำได้เปรียบเทียบการเลือกคู่ค้าด้วยวิธีเดิมกับวิธี Fuzzy TOPSIS และ MCGP ดังภาพที่ 3 และภาพที่ 4



ภาพที่ 3: เปรียบเทียบปริมาณการสั่งซื้อระหว่างวิธีการเดิมกับวิธี Fuzzy TOPSIS และ MCGP



ภาพที่ 4: เปรียบเทียบมูลค่าการสั่งซื้อระหว่างวิธีการเดิมกับวิธี Fuzzy TOPSIS และ MCGP

ตารางที่ 10: ผลประเมินความพึงพอใจเปรียบเทียบวิธีการเลือกคู่ค้าแบบเดิมกับการเลือกคู่ค้าโดยใช้วิธีการตัดสินแบบหลายเกณฑ์ด้วย Fuzzy TOPSIS และ MCGP

รายการประเมิน	ระดับความพึงพอใจ	
	$\tilde{X}$	แปลผล
1. วิธีการเลือกคู่ค้าแบบเดิม	3.33	ปานกลาง
2. วิธีการตัดสินแบบหลายเกณฑ์ด้วยวิธี Fuzzy TOPSIS และ MCGP	4.67	ดี
สรุปผลการประเมิน	4.00	ดี

จากตารางที่ 10 ผลประเมินความพึงพอใจเปรียบเทียบวิธีการเลือกคู่ค้าแบบเดิมกับการเลือกคู่ค้าโดยใช้วิธีการตัดสินแบบหลายเกณฑ์ด้วยวิธีฟัซซี่ท็อปซิสและวิธีการเขียนโปรแกรมหลายเป้าหมาย โดยภาพรวมมีค่าเฉลี่ย 4.00 สรุปได้ว่าผลการประเมินความพึงพอใจงานวิจัยอยู่ในเกณฑ์ ดี

5. สรุป

จากการวิจัยพบว่าวิธี Fuzzy TOPSIS ได้รับค่า CC ดังนี้ คู่ค้ารายที่ 3 = 0.72 คู่ค้า รายที่ 2 = 0.58 คู่ค้า รายที่ 3 = 0.55 นำค่า *CC* ที่ได้รับคำนวณร่วมกับวิธี MCGP ได้ค่าดังนี้ 1. จำนวนรวมการสั่งซื้อสูงสุดวิธี MCGP ได้ค่าปริมาณการสั่งซื้อที่น้อยกว่าวิธีเดิมจาก 200 ชิ้นเป็น 170 ชิ้น 2. ต้นทุนรวมการสั่งซื้อต่ำสุดวิธี MCGP ได้มูลค่าต้นทุนรวมการสั่งซื้อที่ต่ำกว่าวิธีเดิมจาก 3,056,790 บาท เป็น 2,520,000 บาท ซึ่งคิดเป็น 17.56% ระดับความพึงพอใจของผู้ใช้งานเปรียบเทียบวิธีการเดิมกับวิธีการใหม่อยู่ในระดับดี แต่อย่างไรก็ตามการหาผลลัพธ์ตามเป้าหมายที่องค์กรกำหนดยังไม่ครอบคลุมทุกเป้าหมายที่ควรให้ความสนใจ เช่น การลดระยะเวลาในการจัดส่งสินค้าล่าช้า การลดปริมาณของเสีย และงานวิจัยนี้ยังไม่ครอบคลุมการทดลองการเปลี่ยนแปลงของผลลัพธ์เมื่อค่า CC มีการเปลี่ยนแปลง เป็นต้น

เอกสารอ้างอิง

[1] กรุงเทพธุรกิจออนไลน์, “จุดฝันฮับเครื่องมือแพทย์อาเซียน,” จาก *http://www.bangkokbiznews.com,* ค้นหาเมื่อวันที่ 31 ตุลาคม 2558.

[2] บ้านเมืองออนไลน์, “ชี้ 5 ปีผลิตเครื่องมือแพทย์แทนนำเข้า,” จาก *http://www.banmuang.co.th,* ค้นหาวันที่ 31 ตุลาคม 2558.

[3] Babak Daneshvar Rouyendegh and Thomy Eko Saputro, “Supplier selection using integrated fuzzy TOPSIS and MCGP: a case study,” *In 5th World Conference on Educational Sciences WCES,* pp. 3977-3970, 2013.

[4] Chang and C.T., “Multi-choice goal programming Omega,” *The international Journal of Management Science,* Vol.35, pp. 389-396, 2007.

[5] Chin-Nung Liao and Hsing-Pei Kao, “An integrated fuzzy TOPSIS and MCGP approach to supplier selection in supply chain management,” *Expert Systems with Applications,* pp. 10803-10811, 2011.

[6] Shao-qi Zhou and Wen-bing Chang and Sheng-han Zhou1 and Wei Liu, “The method of risk evaluation for equipment development based on triangular fuzzy number and TOPSIS,” *In 26th Chinese Control and Decision Conference (CCDC),* pp. 2272-2276, 2014.

[7] O.Jadidi and T.S. Hong and F.Firouzi and R.M. Yusuff and N. Zulkifli, “TOPSIS and fuzzy multi-objective model integration for supplier selection problem,” *Journal of Achievements in Materials and Manufacturing Engineering,* vol. 31, pp. 762 – 769, 2008.

การท่องเที่ยวเสมือนจริงบนโทรศัพท์เคลื่อนที่: พุทธมณฑลอีสานของจังหวัดมหาสารคาม
Realism Travel on Mobile Phone: Isan Buddhist Diocese of Mahasarakham

ปรัชญา สุธีร์ (Pratya Sutee)[1] เจษฎา มีลา (Jetsada Meela)[2] และจันทิมา พลพินิจ (Jantima Polpinij)[3]
ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาการสารสนเทศ มหาวิทยาลัยมหาสารคาม
[1]powerultimate7@gmail.com , [2]jetsada.mee@msu.ac.th, [3]jantima.p@msu.ac.th

บทคัดย่อ

เอกสารฉบับนี้นำเสนอแอพพลิเคชั่นเพื่อการเที่ยวชมสถานที่แบบความเป็นจริงเสมือน ที่สามารถใช้ได้บนโทรศัพท์เคลื่อนที่แบบแอนดรอยด์ แนวคิดของแอพพลิเคชั่นคือ "การเดินทางท่องเที่ยวจริงผ่านโทรศัพท์เคลื่อนที่" ซึ่งแอพพลิเคชั่นนี้จะเล่นกับการควบคุมผ่านบลูทูธ ชุดกล้องสำหรับดูภาพแบบความเป็นจริงเสมือน และก้านควบคุม แอพพลิเคชั่นที่นำเสนอจะเกี่ยวข้องกับพุทธมณฑลอีสานของจังหวัดมหาสารคาม ได้แก่ พระธาตุนาดูน กู่สันตรัตน์ และกู่น้อย ซึ่งสถานที่เหล่านี้อยู่ในอำเภอนาดูน จังหวัดมหาสารคาม

คำสำคัญ: การท่องเที่ยว ความเป็นจริงเสมือน แอนดรอยด์ โทรศัพท์เคลื่อนที่ พุทธมณฑลอีสาน

Abstract

This work proposes a sightseeing application of Virtual Reality (VR) used on Android-based mobile phone. The idea behind this application is called "Realism Travel by Mobile Phone". This application plays with Bluetooth remote control, VR Glasses Headset, and joystick. The proposed application is related to Isan Buddhist Diocese of Mahasarakham such as Phra-That-Na-Doon, Ku-Santarattana, and Ku-Noi. These places are located in Na-Doon District, Mahasarakham province.

Keyword: Sightseeing Application, Virtual Reality, Isan Buddhist Diocese, Mobile phone, Android.

1. บทนำ

ความเป็นจริงเสมือน (Virtual Reality: VR) [1], [2] เป็นสถานการณ์จำลองหรือสภาพแวดล้อมจำลองที่สร้างขึ้นมาเหนือจริงด้วยการใช้คอมพิวเตอร์จำลองขึ้นมา (Computer Stimulation) ความเป็นจริงเสมือนส่วนมากจะเป็นเรื่องเกี่ยวกับสิ่งที่เกี่ยวกับการมองเห็น ที่แสดงทั้งบนจอภาพหรืออุปกรณ์แสดงผลสามมิติ โดยผู้ใช้สามารถโต้ตอบกับสิ่งแวดล้อมเสมือนจริงได้โดยการใช้อุปกรณ์เสริมมาตรฐานทั่วไป เช่น แป้นพิมพ์ หรือ เมาส์ หรือการเข้าถึงความเป็นจริงเสมือนด้วยการใช้อุปกรณ์หลายภาวะ เช่น ถุงมือรับรู้ (sensor glove) คันบังคับหลายทิศทาง, แว่นตาสามมิติ (VR glasses) ถุงมือสามมิติ (Cyber Glove) เมาส์สามมิติ (3D mouse) และคทาสามมิติ (Wanda) เป็นต้น

ปัจจุบันมีการนำเทคโนโลยี VR มาประยุกต์ใช้ในหลายๆ ด้าน เช่น เกม การศึกษา การฝึกอบรม รวมถึงด้านการท่องเที่ยว โดยเฉพาะในด้านการท่องเที่ยวได้มีการประยุกต์ใช้เทคโนโลยี VR บนแอพพลิเคชั่นต่างๆ เพื่อสร้างแรงจูงใจในการท่องเที่ยวหรือการศึกษาในแหล่งท่องเที่ยวนั้นๆ เพราะการนำเสนอแหล่งท่องเที่ยวบนแอพพลิเคชั่นในรูปแบบ VR จะช่วยทำให้กับนักท่องเที่ยวสามารถศึกษาแหล่งท่องเที่ยวที่กำลังสนใจได้เสมือนตนเองกำลังท่องเที่ยวอยู่ในสถานที่นั้นๆ ซึ่งอาจจะทำให้นักท่องเที่ยวบางคนสนใจที่จะเดินทางไปสัมผัสแหล่งท่องเที่ยวนั้นจริงๆ [3] เช่น Roller Coaster VR attraction [4] คือ แอปพลิเคชั่น VR ที่ใช้ในการเที่ยวชมสถานที่ ในรูปแบบโมเดล 3 มิติ ซึ่งผู้ใช้จะไม่สามารถเคลื่อนที่ไปมา ทำได้เพียงแค่มองไปรอบๆในขณะที่ รถไฟเหาะกำลังเคลื่อนที่ไปตามราง นอกจากนี้ใน Jurrasic VR [5] คือ แอปพลิเคชั่น VR ที่ใช้ในการเที่ยวชม

สถานที่ ในรูปแบบโมเดล 3 มิติ ซึ่งผู้ใช้สามารถเคลื่อนที่ไปมาภายในแอพพลิเคชั่น ซึ่งแอพพลิเคชั่นเกิดขึ้นใน Unity

จากเหตุผลข้างต้น งานวิจัยนี้จึงนำเสนอการประยุกต์ใช้เทคนิค VR กับแอพพลิเคชั่นบนมือถือสำหรับแนะนำสถานที่ท่องเที่ยวในไทยเพื่อเป็นการส่งเสริมการท่องเที่ยว โดยเป็นการจำลองด้วยโมเดล 3 มิติ ที่สามารถชมผ่านแว่นตา 3 มิติ และควบคุมการเคลื่อนที่ในสภาวะเหมือนจริงด้วยรีโมทแบบบลูทูธ นอกจากนี้แอพพลิเคชั่นยังมีเสียงอธิบายเกี่ยวกับสถานที่นั้นๆ ประกอบ โดยสถานที่ท่องเที่ยวที่จะแนะนำบนแอพพลิเคชั่นคือ "*พุทธมณฑลอีสาน จังหวัดมหาสารคาม*" ซึ่งประกอบไปด้วยพระธาตุนาดูน กู่สันตรัตน์ และกู่น้อย โดยสถานที่ดังกล่าวเป็นสถานที่ท่องเที่ยวเชิงวัฒนธรรมที่สำคัญของจังหวัดมหาสารคาม

2. ทบทวนวรรณกรรม

2.1 ความเป็นจริงเสมือน (virtual reality: VR)

ความเป็นจริงเสมือน [1], [2], [6] เป็นสภาพแวดล้อมที่จำลองโดยคอมพิวเตอร์ สภาพแวดล้อมในความเป็นจริงเสมือนส่วนมากจะเป็นเรื่องเกี่ยวกับ สิ่งที่เกี่ยวกับการมองเห็น แสดงทั้งบนจอคอมพิวเตอร์ หรือ อุปกรณ์แสดงผลสามมิติ โดยผู้ใช้สามารถโต้ตอบกับสิ่งแวดล้อมเสมือนได้ทั้งการใช้อุปกรณ์นำเข้ามาตรฐานเช่น แป้นพิมพ์ หรือ เมาส์ สภาพแวดล้อมจำลองยังสามารถทำให้คล้ายกับโลกจริงได้ เช่น การจำลองการฝึกนักบิน สามารถทำให้แตกต่างจากความเป็นจริงก็ได้เช่น เกมความเป็นจริงเสมือน ในทางปฏิบัติแล้วเป็นเรื่องยากมากในการสร้างประสบการณ์ความเป็นจริงเสมือนที่เหมือนจริงมากๆ เนื่องจากข้อจำกัดทางเทคนิคหรือกำลังการประมวลผล ความละเอียดของภาพ อย่างไรก็ตามข้อจำกัดดังกล่าวคาดว่าจะแก้ไขได้ในอนาคตอันใกล้เนื่องจากเทคโนโลยีการสื่อสารภาพและข้อมูล รวมถึงกำลังของหน่วยประมวลผลนั้นพัฒนาขึ้นอย่างรวดเร็วและมีประสิทธิภาพ

โดยทั่วไป ระบบ VR แบ่งตามลักษณะมีพื้นฐานบนวิธีที่ติดต่อกับผู้ใช้ ได้แก่

1) Desktop VR หรือ Window on World Systems (WoW) เป็นระบบความจริงเสมือนที่ใช้จอภาพของคอมพิวเตอร์ในการแสดงผล

2) Video Mapping เป็นการนำวิดีโอมาเป็นอุปกรณ์หรือเครื่องมือนำ เข้าข้อมูลของผู้ใช้และใช้กราฟิกคอมพิวเตอร์นำเสนอการแสดงผลในโมเดลแบบสองมิติหรือสามมิติโดยผู้ใช้จะเห็นตัวเองและเปลี่ยนแปลงตัวเองจากจอภาพ

3) Immersive Systems เป็นระบบความจริงเสมือนสำหรับผู้ใช้ส่วนบุคคล โดยผู้ใช้นำอุปกรณ์ประเภทจอภาพสวมศีรษะ (HMD) ได้แก่ กล้อง VR หรือแว่น VR มาใช้จำลองภาพและการได้ยิน

4) Telepresence เป็นระบบเสมือนจริงที่มีการนำอุปกรณ์ตรวจจับสัญญาณระยะไกลที่อาจติดตั้งกับหุ่นยนต์เชื่อมต่อการใช้งานกับผู้ใช้

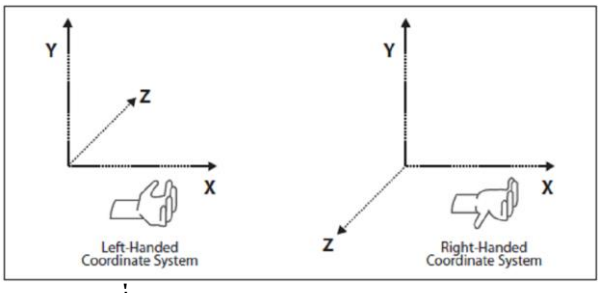
5) Augmented หรือ Mixed Reality Systems เป็นการผสมผสานระหว่าง Telepresence ระบบความจริงเสมือน และเทคโนโลยีภาพเพื่อสร้างสิ่งที่เสมือนจริงให้กับผู้ใช้

2.2 โมเดล 3 มิติและเทคนิคที่นำมาใช้

โมเดล 3 มิติ [7] คือ สิ่งของที่มีรูปทรงเป็นสามมิติ คือทั้งขนาดของ ความกว้าง (width) ความลึก (depth) และความสูง (height) โมเดลที่สร้างด้วยโปรแกรมคอมพิวเตอร์ สามารถแสดงผลให้ใกล้เคียงกับวัตถุจริงๆ โดยให้แสงและเงาที่ทำให้เห็นความลึกของมิติ Stereoscopy คือเทคนิคการสร้างภาพสามมิติเป็นเทคนิคที่เกี่ยวกับการมองของมนุษย์ ทำให้คนมองแล้วรู้สึกถึงความลึกได้ด้วยตา

ในการสร้างโมเดลจะอาศัยแนวคิดเกี่ยวกับระบบแกน 3 มิติ (3D Coordinate System) ซึ่งระบบแกน 3 มิติในคอมพิวเตอร์จะมีความคล้ายคลึงกับระบบแกน 3 มิติ ในวิชาคณิตศาสตร์หรือฟิสิกส์ คือ แกน X จะมีค่าเป็นบวกไปทางขวามือ และแกน Y จะมีค่าเป็นบวกไปทางด้านบนซึ่งมีทิศตรงข้ามกับแกน Y ในระบบแกน 2 มิติ ของจอภาพ แต่ในส่วนของแกน Z จะมี 2 แบบ ซึ่งมีทิศทางตรงกันข้าม โดยขึ้นอยู่กับประเภทของระบบแกน 3 มิติ

ระบบแกน 3 มิติด้านซ้ายมือ (Left-handed 3D Coordinate System) จะมีแกน Z ชี้ไปในทิศทางตามกฎมือซ้าย คือ แกน Z จะชี้เข้าไปในจอ หรือ มีค่าเป็นบวกในทิศลึกเข้าไปในจอ ใน DirectX จะใช้แกน 3 มิติ สำหรับระบบแกน 3 มิติด้านขวามือ (Right-handed 3D Coordinate System) จะมีแกน Z ชี้ไปในทิศทางตามกฎมือขวา คือ แกน Z จะชี้ออกมาจากจอภาพหรือ มีค่าเป็นบวกในทิศชี้ออกมาจากจอภาพ สำหรับใน XNA และ OpenGL [10]



ภาพที่ 1: การเปรียบเทียบระหว่างระบบแกน 3 มิติ [7]

เมื่อสร้างโมเดลด้วยโปรแกรมสร้างงาน 3 มิติ ให้เกิดโครงสร้างแล้ว ก็จะมีการลงรายละเอียดพร้อมตกแต่งโมเดล โดยการใส่พื้นผิวและลวดลายให้กับวัตถุที่เพิ่งสร้างตามที่ได้ออกแบบไว้ สร้างฉาก แสงเงา มุมกล้อง นอกจากนี้อาจจะต้องมีการสร้างฉากและสภาพแวดล้อมของวัตถุด้วย รวมถึงการจัดไฟแสงเงา และมุมมองให้เข้ากับฉากที่สร้างขึ้น เพื่อให้ฉากมีความสมจริงตามเรื่องราว (Story board) ที่ได้วางเอาไว้

2.3 พระธาตุนาดูน กู่สันตรัตน์ และกู่น้อย [8]

พระธาตุนาดูน เป็นพุทธมณฑลอีสานที่สร้างเมื่อ พ.ศ. 2528 เพื่อบรรจุพระบรมสารีริกธาตุขององค์สมเด็จพระสัมมาสัมพุทธเจ้า ตั้งอยู่ที่อำเภอนาดูน จังหวัดมหาสารคาม และต่อมาคณะรัฐมนตรีได้มีมติ ให้เป็นพุทธมณฑลอีสาน เมื่อ พ.ศ.2530 เพื่อเป็นศูนย์กลางส่งเสริมกิจการพระพุทธศาสนาและศิลปวัฒนธรรมของภาคตะวันออกเฉียงเหนือ



ภาพที่ 2: พระธาตุนาดูน

สำหรับกู่สันตรัตน์ ตั้งอยู่ที่ตำบลกู่สันตรัตน์ อำเภอนาดูน จังหวัดมหาสารคาม สร้างด้วยศิลาแลงเป็นศิลปะขอมแบบบายัน มีรูปลักษณะปราสาทหินที่มีปรางค์ประธานเป็นรูปสี่เหลี่ยมจัตุรัสมีมุขด้านหน้ายื่นไปทางทิศตะวันออก มีบรรณาลัยซึ่งเป็นที่เก็บคัมภีร์ทางศาสนาตั้งอยู่ทางทิศตะวันออกเฉียงใต้หันหน้าเข้าหาปรางค์ประธาน อาคารทั้ง 2 ล้อมด้วยกำแพงศิลาแลงซึ่งสร้างยังไม่เสร็จเรียบร้อยอีกชั้นหนึ่ง กู่สันตรัตน์สร้างขึ้นมาด้วยมีวัตถุประสงค์เพื่อประดิษฐานรูปเคารพสำหรับประกอบพิธีกรรมทางศาสนา และเป็นโรคยาศาล คือเป็นที่พักรักษาพยาบาลคนเจ็บป่วย



ภาพที่ 3: กู่สันตรัตน์

กู่น้อย ถูกสร้างขึ้นในราวพุทธศตวรรษ ที่ 17 (ประมาณ พ.ศ.1700) สมัยพระเจ้าชัยวรมันที่ 7 แห่งเขมร เพื่อเป็นศาสนสถานเนื่องในศาสนาฮินดูเป็นหลักฐานสำคัญทางโบราณคดีของเมืองโบราณนครจัมปาศรี ช่วงสมัยทวารวดี ลักษณะการก่อสร้างเป็นศิลปกรรมแบบปาปวน และนครวัดของเขมร มีโบราณวัตถุพบอยู่ในบริเวณกู่น้อยแห่งนี้ ได้แก่ ประติมากรรมหินทราย พระศิวะ พระนารายณ์ และเศียรทวารบาล



ภาพที่ 4: กู่น้อย

3. ขั้นตอนการดำเนินงาน

ในส่วนนี้จะอธิบายในส่วนของขั้นตอนการดำเนินงาน โดยจะแบ่งออกเป็น 4 ส่วน ที่สามารถแสดงรายละเอียดดังนี้คือ

3.1 การเก็บข้อมูล (Data Collection)

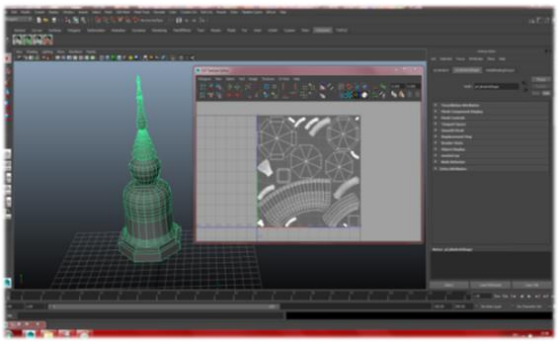
เป็นขั้นตอนของการลงพื้นที่เพื่อเก็บรวบรวมข้อมูลทั้งในรูปแบบการเก็บภาพนิ่ง ภาพวีดีทัศน์ และเอกสารที่เกี่ยวข้องกับประวัติความเป็นมาของพระธาตุนาดูน กู่สันตรัตน์ และกู่น้อย

3.2 การสร้างสตอรี่บอร์ด (Story-board Creation)

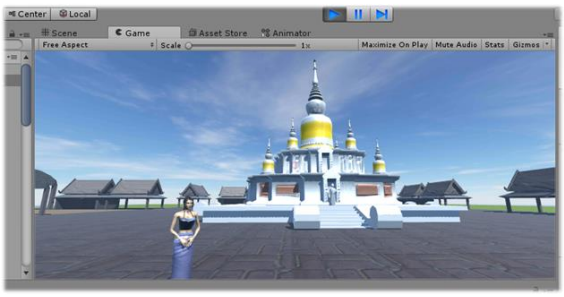
เป็นการออกแบบและสร้างเรื่องราวในแอพพลิเคชั่นที่จะนำเสนอ เพื่อให้เห็นขั้นตอนในการเล่น/ใช้แอพพลิเคชั่นการท่องเที่ยวเสมือนจริงบนโทรศัพท์เคลื่อนที่

3.3 การออกแบบและการสร้างโมเดล 3 มิติ

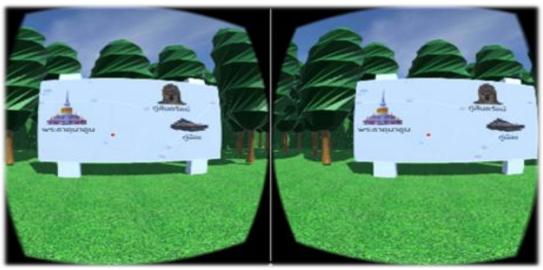
ในขั้นตอนนี้คือขั้นตอนในการสร้างโมเดล 3 มิติ ฉาก แผนที่ (Map) และผู้นำเที่ยว (Guide) ในแอพพลิเคชั่น โดยโมเดลต่างๆ ที่นำเสนอในงานวิจัยฉบับนี้ ได้พัฒนาด้วยโปรแกรม Maya 2014 โมเดลที่จะนำมาใช้งานจะเป็นแบบโพริกอนต่ำ (Low Polygon) เพื่อลดปัญหาการทำงานช้าและการกระตุกเมื่อนำไปใช้งานจริงบนโทรศัพท์เคลื่อนที่



ภาพที่ 5: ตัวอย่างโมเดล 3 มิติและการใส่ UV Texture



ภาพที่ 6: โมเดล 3 มิติของพระธาตุนาดูน



ภาพที่ 7: แผนที่ในแอพพลิเคชั่น

3.4 การพัฒนาระบบ

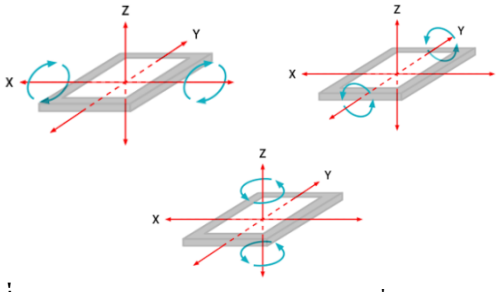
ในส่วนนี้จะอธิบายการพัฒนาระบบในส่วนที่สำคัญๆ ได้ได้แก่ การควบคุมการหมุนในระบบ VR, การใช้ Cardboard และการทำ Trigger โดยสามารถอธิบายได้ดังต่อไปนี้

1) การควบคุมการมองภาพแบบ VR

การควบคุมการมองภาพแบบ VR จะอาศัย Google Cardboard [9] ซึ่งเป็นเครื่องมือที่ใช้ในแอปพลิเคชั่นแบบ VR เพื่อเคลื่อนที่และมุมมองของกล้องในระบบ (ซึ่งมีเลนส์แบบ “เลนส์นูน” ที่จะช่วยปรับระยะภาพให้สามารถมองได้ชัดเจนเมื่อหน้าจออยู่ใกล้ๆสายตา ดังนั้นเมื่อผู้ใช้มองเข้าไปในเลนส์ระยะภาพที่เห็นที่หน้าจอก็จะชัดเจนพอดี) ซึ่ง Cardboard API สามารถเรียกใช้ จากโปรแกรม Unity3D ที่ใช้เพื่อการพัฒนาแอพพลิเคชั่น

2) การควบคุมการหมุนในระบบ VR

ในการเคลื่อนที่ของมุมกล้องในระบบ VR จะใช้อนาล็อกของคอนโทรลเลอร์ ซึ่งระบบจะทำการตรวจสอบทิศทางที่ผู้ใช้ได้ควบคุม เพื่อนให้มุมกล้องเคลื่อนที่ไปในทิศทางที่ผู้ใช้ต้องการ ถ้าดูจากภาพที่ 8 แกน x และ y จะพบว่าการหมุนในแกน y ก็คือการเอียงเครื่องในแกน x และการหมุนในแกน x ก็คือการเอียงเครื่องในแกน x แต่ความแตกต่างอยู่ที่แกน z ถ้าเป็นของ Accelerometer แกน z จะมีการเปลี่ยนแปลงไปตามแกน x และแกน y แต่ Gyroscope การหมุนแกน z จะเป็นอิสระจากแกนอื่น

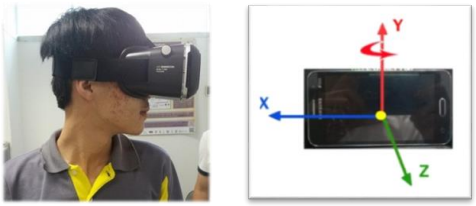


ภาพที่ 8: การแสดงการทำงานเซ็นเซอร์โทรศัพท์ที่เป็นแบบ Gyroscope

สำหรับการหมุนมุมกล้องในระบบ VR จะใช้ Gyro Sensor โทรศัพท์ ซึ่งระบบจะทำการตรวจสอบทิศทางที่ผู้ใช้สัมผัสจากค่า *x*, *y*, *z* เพื่อใช้ในการหมุนมุมกล้องจากจุดกึ่งกลางหน้าจอโทรศัพท์ไปตามที่ผู้ใช้ต้องการ เครื่องมือสำคัญบนโทรศัพท์เคลื่อนที่ที่ช่วยในการทำงานของระบบ VR คือ Gyroscope Sensor [10] ซึ่งเป็นเซ็นเซอร์ที่มีไว้สำหรับตรวจจับลักษณะการหมุนของสมาร์ทโฟน โดยเป็นการตรวจจับแบบ 3 แกน (3-Axes) เช่นเดียวกันกับ Accelerator Sensor แต่จะมีความถูกต้อง และแม่นยำมากกว่า



ภาพที่ 9: ตัวอย่างการมองในระบบ VR และแกนค่า *x*, *y*, *z*



ภาพที่ 10: ตัวอย่างการหันด้านซ้ายในระบบ VR และการหมุนแกน *y*

3) การสร้าง Trigger

Trigger [11] คือเครื่องมือที่ Engine นั้นใช้ในการตรวจสอบว่าผู้ใช้ได้กระทำสิ่งใด ทำให้ผู้พัฒนาสามารถใส่การการทำงานของระบบในส่วนๆนั้นได้อย่างแม่นยำ Trigger นั้นมีหลายรูปแบบ มีแต่ตั้งแต่ จุดเล็กๆไปจนถึง พื้นที่ขนาดใหญ่ และมีการตรวจสอบการทำงานในหลายๆ รูปแบบ เช่น การมองไปที่ Trigger จะทำให้จุดโฟกัสและ Trigger นั้นชนกันหรือการเคลื่อนที่ ไปยัง Trigger นั้นจะทำให้ตัวละครที่ผู้ใช้บังคับชนกับ Trigger ภายใน Engine จึงทำให้ผู้พัฒนาสามารถใส่การทำงานของสิ่งต่างๆ ได้เมื่อ Trigger ทำงาน ซึ่งในแอพพลิเคชั่นนี้จะใช้การมองของผู้เล่นเป็นหลักในการอธิบายข้อมูลและสถานที่ต่างๆ ซึ่งผู้ใช้นั้นไม่สามารถมองเห็น Trigger ได้ แต่ผู้ใช้สามารถรับรู้ได้ว่ามี Trigger โดยผู้สร้างจะทำสัญลักษณ์ไว้ เมื่อจุดโฟกัสนั้นไม่ได้อยู่ในตำแหน่งเดียวกับ Trigger ที่เรามี จุดโฟกัสนั้นจะเป็นสีแดง ดังภาพที่ 11



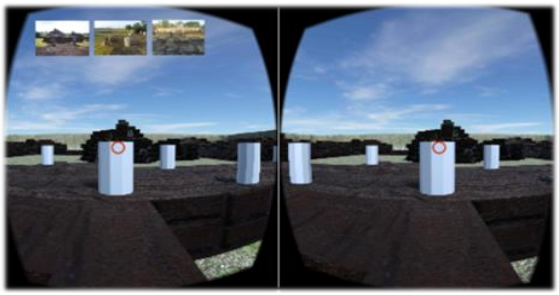
ภาพที่ 11: จำลองรูปแบบ Trigger

ในการออกแบบสำหรับงานนี้ เมื่อผู้ใช้หันหน้าไปในทิศทางที่มี Trigger โดยที่ Trigger จะไม่สามารถมองเห็นได้จะมีลักษณะโปร่งใสและมีจุดโฟกัสตรงกับ Trigger (โดยสังเกตที่จุดโฟกัสขยายขึ้น) แสดงว่าผู้ใช้สามารถเรียกใช้ Trigger นั้นๆ ได้ (เช่น การแสดงเสียงพูดเพื่ออธิบายเกี่ยวกับตำแหน่งนั้นๆ)

สำหรับในแอพพลิเคชั่นจริง รูปแบบการเคลื่อนที่แบบ VR ไปยัง Trigger จะใช้ระบบ Gyroscope Sensor ในการควบคุมทิศทางไปยังตำแหน่งที่มี Trigger อยู่ ส่วนการควบคุมการเดินจะใช้ Joy Strick เป็นตัวควบคุมการเคลื่อนที่ไปยังตำแหน่งต่างๆ ที่ต้องการ (ดังภาพที่ 12-13)



ภาพที่ 12: แสดงจุดโฟกัสที่ตรงกับ Trigger



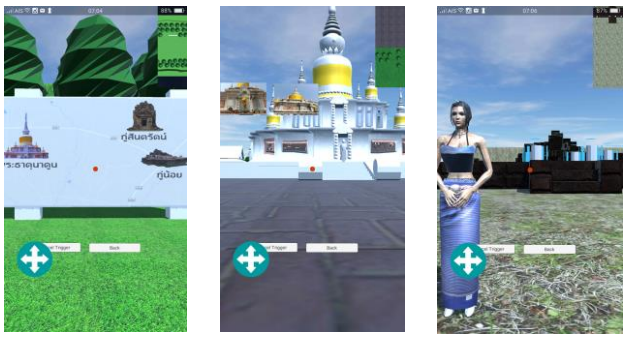
ภาพที่ 13: รูปแบบการมอง Trigger ในแอพพลิเคชั่น

4. การทดสอบการใช้งานแอพพลิเคชั่น

ระบบการท่องเที่ยวเสมือนจริงบนโทรศัพท์เคลื่อนที่เพื่อศึกษาพุทธมณฑลอีสานของจังหวัดมหาสารคาม ที่พัฒนาขึ้นสามารถเล่นบนโทรศัพท์เคลื่อนที่แบบแอนดรอยส์ (Android) ที่สามารถเล่นด้วยการใช้อุปกรณ์เสริมคือ แว่น VR และรีโมทคอลโทรลแบบบลูทูธ (Bluetooth Remote Control)

โดยระบบฯ นี้จะมี 2 สองคือ โหมดการเล่นที่ไม่ใช่ระบบ VR และโหมด VR ซึ่งแสดงสถานที่ท่องเที่ยงหลักๆ 3 แห่งใน

จังวัดมหาสารคามคือ พระธาตุนาดูน กู่สันตรัตน์ และกู่น้อย ตัวอย่างภาพในระบบฯ สามารถแสดงในภาพที่ 14



ภาพที่ 14: ตัวอย่างภาพในระบบการท่องเที่ยวเสมือนจริงบนโทรศัพท์เคลื่อนที่เพื่อศึกษาพุทธมณฑลอีสาน จ. มหาสารคาม

ในการทดสอบการใช้งานจากผู้ใช้งานที่มีความแตกต่างในเรื่องเพศและอายุ โดยการเลือกตัวอย่างสุ่มอย่างง่าย (Simple Random Sampling: SRS) แบบวิธีจับฉลาก (Lottery Method) ซึ่งในงานนี้จะเลือกตัวอย่างหรือตัวแทนจำนวน 30 คนจากประชากรขนาด 50 คนที่มีประสบการณ์ในการเล่มเกมบนโทรศัพท์เคลื่อนที่และการใช้ Joy Stick ซึ่งผู้ทดสอบ 1 คนจะใช้เวลาในการเล่น 30 นาที ภายหลังจากการทดสอบพบว่ามีความพึงพอใจในการใช้งานเฉลี่ยอยู่ที่ร้อยละ 88.06 โดยความพึงพอใจมากที่สุดคือ ด้านความคิดสร้างสรรค์และการประยุกต์เทคโนโลยีคิดเป็นร้อยละ 95.77 ลำดับที่สองคือการออกแบบองค์ประกอบในแอพพลิเคชันคิดเป็นร้อยละ 85.65 และลำดับที่สามคือการใช้งานง่าย (User friendly) คิดเป็นร้อยละ 82.78

5. สรุป

เอกสารฉบับนี้นำเสนอการพัฒนาจำลองสถานที่ท่องเที่ยวความเป็นจริงเสมือนในรูปแบบ 3 มิติ ที่เรียกว่า “Realism Travel by Mobile Phone” ซึ่งมีรูปแบบจำลองสถานที่ท่องเที่ยวเป็นแบบผสมผสานระหว่างการเดินเที่ยวชมบริเวณโดยรอบบริเวณสถานที่ท่องเที่ยว และการเรียนรู้ศึกษาความเป็นมาของสถานที่ท่องเที่ยว ที่จะทำให้ผู้เล่นเข้าไปอยู่ในสภาพแวดล้อมเหมือนจริงที่จะได้ทั้งความรู้และความเพลิดเพลินในสถานที่ท่องเที่ยว ภายในระบบจะมีระบบมัคคุเทศก์เพื่อแนะนำสถานที่ท่องเที่ยว ระบบมัคคุเทศก์จัดทำขึ้นเพื่อให้แนะนำสถานที่ท่องเที่ยว ระบบที่พัฒนาขึ้นจะสามารถเล่นบนโทรศัพท์เคลื่อนที่แบบแอนดรอยส์ที่สามารถเล่นด้วยการใช้อุปกรณ์เสริมคือแว่น VR และรีโมทคอลโทรลแบบบลูทูธ

เอกสารอ้างอิง

[1] John Vince, *Introduction to Virtual Reality*. 1st ed. Springer, 2004.

[2] Gerard Kim, *Designing Virtual Reality Systems: The Structured Approach*, Springer, 2005.

[3] D. Cunningham and M. Krishack, “Virtual reality: a holistic approach to rehabilitation”, *Stud Health Technol Inform*, vol. 62, pp. 90-93, 1999.

[4] E. Grabianowski. “How Brain-computer Interfaces Work”. 2554 [Accessed on 20 January 2017]. http://computer.howstuffworks.com/brain-computer-interface.html.

[5] Ampaipp.ธุรกิจท่องเที่ยวใช้เทคโนโลยี VR โปรโมทสถานที่ก่อนลูกค้าซื้อทัวร์.[อินเทอร์เน็ต].22 มิถุนายน 2558. [เข้าถึงเมื่อ 10 พฤศจิกายน 2558]. https://www.blognone.com/node/69624.

[6] G. Girraldi, R. Silva, and J.C. de Oliveira, “Introduction to Virtual Reality”, LNCC Research Report, National Laboratory for Scientific Computation, Quintadinha, Brazil, vol. 6, 2003.

[7] S. Schnittert, R. Winz, E. von Lieres, “Development of a 3D Model for Packed Bed Liquid Chromatography in Micro-columns”, *The 3rd UKSim European Symposium on Computer Modeling and Simulation*, 2009.

[8] ศูนย์ข้อมูลกลางทางวัฒนธรรม.กู่น้อย. [อินเทอร์เน็ต].6 มกราคม 2555.[เข้าถึงเมื่อ 16 กันยายน 2559]. http://www.mculture.in.th/moc_new/album/120619.

[9] H. Karaouni and A. Mago, “Accessible Positional Tracking for Mobile VR”, Accessed on 15 January 2017, http://stanford.edu/class/ee267/Spring2016/report_mago_karaouni.pdf

[10] A. Broquetas, A. Comerón, A. Gelonch, J.M. Fuertes, J.A. Castro, D. Felip, M.A. López, and J.A. Pulido, “Track Detection in Railway Sidings Based on MEMS Gyroscope Sensors”, *Sensors*, vol. 12, pp. 16228-16249, 2012.

[11] A. Hanzlik, “Simulation-based Application Software Development in Time-Triggered Communication Systems”, *International Journal of Software Engineering & Applications* (*IJSEA*), vol.4, no.2, 2013.

หลักฐานเชิงประจักษ์เกี่ยวกับความสามารถในการบำรุงรักษาซอฟต์แวร์ในโครงการโอเพนซอร์ส

The Empirical Evidence on Software Maintainability in Open Source Projects

ธัญญรัตน์ กิจพาณิชย์ (Thanyarat Kitpanich)[1] และอซีส นันทอมรพงศ์ (Aziz Nanthaamornphong)[2]
สาขาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีและสิ่งแวดล้อม มหาวิทยาลัยสงขลานครินทร์ วิทยาเขตภูเก็ต
[1]thanyarat.ki@phuket.psu.ac.th, [2]aziz.n@phuket.psu.ac.th

บทคัดย่อ

ในหลายปีที่ผ่านมาซอฟต์แวร์โอเพนซอร์สได้ถูกนำไปใช้งานกันอย่างกว้างขวาง อย่างไรก็ตามการพัฒนาในโครงการโอเพนซอร์สยังพบปัญหาเกี่ยวกับคุณภาพของซอฟต์แวร์ เช่น ความมั่นคงและความสามารถในการบำรุงรักษา โดยทั่วไปวิศวกรซอฟต์แวร์จะให้ความสำคัญกับทางด้านความสามารถในการบำรุงรักษาซอฟต์แวร์เพราะเป็นคุณลักษณะที่สามารถลดค่าใช้จ่ายและช่วยเพิ่มผลผลิตของการพัฒนาซอฟต์แวร์ เพื่อที่จะทำความเข้าใจวิธีการที่นักพัฒนาในโครงการโอเพนซอร์สใช้ปรับปรุงซอร์สโค้ดทางด้านการบำรุงรักษาซอฟต์แวร์ งานวิจัยนี้มีจุดมุ่งหมายที่จะตอบคำถามวิจัย คือ นักพัฒนาให้ความสนใจกับความสามารถในการบำรุงรักษาภายใต้การตรวจทานโค้ดของโครงการโอเพนซอร์สหรือไม่? คณะผู้วิจัยได้วิเคราะห์ข้อเสนอแนะเกี่ยวกับการบำรุงรักษา โดยตรวจสอบจากโครงการโอเพนซอร์ส Eclipse และ Qt จากผลการศึกษาพบว่า ค่าเฉลี่ยของจำนวนข้อเสนอแนะที่นักพัฒนาในโครงการโอเพนซอร์สได้ทำการแก้ไขโค้ดด้านการบำรุงรักษามีเพียงร้อยละ 21.15 และ 10.73 ตามลำดับ ซึ่งถือว่าเป็นจำนวนที่ค่อนข้างน้อย อย่างไรก็ตามนักพัฒนามีแนวโน้มให้ความสนใจกับการแก้ไขปัญหาด้านคุณภาพของซอร์สโค้ดเพิ่มมากขึ้นทุกปี คณะผู้วิจัยคาดหวังว่างานวิจัยนี้จะเป็นการเพิ่มหลักฐานเชิงประจักษ์เกี่ยวกับคุณภาพของโครงการซอฟต์แวร์โอเพนซอร์ส โดยเฉพาะทางด้านการบำรุงรักษาซอฟต์แวร์

คำสำคัญ: วิศวกรรมซอฟต์แวร์ ซอฟต์แวร์โอเพนซอร์ส การบำรุงรักษาซอฟต์แวร์ การตรวจทานโค้ด

Abstract

Recently, open source software (OSS) applications have been widely adopting. However, the OSS projects have problems in the quality of the software, such as security and maintainability. Generally, software engineers focus on the software maintainability because this quality attribute can reduce the cost and increase the productivity of software development. To better understand how the OSS developers improve the source code based on software maintenance; this research aims to answer this main question: "Are the developers interested in the maintainability under the peer code review of the open source software projects?" We then analyzed the code review's comments related the maintenance by examining Eclipse and Qt projects. We found that the average numbers of maintenance based comments that were changed by the OSS developers are 21.15% and 10.73%, respectively. These amounts are considered slightly low. However, the developers tend to pay more attention to the software maintainability. Finally, we expect that this research will increase the empirical evidence about the quality of OSS projects, particularly maintainability.

Keyword: Software Engineering, Open Source Software, Software Maintenance, Code Review.

1. บทนำ

ในปัจจุบันองค์กรและอุตสาหกรรมเชิงพาณิชย์ทั่วโลกได้มีการนำซอฟต์แวร์โอเพนซอร์สมาประยุกต์ใช้งานกันอย่างกว้างขวาง เพราะซอฟต์แวร์โอเพนซอร์สได้รับการพัฒนาโดย

วิศวกรซอฟต์แวร์ที่มีความเชี่ยวชาญและประสบการณ์ทางด้านการพัฒนาระบบ จึงทำให้ซอฟต์แวร์โอเพนซอร์สค่อนข้างได้รับความนิยมและความน่าเชื่อถือในแง่ความถูกต้องของฟังก์ชันการทำงาน อีกทั้งยังช่วยลดค่าใช้จ่ายในการลงทุนด้านซอฟต์แวร์ แต่จากงานวิจัยก่อนหน้านี้ [1] แสดงให้เห็นว่าการพัฒนาซอฟต์แวร์ในโครงการโอเพนซอร์สยังขาดกระบวนการหรือขั้นตอนการดำเนินงานอย่างเป็นระบบและขาดเอกสารเกี่ยวกับการพัฒนาระบบ เช่น เอกสารความต้องการเชิงซอฟต์แวร์ การออกแบบ และการทดสอบ ด้วยสาเหตุนี้จึงทำให้โครงการซอฟต์แวร์โอเพนซอร์สบางโครงการมีคุณภาพของซอร์สโค้ดไม่ดีเท่าที่ควร [2] โดยเฉพาะคุณภาพของซอร์สโค้ดทางด้านความมั่นคงและการบำรุงรักษาซอฟต์แวร์

การบำรุงรักษาซอฟต์แวร์เป็นหนึ่งในกระบวนการที่สำคัญของการพัฒนาซอฟต์แวร์ เนื่องจากนักพัฒนาส่วนใหญ่ต้องใช้เวลาประมาณ 40-50% เพื่อค้นหาข้อบกพร่องและข้อผิดพลาดที่เกิดขึ้นระหว่างการพัฒนาซอฟต์แวร์ [3] หรือหลังจากส่งมอบ อีกทั้งต้องเสียค่าใช้จ่ายในกิจกรรมการบำรุงรักษาสูงถึง 40%-80% (เฉลี่ย 60%) [4] ดังนั้นเพื่อที่จะแก้ปัญหาทางด้านคุณภาพของซอฟต์แวร์ โครงการซอฟต์แวร์โอเพนซอร์สหลายโครงการ [5] ได้มีการกำหนดแนวทางในการพัฒนาและปรับปรุงโค้ด โดยโค้ดที่มีการปรับปรุงแก้ไขจากนักพัฒนาจะต้องได้รับการตรวจทานจากผู้ตรวจทาน ซึ่งก็คือนักพัฒนาที่ไม่ใช่ผู้แก้ไขโค้ดนั้นเอง กระบวนการนี้ถูกเรียกว่า การตรวจทานโค้ด (Peer code review หรือเรียกอีกชื่อว่า Modern code review)

การตรวจทานโค้ดถือว่าเป็นส่วนที่สำคัญของกระบวนการพัฒนาซอฟต์แวร์เพราะเป็นวิธีการหนึ่งที่ได้รับการยอมรับทางด้านวิศวกรรมซอฟต์แวร์ในด้านการควบคุมคุณภาพของซอฟต์แวร์ [6] โดยสิ่งที่สำคัญของการตรวจทานโค้ด คือ ข้อเสนอแนะ (Comment) ที่ได้รับจากผู้ตรวจทานโค้ด ซึ่งข้อเสนอแนะดังกล่าวจะชี้ให้เห็นถึงข้อบกพร่องที่ค้นพบในซอร์สโค้ด พร้อมทั้งให้คำแนะนำสำหรับวิธีการแก้ไขปัญหาที่จะช่วยให้นักพัฒนาสามารถแก้ไขโค้ดให้มีคุณภาพที่สูงขึ้นก่อนที่จะถูกนำไปใช้งาน รวมถึงช่วยทำให้ซอร์สโค้ดเป็นมาตรฐานที่ทุกคนสามารถอ่านและทำความเข้าใจเกี่ยวกับระบบเพื่อให้สาธารณชนนำไปพัฒนาต่อยอดในธุรกิจเชิงพาณิชย์ได้

จากการทบทวนวรรณกรรมที่เกี่ยวข้องกับการตรวจทานโค้ดในโครงการโอเพนซอร์ส คณะผู้วิจัยพบว่า ปัจจุบันยังไม่มีงานวิจัยที่ศึกษาว่า นักพัฒนาให้ความสนใจกับความสามารถในการบำรุงรักษาซอฟต์แวร์ (Software Maintainability) ภายใต้การตรวจทานโค้ดในโครงการโอเพนซอร์สหรือไม่ ดังนั้นงานวิจัยนี้คณะผู้วิจัยมีความต้องการที่จะวิเคราะห์ข้อเสนอแนะที่เกิดขึ้นในกระบวนการตรวจทานโค้ด โดยมีวัตถุประสงค์ ดังนี้

1) ศึกษาความสัมพันธ์ระหว่างข้อเสนอแนะที่เกี่ยวข้องกับการบำรุงรักษาที่ได้รับจากผู้ตรวจทานโค้ดกับการแก้ไขปรับปรุงโค้ดทางด้านความสามารถในการบำรุงรักษา

2) ค้นหาคุณลักษณะทางด้านความสามารถในการบำรุงรักษาซอฟต์แวร์ที่นักพัฒนาให้ความสนใจเป็นพิเศษ

โดยข้อมูลของโครงการซอฟต์แวร์โอเพนซอร์สที่คณะผู้วิจัยเลือกนำมาศึกษามีด้วยกัน 2 โครงการ นั่นคือ Eclipse (www.eclipse.org) และ Qt (www.qt.io)

คณะผู้วิจัยคาดว่าผลที่ได้จากงานวิจัยนี้จะเป็นการเพิ่มหลักฐานเชิงประจักษ์เกี่ยวกับคุณภาพทางด้านการบำรุงรักษาของโครงการซอฟต์แวร์โอเพนซอร์สแก่ชุมชนนักวิจัยทางด้านวิศวกรรมซอฟต์แวร์ที่ให้ความสนใจในโครงการซอฟต์แวร์โอเพนซอร์ส และเป็นการเสนอแนะแนวทางแก่นักพัฒนาซอฟต์แวร์ในโครงการโอเพนซอร์สได้ตระหนักถึงความสำคัญของการบำรุงรักษาซอฟต์แวร์ ก่อนที่จะทำการแก้ไขหรือปรับปรุงคุณภาพของซอร์สโค้ด

2. ทฤษฎีและวรรณกรรมที่เกี่ยวข้อง

ในส่วนนี้จะเป็นการอธิบายทฤษฎีและวรรณกรรมที่เกี่ยวข้อง เพื่อเป็นความรู้พื้นฐานที่จะช่วยให้เข้าใจงานวิจัยนี้ได้ดียิ่งขึ้น โดยมีหัวข้อดังต่อไปนี้

2.1 การตรวจทานโค้ดในโครงการซอฟต์แวร์โอเพนซอร์ส

ซอฟต์แวร์โอเพนซอร์สเป็นซอฟต์แวร์ที่อนุญาตให้ผู้ใช้งานหรือนักพัฒนาสามารถเข้าถึงแหล่งเก็บข้อมูลซอร์สโค้ด (Code Repository) [7] เพื่อที่จะทำการแก้ไขปรับปรุงซอร์สโค้ดให้มีคุณภาพที่สูงขึ้น ด้วยเหตุนี้ทางชุมชนวิศวกรในโครงการโอเพนซอร์สจึงได้มีการใช้การตรวจทานโค้ดเป็นสื่อกลางในการติดต่อสื่อสารและแลกเปลี่ยนความรู้เกี่ยวกับการพัฒนาซอฟต์แวร์ โดยมีวัตถุประสงค์ที่สำคัญคือ การตรวจสอบ

เพื่อให้แน่ใจว่าการเปลี่ยนแปลงโค้ดจะทำให้เกิดข้อบกพร่องลดน้อยลงและไม่ส่งผลต่อการบำรุงรักษาในระยะยาว สำหรับงานวิจัยนี้ คณะผู้วิจัยได้ทำการศึกษาข้อเสนอแนะที่ได้มาจากเครื่องมือในการตรวจทานโค้ดที่มีชื่อว่าเกอริต (Gerrit) ซึ่งเป็นระบบที่ให้บริการเกี่ยวกับกระบวนการตรวจทานโค้ด (www.gerritcodereview.com) ที่สามารถใช้งานได้อย่างอิสระและไม่เสียค่าใช้จ่าย อีกทั้งยังเป็นคลังเก็บข้อมูลเกี่ยวกับข้อเสนอแนะของโครงการซอฟต์แวร์โอเพนซอร์สจำนวนมาก

เนื่องจากโครงการโอเพนซอร์สมีแหล่งข้อมูลที่สามารถเข้าถึงได้ง่ายจึงทำให้งานวิจัยหลายงานได้นำโครงการโอเพนซอร์สมาศึกษาในแง่มุมต่างๆ ได้แก่ งานวิจัยของ Baysal และคณะ [8] ได้ทำการศึกษาปัจจัยที่ส่งผลต่อการถูกปฏิเสธการแก้ไขข้อบกพร่องของโปรแกรม (Patch) ในขณะที่ Tao และคณะ [9] ได้นำเสนอแนวทางที่จะช่วยให้นักพัฒนาเขียนโปรแกรมที่ใช้แก้ไขข้อบกพร่องของซอฟต์แวร์ให้ได้รับการยอมรับโดยผู้ตรวจทานโค้ด สำหรับงานวิจัยที่ศึกษาเกี่ยวกับการตรวจทานโค้ด ได้แก่ Bosu และคณะ [10] ได้ทำการวิเคราะห์ข้อเสนอแนะของโครงการของไมโครซอฟต์ 5 โครงการที่มีข้อเสนอแนะจำนวน 1.5 ล้านที่ได้จากเครื่องมือการตรวจทานโค้ดที่ชื่อว่า CodeFlow โดยผลจากงานวิจัยนี้จะเป็นจำแนกลักษณะของข้อเสนอแนะที่มีประโยชน์ เพื่อที่จะช่วยให้นักพัฒนาสามารถทำการแก้ไขปรับปรุงโค้ดตามคำแนะนำที่เป็นประโยชน์ ส่วนงานวิจัยของ Jacek และคณะ [11] ได้ทำการวิเคราะห์ข้อเสนอแนะที่ได้รับจากเครื่องมือ CodeFlow ซึ่งจากการศึกษาพบว่า จำนวนข้อเสนอแนะทั้งหมดเกี่ยวกับการบำรุงรักษามีจำนวนมากถึง 50%

นอกจากงานวิจัยข้างต้น คณะผู้วิจัยพบว่างานวิจัยส่วนใหญ่จะให้ความสำคัญกับกระบวนการพัฒนาซอฟต์แวร์และการบำรุงรักษาโดยมุ่งเน้นการศึกษาที่เกี่ยวข้องกับการค้นหาข้อบกพร่องที่เกิดขึ้นระหว่างการพัฒนาซอฟต์แวร์หรือสาเหตุที่ทำให้เกิดปัญหาทางด้านคุณภาพของซอร์สโค้ด [12] เพื่อที่จะร่วมมือกันแก้ไขข้อผิดพลาดก่อนส่งมอบซอฟต์แวร์ให้แก่ลูกค้าหรือนำไปใช้งานจริง

2.2 ความสามารถในการบำรุงรักษา (Maintainability)

ในงานวิจัยนี้ คณะผู้วิจัยสนใจที่จะทำการศึกษาเกี่ยวกับความสามารถในการบำรุงรักษา ซึ่งเป็นหนึ่งในคุณลักษณะที่มีความสำคัญต่อการเพิ่มคุณภาพของซอฟต์แวร์ และช่วยลดค่าใช้จ่ายในการบำรุงรักษาซอฟต์แวร์ โดย ISO/IEC 25010 [13] ได้นิยามคุณลักษณะ “ความสามารถในการบำรุงรักษา” ไว้ว่าเป็น “คุณลักษณะที่มีความสามารถในการแก้ไขเพื่อปรับปรุงโค้ดให้ดีขึ้น การปรับเปลี่ยนหรือเปลี่ยนแปลงในสภาพแวดล้อมและตามความต้องการของซอฟต์แวร์ที่ได้กำหนดไว้” สาเหตุที่ทางคณะผู้วิจัยได้เลือกศึกษา ISO/IEC 25010 เพราะเป็นมาตรฐานทางด้านคุณภาพของผลิตภัณฑ์ซอฟต์แวร์ที่องค์กรธุรกิจทั่วโลกให้ความสำคัญ โดยจะมุ่งเน้นทางด้านระบบและความต้องการเชิงคุณภาพของซอฟต์แวร์ รวมถึงการประเมินผลคุณภาพของซอฟต์แวร์ ซึ่ง ISO/IEC 25010 ได้กำหนดคุณลักษณะย่อยของความสามารถในการบำรุงรักษาด้วยกัน 5 คุณลักษณะดังนี้

1) โปรแกรมเชิงโครงสร้างโมดูล (Modularity) เป็นคุณลักษณะที่เมื่อมีการเปลี่ยนแปลงส่วนประกอบย่อยหนึ่งๆ จะทำให้มีผลกระทบต่อส่วนประกอบอื่นๆ น้อยที่สุด

2) ความสามารถในการใช้งาน (Reusability) เป็นคุณลักษณะที่สามารถนำส่วนประกอบย่อยของระบบเดิมมาใช้ในการสร้างระบบใหม่

3) ความสามารถในการวิเคราะห์ (Analysability) เป็นคุณลักษณะของการเปลี่ยนแปลงการทำงานของระบบ รวมถึงวิเคราะห์เพื่อหาข้อบกพร่องหรือสาเหตุของความล้มเหลว และระบุส่วนที่ต้องทำการแก้ไขปรับปรุง

4) ความสามารถในการแก้ไข (Modifiability) เป็นคุณลักษณะของระบบที่สามารถแก้ไขโค้ด โดยปราศจากข้อบกพร่องหรือเป็นการลดคุณภาพของซอฟต์แวร์ที่มีอยู่

5) ความสามารถในการทดสอบ (Testability) เป็นคุณลักษณะของประสิทธิภาพและประสิทธิผลของซอฟต์แวร์และองค์ประกอบในซอฟต์แวร์ที่จะถูกทดสอบ

คณะผู้วิจัยจะนำคุณลักษณะทั้ง 5 ประเภทที่ได้กล่าวไว้ข้างต้นไปเป็นชุดคำหลักเบื้องต้นในการดำเนินงานวิจัย เพื่อค้นหาข้อเสนอแนะที่เกี่ยวข้องกับการบำรุงรักษาซอฟต์แวร์โดยรายละเอียดจะกล่าวถึงในหัวข้อที่ 3

3. ขั้นตอนวิธีการดำเนินการวิจัย

การศึกษาวิจัยนี้เป็นการวิเคราะห์ข้อเสนอแนะที่เกิดขึ้นในกระบวนการตรวจทานโค้ด โดยอาศัยเทคนิคการทำเหมืองข้อความ (Text Mining) ซึ่งเป็นกระบวนการเพื่อค้นหาความรู้ในฐานข้อมูลเอกสารหรือหาความสัมพันธ์ที่ซ่อนอยู่ในชุดข้อความนั้น ซึ่งมีขั้นตอนการดำเนินงานดังนี้

3.1 การพัฒนาซอฟต์แวร์ที่ช่วยจัดการข้อมูล

ในงานวิจัยนี้ คณะผู้วิจัยได้ทำการสำรวจแหล่งเก็บข้อมูลเกี่ยวกับข้อเสนอแนะของโครงการโอเพนซอร์สที่สามารถเข้าถึงข้อมูลได้ง่าย โดยคณะผู้วิจัยได้เลือกที่จะศึกษาแหล่งเก็บข้อมูลการตรวจทานโค้ดที่ชื่อว่า ระบบเกอริต ซึ่งมีข้อมูลของโครงการ Qt และ Eclipse ถูกจัดเก็บอยู่ในระบบ สาเหตุหลักที่เลือกศึกษาโครงการโอเพนซอร์สดังกล่าว เพราะทั้ง 2 โครงการเป็นโครงการที่ยังได้รับการพัฒนาอยู่ตลอดจนถึงปัจจุบัน อีกทั้งยังเป็นโครงการที่ได้รับความสนใจและนิยมนำมาศึกษาในงานวิจัยทางด้านวิศวกรรมซอฟต์แวร์

สำหรับการสืบค้นและจัดการข้อมูลเกี่ยวกับข้อเสนอแนะของโครงการโอเพนซอร์ส ทางคณะผู้วิจัยได้ทำการพัฒนาแอปพลิเคชันด้วยภาษาจาวา เพื่อสืบค้นข้อมูลจากระบบเกอริตแบบอัตโนมัติ โดยมีกำหนดเงื่อนไข เพื่อดึงข้อมูลในช่วงเวลาที่ต้องการศึกษา นั่นคือ ข้อเสนอแนะของโครงการ Qt และ Eclipse ในช่วงปี 2012 จนถึง 2016 จากระบบเกอริตผ่าน REST API และจัดเก็บข้อมูลลงฐานข้อมูล MySQL

3.2 การวิเคราะห์ข้อมูลจากระบบเกอริต

สำหรับการวิเคราะห์ข้อมูลจะมี 4 ขั้นตอนย่อยดังต่อไปนี้

1) การสร้างคำหลัก (Keyword) คณะผู้วิจัยได้ทำการสร้างชุดคำหลัก โดยนำมาจาก ISO/IEC 25010 ที่มีการระบุถึงคุณลักษณะของ “ความสามารถในการบำรุงรักษา” ซึ่งมีด้วยกัน 5 คุณลักษณะที่ได้มีการอธิบายไว้ในหัวข้อก่อนหน้านี้ ได้แก่ Modularity Reusability Analysability Modifiability และ Testability ในเบื้องต้นคณะผู้วิจัยจะใช้คำหลักทั้ง 5 คำนี้มาทำการวิเคราะห์ เนื่องจากคำหลักเหล่านี้เป็นคุณลักษณะที่จะส่งผลคุณภาพของซอฟต์แวร์ โดยปกติแล้วประโยคในข้อเสนอแนะของผู้ตรวจทานโค้ดจะไม่มีการใช้คำที่ตรงตามคำหลักที่คณะผู้วิจัยได้ระบุ ดังนั้นจึงต้องมีการจัดกลุ่มคำที่มีความหมายเหมือน (Synonyms) กับคำหลักแต่ละคำ โดยใช้ซอฟต์แวร์ word-net [14] และพจนานุกรมภาษาอังกฤษในการค้นหาคำเหล่านี้จากฐานข้อมูลที่รวบรวมคำศัพท์ภาษาอังกฤษ

2) การทำความสะอาดข้อมูล (Data Cleaning) ในขั้นตอนนี้จะเป็นการลบข้อความที่ไม่มีประโยชน์ในการค้นหาข้อเสนอแนะเกี่ยวกับการบำรุงรักษาออกจากฐานข้อมูล เช่น คำหยุด (Stop-word) โดยคณะผู้วิจัยจะใช้โปรแกรม R ซึ่งเป็นโปรแกรมที่ใช้ในการประมวลผลข้อมูลเชิงสถิติที่มี text mining (tm) package ในการทำความสะอาดข้อมูลและทำการแปลงประโยคหรือคำให้อยู่ในรูปแบบรากศัพท์ (Stemming) เพื่อลดเวลาสำหรับการประมวลผลในการค้นหาคำหลัก

3) ค้นหาคำหลักที่ปรากฏอยู่ในข้อเสนอแนะ (Keywords Extraction) สำหรับขั้นตอนนี้จะเป็นการพัฒนาชุดคำสั่งในโปรแกรม R เพื่อทำการค้นหาคำหลักแต่ละคำที่ปรากฏอยู่ในประโยคของข้อเสนอแนะที่จัดเก็บอยู่ในฐานข้อมูล พร้อมทั้งบันทึกผลลัพธ์ของการค้นหา

4) ลบข้อเสนอแนะที่ไม่ได้รับการแก้ไขโค้ด (Unrelated Comment Removal) หลังจากที่ได้ทำการค้นหาข้อเสนอแนะที่เกี่ยวข้องกับการบำรุงรักษาแล้ว คณะผู้วิจัยจะทำการลบข้อเสนอแนะทางด้านการบำรุงรักษาที่นักพัฒนาไม่ได้ทำการแก้ไขโค้ดออกจากฐานข้อมูล เพื่อให้ง่ายและสะดวกต่อการวิเคราะห์ผลลัพธ์ ด้วยวิธีการนี้จะทำให้ผู้วิจัยทราบว่ามีข้อเสนอแนะไหนที่ได้รับการแก้ไข

4. ผลการดำเนินการวิจัย

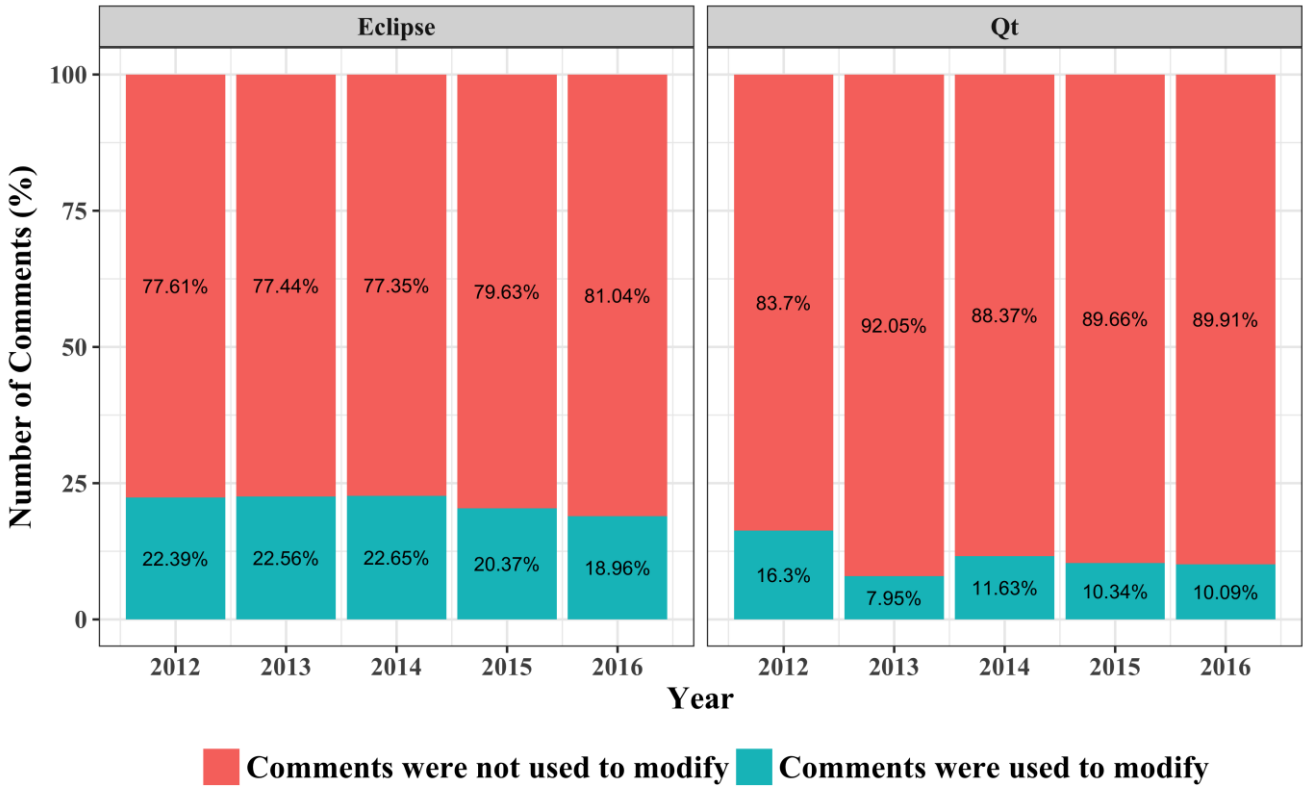
งานวิจัยนี้ได้ทำการศึกษาข้อเสนอแนะในโครงการโอเพนซอร์ส 2 โครงการ โดยผลจากการวิเคราะห์มีรายละเอียดดังนี้

4.1 ผลของความสัมพันธ์ระหว่างข้อเสนอแนะที่เกี่ยวข้องกับการบำรุงรักษาและการแก้ไขโค้ดทางด้านการบำรุงรักษา

คณะผู้วิจัยได้รวบรวมข้อเสนอแนะของโครงการโอเพนซอร์ส 2 โครงการในช่วงปี 2012 - 2016 โดยโครงการ Qt และ Eclipse มีข้อเสนอแนะทั้งหมด 309,396 และ 115,896 ตามลำดับ ในงานวิจัยนี้ คณะผู้วิจัยได้ทำการค้นหาข้อเสนอแนะเกี่ยวกับความสามารถในการบำรุงรักษา 5 ประเภท โดยผลลัพธ์จากการค้นหาด้วยคำหลักปรากฏว่ามีจำนวนข้อเสนอแนะเกี่ยวกับการบำรุงรักษาของโครงการ Qt และ Eclipse เป็นจำนวนด้วยกันทั้งสิ้น 29,840 และ 27,527 ตามลำดับ

ตารางที่ 1: จำนวนข้อเสนอแนะทางด้านความสามารถในการบำรุงรักษาที่ได้รับการแก้ไข

Maintainability	Eclipse	Qt
Modularity	768	354
Reusability	230	188
Analysability	3	1
Modifiability	1,961	1,163
Testability	2,859	1,496



ภาพที่ 1: อัตราส่วนการแก้ไขโค้ดด้านความสามารถในการบำรุงรักษา



ภาพที่ 2: แนวโน้มของข้อเสนอแนะที่ได้รับการแก้ไขโค้ดเกี่ยวกับการบำรุงรักษาซอฟต์แวร์ของโครงการ Eclipse และ Qt

เมื่อคณะผู้วิจัยได้ทำการตรวจสอบความสัมพันธ์ระหว่างข้อเสนอแนะเกี่ยวกับการบำรุงรักษาและจำนวนการแก้ไขโค้ดทางด้านการบำรุงรักษาในโครงการโอเพนซอร์ส ซึ่งผลลัพธ์ของการค้นหาจะแสดงดังตารางที่ 1 โดยจำนวนข้อเสนอแนะที่นักพัฒนาได้ทำการแก้ไขโค้ดเกี่ยวกับความสามารถในการบำรุงรักษาซอฟต์แวร์ทั้ง 5 ประเภทที่ค้นพบในโครงการ Eclipse มีจำนวน 5,821 และโครงการ Qt มีจำนวน 3,202

สำหรับภาพที่ 1 เป็นการแสดงอัตราส่วนของข้อเสนอแนะทางด้านการบำรุงรักษาที่ได้รับการแก้ไขโค้ดในแต่ละปี โดยผลลัพธ์ที่แสดงดังภาพที่ 1 ทำให้สามารถวิเคราะห์ได้ว่าจำนวนการแก้ไขโค้ดที่เกี่ยวข้องกับคุณลักษณะทางด้านการบำรุงรักษาทั้ง 5 ประเภทที่ค้นพบในโครงการ Eclipse และ Qt เป็นระยะเวลา 5 ปี นั่นคือ ในช่วงปี 2012-2016 คิดเป็นค่าเฉลี่ยได้เพียงร้อยละ 21.15 และ 10.73 ซึ่งถือว่าทั้ง 2โครงการมีจำนวนข้อเสนอแนะที่ได้รับการแก้ไขโค้ดทางด้านการบำรุงรักษาซอฟต์แวร์ค่อนข้างน้อย เมื่อเทียบกับจำนวนข้อเสนอแนะทางด้านการบำรุงรักษาที่ได้รับจากผู้ตรวจทานโค้ดทั้งหมด

แม้ว่าจำนวนของการแก้ไขโค้ดทางด้านการบำรุงรักษาจะค่อนข้างน้อย แต่จากการศึกษาแนวโน้มความสนใจของนักพัฒนาที่มีต่อการบำรุงรักษาซอฟต์แวร์แสดงให้เห็นว่านักพัฒนาในชุมชนโอเพนซอร์สให้ความสำคัญกับการบำรุงรักษาเพิ่มมากขึ้น สังเกตได้จากจำนวนข้อเสนอแนะทางด้านการบำรุงรักษาที่ได้รับการแก้ไขในแต่ละโครงการที่มีจำนวนเพิ่มขึ้นตั้งแต่ปี 2012 และมีแนวโน้มที่จะเพิ่มขึ้นอีกในอนาคตดังแสดงในภาพที่ 2 โดยจากแนวโน้มดังกล่าวทำให้คณะผู้วิจัยเชื่อว่า เมื่อมีการแก้ไขหรือปรับปรุงซอฟต์แวร์บ่อยครั้งก็ยิ่งทำให้ระบบมีแนวโน้มที่จะเกิดความซับซ้อนและมีความเสี่ยงที่จะเกิดข้อผิดพลาดหรือข้อบกพร่องมากขึ้นตามไปด้วย อีกทั้งยังส่งผลต่อการทดสอบระบบและการบำรุงรักษาในระยะยาว โดยสังเกตได้จากเส้นกราฟการบำรุงรักษาประเภท Modifiability และ Testability ของทั้ง 2 โครงการที่แสดงให้เห็นว่านักพัฒนาได้ทำการแก้ไขโค้ดมากกว่าประเภทอื่นๆ

นอกจากนี้ คณะผู้วิจัยได้นำผลลัพธ์จากการศึกษาข้างต้นมาทำการวิเคราะห์ โดยใช้สถิติสหสัมพันธ์ (Correlation) เพื่อหาความสัมพันธ์ระหว่างช่วงปีของข้อมูลกับจำนวนข้อเสนอแนะเกี่ยวกับการบำรุงรักษาที่ได้รับการแก้ไขโค้ดของทั้ง 2 โครงการ พบว่าค่าสัมประสิทธิ์สหสัมพันธ์เพียร์สันของโครงการ Eclipse และ Qt เท่ากับ 0.895 และ 0.887 ตามลำดับ โดยค่าดังกล่าวจะแสดงให้เห็นว่าจำนวนปีของการค้นหากับจำนวนข้อเสนอแนะทางด้านการบำรุงรักษามีความสัมพันธ์ระดับสูงไปในทิศทางเดียวกัน จึงอาจจะกล่าวได้ว่าเมื่อจำนวนปีเพิ่มมากขึ้น จำนวนข้อเสนอแนะเกี่ยวกับการบำรุงรักษาซอฟต์แวร์ของทั้ง 2 โครงการก็จะมีจำนวนเพิ่มมากขึ้นด้วยเช่นกัน

4.2 ประเภทของความสามารถในการบำรุงรักษาที่นักพัฒนาให้ความสนใจ

จากผลลัพธ์ที่ได้แสดงในหัวข้อ 4.1 สามารถนำมาเป็นหลักฐานเชิงประจักษ์ได้ว่า ชุมชนนักพัฒนาในโครงการโอเพนซอร์สให้ความสนใจเกี่ยวกับความสามารถในการบำรุงรักษาเป็นพิเศษ คือ ประเภท Testability ที่มีอัตราการแก้ไขโค้ดตามข้อเสนอแนะทางด้านการบำรุงรักษาในโครงการ Eclipse และ Qt มากที่สุดถึงร้อยละ 49.12% และ 46.72% เมื่อเทียบกับจำนวนข้อเสนอแนะที่ได้รับการแก้ไขโค้ดทางด้านการบำรุงรักษาทั้ง 5 คุณลักษณะ โดยประเภทของการบำรุงรักษาที่นักพัฒนาสนใจรองลงมาคือ Modifiability Modularity Reusability และ Analysability ตามลำดับ จากข้อมูลดังกล่าวทำให้สามารถวิเคราะห์ได้ว่า ในปัจจุบันนักพัฒนาอาจจะให้ความสำคัญกับการบำรุงรักษาทางด้านความสามารถในการทดสอบเป็นพิเศษ เพราะเป็นการทดสอบสภาพการคงอยู่และคุณภาพบางประการในระบบ เช่น การทดสอบว่าการประมวลผลของโปรแกรมสามารถทำงานได้สามพันล้านคำสั่งต่อวินาทีได้จริงหรือไม่

5. บทสรุปและงานในอนาคต

งานวิจัยนี้เป็นการวิเคราะห์ข้อเสนอแนะของโครงการโอเพนซอร์ส Eclipse และ Qt ภายใต้กระบวนการตรวจทานโค้ด ซึ่งจากการวิเคราะห์ข้อมูลดังกล่าวทำให้ทราบว่านักพัฒนาในโครงการโอเพนซอร์สทำการแก้ไขโค้ดทางด้านการบำรุงรักษาค่อนข้างน้อย เมื่อเทียบกับจำนวนทั้งหมด แต่จากการศึกษาพบว่า นักพัฒนาในโครงการโอเพนซอร์สมีแนวโน้มในการปรับปรุงคุณภาพของซอร์สโค้ดเพิ่มมากขึ้น โดยจะสังเกตได้จากผลการวิจัยที่แสดงให้เห็นถึงแนวโน้มของการแก้ไขโค้ดตามข้อเสนอแนะเกี่ยวกับการบำรุงรักษา

ซอฟต์แวร์ที่ค่อนข้างมีอัตราที่เพิ่มสูงขึ้นในแต่ละปีโดยเฉพาะทางด้านของการบำรุงรักษาซอฟต์แวร์ ผู้วิจัยมีความเห็นว่านักพัฒนาที่สนใจจะเข้าร่วมการพัฒนาในโครงการโอเพนซอร์สควรให้ความสำคัญกับทางด้านของการบำรุงรักษาซอฟต์แวร์เพราะสามารถช่วยป้องกันผลกระทบจากการแก้ไขโค้ดระหว่างการพัฒนาและอำนวยความสะดวกทางด้านการบำรุงรักษาในอนาคตที่จะช่วยลดเวลาและค่าใช้จ่ายในกระบวนการพัฒนาซอฟต์แวร์

ในอนาคตทางคณะผู้วิจัยจะนำงานวิจัยนี้เป็นแนวทางในการค้นหาคุณลักษณะที่เกี่ยวข้องกับทางด้านความสามารถในการบำรุงรักษาซอฟต์แวร์ที่พบในโครงการโอเพนซอร์สเพิ่มเติมจากเดิมที่มีอยู่ เพื่อเป็นการเพิ่มหลักฐานเชิงประจักษ์และครอบคลุมคำนิยามของ “ความสามารถในการบำรุงรักษา” รวมถึงศึกษาโครงการโอเพนซอร์สอื่นๆ พร้อมทั้งเพิ่มช่วงระยะเวลาที่นานขึ้น เพื่อดูแนวโน้มและแนวทางการให้ข้อเสนอแนะของผู้ตรวจทานโค้ดที่ถูกต้องและแม่นยำมากขึ้น แต่อย่างไรก็ตามข้อมูลที่ได้แสดงในงานวิจัยนี้ก็เพียงพอที่จะเป็นแนวทางในการพัฒนาและปรับปรุงคุณภาพของซอฟต์แวร์โอเพนซอร์สให้ดียิ่งขึ้น

เอกสารอ้างอิง

[1] V. Tiwari, and R.K. Pandey. “Open Source Software and Reliability Metrics,” *The International Journal of Advanced Research in Computer and Communication Engineering*, vol. 1, no. 10, pp. 808-815, December 2012.

[2] B. Norick, J. Krohn, E. Howard, B. Welna, and C. Izurieta. “Effects of the Number of Developers on Code Quality in Open Source Software: A Case Study,” in *Proceedings of the 10th ACM-IEEE International Symposium on Empirical Software Engineering and Measurement*, article no. 62, 1 page, September 2010.

[3] G. S. Walia and J. C. Carver. “Using error information to improve software quality,” in *Proceedings of the International Symposium on Software Reliability Engineering Workshops*, 2013, pp. 107.

[4] Robert L. Glass. “Frequently Forgotten Fundamental Facts about Software Engineering,” *IEEE Software*, vol. 18, no. 3, pp. 112-111, May/June 2001.

[5] Mockus, A., Fielding, R.T., and Herbsleb, J. “Two Case Studies of Open Source Software Development: Apache and Mozill,” *Transactions on Software Engineering and Methodology*, vol. 11, no. 3, pp. 309–346, July 2002.

[6] R. Baker. “Code reviews enhance software quality,” in *Proceedings of the 19th International Conference on Software engineering*, 1997, pp. 570–571.

[7] P. C. Rigby, D. M. German, and M.-A. Storey. “Open source software peer review practices: A case study of the apache server,” in Proceedings of the 30th International Conference on Software Engineering (ICSE’08), ACM, 2008, pp. 541–550.

[8] Baysal, O., Kononenko, O., Holmes, R., and Godfrey, M. W. “The influence of non-technical factors on code review,” in Proceeding of the 20th Working Conference on Reverse Engineering (WCRE), IEEE, 2013, pp. 122-131.

[9] Tao, Yida, DongGyun Han, and Sunghun Kim. “Writing acceptable patches: an empirical study of open source project patches,” in Proceeding of the International Conference on Software Maintenance and Evolution (ICSME), IEEE, 2014, pp. 271-280.

[10] A. Bosu, M. Greiler, and C. Bird. “Characteristics of useful code reviews: an empirical study at Microsoft,” in Proceedings of the 12th Working Conference on Mining Software Repositoriesd IEEE, 2015, pp. 146-156.

[11] J. Czerwonka, M. Greiler and J.Tilford. “Code Reviews Do Not Find Bugs : How the Current Code Review Best Practice Slows Us Down,” in Proceedings of the 37th International Conference on Software Engineering (ICSE '15), ACM/IEEE, 2015, pp. 27-28.

[12] R. Rana and M. Staron. “When Do Software Issues and Bugs get Reported in Large Open Source Software Project?” in Proceedings International Conference on Software Measurement (IWSM-Mensura), 2015, pp. 1-14.

[13] ISO/IEC 25010. “Systems and software engineering -- Systems and software Quality Requirements and Evaluation (SQuaRE) -- System and software quality models,” in ISO/IEC 25010, ed. IEC, 2011.

[14] Fellbaum, Christiane. WordNet. Blackwell Publishing Ltd, 1998.

กฎการแปลงโมเดิร์นฟอร์แทรนสำหรับยูเอ็มแอลซีเควนซ์ไดอะแกรม
Modern Fortran Transformation Rules for UML Sequence Diagrams

อนวัช เล่ห์ทองคำ (Anawat Leatongkam)[1] และอซีส นันทอมรพงศ์ (Aziz Nanthaamornphong)[2]
สาขาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีและสิ่งแวดล้อม มหาวิทยาลัยสงขลานครินทร์ วิทยาเขตภูเก็ต
[1]anawat.le@phuket.psu.ac.th, [2]aziz.n@phuket.psu.ac.th

บทคัดย่อ

ปัจจุบันวิศวกรรมย้อนกลับได้ถูกนำไปใช้อย่างแพร่หลาย โดยเป็นกระบวนการสกัดแนวคิดของระบบ และข้อมูลการออกแบบออกจากระบบ งานวิจัยนี้ให้ความสำคัญกับเครื่องมือด้านวิศวกรรมย้อนกลับที่มีชื่อว่า ForUML ซึ่งสามารถสร้างยูเอ็มแอลไดอะแกรมขึ้นมาจากซอร์สโค้ด โมเดิร์นฟอร์แทรน ซึ่งยังคงเป็นภาษาที่มีนักพัฒนาซอฟต์แวร์ทางด้านวิทยาศาสตร์และวิศวกรรมศาสตร์ใช้อยู่ในปัจจุบัน ในเวอร์ชันแรกของ ForUML สามารถแสดงได้เฉพาะคลาสไดอะแกรมเท่านั้น ซึ่งคลาสไดอะแกรมนั้นจะแสดงให้เห็นถึงโครงสร้าง และความสัมพันธ์ระหว่างคลาสในระบบ ถึงอย่างไรก็ตามคลาสไดอะแกรมก็ยังไม่เพียงพอต่อการวิเคราะห์และทำความเข้าใจระบบ โดยเฉพาะพฤติกรรม และการปฏิสัมพันธ์ระหว่างคลาสในระบบ ซึ่งข้อมูลเหล่านี้สามารถถูกแสดงได้โดยยูเอ็มแอลซีเควนซ์ไดอะแกรม ดังนั้นในงานวิจัยนี้ผู้วิจัยได้นำเสนอแนวคิดการออกแบบกฎการแปลงซอร์สโค้ดภาษาฟอร์แทรนเป็นยูเอ็มแอลซีเควนซ์ไดอะแกรม เพื่อสามารถนำกฎไปพัฒนาต่อยอดความสามารถของเครื่องมือ ForUML ซึ่งจะช่วยให้นักพัฒนาสามารถทำความเข้าใจระบบที่พัฒนาด้วยฟอร์แทรน อีกทั้งยังช่วยส่งเสริมการพัฒนา การบำรุงรักษา กระบวนการตัดสินใจ และการสื่อสารของชุมชนนักพัฒนาทางด้านวิทยาศาสตร์ และวิศวกรรมศาสตร์ทั่วโลก

คำสำคัญ: วิศวกรรมย้อนกลับ ฟอร์แทรน แผนภาพยูเอ็มแอลซีเควนซ์ วิศวกรรมซอฟต์แวร์

Abstract

Recently, reverse engineering has been widely adopted as a valuable process for extracting system abstractions and design information from existing software systems. The proposed research will focus on ForUML, a reverse engineering tool developed to extract UML diagrams from modern, object-oriented Fortran code, which are still used by scientist and engineering application developers. The first version of ForUML produces only UML class diagrams, which provide a useful window into the static structure of a program, including the make-up of each class and the relationships between classes. Rather than visualizing class diagrams, the developers need to understand class behavior and interactions between classes. UML sequence diagrams provide such important algorithmic information. Therefore, we proposed to design rules for transforming object-oriented Fortran into UML sequence diagrams with the goal to extend the ability of ForUML. The proposed capability will enable the visualization of modern Fortran software behavior and algorithmic structures. This ability would enhance the development, maintenance practices, decision processes, and communications in the scientific software community worldwide.

Keyword: Reverse Engineering, Fortran, UML Sequence Diagram, Software Engineering

1. บทนำ

ในปัจจุบันกระบวนการวิศวกรรมย้อนกลับ (Reverse Engineering) เป็นที่รู้จักกันอย่างแพร่หลายโดยเฉพาะอย่างยิ่งในกลุ่มของนักพัฒนาซอฟต์แวร์ การทำวิศวกรรมย้อนกลับในด้านวิศวกรรมซอฟต์แวร์นั้นจะเกี่ยวข้องกับการอ่านและทำความเข้าใจซอร์สโค้ด เพื่อนำมาใช้ในการทำความเข้าใจของระบบซอฟต์แวร์ ซึ่งในกระบวนการพัฒนาซอฟต์แวร์ถ้าหากระบบซอฟต์แวร์มีขนาดใหญ่ หรือมีจำนวนบรรทัดของโค้ดมาก อาจจะก่อให้เกิดปัญหาเรื่องความซับซ้อนของระบบ จึงทำให้ยากต่อการอ่านและทำความเข้าใจซอร์สโค้ด ดังนั้นกระบวนการวิศวกรรมย้อนกลับจึงได้มีส่วนช่วยให้นักพัฒนาสามารถมองเห็นถึงภาพรวมของระบบได้ง่าย เพื่อที่จะได้ทำการบำรุงรักษา และปรับปรุงซอฟต์แวร์ อย่างไรก็ตาม

กระบวนการวิศวกรรมย้อนกลับของระบบซอฟต์แวร์ที่มีขนาดใหญ่ หรือซับซ้อนนั้นเป็นเรื่องที่ยากและมีความท้าทายเป็นอย่างยิ่ง [1] หนึ่งในความท้าทายของกระบวนการวิศวกรรมย้อนกลับ นั่นคือ การสร้างมุมมองที่สื่อความหมายของสิ่งที่เป็นนามธรรมหรือสิ่งที่ไม่มีรูปร่าง เช่น ซอร์สโค้ด เพื่อให้อยู่ในรูปแบบที่สามารถอ่านและทำความเข้าใจความซับซ้อนได้ [2] เช่น ยูเอ็มแอล (Unified Modeling Language: UML)

จากงานวิจัยก่อนหน้านี้ได้มีการพัฒนาเครื่องมือที่มีชื่อว่า ForUML [3] ซึ่งมีคุณสมบัติในการแปลงโค้ดภาษาฟอร์แทรนไปเป็นยูเอ็มแอลคลาสไดอะแกรม (UML Class Diagram) โดยยูเอ็มแอลคลาสไดอะแกรม จะเป็นแผนภาพที่แสดงให้เห็นถึงโครงสร้างการทำงานของคลาส ซึ่งจะบอกว่าแต่ละคลาสมีหน้าที่การทำงานที่สัมพันธ์กับคลาสอื่นอย่างไร สำหรับภาษาฟอร์แทรนนั้น ปัจจุบันได้ถูกพัฒนาเป็นภาษาโปรแกรมเชิงวัตถุคล้ายกับภาษาจาวา (Java) หรือภาษาซีพลัสพลัส (C++) ซึ่งเป็นที่นิยมใช้สำหรับการพัฒนาซอฟต์แวร์ในด้านที่เกี่ยวข้องกับวิทยาศาสตร์ และวิศวกรรมศาสตร์ (Scientific Software) เช่น การพยากรณ์อากาศ ดาราศาสตร์ และการแพทย์ ซึ่งการพัฒนาซอฟต์แวร์ทางด้านนี้ยังขาดเครื่องมือที่ดีในการพัฒนาซอฟต์แวร์ควบคู่กันไป [4] อีกทั้งการพัฒนาซอฟต์แวร์ในด้านนี้ส่วนใหญ่จะพัฒนาขึ้นด้วยการลองผิดลองถูก และการศึกษาค้นคว้าด้วยตัวเอง เนื่องจากนักพัฒนาซอฟต์แวร์ในด้านนี้เป็นนักวิทยาศาสตร์ และวิศวกร ที่มีความรู้เพียงแค่พื้นฐานทั่วไปเกี่ยวกับการเขียนโค้ด จึงทำให้เกิดข้อจำกัดในการเขียนโค้ดในเชิงลึก [5]

อย่างไรก็ตามเครื่องมือ ForUML ที่ถูกพัฒนาขึ้นในเวอร์ชันแรกนั้นยังมีขีดความสามารถที่จำกัด เนื่องจากเครื่องมือนี้สามารถแสดงได้เฉพาะคลาสไดอะแกรมเพียงเท่านั้น โดยแผนภาพนี้จะเป็นแบบจำลองในมุมมองเชิงโครงสร้างซึ่งจะแสดงโครงสร้างของระบบที่ไม่บ่งบอกถึงพฤติกรรมการทำงานและไม่มีการระบุขั้นตอนการดำเนินงาน หรือลำดับการทำงานก่อนหลัง ดังนั้นจึงไม่เพียงพอต่อการวิเคราะห์ และทำความเข้าใจระบบ นอกจากนี้ผู้ใช้งานเครื่องมือได้ให้ข้อเสนอแนะสำหรับการเพิ่มคุณสมบัติหรือความสามารถใหม่ ๆ เช่น การสร้างยูเอ็มแอลซีเควนซ์ไดอะแกรม (UML Sequence Diagram) ซึ่งเป็นแบบจำลองในมุมมองเชิงพฤติกรรมที่แสดงถึงลำดับการทำงานภายในระบบ เพื่อให้สามารถมองเห็นถึงภาพรวมของระบบที่พัฒนาขึ้นมาจากภาษาฟอร์แทรน รวมถึงสามารถช่วยในการตัดสินใจเกี่ยวกับกระบวนการพัฒนาซอฟต์แวร์และการบำรุงรักษาซอฟต์แวร์ได้ดียิ่งขึ้น

งานวิจัยนี้มีวัตถุประสงค์ที่จะนำเสนอกฎการแปลงซอร์สโค้ดภาษาฟอร์แทรนเป็นยูเอ็มแอลซีเควนซ์ไดอะแกรม ซึ่งในปัจจุบันยังไม่มีผู้คิดค้นกฎการแปลงดังกล่าว ผู้วิจัยคาดว่างานวิจัยนี้จะเป็นประโยชน์ในการพัฒนาต่อยอดความสามารถของเครื่องมือ ForUML ให้สามารถสร้างยูเอ็มแอลซีเควนซ์ไดอะแกรมได้ ซึ่งผู้วิจัยเชื่อว่าการมีเอกสารการออกแบบในหลายมุมมอง จะช่วยให้นักพัฒนาสามารถวิเคราะห์ และทำความเข้าใจในระบบซอฟต์แวร์ได้ดียิ่งขึ้น อีกทั้งยังมีส่วนช่วยในกระบวนการพัฒนาซอฟต์แวร์และบำรุงรักษาซอฟต์แวร์ [6]

2. ทฤษฎีและวรรณกรรมที่เกี่ยวข้อง

ในบทนี้จะเป็นการอธิบายทฤษฎีและวรรณกรรมที่เกี่ยวข้อง ได้แก่ ภาษาฟอร์แทรน เครื่องมือ ForUML และวิศวกรรมย้อนกลับ โดยมีรายละเอียดดังต่อไปนี้

2.1 ภาษาฟอร์แทรน (Fortran)

ปัจจุบันภาษาฟอร์แทรนได้ถูกพัฒนาให้มีลักษณะเชิงวัตถุ (Object-Oriented) หรือเรียกอีกอย่างว่า โมเดิร์นฟอร์แทรน (Modern Fortran) [7] เพื่อให้รองรับการพัฒนาซอฟต์แวร์ที่มีความซับซ้อน และเป็นการนำหลักการทางด้านวิศวกรรมซอฟต์แวร์มาประยุกต์ใช้ให้ซอฟต์แวร์มีประสิทธิภาพมากยิ่งขึ้น ดังจะเห็นได้ว่ามีนักวิทยาศาสตร์ และวิศวกรจำนวนมากให้ความสนใจและมีการนำโมเดิร์นฟอร์แทรนมาใช้ในการพัฒนาซอฟต์แวร์ [8], [9] นอกจากนี้ในปัจจุบันผู้ผลิตคอมไพเลอร์สำหรับภาษาฟอร์แทรนจำนวนมากได้พัฒนาให้คอมไพเลอร์มีความสามารถรองรับโมเดิร์นฟอร์แทรน เช่น Numerical Algorithm Group (NAG) และ Intel Fortran

สำหรับโมเดิร์นฟอร์แทรนนั้นมีคุณสมบัติที่สำคัญของภาษาเชิงวัตถุ ได้แก่ การสืบทอดคุณสมบัติ (Inheritance) การพ้องรูป (Polymorphism) การจัดสรรแบบพลวัต (Dynamic Type Allocation) และการกำหนดขอบเขตการทำงาน (Type-Bound Procedures) เนื่องจากโมเดิร์นฟอร์แทรนยังค่อนข้างเป็นเรื่องใหม่ในโลกของการพัฒนาซอฟต์แวร์เชิงวัตถุ (Object-Oriented Programming) จึงทำให้เครื่องมือที่เกี่ยวข้องยังมี

น้อย และยังไม่ได้นำหลักการทางด้านวิศวกรรมซอฟต์แวร์เข้ามาช่วยมากนัก หากเปรียบเทียบกับภาษาเชิงวัตถุอื่น ๆ เช่น ภาษาจาวา หรือภาษาซีพลัสพลัส โดยเฉพาะเครื่องมือในกลุ่มของการทำความเข้าใจโปรแกรม (Program Comprehension Tool) ซึ่งเป็นเครื่องมือที่จะช่วยให้นักพัฒนา และนักออกแบบซอฟต์แวร์สามารถทำความเข้าใจโค้ด หรือระบบซอฟต์แวร์ได้ง่ายขึ้น

2.2 วิศวกรรมย้อนกลับ (Reverse Engineering)

การทำวิศวกรรมย้อนกลับสำหรับซอฟต์แวร์โดยส่วนใหญ่จะเกี่ยวข้องกับซอร์สโค้ด ซึ่งจะเป็นกระบวนการวิเคราะห์ส่วนต่างๆ ของซอร์สโค้ด เพื่อทำความเข้าใจซอร์สโค้ดโดยทั่วไปมักจะใช้สำหรับการวิเคราะห์รหัสไบนารี่ (Binary Code) สำหรับตัวอย่างของซอฟต์แวร์ที่สามารถทำการวิศวกรรมย้อนกลับจากรหัสไบนารี่กลับมาอยู่ในรูปของซอร์สโค้ด (Decompile) นั้นได้แก่ Jad [10] ซึ่งเป็นซอฟต์แวร์ที่สามารถแปลงรหัสไบนารี่ของภาษาจาวา เช่น ไฟล์ที่มีนามสกุล .class เพื่อให้กลับมาอยู่ในรูปของซอร์สโค้ดที่นักพัฒนาซอฟต์แวร์สามารถอ่านและทำความเข้าใจได้

Periklis Andritsos และ Renee J. Miller [11] ได้กล่าวถึงระบบซอฟต์แวร์โดยส่วนใหญ่เมื่อเริ่มมีอายุมากขึ้น การที่จะทำความเข้าใจ และบำรุงรักษานั้นจึงเป็นเรื่องยาก บางครั้งอาจทำให้ระบบไม่มีประสิทธิภาพ และมีค่าใช้จ่ายเพิ่มสูงขึ้น ดังนั้นชุมชนวิศวกรรมซอฟต์แวร์จึงให้ความสำคัญกับการสร้างเครื่องมือที่ช่วยให้นักวิเคราะห์ระบบสามารถมองเห็นถึงโครงสร้างของระบบดังกล่าว

2.3 ยูเอ็มแอลเมต้าโมเดล (UML Meta Model)

ยูเอ็มแอลเป็นภาษาที่ใช้สำหรับอธิบายโมเดลต่าง ๆ ซึ่งเป็นมาตรฐานในการสร้างโมเดลเชิงวัตถุ หรือเป็นมาตรฐานสำหรับสร้างแบบพิมพ์เขียวให้กับระบบซอฟต์แวร์ โดยสามารถใช้ยูเอ็มแอลในการสร้างมุมมองของระบบ กำหนดรายละเอียดของระบบ และพัฒนาระบบขึ้นมาได้ สำหรับเมต้าโมเดลจะเป็นโมเดลที่ใช้สำหรับกำหนดรายละเอียดของโมเดล ซึ่งจะเป็นตัวแทนแสดงถึงรายละเอียดโมเดลนั้น ๆ ในงานวิจัยนี้มีการนำเอกสารเอ็กซ์เอ็มไอมาแสดงเป็นตัวแทนของยูเอ็มแอลซีเควนซ์ไดอะแกรม ยูเอ็มแอลเมต้าโมเดลมีวัตถุประสงค์เพื่อกำหนดคำจำกัดความของวากยสัมพันธ์ (Syntax) และความหมายสำหรับโครงสร้าง หรือองค์ประกอบในยูเอ็มแอลโมเดล ซึ่งเมต้าโมเดลจะช่วยให้นักพัฒนาเข้าใจถึงความหมายของโมเดลที่ต้องการจะสื่อได้ถูกต้องตรงกัน และสามารถสร้างแบบจำลองต่าง ๆ ได้ตรงตามมาตรฐานของยูเอ็มแอล

2.4 เครื่องมือ ForUML

ForUML [3] เป็นเครื่องมือที่สามารถช่วยเรื่องวิศวกรรมย้อนกลับจากซอร์สโค้ดที่พัฒนาขึ้นด้วยโมเดิร์นฟอร์แทรนเป็นยูเอ็มแอลไดอะแกรม ซึ่งเครื่องมือนี้ได้ถูกพัฒนาขึ้นมาโดยหนึ่งในคณะผู้วิจัยของงานวิจัยนี้ และได้ถูกนำไปใช้งานจริง [12] สำหรับขั้นตอนการแปลงซอร์สโค้ดไปเป็นยูเอ็มแอลไดอะแกรมนั้นได้มีการคิดค้นโมเดลขึ้นมาเอง ซึ่งโมเดลที่คิดค้นนี้มีรากฐานแนวคิดมาจากโครงสร้างจำลอง (Meta-Model) ของภาษาเชิงวัตถุที่ถูกพัฒนาขึ้นมาโดย Lethbridge และคณะ [13] ซึ่งทีมของ Lethbridge ได้ทำการออกแบบโครงสร้างที่เรียกว่า The Dagstuhl Middle Metamodel (DMM) ซึ่งเป็นโครงสร้างจำลองที่ได้รับการพิสูจน์ในทางปฏิบัติแล้วว่าเป็นประโยชน์ต่อกระบวนการวิศวกรรมย้อนกลับซอฟต์แวร์ โดย ForUML จะประยุกต์ใช้โครงสร้างดังกล่าวกับซอร์สโค้ดของภาษาฟอร์แทรน สำหรับกระบวนการแปลงนั้นมีทั้งหมด 4 ขั้นตอน ซึ่งมีรายละเอียดในแต่ละขั้นตอนดังนี้

1) Parsing เป็นกระบวนการวิเคราะห์โค้ดออกเป็นส่วนย่อยๆ (Element) โดยใช้ไลบรารี่ Open Fortran Parser (OFP) ซึ่งในกระบวนการตัดคำนี้จะอาศัยไฟล์ไวย์กรณ์ (Grammar) และวากยสัมพันธ์ ของภาษาฟอร์แทรนที่ได้ถูกพัฒนาไว้แล้วในไลบรารี่ OFP โดยกระบวนการนี้จะเป็นการตรวจสอบความถูกต้องของโค้ดที่ผู้ใช้ป้อนให้กับระบบ เพื่อป้องกันไม่ให้เกิดข้อผิดพลาดในขั้นตอนถัดไป

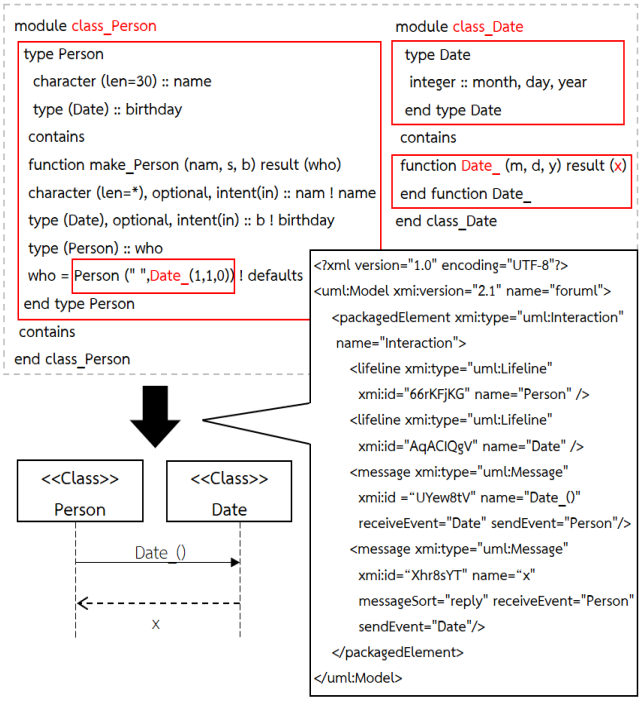
2) Extraction เป็นการค้นหาความสัมพันธ์ระหว่างส่วนย่อย ๆ ที่ได้มาจากขั้นตอนที่ 1

3) Generating เป็นการรวบรวมเอาส่วนย่อย ๆ และความสัมพันธ์ที่ค้นหาได้จากขั้นตอนที่ 1 และ 2 มาสร้างเป็นเอกสารที่อยู่ในรูปแบบเอ็กซ์เอ็มไอ (XML Metadata Interchange: XMI) ซึ่งเอกสารเอ็กซ์เอ็มไอนี้จะเก็บข้อมูลที่จำเป็นเพื่อใช้ในการแสดงยูเอ็มแอลคลาสไดอะแกรม

4) Importing ในขั้นตอนนี้ ForUML จะนำเอาเอกสารเอ็กซ์เอ็มไอ ที่ได้ในขั้นตอนที่ 3 มาแปลงเข้าไปใน ArgoUML เพื่อแสดงยูเอ็มแอลคลาสไดอะแกรม

3. วิธีการดำเนินการวิจัย

การศึกษาวิจัยนี้มีวัตถุประสงค์ที่จะทำการออกแบบกฎการแปลงซอร์สโค้ดภาษาฟอร์แทรนเป็นยูเอ็มแอลซีเควนซ์ไดอะแกรม ซึ่งเป็นแนวคิดที่ได้มาจากการนำมาตรฐานของยูเอ็มแอลซีเควนซ์ไดอะแกรมตามเอกสารรายละเอียดของยูเอ็มแอล (UML Specification) [14] และกฎการแปลงยูเอ็มแอลซีเควนซ์ไดอะแกรมที่มาจากวรรณกรรมที่เกี่ยวข้อง [15], [16], [17] มาประยุกต์ใช้เพื่อสร้างกฎการแปลงซอร์สโค้ดภาษาฟอร์แทรนเป็นยูเอ็มแอลซีเควนซ์ไดอะแกรมขึ้นมาใหม่

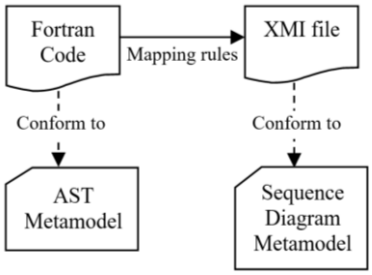


ภาพที่ 1: ตัวอย่างของเอกสารเอ็กซ์เอ็มไอสำหรับ ยูเอ็มแอลซีเควนซ์ไดอะแกรมของโปรแกรมในภาษาฟอร์แทรน

การออกแบบกฎการแปลงซอร์สโค้ดภาษาฟอร์แทรนเป็นยูเอ็มแอลซีเควนซ์ไดอะแกรมจะเริ่มต้นจากการศึกษารายละเอียดของเอกสารเอ็กซ์เอ็มไอ ซึ่งเป็นรูปแบบมาตรฐานของ OMG (Object Management Group) ที่แสดงให้เห็นถึงรายละเอียดข้อมูลของยูเอ็มแอลโมเดล โดยจะเป็นรูปแบบเพื่อใช้สำหรับอธิบายความหมายของกฎต่าง ๆ ในการแปลงเป็นยูเอ็มแอลซีเควนซ์ไดอะแกรม และเพื่อใช้เป็นมาตรฐานกลางในการแลกเปลี่ยนข้อมูลระหว่างกัน สำหรับตัวอย่างเอกสารเอ็กซ์เอ็มไอที่มีข้อมูลของยูเอ็มแอลซีเควนซ์ไดอะแกรมจะแสดงในภาพที่ 1 ซึ่งจะประกอบด้วย

1) xmi:type="uml:Lifeline" จะเป็นรูปแบบที่ไว้กำหนดรายละเอียดของไลฟ์ไลน์ โดยมี xmi:id="66rKFjKG" ที่เป็นหมายเลขกำกับของไลฟ์ไลน์ และ name="Person" ที่แสดงชื่อของไลฟ์ไลน์

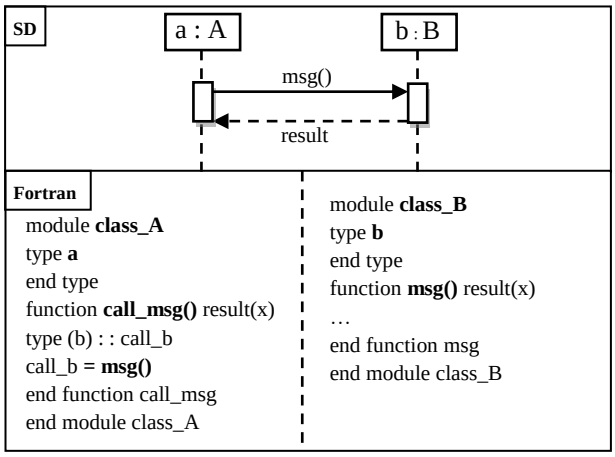
2) xmi:type="uml:Message" จะเป็นรูปแบบที่ไว้กำหนดรายละเอียดของสาร โดยมี xmi:id="Xhr8sYT" เป็นหมายเลขกำกับของสาร messageSort="reply" แสดงประเภทของสาร name="Person" แสดงชื่อของสาร receiveEvent="Person" แสดงไลฟ์ไลน์ที่รับสาร และ sendEvent="Date" แสดงไลฟ์ไลน์ที่ส่งสาร



ภาพที่ 2: กระบวนการแปลงโมเดิร์นฟอร์แทรนเป็นยูเอ็มแอล

สำหรับการออกแบบกฎการแปลงโมเดิร์นฟอร์แทรนเป็นยูเอ็มแอลซีเควนซ์ไดอะแกรมนั้นจะมีขั้นตอนหลักคือ การหาความสัมพันธ์ของ Abstract Syntax Tree (AST) Metamodel ของภาษาฟอร์แทรนและเอกสารเอ็กซ์เอ็มไอ โดยเมต้าโมเดลของทั้งสองจะเป็นโมเดลที่ใช้เป็นตัวแทนของโมเดลหลักดังแสดงในภาพที่ 2

ในส่วนของกฎการแปลงซอร์สโค้ดภาษาฟอร์แทรนเป็นยูเอ็มแอลซีเควนซ์ไดอะแกรม เพื่อให้ง่ายต่อการอธิบาย ผู้วิจัยจะใช้ ภาพที่ 3 สำหรับอธิบายกฎการแปลงที่สร้างขึ้นมา โดยการสร้างกฎการแปลงนั้นผู้วิจัยได้เลือกใช้ภาษา Atlas Transformation Language (ATL) ซึ่งเป็นภาษาที่นิยมใช้สำหรับการแปลงแบบจำลอง โดยผู้วิจัยได้นำส่วนที่สำคัญของกฎการแปลงที่ได้ทำการออกแบบมาแสดงดังนี้



ภาพที่ 3: ตัวอย่างใช้อธิบายกฎการแปลงซอร์สโค้ดภาษาฟอร์แทรนเป็นยูเอ็มแอลซีเควนซ์ไดอะแกรม

1) กฎสำหรับการสร้างไลฟ์ไลน์ เป็นกฎที่ใช้ในการเชื่อมโยงระหว่างไลฟ์ไลน์ของยูเอ็มแอลซีเควนซ์ไดอะแกรมและชื่อคลาสของซอร์สโค้ดภาษาฟอร์แทรน (แสดงในภาพที่ 4)

```
(SD) Lifeline name = (Fortran) Type name
Ex: Lifeline a = Type a
-----------------------------------------
rule Type2Lifeline {
from
    f: Fortran!Type
to
    l: SD!Lifeline (name <- f.name)
}
```

ภาพที่ 4: กฎสำหรับการสร้างไลฟ์ไลน์

2) กฎสำหรับการสร้างสารระหว่างไลฟ์ไลน์ เป็นกฎที่ใช้ในการเชื่อมโยงระหว่างสารของยูเอ็มแอลซีเควนซ์ไดอะแกรมและชื่อเมธอดของภาษาฟอร์แทรน (แสดงในภาพที่ 5)

```
(SD) Message name = (Fortran) Function/ Subroutine name
Ex: Message msg = Function msg
-----------------------------------------
rule Procedures2Message {
from
    f: Fortran!Procedures
to
    m1: SD!SynchronousMessage(name <- f.FunctionName),
    m2: SD!SynchronousMessage(name <- f.SubroutineName),
    m3: SD!AsynchronousMessage(name <- f.SubroutineName)
}
```

ภาพที่ 5: กฎสำหรับการสร้างสารระหว่างไลฟ์ไลน์ของสารแบบ Synchronous และ Asynchronous

3) กฎสำหรับกำหนดการรับและส่งของสารที่เกิดขึ้น เป็นกฎที่ใช้ในการกำหนดการรับสารของไลฟ์ไลน์ และการส่งสารของไลฟ์ไลน์ (แสดงในภาพที่ 6)

4) กฎสำหรับกำหนดการเริ่มต้นและสิ้นสุดการดำเนินงานของสารบนไลฟ์ไลน์ที่เกิดขึ้น เป็นกฎที่ใช้ในการกำหนดจุดเริ่มต้นของสารที่ส่งไปยังไลฟ์ไลน์ และจุดสิ้นสุดของสารบนไลฟ์ไลน์ (แสดงในภาพที่ 6)

5) กฎสำหรับกำหนดการดำเนินงานของสารบนไลฟ์ไลน์ เป็นกฎที่ใช้ในการกำหนดการดำเนินงานที่มีการเกิดขึ้นบนไลฟ์ไลน์นั้น (แสดงในภาพที่ 6)

```
Lifeline name {[Maping name] !/?(send/receive) [Message name];}
Ex: Lifeline a {ab_msg!msg; ab_result?result;}
    Lifeline b {ab_msg?msg; ab_result!result;}
-----------------------------------------
rule Interaction {
from
    f: Fortran!Operation
to
    s: SD!SendOperationEven( id<- f.MessageSendOrder+1,
       covered<- f.TypeName, message<-f.ProceduresName),
    r: SD!ReceiveOperationEven(id<- f.MessageReceiveOrder+1,
       covered <- f.TypeName, message<-f.ProceduresName),
    m: SD!MessageOccurrence (name <- f.ProceduresName,
       send <- s, receive <- r),
    e: SD!ExecuteOccurrence (id<- f.LifelineExecuteOrder+1,
       covered <- f.TypeName),
   se: SD!SendExecuteSpecification(id<- f.SendExecuteOrder+1,
       covered <- f.TypeName, start <- s, finish <- e),
   re: SD!ReceiveExecuteSpecification(
       id<- f. ReceiveExecuteOrder+1,
       covered <- f.TypeName, start <- r, finish <- e)
}
```

ภาพที่ 6: กฎสำหรับการสื่อสารกันระหว่างสารและไลฟ์ไลน์

6) กฎสำหรับการสร้างกรอบของแผนภาพ เพื่อแสดงเงื่อนไข ทางเลือกที่มีหลายเงื่อนไข และการทำซ้ำ เป็นกฎที่ใช้ในการกำหนดกรอบของแผนภาพที่เกิดขึ้น (แสดงในภาพที่ 7)

```
rule Alt {
from
    f: Fortran!Statement(f.Isif_else)
to
    a: SD!CombinedFragment( interactionOperator <- 'Alt')
}
```

ภาพที่ 7: กฎสำหรับการสร้างกรอบของแผนภาพในส่วนที่แสดงเงื่อนไข

4. ผลการดำเนินการวิจัย

จากการออกแบบกฎการแปลงซอร์สโค้ดภาษาฟอร์แทรนเป็นยูเอ็มแอลซีเควนซ์ไดอะแกรมข้างต้นผู้วิจัยได้นำผลลัพธ์เสนอแก่ผู้เชี่ยวชาญด้านภาษาโมเดิร์นฟอร์แทรน ซึ่งเป็นผู้ก่อตั้งสถาบันที่เปิดให้การอบรมรวมถึงที่ปรึกษาทางด้านโมเดิร์นฟอร์แทรน และการพัฒนาซอฟต์แวร์ด้านวิทยาศาสตร์ [10] ซึ่งมีที่ตั้งอยู่ในประเทศสหรัฐอเมริกา ผลจากการตรวจสอบโดยผู้เชี่ยวชาญนั้นพบว่ากฎที่ได้การออกแบบนั้นมีความถูกต้อง และเพื่อตรวจสอบถึงความถูกต้องของกฎการแปลง นอกจากนี้ผู้วิจัยได้ให้ผู้เชี่ยวชาญพิจารณาข้อมูลเปรียบเทียบกับภาษาเชิงวัตถุภาษาอื่น ได้แก่ การเปรียบเทียบวากยสัมพันธ์ของภาษาเชิงวัตถุ เพื่อหาความสัมพันธ์ที่มีความคล้ายคลึงกันในการแสดงสัญลักษณ์แต่ละสัญลักษณ์ของยูเอ็มแอลซีเควนซ์ไดอะแกรม ซึ่งผู้วิจัยได้ทำการเปรียบเทียบกับภาษาจาวา ซึ่งเป็นภาษาโปรแกรมเชิงวัตถุที่นิยมในปัจจุบันดังแสดงในตารางที่ 1 โดยแบ่งออกเป็น 2 ส่วนตามลักษณะการทำงานของสัญลักษณ์ประกอบด้วย

ตารางที่ 1: เปรียบเทียบการแปลงซอร์สโค้ดเป็นยูเอ็มแอลซีเควนซ์ไดอะแกรมระหว่างภาษาจาวาและฟอร์แทรน

Rule	Java Syntax	Modern Fortran Syntax	Notations
Lifeline	public class MyClass	type MyClass	**Instance:Class**
Messages			
Create Message	MyClass my = new MyClass();	type(MyClass) ::my	**create**
Reply Message	public int getID() { return id; }	function getID() result(id)	
Synchronous Message	my.getID();	call my%getID() my = getID();	
Asynchronous Message	my.getID();	call my%getID()	

1) Lifeline แสดงถึงตัวแทนของคลาส ประกอบด้วยชื่อของอินสแตนซ์ (Instance) และชื่อของคลาส

2) Messages แสดงถึงสารที่ส่งระหว่างไลฟ์ไลน์ ประกอบด้วย สารที่ใช้สร้างวัตถุ สารที่ใช้ตอบกลับ สารแบบซินโครนัส และสารแบบอะซิงโครนัส

สำหรับในส่วน Interaction Fragment ซึ่งเป็นส่วนที่สื่อสารกันระหว่างสารและไลฟ์ไลน์ ประกอบด้วย การรับและส่งของสารที่เกิดขึ้น การเริ่มต้นและสิ้นสุดการดำเนินงานของสารบนไลฟ์ไลน์ที่เกิดขึ้น ข้อกำหนดการดำเนินงานของสารบนไลฟ์ไลน์ และกรอบของแผนภาพ สามารถแสดงสัญลักษณ์ที่เหมือนกับในภาษาจาวาได้ดังตารางที่ 2

ตารางที่ 2: สัญลักษณ์ของยูเอ็มแอลในส่วน Interaction Fragment

Rule	Fortran Semantics	Notations
Interaction Fragment		
Message Occurrence	Send and Receive Occurrence	
Execution Specification	Start and Finish Occurrence	start finish
Execution Occurrence	Activation	
Combined Fragment	Loops, Branches, and Other Alternatives	

เพื่อเป็นการตรวจสอบว่ากฎการแปลงที่ได้สร้างขึ้นสามารถนำไปประยุกต์ใช้กับภาษาฟอร์แทรนได้จริง ผู้วิจัยได้ทำการพัฒนาโปรแกรมขึ้นมาใช้ในการทดสอบ โดยประยุกต์ใช้กฎการแปลงที่ได้ทำการออกแบบไว้กับซอร์สโค้ดในภาษาฟอร์แทรนที่นำมาจาก [18] (http://research.te.psu.ac.th/aziz/FortranCode/page1.html) เพื่อสร้างเอกสารเอ็กซ์เอ็มไอ โดยเอกสารเอ็กซ์เอ็มไอที่ได้จะสอดคล้องกับรายละเอียดดังแสดงในตารางที่ 1 และตารางที่ 2 โดยผลลัพธ์จะถูกแสดงดังภาพที่ 8

```
<?xml version="1.0" encoding="UTF-8"?>
<uml:Model xmlns:uml="http://schema.omg.org/spec/UML/2.1.1"
xmlns:xmi="http://schema.omg.org/spec/XMI/2.1" xmi:version="2.1" >
  <packagedElement xmi:type="uml:Interaction" xmi:id="maOS0UN7r9">
    <lifeline xmi:id="50q3889y9i" name="Program main" coveredBy="0ieQJt2B9j
       1jmE1ni4I2 ..."/>
    <lifeline xmi:id="maOS0UN7r9" name="Person" coveredBy="17h0ww6f9T
       6b2ILEJ6MU ..."/>
    <lifeline xmi:id="O3Zh3E7go1" name="Date" coveredBy="W3YJ9uK3HA"/>
    <lifeline xmi:id="P9oGkr9518" name="Student" coveredBy="1eZ1I8ooO4
       g126s2IYv8 ..."/>
    <fragment xmi:type="uml:MessageOccurrenceSpecification" xmi:id="0ieQJt2B9j"
       covered="1OA68Xe0xM"
       message="D1I7wX8K4m"/>
     ...
    <fragment xmi:type="uml:BehaviorExecutionSpecification" xmi:id="nE4aoW22V7"
       covered="3RzudlQI8g"
       start="6b2ILEJ6MU" finish="_K2a3rfZ95N"/>
     ...
    <fragment xmi:type="uml:ExecutionOccurrenceSpecification" xmi:id="vazn0ntg3a"
       covered="3ewfRdbxIp"/>
     ...
    <message xmi:type="uml:Message" xmi:id="D1I7wX8K4m" name="create"
       messageSort="createMessage"
       receiveEvent="17h0ww6f9T" sendEvent="0ieQJt2B9j"/>
    <message xmi:type="uml:Message" xmi:id="XBnOh1Ful7" name="make_Person"
       receiveEvent="6b2ILEJ6MU" sendEvent="1jmE1ni4I2"/>
    <message xmi:type="uml:Message" xmi:id="b3xlhqo5fg" name="Person_"
       receiveEvent="Zcm3918AOX" sendEvent="5ZgCNZRsy1"/>
    <message xmi:type="uml:Message" xmi:id="9sN4ex8845" messageSort="reply"
       receiveEvent="yHfYs8zF0J" sendEvent="W3YJ9uK3HA"/>
     ....
  </packagedElement>
</uml:Model>
```

ภาพที่ 8: ผลลัพธ์การแปลงของซอร์สโค้ดจากระบบที่ได้นำมาทดสอบ

เนื่องจากข้อจำกัดของพื้นที่ ภาพที่ 8 แสดงเนื้อหาบางส่วนของเอกสารเอ็กซ์เอ็มไอที่ได้จากการประยุกต์กฎที่สร้างขึ้น โดยจะประกอบด้วย Lifeline ได้แก่ Program main Person Date และ Student สำหรับในส่วนของ Fragment สามารถแบ่งออกได้ 2 ส่วนคือ ส่วนที่กำหนดการสื่อสารระหว่างสารและไลฟ์ไลน์ ได้แก่ Message Occurrence Specification Behavior Execution Specification และ Execution Occurrence Specification และส่วนสำหรับการสร้างกรอบของแผนภาพ ได้แก่ Alternatives Option และ Loop โดยซอร์สโค้ดที่นำมาทดสอบจะไม่มีรายละเอียดของส่วนการสร้างกรอบของแผนภาพ สำหรับส่วนสุดท้ายคือ Message ดังที่แสดงเช่น create จะเป็นสารประเภท Create Message ซึ่งจะแสดงถึงการสร้างออบเจ็กต์ของคลาส โดยจะมีการกำหนดไลฟ์ไลน์ที่ทำการรับและส่งสาร เป็นต้น ในขั้นตอนสุดท้ายผู้วิจัยได้ตรวจสอบความถูกต้องด้วยวิธีการตวรจสอบโดยจะพิจารณาจากการเปรียบเทียบระหว่างเอกสารเอ็กซ์เอ็มไอที่ได้มาจากโปรแกรมกับซอร์สโค้ดฟอร์แทรนแต่ละส่วนซึ่งการเปรียบเทียบนี้ให้ผลถูกต้องในทุกส่วน

5. อภิปรายผลและสรุปผล

งานวิจัยนี้นำเสนอ แนวคิดการออกแบบกฎการแปลงโมเดิร์นฟอร์แทรนเป็นยูเอ็มแอลซีเควนซ์ไดอะแกรม โดยมีจุดมุ่งหมายเพื่อนำกฎที่ได้ไปประยุกต์ใช้ในการพัฒนาเครื่องมือสำหรับแปลงโมเดิร์นฟอร์แทรนเป็นยูเอ็มแอลซีเควนซ์ไดอะแกรม จากการออกแบบกฎการแปลงข้างต้น ผู้วิจัยได้นำไปเปรียบเทียบกับภาษาจาวา ซึ่งเป็นภาษาเชิงวัตถุที่สมบูรณ์ ในขณะที่ภาษาฟอร์แทรนเริ่มมีการพัฒนาเป็นภาษาเชิงวัตถุในภายหลัง แต่เมื่อมีการนำมาเปรียบเทียบกัน ทั้ง 2 ภาษากลับมีความสามารถที่คล้ายคลึงกัน แต่จะแสดงในลักษณะที่แตกต่างกัน เช่น ชื่อของคลาสและเมธอดของทั้ง 2 ภาษาจะมีความแตกต่างกัน

ในอนาคตผู้วิจัยจะนำกฎการแปลงที่นำเสนอในงานวิจัยนี้ไปใช้สำหรับพัฒนาความสามารถของเครื่องมือ ForUML เพื่อใช้ในการสร้างซีเควนซ์ไดอะแกรม

เอกสารอ้างอิง

[1] M. Lanza and S. Ducasse, “Polymetric views-a lightweight visual approach to reverse engineering,” *IEEE Transactions on Software Engineering*, vol. 29, no. 9, pp. 782–795, Sep. 2003.

[2] T. Systa, “On the relationships between static and dynamic models in reverse engineering java software,” in *Proceedings of Reverse Engineering Sixth Working Conference*, 1999, pp. 304–313.

[3] A. Nanthaamornphong, J. Carver, K. Morris, and S. Filippone, “Extracting uml class diagrams from object-oriented fortran: Foruml,” *Scientific Programming, Article ID 421816*, vol. 2015, pp. 1–15, Jan. 2015.

[4] J. C. Carver, R. P. Kendall, S. E. Squires, and D. E. Post, “Software development environments for scientific and engineering software: A series of case studies,” in *Proceedings of the 29th International Conference on Software Engineering (ICSE’07)*, 2007, pp. 550–559.

[5] J. C. Carver, “Report: the second international workshop on software engineering for CSE,” *Computing in Science & Engineering*, vol. 11, no. 6, pp. 14–19, Nov. 2009.

[6] B. Dobing and J. Parsons, “How UML is used,” *Communications of the ACM*, vol. 49, no. 5, pp. 109–113, May 2006.

[7] N. S. Clerman and W. Spector, *Modern Fortran: style and usage*. Cambridge University Press, 2011.

[8] D. Barbieri, V. Cardellini, S. Filippone, and D. Rouson, “Design patterns for scientific computations on sparse matrices,” in *Proceedings of European Conference on Parallel*, 2011, pp. 367–376.

[9] K. Morris, D. W. Rouson, M. N. Lemaster, and S. Filippone, “Exploring capabilities within ForTrilinos by solving the 3D Burgers equation,” *Scientific Programming*, vol. 20, no. 3, pp. 275–292, Jul. 2012.

[10] A. Rukin, “Java decompilers,” *(2016, Nov 25).* [Online]. Available: http://www.javadecompilers.com/

[11] P. Andritsos and R. J. Miller, “Reverse engineering meets data analysis,” in *Proceedings of the 9th International Workshop on Program Comprehension*, 2001, pp. 157–166.

[12] D. ROUSON, “Sourcery Institute,” *(2016, Nov 25)*. [Online]. Available: http://www.sourceryinstitute.org/software.html

[13] T. C. Lethbridge, S. Tichelaar, and E. Plödereder, “The dagstuhl middle metamodel: A schema for reverse engineering,” *Electronic Notes in Theoretical Computer Science*, vol. 94, pp. 7–18, May 2004.

[14] OMG, “UML specification v2.5,” *(2017, Jan 2).* [Online]. Available: http://www.omg.org/spec/UML/2.5/PDF

[15] P. Sawprakhon and Y. Limpiyakorn, “Sequence Diagram Generation with Model Transformation Technology,” in *Proceedings of the International MultiConference of Engineers and Computer Scientists*, 2014, pp. 12–14.

[16] C. Li, L. Dou, and Z. Yang, “A metamodeling level transformation from UML sequence diagrams to Coq.,” in *Proceedings of International Conference on Information and Communication Technology for Competitive Strategies*, 2014, pp. 147–157.

[17] E. Merah, “Design of ATL Rules for Transforming UML 2 Sequence Diagrams into Petri Nets,” *International Journal of Computer Science and Business Informatics*, vol. 8, no. 1, pp. 1–21, Jan. 2014.

[18] J. Akin, “Object oriented programming via Fortran 90,” *Engineering Computations*, vol. 16, 1999, pp. 26–48.

ระบบสารสนเทศสำหรับการประกันคุณภาพการศึกษาระดับหลักสูตร บนคลาวด์คอมพิวติง กรณีศึกษามหาวิทยาลัยราชภัฏรำไพพรรณี

Information System for Quality Assurance in Program Level Based on Cloud ComputingCase Study : Rambhai Barni Rajabhat University

นิทัศน์ นิลฉวี (Nitat Ninchawee)[1] และวันดี โชคช่วยพัฒนากิจ (Wandee Chokchueypattanakit)[2]

สาขาวิชาเทคโนโลยีสารสนเทศ คณะวิทยาการคอมพิวเตอร์และเทคโนโลยีสารสนเทศ มหาวิทยาลัยราชภัฏรำไพพรรณี

[1]Nitat.n@rbru.ac.th [2]Wandee.c@rbru.ac.th

บทคัดย่อ

งานวิจัยชิ้นนี้มีวัตถุประสงค์เพื่อพัฒนาระบบสารสนเทศสำหรับการประกันคุณภาพการศึกษาระดับหลักสูตร ซึ่งใช้กรอบแนวคิดเซนด์เฟรมเวิร์ค 3 สำหรับภาษาพีเอชพี ฐานข้อมูลมายเอสคิวแอลในการพัฒนา และเลือกใช้การบริการเว็บผ่านคลาวด์คอมพิวติงของแอมะซอนเว็บเซอร์วิสอีซีทู เพื่อหลีกเลี่ยงปัญหาความไม่สามารถเข้ากันได้ของซอฟต์แวร์ที่ติดตั้งอยู่ในเครื่องให้บริการเว็บของมหาวิทยาลัย โดยมีวิธีการวิจัย 4 ขั้นตอนได้แก่ 1) การศึกษาความเป็นไปได้ 2) การวิเคราะห์ปัญหาและการออกแบบระบบ 3) พัฒนาระบบ 4) การทดสอบประสิทธิภาพของระบบ และได้ประเมินประสิทธิภาพของระบบจากผู้ใช้งาน จำนวน 10 ท่านได้ค่าเฉลี่ยเท่ากับ 3.75 และค่าเบี่ยงเบนมาตรฐานเท่ากับ 0.56 พบว่าประสิทธิภาพของระบบอยู่ในระดับดี

คำสำคัญ: คลาวด์คอมพิวติง ระบบสารสนเทศ การประกันคุณภาพการศึกษาระดับหลักสูตร

Abstract

The objective of this research was to implement the information system for quality assurance in program level. The system was implemented using Zend framework 3, PHP as a programming language, MySQL as a database, Amazon Web Service EC2 as a web server in order that software incompatible with recent web server would be avoided. The steps for conducting the research were 1) Feasibility study 2) Analysis and Design 3) Implementation and 4) Evaluation. Additionally, the implemented system was evaluated its proficiency by testing with ten users mean equal 3.75 with standard deviation at 0.56 and the result was in good level.

Keyword: Cloud Computing, Information System, Quality Assurance in Program Level

1. บทนำ

การประกันคุณภาพการศึกษาเป็นการดำเนินการเพื่อสร้างความเชื่อมั่นในคุณภาพของการจัดการศึกษาและคุณภาพของผลผลิตทางการศึกษา ซึ่งตั้งแต่ปีการศึกษา 2542 เป็นต้นมาสถาบันอุดมศึกษาได้นำระบบประกันคุณภาพมาใช้เป็นเครื่องมือในการกำกับการบริหารงานของมหาวิทยาลัย ตามพระราชบัญญัติการศึกษาแห่งชาติ พ.ศ. 2542 มาตรา 47 โดยเกณฑ์มาตรฐานเมื่อใช้ไประยะเวลาหนึ่งจะมีการปรับเปลี่ยนเพื่อความเหมาะสมในการประเมินมาตรฐานการศึกษาของสถาบันอุดมศึกษา โดยเกณฑ์มาตรฐานปีการศึกษา 2553 ได้แบ่งเป็น 9 องค์ประกอบซึ่งประกอบด้วย 1) แผน 2) การเรียนการสอน 3) กิจกรรมนักศึกษา 4) การวิจัย 5) การบริการวิชาการ 6) ทำนุบำรุงศิลปวัฒนธรรม 7) การบริหารจัดการ 8) การเงิน 9) การประกันคุณภาพ ซึ่งจะใช้เกณฑ์มาตรฐานเดียวกันกับหน่วยงานทั้งระดับคณะและสถาบัน

ในปีการศึกษา 2557 สำนักงานคณะกรรมการการอุดมศึกษาได้มีการปรับเปลี่ยนเกณฑ์มาตรฐาน [1] โดยเกณฑ์มาตรฐานการประเมินได้แบ่งเป็นระดับหลักสูตร คณะ และสถาบัน ซึ่งแต่ละระดับมีองค์ประกอบต่างกันดังนี้ ระดับหลักสูตรมี 6 องค์ประกอบ คือ 1) การกำกับมาตรฐาน 2) บัณฑิต 3) นักศึกษา 4) อาจารย์ 5) หลักสูตร การเรียนการสอน การประเมิน

ผู้สอน 6) สิ่งสนับสนุนการเรียนรู้ ในระดับคณะและสถาบันมี 5 องค์ประกอบ คือ 1) การผลิตบัณฑิต 2) การวิจัย 3) การบริการวิชาการ 4) การทำนุบำรุงศิลปะและวัฒนธรรม 5) การบริหารจัดการ โดยการประเมินตั้งแต่รอบปีการศึกษา 2557 จะมุ่งเน้นการประเมินในระดับหลักสูตร ทำให้หลักสูตรจะต้องมีการรวบรวมเอกสารหลักฐาน และข้อมูลเพื่อใช้ในการประเมินการประกันคุณภาพเมื่อสิ้นสุดปีการศึกษา

จากปัญหาดังกล่าวผู้วิจัยจึงได้นำเสนอระบบสารสนเทศสำหรับงานประกันคุณภาพการศึกษาระดับหลักสูตร ที่ครอบคลุมเกณฑ์มาตรฐานของหลักสูตรของสำนักงานคณะกรรมการการอุดมศึกษา ปีการศึกษา 2557 เพื่อใช้ในการจัดเก็บรวบรวมได้อย่างถูกต้องแม่นยำ สามารถประมวลผลและวิเคราะห์ข้อมูลได้อย่างรวดเร็ว โดยใช้กรอบแนวคิดสำหรับภาษาพีเอชพี เซนด์เฟรมเวิร์ค เวอร์ชัน 3 (Zend Framework 3) ฐานข้อมูลมายเอสคิวแอล (MySQL) ในการพัฒนาระบบสารสนเทศ และต้องการให้สามารถเรียกใช้งานผ่านเครือข่ายอินเทอร์เน็ต แต่เนื่องจากมหาวิทยาลัยราชภัฏรำไพพรรณีมีข้อจำกัดในเรื่องทรัพยากร เครื่องที่ให้บริการเว็บยังใช้ตัวแปลภาษาพีเอชพีเวอร์ชัน 5.2 ซึ่งเป็นเวอร์ชันเก่าและยังไม่สามารถอัพเกรดเวอร์ชันของภาษาพีเอชพีเป็นเวอร์ชันใหม่ได้เพราะจะทำให้มีผลกระทบกับระบบเดิมของมหาวิทยาลัย ผู้วิจัยจึงใช้การจำลองเครื่องบริการเว็บผ่านคลาวด์คอมพิวติงของแอมะซอนเว็บเซอร์วิสอีซีทู (Amazon Web Service EC2) เพื่อให้สามารถรองรับตัวแปลภาษาพีเอชพีเวอร์ชันใหม่ได้โดยไม่ต้องลงทุนด้านฮาร์ดแวร์

1.1 วัตถุประสงค์งานวิจัย

1.1.1 เพื่อพัฒนาระบบสารสนเทศเพื่อรวบรวมสารสนเทศสำหรับการประกันคุณภาพระดับหลักสูตรผ่านคลาวด์คอมพิวติง โดยใช้มหาวิทยาลัยราชภัฏรำไพพรรณีเป็นกรณีศึกษา

1.1.2 เพื่อประเมินประสิทธภาพของระบบสารสนเทศสำหรับการประกันคุณภาพระดับหลักสูตรผ่านการจำลองเครื่องบริการเว็บบนคลาวด์คอมพิวติง โดยเลือกใช้มหาวิทยาลัยราชภัฏรำไพพรรณีเป็นกรณีศึกษา

2. วรรณกรรมและงานวิจัยที่เกี่ยวข้อง

2.1 วรรณกรรมที่เกี่ยวข้อง

2.1.1 การประกันคุณภาพการศึกษา

การประกันคุณภาพการศึกษา [1] หมายถึง การบริหารจัดการและดำเนินกิจกรรมตามพันธกิจปกติของสถานศึกษาเพื่อพัฒนาคุณภาพของผู้เรียนอย่างต่อเนื่อง สร้างความมั่นใจให้ผู้รับบริการทางการศึกษา ทั้งผู้รับบริการโดยตรง ได้แก่ ผู้เรียน ผู้ปกครอง และผู้รับบริการทางอ้อม ได้แก่ สถานประกอบการ ประชาชน และสังคมโดยรวม โดยในการประกันคุณภาพการศึกษา ปีการศึกษา 2557 สำนักงานคณะกรรมการอุดมศึกษาได้มีการกำหนดเกณฑ์มาตรฐานสำหรับการประเมินในหน่วยงานแต่ละระดับแบ่งตามองค์ประกอบได้ดังนี้

- ระดับหลักสูตรมี 6 องค์ประกอบ คือ 1) การกำกับมาตรฐาน 2) บัณฑิต 3) นักศึกษา 4) อาจารย์ 5) หลักสูตร การเรียนการสอน การประเมินผู้สอน 6) สิ่งสนับสนุนการเรียนรู้
- ระดับคณะและสถาบันมี 5 องค์ประกอบ คือ 1) การผลิตบัณฑิต 2) การวิจัย 3) การบริการวิชาการ 4) การทำนุบำรุงศิลปะและวัฒนธรรม 5) การบริหารจัดการ

2.1.2 คลาวคอมพิวติง (Cloud Computing)

คลาวด์คอมพิวติง [2] คือ การปรับระบบการประมวลผลแบบเดิมที่ประมวลผลบนเครื่องคอมพิวเตอร์ส่วนบุคคล หรือเครื่องคอมพิวเตอร์แม่ข่ายมาเป็นการประมวลผลผ่านการทำงานของเครื่องคอมพิวเตอร์แม่ข่ายจำนวนมากที่เชื่อมโยงกันผ่านเครือข่ายอินเทอร์เน็ต โดยกลุ่มเครื่องคอมพิวเตอร์แม่ข่ายเหล่านั้นจะได้รับการควบคุมและจัดการรูปแบบ ซึ่งผู้ใช้บริการสามารถกำหนดประสิทธิภาพให้กับผู้รับบริการตามเงื่อนไขที่ตกลงระหว่างกัน

2.1.3 แอมะซอนเว็บเซอร์วิสอีซีทู (Amazon Web Service Elastic Computer Cloud)

แอมะซอนเว็บเซอร์วิสอีซีทู [2] เป็นการให้บริการคลาวด์คอมพิวติงประเภทหนึ่งของแอมะซอนที่ทำให้ผู้ใช้สามารถสร้างตัวจำลองเครื่องบริการเว็บหรือฐานข้อมูลที่ผู้ใช้สามารถกำหนดทรัพยากรที่ต้องการใช้ได้ เช่นการเลือกขนาดของซีพียู หน่วยความจำ ขนาดของดิสก์ในการเก็บข้อมูลและสามารถปรับเปลี่ยนได้อย่างรวดเร็ว เป็นต้น

2.2 งานวิจัยที่เกี่ยวข้อง

สาธิต และวชิรศักดิ์ [3] ได้พัฒนาระบบฐานข้อมูลเพื่อสนับสนุนการซ่อมบำรุงและการบริหารจัดการเครือข่าย กรณีศึกษา ระบบเครือข่ายไร้สายของมหาวิทยาลัยวลัยลักษณ์ เพื่อใช้ในการสนับสนุนงานซ่อมบำรุง และบริหารจัดการโดยใช้แบบจำลอง MVC ในการพัฒนาระบบ ซึ่งความสามารถของระบบประกอบด้วยความสามารถในการจัดเก็บข้อมูลอุปกรณ์ ข้อมูลผู้ผลิต ประวัติการติดตั้งอุปกรณ์ ประวัติการใช้งาน แผนที่จุดติดตั้งอุปกรณ์ การตั้งค่าอุปกรณ์ รวมถึงข้อมูลประวัติการใช้งาน (Syslog) เพื่อให้สามารถรวบรวมข้อมูลเหล่านี้มาใช้ในการวิเคราะห์ ตรวจสอบ และยังมีการจัดทำช่องทางในการติดต่อกับผู้ดูแลระบบผ่านเฟซบุค จดหมายอิเล็กทรอนิกส์ และเว็บแอปพลิเคชัน ทำให้ผู้ดูแลระบบสามารถให้บริการในการซ่อมบำรุงเครือข่ายได้อย่างสะดวกและรวดเร็วยิ่งขึ้น

นวลศรีและอภิสิทธิ์ [4] ได้นำพัฒนาแบบจำลองของการบูรณาการข้อมูลอาจารย์ โดยเน้นที่องค์กรในระดับอุดมศึกษา สำหรับนำมาใช้ในการพัฒนาเว็บแอปพลิเคชันที่สามารถเชื่อมโยงกับระบบสารสนเทศอื่น ๆ โดยใช้เทคโนโลยีเว็บเซอร์วิส และนำเซอร์วิสดังกล่าวมาเชื่อมโยงเว็บไซต์หลักของสถาบันซึ่งมีความสามารถในการปรับเปลี่ยนข้อมูลได้ทันทีที่มีการเปลี่ยนแปลงข้อมูลต้นทาง โดยระบบได้นำไปใช้งานจริงกับวิทยาลัยพยาบาล และวิทยาลัยสาธารณสุข สังกัดกระทรวงสาธารณสุขทั่วประเทศ จำนวน 39 วิทยาลัย

เย็นใจและมณเฑียร [5] ได้พัฒนาระบบสารสนเทศเพื่อการบริหารจัดการวางแผน ติดตามโครงการและงบประมาณ โดยใช้บริษัท บางกอก อินโนเวชั่น เทคโนโลยี จำกัด เป็นกรณีศึกษา ซึ่งมีวัตถุประสงค์ในการพัฒนาเพื่อบริหารจัดการงบประมาณขององค์กร ติดตามความคืบหน้าของโครงการที่กำลังดำเนินการในแต่ละปี ในการพัฒนาระบบใช้ ภาษา C# ในโปรแกรม Microsoft Visual Studio 2015 และฐานข้อมูล Microsoft SQL Server 2014 เป็นฐานข้อมูลในการพัฒนาโปรแกรมในแพลตฟอร์มวินโดวส์ โดยระบบได้รับการประเมินประสิทธิภาพของระบบ โดยใช้ผู้เชี่ยวชาญด้านคอมพิวเตอร์ 2 คน ผู้เชี่ยวชาญด้านการวางแผนและงบประมาณ จำนวน 3 คน และผู้ใช้ในองค์กรจำนวน 30 คน ซึ่งผลการประเมินความคิดเห็นจากผู้เชี่ยวชาญได้ค่าเฉลี่ย 4.01 และค่าเบี่ยงเบนมาตรฐาน 0.01 ในส่วนของผู้ใช้ได้ค่าเฉลี่ย 3.88 และค่าเบี่ยงเบนมาตรฐาน 0.16 ซึ่งระบบสารสนเทศเพื่อการบริหารจัดการวางแผน ติดตามโครงการ และงบประมาณ กรณีศึกษา บริษัท บางกอก อินโนเวชั่น เทคโนโลยี จำกัด มีคุณภาพโดยรวมอยู่ในระดับดี

ศศิประภาและศักดิ์ชาย [6] ได้พัฒนาเว็บแอปพลิเคชันสำหรับจองสถานที่ทำงานโดยใช้หลักการออกแบบเว็บไซต์แบบเรสปอนซีฟ (Responsive) และค้นหาสถานที่ทำงานของธุรกิจเกิดใหม่ (Startup) โดยใช้เทคนิค อีลาสติกเซิร์ช (Elastic Search) มาใช้ประยุกต์ในการตัดคำ เพื่อใช้ในการค้นคืนสารสนเทศภาษาไทย ซึ่งระบบได้รับการประเมินประสิทธิภาพจากผู้เชี่ยวชาญจำนวน 5 คน ได้มีประสิทธิภาพอยู่ในระดับดี สามารถนำไปใช้งานได้

บุรินทร์ [7] ได้พัฒนาระบบตรวจเยี่ยมการประกันคุณภาพการศึกษาระดับหลักสูตรและคณะ ซึ่งพัฒนาระบบให้เป็นเว็บแอปพลิเคชันที่ใช้ตัวแปลภาษาพีเอชพี โดยใช้มายเอสคิวแอลเป็นฐานข้อมูล และใช้ตัวแบบวงจรชีวิตแบบก้นหอย ควบคู่กับการใช้ตัวต้นแบบทูน่าในการจัดการความรู้ในการพัฒนาระบบ โดยใช้มหาวิทยาลัยเนชั่นเป็นกรณีศึกษา ซึ่งได้รับการประเมินโดยผู้เชี่ยวชาญอยู่ในระดับมาก และมีผลประเมินความพึงพอใจของผู้ใช้งานระบบตรวจเยี่ยมอยู่ในระดับมาก

ทัศนีย์ [8] ได้เสนอกรอบแนวคิดกรอบแนวคิดการจัดการและพัฒนาระบบการจัดการสารสนเทศผ่านคลาวด์คอมพิวติง ตามกรอบมาตรฐานคุณวุฒิระดับอุดมศึกษาแห่งชาติ ซึ่งประกอบด้วยโมดูลดังนี้ 1) โมดูลการจัดการเนื้อหาเอกสาร 2) โมดูลผู้ใช้งานระบบ 3) โมดูลการทำงานร่วมกัน 4) โมดูลการเผยแพร่เอกสาร และ 5) โมดูลการสร้างเว็บไซต์สำหรับทำงานร่วมกัน ซึ่งผลการใช้งานระบบโดยผู้ใช้งานระบบจำนวน 54 คน ได้ผลการประเมินอยู่ในระดับมาก

เมื่อศึกษาวรรณกรรมและงานวิจัยที่เกี่ยวข้อง ผู้วิจัยจึงได้เลือกใช้พัฒนาระบบสารสนเทศที่สามารถใช้งานผ่านเครือข่ายอินเทอร์เน็ตโดยเลือกใช้การสร้างบริการที่เกี่ยวข้องกับการให้บริการเว็บบนคลาวด์คอมพิวติงของแอมะซอนเว็บเซอร์วิส และเลือกใช้เซนด์เฟรมเวิร์ค 3 สำหรับภาษาพีเอชพีซึ่งใช้แนวคิด MVC เป็นแนวคิดหลักในการพัฒนาระบบสารสนเทศ

3. วิธีการดำเนินงาน

ในการวิจัยครั้งนี้ผู้วิจัยได้กำหนดวิธีการดำเนินงานวิจัยแบ่งเป็น 4 ขั้นตอนดังนี้ 1) การศึกษาความเป็นไปได้ 2) การวิเคราะห์และออกแบบระบบ 3) พัฒนาระบบ 4) การประเมินประสิทธิภาพของระบบ

3.1 การศึกษาความเป็นไปได้

ในขั้นตอนนี้ผู้วิจัยได้ทำการศึกษาขั้นตอนการดำเนินงานประกันคุณภาพการศึกษาหลักสูตร โดยใช้มหาวิทยาลัยราชภัฏรำไพพรรณีเป็นกรณีศึกษา ดำเนินการโดยศึกษาเกณฑ์การประกันคุณภาพการศึกษาระดับหลักสูตรฉบับปี 2557 รวมถึงวิธีการดำเนินงานประกันคุณภาพระดับหลักสูตรของมหาวิทยาลัยราชภัฏรำไพพรรณี ซึ่งแต่ละหลักสูตรต้องดำเนินการเก็บรวบรวมข้อมูลเอง โดยลักษณะของข้อมูลที่นำมาใช้ประกอบการรายงานผลการดำเนินประกันคุณภาพระดับหลักสูตร รูปแบบของข้อมูลที่ทั้ง ตัวบ่งชี้ที่ต้องรายงานแนวโน้มของผลการดำเนินการเปรียบเทียบผลย้อนหลัง ตัวบ่งชี้ที่ต้องรวบรวมผลการดำเนินงานตลอดปีการศึกษา ดังนั้นการพัฒนาระบบสารสนเทศจึงเป็นเครื่องที่ช่วยให้การรวบรวมวิเคราะห์และนำเสนอข้อมูลสามารถทำได้สะดวกขึ้นในรูปแบบการนำเสนอในลักษณะของ ตาราง กราฟ เป็นต้น โดยเครื่องมือที่นำมาใช้ในการพัฒนาระบบสารสนเทศโดยต้องการหลีกเลี่ยงปัญหาของระบบสารสนเทศใหม่ที่ขัดแย้งกับระบบสารสนเทศที่ใช้งานอยู่บนเครื่องบริการเว็บเดิมจึงได้ใช้การสร้างอินสแตนบนการบริการคลาวด์คอมพิวติงของแอมะซอนเว็บเซอร์วิสเพื่อให้สามารถใช้งานตัวแปลภาษาพีเอชพีเวอร์ชัน 5.6 ขึ้นไปเพื่อให้สามารถใช้งานเซนด์เฟรมเวิร์ค 3 ซึ่งช่วยให้การพัฒนาระบบสารสนเทศที่สามารถใช้งานได้จริง

3.2 การวิเคราะห์ปัญหาและออกแบบระบบ

จากการศึกษาเกณฑ์มาตรฐานตามเกณฑ์ประกับคุณภาพระดับหลักสูตรของคณะกรรรมการการอุดมศึกษาแห่งชาติ ฉบับปีการศึกษา 2557 และขั้นตอนการดำเนินงานประกันคุณภาพในระดับของมหาวิทยาลัยราชภัฏรำไพพรรณีเพื่อนำมาสร้างฐานข้อมูลที่เหมาะกับสำหรับการจัดทำระบบสารสนเทศสำหรับการประกันคุณภาพ โดยส่วนของข้อมูลรายการหลักฐานการดำเนินงานของหลักสูตรแยกตามองค์ประกอบสามารถแจกแจงได้ดังตารางที่ 1

ตารางที่ 1: รายการหลักฐานสำหรับการรายงานผลการดำเนินงานแต่ละองค์ประกอบของเกณฑ์มาตรฐานระดับหลักสูตร คณะกรรมการการอุดมศึกษาแห่งชาติ ปีการศึกษา 2557

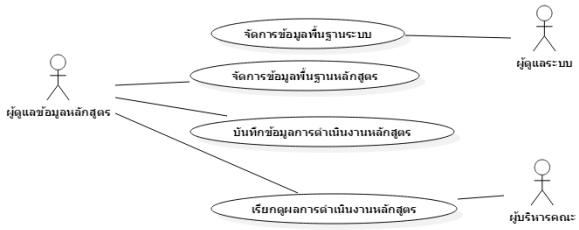
องค์ประกอบ	รายการหลักฐาน
1. การกำกับมาตรฐาน	- รายชื่อกรรมการรับผิดชอบหลักสูตรในแต่ละปีการศึกษา - รายการผลงานวิชาการ 5 ปีย้อนหลัง
2. บัณฑิต	- ข้อมูลการมีงานทำของบัณฑิต - ผลการประเมินความพึงพอใจของผู้ใช้บัณฑิต - รายชื่อสถานประกอบการ
3. นักศึกษา	- รายชื่อกิจกรรมเสริมสร้างคุณลักษณะนักศึกษา - ผลงาน/รางวัลนักศึกษา - ข้อมูลนักศึกษาแรกเข้า - ข้อมูลนักศึกษาที่สำเร็จการศึกษา
4. อาจารย์	- ข้อมูลวุฒิการศึกษาอาจารย์ - ข้อมูลตำแหน่งทางวิชาการ - ข้อมูลผลงานวิชาการในรอบปี
5. หลักสูตร การเรียนการสอน การประเมินผู้เรียน	- วาระการประชุมหลักสูตร - ผลการดำเนินงานตามกรอบมาตรฐานคุณวุฒิแห่งชาติ
6. สิ่งสนับสนุนการเรียนรู้	- ข้อมูลรายการวัสดุ ครุภัณฑ์ หนังสือ สิ่งสนับสนุนการเรียนรู้

ในส่วนของตัวบ่งชี้ประเภทผลลัพธ์ในการรายงานผลต้องมีการเปรียบเทียบผลการดำเนินงานย้อนหลังเพื่อเปรียบเทียบให้เห็นแนวโน้มของผลการดำเนินงานดังตารางที่ 2

ตารางที่ 2: รายการหลักฐานตัวบ่งชี้ประเภทผลลัพธ์ที่ต้องแสดงแนวโน้มการดำเนินงาน

ตัวบ่งชี้	รายการหลักฐาน
3.3 ผลที่เกิดกับนักศึกษา	- การคงอยู่ - การสำเร็จการศึกษา - ความพึงพอใจและผลการจัดการข้อร้องเรียนของนักศึกษา
4.3 ผลที่เกิดกับอาจารย์	- การคงอยู่ของอาจารย์ - ความพึงพอใจของอาจารย์

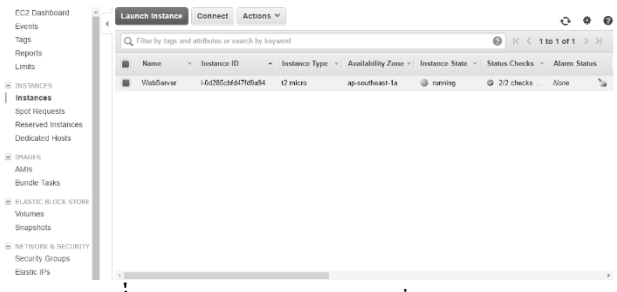
การวิเคราะห์และการออกแบบฐานข้อมูลสำหรับระบบสารสนเทศการประกันคุณภาพระดับหลักสูตรสามารถอธิบายความสามารถของระบบได้จากแผนภาพยูสเคสดังภาพที่ 1



ภาพที่ 1: แผนภาพยูสเคสของระบบ

3.3 พัฒนาระบบ

3.3.1 การพัฒนาเครื่องบริการเว็บผ่านคลาวด์คอมพิวติง ผู้วิจัยใช้บริการคลาวด์คอมพิวติงของแอมะซอนเว็บเซอร์วิส เพื่อสร้างการจำลองเครื่องบริการเว็บเพื่อให้สามารถรองรับ การสร้างเครื่องบริการเว็บที่รองรับตัวแปลภาษาพีเอชพีเวอร์ชัน 5.6 ขึ้นไป โดยได้ทำการสร้างอินสแตนและติดตั้งโปรแกรมบริการเว็บอาปาเช ตัวแปลภาษาพีเอชพี ฐานข้อมูลมายเอสคิวแอล โปรแกรมคอมโพสเซอร์ (Composer) ซึ่งเป็นเครื่องมือสำหรับติดตั้งส่วนขยายของภาษาพีเอชพี และกรอบแนวคิดสำหรับภาษาพีเอชพีเซนด์เฟรมเวิร์ค 3 ดังภาพที่ 2



ภาพที่ 2: การสร้างอินสแตนสำหรับเครื่องบริการเว็บบนแอมะซอนเว็บเซอร์วิส

3.3.2 การพัฒนาระบบสารสนเทศ ผู้วิจัยใช้กรอบแนวคิดสำหรับภาษาพีเอชพีเซนด์เฟรมเวิร์ค เวอร์ชัน 3 ในการพัฒนา ซึ่งมีการใช้แพทเทิร์นแบบ MVC โดยมีการแบ่งส่วนของโปรแกรมที่ในการพัฒนาแยกเป็นโมดูล แต่ละโมดูลจะประกอบด้วยส่วนประกอบสำคัญ 3 ส่วนดังนี้ 1) โมเดล (Model) เป็นส่วนของข้อมูลที่ใช้ในระบบ ซึ่งทำหน้าที่ในการเชื่อมต่อฐานข้อมูล 2) วิว (View) เป็นส่วนของการแสดงผล และทำหน้าที่ติดต่อกับผู้ใช้ 3) คอนโทรลเลอร์ (Controller) เป็นส่วนควบคุมและทำหน้าที่กับข้อมูลที่ได้รับจากผู้ใช้ เพื่อนำมาประมวลผลเพื่อส่งผลลัพธ์ให้กับส่วนแสดงผล

4. ผลการดำเนินงาน

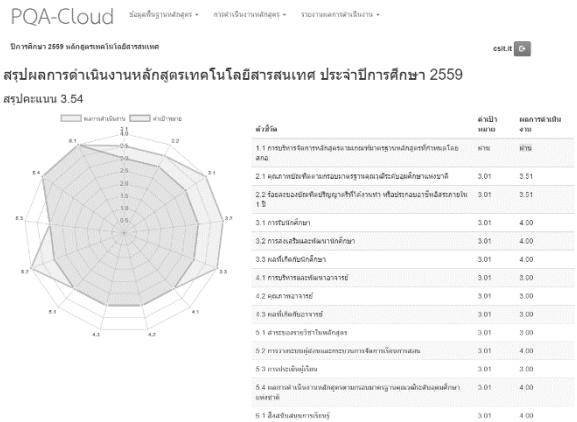
4.1 ผลการพัฒนาระบบ

ผลการพัฒนาระบบสารสนเทศสำหรับการประกันคุณภาพการศึกษาระดับหลักสูตรด้วยคลาวด์คอมพิวติง กรณีศึกษามหาวิทยาลัยราชภัฏรำไพพรรณี ได้มีการแบ่งผู้ใช้งานเป็นสามกลุ่มประกอบด้วย 1) ผู้ดูแลระบบ ซึ่งจะทำหน้าที่ในการจัดการข้อมูลพื้นฐานของระบบ 2) ผู้ดูแลข้อมูลหลักสูตร 3) ผู้บริหารคณะ ในการใช้งานระบบผู้ดูแลข้อมูลหลักสูตรมีหน้าที่บันทึกข้อมูลพื้นฐานและกิจกรรมที่หลักสูตรดำเนินงานเข้าในระบบ ดังภาพที่ 3



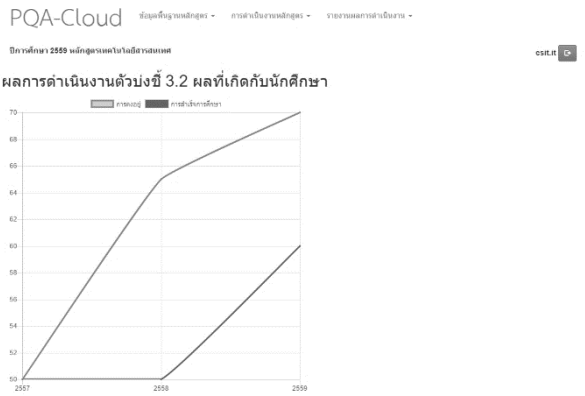
ภาพที่ 3: การบันทึกการดำเนินงานของหลักสูตร

เมื่อหลักสูตรมีการบันทึกกิจกรรมที่ดำเนินการระบบดำเนินการปรับปรุงผลการดำเนินงานลงในฐานข้อมูลและสามารถเรียกดูสรุปผลข้อมูลการดำเนินงานของหลักสูตรในรอบปีการศึกษาผ่านทางการใช้แผนภูมิและตาราง ดังภาพที่ 4



ภาพที่ 4: รายงานผลสรุปผลการดำเนินการหลักสูตรในภาพรวม

ในการแสดงผลการดำเนินงานของตัวบ่งชี้ประเภทผลลัพธ์ เพื่อให้สามารถแสดงแนวโน้มของการดำเนินงานได้อย่างชัดเจนจึงใช้แผนภูมิเส้นการนำเสนอ ดังภาพที่ 5



ภาพที่ 5: แผนภูมิเส้นแสดงผลการดำเนินงานตัวบ่งชี้ประเภทผลลัพธ์

4.2 ผลการประเมินระบบ

ผลการประเมินประสิทธิภาพของระบบโดยผู้ใช้งานระบบจำนวน 10 คน ในการประเมินประสิทธิภาพที่มีหัวข้อในการประเมิน 3 ด้าน ได้แก่ ด้านความสามารถของระบบ (Functional Requirement Test) ด้านการทำงานได้ตามฟังก์ชัน (Functional Test) และด้านความง่ายต่อการใช้งานระบบ (Usability Test) ได้ค่าเฉลี่ย 3.75 และมีส่วนเบี่ยงเบนมาตรฐาน 0.56 ซึ่งอยู่ในระดับดี ดังตารางที่ 2

ตารางที่ 2: ผลการประเมินความพึงพอใจระบบสารสนเทศสำหรับการประกันคุณภาพระดับหลักสูตร

คำถามประเมิน	ระดับความพึงพอใจ		
	ค่าเฉลี่ย	ส่วนเบี่ยงเบนมาตรฐาน	การแปลผล
ด้าน Function Requirement Test	3.88	0.51	ดี
ด้าน Function Test	3.76	0.53	ดี
ด้าน Usability Test	3.61	0.65	ดี
ผลสรุปการประเมิน	3.75	0.56	ดี

5. สรุป

งานวิจัยนี้เป็นการพัฒนาระบบสารสนเทศสำหรับการประกันคุณภาพระดับหลักสูตรบนคลาวด์คอมพิวติง กรณีศึกษามหาวิทยาลัยราชภัฏรำไพพรรณี เพื่อให้หลักสูตรสามารถจัดเก็บข้อมูลพื้นฐานที่มีความจำเป็นสำหรับการประกันคุณภาพการศึกษาในระดับหลักสูตร และเพื่อให้มีความสะดวกในการเข้าถึงข้อมูลจึงเลือกใช้การบริการจำลองเครื่องบริการเว็บของแอมะซอนเว็บเซอร์วิสอีซีทู ซึ่งได้รับการประเมินประสิทธิภาพจากผู้ใช้จำนวน 10 คนในระดับดี ซึ่งสามารถสรุปได้ว่าระบบที่พัฒนาสามารถนำไปใช้งานได้จริง

ข้อเสนอแนะในการพัฒนางานวิจัยนี้คือ เพิ่มความสามารถทำงานร่วมกันตามเวลาจริง และเพิ่มการเรียนรู้ของระบบในการเลือกตัวบ่งชี้ที่เหมาะสมกับกิจกรรมต่าง ๆ ที่หลักสูตรดำเนินงาน

เอกสารอ้างอิง

[1] สำนักงานคณะกรรมการการอุดมศึกษา, *คู่มือการประกันคุณภาพการศึกษาภายในสถานศึกษาระดับอุดมศึกษา พ.ศ. 2557*, กระทรวงศึกษาธิการ, กรุงเทพฯ, 2557.

[2] Adam Shackelford, *Beginning Amazon Web Services with Node.js*, Apress: New York, 2015.

[3] สาธิต ขวัญชุม และ วชิรศักดิ์ โภคากรณ์, "ระบบฐานข้อมูลเพื่อสนับสนุนการซ่อมบำรุงและการบริหารจัดการเครือข่าย กรณีศึกษาระบบเครือข่ายไร้สายมหาวิทยาลัยวลัยลักษณ์" *การประชุมวิชาการระดับชาติด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศ* ครั้งที่ 12. 7-8 กรกฎาคม 2559.

[4] นวลศรี เด่นวัฒนา และอภิสิทธิ์ แสงใส, "ระบบจัดการแฟ้มผลงานอัจฉริยะของอาจารย์ สำหรับสถาบันอุดมศึกษา" *การประชุมวิชาการระดับชาติด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศ* ครั้งที่ 12. 7-8 กรกฎาคม 2559.

[5] เย็นใจ ใช้บางยาง, มณเฑียร รัตนศิริวงศ์วุฒิ, " ระบบสารสนเทศเพื่อการบริหารจัดการวางแผน ติดตามโครงการ และงบประมาณ กรณีศึกษาบริษัท บางกอก อินโนเวชัน เทคโนโลยี จำกัด" *การประชุมวิชาการระดับชาติด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศ* ครั้งที่ 12. 7-8 กรกฎาคม 2559.

[6] ศศิประภา ใจใหญ่ และศักดิ์ชาย ตั้งวรรณวิทย์, "การพัฒนาเว็บ Responsive จองสถานที่ทำงานสำหรับ Startup และการค้นคืนสารสนเทศ" *การประชุมวิชาการระดับชาติด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศ* ครั้งที่ 12. 7-8 กรกฎาคม 2559.

[7] บุรินทร์ รุจจนพันธ์, "การพัฒนาระบบตรวจเยี่ยมการประกันคุณภาพการศึกษาระดับหลักสูตร โดยตัวแบบวงจรชีวิตแบบก้นหอย" *การประชุมวิชาการระดับชาติด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศ* ครั้งที่ 12. 7-8 กรกฎาคม 2559.

[8] ทัศนีย์ รอดมั่นคง, "การพัฒนาระบบการจัดการสารสนเทศผ่านคลาวด์คอมพิวติงตามกรอบมาตรฐานคุณวุฒิระดับอุดมศึกษาแห่งชาติ" วิทยานิพนธ์ปรัชญาดุษฎีบัณฑิต, สาขาวิชาเทคโนโลยีสารสนเทศและการสื่อสารเพื่อการศึกษา คณะครุศาสตร์อุตสาหกรรม มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2558.

การปรับปรุงกระบวนการพัฒนาซอฟต์แวร์เชิงบุคคล: เว็บแอปพลิเคชัน AppDoSI
Personal Software Process Improvement: AppDoSI Web-based Application

นวลศรี เด่นวัฒนา (Nuansri Denwattana)[1] และอภิสิทธิ์ แสงใส (Apisit Saengsai)[2]

[1]คณะวิทยาการสารสนเทศ มหาวิทยาลัยบูรพา จังหวัดชลบุรี

[2]ห้องปฏิบัติการวิจัยระบบสารสนเทศ มหาวิทยาลัยบูรพา

[1]nuansri@buu.ac.th, [2]iserl.center@gmail.com

บทคัดย่อ

การปรับปรุงกระบวนการในการพัฒนาซอฟต์แวร์เชิงบุคคลเกี่ยวข้องกับการจัดการเวลา การประมาณการ การวางแผน การวัดและการประเมินผล ตลอดจนการจัดการคุณภาพซอฟต์แวร์ จำเป็นต้องอาศัย แบบฟอร์ม หรือเครื่องมือที่ถูกออกแบบมาโดยเฉพาะ จากผลงานเผยแพร่ส่วนใหญ่ที่ปรากฏพบว่าเครื่องมือส่วนมากจัดเป็นเครื่องมือแบบ Stand-alone สำหรับเครื่องคอมพิวเตอร์ส่วนบุคคลเท่านั้น ไม่เหมาะสำหรับการประมวลผลและการรายงานผลข้อมูลจำนวนมาก วัตถุประสงค์ของงานวิจัยนี้ เพื่อพัฒนาเครื่องมือสำหรับใช้ปรับปรุงกระบวนการการพัฒนาซอฟต์แวร์เชิงบุคคล โดยใช้ซอฟต์แวร์โอเพนซอร์ส ในลักษณะเว็บแอปพลิเคชัน ซึ่งอาศัยกระบวนการในการพัฒนาซอฟต์แวร์เชิงบุคคล (Personal Software Process: PSP) โดยเครื่องมือประกอบไปด้วยฟังก์ชันงานหลักคือ การตั้งค่าข้อมูลพื้นฐาน การตั้งค่าโครงการซอฟต์แวร์ การจัดการกระบวนการพัฒนาซอฟต์แวร์ส่วนบุคคล และการวิเคราะห์และรายงานสมรรถนะการพัฒนาซอฟต์แวร์ ผู้วิจัยได้นำเครื่องมือดังกล่าวไปใช้งานจริงกับนิสิตหลักสูตรวิศวกรรมซอฟต์แวร์ มหาวิทยาลัยบูรพา พบว่าอาจารย์และนิสิตมีความพึงพอใจต่อการใช้งานอยู่ในระดับดี (4.25/5) นอกจากนี้รายงานสมรรถนะส่วนบุคคลของนิสิต ยังแสดงให้เห็นว่า นิสิตมีสมรรถนะในการพัฒนาซอฟต์แวร์ที่ดีขึ้น

คำสำคัญ: กระบวนการในการพัฒนาซอฟต์แวร์เชิงบุคคล การปรับปรุงกระบวนการในการพัฒนาซอฟต์แวร์เชิงบุคคล โอเพนซอร์สซอฟต์แวร์

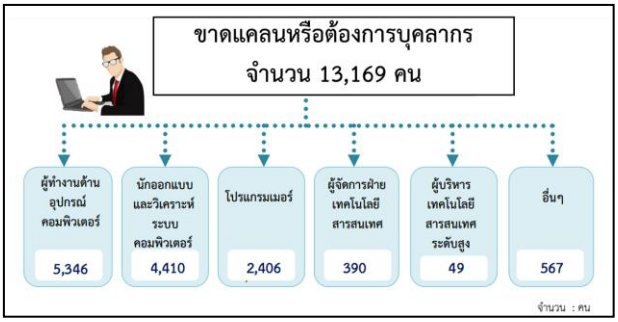
Abstract

The Personal Software Process (PSP) improvement can be used for most aspects of software work, including time scheduling, estimation, planning, and measurement, evaluation and software quality. The PSP process consists of various type of forms and specific tools. Standalone Application was the most widely PSP's tools nowadays. However, that is not appropriate for large scale data processing and generate reports from the big amount of data. The objective of this research was to develop open source web-based application for the Personal Software Process (PSP) improvement. The application consists of 4 sub-systems, that are application setting, software project setting, personal software management, and data analysis and software development performance reports. The sample of the research is software engineering students from Burapha University. The result according to user satisfaction is 4.25/5. In addition, from the result of personal performance report show that the performance in software development of the students are improved.

Keyword: Personal software process, Personal software process improvement, Open source software

1. บทนำ

ในปัจจุบันแทบทุกองค์กรได้มีการนำเทคโนโลยีสารสนเทศเข้ามามีส่วนร่วมในการดำเนินงาน เพื่อสร้างความได้เปรียบทั้งในเชิงธุรกิจ และการปฏิบัติงาน แต่อย่างไรก็ตาม จากผลการสำรวจของสำนักงานสถิติแห่งชาติ ประเทศไทย ในปี พ.ศ. 2558 พบว่าสถานประกอบการยังขาดแคลนบุคลากรด้านเทคโนโลยีสารสนเทศจำนวนมาก [1] ดังภาพที่ 1



ภาพที่ 1: จำนวนบุคลากรที่ปฏิบัติหน้าที่ด้านเทคโนโลยีสารสนเทศที่สถานประกอบการรายงานว่าขาดแคลน

จากภาพที่ 1 แสดงการขาดแคลนบุคลากรที่ปฏิบัติหน้าที่ด้านเทคโนโลยีสารสนเทศของสถานประกอบการ โดยมุ่งประเด็นไปที่นักออกแบบและวิเคราะห์ระบบสารสนเทศ นักวิเคราะห์ระบบคอมพิวเตอร์ และโปรแกรมเมอร์ พบว่ามีการขาดแคลนถึง 6,816 คน ซึ่งสาเหตุหนึ่งของการขาดแคลนคือการย้ายสายงาน เพราะเนื่องจากสายงานนี้ต้องศึกษาหาความรู้ใหม่ๆ อยู่เสมอ ต้องมีวินัยในตัวเอง มีการฝึกฝนและพัฒนาตนเองอยู่ตลอดเวลา

บทความนี้เป็นการนำเสนอเว็บแอปพลิเคชัน AppDoSI ที่ช่วยในการฝึกฝน และปรับปรุงกระบวนการในการพัฒนาซอฟต์แวร์ โดยใช้หลักการกระบวนการพัฒนาซอฟต์แวร์เชิงบุคคล (Personal Software Process: PSP) ซึ่งถูกเสนอโดยสถาบันวิศวกรรมซอฟต์แวร์ (Software Engineering Institute: SEI) มหาวิทยาลัย Carnegie Mellon ประเทศสหรัฐอเมริกา เว็บแอปพลิเคชันดังกล่าวสามารถเข้าใช้งานผ่านอินเทอร์เน็ต และเว็บเบราว์เซอร์ ซึ่งสะดวกในการใช้งานมากกว่าเครื่องมือแบบ Stan-alone พร้อมทั้งสามารถวิเคราะห์และรายงานข้อมูลจำนวนมากในรูปแบบที่เหมาะสม

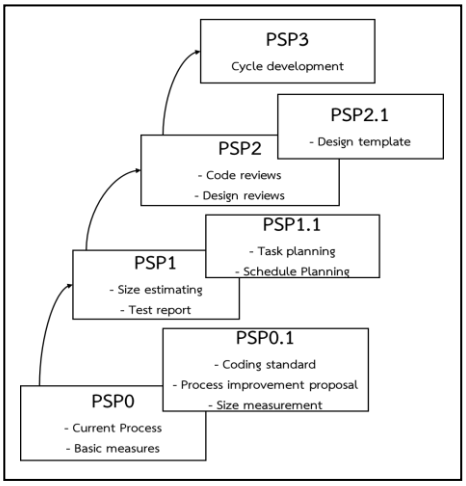
ลำดับในการจัดเรียงเนื้อหาในบทความนี้ ประกอบด้วยส่วนที่ 1 และส่วนที่ 2 จะนำเสนอเกี่ยวกับแนวคิดเรื่องกระบวนการพัฒนาซอฟต์แวร์เชิงบุคคล และหลักการอื่นๆ ที่เกี่ยวข้อง จากนั้นในส่วนที่ 3 จะนำเสนอเว็บแอปพลิเคชัน AppDoSI รวมถึงภาพรวมการทำงาน และลำดับการทำงานของระบบ ในส่วนของผลลัพธ์ในการพัฒนาจะแสดงในส่วนที่ 4 และสุดท้ายสรุปและแผนงานในอนาคตจะแสดงในส่วนที่ 5

2. กรอบแนวคิด หลักการ และทฤษฎี ที่เกี่ยวข้องที่เกี่ยว

เพื่อให้เข้าใจถึงส่วนที่เกี่ยวข้องกับเครื่องมือที่นำเสนอ ในส่วนนี้จะเป็นการอธิบายกรอบแนวคิดของกระบวนการในการพัฒนาซอฟต์แวร์เชิงบุคคล (Personal Software Process: PSP) และการปรับปรุงกระบวนการซอฟต์แวร์ (Software Process Improvement) ที่ใช้เป็นแนวทางในงานวิจัยนี้

2.1 กระบวนการในการพัฒนาซอฟต์แวร์เชิงบุคคล (Personal Software Process: PSP)

[2], [3] PSP เป็นกระบวนการที่สนับสนุนนักพัฒนาซอฟต์แวร์ หรือนักวิศวกรรมซอฟต์แวร์ ให้พัฒนาคุณภาพกระบวนการพัฒนาซอฟต์แวร์ของตนเอง โดยการควบคุม บริหารจัดการ และปรับปรุงสมรรถนะการพัฒนาซอฟต์แวร์ของตนเองให้มีคุณภาพมากขึ้น โดยอาศัยกรอบการดำเนินงานจากแบบฟอร์ม คำแนะนำ และกระบวนการที่ใช้ในการพัฒนาซอฟต์แวร์อยู่ในปัจจุบัน



ภาพที่ 2: ระดับของกระบวนการ PSP

จากภาพระดับของกระบวนการ PSP ที่ปรากฏในภาพที่ 2 ซึ่งประกอบด้วย PSP0, PSP0.1, PSP1, PSP1.1, PSP2, PSP2.1 และ PSP3 มีรายละเอียดในภาพรวมดังต่อไปนี้

- **PSP0**: เกี่ยวข้องกับกระบวนการรายบุคคลเบื้องต้นและแนะนำให้รู้จักวัฏจักร 6 ขั้นตอน (Phase) ในการพัฒนาซอฟต์แวร์ได้แก่ วางแผน (Planning) ออกแบบ (Design) การเขียนโปรแกรม (Code) คอมไพล์ (Compile) ทำทดสอบ (Test) และการสรุปปิดงานและทำรายงานต่างๆ (Post mortem: PM)
- **PSP0.1**: เกี่ยวกับการวัด (Measurement) ในเรื่องต่างๆ ในการพัฒนาซอฟต์แวร์ เช่น การวัดขนาดของ

ซอฟต์แวร์ (Size measurement) และการวัดเวลาในการพัฒนาซอฟต์แวร์ (Time measurement) เป็นต้น นอกจากนี้ยังเสนอให้ทำ Coding standard โดยกำหนดรูปแบบของการเขียนโปรแกรม ทำให้เกิดความคงเส้นคงวา ช่วยในการวัดขนาดของซอฟต์แวร์

- **PSP1**: เน้นเทคนิคการจัดการโครงการรายบุคคล โดยแนะนำการประมาณขนาดของซอฟต์แวร์ (Size estimate) ที่คาดว่าจะใช้ โดยวิธี PROBE (Proxy-Based estimating)
- **PSP1.1**: เกี่ยวข้องกับการวางแผนงาน (Task planning) การวางแผนตารางเวลา (Schedule planning) และการติดตาม (Tracking)
- **PSP2**: มีการเพิ่มขั้นตอน (Phase) การทบทวนการออกแบบ (Design review) และการทบทวนการเขียนโปรแกรม (Code review) เพื่อช่วยให้พบข้อบกพร่องได้เร็วขึ้นในกระบวนการ ทำให้การแก้ปัญหาได้เร็วและประหยัดเวลา
- **PSP2.1**: ชี้ให้เห็นถึงความจำเป็นของการออกแบบ โดยเพิ่มส่วนของ Design notation ที่มีการกำหนดแม่แบบ (Template) ของการออกแบบจำนวน 4 แบบ ดังต่อไปนี้ ข้อกำหนดทางการปฏิบัติ (The Operation Specification Template: OST) ข้อกำหนดทางฟังก์ชัน (The Functional Specification Template: FST) ข้อกำหนดทางสถานะ (The State Specification Template: SST) และข้อกำหนดทางตรรกะ (The Logic Specification Template: LST)

2.2 การปรับปรุงกระบวนการซอฟต์แวร์ (Software Process Improvement)

เกี่ยวข้องกับกระบวนการพัฒนาซอฟต์แวร์ เป็นการเปลี่ยนแปลงกระบวนการของซอฟต์แวร์ เพื่อให้สอดคล้องกับองค์กร โดยจะมีการเลือกใช้หรือกำหนดแบบจำลองในการพัฒนาซอฟต์แวร์ขององค์กร และมีการปรับปรุงอย่างต่อเนื่อง

[4], [5] ซึ่งในปัจจุบันแบบจำลองที่เกี่ยวข้องกับการปรับปรุงกระบวนการการพัฒนาซอฟต์แวร์ที่เป็นที่นิยมคือ CMMI (Capability Maturity Model Integration) ของสถาบันวิศวกรรมซอฟต์แวร์ (Software Engineering Institute: SEI) โดยใช้เป็นเครื่องมือในการประเมินขีดความสามารถด้านกระบวนการขององค์กรว่าอยู่ในระดับใด และใช้เป็นเครื่องมือในการวางแผนปรับปรุงองค์กรให้มีขีดความสามารถในการทำงานเพิ่มขึ้นเป็นลำดับ ซึ่ง CMMI ประกอบไปด้วย

- **Process Area (PA)** เป็นกลุ่มกระบวนการหลักที่ใช้ในการทำงานแต่ละระดับ ตามลักษณะการดำเนินงานและเป้าหมายที่ได้กำหนดไว้
- **Specific Goal (SG)** เป็นกระบวนการในการกำหนดเป้าหมายเฉพาะเจาะจงของแต่ละกระบวนการ โดยใช้เป็นแนวทางในการปรับปรุงกระบวนการ
- **Specific Practices (SP)** เป็นกระบวนการการปฏิบัติงานที่มีความเฉพาะเจาะจง โดยใช้เป็นแนวทางในการปฏิบัติงาน
- **Generic Goal (GG)** เป็นกระบวนการกำหนดเป้าหมายทั่วไป รวมถึงผลลัพธ์ของแต่ละกระบวนการ
- **Generic Practices (GP)** เป็นกระบวนการการปฏิบัติงานทั่วไป รวมถึงผลลัพธ์ของแต่ละการปฏิบัติงาน
- **Sub Practice** เป็นรายละเอียดเพื่อชี้แนะการปฏิบัติงานย่อยของแต่ละกลุ่มกระบวนการ

นอกจากนี้ CMMI ยังแบ่งระดับวุฒิภาวะของการดำเนินงานออกเป็น 5 ระดับ ดังต่อไปนี้

CMMI Level 1 (Initial) ระดับวุฒิภาวะระดับที่ 1 การทำงานปัจจุบันมันจะดำเนินงานตามความพึงพอใจและไม่เป็นขั้นตอน ความสำเร็จขององค์กรมักจะขึ้นกับความสามารถและความเก่งของบุคลภายในองค์กร ไม่มีกระบวนการ หรือระเบียบในการปฏิบัติงาน

CMMI Level 2 (Managed) ระดับวุฒิภาวะระดับที่ 2 องค์กรมีการจัดการและวางแผนกระบวนการทำงาน ปฏิบัติตามมาตรฐาน และมีการควบคุม เริ่มมีนโยบายในการบริหารจัดการโครงการซอฟต์แวร์ และทุกกระบวนการจะถูกจัดเก็บไว้ซึ่งเมื่อมีการปฏิบัติทุกกระบวนการเหล่านั้นจะถูกนำออกมาดำเนินงานตามแผนที่ได้กำหนดไว้

CMMI Level 3 (Defined) ระดับวุฒิภาวะระดับที่ 3 องค์กรสร้างมาตรฐานเพื่อให้แน่ใจว่าจะบรรลุเป้าหมายของแผนที่กำหนดไว้ในระดับวุฒิภาวะระดับที่ 2 ซึ่งมีการปรับปรุงมาตรฐานอย่างต่อเนื่อง รวมไปถึงการอธิบายรายละเอียดของแต่ละกระบวนการในเชิงลึก เพื่อสร้างความเข้าใจที่มากขึ้น

CMMI Level 4 (Quantitatively Managed) ระดับวุฒิภาวะระดับที่ 4 องค์กรบรรลุเป้าหมายหลักทั้งหมดที่ได้กำหนดไว้ โดยการสร้างบรรทัดฐานในการจัดการจุดหมายในเชิงปริมาณ ทำให้เข้าใจเกี่ยวกับคุณภาพ และการทำงานของกระบวนการ โดยนำหลักการทางสถิติมาใช้เพื่อวิเคราะห์การทำงาน เพื่อใช้ในการทำนายหรือค้นหาสาเหตุการเปลี่ยนแปลงของกระบวนการทั้งในทางที่ดีขึ้นและแย่ลง

CMMI Level 5 (Optimizing) ระดับวุฒิภาวะระดับที่ 5 องค์กรมีการปรับปรุงกระบวนการอย่างต่อเนื่องโดยมีลักษณะคล้ายกับระดับวุฒิภาวะระดับที่ 4 แต่จะครอบคลุมในระดับสาเหตุของปัญหาทั้งหมด

2.3 The Software Process Dashboard: PDash

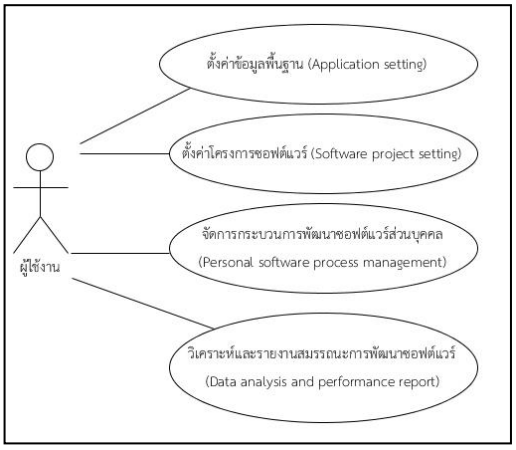
PDash เป็นเครื่องมือที่ใช้สำหรับกระบวนการพัฒนาซอฟต์แวร์เชิงบุคคล ถูกพัฒนาขึ้นโดย SEI และเริ่มใช้งานในปี พ.ศ. 2548 จัดเป็นซอฟต์แวร์โอเพนซอร์ส ประเภท Stand-alone การทำงานหลักของระบบครอบคลุมเรื่องของการวางแผน การติดตามการทำงาน การวิเคราะห์ข้อมูล และการรายงานผลการดำเนินงาน ที่เกี่ยวข้องกับการพัฒนาซอฟต์แวร์เชิงบุคคล [6]

3. เว็บแอปพลิเคชัน AppDoSI

เนื่องจากการทำงานของระบบมีลักษณะการทำงานที่ต่อเนื่องกันเป็นลำดับ เนื้อหาในส่วนนี้จึงเป็นการอธิบายรายละเอียดการทำงานหลักของระบบ และภาพรวมลำดับการทำงานของระบบ รวมไปถึงการออกแบบระบบโดยใช้แผนภาพต่างๆ ที่เกี่ยวของกับระบบ

3.1 ฟังก์ชันการทำงานหลักของระบบ

มอดูลการทำงานหลักของระบบมีจำนวนทั้งสิ้น 4 ส่วน ดังแสดงในภาพที่ 3



ภาพที่ 3: Use-case diagram

จากภาพที่ 3 แสดงแผนภาพ Use-case diagram มอดูลหลักของระบบ ซึ่งประกอบไปด้วย การตั้งค่าข้อมูลพื้นฐาน (Application setting) การตั้งค่าโครงการซอฟต์แวร์ (Software Project setting) การจัดการกระบวนการพัฒนาซอฟต์แวร์ส่วนบุคคล (Personal software process management) และการวิเคราะห์และรายงานสมรรถนะการพัฒนาซอฟต์แวร์ (Data analysis and performance report)

1) การตั้งค่าข้อมูลพื้นฐาน (Application setting): เกี่ยวข้องกับการจัดการข้อมูลเบื้องต้นสำหรับใช้เป็นข้อมูลในการดำเนินงานต่างๆ ภายในระบบ เพื่อให้ระบบเกิดความสม่ำเสมอและเป็นไปในทิศทางเดียวกัน เช่น การกำหนดประเภทข้อบกพร่อง การกำหนดขั้นตอนการพัฒนา (Phases) การกำหนดรูปแบบในการพัฒนา เป็นต้น

2) การตั้งค่าโครงการซอฟต์แวร์ (Software project setting): เกี่ยวข้องกับการสร้างและกำหนดโครงการซอฟต์แวร์ (Software Project) หรืองานที่ต้องการพัฒนา โดยในแต่ละโครงการซอฟต์แวร์ สามารถแบ่งส่วนการพัฒนาออกเป็นส่วนย่อยๆ รวมไปถึงยังสามารถเลือกระดับของกระบวนการ PSP ซึ่งประกอบไปด้วย PSP0, PSP0.1, PSP1, PSP1.1, PSP2, PSP2.1 และ PSP3 (ดังแสดงในภาพที่ 2) ซึ่งจะส่งผลต่อมอดูลการจัดการกระบวนการพัฒนาซอฟต์แวร์ส่วนบุคคลตามระดับที่เลือก

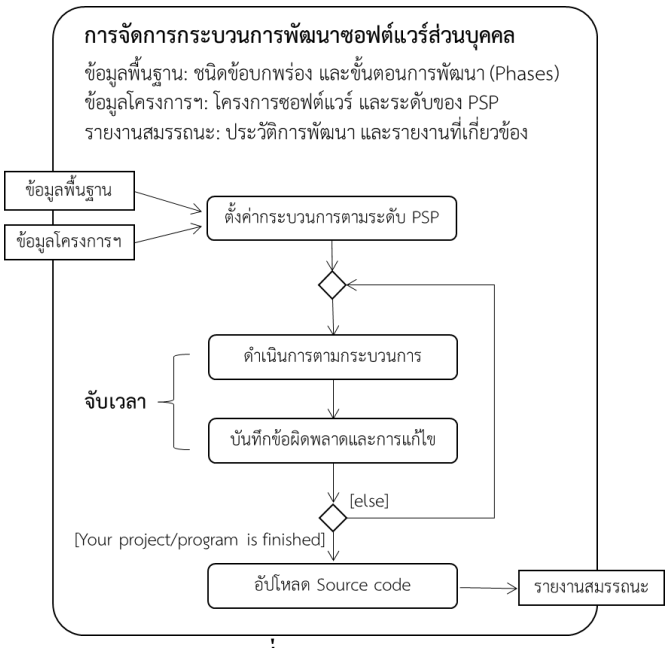
3) การจัดการกระบวนการพัฒนาซอฟต์แวร์ส่วนบุคคล (Personal software process management): การทำงานในมอดูลนี้จะขึ้นอยู่กับการเลือกระดับของ PSP ในมอดูลการตั้งค่าโครงการซอฟต์แวร์ก่อนหน้านี้ หากเลือก PSP ในระดับใด ระบบจะนำกระบวนการที่สอดคล้องกับ PSP ในระดับดังกล่าวมาให้ผู้ใช้ได้ดำเนินตามกระบวนการนั้น กล่าวคือ หากผู้ใช้เลือก PSP ในระดับ 0 ระบบจะเลือกกระบวนการที่เกี่ยวข้องกับกระบวนการรายบุคคลเบื้องต้นและแนะนำให้รู้จักวัฏจักร 6 ขั้นตอน (Phase) ในการพัฒนาซอฟต์แวร์ได้แก่ วางแผน ออกแบบ การเขียนโปรแกรม คอมไพล์ ทำทดสอบ และการสรุปปิดงานและทำรายงานต่างๆ และเช่นเดียวกัน หากผู้ใช้เลือก PSP ในระดับ 0.1 ระบบจะเพิ่มกระบวนการที่เกี่ยวข้องกับการวัด ในเรื่องต่างๆ ในการพัฒนาซอฟต์แวร์ เช่น การวัดขนาดของซอฟต์แวร์ และการวัดเวลาในการพัฒนาซอฟต์แวร์ให้ผู้ใช้ได้ดำเนินการ นอกจากนี้ในระหว่างดำเนินการตาม

กระบวนการที่เลือก ระบบจะให้ผู้ใช้จับเวลา และให้ผู้ใช้บันทึกข้อบกพร่องที่พบตลอดจนกว่าจะเสร็จสิ้นการพัฒนาซอฟต์แวร์ โดยระบบจะให้อัปโหลด Source code เพื่อทำการเก็บประวัติในการพัฒนาและนำไปวิเคราะห์ผลในรายงานต่างๆ ต่อไป

4) การวิเคราะห์และรายงานสมรรถนะการพัฒนาซอฟต์แวร์ (Data analysis and performance report): ในมอดูลนี้ระบบจะทำการประมวลผลข้อมูลต่างๆ ที่ผู้ใช้ได้บันทึกไว้ในแต่ละโครงการ โดยเน้นไปที่การแสดงรายงานสรุปสมรรถนะในมิติต่างๆ ด้วยลักษณะของแผงควบคุมข้อมูลส่วนบุคคล (Personal Dashboard) จำแนกตามโครงการซอฟต์แวร์ โดยมีการแสดงผลในรูปแบบของกราฟ (Chart-like) ตาราง (Tabular) และการ์ดข้อความ (Card) ตัวอย่างเช่น กราฟรายงานเวลาทั้งหมดที่ใช้ในการพัฒนา กราฟความแม่นยำในการประมาณการเวลา กราฟแสดงเวลาทั้งหมดในขั้นตอนต่างๆ (Phases) กราฟความแม่นยำในการประมาณการขนาด กราฟแสดงข้อบกพร่องที่พบและการกำจัดข้อบกพร่องในขั้นตอนต่างๆ (Phases) เป็นต้น

3.2 ภาพรวมลำดับการทำงานของระบบ

เนื่องจากระบบมีขั้นตอนการดำเนินงานอย่างเป็นลำดับ ดังนั้นหากมีการดำเนินการข้าม หรือละเลยการบันทึกข้อมูลแล้วนั้น จะส่งผลกระทบถึงการวิเคราะห์ผล นำไปสู่การรายงานผลที่ผิดพลาด ในหัวข้อนี้จะนำเสนอภาพรวมลำดับการทำงานของระบบดังภาพที่ 4



ภาพที่ 4: Activity Diagram

จากภาพที่ 4 แสดงลำดับการทำงานโดยรวมของระบบ: การจัดการกระบวนการพัฒนาซอฟต์แวร์เชิงบุคคล โดยเริ่มต้นจากผู้ใช้งานต้องกำหนดข้อมูลพื้นฐาน โดยการกำหนดนี้จะถูกนำไปใช้ในทุกๆ โครงการซอฟต์แวร์ และผู้ใช้ต้องสร้างโครงการซอฟต์แวร์ โดยสามารถแบ่งเป็นส่วนย่อยๆ สำหรับการพัฒนาแบบแบ่งส่วนการพัฒนา ตลอดจนผู้ใช้ต้องเลือก PSP ในระดับต่างๆ โดยมีข้อแม้ว่า หากในส่วนงานก่อนหน้าเลือก PSP ในระดับใดๆ ส่วนงานถัดมาไม่สามารถเลือกระดับ PSP ที่น้อยกว่าระดับนั้นๆ ได้ จากนั้นให้ผู้ใช้ทำการเริ่มกระบวนการพัฒนาซอฟต์แวร์เชิงบุคลตามระดับ PSP ที่เลือกไว้ โดยตลอดระยะเวลาในการดำเนินงาน ระบบจะให้ผู้ใช้จับเวลา และให้ผู้ใช้บันทึกข้อผิดพลาดที่พบทั้งหมด ตลอดจนการแก้ไขข้อผิดพลาดนั้น ผู้ใช้สามารถหยุด หรือพักการพัฒนาได้ โดยให้กดปุ่มหยุดการพัฒนา และสามารถเริ่มต้นการพัฒนาต่อได้เมื่อกดปุ่มเริ่มพัฒนาต่อ และเมื่อผู้ใช้พัฒนาซอฟต์แวร์จนสำเร็จระบบจะให้ผู้ใช้ต้องอัปโหลดไฟล์ Source code เพื่อใช้เปรียบเทียบและวิเคราะห์สมรรถนะในส่วนของการรายงานสมรรถนะในลำดับถัดไป

4. ผลการดำเนินงาน

การพัฒนาเว็บแอปพลิเคชัน AppDoSI ได้ประยุกต์ใช้หลักการพัฒนาซอฟต์แวร์แบบ Agile และนำไปใช้จริงกับนิสิตระดับปริญญาตรี ในรายวิชาที่เกี่ยวข้องกับการพัฒนาซอฟต์แวร์ โดยเนื้อหาในส่วนนี้จะกล่าวถึงเครื่องมือที่ใช้ในการพัฒนา ผลการนำเว็บแอปพลิเคชัน AppDoSI ไปใช้งานจริง และตัวอย่างหน้าจอระบบ

4.1 เครื่องมือที่ใช้ในการพัฒนา

เว็บแอปพลิเคชัน AppDoSI พัฒนาโดยใช้เครื่องมือที่เป็นโอเพนซอร์สซอฟต์แวร์ ใช้ภาษา PHP เวอร์ชัน 5.4 ฐานข้อมูลที่ใช้คือ MySQL เวอร์ชัน 5.5 และทำการพัฒนาระบบฯ บนระบบปฏิบัติการ CentOS เวอร์ชัน 6.7

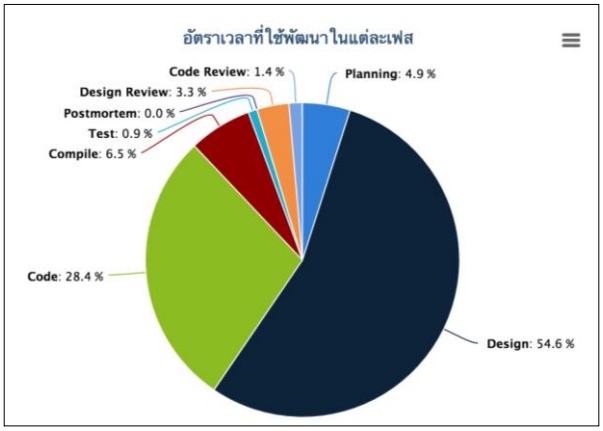
4.2 ผลการนำเว็บแอปพลิเคชัน AppDoSI ไปใช้งานจริง

ระบบดังกล่าวได้นำไปใช้งานและทำการประเมินโดยใช้แบบสอบถามความพึงพอใจการใช้งานระบบ ณ หลักสูตรวิทยาศาสตรบัณฑิต สาขาวิชาวิศวกรรมซอฟต์แวร์ คณะวิทยาการสารสนเทศ มหาวิทยาลัยบูรพา โดยครอบคลุมนิสิตชั้นปีที่ 2 ถึงชั้นปีที่ 3 จำนวนกว่า 300 คน ที่ได้ลงทะเบียนรายวิชา 1) 888248 ปฏิบัติการกระบวนการพัฒนาซอฟต์แวร์เชิงบุคคล

2) 888249 ค่ายฝึกพัฒนาซอฟต์แวร์ 3) 888352 ปฏิบัติการกระบวนการพัฒนาซอฟต์แวร์เชิงกลุ่มงาน 4) 888321 การบริหารโครงการซอฟต์แวร์ พบว่าอาจารย์และนิสิตเคยใช้งานระบบมีความพึงพอใจต่อการใช้งานอยู่ในระดับดี (4.25/5)

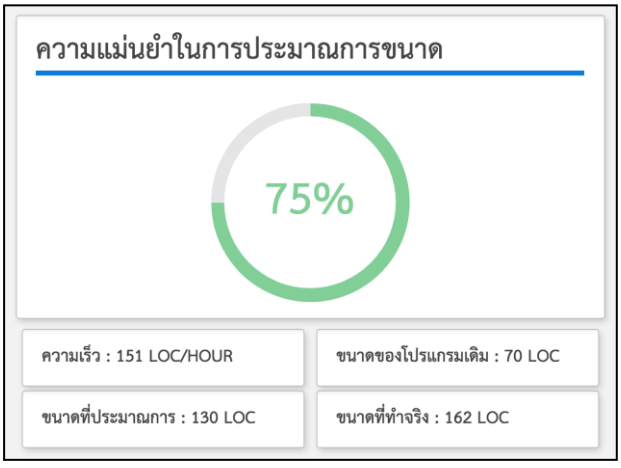
4.3 ตัวอย่างหน้าจอเว็บแอปพลิเคชัน AppDoSI

เนื่องจากพื้นที่ในการนำเสนอมีจำกัด ในส่วนนี้จะแสดงตัวอย่างภาพหน้าจอมอดูลการวิเคราะห์และรายงานสมรรถนะการพัฒนาซอฟต์แวร์ดังภาพที่ 5 – ภาพที่ 6



ภาพที่ 5: กราฟวงกลมแสดงอัตราเวลาที่ใช้พัฒนาในแต่ละเฟส

จากภาพที่ 5 แสดงอัตราของเวลา โดยจำแนกตามเฟสในรูปแบบกราฟวงกลม ซึ่งแสดงให้เห็นว่านิสิตคนดังกล่าวใช้เวลาในการออกแบบ (Design) ไปถึง 54.6 % ของเวลาทั้งหมด รองลงมาคือการเขียนซอฟต์แวร์ (Code) ใช้เวลาไป 28.4 % แสดงให้เห็นว่านิสิตใช้เวลาในการออกแบบเป็นจำนวนมาก ในการพัฒนาซอฟต์แวร์ครั้งถัดไปนิสิตควรพิจารณาเรื่องการออกแบบให้เหมาะสมมากยิ่งขึ้น



ภาพที่ 6: กราฟวงกลมแสดงอัตราเวลาที่ใช้พัฒนาในแต่ละเฟส

จากภาพที่ 6 แสดงกราฟวงกลม และกล่องข้อความ แสดงความแม่นยำในการประมาณการขนาดซอฟต์แวร์ และความเร็วในการพัฒนา ซึ่งนิสิตประมาณการขนาดซอฟต์แวร์ได้ใกล้เคียงถึง 75 % เมื่อเทียบกับขนาดซอฟต์แวร์จริง อีกทั้งข้อมูลในภาพยังแสดงถึงความเร็วของนิสิตในการเขียนซอฟต์แวร์ อยู่ที่ 151 บรรทัดต่อ 1 ชั่วโมง

5. สรุปและแผนงานในอนาคต

บทความนี้เป็นการนำเสนอเว็บแอปพลิเคชัน AppDoSI ซึ่งแสดงให้เห็นถึงภาพรวมการทำงานของระบบ และลำดับการทำงานของระบบ ตลอดจนการนำระบบไปใช้งานจริงกับนิสิต นอกจากนี้ผู้วิจัยกำลังพัฒนาระบบในส่วนต่างๆ เพิ่มขึ้น อาทิ เช่น การแสดงสมรรถนะการประมาณการด้วยเทคนิค PROxy-Based Estimating (PROBE) การปรับปรุงการนับจำนวนบรรทัด (Line Of Code) ด้วยไลบรารี่ CLOC อีกทั้งยังได้รับความต้องการจากอาจารย์ผู้สอนหลักสูตรวิศวกรรมซอฟต์แวร์ คณะวิทยาการสารสนเทศ มหาวิทยาลัยบูรพา ให้พัฒนาระบบในส่วนการติดตามสมรรถนะนิสิตโดยอาจารย์สามารถติดตามนิสิตได้ทั้งรายบุคคล และภาพรวมในชั้นเรียน ในท้ายนี้ คณะผู้วิจัยขอบคุณทีมงานจากห้องปฏิบัติการวิจัยระบบสารสนเทศ มหาวิทยาลัยบูรพา นิสิตผู้มีส่วนร่วมในการพัฒนาซอฟต์แวร์นี้ผ่านรายวิชาโครงงานวิศวกรรมซอฟต์แวร์ และคณาจารย์สาขาวิชาวิศวกรรมซอฟต์แวร์ คณะวิทยาการสารสนเทศ มหาวิทยาลัยบูรพา ที่มีส่วนร่วมในความสำเร็จดังกล่าว

เอกสารอ้างอิง

[1] Economic and Social Statistics Bureau, “The 2015 establishment on the use of information and communication technology,” *National Statistical Office of Thailand*, pp. 16 -18, 2015.

[2] Watt S. Humphrey, *PSP: A self-improvement process for software engineers*. Addison-Wesley, 2005.

[3] Watt S. Humphrey, “The Personal Software Process,” *Carnegie Mellon University*, pp. 1-55, November 2000.

[4] Software Engineering Institute, “CMMI-DEV,” *Carnegie Mellon University*, pp. 16-47, August 2006.

[5] M.B. Chrissis, M.Konrad, and S.Shrum, *CMMI Guidelines for Process Integration and Product Improvement*, Addison-Wesley: Boston, 2003.

[6] The Software Process Dashboard, “Process Dashboard” [ออนไลน์], สืบค้นวันที่ 16 พฤษภาคม พ.ศ. 2560, แหล่งที่มา: http://www.processdash.com/

สาเหตุของสินค้าขาดมือและการแบ่งกลุ่ม ABC แบบหลายเกณฑ์ กรณีศึกษาธุรกิจค้าปลีก
Stock-out Causes and Multiple-Criteria ABC Analysis: A Case Study of Retail Businesses

วรรณวิสา ครุฑไทย (Wanvisa Krutthai)[1] ฉัฐไชย์ ลีนาวงศ์ (Chatchai Leenawong) [2]
และบัณฑิต วรรธนาภา (Bunthit Watanapa)[3]
[1,3]สาขาวิชาระบบสารสนเทศทางธุรกิจ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี

บทคัดย่อ

งานวิจัยนี้ทำการศึกษาเกี่ยวกับสาเหตุหลักของสินค้าขาดมือในธุรกิจค้าปลีก มีวัตถุประสงค์เพื่อลดจำนวนสินค้าขาดมือและเพิ่มประสิทธิภาพการจัดการสินค้า ซึ่งเป็นปัจจัยสำคัญของการจัดการซัพพลายเชนโดยรวม จากการศึกษาพบว่าสาเหตุหลักของการเกิดสินค้าขาดมือ ได้แก่ การไม่จัดลำดับความสำคัญของสินค้าและข้อมูลการพยากรณ์อุปสงค์ที่ผิดพลาด ดังนั้น คณะผู้วิจัยจึงได้นำเสนอหลักการแบ่งกลุ่มสินค้าเพื่อจัดลำดับความสำคัญ ด้วยการแบ่งกลุ่มแบบเอบีซีทั้งแบบเกณฑ์เดียวและแบบหลายเกณฑ์ (Single and Multiple Criteria ABC Analysis) ตามทฤษฎีของพาเรโต โดยเกณฑ์ที่สนใจคือ มูลค่าสินค้าต่อหน่วย และความถี่ของสินค้าขาดมือใน 1 ปี การจัดกลุ่มสินค้าตามหลักการ ABC แบบหลายเกณฑ์นี้ทำให้สามารถพิจารณาถึงปัจจัยที่ส่งผลต่อการจัดลำดับความสำคัญของสินค้ามากกว่า 1 ปัจจัย และทำให้ทราบว่าสินค้ากลุ่มไหนต้องดูแลเป็นพิเศษ เพื่อป้องกันไม่ให้เกิดภาวะสินค้าขาดมือ ยิ่งไปกว่านั้น การแบ่งกลุ่มสินค้ายังช่วยลดงานในการพยากรณ์ให้น้อยลง และสามารถบริหารจัดการสินค้าให้ดียิ่งขึ้นต่อไป

คำสำคัญ: สินค้าขาดมือ ธุรกิจค้าปลีก การแบ่งกลุ่มเอบีซีแบบหลายเกณฑ์

Abstract

In this research, major causes of stock-outs in retail businesses are investigated. The objective is to reduce shortages and simultaneously improve supply management's efficiency. Major causes of short supply include not classifying the products according to some crucial criteria and inaccurate demand forecasting. The ABC Analysis, both Single- and Multiple-criteria, based on Pareto's law is adapted. The criteria of interest are the product's unit value and the frequency of stock-outs within the year. The advantage of doing the multiple criteria is that more than one affecting factors can be considered at the same time. Consequently, the business owners can know exactly which products should be looked after more than the others. It also reduces the demand on forecasting work while keeping the inventory management in order.

Keyword: Multi-criteria ABC Analysis, Stock-out, Decision Support, Supply Chain Management.

1. บทนำ

ปัญหาการขาดแคลนสินค้าในธุรกิจค้าปลีก (Retailer) นั้นสะท้อนให้เห็นถึงความผิดพลาดบางกระบวนการในระบบห่วงโซ่อุปทาน (Supply Chain) แล้วยังเป็นปัญหาที่สำคัญของร้านค้าปลีกอีกด้วย เนื่องจากการขาดแคลนสินค้าในขณะที่ผู้บริโภคมีความต้องการสินค้า จะส่งผลให้สูญเสียโอกาสทางการขายและส่งผลต่อความพึงพอใจของลูกค้าอีกด้วย [2]

การจัดการห่วงโซ่อุปทาน (Supply Chain Management) เป็นการบริหารกระบวนการทางธุรกิจตั้งแต่ต้นน้ำ (Upstream) ไปจนถึงปลายน้ำ (Downstream) เพื่อให้เกิดประสิทธิภาพสูงสุด รวมถึงการตอบสนองความต้องการของลูกค้า [7] โดยโครงสร้างห่วงโซ่อุปทานประกอบไปด้วยซัพพลายเออร์ ผู้ผลิต ศูนย์กระจายสินค้า ผู้ค้าปลีก และลูกค้า ซึ่งจะเห็นได้ว่าผู้ค้าปลีกเป็นส่วนที่ติดต่อกับลูกค้าโดยตรง หากผู้ค้าปลีกไม่สามารถตอบสนองความต้องการของลูกค้าได้ทันก็จะส่งผลกระทบต่อ

ระบบห่วงโซ่อุปทาน เมื่อไม่มีลูกค้า ร้านค้าจึงขายสินค้าไม่ได้ ส่งผลให้ผู้ผลิตลดการผลิตลง ดังนั้นจึงต้องมีการจัดกลุ่มความสำคัญของสินค้าและบริหารจัดการสินค้าคงคลังเพื่อป้องกันไม่ให้เกิดสินค้าขาดมือในขณะที่ลูกค้ามีความต้องการซื้อสินค้า

งานวิจัยนี้ทำการศึกษาเกี่ยวกับสาเหตุหลักของสินค้าขาดมือในธุรกิจค้าปลีก เมื่อสามารถระบุปัญหาได้จะนำไปสู่วิธีการป้องกันและแก้ไขปัญหาที่ถูกต้อง ซึ่งจะสามารถช่วยให้ธุรกิจค้าปลีกลดค่าเสียโอกาสจากการสูญเสียการขาย แก้ไขปัญหาได้ตรงจุด และยังเป็นการเพิ่มความพึงพอใจให้กับลูกค้าอีกด้วย จากการศึกษาพบว่าปัญหาอย่างหนึ่งที่พบบ่อยคือการไม่แบ่งกลุ่มสินค้า แยกแยะตามความสำคัญด้วยเกณฑ์บางอย่าง รวมไปถึงการพยากรณ์ผิดพลาด คณะผู้วิจัยจึงได้นำเสนอหลักการแบ่งกลุ่มสินค้าคงคลังโดยอาศัยทฤษฎีของพาเรโต โดยวิเคราะห์การแบ่งกลุ่มแบบเอบีซีเข้ามาช่วยจัดกลุ่มความสำคัญของสินค้า ซึ่งจะนำเสนอทั้งแบบเกณฑ์เดียวและแบบหลายเกณฑ์ โดยจะใช้เกณฑ์มูลค่าของสินค้าต่อหน่วย และเกณฑ์ความถี่หรือจำนวนเดือนที่สินค้าเกิดการขาดมือใน 1 ปี การจัดกลุ่มสินค้านี้จะสามารถทำให้ทราบว่าสินค้ากลุ่มไหนเป็นสินค้าสำคัญที่ควรดูแลไม่ให้เกิดภาวะสินค้าขาดมือ และเพื่อการบริหารจัดการสินค้าให้มีคุณภาพยิ่งขึ้น [8]

2. ทฤษฎีพื้นฐานและงานวิจัยที่เกี่ยวข้อง

การขาดแคลนสินค้าบนชั้นวางหรือสินค้าขาดมือ หมายถึงสถานการณ์ที่ผู้บริโภคไม่สามารถซื้อสินค้าที่ต้องการได้เนื่องจากไม่มีสินค้า ส่งผลให้เกิดความไม่พึงพอใจและเพิ่มการสูญเสียยอดขาย ซึ่งเหตุการณ์สินค้าขาดมือในร้านค้าปลีกเกิดจากความผิดพลาดของบางกระบวนการในระบบห่วงโซ่อุปทาน ได้แก่ การขาดแคลนสินค้าบนชั้นวาง การเติมเต็มสินค้าไม่ถูกต้อง การประมาณการณ์ผิด และการกระจายสินค้าไม่มีประสิทธิภาพ เนื่องจากไม่แบ่งกลุ่มสินค้าแยกแยะตามความสำคัญ [2]

Adusei and Awunyo-Vitor (2014) ศึกษาปัจจัยของสินค้าขาดมือของร้านค้าปลีกในประเทศกานา พบว่าสินค้าขาดมือเกิดจากการเติมเต็มสินค้าและการสั่งซื้อผิดพลาด การพยากรณ์ความต้องการผิดพลาด และความล่าช้าจากซัพพลายเออร์ [3]

Luis (2009) ศึกษาสาเหตุหลักของสินค้าขาดมือของซูเปอร์มาร์เก็ตในรัฐเซาเปาลู ประเทศบราซิล พบว่าสาเหตุหลักของสินค้าขาดมือเกิดจากสินค้ากลุ่มขายดีไม่ได้ถูกเติมบนชั้นวางสินค้า การขาดแคลนชั้นวางสินค้า ซัพพลายเออร์ส่งสินค้าล่าช้า ส่งสินค้าไม่ถูกต้อง [1]

Ehrenthal and Stölzle (2013) ศึกษาปัญหาในร้านค้าปลีก พบว่าสาเหตุของสินค้าขาดมือ ได้แก่ การไม่แยกแยะสินค้าตามความสำคัญ ทำให้การเติมเต็มสินค้าล่าช้า จัดวางสินค้าผิดตำแหน่ง การพยากรณ์ผิดพลาด การจัดส่งผิดพลาด การรับสินค้าผิดพลาด การเจ็บป่วยของพนักงาน พนักงานไม่แจ้งเมื่อสินค้าหมด [4]

Amy, ZZ. (2002) วิจัยการออกแบบคลังสินค้าเพื่อเพิ่มประสิทธิภาพการจัดวางสินค้า กรณีศึกษาบริษัท Brierly Lombard and CompanyInc. (B&L) โดยใช้วิธีการแบ่งกลุ่มแบบเอบีซีตามเกณฑ์ความเคลื่อนไหวของสินค้า แบ่งสินค้าออกเป็น 3 กลุ่ม ได้แก่ กลุ่ม A, B, และ C ซึ่งมีการเคลื่อนไหวเร็วที่สุดไปจนถึงช้าที่สุดตามลำดับ และทำการวางสินค้ากลุ่ม A ให้อยู่ในตำแหน่งใกล้ประตูมากที่สุด ถัดไปเป็นสินค้ากลุ่ม B และกลุ่ม C ตามลำดับ ผลจากการศึกษาพบว่าการปรับตำแหน่งในการจัดเก็บสินค้าสามารถลดระยะทางลงได้ถึง 554 วินาที และใช้เวลาน้อยกว่า 60 วินาทีในการเข้าถึงสินค้ากลุ่ม A ทำให้จัดส่งสินค้าให้กับลูกค้าได้อย่างรวดเร็วตรงเวลา [6]

3. วิธีการดำเนินงาน

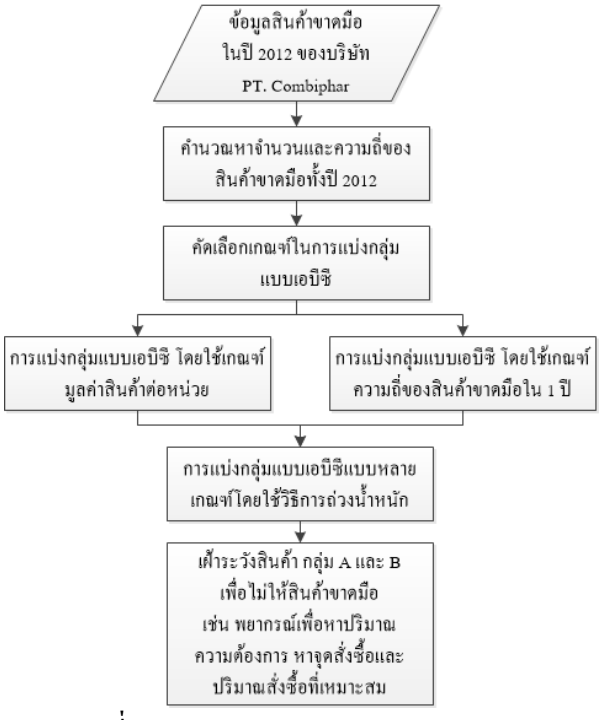
จากการศึกษาทฤษฎีที่เกี่ยวข้องข้างต้น พบว่าสาเหตุหลักอย่างหนึ่งของสินค้าขาดมือในร้านค้าปลีกคือการไม่ได้จัดลำดับความสำคัญของสินค้าแต่ละชนิด ส่งผลให้การพยากรณ์ความต้องการสินค้าผิดพลาด และเกิดภาวะสินค้าขาดมือ คณะผู้วิจัยจึงได้นำหลักการวิเคราะห์การแบ่งกลุ่มเอบีซีแบบหลายเกณฑ์เข้ามาช่วยจัดกลุ่มความสำคัญของสินค้า เพื่อให้ทราบว่าสินค้ากลุ่มไหนเป็นสินค้าสำคัญที่ควรดูแลติดตามอย่างใกล้ชิดเพื่อป้องกันการเกิดสินค้าขาดมือ

ในการศึกษานี้ได้นำข้อมูลสินค้าขาดมือในปี 2012 เป็นระยะเวลา 12 เดือน ของบริษัท PT. Combiphar ประเทศอินโดนีเซีย [5] มาใช้เป็นกรณีศึกษา โดยทำการวิเคราะห์การแบ่งกลุ่มเอบีซีแบบเกณฑ์เดียวและแบบหลายเกณฑ์ โดยแบ่ง

ตารางที่ 1: Stock-Out Products 2012

Product	Jan	Feb	March	April	May	June	July	Aug	Sep	Oct	Nov	Dec	Total QTY	FREQ	Value (RP)
CTS3	37,000	13,150		2,921		23,887		28,700	8,928	13,450			128,036	7	3,424,963,000
LCC1		6,818	69,888	17,918									94,624	3	2,732,170,000
LCC4		57,090		53,363									110,453	2	1,988,154,000
TYA3				2,254	7,080	10,980	45,219						65,533	4	1,753,007,750
DBM1		81						320	160	1,040	960		2,561	5	1,526,868,200
LCM2									13	127	768	5,376	6,284	4	1,474,648,800
PTN3	1,687	3,685		1,999	345								7,716	4	1,097,556,900
LCC8		9,796	9,120	1,932		1							20,849	4	898,507,000
LVR2		4,372	10,600		4,120								19,092	3	840,048,000
RHI4	5,054	1,697	3,080	4,487		4,400		4,700					23,418	6	796,212,000
OBA8				111,698									111,698	1	755,078,480
PHX5				7,005		24	18,912			49	4,872		30,862	5	740,688,000
DNM2							21,058	51,400					72,458	2	724,580,000
PSY1							1,960	1,830					3,790	2	696,412,500
LAC1			52	10,680									10,732	2	665,384,000
FSC1						2,472	15,847						18,319	2	632,005,500
OBJ6						84,635							84,635	1	550,127,500
FTJ4	129				7,503	1,959							9,591	3	527,505,000
DBX1				1,145	3,612							539	5,296	3	522,762,200
NPN1				193	924	1,134		145					2,396	4	503,160,000

ระดับความสำคัญจากมากไปน้อยคือ A, B และ C ตามลำดับ เพื่อจัดลำดับความสำคัญของสินค้าที่ควรดูแลติดตามอย่างใกล้ชิดเพื่อป้องกันการเกิดสินค้าขาดมือ ดังตารางที่ 1 ซึ่งแสดงถึงจำนวนสินค้าขาดมือในแต่ละเดือน จำนวนรวมของสินค้าขาดมือใน 1 ปี จำนวนครั้งของสินค้าขาดมือใน 1 ปี และมูลค่ารวมของสินค้าขาดมือแต่ละชนิดใน 1 ปี โดยมีกระบวนการดำเนินงานการจัดกลุ่มแบบเอบีซี ดังภาพที่ 1



ภาพที่ 1: กระบวนการดำเนินงานการจัดกลุ่มแบบเอบีซี

จากภาพที่ 1 แสดงการดำเนินงานจัดกลุ่มแบบเอบีซีดังนี้

1) ศึกษาชุดข้อมูลสินค้าขาดมือ
2) คัดเลือกเกณฑ์ในการแบ่งกลุ่มเอบีซี โดยคณะผู้วิจัยได้เลือกเกณฑ์มูลค่าสินค้าต่อหน่วย และเกณฑ์ความถี่ของสินค้าขาดมือใน 1 ปี
3) ทำการแบ่งกลุ่มแบบเอบีซีโดยใช้เกณฑ์มูลค่าสินค้าต่อหน่วย และเกณฑ์ความถี่ของสินค้าขาดมือใน 1 ปี
4) นำผลลัพธ์ที่ได้จากข้อ 3) มากำหนดน้ำหนักความสำคัญของแต่ละเกณฑ์ และทำการแบ่งกลุ่มเอบีซีแบบหลายเกณฑ์โดยใช้วิธีการถ่วงน้ำหนัก
5) เฝ้าระวังสินค้ากลุ่ม A และ B ที่ได้จากการจัดกลุ่มในข้อ 4) เพื่อไม่ให้เกิดสินค้าขาดมือ เช่น พยากรณ์เพื่อหาปริมาณความต้องการ หาจุดสั่งซื้อและปริมาณสั่งซื้อที่เหมาะสม เป็นต้น

3.1 การแบ่งกลุ่มแบบเอบีซีโดยเกณฑ์มูลค่าสินค้าต่อหน่วย

เป็นการแบ่งกลุ่มสินค้าตามมูลค่าของสินค้าขาดมือ สินค้าที่มีมูลค่าสูงจะส่งผลกระทบกับรายได้ของบริษัทเป็นอย่างมาก จัดเป็นสินค้าที่มีความสำคัญ จำเป็นต้องดูแลติดตามอย่างใกล้ชิด

คณะผู้วิจัยได้นำข้อมูลสินค้าขาดมือดังตารางที่ 1 มาทำการคำนวณเพื่อแบ่งกลุ่มสินค้า โดยมีขั้นตอนดังนี้

1) คำนวณหาค่ามูลค่าสินค้าต่อหน่วย (Unit Value) โดย Unit Value = Value(RP) / Total QTY

2) เรียงลำดับข้อมูลตามมูลค่าสินค้าต่อหน่วยจากมากไปน้อย
3) คำนวณหาค่าเปอร์เซ็นต์ของมูลค่าสินค้าต่อหน่วย (Percent Unit Value)
4) คำนวณหามูลค่าสะสมของค่าเปอร์เซ็นต์ของมูลค่าสินค้าต่อหน่วย (Cumulative Percent Unit Value)
5) นำค่าเปอร์เซ็นต์สะสมของมูลค่าสินค้าต่อหน่วยเทียบเกณฑ์เพื่อแบ่งกลุ่ม A, B และ C ตามทฤษฎีของพาเรโต โดยกำหนดให้ $0 \leq A \leq 75$, $75 < B \leq 90$ และ $90 < C \leq 100$

เมื่อทำการคำนวณตามขั้นตอนแล้ว จะสามารถแสดงผลลัพธ์ที่ได้จากการจัดกลุ่มด้วยเกณฑ์มูลค่าสินค้าต่อหน่วยดังตารางที่ 2 ซึ่งจะเห็นว่าจากสินค้าทั้งหมด 20 รายการ เมื่อทำการจัดกลุ่มด้วยเกณฑ์มูลค่าสินค้าต่อหน่วย จะมีสินค้ากลุ่ม A จำนวน 5 รายการ ได้แก่ DBM1, LCM2, NPN1, PSY1, PTN3 มีสินค้ากลุ่ม B จำนวน 5 รายการ ได้แก่ DBX1, LAC1, FTJ4, LVR2, LCC8 และมีสินค้าที่เหลือเป็นกลุ่ม C จำนวน 10 รายการ

ตารางที่ 2: การแบ่งกลุ่มเอบีซีโดยเกณฑ์มูลค่าสินค้า

Product	Unit Value	Percent Unit Value	Cumulative unit value	Unit Value Rank
DBM1	596200.00	31.62%	31.62%	A
LCM2	234667.22	12.44%	44.06%	A
NPN1	210000.00	11.14%	55.19%	A
PSY1	183750.00	9.74%	64.94%	A
PTN3	142244.28	7.54%	72.48%	A
DBX1	98708.87	5.23%	77.72%	B
LAC1	62000.00	3.29%	81.00%	B
FTJ4	55000.00	2.92%	83.92%	B
LVR2	44000.00	2.33%	86.25%	B
LCC8	43095.93	2.29%	88.54%	B
FSC1	34500.00	1.83%	90.37%	C
RHI4	34000.00	1.80%	92.17%	C
LCC1	28873.96	1.53%	93.70%	C
CTS3	26750.00	1.42%	95.12%	C
TYA3	26750.00	1.42%	96.54%	C
PHX5	24000.00	1.27%	97.81%	C
LCC4	18000.00	0.95%	98.77%	C
DNM2	10000.00	0.53%	99.30%	C
OBA8	6760.00	0.36%	99.66%	C
OBJ6	6500.00	0.34%	100.00%	C

3.2 การแบ่งกลุ่มแบบเอบีซีโดยเกณฑ์ความถี่ของสินค้าขาดมือใน 1 ปี

เป็นการแบ่งกลุ่มสินค้าตามจำนวนครั้งที่สินค้าขาดมือในรอบ 1 ปี สินค้าที่มีความถี่ของสินค้าขาดมือมากจัดเป็นสินค้าขายดี จึงเป็นสินค้าที่มีความสำคัญ จำเป็นต้องดูแลติดตามเช่นกัน คณะผู้วิจัยนำข้อมูลสินค้าขาดมือดังตารางที่ 1 มาทำการคำนวณเพื่อแบ่งกลุ่มสินค้า โดยมีขั้นตอนดังนี้

1) คำนวณหาค่าอัตราการเกิดสินค้าขาดมือต่อปีโดย Stock-outs Ratio = No. of Stock-outs months / 12
2) จัดเรียงลำดับข้อมูลตามอัตราการเกิดสินค้าขาดมือต่อปีจากมากไปน้อย
3) คำนวณหาค่าเปอร์เซ็นต์อัตราการเกิดสินค้าขาดมือต่อปี (Percent Stock-outs Ratio)
4) คำนวณหามูลค่าสะสมของค่าเปอร์เซ็นต์อัตราการเกิดสินค้าขาดมือต่อปี (Cumulative Percent Stock-outs Ratio)
5) แบ่งกลุ่ม A, B และ C ตามมูลค่าสะสมของค่าเปอร์เซ็นต์อัตราการเกิดสินค้าขาดมือต่อปี โดยกำหนดให้ $0 \leq A \leq 65$, $65 < B \leq 85$, และ $85 < C \leq 100$

ตารางที่ 3: การแบ่งกลุ่มเอบีซีโดยเกณฑ์ความถี่ของสินค้าขาดมือใน 1 ปี

Product	No. of stock months	Stock ratio	Percent Stock ratio	Cumulative % Stock ratio	FREQ Rank
CTS3	7.00	0.58	10.45%	10.45%	A
RHI4	6.00	0.50	8.96%	19.40%	A
DBM1	5.00	0.42	7.46%	26.87%	A
PHX5	5.00	0.42	7.46%	34.33%	A
TYA3	4.00	0.33	5.97%	40.30%	A
LCM2	4.00	0.33	5.97%	46.27%	A
PTN3	4.00	0.33	5.97%	52.24%	A
LCC8	4.00	0.33	5.97%	58.21%	A
NPN1	4.00	0.33	5.97%	64.18%	A
LCC1	3.00	0.25	4.48%	68.66%	B
LVR2	3.00	0.25	4.48%	73.13%	B
FTJ4	3.00	0.25	4.48%	77.61%	B
DBX1	3.00	0.25	4.48%	82.09%	B
LCC4	2.00	0.17	2.99%	85.07%	C
DNM2	2.00	0.17	2.99%	88.06%	C
PSY1	2.00	0.17	2.99%	91.04%	C
LAC1	2.00	0.17	2.99%	94.03%	C
FSC1	2.00	0.17	2.99%	97.01%	C
OBA8	1.00	0.08	1.49%	98.51%	C
OBJ6	1.00	0.08	1.49%	100.00%	C

เมื่อทำการคำนวณตามขั้นตอนแล้ว จะสามารถแสดงผลลัพธ์ที่ได้จากการจัดกลุ่มด้วยเกณฑ์ความถี่ของสินค้าขาดมือใน 1 ปี ดัง

ตารางที่ 3 ซึ่งจะเห็นว่าจากสินค้าทั้งหมด 20 รายการ เมื่อทำการจัดกลุ่มด้วยเกณฑ์ความถี่ของสินค้าขาดมือใน 1 ปี จะมีสินค้ากลุ่ม A จำนวน 9 รายการ ได้แก่ CTS3, RHI4, DBM1, PHX5, TYA3, LCM2, PTN3, LCC8, NPN1 มีสินค้ากลุ่ม B จำนวน 4 รายการ ได้แก่ LCC1, LVR2, FTJ4, DBX1 และสินค้าที่เหลือเป็นกลุ่ม C จำนวน 7 รายการ

3.3 การแบ่งกลุ่มเอบีซีแบบหลายเกณฑ์โดยใช้วิธีการถ่วงน้ำหนัก

จากการแบ่งกลุ่มเอบีซีแบบเกณฑ์เดียวทั้งสองแบบข้างต้น ทำให้สินค้าบางตัวอยู่กันคนละกลุ่ม เมื่อพิจารณาด้วยเกณฑ์คนละแบบ เช่น สินค้า PSY1 เมื่อพิจารณาด้วยเกณฑ์มูลค่าสินค้าจะพบว่าอยู่ในกลุ่ม A แต่เมื่อพิจารณาด้วยเกณฑ์ความถี่ของสินค้าขาดมือ สินค้า PSY1 จะอยู่ในกลุ่ม C

เพื่อเพิ่มความถูกต้องในการแบ่งกลุ่มสินค้า ให้แต่ละสินค้าอยู่ในกลุ่มใดกลุ่มหนึ่งเพียงกลุ่มเดียว คณะผู้วิจัยจึงเลือกใช้วิธีการแบ่งกลุ่มเอบีซีแบบหลายเกณฑ์ โดยใช้วิธีการถ่วงน้ำหนักเข้ามาช่วยในการแบ่งกลุ่ม โดยกำหนดระดับความสำคัญให้กับเกณฑ์แต่ละเกณฑ์ โดยให้เกณฑ์มูลค่าสินค้าต่อหน่วยมีความสำคัญมากกว่าเกณฑ์ความถี่ของสินค้าขาดมือใน 1 ปี เพียงเล็กน้อย นั่นคือ กำหนดให้น้ำหนักของเกณฑ์มูลค่าสินค้าเท่ากับ 0.6 และน้ำหนักเกณฑ์ความถี่ของสินค้าขาดมือเท่ากับ 0.4 ด้วยวิธีนี้ จึงทำให้สามารถรวมความสำคัญของทั้งสองเกณฑ์ไว้ด้วยกัน และทำให้ได้กลุ่มเอบีซีใหม่กับสินค้าแต่ละตัวเพียงกลุ่มเดียว โดยมีการคำนวณตามขั้นตอนดังนี้

1) นำกลุ่มผลลัพธ์ที่ได้จาก 3.1 และ 3.2 แปลงเป็นคะแนน โดยกำหนดให้สินค้ากลุ่ม A = 4.0 สินค้ากลุ่ม B = 3.0 และสินค้ากลุ่ม C = 2.0
2) คำนวณค่า = (0.6)(คะแนนที่แปลงโดยเกณฑ์มูลค่าสินค้าต่อหน่วย) + (0.4)(คะแนนที่แปลงโดยเกณฑ์ความถี่ของสินค้าขาดมือใน 1 ปี)
3) นำคะแนนที่ได้จากขั้นที่ 2) เทียบเกณฑ์เพื่อแบ่งกลุ่ม A, B และ C ตามทฤษฎีของพาเรโต โดยกำหนดให้ $0 < C \leq 2.0$, $2.0 < B \leq 3.0$, และ $3.0 < A \leq 4.0$

เมื่อทำการคำนวณตามขั้นตอนแล้ว จะสามารถแสดงผลลัพธ์ที่ได้จากการจัดกลุ่มด้วยเกณฑ์มูลค่าสินค้าต่อหน่วย ดังตารางที่ 4 ซึ่งจะเห็นว่าจากสินค้าทั้งหมด 20 รายการ เมื่อทำการจัดกลุ่มแบบหลายเกณฑ์ จะมีสินค้ากลุ่ม A จำนวน 6 รายการ ได้แก่ DBM1, LCM2, PTN3, LCC8, PSY1, NPN1 มีสินค้ากลุ่ม B จำนวน 8 รายการ ได้แก่ CTS3, LCC1, TYA3, LVR2, RHI4, PHX5, LAC1, FTJ4, DBX1 และสินค้าที่เหลือเป็นกลุ่ม C จำนวน 5 รายการ

ตารางที่ 4: การแบ่งกลุ่มเอบีซีแบบหลายเกณฑ์โดยใช้วิธีการถ่วงน้ำหนัก

Product	Unit Value Weighted = 0.6		FREQ Weighted = 0.4		Weight Point	Multiple Rank
	Grp.	Point.	Grp.	Point.		
CTS3	C	2.00	A	4.00	2.80	B
LCC1	C	2.00	B	3.00	2.40	B
LCC4	C	2.00	C	2.00	2.00	C
TYA3	C	2.00	A	4.00	2.80	B
DBM1	A	4.00	A	4.00	4.00	A
LCM2	A	4.00	A	4.00	4.00	A
PTN3	A	4.00	A	4.00	4.00	A
LCC8	B	3.00	A	4.00	3.40	A
LVR2	B	3.00	B	3.00	3.00	B
RHI4	C	2.00	A	4.00	2.80	B
OBA8	C	2.00	C	2.00	2.00	C
PHX5	C	2.00	A	4.00	2.80	B
DNM2	C	2.00	C	2.00	2.00	C
PSY1	A	4.00	C	2.00	3.20	A
LAC1	B	3.00	C	2.00	2.60	B
FSC1	C	2.00	C	2.00	2.00	C
OBJ6	C	2.00	C	2.00	2.00	C
FTJ4	B	3.00	B	3.00	3.00	B
DBX1	B	3.00	B	3.00	3.00	B

3.4 ผลการดำเนินงาน

จากการศึกษางานวิจัยที่เกี่ยวข้องพบว่าสาเหตุของสินค้าขาดมือในร้านค้าปลีกที่พบบ่อยคือการไม่แบ่งกลุ่มสินค้าแยกแยะตามความสำคัญด้วยเกณฑ์บางอย่าง รวมไปถึงการพยากรณ์ผิดพลาด คณะผู้วิจัยจึงนำเสนอการแบ่งกลุ่มแบบเอบีซีทั้งแบบเกณฑ์เดียวและแบบหลายเกณฑ์เพื่อให้ทราบว่าสินค้าชนิดใดมีความสำคัญและจำเป็นที่จะต้องดูแลติดตามอย่างใกล้ชิด โดยจากผลการดำเนินงานพบว่าเมื่อใช้เกณฑ์มูลค่าสินค้าต่อหน่วยจะมีสินค้ากลุ่ม A จำนวน 5 รายการ และสินค้ากลุ่ม B จำนวน 5 รายการ เมื่อใช้เกณฑ์ความถี่ของสินค้าขาดมือใน 1 ปี จะมีสินค้ากลุ่ม A จำนวน 9 รายการ และสินค้ากลุ่ม B จำนวน 4 รายการ แต่เมื่อนำทั้งสองเกณฑ์มาคำนวณโดยใช้วิธีการถ่วงน้ำหนัก โดยกำหนดให้น้ำหนักของเกณฑ์มูลค่าสินค้าเท่ากับ 0.6 และน้ำหนักเกณฑ์ความถี่ของสินค้าขาดมือเท่ากับ 0.4 พบว่ามีสินค้า

กลุ่ม A จำนวน 6 รายการ และสินค้ากลุ่ม B จำนวน 8 รายการ ดังแสดงในตารางที่ 4

จะเห็นว่าการใช้การแบ่งกลุ่มเอบีซีแบบหลายเกณฑ์จะทำให้ได้พิจารณาถึงปัจจัยที่ส่งผลต่อการจัดลำดับความสำคัญของสินค้ามากกว่า 1 ปัจจัย ทำให้แต่ละสินค้าอยู่ในกลุ่มใดกลุ่มหนึ่งเพียงกลุ่มเดียว เช่น สินค้า PSY1 เมื่อพิจารณาด้วยเกณฑ์มูลค่าสินค้าจะพบว่าอยู่ในกลุ่ม A แต่เมื่อพิจารณาด้วยเกณฑ์ความถี่ของสินค้าขาดมือ สินค้า PSY1 จะอยู่ในกลุ่ม C และเมื่อใช้วิธีการแบ่งกลุ่มเอบีซีแบบหลายเกณฑ์จะทำให้สามารถจัดกลุ่มสินค้า PSY1 ให้อยู่ในกลุ่ม A เพียงกลุ่มเดียวได้ และยังทำให้ทราบว่าสินค้ากลุ่มใดต้องดูแลเป็นพิเศษ โดยสินค้ากลุ่ม A และกลุ่ม B เป็นกลุ่มสินค้าที่ควรให้ความสำคัญดูแลเป็นพิเศษทั้งในด้านการพยากรณ์อุปสงค์และการจัดการสินค้าคงคลังให้มีประสิทธิภาพ โดยไม่ต้องกระทำกับสินค้าทุกชนิดที่มีซึ่งจะเป็นการเสียเวลาและเปลืองทรัพยากรเป็นอย่างยิ่ง โดยเฉพาะในกรณีที่ธุรกิจมีสินค้าที่หลากหลาย

4. บทสรุปและข้อเสนอแนะ

งานวิจัยนี้ทำการศึกษาสาเหตุหลักที่ทำให้เกิดสินค้าขาดมือ โดยพบว่าสาเหตุของสินค้าขาดมือในร้านค้าปลีกที่พบบ่อยคือการไม่แบ่งกลุ่มสินค้าแยกแยะตามความสำคัญด้วยเกณฑ์บางอย่าง รวมไปถึงการพยากรณ์ผิดพลาด ซึ่งภาวะสินค้าขาดมือนี้จะส่งผลกระทบต่อรายได้ของบริษัท ภาพลักษณ์ของธุรกิจ และยังส่งผลกระทบต่อความพึงพอใจของลูกค้า ทั้งยังสะท้อนให้เห็นถึงความผิดพลาดบางกระบวนการในระบบห่วงโซ่อุปทานอีกด้วย ในงานวิจัยได้เน้นถึงการจัดความสำคัญของสินค้าคงคลังโดยใช้การแบ่งกลุ่มแบบเอบีซีเข้ามาช่วยจัดกลุ่มสินค้า เพื่อให้ทราบถึงรายการสินค้าที่จะต้องดูแลเป็นพิเศษ ได้แก่สินค้าในกลุ่ม A และ กลุ่ม B ซึ่งเหมาะสมเป็นอย่างยิ่งกับธุรกิจค้าปลีกที่มักมีสินค้าหลากหลายรายการ ทำให้ช่วยประหยัดต้นทุนในการดำเนินงาน ทำให้ธุรกิจสามารถแข่งขันในตลาดเสรีได้เป็นอย่างดี

อย่างไรก็ดีเมื่อทำการจัดกลุ่มความสำคัญของสินค้าแล้ว ควรนำสินค้าเหล่านี้ไปทำการพยากรณ์เพื่อหาปริมาณความต้องการที่แท้จริง แล้วจึงคำนวณหาจุดสั่งซื้อ และปริมาณสั่งซื้อที่เหมาะสมที่สุด เพื่อการบริหารจัดการให้มีคุณภาพมากยิ่งขึ้น ทั้งในด้านต้นทุนการจัดการ และการตอบสนองความต้องการของลูกค้า

เอกสารอ้างอิง

[1] Luis Henrique Rigato Vasconcellos, and Mauro Sampaio, “The Stockouts Study: an Examination of the Extent and the Causes in the São Paulo Supermarket Sector,” *BAR, Curitiba*, v.6, n.3, art.6, p.263-279, July/Sept. 2009.

[2] Battista, C., Falsini, D., Scarabotti, L., Schiraldi, M.M. “Quantifying shelf-out-of-stock in fashion & apparel retail stores,” *Breaking down the barriers between research and industry*, Abano Terme, Padua (Italy), 14-16 September 2011.

[3] Charles Adusei and Dadson Awunyo-Vitor, “Determinants of Stock-Out in Retail Shops in Ghana,” *Modern Economy*, 5, 1240-1252, 2014.

[4] Joachim C.F. Ehrenthal and Wolfgang Stölzle, “An-examination-of-the-causes-for-retail-stockouts,” *International Journal of Physical Distribution & Logistics Management*, Vol. 43 Iss 1 pp. 54 - 69, 2013.

[5] Mekel Christine, Anantadjaya Samuel P. D. and Lahindah Laura, “Stock Out Analysis: An Empirical Study on Forecasting, Re-Order Point and Safety Stock Level at PT. Combiphar, Indonesia,” *Review of Integrative Business and Economics Research*, Vol. 3(1), p.52-64, 2014.

[6] Amy, Z.Z., Micheal, M. and Nicholas, F., “Designing an efficient warehouse layout to facilitate the order-filling process: An industrial distributor’s experience”, *Journal of Production and Inventory Management*, Vol.43, No.3-4, p.83-88, 2002.

[7] ชวิศ บุญมี และ ชมพูนุท เกษมเศรษฐ์, “การประยุกต์ใช้เทคนิคการหาค่าดีที่สุดสาหรับการกระจายสินค้า กรณีศึกษาโรงงานผลิตผลไม้กระป๋อง” *วารสารไทยการวิจัยดำเนินงาน* ปีที่ 2 ฉบับที่ 2กรกฎาคม-ธันวาคม 2557.

[8] กัญญาณี แซ่ฉั่ว, จุฑารัตน์ จันชัยภูมิ, ณภัทร วิไลโรจน์วรกุล และ ฉัฐไชย์ ลีนาวงศ์, “การหาปริมาณการสั่งซื้อเมื่อพิจารณาการแบ่งกลุ่ม ABC แบบถ่วงน้ำหนักหลายเกณฑ์สินค้าคลังสำรอง และการพยากรณ์อุปสงค์” คณะวิทยาศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง.

การเปรียบเทียบการยอมรับเฟสบุคและไลน์เพื่อใช้ในการสื่อสารในชั้นเรียน กรณีศึกษา คณะวิทยาการคอมพิวเตอร์ มหาวิทยาลัยราชภัฏอุบลราชธานี

The Adoption of Facebook vs Line for Classroom Communication: A Case Study of Faculty of Computer Science, Ubon Ratchathani Rajabhat University

ชาญศักดิ์ ศรีสวัสดิ์สกุล (Charnsak Srisawatsakul)[1] และวรัณศณางค์ บุณฑริก (Waransanang Boontarig)[2]

[1,2]สาขาวิชาคอมพิวเตอร์และเทคโนโลยีสารสนเทศเพื่อการศึกษา คณะวิทยาการคอมพิวเตอร์ มหาวิทยาลัยราชภัฏอุบลราชธานี

[1]charnsak@charnsak.com,[2]waransanang.b@ubru.ac.th

บทคัดย่อ

ในช่วงทศวรรษที่ผ่านมาระบบเครือข่ายสังคมออนไลน์ได้ถูกนำไปใช้ในการติดต่อสื่อสารในหลากหลายรูปแบบ รวมถึงการใช้งานในห้องเรียนได้รับความนิยมมากขึ้น แต่งานวิจัยที่พูดถึงการเลือกใช้เครื่องมือสื่อสังคมออนไลน์เพื่อการสื่อสารในชั้นเรียนนั้นยังมีอยู่ไม่มาก ดังนั้นงานวิจัยชิ้นนี้จึงมีวัตถุประสงค์ที่จะทำการเปรียบเทียบการยอมรับการใช้งาน Facebook และ Line เพื่อการสื่อสารในชั้นเรียน โดยเป็นการศึกษานำร่องกับกลุ่มตัวอย่าง 2 กลุ่ม แต่ละกลุ่มตัวอย่างจะใช้งานเครือข่ายที่แตกต่างกันเพื่อการสื่อสารในชั้นเรียนเป็นเวลา 3 เดือน จากนั้นจึงตอบแบบสอบถามเพื่อเก็บข้อมูล ผลที่ได้พบว่ากลุ่มตัวอย่างจะใช้งาน Facebook เพื่อการสื่อสารในชั้นเรียนมากกว่า Line โดยมีค่าความถดถอยระหว่างความตั้งใจต่อการใช้งานจริงของ Facebook อยู่ที่ 0.917 และค่าความถดถอยระหว่างความตั้งใจต่อการใช้งานจริงที่ของ Line อยู่ที่ 0.570 ผลที่ได้ถูกนำไปสร้างเป็นข้อแนะนำสำหรับผู้สอนเพื่อประกอบการใช้เครือข่ายสังคมออนไลน์เพื่อการสื่อสารในชั้นเรียนที่จะทำให้นักศึกษายอมรับได้มากขึ้นต่อไป

คำสำคัญ: เครือข่ายสังคมออนไลน์, การยอมรับเทคโนโลยี, เทคโนโลยีการศึกษา, กรอบการยอมรับเทคโนโลยี, ห้องเรียน

Abstract

In the last decade, Online Social Network became an important part in different field. One of the most famous filed is education. However, researchers did not pay much intention on this topic. Hence, there are still gaps in the research on the usage of online social network in the context of classroom communication. This study compare the usage of Facebook and Line in the context of classroom communication. There are 2 groups of participants in this study. Each group of student will use difference online social networks in the classroom communication for 3 months. The questionnaire was used to collected data from each group. The results show that participants have more intention to use Facebook than Line. The beta of multiple linear regression from intention to use of Facebook was 0.917 and Line 0.570. The recommendations were created for anyone who need to use online social network for classroom communication which may help them gain more adoption rate.

Keyword: Online Social Network, Adoption, Education, Technology Acceptance Model, Classroom

1. บทนำ

ปัจจุบันรูปแบบการเรียนการสอนระดับบัณฑิตศึกษาได้มีการเปลี่ยนแปลงไปมาก เทคโนโลยีสารสนเทศได้เข้ามามีบทบาทสำคัญในการเรียนการสอนมากยิ่งขึ้น หนึ่งในเทคโนโลยีที่ได้รับความนิยมนำมาใช้ในห้องเรียนมากขึ้นก็คือเทคโนโลยีเครือข่ายสังคมออนไลน์ จากงานวิจัยสำนักงานพัฒนาธุรกรรมอิเล็กทรอนิกส์ กระทรวงเทคโนโลยีสารสนเทศและการสื่อสาร [1] สามารถสรุปได้ว่าผู้ใช้อินเทอร์เน็ตในไทยใช้เวลากว่า 85.6% ในการใช้งานเครือข่ายสังคมออนไลน์

ที่สำคัญผู้ใช้ที่เป็นนักศึกษาในกลุ่มอายุ 15-19 ปี และ 20-24 ปี ใช้เวลาบนโลกอินเทอร์เน็ตมากกว่า 52 ชั่วต่อสัปดาห์เลยทีเดียว ด้วยเหตุนี้ทำให้การใช้ระบบเครือข่ายสังคมออนไลน์เพื่อสื่อสารในชั้นเรียนเป็นไปอย่างแพร่หลายมากยิ่งขึ้น แต่ทว่าเครือข่ายสังคมออนไลน์ที่ได้รับความนิยมนั้นมีหลายประเภทด้วยกันเช่น Facebook Line Youtube Twitter ผู้สอนแต่ละแห่งก็เลือกใช้งานเครือข่ายเหล่านี้แตกต่างกัน ยังไม่มีการเปรียบเทียบว่า การเลือกใช้เครือข่ายสังคมออนไลน์ตัวใดที่จะสามารถทำให้นักศึกษาเกิดการยอมรับในการใช้งานได้ดีที่สุด

ดังนั้นวัตถุประสงค์ของในงานวิจัยชิ้นนี้จึงจะเปรียบเทียบการยอมรับการใช้งานระบบเครือข่ายสังคมออนไลน์เพื่อสื่อสารในชั้นเรียน โดยเลือกเปรียบเทียบเครือข่ายสังคมออนไลน์ที่ได้รับความนิยมมากที่สุดของไทย 2 อันดับแรก นั่นคือ Facebook และ Line [1] เพื่อหาคำตอบว่านักศึกษาจะยอมรับระบบเครือข่ายสังคมออนไลน์เพื่อสื่อสารในชั้นเรียนชนิดใดมากกว่ากัน ผลที่ได้จะช่วยให้ผู้สอนสามารถเลือกใช้งานระบบเครือข่ายสังคมออนไลน์มีการยอมรับสูงสุดเพื่อที่จะทำให้ผู้เรียนมีความต้องการใช้งานมากขึ้น ซึ่งอาจจะส่งผลต่อประสิทธิภาพในการเรียนการสอนต่อไปได้

เนื่องจาก Facebook และ Line นั้นเป็นระบบเครือข่ายสังคมออนไลน์ที่ต่างชนิดกันดังนั้นขอบเขตการเปรียบเทียบการใช้งานจึงเน้นเฉพาะระบบการทำงานที่เหมือนกันในการติดต่อในชั้นเรียนเท่านั้นเช่น ระบบไทม์ไลน์ ระบบสนทนาแบบกลุ่ม เป็นต้น และงานวิจัยนี้ไม่ได้สนใจตัวอุปกรณ์ที่กลุ่มตัวอย่างใช้เข้าสู่เครือข่ายสังคมออนไลน์ เช่น สมาร์ทโฟน หรือคอมพิวเตอร์ซึ่งอาจจะมีผลต่อการใช้งานกับผู้ใช้บางคน

2. วรรณกรรมที่เกี่ยวข้อง

2.1 ระบบเครือข่ายสังคมออนไลน์กับการศึกษา

Boyd และ Ellision [2] ได้ให้ความหมายของระบบเครือข่ายสังคมออนไลน์เอาไว้ว่า ระบบเครือข่ายสังคมออนไลน์หมายถึง เว็บไซต์ที่ให้บริการแก่ผู้ใช้ โดยที่จะต้องมีความสามารถดังนี้

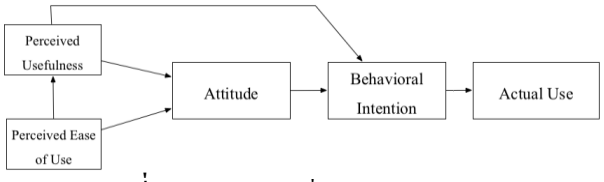
1) มีส่วนสำหรับแสดงผลข้อมูลประวัติส่วนตัวของผู้ใช้ ทั้งแบบเปิดเผยต่อสาธารณะหรือกึ่งสาธารณะ
2) มีระบบการจัดการรายชื่อผู้ใช้อื่นๆที่ผู้ใช้นั้นๆเป็นเพื่อนอยู่ หรือมีการติดต่อกันได้ ภายในระบบเดียวกันหรือต่างระบบก็ได้
3) สามารถเรียกดูรายชื่อผู้ใช้คนอื่นๆที่มีการติดต่อกันได้ต่อไปเรื่อยๆ ภายในระบบเดียวกัน อย่างไรก็ดีความสามารถของระบบเหล่านี้จะแตกต่างกันไปขึ้นอยู่กับการออบแบบและจุดเด่นของแต่ละระบบนั่นเอง

ในปัจจุบันนี้ ผู้ให้บริการเครือข่ายสังคมออนไลน์ได้มีการให้บริการระบบของตัวเองบนโทรศัพท์มือถือ หรือสมาร์ทโฟน โดยอาจจะให้บริการทั้งแบบโมบายเว็บไซด์หรือเนทีฟแอพพลิเคชั่น ซึ่งเรียกโดยรวมว่าเป็น ระบบเครือข่ายสังคมออนไลน์บนอุปกรณ์เคลื่อนที่ (Mobile Social Network) ซึ่งระบบนี้จะทำการตกแต่งหน้าจอของระบบเครือข่ายสังคมออนไลน์ของตัวเองให้เหมาะกับการแสดงผลบนจออุปกรณ์มือถือ เพื่อเพิ่มประสบการณ์ใช้งานที่ดีของผู้ใช้และความรวดเร็วในการใช้งานเช่นกันทำให้ผู้ใช้สามารถเข้าถึงระบบเครือข่ายสังคมออนไลน์ได้ง่ายทุกที่ทุกเวลา

ในมุมของระบบเครือข่ายสังคมออนไลน์กับการศึกษานั้น Greenhow [3] ได้ทำการศึกษาเกี่ยวกับการใช้งานระบบเครือข่ายสังคมออนไลน์และการศึกษา พบว่าระบบเครือข่ายสังคมออนไลน์นั้นมีบทบาทในการสนับสนุนการเรียนการสอนทั้งในเรื่องของการลดความตึงเครียดของการเรียน เพิ่มความคิดสร้างสรรค์ และช่วยในการทำกิจกรรมหลังเรียน เป็นต้น แต่การศึกษาของ Greenhow นั้นทำขึ้นเมื่อปี 2011 ซึ่งในยุคนั้นยังไม่มีตัวเลือกของระบบเครือข่ายสังคมออนไลน์ที่มีคุณสมบัติแตกต่างกันเช่นปัจจุบันนี้ และ Greenhow ก็ยังไม่ได้ให้ความสนใจในกรณีที่ใช้ระบบเครือข่ายสังคมออนไลน์เพื่อการสื่อสารในชั้นเรียนอีกด้วย

2.2 การยอมรับเทคโนโลยี

ในงานวิจัยนี้ได้นำเอา กรอบแนวคิดการยอมรับเทคโนโลยี (Technology Acceptance Model) [4] ซึ่งเป็นกรอบแนวคิดที่ได้รับความนิยมอย่างแพร่หลาย สำหรับการยอมรับและใช้งานเทคโนโลยีต่าง มาประยุกต์ใช้เพื่อวัดระดับการยอมรับการใช้งาน Facebook และ Line สำหรับการติดต่อสื่อสารในชั้นเรียน



ภาพที่ 1: กรอบแนวคิดเพื่อการยอมรับเทคโนโลยี

กรอบแนวคิดการยอมรับเทคโนโลยีนั้น (รูปที่ 1) ถูกสร้างขึ้นมาเพื่อคาดเดาการยอมรับเทคโนโลยีของผู้ใช้ ซึ่งทฤษฎีนี้ถูกพัฒนาขึ้นโดย [4] โดยมีตัวแปรที่เกี่ยวข้องดังนี้

1) Perceived Usefulness การรับรู้ถึงประโยชน์ หมายถึง บุคคลจะยอมรับเทคโนโลยีนั้น ถ้าสามารถรับรู้ได้ว่าเทคโนโลยีนั้นมีประโยชน์ต่อการทำงานของตัวเอง
2) Perceived Ease Of Use การรับรู้ถึงความง่ายในการใช้งาน หมายถึง บุคคลจะยอมรับการใช้งานเทคโนโลยีหรือไม่นั้น ขึ้นอยู่กับการรับรู้ถึงความง่ายในการใช้งาน ไม่ต้องเสียเวลาและแรงในการเรียนรู้การใช้งานมากเกินไป
3) Attitude หมายถึงทัศนะคติ ความรู้สึกที่ตัวบุคคลนั้นมีอยู่ต่อเทคโนโลยีที่จะใช้งาน
4) Bahavioral Intention หมายถึง ความตั้งใจที่จะใช้เทคโนโลยี
5) Actual Use หมายถึง การใช้งานเทคโนโลยีจริงๆ อย่างต่อเนื่อง

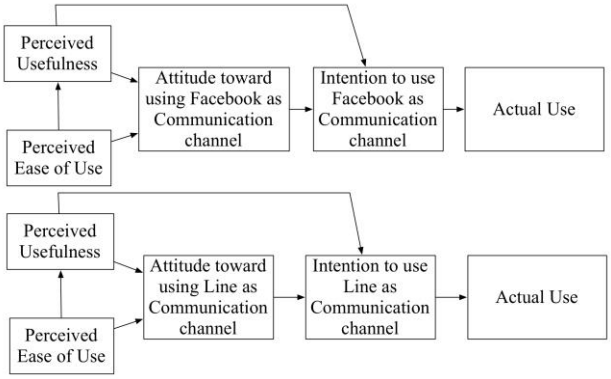
โดยทฤษฎีนี้สามารถอธิบายได้ว่า การจะเกิดการใช้งานเทคโนโลยีจริงอย่างต่อเนื่องนั้น ผู้ใช้จะต้องมีความตั้งใจที่จะใช้งานก่อน การที่เราจะมีความตั้งใจได้นั้น ก็ต้องมีทัศนคติที่ดีก่อน และ การที่บุคคลจะมีทัศนคติที่ดีต่อเทคโนโลยีได้นั้น เทคโนโลยีนั้นจะต้องมีประโยชน์ และใช้งานง่าย นั่นเอง

3. วิธีดำเนินงานวิจัย

3.1 กรอบแนวคิดของงานวิจัย

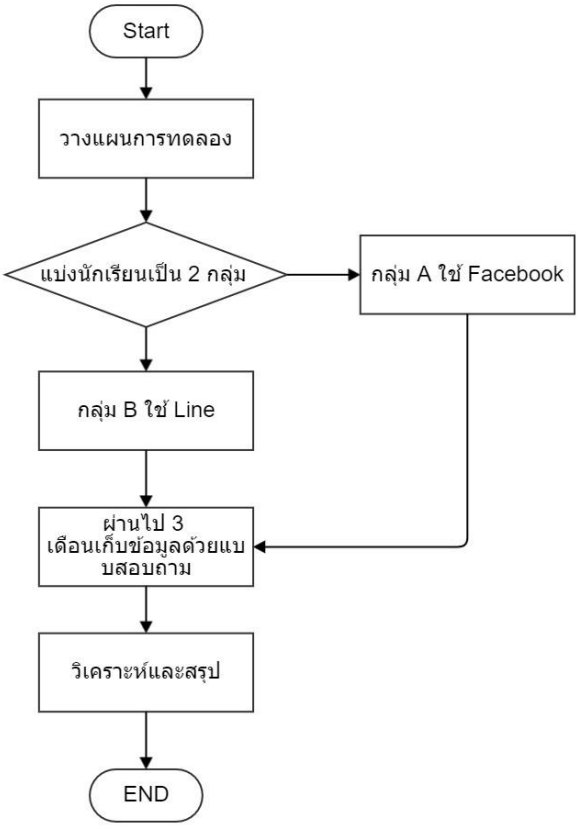
กรอบแนวคิดของงานวิจัยนี้ปรากฏในภาพที่ 2 ซึ่งเป็นการประยุกต์ใช้ กรอบแนวคิดการยอมรับเทคโนโลยี [4] มาใช้ในบริบทของการยอมรับเครือข่ายสังคมออนไลน์เพื่อการสื่อสารในชั้นเรียนดังนั้น ตัวแปรอิสระและตัวแปรตามจึงประกอบไปด้วย ตัวแปรดังต่อไปนี้

1. Perceived Usefulness – การรับรู้ถึงประโยชน์ของเครือข่ายสังคมออนไลน์เพื่อการสื่อสารในชั้นเรียน
2. Perceived Ease of Use – การรับรู้ถึงความง่ายในการใช้งานเครือข่ายสังคมออนไลน์เพื่อการสื่อสารในชั้นเรียน
3. Attitude Toward – ทัศนคติที่มีต่อการใช้งานเครือข่ายสังคมออนไลน์เพื่อการสื่อสารในชั้นเรียน
4. Intention to Use – ความตั้งใจที่จะใช้เครือข่ายสังคมออนไลน์เพื่อการสื่อสารในชั้นเรียน
5. Actual Use – การใช้งานเครือข่ายสังคมออนไลน์เพื่อการสื่อสารในชั้นเรียนจริง



ภาพที่ 2: กรอบแนวคิดของงานวิจัย

งานวิจัยนี้มีวิธีดำเนินงานตาม Flowchart ที่แสดงไว้ในภาพที่ 3



ภาพที่ 3: Flowchart แสดงขั้นตอนการทำงาน

3.3 สมมุติฐานการวิจัย

H1: ผลกระทบของความตั้งใจในการใช้งานเครือข่ายสังคมออนไลน์เพื่อการสื่อสารในชั้นเรียนต่อการใช้งานจริงของเครือข่ายสังคมออนไลน์เพื่อการสื่อสารในชั้นเรียนระหว่าง Facebook และ Line มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ

H2: ผลกระทบของทัศนคติต่อการใช้งานเครือข่ายสังคมออนไลน์เพื่อการสื่อสารในชั้นเรียนต่อ ความตั้งใจในการใช้งานเครื่องสังคมออนไลน์เพื่อการสื่อสารในชั้นเรียน ระหว่าง Facebook และ Line มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ

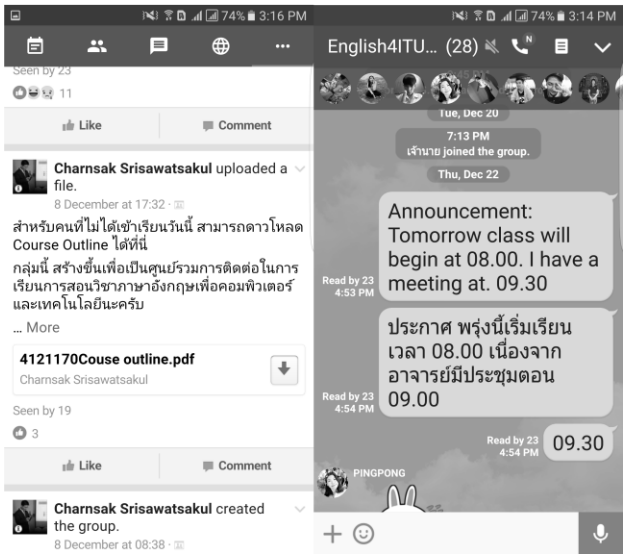
H3: ผลกระทบของการรับรู้ถึงประโยชน์ของเครือข่ายสังคมออนไลน์เพื่อการสื่อสารในชั้นเรียนและการรับรู้ถึงความง่ายในการใช้งานเครือข่ายสังคมออนไลน์เพื่อการสื่อสารในชั้นเรียน ต่อทัศนคติต่อการใช้งานเครือข่ายสังคมออนไลน์เพื่อการสื่อสารในชั้นเรียน ระหว่าง Facebook และ Line มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ

H4: ผลกระทบของการรับรู้ถึงประโยชน์ของเครือข่ายสังคมออนไลน์เพื่อการสื่อสารในชั้นเรียนต่อ ความตั้งใจในการใช้งานเครือข่ายสังคมออนไลน์เพื่อการสื่อสารในชั้นเรียน ระหว่าง Facebook และ Line มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ

H5: ผลกระทบของการรับรู้ถึงความง่ายในการใช้งานเครือข่ายสังคมออนไลน์เพื่อการสื่อสารในชั้นเรียน ต่อ การรับรู้ถึงประโยชน์ของเครือข่ายสังคมออนไลน์เพื่อการสื่อสารในชั้นเรียนระหว่าง Facebook และ Line มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ

3.2 ประชากรและกลุ่มตัวอย่าง

ประชากรของงานวิจัยนี้คือนักศึกษาคณะวิทยาการคอมพิวเตอร์ มหาวิทยาลัยราชภัฏอุบลราชธานี อย่างไรก็ดีงานวิจัยชิ้นนี้ยังเป็นการศึกษานำร่อง จึงได้เลือกกลุ่มตัวอย่างด้วยวิธีการเลือกตัวอย่างแบบเจาะจง (Purposive Sampling) ดังนั้น กลุ่มตัวอย่างของงานวิจัยนี้คือ นักศึกษาชั้นปีที่ 1 สาขาวิทยาการคอมพิวเตอร์ (กลุ่ม A) และสาขามัลติมีเดียและเทคโนโลยีอนิเมชั่น (กลุ่ม B) โดยได้เลือกทำการทดลองในวิชาภาษาอังกฤษเพื่องานด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศ ซึ่งเป็นวิชาที่นักศึกษาทั้ง 2 ห้องเรียนเหมือนกัน โดยมีการควบคุมตัวแปรภายนอกอื่นๆให้เหมือนกัน คือ มีอาจารย์ผู้สอนคนเดียวกัน เนื้อหาที่สอนเหมือนกัน ใช้เอกสารชุดเดียวกัน ที่สำคัญ ทั้ง 2 ห้องจะได้รับข้อมูลข่าวสารที่เหมือนกัน แต่นักศึกษากลุ่ม A จะได้รับข้อมูลผ่านทาง Facebook และ กลุ่ม B จะได้รับข้อมูลผ่านทาง Line เพื่อป้องกันไม่ให้เกิดความไม่ยุติธรรมในการเรียนอันเนื่องมาจากการทดลอง ภาพที่ 4 แสดงหน้าจอตัวอย่างการใช้งาน Facebook และ Line เพื่อการสื่อสารในห้องเรียน



ภาพที่ 4: ตัวอย่างการใช้ Facebook และ Line เพื่อสื่อสารในห้องเรียน

3.3 การเก็บข้อมูล

หลังจากที่ได้ใช้งานเครือข่ายสังคมออนไลน์เพื่อการสื่อสารในห้องเรียนไปแล้วไปแล้ว ประมาณ 3 เดือนจึงได้ทำการเก็บข้อมูลจากนักศึกษาในแต่ละกลุ่มโดยใช้แบบสอบถามออนไลน์ โดยในแบบสอบถามมีทั้งหมด 17 ข้อ แบ่งเป็น 2 ส่วน ส่วนแรกมี 2 ข้อเก็บข้อมูลส่วนตัวเกี่ยวกับผู้ตอบแบบสอบถาม ประกอบไปด้วย เพศและอายุ ส่วนที่ 2 มี 15 ข้อจะเป็นคำถามที่ถามถึงการยอมรับเทคโนโลยี โดยอ้างอิงคำถามมาจาก โมเดลการยอมรับ เทคโนโลยี (TAM) [4] โดยคำถามทั้ง 15 ข้อนี้ได้ถูกทำการทดสอบความเชื่อมันโดยใช้ Cronbach's alpha โดยคำถามทั้งหมดได้ค่า Cronbach's alpha = 0.908 ส่วนค่า Cronbach's alpha ของตัวแปรต่างๆ แสดงในตารางที่ 1

ตารางที่ 1: แสดงค่า Cronbach's alpha ของตัวแปรทั้งหมด

ตัวแปร	จำนวนคำถาม	Cronbach's alpha
การรับรู้ถึงประโยชน์ (Perceived Usefulness - PU)	5	0.903
การรับรู้ถึงความง่าย (Perceived Ease of Use - PE)	5	0.749
ทัศนคติต่อการใช้งาน (Attitude - ATT)	2	0.785
ความตั้งใจในการใช้งาน (Intention to Use - INT)	2	0.834

จะเห็นได้ว่าค่า Cronbach's alpha ของทั้งแบบสอบถามและของแต่ละตัวแปร มีค่ามากกว่า 0.7 ทำให้แบบสอบถามนี้มีค่าความเชื่อมั่นเพียงพอที่จะใช้ในการวิเคราะห์ผลที่ได้ต่อไป

3.4 การทดสอบสมมุติฐานและการวิเคราะห์ข้อมูล

สำหรับการทดสอบสมมุติฐานและวิเคราะห์ผลของปัจจัยในการยอมรับ Facebook และ Line เพื่อการสื่อสารในชั้นเรียนนั้นจะใช้เทคนิคการการวิเคราะห์ความถดถอยอย่างง่าย และเทคนิคการวิเคราะห์ความถดถอยเชิงพหุดังนี้

ตัวอย่างที่ใช้ในการทดสอบผลของตัวแปรการรับรู้ถึงประโยชน์ของเครือข่ายสังคมออนไลน์เพื่อการสื่อสารในชั้นเรียน (PU) และ การรับรู้ถึงความง่ายในการใช้งานเครือข่ายสังคมออนไลน์เพื่อการสื่อสารในชั้นเรียนว่ามีผลต่อ (PE) ทัศนคติที่มีต่อการใช้งานเครือข่ายสังคมออนไลน์เพื่อการสื่อสารในชั้นเรียน (ATT) จะใช้เทคนิคการวิเคราะห์ความถดถอยเชิงพหุที่มีสมการดังนี้

$$ATT = B_0 + B_{PU}PU + B_{PE}PE$$

โดยที่

ATT คือ ตัวแปรตามในที่นี้คือ ทัศนคติที่มีต่อการใช้งานเครือข่ายสังคมออนไลน์เพื่อการสื่อสารในชั้นเรียน

B_0 คือ ค่า ATT–Intercept เป็นค่าของ ATT เมื่อค่าของ PU เป็น 0 ซึ่งจะได้ผลคือ $ATT = B_0 + B_{PE}PE$

B_{PU} คือ ค่าสัมประสิทธิ์การถดถอย (Regression Coefficients) ของ PU

PU คือ ตัวแปรอิสระในที่นี้คือรับรู้ถึงประโยชน์ของเครือข่ายสังคมออนไลน์เพื่อการสื่อสารในชั้น

B_{PE} คือ ค่าสัมประสิทธิ์การถดถอย (Regression Coefficients) ของ PE

PE คือ ตัวแปรอิสระในที่นี้คือรับรู้ถึงความง่ายในการใช้งานของเครือข่ายสังคมออนไลน์เพื่อการสื่อสารในชั้นและใช้สูตรคำนวณเดียวกันนี้ในการหาผลกระทบของตัวแปรอิสระและตัวแปรตามอื่นๆทั้งของ Facebook และ Line

4. ผลการดำเนินงาน

4.1 สถิติเชิงพรรณาของกลุ่มตัวอย่าง

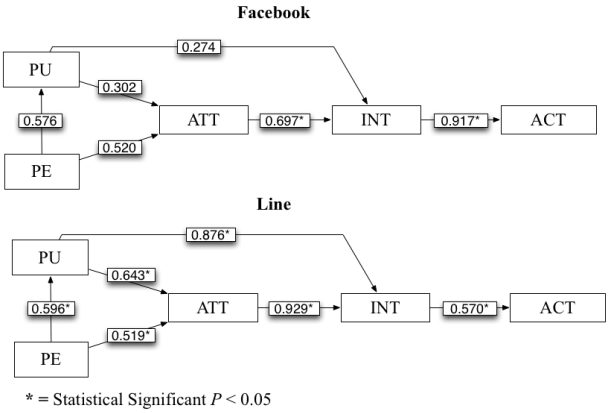
จากการเก็บข้อมูลกลุ่ม A ได้ข้อมูลทั้งหมด 32 คน และกลุ่ม B ได้ข้อมูลทั้งหมด 38 คนทั้ง 2 กลุ่มได้ผ่านเข้าสู่ขั้นตอนการกรองข้อมูล โดยคัดกรองข้อมูลที่ไม่สมบูรณ์ เช่น ตอบแบบสอบถามไม่ครบ หลังจากผ่านการคัดกรองข้อมูลแล้วข้อมูลที่ใช้งานได้ในกลุ่ม A 25 คน และกลุ่ม B 27 คน ตารางที่ 2 แสดงสถิติเชิงพรรณาของทั้ง 2 กลุ่ม กลุ่มตัวอย่างที่ตอบแบบสอบถามนี้อยู่ในช่วงอายุ 13-20 ปีทั้งหมด

ตารางที่ 2: สถิติเชิงพรรณาของกลุ่มตัวอย่าง

เพศ	กลุ่ม A (Facebook)		กลุ่ม B (Line)	
	ความถี่	เปอร์เซ็นต์	ความถี่	เปอร์เซ็นต์
ชาย	15	60	20	74.1
หญิง	10	40	7	25.9
รวม	**25**	100	**27**	100

4.2 ผลจากเทคนิคการวิเคราะห์ความถดถอยเชิงพหุ

การวิจัยนี้ได้ใช้ซอฟต์แวร์ทางด้านสถิติ โดยเฉพาะ (Statistical Package for the Social Sciences - SPSS) ผลที่จะได้จากเทคนิคการวิเคราะห์ความถดถอยเชิงพหุ แสดงในภาพที่ 4



ภาพที่ 4: ค่าสัมประสิทธิ์การถดถอยจากการวิเคราะห์

สำหรับ Facebook การรับรู้ถึงประโยชน์ของFacebook เพื่อการสื่อสารในชั้นเรียนและ การรับรู้ถึงความง่ายในการใช้งาน Facebook นั้นไม่มีผลอย่างมีนัยยะสำคัญทางสถิติ ต่อทัศนคติในการใช้งาน Facebook แต่ทัศนคติในการใช้งาน Facebook เพื่อการสื่อสารในชั้นเรียนนั้นมีผลต่อความตั้งใจในการใช้งาน Facebook เพื่อการสื่อสารในชั้นเรียนอย่างมีนัยสำคัญทางสถิติ (B = 0.697) และเช่นเดียวกัน ความตั้งใจในการใช้งาน Facebook เพื่อการสื่อสารในชั้นเรียนมีผลต่อการใช้งานจริงของFacebook เพื่อการสื่อสารในชั้นเรียนอย่างมีนัยสำคัญทางสถิติ (B = 0.917) ในทางตรงข้ามสำหรับ Line นั้นตัวแปรต้นมีผลกระทบอย่างมีนัยสำคัญทางสถิติ โดยค่าสัมประสิทธิ์การถดถอยที่มีค่าสูงที่สุดคือ ทัศนคติในการใช้งาน Line เพื่อการสื่อสารในชั้นเรียน มีผลต่อความตั้งใจในการใช้งาน Line เพื่อการสื่อสารในชั้นเรียน

4.3 สรุปผลการทดสอบสมมุติฐาน

จากผลการทดลองทั้งหมด เราสามารถนำมาสรุปผลของสมมุติฐานที่ได้ตั้งไว้ โดยได้ผลลัพธ์ดังตารางที่ 3 โดยแสดงเฉพาะความสัมพันธ์ที่มีนัยสำคัญทางสถิติเท่านั้น

ตารางที่ 3: สรุปผลการทดสอบสมมุติฐาน

H	ตัวแปร	Facebook	Line
H1	INT --> ACT	ยอมรับ	ยอมรับ
H2	ATT --> INT	ยอมรับ	ยอมรับ
H3	PU,PE --> ATT	ปฏิเสธ	ยอมรับ
H4	PU --> INT	ปฏิเสธ	ยอมรับ
H5	PE --> PU	ปฏิเสธ	ยอมรับ

5. บทสรุปและข้อเสนอแนะ

5.1 สรุปผล

จากผลการทดลองกลุ่มตัวอย่างมีทัศนคติในการใช้งาน Line ดีกว่า Facebook ทำให้กลุ่มตัวอย่างมีความต้องการใช้ Line เพื่อการสื่อสารในชั้นเรียนมากกว่าด้วย แต่เป็นที่น่าแปลกใน หากดูที่ผลของความตั้งใจ ต่อการใช้งานจริงๆนั้น Facebook กลับมีผลมากกว่า กล่าวคือ กลุ่มตัวอย่างที่มีความตั้งใจจะใช้ Facebook เพื่อการสื่อสารในชั้นเรียนนั้น กลับมีการใช้งานจริงๆ มากกว่า Line ซึ่งอาจจะเป็นเพราะมีตัวแปรด้านอื่นๆที่มีผลชัดเจนกว่าในการใช้งาน Facebook

การรับรู้ถึงความมีประโยชน์และการรับรู้ถึงความง่ายในการใช้งานไม่ได้ผลต่อทัศนติสำหรับ Facebook เลยอาจจะอธิบายได้ว่าทุกคนคุ้นเคยกับการใช้งาน Facebook อยู่แล้ว ทำให้มีทัศนคติที่ดี ที่เกิดจากตัวแปรอิสระอื่นๆที่ไม่ได้ทดลองในงานวิจัยชิ้นนี้ แต่ในทางกลับกัน สำหรับกลุ่มที่ใช้ Line นั้นเรื่องของ ประโยชน์และการใช้งานที่ง่ายจะมีส่วนให้เกิดทัศนคติที่ดีต่อการใช้งาน Line เพื่อการสื่อสารในชั้นเรียนมากกว่านั้น การรับรู้ถึงประโยชน์ จะส่งผลให้เกิดความตั้งใจในการใช้งาน Line เพื่อการสื่อสารในชั้นเรียนอีกด้วย

5.2 ข้อเสนอแนะสำหรับผู้สอน

จากผลการทดลองสามารถสร้างเป็นข้อแนะนำสำหรับอาจารย์ผู้สอนได้ กล่าวคืออาจารย์ผู้สอนควรจะเลือก Facebook เพื่อการสื่อสารในชั้นเรียนมากกว่า เนื่องจากมีการใช้งานจริงที่สูงกว่า แต่ถ้าผู้สอนต้องการที่จะใช้ Line ในการสอนนั้นควรสร้างการรับรู้ถึงประโยชน์ และความง่ายในการใช้งานให้มากยิ่งขึ้น เช่น การใช้สติกเกอร์ช่วย เป็นต้น ก็อาจจะช่วยเพิ่มการยอมรับการใช้งาน Line เพื่อการสื่อสารในชั้นเรียนได้

5.3 ข้อจำกัดและงานวิจัยในอนาคต

ข้อจำกัดของงานวิจัยชิ้นนี้อันดับแรกเป็นงานวิจัยนำร่อง อันดับต่อมาคือการที่กลุ่มตัวอย่างเป็นนักศึกษาในคณะวิทยาการคอมพิวเตอร์ ซึ่งอาจจะมีการยอมรับสูง เนื่องจากมีความชำนาญในการสื่อสารผ่านเครือข่ายสังคมออนไลน์อยู่แล้ว อันดับสุดท้ายคือการยอมรับการใช้งานเครือข่ายสังคมออนไลน์เพื่อการสื่อสารในชั้นเรียนอาจจะแตกต่างกันไปในรายวิชาที่แตกต่างกัน งานวิจัยในอนาคตจึงควรให้ความสำคัญกับเรื่องกลุ่มตัวอย่าง หรือทำการเปรียบเทียบภายในเครือข่ายสังคมออนไลน์เดียวกัน แต่ใช้งานต่างลักษณะกัน เป็นต้น

เอกสารอ้างอิง

[1] Electronic Transactions Development Agency, *Thailand Internet User Profile 2016*. 2016.

[2] D. M. Boyd and N. B. Ellison, “Social Network Sites: Definition, History, and Scholarship,” *J. Comput. Commun.*, vol. 13, no. 1, pp. 210–230, Oct. 2007.

[3] C. Greenhow, Ed., “Online social networks and learning,” *Horiz.*, vol. 19, no. 1, pp. 4–12, 2011.

[4] F. D. Davis, “Perceived usefulness, perceived ease of use, and user acceptance of information technology,” *MIS Q.*, pp. 319–340, 1989.

การพัฒนาบทเรียนคอมพิวเตอร์ช่วยสอนผ่านเว็บด้วยเทคนิคการเรียนรู้แบบโครงงานเป็นฐาน ตามแผนการสอนฐานสมรรถนะแบบผสมผสานด้วย MIAP รายวิชาเทคโนโลยีสารสนเทศและการสื่อสาร

The Development of WBI by Project-Based Learning Techniques for Competency-Based Lesson Plan on Blended Learning by MIAP of Information and Communication Technology Course

พงษ์ดนัย จิตตวิสุทธิกุล (Phongdanai Jittavisuttikul)[1] และกฤช สินธนะกุล (Krich Sintanakul)[2]

[1]สาขาวิชาคอมพิวเตอร์ศึกษา คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏนครปฐม

[2]ภาควิชาคอมพิวเตอร์ศึกษา คณะครุศาสตร์อุตสาหกรรม มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

[1]vazabizatan@gmail.com, [2]krich.s@fte.ac.th

บทคัดย่อ

การวิจัยครั้งนี้เป็นการวิจัยเชิงทดลอง โดยมีวัตถุประสงค์เพื่อพัฒนาบทเรียนคอมพิวเตอร์ช่วยสอนผ่านเว็บด้วยเทคนิคการเรียนรู้แบบโครงงานเป็นฐาน ตามแผนการสอนฐานสมรรถนะแบบผสมผสานด้วย MIAP รายวิชาเทคโนโลยีสารสนเทศและการสื่อสารที่มีประสิทธิภาพเหมาะสมที่จะนำไปใช้ในการเรียนการสอน ซึ่งประกอบด้วยการเรียนแบบ online บนระบบ NPRU LMS และแบบ offline ที่ผู้เรียนจะต้องฝึกทักษะปฏิบัติภายในห้องเรียน ตามเทคนิค MIAP เพื่อให้เกิดสมรรถนะสามารถปฏิบัติชิ้นงานได้ภายใต้เทคนิคการเรียนรู้แบบโครงงานเป็นฐาน โดยมุ่งหวังให้ผู้เรียนมีผลสัมฤทธิ์ทางการเรียนที่ดีขึ้นและสามารถจัดทำโครงงานที่มีคุณภาพในภาพรวมอยู่ในระดับดี กลุ่มเป้าหมาย คือ นักศึกษาระดับปริญญาตรี มหาวิทยาลัยราชภัฏนครปฐม ที่ลงทะเบียนเรียนรายวิชาเทคโนโลยีสารสนเทศและการสื่อสาร ในปีการศึกษาที่ 2/2559 จำนวน 49 คน ซึ่งได้มาจากการสุ่มอย่างง่ายด้วยวิธีจับฉลากห้องเรียน

ผลการวิจัยพบว่า บทเรียนคอมพิวเตอร์ช่วยสอนผ่านเว็บที่พัฒนาขึ้นประกอบด้วยเนื้อหา 2 ส่วน คือ ภาคทฤษฎี และภาคปฏิบัติ นำเสนอเนื้อหาภาคปฏิบัติด้วยบทเรียนแบบวิดีโอสาธิตบน EDpuzzle ซึ่งมีประสิทธิภาพเท่ากับ 83.06/81.02 ส่งผลทำให้ผู้เรียนมีผลสัมฤทธิ์ทางการเรียนหลังเรียนสูงกว่าก่อนเรียนอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 และคุณภาพโครงงานของผู้เรียนในภาพรวมอยู่ในระดับดี (ค่าเฉลี่ย 4.00)

คำสำคัญ: ฐานสมรรถนะ โครงงานเป็นฐาน ผสมผสาน

Abstract

This research was an experimental research with the purposes were to develop WBI by project-based learning techniques for competency-based lesson plan on blended learning by MIAP of Information and Communication Technology course that proper effective for use in teaching which consists online learning on NPRU LMS and offline learning that students have to train practice skills in classroom according to MIAP technique to have competency for performing the work under project-based learning techniques by expecting students to achieve better academic achievement and can make project with total quality at a good level. The target group was 49 undergraduate students of Nakhon Pathom Rajabhat University, who registered in the 2nd semester of the academic year 2016, which derived from drawing a simple random classroom sampling.

The result findings showed the developed WBI consists of content two parts were theoretical and practical by presenting practical content with demonstration video lessons on EDpuzzle effective was 83.06/81.02 that have effect on the student's academic achievements after learned higher than before learned was significantly at .05 levels and the total quality of student' projects was a good level ($\bar{X}$ = 4.00).

Keyword: Competency-Based, Project-Based, Blended

1. บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

คุณภาพแรงงานเป็นสิ่งที่จำเป็นเพราะนอกจากการปลี่ยนแปลงทางเทคโนโลยี โลกาภิวัฒน์ และการแข่งขันทางเศรษฐกิจแล้ว แนวโน้มประชากรไทยที่หยุดการเพิ่มด้านปริมาณ ทำให้ต้องตระหนักถึงความสำคัญด้านคุณภาพมากยิ่งขึ้น [1] ซึ่งคุณภาพดังที่กล่าวมานั้นหมายรวมถึงสมรรถนะหรือความสามารถในด้านต่าง ๆ ที่มีความสำคัญต่อการปฏิบัติงานของแรงงานในสถานประกอบการ ประกอบด้วยความรู้ ทักษะ และคุณลักษณะของเจตคติ ที่รวมเรียกว่า สมรรถนะ (Competency) [2] ดังนั้นจึงมีการจัดทำแผนการสอนฐานสมรรถนะนี้ขึ้น โดยมุ่งหวังให้ผู้เรียนมีความรู้ ทักษะ และเจตคติที่ดีเกี่ยวกับความสามารถพื้นฐานในการใช้งานเทคโนโลยีสารสนเทศและการสื่อสารเบื้องต้น เพื่อที่จะเป็นแรงงานที่มีคุณภาพต่อไปในอนาคตเมื่อต้องประกอบอาชีพ

ตามหลักสูตรของมหาวิทยาลัยราชภัฏนครปฐม นักศึกษาทุกคนจะต้องลงทะเบียนเรียนในรายวิชาเทคโนโลยีสารสนเทศและการสื่อสาร ซึ่งเป็นวิชาที่ประกอบด้วยภาคทฤษฎีและภาคปฏิบัติ โดยในภาคปฏิบัติประกอบด้วยเนื้อหาเกี่ยวกับการใช้โปรแกรมประยุกต์พื้นฐานได้แก่ ด้านการประมวลผลคำ ด้านตารางคำนวณ ด้านการนำเสนอ และด้านการสื่อสารผ่านเครือข่าย จากผลการเรียนภาคปฏิบัติในภาคเรียนที่ 1/2559 พบว่า ในส่วนของภาคปฏิบัติการใช้โปรแกรมประยุกต์พื้นฐานด้านการประมวลผลคำผู้เรียนส่วนใหญ่ไม่สามารถทำชิ้นงานได้ไม่เป็นไปตามสมรรถนะที่คาดหวัง โดยเกิดจากการที่ผู้เรียนไม่สามารถจดจำขั้นตอนการปฏิบัติงานได้หมด ซึ่งในปัจจุบันยังขาดสื่อการเรียนการสอนให้ผู้เรียนได้ศึกษา ทบทวนเพิ่มเติมนอกเหนือจากเวลาเรียน แนวทางในการแก้ไขปัญหานี้ อาจนำเทคโนโลยีสารสนเทศมาใช้เพื่อการศึกษาให้ทันยุคทันสมัย ในรูปแบบของบทเรียนคอมพิวเตอร์ช่วยสอนผ่านเว็บด้วยเทคนิคการจัดการเรียนรู้แบบโครงงานเป็นฐาน (Project-Based Learning) ซึ่งเป็นการเรียนรู้ที่มุ่งพัฒนาผู้เรียนแต่ละคนให้สามารถสร้างความรู้ผ่านกระบวนการคิดและลงมือปฏิบัติสร้างสรรค์ผลงานด้วยตนเอง อันจะนำไปสู่การเรียนรู้ที่เกิดขึ้นโดยผู้เรียนเป็นผู้กระทำกิจกรรม [3] ผ่านการเรียนรู้แบบผสมผสานด้วยเทคนิค MIAP ซึ่งเป็นกระบวนการสอนแบบผสมผสานระหว่างการเรียนแบบเผชิญหน้าในห้องเรียน (Face to Face) กับการเรียนแบบออนไลน์ (Online Learning) เข้าด้วยกัน โดยการนำข้อดีของการเรียนทั้งสองแบบมาเสริมเติมเต็มซึ่งกันและกัน ตามวิธีการสอนด้วยเทคนิค MIAP

ด้วยเหตุผลที่กล่าวมาข้างต้น ผู้วิจัยจึงเห็นควรที่จะพัฒนาบทเรียนคอมพิวเตอร์ช่วยสอนผ่านเว็บด้วยเทคนิคการเรียนรู้แบบโครงงานเป็นฐาน ตามแผนการสอนฐานสมรรถนะแบบผสมผสานด้วย MIAP รายวิชาเทคโนโลยีสารสนเทศและการสื่อสาร มาใช้ในการจัดการเรียนการสอนเพื่อสนับสนุนการเรียนที่เน้นฐานสมรรถนะเกี่ยวกับการใช้โปรแกรมประยุกต์พื้นฐาน โดยมุ่งหวังให้ผู้เรียนแต่ละคนสามารถสร้างความรู้ผ่านกระบวนการคิด และลงมือปฏิบัติด้วยตนเอง อีกทั้งยังเป็นการทบทวนให้ผู้เรียนเกิดการเรียนรู้อย่างมีประสิทธิภาพ ส่งผลให้ผู้เรียนมีผลสัมฤทธิ์ทางการเรียนที่ดีขึ้น และมีสมรรถนะในการใช้งานโปรแกรมประยุกต์พื้นฐาน สามารถนำไปประยุกต์ใช้ในการทำงานในอนาคตต่อไป

1.2 วัตถุประสงค์ของงานวิจัย

1.2.1 เพื่อพัฒนาบทเรียนคอมพิวเตอร์ช่วยสอนผ่านเว็บด้วยเทคนิคการเรียนรู้แบบโครงงานเป็นฐาน ตามแผนการสอนฐานสมรรถนะแบบผสมผสานด้วย MIAP รายวิชาเทคโนโลยีสารสนเทศและการสื่อสาร

1.2.2 เพื่อหาประสิทธิภาพของบทเรียนคอมพิวเตอร์ช่วยสอนผ่านเว็บที่พัฒนาขึ้น

1.2.3 เพื่อเปรียบเทียบผลสัมฤทธิ์ทางการเรียนของผู้เรียน ก่อนและหลังเรียนด้วยบทเรียนคอมพิวเตอร์ช่วยสอนผ่านเว็บที่พัฒนาขึ้น

1.2.4 เพื่อหาคุณภาพของโครงงานที่ผู้เรียนจัดทำขึ้น

2. วรรณกรรมที่เกี่ยวข้อง

2.1 สมรรถนะ

“สมรรถนะ” แปลมาจากคำว่า Competency หมายถึงความสามารถในการปฏิบัติงานหรือการประกอบอาชีพได้เป็นอย่างดี ประกอบด้วย 3 องค์ประกอบ ได้แก่ 1) ความรู้ ความเข้าใจในงานที่ทำ 2) ทักษะในการทำงานให้บรรลุผลสำเร็จ และ 3) กิจนิสัยหรือเจตคติในการทำงาน การฝึกอบรมบุคคลให้มีสมรรถนะจึงเป็นสาระสำคัญของการจัดการเรียนการสอน [4]

2.2 บทเรียนคอมพิวเตอร์ช่วนสอนผ่านเว็บ

บทเรียนคอมพิวเตอร์ช่วยสอนที่นำเสนอผ่านเครือข่ายคอมพิวเตอร์ โดยใช้เว็บเบราเซอร์เป็นตัวจัดการ ดังนั้นจึงมีความแตกต่างกับคอมพิวเตอร์ช่วยสอนธรรมดาอยู่บ้างในส่วนของการใช้งาน ได้แก่ ส่วนของระบบการติดต่อกับผู้ใช้ ระบบการนำเสนอบทเรียน ระบบการสืบท่องข้อมูล และระบบการจัดการบทเรียน [5]

2.3 การเรียนรู้แบบโครงงานเป็นฐาน

สุชาติ [6] กล่าวถึงขั้นตอนการทำโครงงานว่าเป็นกิจกรรมที่ต้องกระทำอย่างต่อเนื่อง ตั้งแต่เริ่มต้นจนกระทั่งเสร็จสิ้นโครงงาน ซึ่งผู้เรียนจะต้องเป็นผู้ดำเนินการเองทั้งสิ้น โดยมีผู้สอนทำหน้าที่เป็นที่ปรึกษา ประกอบด้วย 6 ขั้นตอน ดังนี้

2.3.1 การคิดและเลือกหัวข้อเรื่อง เป็นการคิดหาหัวข้อเรื่องที่จะทำโครงงาน ซึ่งได้มาจากปัญหา คำถาม หรือความอยากรู้อยากเห็นในเรื่องต่าง ๆ ของผู้เรียนเอง

2.3.2 การศึกษาเอกสารที่เกี่ยวข้อง เป็นการศึกษาเอกสารที่เกี่ยวข้อง ซึ่งจะช่วยทำให้ได้แนวคิดในการกำหนดขอบข่ายหรือเค้าโครงของเรื่องที่จะศึกษาชัดเจนยิ่งขึ้น

2.3.3 การเขียนเค้าโครงของโครงงาน เป็นการสร้างแผนที่ความคิดของโครงงานที่วิเคราะห์ไว้มาจัดทำรายละเอียดเพื่อแสดงแนวคิด แผน และขั้นตอนการทำโครงงาน

2.3.4 การปฏิบัติโครงงาน เป็นการดำเนินงานหลังจากที่โครงงานได้รับความเห็นชอบจากผู้สอน ตามแผนงานที่กำหนดไว้ในเค้าโครงของโครงงาน ซึ่งถือว่าเป็นการเรียนรู้เนื้อหา และฝึกทักษะต่าง ๆ ตามที่ระบุไว้ในจุดประสงค์การเรียนรู้

2.3.5 การเขียนรายงาน เป็นการสรุปรายงานผลการดำเนินงานโครงงาน เพื่อให้ผู้อื่นได้ทราบถึงแนวคิด วิธีการดำเนินงาน ผลที่ได้รับ ตลอดจนข้อสรุป ข้อเสนอแนะต่าง ๆ

2.3.6 การแสดงผลงาน เป็นการนำเสนอผลการดำเนินงานโครงงานทั้งหมดให้ผู้อื่นได้ทราบ ซึ่งผลผลิตที่ได้อาจมีลักษณะเป็นเอกสาร รายงาน ชิ้นงาน แบบจำลอง ฯลฯ

2.4 การเรียนแบบผสมผสาน

เป็นรูปแบบการจัดการเรียนการสอนที่ผสมผสานการเรียนบนเว็บและการเรียนในห้องเรียน ซึ่งมีความยืดหยุ่น และตอบสนองต่อความแตกต่างระหว่างบุคคลของผู้เรียน ทั้งด้านรูปแบบการเรียน การคิด และทักษะของผู้เรียนแต่ละคน [7]

2.5 การเรียนการสอนแบบ MIAP

สุชาติ [8] ได้กล่าวว่า บุคคลจะเกิดการเรียนรู้ได้ ก็ต่อเมื่อได้ผ่านกระบวนการเรียนรู้ ซึ่งแบ่งออกเป็น 4 ขั้นตอน ดังนี้

2.5.1 ขั้นสนใจปัญหา (Motivation) การเรียนรู้จะเกิดขึ้นเมื่อผู้เรียนมีความพร้อมที่จะเรียน ตั้งใจและสนใจที่จะเรียน ต้องการทำอะไรบางอย่างที่แปลกใหม่หรือไม่เคยทำได้มาก่อน เมื่อประสบปัญหาจึงเกิดความสนใจที่จะแก้ปัญหานั้น

2.5.2 ขั้นศึกษาข้อมูล (Information) เมื่อผู้เรียนประสบปัญหาย่อมมีความต้องการหรือสนใจที่จะแก้ปัญหานั้น แต่ด้วยเหตุที่เป็นปัญหาใหม่ ไม่เคยรู้ หรือทำได้มาก่อน จึงต้องการศึกษาข้อมูล และทำการเก็บรวบรวมข้อมูลหรือเนื้อหาเรื่องราวต่าง ๆ เพื่อที่จะได้นำไปใช้ในการแก้ปัญหานั้น

2.5.3 ขั้นพยายาม (Application) เพื่อให้เกิดความมั่นใจได้ว่า ข้อมูลหรือเนื้อหาเรื่องราวที่ได้จากการศึกษามานั้นถูกต้องเพียงพอสำหรับการแก้ปัญหาจริง ย่อมต้องการฝึกหัดทดลองใช้งานข้อมูลหรือเนื้อหานั้นในการแก้ปัญหาจริง

2.5.4 ขั้นสำเร็จผล (Progress) การได้พยายามนำข้อมูลหรือเนื้อหามาใช้แก้ปัญหานั้น ๆ ย่อมจะได้ผลลัพธ์ออกมา ถ้าหากข้อมูลที่ศึกษามานั้นมีความถูกต้องและมีปริมาณที่มากเพียงพอก็จะแก้ปัญหาดังกล่าวให้สำเร็จลงได้ ซึ่งจะทำให้ผู้เรียนเกิดการเรียนรู้ขึ้น แต่หากแก้ปัญหาไม่ได้หรือไม่สำเร็จ ก็จะต้องทบทวนย้อนขั้นตอนของกระบวนการอีกครั้ง

2.6 เกณฑ์การให้คะแนนแบบ Scoring Rubrics

การประเมินผลการเรียน ผู้สอนจำเป็นต้องมีเครื่องมือในการให้คะแนน เพื่ออธิบายความสัมฤทธิ์ผลของผู้เรียน ซึ่งการวัดและประเมินผลกำลังเปลี่ยนแปลงสู่การประเมินตามสภาพจริงและการประเมินอิงการปฏิบัติมากขึ้น ผู้ประเมินควรกำหนดเกณฑ์การประเมินให้ผู้เรียนได้ทราบก่อนการทำงาน เกณฑ์การประเมินนี้จะระบุคุณภาพที่ต้องการให้ผู้เรียนกระทำ โดยแต่ละคะแนนบน Rubrics จะสอดคล้องกับการตอบสนองนั้น ๆ ดังนั้นการกำหนดเกณฑ์การให้คะแนนอย่างชัดเจนจะทำให้ผู้เรียนได้รู้ว่าผู้สอนต้องการอะไร และจะต้องทำอย่างไรเพื่อที่จะทำให้ผลงานได้คะแนนในระดับที่ต้องการ [9]

2.7 งานวิจัยที่เกี่ยวข้อง

พงษ์ดนัยและจิรพันธุ์ [10] ได้พัฒนาบทเรียนคอมพิวเตอร์ช่วยสอนผ่านเว็บตามฐานสมรรถนะรายวิชาทฤษฎี

ภาพเคลื่อนไหวร่วมกับเทคนิคการเรียนรู้แบบโครงงานเป็นฐาน ผลการวิจัยสรุปได้ว่า บทเรียนคอมพิวเตอร์ช่วยสอนผ่านเว็บที่พัฒนาขึ้นมีประสิทธิภาพตามเกณฑ์ที่กำหนด ส่งผลทำให้ผู้เรียนมีผลสัมฤทธิ์ทางการเรียนหลังเรียนสูงกว่าก่อนเรียนอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 และผลการประเมินคุณภาพโครงงานของผู้เรียนในภาพรวมอยู่ในเกณฑ์ดี แต่รูปแบบที่พัฒนาขึ้นนี้ ยังขาดในส่วนของกระบวนการติดตาม ตรวจสอบว่าผู้เรียนสามารถปฏิบัติชิ้นงานได้หรือไม่ ผู้วิจัยจึงเพิ่มส่วนของกระบวนการเรียนการสอนแบบผสมผสานด้วย MIAP เพื่อเป็นการฝึกทักษะปฏิบัติงานให้ผู้เรียนเกิดสมรรถนะที่คาดหวัง

ประสิทธิชัย ปณิดา และปรัชญนันท์ [11] ได้พัฒนาเว็บฝึกอบรมแบบผสมผสานด้วยกระบวนการเรียนการสอนรู้แบบ MIAP เรื่องการสร้างบทเรียนออนไลน์ ผลการวิจัยสรุปได้ว่า เว็บฝึกอบรมที่พัฒนาขึ้นมีคุณภาพเหมาะสม สามารถทำให้ผู้อบรมมีสมรรถนะเป็นไปตามเกณฑ์ มีผลสัมฤทธิ์สูงกว่าก่อนอบรมอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 และผู้อบรมมีความพึงพอใจอยู่ในระดับมากที่สุด โดยรูปแบบที่พัฒนาขึ้นนี้เป็นการสร้างสมรรถนะในการสร้างบทเรียนออนไลน์ให้กับผู้อบรม แต่ยังขาดในส่วนของการให้ผู้อบรมสร้างสรรค์ชิ้นงานของตนเองขึ้น ผู้วิจัยจึงเห็นควรที่จะเพิ่มส่วนของการจัดทำโครงงาน เพื่อเป็นการวัดคุณภาพของโครงงานที่ผู้อบรมได้จัดทำขึ้นตามเทคนิคการเรียนรู้แบบโครงงานเป็นฐาน

3. วิธีการดำเนินการวิจัย

3.1 กำหนดกลุ่มเป้าหมายและแบบแผนการทดลอง

นักศึกษาระดับปริญญาตรี มหาวิทยาลัยราชภัฏนครปฐม ที่ลงทะเบียนเรียนรายวิชาเทคโนโลยีสารสนเทศและการสื่อสารในปีการศึกษาที่ 2/2559 จำนวน 49 คน ซึ่งได้มาจากการสุ่มอย่างง่ายด้วยวิธีจับฉลากห้องเรียน โดยกำหนดแบบแผนการทดลองแบบกลุ่มเดียวสอบก่อน-สอบหลัง

3.2 พัฒนาบทเรียนตามขั้นตอน ADDIE Model

3.2.1 ศึกษาและวิเคราะห์เนื้อหารายวิชา (Analysis)

ศึกษาเนื้อหา คำอธิบายรายวิชา วัตถุประสงค์ทั่วไป และกำหนดเป็นรายการงานตามฐานสมรรถนะที่ไม่เป็นไปตามที่คาดหวังในส่วนของปฏิบัติการใช้โปรแกรมประยุกต์พื้นฐานด้านการประมวลผลคำ จำนวน 1 งาน คือ การสร้างแผ่นพับแบบ 3 ทบ ด้วยโปรแกรม Microsoft Word 2013 จากนั้นกำหนดวัตถุประสงค์เชิงพฤติกรรม วิเคราะห์ลำดับขั้นตอนการปฏิบัติงาน สร้างแผนภาพปะการัง เพื่อแสดงความสัมพันธ์ของเนื้อหา และเขียน Network Diagram เพื่อแสดงลำดับของเนื้อหา ต่อไปทำการออกแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนแบบปรนัย 4 ตัวเลือก แล้วนำไปประเมินความสอดคล้อง (IOC) และความเป็นปรนัยโดยผู้เชี่ยวชาญด้านเนื้อหา จากนั้นดำเนินการปรับปรุงแก้ไขตามข้อเสนอแนะก่อนนำไปหาคุณภาพของแบบทดสอบ โดยได้ข้อสอบที่มีคุณภาพเหมาะสมจำนวน 20 ข้อ จากนั้นวิเคราะห์ผู้เรียนและเทคโนโลยีที่จะใช้ในการพัฒนาบทเรียน รวมถึงศึกษาเทคนิคการสอนแบบโครงงานเป็นฐาน การเรียนแบบผสมผสาน เทคนิคการเรียนการสอนแบบ MIAP และเกณฑ์การประเมินแบบรูบริค เพื่อกำหนดกิจกรรมการเรียนการสอน การวัดประเมินผลให้มีความเหมาะสมสอดคล้องกัน และเนื้อหาอื่น ๆ ที่เกี่ยวข้อง แล้วรวบรวมจัดทำเป็นแผนการสอนฐานสมรรถนะแบบผสมผสานด้วย MIAP โดยผู้วิจัยได้เลือกนำเสนอบทเรียนบนเว็บผ่านระบบบทเรียนออนไลน์ของมหาวิทยาลัยราชภัฏนครปฐม

3.2.2 ออกแบบบทเรียน (Design)

ก) ออกแบบรูปแบบระบบบทเรียน ประกอบด้วย 3 ส่วน คือ 1) เทคนิคการเรียนรู้แบบโครงงานเป็นฐาน 2) ฐานสมรรถนะแบบผสมผสานด้วย MIAP 3) การสนับสนุนแบบมีส่วนร่วม

ข) ออกแบบการจัดกิจกรรมการเรียนรู้ดังภาพที่ 1

ค) ออกแบบประเมินคุณภาพบทเรียน

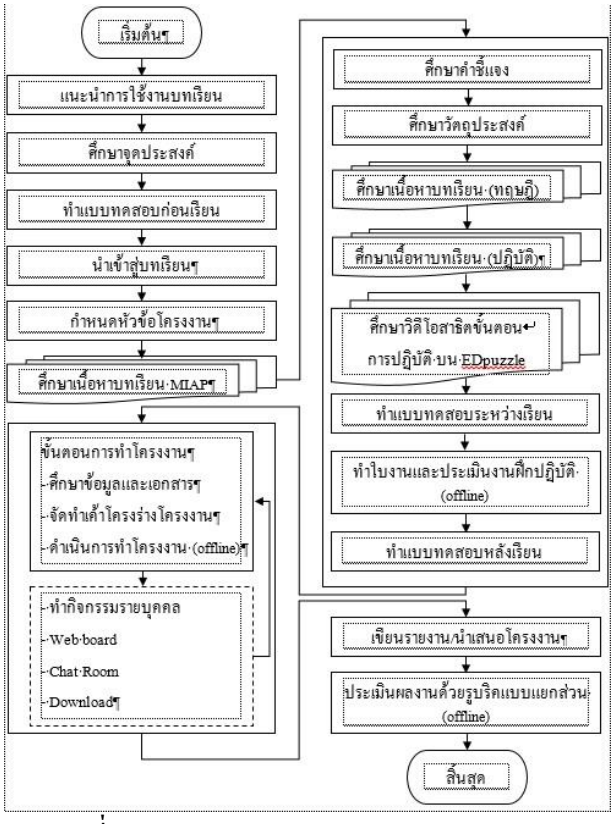
ง) ออกแบบประเมินคะแนนรูบริคแบบแยกส่วน ซึ่งกำหนดเกณฑ์การประเมินคุณภาพโครงงานออกเป็น 2 ด้าน คือ ด้านคุณภาพผลงาน และด้านเจตคติในการทำงาน

จ) ออกแบบหน้าจอและเขียน Storyboard

3.2.3 พัฒนาบทเรียน (Development)

ก) เตรียมการ เป็นการเตรียมวัสดุต่าง ๆ เช่น ข้อความ ภาพนิ่ง ภาพเคลื่อนไหว เสียงบรรยาย และวีดีโอสาธิต

ข) สร้างเอกสารประกอบบทเรียน ได้แก่ ใบเนื้อหา ใบลำดับขั้นตอน ใบงาน ใบสั่งงาน ใบประเมินผลการปฏิบัติ แบบฟอร์มเสนอโครงงาน และใบประเมินผลงานแบบรูบริค



ภาพที่ 1: รูปแบบกิจกรรมการเรียนรู้แบบโครงงานเป็นฐานร่วมกับการเรียนแบบผสมผสานด้วย MIAP

ค) สร้างบทเรียน เป็นการนำข้อมูลที่เตรียมการมาสร้างบทเรียนคอมพิวเตอร์ช่วยสอน จากนั้นติดตั้งบนระบบบทเรียนออนไลน์ และทำการสร้างกิจกรรมการเรียนตามเทคนิคโครงงานเป็นฐาน

3.2.4 การทดลองใช้ (Implementation)

ก) ทดลองใช้โดยผู้พัฒนาและปรับปรุงแก้ไข

ข) ทดลองใช้กับกลุ่มย่อย ซึ่งไม่ใช่กลุ่มเป้าหมาย แต่มีลักษณะเทียบเคียงกับกลุ่มเป้าหมายซึ่งก็คือนักศึกษาระดับปริญญาตรี ที่ลงทะเบียนในรายวิชาเทคโนโลยีสารสนเทศและการสื่อสารในปีการศึกษาที่ 2/2559 จำนวน 9 คน แบ่งเป็น 3 กลุ่ม ได้แก่ กลุ่มเก่ง กลุ่มกลาง กลุ่มอ่อน กลุ่มละ 3 คน แล้วสอบถามความคิดเห็นและนำผลที่ได้ไปปรับปรุงแก้ไขบทเรียน

ค) ทดลองใช้โดยผู้เชี่ยวชาญด้านเนื้อหา และด้านเทคนิควิธีการ ด้านละ 3 ท่าน เพื่อประเมินคุณภาพบทเรียนที่พัฒนาขึ้นด้วยแบบประเมินคุณภาพบทเรียน โดยมีผลการประเมินอยู่ในระดับดีมาก จากนั้นปรับปรุงแก้ไขบทเรียนตามข้อเสนอแนะ

ง) ทดลองใช้บทเรียนกับกลุ่มเป้าหมาย ตามแบบแผนการทดลองแบบกลุ่มเดียวสอบก่อน-สอบหลัง

3.2.5 ประเมินผลข้อมูล (Evaluation)

ประเมินผลข้อมูลที่ได้จากการดำเนินการทดลองกับกลุ่มเป้าหมายด้วยสถิติ ได้แก่ ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน ประสิทธิภาพ E1/E2 และ t-test แบบ Dependent รวมทั้งการประเมินคุณภาพโครงงานของผู้เรียนด้วยรูบริคแบบแยกส่วน

4. ผลการดำเนินการวิจัย

4.1 ผลการพัฒนาบทเรียน

บทเรียนคอมพิวเตอร์ช่วยสอนผ่านเว็บที่พัฒนาขึ้นประกอบด้วยเนื้อหาจำนวน 1 หน่วย คือ การสร้างแผ่นพับแบบ 3 ทบ ด้วยโปรแกรม Microsoft Word 2013 ประกอบด้วยเนื้อหา 3 ส่วน ได้แก่ 1) เนื้อหาความรู้ทฤษฎี 2) เนื้อหาความรู้ทักษะปฏิบัติ และ 3) วีดีโอสาธิตขั้นตอนการปฏิบัติงาน



ภาพที่ 2: หน้าจอการนำเสนอเนื้อหาบทเรียนบน NPRU LMS และวิดีโอสาธิตบน EDpuzzle

4.2 ผลการหาประสิทธิภาพของบทเรียน

ประสิทธิภาพของบทเรียนเท่ากับ 83.06/81.02 สูงกว่าเกณฑ์ที่กำหนด 80/80 สรุปได้ว่าบทเรียนที่พัฒนาขึ้น สามารถนำมาใช้ในการจัดการเรียนการสอนได้ ดังตารางที่ 1

ตารางที่ 1: ผลการหาประสิทธิภาพของบทเรียน

แบบทดสอบ	คะแนนเฉลี่ย	S.D.	ประสิทธิภาพ
ระหว่างเรียน (E1)	16.61	1.51	83.06
หลังเรียน (E2)	16.20	1.53	81.02

4.3 ผลการเปรียบเทียบผลสัมฤทธิ์ทางการเรียน

ผู้เรียนมีผลสัมฤทธิ์ทางการเรียนสูงขึ้นอย่างมีนัยสำคัญที่ระดับ .05 โดยการทดสอบค่าที ดังตารางที่ 2

ตารางที่ 2: ผลการเปรียบเทียบผลสัมฤทธิ์ทางการเรียน

แบบทดสอบ	n	$\bar{x}$	S.D.	$t_{คำนวณ}$	$t_{ตาราง}$
ก่อนเรียน	49	9.08	1.72	*30.6191	1.6772
หลังเรียน	49	16.20	1.53		

*มีนัยสำคัญทางสถิติที่ระดับ .05

4.4 ผลการหาคุณภาพโครงงานของผู้เรียน

คุณภาพโครงงานของผู้เรียนในภาพรวมอยู่ในระดับดี โดยการประเมินด้วยรูบริคแบบแยกส่วน ดังตารางที่ 3

ตารางที่ 3: ผลการประเมินคุณภาพโครงงาน

รายการประเมิน	ค่าเฉลี่ย
คุณภาพผลงาน (วัดได้)	
- สร้างแผ่นพับได้ถูกต้องตามรูปแบบที่กำหนด	4.01
- มีหัวข้อเนื้อหาครบถ้วนตามที่กำหนด	3.95
- มีส่วนประกอบครบถ้วนสมบูรณ์ตามเงื่อนไขที่กำหนด	3.96
คุณภาพผลงาน (วัดไม่ได้)	
- ความสวยงาม	3.88
- ความเรียบร้อย	3.99
- ความประณีต	3.98
เจตคติ (กิจนิสัยฯ) ในการทำงาน	
- ความรับผิดชอบ (อ้างอิง)	4.09
- ความรอบคอบ	3.96
- ความตรงต่อเวลา	4.09
ภาพรวม	4.00

5. อภิปรายผลการวิจัย

จากผลการวิจัยพบว่า ผู้เรียนมีผลสัมฤทธิ์ทางการเรียนสูงขึ้นอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 เนื่องจากมีกระบวนการพัฒนาบทเรียนอย่างเป็นระบบตาม ADDIE Model ส่งผลให้บทเรียนมีประสิทธิภาพเหมาะสมสามารถนำไปใช้ในการเรียนการสอนได้ รวมทั้งยังตอบสนองต่อปัญหาการขาดสื่อการเรียนการสอนให้ผู้เรียนทบทวนเพิ่มเติมนอกเหนือจากเวลาเรียนได้ โดยการจัดกิจกรรมการเรียนรู้ด้วยเทคนิคโครงงานเป็นฐานนั้น ส่งเสริมให้ผู้เรียนเกิดทักษะในรายวิชาที่ศึกษาและสิ่งที่ตนสนใจ ผ่านการเรียนรู้แบบผสมผสานด้วย MIAP ทั้งในรูปแบบเผชิญหน้าในห้องเรียนและการเรียนแบบออนไลน์เข้าด้วยกัน ซึ่งเน้นให้เกิดการเรียนรู้ผ่านการฝึกปฏิบัติจนเกิดความชำนาญ ดังจะเห็นได้จากคุณภาพโครงงานที่ผู้เรียนได้จัดทำขึ้นในภาพรวมอยู่ในระดับดี

6. ข้อเสนอแนะ

ผู้วิจัยเห็นว่าการวิจัยในครั้งต่อไป ควรจะมีการพัฒนาในส่วนของการฝึกปฏิบัติให้เป็นโปรแกรมสำเร็จรูปที่สามารถตรวจจับการฝึกปฏิบัติของผู้เรียน โดยสามารถแนะนำผู้เรียนเมื่อไม่สามารถที่จะปฏิบัติชิ้นงานต่อไปได้

เอกสารอ้างอิง

[1] สราวุธ ไพฑูรย์พงษ์, *คุณภาพแรงงานไทย (An Overview of Thailand's Quality of Labour)*, กรุงเทพฯ: สถาบันวิจัยเพื่อการพัฒนาประเทศไทย, 2551.

[2] David C. Mcclelland, "Testing for competence rather than for intelligence," *American Psychologist*, vol. 28, no.1, pp. 1-14, Jan 1973.

[3] ทิศนา แขมมณี, *ศาสตร์การสอน: องค์ความรู้เพื่อการจัดการกระบวนการเรียนรู้ที่มีประสิทธิภาพ*, พิมพ์ครั้งที่ 18, กรุงเทพฯ: สำนักพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย, 2557.

[4] ธำรง บัวศรี, *ทฤษฎีหลักสูตร การออกแบบและพัฒนา*, พิมพ์ครั้งที่ 4 กรุงเทพฯ: โรงพิมพ์คุรุสภาลาดพร้าว, 2542.

[5] มนต์ชัย เทียนทอง, *การออกแบบและพัฒนาคอร์สแวร์ สำหรับบทเรียนคอมพิวเตอร์ช่วยสอน*, กรุงเทพฯ: ศูนย์ผลิตตำราเรียน สถาบันเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2545.

[6] สุชาติ วงศ์สุวรรณ, *การเรียนรู้สำหรับศตวรรษที่ 21: การเรียนรู้ที่ผู้เรียนเป็นผู้สร้างความรู้ด้วยตนเอง*, กรุงเทพฯ: กรมวิชาการ, 2542.

[7] Bonk, C.J., & Graham, C.R, *Handbook of blended learning: Global perspectives*, San Francisco: Pfeiffer Publishing, 2004.

[8] สุชาติ ศิริสุขไพบูลย์, *เทคนิคและวิธีการสอนวิชาชีพ MIAP*, กรุงเทพฯ: สถาบันเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2549.

[9] เอกรินทร์ สี่มหาศาล และ สุปรารถนา ยุกตะนันทน์, *การออกแบบเครื่องมือวัดและประเมินตามสภาพจริง*, กรุงเทพฯ: บุ๊คพอยท์, 2546.

[10] พงษ์ดนัย จิตตวิสุทธิกุล และ จิรพันธุ์ ศรีสมพันธุ์, "การพัฒนาบทเรียนคอมพิวเตอร์ช่วยสอนผ่านเว็บตามฐานสมรรถนะรายวิชาทฤษฎีภาพเคลื่อนไหวร่วมกับเทคนิคการเรียนรู้แบบโครงงานเป็นฐาน", *The 8th National Conference on Technical Education (NCTechEd08)*, pp. 229-234, Nov 26, 2015.

[11] ประสิทธิชัย มั่งมี, ปณิตา วรรณพิรุณ และปรัชญนันท์ นิลสุข, "การพัฒนาเว็บฝึกอบรมแบบผสมผสานด้วยกระบวนการเรียนการสอนรูปแบบ MIAP เรื่อง การสร้างบทเรียนออนไลน์", *Technical Education Journal King Mongkut's University of Technology North Bangkok*, vol. 5, no. 2, pp. 58-66, 2014.

การพัฒนาระบบบริหารจัดการบทเรียนแสวงรู้
The Development of the Web Quest Online System

ปฐมาวดี เกรยรัมย์ (Pattamavadee Kroeiram)[1] และจารี ทองคำ (Jaree Thongkam)[2]
สาขาเทคโนโลยีสารสนเทศคณะวิทยาการสารสนเทศ มหาวิทยาลัยมหาสารคาม
[1]an_lukpla@hotmail.com, [2] jaree.thongkam@gmail.com

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาระบบบริหารจัดการบทเรียนแสวงรู้ เพื่อสนับสนุนการนาบทเรียนไปใช้ในการเรียนการสอน เพื่อให้นักศึกษาได้แสวงรู้ด้วยตนเอง ระบบถูกพัฒนาขึ้นด้วยโปรแกรม PHP และมี MySQL เป็นฐานข้อมูล ระบบนี้ได้ถูกนาไปใช้จริงโดยนักศึกษาภาควิชาเทคโนโลยีสารสนเทศเพื่องานสื่อสารมวลชน สาขาวิชาเทคโนโลยีมัลติมีเดีย มหาวิทยาลัยเทคโนโลยีราชมงคลอีสาน ผู้วิจัยได้ใช้หลักการของ black box ในการประเมินประสิทธิภาพและประสิทธิผลของระบบ และนอกจากนี้ยังได้ประเมินความพึงพอใจของอาจารย์ผู้ใช้งานระบบ โดยแบ่งเป็น 4 ด้าน ได้แก่ 1) ด้านความต้องการของผู้ใช้ระบบ 2) ด้านการทางานตามฟังก์ชันของระบบ 3) ด้านทดสอบการใช้งาน และ 4) ด้านทดสอบความปลอดภัย ผลการวิจัยพบว่า ระบบบริหารจัดการบทเรียนแสวงรู้สามารถทางานได้ทุกฟังก์ชันเป็นอย่างดี ผลการประเมินความพึงพอใจของผู้ใช้งานระบบพบว่ามีค่าเฉลี่ยในระดับดีมากทุกด้าน โดยมีค่าเฉลี่ย 4.57, 4.59, 4.54 ,4.62 ตามลาดับ ซึ่งผลของความพึงพอใจที่มีต่อระบบในภาพรวมอยู่ในระดับดีมาก คือ 4.58 แสดงให้เห็นว่าระบบที่พัฒนาขึ้นมามีคุณภาพสามารถนาไปใช้งานได้

คำสำคัญ: บทเรียนแสวงรู้, การเรียนทางอิเล็กทรอนิกส์, การเรียนรู้ออนไลน์

Abstract

This research aims to develop the Web Quest Online System to promote the lesson in teaching students for self-learning. This system was developed using PHP program and MySQL as a database management. This system is actually used by the students who major in Multimedia Technology, Department of Multimedia technology, Rajamangala University of Technology Isan. The system was evaluated using black box for system performance and effectiveness. Also, the teacher's satisfaction was evaluated with 4 phase including 1) requirement test2) working function test, 3) usability test, and security test. The results showed that every function of the Web Quest Online System successfully worked. In the part of security test, the user's satisfy with the system with highest level score. The users satisfied in the part of functional requirement test, functional test, and usability test with the highest level score of the average 4.57, 4.59, 4.54 , 4.62, respectively. In the overall satisfaction, users provided with the highest level score up to 4.58 that means the system has a good quality and functionality.

Keyword: Web Quest, E-Learning, Online Learning

1. บทนำ

การเรียนที่ส่งเสริมทักษะการคิดขั้นสูงเป็นทักษะที่จำเป็นต่อรูปแบบการเรียนในปัจจุบันเพราะการเปลี่ยนแปลงของเทคโนโลยีที่รวดเร็วและปริมาณข้อมูลสารสนเทศที่มีอยู่บนโลกนี้อย่างมหาศาล ถึงแม้ว่าปัจจุบันจะมีเทคโนโลยีสารสนเทศอยู่เป็นจำนวนมาก [1]พร้อมจะให้ผู้เรียนได้ศึกษาหาความรู้ได้ทันที แต่ใช่ว่าจะมีคุณภาพให้ข้อมูลที่ถูกต้องและผู้เรียนมีความรู้สึกว่ามีความหมายต่อการเรียนรู้ซึ่งการพัฒนาผู้เรียนจำเป็นต้องส่งเสริมทักษะกระบวนการคิดขั้นสูงและการพิจารณาความน่าเชื่อถือของข้อมูลแหล่งข้อมูลควบคู่กันไป การใช้กิจกรรมการเรียนด้วยเว็บเควสท์ประกอบด้วยส่วนสำคัญ

6 ส่วน ดังนี้ 1) ส่วนนำ 2) ภารกิจ 3) การชี้แหล่งความรู้ 4) กระบวนการ 5) ประเมินผล และ 6) สรุป [2] โดยขั้นตอนของการเรียนในแต่ละส่วนส่งเสริมการใช้ข้อมูลสารสนเทศที่มีอยู่บนอินเทอร์เน็ต

โดยกิจกรรมการเรียนด้วยเว็บเควสท์เป็นการออกแบบสภาพแวดล้อมใหม่ในการเรียนในยุคสารสนเทศ ที่ผู้เรียนต้องมีทักษะการเลือกสรรข้อมูลการประยุกต์ใช้การประเมินค่า และการสร้างสรรค์ ต่อจากนั้นต้องทำการวิเคราะห์ สังเคราะห์เพื่อให้ได้แนวทางหรือคำตอบที่ดีที่สุด การผ่านกระบวนการการทำงานของสมองหรือระบบความคิดดังกล่าวบ่อยๆครั้งส่งผลให้ผู้เรียนปฏิบัติจนเกิดเป็นนิสัยและแสดงออกทุกครั้งเมื่อเผชิญกับปัญหาจากการเรียน การใช้ชีวิตประจำวันจึงเรียกกระบวนการนั้นว่า ทักษะการคิดขั้นสูง [3] ปัจจุบันเทคโนโลยีออนไลน์ เข้ามามีบทบาทในชีวิตประจำวันมากยิ่งขึ้น ใน การพัฒนาระบบออนไลน์ภาษาที่ใช้กันส่วนใหญ่คือ ภาษา PHP ซึ่งเป็นภาษาที่ได้รับความนิยมที่สุด เพราะการแสดงผลของ PHP จะปรากฏในลักษณะ HTML [4] นอกจากนี้ PHP ยังเป็นภาษาที่เรียนรู้และเริ่มต้นได้ไม่ยาก โดยมีเครื่องมือช่วยเหลือและคู่มือที่สามารถหาอ่านได้ฟรีบนอินเทอร์เน็ต ความสามารถการประมวลผลหลักของ PHP [4] ภาษาที่ใช้ในระบบจัดการฐานข้อมูล MySQL คือ ภาษาระบบจัดการฐานข้อมูล MySQL มีจุดเด่นที่ความเร็วในการจัดการ มีความน่าเชื่อถือและใช้งานง่าย สามารถเชื่อมโยงฐานข้อมูลเข้ากับโปรแกรมประยุกต์อื่นได้ง่ายและรวดเร็ว คุณลักษณะเด่นของระบบจัดการฐานข้อมูล

ดังนั้นผู้วิจัยจึงมีแนวคิดที่จะพัฒนาระบบบริหารจัดการบทเรียนแสวงรู้เพื่อส่งเสริมให้นักศึกษาได้มีทักษะ การคิดขั้นสูงรายวิชาเทคโนโลยีสารสนเทศเพื่องานสื่อสารมวลชนของนักศึกษาสาขาวิชาเทคโนโลยีมัลติมีเดียมหาวิทยาลัยเทคโนโลยีราชมงคลความสำคัญในการเรียนเป็นเครื่องมือในการเรียนรู้สามารถติดต่อสื่อสารใช้งานกับแหล่งที่มาของข้อมูลเครือข่ายมีกระบวนการวิเคราะห์สังเคราะห์ความรู้ ที่เน้นความคิดสร้างสรรค์ และประเมินผลโดยผ่านกิจกรรมแบบเว็บเควสท์ด้วยการนำเอา ภาษา PHP และระบบจัดการฐานข้อมูลมาใช้ในเก็บข้อมูลต่างๆ ในการวัดประสิทธิภาพของระบบนี้ ผู้วิจัยได้นำเอาแบบสอบถามความพึงพอใจการใช้ระบบ มาใช้ในการวิเคราะห์ประสิทธิภาพของระบบด้วยค่าเฉลี่ย ในระดับ 3.5 จาก 5 ระดับ

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 เว็บเควสท์

WebQuest [2] คือ กิจกรรมการเรียนการสอนที่เน้นการแสวงรู้ โดยใช้เทคโนโลยีสารสนเทศเป็นฐานในการปฏิสัมพันธ์กับผู้เรียน บนแหล่งต่างๆ บนอินเตอร์เน็ตและอาจเสริมด้วยระบบการประชุมทางไกล WebQuest ได้รับการออกแบบที่จะใช้เวลาของผู้เรียนอย่างมีประสิทธิภาพ เน้นการใช้สารสนเทศมากกว่าการแสวงหาสารสนเทศ สนับสนุนผู้เรียนในการเรียนรู้ขั้นตอนการคิดอย่างวิเคราะห์ สังเคราะห์และการประเมินค่า WebQuest จะส่งเสริมให้ผู้เรียนใช้จินตนาการและทักษะการแก้ปัญหา คำตอบสุดท้าย แนวทางการแก้ปัญหา ดังนั้นผู้เรียนจึงต้องค้นพบและสร้างองค์ความรู้ด้วยตนเอง หรือในกลุ่มของกลุ่มผู้เรียน ผู้เรียนจะท่องไปใน World Wide Web ที่กำหนดไว้อย่างความหมาย

Web Quest มีองค์ประกอบ อยู่ 6 ส่วน คือ ส่วนนำ (Introduction) ภารกิจ (Task) การชี้แหล่งความรู้ (Resources) กระบวนการ (Process) ประเมินผล (Evaluation) สรุป (Conclusion)

2.2 การสร้างเว็บเควสท์

การสร้างเว็บเควสท์ คือการจัดสภาพแวดล้อมในการเรียนรู้ของผู้เรียน โดยมีการออกแบบวางแผน พัฒนาบทเรียน มีขั้นตอน 7 ขั้นตอนดังนี้ 1. เลือกหัวข้อ 2. เลือกการออกแบบ 3. เลือกเครื่องมือในการพัฒนา 4. สร้างการประเมินผล 5. พัฒนากระบวนการ 6. ทดสอบ / เผยแพร่ 7. ประเมินผลเว็บเควสท์

2.3 ระบบ LMS

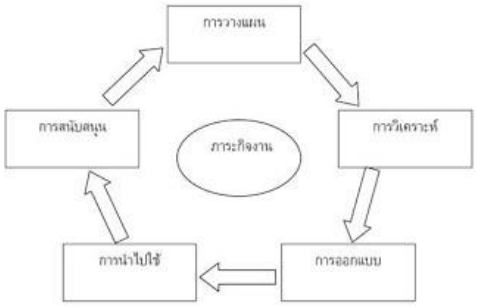
LMS [5] เป็นคำที่ย่อมาจาก Learning Management System หรือระบบการจัดการเรียนรู้เป็นซอฟต์แวร์ที่ทำหน้าที่บริหารจัดการเรียนการสอนผ่านเว็บจะประกอบด้วยเครื่องมืออำนวยความสะดวกให้แก่ผู้สอน ผู้เรียน ผู้ดูแลระบบ โดยที่ผู้สอนนำเนื้อหาและสื่อการสอนขึ้นเว็บไซต์รายวิชาตามที่ได้ขอให้ระบบจัดไว้ให้ได้โดยสะดวก ผู้เรียนเข้าถึงเนื้อหา กิจกรรมต่างๆ ได้

โดยผ่านเว็บ ผู้สอนและผู้เรียนติดต่อสื่อสารได้ผ่านทางเครื่องมือการสื่อสารที่ระบบจัดไว้ให้

2.4 วงจรการพัฒนาระบบ

อำไร วรรณสินธุ์ [6] ได้กล่าวว่าการวิเคราะห์ระบบเป็นการศึกษาระบบงานเดิมที่เกิดปัญหาขึ้น เพื่อแก้ไขปรับปรุงและออกแบบระบบงานใหม่ให้มีทิศทางที่ดีขึ้น

วงจรการพัฒนาระบบ (System Development Life Cycle) การพัฒนาระบบสารสนเทศเป็นกระบวนการในการสร้างระบบสารสนเทศขึ้นมาเพื่อใช้สำหรับแก้ปัญหา หรือสร้างมูลค่าเพิ่มให้กับธุรกิจและด้วยระบบสารสนเทศในยุคปัจจุบันและในอนาคตจะมีความซับซ้อนยิ่งขึ้นและมีขนาดใหญ่ ดังนั้นโครงการพัฒนาระบบสารสนเทศจึงจำเป็นต้องได้รับการวางแผนที่ดี และหากเป็นโครงการขนาดใหญ่ยิ่งสมควรได้รับการเอาใจใส่เป็นพิเศษ ถึงแม้ว่าทีมงานจะเป็นผู้ที่มีประสบการณ์ก็ตามการพัฒนาซอฟต์แวร์ตามปกติจะประกอบด้วยกลุ่มกิจกรรม 3 ส่วน คือ การวิเคราะห์ (Analysis) การออกแบบ (Design) และการนำไปใช้ (Implementation) ซึ่งกิจกรรมทั้งสามนี้สามารถใช้งานได้ดีกับโครงการซอฟต์แวร์ขนาดเล็ก ในขณะที่โครงการซอฟต์แวร์ขนาดใหญ่มักจำเป็นต้องใช้แบบแผนการพัฒนาซอฟต์แวร์ตามแนวทางของ SDLC จนครบทุกกิจกรรม



ภาพที่ 1 System Development Life Cycle : SDLC

2.5 การทดสอบระบบด้วยวิธี Black Box

พรภัทรา ภัทรจารี [7] ได้กล่าวถึงการทดสอบแบบกล่องดำ Black Box Testing คือการทดสอบแบบกล่องดำเป็นการทดสอบเชิงพฤติกรรมจะใช้มุมมองภายนอกของระบบในการสร้างกรณีทดสอบ เป็นการทดสอบโดยไม่คำนึงถึงคำสั่งภายในโปรแกรม เป็นการทดสอบโดยดูค่า Output จาก Input ที่ให้กับโปรแกรมต้องมีความสอดคล้องกัน หรืออาจกล่าวได้ว่ามองซอฟต์แวร์เหมือนกล่องดำที่รับข้อมูลเข้าและให้ผลลัพธ์ออกมาเท่านั้นซึ่งจะไม่พิจารณาถึงโครงสร้างของโปรแกรม

2.6 งานวิจัยที่เกี่ยวข้อง

งานวิจัยที่เกี่ยวข้องกับการค้นคว้าอิสระนี้ได้แก่ งานของชาคริต [8] ได้ศึกษาผลของการใช้บทเรียนออนไลน์แบบเว็บเควสท์ต่อผลสัมฤทธิ์ทางการเรียนและการปฏิสัมพันธ์ระหว่างผู้เรียนกับผู้เรียนในบทเรียนออนไลน์วิชาถ่ายภาพทางการศึกษา พบว่าผลสัมฤทธิ์ทางการเรียนก่อนเรียนและหลังเรียนด้วยบทเรียนออนไลน์แบบเว็บเควสท์แตกต่างกันที่ระดับนัยสำคัญ 0.01 และปฏิสัมพันธ์ในการเรียนวิชาการถ่ายภาพทางการศึกษาสถาบันเทคโนโลยีพระจอมเกล้าพระนครเหนืออยู่ในระดับค่อนข้างดี

งานวิจัยของสุภาวดี [9] ได้ทำการศึกษาการสร้างและหาคุณภาพการเรียนการสอนเทคโนโลยีสารสนเทศและการสื่อสารข้อมูลและระบบเครือข่ายคอมพิวเตอร์ ด้วยบทเรียนออนไลน์แบบเว็บเควสท์ของนักเรียนชั้นมัธยมศึกษาปีที่ 2 มีประสิทธิภาพเท่ากับสูงกว่าเกณฑ์มาตรฐานที่กำหนดไว้ 80/80 ผลสัมฤทธิ์ทางการเรียนของผู้เรียนก่อนและหลังได้รับการเรียนด้วยบทเรียนแบบเว็บเควสท์โดยรวมมีผลสัมฤทธิ์ทางการเรียนสูงกว่าก่อนเรียนอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ผลวิเคราะห์ความพึงพอใจโดยรวมอยู่ในระดับมาก

งานวิจัยของSegers และ Verhoeveen [10] ได้ทำการศึกษาผลกระทบของการเรียนรู้ด้วยระบบแสวงรู้ในระดับประถมในเนอเธอแลนด์ ซึ่งบทเรียนแสวงรู้ได้ถูกใช้ด้วยกันกับเว็บเพจเพื่อช่วยให้นักเรียนได้มีความรู้เพิ่มขึ้น ผลปรากฏว่าการใช้บทเรียนแสวงรู้ช่วยเพิ่มทักษะการเรียนรู้ให้นักเรียนมีการเรียนรู้ที่เพิ่มขึ้นและทำให้เด็กมีทักษะในการใช้ภาษาได้ดีขึ้น

งานวิจัยของAverkieva,Chayka และGlushkov [11] ได้ทำการพัฒนาระบบบทเรียนแสวงรู้ในการสอนวิชาภาษาอังกฤษ ผลปรากฏว่า การใช้บทเรียนแสวงรู้ในการเรียนภาษาอังกฤษของนักเรียนสามารถเพิ่มแรงจูงใจในการเรียนภาษาอังกฤษ เพิ่มกระบวนการคิดของนักเรียนซึ่งคณะวิจัยได้ทำการวัดประสิทธิภาพของระบบด้วยการใช้แบบทดสอบความรู้ทำให้นักเรียนสามารถที่จะแปลข้อมูลข้อเท็จจริงและคำศัพท์ได้ดียิ่งขึ้น

งานวิจัยของ Bernardo [12] ได้ทำการนำบทเรียนแสวงรู้มาเป็นเครื่องมือใช้ในการสอยกับนักเรียนออสเตเลียเป็นกลุ่มวิชาทฤษฎีการเรียนรู้ทางสังคม การพัฒนาทางปัญญา ทฤษฎีการประมวลผลข้อมูล ผลปรากฏว่าการนำบทเรียนแสวงรู้มาใช้สามารถช่วยให้นักเรียนเพิ่มรายละเอียดในการร่วมมือกันเป็นอย่างดี มีมนุษย์สัมพันธ์ที่ดีในกลุ่มทำให้เพิ่มขีดความสามารถและส่งงานได้ตรงเวลา ทำให้นักเรียนมีความสนใจในบทเรียนมากยิ่งขึ้น

จากงานวิจัยดังกล่าวเหล่านี้สามารถนำมาใช้ในการพัฒนาระบบบริหารจัดการบทเรียนแสวงรู้

3. ขั้นตอนวิธีการดำเนินการวิจัย

3.1 การดำเนินการพัฒนาระบบ

ผู้ศึกษาได้ทำการพัฒนาระบบบริหารจัดการบทเรียนแสวงรู้โดยการใช้หลักการวงจรการพัฒนาระบบ ดังนี้

กลุ่มประชากร คือ อาจารย์ผู้สอนสาขาวิชาเทคโนโลยีมัลติมีเดีย คณะวิศวกรรมศาสตร์และสถาปัตยกรรมศาสตร์ มหาวิทยาลัยเทคโนโลยีราชมงคลอีสานจำนวน 7 คน

กลุ่มตัวอย่าง คือ อาจารย์ผู้สอนสาขาวิชาเทคโนโลยีมัลติมีเดีย คณะวิศวกรรมศาสตร์และสถาปัตยกรรมศาสตร์ มหาวิทยาลัยเทคโนโลยีราชมงคลอีสานจำนวน 7 คน โดยการเลือกแบบเจาะจง (Purposive Sampling)

1.วางแผนโครงการ (Project Planning Phase) ศึกษาหลักการของระบบ LMS มีเครื่องมือและส่วนประกอบที่สำคัญสำหรับผู้สอน ผู้เรียนและผู้ดูแลระบบอย่างไรบ้าง ศึกษาหลักการของ Web Quest มีหลักการทำงานทั้งหมด 6 ขั้นตอน คือ ส่วนนำ ภารกิจ ชี้แหล่งความรู้ กระบวนการ การประเมินผลและสรุป

2.การวิเคราะห์ (Analysis Phase) ได้ทำการวิเคราะห์ระบบโดยการกำหนดขอบเขตของการพัฒนาระบบบริหารจัดการบทเรียนแสวงรู้โดยแบ่งออกเป็นทั้งหมด 3 ส่วนคือส่วนของผู้ดูแลระบบ ส่วนของอาจารย์ผู้สอนและส่วนของนักศึกษา

3.การออกแบบ (Design Phase) ทำการออกแบบหน้าเว็บของระบบบริหารจัดการบทเรียนแสวงรู้ ออกแบบฐานข้อมูลโดยใช้ MySQL การพัฒนาระบบบริหารจัดการบทเรียนแสวงรู้ในครั้งนี้ผู้พัฒนาจะใช้ภาษา PHP ในการพัฒนาระบบ

4.การนำระบบไปใช้ (Implementation Phase) ในการทดสอบ อาจารย์ผู้สอนจะทำการระบบด้วยวิธี Black Box Testing โดยจะดูค่า Output จาก Input ที่ให้กับโปรแกรมรับข้อมูลเข้าและให้ผลลัพธ์ออกมาเท่านั้นทดสอบที่ความต้องการของโปรแกรมนั้นๆ โดยให้อาจารย์ผู้สอนตอบแบบสอบถามความพึงพอใจในการใช้งานระบบบริหารจัดการบทเรียนแสวงรู้

5.การบำรุงรักษา (Maintenance Phase) ปรับปรุงแก้ไขระบบหลังจากที่ได้มีการติดตั้งและใช้งาน

3.2 การสร้างเครื่องมือประเมินประสิทธิภาพระบบ

เครื่องมือที่ใช้ในการประเมินหาประสิทธิภาพของระบบที่พัฒนาขึ้นในครั้งนี้คือ แบบประเมินประสิทธิภาพของระบบบทเรียนแสวงรู้ กรณีศึกษา วิชาเทคโนโลยีสารสนเทศเพื่องานสื่อสารมวลชน สาขาวิชาเทคโนโลยีมัลติมีเดีย มหาวิทยาลัยเทคโนโลยีราชมงคลอีสาน ดังนี้

1.สร้างเครื่องมือวิจัยฉบับร่างเสนอคณะกรรมการที่ปรึกษาวิทยานิพนธ์ เพื่อตรวจสอบความเที่ยงตรงเชิงประจักษ์ (Face Validity) และให้ข้อเสนอแนะในการปรับปรุงแก้ไข

2.นำแบบสอบถามที่ปรับปรุงแก้ไขตามข้อเสนอแนะแล้วให้ผู้เชี่ยวชาญเพื่อพิจารณาและตรวจสอบความเที่ยงตรงเชิงเนื้อหา (Content Validity) จำนวน 3 ท่าน

3.นำแบบสอบถามที่ปรับปรุงให้มีความสมบูรณ์แล้วไปทดลองใช้ (Try out) ที่ไม่ใช่กลุ่มตัวอย่างจำนวน 30 คน เพื่อดูว่าแบบสอบถามนั้นผู้ดูแลระบบ อาจารย์ผู้สอน และนักศึกษาสามารถตอบข้อคำถามได้ครบถ้วนและความสมบูรณ์ในการตอบแบบสอบถาม

4.วิเคราะห์หาค่าความเชื่อมั่นของแบบสอบถามความพึงพอใจ มาหาค่าความเชื่อมั่น ตามวิธีของครอนบาค โดยใช้สูตรสัมประสิทธิ์แอลฟา (-Coefficient) ซึ่งได้ค่าความเชื่อมั่นทั้งฉบับเท่ากับ 0.8868

5.จากนั้นแบบสอบถามที่ปรับปรุงแก้ไขได้ไปจัดพิมพ์ฉบับสมบูรณ์ และนำไปใช้เก็บรวบรวมข้อมูลกับกลุ่มตัวอย่างต่อไป

รายละเอียดเครื่องมือทดสอบประสิทธิภาพของระบบบริหารจัดการบทเรียนแสวงรู้ ผู้วิจัยใช้การทดสอบระบบด้วยวิธีการทดสอบแบบแบล็คบอกซ์ เพื่อทดสอบความผิดพลาด

ของระบบแบ่งออกเป็น 4 ด้าน ซึ่งแบ่งการประเมินออกเป็น 4 หัวข้อหลัก คือ

1.ด้านทดสอบความต้องการ (Functional Requirement Test)

2.ด้านทดสอบฟังก์ชันการทำงาน (Functional Test)

3.ด้านทดสอบการใช้งาน (Usability Test)

4.ด้านทดสอบความปลอดภัย (Security Test)

แบบทดสอบหลังเรียนจำนวน 50 ข้อ ให้แก่นักศึกษาในรายวิชาเทคโนโลยีสารสนเทศเพื่องานสื่อสารมวลชน สาขาวิชาเทคโนโลยีมัลติมีเดีย มหาวิทยาลัยเทคโนโลยีราชมงคลอีสานจำนวน 30 คน

3.3 วิเคราะห์ข้อมูล

การวิเคราะห์ข้อมูลในงานวิจัยนี้ได้ใช้ Black Box ในการทดสอบระบบและใช้ค่าสถิติร้อยละ ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐานในการนำเสนอระดับความพึงพอใจของผู้ใช้งานต่อระบบบริหารจัดการบทเรียนแสวงรู้นี้ด้วย spss

การหาค่าเฉลี่ย (Mean)

$$\bar{x} = \frac{\sum x}{n}$$

เมื่อ x แทน คะแนนเฉลี่ย

$\sum x$ แทน ผลรวมของคะแนนทั้งหมด ($\sum$ คือซิกมา)

n แทน จำนวนข้อมูล

การหาค่าเบี่ยงเบนมาตรฐาน (Standard Deviation : S.D.)

$$S.D. = \sqrt{\frac{n\sum x^2 - (\sum x)^2}{n(n-1)}}$$

เมื่อ S.D. คือ ค่าเบี่ยงเบนมาตรฐาน

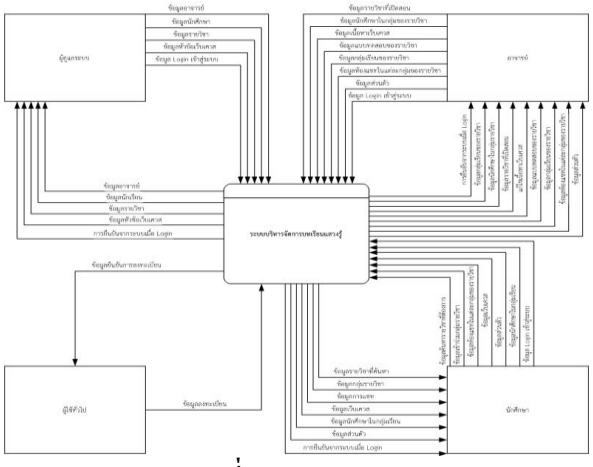
x คือ ข้อมูลแต่ละจำนวน

n คือ จำนวนข้อมูลทั้งหมด

4. การดำเนินการศึกษาและการพัฒนาระบบ

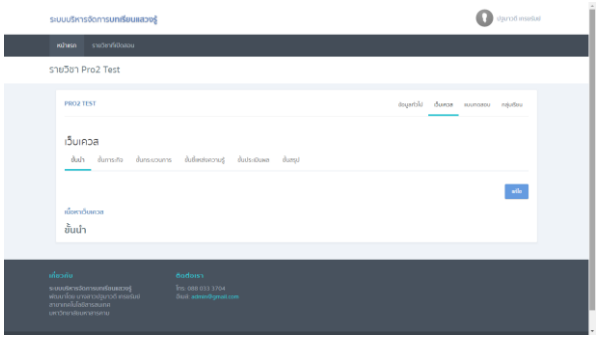
4.1 วิเคราะห์ระบบ

ในงานวิจัยนี้ได้ทำการวิเคราะห์ระบบบริหารจัดการบทเรียนแสวงรู้ได้ดังภาพที่ 2

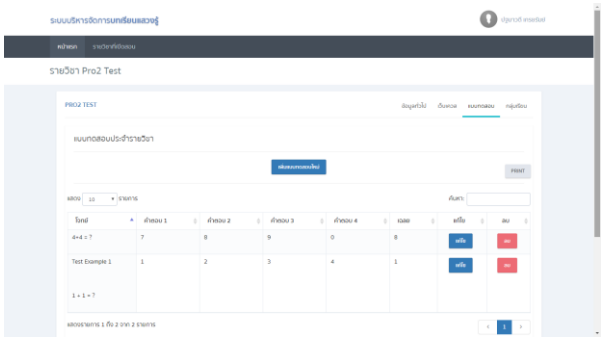


ภาพที่ 2 Context Diagram

จากภาพที่ 2 ระบบบริหารจัดการบทเรียนแสวงรู้ได้แบ่งผู้ใช้ระบบออกเป็น 3 ส่วน คือ ผู้ดูแลระบบ อาจารย์ และนักศึกษา ซึ่งผู้ดูแลระบบสามารถจัดการเพิ่ม ลบ แก้ไขจัดการระบบบริหารจัดการบทเรียนแสวงรู้ได้ ส่วนของอาจารย์สามารถเพิ่ม ลบ แก้ไขเนื้อหารายวิชาต่างๆ กลุ่มผู้เรียนและจำนวนของผู้เรียนได้ และส่วนของนักศึกษาสามารถเข้ามาทำกิจกรรมตามที่ได้กำหนดไว้ได้อย่างครบถ้วนทุกขั้นตอน



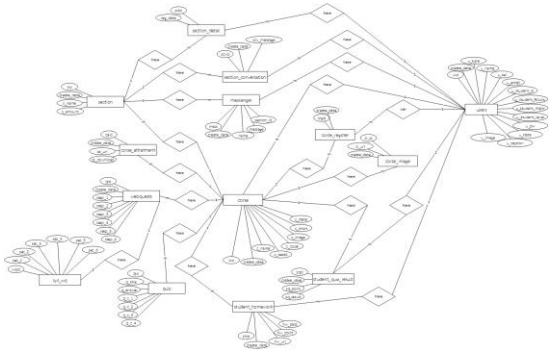
ภาพที่ 3 หน้าจอการจัดการเว็บเควส 6 ขั้นตอน ในส่วนของอาจารย์



ภาพที่ 4 หน้าจอการจัดการแบบทดสอบ ในส่วนของอาจารย์

4.2 วิเคราะห์ฐานข้อมูล

ในงานวิจัยนี้ได้ทำการวิเคราะห์ฐานข้อมูลระบบบริหารจัดการบทเรียนแสวงรู้ได้ดังภาพที่ 5



ภาพที่ 5 ER-Diagram

จากภาพที่ 5 เป็นการออกแบบฐานข้อมูลของระบบบริหารจัดการบทเรียนแสวงรู้

4.3 ประเมินความพึงพอใจ

ในงานวิจัยนี้ได้แบ่งการประเมินความพึงพอใจออกเป็น 4 ด้าน คือ

1. ด้านทดสอบความต้องการ (Functional Requirement Test)
2. ด้านทดสอบฟังก์ชันการทำงาน (Functional Test)
3. ด้านทดสอบการใช้งาน (Usability Test)
4. ด้านทดสอบความปลอดภัย (Security Test)

หลังใช้งานระบบบริหารจัดการบทเรียนแสวงรู้แล้วได้มีการประเมินระบบโดยใช้แบบสอบถามความพึงพอใจที่มีต่อระบบ พบว่า ผู้ใช้งานระบบมีความพึงพอใจต่อระบบในระดับมากดังตารางที่ 1

ตารางที่ 1 ผลประเมินโดยผู้เชี่ยวชาญ

ลำดับ	หัวข้อที่ประเมิน	ค่าเฉลี่ย	ส่วนเบี่ยงเบนมาตรฐาน
1	Functional Requirement Test	4.53	0.58
2	Functional Test	4.60	0.58
3	Usability Test	4.60	0.58
4	Security Test	4.87	0.12
รวมทุกหัวข้อ		4.65	0.47

ผลสัมฤทธิ์ของผู้เรียนในการเรียนผ่านบทเรียนแสวงรู้โดยใช้แบบทดสอบหลังเรียนจำนวน 50 ข้อ ให้แก่นักศึกษาในรายวิชาเทคโนโลยีสารสนเทศเพื่องานสื่อสารมวลชน จำนวน 30 คนได้ผลดังนี้ ($\overline{X}$ = 40.93 S.D. = 3.28)

5. สรุปผลการวิจัย

ผลการวิจัยพบว่า การพัฒนาระบบบริหารจัดการบทเรียนแสวงรู้กรณีศึกษาวิชาเทคโนโลยีสารสนเทศเพื่องานสื่อสารมวลชน สาขาวิชาเทคโนโลยีมัลติมีเดีย มหาวิทยาลัยเทคโนโลยีราชมงคลอีสาน เป็นที่ยอมรับของผู้ใช้ระบบ ผลการประเมินความพึงพอใจต่อระบบด้านทดสอบความปลอดภัย (Security Test) มีความพึงพอใจในระดับมากที่สุด รองลงมาได้แก่ด้านความต้องการของผู้ใช้ระบบ (Functional Requirement Test) ด้านการทำงานตามฟังก์ชันของระบบ (Functional Test) มีความพึงพอใจในระดับมาก ด้านทดสอบการใช้งาน (Usability Test) มีความพึงพอใจในระดับมาก) มีความพึงพอใจในระดับมากตามลำดับ ซึ่งผลของความพึงพอใจที่มีต่อระบบในภาพรวมอยู่ในระดับมาก แสดงให้เห็นว่าระบบที่พัฒนาขึ้นมามีคุณภาพสามารถนำไปใช้งานได้

จากการพัฒนาระบบบริหารจัดการบทเรียนแสวงรู้ระบบนี้ช่วยให้เกิดความสะดวกสบายในการนำมาเป็นสื่อการเรียนการสอน และสะดวกในการติดต่อสื่อสารในการทำกิจกรรมการเรียนระหว่างผู้สอนและผู้เรียน สามารถรวบรวมแหล่งความรู้ที่มีประโยชน์ต่อการเรียนเพื่อให้ผู้เรียนนำไปประยุกต์ใช้ได้จริง

เอกสารอ้างอิง

[1] ศ. ทองขาว และคณะ, คอมพิวเตอร์และเทคโนโลยีสารสนเทศ *Computing essentials 2013*. กรุงเทพฯ: สำนักพิมพ์แมคกรอ-ฮิล, 2556.

[2] ว. อติศัพท์, "Web Quest : การเรียนรู้ที่เน้นผู้เรียนเป็นศูนย์กลางบน World Wide Web," วารสารวิทยบริการ มหาวิทยาลัยสงขลานครินทร์, vol. 4, pp. 52-61, พ.ค.-ส.ค. 2546.

[3] ส. ปรางสร and อ. แพทย์คชา, "กิจกรรมการเรียนด้วยเว็บเควสท์กับทักษะการคิดขั้นสูง," วารสารนักบริหาร, vol. 33, pp. 82-92, กรกฎาคม-กันยายน 2556.

[4] ค. ชาญศิลป์, คู่มือเรียน *WEB Programming* ด้วย *PHP MySQL* และ *AJAX*. โปรวิชั่น: กรุงเทพฯ, 2553.

[5] อ. รัตนถิรกุล, สร้างระบบ *e-Learning* ด้วย *moodle* ฉบับสมบูรณ์.

[6] อำไพ วรรณสินธุ์. (2553, สืบค้น 10 ธันวาคม 2553). เอกสารประกอบการสอน รหัสวิชา *4122502* การวิเคราะห์และออกแบบระบบ *1* Available: สืบค้นจาก: http://www.bcoms.net/system_analysis/index.aspoption=com_rokdownloads&view=file&Itemid=59&id=17:

[7] พรภัทรา ภัทรจารี. (2553, 15 กรกฎาคม). เอกสารประกอบรายวิชา *4153501* วิศวกรรมซอฟต์แวร์ เรื่องกลวิธีการทดสอบซอฟต์แวร์. Available: http://rbu.rbru.ac.th/sathit/4153501.htm.

[8] ชาคริต อนันตวัฒนวงศ์, "ผลของการใช้บทเรียนออนไลน์แบบเว็บเควสท์ ต่อผลสัมฤทธิ์ทางการเรียนและปฏิสัมพันธ์ในการเรียนวิชาการถ่ายภาพทางการศึกษา," สถาบันเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2549.

[9] สุภาวดี บุญแทน, "การสอนแบบการใช้ปัญหาเป็นฐานเว็บเควสท์เรื่องการสื่อสารข้อมูลและเครือข่ายคอมพิวเตอร์ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่2," การประชุมหาดใหญ่วิชาการ, vol. 4, pp. 418-425, 10 พฤษภาคม 2556 2556.

[10] E. Segers and L. Verhoeven, "Learning in a sheltered Internet environment: The use of WebQuests," *Learning and Instruction,* vol. 19, pp. 423-432, 2009.

[11] L. Averkieva, Y. Chayka, and S. Glushkov, "Web Quest as a Tool for Increasing Students' Motivation and Critical Thinking Development," *Procedia-Social and Behavioral Sciences,* vol. 206, pp. 137-140, 2015.

[12] D. V. Bernrdo, "Comparative," *International,* vol. 2, pp. 55-61, 1012.

โยคะเพื่อสุขภาพที่ดี : แอปพลิเคชันบนไอโอเอสสำหรับการฝึกโยคะส่วนบุคคล
Yoga for Good Health : An iOS Application for Personal Yoga Training

สิริภา จันทะบูลย์ (Siripa Jantabon)[1] และ วรสิทธิ์ ชูชัยวัฒนา (Worasit Choochaiwattana)[2]
[1]สาขาวิชาการออกแบบดิจิทัลเชิงสร้างสรรค์
[2]สาขาวิชาวิศวกรรมเว็บและการพัฒนาแอปพลิเคชันบนอุปกรณ์พกพา
วิทยาลัยครีเอทีฟดีไซน์ แอนด์ เอ็นเตอร์เทนเมนต์เทคโนโลยี มหาวิทยาลัยธุรกิจบัณฑิตย์
[1]siripa.jan@dpu.ac.th, [2]worasit.cha@dpu.ac.th

บทคัดย่อ

บทความนี้มีจุดมุ่งหมายเพื่อพัฒนาแอปพลิเคชันบนไอโอเอส สำหรับการฝึกโยคะส่วนบุคคล ส่งเสริมให้เกิดวิธีการที่น่าสนใจและมีความสะดวกสบายสำหรับบุคคลทั่วไปที่ใส่ใจสุขภาพ ใช้ในการบริหารกายโยคะเพื่อสุขภาพเมื่อพวกเขาต้องการ ในการประเมินประสิทธิผลของแอปพลิเคชัน ผู้ประเมินจำนวน 60 ท่าน ได้เข้ามามีส่วนร่วมโดยมีผู้มีประสบการณ์การเล่นโยคะจำนวน 30 คน จากผลการทดลองพบว่าผู้ใช้สามารถใช้บริหารกายโยคะ และสนใจการออกกำลังกายเพื่อสุขภาพมากยิ่งขึ้น ก่อให้เกิดประโยชน์เมื่อใช้แอปพลิเคชันอย่างเป็นประจำ

คำสำคัญ: แอปพลิเคชันเพื่อสุขภาพ แอปพลิเคชันบนอุปกรณ์พกพา การฝึกโยคะ

Abstract

This paper aims at developing an iOS Application for personal Yoga training. It provides people, who are health conscious, an interesting and convenient way to do Yoga whenever they want. To evaluate a performance of the application, 60 participants were recruited. Half of the participants are experienced Yoga lover. The evaluation result showed that the application could be used for personal Yoga training. It persuades its users to start exercising for their good health.

Keywords: Application for Health, Mobile Application, Yoga.

1. บทนำ

ปัจจุบันคนไทยพบปัญหาภาวะน้ำหนักเกิน และเป็นโรคอ้วนมากขึ้นซึ่งส่งผลกระทบเป็นวงกว้าง[1] ทั้งทางด้านสาธารณสุข สังคม และเศรษฐกิจ โดยเป็นปัจจัยเสี่ยงในการเกิดโรคเรื้อรังต่าง ๆ กรมอนามัยกระทรวงสาธารณสุข ระบุผลการสำรวจพบว่าคนไทยมีปัญหาภาวะโภชนาการเกินและกลายเป็นโรคอ้วนเพิ่มสูงสุดในรอบ 10 ปี โดยเด็กแรกเกิด – 12 ปี อ้วนลงพุงร้อยละ 40 รองลงมา คือ อายุช่วง 40 – 50 ปี อ้วนเพิ่มขึ้นเกือบ 2 เท่า ผลสำรวจล่าสุดมีคนไทยอายุ 15 ปีขึ้นไป โรคอ้วนติดอันดับ 5 ของเอเชียแปซิฟิก โดยมีคนอ้วนมากถึง 17 ล้านคนทั่วประเทศและพบว่ากลุ่มที่มีปัญหามากที่สุดอยู่ในวัยทำงานและเป็นผู้หญิงร้อยละ 48

การออกกำลังกายจึงเป็นกิจกรรมที่กระทำแล้วทำให้ร่างกายมีสุขภาพที่ดี แข็งแรงและลดความเสี่ยงต่อโรคเรื้อรัง แต่การออกกำลังกายเป็นการแข่งขันกับตัวเอง หลายคนก่อนจะออกกำลังกายมักอ้างเหตุผลของการไม่ออกกำลังกาย เช่น ไม่มีเวลาหรือไม่มีสถานที่ในการออกกำลังกาย ในขณะที่โยคะเป็นกิจกรรมการออกกำลังกายหรือกิจกรรมเพื่อบริหารกายและจิตที่ได้รับความนิยมในปัจจุบันซึ่งมีท่าของการบริหารร่างกายทุกสัดส่วนเข้าถึงสรีระได้เป็นอย่างดี อีกทั้งยังส่งผลต่อสภาพจิตใจทำให้มีสมาธิและสุขภาพดีขึ้น

ดังนั้น ผู้วิจัยได้เล็งเห็นถึงผลประโยชน์จากการฝึกโยคะจึงได้ทำการพัฒนาแอปพลิเคชันบนอุปกรณ์เคลื่อนที่ (Mobile Application) เพื่อสุขภาพที่ช่วยอำนวยความสะดวกต่อการใช้ชีวิตประจำวันสำหรับผู้ที่สนใจการบริหารกายโยคะได้ฝึกการเคลื่อนไหวซึ่งมีท่ากายบริหารในรูปแบบภาพเคลื่อนไหวให้ผู้ฝึกสามารถทำตามท่าทางได้ และคำนวณปริมาณการเผาผลาญ

พลังงาน หลังจากบริหารร่างกายตลอดจนแสดงรายละเอียดแต่ละท่ากายบริหารให้ผู้ใช้งานได้ทราบ

ในส่วนที่เหลือของบทความนี้ประกอบไปด้วย 2. วรรณกรรมที่เกี่ยวข้อง 3. วิธีการดำเนินการวิจัย 4. ผลการดำเนินงานการวิจัย และ 5. สรุปผลงานวิจัย

2. วรรณกรรมที่เกี่ยวข้อง

2.1 ทฤษฎีที่เกี่ยวข้อง

สถาบันโยคะวิชาการมูลนิธิหมอชาวบ้าน [2] ได้ให้คำจำกัดความไว้ว่า *"โยคะ หมายถึง ศาสตร์ที่ว่าด้วยการฝึกฝนตนเองอันมีรากฐานมาจากอินเดียโบราณ เป้าหมายของโยคะ คือ พัฒนามนุษย์ในทุกมิติ เช่น กาย จิต บุคลิกภาพ ฯลฯ อย่างเป็นองค์รวม การสู่แก่นของโยคะนั้น ประกอบด้วยเทคนิคอันหลากหลาย โดยมีลักษณะร่วมของเทคนิคโยคะทั้งหมดล้วนเป็นเรื่อง กาย – จิต สัมพันธ์ แม้แต่ละเทคนิคจะมีกระบวนการช่วงท่าทางต่างกัน แต่ละเทคนิคก็ล้วนสนับสนุนซึ่งกันและกัน"*

โดยปกติแล้วการวัดอัตราพลังงานหน่วยเป็นแคลอรี่ โดยผู้ชายต้องการแคลอรี่เฉลี่ยต่อวันอยู่ที่ 1,800 – 2,500 กิโลแคลอรี่ ในขณะที่ผู้หญิงต้องการ 1,500 – 2,000 กิโลแคลอรี ต่อวัน [3]

2.2 วิจัยและแอปพลิเคชันที่เกี่ยวข้อง

งานวิจัยแนวโน้มธุรกิจในกลุ่มสมาร์ทโฟน โทรศัพท์มือถือเป็นปัจจัยหนึ่งที่มีความสำคัญอย่างมากในการติดต่อ สื่อสาร ทั้งในกลุ่มประชาชนทั่วไป นักเรียน นักศึกษา และโดยอย่างยิ่งกับกลุ่มนักธุรกิจที่ต้องมีการติดต่อสื่อสารเพื่อสร้างความสะดวกรวดเร็ว การใช้โทรศัพท์มือถือในประเทศไทยจึงได้รับความนิยมเพิ่มขึ้น การเติบโตของโทรศัพท์มือถือในประเทศไทยมีอัตราการเติบโตที่รวดเร็วมากอย่างก้าวกระโดดในปัจจุบัน ความสามารถของโทรศัพท์มือถือ มิได้หยุดอยู่ที่การสนทนาโต้ตอบกันเท่านั้น แต่ยังสามารถทำหน้าที่เหมือนคอมพิวเตอร์ส่วนตัวได้ไม่ว่าการรับ – ส่งอีเมล์ , การท่องโลกอินเทอร์เน็ต, ปฏิทิน การดูสื่อออนไลน์ เป็นต้น [4] [5] ซึ่งความสามารถพิเศษของการทำงานต่าง ๆ เหล่านี้ได้ถูกพัฒนาไว้ในโทรศัพท์มือถือที่เรียกว่า "สมาร์ทโฟน" ที่สามารถตอบโจทย์ไลฟ์สไตล์ของคนยุคใหม่มากขึ้น และเพื่อให้เห็นความแตกต่างของงานที่นำเสนอในบทความนี้ ในส่วนนี้ขอนำเสนอความแตกต่างของแอปพลิเคชันที่เกี่ยวข้องกับการบริหารกายโยคะต่าง ๆ ที่มีคุณสมบัติและฟังก์ชันการทำงานของระบบเมื่อเปรียบเทียบกับแอปพลิเคชันที่ผู้วิจัยพัฒนา

ตารางที่ 1: แสดงชื่อแอปพลิเคชันที่เกี่ยวข้อง คุณสมบัติ และฟังก์ชันการทำงานของระบบ

คุณสมบัติฟังก์ชัน	รายชื่อแอปพลิเคชัน				
	โยคะแบบง่าย	All in Yoga	Authentic Yoga With Deepak Chopra And Tara Stiles	Yoga Break	แอปพลิเคชันโดยผู้วิจัย
ภาพประกอบภาพโยคะ	✓	✓	✓	✓	✓
การคำนวณการเผาผลาญพลังงาน	-	-	-	-	✓
ภาพเคลื่อนไหวท่าโยคะ	✓	✓	✓	✓	✓
ระบบการบันทึกรายการโยคะ	-	✓	-	-	✓

ตารางที่ 1 แสดงให้เห็นว่าแอปพลิเคชันโดยผู้วิจัยได้มีการเพิ่มคุณสมบัติ และฟังก์ชันการทำงานของระบบในส่วนของการคำนวณการเผาผลาญพลังงาน และระบบการบันทึกรายการโยคะ สำหรับการการคำนวณการเผาผลาญพลังงานนั้น ในงานวิจัยชิ้นนี้ ได้นำเอาข้อมูลค่าเฉลี่ยของการเผาผลาญมาจาก [6] ซึ่งนำมาคำนวณกับเวลาในการเล่นโยคะของท่าต่าง ๆ

อย่างไรก็ตามจากการศึกษางานวิจัย พบว่ายังไม่มีการจำลองสถานการณ์ เพื่อให้ทราบถึงประสิทธิภาพในการทำงานของวิธีที่เสนอได้ ทำให้ไม่สามารถทราบถึงคุณภาพที่จะนำมาใช้งาน ดังนั้น ในบทความนี้จึงได้มีการประเมินผลของระบบโดยพิจารณาจากความพึงพอใจของกลุ่มตัวอย่างที่ทดลองใช้งานระบบ เพื่อนำมาวิเคราะห์และแปรผลการทดสอบระบบ ผู้ทดลองได้ทำการตอบแบบสอบถามที่ใช้ในการประเมินความเหมาะสม และความพึงพอใจของระบบ Yoga for Health

3. วิธีการดำเนินการวิจัย

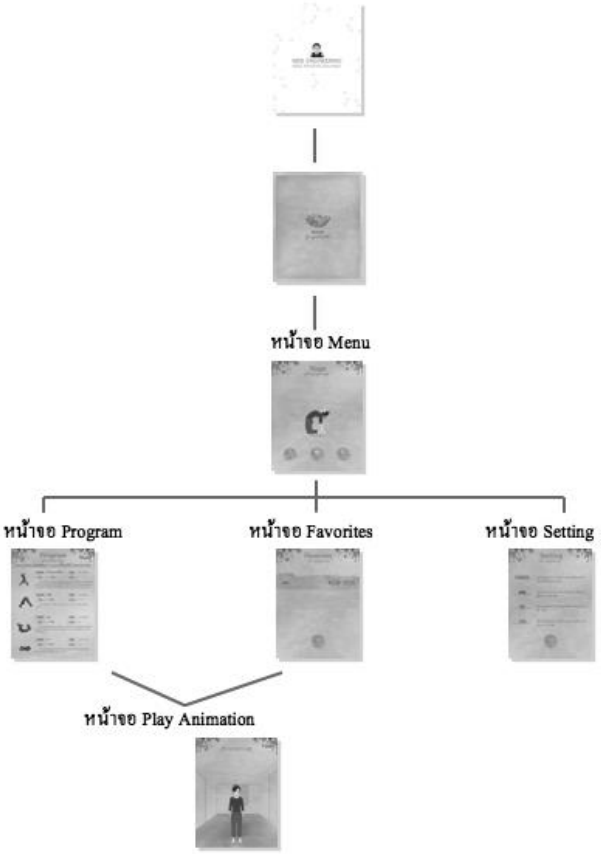
3.1 การวิเคราะห์ปัญหาและศึกษาข้อมูล

ผู้วิจัยได้ทำการศึกษาปัญหาและอุปสรรคในการพัฒนาและรวบรวมข้อมูล ทฤษฎีที่เกี่ยวข้อง ข้อมูลการออกแบบแอปพลิเคชัน และศึกษาเครื่องมือที่ใช้ในการพัฒนาแต่ละส่วน ได้แก่ การศึกษาข้อมูลการใช้งานระบบปฏิบัติการ Mac OS, การใช้เครื่องมือการพัฒนา Xcode SDK, การดำเนินการพัฒนา iOS Application, การออกแบบ Graphic, วิเคราะห์ข้อมูลที่ใช้ในการแก้ปัญหา และวิเคราะห์ข้อมูลต่าง ๆ รวมถึงทางเลือกหาช่องทางในการพัฒนาแอปพลิเคชันต่อไป

การเก็บรวบรวมข้อมูลรูปแบบการบริหารร่างกายโยคะในส่วนการเคลื่อนไหวร่างกายตามส่วนต่าง ๆ ได้ทำการรวบรวมข้อมูลจากหนังสือ อินเทอร์เน็ต กลุ่มเป้าหมายที่ต้องการฝึกโยคะ รวมถึงอาจารย์ผู้ฝึกสอนโยคะ และข้อมูลแสดงการเผาผลาญพลังงานร่างกาย (Kcal)

3.2 การวิเคราะห์ออกแบบการใช้งาน

เริ่มต้นจากออกแบบโครงสร้างการทำงานของระบบ (System Architecture) เพื่อให้เห็นโครงสร้างการทำงานของแอปพลิเคชันในภาพรวมทั้งหมด ดังแสดงรายละเอียดไว้ในภาพที่ 1



ภาพที่ 1: แสดงถึงการออกแบบโครงสร้างการทำงานของระบบ Yoga for Good Health

ภาพที่ 1: แสดงให้เห็นถึงการออกแบบรายละเอียดโครงสร้างของระบบแอปพลิเคชัน Yoga for Good Health ซึ่งประกอบไปด้วยหน้าจอการทำงานต่าง ๆ

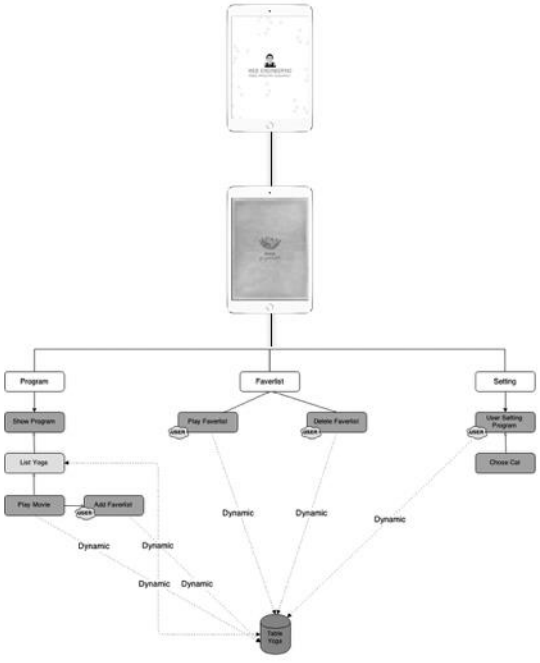
3.3 เครื่องมือที่ใช้ในการพัฒนาระบบ

ผู้วิจัยได้มีการใช้เครื่องมือและเทคโนโลยีในการพัฒนาระบบ ได้แก่ การใช้ Xcode [7] เป็นเครื่องมือพัฒนาแอปพลิเคชันบนระบบปฏิบัติการ MacOS เก็บข้อมูลที่ SQLite Database ในส่วนการออกแบบได้ทำการใช้โปรแกรม Photoshop, Illustrator และ Flash ในการสร้างสื่อกราฟิก 2 มิติ และสื่อ Animation

4. ผลการดำเนินงาน

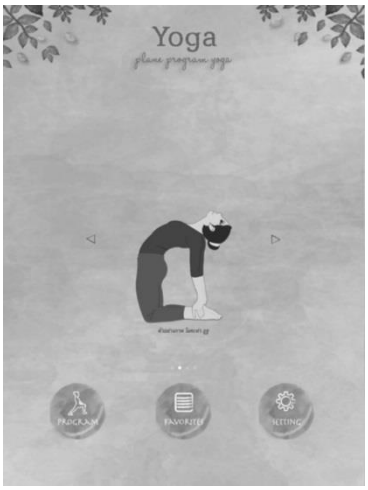
4.1 ผลการออกแบบและพัฒนาระบบ

ผู้วิจัยได้มีการออกแบบและพัฒนาระบบเพื่อใช้อธิบายโครงสร้างของการออกแบบและพัฒนาระบบออกเป็นส่วน ๆ อยู่ในรูปแบบของผังที่ใช้แสดงโครงสร้าง เช่น แผนผังโครงสร้างรูปแบบการใช้งาน แผนผังโครงสร้างระบบฐานข้อมูล แผนผังโครงสร้างสารสนเทศในการใช้งานระบบ แผนผังโครงสร้างการทำงานของระบบที่พร้อมใช้งานและผลการประเมินความพึงพอใจในการใช้งานระบบ และนำผลจากการประเมินมาวิเคราะห์เพื่อหาข้อสรุปความพึงพอใจในการใช้งานระบบ



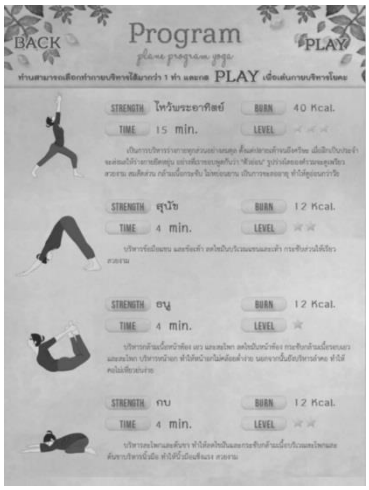
ภาพที่ 2: Information Design : Blueprint ของระบบ Yoga for Good Health

จากภาพที่ 2 แสดงถึงการออกแบบแผนผังโครงสร้างสารสนเทศในการใช้งานระบบ Yoga for Good Health มีการออกแบบการทำงานและการเข้าสู่ข้อมูลของระบบซึ่งประกอบไปด้วยหน้าจอการทำงานต่าง ๆ ได้แก่ หน้าจอแรกและหน้าจอที่สองในการใช้งาน, หน้าจอ Program , หน้าจอ Favorites และหน้าจอ Setting



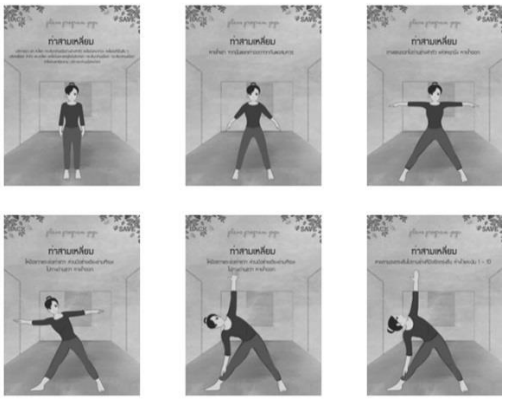
ภาพที่ 3: แสดงหน้าจอ Menu คือ หน้าจอที่แสดงเมนูทั้งหมดในแอปพลิเคชัน

จากภาพที่ 3 แสดงตัวอย่างท่ากายบริหารโยคะและแสดง Menu ทั้งหมด 3 เมนู ได้แก่ Program, Favorites และ Setting



ภาพที่ 4: แสดงหน้าจอ Program คือ หน้าจอที่แสดงผลรายการท่ากายบริหารโยคะ

จากภาพที่ 4 เมื่อผู้ใช้งานทำการเลือกเมนู Program หน้าจอจะแสดงผลรายการท่ากายบริหารโยคะ โดยแสดงรูปภาพประกอบ, ชื่อกายบริหาร, การเผาผลาญพลังงาน(Kcal), เวลา, ระดับความยากง่าย และคุณประโยชน์ของท่ากายบริหารแต่ละท่า ซึ่งผู้ใช้สามารถเลือกท่ากายบริหารได้มากกว่า 1 ท่ากายบริหารและทำการกด Play เพื่อทำการเข้าสู่ Play Animation



ภาพที่ 5: แสดงหน้าจอ Play Animation

จากภาพที่ 5 แสดงหน้าจอ Play Animation เป็นหน้าจอที่แสดงภาพเคลื่อนไหวของแต่ละท่ากายบริหารที่ผู้ใช้ได้ทำการเลือกจากหน้า Program เมื่อทำกายบริหารจบตาม List ที่เลือกจะมีข้อความแจ้งเตือนปริมาณการเผาผลาญพลังงาน(Kcal)

4.2 ผลการประเมินความพึงพอใจของระบบ

การประเมินผลระบบพิจารณาจากความพึงพอใจของกลุ่มตัวอย่างที่ทดลองใช้งานระบบเพื่อนำมาวิเคราะห์และแปรผลการทดสอบระบบ โดยใช้แบบสอบถามเพื่อประเมินระดับความเหมาะสม/ความพึงพอใจของระบบ โดยมีผู้ตอบแบบสอบถามจำนวน 60 คน เป็นผู้มีประสบการณ์การเล่นโยคะ 30 คน และผู้ที่ไม่มีประสบการณ์การเล่นโยคะ 30 คน เป็นผู้ทดลองใช้งานทั่วไป เลือกวิธีการสุ่มตัวอย่างโดยไม่ตั้งใจ (Accidental Sampling) และกำหนดเกณฑ์ในการประเมินผลความเหมาะสม/ความพึงพอใจเป็นมาตราส่วนในการประมาณค่าเป็น Rating Scale เป็น 5 ระดับ เกณฑ์การประเมินจะพิจารณาจากคะแนนเฉลี่ยของ ความเหมาะสม/ความพึงพอใจ ซึ่งการวิเคราะห์ข้อมูลใช้สถิติเชิงพรรณนา (Descriptive Statistics) ในการวัดค่าของข้อมูลโดยใช้ค่าเฉลี่ยเลขคณิตหรือค่าเฉลี่ย (Mean) และวัดการกระจายของข้อมูลโดยใช้ค่าเบี่ยงเบนมาตราฐาน (Standard Deviation)

ผลการวิเคราะห์ข้อมูลความเหมาะสม/ความพึงพอใจ Yoga for Good Health : แอปพลิเคชันโยคะเพื่อสุขภาพที่ดี

ตารางที่ 2: แสดงจำนวนความถี่และร้อยละของข้อมูลความเหมาะสม/ความพึงพอใจในข้อมูลรายละเอียดท่าโยคะมีความถูกต้องครบถ้วน และน่าเชื่อถือ

ข้อมูลรายละเอียดท่าโยคะ มีความถูกต้อง ครบถ้วนน่าเชื่อถือ	คนเล่นโยคะ		คนไม่เล่นโยคะ	
	คน	ร้อยละ	คน	ร้อยละ
น้อยที่สุด	0	0	0	0
น้อย	0	0	0	0
ปานกลาง	2	6.6	4	3.13
มาก	16	3.53	16	3.53
มากที่สุด	12	40	10	3.33
รวม	**30**	**100**	**30**	**100**

จากตารางที่ 2 พบว่าผู้ตอบแบบสอบถามที่มีประสบการณ์การเล่นโยคะและผู้ที่ไม่มีประสบการณ์การเล่นโยคะทั้งหมด 60 คน ส่วนใหญ่มีความพึงพอใจต่อข้อมูลรายละเอียดท่าโยคะ มีความถูกต้อง ครบถ้วน น่าเชื่อถือในระดับมาก คิดเป็นร้อยละ 53.3

ตารางที่ 3: แสดงจำนวนความถี่และร้อยละของข้อมูลความเหมาะสม/ความพึงพอใจในการแสดง Animation มีความต่อเนื่องของท่าได้อย่างถูกต้อง

การแสดง Animation มีความต่อเนื่องของท่าได้อย่างถูกต้อง	คนเล่นโยคะ		คนไม่เล่นโยคะ	
	คน	ร้อยละ	คน	ร้อยละ
น้อยที่สุด	0	0	0	0
น้อย	1	3.3	1	3.3
ปานกลาง	2	6.6	3	10
มาก	17	6.56	17	6.56
มากที่สุด	10	3.33	9	30
รวม	**30**	**100**	**30**	**100**

จากตารางที่ 3 พบว่าผู้ตอบแบบสอบถามที่มีประสบการณ์การเล่นโยคะและผู้ที่ไม่มีประสบการณ์การเล่นโยคะทั้งหมด 60 คน ส่วนใหญ่มีความพึงพอใจต่อการแสดง Animation มีความต่อเนื่องของท่าได้อย่างถูกต้องในระดับมาก คิดเป็นร้อยละ 56.6

ตารางที่ 4: แสดงจำนวนความถี่และร้อยละของข้อมูลความเหมาะสม/ความพึงพอใจในการออกแบบความน่าสนใจ ความสร้างสรรค์ของแอปพลิเคชัน

การออกแบบ ความน่าสนใจ ความสร้างสรรค์ของแอปพลิเคชัน	คนเล่นโยคะ		คนไม่เล่นโยคะ	
	คน	ร้อยละ	คน	ร้อยละ
น้อยที่สุด	0	0	0	0
น้อย	0	0	1	3.3
ปานกลาง	2	6.6	6	20
มาก	13	3.43	15	50
มากที่สุด	15	50	8	6.26
รวม	**30**	**100**	**30**	**100**

จากตารางที่ 4 พบว่าผู้ตอบแบบสอบถามที่มีประสบการณ์การเล่นโยคะจำนวนทั้งสิ้น 30 คน ส่วนใหญ่มีความพึงพอใจในการอ อกแบบความน่าสนใจ ความสร้างสรรค์ของแอปพลิเคชันในระดับมากที่สุด คิดเป็นร้อยละ 50

และพบว่าผู้ตอบแบบสอบถามที่ไม่มีประสบการณ์การเล่นโยคะทั้งหมด 30 คน ส่วนใหญ่มีความพึงพอใจในการออกแบบความน่าสนใจ ความสร้างสรรค์ของแอปพลิเคชันระดับมาก คิดเป็นร้อยละ 50

ตารางที่ 5: ความเหมาะสม/ความพึงพอใจทางด้านข้อมูล จำนวน 4 ข้อคำถาม สำหรับผู้ที่มีประสบการณ์เล่นโยคะ

ความพึงพอใจด้านข้อมูล	$\bar{X}$	S.D.	ระดับความเหมาะสม/ความพึงพอใจ
ข้อมูลรายละเอียดท่าโยคะ มีความถูกต้องครบถ้วน และน่าเชื่อถือ	33.4	61.0	มาก
การจัดหมวดหมู่การใช้งานของข้อมูลมีความถูกต้อง ชัดเจน	33.4	76.0	มาก
สามารถนำข้อมูลท่าโยคะจากแอปพลิเคชันไปใช้เป็นประโยชน์ได้	20.4	71.0	มาก
การแสดงผลท่าเคลื่อนไหว Animation มีความถูกต้องสอดคล้องกับรายการท่าโยคะที่ผู้ใช้ได้ทำการเลือก	13.4	78.0	มาก
เฉลี่ยรวม	**25.4**	**71.0**	**มาก**

จากตารางที่ 5 แสดงความเหมาะสม/ความพึงพอใจด้านข้อมูล จำนวน 4 ข้อคำถาม ในความคิดเห็นของผู้ตอบแบบสอบถามที่มีประสบการณ์การเล่นโยคะ พบว่าให้ความพึงพอใจโดยรวมอยู่ในระดับมาก โดยมีความเฉลี่ยอยู่ที่ 4.25

ตารางที่ 6: ความเหมาะสม/ความพึงพอใจทางด้านข้อมูล จำนวน 4 ข้อคำถาม สำหรับผู้ที่ไม่มีประสบการณ์การเล่นโยคะ

ความพึงพอใจด้านข้อมูล	$\bar{X}$	S.D.	ระดับความเหมาะสม/ความพึงพอใจ
ข้อมูลรายละเอียดท่าโยคะ มีความถูกต้องครบถ้วน และน่าเชื่อถือ	20.4	66.0	มาก
การจัดหมวดหมู่การใช้งานของข้อมูลมีความถูกต้อง ชัดเจน	03.4	72.0	มาก
สามารถนำข้อมูลท่าโยคะจากแอปพลิเคชันไปใช้เป็นประโยชน์ได้	57.4	57.0	มากที่สุด
การแสดงผลท่าเคลื่อนไหว Animation มีความถูกต้องสอดคล้องกับรายการท่าโยคะที่ผู้ใช้ได้ทำการเลือก	27.4	78.0	มาก
เฉลี่ยรวม	**27.4**	**68.0**	**มาก**

จากตารางที่ 6 แสดงความเหมาะสม/ความพึงพอใจด้านข้อมูล จำนวน 4 ข้อคำถาม ในความคิดเห็นของผู้ตอบแบบสอบถามที่ไม่มีประสบการณ์การเล่นโยคะ พบว่าให้ความพึงพอใจโดยรวมอยู่ในระดับมาก โดยมีความเฉลี่ยอยู่ที่ 4.27

จากการประเมินความเหมาะสม/ความพึงพอใจต่อการใช้งานระบบทั้ง 3 ด้านได้สรุปเป็น 2 กลุ่ม คือ ผู้ที่มีประสบการณ์การเล่นโยคะ จำนวน 30 คน และ ผู้ที่ไม่มีประสบการณ์การเล่นโยคะ จำนวน 30 คน รวมทั้งสิ้น 60 คน สามารถสรุปได้ดังนี้

ผู้ที่มีประสบการณ์การเล่นโยคะมีความคิดเห็นต่อ Yoga for Good Health: แอปพลิเคชันสำหรับการฝึกโยคะส่วนบุคคล ในด้านประสิทธิภาพให้ความพึงพอใจอยู่ในระดับมาก โดยมีค่าเฉลี่ยอยู่ที่ 4.19 จากคะแนนเต็ม 5 และมีค่าเบี่ยงเบนมาตรฐานเฉลี่ยอยู่ที่ 0.71 ด้านข้อมูลให้ความพึงพอใจอยู่ในระดับมาก โดยมีค่าเฉลี่ยอยู่ที่ 4.25 จากคะแนนเต็ม 5 และมีค่าเบี่ยงเบนมาตรฐาน เฉลี่ยอยู่ที่ 0.70 และด้านรูปแบบให้ความพึงพอใจในระดับมาก โดยมีค่าเฉลี่ยอยู่ที่ 4.37 จากคะแนนเต็ม 5 และมีค่าเบี่ยงเบนมาตรฐาน เฉลี่ยอยู่ที่ 0.67 รวมทั้ง 3 ด้าน หลังผู้ทดลองใช้งานมีความพึงพอใจต่อการใช้งานระบบทั้ง 3 ด้านอยู่ในระดับมาก มีค่าเฉลี่ยอยู่ที่ 4.27 สูงกว่าคะแนนเฉลี่ยที่ตั้งไว้ 3.5 คะแนนขึ้นไป ส่วนค่าเบี่ยงเบนมาตรฐาน มีค่าเฉลี่ยอยู่ที่ 0.70

ผู้ที่ไม่มีประสบการณ์การเล่นโยคะมีความคิดเห็นต่อ Yoga for Good Health: แอปพลิเคชันสำหรับการฝึกโยคะส่วนบุคคล ในด้านประสิทธิภาพให้ความพึงพอใจอยู่ในระดับมาก โดยมีค่าเฉลี่ยอยู่ที่ 4.24 จากคะแนนเต็ม 5 และมีค่าเบี่ยงเบนมาตรฐานเฉลี่ยอยู่ที่ 0.69 ด้านข้อมูลให้ความพึงพอใจอยู่ในระดับมาก โดยมีค่าเฉลี่ยอยู่ที่ 4.27 จากคะแนนเต็ม 5 และมีค่าเบี่ยงเบนมาตรฐาน เฉลี่ยอยู่ที่ 0.68 และด้านรูปแบบให้ความพึงพอใจในระดับมาก โดยมีค่าเฉลี่ยอยู่ที่ 4.15 จากคะแนนเต็ม 5 และมีค่าเบี่ยงเบนมาตรฐาน เฉลี่ยอยู่ที่ 0.83 รวมทั้ง 3 ด้าน หลังผู้ทดลองใช้งานมีความพึงพอใจต่อการใช้งานระบบทั้ง 3 ด้านอยู่ในระดับมาก มีค่าเฉลี่ยอยู่ที่ 4.22 สูงกว่าคะแนนเฉลี่ยที่ตั้งไว้ 3.5 คะแนนขึ้นไป ส่วนค่าเบี่ยงเบนมาตรฐาน มีค่าเฉลี่ยอยู่ที่ 0.74

ผู้ตอบแบบสอบถามทั้งหมด 53 คน ให้ความคิดเห็นว่าจะ Download มาใช้งาน คิดเป็นร้อยละ 88.3 ด้วยเหตุผลเพื่อใช้ในการฝึกเล่นโยคะเบื้องต้น ส่งเสริมต่อสุขภาพกายและสุขภาพจิตสนใจการออกกำลังกายเพื่อสุขภาพมากยิ่งขึ้น โดยผู้ใช้งานสามารถออกกำลังกายโยคะในท่าทางต่าง ๆ รวมถึงทราบผลการเผาผลาญพลังงานหลังออกกำลังกายอีกทั้งยังทราบคุณประโยชน์โยคะแต่ละท่าทาง ส่งผลให้เป็นช่องทางในการศึกษาข้อมูลโยคะมากขึ้น ซึ่งหากแอปพลิเคชัน Yoga for Good Health อยู่ใน App Store ผู้ตอบแบบสอบถามให้เหตุผลในการ Download ได้แก่ แอปพลิเคชันสามารถใช้เป็นแนวทางการศึกษาในการฝึกเล่นโยคะเบื้องต้น และอธิบายถึงคุณประโยชน์ของท่ากายบริหารโยคะที่สามารถปฏิบัติตามได้มากกว่า 1 ท่า และบันทึกรายการ Favorite ได้เมื่อกลับเข้ามาใช้งานแอปพลิเคชันอีกครั้ง รวมถึงแอปพลิเคชันสามารถใช้งานง่าย มีความสวยงาม น่าสนใจ ทั้งนี้สามารถใช้งานในช่วงเวลาและสถานที่ที่สะดวก สรุปได้ว่า Yoga for good Health แอพพลิคชันสำหรับการฝึกโยคะส่วนบุคคลได้บรรลุตามวัตถุประสงค์ที่ตั้งไว้โดยผู้ใช้สามารถใช้บริหารร่างกายโยคะเพื่อสุขภาพ ผู้ใช้งานหันมาสนใจการออกกำลังกายเพื่อสุขภาพยิ่งขึ้น และก่อให้เกิดประโยชน์สูงสุดในการใช้งานในชีวิตประจำวัน

5. สรุป

บทความนี้ได้นำเสนอการพัฒนาแอปพลิเคชันสำหรับการฝึกโยคะส่วนบุคคล (Yoga for Good Health) มีวัตถุประสงค์เพื่อพัฒนาแอปพลิเคชันบนอุปกรณ์เคลื่อนที่ใช้ในการบริหารร่างกายโยคะเพื่อสุขภาพ รวมถึงส่งเสริมให้บุคคลทั่วไปหันมา

เอกสารอ้างอิง

[1] กระทรวงสาธารณสุข, “ผลสำรวจคนไทยมีปัญหาภาวะโภชนาการเกินและกลายเป็นโรคอ้วน” สืบค้น 3 มีนาคม พ.ศ. 2553, จาก www.happymeny.com/Forum, 20-10-255

[2] สถาบันโยคะวิชาการ มูลนิธิหมอชาวบ้าน, “โยคะ การฝึกทำทุกวัน” สืบค้น 1 มีนาคม 2557, จาก http://www.thaiyogainstitute.com/

[3] วิกิพีเดียสารานุกรมเสรี, “แคลอรี่” สืบค้น 1 มีนาคม 2557, จาก http://th.wikipedia.org/wiki/แคลอรี่

[4] บุษรา ประกอบธรรม, “แนวโน้มธุรกิจในกลุ่มสมาร์ทโฟนปี 2010” *วารสารนักบริหาร* ปีที่ 30 ฉบับที่ 2 พ.ศ. 2553 หน้า 197 – 202.

[5] ศูนย์วิจัยกสิกรไทย, “โทรศัพท์เคลื่อนที่ปี 2554 :สมาร์ทโฟนแรงต่อเนื่องรุกสู่ตลาดระดับล่าง..” สืบค้น 5 มกราคม 2556, จาก http://corehoononline.com/images/stories/pdf/3040-p.doc

[6] กระทรข้อมูลแสดงการเผาผลาญพลังงาน” สืบค้น 5 มีนาคม พ.ศ. 2553, จากhttp://www.self.com/calculatorsprograms/calculators/calriesburned/yoga

[7] ขจรศักดิ์ สังข์เจริญ, *การเขียน iPhone Apps สำหรับผู้เริ่มต้น* บริษัทเอเชียเพรส(1989) จำกัด พ.ศ.2555

โมเดลการเรียนรู้แบบยูบิควิตัสบนระบบคลาวด์สำหรับมุ่งสู่การศึกษาไทย 4.0
Ubiquitous Learning Model on Cloud Systems for Thai Education 4.0

ณัตตยา จิตผ่องแท้ (Nattaya Jitpongtae)[1] นิเวศ จิระวิชิตชัย (Nivet Chirawivhitchai)[2] และ
ประสงค์ ปราณีตพลกรัง (Prasong Praneetpolgrang)[3]
หลักสูตรวิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ บัณฑิตศึกษา มหาวิทยาลัยศรีปทุม
[1]nattaya200972@gmail.com, [2]nivet.ch@spu.ac.th, [3]prasong.pr@spu.ac.th

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อ 1) ออกแบบโมเดลการเรียนรู้แบบยูบิควิตัสบนคลาวด์สำหรับมุ่งสู่การศึกษาไทย 4.0 และ 2) ประ เมินความเหมาะสมของโมเดลการเรียนรู้แบบยูบิควิตัสบนคลาวด์ ผู้วิจัยได้ใช้กลุ่มตัวอย่างจากผู้เรียนที่เป็นนักศึกษาระดับประกาศนียบัตรวิชาชีพชั้นสูง สาขาวิชาช่างไฟฟ้ากำลัง จำนวน 30 คน วิธีดำเนินการวิจัย เป็นการ ศึกษาเอกสาร ทฤษฎี และ งานวิจัยที่เกี่ยวข้อง การสัมภาษณ์เชิงลึกผู้เชี่ยวชาญ รวมทั้งออกแบบและพัฒนาระบบการเรียนรู้แบบยูบิควิตัสบนคลาวด์โดยใช้กูเกิ้ลคลาสรูมเป็นเครื่องมือสนับสนุน สถิติที่ใช้ในการวิเคราะห์ข้อมูลเป็นค่าเฉลี่ย และ ค่าส่วนเบี่ยงเบนมาตรฐาน ผลการวิจัยพบว่า การประเมินความเหมาะสมของโมเดลการเรียนรู้แบบยูบิควิตัสโดยผู้เชี่ยวชาญ มีค่าเฉลี่ยความเหมาะสมอยู่ในระดับมาก ($\bar{X}$ = 4.48, S.D = 0.58) ในขณะที่ ผู้เรียนซึ่งเป็นผู้ใช้ระบบ มีการยอมรับ พึงพอใจและเห็นว่าโมเดลการเรียนรู้แบบยูบิควิตัสบนคลาวด์นี้ มีความเหมาะสมต่อนักศึกษาและบริบทของการเรียนรู้สมัยใหม่อยู่ในระดับมาก ($\bar{X}$ = 4.43, S.D.= 0.65) ทั้งนี้โมเดลการเรียนรู้ใหม่ที่นำเสนอจะเป็นตัวเร่งให้มีการเรียนการสอนที่มีประสิทธิภาพและมีความพร้อม อันส่งผลให้การมุ่งสู่การศึกษาไทย 4.0 ได้รวดเร็วมากยิ่งขึ้น

คำสำคัญ : โมเดลการเรียนรู้, การเรียนแบบยูบิควิตัส, ระบบคลาวด์, การศึกษาไทย 4.0

Abstract

The purposes of the study were as follows: 1) to design ubiquitous learning model on cloud systems for Thai education 4.0 and 2) to evaluate the suitability of ubiquitous learning model on cloud. The subjects of the study were 30 diploma students of Kanchanaburi Technical College. This study was used the research methodology by studying research problems, literature review, and in-depth interview with the expert included designing and developing ubiquitous learning model on cloud systems supported by google class room. The collected data was analyzed by a computer program, mean, and standard deviation. The findings were as follows : 1) with evaluating the suitability of ubiquitous learning model on cloud system for Thai education 4.0 by the experts were in high level ($\bar{X}$ =4.48, S.D. = 0.58) and 2) with evaluating suitability of using the ubiquitous learning model on cloud by the students were also in high level ($\bar{X}$ =4.43, S.D.= 0.65) However, the research revealed that the new learning model could accelerated efficient teaching and learning and ready to leverage education system towards Thai education 4.0.

Keywords : Learning Model, Ubiquitous Learning, Cloud Systems, Thai Education 4.0

1. บทนำ

ระบบการเรียนการสอนในทศวรรษที่ผ่านมาครูจะนำเทคโนโลยีและสื่อการเรียนการสอนที่ทันสมัยในขณะหนึ่งๆ นั้นมาใช้เพื่อให้เกิดการเรียนรู้ที่ใหม่และทันสมัยอยู่ตลอดเวลา ครูจะนำอีเลิร์นนิงมาใช้กับการเรียนการสอนเพื่อช่วยให้เกิดปฏิสัมพันธ์ระหว่างผู้สอนและผู้เรียนได้ร่วมเรียนรู้ด้วยกัน เพื่อให้ครบตามข้อกำหนดของหน่วยการเรียนรู้ในแต่ละวิชาตามแผนการเรียนรู้ที่กำหนดไว้และนักเรียนมีผลการเรียนรู้ตามวัตถุประสงค์ที่ตั้งไว้ แต่การใช้อีเลิร์นนิงที่ผ่านมามีข้อจำกัดของเทคโนโลยีที่ไม่สามารถติดตามตรวจสอบรายงานผลการใช้งานได้ทันที ในยุคการศึกษา 4.0 มีเทคโนโลยีรูปแบบใหม่ถูกนำมาใช้สนับสนุนระบบการจัดการเรียนรู้เพื่อตอบสนองความต้องการของผู้ใช้งานที่เป็นผู้เรียนในยุคปัจจุบันที่มีอุปกรณ์รองรับการใช้งานเชื่อมต่อกับเน็คเวิร์คเพื่อสื่อสารกับสังคมภายนอก ดังนั้นจึงนำสมาร์ตโฟนและอุปกรณ์พกพารวมทั้งเดสท็อป มาใช้เพื่อการเรียนรู้ในยุคการศึกษา 4.0 ให้เกิดการเรียนรู้ที่รวดเร็วขึ้นและสามารถเรียนรู้ที่ใดก็ได้ ไม่จำกัดสถานที่ใช้ได้ทุกแห่งที่มีสัญญาณอินเทอร์เน็ต ซึ่งสอดคล้องกับนโยบายรัฐบาลในยุค 4.0 ปัจจุบันมีการให้ใช้ซอฟแวร์ระบบและทรัพยากรของเครื่องคอมพิวเตอร์ฟรีผ่านอินเทอร์เน็ตบนระบบคลาวด์ ทำให้ผู้สอนได้ใช้เป็นช่องทางในการพัฒนาระบบสนับสนุนการจัดการเรียนรู้ได้สะดวก และประหยัดค่าใช้จ่าย อันส่งผลให้การเรียนการสอนได้มีความทันสมัยมากขึ้น ทั้งนี้ผู้สอนและผู้เรียนสามารถเข้าถึงข้อมูลจากที่ใดก็ได้โดยเชื่อมต่อเว็บเบราว์เซอร์ หรือ ไครเอ็นท์แอปพลิเคชันบนสมาร์ตโฟนหรืออุปกรณ์พกพาได้ ผู้เรียนในปัจจุบันมีอุปกรณ์ดังกล่าวจึงสามารถนำอุปกรณ์มาใช้กับระบบการเรียนการสอนในรูปแบบใหม่ด้วยอุปกรณ์เทคโนโลยีที่มีใช้งานจนเกิดความเคยชินแล้ว เพื่อพัฒนาไปสู่การเป็นคนไทย 4.0 ในยุคการศึกษาไทย 4.0 ด้วย

วัตถุประสงค์การวิจัยครั้งนี้ 1) ออกแบบโมเดลการเรียนรู้แบบยูบิควิตัสบนระบบคลาวด์สำหรับมุ่งสู่การศึกษาไทย 4.0 และ 2) ทำการประเมินความเหมาะสมของโมเดลการเรียนรู้แบบยูบิควิตัสบนระบบคลาวด์ ด้วยผู้เชี่ยวชาญระบบผู้เชี่ยวชาญทางด้านการศึกษา และผู้สอนสังกัดสำนักงานคณะกรรมการการอาชีวศึกษา รวมจำนวน 9 ท่าน และประเมินด้วยผู้ใช้งานที่ผู้เรียนจากกลุ่มตัวอย่างจำนวน 30 คน เพื่อหาค่าความเหมาะสมของโมเดลการเรียนรู้ยูบิควิตัสบนระบบคลาวด์

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 ทฤษฎีที่เกี่ยวข้อง

การประมวลผลแบบคลาวด์ : สถาบันมาตรฐานและเทคโนโลยีแห่งชาติ (National Institute of Standards and Technology: NIST) [1] ให้นิยามไว้ว่า เป็นเทคโนโลยีซึ่งสามารถใช้งานร่วมกันได้ทุกที่ ทุกเวลาด้วยอุปกรณ์ต่างๆ ทำให้สามารถเข้าถึงข้อมูลผ่านทางอินเทอร์เน็ต ทั้งนี้เพื่อแบ่งปันทรัพยากรไอทีได้ตามความต้องการ ดังนั้นคลาวด์จึงเปรียบเสมือนแหล่งที่เก็บวัตถุดิบที่เป็นทรัพยากรทางด้านเทคโนโลยีขนาดใหญ่ ที่ทุกคนในโลกนี้สามารถเข้าไปร่วมกันใช้งานได้โดยไม่จำกัดพื้นที่และเวลา โดยนำทรัพยากรที่มีอยู่ในคลาวด์ไปใช้งานและก่อเกิดประโยชน์สูงสุดต่อองค์กรและตัวบุคคลเอง ตามแนวคิดของ P. Lefrere [2] ที่ประยุกต์ใช้ การเรียนรู้แบบยูบิควิตัส

การศึกษา 4.0 เป็นการมุ่งให้ความรู้ไปที่ตัวผู้เรียนรายบุคคล ก่อให้เกิดการคิดวิเคราะห์จนนำไปประยุกต์ใช้ในการสร้างสิ่งประดิษฐ์และต่อยอดไปจนเกิดเป็นนวัตกรรมจากรายบุคคลหรือการรวมตัวกันเป็นทีม [3] ในปัจจุบันจะพบว่ามีการเรียนการสอน ในลักษณะนี้อยู่ในสถานศึกษาโดยทั่วไป การรวมตัวกันในการสร้างนวัตกรรมจึงเป็นเป้าประสงค์หลักในการให้สถานศึกษาเป็นแหล่งสร้างองค์ความรู้

นอกจากนั้น Arthur M. Harkins [3] ได้นิยามการศึกษา 4.0 ไว้ว่า เป็นยิ่งกว่าการศึกษา หมายความว่า เป็นการสร้างความรู้จากรายบุคคลและบุคคลที่รวมตัวเป็นทีมที่มีนวัตกรรมเป็นจุดเด่น เทคโนโลยีจะเปลี่ยนแปลงตามตัวผู้ เรียน ขยายองค์ความรู้ให้เกิดได้ทุกที่ทุกเวลาทั้งในชีวิตประจำวัน การเรียนและการทำงาน ในโลกไร้พรมแดนที่มีการเชื่อมเป็น เครือข่าย โรงเรียนเป็นที่สร้างนวัตกรรมอย่างต่อเนื่อง ครูคือแหล่งสร้างนวัตกรรม ในปัจจุบันครูจะนำการประมวลผลแบบคลาวด์ มา

ประยุกต์ใช้ในการเรียนการสอนร่วมกับสร้างนวัตกรรมไปพร้อมกันเพื่อให้เกิดการเรียนรู้ต่อผู้เรียนจนสร้างอาชีพได้

2.2 งานวิจัยที่เกี่ยวข้อง

นวพรรษ และคณะ [4] ได้นำเสนอการวิจัยเรื่อง “Ubiquitous Learning อัจฉริยะแห่งการล่วงรู้บริบท” โดยใช้ทฤษฎีของ Paul Lefrere [2] ที่ได้ศึกษาการออกแบบ U-Learning เพื่อการเรียนการสอนแบบร่วมมือกันและเน้นผลด้านทักษะ ทั้งนี้ผู้วิจัยได้ออกแบบ Frame Model สำหรับประยุกต์ใช้กับการเรียนการสอนแบบ U-Learning ได้อย่างลงตัวและเป็นระบบด้วย 6 โมดูล

วิชญา และคณะ [5] วิจัยรูปแบบการจัดการเรียนการสอนบนเทคโนโลยีคลาวด์เพื่อสนับสนุนการเรียนการสอนด้านคอมพิวเตอร์ โดยดำเนินการพัฒนาระบบการจัดการเรียนการสอนได้แก่ จัดทำสื่อลงคลาวด์ จัดการคลังข้อสอบ สังคมเครือข่าย จัดการห้องเรียน ระบบสนับสนุนผู้เรียน ระบบจัดการประเมินด้วยแนวคิดการพัฒนาในรูปแบบระบบประมวลผลแบบคลาวด์และเครือข่ายเชื่อมโยงการกระจายทรัพยากร การวิจัยระบบการจัดการเรียนรู้ยูบิควิตัสแบบเสริมศักยภาพของ

นภดล และคณะ[6] ได้ใช้ทฤษฎีการเรียนรู้ที่เหมาะสมสำหรับการนำมาใช้กับสภาพแวดล้อมการเรียนรู้แบบ ยูบิควิตัส และพัฒนาระบบการจัดการเรียนรู้ ผลการวิจัยพบว่าระบบการเรียนรู้ยูบิควิตัสแบบเสริมศักยภาพโดยใช้ปัญหาเป็นหลัก ระบบมีประสิทธิภาพอยู่ในระดับมากที่สุด ระบบทำให้ผู้เรียนสามารถเรียนรู้ได้ทุกที่ทุกเวลา

วิวัฒน์ [7] ได้กล่าวอ้างถึงผลการศึกษาระบบประมวลผลแบบแบบคลาวด์ของบริษัทไมโครซอฟท์ในงานทางการศึกษาว่าระบบคลาวด์เป็นทางเลือกใหม่ในการสร้างความสมดุลให้กับการจัดการระบบการใช้เทคโนโลยีสารสนเทศ และ การวิจัยครูเพื่อศิษย์ของพิจารณ์ [8] โดยผู้วิจัยได้สร้างห้องเรียนกลับทาง ที่มีแนวคิดว่าการเรียนรู้แบบนี้คือการเรียนรู้ที่เป็นไปตามธรรมชาติของมนุษย์ เป็นการนำเอาทฤษฎีการเรียนรู้แบบเทคโนโลยีทางการศึกษาผสมเข้ากับเทคโนโลยีสมัยใหม่ ช่วยให้นักเรียนได้เรียนรู้ตามธรรมชาติของมนุษย์

นรังสรรค์ และวรรณชัย [9] ได้วิจัยการพัฒนารูปแบบการจัดการห้องเรียนไซเบอร์บนระบบประมวลผลแบบคลาวด์ด้วยหลักจัดการเรียนแบบร่วมมือ พบว่ารูปแบบอยู่ในระดับดี ผลสัมฤทธิ์ทางการเรียนของนักศึกษาที่เรียนจากรูปแบบนี้ดีขึ้น ความสามารถในการเรียนแบบร่วมมือและความพึงพอใจของนักศึกษามีค่าเฉลี่ยอยู่ในระดับดี

สุรศักดิ์ และวาสนา [10] ได้วิจัยพัฒนาระบบการสอนแบบยูบิควิตัสสำหรับนักเรียนระดับมัธยมศึกษาตอนปลายภาคเหนือตอนบน โดยนำแนวคิดทฤษฎีวิธีระบบ (Systems Approach) โดยยึดหลักการและวิธีระบบ ให้สอดคล้องกับหลักการเรียนรู้ที่เกิดขึ้นได้ทุกที่ทุกเวลา เพื่อให้เหมาะสมและสอดคล้องกับบริบทของการจัดการเรียนการสอนระดับการศึกษาขั้นพื้นฐาน สำหรับงาน วิจัยศึกษาความคิดเห็นต่อการใช้สื่อสังคมออนไลน์บนระบบคลาวด์และโมบายแอปพลิเคชัน การจัดการเรียนการสอนของนักศึกษาระดับปริญญาเอกของเจือจันทน์ และคณะ [11] ได้ใช้แนวทางงานวิจัยของ Gartner พบว่าเทคโนโลยีการประมวลผลแบบคลาวด์ มีแนวโน้มความสำคัญเป็นอันดับหนึ่งของเทคโนโลยีสารสนเทศ การประมวลผลจะใช้รูปแบบบริการบนพื้นฐานการจัดสรรทรัพยากรตามการที่ใช้งานจริง

วาฤทธิ และนมน [12] ได้วิจัยการออกแบบรูปแบบการเรียนรู้ร่วมกันผ่านเทคโนโลยีคลาวด์ เพื่อส่งเสริมความคิดสร้างสรรค์ของนักศึกษาระดับอุดมศึกษา โดยนำแนวคิดการสร้างสภาพแวดล้อมจัดกิจกรรมการเรียนการสอนที่เป็นงานวิจัยของ “Google Collaborative Platform” ที่เป็นปัจจัยในการสร้างสภาพแวดล้อมของ อีเลิร์นนิง ซึ่งผลการวิจัยมีความสอดคล้องไปแนวทางของงานวิจัยที่ได้ศึกษา ธนยศ [13] ศึกษาในเรื่องไคลเอนต์คลาวด์ของคอมพิวเตอร์บนการประมวลผลแบบคลาวด์สำหรับแอปพลิเคชันที่ส่งให้หรือกำหนดการออกแบบสำหรับการรับบริการแบบคลาวด์ ได้แก่ มือถือ Android, Iphone, Windows Mobile

นาวิน[14] ศึกษาการประมวลผลแบบคลาวด์ในระดับการศึกษาระดับสูงประกอบด้วย การพัฒนาฐานความรู้เกี่ยวกับระบบคลาวด์ การประเมินขั้นตอนปัจจุบันของมหาวิทยาลัยจากมุมมองของความจำเป็นด้านไอที โครงสร้างและการใช้งาน การ

ทดลองใช้การประมวลผลแบบคลาวด์ การเลือกใช้ระบบคลาวด์ที่เหมาะสมกับการใช้งานและการจัดการ ผลการวิจัยพบว่าค่าใช้จ่ายลดลงอย่างมีนัยสำคัญเนื่องจากดำเนินการแก้ปัญหาที่นำคลาวด์มาใช้งาน กรรวิภา และปณิตา [15] ใช้เทคนิค Black Box Test มาทดสอบระบบ การทดสอบจะไม่คำนึงถึงคำสั่งภายในระบบ ทดสอบฟังก์ชันต่าง ๆ ตามเงื่อนไขที่มี โดยดูค่าเอาต์พุต และอินพุต ที่ให้กับระบบต้องมีความสอดคล้องกันโดยแบ่งการทดสอบการหาประสิทธิภาพของระบบ เป็น 4 ด้านคือ 1. ด้านความสามารถในการทำงานของระบบ 2.ด้านความถูกต้องในการทำงานของระบบ 3.ด้านความสะดวกและง่ายต่อการใช้งานระบบ และ 4.ด้านการรักษาความมั่นคงปลอดภัยของระบบ ผลของระบบที่พัฒนาขึ้นมีประสิทธิภาพและเหมาะสมที่จะนำมาใช้ในการแลกเปลี่ยนเรียนรู้

กนกวรรณ และคณะ [16] ได้นำเสนอแบบจำลองสำหรับ e-Learning Application Model บน การประมวลผลแบบคลาวด์ พบว่า การประมวลผลแบบคลาวด์ สามารถประยุกต์ใช้ในสถาบันการศึกษาโดยสามารถตอบสนองต่อจุดประสงค์ของของ อีเลิร์นนิง เช่น มีพื้นที่เก็บข้อมูล สามารถแบ่ง ปันข้อมูลได้

ทฤษฎีและงานวิจัยที่ได้ศึกษามานั้น ล้วนแต่นำมาเพื่อประยุกต์ใช้และต่อยอดให้แก่การวิจัยที่กำลังดำเนินการ ส่งผลในด้านดีต่อการจัด การเรียนการสอนให้แก่ผู้เรียนในสายอาชีวศึกษาซึ่งเป็นการผลิตผู้เรียนเพื่อประกอบอาชีพ การเรียนรู้ของผู้เรียนที่มีทั้งหลัก สูตรประกาศนียบัตรวิชาชีพ หลักสูตรประกาศนียบัตรวิชาชีพชั้นสูง หลักสูตรระยะสั้น หลักสูตรทวิภาคี หลักสูตรทวิศึกษา และหลักสูตรหลากหลาย มุ่งผลิตกำลังคนเพื่อตอบสนองการใช้งานด้านฝีมือแรงงานให้แก่สถานประกอบการทั้งขนาดเล็ก ขนาดกลางและขนาดใหญ่

การวิจัยดังกล่าวเป็นการเตรียมความพร้อมให้แก่ผู้เรียนที่ต้องนำไปใช้ในขณะที่เข้ารับการฝึกงานก่อนจบการศึกษาในสถานประกอบการ และนำไปใช้ในการประกอบอาชีพเมื่อเข้าสู่สถานประกอบการในฐานะของพนักงานของสถานประกอบการ จึงจำเป็นที่ผู้เรียนต้องเตรียมความพร้อมให้ทันต่อการเปลี่ยนแปลงในยุคการศึกษา 4.0 ตามยุทธศาสตร์การพัฒนาดิจิทัลเพื่อเศรษฐกิจและสังคม กระทรวงดิจิทัลเพื่อเศรษฐกิจและสังคม [17] ด้วยการพัฒนากำลังคนให้พร้อมเข้าสู่ยุคเศรษฐกิจและสังคมดิจิทัล สร้างและพัฒนาบุคลากรผู้ทำงานให้มีความสามารถในการสร้างสรรค์ และใช้เทคโนโลยีดิจิทัลอย่างชาญฉลาดในการประกอบอาชีพรวมถึงพัฒนาทักษะด้านเทคโนโลยีดิจิทัลในบุคลากรภาครัฐ ภาคเอกชนทั้งที่ประกอบอาชีพในสาขาเทคโนโลยีดิจิทัลโดยตรง และทุกสาขาอาชีพให้มีความรู้ความสามารถและความเชี่ยวชาญตามระดับมาตรฐานสากล เพื่อสร้างให้เกิดการจ้างงานที่มีคุณค่าสูง รองรับการพัฒนาประเทศในยุคเศรษฐกิจและสังคมที่ใช้เทคโนโลยีดิจิทัลเป็นปัจจัยหลักในการขับเคลื่อนให้สอดรับกับนโยบายชาติในการปฏิรูปประเทศไทยไปสู่ความมั่นคง มั่งคั่งและยั่งยืน สอดคล้องกับงานวิจัยการออกแบบโมเดลการเรียนรู้แบบยูบิควิตัสบนคลาวด์สำหรับมุ่งสู่การศึกษาไทย 4.0

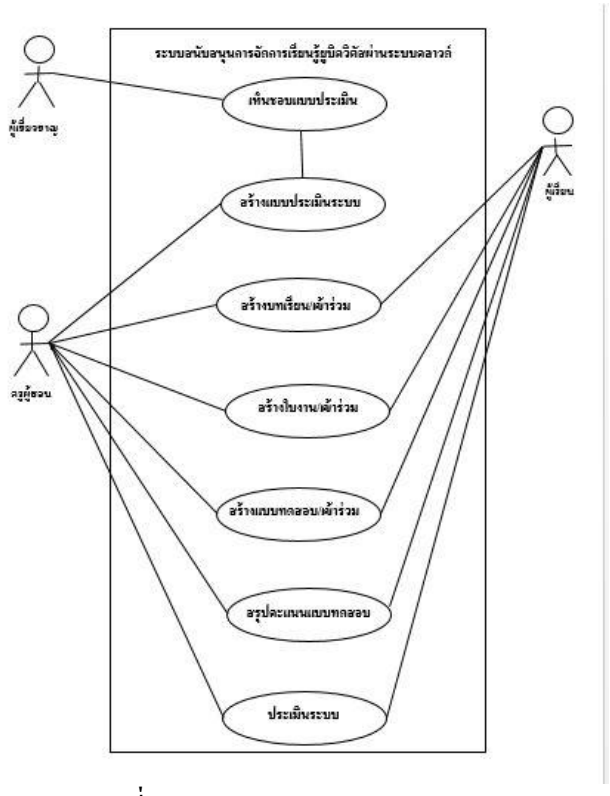
3. วิธีดำเนินการวิจัย

3.1 ขั้นตอนและวิธีดำเนินการวิจัย

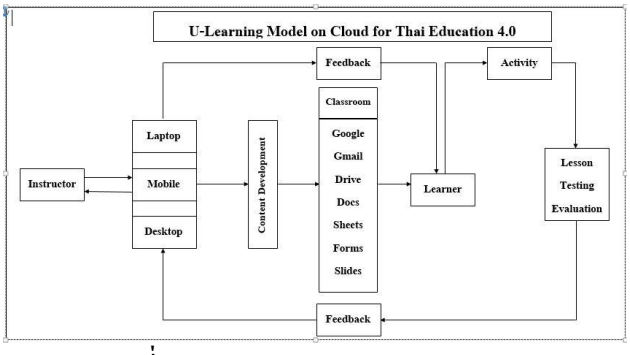
ผู้วิจัยได้กำหนดขั้นตอนการดำเนินการไว้ 1) ศึกษาทฤษฎีและงานวิจัยที่เกี่ยวข้อง 2) กำหนดประชากรและกลุ่มตัวอย่างเป็น 2 กลุ่มได้แก่ กลุ่มผู้ประเมินโมเดลและกลุ่มผู้ใช้งานระบบตามโมเดลที่ได้ออกแบบไว้ 3) ทำการออกแบบโมเดลการเรียนรู้ยูบิควิตัสเลิร์นนิงบนระบบคลาวด์สำหรับมุ่งสู่การศึกษาไทย 4.0 4) สร้างแบบประเมินโมเดลการเรียนรู้ยูบิควิตัสเลิร์นนิงบนระบบคลาวด์สำหรับผู้เชี่ยวชาญ 5) สร้างแบบประเมินผลสำหรับผู้ใช้ระบบการเรียนรู้ยูบิควิตัสเลิร์นนิงบนคลาวด์ และ 6) เก็บแบบประเมินผล ทำการวิเคราะห์และสรุปผล

ผู้วิจัย ได้ศึกษาแนวคิดและทฤษฎีโดยทำการศึกษารวบรวมข้อมูลและสังเคราะห์ข้อมูลระบบการจัดการเรียนรู้ และระบบสนับสนุนการเรียนรู้ในรูปแบบต่าง ๆ ที่เกี่ยวข้องกับยูบิควิตัสเลิร์นนิง ระบบคลาวด์ รวมทั้งงานวิจัยในอดีตและปัจจุบันในขณะ เดียวกัน ได้ดำเนินการสำรวจสภาพการใช้งานบนสมาร์ตโฟน หรือ อุปกรณ์พกพาของนักศึกษาว่ามีแอปพลิเคชันของกูเกิลคลาสรูมที่สามารถใช้ได้กับอุปกรณ์ iOS(R) Android ผ่าน Chrome, Firefox(R), Internet Explorer(R), หรือ Safari(R) ทั้งนี้ เพื่อความพร้อมในการใช้งานของผู้ใช้ระบบ

3.1.1 ออกแบบโมเดลการเรียนรู้ยูบิควิตัสเลิร์นนิงบนระบบคลาวด์สำหรับมุ่งสู่การศึกษาไทย 4.0 ดังภาพที่ 1 และ 2



ภาพที่ 1: Use-case แสดงความสัมพันธ์ระหว่างผู้ใช้กับระบบการเรียนรู้ที่ได้ออกแบบ



ภาพที่ 2: โมเดลการเรียนรู้แบบยูบิควิตัสบนคลาวด์

3.2 กำหนดประชากรและกลุ่มตัวอย่าง

ประชากรที่ใช้ในการวิจัยครั้งนี้ได้แก่ 1. ผู้เชี่ยวชาญจำนวน 9 ท่าน จำแนกเป็นผู้เชี่ยวชาญ 3 ด้าน ด้านละ 3 ท่าน ดังนี้ ผู้เชี่ยวชาญด้านการเรียนการสอนในสาขาวิชาที่เกี่ยวข้องหรือครูผู้สอนสาขาวิชาคอมพิวเตอร์ธุรกิจ จำนวน 3 ท่าน ผู้เชี่ยวชาญด้านระบบเทคโนโลยีสารสนเทศ จำนวน 3 ท่าน และผู้เชี่ยวชาญด้านการศึกษา จำนวน 3 ท่าน 2. ประชากรที่กำหนดนักศึกษาระดับประกาศนียบัตรวิชาชีพชั้นสูง วิทยาลัยเทคนิคกาญจนบุรีทุกสาขาวิชาที่ลงทะเบียนเรียนวิชาเทคโนโลยีสารสนเทศเพื่อการจัดการอาชีพ ภาคเรียนที่ 2 ปีการศึกษา 2559 กลุ่มตัวอย่างที่นำมาใช้ได้แก่นักศึกษาระดับประกาศนียบัตรวิชาชีพชั้นสูงสาขาวิชาช่างไฟฟ้ากำลัง วิทยาลัยเทคนิคกาญจนบุรี จำนวน 30 คน

3.3 เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัยได้แก่ แบบประเมินโมเดลการเรียนรู้ยูบิควิตัสเลิร์นนิงบนคลาวด์ แบบประเมินผลสำหรับผู้ใช้การเรียนรู้ยูบิควิตัสเลิร์นนิงบนคลาวด์ ชุดเครื่องมือเพิ่มประสิทธิภาพการทำงานกูเกิ้ลคลาสรูม บทเรียนวิชาเทคโนโลยีสารสนเทศเพื่อการจัดการเรียนรู้จำนวน 5 หน่วยการเรียนรู้ และการวิเคราะห์ข้อมูลทางสถิติด้วยโปรแกรมคอมพิวเตอร์สำเร็จรูป SPSS

3.4 การเก็บรวบรวมข้อมูล

ผู้วิจัยได้ดำเนินการเก็บข้อมูลจากกลุ่มตัวอย่างกลุ่มแรกได้แก่ผู้เชี่ยวชาญ จำนวน 9 ท่าน ทำการประเมินผลความเห็นของผู้เชี่ยวชาญ และดำเนินการเก็บผลการประเมินจากกลุ่มตัวอย่างที่สอง เป็นผู้เข้าใช้ระบบทำการทดสอบระบบจริงโดยผู้เรียนจำนวน 30 คน เก็บรวบรวมประเมินผลบนระบบนำไปวิเคราะห์ข้อมูลทางสถิติด้วยโปรแกรมวิเคราะห์ข้อมูลทางสถิติและหาค่าเฉลี่ย และค่าส่วนเบี่ยงเบนมาตรฐาน การวิเคราะห์ข้อมูลจากแบบประเมินโดยผู้เชี่ยวชาญ และผู้ใช้ระบบจริงที่เป็นกลุ่มตัวอย่างนั้นจะใช้ระยะเวลาการดำเนินการในภาคเรียนที่ 2 ปีการศึกษา 2559 จำนวน 18 สัปดาห์ ทั้งนี้ เพื่อให้การประเมินความเหมาะสมของโมเดลการเรียนรู้แบบยูบิควิตัสบนระบบคลาวด์สำหรับมุ่งสู่การศึกษาไทย 4.0 สามารถสำเร็จสมบูรณ์ได้

4. ผลการวิจัย

ผลการออกแบบโมเดลการเรียนรู้แบบยูบิควิตัสบนระบบคลาวด์สำหรับมุ่งสู่การศึกษาไทย 4.0 ประกอบด้วย โมดูลย่อยจำนวน 8 โมดูลได้แก่ โมดูลการจัดการชั้นเรียน โมดูลการ

สื่อสารและทำงานร่วมกัน โมดูลการเข้าถึงทุกที่ทุกเวลา โมดูลปรับปรุงการสื่อสารกับผู้เรียน โมดูลประหยัดเวลา โมดูลงานเฉพาะคน โมดูลงานจัดการคลังข้อสอบ โมดูลการวัดและประเมินผลการเรียน

การประเมินความเหมาะสมของโมเดลการเรียนรู้แบบยูบิควิตัสโดยผู้เชี่ยวชาญ จำนวน 9 ท่าน รวม 10 ด้าน ด้านที่มีความเหมาะสมมากที่สุด 5 ด้าน ได้แก่ด้านประโยชน์การใช้งาน ด้านการสื่อสารข้อมูล ด้านการใช้ทรัพยากร ด้านการทำงานร่วมกัน ด้านประสิทธิภาพการทำงานของระบบ ซึ่งสอดคล้องกับแนวทฤษฎี ด้านที่มีความเหมาะสมมาก 5 ด้าน ได้แก่ด้านความมั่นคงปลอดภัย ด้านการจัดการชั้นเรียน ด้านความสามารถในการเข้าถึงข้อมูล ด้านแอปพลิเคชันที่ใช้งานร่วมกับอุปกรณ์ ด้านสภาพแวดล้อมการใช้งาน ค่าความเหมาะสมเฉลี่ยรวมเท่ากับ 4.48 มีความเหมาะสมระดับมาก ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.58 มีค่าความเชื่อมั่นปานกลาง แสดงว่าผู้เชี่ยวชาญมีความคิดเห็นที่มีแนวโน้มไปทางเดียวกันยอมรับโมเดลการเรียนรู้แบบยูบิควิตัสบนคลาวด์มีความเหมาะสมในระดับมาก สอดคล้องกับ Paul Lefrere [4] ที่ได้ศึกษาการออกแบบ U-Learning เพื่อการเรียนการสอนแบบร่วมมือกันและเน้นผลด้านทักษะ โดยเขาได้ออกแบบ Frame Model สำหรับประยุกต์ใช้กับการเรียนการสอนแบบ U-Learning อย่างลงตัวและเป็นระบบตามตารางที่ 1

ตารางที่ 1: การประเมินความเหมาะสมของโมเดลการเรียนรู้แบบยูบิควิตัสโดยผู้เชี่ยวชาญ

หัวข้อการประเมินความเหมาะสมของโมเดลการเรียนรู้แบบยูบิควิตัส	ผลการประเมิน		แปลความระดับความเหมาะสม
	$\bar{X}$	S.D.	
1. ด้านความมั่นคงปลอดภัย	3.88	0.92	มาก
2. ด้านประโยชน์การใช้งาน	4.55	0.52	มากที่สุด
3.ด้านการจัดการชั้นเรียน	4.44	0.72	มาก
4.ด้านการสื่อสารข้อมูล	4.66	0.50	มากที่สุด
5.ด้านการใช้ทรัพยากร	4.55	0.52	มากที่สุด
6.ด้านการทำงานร่วมกัน	4.66	0.50	มากที่สุด
7.ด้านความสามารถในการเข้าถึงข้อมูล	3.88	0.92	มาก
8.ด้านแอปพลิเคชั่นที่ใช้งานร่วมกับอุปกรณ์	4.44	0.72	มาก
9.ด้านสภาพแวดล้อมการใช้งาน	4.33	0.70	มาก
10.ด้านประสิทธิภาพการทำงานของระบบ	4.66	0.50	มากที่สุด
ความเหมาะสม	4.48	0.58	มาก

การประเมินความเหมาะสมของโมเดลการเรียนรู้แบบยูบิควิตัสโดยผู้ใช้ระบบ(ผู้เรียน) จำนวน 30 คน รวม 10 ด้าน ด้านที่มีระดับความเหมาะสมมากที่สุด 6 ด้าน ได้แก่ ทำงานได้ตรงเวลา ใช้ประโยชน์ทรัพยากรได้คุ้มค่า มีความสามารถในการใช้งาน มีความถูกต้อง ทำหน้าที่ตามที่กำหนดไว้ ความสมบูรณ์ของระบบ และมีความเหมาะสมต่อการใช้งาน รองลงมาระดับความเหมาะสมมาก 3 ด้าน ได้แก่ มีความมั่นคงปลอดภัย การติดตามความก้าวหน้า มีความยืดหยุ่นสามารถใช้งานได้ทุกที่

ค่าความเหมาะสมเฉลี่ยรวมเท่ากับ 4.43 ส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 0.65 แสดงว่าโมเดลการเรียนรู้แบบยูบิควิตัสบนคลาวด์ผู้เรียนยอมรับในโมเดลการเรียนรู้ยูบิควิตัส มีความเหมาะสมอยู่ในระดับมาก ดังตารางที่ 2

ตารางที่ 2: การประเมินความเหมาะสมของโมเดลการเรียนรู้แบบยูบิควิตัสโดยผู้ใช้ระบบ(ผู้เรียน)

หัวข้อการประเมินความเหมาะสมของโมเดลการเรียนรู้แบบยูบิควิตัส	ผลการประเมิน		แปลความระดับความเหมาะสม
	$\bar{X}$	S.D.	
1. ทำงานได้ตรงเวลา	4.63	0.55	มากที่สุด
2. ใช้ประโยชน์ทรัพยากรได้คุ้มค่า	4.60	0.56	มากที่สุด
3. มีความสามารถในการใช้งาน	4.63	0.61	มากที่สุด
4. มีความถูกต้อง	4.73	0.52	มากที่สุด
5. มีความมั่นคงปลอดภัย	3.93	0.63	มาก
6. ทำหน้าที่ตามที่กำหนดไว้	4.53	0.50	มากที่สุด
7. ความสมบูรณ์ของระบบ	4.66	0.54	มากที่สุด
8. การติดตามความก้าวหน้า	4.10	0.75	มาก
9. มีความยืดหยุ่นสามารถใช้งานได้ทุกที่	3.93	0.63	มาก
10.มีความเหมาะสมต่อการใช้งาน	4.63	0.55	มากที่สุด
ความเหมาะสม	4.43	0.65	มาก

5. สรุปผลการวิจัย

โมเดลการเรียนรู้แบบยูบิควิตัสบนระบบคลาวด์สำหรับมุ่งสู่การศึกษา 4.0 เป็นโมเดลที่มีความเหมาะสมในการนำไปใช้จริง ผู้สอนและผู้เรียนสามารถติดต่อสื่อสารถึงกันผ่านเทคโนโลยีคลาวด์ด้วยต้นทุนที่ต่ำ และรองรับการใช้งานของผู้เรียนในยุคการศึกษา 4.0 ได้อย่างมีประสิทธิภาพ ซึ่งผลการประเมินความเหมาะสมของระบบจากทั้งผู้เชี่ยวชาญและผู้เรียนบ่งบอกว่ามีความเหมาะสมต่อการใช้งานอยู่ในระดับมาก ระบบนี้สามารถนำไปใช้เพื่อพัฒนาการศึกษาให้เป็นการศึกษาไทย 4.0 ได้

6. การอภิปรายผล

การวิเคราะห์ผลการวิจัยพบว่าการออกแบบในส่วนของผู้ใช้ระบบประกอบด้วยผู้สอนและผู้เรียนโดยมีผู้เชี่ยวชาญให้คำแนะนำและเห็นด้วยกับโมเดลและมีความสอดคล้องกับบริบทของการจัดการเรียนการสอนในสภาพแวดล้อมที่แตกต่างกันให้สามารถทำงานประสานกันได้

อย่างไรก็ดี ใน 12 ด้านแรกที่ควรนำมาใช้ในการจัดการเรียนรู้ซึ่งมีความเหมาะสมมากที่สุด ได้แก่ ด้านการสื่อสารข้อมูล ด้านการใช้ทรัพยากร ด้านการทำงานร่วมกัน ด้านประสิทธิภาพ การทำงานของระบบ ทำงานได้ตรงเวลา ใช้ประโยชน์ทรัพยากรได้คุ้มค่า มีความสามารถในการใช้งาน มีความถูกต้อง ทำหน้าที่ตามที่กำหนดไว้ ความสมบูรณ์ของระบบและ มีความเหมาะสมต่อการใช้งาน เป็นต้น ทั้งนี้ ผลการวิจัยยังได้สอดคล้องกับแนวคิดของ Paul Lefrere [2] ที่ได้ศึกษาการออกแบบ U-Learning สำหรับประยุกต์ใช้กับการเรียนการสอนแบบยูบิควิตัสอย่างลงตัวและเป็นระบบ

เอกสารอ้างอิง

[1] Peter Mell, Timothy Grance, The NIST Definition of Cloud Computing : Recommendations of the National Institute of Standards and Technology, National Institute of Standards and Technology, 2011.

[2] P. Lefrere, *Activity-based scenarios for and approaches To ubiquitous e-Learning,* in Springer, London, UK, 2007.

[3] ไพฑูรย์ สินลารัตน์ สินธะวา คามดิษฐ์ เฉลิมชัย มนูเสวต วาสนา วิสฤตาภา และนักรบ หมื่แสน, *การศึกษา 4.0 เป็นยิ่งกว่าการศึกษา,* โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย: กรุงเทพฯ, 2559

[4] นวพรรษ เพชรมณี และปรัชญนันท์ นิลสุข, "Ubiquitous Learning อัจฉริยะแห่งการล่วงรู้บริบท," *วารสารวิทยบริการ มหาวิทยาลัยสงขลานครินทร์* ปีที่ 21 ฉบับที่ 1 พ.ศ. 2553.

[5] วิชญา รุ่นสุวรรณ์ ดวงกมล โพธิ์นาค และปรวัตน์ วิสูตรศักดิ์, "รูปแบบระบบการจัดการเรียนการสอนบนเทคโนโลยีก้อนเมฆเพื่อสนับสนุนการเรียนการสอนด้านคอมพิวเตอร์," *วารสารปัญญาภิวัตน์* ฉบับที่ 7 พ.ศ. 2557

[6] นภดล ผู้มีจรรยา ปณิตา วรรณพิรุณ และปรัชญนันท์ นิลสุข, "ระบบการจัดการเรียนรู้ยูบิควิตัสแบบเสริมศักยภาพ," *การประชุมวิชาการระดับชาติโสตฯ-เทคโนฯ สัมพันธ์แห่งประเทศไทย* ครั้งที่29 พ.ศ. 2558.

[7] วิวัฒน์ มีสุวรรณ์, "ระบบประมวลผลแบบกลุ่มเมฆในงานทางการศึกษา," *วารสารศึกษาศาสตร์มหาวิทยาลัยนเรศวร* ปีที่ 16 ฉบับที่ 1 พ.ศ 2557.

[8] พิจารณ์ พานิช, "ครูเพื่อศิษย์ห้องเรียนกลับทาง," *มูลนิธิสยามกัมมาจล* ครั้งที่ 2 พ.ศ. 2556 หน้า 54.

[9] นรังสรรค์ วิไลสกุลยง และวรรณชัย วรรณสวัสดิ์ "การพัฒนารูปแบบการจัดการห้องเรียนไซเบอร์บนระบบประมวลผลกลุ่มเมฆด้วยหลักจัดการเรียนแบบร่วมมือ," *วารสารพัฒนาเทคนิคศึกษา มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ* ปีที่ 2017 ฉบับที่ 101.

[10] สุรศักดิ์ ปาเฮ และวาสนา ทวีกุลทรัพย์ "การพัฒนาระบบการสอนภควันตภาพสำหรับนักเรียนระดับมัธยมศึกษาตอนปลายภาคเหนือตอนบน," *วารสารวิชาการมหาวิทยาลัยธนบุรี* ปีที่ 10 ฉบับที่ 23.

[11] เจือจันทน์ วัฒเกเจริญ ไพศาล จันทรังษี และปณิตา วรรณพิรุณ "ความคิดเห็นต่อการใช้สื่อสังคมออนไลน์ คลาวด์คอมพิวติง และโมบายแอพพิเคชั่น เพื่อการศึกษาของนักศึกษาปริญญาเอก," *วารสารพยาบาลศาสตร์ มหาวิทยาลัยมหิดล* ปีที่ 33 ฉบับพิเศษที่ 1.

[12] วาฤทธิ์ กันแก้ว และณมน จีรังสุวรรณ "การออกแบบรูปแบบการเรียนรู้ร่วมกันผ่านเทคโนโลยีคลาวด์เพื่อส่งเสริมความคิดสร้างสรรค์ของนักศึกษาระดับอุดมศึกษา," *วารสารวิชาการครุศาสตร์อุตสาหกรรม* ปีที่ 6 ฉบับที่ 1.

[13] ธนยศ สิริโชดก "แนวคิดในการจัดสภาพแวดล้อมทางการเรียนการสอนแบบร่วมมือกัน เรียนรู้บนคลาวด์คอมพิวติง," *วารสารอาชีวะและเทคนิคศึกษา* ปีที่ 2 ฉบับที่ 3.

[14] นาวิน คงรักษา "คลาวน์คอมพิวติ้งกับการจัดการเรียนการสอนในศตวรรษที่ 21," *วารสารอาชีวะและเทคนิคศึกษา มหาวิทยาลัยพระจอมเกล้าพระนครเหนือ* ปีที่ 4 ฉบับที่ 7 พ.ศ. 2557.

[15] กรรวิภา หวังทอง และปณิตา วรรณพิรุณ "การพัฒนาระบบสารสนเทศเพื่อการแลกเปลี่ยนเรียนรู้ผ่านการประมวลผลแบบกลุ่มเมฆ," *การประชุมวิชาการระดับชาติด้านอีเลิร์นนิง* มหาวิทยาลัยไซเบอร์ไทยประจำปี *พ.ศ. 2556*, 5-6 สิงหาคม 2556, 16 - 22.

[16] กนกวรรณ อัจฉริยะชาญฉิช ณัฐพงศ์ ใจวงศ์ และณัฐชนน สิริปูชกะ "การนำเสนอแบบจำลองเพื่อวิเคราะห์ปัจจัยที่มีอิทธิพลต่อความตั้งใจที่จะใช้ e-Learning บน Cloud Computing," *การประชุมวิชาการระดับชาติด้านอีเลิร์นนิง* มหาวิทยาลัยไซเบอร์ไทยประจำ ปี *พ.ศ. 2556*" 5-6 สิงหาคม 2556, 16 - 22.

[17] กระทรวงดิจิทัลเพื่อเศรษฐกิจและสังคม, *แผนพัฒนาดิจิทัลเพื่อเศรษฐกิจและสังคม,* พิมพ์ครั้งที่ 1 พ.ศ.2559 หน้า 4.

คุณลักษณะสำคัญของสื่อดิจิทัลใหม่ที่มีผลต่อการรับรู้ของนักศึกษาและกระบวนการศึกษาไทย

Key Features in New Digital Media Affecting Student Perception and Thai Education Processes

บุญชม สุดจิตต์[1] (Boonchom Sudjit)[1] และ ประสงค์ ปราณีตพลกรัง[2] (Prasong Praneetpolgrang)[2]
บัณฑิตศึกษา คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยศรีปทุม
[1] sudjit2009@hotmail.com, [2] prasong.pr@spu.ac.th

บทคัดย่อ

ปัจจุบันรัฐบาลให้ความสำคัญต่อสื่อดิจิทัลใหม่ ซึ่งส่งผลในการรับรู้ข่าวสารเชิงรุกในกระบวนการศึกษาไทย ด้วยรูปแบบที่มีความหลากหลายจึงส่งผลกระทบโดยตรงแก่นักศึกษาและ ทำให้การขัดกรองข้อมูลข่าวสารของสื่อดิจิทัลใหม่ที่ถูกต้องไม่สามารถทำได้ 100% ดังนั้น การศึกษาของประเทศไทยถือได้ว่ามีความสำคัญระดับต้นๆ งานวิจัยนี้จึงมีวัตถุประสงค์ เพื่อศึกษาคุณลักษณะสำคัญของสื่อดิจัลใหม่ที่มีผลต่อการรับรู้ของนักศึกษาและกระบวนการศึกษาไทย โดยใช้เทคนิคเฟรมเวิร์กบนพื้นฐานคุณลักษณะสำคัญ 5 ประการ

ผู้วิจัยได้ใช้การวิจัยเชิงคุณภาพโดยเริ่มจากการสัมภาษณ์เจาะลึกแบบมีโครงสร้างกับผู้เชี่ยวชาญด้านสื่อโฆษณาดิจิทัล ผู้เชี่ยวชาญด้านการศึกษา และผู้เชี่ยวชาญด้านสื่อสารมวลชน ผลการวิจัย พบว่า คุณลักษณะสำคัญของสื่อดิจัลใหม่ที่มีผลต่อการรับรู้ของนักศึกษาและกระบวนการศึกษาไทย จากผู้เชี่ยวชาญ โดย มีประเด็นด้านคุณลักษณะตัวอักษรที่ส่งผลกระทบมากถึงร้อยละ 80 ตามด้วยคุณลักษณะวีดิทัศน์ร้อยละ 20 ส่วนคุณลักษณะอื่นๆ ไม่ได้มีผลกระทบ เมื่อนำคุณลักษณะสำคัญของสื่อดิจิทัลใหม่มาให้กับกลุ่มตัวอย่างจำนวน 50 คน พบว่าคุณลักษณะตัวอักษรที่ส่งผลกระทบมากตามด้วยคุณลักษณะวีดิทัศน์ สรุปได้ว่า ผู้เชี่ยวชาญและกลุ่มตัวอย่างให้คุณลักษณะสำคัญเป็นไปในทิศทางเดียวกัน

คำสำคัญ: คุณลักษณะสำคัญ, สื่อดิจิทัล, การศึกษา

Abstract

Currently, the government has paid an attention to a perception of an aggressive news of Thailand educational process in the form of diverse new digital media as first priority. News perception of Thai teenagers is relatively quick and directly affects on Thai educational process. It is therefore necessary to screen these information of digital media news before being transmitting. However, it can not be done 100%. Therefore, the impact of transfering the news to society and education of Thailand is also much importance. The purpose of this research is to study the key features and the impact of new digital media effecting on Thai educational process by using framework technique based on 5 features. Researcher has applied qualitative research by beginning from structured in-depth interviewing with digital advertising experts, educational experts and mass media experts. The research result reveals that the most key feature and the impact of new digital media effecting on thai education process is feature of letter accounting for 80% followed by a feature of video accounting for 20 percent. In addition, other features have not much impact. When these key features of the new digital media is adopted with 50 sample group, it is found that the most impact feature is letter and the secondary priority is the video feature. It can conclude that both the experts and the sample group provide data showing that the key feature is the same direction.

Keywords: Key Features, Digital Media, Education

1. บทนำ

สื่อดิจิทัลใหม่มีความหลากหลายมากยิ่งขึ้น ทำให้ได้รับการตอบสนองถึงรูปแบบของคุณลักษณะที่แตกต่างของทุกมิติ ด้วยการใช้เทคโนโลยีทันสมัยต่อการรับและส่งข้อมูล ทำให้การเผยแพร่รับรู้ข่าวสารมีความรวดเร็วซึ่งปัจจัยดังกล่าวผู้ผลิตสื่อดิจิทัลใหม่จึงพยายามลดต้นทุนการผลิตป้ายสื่อโฆษณาให้มีจำนวนน้อยลงโดยเฉพาะที่ติดตั้งกลางแจ้ง ขณะเดียวกันก็หาช่องทางการโฆษณาโดยใช้รูปแบบสื่อดิจิทัลใหม่เข้าสู่ภาคการศึกษาโดยเจาะกลุ่มเป้าหมายไปที่นักศึกษาเป็นสำคัญทำให้เกิดความรวดเร็วบริษัทเอกชนเร่งเห็นช่องทางด้านการตลาด มุ่งเป้าการตลาดเชิงรุกไปที่กลุ่มนักศึกษาเพื่อแทรกไว้ในอีกหนึ่งมิติของกระบวนการศึกษาที่สามารถทำเงินได้เป็นจำนวนมาก จึงเกิดผลกระทบตามมาทางด้านการศึกษาแบบไม่รู้ตัว การส่งข้อมูลและรับข้อมูลในคุณลักษณะต่างๆ มีทั้งบวกและลบในรูปแบบของเทคโนโลยีสารสนเทศ ดังนั้นรัฐบาลจึงเห็นความสำคัญในมิติที่แตกต่างดังกล่าวจึงมีการระดมสมองเพื่อหาแนวทางป้องกันต่อการขับเคลื่อนของกระบวนการศึกษาไทยสู่การพัฒนาประเทศชาติให้มีความมั่นคง มั่งคั่ง และ ยั่งยืน

การใช้เทคโนโลยีสารสนเทศรูปแบบสื่อดิจิทัลใหม่ ในกระบวนการศึกษามีจุดมุ่งหมายเพื่อการพัฒนาสร้างสรรค์ให้ได้มากที่สุดเป็นประโยชน์ต่อกลุ่มนักศึกษาในองค์ความรู้ต่อการรับข้อมูลและส่งข้อมูลบนพื้นฐานของความถูกต้องในข่าวสาร การออกแบบพัฒนาเทคโนโลยีสารสนเทศที่ผลิตสื่อดิจิทัลใหม่นั้นหลายสถาบันได้ตื่นตัวในเรื่องการพัฒนาบุคลากรด้านความรู้ที่จะนำมาบูรณาการให้เกิดความน่าสนใจมากยิ่งขึ้น

จากประเด็นปัญหาดังกล่าวผู้วิจัยจึงได้ทำการศึกษาวิจัย เพื่อหาว่ามีตัวแปรหรือปัจจัยของคุณลักษณะใดบ้างที่ส่งผลต่อคุณลักษณะสำคัญของสื่อดิจิทัลใหม่ที่มีผลต่อการรับรู้ของนักศึกษา จากนั้นจึงนำผลที่ได้ไปศึกษาและพัฒนาให้เกิดประโยชน์อย่างแท้จริง อันจะนำไปสู่การรับรู้ของนักศึกษา เพื่อความยั่งยืนของการพัฒนาสำหรับสนับสนุนในสถาบันและกระบวนการศึกษาในการพัฒนาสื่อดิจิทัลใหม่ และการสร้างเนื้อหาที่ทำให้เกิดความน่าสนใจต่อภาคการศึกษาของไทยต่อไป

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 สื่อดิจิทัลใหม่

การนำเอาข้อความ กราฟิก ภาพเคลื่อนไหว เสียง และวีดิทัศน์ โดยอาศัยช่องทางเทคโนโลยีสารสนเทศและการสื่อสารเช่น โปรแกรมประยุกต์เป็นส่วนสนับสนุนสื่อโฆษณาที่ติดตั้งภายในและภายนอกสถานที่เช่น บนท้องถนน ภายนอกอาคาร ส่งผลด้านแสง สี เสียง รูปแบบของป้ายสื่อโฆษณาดิจิทัล หรือป้ายตัวอักษรขนาดใหญ่ที่มีคุณลักษณะเฉพาะส่งผลกระทบต่อกลุ่มเป้าหมายด้านการตลาด ทำให้เกิดกิจกรรมเชิงรุกต่อการตัดสินใจของผู้รับสื่อไปถึงผู้ส่งสื่อแบบไม่รู้ตัวในกระบวนการศึกษาไทย

2.2 ปฏิบัติการข่าวสาร (Information Operations: IO)

การปฏิบัติการข่าวสาร [3] หรือที่เรียกกันว่า ไอโอ (IO) เริ่มรู้จักกันอย่างเป็นรูปธรรมช่วงหลังสิ้นสุดยุคสงครามเย็นของรูปแบบการปฏิบัติการ ฝ่ายทหารในประเทศต่างๆ และประเทศไทยซึ่งไม่ได้เรียกชื่อเจาะจงว่าเป็นการปฏิบัติการข่าวสาร โดยตรงมีผู้ตีความว่า ปฏิบัติการข่าวสาร หมายถึง การปฏิบัติการจิตวิทยา การโฆษณา ชวนเชื่อ การลวงทางทหาร ซึ่งมีการเปลี่ยน แปลงไปตามสภาพแวดล้อมทางด้านยุทธศาสตร์และพัฒนาการของเครื่องมือหรือสิ่งอำนวยความสะดวกที่นำมาใช้ในยุคของการพัฒนาเทคโนโลยีสารสนเทศอย่างต่อเนื่อง เช่น การใช้ดาวเทียม การใช้เครือข่ายทางคอมพิวเตอร์ การใช้ไวรัสคอมพิวเตอร์ หรือสื่อดิจิทัลชนิดใหม่ๆ ทั้งทางวิทยุ โทรทัศน์ โทรศัพท์เคลื่อนที่ ฯลฯ ทำให้การปฏิบัติการข่าวสารเป็นเครื่องมือสำคัญของการปฏิบัติการทางทหารเพราะการนำเอาเทคโนโลยีนำมาประยุกต์ใช้ในสนามรบสามารถเผยแพร่ข้อมูลและการควบคุมพื้นที่ในสนามรบมีประสิทธิภาพอย่างมากยิ่งขึ้น ที่จะนำไปใช้สนับสนุนการตัดสินใจในการทำสงครามจำเป็นต้องมีการปฏิบัติการพิสูจน์ทราบเกี่ยวกับฝ่ายตรงข้าม 3 ประการ ได้แก่ 1.ศักยภาพของฝ่ายตรงข้ามในการปฏิบัติการ 2.เจตนารมณ์ในการต่อสู้ และ 3.กระบวนการทางความคิดของฝ่ายตรงข้าม

ผู้วิจัยได้เร่งเห็นการเปลี่ยนแปลงการปฏิบัติการข่าวสารเชิงรุกของเทคโนโลยีสารสนเทศที่มีการพัฒนามากขึ้น จึงค้นคว้ารูปแบบทฤษฎีและงานวิจัยที่เกี่ยวข้องมิติต่างๆในคุณลักษณะที่

สำคัญ เพื่อนำมาพัฒนาให้เหมาะสมกับสภาวะแวดล้อมในปัจจุบัน

2.3 กรอบแนวคิดของแซคแมน

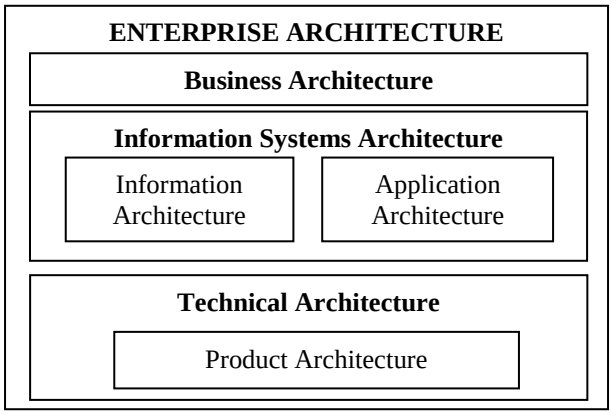
การบริหารบริษัทหรือองค์กรเพื่อที่จะทำการพัฒนาจำเป็นต้องมีระบบแบบแผนอย่างมีขั้นตอนในการประยุกต์ใช้ของการทำงานซึ่งจะทำให้เกิดการเชื่อมโยงที่รวดเร็วนำไปสู่การพัฒนา แนวคิดหรือกระบวนการในการพัฒนาองค์กรสู่เป้าหมายที่มีความชัดเจนในรูปแบบของสถาปัตยกรรมองค์กร (Enterprise Architecture) กรอบแนวคิดของแซคแมน [4], [5], [6] มีลักษณะรูปแบบเป็นเมทริกซ์ขนาด 6 คูณ 6 โดยในส่วนของคอลัมน์ทั้ง 6 จะประกอบได้แก่ What, How, Where, Who, When และWhy ในส่วนของแถวทั้ง 6 แถวก็จะเป็นมุมมองจากผู้ที่เกี่ยวข้องกับกรอบแนวคิดนั้น ๆ อันประกอบไปด้วย Scope, Business Model, System Model, Technology Model, Detailed Representation และ Real System กรอบแนวคิดของแซคแมนเป็นกรอบแนวคิดที่เข้าใจได้ง่าย ไม่มีกระบวนการ รูปแบบหรือสัญลักษณ์เฉพาะ ทั้งนี้เป็นเพียงแค่ตารางธรรมดา ไม่สามารถเพิ่มหรือลดตารางได้ ทำให้จุดแข็งของกรอบแนวคิดคือความเป็นตารางธรรมดา ทำให้ครอบคลุมสิ่งสำคัญทั้งหมดที่เราให้ความสนใจและไม่ก่อให้เกิดความซ้ำซ้อน ดังภาพที่ 1

	What	How	Where	Who	When	Why
Scope (Planner)	List of Things Important to the Business	List of Processes the Business Performs	List of Locations in which the Business Operates	List of Organizations important to the Business	List of Events Significant to the Business	List of Business Goals/Strategies
	- List	- Hierarchical list or tree	- Hierarchical list or tree	- List	- List	- Indented list
Enterprise Model (Owner)	Semantic Model	The Business Process Model	The Business Logistics System	Work Flow Model	Master Schedule	Business Plan
	- Entities Diagram - Entities Dictionary	- Entities VS. Processes Matrix - Processes Dictionary - Activity Diagram	- Functional Decomposition	- Organization Chart - Processes VS. Organization Matrix	- Business Execution Plan	- Business Table
System Model (Designer)	Logical Data Model	Application Architecture	The Distributed Systems Architecture	Human Interface Architecture	Processing Structure	Business Rules
	- Classes Diagram	- Systems VS. Processes Matrix - Systems VS. Entities Matrix - Systems Dictionary	- Systems Diagram	- Systems VS. Roles Matrix	- State Diagram	- Systems VS. Business Rules Matrix
	Data	Function	Network	People	Time	Motivation

ภาพที่ 1: แสดงมุมมองของมิติต่างๆ จากด้านบนลงล่างในแต่ละเซลล์

แซคแมน [4] เป็นกรอบแนวคิด ไม่มีกระบวนการรูปแบบ และเป็นตารางธรรมดา ไม่สามารถเพิ่มหรือลดตาราง ทำให้จุดแข็งของกรอบแนวคิดคือ ครอบคลุมสิ่งที่สำคัญทั้งหมดที่เราให้ความสนใจและไม่ก่อให้เกิดความซ้ำซ้อน

ใน EA Community (www.eacommunity.com) EA เป็นกรอบการทำงานหรือที่เรียกว่า Blueprint วัตถุประสงค์ทางธุรกิจปัจจุบันและอนาคตได้อย่างไร ดังนั้นแต่ละกลยุทธ์จะมีการแบ่งตามหลักการของสถาปัตยกรรมองค์กร มีกรอบทำงานที่ทำให้เกิดความชัดเจนยิ่งขึ้น ดังภาพที่ 2



ภาพที่ 2: แสดงโครงสร้างสถาปัตยกรรมองค์กร

2.3.1 สถาปัตยกรรมธุรกิจ (Business Architecture) คือการให้เห็นความต้องการกับระบบสารสนเทศในการสนับสนุนทางธุรกิจ

2.3.2 สถาปัตยกรรมงานประยุกต์ (Application Architecture) กรอบที่มุ่งเน้นการพัฒนาในองค์กรธุรกิจ

2.3.3 สถาปัตยกรรมสารสนเทศ (Information Architecture) อธิบายตัวแบบของสารสนเทศ

2.3.4 สถาปัตยกรรมเชิงเทคนิค (Technical Architecture) คือ วิธีการรูปแบบพื้นฐานของการสนับสนุนโปรแกรมประยุกต์ ได้ระบุอยู่ในระดับที่ 3 ของภาพรวม

2.3.5 สถาปัตยกรรมผลิตภัณฑ์ (Product Architecture) คือส่วนหนึ่งของส่วน สถาปัตยกรรมเชิงเทคนิค ที่กำหนดความเป็นมาตรฐานของเทคโนโลยีและผลิตภัณฑ์ต่างๆนั้นเอง

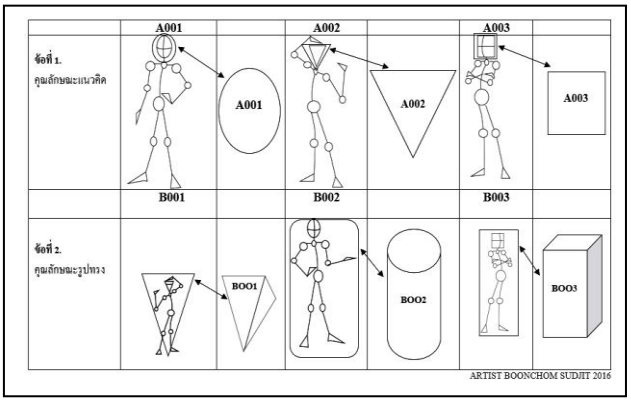
งานวิจัยชิ้นนี้ จึงต้องการศึกษาการสร้างสถาปัตยกรรมองค์กรโดยใช้ Zachman Framework เพื่อเป็นฐานความคิดสำหรับคุณลักษณะสำคัญและผลกระทบของสื่อดิจิทัลใหม่ที่มีผลต่อกระบวนการศึกษาไทยหรือองค์กรเพื่อหามิติที่สำคัญให้เกิดความชัดเจนยิ่งขึ้นสูงการพัฒนาในอนาคต

2.4 งานวิจัยที่เกี่ยวข้อง

งานวิจัยของ ลลิตา วงแหวน [1] “ระบบจำแนกระดับอารมณ์สำหรับข้อความภาษาไทย” ที่พัฒนาระบบสำหรับวิเคราะห์อารมณ์จาก ข้อความตัวอักษรภาษาไทยที่ปรากฏบนสื่อสังคมออนไลน์ ระบบจะวิเคราะห์ข้อความภาษาไทย และสามารถพิจารณา แนวโน้มของอารมณ์ในเชิงบวกหรือเชิงลบ โดยจำแนกออกมา เป็น 3 กลุ่มอารมณ์ 11 ระดับอารมณ์ ซึ่งระบบดังกล่าวสามารถ ที่จะนำไปใช้เพื่อวิเคราะห์ความเสี่ยงของโรคทางจิตวิทยา ได้แก่ โรคซึมเศร้าและโรคอารมณ์สองขั้ว

งานวิจัยของบุญชม สุดจิตต์ และประสงค์ ปราณีตพลกรัง [2] นำเสนอกรอบการออกแบบตัวการ์ตูนสำหรับสนับสนุนในการพัฒนาสื่อดิจิทัล ทั้งนี้เพื่อสร้างกรอบการออกแบบตัวการ์ตูนสำหรับสนับสนุนในการพัฒนาสื่อดิจิทัล โดยใช้เทคนิคของเฟรมเวิร์กที่พัฒนาบนพื้นฐานตามคุณลักษณะ 3 ประการ ได้แก่ ด้านแนวคิด ด้านรูปทรง และ ด้านขนาด

ผลการประเมินคุณลักษณะสำคัญของกรอบการออกแบบตัวการ์ตูนจากผู้เชี่ยวชาญ มีประเด็นด้านแนวคิดของกรอบการออกแบบตัวการ์ตูนที่ได้รับการเลือกมากถึงร้อยละ 90 ตามด้วยด้านรูปทรงได้รับการเลือกร้อยละ 10 ในขณะที่ประเด็นด้านขนาดไม่ได้รับการเลือก เมื่อนำคุณลักษณะสำคัญของกรอบการออกแบบตัวการ์ตูนมาให้กับกลุ่มผู้ชมภาพยนตร์แอนิเมชันสื่อดิจิทัล จำนวน 50 คน พบว่าให้ลักษณะความสำคัญในระดับมากที่สุดได้แก่ คุณลักษณะด้านแนวคิด และ คุณลักษณะด้านรูปทรง ในขณะที่ คุณลักษณะด้านขนาดนั้น ผู้ตอบแบบสอบถามให้น้ำหนักค่าเฉลี่ยอยู่ในระดับปานกลาง ดังรูปภาพที่ 3



ภาพที่ 3: ตัวอย่างการออกแบบตัวการ์ตูนแอนิเมชัน

E. Christofides, A. Muise, and S. Desmarais [7] “Information Disclosure and Control on Facebook: Are They Two Sides of the Same Coin or Two Different Processes?,” ได้ทำการวิจัย กลุ่มวัยรุ่นที่มีระหว่าง 18 -24 ปี ถือได้ว่าเป็นกลุ่มเสี่ยงจากภัยสังคมออนไลน์และโรคที่เกิดจากการใช้ระบบแบบกล้มหน้าส่งผลกระทบโดยตรงทั้งทางด้านสังคมการติดต่อสื่อสารในชีวิตประจำวัน ทำให้ไม่สนใจสิ่งรอบข้าง ขาดทักษะทางสังคม ทางอารมณ์ คนที่อยู่ใกล้ๆ มักจะรู้สึกหงุดหงิดเพราะมัวแต่สนใจสิ่งต่างๆ บนสังคมออนไลน์ ผลกระทบทางด้านร่างกาย อาจทำให้เกิดโรค เช่น ปวดบริเวณข้อมือ คอ รู้สึกเหนื่อย ปวดตา รวมไปถึงผลกระทบในกิจกรรมร่วมภายในบ้านหรือครอบครัวอีกด้วย

งานวิจัย Georgeta Ciobanu, Adina Palea, Annamaria Kilyeni [8] ได้นำเสนอ “The New Digital Media in Educating Computer Engineering Students for in-Company Communication” งานวิจัยนี้ต้องการสำรวจประโยชน์ของการใช้เทคโนโลยีดิจิทัลใหม่ในการให้ความรู้แก่นักศึกษาสาขาวิชาวิศวกรรมคอมพิวเตอร์ เพื่อพัฒนาทักษะการสื่อสารในระดับมืออาชีพ โดยใช้เทคนิคพื้นฐานการติดต่อสื่อสารในปัจจุบัน กำหนดประโยชน์สร้างกรอบการทำงานสำหรับสนับสนุนวัตถุประสงค์ เพื่อทำการศึกษาวิเคราะห์นำมาประยุกต์ใช้เทคโนโลยีใหม่ๆ ของหลักสูตรกรณีศึกษาเป็นหลักสูตรการสื่อสารสำหรับนักศึกษาชั้นปีที่ 2 ที่มหาวิทยาลัย Politehnica ของ Timisoara ทำให้นักศึกษาสามารถนำเทคโนโลยีสารสนเทศมาประยุกต์ใช้ในการศึกษาค้นคว้าเพื่อองค์ความรู้จาก Google Sites และประเมินได้อย่างมีประสิทธิภาพอีกด้วย ทำให้ผู้วิจัยเกิดแนวคิดของการบูรณาการด้านเทคโนโลยีสารสนเทศของกรอบแนวคิดด้านการศึกษาการพัฒนาในคุณลักษณะที่สำคัญเพื่อนำมาประยุกต์ใช้

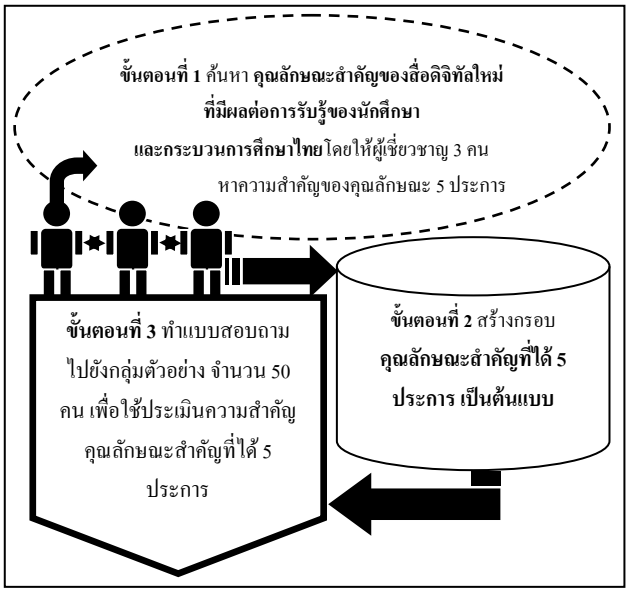
งานวิจัย Lu Geng [11] ได้นำเสนอ “Teaching Exploration and Reform of Program Design Course for Digital Media Art Students” งานวิจัยนี้มี วัตถุประสงค์ เพื่อสร้างภาพเคลื่อนไหวจากมุมมองของการออกแบบระบบการออกแบบแอนิเมชันแบบโต้ตอบเกี่ยวข้องกับการออกแบบตัวละครฉากการออกแบบโหมดโต้ตอบกฎของเกมการออกแบบในแต่ละระดับ ผลที่ได้คือ นักเรียนสามารถเรียนรู้โดยการทำ โครงงานเหล่านี้กระตุ้นความคิดริเริ่มการเรียนรู้ของนักเรียนและเข้าใจ

กระบวนการแก้ปัญหาของโครงงานจากประสบการณ์ การประเมินผลหลักสูตรคือการใช้ความรู้ที่ได้เรียนรู้เพื่อทำโครงการภาพเคลื่อนไหวแบบโต้ตอบที่สมบูรณ์แบบในกลุ่ม กลุ่มจะต้องส่งเอกสารการออกแบบขั้นสุดท้ายซอร์สโค้ดไฟล์ที่ปฏิบัติการได้ คะแนนเป็นตัวชี้วัด ด้วยประเด็นนี้ ผู้วิจัยเกิดแนวคิดในการสร้างแรงดึงดูดความสนใจ และการสร้างสรรค์ผลงานในเชิงความคิดแบบบูรณาการกรอบศิลปะกับเทคโนโลยีสารสนเทศในอนาคต [12], [13], [14], [15], [16], [17]

3. วิธีดำเนินการวิจัย

3.1 ขั้นตอนการวิจัย

การศึกษาและวางแผนดำเนินการวิจัย ผู้วิจัยได้ใช้การวิจัยเชิงคุณภาพ (Qualitative Research) โดยการสัมภาษณ์เจาะลึก (In-depth Interview) ทำการบันทึกในการเก็บข้อมูล โดยมีระเบียบวิธีวิจัยตามแนวทางการวิจัยดังภาพที่ 4



ภาพที่ 4 : ขั้นตอนและวิธีดำเนินการวิจัย

3.2 กำหนดประชากรและกลุ่มตัวอย่าง

ในงานวิจัยนี้ได้กำหนดประชากรและกลุ่มตัวอย่าง ประกอบด้วย ผู้เชี่ยวชาญด้านสื่อโฆษณาดิจิทัล ผู้เชี่ยวชาญด้านการศึกษา และผู้เชี่ยวชาญด้านสื่อสารมวลชน และกลุ่มวัยรุ่นในระบบการศึกษาที่ใช้สื่อดิจิทัล จำนวน 50 คน

3.3 เครื่องมือที่ใช้ในการวิจัย

1. แบบสัมภาษณ์ (Interview)

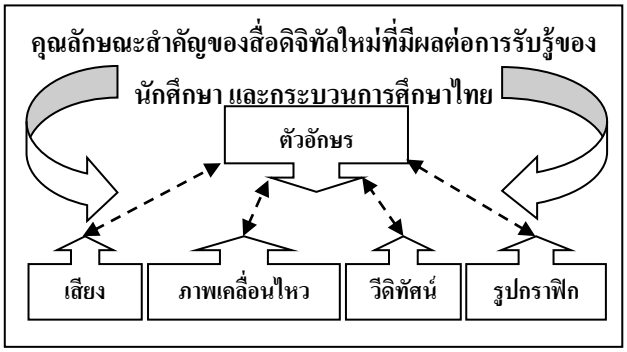
ผู้วิจัยนำแบบสอบถามและแบบสัมภาษณ์ปลายเปิดที่ได้ทำการปรับปรุงแล้วไปเก็บรวบรวมข้อมูล ดังนี้

1.1 ผู้วิจัยทำการนัดหมายผู้เชี่ยวชาญทั้ง 3 ท่าน เพื่อทำการเก็บข้อมูลโดยใช้เครื่องมือที่ผู้วิจัยสร้างขึ้นมา

1.2 ผู้วิจัยทำการเก็บข้อมูลด้วยตัวเองโดยการสัมภาษณ์ในคุณลักษณะทั้ง 5 ข้อจากผู้เชี่ยวชาญด้วยคำถามแบบมีโครงสร้าง

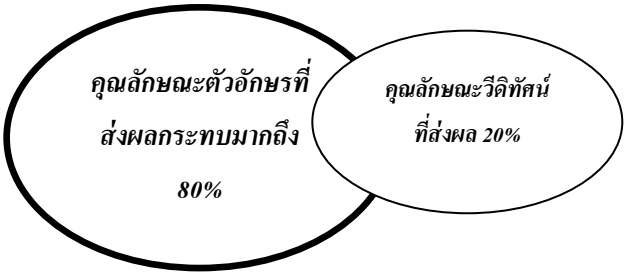
3.4 วิธีการเก็บรวบรวมข้อมูล

ผู้วิจัยส่งแบบสอบถามคุณลักษณะสำคัญของสื่อดิจิทัลใหม่ที่มีผลต่อการรับรู้ของนักศึกษาให้กลุ่มตัวอย่างแบบเจาะจง จำนวน 50 คน ได้แก่คุณลักษณะดังภาพที่ 5



ภาพที่ 5: คุณลักษณะสำคัญของสื่อดิจิทัลใหม่ที่มีผลต่อการรับรู้ของนักศึกษาและกระบวนการศึกษาไทย

โดยที่ผู้วิจัยได้ทำการรวบรวมแบบสอบถามด้วยตนเอง ทั้งนี้เพื่อให้ทราบว่าคุณลักษณะสำคัญของสื่อดิจิทัลใหม่ที่มีผลต่อการรับรู้ของนักศึกษาและกระบวนการศึกษาไทย ใดบ้างที่มีความสำคัญมากที่สุด ดังแสดงรายละเอียดไว้ในภาพที่ 6



ภาพที่ 6: คุณลักษณะสำคัญของสื่อดิจิทัลใหม่ที่มีผลต่อการรับรู้ของนักศึกษาและกระบวนการศึกษาไทย 2 อันดับแรกที่สำคัญ

3.5 การวิเคราะห์ข้อมูล

ผู้วิจัยทำการวิเคราะห์กลุ่มตัวอย่างจำนวน 50 คน โดยใช้โปรแกรมสำเร็จรูปทางสถิติ ทำการวิเคราะห์ข้อมูลและใช้สถิติเชิงพรรณา ได้แก่ ค่าร้อยละ ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน ผลการวิเคราะห์ข้อมูลได้แสดงไว้ในตารางที่ 1

จากตารางที่ 1 พบว่าคุณลักษณะสำคัญของสื่อดิจิทัลใหม่ที่มีผลต่อการรับรู้ของนักศึกษาและกระบวนการศึกษาไทย ในอันดับที่ 1 ได้แก่ คุณลักษณะตัวอักษร โดยมีค่าเฉลี่ยอยู่ที่ระดับมากที่สุด ($\bar{X}$ = 4.80) ด้วยสาเหตุการพิมพ์ข้อความที่มีการส่งออกและรับข้อความอีกทั้งมีความเร็วต่อสื่อดิจิทัลใหม่ ส่วนอันดับที่ 2 คือ คุณลักษณะวีดิทัศน์โดยมีค่าเฉลี่ยอยู่ในระดับมากที่สุด ($\bar{X}$ = 4.64) ซึ่งส่งผลต่อความเร็วของการดาวน์โหลดและอัพโหลด ทำให้ช้ากว่าตัวอักษรต่อการรับรู้เช่นกัน

ตารางที่ 1: ผลการวิเคราะห์คุณลักษณะสำคัญของสื่อดิจิทัลใหม่ที่มีผลต่อการรับรู้ของนักศึกษาและกระบวนการศึกษาไทย

คุณลักษณะ	$\bar{x}$	S.D.	ลักษณะความสำคัญ
1. ตัวอักษร	4.80	0.47	มากที่สุด
2. วีดิทัศน์	4.64	0.41	มากที่สุด
3. รูปกราฟิก	2.80	0.57	ปานกลาง
รวม	**4.08**	**0.08**	**มาก**

ในขณะที่ คุณลักษณะรูปกราฟิก ผู้ตอบแบบสอบถามให้ค่าเฉลี่ยลักษณะความสำคัญอยู่ในระดับปานกลาง ($\bar{X}$ = 2.80)

4. ผลการวิจัย

ผลการวิจัย พบว่าคุณลักษณะสำคัญของสื่อดิจิทัลใหม่ที่มีผลต่อการรับรู้ของนักศึกษาและกระบวนการศึกษาไทย จากผู้เชี่ยวชาญ มีประเด็นคุณลักษณะตัวอักษรที่ได้รับการเลือกมากถึงร้อยละ 80 ส่วนในคุณลักษณะวีดิทัศน์ได้รับการเลือกมากถึงร้อยละ 20 และในประเด็นได้รับการเลือกสุดท้าย เมื่อนำคุณลักษณะสำคัญของสื่อดิจิทัลใหม่ กับกลุ่มตัวอย่างแบบเจาะจง จำนวน 50 คน เป็นเพศชาย จำนวน 35 คน คิดเป็นร้อยละ 70 และเพศหญิงจำนวน 15 คน คิดเป็นร้อยละ 30 และวิเคราะห์ กลุ่มตัวอย่างให้ความสำคัญอันดับที่ 1 ได้แก่ คุณลักษณะตัวอักษร ระดับมากที่สุด รองลงมาคือ คุณลักษณะวีดิทัศน์ และให้ความสำคัญในระดับปานกลาง ได้แก่ คุณลักษณะรูปกราฟิก

ผลจากการค้นในครั้งนี้ สามารถนำมาเป็นกรอบแนวคิดในการป้องกัน คุณลักษณะสำคัญของสื่อดิจิทัลใหม่ที่มีผลต่อการรับรู้ของนักศึกษาและกระบวนการศึกษาไทยการแยกแยะในขบวนการคิดกรองข่าวสาร และเพื่อทราบถึงช่องทางเทคโนโลยีสารสนเทศในสื่อดิจิทัลมาสู่กลุ่มวัยรุ่นไทยให้เกิดความชัดเจนยิ่งขึ้นเพื่อการพัฒนาและกรองข้อมูลข่าวสาร

เมื่อวิเคราะห์คุณลักษณะที่มีความสำคัญในแต่ละมิติของคุณลักษณะ เช่น คุณลักษณะตัวอักษร และวีดิทัศน์ นั้นได้สร้างความรวดเร็ว ในช่องทางการตลาดและด้านการศึกษาบนสื่อดิจิทัลที่มีความน่าสนใจ ถือได้ว่าเป็นสิ่งใหม่

การพัฒนากระบวนการสื่อดิจิทัลต้องให้ความสำคัญในคุณลักษณะตัวอักษรมากกว่าคุณลักษณะอื่น ๆ ทั้งนี้เนื่องจากผลการวิเคราะห์ที่ได้จากความคิดเห็นของผู้เชี่ยวชาญ รวมทั้ง จากกลุ่มตัวอย่างที่รับสื่อดิจิทัล จำนวน 50 คน พบว่า ตรงกันในระดับมากสำหรับการให้ความสำคัญในคุณลักษณะตัวอักษร ในขณะที่ ในภาพรวมของคุณลักษณะโดยรวมมีค่าเฉลี่ยอยู่ในระดับมาก ($\bar{X}$ = 4.08) และมีค่าเบี่ยงเบนมาตรฐาน 0.08 แสดงถึงความสำคัญของตัวแปรหรือคุณลักษณะสำคัญและผลกระทบของสื่อดิจิทัลใหม่ที่มีผลต่อกระบวนการศึกษาไทย ตามประเด็นที่วิจัยได้นั้นมีความสำคัญอยู่ในระดับมาก

5. การอภิปรายผล

คุณลักษณะสำคัญของสื่อดิจิทัลใหม่ที่มีผลต่อการรับรู้ของนักศึกษาและกระบวนการศึกษาไทย ในงานวิจัยนี้ทำให้ผู้วิจัยทราบถึงตัวแปรที่ส่งผลต่อคุณลักษณะที่สำคัญ เพื่อนำมาใช้ในกระบวนการศึกษาไทยในคุณลักษณะสำคัญของสื่อดิจิทัลใหม่ 2 ประการ คุณลักษณะข้อที่ 1. ตัวอักษร กลุ่มผู้ใช้ ได้ให้ค่าระดับความสำคัญมากที่สุด ส่วนคุณลักษณะ ข้อที่ 2. วีดิทัศน์ กลุ่มผู้ใช้ ได้ให้ค่าระดับความสำคัญมากที่สุด เช่นกัน และคุณลักษณะข้อที่ 3. รูปกราฟิก ให้ความสำคัญในระดับปานกลาง ทำให้ทราบได้ว่าในคุณลักษณะสำคัญของสื่อดิจิทัลใหม่ที่มีผลต่อการรับรู้ของนักศึกษาและกระบวนการศึกษาไทยนั้น

ในกระบวนการของการศึกษาควรใช้คุณลักษณะสำคัญที่ค้นพบขึ้นใหม่นี้ นำไปใช้สำหรับสนับสนุนพัฒนาสื่อดิจิทัลใหม่ทางด้านการศึกษา และสร้างระบบป้องกันเพื่อให้เกิดประโยชน์เผยแพร่พร้อมทั้งการควบคุมจะทำได้ดีขึ้นในอนาคต

6. เอกสารอ้างอิง

[1] ลลิตา วงแหวน, ณัฐวรา คีรีเดช , ศิมาภรณ์ พันธุ์เผือก ,และณัฐพงศ์ ทองเทพ (2559). “ระบบจำแนกระดับอารมณ์สำหรับข้อความภาษาไทย”การประชุมวิชาการ ครั้งที่ 12 ด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศระดับชาติ

[2] บุญชม สุดจิตต์,ประสงค์ ปราณีตพลกรัง, “กรอบการออกแบบตัวการ์ตูนสำหรับสนับสนุนในการพัฒนาสื่อดิจิทัล”, การประชุมทางวิชาการระดับชาติด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศ ครั้งที่ 12, ณ โรงแรมเซ็นทาราและคอนเวนชันเซ็นเตอร์ ขอนแก่น, จังหวัดขอนแก่น,2559.

[3] น้าชาติ ประชาชื่น, “ปฏิบัติการข่าวสาร: IO”, ปีที่ 23 ฉบับที่ 8238 ข่าวสดรายวัน, 19 มิถุนายน พ.ศ. 2556

[4] Carla, M, A Method to Define an Enterprise Architecture using the Zachman Framework. Nicosia, Cyprus.2004.

[5] John A. Zachman, “A framework for information systems architecture,” IBM Systems Journal 26, No. 3, 276-292,1987.

[6] J. F. Sowa, and J.A. Zachman, “Extending and formalizing the framework for information systems architecture,” IBM Systems Journal 31, No. 3, 590-616, 1992.

[7] E. Christofides, A. Muise, and S. Desmarais, (2009), “Information Disclosure and Control on Facebook: Are They Two Sides of the Same Coin or Two Different Processes?,” CyberPsychology & Behavior, vol. 12, no. 3, pp. 341–345, Jun. 2009.

[8] Georgeta Ciobanu, Adina Palea, Annamaria Kilyeni, (2011), “The New Digital Media in Educating Computer Engineering Students for in-Company Communication,” International Conference on Computer as a Tool (EUROCON), IEEE 2011

[9] Ji T.-L., Sundareshan M. K., and Roehrig H., “Adaptive Image Contrast Enhancement Based on Human Visual”, IEEE Transactions on Medical Imaging, 13(4):573-586, 1994.

[10] H. J. Scholl, M. B. Eisenberg, L. Dirks, and Timothy S Carlson, “The TEDS framework for assessing information systems from a human actors’ perspective: Extending and repurposing Taylor's Value-Added Model,” Journal of the American Society for Information Science and Technology, vol. 64, no. 4, pp. 789–804, 2011.

[11] Lu Geng, (2015), “Teaching Exploration and Reform of Program Design Course for Digital Media Art Students,” The 10th International Conference on Computer Science & Education (ICCSE 2015) July 22-24, Fitzwilliam College, Cambridge University, UK 2015.

[12] D. Liu Yang, "The Sense of the Originality and the Cartoon Characters in the Art of Television Animation", Shangdong Normal University, 2008.

[13] M. Gao Yang, "Characters in the Cartoon and the Modern Design", Fujian Arts Publishing House 2007.

[14] Mahmoud AbuShawali1, Yan Peng Lim2, Ahmad Radzi Bedu3, "The Impact of New Digital Media on Graphic Designer in Jordan", International Conference on Informatics and Creative Multimedia, 2013.

[15] Solvita Zarina, "Case Studies of Digital and Media Art", 20th International Conference Information Visualisatio, 2016.

[16] Indrawati1and Nadia Primasari2, "Digital Advertising Media Adoption in Consumer Goods Industry"(An Indonesian Perspective), Fourth International Conference on Information and Communication Technologies (ICoICT) ,2016.

[17] Zhou Feng, "Study of New Digital Media Arts", Computer-Aided Industrial Design & Conceptual Design, CAID & CD. 10th International Conference on, IEEE 2009.

การเพิ่มประสิทธิผลการเรียนรู้ของเจนเนอเรชันแซดผ่านการเรียนรู้แบบไมโคร
Learning Effectiveness for Generation Z using Micro Learning

ธารีย์ฉิชา ลีพีรวิทิต (T. Leepeeravitit)[1] กว้าน สีตะธนี (K. Sitathani)[2] และวิเชียร ชุติมาสกุล (W. Chutimaskul)[3]
กลุ่มวิจัยวิศวกรรมความต้องการ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี
[1]57130700331@st.sit.kmutt.ac.th, [2]skwanid@gmail.com, [3]wichian@sit.kmutt.ac.th

บทคัดย่อ

เจนเนอเรชันแซดเป็นกลุ่มที่เติบโตมาพร้อมกับเทคโนโลยีสารสนเทศ ทำให้เลี่ยงการเรียนรู้แบบการบรรยายเดิมๆ การเรียนรู้แบบไมโครจึงเข้ามามีบทบาทสำคัญ งานวิจัยนี้ได้นำการเรียนรู้แบบไมโครมาใช้ในการเรียนรู้ของเจนเนอเรชันแซด เพื่อเพิ่มประสิทธิผลการเรียนรู้ โดยกลุ่มที่หนึ่งเป็นการเรียนรู้แบบไมโคร กลุ่มที่สองเป็นการเรียนรู้แบบดั้งเดิม ผลการวิจัยที่ได้ คือ กลุ่มตัวอย่างที่ทำการเรียนรู้แบบไมโครมีประสิทธิผลการเรียนรู้ที่สูงกว่ากลุ่มตัวอย่างที่ทำการเรียนรู้แบบดั้งเดิม อย่างมีนัยสำคัญทางสถิติที่ 0.05 และผลความพึงพอใจในการเรียนรู้แบบไมโครอยู่ในระดับมากที่สุด ซึ่งแสดงให้เห็นว่ารูปแบบและปัจจัยในการเรียนรู้แบบไมโครสามารถตอบสนองและตรงต่อความต้องการในการเรียนรู้ของผู้เรียนในเจนเนอเรชันแซด

คำสำคัญ: การเพิ่มประสิทธิผลการเรียนรู้ การเรียนรู้แบบดั้งเดิม การเรียนรู้แบบไมโคร เจนเนอเรชันแซด

Abstract

Generation Z is grown up with information technology. They tend to avoid a traditional teaching and learning method. Micro learning plays a significant role. This research therefore used Micro learning on a Generation Z to increase their learning effectiveness. The former with the Micro learning method and the latter with the Traditional learning method. The result revealed that the former performed more effectively in learning than the latter at the significance of 0.05 and the satisfaction on the Micro learning method was at the highest. Therefore, the Micro learning method supported and met the needs of Generation Z.

Keyword: Generation Z, Learning Effectiveness, Micro Learning, Traditional Learning.

1. บทนำ

ปัจจุบันเทคโนโลยีสารสนเทศกลายเป็นเครื่องมือสำคัญในการขับเคลื่อนความเจริญของโลกทุกด้าน รวมถึงในด้านการศึกษา ซึ่งการศึกษานับว่าเป็นปัจจัยสำคัญที่มีส่วนช่วยในการพัฒนาสังคมและประเทศ [1] จึงจำเป็นต้องมีการปรับตัวและเปลี่ยนแปลงจากการศึกษาแบบดั้งเดิมหรือที่เรารู้จักกันในรูปแบบการบรรยาย การศึกษาในศตวรรษที่ 21 จึงได้นำเทคโนโลยีสารสนเทศมาใช้ในการพัฒนาการศึกษา ประกอบกับกลุ่มผู้เรียนที่เปลี่ยนเจนเนอเรชัน

เจนเนอเรชันแซดเป็นกลุ่มคนที่อยู่ในวัยเรียนในปัจจุบัน ซึ่งเติบโตมาพร้อมกับเทคโนโลยีสารสนเทศที่ทันสมัยประกอบกับช่วงความสนใจอย่างต่อเนื่องของคนในปัจจุบันมีจำกัด [2],[3] ทำให้คนกลุ่มนี้หลีกเลี่ยงการเรียนรู้แบบดั้งเดิม ดังนั้นจึงได้มีการเสนอการเรียนรู้แบบไมโครขึ้น

การเรียนรู้แบบไมโครเป็นการเรียนรู้ที่ไม่เป็นทางการ ใช้เวลาในการเรียนรู้เพียงไม่เกิน 15 นาที สามารถเรียนรู้ผ่านอุปกรณ์อิเล็กทรอนิกส์ เช่น สมาร์ตโฟน แท็ปเล็ต ซึ่งสามารถพกพาได้สะดวก ทำให้การเรียนรู้สามารถทำได้ทุกที่ทุกเวลา [4] ด้วยรูปแบบและลักษณะของการเรียนรู้แบบไมโครสามารถตอบสนองและตรงต่อความต้องการของผู้เรียนที่อยู่ในกลุ่ม

เจนเนอเรชันแซดได้ นับว่าการเรียนรู้แบบไมโครเป็นการเรียนรู้ที่ช่วยสนับสนุนการเรียนรู้ในระยะยาวหรือการเรียนรู้ตลอดชีวิต

2. ทฤษฎี และงานวิจัยที่เกี่ยวข้อง

2.1 เจนเนอเรชัน (Generation)

เจนเนอเรชันมีแนวคิดที่ว่า กลุ่มคนที่มีอายุใกล้เคียงกันจะเติบโตมาในสภาพแวดล้อมที่คล้ายกัน ทำให้มีทัศนคติและลักษณะนิสัยที่เป็นเฉพาะตัวของคนรุ่นนั้น โดยเจนเนอเรชันจะแบ่งตามปีเกิด [5],[6] โดยมีรายละเอียดดังต่อไปนี้

เจนเนอเรชันวาย (Generation Y)

เจนเนอเรชันวายเป็นกลุ่มประชากรที่เกิดในช่วง ค.ศ. 1982 - 1994 [7] โดยลักษณะนิสัย เป็นคนที่ไม่ชอบทำตามกฎระเบียบ ความคาดหวังสูง เปิดใจยอมรับและปรับตัวให้เข้ากับสถานการณ์ที่มีความแตกต่าง [5]

เจนเนอเรชันแซด (Generation Z)

เจนเนอเรชันแซดเป็นกลุ่มประชากรที่เกิดตั้งแต่ปี ค.ศ. 1995 เป็นต้นไป [8] คนกลุ่มนี้เติบโตมากับเทคโนโลยีสารสนเทศที่ทันสมัย โดยลักษณะนิสัยเป็นคนชอบความสะดวกสบาย ไม่ชอบพิธีการ สามารถทำอะไรหลาย ๆ อย่างได้ในเวลาเดียวกัน ชอบความท้าทาย เป็นคนที่มีสมาธิสั้น ยอมรับการเปลี่ยนแปลงใหม่ ๆ โดยในการสื่อสารระหว่างคนกลุ่มนี้ส่วนใหญ่เป็นการสื่อสารผ่านข้อความบนหน้าจอมือถือแทนการพูด และมีแนวทางเป็นของตัวเองชัดเจน [9]

2.2 รูปแบบของการเรียนรู้

รูปแบบของการเรียนรู้ที่นำมาใช้ในการวิจัยนี้ มีดังนี้

การเรียนรู้แบบดั้งเดิม (Traditional Learning)

การเรียนรู้แบบดั้งเดิม คือ การเรียนรู้ในรูปแบบของการบรรยาย โดยมีครูเป็นศูนย์กลางการเรียนรู้ ผู้เรียนเรียนกันเป็นกลุ่มใหญ่ มีชั่วโมงหรือเวลาในการเรียนการสอนถูกกำหนดอย่างชัดเจนตามหลักสูตรที่กำหนด มีตัวชี้วัดคือข้อสอบและผลการเรียน การเรียนมุ่งเน้นผลการสอบเป็นหลัก การวัดผลการให้คะแนนและการตัดเกรดสามารถทำได้ง่าย การเรียนรู้เกิดขึ้นในห้องเรียน/สถานศึกษา [10],[11]

การเรียนรู้แบบไมโคร (Micro Learning)

การเรียนรู้แบบไมโครเป็นการเรียนรู้ที่ไม่เป็นทางการ ด้วยลักษณะของการเรียนรู้แบบไมโคร ได้แก่ ผู้เรียนใช้เวลาสั้น ๆ ในการเรียนรู้ (ไม่เกิน 15 นาที), เนื้อหามีขนาดเล็ก โดยมีหัวข้อการเรียนรู้ที่เป็นหัวข้อย่อยหรือเป็นตอน ๆ โดยแต่ละตอนมีขนาดที่จำกัด, สามารถเรียนรู้ได้ตรงประเด็นตามที่ผู้เรียนต้องการ, รูปแบบในการเรียนรู้มีความหลากหลาย ได้แก่ เกม วิดีโอ [12],[13],[14] ซึ่งเป็นวิธีการเรียนรู้ที่สามารถช่วยแก้ปัญหาในเรื่องของความสนใจของผู้เรียนที่มีน้อยลงได้ เนื่องจากคนเรามีช่วงความสนใจที่จำกัด [2],[3] โดยการเรียนรู้แบบไมโครสามารถเข้าถึงได้ง่ายผ่านทางอุปกรณ์ต่าง ๆ เช่น สมาร์ตโฟน แท็บเล็ต ทำให้ผู้เรียนสามารถเรียนรู้ได้ทุกที่ทุกเวลา นอกจากนี้ผู้เรียนยังสามารถจัดการ/วางแผนในการเรียนรู้ได้ง่าย เรียนซ้ำได้บ่อยครั้งเท่าที่ต้องการ และสามารถเรียนรู้ได้ในเวลาที่จำกัด เช่น เวลารอรถโดยสาร เวลาพักกลางวัน การเรียนรู้แบบไมโครช่วยให้การเรียนรู้สามารถทำได้อย่างง่ายดาย นอกจากจะส่งผลให้การเรียนรู้ทำได้อย่างรวดเร็ว มีความสนุกสนาน ยังทำให้ผู้เรียนได้รับประสบการณ์จากการเรียนรู้ [4],[15],[16],[17],[18] การเรียนรู้แบบไมโครเหมาะสำหรับการเรียนรู้ของผู้เรียนในยุคปัจจุบันที่มีเทคโนโลยีสารสนเทศมาสนับสนุน ซึ่งจะช่วยส่งเสริมการเรียนรู้ในระยะยาวหรือการเรียนรู้ตลอดชีวิต

2.3 งานวิจัยที่เกี่ยวข้อง

ด้วยคุณลักษณะและข้อดีของการเรียนรู้แบบไมโครพบว่างานวิจัยหลาย ๆ งาน นำการเรียนรู้แบบไมโครมาใช้ในการเรียนรู้มากขึ้น โดยจากงานวิจัย [19] ได้ออกแบบเกมในการเรียนรู้ภาษาอาหรับ มีแพลตฟอร์มหนึ่งในคุณสมบัติของการเรียนรู้แบบไมโครมาใช้ในการออกแบบและพัฒนาเกม พบว่า 92 % ของกลุ่มตัวอย่างมีความพึงพอใจและมีความต้องการที่จะใช้งานต่อ นอกจากนี้ยังได้นำการเรียนรู้แบบไมโครไปใช้ในการอบรมในบริษัทต่าง ๆ เพื่อพัฒนาบุคลากรในองค์กรให้มีประสิทธิภาพมากขึ้น โดยได้มีการทำสื่อการอบรมโดยใช้เว็บ 2.0 สร้างแพลตฟอร์มการฝึกอบรมแบบไม่เป็นทางการที่ใช้เวลาในการเรียนรู้เพียงสั้น ๆ สามารถเรียนรู้ที่ไหนก็ได้ ทำให้พนักงานสามารถพูดคุย มีการแบ่งปันความรู้ ความคิด

ค้นหาข้อมูล จากเพื่อนร่วมงานหรือคนอื่น ที่อยู่รอบ ๆ ซึ่งองค์กรสามารถใช้ประโยชน์จากทรัพยากรที่มีอยู่ในการสร้างสถานที่ฝึกอบรมแบบไม่เป็นทางการเพื่อส่งเสริมการเรียนรู้ด้วยตนเองและการเรียนรู้ทางสังคมของพนักงาน ผลที่ได้ คือ ช่วยลดค่าใช้จ่ายในการอบรมพนักงาน [20]

3. วิธีดำเนินการวิจัย

การวิจัยครั้งนี้ มีวิธีการดำเนินการวิจัยตามขั้นตอนดังนี้

3.1 ประชากรและกลุ่มตัวอย่าง

การวิจัยครั้งนี้ได้กำหนดประชากรและกลุ่มตัวอย่างดังนี้

ประชากร คือ กลุ่มคนที่อยู่ในเจนเนอเรชันแซด มีอายุระหว่าง 7 – 22 ปี ในประเทศไทย

กลุ่มตัวอย่างเป็นนักเรียนระดับชั้นมัธยมศึกษาชั้นปีที่ 3 กรณีศึกษาโรงเรียนเบญจมราชาลัย ในพระบรมราชูปถัมภ์ (เพศหญิง) จำนวน 120 คน เป็นการสุ่มตัวอย่างแบบเจาะจง

3.2 เครื่องมือที่ใช้ในงานวิจัย

เครื่องมือที่ใช้ในการศึกษาวิจัยครั้งนี้ แบ่งออกเป็น

การเรียนรู้และเนื้อหาที่ใช้สำหรับทดสอบหาประสิทธิผลการเรียนรู้ที่เพิ่มขึ้นของเจนเนอเรชันแซด ได้แก่

การเรียนรู้แบบไมโคร เป็นการเรียนรู้ในรูปแบบสื่อประสม (Multimedia) ใช้เวลาในการเรียนรู้ไม่เกิน 10 นาที กลุ่มตัวอย่างทำการเรียนรู้ผ่านอุปกรณ์อิเล็กทรอนิกส์ คือ สมาร์ตโฟน (1 คนต่อ 1 เครื่อง)

การเรียนรู้แบบดั้งเดิมเป็นการเรียนรู้แบบบรรยาย ในห้องเรียน โดยมีครูเป็นผู้ที่ให้ความรู้ และกลุ่มตัวอย่างจะเป็นผู้เรียน (ครู 1 คนต่อนักเรียนหลายคน) โดยจะเรียนจนจบครบตามเนื้อหาที่กำหนด (ไม่กำหนดเวลาที่ใช้ในการเรียน)

เนื้อหาและหัวข้อที่ใช้ในการสร้างบทเรียนสำหรับการเรียนรู้ในงานวิจัยนี้ คือ วิชาวิทยาศาสตร์ เรื่อง อิเล็กทรอนิกส์เบื้องต้น

แบบทดสอบเพื่อหาผลสัมฤทธิ์ทางการเรียน (ก่อน/หลังทำการเรียนรู้) เป็นแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก จำนวน 20 ข้อ การให้คะแนนการทำแบบทดสอบ คือ ตอบถูกให้ 1 คะแนน ตอบผิด/ไม่ตอบ/ตอบมากกว่า 1 ข้อให้ 0 คะแนน

แบบประเมินความพึงพอใจในการเรียนรู้แบบไมโคร แบ่งออกเป็น 3 ด้าน คือ ด้านเนื้อหา ด้านสื่อ ด้านผู้เรียน

3.3 การดำเนินการเก็บรวบรวมข้อมูล

งานวิจัยนี้ ผู้วิจัยได้ดำเนินการเก็บรวบรวมข้อมูลดังต่อไปนี้

1. แบ่งกลุ่มตัวอย่าง ออกเป็น 2 กลุ่ม กลุ่มละ 20 คน คือ
   กลุ่มทดลอง ทำการเรียนรู้โดยการเรียนรู้แบบไมโคร
   กลุ่มควบคุม ทำการเรียนรู้โดยการเรียนรู้แบบดั้งเดิม
2. ผู้วิจัยอธิบายวัตถุประสงค์และวิธีดำเนินการทดลองการเรียนรู้ให้กลุ่มตัวอย่างทราบ
3. กลุ่มตัวอย่างทั้ง 2 กลุ่มทำแบบทดสอบก่อนการเรียนรู้
4. กลุ่มตัวอย่างทำการเรียนรู้ตามที่ได้แบ่งกลุ่มไว้ คือ
   กลุ่มทดลอง ทำการเรียนรู้โดยการเรียนรู้แบบไมโคร
   กลุ่มควบคุม ทำการเรียนรู้โดยการเรียนรู้แบบดั้งเดิม
   และกลุ่มตัวอย่างทั้ง 2 กลุ่มทำแบบทดสอบหลังการเรียนรู้
5. กลุ่มตัวอย่างที่ทำการเรียนรู้แบบไมโคร ทำแบบประเมินความพึงพอใจในการเรียนรู้แบบไมโคร
6. ทดลองซ้ำ (ข้อ 1 - 5) อีก 2 ครั้ง โดยเปลี่ยนกลุ่มตัวอย่าง (เป็นการสุ่มตัวอย่างแบบเจาะจง) ทดลองทั้งสิ้นรวม 3 ครั้ง
7. นำผลการทดสอบก่อน/หลังการเรียนรู้ และผลการตอบแบบประเมินความพึงพอใจในการเรียนรู้แบบไมโครลงในโปรแกรมสำเร็จรูปทางสถิติ (SPSS) เพื่อวิเคราะห์ต่อไป

4. ผลการวิจัย

การวิจัยในครั้งนี้ จำแนกผลการวิจัย ได้ดังนี้

4.1 ผลการวิเคราะห์ประสิทธิผลการเรียนรู้ก่อนและหลังทำการเรียนรู้ของกลุ่มตัวอย่างที่ทำการเรียนรู้แบบดั้งเดิม

ผลการวิเคราะห์จากการทดลองทั้ง 3 ครั้งพบว่า ประสิทธิผลทางการเรียนรู้ของกลุ่มตัวอย่างที่ทำการเรียนรู้แบบดั้งเดิม หลังทำการเรียนรู้สูงกว่าก่อนทำการเรียนรู้อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05

4.2 ผลการวิเคราะห์ประสิทธิผลการเรียนรู้ก่อนและหลังทำการเรียนรู้ของกลุ่มตัวอย่างที่ทำการเรียนรู้แบบไมโคร

ผลการวิเคราะห์จากการทดลองทั้ง 3 ครั้งพบว่า ประสิทธิผลทางการเรียนรู้ของกลุ่มตัวอย่างที่ทำการเรียนรู้แบบไมโคร หลังทำการเรียนรู้สูงกว่าก่อนทำการเรียนรู้อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.01

4.3 ประสิทธิผลการเรียนรู้ก่อนทำการเรียนรู้ (Pre - Test) ของกลุ่มตัวอย่างทั้ง 2 กลุ่ม

ผลการวิเคราะห์จากการทดลองทั้ง 3 ครั้งพบว่า ประสิทธิผลทางการเรียนรู้ของกลุ่มตัวอย่างที่ทำการเรียนรู้แบบดั้งเดิม มีประสิทธิผลการเรียนรู้สูงกว่ากลุ่มที่ทำการเรียนรู้แบบไมโครอย่างไม่มีนัยสำคัญทางสถิติ จำนวน 2 ครั้ง และ กลุ่มที่ทำการเรียนรู้แบบไมโคร มีประสิทธิผลการเรียนรู้สูงกว่ากลุ่มที่ทำการเรียนรู้แบบดั้งเดิมอย่างไม่มีนัยสำคัญทางสถิติ จำนวน 1 ครั้ง หมายถึง ความรู้พื้นฐานของกลุ่มตัวอย่างทั้ง 2 กลุ่ม มีพื้นฐานความรู้ที่เท่ากัน หรือไม่แตกต่างกัน

4.4 ประสิทธิผลการเรียนรู้หลังทำการเรียนรู้ (Post - Test) ของกลุ่มตัวอย่างทั้ง 2 กลุ่ม

ผลการวิเคราะห์แสดงได้ดังตารางต่อไปนี้

ตารางที่ 1: แสดงค่าประสิทธิผลการเรียนรู้หลังทำการเรียนรู้ของกลุ่มตัวอย่างทั้ง 2 กลุ่ม

การทดลองครั้งที่	การทดสอบ (กลุ่มที่ทำการเรียนรู้)	N	$\bar{x}$	S.D.	t	Sig.
1	แบบดั้งเดิม	20	15.00	1.892	2.208	.033*
	แบบไมโคร	20	16.25	1.682		
2	แบบดั้งเดิม	20	11.40	2.415	2.640	.012*
	แบบไมโคร	20	13.25	1.997		
3	แบบดั้งเดิม	20	10.90	2.882	2.063	.046*
	แบบไมโคร	20	12.75	2.789		

*นัยสำคัญทางสถิติที่ระดับ 0.05

จากการทดลองทั้ง 3 ครั้ง พบว่าประสิทธิผลทางการเรียนรู้หลังทำการเรียนรู้ กลุ่มที่ทำการเรียนรู้แบบไมโครมีประสิทธิผลการเรียนรู้สูงกว่ากลุ่มที่ทำการเรียนรู้แบบดั้งเดิมอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 ซึ่งสอดคล้องกับสมมติฐานที่ตั้งไว้

4.5 ผลการวิเคราะห์ความพึงพอใจในการเรียนรู้แบบไมโคร

การวิจัยนี้แบ่งคะแนนออกเป็น 5 ระดับ ดังนี้

5.00 – 4.50	หมายถึง ระดับมากที่สุด
4.49 – 3.50	หมายถึง ระดับมาก
3.49 – 2.50	หมายถึง ระดับปานกลาง
2.49 – 1.50	หมายถึง ระดับน้อย
1.49 – 1.00	หมายถึง ระดับน้อยที่สุด

ผลการวิเคราะห์พบว่า ด้านเนื้อหา ด้านสื่อการสอนและด้านผู้เรียน ทั้ง 3 ด้านได้รับความพึงพอใจระดับมากที่สุด และความพึงพอใจในภาพรวมของการเรียนรู้แบบไมโคร ได้รับความพึงพอใจระดับมากที่สุด (4.65)

5. สรุปและอภิปรายผล

จากผลการวิจัยสามารถสรุปและอภิปรายผลได้ดังนี้

5.1 การศึกษาประสิทธิผลการเรียนรู้ก่อนและหลังทำการเรียนรู้ของกลุ่มตัวอย่างที่ทำการเรียนรู้แบบดั้งเดิม

จากการศึกษาพบว่าประสิทธิผลการเรียนรู้หลังทำการเรียนรู้ของผู้เรียนที่ทำการเรียนรู้แบบดั้งเดิมสูงกว่าก่อนทำการเรียนรู้ อย่างมีนัยสำคัญทางสถิติ สาเหตุอาจเป็นเพราะ การเรียนรู้แบบดั้งเดิมเป็นการให้ความรู้ในรูปแบบบรรยาย ซึ่งในปัจจุบันใช้การสอนในรูปแบบนี้เป็นหลัก ผู้เรียนได้รับความรู้จากการเรียน ทำให้ผู้เรียนสามารถทำข้อสอบได้ ประสิทธิผลการเรียนรู้หลังทำการเรียนรู้จึงสูงกว่าก่อนทำการเรียนรู้

5.2 การศึกษาประสิทธิผลการเรียนรู้ก่อนและหลังทำการเรียนรู้ของกลุ่มตัวอย่างที่ทำการเรียนรู้แบบไมโคร

จากการศึกษาพบว่าประสิทธิผลการเรียนรู้หลังทำการเรียนรู้ของผู้เรียนที่ทำการเรียนรู้แบบไมโครสูงกว่าก่อนทำการเรียนรู้ อย่างมีนัยสำคัญทางสถิติ สาเหตุอาจเป็นเพราะ การเรียนรู้แบบไมโครมีรูปแบบในการเรียนรู้ที่เป็นประโยชน์และตรงต่อความต้องการของผู้เรียน ทำให้ผู้เรียนสามารถรับความรู้และเข้าใจในบทเรียน ส่งผลให้ผู้เรียนมีความรู้เพิ่มขึ้นจึงสามารถทำข้อสอบได้ ประสิทธิผลการเรียนรู้หลังทำการเรียนรู้จึงสูงกว่าก่อนทำการเรียนรู้

5.3 การศึกษาเปรียบเทียบประสิทธิผลการเรียนรู้ก่อนทำการเรียนรู้ของกลุ่มตัวอย่างทั้ง 2 กลุ่ม

จากการศึกษาพบว่าประสิทธิผลการเรียนรู้ก่อนทำการเรียนรู้ของผู้เรียนที่ทำการเรียนรู้แบบดั้งเดิมสูงกว่าผู้เรียนที่ทำการเรียนรู้แบบไมโคร อย่างไม่มีนัยสำคัญทางสถิติ สาเหตุอาจเป็นเพราะ นักเรียนทั้ง 2 กลุ่ม ยังไม่เคยเรียนในบทเรียนนี้ ทำให้ไม่มีความรู้ความเข้าใจในบทเรียน ซึ่งแสดงให้เห็นว่านักเรียนทั้ง 2 กลุ่ม มีพื้นฐานความรู้ที่เท่ากันหรือไม่แตกต่างกัน

5.4 การศึกษาเปรียบเทียบประสิทธิผลการเรียนรู้หลังทำการเรียนรู้ของกลุ่มตัวอย่างทั้ง 2 กลุ่ม

จากการศึกษาพบว่า ประสิทธิผลการเรียนรู้หลังทำการเรียนรู้ของผู้เรียนที่ทำการเรียนรู้แบบไมโครสูงกว่าผู้เรียนที่ทำการเรียนรู้แบบดั้งเดิม อย่างมีนัยสำคัญทางสถิติ ซึ่งเป็นไปตามสมมติฐานที่ตั้งไว้ แสดงว่าการนำการเรียนรู้แบบไมโครมาใช้ในการเรียน ทำให้ผู้เรียนมีประสิทธิผลทางการเรียนสูงขึ้น ซึ่งสอดคล้องกับงานวิจัย [19] สาเหตุอาจเป็นเพราะ ลักษณะ/รูปแบบการเรียนรู้แบบไมโคร การเรียนรู้แบบไมโคร ช่วยเพิ่มทักษะและลดช่องว่างระหว่างการเรียนรู้ของผู้เรียน มีรูปแบบการเรียนรู้ที่หลากหลาย ผู้เรียนสามารถได้รับความรู้ได้อย่างรวดเร็วและตรงประเด็น การเรียนรู้แบบไมโครสามารถทำให้เกิดความเข้าใจได้อย่างเป็นขั้นเป็นตอน สามารถเรียนรู้ในการเรียนรู้ที่สั้น ๆ ได้ง่ายและได้ผลดีกว่าการเรียนรู้แบบดั้งเดิมที่เรียนรู้แบบมาก ๆ ในครั้งเดียว การเรียนรู้แบบไมโครเหมาะสำหรับหลาย ๆ สถานการณ์ เนื่องจากสามารถเปลี่ยนแปลงข้อมูลได้อย่างรวดเร็ว ใช้ทรัพยากรที่มีอยู่ได้อย่างอิสระ มีเทคโนโลยีสารสนเทศมาสนับสนุน

5.5 การศึกษาความพึงพอใจในการเรียนรู้แบบไมโคร

จากการศึกษาพบว่ารูปแบบและคุณสมบัติของการเรียนรู้แบบไมโครตอบโจทย์กับกลุ่มผู้เรียนที่อยู่ในเจนเนอเรชันแซด โดยแบ่งออกเป็น 3 ด้าน ดังนี้

ด้านเนื้อหาพบว่าเนื้อหาตรงต่อความต้องการของผู้เรียน ปริมาณของเนื้อหาและเวลาที่ใช้ในการเรียนมีความเหมาะสม มีความชัดเจนในการอธิบายเนื้อหาและลำดับขั้นตอนในการนำเสนอเนื้อหามีความเหมาะสม ทำให้ผู้เรียนสามารถเลือกหัวข้อที่สนใจในการเรียนรู้ได้โดยตรง ลำดับความเข้าใจในบทเรียนได้ง่ายขึ้น

ด้านสื่อการสอนพบว่าการเรียนรู้แบบไมโครทำให้ผู้เรียนเห็นภาพประกอบในการเรียน ทำให้ผู้เรียนสามารถเข้าใจในบทเรียนได้ง่ายบทเรียนมีความน่าสนใจ ใช้งานง่ายและสะดวก นอกจากนี้ยังส่งเสริมให้ผู้เรียนเรียนรู้ด้วยตนเอง ทำให้ผู้เรียนมีความสนใจในการเรียนรู้มากขึ้น มีความเข้าใจและเห็นภาพในเรื่องที่จะเรียน สามารถเรียนในระยะเวลาที่ผู้เรียนพึงพอใจ ลดความเบื่อหน่ายในการเรียน

ด้านผู้เรียน พบว่า ผู้เรียนสามารถเข้าถึงบทเรียนได้ง่าย สะดวก สามารถเรียนซ้ำได้เมื่อไม่เข้าใจ การเรียนรู้แบบไมโครทำให้ผู้เรียนเข้าใจบทเรียนมากขึ้น สามารถจัดเวลา/ตารางในการเรียนรู้ได้ สามารถเรียนรู้ได้ทุกที่ทุกเวลาผ่านอุปกรณ์อิเล็กทรอนิกส์ เช่น สมาร์ตโฟน แท็ปเล็ต ได้ง่าย นอกจากนี้ผู้เรียนยังสามารถแบ่งและจัดสรรเวลาเรียนได้อย่างมีประสิทธิภาพ

6. ข้อเสนอแนะ

การวิจัยครั้งนี้มีข้อเสนอแนะเพื่อเป็นแนวทางในการวิจัยครั้งต่อไป ดังนี้

1. ควรศึกษาเพิ่มเติมถึงปัจจัยอื่น ๆ ที่มีผลต่อการเรียนรู้
2. ควรศึกษาเพิ่มเติมถึงรายวิชาอื่น ๆ ที่นำมาใช้ในการเรียนรู้ เพื่อให้เห็นผลว่าการเรียนรู้แบบไมโครเหมาะสมกับการเรียนรู้แบบใด
3. ควรศึกษาถึงสื่ออิเล็กทรอนิกส์ตัวอื่น ๆ เพิ่มเติม เพื่อหาลักษณะการใช้งานของผู้เรียน
4. ควรกำหนดขอบเขตของประชากรและกลุ่มตัวอย่างให้ครอบคลุมประชากรที่อยู่ในกลุ่มที่อยู่ในช่วงของเจนเนอเรชันแซดเพิ่มเติม เพื่อให้เห็นผลในมุมกว้างมากขึ้น

เอกสารอ้างอิง

[1] ประยุทธ์ จันทร์โอชา, 2559, *ปฏิรูปการศึกษาเพื่ออนาคตประเทศไทย มั่นคงมั่งคั่งยั่งยืน นโยบายด้านการศึกษาของนายกรัฐมนตรี (พลเอกประยุทธ์ จันทร์โอชา)*, พิมพ์ครั้งที่ 2, สำนักงานเลขาธิการสภาการศึกษา, กรุงเทพฯ, หน้า 1 - 35.

[2] Statistic Brain, *Attention Span Statistics*, statistic brain [Online], 2015, Available: http://www. statistibrain.com/ attention- span-statistics/ [2015, December 25].

[3] Watson, L., *Humans have shorter attention span than goldfish, thanks to smartphones*, telegraph [Online], 2015, Available: http://www.telegraph.co.uk/news/ science/ science-news/11607315/Humans-have-shorter-attention-span-than-goldfish-thanks-to-smartphones.html [2016, January 11].

[4] Hug, T., Innsbruck, *Micro Learning and Narration Exploring possibilities of utilization of narrations and storytelling for the designing of "micro units" and didactical micro-learning arrangements*, 4th, May 6 - 8, 2005, MIT, Cambridge (MA), USA, p. 4, 2005.

[5] เดชา เดชะวัฒนไพศาล, กฤษยา นุ่มพยา, จีราภา นวลลักษณ์และชนพัฒน์ ปลื้มบุญ, "การศึกษาเจนเนอเรชันเอ็กซ์และเจนเนอเรชันวายในมุมมองต่อคุณลักษณะของตนเองและความคาดหวังต่อคุณลักษณะของเจนเนอเรชันอื่น" *จุฬาลงกรณ์ธุรกิจปริทัศน์* ปีที่ 36 ฉบับที่ 141 ก.ค.- ก.ย. 57 หน้า 4 - 7.

[6] มนสิการ กาญจนะจิตรา, "ต่างรุ่นต่างมุมกับทฤษฎีเจเนอเรชัน" *จดหมายข่าวประชากรและการพัฒนา มหาวิทยาลัยมหิดล* ปีที่ 35 ฉบับที่ 6 หน้า 7.

[7] McDonald, K., *"How to Market to People Not Like You: "Know It or Blow It" Rules for Reaching Diverse Customers"*, John Wiley & Sons, Incorporated, p. 86-89, 2011.

[8] Schroer, W.J., *"Generations X, Y, Z and the Others - Cont'd"*, socialmarketing [Online], 2014, Available: http://www.socialmarketing.org/newsletter/features/generation3.htm [2015, November 14].

[9] มนัสนันท์ หัตถศักดิ์, 2557, *"การสร้างความรู้ความเข้าใจกระบวนการเรียนการสอน"*, คณะศึกษาศาสตร์ มหาวิทยาลัยเกษตรศาสตร์, หน้า 8 - 40.

[10] Gharibpoor, M., Sargazi S., Aref M., "Efficiency evaluation of e-learning compared to traditional education in human resource development (Case Study: Small and medium enterprises in Shiraz)", *7th International Conference on e-Commerce in Developing Countries with focus on e-Security*, Kish Island, Iran, pp. 1 – 5, 2013.

[11] Alaneme, G.C., Olayiwola, P.O., Reju, C.O., "Combining Traditional Learning and the E-Learning Methods in Higher Distance Education: Assessing Learners' Preference", *4th International Conference on Distance Learning and Education (ICDLE)*, p. 187, 2010.

[12] Mest, *Microlearning*, edutechwiki [Online], 2007, Available: http://edutechwiki.unige.ch/en/ Microlearning #Definition [2015, December 1].

[13] Buchem, I., *Microlearning: a strategy for ongoing professional development*, openeducationeueopa [Online] , 2010, Available: www.elearningpapers.eu [2015, December 1].

[14] Malamed, C., *Is Microlearning The Solution You Need?*, theeLearningCoach [Online], 2015, Available: http://theelearningcoach.com/elearning2-0/what-is-microlearning/ [2015, December 1].

[15] Epignosis LLC, *E-LEARNING CONCEPTS, TRENDS, APPLICATIONS*, talentlms[Online], 2014, Available: http://www.talentlms.com/elearning/elearning-101-jan2014-v1.1.pdf [2015, November 14].

[16] Eades, J., *Why Microlearning is HUGE and how to be a part of it*, elearningindustry [Online], 2014, Available: http://elearningindustry.com/why-microlearning-is-huge [2015, December 1].

[17] Eades, J., *5 Microlearning Commandments To Create You Own Microlearning*, eLearningindustry [Online], 2014, Available: http://elearningindustry.com/why-microlearning-is-huge [2015, November 14].

[18] Kar, S., *Microlearning Pros and Cons*, eLearning industry [Online], 2014, Available: http://elearningindustry.com/ microlearning-pros-and-cons [2015, January 14].

[19] Erradi, A., Almerekhi, H., Nahia, S., "Game-based micro-learning approach for language vocabulary acquisition using LingoSnacks", *IEEE Xplore Digital Library* [Electronic], 2013, pp. 235 – 237,Available: http://ieeexplore.ieee.org/ stamp/stamp.jsp?tp=&arnumber=6601916 [2015, December 1].

[20] Zhang, X., Ren, L., "Design for Application of Micro Learning to Informal Training in Enterprise", *IEEE Xplore Digital Library* [Electronic], 2011, pp. 2024-2027, Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=6011235 [2016, January 11].

สื่อการเรียนบนอุปกรณ์เคลื่อนที่สำหรับการเรียนภาษาจีนพื้นฐานด้วยพินอิน
Mobile Media Learning in Basic Chinese Based on Pinyin

สุวิมล ห้วงเกษม (Suwimon Huangkasem)[1] เดชาวุฒิ วานิชสรรพ์ (Dechawut Wanichsan)[2]
พรฤดี สิทธิพงศ์ (Pornrudee Sittipong)[3] นฤมล อรชร (Narumon Orachon)[4] และรุจิกา บุญเชิด (Rujika Booncherd)[5]
คณะครุศาสตร์ [2]คณะวิทยาการคอมพิวเตอร์และเทคโนโลยีสารสนเทศ [5]คณะมนุษยศาสตร์และสังคมศาสตร์ มหาวิทยาลัยราชภัฏรำไพพรรณี
[1]15515261045@rbru.ac.th, [2]Dechawut.w@rbru.ac.th, [3]Ayzlor@gmail.com
[4]Narumon1019@gmail.com, [5]Rujika.b@rbru.ac.th

บทคัดย่อ

วัตถุประสงค์ของงานวิจัยชิ้นนี้คือการพัฒนาสื่อการสอนบนอุปกรณ์เคลื่อนที่สำหรับการเรียนพื้นฐานภาษาจีนด้วยพินอิน เพื่อให้ผู้เรียนได้พัฒนาทักษะในการใช้ภาษาจีนของผู้ใช้ด้านการฟัง พูด อ่าน และเขียน โดยมีภาพ ตัวอักษร และเสียงจากคนจีนเจ้าของภาษาจำนวน 1,783 ไฟล์เสียงเป็นสื่อกลางในการเรียนรู้ เนื้อหาทั้งหมดในบทเรียนแบ่งออกเป็นหกส่วน ได้แก่ หลักการ พินอิน ผสมคำ ตัวอักษรภาษาจีน คำศัพท์ และเกม ซึ่งมีเนื้อหาครอบคลุมตั้งแต่ความรู้พื้นฐาน ตลอดจนผู้ใช้สามารถฝึกทักษะการเขียนอักขระจีน ทักษะการฟัง การออกเสียงได้เป็นอย่างดี สื่อการสอนที่พัฒนาขึ้นสามารถทำงานได้จริงบนระบบปฏิบัติการ ไอโอเอส แอนดรอยด์และวินโดวส์ ผลการทดลองแสดงให้เห็นว่าผู้ใช้งานสามารถพัฒนาผลการเรียน และมีความพึงพอใจอยู่ในระดับดีมาก

คำสำคัญ: พินอิน ภาษาจีน สัทอักษร สื่อการเรียนบนอุปกรณ์เคลื่อนที่ แอนดรอยด์

Abstract

The main objective of this research was to develop a mobile media learning in basic Chinese based on Pinyin system in order that learners would enhance multiple skills, which is listening, speaking, reading and writing. It comprised several media such as images, alphabet, and 1,783 sounds pronounced by Chinese native speaker. The media were separated into six sections, which is a principle, pinyin, constructing Chinese alphabet, vocabulary and game. Its content covered from basic lessons to difficult lessons in order to help the learner to practice writing Chinese alphabet, listening and spelling words as well. The media could be practically used on several operating system such as iOS, Android, and Windows. The experimental results reveal that after using the media, the learners could improve learning achievement, and satisfaction level is very high.

Keyword: Pinyin, Chinese, Phonetic Alphabet, Mobile Media Learning, Android

1. บทนำ

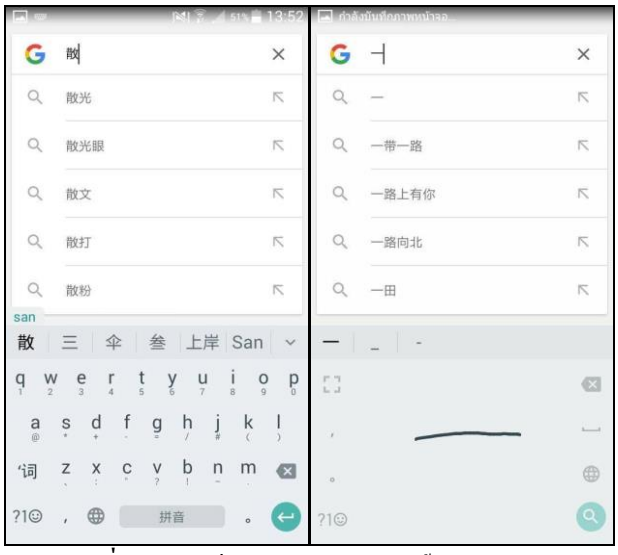
ภาษาจีนเป็นภาษาสากลของโลก ซึ่งมีผู้ใช้ถึงหนึ่งในห้าของประชากรโลก [1] และยังเป็นหนึ่งในภาษาสำคัญที่ใช้ในการประชุมต่าง ๆ ขององค์การสหประชาชาติ สังเกตได้จากหน้าแรกของเว็บไซต์ www.un.org ซึ่งมีภาษาให้เลือกใช้สำหรับการเข้าถึงข้อมูลหกภาษาอันประกอบด้วยภาษาอาหรับ จีน อังกฤษ ฝรั่งเศส รัสเซีย และสเปน [2] นอกจากนี้ภาษาจีนยังเป็นภาษาแม่ของอีกหลายประเทศ รวมถึงประเทศที่เป็นสมาชิกประชาคมอาเซียนอย่างประเทศสิงคโปร์อีกด้วย เพื่อความเป็นเอกภาพ รัฐบาลจีนจึงกำหนดให้ภาษาจีนกลาง เป็นภาษากลางในการติดต่อสื่อสาร โดยผู้ที่สนใจศึกษาสามารถเรียนผ่าน “สัทอักษรภาษาจีน” ซึ่งเป็นการนำตัวอักษรหรือสัญลักษณ์มากำหนดเพื่อแทนเสียงต่าง ๆ แบ่งออกเป็นสองประเภทหลัก ๆ คือ จู้อิน (注音符号) และพินอิน (汉语拼音) ซึ่งงานวิจัยนี้กล่าวถึงการเรียนภาษาจีนโดยใช้พินอิน

จากการเปลี่ยนแปลงทางสภาพสังคมโลก ผู้วิจัยจึงได้พัฒนาสื่อการเรียนบนอุปกรณ์เคลื่อนที่สำหรับการเรียนภาษาจีนพื้นฐานด้วยพินอิน เพื่อต้องการพัฒนาความสามารถในการใช้ภาษาจีนของคนไทย ทั้งทักษะในด้านการฟัง การพูด การอ่าน การเขียน

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 พินอิน

พินอิน หรือ ฮั่นหยวี่พินอิน คือ ระบบในการถอดเสียงภาษาจีนมาตรฐาน ด้วยตัวอักษรละติน ความหมายของพินอินคือ "การรวมเสียงเข้าด้วยกัน" โดยนัยก็คือ การเขียนแบบสัทศาสตร์ การสะกด การถอดเสียง หรือการทับศัพท์ พินอินเป็นการทับศัพท์ด้วยอักษรโรมัน (Romanization) มิใช่การถอดเสียงแบบภาษาอังกฤษ (Anglicization) นั่นคือ การกำหนดให้ใช้ตัวอักษรตัวหนึ่ง สำหรับแทนเสียงหนึ่ง ๆ ในภาษาจีนไว้อย่างตายตัว พินอินถูกใช้อย่างกว้างขวางในจีนแผ่นดินใหญ่ มาเลเซีย สิงคโปร์ และไต้หวัน แบบอักษรพินอินได้รับการจัดระบบมาตรฐาน และได้รับการยอมรับจากไอเอสโอ (International Organization for Standardization: ISO) เป็นแบบอักษรจีนสมัยใหม่ในปี 1982 ด้วยมาตรฐาน ISO7098 [3] นอกจากนี้ในปัจจุบันบริษัทกูเกิล (Google) ได้เล็งเห็นความสำคัญของภาษาจีน และได้พัฒนาโปรแกรมเสริมสำหรับการพิมพ์ภาษาจีนบนระบบปฏิบัติการแอนดรอยด์ ที่มีชื่อว่า กูเกิ้ล พินอิน อินพุต (Google Pinyin Input) [4] โดยระบบดังกล่าวสนับสนุนการป้อนข้อมูลที่มีความหลากหลายทั้งการป้อนข้อมูลด้วยการเลือกเส้นขีดตัวอักษรแบบอัจฉริยะ ด้วยการพิมพ์พินอินดังแสดงด้วยตัวอย่างหน้าจอในภาพที่ 1 ด้านซ้าย รวมถึงสามารถนำข้อมูลเข้าด้วยการเขียนด้วยลายมือได้ด้วย ดังแสดงด้วยตัวอย่างหน้าจอในภาพที่ 1 ด้านขวา



ภาพที่ 1: หน้าจอที่นำเข้าอักขระจีนด้วยกูเกิ้ล พินอิน อินพุต

2.2 งานวิจัยที่เกี่ยวข้อง

วลงกรณ์ และคณะ [5] ได้พัฒนาเกม White Horse เพื่อการเรียนรู้ภาษาอังกฤษขั้นพื้นฐานบนระบบปฏิบัติการแอนดรอยด์ โดยมีแนวคิดการเรียนรู้ภาษาอังกฤษที่ว่าด้วยการอ่านออกเสียงและสะกดคำ (Phonics) มาผสมผสานกับรูปภาพและเทคนิคการสร้างเกมเพื่อการเรียนรู้สำหรับเด็ก เพื่อส่งเสริมทักษะภาษาอังกฤษด้านการฟัง การพูด การอ่าน การเขียน การสะกดคำ ผลการทดลองพบว่า นักเรียนมีความพึงพอใจต่อเกมในระดับมาก และสามารถนำเกมไปใช้เป็นเกมเสริมการเรียนรู้ได้

Lu Sheng [6] ได้ศึกษาระบบเสียงภาษาไทยกับภาษาจีน พบว่า ระบบเสียงทั้งสองภาษาจะต่างกันมาก หากนักศึกษามีความรู้พื้นฐานเกี่ยวกับสัทศาสตร์ จะมีส่วนช่วยในการออกเสียงได้อย่างถูกต้อง

รัตติยา จันทร์เอียด [7] ได้ทำการวิจัยในชั้นเรียนเรื่องการอ่านสะกดคำในภาษาจีนโดยใช้แบบทดสอบ โดยสร้างชุดฝึกทักษะการอ่านสะกดพินอิน ผลการวิจัยพบว่า นักเรียนมีความกระตือรือร้นที่จะเรียน มีความสนุกสนานขณะเรียน เกิดทักษะการเรียนรู้มากขึ้นกว่าเดิม และการทดสอบหลังเรียนมีคะแนนสูงกว่าก่อนเรียน

เดชาวุฒิ และคณะ [8] ได้พัฒนาขั้นตอนวิธีสำหรับการตรวจสอบการเขียนอักษรไทยบนแท็บเลตพีซี โดยขั้นตอนวิธีดังกล่าวสามารถตรวจทิศทางการเขียนอักษระไทยในตำแหน่งและทิศทางที่ถูกต้อง รวมถึงมีวิธีการคิดคะแนนความคล้ายของตัวอักษรที่วาดขึ้นกับตัวอักษรต้นแบบอีกด้วย

ทฤษฎีและงานวิจัยที่เกี่ยวข้องข้างต้น ได้ให้แนวคิดพื้นฐานกับผู้วิจัยเพื่อพัฒนาสื่อการสอนภาษาจีนที่มีส่วนประกอบต่าง ๆ ดังนี้ 1) ทำงานได้บนอุปกรณ์เคลื่อนที่ เพื่อความสะดวกในการพกพาสื่อการสอนของผู้ใช้งาน 2) เนื้อหาในสื่อเป็นภาษาจีนบนพินอิน เนื่องจากได้รับยอมรับอย่างกว้างขวาง แม้กระทั่งกูเกิ้ลยังสร้างระบบป้อนข้อมูลด้วยการใช้หลักการ พินอิน การที่ผู้ใช้งานมีความรู้ด้านภาษาจีน พินอิน รวมถึงเส้นขีดต่าง ๆ จะทำให้ผู้ใช้สามารถป้อนข้อมูลบนอุปกรณ์ที่รองรับการป้อนข้อมูลโดยใช้หลักการภาษาจีน พินอินได้อย่างถูกต้องและง่ายดาย 3) สื่อการสอนจะประยุกต์ขั้นตอนวิธีสำหรับการใช้ตรวจสอบความถูกต้องของการเขียนอักขระจีน เพื่อผู้เรียนจะได้รับความรู้ที่ถูกต้องเกี่ยวกับทิศทางและรูปทรง

การเขียนอักขระจีนที่ถูกต้อง เนื่องจากทิศทางการเขียนอักษรจีนมีความสำคัญมาก ตัวอย่างเช่น การวาดเส้นขีดผิดทิศทาง แม้ว่าผลลัพธ์สุดท้ายจะได้รูปทรงเดียวกัน ทำให้กูเกิ้ล พินอิน อินพุต ไม่สามารถตรวจสอบว่าผู้ใช้ป้อนอักขระตัวใดเข้าไป 4) มีหลักการของสัทศาสตร์เป็นส่วนหนึ่งของเนื้อหา เพราะมีส่วนช่วยในการออกเสียงคำโดยตรง

3. วิธีดำเนินการวิจัย

3.1 การออกแบบและพัฒนาสื่อการเรียน

สื่อการเรียนบนอุปกรณ์เคลื่อนที่จะได้รับพัฒนาขึ้นโดยใช้ภาษาแอ็กชันสคริปต์ 3.0 (Action Script) ด้วยเหตุผลที่ว่าเมื่อพัฒนาเสร็จแล้วจะสามารถทำงานได้บนอุปกรณ์เคลื่อนที่หลายประเภทได้แก่ คอมพิวเตอร์โน้ตบุ๊ค (Computer Notebook) ที่ทำงานบนระบบปฏิบัติการวินโดวส์ (Windows) และโทรศัพท์มือถือหรือแท็บเล็ตพีซีที่ทำงานบนระบบปฏิบัติการแอนดรอยด์ (Android) และไอโอเอส (iOS)

3.2 การออกแบบเนื้อหาในสื่อการเรียน

เนื้อหาเกี่ยวกับภาษาจีนเบื้องต้นที่อยู่ในสื่อการเรียนแบ่งออกเป็นหกเรื่อง ได้แก่ หลักการ พินอิน การผสมคำ เส้นขีดตัวอักษรในภาษาจีน คำศัพท์ภาษาจีนที่เกี่ยวกับชีวิตประจำวันและเกม โดยผู้พัฒนาสื่อการเรียนจำเป็นต้องศึกษาสัทอักษร (Phonetic Alphabet) เพื่อผนวกเข้าเป็นส่วนหนึ่งของสื่อ เนื่องจากเนื้อหาในส่วนนี้ช่วยในการออกเสียงคำศัพท์โดยตรง นอกจากนี้งานวิจัยนี้ได้พัฒนาขั้นตอนวิธีสำหรับการวาดอักษรจีน จากขั้นตอนวิธีของเดชาวุฒิ และคณะ [8] เพื่อพัฒนาให้สื่อสามารถตรวจสอบความถูกต้องของการลากเส้นขีดตัวอักษร ทำให้สื่อการสอนการตรวจสอบทิศทางการเขียนอักษรมีความถูกต้อง เสียงที่ใช้ในสื่อการสอนทั้งหมดจะได้รับการบันทึกจากชาวจีนเจ้าของภาษา โดยจะบันทึกทั้งในส่วนของพินอิน คำศัพท์ และการผสมคำ

3.3 เครื่องมือที่ใช้วัดประสิทธิภาพของสื่อการเรียน

นอกเหนือจากสื่อการเรียนบนอุปกรณ์เคลื่อนที่ที่พัฒนาขึ้น เครื่องมือที่ใช้ในการวัดประสิทธิภาพสื่อการเรียนจากผู้ใช้งานได้แก่

3.3.1 แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียน เป็นข้อสอบจำนวน 40 ข้อ แต่ละข้อมีสามตัวเลือก โดยข้อสอบวัดผลสัมฤทธิ์ทางการเรียนจะได้รับการสลับข้อให้มีลำดับข้อที่ไม่เหมือนกันเพื่อใช้เป็นแบบทดสอบก่อนเรียน และแบบทดสอบหลังเรียน

3.3.2 แบบสอบถามวัดความพึงพอใจ มีจำนวน 12 ข้อ โดยมุ่งเน้นเพื่อวัดความพึงพอใจในสามด้าน ได้แก่ ด้านการออกแบบสื่อ ด้านเนื้อหา และด้านการใช้งาน

โดยเครื่องมือทั้งในข้อ 3.3.1 และ 3.3.2 ได้รับการตรวจสอบความถูกต้องจากคณาจารย์ผู้สอนในสาขาวิชาภาษาจีนจำนวนสามท่าน แต่ละท่านมีคุณวุฒิขั้นต่ำระดับปริญญาโท และมีประสบการณ์สอนในระดับอุดมศึกษาไม่น้อยกว่าสามปี เมื่อเครื่องมือเสร็จสมบูรณ์ จึงนำไปทดลองใช้กับผู้ร่วมวิจัยจำนวน 30 คน ที่มีโอกาสศึกษาหรือนำภาษาจีนกลางไปใช้ในชีวิตประจำวันช่วงอายุ 7-40 ปี ซึ่งอยู่ในวัยเด็กถึงวัยผู้ใหญ่บริเวณตัวเมืองจังหวัดจันทบุรี โดยแบ่งช่วงอายุเป็นหกช่วง ได้แก่ ช่วงอายุ 7-12, 13-18, 19-23, 24-29, 30-35 และ 36-40 ปี แต่ละช่วงจะมีผู้ร่วมวิจัยจำนวนห้าคน

4. ผลการดำเนินการวิจัย

4.1 สื่อการสอนที่พัฒนาขึ้น

สื่อการสอนที่พัฒนาขึ้นสามารถทำงานได้บนระบบปฏิบัติการวินโดวส์ แอนดรอยด์ และไอโอเอส โดยหน้าจอหลักของสื่อการเรียนแสดงดังภาพที่ 2



ภาพที่ 2: หน้าเมนูของสื่อการเรียนรู้

จากภาพที่ 2 หน้าจอดังกล่าวมีปุ่มให้เลือกกดอยู่ 6 รายการ ได้แก่ หลักการ พินอิน ผสมคำ ตัวอักษรภาษาจีน คำศัพท์ และเกม เมื่อผู้ใช้งานคลิกปุ่ม “หลักการ” ข้อมูลทั่วไปเกี่ยวกับหลักการของภาษาจีนที่จำเป็นต้องทราบจะแสดงออกทางหน้าจอตามลำดับอันได้แก่ ข้อมูลเกี่ยวกับการจำแนกเสียงตาม

แหล่งกำเนิดเสียง การใส่เครื่องหมายวรรณยุกต์ กฎการเขียนสัทอักษรจีน และกฎการอ่านวรรณยุกต์ โดยภาพที่ 3 แสดงหน้าจอการจำแนกเสียงตามแหล่งกำเนิดเสียง จากหน้าจอดังกล่าวผู้ใช้งานสามารถกดปุ่มถัดไปเพื่อไปยังหน้าอื่น หรือกดปุ่มโฮมเพื่อกลับสู่เมนูหลัก



ภาพที่ 3: หน้าเนื้อหาการออกเสียงตามแหล่งกำเนิด

จากหน้าเมนูหลัก เมื่อผู้ใช้เลือกเมนูพินอิน จะมีเมนูย่อยสี่เมนู ได้แก่ พยัญชนะ สระเดี่ยว สระผสม และวรรณยุกต์ ให้ผู้ใช้เลือกเมนูที่ต้องการเรียนรู้ ตัวอย่างเช่นหากผู้ใช้เลือกเมนูพยัญชนะ จะแสดงหน้าเมนูพยัญชนะขึ้นมา ตัวอย่างหน้าจอพยัญชนะ “ป๋ว” ที่ประกอบด้วยพินอิน คำอ่าน แสดงดังภาพที่ 4 หากกดที่ลำโพง ผู้ใช้สามารถฟังเสียงพินอินได้



ภาพที่ 4: หน้าเนื้อหาพินอิน ของพยัญชนะ “ป๋ว”

จากนั้นหน้าเมนูหลัก เมื่อผู้ใช้เลือกเมนูผสมคำ โปรแกรมจะแสดงเมนูย่อยพยัญชนะ เพื่อให้ผู้ใช้เลือกว่าต้องการเรียนรู้พยัญชนะตัวใด และเมื่อกดเลือกพยัญชนะแล้ว จะปรากฏเนื้อหา ดังตัวอย่างหน้าจอที่เลือกการผสมคำพยัญชนะ “ป๋ว” แสดงดังภาพที่ 5 ในหน้าจอนี้ผู้ใช้สามารถที่เซลล์ของตารางเพื่อฟังเสียงแต่ละโทนเสียงได้



ภาพที่ 5: หน้าการผสมคำของพยัญชนะ “ป๋ว”

จากหน้าเมนูหลัก เมื่อผู้ใช้เลือกเมนูตัวอักษรภาษาจีน จะมีเมนูย่อยให้เลือกอีกสี่เมนู ดังหน้าจอในภาพที่ 6 ได้แก่ 1) เส้นขีดตัวอักษร 32 รูปแบบ 2) ลำดับขีดตัวอักษร 8 รูปแบบ 3) โครงสร้างของตัวอักษรจีน 17 รูปแบบ และฝึกเขียน ซึ่งปุ่มโฮมตรงกลางสามารถกดเพื่อกลับไปยังเมนูหลัก



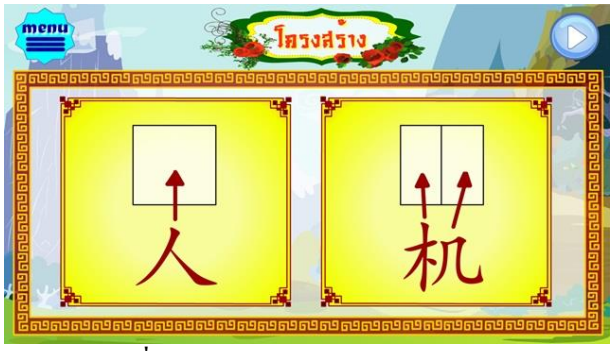
ภาพที่ 6: หน้าเนื้อหาการออกเสียงตามแหล่งกำเนิด

จากหน้าจอในภาพที่ 6 เมื่อผู้ใช้เลือกเมนูเส้นขีดตัวอักษร จะแสดงเนื้อหาประกอบด้วย วิธีการเขียน ชื่อภาษาจีน พินอิน สามารถกดที่ลำโพงเพื่อฟังเสียงได้ รวมถึงแสดงตัวอย่างตัวอักษรที่มีเส้นขีดนั้นประกอบอยู่ด้วย หากผู้ใช้เลือกเมนูลำดับขีดตัวอักษร จะแสดงเนื้อหาประกอบด้วย พื้นฐานการขีดเส้นต่าง ๆ ในภาษาจีน 8 รูปแบบ หากผู้ใช้เลือกเมนูโครงสร้างของตัวอักษรจีน จะแสดงรูปของโครงสร้างทั้ง 17 รูปแบบ และหากผู้ใช้เลือกเมนูฝึกเขียน โปรแกรมจะแสดงเนื้อหาสำหรับการฝึกเขียน โดยผู้ใช้ต้องเป็นผู้เขียนเส้นขีดอักษรตาม จากนั้นกดปุ่ม “ส่ง” เพื่อให้โปรแกรมคำนวณคะแนนการเขียน ซึ่งในทุกเมนูสามารถกดปุ่มเมนูมุมบนซ้ายมือ เพื่อกลับไปเลือกเมนูย่อยทั้ง 4 เมนู หรือกดปุ่มถัดไปเพื่อแสดงเนื้อหาอื่น ๆ ใน

เมนูแต่ละเมนูได้ ดังตัวอย่างหน้าจอ แสดงดังภาพที่ 7-9 ตามลำดับ



ภาพที่ 7: หน้าต่างเนื้อหาเมนูลำดับขีด



ภาพที่ 8: หน้าต่างเนื้อหาเมนูโครงสร้างของตัวอักษร



ภาพที่ 9: หน้าต่างเนื้อหาการฝึกเขียน

เมื่อผู้ใช้เลือกเมนูคำศัพท์ โปรแกรมจะแสดงเมนูย่อยคำศัพท์เพื่อให้ผู้ใช้เลือกหมวดคำศัพท์ที่ต้องการศึกษา ประกอบด้วย 11 หมวด ได้แก่ ผลไม้ สัตว์ ตัวเลข ห้องเรียน อาชีพ ของใช้ในชีวิตประจำวัน ร่างกาย ครอบครัว สี เครื่องแต่งกาย และอื่น ๆ (ผัก สถานที่ อาหาร ฤดูกาลและสภาพอากาศ ยานพาหนะ) เมื่อผู้ใช้เลือกหมวดแล้ว จะปรากฏเมนูย่อยของแต่ละหมวด จากนั้นให้ผู้ใช้เลือกคำศัพท์ที่สนใจ โดยคำศัพท์ประกอบด้วย รูป ชื่อภาษาจีน พินอิน คำแปล และผู้ใช้สามารถกดที่ลำโพงเพื่อฟังเสียงได้ ดังตัวอย่างหน้าจอแสดงคำศัพท์หมวด “เครื่องแต่งกาย” แสดงภาพที่ 10



ภาพที่ 10: หน้าต่างเนื้อหาคำศัพท์ “เน็คไท”

หากผู้ใช้เลือกเมนูเกม โปรแกรมจะแสดงเมนูเกมประกอบด้วยสองเมนูได้แก่ 1) พินอินหรรษาและ 2) รอยขีดความจำ เนื้อหา ประกอบด้วย คำถามแบบ 3 ตัวเลือก มีทั้งกดฟังเสียง และเลือกตอบพินอิน ให้ผู้ใช้เลือกคำตอบที่ถูกต้องที่สุด โดยคำถามมีทั้งหมดเกมละ 30 ข้อ โปรแกรมจะทำการสุ่มให้ผู้ใช้เอง เมื่อเล่นเกมครบทั้ง10 ข้อในแต่ละเกมแล้ว โปรแกรมจะแสดงคะแนนออกทางจอภาพ

4.2 ผลสัมฤทธิ์ทางการเรียนของผู้ใช้งาน

งานวิจัยชิ้นนี้ได้วัดผลสัมฤทธิ์ทางการเรียนของผู้ใช้งานซึ่งเป็นบุคคลทั่วไปที่มีโอกาสศึกษาหรือนำภาษาจีนกลางไปใช้ในชีวิตประจำวัน โดยผลการทดลองแสดงดังตารางที่ 1 ที่แสดงค่าการเปรียบเทียบผลสัมฤทธิ์ทางการเรียนระหว่างการทำแบบทดสอบก่อนเรียนและหลังเรียน ผลการทดลองแสดงให้เห็นว่าผู้ใช้งานมีผลสัมฤทธิ์ทางการเรียนโดยได้รับคะแนนเฉลี่ยหลังเรียน 32.87 คะแนน ซึ่งสูงกว่าคะแนนเฉลี่ยก่อนเรียน 16.40 คะแนนอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05

ตารางที่ 1: ผลสัมฤทธิ์ทางการเรียนก่อนเรียนและหลังเรียน

แบบทดสอบ	คะแนนเต็ม	M	SD	*t*	*p*
ก่อนเรียน	40	16.40	3.874	26.207	.000*
หลังเรียน	40	32.87	2.285		

$p^* \leq 0.05$

4.3 ระดับความพึงพอใจของผู้ใช้งาน

หลังจากทดสอบผลสัมฤทธิ์ทางการเรียนของผู้ใช้งานงานวิจัยชิ้นนี้ยังได้ศึกษาความพึงพอใจจากผู้ใช้งานกลุ่มดังกล่าว ผลที่ได้แสดงดังตารางที่ 2

ตารางที่ 2: ความพึงพอใจของผู้ใช้งาน

รายการประเมิน	ค่าเฉลี่ย	SD	ระดับ
ด้านการออกแบบ			
1. มีสีสันสวยงาม เป็นระเบียบเรียบร้อย	4.57	0.504	ดีมาก
2. ขนาดตัวอักษร ภาพ มีความเหมาะสม	4.43	0.568	ดี
3. ปุ่มมีการจัดวางเหมาะสม	4.50	0.509	ดีมาก
ด้านเนื้อหา			
1. เนื้อหาเข้าใจง่าย ไม่สับสน	4.37	0.556	ดี
2. เนื้อหาเหมาะสมกับผู้ที่เริ่มศึกษาภาษาจีน	4.67	0.547	ดีมาก
3. ความชัดเจนของเสียงบรรยาย	4.20	0.761	ดี
ด้านการใช้งาน			
1. ประสิทธิภาพของสื่อการเรียนรู้	4.40	0.498	ดี
2. ผู้ใช้ได้รับความรู้เพิ่มขึ้น	4.77	0.43	ดีมาก
3. ใช้งานได้ง่าย ไม่ซับซ้อน	4.60	0.498	ดีมาก
4. สามารถนำความรู้ไปใช้ได้จริง	4.50	0.509	ดีมาก
5. มีความสนุกสนาน เพลิดเพลิน	4.50	0.509	ดีมาก
6. ความพึงพอใจในภาพรวม	4.53	0.507	ดีมาก
รวม	4.50	0.533	ดีมาก

ตารางที่ 2 แสดงให้เห็นว่าผลจากการประเมินความพึงพอใจเฉลี่ยของผู้ใช้งานอยู่ในระดับดีมาก โดยมีค่าเฉลี่ย 4.5033 และค่าส่วนเบี่ยงเบนมาตรฐาน 0.5333 ตามลำดับ

5. สรุป

งานวิจัยชิ้นนี้ได้พัฒนาสื่อการสอนบนอุปกรณ์เคลื่อนที่สำหรับการเรียนภาษาจีนเบื้องต้นด้วยพินอิน ที่จัดทำขึ้นสำหรับบุคคลทั่วไปที่สนใจศึกษา ซึ่งมีเนื้อหาประกอบด้วยพินอินทั้งหมด จำนวน 60 ตัว การผสมคำทั้งหมด จำนวน 1,484 คำ จำนวนคำศัพท์ทั้งหมด 207 คำ มีการแบ่งหมวดหมู่ของคำศัพท์เป็น 11 หมวด โดยคำศัพท์แต่ละคำ ประกอบด้วย รูปภาพ ชื่อภาษาจีน พินอิน คำแปล พร้อมกับปุ่มกดฟังเสียง เส้นขีดตัวอักษรภาษาจีนทั้งหมดจำนวน 32 รูปแบบ ลำดับขีดตัวอักษรแปดรูปแบบ โครงสร้างของตัวอักษรจีน 17 รูปแบบ และเกมฝึกทักษะแบบสามตัวเลือก จำนวนสองเกม ได้แก่ พินอินหรรษาและรอยขีดความจำ โดยแต่ละเมนูจะมีคำแนะนำก่อนเข้าสู่บทเรียน ทำให้ผู้ใช้งานเข้าใจและไม่เกิดความสับสนในระหว่างการเรียนรู้ ผลการทดลองแสดงให้เห็นว่าผู้เรียนมีผลสัมฤทธิ์ทางการเรียนสูงขึ้นกว่าก่อนเรียนอย่างมีนัยสำคัญทางสถิติ และมีความพึงพอใจต่อการใช้งานสื่อการเรียนรู้อยู่ในระดับดีมาก อย่างไรก็ตามจากการสัมภาษณ์เพิ่มเติมหลังการใช้งานสื่อการเรียน ผู้เรียนจำนวนหนึ่งต้องการให้พัฒนาเนื้อหาเพิ่มเติมเกี่ยวกับประโยคสนทนาที่ใช้บ่อยในชีวิตประจำวัน เนื่องจากผู้ใช้งานมีสายอาชีพที่จำเป็นต้องพบปะพูดคุยกับคนจีน หากผู้ใช้สนใจนำสื่อการเรียนรู้ ไปใช้เป็นสื่อหรืออุปกรณ์ในการเรียนการสอน สามารถติดต่อมาทางคณะผู้จัดทำได้ ทางคณะผู้จัดทำจะแนะนำการใช้งานให้กับผู้ที่สนใจอย่างเต็มที่

เอกสารอ้างอิง

[1] worldmeters.info. “ประชากรโลก”. Internet: www.worldometers.info/world-population. [Mar. 15, 2016].

[2] United Nation. “ภาษาทางการของสหประชาชาติ”. Internet: www.un.org. [Mar. 15, 2016].

[3] ISO. “มาตรฐานพินอิน (ISO7098)”. Internet: www.iso.org/standard/61420.html. [Mar. 30, 2016].

[4] google. “Google Pinyin Input”. Internet: https://play.google.com/store/apps/details?id=com.google.android.inputmethod.pinyin. [Mar. 30, 2016].

[5] วลงกรณ์ และคณะ, “การพัฒนาเกม White Horse เพื่อการเรียนรู้ภาษาอังกฤษขั้นพื้นฐานบนระบบปฏิบัติการแอนดรอยด์” *การประชุมวิชาการระดับชาติด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศ ครั้งที่ 10 (NCCIT 2014),* ภูเก็ต, 8-9 พฤษภาคม 2557.

[6] Lu Sheng, *การพัฒนารูปแบบการสอนการออกเสียงภาษาจีน โดยใช้สัทอักษรปฏิสัมพันธ์สำหรับนักศึกษาจีน.* วิทยานิพนธ์ปรัชญาดุษฎีบัณฑิต สาขาวิชาเทคโนโลยีทางการศึกษา. บัณฑิตวิทยาลัย มหาวิทยาลัยบูรพา, 2553.

[7] รัตติยา จันทร์เอียด, *การสะกดคำในภาษาจีนโดยใช้แบบทดสอบ.* วิจัยในชั้นเรียน โรงเรียนปราษฎร์พิทยา อำเภอ เกาะสมุย จังหวัดสุราษฎร์ธานี, 2553.

[8] เดชาวุฒิ และคณะ, “ขั้นตอนวิธีสำหรับพัฒนาแบบฝึกหัดการเขียนอักษรไทยที่ใช้งานบนแท็บเล็ตพีซี” *การประชุมวิชาการระดับชาติด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศ ครั้งที่ 11 (NCCIT 2015),* กรุงเทพมหานครฯ, 2-3 กรกฎาคม 2558.

การพัฒนาบทเรียนคอมพิวเตอร์มัลติมีเดีย เรื่อง องค์ประกอบศิลป์เบื้องต้น สำหรับนักศึกษาสาขาวิชาเทคโนโลยีสารสนเทศ มหาวิทยาลัยราชภัฏนครปฐม

The Development of the Computer Multimedia Instruction on Compositions in Visual Arts for Information Technology Undergraduate Students of Nakhon Pathom Rajabhat University

ศัลยพงศ์ วิชัยดิษฐ (Salyapong Wichaidit)[1] และกุลยา เจริญมงคลวิไล (Kunlaya Charoenmongkonvilai1)[2]
สาขาวิชาเทคโนโลยีมัลติมีเดีย คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏนครปฐม
[1]Salyapong@webmail.npru.ac.th, [2]Kunlayacha@webmail.npru.ac.th

บทคัดย่อ

งานวิจัยนี้ มีวัตถุประสงค์เพื่อ 1) พัฒนาบทเรียนคอมพิวเตอร์มัลติมีเดีย เรื่อง องค์ประกอบศิลป์ สำหรับนักศึกษาสาขาวิชาเทคโนโลยีสารสนเทศ 2) เปรียบเทียบผลสัมฤทธิ์ทางการเรียนก่อนเรียนและหลังเรียน จากการเรียนด้วยบทเรียนคอมพิวเตอร์มัลติมีเดีย 3) ศึกษาความพึงพอใจของผู้เรียนที่มีต่อบทเรียนคอมพิวเตอร์มัลติมีเดีย กลุ่มตัวอย่างในการศึกษาครั้งนี้คือ นักศึกษาสาขาวิชาเทคโนโลยีสารสนเทศ (เทคโนโลยีมัลติมีเดีย)ซึ่งได้มาจากการเลือกแบบเฉพาะเจาะจง จำนวน 36 คน เครื่องมือที่ใช้ในการวิจัยครั้งนี้ประกอบด้วย 1) บทเรียนคอมพิวเตอร์มัลติมีเดีย เรื่อง องค์ประกอบศิลป์ สำหรับนักศึกษาสาขาวิชาเทคโนโลยีสารสนเทศ 2) แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียน 3) แบบสอบถามความพึงพอใจ สถิติที่ใช้ในการวิเคราะห์ข้อมูล ได้แก่ ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน และการทดสอบค่าที (t-test) จากผลการวิจัยพบว่า ผลสัมฤทธิ์ทางการเรียนของผู้เรียนที่เรียนด้วยบทเรียนคอมพิวเตอร์มัลติมีเดียหลังเรียน ($\bar{X}$ = 4.17 ,S.D. = 1.13) สูงกว่าก่อนเรียน ($\bar{X}$ = 8.72 ,S.D. = 0.85) และผลการศึกษาความพึงพอใจของผู้เรียนหลังเรียนด้วยบทเรียนคอมพิวเตอร์มัลติมีเดียอยู่ในระดับมาก ($\bar{X}$ = 4.35 ,S.D. = 0.073)

คำสำคัญ: บทเรียนคอมพิวเตอร์มัลติมีเดีย

Abstract

This research aims to: 1) develop multimedia computer multimedia subject matter For students in information technology 2) compare learning achievement before and after class 3) To study students' satisfaction with multimedia computer lessons. The sample in this study was Students in Information Technology (Multimedia technology) was selected from 36 specific types. The tools used in this research were 1) multimedia computer lesson on art composition For students in Information Technology 2) Achievement Test 3) Satisfaction Questionnaire Statistics used in data analysis were mean, standard deviation And t-test. The learning achievement of students learning with multimedia computer lessons after learning (= 4.17, SD = 1.13) was higher than before learning (= 8.72, SD = 0.85). Multimedia was at a high level (= 4.35, SD = 0.073).

Keyword: Computer Multimedia Instruction

1. บทนำ

บทบาทของเทคโนโลยีคอมพิวเตอร์ในปัจจุบันได้นำมาใช้ในการจัดการเรียนการสอนเพื่อให้ผู้เรียนมีทักษะความสามารถในการใช้เทคโนโลยีเพื่อการพัฒนาตนเอง ในการทำงาน การ

แก้ปัญหาอย่างสร้างสรรค์ และมีคุณธรรมจึงได้มีการพัฒนาบทเรียนขึ้นมาหลากหลายจนเข้าสู่รูปแบบสื่อมัลติมีเดียคอมพิวเตอร์เป็นเครื่องมือที่มีการนำมาใช้ในรูปแบบของบทเรียนคอมพิวเตอร์มัลติมีเดีย ทำให้การเรียนการสอนมีการลักษณะปฏิสัมพันธ์ช่วยให้ผู้เรียนสามารถเรียนรู้ สามารถทำกิจกรรมรวมถึงดูสื่อต่างๆ ได้ด้วยตนเองสื่อที่นำมารวมในมัลติมีเดียช่วยให้เกิดความหลากหลาย เพิ่มความสนุกสนานความน่าสนใจในการเรียนรู้มากขึ้น การนำคอมพิวเตอร์มาใช้ในลักษณะของการนำเสนอที่น่าสนใจ บทเรียนคอมพิวเตอร์มัลติมีเดีย ประกอบด้วยข้อความ ภาพนิ่ง ภาพเคลื่อนไหว ที่ทำงานเป็นลำดับขั้นกระบวนการ สื่อในรูปแบบมัลติมีเดียถือว่าเป็นสื่อการเรียนรู้ที่มีประสิทธิภาพ [2] การใช้บทเรียนคอมพิวเตอร์ในรูปแบบของมัลติมีเดียนั้น ได้รับความนิยมอย่างกว้างขวาง และแพร่หลายในปัจจุบัน [3]

จากเหตุผลข้างต้น ผู้วิจัยเล็งเห็นถึงความสำคัญในการนำสื่อมัลติมีเดียมาใช้พัฒนาการเรียนการสอน จึงได้พัฒนาบทเรียนคอมพิวเตอร์มัลติมีเดีย เรื่อง องค์ประกอบศิลป์เบื้องต้น สำหรับนักศึกษาสาขาวิชาเทคโนโลยีสารสนเทศ มหาวิทยาลัยราชภัฏนครปฐม เพื่อช่วยสนับสนุนการเรียนรู้ของผู้เรียนให้มีประสิทธิภาพมากยิ่งขึ้น

2.งานวิจัยที่เกี่ยวข้อง

2.1 งานวิจัยในประเทศ

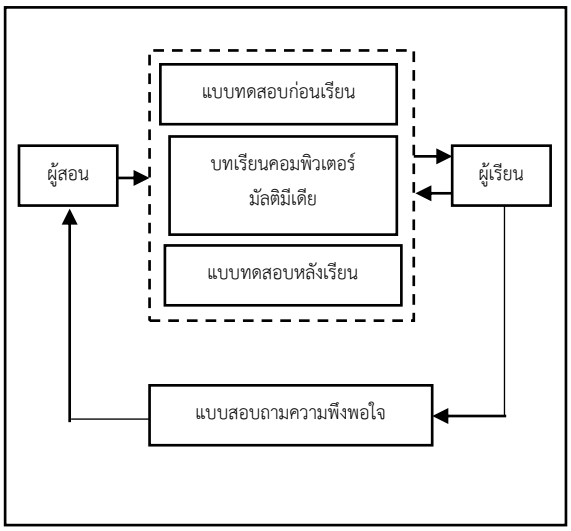
ถนอมพร ตันพิพัฒน์ (2549)ได้ทำการศึกษาวิจัยเพื่อพัฒนาบทเรียนคอมพิวเตอร์มัลติมีเดียการพัฒนากิจกรรมสำหรับหลักสูตรอินเทอร์เน็ตในห้องเรียน ผลการทดลองพบว่าบทเรียนที่สร้างขึ้นมีประสิทธิภาพตามเกณฑ์ 88/85 แสดงว่าบทเรียนที่สร้างขึ้นสามารถนำไปใช้ในการเรียนการสอนได้ [4]

กัลยา และนงลักษณ์ (2547) ได้ศึกษาเกี่ยวกับการสร้างและหาประสิทธิภาพของบทเรียนคอมพิวเตอร์ช่วยสอนมัลติมีเดีย ผลการทดลองพบว่า บทเรียนคอมพิวเตอร์ช่วยสอนที่สร้างขึ้นมีค่าประสิทธิภาพ 91.06/87.79 ซึ่งสูงกว่าเกณฑ์ 80/80 ที่กำหนดไว้แสดงว่าบทเรียนคอมพิวเตอร์ช่วยสอนที่สร้างขึ้น มีโต้ตอบกันได้ระหว่างผู้เรียนกับคอมพิวเตอร์ นอกจากนี้บทเรียนคอมพิวเตอร์มัลติมีเดียยังมีความสามารถในการตอบสนองต่อข้อมูลที่ผู้เรียนป้อนเข้าไปได้ทันที ซึ่งเป็นการเสริมแรงให้ผู้เรียนเกิดการเรียนรู้ [1] การใช้มัลติมีเดียในประสิทธิภาพเหมาะสมและสามารถนำไปใช้ในการเรียนการสอนได้ [5]

จากการศึกษางานวิจัย ผู้วิจัยสรุปได้ว่า สื่อมัลติมีเดียมีส่วนช่วยให้ผลสัมฤทธิ์ทางการเรียนของผู้เรียนสูงขึ้น ทำให้ผู้เรียนเกิดความสนใจ สามารถเลือกเรียนตามความสามารถของตนเอง และสื่อมัลติมีเดียยังสามารถปรับปรุงพัฒนาให้เข้ากับยุคสมัยของเทคโนโลยีการเรียนการสอนในปัจจุบันให้เกิดประสิทธิภาพมากยิ่งขึ้น

3. ขั้นตอนวิธีดำเนินการวิจัย

3.1 กรอบแนวคิดการวิจัย



ภาพที่ 1: กรอบแนวคิดการวิจัย

3.2 ขั้นตอนการดำเนินการ

ผู้วิจัยได้ดำเนินการทดลองเพื่อหาประสิทธิภาพของบทเรียนคอมพิวเตอร์มัลติมีเดีย เรื่อง องค์ประกอบศิลป์เบื้องต้น โดยให้กลุ่มตัวอย่างทำแบบทดสอบก่อนเรียนและหลังเรียน เพื่อตรวจสอบบทเรียนให้มีประสิทธิภาพตามเกณฑ์ 80/80 และหาผลสัมฤทธิ์การเรียนรู้

ประชากรที่ใช้ในการศึกษา คือ นักศึกษามหาวิทยาลัยราชภัฏนครปฐม ประจำปีการศึกษา 2/2559

กลุ่มตัวอย่างที่ใช้ในการศึกษา คือ นักศึกษามหาวิทยาลัยราชภัฏนครปฐม สาขาวิชาเทคโนโลยีสารสนเทศ (มัลติมีเดีย) จำนวน 36 คน ใช้วิธีการเลือกแบบเจาะจง

ศึกษาแนวคิด ทฤษฎี เอกสารที่เกี่ยวข้อง กำหนดแบบแผนการวิจัย

ก. ผู้สอนให้ผู้เรียนทำแบบทดสอบก่อนเรียน ผ่านบทเรียนคอมพิวเตอร์มัลติมีเดีย

ข. ผู้สอนให้ผู้เรียนศึกษาเนื้อหาการเรียนรู้เรื่อง องค์ประกอบศิลป์เบื้องต้น โดยใช้สื่อบทเรียนคอมพิวเตอร์มัลติมีเดีย

ค. ผู้สอนให้ผู้เรียนทำแบบทดสอบหลังเรียนผ่านบทเรียนคอมพิวเตอร์มัลติมีเดีย

ง. เปรียบเทียบผลสัมฤทธิ์ทางการเรียนก่อนเรียน และหลังเรียนผ่านบทเรียนคอมพิวเตอร์มัลติมีเดีย

จ. ผู้สอนให้ผู้เรียนทำแบบสอบถามความพึงพอใจหลังการใช้สื่อบทเรียนคอมพิวเตอร์มัลติมีเดีย เรื่อง องค์ประกอบศิลป์เบื้องต้น

3.3 เครื่องมือที่ใช้ในการวิจัย

การวิจัยเรื่อง การพัฒนาบทเรียนคอมพิวเตอร์มัลติมีเดีย เรื่อง องค์ประกอบศิลป์เบื้องต้น สำหรับนักศึกษาสาขาวิชาเทคโนโลยีสารสนเทศ มหาวิทยาลัยราชภัฏนครปฐม

3.3.1 บทเรียนคอมพิวเตอร์มัลติมีเดีย เรื่ององค์ประกอบศิลป์เบื้องต้น โดยคำนึงถึงลำดับการเรียนก่อนหลัง หรือเรียนแบบคู่ขนานกันได้โดยไม่เกี่ยวข้องกัน นำหัวข้อต่าง ๆ ที่ได้กำหนดมาแบ่งเป็นหน่วยการเรียนและกำหนดวัตถุประสงค์เชิงพฤติกรรมของแต่ละหน่วยให้สอดคล้องกับเนื้อหาที่จะนำเสนอ พัฒนากรอบเนื้อหาบทเรียน เขียนรายละเอียดเนื้อหา จัดลำดับเนื้อหาตามคำอธิบายรายวิชานำเนื้อหาทั้งหมดให้ผู้เชี่ยวชาญตรวจสอบความถูกต้อง และความเที่ยงตรงเชิงเนื้อหา (IOC) นำเนื้อหาที่ได้เตรียมไว้เข้าไปพัฒนาในซอฟต์แวร์ iSpring Suites และ Adobe Photoshop

3.3.2 แบบทดสอบก่อนเรียนและหลังเรียน ใช้เป็นแบบทดสอบสำหรับวัดผลการเรียนรู้ก่อนเรียนและหลังเรียนจากการใช้บทเรียนคอมพิวเตอร์มัลติมีเดีย ดำเนินการสร้างแบบทดสอบเพื่อประเมินคุณภาพ หาประสิทธิภาพนำแบบทดสอบไปทดลองกับกลุ่มตัวอย่างจากนั้นนำแบบทดสอบที่ผ่านกระบวนการทดสอบแล้วมาวิเคราะห์ เพื่อหาค่าความยากง่าย (P) อำนาจจำแนก (D) และความเชื่อมั่นของแบบทดสอบ โดยพิจารณาตามเกณฑ์ข้อสอบเลือกเฉพาะข้อที่มีระดับความยากง่ายอยู่ระหว่าง 0.20 – 0.80 และคัดเลือกเฉพาะข้อสอบที่มีค่าอำนาจจำแนก 0.20 ขึ้นไปวิเคราะห์ค่าความเชื่อมั่นโดยคำนวณหาค่าสัมประสิทธิ์ความเชื่อมั่นตามสูตร คูเดอร์ริชาร์ดสัน โดยค่าความเชื่อมั่นที่อยู่ในเกณฑ์ใช้ได้จะต้องมีค่าไม่ต่ำกว่า 0.60 คัดเลือกมาใช้ตามวัตถุประสงค์ของการวิจัย

3.3.3 แบบสอบถามความพึงพอใจ ของนักศึกษาสาขาวิชาเทคโนโลยีสารสนเทศ(มัลติมีเดีย) ที่มีต่อบทเรียนคอมพิวเตอร์มัลติมีเดีย แบบวิธี Likert มีเกณฑ์การให้คะแนน ดังนี้ [6]

ระดับความเหมาะสมมากที่สุด	5	คะแนน
ระดับความเหมาะสมมาก	4	คะแนน
ระดับความเหมาะสมปานกลาง	3	คะแนน
ระดับความเหมาะสมน้อย	2	คะแนน
ระดับความเหมาะสมน้อยที่สุด	1	คะแนน

การแปลความหมายค่าเฉลี่ยใช้เกณฑ์การแปลความหมาย ดังนี้

4.51 – 5.00	มีระดับความพึงพอใจมากที่สุด
3.51 – 4.50	วามพึงพอใจมากมีระดับค
2.51 – 3.50	มีระดับความพึงพอใจปานกลาง
1.51 – 2.50	มีระดับความพึงพอใจน้อย
1.00 – 1.50	มีระดับความพึงพอใจน้อยที่สุด

3.4 การวิเคราะห์ข้อมูล

นำผลคะแนนก่อนเรียนและหลังเรียนมาหาค่าทางสถิติ ได้แก่ ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน และเปรียบเทียบผลสัมฤทธิ์ทางการเรียนก่อนเรียนและหลังเรียน ด้วยวิธีสถิติทดสอบ t-test dependent [7] ดังนี้

$$T = \frac{\sum D}{\frac{\sqrt{n\sum D^2-(\sum D)^2}}{n-1}}$$

เมื่อ D = ความแตกต่างของคะแนนแต่ละคู่

N = จำนวนคู่

หาค่าเฉลี่ย (Mean) [8]

$$\bar{x} = \sum_{N} x$$

เมื่อ $\bar{X}$ = ค่าเฉลี่ย

$\sum x$ = ผลรวมของคะแนนทั้งหมด

N = จำนวนข้อมูลทั้งหมด

หาค่าเบี่ยงเบนมาตรฐาน

$$S.D. = \sqrt{\frac{N\sum x2(\sum x)2}{N(N1)}}$$

เมื่อ S.D. = ค่าความเบี่ยงเบนมาตรฐาน

$\sum x^2$ = ผลรวมของคะแนนยกกำลังสอง

$(\sum x)^2$ = กำลังสองของคะแนนผลรวม

N = จำนวนข้อมูลทั้งหมด

นำผลแบบสอบถามความพึงพอใจของผู้เรียนหลังเรียนด้วยบทเรียนคอมพิวเตอร์มัลติมีเดีย มาหาค่าทางสถิติ ได้แก่ ค่าเฉลี่ยและค่าส่วนเบี่ยงเบนมาตรฐาน

4. ผลการดำเนินการวิจัย

4.1 เปรียบเทียบผลสัมฤทธิ์ทางการเรียนของผู้เรียนด้วยบทเรียนคอมพิวเตอร์มัลติมีเดีย ก่อนเรียนและหลังเรียน

ตารางที่ 1: ผลการเปรียบเทียบผลสัมฤทธิ์ทางการเรียน

ลำดับ	N	คะแนนก่อนเรียน (10)		คะแนนหลังเรียน (10)		t-test	Sig.
		$\bar{X}$	S.D.	$\bar{X}$	S.D.		
1	36	4.17	1.13	8.72	0.85	22.91	.000*

จากตารางที่ 1 พบว่าผู้เรียนที่เรียนด้วยบทเรียนคอมพิวเตอร์มัลติมีเดีย มีคะแนนเฉลี่ยก่อนเรียนเท่ากับ 4.17 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 1.13 มีค่าคะแนนเฉลี่ยหลังเรียนเท่ากับ 8.72 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.85 ซึ่งมีค่าคะแนนเฉลี่ยสูงกว่าก่อนเรียน เมื่อเปรียบเทียบผลสัมฤทธิ์ทางการเรียนก่อนเรียนและหลังเรียนของผู้เรียน โดยทำการทดสอบค่าเฉลี่ยด้วยสถิติทดสอบ t-test พบว่า มีค่า t เท่ากับ 22.91 และค่า Sig เท่ากับ .000* ซึ่งน้อยกว่าระดับนัยสำคัญที่ 0.05 แสดงว่าผลคะแนนเฉลี่ยของหลังเรียนสูงกว่าก่อนเรียน ซึ่งผลสัมฤทธิ์ทางการเรียนของผู้เรียนที่เรียนด้วยบทเรียนคอมพิวเตอร์มัลติมีเดียอยู่ในระดับสูง

4.2 ความพึงพอใจของผู้เรียนหลังเรียนด้วยบทเรียนคอมพิวเตอร์มัลติมีเดีย

ตารางที่ 2: ผลความพึงพอใจของผู้เรียนหลังเรียน

ข้อที่	หัวข้อ	ระดับความพึงพอใจ		
		$\bar{X}$	S.D.	แปลผล
1	บทเรียนคอมพิวเตอร์มัลติมีเดียมีความน่าสนใจกว่าสื่ออื่น ๆ	4.25	0.76	มาก
2	บทเรียนคอมพิวเตอร์มัลติมีเดียใช้งานสะดวกเหมาะสมกับการจัดการเรียนการสอน	4.36	0.59	มาก
3	บทเรียนคอมพิวเตอร์มัลติมีเดีย มีสื่อภายในที่หลากหลาย	4.30	0.66	มาก
4	บทเรียนคอมพิวเตอร์มัลติมีเดีย มีเนื้อหาครบถ้วน	4.33	0.71	มาก
5	บทเรียนคอมพิวเตอร์มัลติมีเดีย ทำให้เกิดแรงจูงใจในการเรียนรู้	4.52	0.60	มากที่สุด
รวม		**4.35**	**0.073**	**มาก**

จากตารางที่ 2 พบว่า โดยรวมผู้เรียนหลังเรียนที่เรียนด้วยบทเรียนคอมพิวเตอร์มัลติมีเดีย มีความพึงพอใจอยู่ในเกณฑ์มากมีค่าเฉลี่ยเท่ากับ 4.35 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.07 ซึ่งผู้เรียนมีความพึงพอใจมากที่สุดจากบทเรียนคอมพิวเตอร์มัลติมีเดีย ทำให้เกิดแรงจูงใจในการเรียนรู้ มีค่าเฉลี่ยเท่ากับ 4.52 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.60 ตามด้วยก็คือ บทเรียนคอมพิวเตอร์มัลติมีเดียใช้งานสะดวกเหมาะสมกับการจัดการเรียนการสอน มีค่าเฉลี่ยเท่ากับ 4.36 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.59 ตามด้วยก็คือ บทเรียนคอมพิวเตอร์มัลติมีเดียมีเนื้อหาครบถ้วน มีค่าเฉลี่ยเท่ากับ 4.33 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.71 ตามด้วยก็คือ บทเรียนคอมพิวเตอร์มัลติมีเดียมีสื่อภายในที่หลากหลาย มีค่าเฉลี่ยเท่ากับ 4.30 ส่วนเบี่ยงเบน

มาตรฐานเท่ากับ 0.66 และบทเรียนคอมพิวเตอร์มัลติมีเดียมีความน่าสนใจกว่าสื่ออื่น ๆ มีค่าเฉลี่ยเท่ากับ 4.25 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.76 ตามลำดับ

5. สรุป

จากผลของการวิจัยการพัฒนาบทเรียนคอมพิวเตอร์มัลติมีเดีย เรื่อง องค์ประกอบศิลป์เบื้องต้น สำหรับนักศึกษาสาขาวิชาเทคโนโลยีสารสนเทศ มหาวิทยาลัยราชภัฏนครปฐม พบว่าผลสัมฤทธิ์ทางการเรียนของผู้เรียนหลังเรียนสูงกว่าก่อนเรียนซึ่งสอดคล้องกับวัตถุประสงค์ของการวิจัยที่กำหนดไว้ ทั้งนี้อาจเนื่องด้วยใช้สื่อมัลติมีเดียที่กอบด้วย ข้อความ รูปภาพ กราฟิก เสียง มาใช้ในการจัดการเรียนการสอน โดยทำการทดสอบค่าเฉลี่ยด้วยสถิติทดสอบ t-test พบว่า มีค่า t เท่ากับ 0.2291 และค่า Sig เท่ากับ .000* ระดับนัยสำคัญที่ 0.05 แสดงว่าผลคะแนนเฉลี่ยของผู้เรียนหลังเรียนสูงกว่าก่อนเรียน และผลสัมฤทธิ์ทางการเรียนหลังเรียนของผู้เรียนที่เรียนด้วยบทเรียนคอมพิวเตอร์มัลติมีเดียนั้นอยู่ในระดับสูงและจากแบบสอบถามความพึงพอใจของผู้เรียนที่เรียนด้วยบทเรียนคอมพิวเตอร์มัลติมีเดียพบว่า ผู้เรียนมีความพึงพอใจอยู่ในเกณฑ์มาก

6. เอกสารอ้างอิง

[1] นงเยาว์ เอี่ยมภาคีนิวัฒน์, "*การพัฒนาบทเรียนคอมพิวเตอร์ช่วยสอนบนอินเทอร์เน็ตเรื่องฐานข้อมูลเบื้องต้น*," คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏนครสวรรค์ นครสวรรค์ ปีที่ 2550.

[2] รุจโรจน์ แก้วอุไร, *"การออกแบบบทเรียนคอมพิวเตอร์ช่วยสอนตามแนวคิดของกาเย่*," กรุงเทพฯ ปีที่ 2548.

[3] จตุรงค์ ทองรวย, "*การพัฒนาบทเรียนคอมพิวเตอร์ช่วยสอน วิชาเทคโนโลยีการถ่ายภาพ*," บัณฑิตมหาวิทยาลัย มหาวิทยาลัยบูรพา ชลบุรี ปีที่ 2547.

[4] ถนอมพร ตันพิพัฒน์ , "*การพัฒนากิจกรรมสำหรับหลักสูตรอินเทอร์เน็ตในห้องเรียน*," สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี ปีที่ 2549.

[5] กัลยา อุบลรัตน์ และนงลักษณ์ ตั้งจิตตชอบ , "*การสร้างและหาประสิทธิภาพบทเรียนคอมพิวเตอร์ช่วยสอนระบบมัลติมีเดีย*," สำนักพัฒนาเทคนิคศึกษา สถาบันเทคโนโลยีพระจอมเกล้าพระนครเหนือ ปีที่ 2547.

[6] ชนาธิป พรกุล, *การสอนกระบวนการคิด ทฤษฎีและการนำไปใช้*, จุฬาลงกรณ์มหาวิทยาลัย กรุงเทพฯ :2554.

[7] ชูศรี วงศ์รัตนะ, *สถิติเพื่อการวิจัย*,โรงพิมพ์เจริญพร พิมพ์ครั้งที่ 3 . กรุงเทพฯ , 2534.

[8] บุญชม ศรีสะอาด, *การวิจัยเบื้องต้น* ,สำนักพิมพ์ สุวีริยาสาส์น มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร พิมพ์ครั้งที่ 2 .กรุงเทพฯ ,2535.

การพัฒนาบทเรียนคอมพิวเตอร์ช่วยสอนผ่านเว็บตามแผนการสอนฐานสมรรถนะตามขั้นตอนการเรียนรู้แบบ MIAP ในรายวิชาการศึกษาวงจรและซ่อมบำรุงไมโครคอมพิวเตอร์ โดยใช้โครงงานเป็นฐานร่วมกับการจัดการเรียนรู้แบบผสมผสาน

The Development of WBI for Competency-Based Teaching Plan MIAP Learning in Circuit Description and Microcomputer Maintenance using PjBL with Blended Learning

วินัย เพ็งภิญโญ (Winai Phengpinyo)[1] และกฤช สินธนะกุล (Krich Sintanakul)[2]

[1]สาขาวิชาคอมพิวเตอร์ศึกษา คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏนครปฐม

[2]ภาควิชาคอมพิวเตอร์ศึกษา คณะครุศาสตร์อุตสาหกรรม มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

[1]Winai36@hotmail, [2]Krich.s@fte.kmutnb.ac.th

บทคัดย่อ

การวิจัยครั้งนี้เป็นการวิจัยเชิงทดลองโดยมีวัตถุประสงค์เพื่อ 1) พัฒนาบทเรียนคอมพิวเตอร์ช่วยสอนผ่านเว็บตามแผนการสอนฐานสมรรถนะ ในรายวิชาการศึกษาวงจรและซ่อมบำรุงไมโครคอมพิวเตอร์ โดยใช้โครงงานเป็นฐานร่วมกับการเรียนรู้แบบผสมผสาน 2) หาประสิทธิภาพของบทเรียนคอมพิวเตอร์ช่วยสอนผ่านเว็บที่พัฒนาขึ้น 3) เปรียบเทียบผลสัมฤทธิ์ทางการเรียนของผู้เรียนก่อนและหลังเรียนด้วยบทเรียนคอมพิวเตอร์ช่วยสอนผ่านเว็บที่พัฒนาขึ้น 4) หาความพึงพอใจของผู้เรียนที่มีต่อบทเรียนคอมพิวเตอร์ช่วยสอนผ่านเว็บ โดยกลุ่มเป้าหมายที่ใช้ในการวิจัยครั้งนี้ คือ นักศึกษาระดับปริญญาตรี ชั้นปีที่ 3 มหาวิทยาลัยราชภัฏนครปฐม จำนวน 30 คน เครื่องมือที่ใช้ในการวิจัยได้แก่ แผนการสอนฐานสมรรถนะตามขั้นตอน MIAP, บทเรียนคอมพิวเตอร์แบบสาธิตร่วมกับการจัดการเรียนการสอนแบบโครงงานเป็นฐานและการจัดการเรียนรู้แบบผสมผสาน, แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนและแบบสอบถามความพึงพอใจ สถิติที่ใช้ในการวิจัยได้แก่ ค่าเฉลี่ย, ส่วนเบี่ยงเบนมาตรฐาน, ประสิทธิภาพ E1/E2 และค่าสถิติที

ผลการวิจัยพบว่า 1) บทเรียนคอมพิวเตอร์ช่วยสอนผ่านเว็บที่ผู้วิจัยได้พัฒนาขึ้นมีประสิทธิภาพเท่ากับ 84.13/82.04 สูงกว่าเกณฑ์ที่ตั้งไว้ 80/80 2) ผลสัมฤทธิ์ทางการเรียนของผู้เรียนหลังเรียนด้วยบทเรียนคอมพิวเตอร์ช่วยสอนผ่านเว็บที่พัฒนาขึ้นสูงกว่าก่อนเรียนอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 และ 3) ผลความพึงพอใจของผู้เรียนที่มีต่อบทเรียนคอมพิวเตอร์ช่วยสอนผ่านเว็บโดยรวมอยู่ในระดับมาก

คำสำคัญ: ฐานสมรรถนะ การจัดการเรียนรู้แบบผสมผสาน

Abstract

The purposes of this research were to: 1) developed a WBI for Competency-Based Teaching Plan in Circuit Description and Microcomputer Maintenance Using PjBL with Blended Learning, 2) find performance of developed WBI 3) compare the students learning achievements before and after learning with the developed WBI, 4) study the students' satisfaction through the WBI. The target group, was 30 undergraduate students of Nakhon Pathom Rajabhat University, which derived from drawing a simple random sampling. The research tools used Competency-Based Teaching Plan MIAP Learning, Achievements Test, and questionnaire. The statistics used mean, standard deviation, E1/E2 and t-test.

The result findings showed: 1) The performance of WBI developed with the 84.13/81.28 which was higher than the criteria defined is 80/80 2) the students learning achievements after learned higher than before learned was significantly at .01 levels; and 3) The results of the students' satisfaction to WBI were at a high level.

Keywords : WBI, Competency-based, Blended Learning.

1. บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

พระราชบัญญัติการศึกษาแห่งชาติ พ.ศ. 2542 แก้ไขเพิ่มเติม (ฉบับที่ 2) พ.ศ. 2545 และ (ฉบับที่ 3) พ.ศ. 2553 ได้กล่าวถึงแนวทางการจัดการศึกษาไว้ในหมวด 4 มาตรา 22 การจัดการศึกษาต้องยึดหลักว่าผู้เรียนทุกคนมีความสามารถเรียนรู้และพัฒนาตนเองได้ และถือว่าผู้เรียนมีความสำคัญที่สุด กระบวนการจัดการศึกษาต้องส่งเสริมให้ผู้เรียนสามารถพัฒนาตามธรรมชาติและเต็มตามศักยภาพ นอกจากนี้ในหมวด 9 ว่าด้วยเทคโนโลยีเพื่อการศึกษา มาตรา 64 รัฐต้องส่งเสริมและสนับสนุนให้มีการผลิต และพัฒนาแบบเรียน ตารา หนังสือทางวิชาการ สื่อสิ่งพิมพ์อื่น วัสดุอุปกรณ์ และเทคโนโลยีเพื่อการศึกษาอื่นโดยเร่งรัดพัฒนาขีดความสามารถในการผลิต จัดให้มีเงินสนับสนุนการผลิตและมีการให้แรงจูงใจแก่ผู้ผลิต และพัฒนาเทคโนโลยีเพื่อการศึกษาทั้งนี้ โดยเปิดให้มีการแข่งขันโดยเสรีอย่างเป็นธรรม และมาตรา 66 ผู้เรียนมีสิทธิได้รับการพัฒนาขีดความสามารถในการใช้เทคโนโลยีเพื่อการศึกษาในโอกาสแรกที่ทาได้ เพื่อให้มีความรู้และทักษะเพียงพอที่จะใช้เทคโนโลยีเพื่อการศึกษาในการแสวงหาความรู้ด้วยตนเองได้อย่างต่อเนื่องตลอดชีวิต[1]

ปัจจุบันการจัดการเรียนการสอนระดับปริญญาตรีในประเทศไทย พบว่ามีหลายประเด็นที่เป็นปัญหาสำคัญ เช่น การเรียนการสอนไม่เอื้อให้นักศึกษาคิดเป็น ทำเป็น แก้ปัญหาเป็น การเรียนการสอนที่เน้นการท่องจำมากกว่าการเน้นให้ผู้เรียนได้คิด ได้ลงมือปฏิบัติ ขาดการแสวงหาความรู้ ขาดการปฏิสัมพันธ์กับเพื่อน ครู และชุมชนสิ่งแวดล้อม เพราะการศึกษาเน้นการฟังบรรยายภายใต้กรอบอันจำกัดของห้องเรียนทำให้เป็นตัวขัดขวางการพัฒนาการเรียนรู้ และการเรียนการสอนส่วนใหญ่ในสถาบันการศึกษาเป็นแบบการเรียนการสอนปกติในห้องคือมีอาจารย์เป็นผู้สอนโดยใช้วิธีบรรยาย (Lecture) ซึ่งได้รับผลกระทบเป็นอย่างมากจากการปฏิรูปการศึกษาที่เน้นผู้เรียนเป็นสำคัญ [2]

การเรียนการสอนผ่านเว็บ (Web-Based Instruction) เป็นการผสมผสานกันระหว่างเทคโนโลยี ปัจจุบันกับกระบวนการออกแบบการเรียนการสอน เพื่อเพิ่มประสิทธิภาพทางการเรียนรู้และแก้ปัญหาในเรื่องข้อจำกัดทางด้านสถานที่และเวลา โดยการสอนบนเว็บจะประยุกต์ใช้คุณสมบัติและทรัพยากรของเวิลด์ ไวด์ เว็บในการจัดสภาพแวดล้อมที่ส่งเสริมและสนับสนุนการเรียนการสอน ซึ่งการเรียนการสอนที่จัดขึ้นผ่านเว็บนี้อาจเป็นบางส่วนหรือทั้งหมดของกระบวนการเรียนการสอนก็ได้ เวิลด์ ไวด์ เว็บ เป็นบริการบนเครือข่ายอินเทอร์เน็ตซึ่งได้รับความนิยมอย่างแพร่หลายในปัจจุบันเริ่มเข้ามาเป็นที่รู้จักในวงการศึกษาในประเทศไทยตั้งแต่ พ.ศ. 2538 ที่ผ่านมาเว็บได้เข้ามามีบทบาทสำคัญทางการศึกษาและกลายเป็นคลังแห่งความรู้ที่ไร้พรมแดน ซึ่งผู้สอนได้ใช้เป็นทางเลือกใหม่ในการส่งเสริมการเรียนรู้เพื่อเปิดประตูการศึกษาจากห้องเรียนไปสู่โลกแห่งการเรียนรู้อันกว้างใหญ่ รวมทั้งการนำการศึกษาไปสู่ผู้ที่ขาดโอกาสด้วยข้อจำกัดทางด้านเวลาและสถานที่ [3]

ด้วยเหตุผลที่กล่าวมาข้างต้น ผู้วิจัยจึงเห็นควรที่จะพัฒนาบทเรียนคอมพิวเตอร์ช่วยสอนผ่านเว็บตามแผนการสอนฐานสมรรถนะ โดยใช้โครงงานเป็นฐานร่วมกับการเรียนรู้แบบผสมผสาน ในรายวิชาการศึกษาวงจรและซ่อมบำรุงไมโครคอมพิวเตอร์ มาใช้ในการจัดการเรียนการสอนเพื่อสนับสนุนการเรียนที่เน้นฐานสมรรถนะพื้นฐานเกี่ยวกับการซ่อมบำรุงไมโครคอมพิวเตอร์ โดยมุ่งหวังให้ผู้เรียนแต่ละคนสามารถสร้างความรู้ผ่านกระบวนการคิด และลงมือปฏิบัติด้วยตนเอง อีกทั้งยังเป็นการฝึกทบทวนให้ผู้เรียนเกิดการเรียนรู้อย่างมีประสิทธิภาพ ส่งผลให้ผลสัมฤทธิ์ทางการเรียนดีขึ้น รวมทั้งยังสามารถนำไปประยุกต์ใช้กับรายวิชาอื่น ๆ ได้

1.2 วัตถุประสงค์ของงานวิจัย

1.2.1 เพื่อพัฒนาบทเรียนคอมพิวเตอร์ช่วยสอนผ่านเว็บตามแผนการสอนฐานสมรรถนะในรายวิชาการศึกษาวงจรและซ่อมบำรุงไมโครคอมพิวเตอร์โดยใช้โครงงานเป็นฐานร่วมกับการเรียนรู้แบบผสมผสาน

1.2.2 เพื่อหาประสิทธิภาพของบทเรียนคอมพิวเตอร์ช่วยสอนผ่านเว็บที่พัฒนาขึ้น

1.2.3 เพื่อเปรียบเทียบผลสัมฤทธิ์ทางการเรียนของผู้เรียนก่อนและหลังเรียนด้วยบทเรียนคอมพิวเตอร์ช่วยสอนผ่านเว็บที่พัฒนาขึ้น

1.2.4 เพื่อหาความพึงพอใจของผู้เรียนที่มีต่อบทเรียนคอมพิวเตอร์ช่วยสอนผ่านเว็บ

2. วรรณกรรมที่เกี่ยวข้อง

2.1 บทเรียนคอมพิวเตอร์ช่วยสอนผ่านเว็บ

การเรียนการสอนผ่านเว็บเป็นการใช้เว็บในการเรียนการสอนโดยอาจใช้เว็บเพื่อนาเสนอบทเรียนในลักษณะสื่อหลายมิติของวิชาทั้งหมดตามหลักสูตร หรือใช้เพียงการเสนอข้อมูลบางอย่างเพื่อประกอบการสอนก็ได้ รวมทั้งใช้ประโยชน์จากคุณลักษณะต่างๆของการสื่อสารที่มีอยู่ในระบบอินเทอร์เน็ต เช่น การเขียนโต้ตอบกันทางไปรษณีย์อิเล็กทรอนิกส์และการพูดคุยสดด้วยข้อความและเสียงมาใช้ประกอบด้วยเพื่อให้เกิดประสิทธิภาพสูงสุด [4]

2.2 การเรียนรู้แบบโครงงานเป็นฐาน

การเรียนรู้โดยใช้โครงงานเป็นฐานหรือการเรียนรู้แบบโครงงาน (Project-based Learning) เป็นการเรียนรู้ที่เชื่อมโยงหลักการพัฒนาการคิดของบลูม (Bloom) ทั้ง 6 ขั้น คือ ความรู้ความจำ (Knowledge) ความเข้าใจ (Comprehension) การนำไปใช้ (Application) การวิเคราะห์ (Analysis) การสังเคราะห์ (Synthesis) และการประเมินค่า (Evaluation) เป็นกระบวนการเรียนรู้ที่เน้นผู้เรียนเป็นสำคัญในทุกขั้นตอนของการเรียนรู้ ตั้งแต่การวางแผนการเรียนรู้ การออกแบบการเรียนรู้ การสร้างสรรค์ ประยุกต์ใช้ผลผลิต และการประเมินงาน [5]

การเรียนรู้โดยใช้โครงงานเป็นฐานเป็นวิธีการจัดการเรียนที่เน้นผู้เรียนเป็นสำคัญวิธีการหนึ่งที่จะช่วยพัฒนาผู้เรียนทั้งด้านความรู้และทักษะการทำงานที่มีการค้นคว้าและการใช้ความรู้ในชีวิตจริงโดยมีตัวผลงานและการแสดงออกถึงศักยภาพจากการเรียนรู้

ขั้นตอนของการจัดการเรียนรู้แบบโครงงานแบ่งออกได้เป็น 3 ขั้นตอนใหญ่ๆ คือ

2.2.1 ระยะที่ 1 การเริ่มต้นโครงงาน ผู้สอนต้องสังเกตหรือสร้างความสนใจให้เกิดขึ้นในตัวผู้เรียน จากนั้นผู้สอนและผู้เรียนร่วมกันกำหนดเรื่องที่ต้องการศึกษา

ในการกำหนดหัวข้อโครงงาน ควรให้ผู้เรียนร่วมกันอภิปรายเกี่ยวกับเรื่องที่ตนเองสนใจ โดยอาจกำหนดหัวข้อโครงงานภายใต้ความสนใจของผู้เรียนทั้งกลุ่มหรือผู้เรียนส่วนใหญ่ จากนั้นผู้เรียนแลกเปลี่ยนความรู้เบื้องต้นเกี่ยวกับหัวข้อโครงงาน

2.2.2 ระยะที่ 2 ขั้นพัฒนาโครงงาน เป็นขั้นตอนที่ผู้เรียนกำหนดคำถามหรือประเด็นปัญหาที่สนใจอยากรู้ ตั้งสมมติฐานมาตอบคำถามเหล่านั้น จากนั้นทดสอบสมติฐานด้วยการลงมือปฏิบัติจนค้นพบคำตอบด้วยตนเอง ซึ่งอาจต้องมีการศึกษาหาความรู้เพิ่มเติมเพื่อตรวจสอบสมมติฐานที่ตั้งไว้

2.2.3 ระยะที่ 3 ขั้นสรุป เมื่อผู้เรียนค้นพบคำตอบของปัญหาและเริ่มแสดงให้เห็นว่าได้สิ้นสุดความสนใจในหัวข้อโครงงานและเริ่มสนใจเรื่องใหม่ ผู้สอนควรแนะนำผู้เรียนเข้าสู่ขั้นตอนการรวบรวมและสรุปผล โดยให้ผู้เรียนเขียนสรุปรายงาน และนำเสนอผลงาน

2.3 สมรรถนะ

สมรรถนะ (Competency) หมายถึง ความสามารถในการทำบางสิ่งได้เป็นอย่างดี ความสามารถในการทำงานหรือการประกอบอาชีพ ซึ่งในการทำงานหรือการประกอบอาชีพนั้นต้องใช้ความสามารถที่มีอยู่ในตัวบุคคล เพื่อจะทำภารกิจของงานนั้น ถ้าบุคคลใดมีความสามารถในการทำงานได้ เรียกว่าเป็นคนที่มีสมรรถนะในการทำงาน และในทางตรงข้ามถ้าบุคคลใดไม่สามารถทำงานได้ก็เรียกว่าเป็นคนไม่มีสมรรถนะ การสร้างเสริมให้คนมีความสามารถในการทำงานเป็นการสร้างสมรรถนะให้เกิดขึ้นใน ตัวบุคคล การฝึกอบรมคนให้มีสมรรถนะสำหรับการทำงานจึงเป็นสาระสำคัญของการจัดการเรียนการสอน [6]

2.4 การเรียนแบบผสมผสาน

มนต์ชัย [7] กล่าวว่า การเรียนรู้แบบผสมผสานหมายถึง การบูรณาการระหว่างการเรียนรู้ แบบเผชิญหน้าในชั้นเรียนโดยมีผู้สอนเป็นผู้กำกับการเรียนรู้แบบออนไลน์ซึ่งใช้เทคโนโลยีคอมพิวเตอร์ และเทคโนโลยีการสื่อสารเพื่อให้การเรียนรู้มีประสิทธิภาพสูงสุด ภายใต้สภาพแวดล้อมของชุมชนแห่งการเรียนรู้ เป็นการใช้ประโยชน์จาก ICT ส่งผ่านความรู้และติดต่อสื่อสารระหว่างผู้เรียนกับ ผู้สอนหรือระหว่างผู้เรียนด้วยกันที่เชื่อมต่อเข้าด้วยกันในระยะไกล

การเรียนรู้แบบผสมผสาน เป็นการบูรณาการร่วมกันโดยการผสมผสานลักษณะการจัดการ เรียนรู้ที่มีบริบทแตกต่างกัน กล่าวคือด้านแรกเป็นลักษณะการเรียนรู้แบบดั้งเดิมในชั้นเรียน (Traditional Classroom) ที่ผู้สอนและผู้เรียนมีปฏิสัมพันธ์กันแบบเผชิญหน้า (Face-to-Face) อีกด้านหนึ่งเป็นลักษณะการ

เรียนรู้ที่นำเอาความก้าวหน้าทางด้านเทคโนโลยีสื่อคอมพิวเตอร์ (Computer-Mediated) แบบออนไลน์เข้ามามีส่วนร่วมในกิจกรรมการเรียนรู้เพื่อให้ผู้เรียนเกิดการเรียนรู้ แม้ว่าไม่ได้อยู่ภายในชั้นเรียน อัตราส่วนที่นิยมในการจัดการเรียนรู้แบบผสมผสานการเรียนรู้โดยใช้สื่อคอมพิวเตอร์ออนไลน์นอกชั้นเรียนประมาณ 30 - 70% ซึ่งการแบ่งช่วงของการ เรียนรู้ภายในชั้นเรียนและภายนอกชั้นเรียนขึ้นอยู่กับความเหมาะสมของเนื้อหา วัตถุประสงค์ ตลอดจน ความพร้อมของอุปกรณ์ ผู้เรียนและผู้สอน

3. วิธีการดำเนินการวิจัย

การพัฒนาบทเรียนคอมพิวเตอร์ช่วยสอนผ่านเว็บตามแผนการสอนฐานสมรรถนะในรายวิชาการศึกษาวงจรและซ่อมบำรุงไมโครคอมพิวเตอร์ โดยใช้โครงงานเป็นฐานร่วมกับการจัดการเรียนรู้แบบผสมผสาน มีวิธีดำเนินการพัฒนาบทเรียนตามขั้นตอนดังนี้

3.1 ประชากรและกลุ่มตัวอย่าง

ประชากร คือ นักศึกษาระดับปริญญาตรีชั้นปีที่ 3 สาขาวิชาคอมพิวเตอร์ศึกษา มหาวิทยาลัยราชภัฏนครปฐม ที่ลงทะเบียนเรียนในภาคเรียนที่ 1 ปีการศึกษา 2559 จำนวนรวม 82 คน

กลุ่มตัวอย่าง คือ นักศึกษาระดับปริญญาตรีชั้นปีที่ 3 สาขาวิชาคอมพิวเตอร์ศึกษา ลงทะเบียนเรียนรายวิชาการศึกษาวงจรและซ่อมบำรุงไมโครคอมพิวเตอร์ ในปีการศึกษาที่ 1/2559 ห้อง 57/12 จำนวน 30 คน ซึ่งได้มาโดยวิธีการสุ่มอย่างง่ายด้วยวิธีการจับฉลาก

3.2 ศึกษาและวิเคราะห์เนื้อหารายวิชา (Analysis)

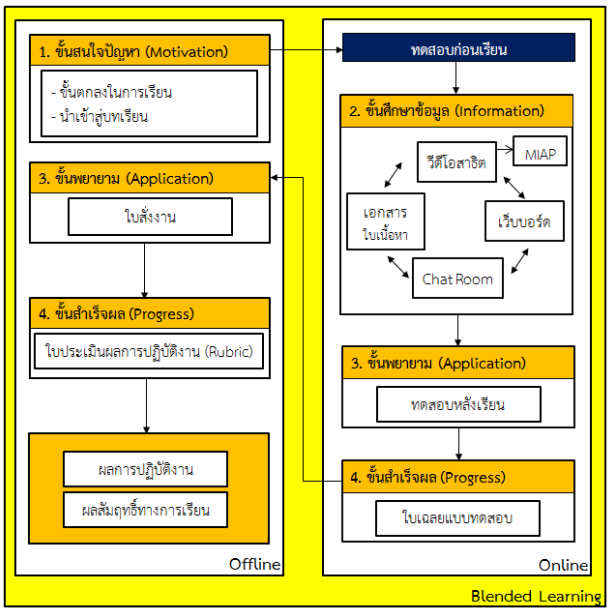
ศึกษาเนื้อหา คำอธิบายรายวิชา วัตถุประสงค์ทั่วไปในรายวิชาการศึกษาวงจรและซ่อมบำรุงไมโครคอมพิวเตอร์ แล้วทำการกำหนดวัตถุประสงค์เชิงพฤติกรรม วิเคราะห์ลำดับขั้นตอนการปฏิบัติงาน ลำดับของเนื้อหา แล้วทำการออกแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียน แล้วนำไปประเมินความสอดคล้อง (IOC) โดยผู้เชี่ยวชาญและหาคุณภาพของแบบทดสอบ จากนั้นวิเคราะห์การสอนแบบโครงงานเป็นฐานการเรียนแบบผสมผสาน การเรียนการสอนแบบ MIAP และเกณฑ์การประเมินแบบรูบริค เพื่อการวัดประเมินผลให้มีความเหมาะสมสอดคล้องกัน

3.2 ออกแบบบทเรียน (Design)

3.3.1 สร้างเอกสารประกอบบทเรียน ได้แก่ ตัวอย่างชิ้นงาน ใบงาน ใบสั่งงาน

3.3.2 เนื้อหาประกอบด้วยเนื้อหาจำนวน 3 หน่วย ได้แก่ 1) ความรู้เบื้องต้นเกี่ยวกับคอมพิวเตอร์ 2) ส่วนประกอบของเครื่องคอมพิวเตอร์ และ 3) การติดตั้งและประกอบเครื่องไมโครคอมพิวเตอร์

3.3.2 ออกแบบรูปแบบกิจกรรมการเรียนรู้ตามแผนการสอนฐานสมรรถนะโดยใช้โครงงานเป็นฐานร่วมกับการจัดการเรียนรู้แบบผสมผสาน ดังภาพที่ 1



ภาพที่ 1: รูปแบบกิจกรรมการเรียนรู้ตามแผนการสอนฐานสมรรถนะโดยใช้โครงงานเป็นฐานร่วมกับการจัดการเรียนรู้แบบผสมผสาน

3.3 พัฒนาบทเรียน (Development)

3.3.1 เตรียมการ เป็นการเตรียม ข้อความ ภาพนิ่ง เสียงบรรยาย และวีดีโอสาธิต

3.3.2 สร้างบทเรียน เป็นการนำข้อมูลที่เตรียมการมาทั้งหมดมาสร้างบทเรียนคอมพิวเตอร์ช่วยสอน จากนั้นติดตั้งบนระบบการจัดการเรียนรู้ผ่านเว็บ และทำการสร้างกิจกรรมการเรียนการสอน

3.4 การทดลองใช้ (Implementation)

3.4.1 ทดลองใช้โดยผู้พัฒนาและนำมาปรับปรุงแก้ไข

3.4.2 ประเมินคุณภาพบทเรียน โดยผู้เชี่ยวชาญด้านเนื้อหาและด้านเทคนิคการผลิตสื่อ ด้านละ 3 ท่าน โดยใช้แบบ

ประเมินคุณภาพของบทเรียน ซึ่งมีผลเฉลี่ยอยู่ในระดับดีมาก และนำข้อเสนอแนะไปปรับปรุงแก้ไขบทเรียน

3.4.3 ทดลองใช้กับกลุ่มย่อย ซึ่งไม่ใช่กลุ่มเป้าหมายจริง แต่มีลักษณะเทียบเคียงกับกลุ่มเป้าหมายจริง แล้วสอบถามความคิดเห็นและนำผลที่ได้ไปแก้ไขปรับปรุงบทเรียน

3.4.4 ทดลองใช้บทเรียนที่พัฒนาขึ้นกับกลุ่มเป้าหมาย คือ นักศึกษาระดับปริญญาตรี ชั้นปีที่ 3 มหาวิทยาลัยราชภัฏนครปฐม ที่ลงทะเบียนเรียนรายวิชาการศึกษาวงจรและซ่อมบำรุงไมโครคอมพิวเตอร์ ในปีการศึกษาที่ 1/2559 จำนวน 30 คน โดยใช้วิธีการเรียนรู้แบบผสมผสาน ด้วยการให้ผู้เรียนเรียนรู้ผ่านเว็บ หลักจากนั้นก็ให้ผู้เรียนได้ฝึกปฏิบัติด้วยเครื่องมือที่เป็นของจริง

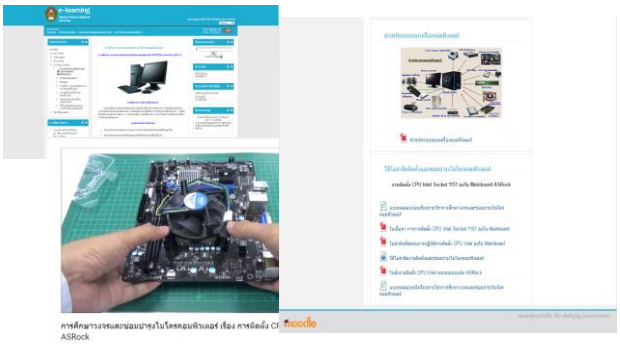
3.5 ประเมินผลข้อมูล (Evaluation)

ประเมินผลข้อมูลที่ได้จากการดำเนินการเก็บข้อมูลกับกลุ่มเป้าหมายโดยให้ผู้เรียนทำแบบทำสอบก่อนเรียนและหลังเรียนจำนวน 20 ข้อ และแบบประเมินผลการปฏิบัติงาน สถิติที่ใช้ได้แก่ ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน ประสิทธิภาพ E1/E2 และ t-test แบบ Dependent

4. ผลการดำเนินการวิจัย

4.1 ผลการพัฒนาบทเรียน

การพัฒนาบทเรียนคอมพิวเตอร์ช่วยสอนผ่านเว็บตามแผนการสอนฐานสมรรถนะในรายวิชาการศึกษาวงจรและซ่อมบำรุงไมโครคอมพิวเตอร์ โดยใช้โครงงานเป็นฐานร่วมกับการจัดการเรียนรู้แบบผสมผสานที่พัฒนาขึ้น



ภาพที่ 2: หน้าจอการนำเสนอเนื้อหาบทเรียนคอมพิวเตอร์ผ่านเว็บและวิดีโอสาธิต

4.2 ผลการหาประสิทธิภาพของบทเรียน

ประสิทธิภาพของบทเรียนเท่ากับ 82.50/81.04 สูงกว่าเกณฑ์ที่กำหนด 80/80 โดยพิจารณาจากผลคะแนนการทำแบบทดสอบระหว่างเรียน (E1) และแบบทดสอบหลังเรียน (E2) สรุปได้ว่าบทเรียนที่พัฒนาขึ้น สามารถนำมาใช้ในการจัดการเรียนการสอนได้ ดังตารางที่ 1

ตารางที่ 1: ผลการหาประสิทธิภาพของบทเรียนคอมพิวเตอร์ผ่านเว็บ

แบบทดสอบ	คะแนนเฉลี่ย	S.D.	ประสิทธิภาพ
ระหว่างเรียน (E1)	24.75	1.74	82.50
หลังเรียน (E2)	18.05	1.96	82.04

4.3 ผลการเปรียบเทียบผลสัมฤทธิ์ทางการเรียน

ผู้เรียนมีผลสัมฤทธิ์ทางการเรียนสูงขึ้นอย่างมีนัยสำคัญที่ระดับ .01 โดยพิจารณาจากการเปรียบเทียบคะแนนการทำแบบทดสอบก่อนเรียน และแบบทดสอบหลังเรียน จำนวน 20 ข้อ คะแนนเต็ม 20 คะแนน โดยการทดสอบค่าที ดังตารางที่ 2

ตารางที่ 2: ผลการเปรียบเทียบผลสัมฤทธิ์ทางการเรียนก่อนเรียนและหลังเรียน

แบบทดสอบ	n	$\bar{X}$	S.D.	$t_{คำนวณ}$	$t_{ตาราง}$
ก่อนเรียน	30	6.60	3.23	*13.8172	2.756
หลังเรียน	30	18.05	1.96		

*มีนัยสำคัญทางสถิติที่ระดับ .01

4.4 หาความพึงพอใจของผู้เรียนที่มีต่อบทเรียนคอมพิวเตอร์ช่วยสอนผ่านเว็บ

ผลการศึกษาความพึงพอใจของผู้เรียนที่มีต่อบทเรียนคอมพิวเตอร์ผ่านเว็บตามแผนการสอนฐานสมรรถนะในรายวิชาการศึกษาวงจรและซ่อมบำรุงไมโครคอมพิวเตอร์ โดยใช้โครงงานเป็นฐานร่วมกับการจัดการเรียนรู้แบบผสมผสานที่พัฒนาขึ้น พบว่าผู้เรียนมีความมีความพึงพอใจโดยรวมอยู่ในระดับมากที่สุด ดังตารางที่ 3

ตางรางที่ 3: ผลความพึงพอใจของผู้เรียนที่มีต่อบทเรียน

รายการ	$\bar{X}$	S.D.	ผลประเมิน
1. การนำเสนอเนื้อหาบทเรียนคอมพิวเตอร์	4.52	0.62	มากที่สุด
2. การออกแบบบทเรียนคอมพิวเตอร์	4.57	0.49	มากที่สุด
3.ประโยชน์จากการเรียนด้วยบทเรียนคอมพิวเตอร์	4.55	0.54	มากที่สุด
รวม	4.55	0.54	มากที่สุด

5. อภิปรายผลการวิจัย

จากผลการวิจัยการพัฒนาบทเรียนคอมพิวเตอร์ช่วยสอนผ่านเว็บตามแผนการสอนฐานสมรรถนะในรายวิชาการศึกษาวงจรและซ่อมบำรุงไมโครคอมพิวเตอร์ โดยใช้โครงงานเป็นฐานร่วมกับการจัดการเรียนรู้แบบผสมผสานที่พัฒนาขึ้น พบว่าประสิทธิภาพทางการเรียนของผู้เรียนสูงขึ้นกว่าเกณฑ์ที่กำหนดไว้ เนื่องมาจากผู้วิจัยได้พัฒนาบทเรียนคอมพิวเตอร์ช่วยสอนผ่านเว็บโดยใช้ฐานสมรรถนะ

ผู้เรียนมีผลสัมฤทธิ์ทางการเรียนหลังเรียนสูงขึ้นกว่าก่อนเรียนอย่างมีนัยสำคัญทางสถิติที่ .01 เนื่องมาจากผู้วิจัยได้ศึกษาจากหนังสือความรู้เกี่ยวกับสื่อการเรียนคอมพิวเตอร์ผ่านเว็บ ศึกษาหลักการวิธีการและรายละเอียดเกี่ยวกับขั้นตอนการจัดการเรียนรู้แบบผสมผสานจากเอกสารการวิจัยต่าง ๆ กำหนดเนื้อหาของบทเรียนและจัดทำแผนการจัดการเรียนรู้แบบผสมผสาน (Blended Learning) ผู้วิจัยได้ดำเนินการตามแผนที่วางไว้จึงทำให้บทเรียนคอมพิวเตอร์ออกมามีคุณภาพ

นอกจากนี้ผู้วิจัยยังได้พัฒนาบทเรียนคอมพิวเตอร์ผ่านเว็บอย่างเป็นระบบตามทฤษฎีของ ADDIE Model 5 ขั้นตอน ส่งผลให้บทเรียนมีประสิทธิภาพเหมาะสมสามารถนำไปใช้ในการเรียนการสอนได้ รวมทั้งยังตอบสนองต่อปัญหาของการวิจัยที่ขาดสื่อการเรียนการสอนให้ผู้เรียนได้ทบทวนเพิ่มเติมนอกเหนือจากเวลาเรียนได้ตามความต้องการ อีกทั้งการจัดกิจกรรมการเรียนรู้ด้วยเทคนิคโครงงานเป็นฐานนั้น ส่งเสริมให้ผู้เรียนได้เรียนรู้วิธีการเรียนรู้ และสามารถรู้จักการแก้ปัญหา สามารถค้นคว้าหาความรู้ได้ด้วยตนเอง เกิดทักษะในรายวิชาที่ศึกษาและสิ่งที่ตนสนใจ ดังจะเห็นได้จากผลการศึกษาความพึงพอใจของผู้เรียนที่มีต่อบทเรียนคอมพิวเตอร์ผ่านเว็บตามแผนการสอนฐานสมรรถนะในรายวิชาการศึกษาวงจรและซ่อมบำรุงไมโครคอมพิวเตอร์ โดยใช้โครงงานเป็นฐานร่วมกับการจัดการเรียนรู้แบบผสมผสานที่พัฒนาขึ้น พบว่าผู้เรียนมีความมีความพึงพอใจโดยรวมอยู่ในระดับมากที่สุด

6. ข้อเสนอแนะ

1. จากผลการวิจัยพบว่าในการเรียนโดยใช้บทเรียนคอมพิวเตอร์แบบผสมผสานผ่านเว็บ ก่อนเรียนผู้สอนควรมีการแนะนำให้ผู้เรียนรู้ถึงวิธีการใช้บทเรียนเพราะผู้เรียนบางคนอาจไม่คุ้นเคยเรียนบทเรียนคอมพิวเตอร์ผ่านเว็บ ผู้สอนควรบอกวัตถุประสงค์ของการเรียนรู้ให้ผู้เรียนทราบ

2. จากการวิจัยพบว่าควรมีการใช้กิจกรรมในรูปแบบของการแข่งขัน เพื่อเป็นการกระตุ้นให้ผู้เรียนเกิดความสนใจ อยากที่จะเรียนรู้และค้นหาด้วยตนเอง

3. จากผลการวิจัยพบว่า ผู้เรียนมีผลสัมฤทธิ์ทางการเรียนสูงขึ้นหลังจากเรียนด้วยบทเรียนคอมพิวเตอร์ผ่านเว็บและมีความคิดเห็นที่ดีต่อการเรียน จึงควรส่งเสริมให้สร้างบทเรียนคอมพิวเตอร์ผ่านเว็บในรายวิชาอื่น ๆ เพื่อที่นักเรียนสามารถที่จะทำไปศึกษาด้วยตนเองตามความสนใจ

เอกสารอ้างอิง

[1] กระทรวงศึกษาธิการ, *พระราชบัญญัติการศึกษาแห่งชาติ. พ.ศ. 2542 และที่แก้ไขเพิ่มเติม (ฉบับที่ 2) พ.ศ. 2545.* กรุงเทพฯ : โรงพิมพ์คุรุสภาลาดพร้าว, 2542.

[2] คณะกรรมการการศึกษาแห่งชาติ, *การปฏิรูปการเรียนรู้ ผู้เรียนสำคัญที่สุด.* กรุงเทพมหานคร : สำนักงานคณะกรรมการการศึกษาแห่งชาติ, 2543.

[3] ถนอมพร เลาหจรัสแสง, *การสอนบนเว็บ (Web-Based Instruction) นวัตกรรมเพื่อคุณภาพการเรียนการสอน.* กรุงเทพฯ : วารสารศึกษาศาสตร์สาร, 2544.

[4] กิดานันท์ มลิทอง, *เทคโนโลยีการศึกษาและนวัตกรรม.* พิมพ์ครั้งที่ 2. กรุงเทพมหานคร: อรุณการพิมพ์, 2543.

[5] ศูนย์ประกันคุณภาพการศึกษา, *การเรียนรู้โดยใช้โครงงานเป็นฐาน.* กรุงเทพฯ : ศูนย์ผลิตตำราเรียน มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2558.

[6] ธำรง บัวศรี, *ทฤษฎีหลักสูตร การออกแบบและพัฒนา.* พิมพ์ครั้งที่ 4 กรุงเทพฯ: โรงพิมพ์คุรุสภาลาดพร้าว, 2542.

[7] มนต์ชัย เทียนทอง, *Blended Learning : การเรียนรู้แบบผสมผสานในยุค ICT(ตอนที่1).* วารสารวิชาการครุศาสตร์อุตสาหกรรม, 2549.

การประเมินระดับความยากง่ายของข้อความด้วยวิธีการสกัดคำยืมจากข้อความภาษาไทย
Text Readability Assessment using Loanword Extraction from Thai Texts

บุรฮาน หวังเหล็ม (Burhan Wanglem)[1] และณัฐพงศ์ ทองเทพ (Nattapong Tongtep)[2]
สาขาวิชาเทคโนโลยีสารสนเทศและการสื่อสาร คณะเทคโนโลยีและสิ่งแวดล้อม มหาวิทยาลัยสงขลานครินทร์ วิทยาเขตภูเก็ต
[1]burhan.w@phuket.psu.ac.th, [2]nattapong.t@phuket.psu.ac.th

บทคัดย่อ

โดยปกติระยะเวลาในการเลือกเอกสารให้เหมาะสมสำหรับผู้อ่านแต่ละคนขึ้นกับจำนวนของเอกสารที่กำลังเลือก หากเอกสารที่กำลังเลือกถูกจัดกลุ่มและเรียงตามระดับความยากง่าย ส่งผลให้ผู้อ่านสามารถเลือกเอกสารที่เหมาะสมกับตนเองได้รวดเร็วยิ่งขึ้น ดังนั้น งานวิจัยนี้ได้นำเสนอเทคนิคสำหรับการจัดกลุ่มเอกสารตามระดับความยากง่ายด้วยวิธีการสกัดคำยืมในเอกสารข้อความภาษาไทย ประกอบด้วยสามเทคนิค คือ 1) การจัดกลุ่มเอกสารที่คำนวณจากสัดส่วนของจำนวนเอกสารทั้งหมด (DoC-A) 2) การจัดกลุ่มเอกสารด้วยอันตรภาคชั้นที่คำนวณจากค่าที่มากที่สุดและน้อยที่สุดที่เกิดขึ้นจริง (DoC-B) 3) การจัดกลุ่มเอกสารด้วยอันตรภาคชั้นที่คำนวณจากค่าที่มากที่สุด (DoC-C) ผลลัพธ์การจัดกลุ่มเอกสารด้วยเทคนิค DoC-A ที่ได้นำเสนอมีความใกล้เคียงกับผลลัพธ์การจัดกลุ่มของผู้อ่านถึงร้อยละ 75

คำสำคัญ: การประเมินระดับความยากง่ายของข้อความ การจัดกลุ่มข้อความ คำยืมในภาษาไทย

Abstract

Normally, duration of selecting an appropriate document for each reader depends on a number of documents. If these documents are grouped and ranked by readability assessment then the reader can quickly find them. In this paper, we propose three document clustering techniques for Thai texts readability by detecting Thai loanword as follows: 1) Document clustering calculated from the proportion of total number of documents (DoC-A) 2) Document clustering using class interval calculated from the actual maximum and minimum values (DoC-B) 3) Document clustering using class interval calculated from the actual maximum value (DoC-C). The results show that our proposed text clustering technique DoC-A is 75% close to the text clustering result from readers.

Keyword: Text readability assessment, Text clustering, Thai loanword.

1. บทนำ

การเลือกเอกสารของผู้อ่านให้เหมาะสมกับตนเอง โดยปกติแล้วผู้อ่านจะต้องอ่านเอกสารแต่ละฉบับเพื่อประเมินความยากง่าย และเพื่อให้ผู้อ่านตัดสินใจเลือกเอกสารให้เหมาะสมกับตนเอง จากกรณีข้างต้นหากมีจำนวนเอกสารเพิ่มมากขึ้น ส่งผลให้ผู้อ่านใช้ระยะเวลาเพิ่มมากขึ้นในการเลือกเอกสาร หากเอกสารเหล่านั้นถูกประเมินระดับความยากง่าย ส่งผลให้ผู้อ่านสามารถเลือกอ่านเอกสารที่เหมาะสมกับตนเองได้รวดเร็วยิ่งขึ้น จากประเด็นดังกล่าวเป็นสิ่งที่ผู้วิจัยหาเทคนิคหรือวิธีการประเมินระดับความยากง่ายของข้อความเพื่อช่วยผู้อ่านเลือกเอกสารให้เหมาะสมกับตนเองและใช้เวลาลดน้อยลงในการเลือกเอกสาร

สำหรับวิธีการประเมินระดับความยากง่ายของข้อความที่ใช้ในปัจจุบันนั้นมีหลากหลาย เช่น Flesh [1] และ Flesh-Kincaid Score [2] ซึ่งใช้วิธีการพิจารณาความซับซ้อนของไวยากรณ์ เช่น การคำนวณจำนวนคำและพยางค์ในประโยค อย่างไรก็ตาม วิธีการดังกล่าวเหมาะสำหรับใช้ในภาษาอังกฤษ แต่มีข้อจำกัดในการประยุกต์ใช้กับภาษาอื่น ๆ เช่น ภาษาไทย เป็นต้น

สำหรับงานวิจัยที่เกี่ยวกับการประเมินความยากง่ายของข้อความในภาษาไทยอยู่ในช่วงเริ่มต้น เนื่องจากภาษาไทย คือ ภาษาระบบการเขียนแบบอักษรพยางค์ (Syllabic Alphabets)

[3] เป็นระบบภาษาที่ยากต่อกระบวนการประมวลผลข้อความ เช่น การหาจุดจบของคำ หรือประโยค เป็นต้น

ในภาษาไทยมีการยืมคำจากภาษาอื่นมาใช้ เช่น บาลี สันสกฤต อังกฤษ เป็นต้น ส่วนหนึ่งจากคำยืมถูกนำมาใช้เฉพาะทาง เช่น คำศัพท์ทางศาสนา คำราชาศัพท์ ชื่อเฉพาะ เป็นต้น [4] ด้วยคุณลักษณะการนำคำยืมมาใช้ สามารถใช้เป็นคุณสมบัติในการประเมินความยากง่ายของข้อความในเอกสารภาษาไทย

เพื่อค้นหาเทคนิคที่มีความใกล้เคียงในการประเมินระดับความยากง่ายสำหรับผู้อ่านมากที่สุด ดังนั้น งานวิจัยนี้ได้นำเสนอเทคนิคการจัดกลุ่มเอกสารสำหรับการประเมินระดับความยากง่ายสำหรับเอกสารข้อความดิจิทัลของภาษาไทย โดยใช้คำยืมในภาษาไทยเป็นพื้นฐานในการประเมินระดับความยากง่ายในเอกสาร ประกอบด้วยสามเทคนิคที่มีความแตกต่างกันในการจัดกลุ่มเอกสาร

2. งานวิจัยที่เกี่ยวข้อง

2.1 การประเมินความยากง่ายของข้อความ (Text readability assessment)

การประเมินระดับความยากง่ายของข้อความในเอกสารเป็นการจัดเรียงตามระดับความยากง่ายของข้อความในเอกสาร เพราะระดับความเข้าใจของผู้อ่านจะขึ้นกับคุณลักษณะทางภาษาของข้อความในเอกสาร ดังนั้น งานวิจัยส่วนใหญ่ที่ผ่านมาจะมุ่งเน้นการประเมินระดับความยากง่ายของข้อความในเอกสาร [5] ตัวชี้วัดความยากง่ายแบบดั้งเดิมที่ใช้กันอย่างแพร่หลาย คือ Flesh-Kincaid score [2] โดยเป็นสมการสำเร็จรูปในการคำนวณ ดังนี้

$$RGFK = 0.93\left(\frac{total\ words}{total\ sentences}\right) + 11.8\left(\frac{total\ syllables}{total\ words}\right) - 15.59 \qquad (1)$$

โดยที่ total words คือ จำนวนคำทั้งหมด

total sentences คือ จำนวนประโยคทั้งหมด

total syllables คือ จำนวนพยางค์ทั้งหมด

ซึ่งสมการแบบดั้งเดิมสามารถนำมาปรับใช้สำหรับภาษากลุ่มระบบการเขียนแบบ Syllabic Alphabets ได้ แต่มีประสิทธิภาพน้อย [6] เนื่องจากสมการถูกออกแบบให้เหมาะสมสำหรับภาษาของทวีปยุโรป เช่น ภาษาอังกฤษ

สำหรับภาษาไทยมีงานวิจัยที่เกี่ยวกับเรื่องนี้เช่นกัน เช่น งานวิจัยของ Daowadung and Chen [7] ได้ประเมินระดับความยากง่ายเพื่อจำแนกเอกสารสำหรับนักเรียนระดับชั้นประถมศึกษาของประเทศไทย โดยคุณสมบัติสำหรับการประเมินระดับความยากง่าย คือ การกำหนดเงื่อนไขของคำ (Term) และใช้ Support Vector Machines (SVM) ในการจำแนกเอกสาร ซึ่งผลการทดลองได้ค่า F-measure ถึงร้อยละ 83 และได้ต่อยอดงานวิจัยข้างต้น [8] โดยสร้างชุดคุณสมบัติสำหรับประเมินระดับความยากง่ายที่แตกต่างกัน ประกอบด้วยสามชุดคุณสมบัติ ดังนี้ 1) Shallow (SL) 2) Language model (LM) และ3) Term frequency (TF) และใช้ SVM สำหรับจำแนกเอกสาร ซึ่งผลการทดลองชุดของ TF มีประสิทธิภาพมากกว่าชุดคุณสมบัติอื่น ๆ แต่ผลของค่าความถูกต้อง (Accuracy) น้อยกว่างานวิจัยก่อนหน้า เนื่องจากมีการกำหนดชุดสำหรับจำแนกแบบMulti-classและต่อมาเป็นงานวิจัยของ Tongtep, et al. [6] ได้นำเสนอเทคนิคการประเมินระดับความยากง่ายเพื่อจำแนกเอกสารภาษาไทยสำหรับนักเรียนระดับชั้นมัธยมตอนต้น ตอนปลาย และนักศึกษาระดับปริญญาตรี จำนวนเจ็ดเทคนิค โดยเทคนิค Character Per Word (CPW) Document Ranking ให้ค่าความแม่นยำสูงถึงร้อยละ 46.53

2.2 คำยืมในภาษาไทย (Thai loanword)

คำยืม คือ การนำคำจากภาษาอื่นมาปรับใช้ ส่งผลให้ภาษาเจริญขึ้นและจำนวนคำเพิ่มขึ้น ทุกภาษาย่อมมีคำยืมปะปน [9] เช่น ภาษาอังกฤษมีการยืมคำจากหลากหลายภาษา เช่น ฝรั่งเศส ละติน กรีก เป็นต้น และภาษาไทยมีคำยืมเช่นกัน เนื่องจากสมัยก่อนประเทศไทยมีความสัมพันธภาพกับต่างชาติมากมาย เช่น กัมพูชา จีน ลาว อังกฤษ ฝรั่งเศส เป็นต้น ทำให้ประเทศไทยได้รับภาษา หรือ คำต่าง ๆ ของชาตินั้นเข้ามาด้วย [10] เช่น กษัตริย์ ขัตติยะ เป็นต้น ลักษณะคำยืมของภาษาจะตรงกันข้ามกับคำไทยแท้ โดยคำไทยแท้มักมีลักษณะเป็นพยางค์เดียว มีความหมายสมบูรณ์ในตัวและมีตัวสะกดตรงตามมาตรา จากลักษณะของคำยืมในข้างต้น สามารถใช้ประเมินระดับความยากง่ายในเอกสารได้ ดังนั้น ผู้วิจัยได้กำหนดให้ระดับความยากง่ายของข้อความขึ้นกับจำนวนคำยืมที่ตรวจพบในเอกสาร คำยืมที่ผู้วิจัยสนใจ ได้แก่ ภาษาบาลี สันสกฤต และคำยืมจากภาษาอื่น ๆ ที่มีเครื่องหมายทัณฑฆาต เช่น คำยืมจากภาษาอังกฤษ เป็นต้น

3. วิธีการดำเนินการวิจัย

ในการดำเนินการวิจัยของผู้วิจัยได้แบ่งออกเป็นสามขั้นตอนหลัก คือ 1) ขั้นตอนการเตรียมข้อมูล 2) ขั้นตอนการจัดกลุ่มข้อมูล 3) ขั้นตอนการประเมินผลการจัดกลุ่มข้อมูล ดังที่ได้แสดงในภาพที่ 1 ซึ่งจะอธิบายในแต่ละขั้นตอนในย่อหน้าถัดไป

3.1 ขั้นตอนการเตรียมข้อมูล

เป็นขั้นตอนที่มีกระบวนการนำข้อมูลเข้ามาเตรียมความพร้อมก่อนเข้าสู่ขั้นตอนการจัดกลุ่ม โดยกระบวนการเริ่มแรกคือ นำเข้าเอกสารจำนวนหัวเรื่องละ 10 ฉบับที่ได้จากการค้นในโปรแกรมค้นหาของ Google ดังที่ได้แสดงในตารางที่ 1 ซึ่งผ่านการระบุคำยืม (คำยืม ในงานวิจัยนี้ประกอบด้วย บาลี สันสกฤต และตัวการันต์) มาตรวจหาตำแหน่งของคำยืม โดย 1 ตำแหน่งที่ตรวจจับได้มีค่าเท่ากับ 1 คะแนน ทำให้ได้คะแนนของคำยืมในแต่ละเอกสาร (การให้คะแนนคำยืม) กระบวนการต่อมาเป็นการหาคะแนนที่ซ้ำกัน หรือเท่ากันของแต่ละเอกสาร เพื่อต้องการที่จะจัดกลุ่มเอกสารที่มีคะแนนเท่ากันให้อยู่ในกลุ่มเดียวกัน ผลลัพธ์ที่ได้ คือ เอกสารที่มีคะแนนเท่ากันจะอยู่ในกลุ่มเดียวกัน (การจัดกลุ่มเอกสารที่มีคะแนนเท่ากัน)

3.2 ขั้นตอนการจัดกลุ่มข้อมูล

เป็นขั้นตอนที่มีกระบวนการจัดกลุ่มให้กับเอกสารโดยเริ่มจากการนำเอกสารที่ผ่านขั้นตอนการเตรียมข้อมูลเข้าสู่กระบวนการจัดกลุ่มเอกสารในแต่ละเทคนิค ซึ่งเทคนิคการจัดกลุ่มเอกสารที่ผู้วิจัยได้พัฒนาและได้นำเสนอในงานวิจัยนี้ประกอบด้วยสามเทคนิค ดังนี้

ตารางที่ 1: ชุดเอกสารที่ใช้สำหรับทดสอบ

หัวเรื่อง	ตัวอย่างคำค้น	จำนวนเอกสาร	จำนวนอักขระ	จำนวนคำยืม
Data mining	การทำเหมืองข้อมูล, data mining คือ	11	27,852	59
iOS operation system	ios คือ, ระบบปฏิบัติการ ios	19	10,336	33

3.2.1 Document Clustering A (DoC-A)

เป็นวิธีการหาเศษส่วนเพื่อกำหนดจำนวนสมาชิก (เอกสาร) ในแต่ละกลุ่มเพื่อให้ได้จำนวนเอกสารในแต่ละกลุ่มใกล้เคียงกันมากที่สุด ดังที่ได้แสดงในภาพที่ 2 โดยมีขั้นตอนและเงื่อนไขดังนี้

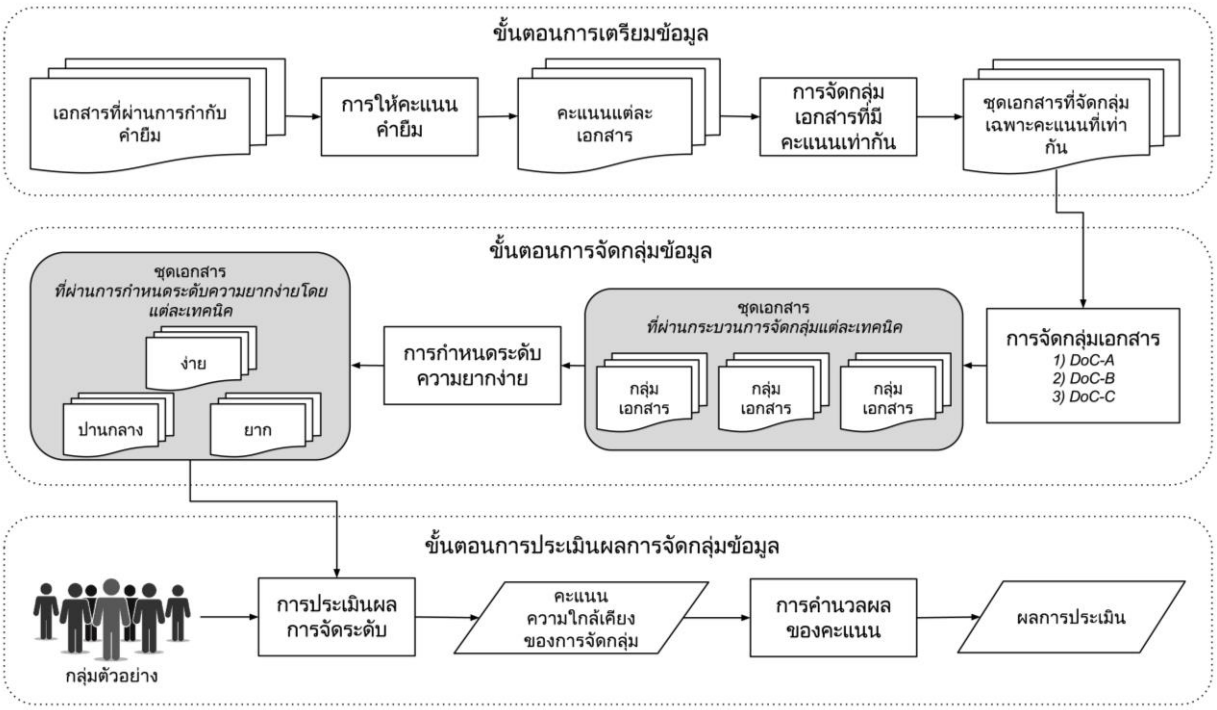
3.2.1.1 จัดเรียงเอกสารตามคะแนนจากน้อยไปมาก

3.2.1.2 คำนวณจำนวนสมาชิกของแต่ละกลุ่มโดยใช้สมการ (2)

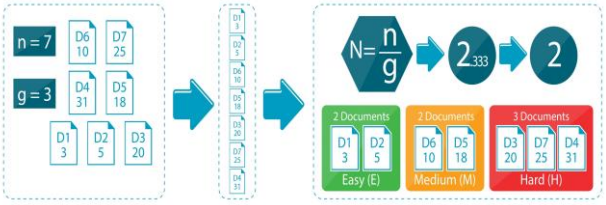
$$N = \frac{n}{g} \quad (2)$$

โดยที่ N คือ จำนวนสมาชิกที่จะต้องกำหนดให้ในแต่ละกลุ่ม

n คือ จำนวนของเอกสารทั้งหมด



ภาพที่ 1: การดำเนินการวิจัย



g คือ จำนวนกลุ่มที่ต้องการแบ่งเอกสาร และในงานวิจัยนี้กำหนดให้ g เท่ากับ 3 กลุ่ม คือ ระดับง่าย (Easy) ระดับปานกลาง (Medium) และระดับยาก (Hard)

3.2.1.3 นำค่า N ที่ได้จากสมการ (2) มาเป็นเกณฑ์

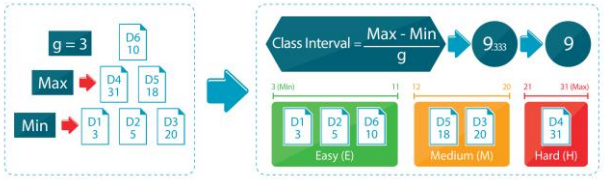
ภาพที่ 3: เทคนิคการจัดกลุ่มเอกสาร DoC-A

ในการกำหนดจำนวนสมาชิกของแต่ละกลุ่ม ซึ่งมีเกณฑ์ดังนี้

(ก) หากค่า N เป็นจำนวนจริงที่น้อยกว่า 0.5 ให้ลดค่า N เป็นจำนวนเต็ม และหากมากกว่าหรือเท่ากับ 0.5 ให้เพิ่มค่า N เป็นจำนวนเต็ม ดังที่ได้แสดงในภาพที่ 2

(ข) ทำการแบ่งจำนวนสมาชิกให้กับกลุ่ม Easy และ Medium ก่อนตามลำดับ โดยแบ่งตามค่า N และสุดท้ายแบ่งให้กับ Hard ตามจำนวนสมาชิกที่คงเหลือ ดังที่ได้แสดงในภาพที่ 2

3.2.2 Document Clustering B (DoC-B) เป็นวิธีการจัดกลุ่มโดยการสร้างอันตรภาคชั้นสำหรับแต่ละกลุ่มจากค่าคะแนนที่มากและน้อยที่สุดของเอกสารที่เกิดขึ้นจริง ดังที่ได้แสดงในภาพที่ 2 โดยมีขั้นตอนและเงื่อนไขดังนี้



ภาพที่ 2: เทคนิคการจัดกลุ่มเอกสาร DoC-B

3.2.2.1 คำนวณจำนวนอันตรภาคชั้นของแต่ละกลุ่มโดยใช้สมการ (3)

$$Class\ Interval = \frac{Max - Min}{g} \qquad (3)$$

โดยที่ Class Interval คือ จำนวนอันตรภาคชั้น

Max คือ คะแนนของเอกสารที่มีค่ามากที่สุด

Min คือ คะแนนของเอกสารที่มีค่าน้อยที่สุด

g คือ จำนวนกลุ่มที่ต้องการแบ่งเอกสาร และในงานวิจัยนี้กำหนดให้ g เท่ากับ 3 กลุ่ม คือ Easy Medium และ Hard

3.2.2.2 หาก Class Interval เป็นจำนวนจริงที่น้อยกว่า 0.5 ให้ลด Class Interval เป็นจำนวนเต็ม และหากมากกว่าหรือเท่ากับ 0.5 ให้เพิ่ม Class Interval เป็นจำนวนเต็ม

3.2.2.3 นำค่า Class Interval สร้างอันตรภาคชั้นสำหรับแต่ละกลุ่ม ซึ่งเขตข้อมูลจะเริ่มต้นที่ Min ถึง Max โดยเริ่มแบ่งช่วงให้กับ Easy และ Medium ให้เท่ากันตามลำดับ และสุดท้ายแบ่งช่วงให้กับ Hard ตามจำนวนช่วงที่เหลือ

3.2.2.4 แบ่งเอกสารที่มีคะแนนของคำยืมตามช่วงของอันตรภาคชั้นที่สร้างขึ้น

3.2.3 Document Clustering C (DoC-C) เป็นวิธีการจัดกลุ่มโดยการสร้างอันตรภาคชั้นสำหรับแต่ละกลุ่มจากค่าคะแนนที่มากที่สุดของเอกสาร ดังที่ได้แสดงในภาพที่ 4 โดยมีขั้นตอนและเงื่อนไขดังนี้



ภาพที่ 4: เทคนิคการจัดกลุ่มเอกสาร DoC-C

3.2.3.1 คำนวณจำนวนอันตรภาคชั้นของแต่ละกลุ่มโดยใช้สมการ (4)

$$Class\ Interval = \frac{Max}{g} \qquad (4)$$

โดยที่ Class Interval คือ จำนวนอันตรภาคชั้น

Max คือ คะแนนของเอกสารที่มีค่ามากที่สุด

g คือ จำนวนกลุ่มที่ต้องการแบ่งเอกสาร และในงานวิจัยนี้กำหนดให้ g เท่ากับ 3 กลุ่ม คือ Easy Medium และ Hard

3.2.3.2 ใช้เกณฑ์เดียวกับข้อ 3.2.2.2

3.2.2.3 ใช้เกณฑ์เดียวกับข้อ 3.2.2.3 แต่เขตข้อมูลเริ่มต้นที่ 0 ถึง Max

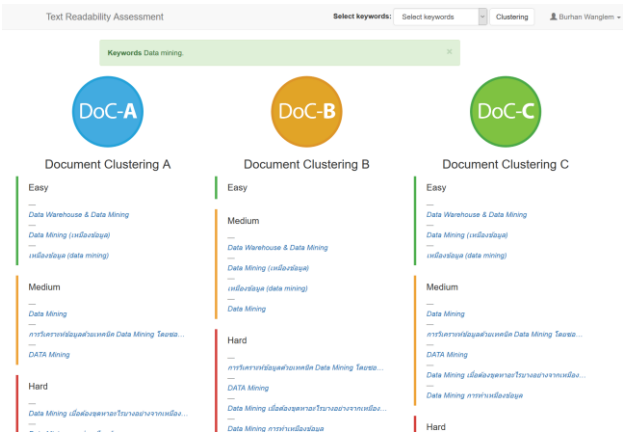
3.2.3.4 ใช้เกณฑ์เดียวกับข้อ 3.2.2.4

โดยผลลัพธ์ที่ได้จากแต่ละเทคนิค คือ เอกสารที่ถูกแบ่งกลุ่มตามจำนวนกลุ่มที่ได้กำหนดเอาไว้ ซึ่งจำนวนกลุ่มที่ได้กำหนดในการวิจัยในครั้งนี้คือ 3 กลุ่ม (การจัดกลุ่มเอกสาร) กระบวนการต่อมา คือ การกำหนดระดับความยากง่ายให้กับ

เอกสารที่ได้ถูกแบ่งกลุ่มในกระบวนการก่อนหน้า ซึ่งในการวิจัยครั้งนี้ได้กำหนดระดับความยากง่ายออกเป็นสามระดับ คือ Easy Medium และ Hard และนิยามของความยากง่าย คือ ระดับความยากง่ายของเอกสารขึ้นกับจำนวนของคำยืมในเอกสาร ผลลัพธ์ที่ได้จากกระบวนการนี้ คือ ชุดเอกสารแต่ละชุดที่ถูกระบุระดับความยากง่าย โดยเอกสาร 1 กลุ่มจะถูกกำหนดระดับความยากง่ายเพียง 1 ระดับเท่านั้น (การกำหนดระดับความยากง่าย)

3.3 ขั้นตอนการประเมินผลการจัดกลุ่มข้อมูล

เป็นขั้นตอนในการประเมินเทคนิคการจัดระดับความยากง่ายของเอกสารแต่ละเทคนิค เนื่องจากผู้วิจัยต้องการหาเทคนิคในการจัดระดับความยากง่ายของเอกสารที่เหมาะสม และตรงต่อระดับความยากง่ายในการอ่านเอกสารของมนุษย์มากที่สุด ด้วยเหตุนี้ผู้วิจัยจึงนำผลลัพธ์การประเมินระดับความยากง่ายของเอกสารที่ได้จากเทคนิคแต่ละเทคนิคมาประเมิน ผู้วิจัยได้สร้างระบบสำหรับประเมินผลของแต่ละเทคนิคดังที่ได้แสดงในภาพที่ 5 คือ ตัวอย่างส่วนต่อประสานกับผู้ใช้ (User interface) โดยเริ่มจากให้กลุ่มตัวอย่างจำนวน 20 คน ประเมินผลของการจัดระดับความยากง่ายของแต่ละเทคนิค โดยให้ระดับคะแนนความใกล้เคียงในการจัดกลุ่ม 1 ถึง 3 โดยกำหนดให้ผู้ประเมินให้ระดับคะแนนแต่ละเทคนิคไม่เท่ากัน ผลลัพธ์ที่ได้จากกระบวนการนี้ คือ ระดับคะแนนความใกล้เคียงในการจัดกลุ่มระหว่างเทคนิคการจัดกลุ่มและกลุ่มตัวอย่าง (การประเมินผลการจัดระดับ) กระบวนการต่อมา คือ การนำผลการประเมินในข้างต้นมาคำนวณหาผลรวมแต่ละเทคนิค และค่าเฉลี่ยทั้งหมด เพื่อนำผลมาสรุป และดูแนวโน้มของผลลัพธ์



ภาพที่ 5: ตัวอย่างส่วนต่อประสานกับผู้ใช้ของระบบประเมินผล

4. ผลการดำเนินงานวิจัย

ผลลัพธ์ของการประเมินความใกล้เคียงกันของผลการจัดกลุ่มเอกสารระหว่างเทคนิคการจัดกลุ่มและกลุ่มตัวอย่างได้แสดงในตารางที่ 2 และ 3 เป็นตารางการสรุปคะแนนความใกล้เคียงในการจัดกลุ่มของกลุ่มตัวอย่างแต่ละบุคคลในแต่ละเทคนิค และแต่ละหัวเรื่อง โดยงานวิจัยนี้ ได้กำหนดหัวเรื่องจำนวนสองหัวเรื่อง คือ Data mining และ iOS operating system ตามลำดับ หัวเรื่องดังกล่าวมีเนื้อหาให้ความรู้แก่ผู้อ่าน ซึ่งเหมาะสำหรับการจัดระดับความยากง่ายของข้อความ นอกจากนี้ ตารางที่ และ 3 ได้แสดงถึงผลรวมของคะแนนความใกล้เคียงในแต่ละเทคนิค โดยเทคนิค DoC-A ได้ให้ผลรวมสูงที่สุดถึง 42 และ 45 ตามลำดับ ส่วนเทคนิค DoC-C ได้ให้ผลรวมรองลงมา คือ 41 และ 44 ตามลำดับ และ เทคนิค DoC-B ได้ให้ผลรวมน้อยที่สุด คือ 36 และ 31 ตามลำดับ

จากภาพที่ 6 ได้แสดงถึงค่าเฉลี่ยของผลความใกล้เคียงของการจัดกลุ่มเอกสารแต่ละหัวเรื่องเป็นร้อยละ ซึ่งเทคนิค DoC-A มีค่าเฉลี่ยสูงที่สุดถึง ร้อยละ 70 และ ร้อยละ 75 จากทั้งสองหัวเรื่อง ส่วนเทคนิค DoC-C และ DoC-B มีค่าเฉลี่ยรองลงมาตามลำดับ

ตารางที่ 2: ผลการประเมินความใกล้เคียงในการจัดกลุ่มในหัวเรื่อง Data mining

เทคนิค	ความถี่ของกลุ่มตัวอย่างจำแนกตามระดับความใกล้เคียงของการจัดกลุ่ม			การคำนวณผล		
	3	2	1	ผลรวม	ฐานนิยม	ค่าเฉลี่ย
DoC-A	6	10	4	42	2	2.10
DoC-B	7	2	11	36	1	1.80
DoC-C	6	9	5	41	2	2.05

ตารางที่ 3: ผลการประเมินความใกล้เคียงในการจัดกลุ่มในหัวเรื่อง iOS operating system

เทคนิค	ความถี่ของกลุ่มตัวอย่างจำแนกตามระดับความใกล้เคียงของการจัดกลุ่ม			การคำนวณผล		
	3	2	1	ผลรวม	ฐานนิยม	ค่าเฉลี่ย
DoC-A	6	13	1	45	2	2.25
DoC-B	4	3	13	31	1	1.55
DoC-C	10	4	6	44	3	2.20

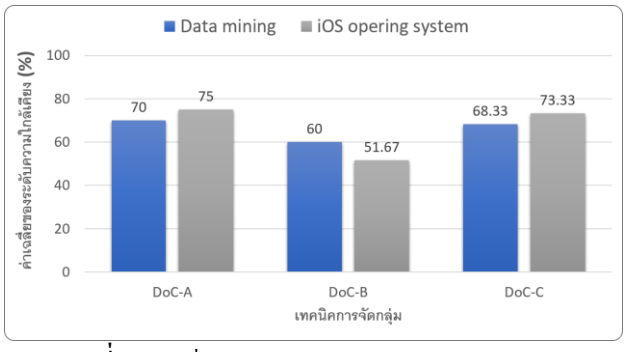
5. อภิปรายผล

ผู้วิจัยได้สรุปประเด็นสำคัญที่เกี่ยวข้องกับงานวิจัยนี้ สามารถแบ่งได้ 3 ประเด็น ดังนี้

1) การใช้คำยืมเป็นชุดคุณสมบัติในการประเมินระดับความยากง่าย ซึ่งแตกต่างกับงานวิจัยก่อนหน้าในหัวข้อ 2.1 ส่งผลให้ผลความใกล้เคียงของงานวิจัยนี้สูงกว่าเมื่อเปรียบเทียบกับผลความถูกต้องของงานวิจัยก่อนหน้า

2) การใช้เทคนิคการจัดกลุ่มเพื่อจัดกลุ่มเอกสารตามระดับความยากง่าย ซึ่งแตกต่างกับงานวิจัยก่อนหน้าในหัวข้อ 2.1 ส่งผลให้ระดับความยากง่ายของงานวิจัยนี้สามารถปรับใช้กับบุคคลทั่วไปได้ อย่างไรก็ตาม เทคนิคในการจัดกลุ่มที่ได้นำเสนอต้องอาศัยจำนวนเอกสารที่มากกว่าหนึ่งฉบับขึ้นไปในการจัดจัดกลุ่มเอกสาร หากภายในชุดเอกสารไม่มีจำนวนคำยืมผลการจัดกลุ่มของเอกสารทั้งหมด คือ Easy

3) สำหรับงานวิจัยนี้เป็นการทดลองตั้งต้น โดยมีหัวเรื่องเอกสารสำหรับการทดลองจำนวน 2 หัวเรื่อง เพื่อให้ครอบคลุมสำหรับหัวเรื่องที่หลากหลาย จำเป็นต้องเพิ่มเอกสารสำหรับการทดลอง ซึ่งผู้วิจัยจะดำเนินการเพิ่มจำนวนหัวเรื่องมากขึ้นในอนาคต



ภาพที่ 6: ค่าเฉลี่ยระดับความใกล้เคียงของการจัดกลุ่มเอกสารในแต่ละหัวเรื่อง

6. สรุปและงานวิจัยในอนาคต

งานวิจัยนี้ได้นำเสนอเทคนิคการจัดกลุ่มเอกสารสามเทคนิค คือ 1) DoC-A 2) DoC-B 3) DoC-C สำหรับประเมินระดับความยากง่ายของข้อความ โดยการตรวจจับคำยืมในเอกสารภาษาไทย ซึ่งทั้งสามเทคนิคเป็นการจัดกลุ่มตามจำนวนกลุ่มที่กำหนด คือ Easy Medium และ Hard ผู้วิจัยได้สร้างระบบเพื่อประเมินผลลัพธ์ของเทคนิคการจัดกลุ่มเอกสารแต่ละเทคนิค และใช้กลุ่มตัวอย่างจำนวน 20 คน ในการประเมินผลความใกล้เคียงในการจัดกลุ่มสำหรับแต่ละเทคนิค ผลลัพธ์ที่ได้ คือ DoC-A สามารถจัดกลุ่มเอกสารที่ใกล้เคียงกับกลุ่มตัวอย่างที่ใช้ทดลองมากที่สุดโดยเฉลี่ยถึง ร้อยละ 70 ในหัวเรื่อง Data mining และ ร้อยละ 75 ในหัวเรื่อง iOS operating system

ผู้วิจัยวางแผนที่จะพัฒนาต่อยอดด้วยการพัฒนาระบบการประเมินระดับความยากง่ายของข้อความในภาษาไทยแบบอัตโนมัติ

เอกสารอ้างอิง

[1] R. Flesch, “A new readability yardstick,” *Journal of applied psychology*, Vol. 32, No. 3, pp. 221, 1948.

[2] J. P. Kincaid, et al. “Derivation of new readability formulas (automated readability index, fog count and flesch reading ease formula) for navy enlisted personnel,” *Naval Technical Training Command Millington TN Research Branch*, No. RBR-8-75, 1975.

[3] S. Ager, 1998. *The online encyclopedia of writing systems & languages.* Retrieved May, 2016, from Omniglot: http://www.omniglot.com/

[4] รำเนา พลอยขาว. 2549. *ภาษาไทย.* สืบค้นเมื่อ พฤษภาคม, 2559, จาก โรงเรียนสมุทรปราการ: http://www.prakan.ac.th/Link-Data/e-learning-49/e-learning/index.html

[5] K. Collins-Thompson, “Computational assessment of text readability: A survey of current and future research,” *ITL-International Journal of Applied Linguistics*, Vol. 165, No. 2, pp. 97-135, 2014.

[6] N. Tongtep, F. Coenen and T. Theeramunkong, “Content-Based Readability Assessment: A Study Using A Syllabic Alphabetic Language (Thai),” *Pacific Rim International Conference on Artificial Intelligence*, Springer International Publishing, pp. 863-870, 2014.

[7] P. Daowadung and Y. H. Chen, “Using word segmentation and SVM to assess readability of Thai text for primary school students,” *Computer Science and Software Engineering (JCSSE), 2011 Eighth International Joint Conference on*, IEEE, pp.170-174, 2011.

[8] Y. H. Chen and P. Daowadung, “Assessing readability of Thai text using support vector machines,” *Maejo International Journal of Science and Technology*, Vol. 9, No. 3, pp. 355, 2015.

[9] สุธิวงศ์ พงศ์ไพบูลย์, *หลักภาษาไทย*, ไทยวัฒนาพานิช: กรุงเทพฯ, 2534.

[10] กำชัย ทองหล่อ, *หลักภาษาไทย*, รวมสานส์: กรุงเทพฯ, 2540.

การจำแนกอารมณ์ของตัวการ์ตูนโดยใช้ค่าความถี่ของทิศทางตามค่าเกรเดียนท์การระบุ
Cartoonized Facial Expression Classification using Histogram of Oriented Gradient (HOG)

อรวรรณ เลียบศิริ (Orawan Liabsiri)[1] และเสาวลักษณ์ วรรธนาภา (Saowaluk Watanapa)[2]
สาขาวิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยธรรมศาสตร์
[1]orawan2529@gmail.com, [2]scw@cs.tu.ac.th

บทคัดย่อ

บทความนี้มีวัตถุประสงค์ เพื่อนำเสนอวิธีการจำแนกอารมณ์จากการแสดงออกทางใบหน้าของตัวการ์ตูน โดยใช้ค่าความถี่ของทิศทางตามค่าเกรเดียนท์ (Histogram of Oriented Gradient: HOG) ในการสกัดคุณลักษณะของภาพ จากนั้นนำไปประมวลผลการจำแนกอารมณ์ด้วยวิธีการต้นไม้การตัดสินใจ (Decision – Tree: DT) ตัวแบบเพื่อนบ้านที่ใกล้เคียงที่สุด k ตัว (k- Nearest Neighbors: k-NN) และซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine: SVM) โดยชุดข้อมูลที่ใช้ในการทดลองประกอบด้วย รูปภาพส่วนที่เป็นบริเวณใบหน้าตัวการ์ตูนจำนวน 60 ภาพ จากอารมณ์ทั้งหมด 3 อารมณ์ ซึ่งผลการทดลองแสดงให้เห็นว่าวิธีการจำแนกประเภทโดยใช้ต้นไม้การตัดสินใจ ให้ค่าความถูกต้องในการจำแนกร้อยละ 53.3 ตัวแบบเพื่อนบ้านที่ใกล้เคียงที่สุด k ตัว โดยกำหนดให้ k=3 ให้ค่าความถูกต้องในการจำแนกร้อยละ 76.7 และการจำแนกประเภทแบบซัพพอร์ตเวกเตอร์แมชชีนให้ค่าความถูกต้องในการจำแนกสูงสุดร้อยละ 90.0

คำสำคัญ: การจำแนกอารมณ์จากการแสดงออกทางใบหน้าขอตัวการ์ตูน ค่าความถี่ของทิศทางตามค่าเกรเดียนท์

Abstract

This paper presents techniques to classify facial expressions of cartoons. The histogram of oriented gradient (HOG) was used to extract features. After that, the decision tree (DT) classifier, the k- nearest neighbors (k-NN) classifier and the support vector machine (SVM) with linear kernel were applied to classify the emotions of cartoons. The experimental dataset consists of 60 images with 3 emotion – based classes. From the experimental results, the decision tree classifier yields 53.3%, the k - nearest neighbors classifier yields 76.7% and the support vector machine yields 90.0% accuracy.

Keyword: Cartoonized Facial Expression Classification, HOG.

1. บทนำ

การแยกแยะอารมณ์จัดเป็นงานจำแนกประเภทวัตถุ (Classification) ในระดับนามธรรม ซึ่งสามารถนำไปใช้ทำในส่วนของดัชนีภาพ (Index) ได้ด้วย การทำดัชนีของภาพนั้น มีอยู่ด้วยกัน 3 ระดับ (1) ระดับพื้นฐาน คือ ดัชนีเกี่ยวกับ สีพื้นผิว และรูปทรง (2) ระดับการแยะแยะความหมายเบื้องต้นของภาพ เช่น การทำดัชนีเกี่ยวกับ ภาพคน ภาพสัตว์ และภาพสิ่งของ และ (3) ระดับนามธรรม เช่น การทำดัชนีเกี่ยวกับภาพสื่อความหมายเกี่ยวกับศาสนาหรือวัฒนธรรมหรือดัชนีเกี่ยวกับอารมณ์ของภาพหรืออารมณ์ตัวละครในภาพ เป็นต้น

การผลิตภาพยนตร์มีความหลากหลาย ไม่ใช่แค่ภาพยนตร์ที่มนุษย์แสดงเท่านั้นแต่มีภาพยนตร์ที่สร้างขึ้นมาจากการ์ตูนแอนนิเมชั่นรวมเข้ามาด้วย จากการศึกษางานวิจัยที่เกี่ยวข้องกับการแยกแยะอารมณ์บนใบหน้าที่เห็นกันทั่วไปงานวิจัยส่วนใหญ่จะแยกแยะอารมณ์บนใบหน้าของมนุษย์เป็นหลัก การแยกแยะอารมณ์บนใบหน้าของตัวการ์ตูนที่พบเห็นก็จะอยู่ในส่วนของงานทางด้านจิตวิทยาเป็นหลัก

จากปัญหาดังกล่าวสามารถแก้ไขได้โดยการนำภาพการ์ตูนแอนนิเมชั่นมาทำการวิเคราะห์โดยใช้ความรู้ด้านการประมวลผลภาพดิจิตอล (Image Processing Digital) มาช่วยในการหาคุณสมบัติและจำแนกประเภทของภาพใบหน้าการ์ตูนว่ากำลังแสดงอารมณ์ประเภทใด เพื่อให้ระบบสามารถนำไป

ประยุกต์ใช้ในการทำดัชนีอารมณ์ของตัวการ์ตูนในภาพยนตร์แทนมนุษย์ได้

2. วรรณกรรมที่เกี่ยวข้อง

ในส่วนนี้จะนำเสนอเกี่ยวกับทฤษฎีและงานวิจัยที่เกี่ยวข้องเกี่ยวกับการจำแนกอารมณ์บนใบหน้า ซึ่งจะอธิบายแยกตามหัวข้อดังนี้

2.1 การแสดงอารมณ์บนใบหน้า (facial expressions)

การแสดงอารมณ์ออกทางสีหน้าของมนุษย์นั้นมีบทบาทสำคัญมากในความสัมพันธ์ระหว่างบุคคล เพราะวิธีการแสดงออกแบบนี้แสดงให้เห็นถึงการถ่ายทอดของข้อมูลจำนวนมาก จากการมองเห็นได้ด้วยสายตามากกว่าคำพูด โดยที่มนุษย์สามารถที่จะรู้จำการแสดงออกทางสีหน้าได้อย่างรวดเร็ว การแสดงออกทางสีหน้าประกอบไปด้วย องค์ประกอบหลัก 3 ประการ ได้แก่ ตา คิ้ว ปาก [1] นักจิตวิทยาชาวอเมริกันอธิบายเกี่ยวกับการแสดงออกทางสีหน้าของอารมณ์ทั้ง 10 ชนิด (1) อารมณ์สับสน (Confusion) จมูกและหน้าผากที่ขมวดอาจจะมีคิ้วยกสูงขึ้น (2) อารมณ์อับอาย (Shame) หลบสายตาและมีลักษณะเศร้าหรือกังวลร่วมด้วย (3) อารมณ์แปลกใจ (Surprise) ดวงตาโตกว่าปกติและปากที่เบิกกว้าง ซึ่งใกล้เคียงกับอารมณ์กลัว (4) อารมณ์มุ่งมั่น (Focus) สายตาที่แน่วแน่และมั่นคงมุ่งเน้นไปที่เป้าหมายที่อยู่ทางด้านหน้ากระพริบตาเล็กน้อย (5) อารมณ์อ่อนเพลีย (Exhaustion) พยายมยกคิ้วให้สูงขึ้น เบิกตาให้กว้างขึ้น (6) อารมณ์ยั่วยวน (Seduction) ดวงตาจับจ้องไปยังบุคคลตรงหน้า ริมฝีปากเม้ม ศีรษะเอียงเล็กน้อย กระพริบตาที่ถี่ขึ้นและยิ้มอ่อน (7) อารมณ์โกรธ (Anger) คิ้วขมวดเปลือกตาตึงและขวาง หัวคิ้วมักจะลดลงเล็กน้อย (8) อารมณ์กลัว (Fear) ตาที่เบิกกว้าง คิ้วที่เฉียง ปากมักจะถูกเปิดออกในระดับหนึ่งซึ่งใกล้เคียงกับอารมณ์ประหลาดใจ (9) อารมณ์เศร้า (Sadness) คิ้วเฉียงและขมวด ปากคว่ำลง สีหน้าไม่สดใส (10) อารมณ์มีความสุข (Happiness) เป็นอารมณ์ที่เป็นสากล รอยยิ้มและดวงตาเป็น รูปพระจันทร์ครึ่งเสี้ยว [2]

2.2 การจำแนกอารมณ์พื้นฐานของมนุษย์จากการแสดงบนใบหน้า (Facial Emotion Classification)

การจำแนกอารมณ์จากการแสดงสีหน้าโดยอัตโนมัติมีขั้นตอนหลักสามขั้นตอน ได้แก่ การตรวจจับบริเวณใบหน้าในภาพ การสกัดคุณลักษณะสำคัญ และการจำแนกด้วยตัวจำแนก (Classifier) แบบต่าง ๆ ในหัวข้อนี้จะกล่าวถึงงานวิจัยที่เกี่ยวข้องกับการจำแนกอารมณ์จากการแสดงสีหน้ามนุษย์

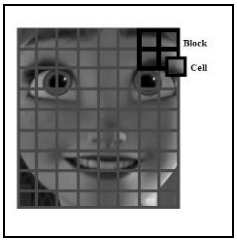
การตรวจจับบริเวณที่เป็นใบหน้าที่ดีต้องสามารถวิเคราะห์เพื่อหาคุณลักษณะเฉพาะบนใบหน้าและนำคุณลักษณะเด่นในสภาพแวดล้อมของบุคคลที่มีความแตกต่างกันออกไปทางสรีระและความยืดหยุ่นกับสภาพแวดล้อมที่แตกต่างกันได้ ในการรู้จำทางอารมณ์ Arumugam [3] ใช้หลักการในการคำนวณหา Statistical Method หลังจากรับภาพเข้ามา โดยใช้วิธีการคำนวณหาทั้งหมดสามวิธี (1) คำนวณหาเวกเตอร์ลักษณะเฉพาะ (Eigen vector) และค่าลักษณะเฉพาะ (Eigen Value) (2) การคำนวณหาแบบเชิงเส้นด้วยเทคนิคของ Fisher's Linear Discriminant และ (3) จำแนกหาคุณลักษณะด้วยลักษณะเฉพาะแบบการกระจายเมทริกซ์ด้วยวิธีการแยกค่าแบบเดี่ยว (Singular Value Decomposition: SVD) หลังจากนั้นนำข้อมูลที่ได้จากการคำนวณทั้งสามวิธีมาทำการจำแนกประเภทข้อมูลด้วยวิธีเครือข่ายใยประสาท (Neural Network) โดยเลือกวิธีการจำแนกแบบ RBF (Radial Basis Function Network) [4]

Chibelushi และ Bourel ทำการวิเคราะห์ในขั้นสูง ซึ่งเกี่ยวกับการทำดัชนีของภาพที่เป็นแบบนามธรรม มาใช้สำหรับการแยกแยะทางอารมณ์ของการแสดงออกทางสีหน้าของมนุษย์ด้วยระบบงานอัตโนมัติ ซึ่งมีเทคนิคในการจำแนกตามขั้นตอนดังนี้ ขั้นตอนแรกรับภาพสองมิติเข้ามาในระบบ จากนั้นนำมาแปลงเป็นภาพระดับเทาและนำภาพที่ได้มาสกัดหาคุณลักษณะที่ช่วยในการหาบริเวณที่เป็นใบหน้า จากคุณลักษณะด้านต่าง ๆ ได้แก่ รูปทรง การเคลื่อนไหว สี ลักษณะพื้นผิวและตำแหน่ง จากนั้นทำการแบ่งส่วนย่อย ได้แก่ คิ้ว ตา และปาก ในขั้นตอนสุดท้ายใช้ของเบย์ (Bayes Classifier) ในการจำแนก [5]

2.3 ค่าความถี่ของทิศทางตามค่าเกรเดียนท์ (Histograms of Oriented Gradients: HOG)

เป็นการดึงลักษณะเด่นของวัตถุออกจากภาพโดยใช้การกระจายตัวความเข้มของเกรเดียนท์หรือทิศทางของเส้นขอบ [6] หลักในการดึงคุณลักษณะของ HOG ทำได้โดยการแบ่งภาพออกเป็นบล็อก (Block) และเซลล์ (Cell) ขนาดเล็ก โดยในแต่ละเซลล์ประกอบด้วยทิศทางของค่าเกรเดียนท์ที่ถูกเก็บใน

รูปแบบของฮิสโตรแกรม ที่อธิบายคุณลักษณะของวัตถุที่อยู่ในเซลล์ [7] ดังแสดงตัวอย่างในภาพที่ 1



ภาพที่ 1: ตัวอย่างการแบ่งภาพเป็นเซลล์และบล็อค

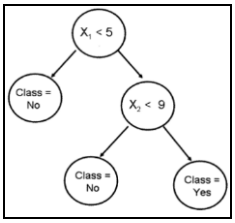
การรู้จำท่ามือภาษามือไทยใช้ HOG ร่วมกับ PCA (Principal Component Analysis) ผลที่การทดลองปรากฏว่าขั้นตอนวิธีการ HOG - PCA - SVM มีประสิทธิภาพการรู้จำสูงสุด ให้ความถูกต้องมากที่สุดถึง 96.00% และมีผลรวมระยะเวลาที่ใช้น้อยสุดคือ 59.84 วินาที [7] การตรวจจับมนุษย์จากการภาพวีดีโอโดยใช้คุณลักษณะ CT (Census Transform) ร่วมกับ HOG, PCA และ MELM-GRBF (A Modified Version of the Extreme Learning Machine for Generalized Radial Basis Function Neural Networks) ให้ความถูกต้องมากที่สุดถึง 97.89% [3] การตรวจจับมนุษย์จากภาพวีดีโอยังสามารถทำการตรวจจับได้จากการใช้ HOG ร่วมกับ IRC (Integral Region Covariance) และใช้ SVM ในการจำแนก ซึ่งให้ความถูกต้องมากที่สุดถึง 98.94% จากชุดข้อมูลที่ใช้ในการทดลอง [4]

2.4 ตัวจำแนก (Classifier)

ตัวจำแนกเป็นอัลกอริธึ่มที่ใช้ในการจำแนกข้อมูลประเภท (Class) ต่าง ๆ ตัวจำแนกมีหลายประเภท ที่นิยมใช้กัน ได้แก่ ต้นไม้ตัดสินใจ (Decision Tree: DT) ตัวจำแนกแบบเพื่อนบ้านที่ใกล้เคียงที่สุด k ตัว (k – Nearest Neighbors: k-NN) และซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine: SVM)

2.4.1 ตัวจำแนกแบบต้นไม้การตัดสินใจ

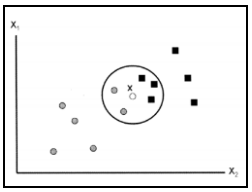
จะทำการจำแนกข้อมูลเป็นรูปแบบต้นไม้การตัดสินใจ แต่ละโหนด (Node) จะประกอบไปด้วยเงื่อนไขที่ใช้ตัวแปรใดตัวแปรหนึ่งของข้อมูลในการตัดสินใจเลือกโหมดลูกเพื่อใช้เป็นโหนดต่อ การตัดสินใจเริ่มจากโหนดราก (Root) ของต้นไม้และไล่ไปยังโหนดลูกจนถึงโหนดใบ เงื่อนไขของแต่ละโหนดสามารถบอกประเภทของข้อมูลนั้นได้ ตัวอย่างการสร้างข้อมูลต้นไม้การตัดสินใจ ดังภาพที่ 2 โดยมีตัวแปร คือ X_1 และ X_2 และข้อมูลอาจจะมีประเภทเป็น Yes หรือ No ก็ได้ [6]



ภาพที่ 2: การจำแนกประเภทของข้อมูลด้วยต้นไม้การตัดสินใจ

2.4.2 ตัวจำแนกแบบตัวแบบเพื่อนบ้านที่ใกล้เคียงที่สุด k ตัว

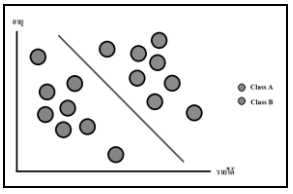
การจำแนกประเภทของข้อมูลนี้ ใช้การเปรียบเทียบข้อมูลกับข้อมูลตัวอย่างที่เก็บไว้ในฐานข้อมูล เมื่อต้องการจำแนกประเภทข้อมูลใหม่ จะค้นหาตัวอย่างของข้อมูลในฐานข้อมูลที่มีลักษณะใกล้เคียงข้อมูลใหม่มากที่สุดจำนวน k ตัว มาใช้ในการตัดสินใจว่าข้อมูลใหม่นี้ควรจะเป็นประเภท ดังภาพที่ 3 [8]



ภาพที่ 3: การจำแนกข้อมูล x โดยใช้วิธี k - NN

2.4.3 ตัวจำแนกข้อมูลแบบซัพพอร์ตเวกเตอร์แมชชีน

ถูกนำมาใช้ในการจำแนกค่าคุณลักษณะของชุดข้อมูลสองกลุ่มที่ออกจากกันโดยการสร้างเส้นแบ่งที่เป็นเส้นตรงเพื่อให้ทราบว่าเส้นตรงที่แบ่งข้อมูลออกจากกันเส้นตรงใดเป็นเส้นตรงที่ดีที่สุด ดังภาพที่ 4 [7],[9]



ภาพที่ 4: แสดงโมเดล SVM ที่แบ่งข้อมูลออกเป็น 2 กลุ่ม

โดยเลือกแบบจำลองเส้นตรง (Linear Model) ที่มีระยะห่างระหว่าง 2 กลุ่มที่ห่างกันมากที่สุด ซึ่งสามารถนำมาเขียนเป็นสมการในการแก้ปัญหา โดยนำข้อมูลมาวางไว้ในพื้นที่คุณลักษณะกลุ่มข้อมูลในรูปแบบเวกเตอร์จะได้ ดังสมการที่ (1)

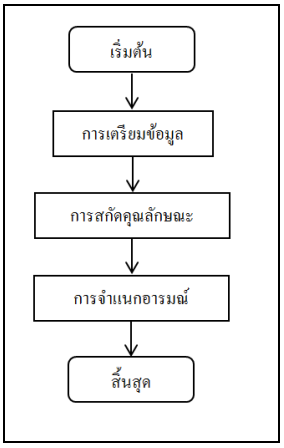
$$X = ((x_1 - y_1),...,(x_i - y_i)) \quad (1)$$

กำหนดสมการตัวบ่งชี้ข้อมูลเส้นแบ่ง (Hyperplane) ด้วยการใช้เส้นขอบ (Margin) ของเส้นแบ่งเป็นเส้นที่สัมผัสกับค่าข้อมูลในพื้นที่คุณลักษณะที่ใกล้สุดเส้นขอบของเส้นทั้งสอง ซึ่งถูกแทนด้วย $w^T x+b \geq y \geq 1$ ถ้า $y = 1$ และ $w^T x+b \leq y \leq -1$ ถ้า $y = -1$ หากเส้นขอบใดมีค่ากว้างมากสุดจะแสดงให้เห็นข้อมูลที่แยกกันได้ชัดเจน ดังนั้นเส้นขอบที่มีความกว้างมากที่สุดจะเป็นเส้นแบ่งข้อมูลที่ดีที่สุด โดยเรียกเส้นแบ่งข้อมูลทั้งสองว่าเส้นขอบ [7]

จากการศึกษางานวิจัยที่เกี่ยวข้อง เลือกอารมณ์พื้นฐานมาใช้ในการทดลอง 3 อารมณ์ ได้แก่ อารมณ์โกรธ อารมณ์เศร้า และอารมณ์มีความสุข เพราะทั้ง 3 อารมณ์นี้เป็นอารมณ์พื้นฐานของมนุษย์และสามารถพบได้มากในภาพยนตร์และแอนนิเมชั่น ส่วนของการจำแนกอารมณ์พื้นฐานของมนุษย์จากการแสดงบนใบหน้านั้นยังไม่มีงานวิจัยใดที่ใช้ค่าความถี่ของทิศทางตามค่าเกรเดียนท์มาใช้ในการจำแนกอารมณ์ ผู้วิจัยจึงได้ออกแบบการทดลองกับภาพใบหน้าการ์ตูนโดยนำเทคนิค HOG มาใช้ในการสกัดคุณลักษณะสำคัญและใช้ตัวจำแนกแบบ DT, k-NN และ SVM มาใช้ในการศึกษาและพัฒนาอัลกอริทึมและสำหรับงานวิจัยนี้ในบทถัดไป

3. วิธีการดำเนินงาน

วิธีการทดลองนั้น สามารถแบ่งออกเป็นสามขั้นตอน ดังนี้ การเตรียมข้อมูล (Processing) การสกัดคุณลักษณะ (Feature Extraction) และการจำแนกข้อมูลโดยการทดลองกับชุดข้อมูลภาพถ่ายจากรูปภาพที่คัดลอกมาจากอินเตอร์เน็ต (Internet) โดยทำการคัดเลือกภาพที่เป็นบริเวณใบหน้าการ์ตูนที่มีการแสดงออกบนของอารมณ์บนใบหน้าการ์ตูน จำนวน 60 ภาพ โดยแบ่งเป็นทั้งหมด 3 อารมณ์ โดยแสดงขั้นตอนการดำเนินงานดังภาพที่ 5



ภาพที่ 5: แสดงขั้นตอนวิธีของการจำแนกอารมณ์ของตัวการ์ตูน

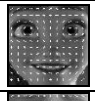
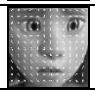
3.1 การเตรียมข้อมูล

งานวิจัยนี้รวบรวมภาพจากการสืบค้นทางอินเตอร์เน็ต มาใช้ในการทดลอง โดยเลือกเฉพาะภาพที่เป็นบริเวณใบหน้าของการ์ตูน โดยตัดภาพใบหน้าจากพื้นหลังด้วยมือ และจำแนกประเภทของอารมณ์ออกเป็น 3 ประเภท ดังแสดงในตารางที่ 1 และ 2 โดยแต่ละอารมณ์จะมี 20 ภาพ รวมทั้งสิ้น 60 ภาพ และปรับขนาดของภาพให้เป็น 100 × 100 พิกเซล เพื่อใช้เป็นชุดข้อมูลสำหรับการทดลอง

3.2 ขั้นตอนการสกัดคุณลักษณะ

ขั้นตอนนี้จะนำชุดภาพที่ได้จากการเตรียมข้อมูลการทดลองมาสกัดหาคุณลักษณะโดยอัตโนมัติด้วยวิธีการหาค่าความถี่ของทิศทางตามค่าเกรเดียนท์ (Histogram of Oriented Gradient) ดังแสดงในตารางที่ 1 โดยจะได้ค่าคุณลักษณะออกมาประมาณ 4,000 คุณลักษณะ

ตารางที่ 1: ตารางแสดงตัวอย่างภาพของการแสดงออกบนใบหน้าของการ์ตูนในแต่ละอารมณ์

ประเภทอารมณ์	ตัวอย่างภาพต้นฉบับ	ตัวอย่างภาพที่ถูกสกัดคุณลักษณะด้วย HOG
โกรธ (Anger)		
มีความสุข (Happiness)		
เศร้า (Sadness)		

3.3 ขั้นตอนการจำแนกประเภท

ในขั้นตอนการจำแนกอารมณ์บนใบหน้าการ์ตูนนั้น ผู้วิจัยได้ทำการทดลองโดยใช้ตัวจำแนก 3 ประเภท คือ DT แบบมาตรฐาน k-NN โดยจากการทดลองเบื้องต้นพบว่ากำหนดค่า k = 3 จะให้ผลการทดลองได้ดีกว่าจำนวนอื่นและใช้ Euclidean Distance เป็นฟังก์ชั่นในการวัดระยะห่างระหว่างประเภท และจำแนกด้วยวิธี SVM โดยใช้เคอร์เนลแบบเชิงเส้น จากนั้นทำการเปรียบเทียบประสิทธิภาพการจำแนกประเภทของทั้ง 3 วิธี ดังแสดงในตารางที่ 2, 3, 4 และ 5 ตามลำดับ

ตารางที่ 2: ผลการจำแนกด้วยต้นไม้ตัดสินใจ

ประเภทที่ถูกต้อง \ ประเภทที่ระบบทำนาย	Anger	Happiness	Sadness	%
Anger	12	2	6	60.0%
Happiness	1	12	7	60.0%
Sadness	5	7	8	40.0%
ค่าเฉลี่ยความถูกต้อง (%)				**53.3%**

ตารางที่ 3: ผลการจำแนกด้วยตัวจำแนกเพื่อนบ้านที่ใกล้เคียงที่สุด k ตัว

ประเภทที่ถูกต้อง \ ประเภทที่ระบบทำนาย	Anger	Happiness	Sadness	%
Anger	17	1	2	85.0%
Happiness	0	18	2	90.0%
Sadness	2	7	11	55.0%
ค่าเฉลี่ยความถูกต้อง (%)				**76.7%**

ตารางที่ 4: ผลการจำแนกด้วยตัวจำแนกแบบซัพพอร์ตเวคเตอร์แมชชีน

ประเภทที่ถูกต้อง \ ประเภทที่ระบบทำนาย	Anger	Happiness	Sadness	%
Anger	19	0	1	95.0%
Happiness	0	18	2	90.0%
Sadness	1	2	17	85.0%
ค่าเฉลี่ยความถูกต้อง (%)				**90.0%**

ตารางที่ 5: ผลการจำแนกด้วยตัวจำแนกแบบ DT, k-NN และแบบ SVM

ตัวจำแนก	ผลการทดลอง
DT	**53.3%**
k-NN (k=3)	**76.7%**
SVM	**90.0%**

4. สรุปผลการทดลอง

การทดลองพบว่าการจำแนกประเภทด้วยวิธีแบบ DT ให้ผลลัพธ์ในการจำแนกร้อยละ 53.3 โดยแสดงผลการทดลองในรูปแบบคอนฟิวชั่นเมทริกซ์ (Confusion Matrix) ตารางที่ 2 พบว่าอารมณ์เศร้าผิดพลาดมากที่สุดโดยทำนายอารมณ์เศร้าเป็นโกรธ 5 ภาพ (ร้อยละ 25) อารมณ์เศร้าเป็นมีความสุขถึง 7 ภาพ (ร้อยละ 35) การจำแนกประเภทด้วยวิธีแบบ k-NN ให้ผลลัพธ์ในการจำแนกร้อยละ 76.7 ดังตารางที่ 3 ซึ่งจะพบว่าระบบทำนายอารมณ์เศร้ามีความผิดพลาดมากที่สุด โดยที่ทำนายอารมณ์เศร้าเป็นโกรธ 2 ภาพ (ร้อยละ 10) อารมณ์เศร้าเป็นมีความสุข 7 ภาพ (ร้อยละ 35) และการจำแนกประเภทด้วยวิธีแบบ SVM สามารถจำแนกอารมณ์แต่ละประเภทได้ดีที่สุด ให้ผลลัพธ์ในการจำแนกสูงถึงร้อยละ 90.0 ดังตารางที่ 4 ซึ่งระบบจำแนกอารมณ์โกรธได้เกือบจะถูกต้องทั้งหมดโดยระบบจำแนกอารมณ์โกรธเป็นเศร้า 1 ภาพ (ร้อยละ 5) กล่าวคือ อารมณ์เศร้า ดังภาพที่ 6 ระบบนำมาจำแนกเป็นอารมณ์โกรธ อารมณ์มีความสุขเป็นเศร้า 2 ภาพ (ร้อยละ 10) คือ อารมณ์เศร้า ดังภาพที่ 7 ระบบนำมาจำแนกเป็นอารมณ์มีความสุข อารมณ์เศร้าเป็นโกรธ 1 ภาพ (ร้อยละ 5) คือ อารมณ์โกรธระบบนำมาจำแนกเป็นอารมณ์เศร้า และอารมณ์เศร้าเป็นอารมณ์มีความสุข 2 ภาพ (ร้อยละ 10) คือ อารมณ์มีความสุขระบบนำมาจำแนกเป็นอารมณ์เศร้า ดังภาพที่ 8 ตามลำดับ



เศร้า เป็น โกรธ

ภาพที่ 6: แสดงผลการทำนายผิดพลาดของ SVM ของอารมณ์โกรธ



เศร้า เป็น มีความสุข

ภาพที่ 7: แสดงผลการทำนายผิดพลาดของ SVM ของอารมณ์มีความสุข



โกรธ เป็น เศร้า มีความสุข เป็น เศร้า

ภาพที่ 8: แสดงผลการทำนายผิดพลาดของ SVM ของอารมณ์เศร้า

จากการทดลองความผิดพลาดในการจำแนกประเภทของตัวจำแนกแบบ SVM ในแต่ละอารมณ์ มีผลมาจากการแสดงออกทางสีหน้าของการ์ตูนนั้นที่มีลักษณะใกล้เคียงกัน จึงส่งผลให้ค่าการกระจายตัวความเข้มของเกรเดียนท์มีค่าใกล้เคียงกันทำให้การจำแนกประเภทอารมณ์ของระบบเกิดความผิดพลาดได้

5. ข้อเสนอแนะ

งานวิจัยนี้นำเสนอวิธีการจำแนกอารมณ์บนใบหน้าการ์ตูน 3 อารมณ์หลัก ได้แก่ อารมณ์ โกรธ มีความสุข และเศร้า โดยใช้คุณสมบัติที่สกัดได้จากเทคนิค HOG และเปรียบเทียบประสิทธิภาพของการจำแนกโดยใช้ตัวจำแนก 3 ประเภท ได้แก่ DT, k-NN และ SVM จากผลการทดลองสรุปได้ว่า SVM ให้ผลการจำแนกดีที่สุดคือ 90.00% แสดงว่าคุณลักษณะที่ได้จากเทคนิค HOG น่าจะใช้ในการจำแนกอารมณ์บนใบหน้าการ์ตูนได้ดี

อย่างไรก็ตาม การทดลองเบื้องต้นยังใช้ชุดข้อมูลขนาดเล็ก ในขั้นถัดไปผู้วิจัยจะทำการทดลองกับชุดข้อมูลที่มีจำนวนมากขึ้น นอกจากนี้ยังอาจปรับปรุงอัลกอริธึมการสกัดคุณลักษณะ HOG จากทั้งภาพใบหน้า เป็นการสกัดคุณลักษณะ HOG แต่ละส่วนของใบหน้าที่มีส่วนสำคัญในการแสดงอารมณ์ ได้แก่ คิ้ว ตา และปาก ซึ่งอาจทำให้ได้คุณลักษณะที่เหมาะสมและช่วยเพิ่มประสิทธิภาพการจำแนกอารมณ์ได้ดีขึ้น

ทั้งนี้เพื่อให้เกิดประโยชน์ในการประยุกต์ใช้ ในการทำดัชนีเพื่อค้นคืนภาพหรือวีดีโอ โดยแบ่งตามอารมณ์บนใบหน้าตัวละครที่เป็นการ์ตูนต่อไป

เอกสารอ้างอิง

[1] Dani Jones, “Tips for Drawing Facial Expression,” 2007.

[2] Arthur Dobrin, “Facial Expressions: Universal vs. Cultural,” Jun 25, 2013.

[3] Devi Arumugam and Purushothaman B.E “Emotion Classification Using Facial Expression,” *International Journal of Advanced Computer Science and Applications*, Vol. 2, No. 7, 2011

[4] Yahia Said, Mohamed Atri and Rached Tourki, “Human Detection Based on Integral Histograms of Oriented Gradients and SVM,” *International Conference, Communications, Computing and Control Applications (CCCA)*, pp. 1- 5, 2011.

[5] Claude C. Chibelushi and Fabrice Bourel “Facial Expression Recognition,” *A Brief Tutorial Overview*, 2012.

[6] ชุมพล เสนาพันธ์ และ คำรณ สุนัต, “การตรวจจับมนุษย์โดยใช้คุณลักษณะตามการแปลงค่าการเก็บรวบรวมข้อมูลร่วมกับค่าความถี่ของค่าทิศทางตามค่าเกรเดียนท์และตัวแบบการเรียนรู้เครื่องจักรเรียนรู้เอ็กซ์ทรีมสำหรับเรเดียลเบสิสทั่วไปรุ่นปรับปรุง” *การประชุม National Conference on Computing and Information Technology (NCCIT) ครั้งที่ 9*, 2556.

[7] อนุสรณ์ อุ่นท้าว และ สมปอง เวฬุวนาธร, “การรู้จำท่ามือภาษาไทยโดยใช้ค่าความถี่ของทิศทางตามค่าเกรเดียนท์ร่วมกับการวิเคราะห์องค์ประกอบหลักและซัพพอร์ตเวกเตอร์แมชชีน.” *การประชุม National Conference on Computing and Information Technology (NCCIT) ครั้งที่ 10*, 2557.

[8] สุรพงค์ เอื้อวัฒนามงคล, “การทำเหมืองข้อมูล Data Mining” *สำนักพิมพ์สถาบันบัณฑิตพัฒนบริหารศาสตร์ สถาบันบัณฑิตพัฒนบริหารศาสตร์* พิมพ์ครั้งที่ 1 พ.ค. 2559 หน้า 35-76.

[9] Christopher D. Manning, Prabhakar Raghavan and Hinrich Schütze, “Introduction to Information Retrieval,” *Cambridge University Press*, 2008.

การคัดเลือกคุณลักษณะแบบกลุ่มสำหรับข้อมูลเบาบาง
Ensemble Feature Selection for Sparse Data

วรรณนภา โพธิ์ผลิ (Wannapa Phopli)[1] สุกัญชลิกา บุญมาธรรม (Sukanchalika Boonmatham)[2]
และพยุง มีสัจ (Phayung Meesad)[3]
[1]สาขาวิชาคอมพิวเตอร์ศึกษา คณะครุศาสตร์ มหาวิทยาลัยราชภัฏธนบุรี
[2]สาขาวิชาคอมพิวเตอร์ประยุกต์ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยราชภัฏเพชรบุรี
[3]ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
wannapa.p@dru.ac.th, sukanchalika.boo@mail.pbru.ac.th, phayung.m@it.kmutnb.ac.th

บทคัดย่อ

บทความนี้นำเสนอการคัดเลือกคุณลักษณะแบบกลุ่มสำหรับข้อมูลแบบเบาบางสูง วิธีการที่นำเสนอเลือกใช้วิธีการโหวตเสียงข้างมากคัดเลือกคุณลักษณะจาก 7 วิธี คือ Information Gain, Information Gain Ratio, Chi Squared, Rule, Deviation, Correlation และ Support Vector Machine ผู้วิจัยใช้ชุดข้อมูลทดสอบจากฐานข้อมูล UCI Machine Learning Database Repository ซึ่งเป็นข้อมูลทางธุรกิจที่มีลักษณะข้อมูลแบบเบาบางสูง ได้แก่ CNAE-9 การทดสอบหาประสิทธิภาพในการจำแนกด้วยเทคนิคการจำแนก k-NN, Naïve Bayes และ Random Forest เพื่อวัดค่าความถูกต้อง

ผลการทดลองเมื่อนำข้อมูลแบบเบาบางเข้าสู่การคัดเลือกคุณลักษณะแบบกลุ่มด้วยวิธีการโหวตแบบเสียงข้างมาก พบว่าวิธีการคัดเลือกคุณลักษณะแบบกลุ่มสามารถลดจำนวนของคุณลักษณะลงได้ดีขึ้น ซึ่งการคัดเลือกคุณลักษณะจำนวน 30% เรียงตามค่าน้ำหนักมากไปหาน้อยของแต่ละวิธีเมื่อผ่านกระบวนการโหวตเลือกคุณลักษณะสามารถลดจำนวนคุณลักษณะจากเดิมลงอีก 9.3% ผลการทดสอบประสิทธิภาพในการจำแนกข้อมูลพบว่าวิธีการ Naïve Bayes ให้ค่าความถูกต้องสูงสุดที่ 93.74% นั่นแสดงให้เห็นว่าวิธีการคัดเลือกคุณลักษณะแบบกลุ่มสามารถลดจำนวนของคุณลักษณะลงในขณะที่สามารถนำมาใช้ในการจำแนกข้อมูลที่ให้ค่าความถูกต้องในการจำแนกสูงขึ้น

คำสำคัญ: การคัดเลือกคุณลักษณะแบบกลุ่ม เหมืองข้อมูล การจำแนกหมวดหมู่ ข้อมูลแบบเบาบาง

Abstract

This paper presents an ensemble feature selection technique for highly sparse data. The proposed method employed seven feature selection techniques: Information Gain, Information Gain Ratio, Chi Squared, Rule, Deviation, Correlation, and Support Vector Machine. The ensemble feature selection were studied with a highly sparse business data set from UCI Machine Learning Database Repository, CNAE-9 data set. Three machine learning models: k-NN, Naïve Bayes, and Random Forest were used as comparative classifiers. Accuracy was used as the performance metric.

For experimental results it is found that the ensemble feature selection method for sparse data could reduce the number of features, dramatically. Top 30% of the features selected based on top weighting from each feature selection methods to input to the ensemble feature selection for voting. The ensemble method can reduce the number of features by additional reduction of 9.3%. For classification performance it shows that the Naïve Bayes method gives the highest accuracy of 93.74%. This illustrates that the ensemble feature selection method can reduce the number of features while enhancing the classification efficiency with higher accuracy.

Keyword: Ensemble Feature Selection, Data Mining, Classification, Sparse Data.

1. บทนำ

การนำข้อมูลสร้างเป็นโมเดลเพื่อใช้ในการทำนายเหตุการณ์ต่างๆ ที่อาจเกิดขึ้นในอนาคต ที่จะให้ค่าความถูกต้องของเหตุการณ์ในการทำนายสูงนั้น เป็นเป้าหมายสำคัญในการพัฒนาประสิทธิภาพของกระบวนการในการทำนาย โดยวิธีการคัดเลือกคุณลักษณะ (Feature Selection) เป็นวิธีการหนึ่งที่นำมาใช้เพื่อลดจำนวนคุณลักษณะจากคุณลักษณะจำนวนมาก เพื่อนำคุณลักษณะเหล่านั้นมาใช้ในการสร้างโมเดล ทำให้ได้โมเดลที่มีประสิทธิภาพในการทำนายที่สูงขึ้น [1] [2] โดยในบทความนี้นำเสนอการคัดเลือกคุณลักษณะแบบกลุ่มสำหรับข้อมูลแบบเบาบางเพื่อให้การคัดเลือกคุณลักษณ์ที่เหมาะสมแล้วนำไปสร้างโมเดลในการทำนายเพื่อให้ค่าความถูกต้องในการทำนายสูง โดยได้ทดลองวิธีการคัดเลือกคุณลักษณะ 7 วิธี ได้แก่ Information Gain, Information Gain Ratio, Chi Squared, Rule, Deviation, Correlation และ SVM ซึ่งเป็นวิธีการคัดเลือกคุณลักษณะในกลุ่มฟิลเตอร์ (Filter Models) มาใช้ในการคัดเลือกคุณลักษณะแบบกลุ่ม โดยนำข้อมูลที่ได้มาทดสอบหาประสิทธิภาพการจำแนกหมวดหมู่ ซึ่งใช้เทคนิคการจำแนกหมวดหมู่ 3 เทคนิค คือ Naïve Bayes, k-NN และ Random Forest และนำค่าที่ได้มาเปรียบเทียบประสิทธิภาพค่าความถูกต้องจากวิธีการเลือกคุณลักษณะดังกล่าว

การนำเสนอเนื้อหาในบทความแบ่งออกเป็น 5 ส่วนเพื่อให้เกิดความชัดเจน ได้แก่ บทนำ วรรณกรรมที่เกี่ยวข้อง วิธีการดำเนินการวิจัย ผลการทดลอง และสรุปผล

2. วรรณกรรมที่เกี่ยวข้อง

2.1 เหมืองข้อมูล (Data Mining)

การทำเหมืองข้อมูลถือได้ว่าเป็นวิธีในการสกัดความรู้จากข้อมูลจำนวนมาก [3] [4] ซึ่งเป็นการจัดการข้อมูลขั้นสูงอย่างหนึ่ง มีลักษณะในการนำชุดฐานข้อมูลขนาดใหญ่ และมีความหลากหลายมาวิเคราะห์ และสืบค้นรูปแบบของกฎที่ซ่อนอยู่ในชุดข้อมูล เพื่อเป็นการแสดงให้ เห็นถึงองค์ความรู้จากการประมวลผลข้อมูล ทั้งนี้ถือได้ว่า กระบวนการเหมืองข้อมูลเป็นกระบวนการสำคัญสำหรับ การสร้างฐานความรู้ (Knowledge Base) เพื่อประยุกต์ใช้ในงานด้านต่างๆ รูปแบบที่สำคัญในการทำเหมืองข้อมูล ประกอบด้วย กฎความสัมพันธ์ การจำแนกหมวดหมู่และการทำนาย การจัดกลุ่มข้อมูล การวิเคราะห์แนวโน้ม และการหาค่าผิดปกติที่เกิดขึ้น [5] สำหรับการจัดทำเหมืองข้อมูลประกอบไปด้วย 5 องค์ประกอบที่สำคัญ ได้แก่ 1) การสกัดข้อมูล การเปลี่ยนแปลงข้อมูล และการโหลดข้อมูล 2) การจัดเก็บและการจัดการข้อมูลในฐานข้อมูล 3) การเข้าถึงข้อมูล 4) การวิเคราะห์ข้อมูลโดยการประยุกต์ใช้ซอฟต์แวร์ 5) นำเสนอและแปลผล [6]

2.2 ข้อมูลแบบเบาบาง (Sparse Data)

ข้อมูลแบบเบาบางคือข้อมูลที่มีลักษณะของจำนวนข้อมูลที่สนใจอยู่น้อย โดยปกติหากข้อมูลถูกทำให้อยู่ในรูปของตัวเลขที่ประกอบด้วย 0 และ 1 โดยที่ 1 คือข้อมูลที่สนใจ ดังนั้นข้อมูลแบบเบาบางจะประกอบไปด้วยข้อมูลที่มี 0 อยู่เป็นจำนวนมาก และมี 1 ซึ่งเป็นข้อมูลที่สนใจอยู่จำนวนน้อย [7]

2.3 การคัดเลือกคุณลักษณะ (Feature Selection)

การคัดเลือกคุณลักษณะเป็นกระบวนการหนึ่งที่สำคัญสำหรับการจัดการความรู้กับชุดข้อมูลมีขนาดใหญ่และมีคุณลักษณะจำนวนมาก ซึ่งชุดข้อมูลที่ใช้สำหรับการจำแนกอาจจะให้ผลลัพธ์ที่ไม่มีประสิทธิภาพ และใช้ทรัพยากรเพิ่มมากขึ้น โดยเฉพาะเวลาในการประมวลผล ซึ่งส่วนมากคุณลักษณะของข้อมูลมักมีความซ้ำซ้อนและไม่สอดคล้องกัน จะส่งผลต่อการจำแนกหมวดหมู่ จึงควรปรับปรุงประสิทธิภาพในการคัดเลือกคุณลักษณะ เพื่อลดซ้ำซ้อนและความไม่สอดคล้องกัน [8] [9]

การคัดเลือกคุณลักษณะในการทำเหมืองข้อมูลมีวัตถุประสงค์ หลักสามประการ คือ 1) เพื่อเพิ่มประสิทธิภาพการทำนาย 2) เพื่อการสังเคราะห์โมเดลได้อย่างรวดเร็ว และ 3) เพื่อลดความซับซ้อนของรูปแบบโมเดล การคัดเลือกคุณลักษณะที่สำคัญมาใช้งาน ซึ่งสามารถแบ่งได้เป็น 3 กลุ่มใหญ่ดังนี้

1) ฟิลเตอร์โมเดล (Filter Models) [10] เป็นการคัดเลือกคุณลักษณะโดยใช้การคำนวณหาค่าน้ำหนัก ซึ่งอาจจะเป็นค่าความสัมพันธ์ระหว่างแต่ละคุณลักษณะและคลาสต่างๆ และเลือกคุณลักษณะ โดยเรียงลำดับตามค่าน้ำหนักที่คำนวณได้แล้วเลือกคุณลักษณะที่มีค่าน้ำหนักมากกว่าที่ต้องการมาใช้งานต่อไป วิธีการนี้ต่างจากวิธีการแรปเปอร์ (Wrapper Models) ที่วิธีการนี้จะไม่มีการสร้างโมเดลเพื่อคัดเลือกคุณลักษณะ เทคนิค

ในการคำนวณค่าน้ำหนักของคุณลักษณะที่นิยมหลายวิธี เช่น Information Gain, Chi-Square หรือ Correlation

2) แรปเปอร์โมเดล (Wrapper Models) [10] เป็นการคัดเลือกคุณลักษณะด้วยการสร้างโมเดลในการจำแนก (Classification Model) ขึ้นมาจากเซตของคุณลักษณะ ที่กำหนดไว้และวัดประสิทธิภาพการทำงานของโมเดลและเลือกเซตของคุณลักษณะที่ทำให้โมเดลมีประสิทธิภาพมากที่สุดมาใช้งาน โดยเริ่มจาก 1) การหาเซตของคุณลักษณะ 2) การประเมินผลการจัดกลุ่มที่มีคุณภาพโดยใช้เซตที่เลือก สุดท้ายทำซ้ำ 1) และ 2) จนกระทั่งได้คุณภาพที่ต้องการ เช่น โมเดลที่ให้ค่าความถูกต้อง (Accuracy) มากที่สุด การคัดเลือกคุณลักษณะด้วยวิธีการนี้แบ่งย่อยได้เป็น 2 รูปแบบ คือ การเพิ่มตัวแปรอิสระ (Forward Selection) และการลดตัวแปรอิสระ (Backward Elimination)

3) เอ็นเซ็มเบลโมเดล (Ensemble Model) ซึ่งเป็นโมเดลที่มีประสิทธิภาพสูง เป็นเทคนิคที่ใช้โมเดลในการจำแนกหลายๆ โมเดลนำมาช่วยในการหาคำตอบ อาจจะได้เป็นด้วยวิธีการหาจากค่าเฉลี่ย (Averaging) หรือการโหวตเสียงข้างมากที่ได้รับความนิยม (Majority vote) [11] หลักการของโมเดลนี้แบ่งออกเป็น 3 วิธีการ คือ วิธีการ Vote Ensemble เป็นการใช้ข้อมูลชุดทดสอบชุดเดียวกันแต่ใช้การสร้างโมเดลโดยการคัดเลือกคุณลักษณะด้วยการโหวตเสียงข้างมาก วิธีการที่ Bootstrap Aggregating (Bagging) โดยการสร้างโมเดลโดยการสุ่มข้อมูลชุดทดสอบออกมาเป็นหลายๆ ชุด แต่ใช้การสร้างโมเดลการจำแนกด้วยเทคนิคเดียวกันทั้งหมด และวิธีการ Random Forest ซึ่งเป็นวิธีการสร้างโมเดลการสุ่มข้อมูลและคุณลักษณะจากชุดทดสอบออกมาเป็นหลายๆ ชุด และการสร้างโมเดลการจำแนกด้วยเทคนิคต้นไม้ตัดสินใจ [12]

2.4 การจำแนกหมวดหมู่ด้วยเทคนิค k-NN, Naive Bayes และ Random Forest

วิธีการเคเนียเรสเนเบอร์ (k-Nearest Neighbor) เป็นหนึ่งในวิธีการที่นิยมใช้กับปัญหาในลักษณะการจำแนกหมวดหมู่ [13] [14] โดยใช้วิธีการคำนวณเพื่อวัดระยะทางระหว่างข้อมูลเพื่อพิจารณาว่าข้อมูลนั้นเข้าใกล้กับจุดใดมากที่สุด โดยคำนวณเพื่อวัดระยะห่างระหว่างข้อมูลตามกลุ่มที่กำหนด หรือกำหนดค่า k ถ้าข้อมูลเข้าใกล้จุด k ใดนั้นหมายความว่าข้อมูลจะอยู่คลาสหรือกลุ่มนั้น โดยเทคนิคในการวัดระยะที่นิยมนำมาใช้ด้วยวิธีการหาระยะทางแบบยุคลิด (Euclidean Distance) โดยการวัดระยะทางว่าอยู่ใกล้กับจุดใดมากที่สุดเพื่อทำนายคลาสของข้อมูล

การเรียนรู้ของเบย์ (Naïve Bayes) เป็นวิธีการจำแนกหมวดหมู่ โดยใช้ค่าความน่าจะเป็นของเหตุการณ์ที่เกิดขึ้นมาคำนวณด้วยวิธีการของ Reverend Thomas ซึ่งวิธีการคำนวณได้ถูกนำเสนอช่วงปี 1800 [15] การจะทำนายกลุ่มของข้อมูลใหม่ จะคำนวณค่าความน่าจะเป็นเพื่อพิจารณาว่าค่าที่ได้ในกลุ่มใดมีค่าสูงกว่าก็จะทำนายเป็นคลาส ตามสมการ (1)

$$P(C|X) = \frac{P(X|C)P(C)}{P(X)} \quad (1)$$

เมื่อ P(C|X) คือ ค่าความน่าจะเป็นของกลุ่ม C ที่จะมีข้อมูล X เกิดขึ้นก่อน และค่า P(X|C) คือ ค่าความน่าจะเป็นของข้อมูล X ที่เป็นกลุ่ม C P(C) เป็นค่าความน่าจะเป็นของกลุ่ม C และ P(X) เป็นค่าความน่าจะเป็นของข้อมูล X

วิธีการจำแนกป่าแบบสุ่ม (Random Forest Classifier) [16] เป็นวิธีแบบ Ensemble Classifier ซึ่งใช้วิธีการในการสร้างต้นไม้ด้วยวิธีเดียวกับการสร้างต้นไม้ตัดสินใจ (Decision tree) โดยมีหลักการเพื่อเลือกคุณลักษณะที่ดีสุด วิธีการสร้างโมเดลจะใช้วิธีการสุ่มคุณลักษณะและสุ่มข้อมูลตัวอย่างจากข้อมูลทั้งหมดที่มีเพื่อนำมาสร้างเป็นต้นไม้ตัดสินใจซึ่งจะสุ่มออกมาหลายๆ ครั้งในการสร้างต้นไม้ จากนั้นนำข้อมูลเข้าทดสอบกับต้นไม้ที่ได้ซึ่งจะทดสอบกับต้นไม้ทุกต้นที่มีอยู่ จากนั้นใช้วิธีการโหวตจากเสียงข้างมาก เพื่อให้ได้คำตอบว่าข้อมูลที่นำมาทดสอบนั้นคือคลาสใด

3. วิธีการดำเนินการวิจัย

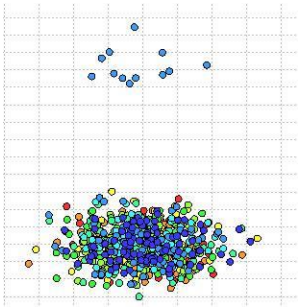
3.1 การศึกษาความต้องการของระบบ

วิธีการคัดเลือกคุณลักษณะแบบกลุ่มเป็นวิธีการหนึ่งที่สามารถนำมาใช้เพื่อเพิ่มประสิทธิภาพของโมเดลในการทำนายได้ โดยบทความนี้ได้ทดลองเปรียบเทียบวิธีการคัดเลือกคุณลักษณะแบบกลุ่มโดยใช้วิธีการในกลุ่มของฟิลเตอร์โมเดล ได้แก่ Information Gain, Information Gain Ratio, Chi Squared, Rule, Deviation, Correlation และ SVM ร่วมกันเพื่อคัดเลือกคุณลักษณะและใช้วิธีการจำแนกหมวดหมู่แบบ Naïve Bayes, k-NN และ Random Forest ในการวัด

ประสิทธิภาพของการจำแนก โดยข้อมูลที่ใช้ในการทดลองเป็นข้อมูลตัวอย่างจากฐานข้อมูลออนไลน์ UCI Machine Learning Database Repository ได้นำข้อมูล CNAE-9 ซึ่งเป็นข้อมูลทางธุรกิจ ที่อยู่ในรูปแบบของเอกสารข้อความ และถูกแปลงให้อยู่ในรูปแบบของตัวเลขเวคเตอร์ ที่มีลักษณะข้อมูลตัวเลข 0 จำนวนมากหรือที่เรียกว่าข้อมูลแบบบาง โดยมีจำนวนทั้งหมด 1,080 เรคคอร์ด 857 คุณลักษณะ แบ่งเป็น 9 คลาส

3.2 การเตรียมข้อมูล (Data Preparation)

เป็นขั้นตอนการเตรียมข้อมูลก่อนที่จะส่งข้อมูลเข้าสู่กระบวนการจำแนก จากข้อมูลข้างต้นได้วิเคราะห์พฤติกรรมของข้อมูล โดยเลือกวิธีการลงจุดแบบกระจายบนกราฟ (Scatter Plot) พบว่าข้อมูลส่วนใหญ่มีการกระจายตัวแบบเกาะกลุ่ม และมีค่าผิดปกติ (Outliers) น้อย



ภาพที่ 1: ตัวอย่างข้อมูลจากการสร้างกราฟแบบ Scatter Plot

3.3 การสร้างโมเดลระบบและการทดสอบประสิทธิภาพ

การสร้างโมเดลและการทดสอบประสิทธิภาพ ได้ทดสอบประสิทธิภาพของการคัดเลือกคุณลักษณะกับเทคนิคการจำแนกหมวดหมู่ โดยแบ่งขั้นตอนการดำเนินงานดังนี้

3.3.1 ชุดข้อมูลสำหรับฝึกและทดสอบ

การทดลองในครั้งนี้ได้นำ CNAE-9 เป็นชุดข้อมูล ซึ่งได้จากฐานข้อมูลออนไลน์ UCI Machine Learning Database Repository [17] จำนวนทั้งสิ้น 1080 เรคคอร์ด 857 คุณลักษณะ โดยได้แบ่งข้อมูลออกเป็นชุดข้อมูลสำหรับฝึก (Training set) จำนวน 799 เรคคอร์ด และใช้ชุดข้อมูลทดสอบ (Test set) จำนวน 281 เรคคอร์ด คิดเป็นเปอร์เซ็นต์ 80 : 20

3.3.2 การคัดเลือกคุณลักษณะแบบกลุ่ม

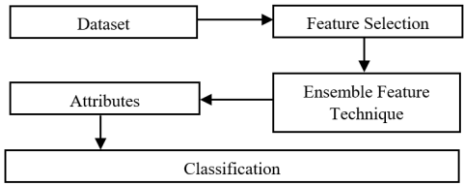
นำชุดข้อมูลมาผ่านกระบวนการคัดเลือกคุณลักษณะด้วยวิธีการทั้ง 7 วิธีการ โดยคัดเลือกคุณลักษณะที่ให้ค่าน้ำหนักสูงสุดเรียงลำดับจากมากไปน้อยจำนวน 5%, 10%, 15%, 20%, 25% และ 30% จากแต่ละวิธีการ จากนั้นนำคุณลักษณะที่ได้มาพิจารณาอีกครั้ง เพื่อเลือกคุณลักษณะที่มีวิธีการคัดเลือกคุณลักษณะเลือกคุณลักษณะนั้นตรงกันในลักษณะ 4 ใน 7 เพื่อใช้ในการจำแนกหมวดหมู่และทดสอบประสิทธิภาพโมเดล

3.3.3 การทดสอบประสิทธิภาพโมเดล

นำข้อมูลที่ผ่านการคัดเลือกแบบกลุ่มเข้าสู่กระบวนการจำแนกหมวดหมู่ โดยใช้การทดสอบข้อมูลแบบข้ามสิบทบ (10-fold Cross-Validation) [18][19] ซึ่งใช้การแบ่งข้อมูลการทดสอบเป็น 10 ชุดย่อย (fold) โดยใช้ชุดข้อมูล 9 ชุด เป็นข้อมูลสำหรับฝึก ส่วนที่เหลืออีก 1 ชุด ไว้เป็นชุดข้อมูลทดสอบ

3.4 วิเคราะห์และสรุปผล

เป็นขั้นตอนสุดท้ายของการดำเนินการทดลองเพื่อการวิเคราะห์ประสิทธิภาพความถูกต้องของโมเดลการทำนาย (Accuracy) รวมถึงค่าความผิดพลาดที่เกิดขึ้น (Root Mean Square Error) และค่าความคลาดเคลื่อนสมบูรณ์ (Absolute Error) ของแต่ละวิธีการคัดเลือกคุณลักษณะ และเทคนิคในการจำแนกที่เลือกใช้ โดยการวิเคราะห์จะสรุปให้อยู่ในรูปของข้อความ (Text Output) และแผนภาพ (Graphics Output)



ภาพที่ 2: ขั้นตอนการดำเนินการทดลอง

4. ผลการทดลอง

จากการทดลองเมื่อนำข้อมูลมาผ่านกระบวนคัดเลือกคุณลักษณะแบบกลุ่มด้วยวิธีการโหวดแบบเสียงข้างมาก จากคัดเลือกคุณลักษณะที่ให้ค่าน้ำหนักสูงสุดเรียงลำดับจากมากไปน้อยจำนวน 5%, 10%, 15%, 20%, 25% และ 30% มาทดสอบเพื่อหาประสิทธิภาพของการจำแนกพบว่า วิธีการคัดเลือกคุณลักษณะแบบกลุ่มนั้นสามารถลดจำนวนของคุณลักษณะลงมากขึ้น ซึ่งการคัดเลือกคุณลักษณะจำนวน 30% เมื่อนำมาผ่านกระบวนการโหวดในการเลือกคุณลักษณะ สามารถลดจำนวนของแอททริบิวต์ลงมากขึ้นจาก 258 เป็น 234 คุณลักษณะ คิด

เป็น 9.3% ของจำนวนคุณลักษณะที่ลดลง การคัดเลือกคุณลักษณะจำนวน 25% สามารถลดลงได้ 12.3% การคัดเลือกคุณลักษณะจำนวน 20% ลดลงได้ 12.3% การคัดเลือกคุณลักษณะจำนวน 15% ลดลงได้ 8.6% การคัดเลือกคุณลักษณะจำนวน 10% ลดลงได้ 22.1% และการคัดเลือกคุณลักษณะจำนวน 5% สามารถลดจำนวนคุณลักษณะลงได้เพิ่มขึ้น 11.7% ซึ่งแสดงไว้ในตารางที่ 1

ตารางที่ 1: จำนวนคุณลักษณะที่ลดลงจากกระบวนการคัดเลือกคุณลักษณะแบบกลุ่มด้วยวิธีการโหวด

ค่าน้ำหนักจากการคัดเลือกคุณลักษณะแบบกลุ่ม (%)	จำนวนคุณลักษณะทั้งหมดจาก 7 วิธี	จำนวนคุณลักษณะ (คัดเลือกแบบกลุ่มโดยโหวตจาก 4 ใน 7)	จำนวนคุณลักษณะที่ลดลง (%)
5	43	38	11.7
10	86	67	22.1
15	129	118	8.6
20	172	151	12.3
25	215	196	12.3
30	258	234	9.3

ซึ่งเมื่อนำไปทดสอบหาประสิทธิภาพในการจำแนกข้อมูล พบว่าที่การคัดเลือกคุณลักษณะแบบกลุ่มจำนวน 30% ให้ค่าความถูกต้องในการจำแนกสูงสุด โดยเมื่อพิจารณาที่เทคนิคการจำแนกข้อมูลด้วย Naïve Bayes ให้ค่าความถูกต้องสูงสุด คือ 93.74% ซึ่งเมื่อพิจารณาที่จำนวนคุณลักษณะที่ถูกคัดเลือก 5% พบว่าให้ค่าความถูกต้องในการจำแนกด้วยวิธีการ k-NN สูงที่สุดคือ 79.85% โดยได้แสดงผลการวัดประสิทธิภาพของการใช้วิธีการคัดเลือกคุณลักษณะแบบกลุ่มด้วยวิธีการโหวดเสียงข้างมาก ซึ่งแสดงไว้ในตารางที่ 2

5. สรุปผลการทดลอง

บทความนี้นำเสนอการคัดเลือกคุณลักษณะแบบกลุ่มสำหรับข้อมูลแบบเบาบางเพื่อเปรียบเทียบจำนวนคุณลักษณะที่สามารถนำมาใช้ในการจำแนกข้อมูลที่ให้ค่าความถูกต้องในการจำแนกสูง โดยใช้ชุดข้อมูลทดสอบ CNAE-9 ซึ่งเป็นข้อมูลทางธุรกิจในลักษณะข้อมูลแบบเบาบางสูง จากฐานข้อมูลออนไลน์ UCI Machine Learning Database Repository โดยได้ทดลองวิธีการคัดเลือกคุณลักษณะ 7 วิธี คือ Information Gain, Information Gain Ratio, Chi Squared, Rule, Deviation, Correlation และ SVM ในการคัดเลือกคุณลักษณะแบบกลุ่มโดยใช้เสียงข้างมาก และทดสอบหาประสิทธิภาพในการจำแนกหมวดหมู่ด้วยเทคนิคการจำแนก k-NN, Naïve Bayes และ Random Forest เพื่อวัดค่าความถูกต้อง (Accuracy) ค่าความคลาดเคลื่อน (Root Mean Square Error) และค่าความคลาดเคลื่อนสมบูรณ์ (Absolute Error)

ผลการทดลองในการคัดเลือกคุณลักษณะ พบว่าการคัดเลือกคุณลักษณะแบบกลุ่มสามารถลดจำนวนของคุณลักษณะลงได้เพิ่มขึ้น และเมื่อพิจารณาที่จำนวนของคุณลักษณะที่ถูกเลือกเพื่อนำมาใช้ทดสอบประสิทธิภาพพบว่าด้วยจำนวนคุณลักษณะที่ถูกเลือกจำนวน 30% ให้ค่าความถูกต้องในการจำแนกสูงสุดด้วยเทคนิค Naïve Bayes โดยมีค่าความถูกต้องในการจำแนกเท่ากับ 93.74% ซึ่งสามารถสรุปได้ว่าวิธีการคัดเลือกคุณลักษณะแบบกลุ่มด้วยวิธีการโหวตเสียงข้างมากสามารถนำมาใช้เพื่อการคัดเลือกคุณลักษณะที่ให้ค่าความถูกต้องในการจำแนกสูง

ตารางที่ 2: จำนวนคุณลักษณะ ค่าความถูกต้องในการจำแนก ค่าความคาดเคลื่อนในการสร้างโมเดล

เทคนิคการจำแนกข้อมูล	ค่าน้ำหนักจากการคัดเลือกคุณลักษณะแบบกลุ่ม (%)	ค่าความถูกต้อง (%)	ค่าความคลาดเคลื่อน	ค่าความคลาดเคลื่อนสมบูรณ์
Naïve Bayes	5	77.35	0.23	0.47
	10	83.73	0.16	0.40
	15	88.61	0.11	0.33
	20	91.87	0.08	0.28
	25	93.12	0.07	0.26
	30	93.74	0.06	0.23
k-NN	5	79.85	0.20	0.45
	10	82.23	0.19	0.42
	15	83.35	0.17	0.41
	20	84.61	0.15	0.39
	25	84.61	0.15	0.39
	30	84.49	0.16	0.39
Random Forest	5	66.20	0.77	0.78
	10	66.84	0.79	0.79
	15	60.31	0.11	0.32
	20	56.71	0.83	0.83
	25	50.31	0.83	0.83
	30	54.81	0.83	0.84

เอกสารอ้างอิง

[1] S. Visalakshi and V. Radha, "A literature review of feature selection techniques and applications: Review of feature selection in data mining," in *2014 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC)*, 2014, pp. 1–6.

[2] R. Kavitha and E. Kannan, "An efficient framework for heart disease classification using feature extraction and feature selection technique in data mining," in *2016 International Conference on Emerging Trends in Engineering, Technology and Science (ICETETS)*, 2016, pp. 1–5.

[3] J. Han, J. Pei, and M. Kamber, *Data Mining: Concepts and Techniques*. Elsevier, 2011.

[4] L. Zhang, Y. Chen, Y. Liang, and N. Li, "Application of Data Mining Classification Algorithms in Customer Membership Card Classification Model," in *2008 International Conference on Information Management, Innovation Management and Industrial Engineering*, 2008, vol. 1, pp. 211–215.

[5] K. Ramamohanarao, J. Bailey, and H. Fan, "Efficient Mining of Contrast Patterns and Their Applications to Classification," in *2005 3rd International Conference on Intelligent Sensing and Information Processing*, 2005, pp. 39–47.

[6] N. Jain and V. Srivastava, "Data Mining Techniques: A Survey Paper," *International Journal of Research in Engineering and Technology*, vol. 02, issue. 11, pp. 116-119, Nov. 2013.

[7] I. H. Wang, S. L. Huang, and K. Y. Lee, "Extracting sparse data via histogram queries," in *2016 54th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, 2016, pp. 39-45.

[8] J. Tang, S. Alelyani, and H. Liu, "Feature selection for classification: A review," in *Data Classification: Algorithms and Applications*, C. Aggarwal, Ed. CRC Press, 2014, p. 37.

[9] B. Azhagusundari and Antony Selvadoss Thanamani, "Feature Selection based on Information Gain," *International Journal of Innovative Technology and Exploring Engineering (IJITEE)*, vol. 2, no. 2, pp. 18–21, Jan. 2013.

[10] C. C. Aggarwal, Ed., "Data Classification," in *Data Classification: Algorithms and Applications*, 1 edition, Boca Raton: Chapman and Hall/CRC, 2014, pp. 42–47.

[11] Q. Shen, R. Diao, and P. Su, in *Feature Selection Ensemble*, A. Voronkov., vol. 10, n.d., pp. 289–306.

[12] เอกสิทธิ์ พัชรวงศ์ศักดา. การวิเคราะห์ข้อมูลด้วยเทคนิคดาต้า ไมน์นิง เบื้องต้น (พิมพ์ครั้งที่ 1). กรุงเทพมหานคร: เอเชีย ดิจิตอลการพิมพ์ จำกัด, 2557.

[13] S. Begum, D. Chakraborty, and R. Sarkar, "Data Classification Using Feature Selection and kNN Machine Learning Approach," in *2015 International Conference on Computational Intelligence and Communication Networks (CICN)*, 2015, pp. 811–814.

[14] İ. G. Tekdemir and B. Alboyacı, "Improvement of power swing detection performance of a distance relay by using k-NN algorithm," in *2015 9th International Conference on Electrical and Electronics Engineering (ELECO)*, 2015, pp. 541–545.

[15] Sfenrianto, I. Purnamasari and R. B. Bahaweres, *"Naïve Bayes Classifier Algorithm and Particle Swarm Optimization for Classification of Cross Selling (Case Study: PT TELKOM Jakarta)*," in *2016 4th International Conference on Cyber and IT Service Management*, 2016, pp. 1–4.

[16] Su Bo, "*Research on the Classification of High Dimensional Imbalanced Data Based on the Optimizational Random Forest Algorithm"* in *2017 9th International Conference on Measuring Technology and Mechatronics Automation*, 2017, *pp*. 228-231.

[17] "UCI Machine Learning Repository: CNAE-9 Data Set." [Online]. Available: https://archive.ics.uci.edu/ml/datasets/CNAE-9. [Accessed: 23-Apr-2017].

[18] M. Vasinek, J. Plato, and V. Snasel, "Limitations on Low Variance k-Fold Cross Validation in Learning Set of Rules Inducers," in *2016 International Conference on Intelligent Networking and Collaborative Systems (INCoS)*, 2016, pp. 207–214.

[19] S. Yadav and S. Shukla, "Analysis of k-Fold Cross-Validation over Hold-Out Validation on Colossal Datasets for Quality Classification," in *2016 IEEE 6th International Conference on Advanced Computing (IACC)*, 2016, pp. 78–83.

การสกัดลักษณะจำเพาะของเสียงสระในภาษาไทย สำหรับการรู้จำเสียงของหุ่นยนต์
Vowel Thai Speech Extraction for Robot Speech Recognition

มนัญชยา อมรพลัง (Mananchaya Amornpalang) [1] และสายยัญ สายยศ (Saiyan Saiyod) [2]
ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น
[1]aoy.mananchaya@gmail.com, [2] saiyan@kku.ac.th

บทคัดย่อ

ในอนาคตหุ่นยนต์ในบ้านจะเข้ามามีบทบาทในชีวิตประจำวันของคนไทยมาก การสั่งงานหุ่นยนต์ด้วยเสียงภาษาไทยจึงมีความจำเป็นที่ช่วยพัฒนาหุ่นยนต์เพื่อให้หุ่นยนต์สามารถจำแนกเสียงภาษาไทยได้ โดยมีการจำแนกแบบออฟไลน์ การสกัดค่าลักษณะจำเพาะของเสียงสระในภาษาไทยจึงเป็นส่วนหนึ่งที่ช่วยพัฒนาให้หุ่นยนต์มีความสามารถในการจำแนกเสียง ซึ่งในการวิเคราะห์เสียงสระนั้นจะพิจารณาจากค่าความถี่ฟอร์แมนต์ของเสียงสระ โดยมีการบันทึกเสียงทั้งเพศชายและเพศหญิงเพื่อมาทำการวิเคราะห์ มีการบันทึกเสียงสระเสียงเดี่ยว ซึ่งแบ่งออกเป็น 2 กลุ่มคือ กลุ่มสระเสียงสั้น (อะ อิ อึ อุ เอะ แอะ โอะ เออะ เอาะ) กลุ่มสระเสียงยาว (อา อี อือ อู เอ แอ โอ เออ ออ) จากนั้นนำเสียงที่ได้ไปหาค่าความถี่ฟอร์แมนต์และหาลักษณะจำเพาะของเสียงสระแต่ละเสียง โดยผลการทดลองวัดความถูกต้องในการจำแนกสระเสียงสั้นเท่ากับร้อยละ 81 และสระเสียงยาวเท่ากับร้อยละ 76 และวัดความถูกต้องของสระเสียงเดี่ยวทั้งหมด 18 เสียง เท่ากับร้อยละ 79

คำสำคัญ: การรู้จำเสียงภาษาไทย, เสียงพูดภาษาไทย, ความถี่ฟอร์แมนต์, หุ่นยนต์, การควบคุมหุ่นยนต์

Abstract

In the future, Robots at home will have a role in the daily lives of Thai people. So that, the robotic voice commanding for Thai language is necessary to develop for Thai classifying ability in offline mode. Extraction of Thai vowels is part of developing a robot capable of voice recognition. The vowels analyses are considered by means of vowel formant frequencies. The recordings both male and female are used to analyze. Recordings of single vowel are divided into two groups which are a short vowel group and a long vowel group. Then, each Thai vowel is analyzed to find formant frequencies and the specific nature. The experimental results show that the classification accuracy of the short vowels and the long vowels are 81% and 76%, respectively. And, the classification accuracy of the 18 single vowels is 79%.

Keyword: Thai Speech Recognition, Thai Speech, Formant frequency, Robot, Robot Controlling

1. บทนำ

เนื่องจากปัจจุบันหุ่นยนต์เป็นเทคโนโลยีอย่างหนึ่งที่มีการพัฒนาอย่างต่อเนื่อง หุ่นยนต์คือ เครื่องจักรใช้งานแทนมนุษย์ ที่ออกแบบให้สามารถใช้งานได้หลากหลายหน้าที่ รวมถึงหุ่นยนต์ที่ใช้ในการทำงานในบ้านหรือเรียกว่าหุ่นยนต์บ้าน ที่สร้างมาเพื่อช่วยอำนวยความสะดวกให้กับมนุษย์ โดยมีการตั้งลำดับการทำงานของหุ่นยนต์ จนถึงการเคลื่อนไหวของหุ่นยนต์ที่ควบคุมโดยมนุษย์หรือการตั้งโปรแกรมการทำงานอัตโนมัติ โดยทั่วไปหุ่นยนต์จะถูกสร้างขึ้น เพื่อให้มนุษย์สามารถใช้งานหุ่นยนต์ได้

สังคมมนุษย์การสื่อสารที่มีสำคัญเกิดจากเสียงที่เปล่งออกมาเป็นภาษาไม่ว่าจะเป็นเสียงพูด เสียงดนตรี ก็ทำให้มนุษย์เกิดความเข้าใจซึ่งกันและกันได้ ในสังคมไทยจะมีภาษาไทยที่ใช้ในการสื่อสาร โดยภาษาไทยแบ่งหน่วยเสียงออกเป็น 3 หน่วยเสียง คือ หน่วยเสียงสระ หน่วยเสียงพยัญชนะ และหน่วยเสียงวรรณยุกต์ โดยผู้เชี่ยวชาญทางวิชาสรีรสัทศาสตร์กล่าวว่าเสียงสระนั้นเกิดความสัมพันธ์กับเสียงอื่นในพยางค์โดยทำหน้าที่เป็นใจกลางของพยางค์ให้เสียงอื่นเกาะหน้าเกาะหลังได้

และในแต่ละพยางค์จะต้องมีเสียงสระเสมอ[1] ดังนั้นเสียงสระจึงเป็นหน่วยเสียงที่สำคัญที่จะทำให้พยางค์ในภาษาไทยเกิดขึ้น

เทคโนโลยีการรู้จำเสียงพูด เป็นระบบที่ช่วยแปลงเสียงพูดให้เป็นข้อความที่ให้ระบบคอมพิวเตอร์สามารถนำไปใช้ในการประมวลผลต่อได้ เพื่อช่วยลดข้อจำกัดในการเข้าถึงเทคโนโลยีสมัยใหม่ซึ่งผู้จัดทำได้นำเทคโนโลยีการรู้จำเสียงพูดมาประยุกต์ใช้กับหุ่นยนต์ ให้หุ่นยนต์นั้นสามารถรับรู้เสียงพูดของมนุษย์ได้

เมื่อหุ่นยนต์ถูกพัฒนาให้สามารถทำงานแทนมนุษย์ได้การสื่อสารระหว่างมนุษย์กับหุ่นยนต์ก็มีความสำคัญ ทำให้หุ่นยนต์สามารถทำงานตามเป้าหมายที่กำหนดได้ การนำเสียงภาษาไทยมาเป็นตัวกลางในการสั่งงานระหว่างมนุษย์กับหุ่นยนต์ จะเป็นสิ่งที่สร้างความสะดวกให้มนุษย์และช่วยให้มนุษย์สั่งงานหุ่นยนต์ได้ หุ่นยนต์สามารถทำงาน ณ ขณะเวลานั้นได้ ส่งผลให้การทำงานของหุ่นยนต์มีคุณภาพและเกิดประโยชน์กับมนุษย์ โดยในงานวิจัยมุ่งเน้นการวิเคราะห์เสียงสระ เนื่องจากเป็นส่วนสำคัญของเสียงภาษาไทย โดยเสียงสระคือเสียงก้องที่เปล่งออกมาโดยให้ลมออกทางช่องปากโดยไม่ถูกลิ้นกักหรือขัดขวาง ดังนั้นเวลาออกเสียงสระจะออกเสียงได้สะดวกและออกเสียงได้นาน ทั้งนี้เพราะคุณสมบัติของเสียงสระมีความดังเด่นกว่าเสียงอื่น ๆ ที่เรียงอยู่ข้างเสมอ[2] โดยในการวิเคราะห์เสียงสระนี้จะพิจารณาจากคุณสมบัติความถี่ฟอร์แมนต์ของเสียง ซึ่งการหาค่าความถี่ฟอร์แมนต์จะมีความแตกต่างกัน เมื่อหาค่าจากเสียงที่มีลักษณะเสียงก้อง ดังนั้นการสกัดลักษณะจำเพาะของเสียงสระในภาษาไทย สำหรับการรู้จำเสียงของหุ่นยนต์ จัดทำเพื่อพัฒนาหุ่นยนต์ให้สามารถจำแนกเสียงสระภาษาไทยได้ ซึ่งเป็นส่วนหนึ่งของการพัฒนาหุ่นยนต์ให้สามารถรับรู้คำสั่งเสียงภาษาไทยจากมนุษย์ได้

2. วรรณกรรมที่เกี่ยวข้อง

2.1 แนวคิดเกี่ยวกับเสียง

เสียงในภาษา เป็นเสียงที่มนุษย์ใช้ในการสื่อสารกันในสังคมให้เกิดความเข้าใจและให้สังคมดำเนินต่อไปได้ ซึ่งคำว่า “เสียงในภาษา” หมายถึง เสียงที่มนุษย์เปล่งออกมา เพื่อใช้สื่อสารความหมายระหว่างมนุษย์ด้วยกัน เพื่อให้เข้าใจความต้องการ ความรู้สึกนึกคิดให้ผู้ฟังเข้าใจซึ่งกันและกัน โดยอวัยวะที่ทำให้เกิดเสียง ได้แก่ ปอด หลอดลม กล่องเสียง เส้นเสียง ลิ้นไก่ ลิ้น ฟัน ปุ่มเหงือก ปาก และจมูก[3]

เสียงสระ หมายถึงเสียงที่เกิดขึ้นจากการทำให้ลมออกจากปอดผ่านหลอดลมและกล่องเสียงที่ลำคอ ออกมาพ้นช่องปากหรือช่องจมูก โดยไม่ถูกสกัดกั้น ณ.ที่หนึ่งที่ใดในช่องทางเดินของลม ในการออกเสียงสระจะมีการออกเสียงได้ยาวนาน เพราะเมื่อออกเสียงสระจะไม่มีการกีดกั้นลมที่ตำแหน่งใด ๆ ในช่องปาก แต่ละเสียงเมื่อนำมาใช้ร่วมกับพยัญชนะก็จะได้คำที่มีความหมายแตกต่างกันไปทั้งสระเสียงสั้น และเสียงยาว ลักษณะของการออกเสียงสระอาศัยลิ้นและริมฝีปาก

ระดับความสูง-ต่ำของเสียง (Pitch) ในการสื่อสารของมนุษย์คำพูดนั้นเป็นสิ่งสำคัญมาก การพูดคุยของมนุษย์มีความแตกต่างของ ระดับเสียงอย่างชัดเจน ซึ่งสังเกตุได้จากระดับเสียงในการพูดของผู้หญิงจะมีระดับซึ่งที่สูงกว่าระดับเสียงพูดของผู้ชาย [4]

2.2 แนวคิดเกี่ยวกับความถี่ฟอร์แมนต์

ความถี่ฟอร์แมนต์ (Formant frequency) คือ ความถี่ที่เกิดจากการสั่งพ้องของเสียงที่กระทบกับอวัยวะภายในช่องปาก ทำให้ความถี่มีการเปลี่ยนแปลง ฟอร์แมนต์เป็นลักษณะที่สำคัญของเสียงพูด เป็นส่วนช่วยให้ออกเสียงได้ชัดเจนมากขึ้น และความถี่ฟอร์แมนต์มีค่าที่แตกต่างกันไปในแต่ละเสียงพูด[5]

โดยในการหาค่าฟอร์แมนต์ใช้วิธีการหาสัมประสิทธิ์การทำนายเชิงเส้น (Linear Predictive Coefficient : LPC) ซึ่งเป็นเทคนิคที่ใช้ได้ดีสำหรับสัญญาณเสียงพูด การประมาณพันธะเชิงเส้นเป็นวิธีการที่วิเคราะห์สัญญาณเสียงที่ผ่านเข้ามาเพื่อหาสัมประสิทธิ์ตัวทำนายไปเปรียบเทียบกับสัญญาณจริง[6] สัมประสิทธิ์การทำนายเชิงเส้นมีสมการ[7] ดังสมการที่ 1

$$\hat{x}(n)= -a(2)x(n-1) - a(3)x(n-2) -\cdots- a(p+1)x(n-p) \quad \ldots(1)$$

โดย p คือลำดับของการทำนายสมการพนุนาม

ถ้าในการหา LPC ไม่ได้ระบุค่าลำดับการทำนายของสมการพหุนาม จะมีการกำหนดค่า p ดังสมการที่ 2 โดยที่ค่า p จะต้องมีค่าน้อยกว่าหรือเท่ากับจำนวนของค่า x [7]

$$p = n(x) - 1 \quad \ldots(2)$$

โดย n(x) คือจำนวนของค่า x จำนวน n ค่า

2.3 แนวคิดเกี่ยวกับ Normalize Data

เป็นวิธีการปรับหน่วยการวัดของข้อมูล ให้อยู่ในหน่วยเดียวกันหรือมาตรฐานเดียวเพื่อให้การเปรียบเทียบระหว่างข้อมูลได้ง่ายขึ้น โดยงานวิจัยใช้การปรับข้อมูลด้วยวิธีการ Normalization to a norm of one เป็นการปรับค่าข้อมูลให้อยู่ระหว่างค่า 0-1 เป็นวิธีที่ใช้กับข้อมูลเวกเตอร์ในหลายมิติ มีการหาค่า norm ของเวกเตอร์[8] ดังสมการที่ 3

$$||y|| = \sqrt{\sum_{i=1}^{n}(x_i^2)} \quad \ldots(3)$$

โดย $||y||$ คือ Norm of Vector

x_i คือ ค่าข้อมูลของเวกเตอร์ที่ i = 1,2,…,n

n คือ ขนาดเวกเตอร์

ดังนั้นการหาค่า Normalization with norm จะได้ดังสมการที่ 4

$$\tilde{a} = \begin{bmatrix} \frac{x_1}{||y||} \\ \frac{x_2}{||y||} \\ \frac{x_3}{||y||} \\ \vdots \\ \frac{X_n}{||y||} \end{bmatrix} \quad \ldots(4)$$

โดย $||y||$ คือ Norm of Vector

x_i คือ ค่าข้อมูลของเวกเตอร์ที่ i = 1,2,…,n

$\tilde{a}$ คือ The normalized vector

2.4 งานวิจัยที่เกี่ยวข้อง

ชัยยงค์ โครินทร์พันธ์และสุกัญญา ผ่านสำแดง (2012) กล่าวถึง วิธีการสังเคราะห์ฟอร์แมนท์เพื่อสังเคราะห์เสียงพูดภาษาไทย วิธีการนี้มีส่วนประกอบที่เป็นปัจจัยสำคัญคือ การนำรูปแบบของการออกเสียงที่เรียกว่ากฎ เริ่มแรกหลักการนี้จะเก็บค่าเสียงพูดในรูปแบบของความถี่ฟอร์แมนท์ ซึ่งประกอบด้วยกลุ่มพยัญชนะ, กลุ่มสระ และกลุ่มตัวสะกด โครงสร้างเหล่านี้จะทำงานร่วมกันกับการดัดแปลงความถี่ฟอร์แมนท์ ซึ่งเรียกว่ากฎการออกเสียงภาษาไทย โดยจะเริ่มต้นด้วยการแทนที่ข้อมูลที่นำเข้าซึ่งเป็นอาเรย์ ของข้อมูลหน่วยเสียงต่าง ๆ ของพยางค์ด้วยความถี่ฟอร์แมนท์ที่ถูกต้อง แล้วผลที่ได้จากการประมวลผลนี้ไปดัดแปลงด้วยการผ่านค่าพารามิเตอร์ การสังเคราะห์เสียงพูดภาษาไทยสำหรับหุ่นยนต์นี้จะทำการทดสอบผลที่ได้โดยใช้ตัวอย่าง 50 ตัวอย่างของหน่วยเสียงภาษาไทย ให้ 10 คนฟังในห้องที่มีการควบคุม และเสียงที่สังเคราะห์ออกมาสามารถฟังแล้วเข้าใจได้ ผลการประเมินเสียงภาษาไทยที่ได้ฟังมีความถูกต้องร้อยละ 71.8 [2]

ไพลิน แสงเพ็ชรและสุพจน์ นิตย์สุวัฒน์ (2012) กล่าวถึงระบบระบุผู้พูดด้วยเสียงโดยใช้โครงข่ายประสาทเทียม มีการประยุกต์ใช้สัมประสิทธิ์เซปสตรัมบนสเกลเมลในการหาคุณลักษณะเด่นของความถี่เสียง เป็นส่วนที่จะทำให้ได้คุณสมบัติเสียงของแต่บุคคลออกมา และก็นำคุณสมบัติเสียงเข้ากระบวนการเรียนรู้ในโครงข่ายประสาทเทียม โดยในการทดสอบผู้วิจัยได้ทำบันทึกเสียงของคน จำนวน 10 คน คนละ 10 ครั้ง การทดสอบปรากฏว่ามีความถูกต้องในการระบุตัวผู้พูดร้อยละ 80 ซึ่งกระบวนการทำงานแบ่งออกเป็น 2 ขั้นตอนคือ

1.ขั้นตอนการเรียนรู้ จะมีการป้อนสัญญาณเสียงที่บันทึกไว้เข้าระบบ จากนั้นสัญญาณเสียงจะถูกนำไปเข้ากระบวนการปรับแต่งเสียงและตัดคำ เพื่อนำมาใช้ในการหาคุณสมบัติของเสียง เมื่อได้คุณสมบัติเสียงของแต่ละบุคคล นำไปเข้ากระบวนการเรียนรู้จำโดยโครงข่ายประสาทเทียม จึงทำให้ได้โมเดลโครงข่ายประสาทเทียมขึ้นมา

2. ขั้นตอนการทดสอบ เป็นขั้นตอนที่มีการทำงานย่อยคล้ายกับขั้นตอนการเรียนรู้ มีการนำสัญญาณเสียงมาหาคุณสมบัติเสียง และนำคุณสมบัติที่ได้มาเทียบกับโมเดลโครงข่ายประสาทเทียมจากนั้นดูผลลัพธ์ที่ออกมาใกล้เคียงกับผู้พูดหรือไม่[9]

3. วิธีการดำเนินงานวิจัย

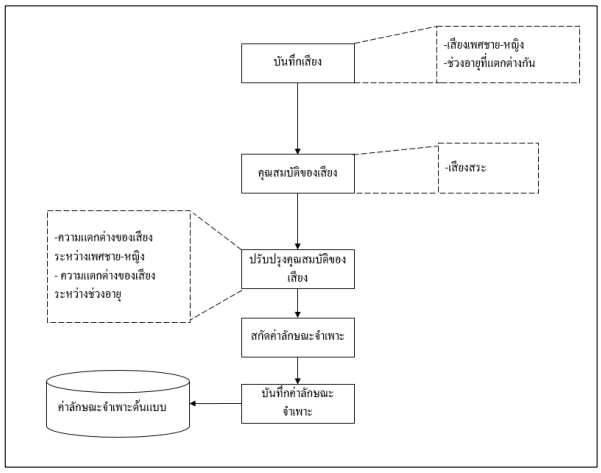
งานวิจัยนี้ใช้เครื่องมือ MATLAB โดยมีวิธีการวิเคราะห์เสียงสระ ดังภาพที่ 1

3.1 บันทึกเสียง

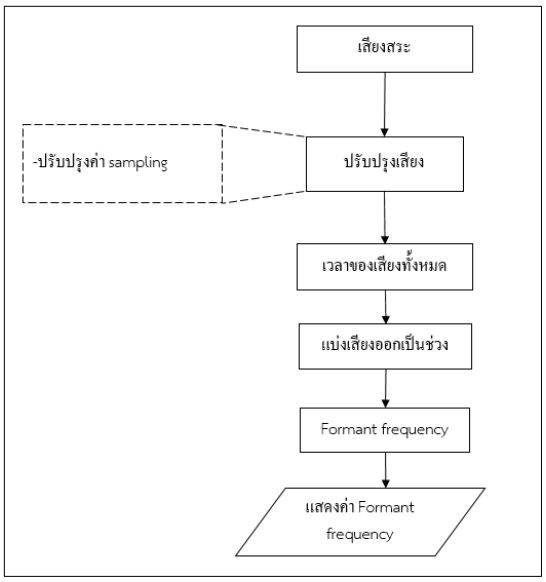
บันทึกเสียง มีการแยกบันทึกเสียงเสียงระหว่างเพศชาย-หญิง ในช่วงอายุที่แตกต่างกัน เพื่อสามารถนำค่าที่ได้จากการวิเคราะห์ไปใช้ได้กับบุคคลทุกเพศทุกวัย โดยการบันทึกใช้ Sample rate 44100 Hz และ Bit Depth 16 bits

3.2 คุณสมบัติของเสียง

เป็นการหาค่าคุณสมบัติของเสียงสระ โดยแสดงขั้นตอนการทำงาน ดังภาพที่ 2



ภาพที่ 1 ภาพรวมระบบในส่วนการวิเคราะห์เสียง



ภาพที่ 2 แผนภาพการหาคุณสมบัติของเสียงสระ

จากภาพที่ 2 เสียงสระ จะมีการวิเคราะห์จากการหาความถี่ฟอร์แมนต์ ซึ่งมีขั้นตอนคือ เมื่อได้เสียงสระ นำมาปรับปรุงเสียง คือการปรับค่า Sample rate ของเสียงที่บันทึกมา ซึ่งได้กำหนดค่า Sample rate เท่ากับ 8000 Hz จากนั้นนำเสียงมาหาระยะเวลาทั้งหมดของเสียง เพื่อนำใช้ในการป้องกันไม่ให้เกิดการแบ่งช่วงเสียงมาวิเคราะห์เกินระยะเวลาของเสียง ในส่วนของการแบ่งช่วงเวลาได้กำหนดระยะเวลาของช่วงอยู่ที่ 0.03 วินาที และในการคำนวณจะแบ่งช่วงเวลาให้ Overlap กัน 0.015 วินาที โดยทำการแบ่งช่วงเสียงเพื่อเพิ่มความละเอียดในการวิเคราะห์ความถี่ฟอร์แมนต์และนำเสียงแต่ละช่วงไปหาความถี่ฟอร์แมนต์ โดยค่าที่ได้ออกมามี 3 ค่า แทนด้วยสัญลักษณ์ f1 คือ ความถี่ฟอร์แมนต์ลำดับที่ 1

f2 คือ ความถี่ฟอร์แมนต์ลำดับที่ 2

f3 คือ ความถี่ฟอร์แมนต์ลำดับที่ 3

เมื่อได้ค่าความถี่ฟอร์แมนต์ครบทุกช่วงแล้วจะแสดงค่าความถี่ฟอร์แมนต์ ดังตารางที่ 1

ตารางที่ 1 แสดงความถี่ฟอร์แมนต์ของสระอะ (เพศชาย)

สระอะ	0.0001-0.03	0.015-0.045	0.03-0.06	0.045-0.075	0.06-0.09	0.075-0.105
f1	743	761	782	766	780	781
f2	1452	1434	1433	1432	1414	1438
f3	2814	2818	2792	2782	2690	2730

จากตารางที่ 1 เป็นข้อมูลของความถี่ฟอร์แมนต์ในแต่ละช่วงเวลาที่แบ่งไว้ ดังนี้นำค่าข้อมูลที่ได้มาหาค่ามัธยฐาน (Median) เพื่อให้ข้อมูลนี้เป็นตัวแทนของเสียงสระ ในการวิเคราะห์ในขั้นต่อไป จากข้อมูลตารางที่ 1 จะได้ค่ามัธยฐานดังตารางที่ 2

ตารางที่ 2 แสดงความถี่ฟอร์แมนต์ของสระอะ (เพศชาย)

สระอะ	ค่ามัธยฐาน
f1	743
f2	1452
f3	2814

3.3 ปรับปรุงคุณสมบัติของเสียง

การปรับปรุงคุณสมบัติของเสียง เมื่อได้ทำการหาค่าคุณสมบัติของเสียงสระจากขั้นที่ 3.2 ครบทุกเสียงที่บันทึกมาทั้งเพศชายและเพศหญิง ขั้นทำการหาค่าเฉลี่ยของข้อมูลของแต่ละเพศ ดังตารางที่ 3 และ 4

ตารางที่ 3 แสดงค่าเฉลี่ยความถี่ฟอร์แมนต์ของสระเสียงสั้น

สระ	ความถี่ฟอร์แมนต์ เพศชาย			ความถี่ฟอร์แมนต์ เพศหญิง		
	f1	f2	f3	f1	f2	f3
อะ	813	1469.5	2612.5	1033.25	1687	2749.75
อิ	301	2366	3285	378.25	2956.25	3547.75
อึ	287.5	1841	2494	421	1664	2967.5
อุ	262	602.5	2476	386.5	563.25	2681.75
เอะ	405.5	2379.5	2807	445.25	2609.5	3265

ตารางที่ 3 แสดงค่าเฉลี่ยความถี่ฟอร์แมนต์ของสระเสียงสั้น (ต่อ)

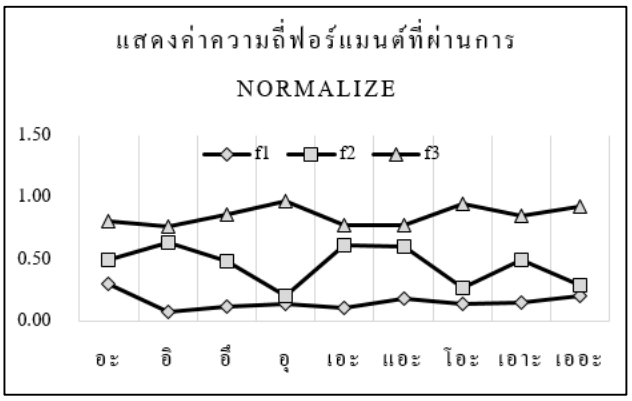
สระ	ความถี่ฟอร์แมนต์ เพศชาย			ความถี่ฟอร์แมนต์ เพศหญิง		
	f1	f2	f3	f1	f2	f3
แอะ	648.5	2042	2700	726	2339.25	3027.25
โอะ	374	701.5	2636	417.5	797.75	2808
เออะ	427	1392.5	2404.5	491.75	1558	2652.25
เอาะ	642.5	921	2626.5	673.75	987	3078.25

ตารางที่ 4 แสดงค่าเฉลี่ยความถี่ฟอร์แมนต์ของสระเสียงยาว

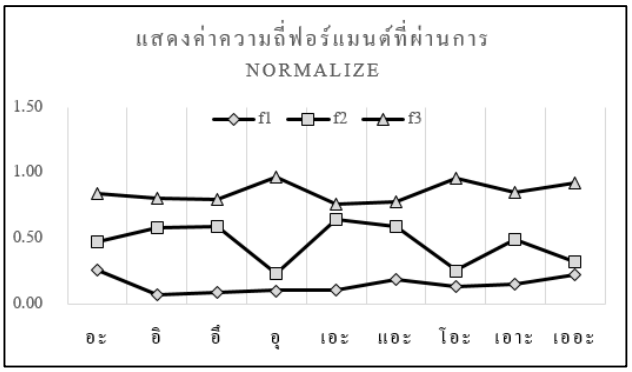
สระ	ความถี่ฟอร์แมนต์ เพศชาย			ความถี่ฟอร์แมนต์ เพศหญิง		
	f1	f2	f3	f1	f2	f3
อะ	765	1481	2658	894.5	1645.75	2743.5
อิ	295	2282	3119	428	2909.5	3559
อึ	315	1981	2615	431.5	1728.75	2879.5
อุ	266.5	609	2469.5	400.25	590	2738.25
เอะ	383	2322	2843	443.5	2608.5	3263.75
แอะ	590.5	2020	2646.5	687.5	2362	3081.5
โอะ	362	676	2703	424	808.25	2915.75
เออะ	441	1622.5	2516	473.75	1503.5	2601.5
เอาะ	571	869.5	2695	600	987	2926.25

3.4 สกัดค่าลักษณะจำเพาะ

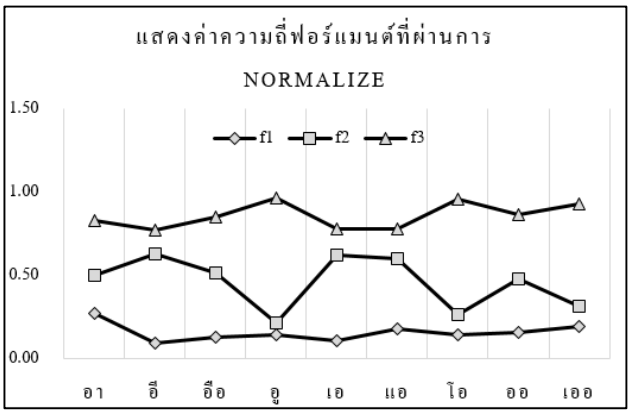
สกัดค่าลักษณะจำเพาะ เป็นการนำค่าคุณสมบัติของเสียงที่ได้ในแต่ละกลุ่ม มาพิจารณาหาค่าลักษณะจำเพาะ เพื่อได้ค่าที่สามารถแยกเสียงสระแต่ละเสียงออกจากกันได้ โดยทำการ Normalize ข้อมูล ดังสมการที่ 3และ 4 ให้มีค่าอยู่ระหว่าง 0-1 เพื่อให้ง่ายต่อการนำไปจำแนก โดยข้อมูลจากตารางที่ 3 และ 4 สามารถนำข้อมูลมา Normalize และ แสดงเป็นกราฟได้ดังภาพที่ 3, 4, 5 และ 6 จะเห็นได้ว่ากราฟมีลักษณะกราฟมีแนวโน้มที่คล้ายกัน ทำให้สามารถนำค่าที่ได้จากการสกัดค่าเฉพาะมาหาค่าเฉลี่ยเพื่อเป็นต้นแบบของสระแต่ละเสียงได้ และเป็นต้นแบบในการเปรียบเทียบในการจำแนกสระภาษาไทย โดยเมื่อได้ค่าลักษณะจำเพาะต้นแบบแล้วได้มีการกำหนดเงื่อนไขในการจำแนกสระแต่ละเสียงดังภาพที่ 7 ที่มีการกำหนดเงื่อนไขจากความถี่ฟอร์แมนต์ลำดับที่ 1 และความถี่ฟอร์แมนต์ลำดับที่ 2 ซึ่งสระเสียงสั้นและสระเสียงยาวใช้เงื่อนไขในการเปรียบเทียบเดียวกันแต่จะระยะเวลาของเสียงสระในการแยกระหว่างสระเสียงสั้นและสระเสียงยาว



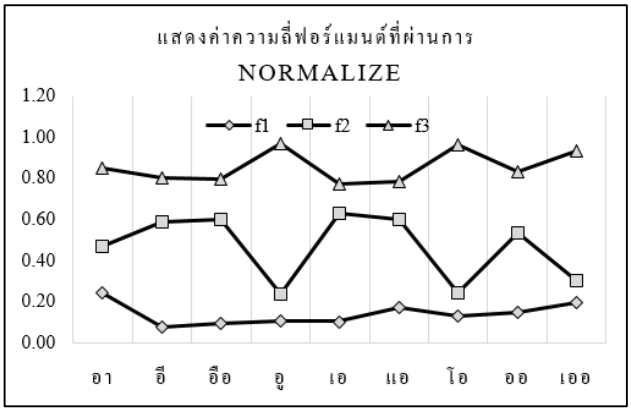
ภาพที่ 3 แสดงค่าความถี่ฟอร์แมนต์ของสระเสียงสั้นที่ผ่านการ Normalize (เพศหญิง)



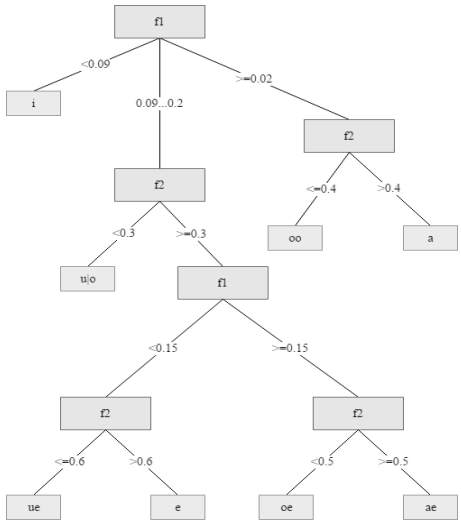
ภาพที่ 4 แสดงค่าความถี่ฟอร์แมนต์ของสระเสียงสั้นที่ผ่านการ Normalize (เพศชาย)



ภาพที่ 5 แสดงค่าความถี่ฟอร์แมนต์ของสระเสียงยาวที่ผ่านการ Normalize (เพศหญิง)



ภาพที่ 6 แสดงค่าความถี่ฟอร์แมนต์ของสระเสียงยาวที่ผ่านการ Normalize (เพศชาย)



ภาพที่ 7 เงื่อนไขในการเปรียบเทียบ (f1:ความถี่ฟอร์แมนต์ลำดับที่ 1 f2: ความถี่ฟอร์แมนต์ลำดับที่ 2)

4. ผลการทดลอง

การทดลองได้บันทึกเสียงสระ 18 สระ (สระเสียงสั้น 9 สระ และสระเสียงยาว 9 สระ) จากเพศชาย 3 คน และเพศหญิง 4 คน ทั้งหมด จำนวน 126 เสียง แบ่งเป็นเสียงผู้ชาย 54 เสียง และเสียงผู้หญิง 72 เสียง ซึ่งจากการทดลองได้ผลการทดลองว่าสามารถจำแนกเสียงของสระเสียงสั้น และสระเสียงยาว ได้ดังตารางที่ 5

5. สรุป

การทำวิจัยนี้มีวัตถุประสงค์เพื่อการจำแนกเสียงสระภาษาไทย สำหรับการนำมาใช้ในการรู้จำเสียงของหุ่นยนต์ โดยมีการวิเคราะห์เฉพาะสระเสียงเดี่ยว จากผลการทดลองงานวิจัยนี้มีการวัดความถูกต้องในการจำแนกเสียงสระเท่ากับร้อยละ 79 ซึ่งการทำงานวิจัยนี้ได้พบปัญญาคือยังมีสระอยู่ 2 กลุ่ม คือ 1.อุ โอะ 2. อู โอ ที่ยังไม่สามารถมีเงื่อนไขการเปรียบเทียบค่าแล้วแยกเสียงในแต่ละกลุ่มออกจากกันได้ และผู้ทำวิจัยจะพัฒนาวิเคราะห์เสียงพยัญชนะและเสียงวรรณยุกต์ เพื่อทำให้หุ่นยนต์สามารถรู้จำเสียงได้

ตารางที่ 5 ผลการทดลอง

สระ	ประสิทธิภาพในการจำแนกเสียง		
	ถูก(ครั้ง)	ผิด(ครั้ง)	ร้อยละ
เสียงสั้น	51	12	81
เสียงยาว	48	15	76
รวม	99	27	79

เอกสารอ้างอิง

[1] อุดม วโรตม์ลิกขิตถ์. *ภาษาศาสตร์เบื้องต้น*. ค้นเมื่อ 15 ตุลาคม 2558, จาก http://e-book.ram.edu/e-book/inside/html/dlbook.asp? code=LI210(50)

[2] ชัยยงค์ โคกินทร์พันธ์, สุกัญญา ผ่านสำแดง. (2556). *การสังเคราะห์เสียงภาษาไทยโดยการสังเคราะห์ฟอร์แมนต์สำหรับหุ่นยนต์*. โครงงานคอมพิวเตอร์ปริญญาวิทยาศาสตรบัณฑิต สาขาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น

[3] เปรมปราณี เทียบตา .[ม.ป.ป.] .*เสียงในภาษา* [ฉบับออนไลน์]โรงเรียน . พิษณุโลกพิทยาคม สำนักงานเขตพื้นที่การศึกษามัธยมศึกษาเขต39

[4] *สีสันหรือคุณภาพของเสียง*. [ม.ป.ป.]. ค้นเมื่อ 15 ตุลาคม 2558, จาก http://dynamic.psu.ac.th/musiclib.psu.ac.th/data/western-musuc/Chapter2/chap2-2.html

[5] กานดา รงค์สวัสดิ์, วัชระ ฉัตรวิริยะ.(2014).*การจำแนกน้ำเสียงภาษาไทยเชิงขอร้องและเชิงสั่งการโดยใช้โครงข่ายประสาทเทียม*. ค้นเมื่อ 25 กันยายน 2559, จาก http://202.44.34.144/nccitedoc/admin/nccit_files/NCCIT-20140511160812.pdf

[6] สุวรรณ รื่นสุคนธ์. (2552). *ประสิทธิภาพของ LPC และ LSP ในการรู้จำแนวทางเดนเสียงพูดภาษาไทยร่วมกับโครงข่ายประสาทเทียม*. การศึกษาหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต สาขาวิศวกรรมไฟฟ้า คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเทคโนโลยีราชมงคลธัญญบุรี

[7] *Linear prediction filter coefficients*. ค้นเมื่อ 25 กันยายน 2559, จาก https://www.mathworks.com/help/signal/ref/lpc.html

[8] Herve Abdi. (2010).*Normalizing Data*. ค้นเมื่อ 25 มีนาคม 2560, จาก https://www.utdallas.edu/~herve/abdi-Normalizing2010-pretty.pdf

[9] ไพลิน แสงเพ็ชร, สุพจน์ นิตย์สุวัฒน์. (2012).*ระบบระบุผู้พูดด้วยเสียงโดยใช้โครงข่ายประสาท-เทียม*.ภาควิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ. ภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์ประยุกต์ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ.

ระบบธุรกิจอัจฉริยะเพื่อวางแผนการขายสำหรับอุตสาหกรรมอสังหาริมทรัพย์
A Business Intelligence System for Selling Plan in Real Estate Industry

ดลฤดี ระพิสุวรรณ (Donruedee Rapisuwan)[1] และศักดิ์ชาย ตั้งวรรณวิทย์ (Sakchai Tangwannawit)[2]

ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ

มหาวิทยาลัยพระจอมเกล้าพระนครเหนือ

[1]s5807021857067@email.kmutnb.ac.th, [2]sakchai.t@it.kmutnb.ac.th

บทคัดย่อ

งานวิจัยเรื่องนี้มีวัตถุประสงค์ในการพัฒนาระบบธุรกิจอัจฉริยะเพื่อการวางแผนการขายสำหรับบริษัทอสังหาริมทรัพย์ โดยใช้ธุรกิจอัจฉริยะ (Business Intelligence) เพื่อช่วยในการจัดเก็บข้อมูลอย่างมีประสิทธิภาพ และการนำเสนอข้อมูลหลักในองค์กรได้อย่างถูกต้อง ครบถ้วน รวดเร็ว และทันต่อเหตุการณ์ ประกอบด้วย การสร้างคลังข้อมูลขนาดเล็ก (Data Mart) การใช้โครงสร้างแบบกาแลคซี่ (Constellation Schema) ที่เชื่อมโยงความสัมพันธ์ของข้อมูลขาย และก่อสร้าง และส่วนการแสดงผลรายงานที่สามารถปรับเปลี่ยนมุมมองในการวิเคราะห์ที่จะช่วยให้ผู้บริหารสามารถตัดสินใจได้เร็ว และมีประสิทธิภาพมากขึ้น ซึ่งผลการประเมินคุณภาพของระบบ จากกลุ่มตัวอย่าง 2 กลุ่ม คือ ผู้เชี่ยวชาญจำนวน 5 คน และผู้ใช้งานทั่วไปจำนวน 10 คน จากผลการประเมินคุณภาพของระบบโดยผู้เชี่ยวชาญได้ค่าเฉลี่ยเท่ากับ 3.87 และค่าเบี่ยงเบนมาตรฐานเท่ากับ 0.52 และผู้ใช้งานทั่วไปพบว่ามีค่าเฉลี่ยเท่ากับ 3.97 และค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.29 ซึ่งสรุปให้เห็นว่าการพัฒนาระบบธุรกิจอัจฉริยะเมื่อนำมาประยุกต์ใช้ในการวางแผนการขายสำหรับบริษัทอสังหาริมทรัพย์สามารถเพิ่มประสิทธิภาพในการทำงานได้ และในระดับดี

คำสำคัญ: ธุรกิจอัจฉริยะ คลังข้อมูลขนาดเล็ก การวางแผนการขาย

Abstract

The objective of this research was to a develop business intelligence system for selling plan in real estate company. That can help to efficient data storage and present the core data in organization accurately, completely and timely. Therefore, the researcher using the principles of data mart, use to constellation schema to link relationship between sales and construction data to create the intelligent report. This is achieved by using database management in the form of a view to analyze for management to make quick decision and more effective. The results of this study found that the developed system reported data of selling plan successfully. In addition, the evaluation of user acceptance from five experts and ten users, the average score of expert was 3.87 with standard deviation of 0.52 and the average score of expert was 3.97 with standard deviation of 0.29. In conclusion, the results showed that users were high satisfied with the system, proving the usability and usefulness.

Keywords: Business Intelligence, Data Mart, Selling Plan

1. บทนำ

ปัจจุบันหลายองค์กรมักประสบปัญหาในการบริหารจัดการข้อมูล ซึ่งปัจจัยที่สำคัญอย่างหนึ่งที่ช่วยในการสนับสนุนการตัดสินใจของผู้บริหารได้อย่างมีประสิทธิภาพ คือ ข้อมูลที่มีอยู่ในองค์กร ดังนั้นในยุคแห่งข้อมูลข่าวสารเช่นนี้ ธุรกิจต่างๆ จึงจำเป็นต้องหาเครื่องมือมาช่วยในการบริหารจัดการข้อมูลให้เปลี่ยนเป็นสารสนเทศที่มีคุณค่าต่อองค์กร [1] เพื่อใช้ประกอบการตัดสินใจ และวางแผนกลยุทธ์ด้านธุรกิจได้อย่างเหมาะสม และทันต่อเหตุการณ์

การพัฒนาระบบธุรกิจอัจฉริยะเพื่อวางแผนการขายสำหรับบริษัทอสังหาริมทรัพย์มีการออกแบบ และพัฒนาคลังข้อมูลขนาดเล็ก ในการจัดการฐานข้อมูลเพื่อให้ได้ข้อมูลที่จะนำไปสร้างความสัมพันธ์กับข้อมูลที่ต้องการ และมีมาตรฐานข้อมูล

เดียวกัน โดยนำเครื่องมือมาช่วยในการวิเคราะห์ข้อมูลในลักษณะการหมุนมิติ ของช่วงเวลา หรือลำดับชั้นของข้อมูล ซึ่งเป็นการประมวลผลเชิงวิเคราะห์แบบออนไลน์ (OLAP) ที่มีความสามารถในการค้นหา และวิเคราะห์ข้อมูลจากคลังข้อมูลขนาดเล็กได้อย่างรวดเร็ว รวมทั้งสามารถแสดงรายงานในรูปแบบ ตาราง หรือรายงานอัจฉริยะ ที่สามารถปรับเปลี่ยนมุมมองในการวิเคราะห์ได้ตามความต้องการ ผู้บริหารจึงมีความจำเป็นต้องมีเครื่องมือที่ช่วยในการวิเคราะห์ ตัดสินใจที่มีประสิทธิภาพ และนำบริษัทไปสู่ความสำเร็จ

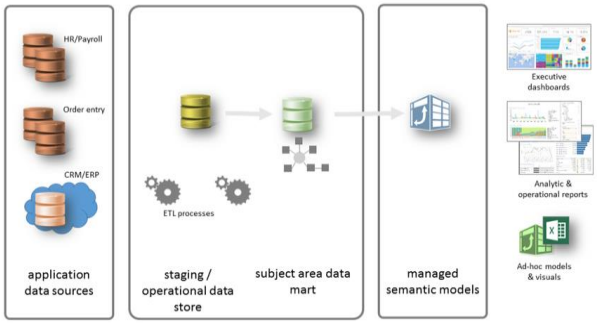
2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 ธุรกิจอัจฉริยะ (Business Intelligence)

ธุรกิจอัจฉริยะ คือ ระบบ หรือกลุ่มซอฟต์แวร์ที่ถูกพัฒนาขึ้นมาสำหรับเป็นเครื่องมือในการวิเคราะห์ข้อมูลที่ได้เก็บรวบรวมไว้ในคลังข้อมูล เพื่อสนับสนุนการวางแผนการตัดสินใจ และการบริหารงานของผู้บริหารผ่านการประมวลผลเชิงวิเคราะห์แบบออนไลน์ ทำให้องค์กรสามารถคาดการณ์พยากรณ์ความต้องการของผู้บริโภคได้อย่างถูกต้อง แม่นยำ ส่งผลให้ประสิทธิภาพการทำงานขององค์กรเพิ่มสูงขึ้น ทำให้สามารถนำไปใช้วิเคราะห์ และวางแผนการณ์ให้กับธุรกิจในด้านต่างๆได้โดยสะดวก เช่น การวิเคราะห์ยอดขายของสินค้าแต่ละชิ้นว่าสินค้าไหนมียอดขายเป็นอย่างไร เพื่อวางแผนการสั่งซื้อ

2.2 คลังข้อมูลขนาดเล็ก (Data Mart)

คลังข้อมูลขนาดเล็ก เป็นการรวมข้อมูลที่เกิดจากข้อมูลการปฏิบัติงาน และมาจากแหล่งข้อมูลอื่นๆที่ได้รับการออกแบบให้รองรับการใช้งานเฉพาะของผู้ใช้ ซึ่งส่วนใหญ่ข้อมูลในคลังข้อมูลขนาดเล็ก มักถูกออกแบบเป็นข้อมูลเฉพาะเจาะจงในแต่ละส่วน หรือแผนกในองค์กร ดังนั้นจึงนับได้ว่าตลาดข้อมูลเป็นส่วนย่อยของคลังข้อมูล และข้อมูลที่เก็บมักเป็นข้อมูลในระดับหน่วย หรือฝ่ายเท่านั้น [2] ดังภาพที่ 1



ภาพที่ 1: สถาปัตยกรรมของคลังข้อมูล

2.3 Dimensional Modeling

การออกแบบโครงสร้างในการจัดเก็บข้อมูลของคลังข้อมูล คือ การจัดรูปแบบของตารางเก็บข้อมูล โดยใช้วิธีดีนอร์มัลไลเซชั่น ซึ่งผลลัพธ์ที่เกิดขึ้นจะเป็นการสร้างฐานข้อมูลให้อยู่ในลักษณะของเชิงมิติ สำหรับองค์ประกอบของฐานข้อมูลหลายมิติมีการออกแบบโครงสร้างของฐานข้อมูลหลายมิติ [3] แบ่งออกเป็น 3 ประเภท คือ

2.3.1 Star Schema เป็นโครงสร้างแบบดาว มีตารางข้อเท็จจริงอยู่ตรงกลาง ส่วนภายนอกถูกล้อมรอบด้วยตารางมิติ วิธีนี้ข้อดีคือ ช่วยเพิ่มความสามารถในการค้นหาได้ง่ายและเร็ว เนื่องจากมีตารางจำนวนน้อยจึงทำให้ลดการตรวจสอบ

2.3.2 Snowflake Schema เป็นโครงสร้างแบบเกล็ดหิมะ คือมีโครงสร้างของตารางข้อเท็จจริงอยู่ตรงกลางเช่นเดียวกันกับ Star Schema แต่ตารางมิติจะมีการเชื่อมโยงไปยังตารางย่อย

2.3.3 Constellation Schema เป็นโครงสร้างแบบกาแลคซี่ จะมีตารางมิติอยู่กึ่งกลาง ซึ่งมักนิยมใช้เมื่อต้องการมุ่งเน้นในการวิเคราะห์ข้อมูลในแต่ละด้านเป็นพิเศษ เพราะตารางข้อเท็จจริงทั้งหมดจะถูกเชื่อมต่อกับตารางมิติ

2.4 งานวิจัยที่เกี่ยวข้อง

การพัฒนาระบบเพื่อวิเคราะห์ความผิดปกติของรายการทางธุรกิจ วัฒนพงศ์ โพธิ์พึ่ง [4]ได้นำเสนอการพัฒนาและออกแบบระบบในการวิเคราะห์ความผิดปกติด้านกระบวนการตรวจสอบความผิดปกติของแต่ละรายการบัญชีธนาคาร และในด้านการจัดการความเสี่ยง เพื่อติดตามพฤติกรรมการทำงานของพนักงานที่ทำรายการผิดปกติ ซึ่งในการนำเสนอข้อมูลการทำรายการบัญชีธนาคารของพนักงานนั้น ใช้รูปแบบโครงสร้างแบบดาว ใช้เครื่องมือ Business Objects BI

การพัฒนาระบบธุรกิจอัจฉริยะเพื่อการวางแผนการขายสินค้าสำหรับบริษัทผู้จัดจำหน่าย วิจิตรา พัชรกาจายกุล [5] ได้นำเสนอการพัฒนาระบบเพื่อสนับสนุนการตัดสินใจและการวางแผนด้านการขายของผู้บริหาร โดยใช้ระบบธุรกิจอัจฉริยะ โดยใช้กรณีศึกษาจากฐานข้อมูลการจำหน่ายสินค้าขนาดใหญ่ North Wind Relation Database ที่ใช้หลักการออกแบบโครงสร้างการเก็บข้อมูลโครงสร้างแบบดาวในการวัดประสิทธิภาพจำนวนสินค้าที่ขายได้ และผลรวมของยอดขายในรายเดือน หรือปี เพื่อการวางแผนการขายสินค้า

ธุรกิจอัจฉริยะสำหรับสนับสนุนการตัดสินใจในการบริหารการเบิกจ่ายวัสดุของธุรกิจเดินเรือทะเล ภัทรศรี วอนขอพร [6] งานวิจัยเพื่อนำเสนอการพัฒนาธุรกิจอัจฉริยะในการตรวจสอบและอนุมัติการเบิกจ่ายวัสดุอุปกรณ์ของบริษัท เพื่อออกรายงานสำหรับผู้บริหารในการเรียกดูสรุป และการออกรายงานสำหรับฝ่ายจัดซื้อ เป็นต้น นอกจากนี้ข้อมูลจากคลังข้อมูลได้นำมาใช้ในการวิเคราะห์หาความสัมพันธ์ในการเบิกวัสดุ เพื่อช่วยในการพิจารณาถึงวัสดุที่มักจะเบิกไปด้วยกัน เพื่อใช้แนะนำผู้เบิกเป็นการป้องกันการลืมเบิก และการพยากรณ์การเบิกวัสดุเพื่อช่วยคาดการณ์แนวโน้มของการเบิกวัสดุในอนาคต โดยออกแบบโครงสร้างแบบดาว

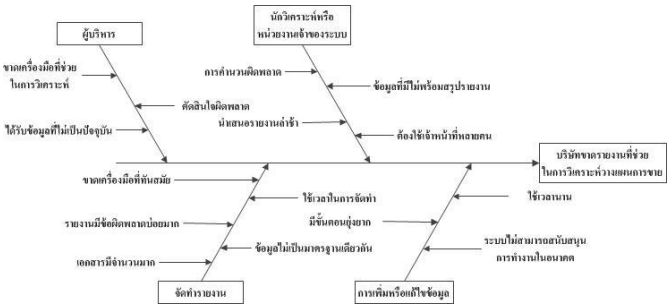
การบูรณาการข้อมูลเชิงพื้นที่กับระบบธุรกิจอัจฉริยะ เอกภพ บุญเพ็ง [7] งานวิจัยเพื่อนำเสนอการพัฒนาระบบคลังข้อมูลสำหรับการประเมินผลการจัดการศึกษาของประเทศโดยใช้ระบบธุรกิจอัจฉริยะ เป็นเครื่องมือในการพัฒนาระบบ โดยผู้บริหารต้องการวิเคราะห์ข้อมูลในหลายมิติ และในขณะเดียวกันก็ต้องการให้แสดงผลข้อมูลบนแผนที่ด้วย เพื่อให้เกิดความชัดเจนในการวิเคราะห์ และวางแผนการพัฒนาระบบต่อยอดจากระบบธุรกิจอัจฉริยะ โดยออกแบบโครงสร้างแบบกาแลคซี่ ซึ่งเป็นระบบที่บูรณาการข้อมูลเชิงพื้นที่กับระบบธุรกิจอัจฉริยะ เพื่อรวมความสามารถเข้าด้วยกัน เพื่อให้ผู้ใช้งานข้อมูลสามารถเรียกดูรายงานเชิงวิเคราะห์ในหลายมิติ และแสดงแผนที่ได้

3. วิธีดำเนินงาน

วิธีการดำเนินงานการพัฒนาระบบธุรกิจอัจฉริยะเพื่อการวางแผนการขายสำหรับบริษัทอสังหาริมทรัพย์ กำหนดวิธีการดำเนินงานตามวงจรการพัฒนาระบบ (SDLC) มี 5 ขั้นตอน

3.1 การศึกษา และรวบรวมข้อมูลระบบ

จากการศึกษาระบบงานเดิม ปัจจุบันใช้การเชื่อมโยงข้อมูลผ่านโปรแกรม Microsoft Excel โดยมีการรวบรวมข้อมูลมากมายที่มาจากหลายระบบจัดทำรายงานนำเสนอผู้บริหาร ส่งผลให้นักวิเคราะห์ต้องเร่งจัดทำรายงาน ซึ่งบางครั้งข้อมูลที่มีอยู่อาจจะยังไม่พร้อมใช้สรุปเป็นรายงานเสนอต่อผู้บริหาร ทำให้ใช้เวลานาน ใช้พนักงานหลายคน และด้วยความเร่งรีบอาจทำให้เกิดการคำนวณตัวเลขผิดพลาดได้ ทำให้การวิเคราะห์และตัดสินใจวางแผนการขายนั้นทำได้ไม่สะดวก และไม่รวดเร็วเท่าที่ควร ทั้งนี้ได้แสดงดังแผนภูมิเหตุ และผลดังภาพที่ 2



ภาพที่ 2: แผนผังสาเหตุและผล

จากประเด็นปัญหาที่ต้องการรายงานในการช่วยตัดสินใจสรุปได้ดังต่อไปนี้

3.1.1 นักวิเคราะห์ขาดเครื่องมือที่สามารถรวบรวมและค้นหาประวัติข้อมูลยอดขาย ยอดโอน และการก่อสร้าง ตามแต่ละช่วงเวลา

3.1.2 นักวิเคราะห์ไม่สามารถสร้างรายงานที่ใช้เสนอต่อผู้บริหารตามรูปแบบที่ได้รับมอบหมายได้

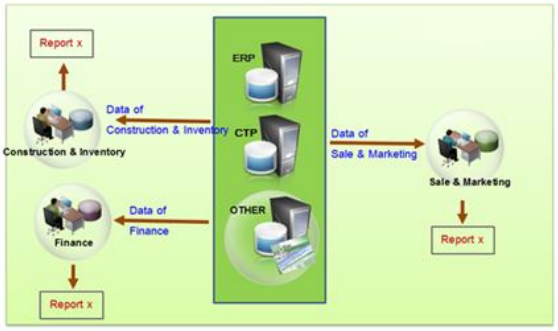
3.1.3 ผู้บริหารขาดเครื่องมือที่ช่วยในการวิเคราะห์ยอดขาย ยอดโอน และก่อสร้าง เพื่อมาวางแผนการขาย

3.1.4 ผู้บริหารไม่สามารถสรุปสถานการณ์การดำเนินงานจริงได้บางครั้งมีความคลาดเคลื่อนของแกนเวลาและตัวเลข

3.2 การวิเคราะห์ และออกแบบระบบ

จากการที่ผู้วิจัยศึกษา และวิเคราะห์การทำงานของระบบเดิม มีโครงสร้างดังภาพที่ 3 จึงได้ออกแบบความต้องการของระบบใหม่สำหรับวิเคราะห์เพื่อวางแผนการขาย โดยรวบรวมข้อมูลจากระบบงานขาย และระบบงานก่อสร้าง รวมถึงแผนการขายรายปี โดยออกแบบการทำงานของระบบออกเป็น 3 ส่วน คือ ส่วนการวิเคราะห์ข้อมูล ส่วนการออกรายงานและ

รูปแบบการนำเสนอ และส่วนของการเตรียมข้อมูลเข้าสู่ระบบ โดยมีขอบเขตของข้อมูลที่มาจากระบบต่างๆ ดังภาพที่ 4



ภาพที่ 3: โครงสร้างระบบงานเดิม



ภาพที่ 4: ขอบเขตของข้อมูลที่ใช้สำหรับวางแผนการขาย

ในส่วนของการออกรายงานการนำเสนอสามารถแบ่งกลุ่มรายงานจากปัญหาและความต้องการออกเป็น 3 กลุ่ม คือ

3.2.1 กลุ่มรายงาน Sale and Transfer โดยมีรายงานสรุปยอดขาย และยอดโอน เป็นการแสดงข้อมูลยอดขาย และยอดโอนตามช่วงเวลา รวมทั้งรายงานสรุปรายได้จริงของบริษัทตามสายงาน และโครงการ

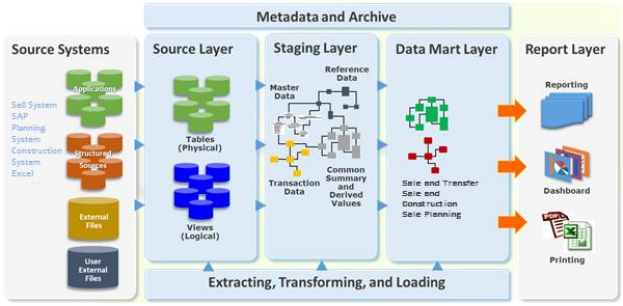
3.2.2 กลุ่มรายงาน Sale and Construction โดยมีรายงานบ้านว่างพร้อมขาย ที่สรุปข้อมูลบ้านในแต่ละสถานะการก่อสร้างตามสายงานและโครงการที่สามารถขาย หรือโอนได้

3.2.3 กลุ่มรายงาน Sale Analysis โดยมีรายงานแสดงยอดขาย และยอดโอนเปรียบเทียบแผนเป็นการวัดประสิทธิภาพด้านยอดขาย และยอดโอนตามช่วงเวลา

3.3 การพัฒนาระบบ

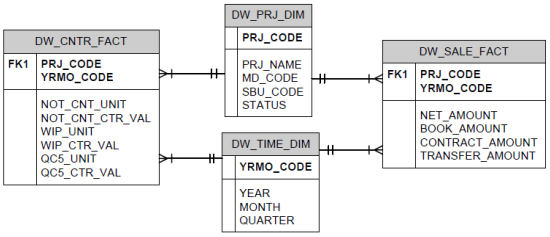
ผู้วิจัยจัดทำคลังข้อมูลสำหรับการวิเคราะห์เพื่อวางแผนการขาย ซึ่งในการพัฒนาจะใช้ SAP Business Intelligence เป็นเครื่องมือสำคัญในการพัฒนา โดยมีกระบวนการดังนี้

กระบวนการอีทีแอลขายและก่อสร้าง (ETL Sale and Construction หรือ SC-ETL) ที่ผู้วิจัยพัฒนาขึ้นนั้นใช้เครื่องมือ Business Object Data Services นำข้อมูลจากแหล่งข้อมูลที่ต้องการทำการแปลงเข้าสู่คลังข้อมูลขนาดเล็ก ดังภาพที่ 5



ภาพที่ 5: ขั้นตอนการทำอีทีแอลข้อมูลขายและก่อสร้าง (SC-ETL)

กระบวนการออกแบบคลังข้อมูลขนาดเล็กโดยใช้โครงสร้างแบบกาแลคซี่ ด้วยเครื่องมือ Universe Design Tools หรือ Information Design Tools ดังภาพที่ 6



ภาพที่ 6: คลังข้อมูลโครงสร้างแบบกาแลคซี่

จากรูปที่ 6 คลังข้อมูลที่ใช้โครงสร้างแบบกาแลคซี่ประกอบด้วยตารางข้อเท็จจริง 2 ตาราง คือ ตารางที่เก็บข้อมูลการขาย การโอน เช่น มูลค่าขายไม่หักส่วนลด มูลค่าขายหักส่วนลดแล้ว มูลค่าที่มีการยกเลิกการขาย และตารางข้อมูลการขายและก่อสร้าง เช่น มูลค่าบ้านที่ผ่านการตรวจสอบคุณภาพ มูลค่าบ้านที่อยู่ระหว่างการก่อสร้าง โดยตารางข้อเท็จจริงทั้ง 2 ตารางมีการใช้ตารางมิติเวลา และตารางมิติโครงการร่วมกัน ซึ่งโครงสร้างการเชื่อมโยงที่ซับซ้อนมีตารางข้อเท็จจริงมากกว่า 1 ตาราง โครงสร้างนี้คล้ายแบบโครงสร้างแบบดาวมาประกอบกัน ข้อดีจึงมีความยืดหยุ่นสูง เนื่องจากมีการใช้ตารางมิติร่วมกัน และสามารถแยกตารางมิติจากกันได้

กระบวนการออกแบบรายงานด้วยเครื่องมือ Web Business Intelligence เป็นเครื่องมือสำหรับทำรายงานจาก

ปัญหา และความต้องการจึงได้พัฒนารายงานเพื่อช่วยในการวิเคราะห์เพื่อวางแผนการขาย โดยแบ่งกลุ่มรายงาน ดังภาพที่ 7



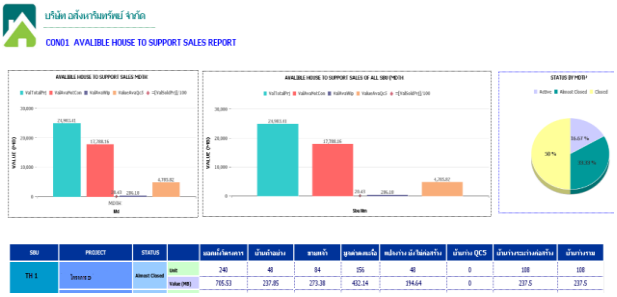
ภาพที่ 7: กลุ่มรายงานสำหรับวางแผนการขาย

กลุ่มรายงาน Sale and Transfer แสดงรายงานเกี่ยวกับยอดขายที่จะเปลี่ยนเป็นรายได้จริง แสดงให้เห็นถึงรายได้สุทธิตามสายงาน โครงการ และตามช่วงเวลา เพื่อดูรายได้ที่เกิดขึ้นตามฤดูกาลแบบรายปี รวมทั้งยอดรายได้รวมของแต่ละปีย้อนหลังที่สามารถวิเคราะห์แนวโน้มการเติบโตของธุรกิจบริษัท ซึ่งข้อมูลที่ได้สามารถนำมาการวิเคราะห์เพื่อสร้าง และปรับปรุงกลยุทธ์ เช่น การเร่งอัตราการขาย การโอนในโครงการที่ขายดีได้ โดยรายงานมีทั้งแผนภูมิเพื่อให้เห็นภาพรวม และตารางแสดงรายละเอียด ดังภาพที่ 8



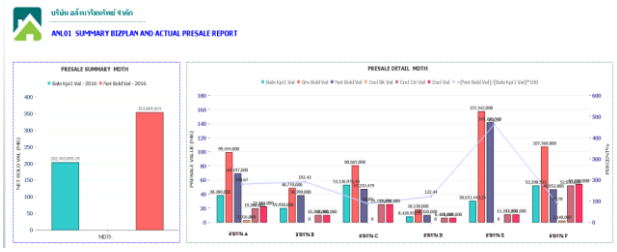
ภาพที่ 8: ตัวอย่างรายงานสรุปยอดขายตามไตรมาส

กลุ่มรายงาน Sale and Construction เนื่องจากการก่อสร้างใช้เวลาค่อนข้างนานซึ่งแตกต่างจากการซื้อสินค้ามาวางขายทั่วไป จึงจำเป็นต้องมีการวางแผนที่จะต้องสอดคล้องกันทั้งขายและก่อสร้าง โดยรายงานที่ได้สามารถนำไปประกอบการวิเคราะห์เพื่อสร้าง และปรับปรุงกลยุทธ์ เช่น การชะลอการก่อสร้างเพื่อให้เกิดบ้านค้างสต็อกน้อยลง หรือ การเร่งยอดขายสำหรับบ้านค้างสต็อก ดังภาพที่ 9



ภาพที่ 9: ตัวอย่างรายงานบ้านว่างเพื่อรองรับการขาย

กลุ่มรายงาน Sale Analysis ช่วยในการวิเคราะห์วางแผนการขายให้เป็นไปตามเป้าทั้งรายเดือน และรายปี โดยนำข้อมูลของการขาย การโอน และการก่อสร้างมาประกอบการตัดสินใจ เช่น ยอดขายต่ำกว่าแผน เร่งยอดขายสำหรับบ้านค้างสต็อก ดังภาพที่ 10



ภาพที่ 10: ตัวอย่างรายงานแสดงยอดขายเปรียบเทียบแผน

3.4 การทดสอบระบบ

ได้ทำการทดสอบโดยนำหลักกระบวนการทดสอบแบบแบล็คบอกซ์ [3] โดยทดสอบความสมบูรณ์ของระบบที่จัดโดยกลุ่มผู้เชี่ยวชาญทางด้านระบบสารสนเทศจำนวน 5 คน และการทดสอบระบบเป็นการทดสอบความสมบูรณ์ของระบบโดยผู้ใช้งานทั่วไปจำนวน 10 คน

ตารางที่ 1: สรุปผลการประเมินความพึงพอใจของผู้ใช้งาน

หัวข้อการประเมินคุณภาพ	Mean	S.D	ระดับคุณภาพ
1. ด้านความสามารถของระบบ	4.11	0.45	ดี
2. ด้านความถูกต้องของระบบ	4.00	0.23	ดี
3. ด้านการใช้งานของระบบ	4.04	0.24	ดี
4. ด้านประสิทธิภาพของระบบ	3.73	0.35	ดี
รวม	3.97	0.29	ดี

3.5 การประเมินระบบ

ผลการประเมินคุณภาพของระบบได้ทำการทดสอบการใช้งานจริงจากแบบประเมิน และแบบสอบถามของกลุ่มผู้เชี่ยวชาญจำนวน 5 คน และผู้ใช้งานทั่วไปจำนวน 10 คน โดยแบ่งการประเมินออกเป็น 4 ด้าน คือ ด้านความสามารถของระบบตรงตามความต้องการ เช่น ความถูกต้องข้อมูลโดยรวมในการรายงานจากข้อมูลที่นำมาวิเคราะห์ระบบ มีความเหมาะสมของการแสดงผลข้อมูลตามที่ต้องการของบริษัท ด้านความถูกต้องของระบบ เช่น ความถูกต้องในการรายงานข้อมูลโดยรวม จากข้อมูลที่นำมาวิเคราะห์ ด้านการใช้งานของระบบ เช่น ความง่ายในการใช้งาน และด้านประสิทธิภาพของระบบ เช่น ความเร็วในการประมวลผลการวิเคราะห์

ผลการประเมินคุณภาพของระบบในทุกด้านดังที่กล่าวมาข้างต้นสรุปได้ว่าจากการประเมินความพึงพอใจโดยผู้เชี่ยวชาญระบบได้ค่าเฉลี่ยเท่ากับ 3.87 และค่าเบี่ยงเบนมาตรฐานเท่ากับ 0.52 และผู้ใช้งานทั่วไปพบว่ามีค่าเฉลี่ยเท่ากับ 3.97 และค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.29 ซึ่งแสดงว่าผู้ใช้งานมีการยอมรับและมีความพึงพอใจในการใช้งานในทุกด้านอยู่ในระดับดี

4. ประโยชน์ของการวิจัย

1. ประโยชน์โดยตรง : ได้ระบบธุรกิจอัจฉริยะสำหรับการวิเคราะห์วางแผนการขาย และช่วยลดเวลาในการจัดทำรายงานเพื่อการวิเคราะห์ข้อมูล

2. ประโยชน์จากการต่อยอด : ได้คลังข้อมูลเพื่อใช้งานในการวิเคราะห์ธุรกิจด้านการขายและการก่อสร้าง ที่สามารถนำเทคโนโลยีมาใช้เพิ่มเติมในการสร้างกฎ และการพยากรณ์ข้อมูลที่จะเกิดขึ้นในอนาคตได้ เช่น การพยากรณ์ยอดขายอีก 5 ปีข้างหน้า เพื่อนำมาวางแผนการก่อสร้าง

5. สรุปผล

งานวิจัยนี้ได้พัฒนาธุรกิจอัจฉริยะเพื่อการวางแผนการขายสำหรับบริษัทอสังหาริมทรัพย์ ในการนำข้อมูลมาวิเคราะห์เพื่อวางแผนการขายให้เป็นไปตามเป้าหมาย ซึ่งประกอบด้วยขั้นตอนการสร้างคลังข้อมูลด้วยเครื่องมือ Business Object Data Services สร้างลูกบาศก์ด้วย Universe Design Tools จากนั้นจึงสร้างรายงานด้วย Web Business Intelligence เพื่อแก้ไขปัญหาในประเด็นต่างๆ กล่าวคือ การออกรายงานสำหรับผู้บริหารดูภาพรวมรายได้แนวโน้มการเติบโตของบริษัท การออกรายงานสำหรับนักวิเคราะห์ที่สามารถนำข้อมูลรายละเอียดไปวางแผน และจัดทำรายงานเสนอที่ประชุม จุดเด่นของงานวิจัยนี้อยู่ที่ระบบสามารถนำข้อมูลที่มีอยู่มาจัดทำรายงานในรูปแบบต่างๆ ที่เหมาะสมกับมุมมองในการวิเคราะห์ แสดงความสัมพันธ์ ที่เชื่อมโยงกันทั้งด้านการขาย สร้าง และโอน โดยใช้หลักการออกแบบโครงสร้างแบบกาแลคซี่เกิดความยืดหยุ่นในการแสดงผลลัพธ์ของการขาย และก่อสร้างที่มีการใช้มิติมุมมองร่วมกัน ทำให้ข้อมูลที่ได้มีความถูกต้อง และตรงตามความต้องการ ก่อให้เกิดประโยชน์ในการวางแผนกลยุทธ์ด้านการขายของธุรกิจอสังหาริมทรัพย์ได้อย่างมีประสิทธิภาพ

เอกสารอ้างอิง

[1] สมเกียรติ ลีลาทวีวุฒิ จักรกฤษณ์ ดวงพัสตรา และวรภัทร ไพรีเกรง, “แบบจำลองข้อมูลของระบบธุรกิจอัจฉริยะประชาคมเศรษฐกิจอาเซียน” หลักสูตรบริหารธุรกิจดุษฎีบัณฑิต มหาวิทยาลัยธุรกิจบัณฑิตย์, 2559.

[2] Manya Sethi, “Data Warehouse and OLAP Technology” MCA Final Year Amity University,Uttar Pradesh, 2012.

[3] Mandeep Kaur Sandhu, Amanjot Kaur and Ramandeep Kaur. “Data Warehouse Schemas” Computer Science Deptt, SDSPM College for Women, Rayya, 2015.

[4] วัฒนพงศ์ โพธิ์พึ่ง, “การพัฒนาระบบเพื่อวิเคราะห์ความผิดปกติของรายการธุรกิจ กรณีศึกษา: รายการทางบัญชีธนาคาร” สารนิพนธ์วิทยาศาสตรมหาบัณฑิต ภาควิชาเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2550.

[5] วิจิตรา พัชรกาจายกุล และนิเวศ จิระวิชิตชัย, “การพัฒนาระบบธุรกิจอัจฉริยะเพื่อการวางแผนการขายสินค้าสำหรับบริษัทผู้จัดจำหน่าย” สารนิพนธ์วิทยาศาสตรมหาบัณฑิต คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยศรีปทุม, 2555.

[6] ภัทรศรี วอนขอพร และสมชาย เล็กเจริญ, “ธุรกิจอัจฉริยะสำหรับสนับสนุนการตัดสินใจในการบริหารการเบิกจ่ายวัสดุของธุรกิจเดินเรือทะเล” สารนิพนธ์วิทยาศาสตรมหาบัณฑิต คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยรังสิต, 2555.

[7] เอกภพ บุญเพ็ง อนุสรณ์ เบญจธนรัตน์ สุรางคนา ธรรมลิขิต และพีระศักดิ์ เพียรประสิทธิ์, “การบูรณาการข้อมูลเชิงพื้นที่กับระบบธุรกิจอัจฉริยะ” สำนักคอมพิวเตอร์ มหาวิทยาลัยบูรพา, 2553.

[8] Emany Sidi, Mohamed El Merouani and El Amin A. Abdelouarit, “Star Schema Advantages on Data Warehouse: Using Bitmap Index and Partitioned Fact Tables” Laboratory modeling and information theory Abdelmalek Essaadi University, Tétouan, Morocco, 2016.

การสร้างเอกสารที่ถูกตัดกลับคืนโดยใช้การทำนายเชิงเส้น
Shredded Document Reconstruction using Linear Prediction

ธนัสนี เพียรตระกูล (Tanasanee Phienthrakul)
ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยมหิดล
tanasanee.phi@mahidol.ac.th

บทคัดย่อ

เอกสารที่เป็นกระดาษเป็นหลักฐานแบบหนึ่ง ซึ่งเมื่อถูกทำลาย การต่อเอกสารกลับคืนมาอาจทำให้ได้หลักฐานเพื่อช่วยในการสืบสวนได้ บทความนี้นำเสนอวิธีการในการต่อเอกสารที่ถูกทำลายโดยการตัดกลับคืนมา โดยการทำนายจุดแบบเชิงเส้นมาช่วยในการเลือกชิ้นส่วนที่เหมาะสมสำหรับการต่อ จุดสีดำที่อยู่ในชิ้นส่วนเดิมจะถูกนำมาใช้ในการทำนายค่าตำแหน่งของจุดสีดำในชิ้นถัดไป ดังนั้นบริเวณขอบของชิ้นส่วนถัดไปจะถูกให้คะแนนตามค่าผลรวมของการทำนายที่เกิดขึ้น คะแนนเหล่านี้จะใช้ในการตัดสินใจเลือกชิ้นส่วนที่เหมาะสม ผลการทดสอบแสดงให้เห็นว่าวิธีการที่นำเสนอนี้ให้ผลการต่อเอกสารกลับคืนที่ดีกว่าการพิจารณาแค่ขอบเพียงอย่างเดียว

คำสำคัญ: การสร้างเอกสารกลับคืน เอกสารที่ถูกตัด การทำนายเชิงเส้น การรวมเชิงเส้น

Abstract

Document paper is a kind of evidence and if it was destroyed, the document reconstruction can assist in the investigation. This paper presents a shredded document reconstruction method using linear prediction as a guideline for selecting the appropriated piece of paper. Black pixels in a piece of paper are used for predicting the next pixels in the next piece. Thus, edge of another piece will be scored by the total of predicting results. These scores are used as a part of decision making to select a suitable part. The experimental results show that this proposed method yields the document reconstructing results that are better than edge considering method.

Keyword: document reconstruction, shredded document, linear prediction, linear combination.

1. บทนำ

แม้ว่าปัจจุบันเอกสารต่างๆจะถูกเก็บไว้ในรูปแบบไฟล์ดิจิตอลมากขึ้น แต่ก็ไม่ได้ลดความสำคัญของเอกสารที่อยู่ในรูปแบบกระดาษลง เนื่องจากเอกสารเหล่านี้นั้นยังคงถูกใช้ในสำนักงานและบริษัทต่างๆเพื่อเป็นหลักฐานของการตกลงแบบตายตัวหรือสามารถตรวจสอบการแก้ไขได้ ต่างกับไฟล์ข้อมูลในคอมพิวเตอร์ที่การเปลี่ยนแปลงเนื้อหาในไฟล์ยากสำหรับบุคคลทั่วไปในการตรวจสอบความเปลี่ยนแปลง ซึ่งทำให้ยังคงมีการใช้เอกสารในรูปแบบกระดาษอยู่ในปัจจุบัน

เอกสารที่เป็นกระดาษนั้นยังถือเป็นหนึ่งในหลักฐานสำคัญในคดีต่างๆ เช่น คดีทุจริต หรือ คดีฉ้อโกง โดยแตกต่างจากไฟล์เอกสารดิจิทัลที่สามารถกำจัดทิ้งโดยง่ายเพียงแค่ลบไฟล์เอกสารนั้นจากคอมพิวเตอร์เท่านั้น แต่สำหรับเอกสารในรูปแบบกระดาษหนึ่งในวิธีการกำจัดเอกสารที่ได้รับความนิยมคือการใส่เอกสารผ่านเครื่องทำลายเอกสาร ทำให้เอกสารกลายเป็นเศษกระดาษฝอย ดังแสดงในภาพที่ 1



ภาพที่ 1: กระดาษฝอยจากเครื่องทำลายเอกสาร [1]

ในการสืบสวนสอบสวน เมื่อพบเอกสารที่ถูกทำลายด้วยวิธีนี้ การที่ผู้สืบสวนจะสามารถอ่านเนื้อหาในเศษกระดาษฝอยได้จะต้องมีการสร้างเอกสารขึ้นมาใหม่โดยการต่อเศษฝอยของเอกสาร แต่การต่อเศษเอกสารนั้นทำได้ยาก เนื่องจากเศษเอกสารที่ผ่านเครื่องทำลายเอกสารนั้นมีรูปแบบขอบที่ตรงเหมือนกันทั้งหมด จึงทำให้เหลือแค่ข้อมูลในแต่ละเศษที่ผู้สืบสวนสามารถใช้ในการต่อเศษกระดาษได้

เครื่องทำลายเอกสารตามสำนักงานนั้นทำลายเอกสารโดยการตัดแนวตั้งให้เศษแต่ละชิ้นมีขนาดความกว้าง 20 มิลลิเมตร และการทำลายเอกสารโดยใส่เอกสารในแนวตั้งนั้นมีผลลัพธ์เป็นเศษฝอยของเอกสารเป็นจำนวนประมาณ 50 ชิ้น นอกจากนั้นในการทำลายเอกสารแต่ละครั้ง มิได้ทำลายทีละแผ่นแต่จะทำลายหลายๆแผ่นต่อเนื่องกันซึ่งทำให้มีชิ้นส่วนที่จะต้องนำมาใช้ในการต่อเป็นจำนวนมาก

การต่อด้วยมือจะมีความแม่นยำที่สูงที่สุด แต่ใช้ทรัพยากรบุคคลและเวลาในการต่อเอกสารสิ้นเปลืองมาก ดังนั้นหากมีเครื่องมือที่ช่วยจัดเรียงชิ้นส่วนเอกสารเพื่อประกอบกันขึ้นมาเป็นเอกสารใหม่ก็จะช่วยลดเวลาและทำให้การสืบสวนสอบสวนสามารถทำได้สะดวกมากขึ้น ดังนั้นวัตถุประสงค์ของงานนี้จึงต้องการพัฒนาซอฟต์แวร์เพื่อลดเวลาในการสร้างเอกสารใหม่จากเศษเอกสาร และเพิ่มความแม่นยำในการสร้างเอกสารใหม่จากเศษเอกสารโดยพิจารณาจากเนื้อหาในเศษเอกสาร

2. งานวิจัยที่เกี่ยวข้อง

ในปี พ.ศ. 2554 สำนักงานโครงการวิจัยขั้นสูงของกระทรวงกลาโหมสหรัฐอเมริกา หรือ DARPA [2] ได้มีการจัดการแข่งขันการสร้างเอกสารใหม่จากเศษเอกสารขึ้นโดยมีเงินรางวัลสำหรับกลุ่มผู้เข้าแข่งขันที่สามารถตอบคำถามที่ทาง DARPA ตั้งไว้สำหรับปริศนาเศษกระเอกสารแต่ละชุด โดยข้อมูลที่ใช้สำหรับตอบคำถามนั้นจะอยู่ในข้อความที่ถูกเขียนลงไปยังแผ่นเอกสารแล้วนำใส่ผ่านเครื่องทำลายเอกสารเพื่อเศษเอกสารที่ได้มาเป็นผลลัพธ์ ซึ่งในการสร้างใหม่นั้นไม่จำเป็นต้องสร้างเอกสารใหม่ทั้งหมด หากแต่สร้างใหม่เท่าที่จำเป็นที่จะสามารถอ่านข้อมูลในเศษเอกสารชุดนั้นๆมาตอบคำถามที่ถูกตั้งไว้

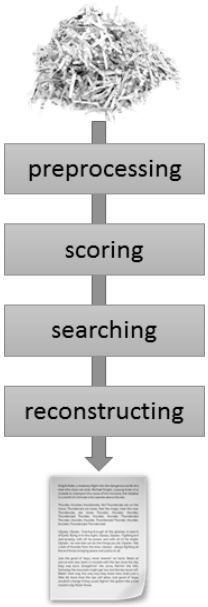
กลุ่มผู้เข้าแข่งขันที่อยู่ใน 10 อันดับแรกนั้นได้ผสมวิธีการจับคู่เศษแบบอัตโนมัติเข้ากับการจับคู่ด้วยมือ [3] เพื่อได้วิธีการสร้างใหม่ของเอกสารที่มีประสิทธิภาพสูงสุด วิธีการแบบผสมหรือแบบกึ่งอัตโนมัตินั้นได้จับคู่เศษเอกสารเบื้องต้นก่อนที่จะส่งต่อให้ผู้ใช้ซึ่งการจับคู่ในระดับต้นนั้นส่งผลให้จำนวนคู่เศษเอกสารลดลงอย่างมากเพื่อเป็นการแบ่งเบาภาระการจับคู่ระดับสูงที่จำเป็นต้องให้ผู้ใช้ในการเข้ามาตัดสินใจ ซึ่งด้วยวิธีกึ่งอัตโนมัตินี้ หน้าที่ของคอมพิวเตอร์นั้นได้แก่การปรับแต่งภาพ (pre-processing) และการหาคุณลักษณะและจับคู่ (feature extraction and matching) และหน้าที่ของผู้ใช้คือการจับคู่ระดับสูงโดยตัดสินจากการมองภาพรวมของการสร้างใหม่จากเศษเอกสาร

ในส่วนของวิธีการที่ใช้ในการจับคู่และจัดเรียงเศษเอกสารนั้น มีทำการเปรียบเทียบพิกเซลต่อพิกเซลโดยตรง (direct pixel-based alignment) [4] โดยจะนำพิกเซลด้านขวาของชิ้นส่วนแรก มาเปรียบเทียบกับพิกเซลด้านซ้ายของอีกเศษชิ้นส่วน ซึ่งจะทำการเปรียบเทียบเพื่อหาคู่ของชิ้นส่วนที่มีค่าความต่างของคู่พิกเซลรวมน้อยที่สุด ในส่วนของค่าความต่างของคู่พิกเซลสามารถคำนวนได้จากจำนวนพิกเซลที่ต้องข้ามก่อนที่จะเจอพิกเซลที่มีคุณลักษณะตรงกัน เช่น มีสีเดียวกัน

ในงานวิจัยของ Huei-Yung Lin และ Wen-Cheng Fan-Chiang [5] ได้นำเสนอการต่อเอกสารโดยใช้เทคนิคด้านการประมวลผลภาพ เพื่อเลือกชิ้นส่วนที่สอดคล้องกัน นอกจากนั้นยังใช้เทคนิคทางกราฟเพื่อช่วยในการจัดเรียง R. Ranca [6], [7] เสนอกรอบงานสำหรับการต่อเอกสารอัตโนมัติ โดยแบ่งการทำงานเป็น 3 ขั้นตอน คือเตรียมภาพ คำนวณคะแนน และค้นหา ซึ่งเป็นกรอบการทำงานที่สามารถนำมาประยุกต์ใช้ในงานนี้ได้เป็นอย่างดี

3. การสร้างเอกสารที่ถูกตัดกลับคืน

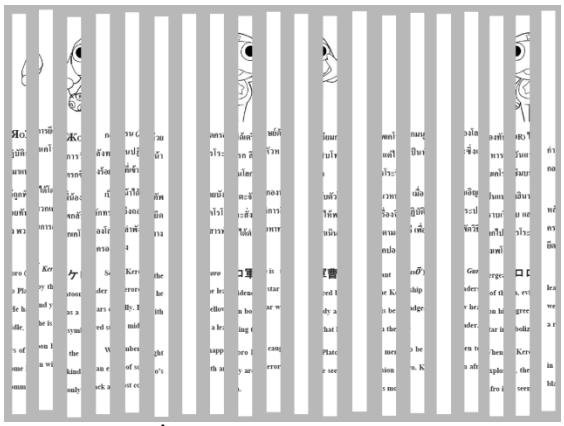
การพัฒนาโปรแกรมเพื่อสร้างเอกสารที่ถูกตัดกลับคืนมาใหม่จากเศษเอกสารนั้น จะทำการแบ่งขั้นตอนการทำงานออกเป็น 4 ขั้นตอน ได้แก่ การปรับแต่งภาพเบื้องต้น (pre-processing) การให้คะแนน (scoring) การค้นหาชิ้นส่วน (searching) และ การสร้างเอกสารกลับคืน (reconstructing) ดังแสดงในภาพที่ 2



ภาพที่ 2: ขั้นตอนการสร้างเอกสารที่ถูกตัดกลับคืน

3.1 การปรับแต่งภาพเบื้องต้น

ในส่วนของการปรับแต่งภาพเบื้องต้นนั้นจะทำการปรับปรุงภาพ ลดสัญญาณรบกวน (noise) ในภาพ และตัดภาพเอกสารแยกเป็นภาพเศษเอกสารในแต่ละชิ้น เพื่อเตรียมความพร้อมก่อนการนำแต่ละชิ้นไปคำนวณในขั้นตอนถัดไป ชิ้นส่วนเหล่านี้จะถูกแปลงให้เป็นภาพขาวดำ ภาพตัวอย่างของชิ้นส่วนเอกสารที่ถูกตัด แสดงในภาพที่ 3

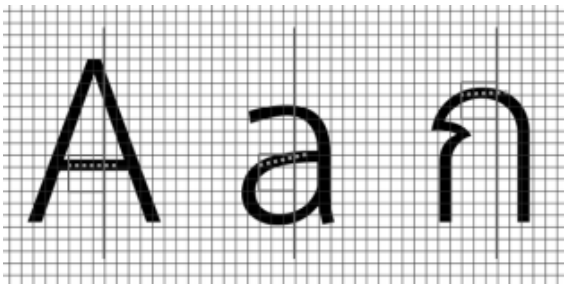


ภาพที่ 3: ตัวอย่างชิ้นส่วนของเอกสารที่ถูกตัด

3.2 การให้คะแนน

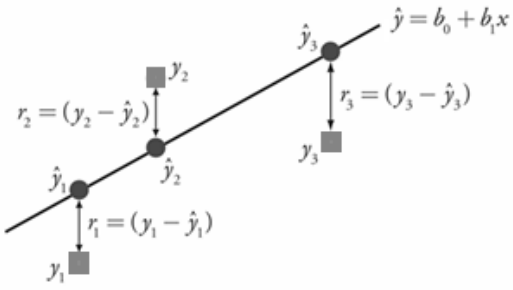
ขั้นตอนถัดมาคือส่วนของการให้คะแนนในแต่ละชิ้น ในกรณีของการต่อเอกสารที่โดนตัดในลักษณะแนวตั้งนี้ จะมีบางชิ้นส่วนที่ไม่มีข้อมูลใดๆอยู่เลย เราจะถือว่าชิ้นส่วนเหล่านั้นเป็นส่วนของขอบ และจะไม่นำมาพิจารณาในขั้นตอนของการต่อเอกสาร ในส่วนของฟังก์ชันการให้คะแนนนั้นจะทำการหาคุณลักษณะของภาพเศษเอกสารเพื่อทำการจับคู่กับเศษเอกสารชิ้นอื่น แล้วประมวลผลลัพธ์เป็นคะแนนความต่อกันได้สนิทของคู่เศษกระดาษนั้นๆ โดยคะแนนนี้จะคำนวณจากขอบทางขวาของชิ้นก่อนหน้า เพื่อหาขอบทางซ้ายของชิ้นถัดไปที่ต่อกันได้เหมาะสมที่สุด ในกรณีของการใช้การทำนายเชิงเส้นที่นำเสนอในบทความนี้ จะเริ่มจากการพิจารณาชิ้นส่วนด้านซ้ายเป็นตารางย่อย (grid) ดังแสดงในภาพที่ 4

จากตารางย่อยพิจารณาค่าพิกเซลของบริเวณที่มีข้อมูล จากนั้นหาแนวโน้มของเส้นว่าควรมีทิศทางขึ้นหรือลงหรือคงที่ โดยดูจากความชันของเส้น ในขั้นตอนนี้เราจะใช้วิธีการถดถอยเชิงเส้น (linear regression) เพื่อหาสมการเส้นตรงที่สอดคล้องกับเส้นภายในภาพชิ้นส่วน



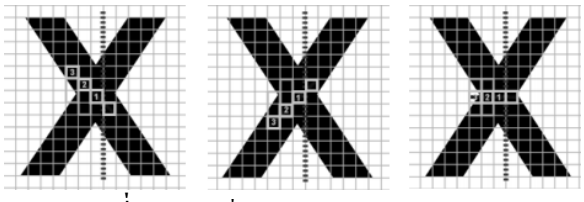
ภาพที่ 4: แบ่งภาพตัวอักษรเป็นตารางย่อย

วิธีการถดถอยเชิงเส้นจะหาสมการเส้นตรงที่สอดคล้องกับข้อมูลโดยให้มีค่าคลาดเคลื่อนน้อยที่สุด ตำแหน่งของพิกเซลสีดำในแนวแกนตั้ง ถูกนำมาใช้ในการคำนวณ เพื่อหาสมการเส้นตรงที่จะนำไปใช้ทำนายตำแหน่งถัดไปของพิกเซลสีดำในชิ้นที่จะนำมาต่อ ภาพที่ 5 แสดงตัวอย่างสมการแส้นตรงที่ประมาณจากการพิจารณาข้อมูลโดยใช้วิธีการถดถอยเชิงเส้น



ภาพที่ 5: วิธีการถดถอยเชิงเส้น

แต่ละพิกเซลที่บริเวณขอบทางขวาของชิ้นส่วนทางซ้าย จะถูกคำนวณหาสมการถดถอยเชิงเส้นในหลายๆทิศทาง เพื่อพยากรณ์ตำแหน่งถัดไป ดังแสดงในภาพที่ 6 ดังนั้นในหนึ่งพิกเซลของชิ้นทางซ้ายอาจต่อเนื่องไปยังชิ้นถัดไปได้ในหลายตำแหน่ง



ภาพที่ 6: ทิศทางที่เป็นไปได้ในการใช้การทำนายเชิงเส้น

ตำแหน่งของพิกเซลถัดไปที่เป็นไปได้จะถูกรวบรวมไว้ทั้งหมด เพื่อใช้บอกว่าที่ตำแหน่งไหนเป็นไปได้ที่จะมีพิกเซลสีดำ ส่วนที่เหลือคือตำแหน่งที่ไม่ควรจะมีพิกเซลสีดำปรากฏ ชิ้นส่วนถัดไปที่จะนำมาต่อก็จะถูกนำพิกเซลที่บริเวณขอบทางซ้ายมาเทียบกับค่าเหล่านี้ เพื่อความสะดวกในการจัดเรียงชิ้นส่วน เราจะคำนวณค่าความเหมาะสมที่จะนำมาต่อกันของชิ้นส่วนแต่ละคู่เตรียมไว้ และเก็บค่าไว้ในเมตริกซ์ขนาด $N \times N$ เมื่อ N เป็นจำนวนชิ้นส่วนที่ต้องการนำมาต่อกัน

3.3 การค้นหาชิ้นส่วน

การค้นหาชิ้นส่วนที่จะนำมาเรียงต่อกัน จะพิจารณาจากค่าในเมตริกซ์ที่ได้สร้างไว้ในขั้นตอนก่อนหน้า โดยเลือกชิ้นส่วนที่จะเป็นขอบด้านซ้ายและด้านขวาขึ้นมาก่อน จากนั้นจึงมองหาชิ้นส่วนที่เหมาะสมที่สุดที่จะนำมาต่อกับชิ้นส่วนซ้ายสุดและขวาสุดนี้ โดยในการทดสอบจะเลือกชิ้นส่วนที่เหมาะสมที่สุดสำหรับการต่อทางซ้ายและขวาสลับกันไป

ในส่วนของการออกแบบโปรแกรม การจัดเรียงเศษเอกสารที่มีจำนวนชิ้นเป็นจำนวนมากหรืออาจเป็นเศษเอกสารที่มาจากเอกสารหลายๆแผ่น ทำให้ค่าที่ได้จากฟังก์ชันการให้คะแนนอาจไม่สามารถตัดสินได้อย่างแม่นยำ ระบบจะทำการเรียงลำดับคะแนนของคู่เอกสารที่ต่อกันได้สนิทจากมากไปหาน้อย เพื่อให้ผู้ใช้งานสามารถเลือกปรับเปลี่ยนตำแหน่งของการจัดเรียงเศษเอกสารได้ โดยใช้เวลาในการเปรียบเทียบและตัดสินใจไม่มาก เพราะโปรแกรมได้ตัดตัวเลือกที่ไม่เหมาะสมออกไปให้บางส่วนแล้ว

3.4 การสร้างเอกสารกลับคืน

ในขั้นตอนนี้จะทำการสร้างเอกสารกลับคืนมา โดยจะบันทึกออกมาเป็นไฟล์ JPG หรือ PNG เพื่อตวามสวยงามของเอกสาร ชิ้นส่วนว่างที่ไม่มีข้อมูลจะถูกนำมาเฉลี่ยไปทางซ้ายและขวาในจำนวนที่ใกล้เคียงกันเพื่อในเนื้อหาในเอกสารอยู่บริเวณตรงกลาง และขนาดหรือสัดส่วนของภาพใกล้เคียงกับเอกสารจริง ในส่วนของโปรแกรมที่ได้ออกแบบ จะมีส่วนทีให้ผู้ใช้งานเลือกปรับเปลี่ยนลำดับชิ้นส่วนก่อนการบันทึก

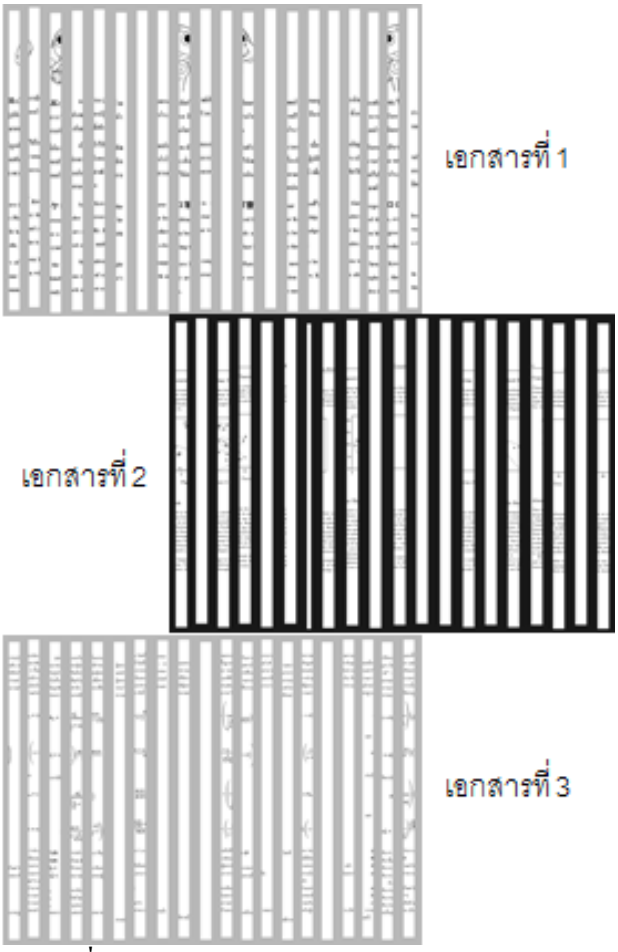
4. การทดสอบและผลการประเมิน

เพื่อทดสอบการต่อเอกสารกลับคืน ในงานนี้จะใช้ภาพชิ้นส่วนเอกสารที่ผ่านการเตรียมโดยทำการตัดและสลับชิ้นส่วนของเอกสารด้วยโปรแกรมตกแต่งภาพ เพื่อลดข้อผิดพลาดในส่วนของการเก็บตัวอย่าง และการประมวลผลภาพเบื้องต้น ภาพชิ้นส่วนเอกสารเหล่านี้จะถูกนำมาตัดเป็นชิ้นส่วนย่อยๆ และทำการจัดเรียงด้วยวิธีการทำนายเชิงเส้นเปรียบเทียบกับการพิจารณาแค่ขอบอย่างเดียว

การวัดความถูกต้องของการต่อชิ้นส่วนจะใช้การตรวจสอบว่าชิ้นส่วนนั้นต่อกับชิ้นส่วนที่ถูกต้องหรือไม่ เช่น ถ้าลำดับของชิ้นส่วนที่ถูกต้องคือ 1 2 3 4 5 แต่ถ้าผลการต่อเอกสารออกมาเป็น 1 2 5 3 4 จะมีชิ้นส่วนที่ผิดอยู่ 2 ชิ้น คือชิ้นหมายเลข 5 และหมายเลข 3 ที่ต่ออยู่กับชิ้นส่วนที่ไม่ถูกต้อง

การทดสอบจะใช้ภาพเอกสารที่ผ่านการตัด จัดเรียงไว้ในลำดับที่ไม่ถูกต้อง และอยู่บนพื้นหลังที่มีสีเข้มกว่า ดังแสดงในภาพที่ 7 ผลการทดสอบแสดงดังในตารางที่ 1 พบว่าเอกสารแรกสามารถต่อได้ถูกต้อง 100% ทั้ง 2 วิธี เนื่องจากเป็นเอกสารที่ตัวอักษรค่อนข้างใหญ่ มีส่วนของภาพที่เป็นลายเส้น ดังแสดงในภาพที่ 8 ทำให้ง่ายในการต่อเอกสารชิ้นนี้

ส่วนเอกสารชิ้นที่ 2 เป็นข้อความและมีภาพประกอบ ซึ่งหากต่อเอกสารโดยพิจารณาแค่ขอบอย่างเดียวให้ความถูกต้องเพียง 25% แต่เมื่อใช้การทำนายเชิงเส้นเข้ามาช่วย ทำให้ความถูกต้องเพิ่มขึ้นเป็น 41.67% จะเห็นว่ายังมีบางส่วนที่ยังมีความผิดพลาดอยู่ เนื่องจากตัวอักษรค่อนข้างเล็ก และภาพประกอบมีลักษณะเป็นจุดและเส้นประอยู่ด้วย

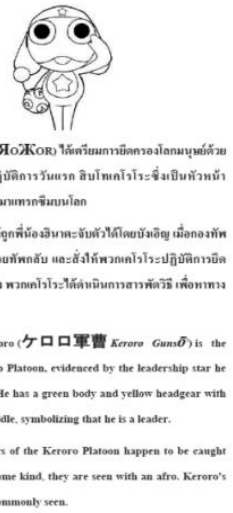


ภาพที่ 7: ภาพชิ้นส่วนเอกสารที่จะนำมาใช้ทดสอบ

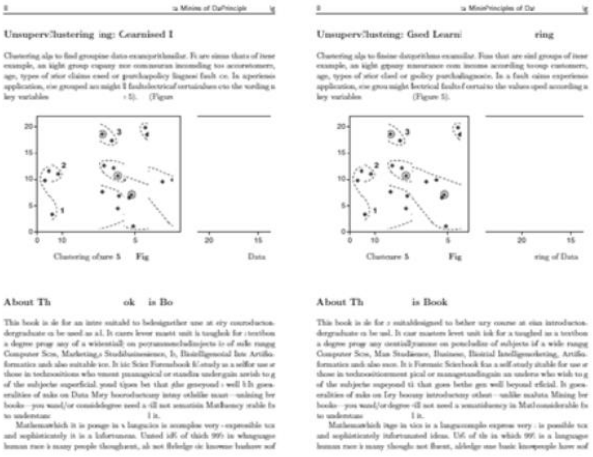
ตารางที่ 1: ผลการทดสอบเปรียบเทียบการใช้การทำนายเชิงเส้นและการพิจารณาแค่ขอบสำหรับการต่อเอกสารกลับคืน

เอกสารที่	จำนวนชิ้นส่วน		ความถูกต้อง (%)	
	ชิ้นส่วนที่มีข้อมูล	ชิ้นส่วนที่ไม่มีข้อมูล	พิจารณาแค่ขอบ	ใช้การทำนายเชิงเส้น
1	16	4	**100**	**100**
2	12	8	25.00	**41.67**
3	18	2	66.67	**88.89**

เอกสารที่ 3 มีลักษณะเป็นสมการและคำอธิบาย มีการใช้วงเล็บขนาดใหญ่ ทำให้การใช้การทำนายเชิงเส้นสามารถช่วยเพิ่มความถูกต้องของการต่อเอกสารได้เป็นอย่างดี เอกสารที่เป็นผลลัพธ์ของการสร้างกลับคืน แสดงในภาพที่ 8-10

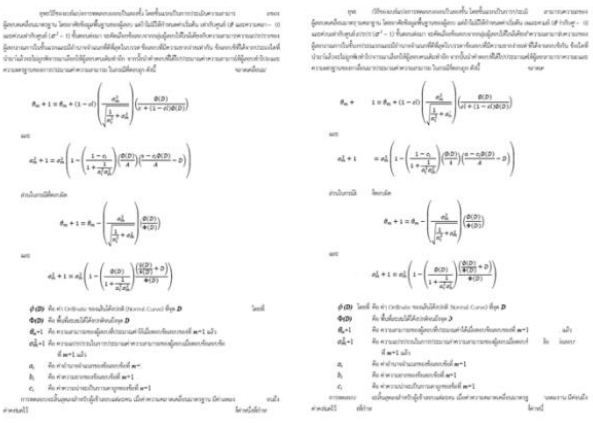


ภาพที่ 8: ผลการสร้างเอกสารที่ 1 กลับคืน ให้ผลเหมือนกันทั้ง 2 วิธี



ก. พิจารณาแค่ขอบ ข. การทำนายเชิงเส้น

ภาพที่ 9: ผลการสร้างเอกสารที่ 2 กลับคืน



ก. พิจารณาแค่ขอบ ข. การทำนายเชิงเส้น

ภาพที่ 10: ผลการสร้างเอกสารที่ 3 กลับคืน

5. สรุป

บทความนี้ได้นำเสนอวิธีการในการต่อเอกสารที่ถูกทำลายโดยผ่านเครื่องตัดเอกสาร เพื่อให้ได้เอกสารที่ใกล้เคียงกับเอกสารเดิมกลับคืนมา โดยใช้การทำนายเชิงเส้นมาช่วยในการกำหนดคะแนนให้กับชิ้นส่วนที่จะนำมาต่อ จากผลการทดสอบแสดงให้เห็นว่าวิธีการที่นำเสนอ สามารถต่อเอกสารกลับคืนมาได้ โดยให้ความถูกต้องสูงกว่าหรือใกล้เคียงกับการต่อชิ้นส่วนที่พิจารณาเพียงแค่ขอบของชิ้นส่วน และสามารถใช้ต่อเอกสารที่มีภาพหรือสมการอยู่ในนั้นได้ ซึ่งจะช่วยทำให้งานกู้คืนเอกสารที่ถูกทำลายสามารถทำได้เวลาที่รวดเร็วขึ้น

วิธีการที่นำเสนอนี้อาจปรับปรุงให้มีประสิทธิภาพที่ดีขึ้น โดยพิจารณาจำนวนพิกเซลเพิ่มมากขึ้น หรือปรับปรุงขั้นตอนในการค้นหาชิ้นส่วนที่เหมาะสม โดยดูจากค่ารวมหลักจากต่อเสร็จแล้ว แทนการพิจารณาทีละชิ้น เพื่อให้ได้คำตอบรวมที่ดีที่สุด และสามารถสร้างเอกสารขึ้นมาใหม่จากเศษของเอกสารที่ผ่านเครื่องทำลายเอกสารแล้วเพื่อช่วยในการสืบสวนสอบสวนหรือ นำเอกสารที่ถูกทำลายไปแล้วกลับคืนมา เพื่อใช้เป็นหลักฐานประกอบการพิจารณาคดีต่อไปได้

เอกสารอ้างอิง

[1] Shredded Paper [online].https://recycling.ncsu.edu/ what-to-do-with/shredded-paper/

[2] DARPA. DARPA’s Shredder Challenge Solved [online]. http://www.darpa.mil/NewsEvents/Releases/ 2011/12/02_.aspx "DARPA Shredder Challenge".

[3] A. Deever and A. Gallagher, "Semi-automatic assembly of real cross-cut shredded documents," Image Processing (ICIP), 2012 19th IEEE International Conference on, pp.233 - 236, 2012. doi: 10.1109/ICIP.2012.6466838

[4] R. Szeliski, "Image alignment and stitching: a tutorial," Foundations and Trends in Computer Graphics and Vision, 2(1), pp.1 - 104, 2006. doi:10.1561/0600000009

[5] H.Y. Lin and W.C. Fan-Chiang, “Reconstruction of shredded document based on image feature matching”, Expert Systems with Applications, 39, 2012, pp. 3324–3332.

[6] R. Ranca, “A Modular Framework for the Automatic Reconstruction of Shredded Documents,” presented at the Twenty-Seventh AAAI Conference on Artificial Intelligence, Bellevue, WA, 2013.

[7] R. Ranca, "Reconstructing Shredded Documents," B.S. thesis, School of Informatics, Univ. Edinburgh, Edinburgh, UK, 2013.

ระบบการค้นคืนเอกสารข้อความภาษาไทยด้วยเสียง
A System of Thai Text Document Retrieval by Speech

ปพนพัชร์ โสภณ (Paphonput Sopon)[1] และจันทิมา พลพินิจ (Jantima Polpinij)[2]
ห้องปฏิบัติการวิจัยเชิงปัญญา (Intellect Laboratory)
ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาการสารสนเทศ มหาวิทยาลัยมหาสารคาม
[1]paphonput@gmail.com, [2]jantima.p@msu.ac.th

บทคัดย่อ

งานวิจัยฉบับนี้นำเสนอระบบค้นคืนเอกสารภาษาไทยด้วยเสียงพูด โดยในระบบนี้ประกอบด้วยการทำงานหลักใน 2 ส่วน คือส่วนของการรู้จำเสียงเพื่อให้ได้โมเดลเสียง และส่วนของการนำโมเดลเสียงไปค้นคืนเอกสาร ในขั้นตอนการรู้จำเสียงพูดนั้นจะเป็นการรู้จำเสียงพูดภาษาไทยแบบคำโดด โดยใช้โมเดลฮิดเด็นมาร์คอฟเป็นเครื่องมือหลักในการรู้จำเสียง ขณะที่ส่วนที่ 2 คือส่วนที่ใช้ในการค้นคืนเอกสารด้วยเสียง ที่ประกอบด้วยองค์ประกอบ 2 ส่วน คือส่วนของการแปลงเสียงพูดเป็นข้อความ และส่วนของการค้นหาเอกสารด้วยคำสำคัญ

คำสำคัญ: การรู้จำเสียง การค้นคืนสารสนเทศด้วยเสียง ภาษาไทย

Abstract

This work presents a system of Thai text document retrieval by speech. The system consists of 2 main components: speech recognition and speech-based text retrieval. The first component is speech recognition that is experimented with the isolated word recognition. The main mechanism of this component is Hidden Markov Models that is applied for recognizing of speech. The second component is speech-based text retrieval that consist of 2 significant components: speech to text and text retrieval by keywords. In the stage of system testing, the relevant document relating to illness and medicine household in everyday life.

Keyword: Speech Recognition, Speech-based Text Retrieval, Thai Language.

1. บทนำ

ปัจจุบันการค้นคืนสารสนเทศ (Information Retrieval) [1] เป็นเครื่องมือที่สำคัญในการสืบค้นข้อมูลของระบบต่างๆ เช่น ห้องสมุดดิจิทัล (Digital Library) หรือ เสิร์ชเอนจิน (Search Engine) โดยทั่วไป ระบบค้นคืนสารสนเทศเกือบทั้งหมดจะขับเคลื่อนด้วยการใช้ข้อความในการค้นคืนสารสนเทศ (text-based information retrieval) [1][2] เนื่องจากข้อมูลที่มีอยู่ในปัจจุบันส่วนใหญ่อยู่ในรูปแบบข้อมูลที่ไม่มีโครงสร้าง (Unstructured data) เช่น ข้อความ (Text) [3][4] ซึ่งระบบการใช้ข้อความในการค้นคืนสารสนเทศ [5] คือ การป้อนคำขอ (query) ด้วย '*คำ*' หรือ '*กลุ่มคำ*' จากนั้นใช้คำเหล่านี้ในการค้นคืนสารสนเทศที่สอดคล้องกับคำขอ

ปัจจุบันมีอุปกรณ์หรือแอพพลิเคชันประเภทที่ไม่ต้องใช้คีย์บอร์ดเพิ่มมากขึ้น การค้นคืนสารสนเทศด้วยเสียง (Speech-based Information Retrieval) [6][7] จึงกลายเป็นเทคโนโลยีสำคัญสำหรับการสืบค้นข้อมูลในปัจจุบัน ทำให้มีการใช้อย่างแพร่หลายมากขึ้น เช่น การค้นหาข้อมูลเอกสารด้วยเสียงบน Google หรือ Smart phone

อย่างไรก็ตาม ยังไม่มีโปรแกรมใดๆ ที่เปิดโอกาสให้ผู้ใช้สามารถดาวน์โหลดลงมาใช้บนเครื่องคอมพิวเตอร์ส่วนบุคคลเพื่อใช้สำหรับการค้นหาเอกสารหรือข้อมูลในเครื่องคอมพิวเตอร์ส่วนบุคคล แม้จะมีผู้พัฒนาโปรแกรมที่ชื่อ Mitini (หรืออาจจะเรียกว่า SIRI สำหรับ Windows) แต่โปรแกรมดังกล่าวเป็นโปรแกรมสำหรับการใช้งานในการค้นคืนเอกสารด้วยเสียงบนเครื่องคอมพิวเตอร์ส่วนบุคคล แต่ภาษาที่ใช้งานคือภาษาอังกฤษเท่านั้น

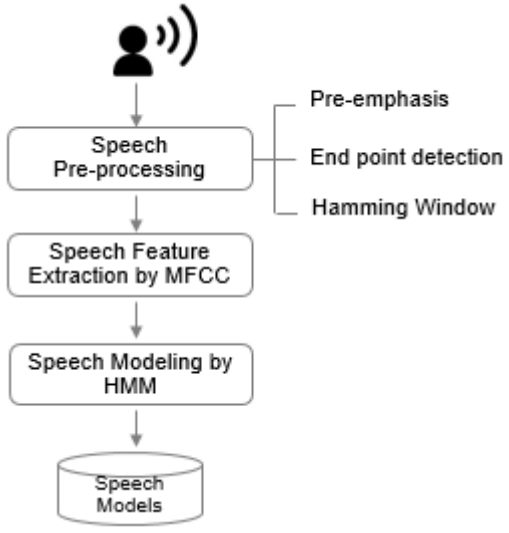
ดังนั้นงานวิจัยฉบับนี้ จึงนำเสนอการวิจัยและพัฒนาระบบต้นแบบในการค้นคืนเอกสารด้วยเสียงในเครื่องคอมพิวเตอร์ส่วนบุคคลสำหรับภาษาไทย โดยระบบนี้จะประกอบด้วยการทำงานหลักๆ ใน 2 ส่วน คือส่วนของการรู้จำเสียงเพื่อให้ได้โมเดลเสียง และส่วนของการนำโมเดลเสียงไปค้นคืนเอกสาร โดยในที่ส่วนที่ 2 ที่ใช้ในการค้นคืนเอกสารด้วยเสียง จะประกอบด้วยองค์ประกอบ 2 ส่วน คือส่วนของการแปลงเสียงพูดเป็นข้อความ (Speech-to-text) และส่วนของการค้นหาเอกสารด้วยคำสำคัญ (Keyword search)

2. ระบบการค้นคืนเอกสารภาษาไทยด้วยเสียง (A System of Thai Text Document Retrieval by Speech)

ในส่วนนี้จะแสดงกรอบการพพัฒนาระบบการค้นคืนเอกสารภาษาไทยด้วยเสียง โดยระบบที่นำเสนอประกอบด้วยองค์ประกอบหลัก 2 ส่วนคือ ส่วนของการรู้จำเสียงเพื่อให้ได้โมเดลเสียง และส่วนของการนำโมเดลเสียงไปค้นคืนเอกสาร

2.1 การรู้จำเสียงพูดภาษาไทย (Thai Speech Recognition)

กระบวนการรู้จำเสียงพูดภาษาไทยสามารถแสดงได้ดังภาพที่ 1 และสามารถอธิบายแต่ละขั้นตอนได้ดังนี้



ภาพที่ 1: ขั้นตอนการรู้จำเสียงพูดภาษาไทย

ขั้นที่ 1: คือการเตรียมข้อมูลเสียง (Speech Pre-processing) เป็นการปรับสัญญาณเสียงที่รับเข้ามา (ไฟล์ .wav) ให้อยู่ในรูปแบบที่พร้อมต่อสกัดและการคัดเลือกคุณลักษณะของเสียง โดยทั่วไปการเตรียมข้อมูลเสียง ประกอบด้วยขั้นตอนย่อยๆ หลายขั้นตอน ได้แก่

(1) Pre-emphasis [8] เป็นขั้นตอนการกำจัดสัญญารบกวน (Noise Removal) โดยธรรมชาติของเสียงเมื่อเสียงเดินทางผ่านอากาศจะมีสัญญาณรบกวน (Noise Signal) ทำให้การประมวลผลเสียงมีความผิดพลาดสูง

(2) การหาจุดสิ้นสุดของเสียงพูด (End Point Detection) เป็นการหาจุดเริ่มต้นและจุดสิ้นสุดของสัญญาณเสียง เพื่อให้ได้ช่วงเสียงที่จะนำไปใช้ในขั้นตอนการหาคุณลักษณะ (Feature Extraction) ในงานวิจัยฉบับนี้ใช้เทคนิคที่เรียกว่า Probability Density Function [9]

(3) การวางกรอบสัญญาณแบบแฮมมิ่ง (Hamming Window) [10] เนื่องจากว่าเสียงมีคุณสมบัติคือการเปลี่ยนแปลงของสัญญาณไม่ขึ้นกับเวลา และมีลักษณะการเปลี่ยนแปลงที่หลากหลาย ดังนั้นจึงจำเป็นต้องวิเคราะห์สัญญาณเสียงทีละช่วงสั้นๆ ที่อยู่ภายในกรอบสัญญาณที่กำหนดขึ้นโดยขนาดของกรอบสัญญาณที่นิยมใช้มีค่าระหว่าง 10-30 มิลลิวินาที และควรวางกรอบสัญญาณถัดไปให้มีความเหลื่อมกันประมาณ 1/3 ของขนาดกรอบสัญญาณ เพื่อให้ข้อมูลที่วิเคราะห์มีความต่อเนื่องกัน สำหรับ Hamming Window เป็นวิธีวางกรอบสัญญาณเสียงที่ทำให้มีความต่อเนื่องของสัญญาณเสียงมากที่สุดและเกิดการสูญหายของสัญญาณเสียงน้อยที่สุด

ขั้นที่ 2: การสกัดและการคัดเลือกคุณลักษณะของเสียง (Speech Feature Selection and Extraction) เป็นการหาสกัดคุณลักษณะของเสียงที่มีความเหมาะสมต่อขั้นตอนของการเรียนรู้ เทคนิคที่ใช้ คือ และสัมประสิทธิ์เซปสตรัมบนสเกลเมล (Mel-Frequency Cepstral Coefficients: MFCC) [11] โดยการหาค่า MFCC เริ่มต้นจากนำสัญญาณเสียงมาผ่านการประมวลผลสัญญาณเสียงพูดเบื้องต้น (Preprocessed speech) ผ่าน Fast Fourier Transform (FFF) จากนั้นส่งสัญญาณเสียงผ่านชุดตัวกรอง (Filter Bank) เพื่อเน้นความถี่ที่อยู่ในช่วงกลางของชุดตัวกรองแต่ละตัวกรอง โดยที่ความถี่กลางของตัวกรองแต่ละอันนั้นเกิดจากการแปลงค่าความถี่ปรกติให้อยู่บนสเกลเมล ต่อมาจึงคำนวณค่าพลังงานเสียงและเมื่อนำค่าลอการิทึมของพลังงาน (Logarithm of Energy) นี้มาผ่านการแปลง

โคซายน์แบบไม่ต่อเนื่อง (Discrete Cosine Transformation: DCT) ก็จะได้ค่า MFCC ในที่สุด

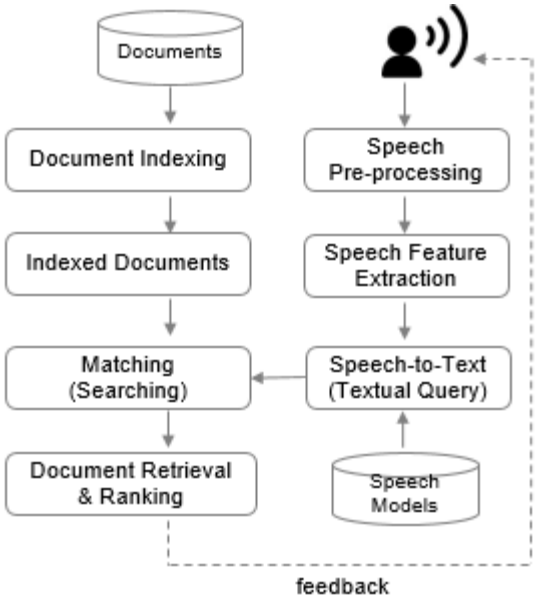
ขั้นที่ 3: การสร้างโมเดลของเสียงพูด (Acoustic Modeling) ซึ่งขั้นตอนสร้างโมเดลเสียงนี้ จะสร้างโมเดลเสียงของแต่ละคำจากคุณลักษณะของเสียงที่สกัดได้ เพื่อกำหนดสัญลักษณ์แทน "*คำ*" หรือ "*เสียงพูด*" ซึ่งในการรู้จำเสียงพูดจะกระทำผ่าน Hidden Markov Model (HMM) [12] เนื่องจาก เหมาะสำหรับการฝึกฝนข้อมูลเสียงที่ใช้งานร่วมกับระบบที่ไม่มีการกำหนดตายช่วย เช่น เสียงพูดขณะที่พูดแบบไม่ขึ้นกับผู้พูด [13] โดยทั่วไปหน่วยเสียงที่เล็กที่สุดในภาษาคือ phoneme ซึ่ง phoneme จะเป็นส่วนที่นิยมใช้ในการรู้จำเสียงพูด ยกตัวอย่าง เช่น คำว่า "ป่วย" ซึ่งสามารถแยก phoneme ได้คือ /ป/ /^/ /ว/ และ /ย^/

ภายหลังการรู้จำเสียงด้วย HMM จะได้โมเดลเสียง (Acoustic Models) เพื่อนำไปใช้ในส่วนของการค้นคืนเอกสารข้อความด้วยเสียง

2.2 การค้นคืนเอกสารด้วยเสียง (Speech-based Thai Text Retrieval)

ในส่วนของการค้นคืนเอกสารด้วยเสียง คือ ส่วนของการการค้นเอกสารข้อความภายในเครื่องคอมพิวเตอร์ส่วนบุคคล โดยใช้เสียงพูดเพื่อให้ได้เอกสารที่ต้องการกลับมา (Selective Searching) ซึ่งแนวคิดวนการค้นคืนเอกสารนั้นจะอาศัยลักษณะเด่นของเนื้อหาของเอกสารเป็นตัวแทนของเอกสาร (Document surrogate) ที่สามารถใช้เพื่อแยกแยะเอกสารที่เกี่ยวข้อง (relevant) และไม่เอกสารที่ไม่เกี่ยวข้อง (Non-relevant) ออกจากกันได้ ภายหลังจากการวิเคราะห์ร่วมกับคำขอ (Query) ที่ผู้ใช้งานป้อนเข้าสู่ระบบ

สำหรับกรอบการดำเนินงานในระบบการค้นคืนเอกสารด้วยเสียง เพื่อการใช้งานบนคอมพิวเตอร์ส่วนบุคคล สามารถแบ่งการการงานรออกเป็น 3 ส่วนใหญ่ๆ คือ ส่วนของการจัดการเอกสาร ส่วนของการแสดงรับคำขอ และส่วนของการค้นคืนเอกสาร ได้ดังภาพที่ 2



ภาพที่ 2: กรอบการค้นคืนเอกสารด้วยเสียง

ส่วนที่ 1: ส่วนของการจัดการเอกสาร

เป็นส่วนที่ใช้ในการดำเนินการเพื่อการจัดการข้อมูลเอกสารให้ง่ายต่อการสืบค้น โดยมี 2 ส่วนย่อยคือ

1. กระบวนการทำดัชนีเอกสาร (Indexing) เพื่อใช้เป็นเครื่องมือสำหรับค้นหาข้อมูล โดยการทำดัชนีเอกสารในงานวิจัยนี้เป็นการใช้คำสำคัญ (Keyword) หรือกลุ่มคำสำคัญในเอกสารเป็นดัชนี
2. การจัดเก็บเอกสารพร้อมดัชนี ซึ่งเอกสารจะมีการนำมาผ่านกระบวนการทำดัชนีเอกสารแล้ว จะมีการบันทึกคำสำคัญไว้ร่วมกับตำแหน่งการจัดเก็บไฟล์เอกสารที่สอดคล้องกับคำสำคัญนั้นๆ ในโครงสร้างข้อมูล XML ดังแสดงในภาพที่ y

ส่วนที่ 2: ส่วนของการแสดงรับคำขอ

ในระบบฯ ที่นำเสนอ ผู้ใช้จะสามารถระบุ "*คำขอ*" แบบ "*เสียงพูด*" จากนั้นระบบฯ ก็จะนำ "*เสียงพูด*" ของผู้ใช้งานมาผ่านขั้นตอนการเตรียมเสียงพูด (Pre-processing) รวมถึงการสกัดคุณลักษณะเสียงพูด (Feature Extraction) เช่นเดียวกับขั้นตอนของการเตรียมเสียงพูดและการสกัดคุณลักษณะของเสียงในส่วนของการสร้างโมเดลเพื่อรู้จำเสียงพูด จากนั้นจะนำคุณลักษณะเสียงที่ได้ ไปเปรียบเทียบกับโมเดลเสียงที่สร้างเอาไว้ หากตรงกันก็จะแสดงผลลัพธ์ออกมาในรูปแบบของ "*คำ*" หรือ "*กลุ่มคำ*" นั่นเป็นการแปลงเสียงเป็นข้อความ หรือ

Speech-to-Text นั่นเอง ซึ่งในการค้นคืนข้อมูลก็จะใช้คำหรือกลุ่มคำที่ได้นี้ เป็นคำขอในการค้นคืนเอกสารข้อความภาษาไทยตามวิธีการสืบค้นเอกสารด้วยคำสำคัญ (Keyword search)

ส่วนที่ 3: ส่วนของการค้นคืนเอกสาร

เมื่อได้คำขอที่เป็นคำหรือกลุ่มคำแล้ว จะมีการนำคำหรือกลุ่มคำเหล่านี้มาใช้เป็นเครื่องมือในการสืบค้นเอกสารที่เกี่ยวข้องหรือที่สอดคล้องกับความต้องการของผู้ใช้งานผ่านกระบวนการ Matching แบบ Exact Match บนรูปแบบการค้นคืนเอกสารแบบ Vector Space Model (VSM) ซึ่งเป็นเวคเตอร์แบบ 2 มิติที่แสดงความสัมพันธ์ระหว่างเอกสารที่มีเนื้อหาแบบคำและคำขอ

สมมติคำขอคือ "*ปวดท้อง*" และในคลังเอกสารมีเอกสารทั้งหมด 3 เอกสาร โดยเอกสารแต่ละฉบับมีดังนี้

เอกสาร#1 : *อาการ, ปวดฟัน*

เอกสาร#2 : *อาการ, ปวดหัว*

เอกสาร#3 : *ปวดหัว, ตัวร้อน*

เมื่อนำเอกสารข้างต้นพร้อมคำสำคัญหรือดัชนีของเอกสารนั้นๆ แสดงในรูปแบบของ VSM จะสามารถแสดงได้ดังตารางที่ 1 โดยค่า '0' ที่ปรากฏในตารางแสดงว่าเอกสารนั้นไม่มีคำที่ตรงกับคำขอ ส่วนค่า '1' ที่ปรากฏในตารางแสดงว่าเอกสารนั้นมีคำที่ตรงกับคำขอ

ตารางที่ 1: เอกสารในการนำเสนอแบบ VSM

เอกสาร/คำขอ	อาการ	ปวดท้อง	ปวดฟัน	เป็นไข้
#1	0	0	0	0
#2	0	1	0	0
#3	0	1	0	0

จากนั้นจะใช้การพิจารณาแบบ *Euclidean Distance* เพื่อพิจารณาความคล้ายคลึงของเอกสารสำหรับการสืบค้นเอกสารในคอมพิวเตอร์ส่วนบุคคลนั้น ซึ่งขั้นตอนนี้คือการค้นคืนเอกสาร หรือ *Document Retrieval* นั่นเอง

ภายหลังจากที่ค้นคืนเอกสารได้ทั้งหมด ระบบฯ จะทำการจัดลำดับของเอกสารที่ค้นคืนได้ (Ranking) โดยการพิจารณาจำนวน "คำขอ" ที่พบในเอกสารที่ค้นคืนได้แต่ละฉบับว่า ฉบับใดมีความถี่ (frequency) ของ "คำขอ" มากที่สุด เอกสารฉบับนั้นๆ ก็จะถูกจัดอันดับของการค้นคืนในลำดับต้นๆ

จากตัวอย่างข้างต้นนี้ เอกสารที่ 2 และ เอกสารที่ 3 จะถูกค้นคืนมาให้ เพราะมีคำที่ตรงกับ "คำขอ"

3. การประเมินระบบ

ในการประเมินระบบการค้นคืนเอกสารข้อความภาษาไทยด้วยเสียงจะกระทำใน 2 ส่วนหลักของระบบคือ การประเมินการรู้จำเสียง และการประเมินการสืบค้นข้อมูลด้วยเสียง

3.1 ชุดข้อมูล

ข้อมูลที่ใช้ในงานวิจัยฉบับนี้มี 2 ชุดข้อมูลคือ ชุดข้อมูลเสียงพูดภาษาไทยเพื่อใช้ในการรู้จำเสียงพูดภาษาไทย และชุดข้อมูลเอกสารภาษาไทยที่จะใช้ทดลองการสืบค้นเอกสารด้วยเสียง ในการรู้จำเสียงจะรู้จำเสียงจากคำโดดจำนวน 7 คำ ได้แก่ ปวด ท้อง เสีย ไข้ หวัด หัว และ เป็น

(1) ชุดข้อมูลเสียงพูดภาษาไทย

ข้อมูลเสียงที่จะใช้ในการสร้างโมเดลเสียง จะเก็บจากคนจำนวน 50 คน โดยเป็นข้อมูลเสียงจากผู้ชายจำนวน 25 คน และข้อมูลเสียงจากผู้หญิงจำนวน 25 คน โดยแต่ละคนจะเก็บข้อมูลเสียงเป็นจำนวน 5 รอบ ซึ่งไฟล์เสียงจะเก็บในรูปแบบของ .wav

(2) ชุดข้อมูลเอกสารภาษาไทย

เอกสารข้อมูลเพื่อใช้ในการทดสอบการสืบค้น จะเป็นเอกสารเกี่ยวกับอาการป่วยในชีวิตประจำวัน ซึ่งจะเก็บอยู่ในรูปแบบของ XML **โดย**คลังเอกสารของเราได้ทำการรวบเอกสารไว้ทั้งสิ้น 200 เอกสาร แยกเป็นเอกสารเกี่ยวกับการเป็นไข้ ปวดท้อง เป็นหวัด และ ปวดหัว จำนวนกลุ่มละ 50 เอกสาร

3.2 การประเมินการรู้จำเสียง

การวัดประสิทธิภาพของการรู้จำเสียงคือการตรวจสอบความถูกต้อง หรือจำนวนความถูกต้องว่าสามารถรู้จำได้ตรงกับที่ต้องการหรือไม่ ขั้นตอนนี้ถือเป็นสิ่งที่จำเป็นมาก เพราะจะเป็นการประเมินโมเดลเสียงที่ได้ว่าสามารถนำไปใช้งานได้หรือไม่ ในงานวิจัยฉบับนี้จะวัดประสิทธิภาพของการรู้จำเสียงด้วยวิธีการที่เรียกว่า Word Error Rate (WER) [14]

$$WER = \frac{insertion + Substitution + Deletion}{Total\ of\ Words\ in\ Transcript}$$

โดยที่

Insertion *คือคำที่ถูกแทรกขึ้นมาจากข้อความเดิม*
Substitution *คือคำที่ถูกแทนที่ไปจากคำเดิม*
Deletion *คือคำที่หายไปจากข้อความเดิม*
Total of Words *คือจำนวนคำทั้งหมดที่รับเข้ามา*

ระดับที่น่าพอใจนั่นคือ มีค่าที่ 5% ซึ่งค่า WER ยิ่งมีค่าน้อยประสิทธิภาพของโมเดลก็จะยิ่งดี [14]

3.3 การประเมินการสืบค้นข้อมูล

ในการสืบค้นเอกสารด้วยเสียงพูด จะทดสอบการสืบค้นด้วย “คำโดด” และ “คำผสม” ที่ไม่เกิน 2 คำโดด เช่น ปวดหัว ปวดท้อง ท้องเสีย เป็นไข้ เป็นหวัด เป็นต้นเป็นการนำเสนอ ในการ

โดยในการทดสอบประสิทธิภาพการรู้จำจะกระทำโดยผู้ทดสอบจำนวน 30 คน แยกเป็นชาย 15 คน และหญิง 15 คน ซึ่งกลุ่มคนเหล่านี้ไม่ใช่กลุ่มคนที่ใช้ในการเก็บข้อมูลสำหรับการสร้างโมเดลเสียง ซึ่งผลการทดสอบในระดับค่า WER อยู่ใน

โดยในการทดสอบ จะทดสอบด้วยค่าความระลึก (Recall) [14] ค่าความแม่นยำ (Precision) [14] และค่าเอฟ (*F*-measure) [14] ได้ผลลัพธ์ดังแสดงในตารางที่ 2

ตารางที่ 2: ผลการค้นคืนเอกสารด้วยเสียง

เสียงพูด	ค่าความระลึก	ค่าความแม่นยำ	F-measure
เป็นไข้	0.79	0.80	0.79
เป็นหวัด	0.75	0.79	0.77
ปวดหัว	0.78	0.80	0.79
ปวดท้อง	0.76	0.82	0.79
ท้องเสีย	0.72	0.75	0.73
ปวด	0.80	0.82	0.81
ไข้	0.79	0.81	0.80
ค่าเฉลี่ย	0.77	0.80	0.78

จากตารางที่ 2 จะเห็นว่า ผลการค้นคืนเอกสารด้วยเสียงอยู่ในระดับที่น่าพอใจ สาเหตุก็เนื่องมาจากโมเดลเสียงที่มีประสิทธิภา ทำให้ในขั้นตอนการเปลี่ยนคำขอที่เป็น “เสียงพูด” ให้เป็น “ข้อความ” ที่จะใช้ในการสืบค้นเอกสารมีประสิทธิภาพด้วย นอกจากนี้ยังได้ทำการเก็บเอกสารทดสอบกระจายไปในส่วนต่างๆ ของเครื่องคอมพิวเตอร์ ก็พบว่าประสิทธิภาพของการสืบค้นเอกสารภาษาไทยด้วยเสียงยังอยู่ในระดับที่น่าพอใจเช่นกัน จากตัวอย่างที่แสดงในภาพที่ 3 เป็นการแสดงตัวอย่างการทำงานของระบบการค้นคืนเอกสารข้อความภาษาไทยด้วยเสียง โดยป้องเสียงพูดคำว่า “เป็นไข้” จากนั้นระบบฯ จะเปลี่ยนเสียงพูดให้เป็น “กลุ่มคำ” เพื่อใช้เป็นคำขอในการสืบค้นเอกสารที่เกี่ยวข้องในคอมพิวเตอร์ โดยระบบฯ จะจัดอันดับเอกสารที่ถูกค้นคืนตามจำนวนความถี่ของคำขอที่พบในแต่ละเอกสาร เอกสารใดมีความถี่ของคำขอมาก ก็จะถูกจัดอันดับให้อยู่ในอันดับของการค้นคืนอยู่ในอันดับแรกๆ

4. สรุปผลงานวิจัย

งานวิจัยฉบับนี้นำเสนอการวิจัยเพื่อพัฒนาระบบการค้นคืนเอกสารภาษาไทยด้วยเสียงพูด โดยระบบที่นำเสนอนี้ประกอบด้วยการทำงานใน 2 ส่วน คือส่วนของการรู้จำเสียงเพื่อให้ได้โมเดลเสียง และส่วนของการนำโมเดลเสียงไปค้นคืนเอกสาร ในขั้นตอนการรู้จำเสียงพูดนั้นจะเป็นการรู้จำเสียงพูดภาษาไทยแบบคำโดด ภายหลังจากที่ผู้ใช้งานป้อนคำขอด้วยเสียง เสียงพูดดังกล่าวจะถูกแปลงให้เป็นข้อความ เรียกว่า แปลงเสียงพูดเป็นข้อความ และในที่สุด เอกสารที่เกี่ยวข้องกับอาการป่วยและยาในชีวิตประจำวันก็จะถูกค้นคืน ตามแนวคิดของการค้นหาเอกสารด้วยคำสำคัญ

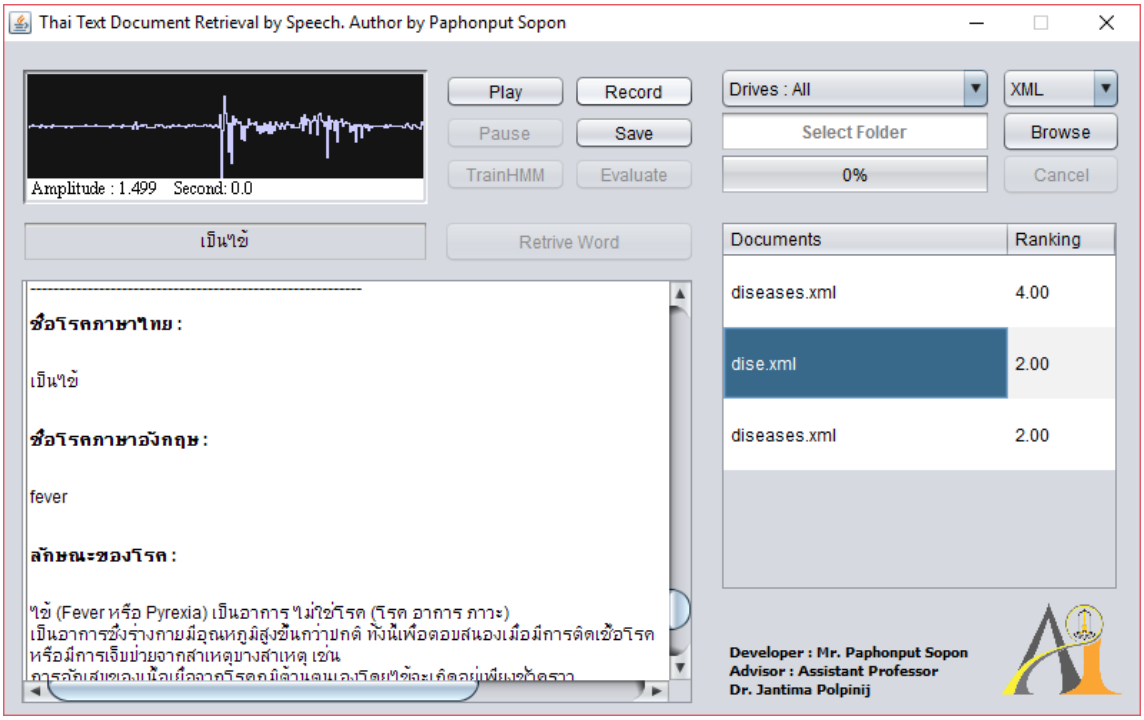
เอกสารอ้างอิง

[1] L.S. Lee and Y.C. Pan, “Voice-based information retrieval — how far are we from the text-based information retrieval?”, IEEE Workshop on Automatic Speech Recognition & Understanding, 2009.

[2] B.R. Bai, B. Chen, and H.M. Wang, “Syllable-Based Chinese Text/Spoken Document Retrieval. Using Text/Speech Queries”, International Journal of Pattern Recognition and Artificial Intelligence, Vol. 14, No. 5, pp. 603-616, 2000.

[3] T.Y. Lui, “Learning to Rank for Information Retrieval”, Database Management & Information Retrieval, Springer-Verlag Berlin Heidelberg, 2011.

[4] D. Sanchez, M.J. Martin-Bautista, I. Blanco and C. Torre, “Text Knowledge Mining: An Alternative to Text Data



ภาพที่ 3: ตัวอย่างการทำงานของระบบการค้นคืนเอกสารข้อความภาษาไทยด้วยเสียง

Mining", IEEE International Conference on Data Mining Workshops, 2008.

[5] R. Papka, J.P. Callan, and A.G. Barto, "Text-based Information Retrieval using Exponentiated Gradient Descent", Advances in Neural Information Processing Systems (NIPS), Vol. 9, 1996.

[6] A. Fujii, K. Ituo, and T. Ishikawa, "A Method for Open-Vocabulary Speech-Driven Text Retrieval", Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP), pp. 188-195, 2002.

[7] X. Anguera, L. Javier, I. Szoke, A. Buzo, and F. Metze, "Query by Example Search on Speech at Mediaeval 2014", Mediaeval, 2014.

[8] R. Vergin and D. O'Shaughnessy, "Pre-emphasis and speech recognition", Canadian Conference on Electrical and Computer Engineering, pp. 1062-1065, 1995.

[9] J.L. Shen, J.W. Hung, and L.S. Lee, "Robust Entropy-based Endpoint Detection for Speech Recognition in Noisy Environments", *The 19th Telecommunications Forum*, 2011.

[10] L.R. Rabiner and B.H. Juang, "Fundamentals of speech recognition", PTR Prentice Hall Englewood Cliffs, 1993.

[11] A. Das, M.R. Jena, and K.K. Barik, "Mel-Frequency Cepstral Coefficient (MFCC) - a Novel Method for Speaker Recognition", Journal of Digital Technologies, Vol. 1, No. 1, pp. , 2015.

[12] M. Gales and S. Young, "The Application of Hidden Markov Models in Speech Recognition", Foundations and Trends in Signal Processing, Vol. 1, No. 3, pp. 195–304, 2007.

[13] B.H. Juang, W. Chou, and C.H. Lee, Fellow Minimum Classification Error Rate Methods for Speech Recognition, IEEE Transactions on Speech and Audio Processing, vol. 5, no. 3, may 1997.

[14] R. Baeza-Yates and B. Ribeiro-Neto, Modern information retrieval. ACM Press, New York. 1999.

การวิเคราะห์การจัดกลุ่มข้อมูลการเคลื่อนไหวของกระดูกสะบักโดยใช้เคมีน
Analysis of Scapular Data Using K-mean Clustering

พรธีรา จันตา และวิสันต์ ตั้งวงษ์เจริญ
ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
เลขที่ 1 ซอยฉลองกรุง1 แขวงลาดกระบัง เขตลาดกระบัง กรุงเทพ 10520
porntheeratukta@gmail.com, ktwisan@kmitl.ac.th

บทคัดย่อ

ปัจจุบันคนมีพฤติกรรมที่เปลี่ยนไปทำให้สะบักเกิดความผิดปกติ เนื่องจากการดำรงชีวิตของคนปัจจุบันสาเหตุที่ทำให้สะบักเกิดความผิดปกติ เกิดจากหลายปัจจัย เช่นการทำงานใช้ที่ใช้กำลัง การเคลื่อนไหวที่ผิดท่าทาง และผู้สูงอายุ ดังนั้นงานวิจัยนี้ได้นำเสนอการวิเคราะห์การจัดกลุ่มการเคลื่อนไหวสะบักคนปกติและผิดปกติ โดยใช้อุปกรณ์ตรวจจับการเคลื่อนไหว Razor-IMU โดยจะส่งข้อมูลผ่านสัญญาณบลูทูธแสดงค่าที่วัดได้เป็นตัวเลขให้เครื่องคอมพิวเตอร์ประมวลผลและแปลงข้อมูลโดยสมการพิกัดเชิงขั้วเพื่อสร้างกราฟออยเลอร์ ซึ่งกราฟดังกล่าวทำให้ไม่สามารถจำแนกข้อมูลที่ชัดเจนได้ จากนั้นได้นำข้อมูลจากกราฟดังกล่าวมาจัดกลุ่มข้อมูลแบบเคมีนเพื่อจำแนกกลุ่มข้อมูลให้เห็นข้อมูลชัดเจนขึ้น และจากการทดลองพบว่าการจัดกลุ่มข้อมูลแบบเคมีนสามารถจำแนกคนปกติและคนผิดปกติได้ งานวิจัยนี้สามารถเป็นเครื่องมือที่จะช่วยแพทย์จำแนกอาการบาดเจ็บสะบักได้รวดเร็วยิ่งขึ้น

คำสำคัญ: การบาดเจ็บสะบัก เคมีน เลเซอร์ไอเอ็มยู ออยเลอร์

Abstract

Nowadays, people have behavior that changes the scapula. Due to the existence of the present, the cause of the scab causes abnormalities. Caused by many factors such as the use of force. Wrong move and the elderly Therefore, this research has presented an analysis of groupings of normal and abnormal scapulae movements. Using the Razor-IMU motion detector, it sends data through Bluetooth signals, displaying numerical values to the computer, processing and converting data by polar coordinates to generate an Euler graph. This makes it impossible to distinguish the clear data, and then the data from the graph is grouped into chemical data to identify the data group more clearly. And from the experiments, it was found that the classification of chemical data can be classified into normal and abnormal groups. This research can be a tool to help physicians identify scapular injuries more quickly.

Keyword: Shoulder injury, K-mean, Razor-IMU, Euler graph.

1. บทนำ

สะบักเป็นจุดสำคัญในการเคลื่อนไหวของแขนบริเวณหัวไหล่ สะบักเป็นเสมือนฐานที่คอยพยุงการเคลื่อนไหวและเป็นจุดที่ยึดติดกับส่วนของเอ็นและกล้ามเนื้อแขน ซึ่งการรักษาการบาดเจ็บของสะบักในอดีตที่ผ่านมา ทำได้เพียงใช้สายตาในการช่วยวิเคราะห์เท่านั้นทำให้ไม่สามารถวินิจฉัยโรคได้อย่างแม่นยำ และไม่สามารถวิเคราะห์ผลออกมาเป็นตัวเลขได้ จึงเป็นที่มาในการจำแนกเพื่อจัดกลุ่มการเคลื่อนไหวสะบักคนปกติและคนผิดปกติ ดังนั้นทางผู้วิจัยได้นำเสนอแนวคิดในการจำแนกกลุ่มของข้อมูลโดยการจัดกลุ่มของข้อมูลเพื่อช่วยให้ผลการรักษาของแพทย์มีความแม่นยำมากขึ้น งานวิจัยนี้จึงเป็นการวิเคราะห์เบื้องต้นเท่านั้น ซึ่งยังคงต้องมีการศึกษาและพัฒนาต่อไปอีกในอนาคต

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 งานวิจัยที่เกี่ยวข้อง

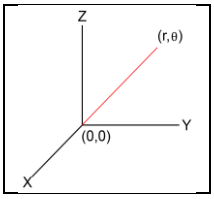
ในปี 2009 แคโรลีน แวน แอนเดล (Carolien van Andel) และคณะได้พัฒนาเครื่องมือที่ใช้ในการวัดการเคลื่อนไหวของสะบักมีชื่อว่าเครื่องมือ AMC [1] เครื่องมือนี้ประกอบด้วยกล้องตรวจจับการเคลื่อนไหว 3 ตัว ซึ่งติดหน้าอก สะบัก มีเครื่องมือย่อยคือเครื่องมือ Scapular Locator (SL) โดยติดที่แขนบนและแขนหน้า เพื่อดูการเคลื่อนไหวของสะบักในแต่ละท่าการหมุนไหล่ ซึ่งในปัจจุบันจะเห็นได้ว่ามีการเริ่มนำเอาเครื่องมือ AMC มาประยุกต์ใช้ในการตรวจวัดอาการบาดเจ็บแต่ละส่วนของร่างกายมากขึ้น

ในปี 2015 นที เหรียญตระกูลชัย วิสันต์ ตั้งวงษ์เจริญ และคณะได้นำเสนอการประยุกต์ใช้เซ็นเซอร์ตรวจจับการเคลื่อนไหว Razor-IMU ร่วมกับแอพพลิเคชั่นบนแท็บเล็ต [2] เพื่อใช้ในการตรวจวัดการเคลื่อนไหวของสะบัก และเป็นการลดการสังเกตหรือการวัดของผู้เชี่ยวชาญโดยตรงในการวัดการเคลื่อนไหวของกระดูกสะบัก ดังนั้นงานวิจัยนี้ช่วยให้ผู้เชี่ยวชาญทำการวิเคราะห์ได้สะดวกมากขึ้น

ในปี 2015 จิตเจียระไน พานิชชาติ และวรางคณา กิ้มปานได้วิจัยเกี่ยวกับการสร้างกราฟออยเลอร์หรือเรียกว่ากราฟเฟส[3] เพื่อสร้างรูปแบบโดยใช้เซ็นเซอร์ Razor-IMU ในการตรวจจับความเคลื่อนไหวของสะบัก และสมการพิกัดเชิงขั้วในการวิเคราะห์รูปแบบคนปกติและคนผิดปกติ เพื่อให้สังเกตง่ายขึ้น

2.2 ทฤษฎีกราฟออยเลอร์ (Euler Graph)

ในทางคณิตศาสตร์ระบบพิกัดเชิงขั้ว [4] คือ ระบบค่าพิกัดสองมิติในแต่ละจุดบนระนาบถูกกำหนดโดยระยะทางจากจุดตรึงและมุมจากทิศทางตรึง จุดตรึง เหมือนจุดกำเนิดของระบบพิกัดคาร์ทีเซียนเรียกว่าขั้ว และลากรังสีจากขั้วเข้ากับทิศทางตรึงคือแกนเชิงขั้ว ระยะทางจากขั้วเรียกว่าพิกัดรัศมีหรือรัศมี และมุมคือพิกัดมุม



รูปที่ 1 การแปลงพิกัดเชิงขั้วเป็นพิกัดคาร์ทีเซียน

ระหว่างระบบพิกัดคาร์ทีเซียนและระบบพิกัดเชิงขั้วสามารถแปลงค่าระหว่างสองพิกัดนี้ได้ โดยใช้ฟังก์ชันทางตรีโกณมิติ sin และ cos ดังสมการที่ (1) และ (2)

$$x = r\cos\theta \quad (1)$$

$$y = r\sin\theta \quad (2)$$

เมื่อ x, y คือ พิกัดคาร์ทีเซียน

r คือ รัศมี

θ คือ มุมที่ได้จากอุปกรณ์ตรวจจับการเคลื่อนไหว

2.3 การจัดกลุ่มข้อมูลแบบเคมีน (K-means Clustering)

การจัดกลุ่มข้อมูลแบบเคมีน [5] คือ กระบวนการจัดกลุ่มข้อมูลขั้นพื้นฐานแบบไม่มีผู้สอน (Unsupervised learning technique) มีเป้าหมายเพื่อการจำแนกกลุ่มข้อมูลที่มีคุณลักษณะคล้ายกันอยู่ในกลุ่มเดียวกัน โดยข้อมูลแต่ละกลุ่มจะถูกเรียกว่าคลัสเตอร์ (Cluster) โดยมีการตัดแบ่งส่วน (Partition) วัตถุออกเป็น K กลุ่ม โดยแทนกลุ่มแต่ละกลุ่มด้วยค่าเฉลี่ยของข้อมูล ซึ่งใช้ค่าเฉลี่ยเป็นจุดศูนย์กลางของกลุ่มในการวัดระยะห่างของข้อมูลในกลุ่มเดียวกัน ระยะห่างของข้อมูลจะมีค่าน้อยถ้าหากข้อมูลอยู่ในกลุ่มเดียว และระยะห่างข้อมูลจะมีค่ามาก ถ้าข้อมูลต่างกลุ่มกัน ในการหาระยะห่างโดยใช้วิธีคิดแบบระยะห่างยูคลิเดียน (Euclidean distance) โดยที่ค่า x แต่ละค่ามีค่า C_i ได้เพียงค่าเดียวเท่านั้น ดังสมการที่ (3)

$$SSE = \sum_{i=1}^{k} \sum_{x \in C_i} dist(c_i, x)^2 \quad (3)$$

เมื่อ x คือ วัตถุที่สนใจ

C_i คือ คลัสเตอร์ลำดับที่ i

c_i คือ จุดกึ่งกลางของคลัสเตอร์ลำดับที่ i

k คือ จำนวนของคลัสเตอร์

และปรับค่า Centroid จากสมการที่ (4)

$$c_i = \frac{1}{m_i} \sum_{x \in C_i} x \quad (4)$$

เมื่อ m_i คือ จำนวนของวัตถุในคลัสเตอร์ลำดับที่ i

2.4 กลศาสตร์ของข้อไหล่ (Shoulder Kinematics)

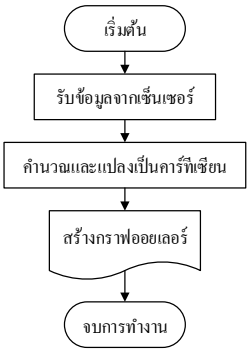
การเคลื่อนไหวของกระดูกสะบักในระนาบส่วนหน้า ได้แก่ การยกขึ้นของกระดูกสะบัก การลดต่ำของกระดูกสะบัก การกางของกระดูกสะบัก การหุบเข้าของกระดูกสะบัก และการหมุนลงของกระดูกสะบัก การเคลื่อนไหวของกระดูกสะบักในระนาบแกน x ได้แก่ การหมุนเข้าในของกระดูกสะบัก การหมุนออกของกระดูกสะบัก การเคลื่อนไหวในระนาบด้านข้าง ทางผู้วิจัยได้ เพื่อศึกษาการเคลื่อนไหวของสะบัก

3. ขั้นตอนการวิจัย

ทำการเก็บข้อมูลผู้ป่วยเพื่อตรวจวัดการเคลื่อนไหวกระดูกสะบักจำนวน 12 คน โดยจำแนกเป็นคนปกติ 6 คนและคนผิดปกติ 6 คน จากการตรวจอาการบาดเจ็บของแพทย์

3.1 ขั้นตอนวิธีการสร้างกราฟออยเลอร์

เมื่อได้รับข้อมูลจากอุปกรณ์ตรวจจับการเคลื่อนไหวของสะบักแล้ว ผู้วิจัยนำข้อมูลที่ได้มาทำการแปลงในสมการพิกัดเชิงขั้วและคาร์ทีเซียนโดยใช้สมการที่ (1) และ (2) เพื่อนำข้อมูลที่ได้ไปสร้างกราฟแบบออยเลอร์

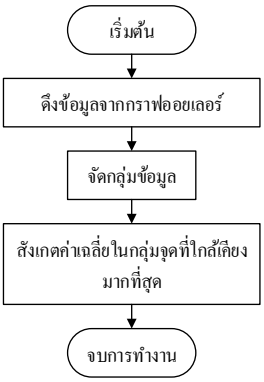


รูปที่ 2 ขั้นตอนการสร้างกราฟออยเลอร์

จากรูปที่ 2 จากการนำกราฟมาจัดกลุ่มข้อมูลจะเห็นได้ว่าข้อมูลดังกล่าวนั้นไม่สามารถแยกได้และสังเกตได้ยากว่ากราฟไหนปกติหรือผิดปกติ

3.2 วิธีการจัดกลุ่มข้อมูลโดยการจัดกลุ่มแบบเคมีน

ผู้วิจัยได้นำข้อมูลที่ได้จากกราฟออยเลอร์ที่ยังไม่สามารถจำแนกข้อมูลของผู้บาดเจ็บได้อย่างชัดเจน ผู้วิจัยได้ทำการจัดกลุ่มข้อมูล เพื่อต้องการดูข้อมูลแต่ละกลุ่มว่ามีความแตกต่างอย่างไรบ้างโดยใช้วิธีการจัดกลุ่มแบบเคมีนและกำหนดค่าในการจัดกลุ่มแต่ล่ะกลุ่มย่อยตั้งแต่ k=3, 4, 5 และ 6 ในแต่ละชุดข้อมูล

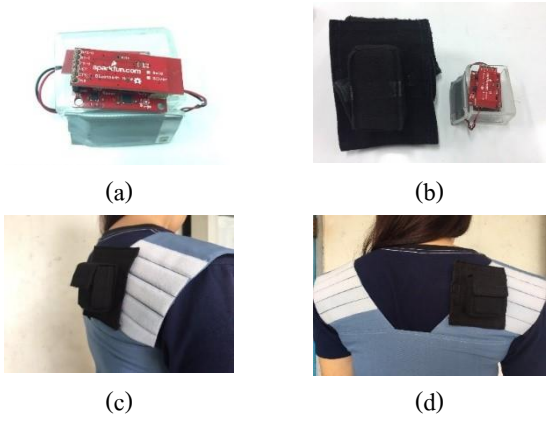


รูปที่ 3 ขั้นตอนการจัดกลุ่มข้อมูลในรูปแบบเคมีน

จากรูปที่ 3 เนื่องจากข้อมูลแบบกราฟเวนน์ออยเลอร์ไม่สามารถรู้จุดที่มีความหนาแน่นของข้อมูลที่ชัดเจน แต่สามารถประเมินความหนาแน่นได้จากสายตาจึงนำกราฟทั้งหมดมาจำแนกกลุ่มข้อมูลเพื่อกำหนดข้อมูลกลุ่มย่อย สำหรับนำไปใช้หาความหนาแน่นเฉพาะกลุ่มที่ผู้วิจัยสนใจ

3.3 อุปกรณ์ตรวจจับการเคลื่อนไหวของสะบัก

อุปกรณ์ตรวจจับการเคลื่อนไหวที่ใช้ในงานวิจัยนี้คือ Razor IMU เป็นเซ็นเซอร์ที่ ภายในวงจรประกอบด้วย 3 ส่วน คือ Gyroscope ทำหน้าที่วัดค่าการเอียงในระนาบ 3 มิติ Accelerometer ทำหน้าที่วัดแรงที่มากระทำกับเซ็นเซอร์และ Magnetometer ทำหน้าที่วัดค่าสนามแม่เหล็กในระนาบ 3 มิติ โดยมีตัวส่งสัญญาณบลูทูธทำหน้าที่ส่งข้อมูลไปยังคอมพิวเตอร์ อุปกรณ์ที่ใช้ในงานวิจัยแสดงดังรูป 4(a) และ 4(b) เป็นอุปกรณ์ตรวจจับการเคลื่อนไหว เสื้อที่ใช้สำหรับติดเซ็นเซอร์ ประกอบกันแสดงในรูป 4(c) และ4(d) เมื่อประกอบอุปกรณ์ตรวจจับการเคลื่อนไหวแล้ว จึงนำมาติดด้านหลังของผู้ทดลองดังรูป 4(c) และ 4(d)



(a) (b)

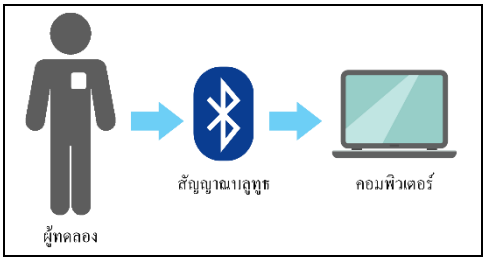
(c) (d)

รูปที่ 4 อุปกรณ์ตรวจจับการเคลื่อนไหวและการติดตั้งอุปกรณ์

3.4 ขั้นตอนการเก็บข้อมูลการเคลื่อนไหวสะบัก

ผู้วิจัยได้ติดอุปกรณ์ตรวจจับการเคลื่อนไหวที่สะบักด้านหลังให้กับผู้ทดลอง โดยอุปกรณ์จะส่งข้อมูลการเคลื่อนไหวของสะบักผ่านสัญญาณบลูทูธไปที่เครื่องคอมพิวเตอร์ ข้อมูลที่ได้จากเซ็นเซอร์จะมีค่า Gyroscope คือค่า x, y และ z และข้อมูลที่ได้จากเซ็นเซอร์ Accelerometer คือค่า yaw, pitch และ roll จากนั้นผู้วิจัยจึงนำข้อมูลที่ได้มาทำการแปลงค่ามุมด้วยสมการพิกัดเชิงขั้ว และนำค่ามุมที่ได้มาสร้างกราฟออยเลอร์ ข้อมูลที่สร้างกราฟดังกล่าวทำให้ลักษณะข้อมูล

ที่ออกมามีการกระจายตัวและทำให้สังเกตยาก ดังนั้นจึงได้มีการนำเอาข้อมูลที่ได้มาทำการจัดกลุ่มของข้อมูลแบบเคมีนเพื่อง่ายในการสังเกตลักษณะของข้อมูลและทราบชุดข้อมูลคนปกติและผิดปกติดังรูปที่ 5



รูปที่ 5 ขั้นตอนการเก็บข้อมูลการเคลื่อนไหวสะบัก

4. วิธีการทดลอง

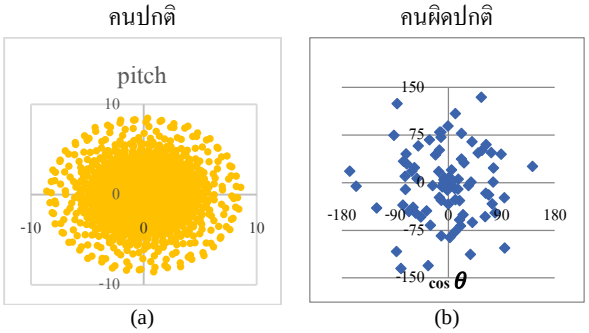
4.1 ขั้นตอนผลการทดลอง

ผู้วิจัยได้ทำการติดอุปกรณ์ตรวจจับการเคลื่อนไหวในรูปที่ 3(c) และ 3(d) และให้ผู้ทดลองเคลื่อนไหวแขนโดยตัวอุปกรณ์จะส่งข้อมูลการเคลื่อนไหวของสะบักผ่านสัญญาณบลูทูธไปยังเครื่องคอมพิวเตอร์จากนั้นผู้วิจัยนำข้อมูลที่ได้มาแปลงด้วยสมการพิกัดเชิงขั้วเพื่อสร้างกราฟออยเลอร์จากนั้นนำข้อมูลดังกล่าวมาจัดกลุ่มย่อยแบบเคมีน

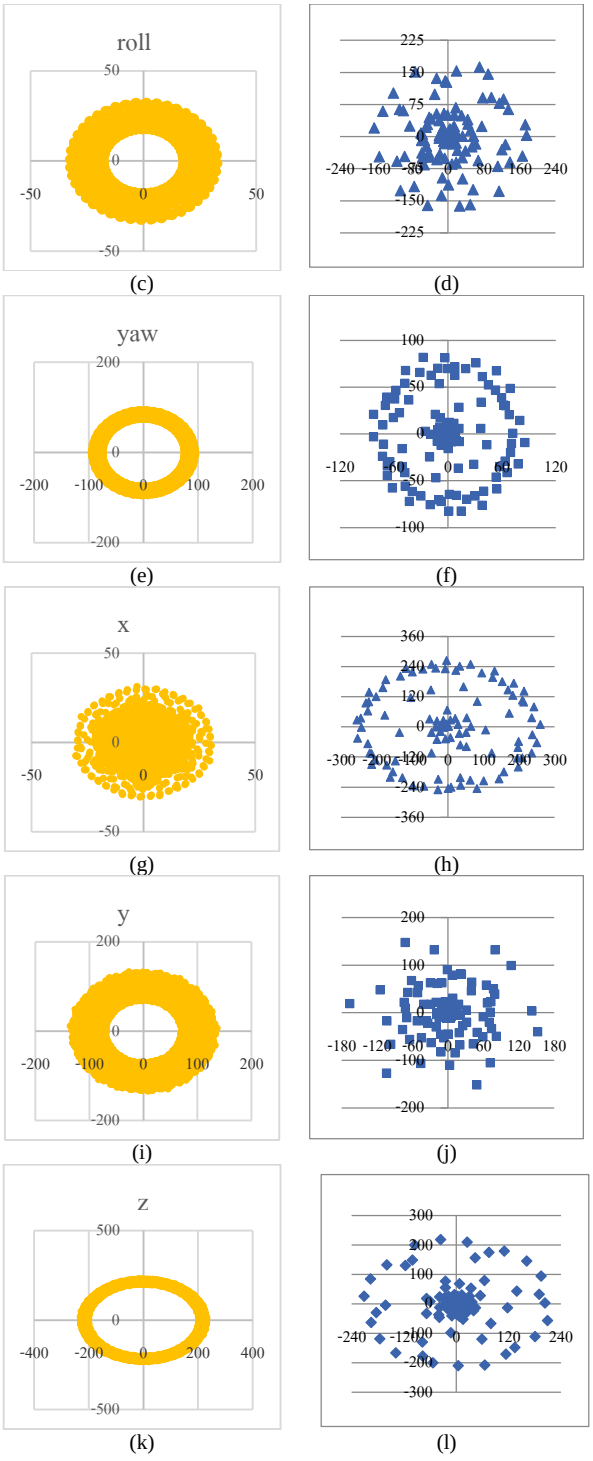
4.2 ผลการทดลอง

4.2.1 ผลการทดลองกราฟแบบออยเลอร์

งานวิจัยนี้ได้ทำการทดลองจากผู้ทดลองจำนวน 12 คนโดยจะแบ่งออกเป็นคนปกติ 6 คนในรูปที่ 6(a), 6(c), 6(e), 6(g), 6(i) และ 6(k) และคนผิดปกติ 6 คนในรูปที่ 6(b), 6(d), 6(f), 6(h), 6(j) และ 6(l) การทดลองนี้จะนำค่าข้อมูลที่ได้จากการแปลงด้วยสมการเชิงขั้วมาทดลองเพื่อสร้างกราฟออยเลอร์



(a) (b)



(c) (d) (e) (f) (g) (h) (i) (j) (k) (l)

รูปที่ 6 ผลการทดลองกราฟออยเลอร์คนปกติและผิดปกติ

จะเห็นว่ากราฟของคนปกติในรูปที่ 6(a), 6(c), 6(e), 6(g), 6(i) และ 6(k) มีลักษณะของข้อมูลกระจายตัวออกจากกลุ่มบางส่วน ในรูปที่6(a) และ 6(g) และคล้ายกับโดนัทในรูป 6(c), 6(e), 6(i) และ 6(k) ส่วนข้อมูลคนผิดปกติรูปที่ 6(b), 6(d), 6(f), 6(h), 6(j) และ 6(l) มีลักษณะของข้อมูลจะมีการกระจายตัวไม่

สม่ำเสมอ และบริเวณตรงกลางมีความหนาแน่นของข้อมูลมากหรือน้อยตามจำนวนข้อมูล

4.2.2 ผลการทดลองการจัดกลุ่มข้อมูลเคมีน

ข้อมูลส่วนนี้ได้จากการนำกราฟในหัวข้อที่ 4.2.1 นำมาจัดกลุ่มข้อมูลแบบเคมีนเนื่องจากข้อมูลแบบกราฟเวนน์ออยเลอร์ไม่สามารถรู้จุดที่มีความหนาแน่นของข้อมูลที่ชัดเจนแต่สามารถประเมินความหนาแน่นได้จากสายตา และข้อมูลบางกลุ่มมีความคล้ายคลึงกัน จึงนำกราฟทั้งหมดมาจัดกลุ่มในรูปแบบเคมีนเพื่อกำหนดข้อมูลกลุ่มย่อย สำหรับนำไปใช้หาความหนาแน่นเฉพาะกลุ่มที่ผู้วิจัยสนใจ ในการทดลองจะแสดงค่าเฉพาะในแกน pitch เนื่องจากเป็นแกนที่สามารถเห็นความแตกต่างได้ชัดเจนที่สุดจากการสร้างกราฟออยเลอร์ ในการจัดกลุ่มผู้วิจัยได้กำหนดค่ากลุ่มข้อมูล k=3, 4, 5 และ 6 โดยเป็นค่าที่ผู้วิจัยทำการสุ่มค่าขึ้นมา จากผลการทดลองผู้วิจัยได้ทำการเลือกค่ากลุ่มข้อมูลคือ k=4 เนื่องจากผู้วิจัยได้ทดลองแล้วว่าเมื่อใช้ค่านี้สามารถได้ผลลัพธ์ที่ดีที่สุด สามารถเห็นความหนาแน่นของข้อมูลในกลุ่มย่อยชัดเจน ข้อมูลมีแนวโน้มไปในทางเดียวกัน ข้อมูลส่วนนี้ได้จากการนำผลการเคลื่อนไหวสะบักจากผู้ทดสอบ 12 คน โดยแบ่งเป็นคนปกติจำนวน 6 คน คนผิดปกติจำนวน 6 คน

โดยมีผลของแบบคนปกติแสดงในรูปที่ 8 และแบบคนผิดปกติแสดงในหัวข้อที่ 7

5. สรุปผลการทดลอง

งานวิจัยนี้เป็นการทดลองเบื้องต้นเพื่อให้ทราบการจัดกลุ่มข้อมูลของคนปกติและผิดปกติโดยใช้การจัดกลุ่มแบบเคมีนในการจำแนกข้อมูล จากการทดลองจะเห็นว่าผู้วิจัยได้ทดลองแล้วว่าเมื่อใช้ค่า k เท่ากับ 4 สามารถได้ผลลัพธ์ที่ดีที่สุด สามารถเห็นข้อมูลในกลุ่มย่อยชัดเจนข้อมูลอยู่ใกล้เคียงกันที่สุด และสามารถนำข้อมูลทั้งของคนปกติและผิดปกติมาช่วยแพทย์ในการวิเคราะห์อาการบาดเจ็บได้ การนำเสนอนี้จึงเป็นเพียงการแยกแยะผลลัพธ์ระหว่างคนปกติและคนผิดปกติเท่านั้น

เอกสารอ้างอิง

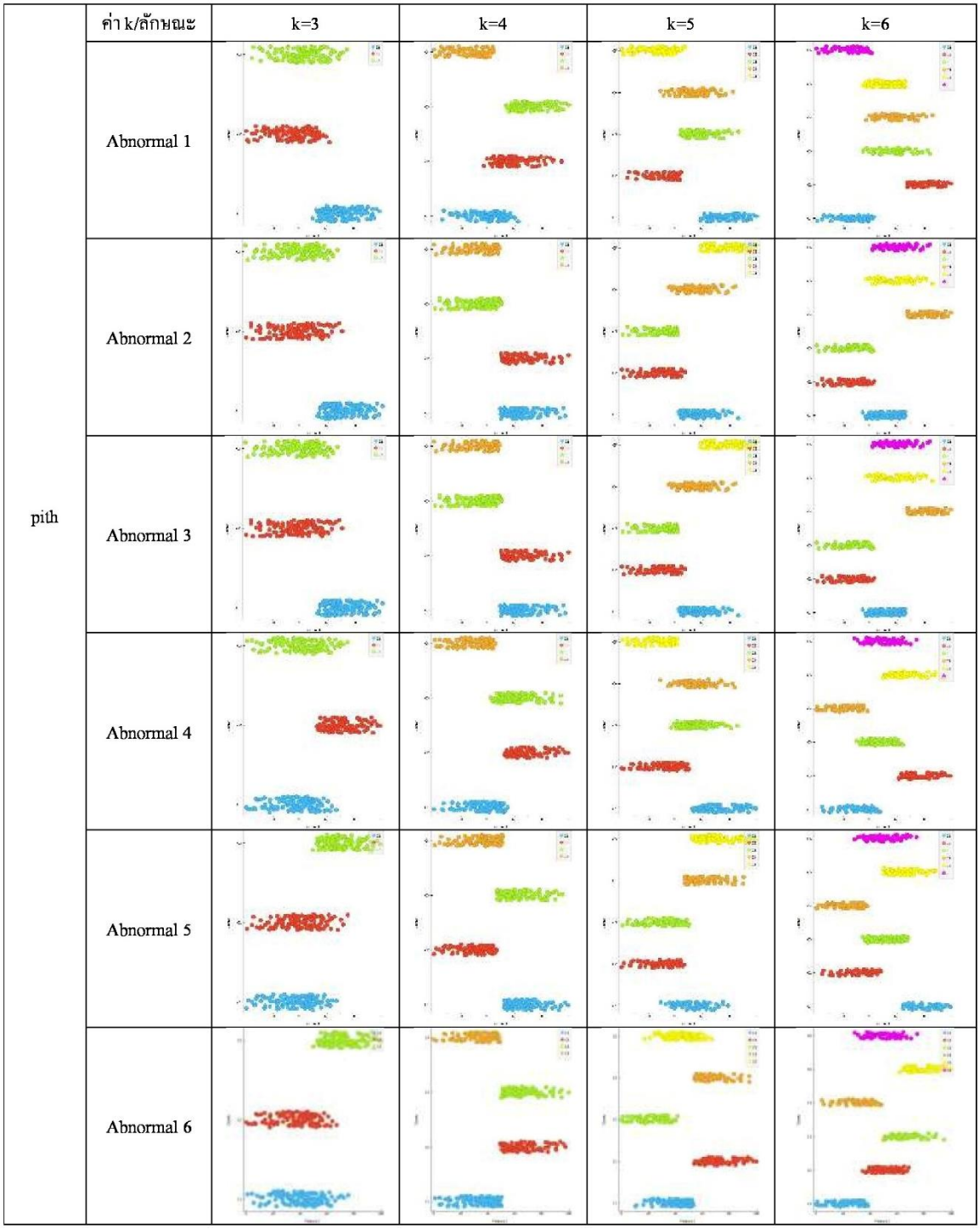
[1] V. A. Carolien, V. H. Kim, E. Marielle, V. DirkJan and H. Jaap, “Recording scapular motion using an acromion marker cluster,” Gait & Posture, vol. 29, pp. 123-128, January 2009.

[2] นที เหรียญตระกูลชัย, วิสันต์ ตั้งวงษ์เจริญ และคณะ.การประยุกต์ใช้เซ็นต์เซอร์ตรวจจับการเคลื่อนไหวร่วมกับแท็บเล็ตสำหรับการตรวจวัดการเคลื่อนไหวของกระดูกสะบักPhanitchart,

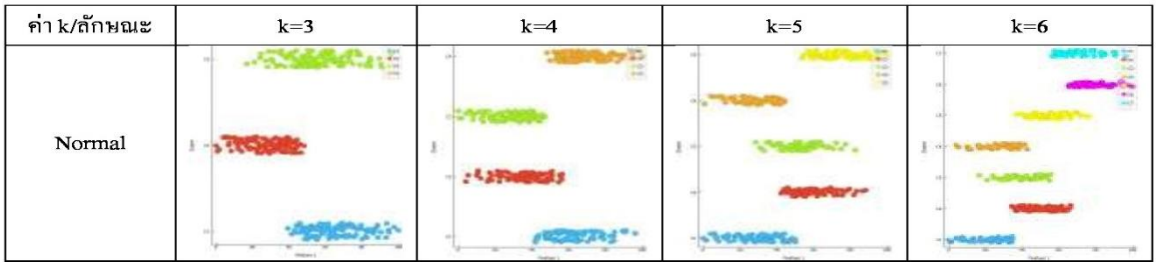
[3] J., &Kimpan, W.(2015). Phase Graphs Applying Used in Analyzing Scapular Patients. Chonburi, Thailand: The 7th conference of Electrical Engineering Network of Rajamangala University of Technology, 137-140.

[4] วารสารวิชาการพระจอมเกล้าพระนครเหนือ ปีที่ 25 ฉบับที่ 3 ก.ย.–ธ.ค. 2558. The Journal of KMUTNB., Vol. 25, No. 3, Sep.–Dec. 2015

[5] C. F. Tsai, C. T. Tsai, C. S. Hung and P. S. Hwang. “Data Mining Techniques for Identifying Students at Risk of Failing a Computer Proficiency Test Required for Graduation.” Australasian Journal of Educational Technology, Vol. 27, No. 3, pp. 481–498, 2011.



รูปที่ 7 ผลการทดลองกราฟเคมีนคนผิดปกติ



รูปที่ 8 ผลการทดลองกราฟเคมีนคนปกติ

การวิเคราะห์ทวิตเตอร์ในเหตุการณ์ สวรรคตของพระบาทสมเด็จพระปรมินทรมหาภูมิพลอดุลยเดช

An Analysis of Twitter in the Passing of His Majesty King Bhumibol Adulyadej

อารยา พุดตาล (Araya Pudtal)[1] และสุกรี สินธุภิญโญ (Sukree Sinthupinyo)[2]
ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
[1]Araya.Pu@student.chula.ac.th, [2]Sukree.S@chula.ac.th

บทคัดย่อ

งานวิจัยนี้ต้องการเก็บรวบรวมและวิเคราะห์เหตุการณ์การสวรรคตของพระบาทสมเด็จพระปรมินทรมหาภูมิพลอดุลยเดชที่เกิดขึ้นในทวิตเตอร์ โดยผ่าน Twitter Advanced Search (https://twitter.com/search-advanced) ร่วมกับไลบรารี Selenium ของภาษาไพทอน โดยค้นหาจากคำสำคัญและแฮชแท็ก เพื่อนำมาวิเคราะห์แนวโน้มและอิทธิพลของการกระจายข้อมูลของทวิตเตอร์ โดยพิจารณาจากจำนวนรีทวีต คำ และแฮชแท็ก พร้อมทั้งวิเคราะห์ความเข้มข้นของการรีทวีต และจัดกลุ่มของทวีตด้วยวิธีการจัดกลุ่มแบบ K-means โดยพิจารณาจากคำ จำนวนของคำ และเนื้อหา เพื่อค้นหาเรื่องราวหรือสิ่งที่ถูกกล่าวถึง ซึ่งแสดงผลไว้ในบทความ และนอกจากนี้ยังพบว่า การกระจายข้อมูลของทวีตที่ถูกรีทวีตเป็นแบบ power-law distribution ด้วย exponent a เท่ากับ 0.64

คำสำคัญ: ทวิตเตอร์ อิทธิพล รีทวีต แฮชแท็ก ความเข้มข้น

Abstract

Our research aims to collect and analyze information from twitter about the passing of His Majesty King Bhumibol Adulyadej by Twitter Advanced Search (https://twitter.com/search-advanced) and Selenium package in Python language. We hence searched by keywords and hashtags in order to analyze trends and influence distribution of information based on the number of retweets, words and hashtags. We analyzed intensity of retweets and cluster of tweets using K-means Clustering based on words, the number of words and contents in order to extract the stories, shown in the same group of tweets. Finally, we could find that the distribution of retweets is power-law with exponent a = 0.64.

Keyword: Twitter, Influence, Retweet, Hashtag, Intensity

1. บทนำ

ทวิตเตอร์ถือเป็นหนึ่งในสื่อสังคมออนไลน์ที่ผู้ใช้สามารถโพสต์ข้อความ รูป หรือลิงก์ต่างๆ ได้ไม่เกิน 140 ตัวอักษร อีกทั้งยังสามารถค้นหาข้อมูลที่น่าสนใจหรือแนวโน้มของเหตุการณ์ที่เกิดขึ้นในช่วงเวลานั้นๆ ได้ เช่น การค้นหาเหตุการณ์ทางธรรมชาติแบบเรียลไทม์ [1] ซึ่งข้อความในทวิตเตอร์จะสะท้อนให้เห็นถึงความหลากหลายของเหตุการณ์ที่จะเกิดขึ้น รวมถึงมุมมองของผู้ใช้ [2] ทวิตเตอร์ไม่เพียงแต่เป็นแหล่งรวบรวมข่าวสาร (Breaking News) แต่ยังสามารถโพสต์แสดงความคิดเห็น และบรรยายเรื่องราวต่างๆ ได้อีกด้วย จากลักษณะเด่นของทวิตเตอร์ที่มีรูปแบบข้อความสั้นๆ และเรียลไทม์ จึงต้องใช้แฮชแท็ก (#) ที่ผู้ใช้สามารถสร้างได้อย่างอิสระ หรือคำสำคัญแทนการบอกเล่าเรื่องราว ทำให้ผู้ใช้สามารถเข้าถึงได้ง่ายและรวดเร็วมากขึ้น โดยจำนวนการโพสต์หรือการถูกรีทวีตจะขึ้นอยู่กับลักษณะของเนื้อหา URLs และแฮชแท็ก [3] เมื่อมีจำนวนเพิ่มมากขึ้นอาจแสดงให้เห็นถึงการปรากฏขึ้นของสิ่งที่น่าสนใจของเหตุการณ์ (Trending Topics) นั้น

ในปี พ.ศ. 2559 ประเทศไทยได้เกิดเหตุการณ์การสูญเสียครั้งใหญ่ ทำให้คนไทยทั่วทั้งประเทศต่างร่ำไห้ เนื่องจากเกิดเหตุการณ์การสวรรคตของพระบาทสมเด็จพระปรมินทรมหาภูมิพลอดุลยเดชในวันที่ 13 ตุลาคม 2559 ซึ่งเป็นเหตุการณ์ที่มีความสำคัญและมีอิทธิพลต่อคนไทยอย่างมาก ผู้วิจัยจึงสนใจที่จะวิเคราะห์แนวโน้มและอิทธิพลของการกระจายข้อมูลของทวิตเตอร์ โดยพิจารณาจากจำนวนรีทวีต คำ และแฮชแท็ก แต่เนื่องจากแฮชแท็กที่เกิดขึ้นนั้นมีความหลากหลายและจำนวน

มาก ผู้วิจัยจึงเลือกเก็บข้อมูลผ่าน Twitter Advanced Search (https://twitter.com/search-advanced) ที่สามารถค้นหาได้ทั้งจากคำและแฮชแท็ก แต่ผลการค้นหาจะแสดงครั้งละประมาณ 20 ทวีต จึงต้องใช้เครื่องมือในการควบคุมหน้าเว็บและอ่านข้อมูลไปพร้อมๆ กัน ในงานวิจัยนี้ ผู้วิจัยเลือกใช้ไลบรารี Selenium ของภาษาไพทอน โดยค้นหาจากคำค้นและแฮชแท็กที่เกี่ยวข้องกับในหลวง รัชกาลที่ 9 ตั้งแต่วันที่ 1 ตุลาคม 2559 – 31 ธันวาคม 2559 พบว่ามี 180,574 ทวีต และ 12,749 แฮชแท็ก ซึ่งในทวีตทั้งหมดนี้ถูกรีทวีตทั้งสิ้น 33,747,331 ครั้ง โดยผู้วิจัยเริ่มวิเคราะห์จากการกระจายข้อมูลของทวีตที่ถูกรีทวีต แฮชแท็ก และคำที่เกิดขึ้น ซึ่งการตัดคำภาษาไทยนั้นจะใช้ไลบรารี PyICU ของภาษาไพทอน และทำการวิเคราะห์ความเข้มข้นของการรีทวีต โดยพิจารณาจากแฮชแท็กและคำที่เกิดขึ้น หลังจากนั้นทำการจัดกลุ่มของทวีต โดยใช้เทคนิค K-means และใช้วิธี Elbow Method ในการกำหนดจำนวนกลุ่ม (k) ที่เหมาะสม เพื่อค้นหาเรื่องราวหรือสิ่งที่ถูกกล่าวถึง

2. งานวิจัยที่เกี่ยวข้อง

จากลักษณะเด่นของทวิตเตอร์ที่เป็นแบบเรียลไทม์ จึงทำให้มีผู้สนใจเก็บรวบรวมข้อมูลทวิตเตอร์ผ่าน Twitter Search API, Twitter API หรือ Twitter developer API เพื่อนำมาวิเคราะห์โครงสร้างและอิทธิพลของการกระจายข้อมูลของทวิตเตอร์ [4] วิเคราะห์รูปแบบการเผยแพร่ข้อมูลบนทวิตเตอร์โดยพิจารณาจากการแชร์หรือการโพสต์ข้อความในหัวข้อที่น่าสนใจ [5] อีกทั้งยังมีการวิเคราะห์ปรากฏการณ์ต่างๆ ที่เกิดขึ้นในทวิตเตอร์โดยพิจารณาจากลักษณะหรือโครงสร้างทางภูมิศาสตร์ [6] และนอกจากนี้ยังมีการวิเคราะห์ข้อมูลทวิตเตอร์จากเหตุการณ์สำคัญต่างๆ [4] เช่น การวิเคราะห์การเกิดแผ่นดินไหวในประเทศญี่ปุ่นโดยพิจารณาจากคำ จำนวนของคำ และเนื้อหาของทวีต เพื่อนำมาสร้างโมเดลเพื่อใช้ในการค้นหาตำแหน่งที่เกิดเหตุการณ์และทำการแจ้งเตือนไปยังผู้ที่ลงทะเบียนไว้ [1] รวมถึงมีการวิเคราะห์พฤติกรรมของผู้ใช้ในระหว่างที่เกิดเหตุการณ์วิกฤตโดยพิจารณาจากการเผยแพร่ข้อมูลผ่านทางเครือข่ายทวิตเตอร์และประเมินความน่าเชื่อถือของแหล่งที่มา [7] อีกทั้งยังมีการวิเคราะห์พฤติกรรมการใช้งานของผู้ใช้ทวิตเตอร์หลังเกิดแผ่นดินไหว Great East Japan ในประเทศญี่ปุ่น ซึ่งพบว่าจำนวนรีทวีตจะเพิ่มสูงขึ้นหลังจากเกิดแผ่นดินไหวทันที [8] และเนื่องจากในประเทศไทยมักมีแนวโน้มที่จะเกิดเหตุการณ์น้ำท่วมในช่วงฤดูฝน จึงทำให้จำนวนทวีตเพิ่มสูงขึ้น โดยตั้งแต่เดือนกันยายนถึงตุลาคม 2554 มีจำนวนทวีตเพิ่มขึ้นถึง 52% และเพิ่มขึ้นอย่างต่อเนื่องจนมีจำนวนมากที่สุดในเดือนพฤศจิกายน 2554 [9] จึงมีงานวิจัยที่ทำการวิเคราะห์และจัดกลุ่มประเภทของทวีตที่เกิดขึ้น และวิเคราะห์อิทธิพลของผู้ใช้ในระหว่างที่เกิดเหตุการณ์นั้น โดยพิจารณาจากจำนวนทวีต จำนวนรีทวีต และผู้ติดตาม แล้วทำการเปรียบเทียบเพื่อหาความน่าเชื่อถือของข้อมูล เนื่องจากข้อมูลเหล่านั้นอาจจะสามารถนำมาใช้เพื่อแจ้งเตือนได้ [10] อีกทั้งยังมีการวิเคราะห์ความวิตกกังวลจากเหตุการณ์การเกิดแผ่นดินไหว Tohuku เหตุการณ์การเกิดสึนามิ และเหตุการณ์การเกิดนิวเคลียร์ เพื่อหาความสัมพันธ์และอิทธิพลของข้อมูลทวิตเตอร์อีกด้วย [11]

3. วิธีการดำเนินงานวิจัย

3.1 การจัดเก็บข้อมูล

Twitter Search API เป็นเครื่องมือหนึ่งที่ใช้ในการดึงข้อมูลผ่านทวิตเตอร์ แต่มีข้อจำกัด คือ ไม่สามารถดึงข้อมูลย้อนหลังเกินกว่า 7 วันได้ และยังมีขีดจำกัดอัตราการใช้งาน (Rate Limits) ของการค้นหาข้อมูลที่มีความซับซ้อน [12] ผู้วิจัยจึงสนใจที่จะใช้วิธีการดึงข้อมูลทวิตเตอร์ผ่าน Twitter Advanced Search (https://twitter.com/search-advanced) ซึ่งสามารถค้นหาข้อมูลได้หลายประเภท เช่น กลุ่มคำ บุคคล สถานที่ ช่วงของวันที่ และข้อความเชิงบวก เชิงลบ รวมถึงข้อความที่ถูกรีทวีต แต่จากการค้นหาดังกล่าว หน้าเว็บจะแสดงครั้งละประมาณ 20 ทวีต เมื่อเลื่อนลงไปจนสุดถึงจะได้รับทวีตถัดไป ผู้วิจัยจึงจำเป็นต้องใช้เครื่องมือในการควบคุมหน้าเว็บและอ่านข้อมูลทวิตเตอร์ไปพร้อมๆ กัน จึงเลือกใช้ไลบรารี Selenium ของภาษาไพทอน โดยเริ่มค้นหาจากคำค้นที่เกี่ยวข้องกับในหลวง คือคำว่า “ในหลวง” และ “ภูมิพล” ตั้งแต่วันที่ 1 ตุลาคม 2559 – 31 ธันวาคม 2559 จากนั้นจึงใช้การค้นหาตามแฮชแท็กที่เกี่ยวข้องกับในหลวงและเป็นแฮชแท็กที่ปรากฏมากที่สุด 10 อันดับแรก หลังจากนั้นจะทำการค้นหาตามแฮชแท็กที่ปรากฏมากที่สุดต่อไปจนกว่าแฮชแท็กนั้นจะไม่เปลี่ยนแปลง ซึ่งในการเก็บข้อมูลครั้งนี้พบว่าเมื่อการค้นหาผ่านไป 3 รอบ ลำดับของแฮชแท็กไม่มีการเปลี่ยนแปลงจึงหยุดค้นหา และในการค้นหา 1 รอบใช้เวลาประมาณ 4 วัน ผู้วิจัยพบว่ามีจำนวนทวีตที่

เกิดขึ้นถึง 180,574 ทวีต และ 12,749 แฮชแท็ก โดยมีแฮชแท็กที่ใช้ในการค้นหาทั้งหมด ดังตารางที่ 1

ตารางที่ 1: แฮชแท็กที่ใช้ในการค้นหาทั้งหมด ในระหว่างวันที่ 1 ตุลาคม 2559 – 31 ธันวาคม 2559

แฮชแท็ก	จำนวน
#เรารักในหลวง	10,637,343
#KingBhumibol	10,395,947
#ขอเป็นข้ารองพระบาททุกชาติไป	6,446,257
#LongLiveTheKing	6,107,321
#ทรงพระเจริญ	2,947,917
#ขอน้อมส่งเสด็จสู่สวรรคาลัย	1,290,406
#kingbhumibol	968,648
#ฉันเกิดในรัชกาลที่๙	756,639
#ขอเป็นข้ารองบาททุกชาติไป	738,952
#สรรเสริญพระบารมี	680,010

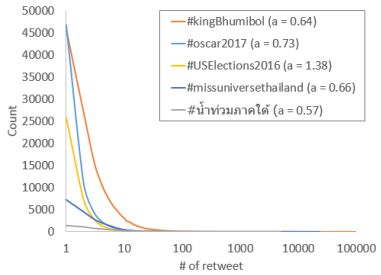
แฮชแท็ก	จำนวน
#KingBhumipol	647,943
#ขอพระองค์เสด็จสู่สวรรคาลัย	623,478
#ฉันเกิดในรัชกาลที่9	475,029
#ทรงพระเจริญยิ่งยืนนาน	416,041
#รัชกาลที่10	383,130
#ในหลวง	348,446
#น้อมส่งเสด็จสู่สวรรคาลัย	275,586
#longlivetheking	274,183
#เราเกิดในรัชกาลที่๙เดอะซีรีส์	202,657
#kingofthailand	187,983

3.2 การตัดคำภาษาไทย

เนื่องจากภาษาไทยเป็นภาษาที่มีความซับซ้อน และไม่มีการเว้นวรรคตอนเหมือนกับภาษาอื่นๆ อีกทั้งยังมีรูปแบบของคำที่หลากหลายและเฉพาะเจาะจง เช่น คำราชาศัพท์ ทำให้ยากต่อการตัดคำ ผู้วิจัยจึงเลือกใช้ไลบรารี PyICU ของภาษาไพทอนในการตัดคำภาษาไทย ซึ่ง PyICU [13] เป็น API สำหรับใช้เชื่อมต่อเข้ากับ ICU C++ API ของ IBM ที่ทำงานเกี่ยวกับ Localization & Internationalization และรองรับภาษาไทย แต่เนื่องจากข้อมูลที่ผู้วิจัยนำมาวิเคราะห์มีคำที่เฉพาะเจาะจงและเป็นคำทับศัพท์ ผู้วิจัยจึงจำเป็นต้องสร้างพจนานุกรมเพิ่มเติมเพื่อรวมคำให้มีความหมายและชัดเจนมากขึ้น และเมื่อต้องการพิจารณาการเกิดขึ้นของคำในแต่ละทวีตจึงจำเป็นต้องกำจัดคำหยุด (Stop Word) หรือคำที่ไม่จำเป็นออก

4. การวิเคราะห์ข้อมูลทวิตเตอร์

4.1 การกระจายข้อมูลทวิตเตอร์



ภาพที่ 1: กราฟเปรียบเทียบจำนวนทวีตที่ถูกรีทวีตของแต่ละแฮชแท็ก

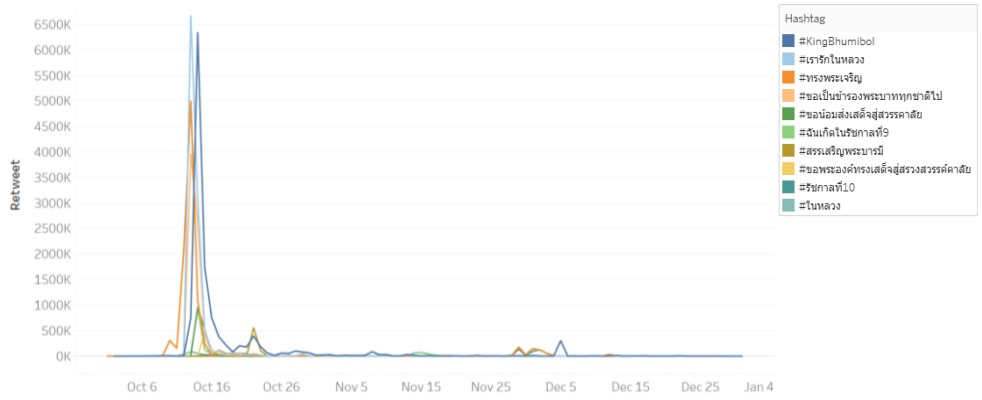
จากภาพที่ 1 แสดงให้เห็นถึงการกระจายข้อมูลของจำนวนทวีตที่ถูกรีทวีต (Retweet Distribution) โดยผู้วิจัยได้ทำการเปรียบเทียบและพิจารณาจากกลุ่มของแฮชแท็ก #KingBhumibol ซึ่งเกิดขึ้นในเดือนตุลาคม 2559 #USElections2016 เกิดขึ้นในเดือนพฤศจิกายน 2559 #น้ำท่วมภาคใต้ เกิดขึ้นในเดือนมกราคม 2560 #missuniversethailand เกิดขึ้นในเดือนมกราคม 2560 และ #oscar2017 เกิดขึ้นในเดือนกุมภาพันธ์ 2560 โดยแกน x คือ จำนวนทวีตที่ถูกรีทวีต และแกน y คือ ความถี่หรือจำนวนครั้ง ซึ่งจะสังเกตเห็นว่า ทวีตที่ถูกรีทวีตน้อยกว่า 20 ครั้ง มีปริมาณสูงกว่าทวีตที่ถูกรีทวีตจำนวนมาก เช่น แฮชแท็ก #KingBhumibol มีทวีตที่ถูกรีทวีตเพียง 1 ครั้ง จำนวนมากถึง 45,864 ทวีต และเมื่อพิจารณาจากกราฟ พบว่าทวิตเตอร์มีการกระจายข้อมูลแบบ power-law distribution โดยข้อมูลจากกลุ่มของแฮชแท็ก #KingBhumibol มี exponent เท่ากับ 0.64

4.2 เหตุการณ์ในช่วงเดือนตุลาคม 2559

เนื่องจากข้อจำกัดของทวิตเตอร์ที่สามารถโพสต์ข้อความได้ไม่เกิน 140 ตัวอักษร จึงจำเป็นต้องใช้แฮชแท็ก (#) ในการแท็กเรื่องราวต่างๆ ที่น่าสนใจ เพื่อช่วยในการรวบรวมเรื่องราวที่เกิดขึ้น และนอกจากนี้แฮชแท็กยังสามารถเป็นตัวช่วยในการคัดกรองหรือแบ่งแยกเหตุการณ์ที่น่าสนใจได้อีกด้วย แต่เนื่องจากแฮชแท็กสามารถสร้างได้อย่างอิสระ จึงส่งผลให้คำที่ใช้มีความหลากหลาย เช่น แฮชแท็ก #KingBhumibol #Rama9 #Kingbhumibhol เมื่อพิจารณาแล้วพบว่าเป็นกลุ่มเดียวกัน ซึ่งหมายถึง #KingBhumibol ผู้วิจัยจึงรวมกลุ่มของแฮชแท็กที่มีความหมายเดียวกันไว้ด้วยกัน จากภาพที่ 2 แสดงให้เห็นว่า ในวันที่ 13 ตุลาคม เป็นวันที่คนไทยทวีตมากที่สุด ด้วยแฮชแท็ก #เรารักในหลวง #ทรงพระเจริญ และ #ขอเป็นข้ารองพระบาททุกชาติไป โดยข้อความที่คนไทยให้ความสนใจและถูกรีทวีตมากที่สุดคือ “ตอนสมเด็จย่าพระประชวร พระองค์มาเล่าให้ลูกๆฟังว่า สมเด็จย่าเคยรับสั่ง ตายแล้วห้ามร้องไห้ เพราะเป็นของธรรมดา คนเราก็ต้องตาย #KingBhumibol” ดังตารางที่ 2

ตารางที่ 2: ข้อความที่ถูกรีทวีตมากที่สุด 10 อันดับแรก

จำนวนรีทวีต	ข้อความ
101,475	ตอนสมเด็จย่าพระประชวร พระองค์มาเล่าให้ลูกๆฟังว่า สมเด็จย่าเคยรับสั่ง ตายแล้วห้ามร้องไห้ เพราะเป็นของธรรมดา คนเราก็ต้องตาย #KingBhumibol
97,143	คลิปหาดูยาก ในหลวงครัสตอบนักข่าวฝรั่งที่ถามท่านเกี่ยวกับการสร้างเขื่อน ดูแล้วประทับใจอย่างสุดซึ้ง cr.เพจองกรณ์เก็บขยะแผ่นดิน #KingBhumibol
93,028	คลิปนี้น่ารักมากๆเหมือนพระองค์จะทรงครัสให้เหล่าราชนิกุลด้านหลังเบาเสียงลูจากพระพักตร์ของพระองค์ ดูจากองค์ภาแม็มปาก #KingBhumibol คำผิดขออภัย
92,103	เปลี่ยนไปแค่กาลเวลา แค่ความ ชอบสองภาพนี้ .. รักยังอยู่(น่า)#KingBhumibol
88,132	สมเด็จพระเจ้าพี่นางเธอไม่ทรงโปรดใส่ชุดของผู้หญิง ในหลวงจึงเสียสละใส่แทน #เรารักในหลวง
86,875	เป็นข่าวพระราชสำนักที่ชอบที่สุดและไม่เคยลืมในหลวงทรงทอดพระเนตและเลือกซื้อสินค้าต่างๆ เป็นการส่วนพระองค์ ด้วยพระพักตร์ที่สดใส #เรารักในหลวง
77,422	ไปเจอมา "พ่อดูเรา บอกให้เราไปถ่ายที่อื่นสำรวมด้วย " อ่านแล้วอึ้มม่ารักกก #ทรงพระเจริญ
74,561	เพื่อนส่งรูปที่ศิริราชมาให้ดูในไลน์ มาร์บพ่อกลับสวรรค์ใช่มั้ยคะ #เรารักในหลวง #ขอเป็นข้ารองพระบาททุกชาติไป
69,500	ภาพยนตร์ประวัติศาสตร์ที่ฝรั่งถ่ายไว้ " ในหลวงของเรา " Cr.M-Musicthai #เรารักในหลวง
69,124	ภาพประวัติศาสตร์ แม่เฒ่าคุ้ม วัย 102 ปี ไปรอเฝ้ารับเสด็จโดยให้หลานอุ้มไส่รถลาก พร้อมดอกบัวสายสีชมพู 3 ดอก ตั้งแต่เช้าจนบ่าย #KingBhumibol



ภาพที่ 2: แฮชแท็กที่เกิดขึ้นในช่วงวันที่ 1 ตุลาคม 2559 – 31 ธันวาคม 2559

และหลังจากวันที่ 13 ตุลาคม มีเรื่องราวต่างๆ เกี่ยวกับพระบาทสมเด็จพระปรมินทรมหาภูมิพลอดุลยเดช มหิตลาธิเบศรรามาธิบดี จักรีนฤบดินทร สยามินทราธิราช บรมนาถบพิตรเกิดขึ้นมากมาย ทำให้เกิดแฮชแท็กต่างๆ เพิ่มขึ้น เช่น วันที่ 14 ตุลาคม เกิดแฮชแท็ก #KingBhumibol วันที่ 22 ตุลาคม เกิดแฮชแท็ก #สรรเสริญพระบารมี และ #KingBhumibol วันที่ 29 พฤศจิกายน และ 2 ธันวาคม เกิดแฮชแท็ก #ทรงพระเจริญ และ #รัชกาลที่10 และวันที่ 5 ธันวาคม เกิดแฮชแท็ก #KingBhumibol

4.3 เหตุการณ์ในวันที่ 13 ตุลาคม 2559

ตั้งแต่เวลา 16 นาฬิกา มีการปรากฏแฮชแท็ก #เรารักในหลวง และ #ทรงพระเจริญ เพิ่มสูงขึ้นเรื่อยๆ จนเมื่อมีการแถลงข่าวอย่างเป็นทางการ เรื่อง พระบาทสมเด็จพระเจ้าอยู่หัวรัชกาลที่ 9 เสด็จสวรรคต ในวันที่ 13 ตุลาคม เวลา 19 นาฬิกา จึงทำให้เกิดแฮชแท็ก #ขอเป็นข้ารองพระบาททุกชาติไป และส่งผลให้มีการใช้แฮชแท็กเพิ่มสูงขึ้นจนถึงจุดสูงสุดในช่วงเวลา 20 นาฬิกา โดยแฮชแท็กที่ปรากฏสูงสุด คือ #ขอเป็นข้ารองพระบาททุกชาติไป #เรารักในหลวง และ #ทรงพระเจริญ ตามลำดับ ดังภาพที่ 3 (ก)

4.4 เหตุการณ์หลังจากวันที่ 13 ตุลาคม 2559

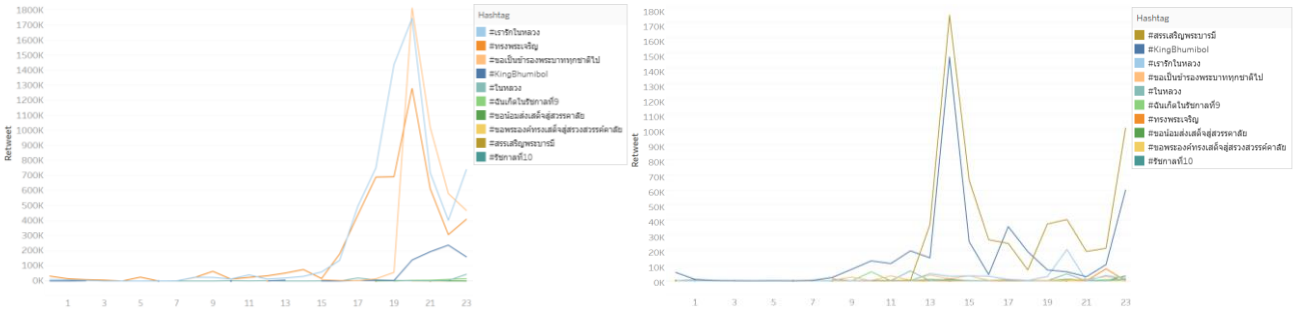
ในวันที่ 22 ตุลาคม มีจุดสูงสุดเกิดขึ้นตั้งแต่เวลา 14 นาฬิกา เป็นต้นไป เนื่องจากเป็นช่วงเวลาที่คนไทยรวมพลังร้องเพลงสรรเสริญพระบารมี ณ ท้องสนามหลวง โดยรอบแรกเกิดขึ้นในเวลา 14 นาฬิกา ทำให้เกิดแฮชแท็ก #สรรเสริญพระบารมี และ #KingBhumibol มากที่สุด ในช่วงที่ 2 เกิดขึ้นในเวลา 17 นาฬิกา ด้วยแฮชแท็ก #KingBhumibol ในช่วงที่ 3 เกิดขึ้นในเวลา 20 นาฬิกา ด้วยแฮชแท็ก #สรรเสริญพระบารมี และ #เรารักในหลวง และในช่วงสุดท้ายเกิดขึ้นในเวลา 23 นาฬิกา ด้วยแฮชแท็ก #สรรเสริญพระบารมี และ #KingBhumibol อีกครั้งหนึ่ง ดังภาพที่ 3 (ข) ซึ่งแสดงให้เห็นว่า ในช่วงเวลา 14 นาฬิกา เป็นช่วงที่ผู้คนให้ความสนใจมากกว่าช่วงเวลาอื่น เนื่องจากเป็นการร้องเพลงสรรเสริญพระบารมีรอบแรกและได้ร้องร่วมกับวงออร์เคสตรา

4.5 ทวีต

เพื่อวิเคราะห์คำต่างๆ ที่เกิดขึ้นในทวีตจากกลุ่มของแฮชแท็กที่เกี่ยวข้องกับในหลวงตั้งแต่วันที่ 1 ตุลาคม 2559 – 31 ธันวาคม 2559 โดยแยกแฮชแท็ก ลิงก์ และรูปภาพออก พบว่าคำที่ปรากฏในทวีตและถูกรีทวีตมากที่สุด คือคำว่า “พ่อ” ตามมาด้วยคำว่า “เกิด” “ท่าน” “ลูก” “พระองค์” และ “รัก” ตามลำดับ ซึ่งปรากฏมากที่สุดในวันที่ 13 ตุลาคม และในวันที่ 22 ตุลาคม คำที่ปรากฏในทวีตและถูกรีทวีตมากที่สุด คือคำว่า “ร้องเพลง” “พ่อ” “สนามหลวง” “คน” และ “ในหลวง” ดังนั้นการปรากฏขึ้นของคำจะแสดงให้เห็นถึงสิ่งที่ถูกกล่าวถึงมากที่สุดในช่วงเวลานั้นๆ ซึ่งสามารถนำมาบอกเล่าเรื่องราวหรือเหตุการณ์ที่เกิดขึ้นได้ เช่น มีกิจกรรมร้องเพลงสรรเสริญพระบารมีที่สนามหลวง ณ วันที่ 22 ตุลาคม เป็นต้น

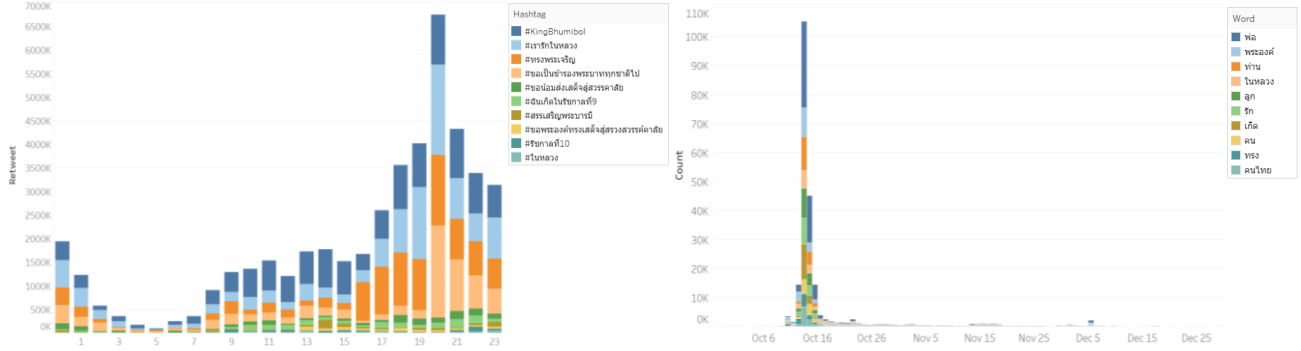
4.6 ความเข้มข้นของการรีทวีต

เมื่อวิเคราะห์ถึงความเข้มข้นของการรีทวีต โดยพิจารณาจากแฮชแท็กที่เกิดขึ้นในแต่ละชั่วโมง พบว่า ตั้งแต่เวลา 16 นาฬิกา ถึง 17 นาฬิกา แฮชแท็กมีปริมาณเพิ่มสูงขึ้นประมาณ 1 ล้านครั้งภายใน 1 ชั่วโมง โดยแฮชแท็กที่มีความเข้มข้นมากที่สุด คือ #ทรงพระเจริญ และในช่วงเวลา 19 นาฬิกา ถึง 20 นาฬิกา ปริมาณแฮชแท็กเพิ่มสูงขึ้นอีกประมาณ 3 ล้านครั้งจนกระทั่งถึงจุดสูงสุด จึงส่งผลให้ในเวลา 20 นาฬิกามีความเข้มข้นมากที่สุด



(ก) (ข)

ภาพที่ 3:แฮชแท็กที่เกิดขึ้นในแต่ละชั่วโมง (ก) ในวันที่ 13 ตุลาคม 2559 (ข) ในวันที่ 22 ตุลาคม 2559



(ก) (ข)

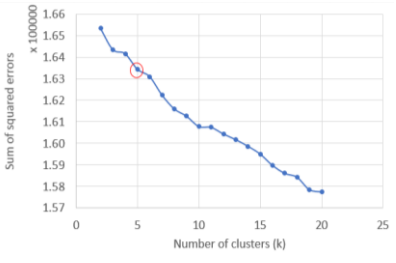
ภาพที่ 4: ความเข้มข้นของการรีทวีต ในระหว่างวันที่ 1 ตุลาคม 2559 – 31 ธันวาคม 2559 โดยพิจารณาจาก (ก) แฮชแท็ก และ (ข) คำในกลุ่มของแฮชแท็กที่เกี่ยวข้องกับในหลวง

ด้วยแฮชแท็ก #ขอเป็นข้ารองพระบาททุกชาติไป #เรารักในหลวง และ #ทรงพระเจริญ ตามลำดับ หลังจากนั้นปริมาณแฮชแท็กจะค่อยๆ ลดลง ทั้งนี้เมื่อทำการเปรียบเทียบข้อมูลจากการวิเคราะห์แฮชแท็ก ดังภาพที่ 2 และการวิเคราะห์ความเข้มข้นของการรีทวีต ดังภาพที่ 4 (ก) พบว่าแฮชแท็กที่ถูกรีทวีตมากที่สุด คือ แฮชแท็ก #เรารักในหลวง แต่แฮชแท็กที่มีความเข้มข้นมากที่สุด คือ แฮชแท็ก #ขอเป็นข้ารองพระบาททุกชาติไป

และเมื่อพิจารณาจากคำในกลุ่มของแฮชแท็กที่เกี่ยวข้องกับในหลวง พบว่า ในวันที่ 13 ตุลาคม 2559 มีคำเกิดขึ้นมากถึง 91,000 ครั้งภายใน 1 วัน จึงส่งผลให้เป็นวันที่มีความเข้มข้นมากที่สุด ด้วยคำว่า “พ่อ” “เกิด” และ “ท่าน” ดังภาพที่ 4 (ข)

5. การจัดกลุ่มของทวีต (Text Clustering)

เพื่อหาความสัมพันธ์หรือความคล้ายคลึงกันของคำที่เกิดขึ้นในแต่ละทวีต และเพื่อพิจารณาถึงสิ่งที่คนไทยกล่าวถึงในช่วงวันที่ 1 ตุลาคม 2559 – 31 ธันวาคม 2559 ผู้วิจัยจึงเลือกใช้เครื่องมือ Weka ในการจัดกลุ่มของคำ และใช้เทคนิค K-means เนื่องจากเป็นวิธีการแบ่งกลุ่มที่ง่ายและรวดเร็วเมื่อใช้กับชุดข้อมูลขนาดใหญ่ ซึ่ง K-means เป็นการเรียนรู้แบบไม่มีผู้สอน (Unsupervised Learning) หลักการทำงาน คือ จะแบ่งข้อมูลออกเป็นกลุ่มย่อยๆ ตามจำนวนกลุ่ม (k) ที่กำหนดหรือสุ่มได้ในแต่ละรอบ เพื่อคำนวณหาระยะห่างระหว่างข้อมูลกับค่ากลาง (Centroid) โดยใช้การวัดระยะห่างแบบยูคลิด (Euclidean distance) และหลังจากนั้นกระทำซ้ำจนกระทั่งค่ากลางไม่เปลี่ยนแปลงหรือเปลี่ยนแปลงน้อยที่สุด [14] ซึ่งในการกำหนดจำนวนกลุ่มหรือค่า k นั้นจะต้องเลือกค่าที่เหมาะสมกับชุดข้อมูล ผู้วิจัยจึงเลือกใช้วิธี Elbow Method ซึ่งพิจารณาจากความแปรปรวนหรือการขึ้นลงของกราฟ โดยจะเลือกค่าที่เริ่มหยุดการเปลี่ยนแปลงหรือมีการเปลี่ยนแปลงเพียงเล็กน้อย [15] ทั้งนี้ผู้วิจัยได้ทำการทดสอบจำนวนค่ากลุ่มหรือ k เท่ากับ 2 และเพิ่มขึ้นทีละ 1 จนกระทั่งถึง 20 กลุ่ม แล้วพิจารณาค่า Sum of Square error (SSE) พบว่าค่ากลุ่ม k ที่เหมาะสม คือ k = 5 เนื่องจากตั้งแต่ 2-4 กลุ่ม ค่า k มีการเปลี่ยนแปลงสูง ดังภาพที่ 5



ภาพที่ 5: ความสัมพันธ์ระหว่างจำนวนกลุ่ม (k) และค่า Sum of Square error (SSE)

เมื่อทำการจัดกลุ่มของคำโดยที่ค่า k = 5 แล้วนั้น ผู้วิจัยจึงนำผลที่ได้มาวิเคราะห์โดยเลือกคำที่มีค่ามากที่สุด 10 อันดับแรกของแต่ละกลุ่ม แล้วนำคำเหล่านั้นไปค้นหาจากทวีตที่เกี่ยวข้องกับในหลวงทั้งหมด เพื่อค้นหาเรื่องราวหรือสิ่งที่ถูกกล่าวถึงในระหว่างวันที่ 1 ตุลาคม 2559 – 31 ธันวาคม 2559 โดยพิจารณาจากทวีตที่มีจำนวนคำในแต่ละกลุ่มปรากฏมากที่สุด และเป็นทวีตที่สั้นที่สุดในกลุ่ม พบว่า ในแต่กลุ่มจะแสดงให้เห็นถึงเหตุการณ์ที่เกิดขึ้นและความรู้สึกที่คนไทยมีต่อพระองค์หลังจากเกิดเหตุการณ์การสวรรคต ซึ่งแสดงให้เห็นถึงความรัก ความทรงจำ ความเศร้าโศกเสียใจ และความภูมิใจที่ได้เกิดในแผ่นดินไทย ดังตารางที่ 4

ตารางที่ 4: คำที่มีค่ามากที่สุด 10 อันดับแรกของแต่ละกลุ่มและตัวอย่างของทวีตที่มีคำในแต่ละกลุ่มปรากฏมากที่สุด

กลุ่ม	คำ	ตัวอย่างทวีต
1	เทวดา สวรรค์ สู่ กลับ บน เสด็จ สวรรคาลัย ท่าน พ่อ สรวงสวรรค์	เทวดาของแผ่นดินไทยเสด็จกลับสู่สรวงสวรรค์แล้วภูมิใจที่ลูกได้เกิดบนแผ่นดินพ่อ #ขอเป็นข้ารองพระบาททุกชาติไป
2	ร้องเพลง สรรเสริญพระบารมี สนามหลวง สรรเสริญ จุดเทียน ชาติ ร้อง 22 ตอน ห้อง	การร้องเพลงสรรเสริญ กำหนดเดิมเวลา14.00 และ 16.00 น . รออีกรอบคือจุดเทียนตอน .. ยกเลิกแล้วนะคะ 22.00 นที่ . ห้องสนามหลวงเลย #สรรเสริญพระบารมี
3	พ่อ รัก ลูก ตลอดไป เหนื่อย ท่าน อยู่ในใจ หนู คนไทย ทุกคน	พ่อเหนื่อยมามากพอแล้ว พักให้สบายนะพ่อต่อไปนี้ไม่มีอีกแล้วพ่อของแผ่นดินไทย แต่ท่านจะยังอยู่ในใจของคนไทยตลอดไป หนู รัก ในหลวง #เรารักในหลวง
4	ในหลวง พระองค์ ทรง มัน คน ท่าน วันนี้ ดู ใจ ร้องไห้	#wiranloveking ในหลวงทรงเปิดไฟในรถให้ทุกคนได้เห็นพระองค์ แล้วโบกมือ ทุกคนร้องไห้ หัวใจเราพองโตมาก ไม่รู้ทำไมเหมือนกัน มันดื้นดันไปหมด (จบ)
5	เกิด ภูมิใจ รัชกาลที่9 ลูก แผ่นดิน คนไทย โชคดี พ่อ ดีใจ บน	ภูมิใจที่เกิดบนแผ่นดินไทย ภูมิใจที่เกิดในรัชกาลที่ 9 ภูมิใจที่ได้เกิดเป็นลูกพ่อ เป็นคนไทย #LongLiveTheKing

6. สรุป

ผู้วิจัยเก็บข้อมูลทวิตเตอร์ผ่าน Twitter Advanced Search (https://twitter.com/search-advanced) ร่วมกับไลบรารี Selenium ของภาษาไพทอน โดยค้นหาจากคำค้นและแฮชแท็กที่เกี่ยวข้องกับในหลวง ตั้งแต่วันที่ 1 ตุลาคม 2559 – 31 ธันวาคม 2559 ได้ทวีตทั้งหมด 180,574 ทวีต และ 12,749 แฮชแท็ก ซึ่งพบว่า การกระจายตัวของทวีตที่ถูกรีทวีตเป็นแบบ power-law distribution โดยทวีตที่ถูกรีทวีตจำนวนน้อย มีปริมาณสูงกว่าทวีตที่ถูกรีทวีตจำนวนมาก และจากการวิเคราะห์ข้อมูลด้วยแฮชแท็กพบว่า ในวันที่ 13 ตุลาคม เป็นวันที่คนไทยทวีตมากที่สุดด้วยแฮชแท็ก #เรารักในหลวง และในเวลา 20 นาฬิกา แฮชแท็ก #ขอเป็นข้ารองพระบาททุกชาติไป ถูกรีทวีตมากที่สุดและยังเป็นแฮชแท็กที่มีความเข้มข้นมากที่สุดด้วย เมื่อพิจารณาจากคำที่เกิดขึ้นในกลุ่มของแฮชแท็กในหลวง พบว่า การปรากฏขึ้นของคำในแต่ละวัน สามารถนำมาวิเคราะห์เพื่อค้นหาเหตุการณ์ที่เกิดขึ้นในแต่ละวันได้ เมื่อทำการจัดกลุ่มของทวีต พบว่า สิ่งที่คนไทยกล่าวถึง คือความรู้สึกและความประทับใจที่คนไทยมีต่อพระองค์

เอกสารอ้างอิง

[1] Sakaki, T., M. Okazaki, and Y. Matsuo, "Earthquake shakes Twitter users: real-time event detection by social sensors," in *Proceedings of the 19th international conference on World wide web*. 2010, ACM: Raleigh, North Carolina, USA. p. 851-860.

[2] Diakopoulos, N., M. Naaman, and F. Kivran-Swaine. "Diamonds in the rough: Social media visual analytics for journalistic inquiry." in *2010 IEEE Symposium on Visual Analytics Science and Technology*. 2010.

[3] Suh, B., et al., "Want to be Retweeted? Large Scale Analytics on Factors Impacting Retweet in Twitter Network," in *Proceedings of the 2010 IEEE Second International Conference on Social Computing*. 2010, IEEE Computer Society. p. 177-184.

[4] Kwak, H., et al., "What is Twitter, a social network or a news media?," in *Proceedings of the 19th international conference on World wide web*. 2010, ACM: Raleigh, North Carolina, USA. p. 591-600.

[5] Zhou, Z., et al., "Information resonance on Twitter: watching Iran," in *Proceedings of the First Workshop on Social Media Analytics*. 2010, ACM: Washington D.C., District of Columbia. p. 123-131.

[6] Java, A., et al., "Why we twitter: understanding microblogging usage and communities," in *Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 workshop on Web mining and social network analysis*. 2007, ACM: San Jose, California. p. 56-65.

[7] Mendoza, M., B. Poblete, and C. Castillo, "Twitter under crisis: can we trust what we RT?," in *Proceedings of the First Workshop on Social Media Analytics*. 2010, ACM: Washington D.C., District of Columbia. p. 71-79.

[8] Miyabe, M., A. Miura, and E. Aramaki, "Use trend analysis of twitter after the great east japan earthquake," in *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work Companion*. 2012, ACM: Seattle, Washington, USA. p. 175-178.

[9] iT24Hrs. "Infographic and statistics of Thai Twitter 2011." 2012; Available from: https://www.it24hrs.com/2012/infographic-thailand-twitter-stats-2011/.

[10] Kongthon, A., et al. "The role of Twitter during a natural disaster: Case study of 2011 Thai Flood." in *2012 Proceedings of PICMET '12: Technology Management for Emerging Technologies*. 2012.

[11] Doan, S., B.-K.H. Vo, and N. Collier, "An Analysis of Twitter Messages in the 2011 Tohoku Earthquake," in *Electronic Healthcare: 4th International Conference, eHealth 2011, Málaga, Spain, November 21-23, 2011, Revised Selected Papers*, P. Kostkova, M. Szomszor, and D. Fowler, Editors. 2012, Springer Berlin Heidelberg: Berlin, Heidelberg. p. 58-66.

[12] Twitter, I. "The Search API." Available from: https://dev.twitter.com/rest/public/search.

[13] Vajda, A. "PyICU 1.9.5." Available from: https://pypi.python.org/pypi/PyICU/.

[14] Liu, S. and Y. Cheng. "Research on K-Means Algorithm Based on Cloud Computing." in *2012 International Conference on Computer Science and Service System*. 2012.

[15] Omran, M.G.H., A.P. Engelbrecht, and A. Salman, "An overview of clustering methods," *Intell. Data Anal.*, 2007. **11**(6): p. 583-605.

ระบบอัจฉริยะสำหรับค้นหาทักษะทางด้านคอมพิวเตอร์
An Intelligent System for Discover Computer Skill

ศิณาภัส จวนงาม (Sinapas Juanngarm)[1] และสรวิชญ์ ยศอาลัย (Sorawit Yotalai)[2]
ภาควิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีและการจัดการอุตสาหกรรม
มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
[1]5606021620033@fitm.kmutnb.ac.th, [2]5606021610054@fitm.kmutnb.ac.th

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาระบบอัจฉริยะสำหรับการค้นหาทักษะทางด้านคอมพิวเตอร์โดยใช้เทคนิคการทำเหมืองข้อมูล เพื่อช่วยให้นักศึกษาสามารถค้นหาทักษะทางด้านคอมพิวเตอร์ได้ด้วยตนเองผ่านเว็บแอปพลิเคชัน และสามารถนำไปใช้ตัดสินใจเพื่อพัฒนาทักษะและความสามารถในด้านคอมพิวเตอร์ได้ถูกต้อง โดยระบบจะทำการจัดกลุ่มทักษะทางด้านคอมพิวเตอร์ด้วยขั้นตอนวิธี k-means หลังจากนั้นในแต่ละกลุ่มจะสร้างกฎความสัมพันธ์ของตัวแปรด้วยขั้นตอนวิธี Apriori สำหรับชุดข้อมูลที่นำมาจากระบบงานทะเบียนนักศึกษาระหว่างปีการศึกษา 2548 – 2557 หลักสูตรเทคโนโลยีบัณฑิต ภาควิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีและการจัดการอุตสาหกรรม มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ วิทยาเขตปราจีนบุรี มีจำนวน 21 แอททริบิวต์ และ 733 ชุดข้อมูล โดยกฎที่ได้จากการทดลองจะถูกนำไปพัฒนาเป็นระบบอัจฉริยะสำหรับค้นหาทักษะทางด้านคอมพิวเตอร์ และนำไปทดสอบกับนักศึกษาระดับชั้นปีที่ 2 ปีการศึกษา 2558 จำนวนทั้งสิ้น 30 คน รวมถึงมีผลการประเมินของผู้เชี่ยวชาญซึ่งมีค่าเฉลี่ยเท่ากับ 4.47 และมีค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.56 ดังนั้น สามารถสรุปได้ว่าระบบที่พัฒนาขึ้นสามารถทำงานได้ตรงกับความต้องการและนำไปประยุกต์ใช้งานได้อย่างมีประสิทธิภาพ

คำสำคัญ: ระบบอัจฉริยะ ทักษะคอมพิวเตอร์ เว็บแอปพลิเคชัน เหมืองข้อมูล

Abstract

The objective of this paper is to develop an Intelligent System for Discover Computer Skill (DiscoverSkill) using the Data Mining technique to help students determine their own computer skills through web applications, and can apply it to make decisions to correctly improve their skills and capabilities in the computer field. The system will manage computer skills into clusters using k-mean algorithm. Then, for each cluster, the dataset used to find the association rule using Apriori algorithm. The dataset is obtained from the Student Registration Office in academic year 2005-2014 of Department of Information Technology of Faculty of Industrial Technology and Management, King Mongkut's University of Technology North Bangkok. The dataset contains 21 attributes and 733 instances. Moreover, the rules obtained can be developed in an intelligence system to predict computer skill. The result from the system improvement will be tested with sophomore students in the academic year 2015 with the total number of 30 people. In addition, the results of the satisfactory evaluation by the computer experts give the mean of 4.35 and the standard deviation of 0.53. Therefore, the developed system was designed and considered good and applicable effectively.

Keyword: Intelligent System, Computer Skill, Web Application, Data Mining.

1. บทนำ

ปัจจุบันได้มีการคิดค้นแนวทางทฤษฎีรวมถึงเทคโนโลยีใหม่ๆเพื่อพัฒนามาตรฐานระบบการศึกษาและนำมาใช้ในการแก้ปัญหาในระบบการศึกษาเป็นจำนวนมาก[1] ตัวอย่างเช่น การนำเทคโนโลยีมาช่วยพัฒนาสื่อการเรียนการสอนและการประเมินผลการเรียน [2] การนำเทคนิคการทำเหมืองข้อมูลมาช่วยในการทำนายประสิทธิภาพของนักศึกษาหรือผลการเรียน [3] การพัฒนาระบบสนับสนุนการตัดสินใจในการเลือกสาขาวิชาเรียน [4] เป็นต้น

การทำเหมืองข้อมูลเป็นอีกวิธีการหนึ่งที่ถูกนำมาใช้ในการพัฒนาคุณภาพการศึกษา โดยเป็นการค้นหารูปแบบและความสัมพันธ์ที่ซ่อนอยู่ในข้อมูลจำนวนมาก ซึ่งในมหาวิทยาลัยส่วนใหญ่มีข้อมูลที่ถูกรวบรวมและจัดเก็บอยู่ในระบบสารสนเทศเป็นจำนวนมาก (เช่น ข้อมูลนักศึกษา ข้อมูลผลการเรียน ข้อมูลหลักสูตร เป็นต้น) โดยข้อมูลส่วนใหญ่จะถูกนำมาใช้ประโยชน์ในรูปแบบรายงานหรือข้อมูลสถิติเท่านั้น ดังนั้น หากมีการนำเทคนิคการทำเหมืองข้อมูลมาใช้ร่วมกับเทคโนโลยีสารสนเทศเพื่อพัฒนาเป็นระบบสารสนเทศอัจฉริยะและนำมาใช้กับนักศึกษา หรือบุคลากรของมหาวิทยาลัย น่าจะทำให้มีการนำข้อมูลมาใช้ให้เกิดประโยชน์มากขึ้น รวมถึงอาจจะช่วยในการพัฒนาคุณภาพการศึกษาได้ในอนาคต

ภาควิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีและการจัดการอุตสาหกรรม มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระ-นครเหนือวิทยาเขตปราจีนบุรี เป็นหน่วยงานหนึ่งที่มีการจัดการเรียนการสอนสาขาวิชาคอมพิวเตอร์ ซึ่งสามารถแบ่งกลุ่มทักษะและความถนัดได้ 3 ด้านหลัก ๆ คือ 1) กลุ่มเครือข่ายและความปลอดภัย 2) กลุ่มซอฟต์แวร์และการประยุกต์ และ 3) กลุ่มแอนิเมชั่น ซึ่งในปัจจุบันนักศึกษาบางส่วนของภาควิชาฯ จะไม่สามารถจำแนกความถนัดของตนเองได้โดยเฉพาะอย่างยิ่งในช่วงชั้นปีที่ 1 ของการศึกษา และเมื่อนักศึกษาไม่สามารถทราบได้แน่ชัดว่าตนเองมีความถนัดในด้านใด อาจจะส่งผลให้การเลือกแนวทางการศึกษาต่อในหลักสูตรทางสาขาด้านคอมพิวเตอร์ในชั้นปีที่ 2 จะไม่ตรงกับความสามารถของนักศึกษาหรือนักศึกษาอาจได้รับการพัฒนาและส่งเสริมได้ไม่เต็มประสิทธิภาพเท่าที่ควร

ดังนั้น จากปัญหาดังกล่าว ผู้วิจัยจึงเกิดแนวคิดในการพัฒนาเว็บแอปพลิเคชันแบบอัจฉริยะเพื่อค้นหาทักษะทางด้านคอมพิวเตอร์ให้แก่นักศึกษา โดยการนำเทคโนโลยีที่มีการทำงานด้วยระบบคอมพิวเตอร์มาประยุกต์ร่วมกับกระบวนการการทำเหมืองข้อมูลเพื่อนำกฎความสัมพันธ์ที่ได้มาสร้างเป็นโมเดลสำหรับการทำนายและหาทักษะทางคอมพิวเตอร์ ซึ่งระบบนี้จะช่วยให้นักศึกษาได้รับข้อมูลมาประกอบการตัดสินใจในการหาความถนัดตามทักษะทางคอมพิวเตอร์ของตนเอง นอกจากนี้ อาจารย์ที่ปรึกษาและภาควิชาฯ สามารถนำข้อมูลดังกล่าวมาใช้ในการพัฒนาและส่งเสริมความถนัดของนักศึกษาให้เหมาะสมกับนักศึกษาได้มากขึ้น

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 การทำเหมืองข้อมูล (Data Mining) [5]

การทำเหมืองข้อมูล คือ กระบวนการในการดึงความรู้ (Knowledge) ออกจากฐานข้อมูลขนาดใหญ่เพื่อนำไปใช้ประโยชน์ในการตัดสินใจและใช้งานในรูปแบบอื่น ๆ ด้วยการวิเคราะห์ค้นหาสารสนเทศหรือความรู้ซึ่งอาจจะไม่ได้ปรากฏอยู่จริงในฐานข้อมูลโดยใช้เทคโนโลยีคอมพิวเตอร์ด้านต่างๆ เช่น ทางด้านปัญญาประดิษฐ์ (Artificial Intelligence) เข้ามาช่วยในการค้นหาสารสนเทศหรือความรู้ การทำเหมืองข้อมูลมีหลายโมเดล แต่ละโมเดลจะมีสูตรทางธุรกิจและผลลัพธ์ในรูปแบบที่แตกต่างกัน เช่น ได้ผลลัพธ์เป็นกฎ หรือเป็นแผนภูมิการตัดสินใจ เป็นต้น

2.2 การจัดกลุ่ม (Clustering) [5]

เป็นการจัดกลุ่มข้อมูลซึ่งมีลักษณะคล้ายกับการแบ่งประเภท แต่จะไม่เหมือนกัน โดยการแบ่งประเภทจะวิเคราะห์ข้อมูลตามต้นแบบ แต่สำหรับการแบ่งกลุ่มเป็นการวิเคราะห์โดยไม่พิจารณาจัดกลุ่มตามประเภทที่มีหรือที่รู้จัก โดยเป้าหมายของการจัดกลุ่มคือ การรวบรวมข้อมูลที่มีคุณสมบัติใกล้เคียงกันไว้ในกลุ่มเดียวกันและแบ่งแยกข้อมูลที่มีความแตกต่างกันมากให้อยู่คนละกลุ่ม ขั้นตอนวิธีที่ได้รับความนิยมสำหรับการจัดกลุ่ม คือ ขั้นตอนวิธี k-means [6]

ขั้นตอนวิธี k-means จะกำหนดให้ k คือจำนวนกลุ่มข้อมูล และ means คือค่าเฉลี่ยของสมาชิกข้อมูลในกลุ่มซึ่งก็คือตัวแทนของกลุ่ม (cluster) ซึ่งตัวแทนของกลุ่มนี้จะถือเป็นศูนย์กลางของคลัสเตอร์นั้น ๆ เมื่อการจัดกลุ่มเสร็จสิ้น จะพบว่าข้อมูลทุกตัวอยู่ในกลุ่มข้อมูลที่เหมาะสม โดยวัดจากค่าระยะห่างระหว่างสมาชิกข้อมูลนั้นกับศูนย์กลางของกลุ่มข้อมูลจะต้องเป็นระยะทางที่สั้นที่สุดเมื่อเทียบกับศูนย์กลางของกลุ่มข้อมูลอื่นๆ

2.3 การหากฎความสัมพันธ์ (Association Rules) [5]

เป็นการค้นหากฎความสัมพันธ์ของข้อมูลโดยค้นหาความสัมพันธ์ของข้อมูลทั้งสองชุดหรือมากกว่าสองชุดขึ้นไปไว้ด้วยกัน ความสำคัญของกฎทำการวัดโดยใช้ข้อมูลสองตัวด้วยกันคือค่าสนับสนุน (Support) ซึ่งเป็นเปอร์เซ็นต์ของการ

ดำเนินการที่กฎสามารถนำไปใช้ หรือเป็นเปอร์เซ็นต์ของการดำเนินการที่กฎที่ใช้มีความถูกต้อง และข้อมูลตัวที่สองที่นำมาใช้วัดคือค่าความมั่นใจ (Confidence) ซึ่งเป็นจำนวนของกรณีที่กฎถูกต้องโดยสัมพันธ์กับจำนวนของกรณีที่กฎสามารถนำไปใช้ได้ ในการหากฎความสัมพันธ์นั้นจะมีขั้นตอนวิธีการหาหลายวิธีด้วยกัน แต่ขั้นตอนวิธีที่เป็นที่รู้จักและใช้อย่างแพร่หลายคือ ขั้นตอนวิธี Apriori [7]

2.4 งานวิจัยที่เกี่ยวข้อง

จากการสำรวจพบว่ามีงานวิจัยที่มีการใช้เทคนิคการทำเหมืองข้อมูลมาใช้เพื่อเป็นส่วนหนึ่งในการพัฒนาระบบที่ใช้ในสถาบันการศึกษาหลายงานวิจัยเช่น ชาลี ดวงแย้มและคณะ[8] ได้พัฒนาระบบสนับสนุนการตัดสินใจเลือกประกอบอาชีพสำหรับผู้สำเร็จการศึกษาระดับปริญญาตรี ด้วยวิธีการเรียนรู้แบบเบย์ พบว่าตัวจำแนกแบบเบย์อย่างง่ายสามารถนำมาใช้ในการทำนายการเลือกอาชีพได้เป็นอย่างดี สามารถช่วยตัดสินใจโดยดูจากข้อมูลการศึกษาและคุณลักษณะส่วนตัว ซึ่งสามารถช่วยในการเลือกแนวทางการประกอบอาชีพได้ตรงตามบุคลิกลักษณะและความต้องการของผู้ใช้ได้ นอกจากนี้ยังมีงานวิจัยของจารุวรรณและคณะ [4] ที่นำเสนอระบบสนับสนุนการตัดสินใจสำหรับการเลือกคณะที่ต้องการศึกษาต่อในมหาวิทยาลัยด้วยการผสมผสานหลาย ๆ โมเดลเข้าด้วยกัน ประกอบไปด้วย Scoring Model, Financial Model และ Mix Model โดยโมเดลทั้งหมดได้ถูกนำไปพัฒนาในรูปแบบเว็บแอปพลิเคชันเพื่อช่วยนักศึกษาในการตัดสินใจเลือกคณะที่ต้องการศึกษาต่อโดยพิจารณาจากคะแนนเฉลี่ยแอดมิชชั่นของแต่ละคณะในมหาวิทยาลัย

สำหรับงานวิจัยที่ใช้เทคนิคการหากฎความสัมพันธ์มาใช้กับข้อมูลนักศึกษาพบว่า มีงานวิจัยเรื่องการหาปัจจัยที่ส่งผลต่อผลการเรียนของนักศึกษาเรียนอ่อน โดยใช้เทคนิคกฎความสัมพันธ์ [9] โดยเทคนิคดังกล่าวถูกนำมาใช้เพื่อทำนายความเสี่ยงของการเรียนอ่อนของนักศึกษาชั้นปีที่ 1 ในคณะต่างๆ ของมหาวิทยาลัยสงขลานครินทร์ งานวิจัยนี้ใช้ข้อมูลของนักศึกษาที่เก็บไว้แล้วในฐานข้อมูลของมหาวิทยาลัย ตั้งแต่ปี พ.ศ. 2550-2552 มาศึกษาโดยผ่านกระบวนการต่างๆของการทำเหมืองข้อมูลและใช้เทคนิคกฎความสัมพันธ์เพื่อหารูปแบบของข้อมูล จากผลการวิจัยพบว่าปัจจัยที่ส่งผลต่อผลการศึกษาของนักศึกษาในแต่ละคณะแตกต่างกันและความถูกต้องของผลลัพธ์ก็แตกต่างกันอันเนื่องมาจากข้อมูลนำเข้าของนักศึกษาในแต่ละคณะต่างกัน ดังนั้นผลลัพธ์และความถูกต้องของกฎความสัมพันธ์จึงขึ้นอยู่กับข้อมูลที่ใช้ในการเรียนรู้เมื่อมองในภาพรวมแล้วพบว่าปัจจัยที่ส่งผลต่อผลการศึกษาของนักศึกษาเรียนอ่อนคือ วิธีการเข้าศึกษา คะแนนสอบเข้าศึกษาของวิชาต่างๆ และเพศของนักศึกษา

นอกจากนี้ยังมีระบบช่วยสอนอัจฉริยะบนเว็บสำหรับการเรียนการเขียนโปรแกรมคอมพิวเตอร์ (BITS) [2] เป็นอีกระบบหนึ่งที่ถูกพัฒนาขึ้นโดยนำเทคนิคเครือข่ายแบบเบย์ (Bayesian network) มาช่วยสนับสนุนการตัดสินใจ โดยระบบ BITS นี้สามารถช่วยนักศึกษาที่มหาวิทยาลัย Regina ในการค้นหาเนื้อหาที่สอดคล้องกับความต้องการของนักศึกษาโดยสามารถเรียงลำดับของเนื้อหาได้อัตโนมัติ ซึ่งผลที่ได้จากการพัฒนาระบบพบว่าสามารถนำมาใช้ช่วยสอนในรายวิชาดังกล่าวได้อย่างมีประสิทธิภาพ

3. วิธีการดำเนินงาน

ในการพัฒนาระบบอัจฉริยะเพื่อค้นหาทักษะทางคอมพิวเตอร์โดยใช้เทคนิคการจัดกลุ่มและการหากฎความสัมพันธ์ ผู้วิจัยได้ทำการรวบรวมข้อมูล ออกแบบ และพัฒนาระบบตามขั้นตอนดังต่อไปนี้

3.1 การวิเคราะห์ระบบ

ผู้วิจัยได้ดำเนินการศึกษาปัญหาและความต้องการของระบบโดยทำการศึกษาเกี่ยวกับข้อมูลรายวิชาที่นักศึกษาชั้นปีที่ 1 และ ชั้นปีที่ 2 ว่ามีลักษณะวิชาเรียนอย่างไรซึ่งจะส่งผลต่อการเกิดทักษะของนักศึกษา และนักศึกษาที่กำลังศึกษาอยู่อาจยังไม่ทราบถึงทักษะความสามารถของตนเองอย่างชัดเจน ทำให้นักศึกษาอาจเกิดปัญหาที่ทำให้ขาดประสิทธิภาพในการเรียน จึงได้ทำการศึกษารายละเอียดเกี่ยวกับรายวิชาที่เกี่ยวข้องในการทำนาย รวมทั้งศึกษาความต้องการของระบบในเรื่องของส่วนประกอบของระบบและผู้ใช้งานระบบ

นอกจากนี้ผู้วิจัยได้ดำเนินการเก็บรวบรวมข้อมูลของระบบโดยทำการเก็บรวบรวมข้อมูลผลการเรียนของนักศึกษาตั้งแต่ชั้นปีที่ 1-2 ตั้งแต่ปีการศึกษา 2548-2557 เพื่อนำมาจัดสร้างเป็นฐานข้อมูลที่ใช้ในการหาทักษะทางด้านคอมพิวเตอร์ โดยรวบรวมข้อมูลจาก ระบบงานทะเบียนสารสนเทศนักศึกษา และ

ดำเนินการการศึกษาเครื่องมือต่าง ๆ ที่ใช้ในการพัฒนาระบบ โดยทำการศึกษาการเขียนโปรแกรมด้วยภาษา JavaScript ศึกษาการใช้งานฐานข้อมูล MongoDB

3.2 ข้อมูลที่ใช้ในการวิจัย

ข้อมูลที่นำมาใช้ในการทำวิจัยคือข้อมูลผลการเรียนของนักศึกษาจากระบบทะเบียนนักศึกษาระหว่างปี พ.ศ. 2548 – 2557 หลักสูตรเทคโนโลยีบัณฑิต ภาควิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีและการจัดการอุตสาหกรรม มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ วิทยาเขตปราจีนบุรี จำนวน 19 แอททริบิวต์ และ 733 ชุดข้อมูล โดยค่าของแต่ละแอททริบิวต์จะอยู่ในรูปของผลการเรียนตั้งแต่ F-A หลังจากนั้นได้มีกระบวนการในการเตรียมข้อมูล (Pre-processing Data) กับข้อมูลที่สูญหาย (Missing Value) โดยการค้นหาข้อมูลจากระบบงานทะเบียน และดำเนินการแปลงข้อมูลให้อยู่ในรูปของกลุ่มรายวิชา ทำให้เหลือแอททริบิวต์ที่จะถูกนำมาพิจารณาในการสร้างโมเดลเพียง 9 แอททริบิวต์ และผลการเรียนที่เกิดขึ้นจะเป็นการหาจากค่าเฉลี่ยของเกรดในแต่ละรายวิชาของแต่ละหมวดวิชา ซึ่งรายละเอียดการจัดกลุ่มรายวิชาและแอททริบิวต์ 9 แอททริบิวต์ แสดงดังตารางที่ 1

3.3 โมเดลที่ใช้สำหรับการค้นหาทักษะทางคอมพิวเตอร์

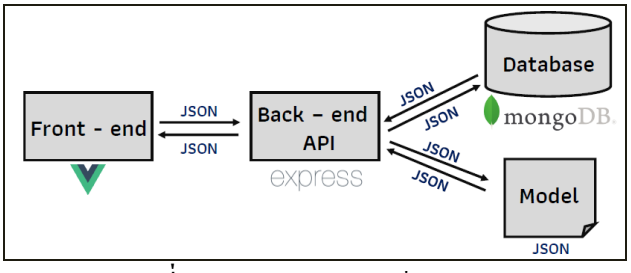
เมื่อได้ข้อมูลที่พร้อมสำหรับการสร้างโมเดลแล้ว ขั้นตอนต่อมานำข้อมูลดังกล่าวมาจัดกลุ่มด้วยขั้นตอนวิธี K-Means โดยกำหนดเป็น 3 กลุ่ม (k=3) โดยค่าของ k พิจารณาจากจำนวนของสาขาตามทักษะทางคอมพิวเตอร์ที่ทางภาควิชาได้จัดการเรียนการสอน ซึ่งเมื่อดำเนินการจัดกลุ่มแล้วจะได้ผลลัพธ์ดังนี้ 1. Cluster 0 (แอนิเมชัน) 256 ข้อมูล 2. Cluster 1 (เครือข่ายและความปลอดภัย) 242 ข้อมูล และ 3. Cluster 3 (ซอฟแวร์และการประยุกต์) 235 ข้อมูล เมื่อได้ข้อมูลจัดกลุ่มทั้ง 3 กลุ่ม จึงนำข้อมูลแต่ละกลุ่มมาทำการหากฎความสัมพันธ์ด้วยวิธีการ Apriori Algorithm ซึ่งได้ผลลัพธ์ดังนี้ 1. Cluster 0 (แอนิเมชัน) ได้กฎความสัมพันธ์ 23 กฎความสัมพันธ์ 2. Cluster 1 (เครือข่ายและความปลอดภัย) ได้กฎความสัมพันธ์ 15 กฎความสัมพันธ์ และ 3. Cluster 3 (ซอฟแวร์และการประยุกต์) ได้กฎความสัมพันธ์ 17 กฎความสัมพันธ์และจะนำกฎความสัมพันธ์ที่ได้นี้นำไปใช้เป็นรูปแบบในการทำนายทักษะทางด้านคอมพิวเตอร์ ค่าประสิทธิภาพที่ได้จากกฎความสัมพันธ์โดยใช้ค่าสนับสนุน (Support) 20% ค่าความเชื่อมั่น (Confidence) 100% และค่าบ่งบอกความสัมพันธ์ (Lift) 100% เพื่อให้กฎความสัมพันธ์ที่ได้มีประสิทธิภาพในการหาทักษะทางคอมพิวเตอร์มากที่สุด

ตารางที่ 1: การจัดกลุ่มตามประเภทของรายวิชา

หมวดวิชา	ชื่อวิชา
1. Program	- Computer Programming - Computer Programming Lab (C++) - Object-Oriented Programming - Object-Oriented Programming Lab (Java) - Web Programming (PHP)
2. Math	- General Mathematics - Discrete Mathematics and Application - Statistics for Engineer
3. English	- English I - English II
4. Algorithm	- Data structure and algorithm
5. Database System	- Database System
6. Data communication	- Data communication and computer network
7. Graphic	- Computer Graphics
8. System Analysis	- System Analysis and Design
9. Computer	- Introduction to Information - Fundamental of Digital - Computer Architecture - Computer Operating Systems

3.4 การออกแบบการทำงานของระบบ

จากการวิเคราะห์ระบบทางผู้วิจัยได้แบ่งระบบออกเป็น 2 ส่วนคือ 1) ประสานงานกับผู้ใช้ (front-end) ที่ทำหน้าที่แสดงผลข้อมูลและประมวลผลข้อมูล โดยนักศึกษาจะกรอกรายละเอียดของข้อมูลเกรดผ่านฟอร์มแล้วทำการยืนยันเพื่อค้นหาทักษะทางคอมพิวเตอร์ที่สอดคล้องกับนักศึกษามากที่สุด และ 2) ส่วนประมวลผล (back – end) คือส่วนของ API ที่ทำหน้าที่ติดต่อกับฐานข้อมูลและโมเดล ในส่วนนี้ระบบจะดำเนินการนำข้อมูลที่นักศึกษากรอกมาดำเนินการเรียนรู้ผ่านโมเดลที่ได้ออกมาในรูปแบบของกฎในแต่ละกลุ่มความถนัด (Cluster) แล้วพิจารณาจากจำนวนของกฎที่สอดคล้องกับข้อมูลของนักศึกษามากที่สุด ซึ่งหากกลุ่มหรือคลัสเตอร์ใดที่มีปริมาณของกฎสอดคล้องกับข้อมูลของนักศึกษามากที่สุดจะพิจารณาเลือกทักษะทางคอมพิวเตอร์ให้กับนักศึกษาคนดังกล่าว การออกแบบโครงสร้างของเว็บแอปพลิเคชันของระบบแสดงดังภาพที่ 1



ภาพที่ 1: โครงสร้างเว็บแอพพลิชั่นของระบบ

3.5 การพัฒนาระบบ

การพัฒนาระบบผู้วิจัยได้พัฒนาลักษณะของเว็บแอปพลิเคชัน โดยพัฒนาโดยใช้ภาษา JavaScript เชื่อมต่อกับฐานข้อมูล MongoDB ที่ใช้สำหรับจัดเก็บข้อมูลของนักศึกษา ในส่วนของข้อมูลกฎความสัมพันธ์ที่ได้จากโมเดลจะถูกจัดเก็บอยู่ในรูปแบบของไฟล์ JSON

3.6 การทดสอบและประเมินคุณภาพของระบบ

หลังจากการพัฒนาระบบเสร็จสิ้น ผู้วิจัยจะทำการทดสอบระบบโดยแบ่งเป็น 2 ส่วนคือ 1) ทำการทดสอบหรือประเมินคุณภาพของระบบด้วยวิธีการแบบแบล็กบ๊อกซ์ (Black Box Testing) ซึ่งเป็นวิธีการทดสอบเพื่อหาข้อผิดพลาดที่เกิดจากการใช้งานระบบ ด้วยกลุ่มนักศึกษาชั้นปีที่ 2 ปีการศึกษา 2558 จำนวนทั้งสิ้น 30 คน 2) ประเมินประสิทธิภาพของระบบโดยผู้เชี่ยวชาญจำนวนทั้งสิ้น 5 คน

4. ผลการดำเนินงาน

4.1 ผลการพัฒนาระบบ

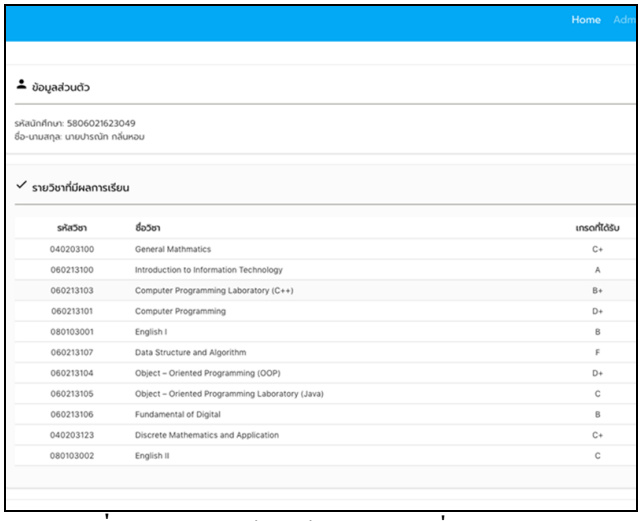
หน้าจอหลักที่สำคัญของระบบมีดังนี้

ก) หน้าจอเริ่มต้นระบบผู้ใช้สามารถกรอกรหัสนักศึกษาเพื่อจะเข้าสู่การทำนายทักษะ



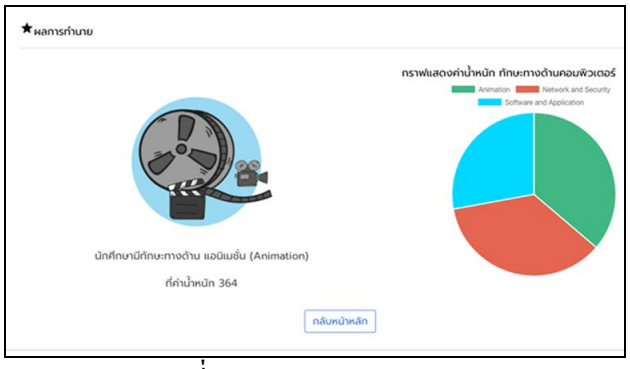
ภาพที่ 2: ตัวอย่างหน้าจอแรกของระบบ

ข) หลังจากที่นักศึกษาได้ทำการกดปุ่มทำนาย ระบบจะทำการดึงข้อมูลจากฐานข้อมูลมาแสดงผลการเรียนของนักศึกษาอัตโนมัติ ในกรณีที่นักศึกษายังไม่มีผลการเรียนในระบบทะเบียนนักศึกษาระบบจะให้นักศึกษาทำการคาดการณ์ผลการเรียนที่คาดว่าจะได้รับในวิชานั้น ๆ ตัวอย่างหน้าจอการแสดงผลการเรียนแสดงดังภาพที่ 3



ภาพที่ 3: ตัวอย่างหน้าจอข้อมูลในกรณีที่มีผลการเรียน

ค) เมื่อนักศึกษากรอกข้อมูลครบถ้วนระบบจะทำการทำนายทักษะทางด้านคอมพิวเตอร์ของนักศึกษาแสดงดังภาพที่ 4



ภาพที่ 4: ตัวอย่างหน้าจอผลการทำนาย

4.2 ผลการประเมิน

ผู้วิจัยได้ดำเนินการประเมินระบบด้านต่างๆ ด้วยผู้เชี่ยวชาญจำนวน 5 ท่าน เพื่อประเมินหาประสิทธิภาพของระบบที่ได้พัฒนาขึ้น โดยผลการประเมินแสดงดังตารางที่ 2

จากตารางที่ 2 พบว่าผลการประเมินคุณภาพของระบบจากผู้เชี่ยวชาญระบบได้รับค่าเฉลี่ย (Mean) ในทุก ๆ ด้านเท่ากับ 4.47 และได้รับค่าส่วนเบี่ยงเบนมาตรฐาน (SD) ในทุก ๆ ด้านเท่ากับ 0.56 ดังนั้นระบบอัจฉริยะในการค้นหาทักษะทางคอมพิวเตอร์ที่พัฒนาขึ้นมีการใช้งานระบบโดยรวมให้คุณภาพอยู่ในระดับดี

ตารางที่ 2: ผลการประเมินคุณภาพของระบบโดยผู้เชี่ยวชาญ

รายการ	$\bar{x}$	S.D.
1.ด้านความสามารถตอบสนองความต้องการของผู้ใช้	4.53	0.45
2. ด้านการติดต่อกับผู้ใช้โปรแกรม	4.95	0.59
3. ด้านความถูกต้องในการทำงานของระบบ	4.64	0.74
4. ด้านการประมวลผลของระบบ	4.72	0.46
5. ด้านการรักษาความปลอดภัยของข้อมูล	3.53	0.56
ผลประเมินเฉลี่ยทั้งหมด	**4.47**	**0.56**

นอกจากนี้ผู้วิจัยได้ดำเนินการวัดประสิทธิภาพของตัวโมเดลกับชุดข้อมูลใหม่ โดยการนำระบบอัจฉริยะนี้ไปทดสอบกับนักศึกษาชั้นปีที่ 2 จำนวนทั้งสิ้น 30 คน พบว่า ระบบสามารถทำนายทักษะทางคอมพิวเตอร์ได้ถูกต้องและสอดคล้องกับความสามารถของนักศึกษาร้อยละ 75.33 ซึ่งผลการทำนายดังกล่าวสามารถนำมาช่วยให้นักศึกษาวางแผนการเรียนตามความถนัดของตนเองได้ในอนาคต

5. สรุปผลงานวิจัย

งานวิจัยนี้ได้นำเสนอระบบอัจฉริยะเพื่อค้นหาทักษะทางคอมพิวเตอร์โดยใช้ 2 เทคนิคผสมผสานกันคือเทคนิคการจัดกลุ่มและเทคนิคการหากฎความสัมพันธ์ร่วมกับเทคโนโลยีทางระบบสารสนเทศ ระบบนี้ได้ถูกพัฒนาขึ้นในลักษณะเว็บแอปพลิเคชันสำหรับนักศึกษาภาควิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีและการจัดการอุตสาหกรรม มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ วิทยาเขตปราจีนบุรี เพื่อนำมาช่วยอำนวยความสะดวกให้กับนักศึกษาในการค้นหาทักษะทางคอมพิวเตอร์ที่สอดคล้องกับผลการเรียนของนักศึกษาได้ด้วยตนเอง โดยระบบจะใช้ข้อมูลผลการเรียนในภาคเรียนที่ 1 ถึงภาคเรียนที่ 3 จากผลการทดลองพบว่า ระบบได้รับการประเมินประสิทธิภาพด้วยผู้เชี่ยวชาญ ซึ่งผลที่ได้พบว่าระบบโดยรวมมีคุณภาพดี โดยมีค่าเฉลี่ยเท่ากับ 4.47 และมีค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.56 นอกจากนี้ผู้วิจัยได้นำระบบไปใช้กับนักศึกษาชั้นปีที่ 2 จำนวน 30 คน เพื่อวัดประสิทธิภาพของโมเดลพบว่าระบบสามารถทำนายได้ถูกต้องร้อยละ 77.53 จากผลการประเมินทั้งหมดสรุปได้ว่า ระบบอัจฉริยะเพื่อค้นหาทักษะทางคอมพิวเตอร์มีประสิทธิภาพดีสามารถนำไปประยุกต์ใช้งานได้จริง

6. กิตติกรรมประกาศ

งานวิจัยนี้ได้รับทุนในการพัฒนาจากการแข่งขันพัฒนาโปรแกรมคอมพิวเตอร์แห่งประเทศไทย ครั้งที่ 19 ศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ

เอกสารอ้างอิง

[1] R. Kresimir, B. G. Marijana, and M. Vlado, "Development of the Intelligent System for the Use of University Information System," *Procedia Engineering,* pp. 402-409. vol. 69, 2014/01/01 2014.

[2] C. J. Butz, S. Hua, and R. B. Maguire, "A Web-Based Intelligent Tutoring System for Computer Programming," in *Web Intelligence, 2004. WI 2004. Proceedings. IEEE/WIC/ACM International Conference on*, pp. 159-165, 2004.

[3] พ. บุตรเอก และ ส. บุญลือ, "การพยากรณ์โอกาสสำเร็จการศึกษาของนักศึกษาโดยใช้ซัพพอร์ตเวอกเตอร์แมชชีน," *Veridian E-Journal Science and Technology Silpakorn University,* หน้า 40-49. vol. 1, 2557.

[4] J. Chanwijit, W. Lomwongpaiboon, O. Dowjam, and P. Tangworakitthaworn, "Decision Support System for Targeting Higher Education," in *2016 Fifth ICT International Student Project Conference (ICT-ISPC)*, pp. 154-157, 2016.

[5] น. สุนทรภิษัช, ปัญญาประดิษฐ์, 2 ed.: ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยเกษตรศาสตร์, 2553.

[6] J. MacQueen, "Some methods for classification and analysis of multivariate observations," in *In 5-th Berkeley Symposium on Mathematical Statistics and Probability}*, ed: University of California Press, 1967, pp. 281--297.

[7] R. Agrawal and R. Srikant, "Fast Algorithms for Mining Association Rules in Large Databases," *presented at the Proceedings of the 20th International Conference on Very Large Data Bases*, 1994.

[8] ช. ดวงแย้ม, ก. เฮงพระพรหม,และ ส. เฮงพระพรหม, "ระบบสนับสนุนการตัดสินใจเลือกประกอบอาชีพสำหรับผู้สำเร็จการศึกษาระดับปริญญาตรี ด้วยวิธีการเรียรู้แบบเบย์," in *The 9th Natoinal Conference on Computing and Information Technology*, 2013.

[9] ณ. สุวรรณโณ และ อ. สิงห์เอี่ยม, "การหาปัจจัยที่ส่งผลต่อความเสี่ยงของนักศึกษาเรียนอ่อนด้วยเทคนิคกฎความสัมพันธ์ กรณีศึกษามหาวิทยาลัยสงขลานครินทร์," วารสารวิทยาการจัดการ, หน้า. 65-79. vol. 28, 2554.

การเพิ่มประสิทธิภาพของแบบจำลองกฎการพยากรณ์ในเหมืองข้อมูล
Enhancing Performance of Rule-based Predicting Models in Data Mining

ดวงกมล โสดาลี (Duangkamol Sodalee)[1] และจารี ทองคำ (Jaree Thongkam)[2]

[1]นิสิต สาขาเทคโนโลยีสารสนเทศ คณะวิทยาการสารสนเทศ มหาวิทยาลัยมหาสารคาม

[2]อาจารย์ที่ปรึกษา กลุ่มสารสนเทศเชิงประยุกต์ สาขาเทคโนโลยีสารสนเทศ คณะวิทยาการสารสนเทศ มหาวิทยาลัยมหาสารคาม

snoopyamp@gmail.com, jaree.thongkam@gmail.com

บทคัดย่อ

เทคนิคในเหมืองข้อมูลโดยส่วนใหญ่สามารถสร้างกฎการจำแนกซึ่งเป็นการสร้างแบบจำลองกฎการพยากรณ์ที่มีประสิทธิภาพ แต่ไม่สามารถสร้างแบบจำลองกฎการพยากรณ์ที่มีประสิทธิภาพจากข้อมูลจริง (real world data) งานวิจัยนี้จึงมีจุดประสงค์เพื่อเพิ่มประสิทธิภาพของแบบจำลองกฎพยากรณ์ (Rule-based) โดยการใช้ Gain Ratio เพื่อเลือกตัวแปรที่สำคัญต่อการพยากรณ์ และคัดเลือกข้อมูลที่มีค่าผิดปกติ (outliers) ออก โดยข้อมูลถูกรวบรวมจากสถานศึกษาอาชีวศึกษาเอกชนในจังหวัดกาฬสินธุ์ จำนวน 12 สถาบัน ในปีการศึกษา 2557-2559 จำนวน 929 ระเบียน ในการนี้เทคนิค Decision Table, Decision Rule, Repeated Incremental Pruning to Produce Error Reduction (RIPPER) และ OneR ได้ถูกนำมาใช้ในการสร้างแบบจำลองเพื่อการพยากรณ์ ในการวัดประสิทธิภาพแบบจำลอง 10-fold cross validation ได้ถูกนำมาใช้ในการแบ่งข้อมูลออกเป็นชุดข้อมูลเรียนรู้ และชุดข้อมูลทดสอบ โดยประสิทธิภาพของแบบจำลองที่สร้างจากแต่ละเทคนิคใช้ ค่าความเที่ยงตรง (Precision) ค่าความระลึก (Recall) และค่า F-measure จากการศึกษาพบว่า เมื่อใช้ Gain Ratio ในการเลือกตัวแปร และคัดเลือกข้อมูลที่มีค่าผิดปกติ (outliers) ออก สามารถสร้างแบบจำลองกฎพยากรณ์ที่ประสิทธิภาพที่มีความถูกต้อง (Precision) สูงขึ้นจากข้อมูลดิบถึง 11.30% ค่าความระลึก (Recall) สูงขึ้น 11.60 % และค่า F-measure สูงขึ้น 11.30%

คำสำคัญ: Gain Ratio, แบบจำลองกฎการพยากรณ์, เหมืองข้อมูล , แบบจำลองการพยากรณ์ตามกฎ

Abstract

Almost every technique in data mining can create effective rule-based prediction models, but they cannot create effective rule-based prediction models from the real-world data. The purpose of this research is to enhance rule-based prediction models using Gain Ratio for selecting variables importance and filtering out the outliers. The data were collected from 12 vocational education institution in Kalasin from year 2557-2559. The data consist of 929 records. In this paper, Decision Table, Decision Rule, Repeated Incremental Pruning to Produce Error Reduction (RIPPER) and OneR are used for building the rule-base prediction models. In order to measure the performance of the rule-base prediction models, 10-fold cross validation is applied to divide the data set into learning and testing sets. The performance of the prediction models created from each technique is measured with precision, recall and F-measure. The experimental results revealed that Gain Ratio can increase performance of rules-based prediction models. After using Gain Ratio for selecting variables and filtering out the outliers, the techniques can create rule-based prediction model which increase precision up to 11.30%, recall up to 11.60 % and F-measure up to 11.30%.

Keyword: Gain Ratio, Rule-based Prediction Models, Data Mining, Rule-based Prediction Models.

1. บทนำ

กฎการจำแนก (Rule-Based Classification) หมายถึง การจำแนกข้อมูลด้วยกฎซึ่งเป็นแบบจำลองที่จะแสดงผลด้วยเซตของกฎที่มีลักษณะแบบ ‘IF-THEN’ โดยกฎหนึ่งๆจะถูกแสดงอยู่ในรูปฟอร์ม ดังนี้

IF condition THEN conclusion

การสร้างแบบจำลองกฎการพยากรณ์ที่มีประสิทธิภาพในการพยากรณ์เป็นงานวิจัยที่ท้าทายในเหมืองข้อมูล เช่น ปวีณา ชัยวนารมย์ [1] ได้ทดสอบเพื่อหาแบบจำลองที่เหมาะสมสำหรับพยากรณ์ความเครียดด้วยเทคนิคการทำเหมืองข้อมูล โดยใช้โปรแกรม WEKA จำนวน 6 อัลกอริทึม คือ Bayesian Network, Naïve Bayesian, Decision Tree:4.5, Decision Table, Partial Rules (PART) และ Multilayer Perceptron (MLP) จากการทดสอบแบบจำลองทั้งหมดพบว่า การสร้างแบบจำลองจากอัลกอริทึม Bayesian Network มีค่าความถูกต้อง ค่าความแม่นยำ ค่าความระลึก และค่าความเหวี่ยงมีค่า เท่ากับ 76.76%, 78.5%,76.7% และ 77.2% อัลกอริทึม Naïve Bayesian มีค่า 77.67%, 79.6%, 77.7% และ 78.3% อัลกอริทึม Decision Tree:4.5 มีค่า 77%,77.1%, 77% และ 76.5% อัลกอริทึม Decision Table มีค่า 58.33%, 58.4%,58.3 และ 57% อัลกอริทึม Partial Rules (PART) มีค่า 69.33%, 69.2%,69.3% และ 69.2%

เสถียร วิชาเรือง [2]ได้ประยุกต์ใช้ 3 เทคนิคในการทำเหมืองข้อมูลในกระบวนการอนุมัติสินเชื่อ คือ c4.5, Decision Rules และ Decision Table ผลการทดลองพบว่า Decision Rules ให้ผลลัพธ์ที่มีความถูกต้องมากที่สุดเพียง 72.70%

จะเห็นว่าแบบจำลองที่สร้างจากข้อมูลจริงโดยส่วนใหญ่จะให้ค่าความถูกต้องที่น้อยกว่า 80% ทำให้แบบจำลองมีความน่าเชื่อถือต่ำ ในการนี้มีนักวิจัยหลายท่านได้นำเอากระบวนการก่อนการสร้างแบบจำลองมาทำการคัดเลือกตัวแปรที่มีประสิทธิภาพ และทำการคัดกรองเอาค่าที่ผิดปกติออก [3] [4]

ดังนั้นในงานวิจัยนี้จึงนำเอา Gain Ratio ซึ่งเป็นหน่วยวัดการสุ่มตัวอย่าง (Entropy) มาใช้ในการเพิ่มประสิทธิภาพของแบบจำลองกฎการพยากรณ์ การสำเร็จการศึกษาระดับประกาศนียบัตรวิชาชีพชั้นสูง ในสถานศึกษาอาชีวศึกษาเอกชน จังหวัดกาฬสินธุ์ ในการสร้างแบบจำลองผู้วิจัยได้ใช้ 4 เทคนิค รวม คือ Decision Table, Decision Rule, RIPPER และ OneR รวมถึงการใช้หลักการ 10-fold cross validation ในการแบ่งกลุ่มข้อมูลเป็นชุดข้อมูลเรียนรู้ และชุดข้อมูลทดสอบ และวัดประสิทธิภาพการพยากรณ์ของแบบจำลองด้วยค่าความเที่ยงตรง (Precision) ค่าความระลึก (Recall) และค่า F-measure

2. งานวิจัยที่เกี่ยวข้อง

2.1 การอาชีวศึกษาและอาชีวะ

การศึกษาวิชาชีพและผลิตกำลังคนที่มีทักษะความสามารถเจตคติ และความรับผิดชอบที่ดีรวมทั้งเป็นการจัดการศึกษาเพื่อประกอบอาชีพให้หลากหลายด้วยการใช้ยุทธวิธีหรือกลยุทธ์ในเชิงรุก เพื่อรับและตอบสนองความต้องการของตลาดแรงงานได้อย่างมีประสิทธิภาพ [5] [6]

2.2 เหมืองข้อมูล

การทำเหมืองข้อมูล (Data Mining) [7] เป็นกระบวนการที่กระทำกับข้อมูลจำนวนมากเพื่อค้นหารูปแบบและความสัมพันธ์ที่ซ่อนอยู่ในชุดข้อมูลนั้น มาสู่การจัดเก็บในรูปแบบฐานข้อมูล ที่สามารถดึงข้อมูลสารสนเทศมาใช้ จนถึงการทำเหมืองข้อมูลที่สามารถค้นพบความรู้ที่ซ่อนอยู่ในข้อมูล ซึ่งในปัจจุบันการทำเหมืองข้อมูลได้ถูกนำไปประยุกต์ใช้ในงานหลายประเภท ทั้งในด้านธุรกิจที่ช่วยในการตัดสินใจของผู้บริหาร ในด้านวิทยาศาสตร์และการแพทย์รวมทั้งในด้านเศรษฐกิจและสังคม

2.3 เทคนิคกฎการจำแนกในเหมืองข้อมูล

1. เทคนิค Decision Table [8] เป็นการจัดกลุ่มในลักษณะตารางการตัดสินใจโดยการนำผลที่ได้จากการเรียนรู้จากข้อมูลที่นำเข้ามาที่มีลักษณะของข้อมูลที่ใกล้เคียงกันให้มาอยู่ในกลุ่มเดียวกัน โดยแอตทริบิวต์สุดท้ายของตารางจะใช้กำหนดกลุ่มที่ต้องการแบ่ง

2. เทคนิค Decision Rule [9] เป็นการจัดกลุ่มโดยให้ผลลัพธ์ที่ได้จะอยู่ในรูปของกฎที่นำเอาข้อมูลแต่ละแอตทริบิวต์มาเป็นเงื่อนไขในการตัดสินใจลักษณะของข้อมูลที่ใกล้เคียงกันให้มาอยู่ในกลุ่มเดียวกัน โดยค่าของข้อมูลของแอตทริบิวต์ที่มีค่าต่อเนื่องจะใช้เครื่องหมาย “<”, “<=”, “=>”, “>” ในการพิจารณาส่วนข้อมูลที่เป็นค่าไม่ต่อเนื่องจะใช้เครื่องหมาย “=” และใช้เงื่อนไข “AND” ในการเชื่อมโยงระหว่างแอตทริบิวต์ จะเชื่อมโยงไปจนถึงแอตทริบิวต์สุดท้ายที่จะแบ่งกลุ่มได้

3. เทคนิค RIPPER เป็นกฎที่เกิดจากการแบ่งกฎการจำแนกประเภทข้อมูลเป็นกลุ่มตามคลาสเป้าหมาย แล้วใช้หลักการตัดกิ่งเพื่อเลือกกฎที่มีประสิทธิภาพมากกว่าอัตราความผิดพลาด (Error Rate) ที่กำหนดมาใช้ในการทำนายข้อมูล

4. เทคนิค OneR [10] เป็นอัลกอริทึมที่มีวิธีการสร้างกฎได้ง่ายและใช้งานง่าย โดยใช้หลักการสร้างกฎการจำแนกประเภทด้วยการเลือกแอตทริบิวต์ที่มีค่าความผิดพลาดน้อยที่สุดเพียงแอตทริบิวต์เดียวมาเป็นตัวทำนายคลาสของข้อมูล ซึ่งจะทำให้ได้รับกฎการจำแนกประเภทข้อมูลที่มีจำนวนน้อย

2.4 Gain Ratio

Gain Ratio [11] เป็นหน่วยวัดการสุ่มตัวอย่าง (Entropy) ซึ่งสามารถใช้ได้ทั้งในกระบวนการหาปัจจัยหรือตัวแปรที่มีความสำคัญในการพยากรณ์ และยังสามารถใช้ในการคัดกรองข้อมูลที่มีค่าผิดปกติอีกด้วย ในการคำนวณหาค่า Gain Ratio สามารถคำนวณโดยการหาค่า Information Gain (Gain (a)) ดังสมการที่ 1 และ splitinfo ดังสมการที่ 2 จะได้ค่า Gain Ratio ดังสมการ 3

$$\text{Gain (a)} = \text{info(D)} - \text{Info}_a\text{(D)} \qquad (1)$$

$$\text{SplitInfo}_a\text{(D)} = -\sum_{j=1}^{v} \frac{|D_j|}{|D|} \log_2 \left(\frac{|D_j|}{|D|}\right. \qquad (2)$$

$$\text{Gain Ratio(a)} = \frac{\text{Gain(a)}}{\text{SplitInfo(a)}} \qquad (3)$$

3. ขั้นตอนวิธีการดำเนินการวิจัย

ในการทำวิจัยคณะผู้วิจัยได้มีแบ่งวิธีการดำเนินการวิจัยออกเป็น 4 ขั้นตอนหลัก 1) การเตรียมข้อมูล 2) กระบวนการก่อนการสร้างแบบจำลอง ประกอบด้วย การแปลงข้อมูล และการวิเคราะห์ปัจจัย 3) การสร้างแบบจำลอง และ 4) การวัดประสิทธิภาพแบบจำลอง

3.1 การเตรียมข้อมูล

การเตรียมข้อมูลในงานวิจัยฉบับนี้ ได้รวบรวมข้อมูลจากสถานศึกษาอาชีวศึกษาเอกชนในจังหวัดกาฬสินธุ์ จำนวน 12 แห่ง ประจำปีการศึกษาที่ 2557-2559 ได้ข้อมูลจำนวน 929 ระเบียน

3.2 การทำกระบวนการก่อนการสร้างแบบจำลอง

ในขั้นตอนนี้ผู้วิจัยได้ทำการแปลงข้อมูลทำการวิเคราะห์ปัจจัยหรือเลือกตัวแปรที่มีความสำคัญในการพยากรณ์ ด้วย Gain Ratio และทำการคัดกรองข้อมูลที่ผิดปกติออกด้วย Gain Ratio เช่นกัน

1.การแปลงข้อมูล (Data transformation) จากข้อมูลที่ได้จากการสำรวจข้อมูลจำนวนนักเรียนจากสถานศึกษาอาชีวศึกษาเอกชนในจังหวัดกาฬสินธุ์ ประจำปีการศึกษา 2557-2559 จำนวน 12 แห่ง มีทั้งหมด 21 ตัวแปร ข้อมูลพื้นฐานนี้เกี่ยวกับนักเรียนแต่ละบุคคล เช่น เพศ, ผลการเรียน, สภาพครอบครัว, จำนวนพี่น้อง และ อาชีพของบิดาและมารดา ฯลฯ ซึ่งข้อมูลเหล่านี้ล้วนแต่มีผลต่อการพยากรณ์การสำเร็จการศึกษา [12] ดังตารางที่ 1

2. การวิเคราะห์ปัจจัยเป็นการคัดเลือกตัวแปรที่สามารถเป็นตัวแทนในการพยากรณ์ซึ่งงานวิจัยนี้ได้นำ Gain Ratio มาใช้ในการวิเคราะห์และได้ทำการตัดตัวแปรที่มีผลต่อการพยากรณ์น้อยออก จำนวน 3 ตัวแปรคือ Age, Number of siblings และ Province คงเหลือ 18 แอตทริบิวต์ ดังนี้ Amphoe, Type of course, Sex, Tambol, Health, Village, Mother Career, Art, Father Career, Thai, Social, Mother Status, Science, Career and Technology, Math, Foreign language, Father Status และ Graduate(Class)

3. ทำการคัดกรองข้อมูลที่ผิดปกติออก ด้วย ต้นไม้ตัดสินใจ (C4.5) ซึ่ง C4.5 ได้ใช้ Gain Ratio ใช้ในการสร้างต้นไม้ ดังนั้น ถ้าข้อมูลใดที่ไม่สามารถจำแนกด้วย C4.5 ได้ถือว่าเป็นข้อมูลที่ผิดปกติ โดยจะถูกนำออกจากชุดข้อมูล จากการเตรียมข้อมูลข้างต้นจะเห็นว่าจำนวนของข้อมูลลดลงเพียง 10.33% สามารถแสดงข้อมูลได้ดังตารางที่ 2

ตารางที่ 1 ตัวแปรในงานวิจัย

ชื่อตัวแปร	ชนิดของตัวแปร	ข้อมูล
Sex	Nominal	Male
		Female
Age	Numeric	Age
Number of siblings	Numeric	Number of siblings
Father Status	Nominal	Family
		Separated
		Divorced
		Deceased
		Married
Father Career	Nominal	Government
		State Enterprises
		Trade
		Agriculture
		Contractors
		Government Employee
		Retired Government
		Priest
		Jobless
		Others
		Deceased
Mother Status	Nominal	Family
		Separated
		Divorced
		Deceased
		Married
Mother Career	Nominal	Government
		State Enterprises
		Trade
		Agriculture
		Contractors
		Government Employee
		Retired Government
		Priest
		Jobless
		Others
		Deceased
Village	Nominal	Village
Tambol	Nominal	Tambol
Amphoe	Nominal	Amphoe
Province	Nominal	Province
Thai	Nominal	Thai
Math	Nominal	Math
Science	Nominal	Science
Social	Nominal	Social
Health	Nominal	Health
Art	Nominal	Art
Career and Technology	Nominal	Career and Technology
Foreign language	Nominal	Foreign language
Type of course	Nominal	Commercial
		industry
Graduate (Class)	Nominal	successfully
		unsuccessful

ตารางที่ 2 จำนวนข้อมูลก่อนการทำเหมืองข้อมูล

คลาส	ข้อมูลจริง	ข้อมูลหลังการเลือกตัวแปร	ข้อมูลหลังการเลือกตัวแปร และคัดกรองค่าผิดปกติ
สำเร็จ	617	617	580
ไม่สำเร็จ	312	312	253
รวม	929	929	833

3.3 การสร้างแบบจำลอง

การสร้างแบบจำลองพยากรณ์การสำเร็จการศึกษาของผู้ที่สนใจเข้าศึกษาต่อสถานศึกษาอาชีวศึกษาเอกชนในจังหวัดกาฬสินธุ์ เพื่อช่วยในการตัดสินว่าเมื่อเข้าศึกษาต่อแล้วจะสามารถสำเร็จการศึกษาได้หรือไม่ด้วยตัวแปรทั้ง 21 ตัวแปรที่ได้จากการสำรวจข้อมูลจำนวนนักเรียนในปีการศึกษา 2557-2559 นั้น ด้วยการทำเหมืองข้อมูลและเทคนิคที่นำมาใช้ในการสร้างแบบจำลองมีจำนวน 4 เทคนิค คือ

1. เทคนิค Decision Table
2. เทคนิค Decision Rule
3. เทคนิค Repeated Incremental Pruning to Produce Error Reduction (RIPPER)
4. เทคนิค OneR

3.4 การวัดประสิทธิภาพของแบบจำลอง

ในการวัดประสิทธิภาพของแบบจำลองนั้น ผู้วิจัยได้ทำการทดลองแบ่งชุดข้อมูลเป็น 2 ชุด คือ ข้อมูลชุดสอน และข้อมูลชุดทดสอบ ด้วยหลักการ 10-fold cross validation และวัดค่าแต่ละเทคนิคด้วย Precision, Recall และ F-measure เพื่อแสดงถึงประสิทธิภาพของแบบจำลอง

การวิเคราะห์ประสิทธิภาพ คือ การวัดประสิทธิภาพการทำงานในแต่ละขั้นตอนวิธี สามารถวัดได้จากผลของการจำแนกกลุ่มข้อมูล โดยค่าของผลลัพธ์ที่ได้จากการจำแนกคือ ค่า True Positive (TP) ค่า True Negative (TN) ค่า False Positive (FP) ค่า False Negative (FN) และค่าความถูกต้อง (Precision) คือ ค่าความถูกต้องในการพยากรณ์ของคลาสของนักเรียนที่สำเร็จ

การศึกษา ต่อจำนวนทั้งหมดในกลุ่มของนักเรียนที่พยากรณ์ว่าสำเร็จการศึกษา ค่าความเที่ยงตรงที่ดีควรเป็นค่าที่สูง สามารถหาได้จากสมการ

$$\text{Precision} = \frac{TP}{TP + FP} \qquad (4)$$

ค่าเรียกคืน (Recall) คือ ค่าความถูกต้องในการพยากรณ์ของคลาสของนักเรียนที่สำเร็จการศึกษา ต่อจำนวนทั้งหมดที่สำเร็จการศึกษาจริง ค่าเรียกคืนที่ดีควรเป็นค่าที่สูง สามารถหาได้จากสมการ

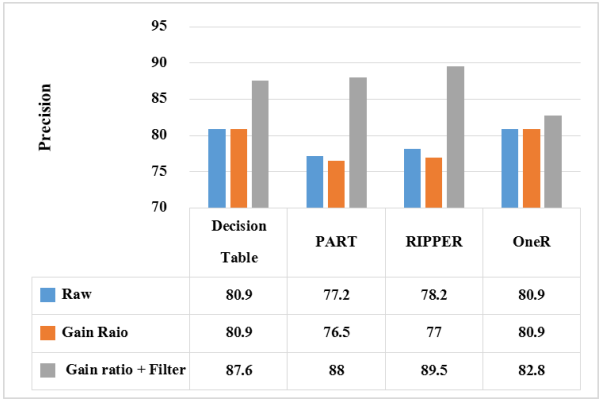
$$\text{Recall} = \frac{TP}{TP + FN} \qquad (5)$$

และค่า F-measure คือ ค่าที่เกิดจากการเปรียบเทียบระหว่างค่าความเที่ยงตรง และเรียกคืน สามารถหาได้จากสมการ

$$\text{F - measure} = \frac{2 \times (\text{Precision} \times \text{Recall})}{\text{Precision} + \text{Recall}} \qquad (6)$$

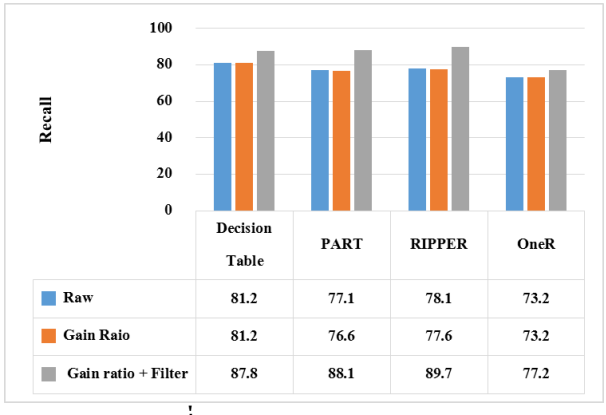
4. ผลการดำเนินการวิจัย

การวัดประสิทธิภาพของแบบจำลองนี้คณะผู้วิจัยได้นำเอาโปรแกรม WEKA เวอร์ชั่น 3.8.1 มาเป็นเครื่องมือมาใช้ในการสร้างแบบลองด้วยเทคนิค ใช้เทคนิค Decision Table, Decision Rule, RIPPER และ OneR โดยใช้ Gain Ratio และ Filter ในการคัดเลือกแอตทริบิวต์ การวัดประสิทธิภาพของแบบจำลองด้วย Precision, Recall และ F-measure ได้ผลการทดลองดัง ภาพที่ 1, ภาพที่ 2 และภาพที่ 3



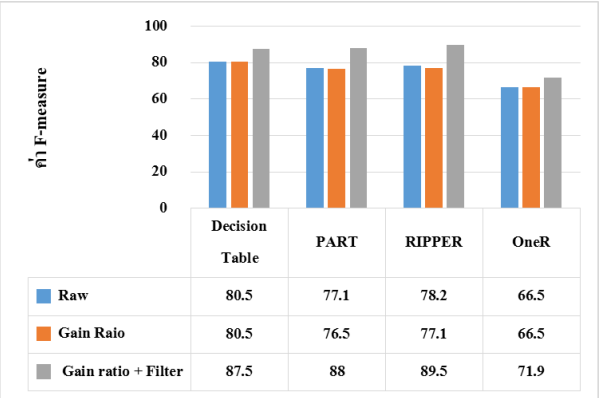
ภาพที่ 1 ค่า Precision ของกฎการพยากรณ์

ภาพที่ 1 แสดงให้เห็นค่า Precision หลังจากคัดเลือกตัวแปรในเทคนิค Decision Table และ OneR มีค่าเท่าเดิม 80.90% ส่วนเทคนิค Decision Rule และ RIPPER มีค่าลดลง 0.7% และ 1.2% และเมื่อใช้ การคัดเลือกตัวแปรและคัดกรองค่าผิดปกติออกแล้วประสิทธิภาพสูงขึ้นอีก 6.7%, 11.5%, 12.5% และ 1.9% ผลพบว่าเทคนิคแสดงให้เห็นค่า Precision ของแบบจำลองได้แก่ เทคนิค RIPPER ให้ค่า Precision มากที่สุด 89.50% ต่อมาเทคนิค Decision Rule ให้ค่า Precision 88% เทคนิค Decision Table ให้ค่า Precision 87.60% เทคนิค OneR ให้ค่า Precision 82.80% ตามลำดับ



ภาพที่ 2 ค่า Recall ของกฎการพยากรณ์

ภาพที่ 2 แสดงให้เห็นค่า Recall หลังจากคัดเลือกตัวแปรในเทคนิค Decision Table และ OneR มีค่าเท่าเดิม 81.20% และ 73.20% ส่วน เทคนิค Decision Rule และ RIPPER มีค่าลดลง 0.5% และเมื่อใช้ การคัดเลือกตัวแปรและคัดกรองค่าผิดปกติออกแล้วประสิทธิภาพสูงขึ้นอีก 6.6%, 11.5%, 12.1% และ 4% ผลพบว่าเทคนิคแสดงให้เห็นค่า Recall ของแบบจำลองได้แก่ เทคนิค RIPPER ให้ค่า Recall มากที่สุด 89.70% ต่อมา เทคนิค Decision Rule ให้ค่า Recall 88.10% เทคนิค Decision Table ให้ค่า Recall 87.8% เทคนิค OneR ให้ค่า Recall 77.20% ตามลำดับ



ภาพที่ 3 ค่า F-measure ของกฎการพยากรณ์

ภาพที่ 3 แสดงให้เห็นค่า F-measure หลังจากคัดเลือกตัวแปรในเทคนิค Decision Table และ OneR มีค่าเท่าเดิม 80.50% และ 66.50% ส่วน เทคนิค Decision Rule และ RIPPER มีค่าลดลง 0.6% และ 1.1% และเมื่อใช้ การคัดเลือกตัวแปรและคัดกรองค่าผิดปกติออกแล้วประสิทธิภาพสูงขึ้นอีก 7%, 11.5%, 12.4% และ 5.4% ผลพบว่าเทคนิคแสดงให้เห็นค่า F-measure ของแบบจำลองได้แก่เทคนิค RIPPER ให้ค่า F-measure มากที่สุด 89.50% ต่อมา เทคนิค Decision Rule ให้ค่า F-measure 88% เทคนิค Decision Table ให้ค่า F-measure 87.5% เทคนิค OneR ให้ค่า F-measure 71.9% ตามลำดับ

5. สรุป

การเพิ่มประสิทธิภาพของแบบจำลองกฎพยากรณ์ (Rule Based) โดยการใช้ Gain Ratio เพื่อเลือกตัวแปร และคัดเลือกข้อมูลที่มีค่าผิดปกติ (outliers) ออก โดยแบบจำลองที่ใช้ในการเปรียบเทียบ 4 เทคนิค ได้แก่ เทคนิค Decision Table, Decision Rule, RIPPER และ OneR ผลการทดลองพบว่า การใช้ Gain Ratio ร่วมกับการคัดกรองข้อมูล สามารถทำให้ประสิทธิภาพของแบบจำลองที่ได้จากเทคนิค RIPPER มีค่า Precision สูงขึ้น 12.50 % ค่า Recall เพิ่มขึ้น 12.10% และค่า F-measure เพิ่มขึ้น 12.40% ซึ่งเทคนิค RIPPER นี้สามารถสร้างแบบจำลองที่มีค่า Precision (89.50%), Recall (89.70%) และ F-measure (89.50%) สูงกว่าเทคนิค Decision Table, Decision Rule และ OneR

เอกสารอ้างอิง

[1] ปวีณา ชัยวนารมย์, "การพัฒนาแบบจำลองเพื่อพยากรณ์การเกิดความเครียดในหลายระดับด้วยเทคนิคการทำเหมืองข้อมูล," in การประชุมวิชาการระดับชาติมหาวิทยาลัยเทคโนโลยีราชมงคลสุวรรณภูมิ ครั้งที่ *1*, 2016, pp. 406-416.

[2] เสถียร วิชาเรือง, "การประยุกต์ใช้เทคนิคการทำเหมืองข้อมูลในกระบวนการอนุมัติสินเชื่อ," มหาวิทยาลัยธุรกิจบัณฑิตย์, 2553.

[3] V. Sukmak and J. Thongkam, "IMPROVING QUALITY OF BREAST CANCER DATA THROUGH PRE-PROCESSING," *KKU Engineering Journal*, vol. 40, pp. 493-504, 2013.

[4] T. Jonsdottira, E. T. Hvannberg, H. Sigurdsson, and S. Sigurdsson, " The feasibility of constructing a predictive outcome model for breast cancer using the tools of data mining," *Expert Systems with Applications*, vol. 34, pp. 108-118, 2008.

[5] เรวดี นามทองดี, "การอาชีวศึกษากับการพัฒนาศักยภาพแรงงานไทยในการเป็นประชาคมอาเซียน," *Veridian E-Journal, Silpakorn University* ฉบับภาษาไทย สาขามนุษยศาสตร์ สังคมศาสตร์ และศิลปะ, vol. 8, pp. 669-681, 2558.

[6] ประชาคม จันทรชิต, "คุณลักษณะที่จำเป็นของผู้สำเร็จการศึกษาในระดับประกาศนียบัตรวิชาชีพ (ปวช.) สาขาวิชาพณิชยการ ที่สถานประกอบการใน 5 จังหวัดชายแดนภาคใต้พึงประสงค์," วิทยาลัยอาชีวศึกษาปัตตานี, ปัตตานี, 2543.

[7] J. Han and M. Kamber, *Data Mining Concepts and Techniques*: Morgan Kaufmann, 2006.

[8] R. Kohavi, "The Power of Decision Tables," presented at the In: 8th European Conference on Machine Lerning, 1995.

[9] E. Frank and I. H, "Witten:Generating Accurate Rule Sets Without Global Optimization," presented at the In: Fifteenth International Conference on Machine Learning, 1998.

[10] R. C. Holte, "Very Simple Classification Rules Perform well on Most Coฏmmonly used datasets," *Machine Learning*, vol. 11, pp. 63-91, 1993.

[11] J. R. Quinlan, *C4.5:Programs for Machine Learning*: Morgan Kaufmann Publishers, 1993.

[12] S. Vilailuck, V. Jaroenpuntaruk, and D. Wichadakul, "Utilizing Data Mining Techniques to Forecast Student Academic Achievement of Kasetsart University Laboratory School Kamphaeng Saen Campus Educational Research and Development Center," *Veridian E-Journal, Science and Technology Silpakorn University*, vol. 2, pp. 1-17, 2015.

การวิเคราะห์ปัจจัยการเกิดโรคไตเรื้อรังของผู้ป่วยโรคเบาหวาน โดยใช้เทคนิคการทำเหมืองข้อมูล

Factor Analysis of Chronic Kidney Disease for Diabetics Using Data Mining Techniques

กฤษดา เครือวัลย์ (Krisada Khrueawan)[1] มัลลิกา วัฒนะ (Monlica Wattana)[2]
จักรชัย โสอินทร์ (Chakchai So-In)[3] นันท์นภัส เบญจมาศ (Nunnapus Benjamas)[4]
ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น
[1]krisada.k@kkumail.com, [2]monlwa@kku.ac.th, [3]chakso@kku.ac.th, [4]nunnapus@kku.ac.th

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อวิเคราะห์ปัจจัยการเกิดโรคไตเรื้อรังของผู้ป่วยเบาหวาน โดยใช้เทคนิคการทำเหมืองข้อมูล โดยใช้ข้อมูลประวัติการรักษาและข้อมูลผลการตรวจทางห้องปฏิบัติการผู้ป่วยเบาหวาน โรงพยาบาลส่งเสริมสุขภาพตำบลในเขตอำเภอชุมพลบุรี จังหวัดสุรินทร์ ปี 2558 ซึ่งข้อมูลที่ได้มีลักษณะไม่สมดุลกันของชุดข้อมูล ดังนั้นใช้เทคนิค SMOTE เพื่อแก้ไขปัญหาการไม่สมดุล ผลการทดลองพบว่าปัจจัยที่ถูกคัดเลือกมีทั้งหมด 10 ปัจจัย ได้แก่ เพศ อายุ ระยะเวลาการเป็นโรคเบาหวาน ระดับน้ำตาลสะสม ระดับน้ำตาลในเลือด ค่าครีตินิน ค่าคลอเรสเตอรอล ค่าไตรกลีเซอไรด์ ค่าไขมันที่มีความหนาแน่นต่ำ และค่าไขมันที่มีความหนาแน่นสูง ผลการทดลองด้วยเทคนิคข่ายงานประสาทเทียมให้ค่าความถูกต้องสูงสุดคือ 97.10% และค่าเฉลี่ยเรขาคณิต (G-Mean) คือ 97% เทคนิคต้นไม้ตัดสินใจให้ค่าความถูกต้องที่ 90.54% และค่าเฉลี่ยเรขาคณิตคือ 91% และเทคนิคการเรียนรู้แบบเบย์ให้ค่าความถูกต้องที่ 85.20% และค่าเฉลี่ยเรขาคณิตคือ 85%

คำสำคัญ: การวิเคราะห์ปัจจัย โรคเบาหวาน การทำเหมืองข้อมูล ต้นไม้ตัดสินใจ การเรียนรู้แบบเบย์ ข่ายงานประสาทเทียม

Abstract

The main objective of this research is to analyze factors of Chronic Kidney Disease in Diabetics by using data mining techniques. The study included the factors in chronic kidney disease of diabetic patients which were performed at Chumphonburi District Health Promoting Hospital, Surin Province during the year 2015. The data given by the subjects is imbalance so SMOTE technique was used to solve the problem of that imbalance data. There are 10 analyzed factors were selected: Age, Gender, Duration of diabetes, HbA1c, FBS, Creatinine, Cholesterol, Triglyceride, LDL and HDL. The results showed that the accuracy of Artificial Neural Network technique was 97.10% and the G-Mean value was 97%. The accuracy of Decision Tree technique was 90.54% and the G-Mean value was 91%. Moreover, the accuracy of Naïve Bayesian technique was 85.20% and the G-Mean value was 85%.

Keyword: Factor Analysis, Diabetics, Data Mining, Decision Tree, Naïve Bayesian Learning, Neural Network.

1. บทนำ

เบาหวานเป็นโรคเรื้อรังและเป็นโรคที่เป็นปัญหาทางสาธารณสุขของประเทศก่อให้เกิดภาวะแทรกซ้อนในหลายระบบของร่างกาย ส่งผลกระทบต่อผู้ป่วยและครอบครัว สิ่งสำคัญของการจัดการโรคเบาหวานคือการค้นหาโรคตั้งแต่ระยะเริ่มแรกและการดูแลรักษาเพื่อชะลอการเกิดภาวะแทรกซ้อน [1] ภาวะโรคแทรกซ้อนของเบาหวานเป็นผลสืบเนื่องมาจากระดับน้ำตาลในกระแสเลือดมีปริมาณสูงขึ้น เมื่อระดับน้ำตาลที่สูงติดต่อกันเป็นเวลานานจะทำลายหลอดเลือด และทำให้เกิดอาการแทรกซ้อนต่อระบบอวัยวะต่างๆ เช่น มีผลต่อเส้นเลือดที่

ไต ทำให้การทำงานของไตผิดปกติ ไตวาย ภาวะไตเสื่อมจากโรคเบาหวาน ผู้ป่วยเบาหวานที่มีระดับน้ำตาลในเลือดสูงเป็นระยะเวลานาน ทำให้เกิดความผิดปกติที่หลอดเลือดไปยังไต ซึ่งเป็นผลให้การทำงานของไตผิดปกติ ไม่สามารถกรองของเสียได้ [2]

ดังนั้นจากปัญหาดังกล่าว งานวิจัยนี้จึงนำเสนอการวิเคราะห์ปัจจัยความเสี่ยงของการเกิดโรคไตเรื้อรังจากเบาหวานโดยเทคนิคการทำเหมืองข้อมูลซึ่งใช้วิธีการจำแนกประเภทข้อมูลประกอบด้วยเทคนิคต้นไม้ตัดสินใจ เทคนิคการเรียนรู้แบบเบย์ และเทคนิคข่ายงานประสาทเทียม เพื่อศึกษาปัจจัยและสร้างแบบจำลอง และทดสอบประสิทธิภาพความถูกต้องโดยวิธี K-fold Cross-validation ในการทดลองใช้ข้อมูลผู้ป่วยเบาหวานโรงพยาบาลส่งเสริมสุขภาพตำบลในเขตอำเภอชุมพลบุรี จังหวัดสุรินทร์

2. วรรณกรรมที่เกี่ยวข้อง

2.1 ต้นไม้ตัดสินใจ (Decision Tree)

การเรียนรู้ของต้นไม้ตัดสินใจ (Decision Tree) [5] เป็นการเรียนรู้โดยการจำแนกประเภท (Classification) ข้อมูลออกเป็นกลุ่ม (Class) ต่างๆ โดยใช้คุณลักษณะ (Attribute) ข้อมูลในการจำแนกประเภทต้นไม้ตัดสินใจที่ได้จากการเรียนรู้ทำให้ทราบว่า คุณลักษณะใดเป็นตัวกำหนดการจำแนกประเภท และคุณลักษณะแต่ละตัวมีความสำคัญมากน้อยต่างกันอย่างไร เพราะฉะนั้น การจำแนกประเภทมีประโยชน์ช่วยให้สามารถวิเคราะห์ข้อมูลและตัดสินใจได้ถูกต้องยิ่งขึ้น

2.2 วิธีการเรียนรู้เบย์อย่างง่าย (Naïve Bayesian Learning)

ตัวจำแนกแบบเบย์อย่างง่าย (Naïve Bayesian Classifier) [5] คือโมเดลการจำแนกประเภทข้อมูล ที่ใช้หลักความน่าจะเป็นซึ่งอยู่บนพื้นฐานของ Bayes' Theorem และสมมติฐานที่ให้การเกิดของเหตุการณ์ต่างๆ เป็นอิสระต่อกัน (Independence) วิธีการของ Naïve Bayesian คือ การใช้วิธีการของเบย์พร้อมสมมติฐานของการเป็นอิสระต่อกันของตัวแปรอิสระทุกตัว

2.3 โครงข่ายประสาทเทียม (Artificial Neural Network)

โครงข่ายประสาทเทียม (Artificial Neural Network) [5] หรือที่มักจะเรียกสั้นๆว่า ข่ายงานประสาท (Neural Network หรือ Neural Net) คือ โมเดลทางคณิตศาสตร์ สำหรับประมวลผลสารสนเทศด้วยการคำนวณแบบคอนเนคชันนิสต์ (Connectionist) เพื่อจำลองการทำงานของเครือข่ายประสาทในสมองมนุษย์ด้วยวัตถุประสงค์ที่จะสร้างเครื่องมือซึ่งมีความสามารถในการเรียนรู้การจดจำรูปแบบ (Pattern Recognition) และการอุปมานความรู้ เช่นเดียวกับความสามารถที่มีในสมองมนุษย์

2.4 การสุ่มเพิ่มจำนวนตัวอย่างกลุ่มน้อย (Synthetic Minority Over-sampling Technique – SMOTE)

เป็นเทคนิคที่ช่วยแก้ปัญหาของข้อมูลที่ไม่สมดุล โดยจะสุ่มเพิ่มจำนวนข้อมูลในกลุ่มเสียงข้างน้อย (Minority Class) ให้มีปริมาณข้อมูลใกล้เคียงหรือเท่ากับข้อมูลในเสียงข้างมาก (Majority Class) เมื่อนำไปจำแนกประเภท ซึ่งช่วยให้ผลลัพธ์ในการสังเคราะห์กลุ่มข้างน้อยดีขึ้น [3]

2.5 การคัดเลือกปัจจัย

Attribute Selection [4] เป็นวิธีการคัดเลือกแอตทริบิวต์ ที่สำคัญในการสร้างโมเดล แบ่งได้เป็น 2 แบบ

1. Filter approach เป็นการคำนวณค่าน้ำหนัก (หรือค่าความสัมพันธ์) ของแต่ละแอตทริบิวต์และเลือกเฉพาะแอตทริบิวต์ที่สำคัญเก็บไว้

2. Wrapper approach เป็นการคำนวณค่าน้ำหนักโดยใช้โมเดลวิธีการจำแนกประเภทข้อมูล เป็นตัววัดประสิทธิภาพของแอตทริบิวต์ ประกอบด้วย 3 วิธี คือ

2.1 Forward Selection เป็นวิธีการเลือกแอตทริบิวต์ใส่เข้าไปหรือถอดออกมาเพื่อสร้างโมเดล และเลือกเซตของแอตทริบิวต์ที่ดีไว้ใช้

2.2 Backward Elimination เป็นวิธีการใช้แอตทริบิวต์ทั้งหมดและตัดแอตทริบิวต์ออกไปทีละ 1 ตัว เพื่อคัดเลือกแอตทริบิวต์ที่มีความสำคัญเก็บไว้ ถ้าแอตทริบิวต์ที่ตัดออกไปให้ค่าประสิทธิภาพ (Performance) ดีขึ้นก็จะตัดแอตทริบิวต์นี้ทิ้ง แต่ถ้าแอตทริบิวต์ที่ตัดออกไปให้ค่าประสิทธิภาพ แย่ลงก็จะเก็บแอตทริบิวต์นี้ไว้

2.3 Evolutionary Selection เป็นวิธีการสุ่มเลือกแอตทริบิวต์ขึ้นมาและวัดประสิทธิภาพ แล้วคัดเลือกแอตทริบิวต์ที่มีประสิทธิภาพและสุ่มเลือกตัวอื่นเพิ่มขึ้นมา

2.6 งานวิจัยที่เกี่ยวข้อง

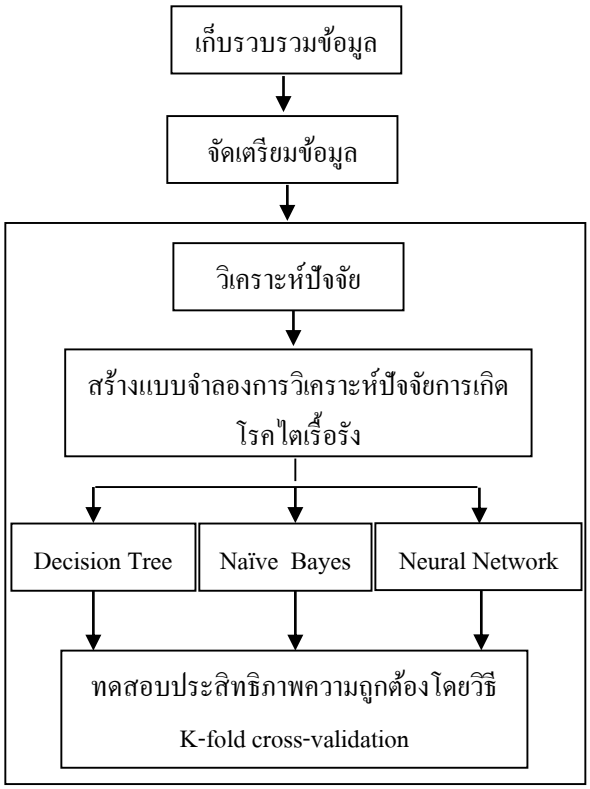
งานวิจัย [10] ได้ทำนายการเกิดโรคไตเรื้อรัง ระยะที่ 3 ถึง 5 โดยใช้วิธีการทำเหมืองข้อมูล เทคนิคต้นไม้ตัดสินใจ ให้ค่าความถูกต้องสูงที่สุด ซึ่งสอดคล้องกับงานวิจัย [6] ได้สร้างแบบจำลองเพื่อพยากรณ์การเกิดแผลที่เท้าของผู้ป่วยโรคเบาหวาน โดยใช้เทคนิคการทำเหมืองข้อมูล เทคนิคต้นไม้ตัดสินใจ ให้ค่าประสิทธิภาพสูงที่สุด ในงานวิจัย [7] ได้วิเคราะห์หาปัจจัยพฤติกรรมที่มีผลต่อการรักษาวัณโรคปอด โดยใช้เทคนิคการทำเหมืองข้อมูล เทคนิคต้นไม้ตัดสินใจ ให้ค่าประสิทธิภาพสูงที่สุด ซึ่งแตกต่างจากงานวิจัย [9] ได้วิเคราะห์การเกิดโรคไตเรื้อรัง โดยใช้เทคนิคการทำเหมืองข้อมูล ด้วยวิธีการจำแนกประเภทข้อมูล เทคนิควิธีการเรียนรู้เบย์อย่างง่าย ให้ค่าความถูกต้องสูงที่สุด สำหรับงานวิจัย [11] ได้พยากรณ์โรคเบาหวานด้วยเทคนิคการหากฎความสัมพันธ์วิธี Apriori และ FP-Growth นอกจากนี้งานวิจัย [8] ได้วิเคราะห์ความเสี่ยงการเป็นโรคไวรัสตับอักเสบซี โดยใช้ต้นไม้ตัดสินใจและทฤษฎีเบย์เซี่ยน ทฤษฎีเบย์เซี่ยน ให้ค่าประสิทธิภาพสูงที่สุด

3. วิธีการดำเนินการวิจัย

งานวิจัยนี้ได้ศึกษาปัจจัยและสร้างแบบจำลองการวิเคราะห์ปัจจัยความเสี่ยงของการเกิดโรคไตเรื้อรังจากเบาหวาน โดยใช้วิธีการจำแนกประเภทข้อมูล ประกอบด้วยเทคนิคต้นไม้ตัดสินใจ เทคนิคการเรียนรู้แบบเบย์ และเทคนิคข่ายงานประสาทเทียม มีขั้นตอนในการดำเนินงานวิจัยดังภาพที่ 1

3.1 การรวบรวมข้อมูล (Data Collection)

เก็บรวบรวมข้อมูลประวัติการรักษาผู้ป่วยเบาหวานและข้อมูลผลการตรวจทางห้องปฏิบัติการผู้ป่วยเบาหวาน จากระบบฐานข้อมูลโรงพยาบาลส่งเสริมสุขภาพตำบลในเขตอำเภอชุมพลบุรี จังหวัดสุรินทร์ ปี 2558 จำนวน 1,249 ราย ผู้ป่วยเบาหวานที่มีภาวะแทรกซ้อนทางไต จำนวน 89 ราย ประกอบด้วยปัจจัยในตารางที่ 1 ซึ่งปัจจัยเหล่านี้ได้รับคำแนะนำจากแพทย์ผู้เชี่ยวชาญแล้วว่าเป็นปัจจัยที่มีผลต่อการเกิดภาวะแทรกซ้อนทางไตของผู้ป่วยเบาหวาน



ภาพที่ 1: ขั้นตอนการดำเนินงานวิจัย

ตารางที่ 1: ปัจจัยการเกิดโรคไตเรื้อรังของผู้ป่วยโรคเบาหวาน

ลำดับ	ปัจจัย	คำอธิบาย
1	Gender	เพศ
2	Age	อายุ
3	HbA1c	ระดับน้ำตาลสะสม
4	Fasting blood sugar (FBS)	ระดับน้ำตาลในเลือด
5	Creatinine	ค่าครีตินิน
6	Cholesterol	ค่าคลอเรสเตอรอล
7	Trigleceride	ค่าไตรกลีเซอไรด์
8	LDL	ค่าไขมันที่มีความหนาแน่นต่ำ
9	HDL	ค่าไขมันที่มีความหนาแน่นสูง
10	Duration of diabetes	ระยะเวลาการเป็นโรคเบาหวาน
11	Class	ภาวะแทรกซ้อน CKD = เกิดโรคไตเรื้อรัง NotCKD = ไม่เกิดโรคไตเรื้อรัง

3.2 การเตรียมข้อมูล (Data Preprocessing)

เนื่องจากข้อมูลผู้ป่วยเบาหวานทั้งหมด 1,249 ราย เป็นข้อมูลที่ยังไม่สมบูรณ์ การทำข้อมูลให้สมบูรณ์โดยนำข้อมูลที่ได้ตรวจสอบข้อมูลที่ขาดหายหรือข้อมูลที่เป็นค่าว่างตั้งแต่ 2 ค่าขึ้นไปทำการลบชุดข้อมูลนั้นทิ้ง ทำให้มีชุดข้อมูลทั้งหมด 681 ชุดข้อมูล แอตทริบิวต์ที่มีค่าข้อมูลขาดหาย ได้แก่แอตทริบิวต์ HbA1c, FBS, Creatinine, LDL, HDL ซึ่งเป็นแอตทริบิวต์ที่เป็นข้อมูลเชิงตัวเลขจะแทนที่ค่าข้อมูลด้วยค่าเฉลี่ย (mean) โดยวิธี Replace Missing Values จากนั้นแปลงค่าข้อมูลให้อยู่ในรูปแบบเดียวกันเพื่อให้คอมพิวเตอร์สามารถนำไปประมวลผลได้ เช่น เพศชาย เป็น Male, เพศญิง เป็น Female และสุ่มเพิ่มกลุ่มตัวอย่างกลุ่มน้อย (SMOTE) ซึ่งเป็นขั้นตอนวิธีที่ใช้ในการแก้ไขปัญหาของชุดข้อมูลที่ไม่สมดุลกัน ซึ่งผู้วิจัยได้ใช้โปรแกรม WEKA เวอร์ชั่น 3.8.1 เพิ่มกลุ่มตัวอย่าง Class CKD (เกิดโรคไตเรื้อรัง) จาก 10ชุดข้อมูล เป็น 640 ชุดข้อมูล โดยเพิ่ม Class CKD ให้มีค่าใกล้เคียง Class NotCKD (ไม่เกิดโรคไตเรื้อรัง) มีจำนวน 671 ชุดข้อมูล ทำให้ชุดข้อมูลที่ใช้ในการทดลองนี้มีทั้งสิ้น 1,311 ชุดข้อมูล ใช้ค่า nearestNeighbors เท่ากับ 5 จะได้ชุดข้อมูลทั้งหมด ชุดที่มีจำนวนเท่าๆกัน 5

3.3 การวิเคราะห์ปัจจัย (Factor Analysis)

ข้อมูลที่ใช้วิเคราะห์ปัจจัยการเกิดโรคไตเรื้อรังของผู้ป่วยโรคเบาหวานทั้งหมด 1,311 ชุดข้อมูล ทำการคัดเลือกแอตทริบิวต์แบบ Wrapper approach ซึ่งเป็นวิธีคัดเลือกแบบคำนวณค่าน้ำหนักโดยใช้โมเดลวิธีการจำแนกประเภทข้อมูลเป็นตัววัดประสิทธิภาพของแอตทริบิวต์ ประกอบด้วย 3 วิธี คือ Forward Selection, Backward Elimination, Evolutionary Selection โดยใช้โปรแกรม RapidMiner Studio เวอร์ชั่น 7.5.0 ในการคัดเลือกแอตทริบิวต์

3.4 การสร้างโมเดล (Modeling)

นำปัจจัยที่ถูกคัดเลือกด้วยวิธี Optimize Weights (Backward) มาสร้างแบบจำลองการวิเคราะห์ปัจจัยความเสี่ยงของการเกิดโรคไตเรื้อรังจากเบาหวาน โดยใช้โปรแกรม RapidMiner Studio เวอร์ชั่น 7.5.0 ใช้วิธีการจำแนกประเภทข้อมูลประกอบด้วย เทคนิคต้นไม้ตัดสินใจ เทคนิคการเรียนรู้แบบเบย์ และเทคนิคข่ายงานประสาทเทียม

3.5 การทดสอบประสิทธิภาพ (Model Evaluation)

นำแบบจำลองการวิเคราะห์ปัจจัยความเสี่ยงของการเกิดโรคไตเรื้อรังจากเบาหวาน ไปทดสอบประสิทธิภาพโดยวิธี K-fold Cross validation กำหนดค่า k เท่ากับ 10 และประเมินค่าความถูกต้อง (Accuracy) โดยใช้ผลจากตาราง ตาราง Confusion Matrix ดังตารางที่ 2

ตารางที่ 2: ตาราง Confusion Matrix ของข้อมูลปัจจัยการเกิดโรคไตเรื้อรังของผู้ป่วยโรคเบาหวาน ซึ่งมี 2 คลาส

predict / actual	CKD	NotCKD
pred. CKD	TP	FP
pred. NotCKD	FN	TN

สำหรับการประเมินค่าความถูกต้อง (Accuracy) ดังสมการที่ 1 นอกจากนี้ยังมีการคำนวณ Geometric Mean (G-Mean) ซึ่งเป็นค่าการสร้างความสมดุลของข้อมูลเพื่อให้การทำนายไม่เอนเอียงไปทางฝั่งที่มีข้อมูลมากกว่า เพื่อให้การทำนายนั้นสามารถใช้ข้อมูลจากทั้ง 2 คลาสมาสร้างรูปแบบการเรียนรู้ได้ดี จากค่าที่ได้จากตาราง Confusion Matrix จึงนำมาหาค่า G-Mean ตามสมการที่ 2 เพื่อใช้ในการวัดความสมดุลของข้อมูลในชุดข้อมูลฝึกสอนและข้อมูลทดสอบ โดยที่ตัวชี้วัดจะแสดงถึงความสมดุลทั้งชุดข้อมูลฝึกสอนและชุดข้อมูลทดสอบ เมื่อนำกลุ่มข้อมูลที่มากและกลุ่มของข้อมูลที่น้อยมาเปรียบเทียบ

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \quad (1)$$

$$G - Mean = \sqrt{Sensitivity * Specificity} \quad (2)$$

$$Sensitivity = \frac{TP}{TP + FN} \quad (3)$$

$$Specificity = \frac{TN}{TN + FP} \quad (4)$$

Sensitivity คือ ค่าที่บอกว่าการคำนวณข้อมูลมีผลลัพธ์ถูกต้องโดยพิจารณาจากข้อมูลของผลลัพธ์เดียวกัน คำนวณได้จากสมการที่ 3 และ Specificity คือค่าที่บอกว่าผลลัพธ์ที่ได้ว่า

ไม่จริงเป็นอัตราส่วนเท่าไรของผลลัพธ์จริงทั้งหมด คำนวณได้จากสมการที่ 4

4. ผลการดำเนินงาน

งานวิจัยนี้ได้ทำการวิเคราะห์ปัจจัยด้วยวิธีการคัดเลือกแอตทริบิวต์ เพื่อเลือกหาชุดปัจจัยที่เหมาะสมที่สุด โดยใช้เทคนิคการทำเหมืองข้อมูล คือ เทคนิคต้นไม้ตัดสินใจ เทคนิคการเรียนรู้แบบเบย์ และเทคนิคข่ายงานประสาทเทียม จากผลการทดสอบวัดประสิทธิภาพเปรียบเทียบของวิธีการคัดเลือกแอตทริบิวต์ต่างๆ พบว่าวิธี Optimize Weights (Backward) ให้ค่าความถูกต้องสูงสุด คัดเลือกได้ทั้งสิ้น 10 ปัจจัย ได้แก่ เพศ อายุ ระยะเวลาการเป็นโรคเบาหวาน ระดับน้ำตาลสะสม ระดับน้ำตาลในเลือด ค่าครีตินิน ค่าคลอเรสเตอรอล ค่าไตรกลีเซอไรด์ ค่าไขมันที่มีความหนาแน่นต่ำ และค่าไขมันที่มีความหนาแน่นสูง ดังตารางที่ 3

ตารางที่ 3: ปัจจัยที่ถูกคัดเลือกด้วยวิธี Optimize Weights (Backward) และค่าน้ำหนัก

ลำดับ	ชื่อปัจจัย	ค่าน้ำหนัก
1	เพศ	1
2	อายุ	1
3	ระยะเวลาการเป็นโรคเบาหวาน	1
4	ระดับน้ำตาลสะสม	1
5	ระดับน้ำตาลในเลือด	1
6	ค่าครีตินิน	1
7	ค่าคลอเรสเตอรอล	1
8	ค่าไตรกลีเซอไรด์	1
9	ค่าไขมันที่มีความหนาแน่นต่ำ	1
10	ค่าไขมันที่มีความหนาแน่นสูง	1

จากตารางที่ 3 นำปัจจัยที่ถูกคัดเลือกด้วยวิธี Optimize Weights (Backward) มาสร้างแบบจำลองการวิเคราะห์ปัจจัยความเสี่ยงของการเกิดโรคไตเรื้อรังจากเบาหวาน ด้วยเทคนิคการทำเหมืองข้อมูล ผลการทดลองพบว่าเทคนิคข่ายงานประสาทเทียมให้ค่าความถูกต้องสูงสุดคือ 97.10% และค่าเฉลี่ยเรขาคณิต คือ 97% เทคนิคต้นไม้ตัดสินให้ค่าความถูกต้องที่ 90.54% และค่าเฉลี่ยเรขาคณิต คือ 91% จากผลการทดลองได้เลือกเทคนิคต้นไม้ตัดสินใจ มาสร้างแบบจำลองการวิเคราะห์ปัจจัยความเสี่ยงของการเกิดโรคไตเรื้อรังจากเบาหวาน ได้ผลลัพธ์โครงสร้างต้นไม้ตัดสินใจและสร้างเป็นกฎการตัดสินใจจะได้กฎซึ่งอยู่ในรูปแบบ If Then ทั้งหมด 14 กฎ ดังตารางที่ 5

ตารางที่ 4: การเปรียบเทียบประสิทธิภาพของแต่ละเทคนิค

เทคนิค	ค่าความถูกต้อง (Accuracy)	ค่า G-Mean
ต้นไม้ตัดสินใจ	90.54%	91%
การเรียนรู้แบบเบย์	85.20%	85%
ข่ายงานประสาทเทียม	97.10%	97%

ตารางที่ 5: ตัวอย่างกฎต้นไม้ตัดสินใจ

กฎต้นไม้ตัดสินใจ
กฎข้อที่ 1 : IF Creatinine > 0.850 and Gender = female and Cholesterol > 151.500 and HDL > 62.500 THEN NotCKD
กฎข้อที่ 2 : IF Creatinine > 0.850 and Gender = female and Cholesterol > 151.500 and HDL ≤ 62.500 and HbA1c > 11.250 THEN NotCKD
กฎข้อที่ 3 : IF Creatinine > 0.850 and Gender = female and Cholesterol > 151.500 and HDL ≤ 62.500 and HbA1c ≤ 11.250 and Cholesterol > 264.500 THEN NotCKD
กฎข้อที่ 4 : IF Creatinine > 0.850 and Gender = female and Cholesterol > 151.500 and HDL ≤ 62.500 and HbA1c ≤ 11.250 and Cholesterol ≤ 264.500 and FBS > 89 and Age > 42.500 and Duration of diabetes > 1.500 and LDL > 179 THEN NotCKD
กฎข้อที่ 5 : IF Creatinine > 0.850 and Gender = female and Cholesterol > 151.500 and HDL ≤ 62.500 and HbA1c ≤ 11.250 and Cholesterol ≤ 264.500 and FBS > 89 and Age > 42.500 and Duration of diabetes > 1.500 and LDL ≤ 179 and LDL > 60.500 and Creatinine > 3.550 THEN NotCKD
กฎข้อที่ 6 : IF Creatinine > 0.850 and Gender = female and Cholesterol > 151.500 and HDL ≤ 62.500 and HbA1c ≤ 11.250 and Cholesterol ≤ 264.500 and FBS > 89 and Age > 42.500 and Duration of diabetes > 1.500 and LDL ≤ 179 and LDL > 60.500 and Creatinine ≤ 3.550 and HDL > 25 THEN CKD
กฎข้อที่ 7 : IF Creatinine > 0.850 and Gender = female and Cholesterol > 151.500 and HDL ≤ 62.500 and HbA1c ≤ 11.250 and Cholesterol ≤ 264.500 and FBS > 89 and Age > 42.500 and Duration of diabetes > 1.500 and LDL ≤ 179 and LDL > 60.500 and Creatinine ≤ 3.550 and HDL ≤ 25 THEN NotCKD

ตารางที่ 5: ตัวอย่างกฎต้นไม้ตัดสินใจ (ต่อ)

กฎต้นไม้ตัดสินใจ
กฎข้อที่ 8 : IF Creatinine > 0.850 and Gender = female and Cholesterol > 151.500 and HDL ≤ 62.500 and HbA1c ≤ 11.250 and Cholesterol ≤ 264.500 and FBS > 89 and Age > 42.500 and Duration of diabetes > 1.500 and LDL ≤ 179 and LDL ≤ 60.500 THEN NotCKD
กฎข้อที่ 9 : IF Creatinine > 0.850 and Gender = female and Cholesterol > 151.500 and HDL ≤ 62.500 and HbA1c ≤ 11.250 and Cholesterol ≤ 264.500 and FBS > 89 and Age > 42.500 and Duration of diabetes ≤ 1.500 THEN NotCKD
กฎข้อที่ 10 : IF Creatinine ≤ 0.850 THEN NotCKD

5. สรุป

งานวิจัยนี้ได้นำเสนอการวิเคราะห์ปัจจัยการเกิดโรคไตเรื้อรังของผู้ป่วยโรคเบาหวาน ด้วยเทคนิคการทำเหมืองข้อมูล คัดเลือกแอตทริบิวต์โดยวิธีคัดเลือกแบบคำนวณค่าน้ำหนักโดยใช้โมเดลการจำแนกประเภทข้อมูลเป็นตัววัดประสิทธิภาพของแอตทริบิวต์ 3 วิธี คือ Optimize Weights (Forward), Optimize Weights (Backward), Optimize Weights (Evolutionary) จากผลการทดลองพบว่าวิธี Optimize Weights (Backward) ให้ค่าความถูกต้องสูงสุด คัดเลือกได้ทั้งสิ้น 10 ปัจจัย ได้แก่ อายุ เพศ ระยะเวลาการเป็นโรคเบาหวาน ระดับน้ำตาลสะสม ระดับน้ำตาลในเลือดค่าครีตินิน ค่าคลอเรสเตอรอล ค่าไตรกลีเซอไรด์ ค่าไขมันที่มีความหนาแน่นต่ำ และค่าไขมันที่มีความหนาแน่นสูง ซึ่งให้ค่าน้ำหนักเท่ากับ 1 และเทคนิคข่ายงานประสาทเทียมให้ค่าความถูกต้องสูงสุดคือ 97.10% และค่าเฉลี่ยเรขาคณิต คือ 97% เทคนิคต้นไม้ตัดสินให้ค่าความถูกต้องที่ 90.54% และค่าเฉลี่ยเรขาคณิต คือ 91%

เอกสารอ้างอิง

[1] สถาบันวิจัยและประเมินเทคโลยีสารสนเทศทางการแพทย์. เอกสารผลงานวิชาการแนวทางการตรวจคัดกรองและดูแลรักษาภาวะแทรกซ้อนทางไตในผู้ป่วยเบาหวานและความดันโลหิตสูง. (ออนไลน์). สืบค้นจาก http://www.dms.moph.go.th/imrta/data/aw20121120-2.pdf [16 ธันวาคม 2558]

[2] สถาบันวิจัยและประเมินเทคโลยีสารสนเทศทางการแพทย์เอกสาร . ผลงานวิชาการแนวทางเวชปฏิบัติการป้องกันดูแลรักษาภาวะแทรกซ้อนจากโรคเบาหวาน(ตา ไต เท้า). (ออนไลน์). สืบค้นจาก http//www.dms.moph.go.th/imrta/images/pdf-cpg/2553/53-3.pdf [16 ธันวาคม 2558]

[3] เอกสิทธิ์ พัชรวงศ์ศักดา .*การวิเคราะห์ข้อมูลด้วยเทคนิคดาต้าไมน์นิงเบื้องต้น*. กรุงเทพฯ: บริษัท เอเชีย ดิจิตอลการพิมพ์ จำกัด. 2557.

[4] เอกสิทธิ์ พัชรวงศ์ศักดา. *เอกสารประกอบการอบรม หลักสูตร Practical Data Minig with RapidMiner Studio 7*. กรุงเทพฯ: บริษัท เอเชีย ดิจิตอลการพิมพ์ จำกัด. 2560.

[5] วรารัตน์ สงฆ์แป้น. *เอกสารประกอบการสอนวิชาการทำเหมืองข้อมูล (Data Mining)*. มหาวิทยาลัยขอนแก่น. 2558.

[6] บรรจบ ดลกุล จารี ทองคำ และ วาทินี สุขมาก, “การสร้างแบบจำลองเพื่อพยากรณ์การเกิดแผลที่เท้าของผู้ป่วยโรคเบาหวานโดยใช้เทคนิคเหมืองข้อมูล,” *วารสารวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยมหาสารคาม* ปีที่ 33 ฉบับที่ 6 พ.ย.-ธ.ค. 2557 หน้า 703-710.

[7] เกศินี ชาลีพร และ สิลดา อินทรโสธรฉันท์, “การวิเคราะห์พฤติกรรมที่มีผลต่อการรักษาวัณโรคปอดโดยใช้เทคนิคการทำเหมืองข้อมูล: กรณีศึกษาสำนักงานสาธารณสุข อำเภอเมืองสกลนคร,” *ICSEC International Computer Science and Engineering Conference*, หน้า 100-105, 2556.

[8] ชลธิศา พลทองมาก และ พุธษดี ศิริแสงตระกูล, “การวิเคราะห์ความเสี่ยงการเป็นโรคไวรัสตับอักเสบซีโดยต้นไม้ตัดสินใจและทฤษฎีเบย์เซียน,” *National Conference on Information Technology*, ปีที่ 7 หน้า 558-63, 2554.

[9] Veenita Kunwar, Khushboo Chandel, A. Sai Sabitha and Abhay Bansai “Chronic Kidney Disease Analysis Using Data Mining Classification Technioques,” *6th International Conference Cloud System and Big Data Engineering (Confluence)*, pp. 300-305, 2016.

[10] Patcharaporn Panwong and Natthakan Iam-On “Predicting Transitional Interval of Kidney Disease Stages 3 to 5 Using Data Mining Method,” *Second Asian Conference on Defence Technology (ACDT)*, 2016.

[11] Sankaranarayanan.S and Pramananda Perumal.T “Diabetic prognosis through Data Mining Methods and Techniques,” *International Conference on Intelligent Computing Applications*, pp. 162-166, 2014.

การจัดเส้นทางการท่องเที่ยวที่ให้ความพึงพอใจสูงสุด ด้วยวิธีพาทิเคิลสวอมออปติไมเซชันและฟัซซีเซต

Travelling Planning for Maximize Satisfaction using Particle Swarm Optimization and Fuzzy Set

ธนกฤษ จันทร์แสง (Tanagrit Chansaeng)[1] พิทา จารุพูนผล (Pita Jarupunphol)[2]
และณรงค์เดช กีรติพรานนท์ (Narongdech Keeratipranon)[3]
[12]สาขาเทคโนโลยีสารสนเทศ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏภูเก็ต
[3]ภาควิชาวิศวกรรมคอมพิวเตอร์และโทรคมนาคม วิทยาลัยนวัตกรรมด้านเทคโนโลยีและวิศวกรรม
มหาวิทยาลัยธุรกิจบัณฑิตย์
[1]tanagrit.c@pkru.ac.th, [2]p.jarupunphol@pkru.ac.th, [3]narongdech.ken@dpu.ac.th

บทคัดย่อ

บทความนี้มีวัตถุประสงค์เพื่อนำเสนอแบบจำลองการค้นหาเส้นทางการท่องเที่ยว โดยใช้ฟัซซีเซตในการจัดประเภทการท่องเที่ยวของแต่ละแหล่งท่องเที่ยว และใช้วิธีพาทิเคิลสวอมออปติไมเซชัน ในการค้นหาเส้นทางการท่องเที่ยวเพื่อให้ได้แหล่งท่องเที่ยวที่ตรงกับประเภทที่นักท่องเที่ยวต้องการมากที่สุดและอยู่ภายใต้เงื่อนไขเวลาในการท่องเที่ยวของนักท่องเที่ยว จากผลการทดลองเปรียบเทียบการค้นหาเส้นทางโดยใช้แบบจำลองที่ใช้ฟัซซีเซตและเซตแบบทวินัยพบว่าแบบจำลองที่นำเสนอสามารถหาเส้นทางที่ตรงตามความต้องการได้อย่างมีประสิทธิภาพ โดยค่าเฉลี่ยความพึงใจอยู่ที่ 4.02 ซึ่งมีระดับความพึงพอใจอยู่ในระดับมาก

คำสำคัญ: วิธีพาทิเคิลสวอมออปติไมเซชัน ฟัซซีเซต เซตแบบทวินัย

Abstract

This paper presents a model of travelling routing for Maximize Satisfaction using fuzzy set to classify types of tourist attractions and applies Particle Swarm Optimization to find the path to all destination which matches to the type of tourist attractions. The criteria of travelling routing consists of the tourist attraction type which travelers want most and travel time. The result of the experiment shows that the travelling routing using the proposed model is efficiency compared with routing without using fuzzy set. The average satisfaction is 4.02

Keyword: Particle Swarm Optimization, Fuzzy Set.

1. บทนำ

ปัจจุบันมีการนำปัญญาประดิษฐ์มาประยุกต์ในงานวิจัยทางด้านวิศวกรรมที่เกี่ยวข้องกับการแก้ปัญหาแบบ NP Hard หรือปัญหาที่ต้องหาค่าคำตอบที่ดีที่สุดอย่างมีประสิทธิภาพ โดยทั่วไปมักใช้เทคนิคการหาคำตอบที่ดีที่สุดโดยใช้วิธีการคำนวณเชิงวิวัฒนาการ หรือ Evolutionary computation [1][2] ในการแก้ปัญหาเส้นทางเดินรถที่มีข้อจำกัดด้านกรอบเวลา (Vehicle Routing Problem with Time Windows) ทั้งนี้ปัญหาการจัดเส้นทางการเดินรถโดยมีข้อจำกัดด้านกรอบเวลาเพียงอย่างเดียวนั้น ถือว่าเป็นปัญหาที่มีความซับซ้อนมากซึ่งจัดอยู่ในประเภท NP-Hard โดยปัญหามีจำนวนตัวแปรมากและความสัมพันธ์ระหว่างตัวแปรค่อนข้างซับซ้อนซึ่งยากต่อการหาคำตอบ [3] จังหวัดภูเก็ตเป็นจังหวัดที่มีนักท่องเที่ยวให้ความนิยมมาท่องเที่ยวเป็นอันดับต้นๆ ของประเทศไทยเนื่องจากมีแหล่งท่องเที่ยวอันสวยงามและหลากหลายประเภท อาทิ แหล่งท่องเที่ยวเชิงนิเวศ แหล่งท่องเที่ยวเชิงวัฒนธรรม แหล่งท่องเที่ยวประเภทชายหาด แหล่งท่องเที่ยวประเภทจุดชมทิวทัศน์ แหล่งท่องเที่ยวเชิงศิลปะวิทยาการ และแหล่งท่องเที่ยวเชิงประวัติศาสตร์ เป็นต้น [4] นักท่องเที่ยวแต่ละคนมีความสนใจการท่องเที่ยวแต่ละประเภทแตกต่างกันบางคนสนใจการ

ท่องเที่ยวเชิงนิเวศหรือเชิงวัฒนธรรม บางคนสนใจการท่องเที่ยวประเภทชายหาดหรือจุดชมทิวทัศน์ หรือบางคนอาจสนใจการท่องเที่ยวเชิงศิลปะวิทยาการเพียงอย่างเดียว ในขณะที่แหล่งท่องเที่ยวหนึ่งๆ อาจถูกจัดอยู่ในประเภทการท่องเที่ยวได้หลายประเภท เช่น ถนนถลางในจังหวัดภูเก็ตที่มีสถาปัตยกรรมชิโนโปตุเกสอันสวยงามและมีร้านค้าเก่าแก่ของจังหวัดตั้งอยู่หลายร้าน สามารถถูกจัดอยู่ในแหล่งท่องเที่ยวเชิงศิลปะวิทยาการหรือแหล่งท่องเที่ยวเชิงวัฒนธรรมหรือทั้งสองอย่างก็ได้ ซึ่งจะเป็นปัญหาสำหรับนักท่องเที่ยวในการเลือกแหล่งท่องเที่ยวต่างๆ ให้ตรงตามความสนใจของตนให้มากที่สุด นอกจากนั้นจังหวัดภูเก็ตมีแหล่งท่องเที่ยวหลายแห่งและมีเส้นทางที่ซับซ้อน ทำให้นักท่องเที่ยวประสบปัญหาเรื่องการวางแผนการเดินทาง เนื่องจากนักท่องเที่ยวมีเวลาในการท่องเที่ยวจำกัด แต่ในขณะเดียวกันก็ต้องการท่องเที่ยวหลายๆ ที่ จากปัญหาดังกล่าวผู้วิจัยจึงมีแนวคิดพัฒนาแบบจำลองการค้นหาเส้นทางการท่องเที่ยวในจังหวัดภูเก็ตโดยใช้วิธีพาทิเคิลสวอมออปติไมเซชันและฟัซซีเซต ในการหาเส้นทางเพื่อให้ได้แผนการท่องเที่ยวที่มีแหล่งท่องเที่ยวที่ตรงตามความต้องการของนักท่องเที่ยวและได้จำนวนแหล่งท่องเที่ยวมากที่สุดภายใต้เงื่อนไขด้านเวลาที่มี

2. วรรณกรรมที่เกี่ยวข้อง

ในการพัฒนาแบบจำลองการค้นหาเส้นทางการท่องเที่ยวในจังหวัดภูเก็ตที่ให้ความพึงพอใจสูงสุดใช้ทฤษฎีหลักสองเรื่อง คือ วิธีพาทิเคิลสวอมออปติไมเซชันและฟัซซีเซต โดยมีรายละเอียดดังนี้

2.1 ฟัซซีเซต

ฟัซซีเซต (Fuzzy set) คือ เซตที่สมาชิกแต่ละตัวในเซตมีค่าความเป็นสมาชิก (Membership value) เพื่อบ่งบอกถึงระดับความเป็นสมาชิกในเซตนั้น ค่าความเป็นสมาชิกจะมีค่าอยู่ช่วง 0 ถึง 1 ซึ่งแตกต่างจากเซตแบบทวินัย (Crisp set) ที่สมาชิกแต่ละตัวจะถูกพิจารณาว่าเป็นหรือไม่เป็นสมาชิกในเซตนั้น หรือกล่าวคือค่าความเป็นสมาชิกของเซตแบบทวินัยจะเป็น 0 หรือ 1 เท่านั้น (Justify) [5]

ฟัซซีเซต A สามารถเขียนอยู่ในรูป $A = \{(x,\mu_A(x)) \mid x \in X\}$ โดยที่ $\mu_A(x)$ คือฟังก์ชันความเป็นสมาชิก (Membership function) ที่ใช้ในการคำนวณค่าความเป็นสมาชิกของ x ในเซต A ฟังก์ชันความเป็นสมาชิกนั้นมีได้หลายแบบ เช่น ฟังก์ชันรูประฆังคว่ำ ฟังก์ชันรูปสี่เหลี่ยมคางหมู และฟังก์ชันรูปสามเหลี่ยม ส่วน X คือ เอกภพสัมพัทธ์ (Universe) ของฟัซซีเซตซึ่งเป็นได้ทั้งแบบต่อเนื่อง (Continuous) และไม่ต่อเนื่อง (Discrete) ตัวอย่างของเอกภพสัมพันธ์ที่ไม่ต่อเนื่องและไม่มีลำดับ (Non-ordered) เช่น เซต B คือ เซตของสถานที่ท่องเที่ยวในจังหวัดภูเก็ตที่น่าสนใจ มีเอกภพสัมพัทธ์คือ X = {หาดป่าตอง วัดฉลอง ถนนถลาง แหลมพรหมเทพ} สามารถเขียนเซต B ได้เป็น B = {(หาดป่าตอง,0.9) (วัดฉลอง,0.5) (ถนนถลาง,0.7) (แหลมพรหมเทพ,0.8)} ตัวเลขที่อยู่หลังสถานที่ท่องเที่ยว คือ ค่าความเป็นสมาชิกในเซต B ของจังหวัดภูเก็ต ในกรณีที่เอกภพสัมพัทธ์เป็นแบบต่อเนื่องค่าความเป็นสมาชิกจะถูกกำหนดจากฟังก์ชันความเป็นสมาชิก แต่ในกรณีที่เอกภพสัมพัทธ์เป็นแบบไม่ต่อเนื่องค่าความเป็นสมาชิกจะถูกกำหนดขึ้นมาเองตามความคิดเห็นของแต่ละบุคคล [6] ซึ่งฟัซซีเซตถูกนำไปประยุกต์ใช้ในงานต่างๆ [7][8]

2.2 วิธีพาทิเคิลสวอมออปติไมเซชัน

Particle Swarm Optimization (PSO) หรือการหาค่าเหมาะสมที่สุดแบบกลุ่มอนุภาคเป็นเทคนิคหนึ่งของการคำนวณเชิงวิวัฒนาการ (Evolutionary Computation) [9][10] ที่ถูกพัฒนาโดย Kennedy และ Eberhart [9] เพื่อค้นหาผลลัพธ์ของปัญหาที่มีความซับซ้อน โดยการคำนวณเชิงวิวัฒนาการมีแนวคิดมาจากทฤษฎีวิวัฒนาการของสิ่งมีชีวิต ซึ่ง Charles Darwin เป็นผู้คิดค้น [11][12] ซึ่ง PSO สามารถมาใช้แก้ปัญหาที่ซับซ้อน (NP-Hard) ได้อย่างมีประสิทธิภาพโดยเฉพาะอย่างยิ่งเมื่อสมการของปัญหาแบบไม่เป็นเชิงเส้น (Non Linear) [10]

ขั้นตอนวิธีของ PSO เลียนแบบมาจากพฤติกรรมทางสังคมของสิ่งมีชีวิต เช่น การเคลื่อนที่ของฝูงนกที่ออกหาอาหารพร้อมๆ กันเป็นฝูง นกแต่ละตัวทำหน้าที่บินค้นหาอาหารและอาจย้ายที่อยู่อาศัยไปพร้อมๆ กัน จุดใดที่มีอาหารอยู่หรือสามารถพักอาศัยได้ หรือแม้แต่จุดที่มีอันตราย นกจะสื่อสารกันในกลุ่มเพื่อแลกเปลี่ยนข้อมูลระหว่างกัน [9][11] นกแต่ละตัวซึ่งเปรียบเป็นอนุภาคแต่ละตัวจะต้องเก็บข้อมูลเวคเตอร์สองเวคเตอร์ คือ เวคเตอร์ระบุตำแหน่งและเวคเตอร์ระบุความเร็วในการบิน ดังนั้นถ้าในหนึ่งอนุภาคแทนด้วยพารามิเตอร์ของ

ปัญหาหนึ่งจะประกอบด้วยตัวแปรซึ่ง คือ เซตข้อมูลของอนุภาคแต่ละตัว ดังสมการ (1) [11][12]

$$PK = \{X, V, Fitness\ Value, gbest, pbest\} \quad (1)$$

เมื่อ

P คือ อนุภาคประกอบด้วย P = {P1, P2, …, PK}

K คือ จำนวนอนุภาค

X คือ ตำแหน่งที่อนุภาคที่อยู่ในปัญหา เขียนในรูปของ Xd = {x1, x2, …,xd} เมื่อ d คือจำนวนมิติของปัญหา

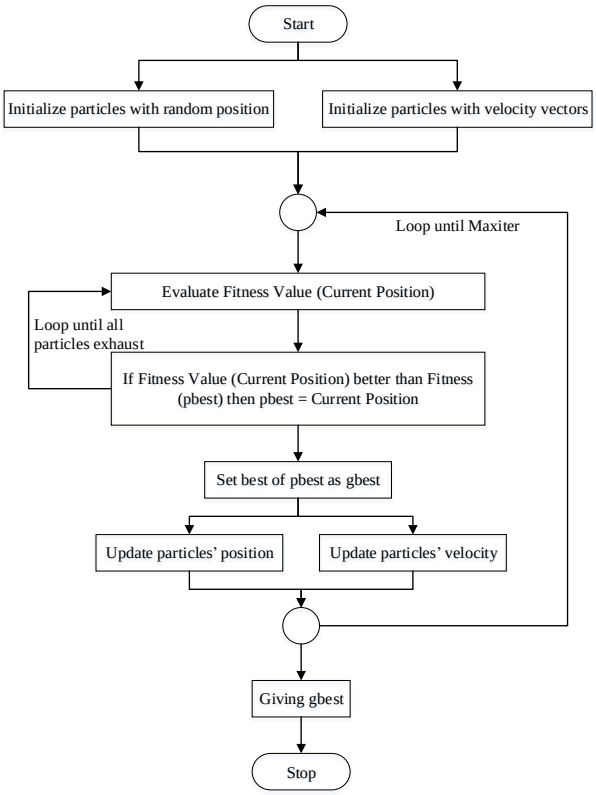
V คือ ค่าความเร็วในการเคลื่อนที่ของอนุภาคประกอบด้วยเวกเตอร์ตามขนาดของมิติของปัญหา Vd = {v1,v2, …,vd}

Fitness Value คือค่าความเหมาะสมของอนุภาคนั้นซึ่งสามารถคำนวณได้จากฟังก์ชันเป้าหมาย ดังนี้

Fitness value = Objective Function (X)

gbest คือ ค่าตำแหน่งที่อนุภาคนั้นได้ค่าความเหมาะสมสูงสุด โดย gbestd = {gbest1, gbest2, …,gbestd}

pbest คือ ค่าความเหมาะสมที่ดีที่สุดที่กลุ่มอนุภาครุ่นปัจจุบันได้ โดย pbestd = {pbest1, pbest2, …, pbestd}



ภาพที่ 1 แผนผังวิธีพาทิเคิลสวอมออปติไมเซชัน [13]

3. แบบจำลองการค้นหาเส้นทางการท่องเที่ยว

การพัฒนาแบบจำลองการค้นหาเส้นทางการท่องเที่ยวนี้ตั้งอยู่บนสมมติฐานที่ว่า นักท่องเที่ยวจะเริ่มออกเดินทางจากที่โรงแรมที่พักและสิ้นสุดการเดินทางที่โรงแรมเดิมอีกครั้งหนึ่ง นักท่องเที่ยวแต่ละคนจะมีความสนใจในประเภทการท่องเที่ยวไม่เหมือนกัน จำนวนประเภทการท่องเที่ยวที่สนใจและเวลาในการท่องเที่ยวไม่เท่ากัน ในการพัฒนาแบบจำลองนั้นจะแบ่งออกเป็นสองส่วน คือ การเตรียมข้อมูลและการค้นหาเส้นทางการท่องเที่ยว แต่ละส่วนมีรายละเอียดดังนี้

3.1 การเตรียมข้อมูล

การเตรียมข้อมูลผู้วิจัยได้ใช้ Google Maps API Web Services ในการหาระยะทางและเวลาที่ใช้ในการเดินทางที่สั้นที่สุดจากแหล่งท่องเที่ยวจุดหนึ่งไปยังอีกจุดหนึ่งเพื่อนำมาคำนวณระยะทางและเวลาโดยรวมของการท่องเที่ยว แสดงตัวอย่างข้อมูลระยะทาง (กิโลเมตร) และเวลาที่ใช้ (นาที) ได้ดังตารางที่ 1 และตารางที่ 2 ตามลำดับ

ตารางที่ 1 ตัวอย่างข้อมูลระยะทาง (กิโลเมตร)

โหนด	01	02	03	04	05
01	0.00	11.30	13.00	22.80	22.90
02	13.20	0.00	11.70	14.20	14.30
03	10.70	11.60	0.00	14.60	14.70
04	18.40	10.90	8.90	0.00	3.20
05	22.90	15.40	13.40	3.00	0.00

ตารางที่ 2 ตัวอย่างข้อมูลระยะเวลาเดินทาง (นาที)

โหนด	01	02	03	04	05
01	0	13	17	26	27
02	15	0	18	20	21
03	14	18	0	25	26
04	22	19	20	0	9
05	26	22	24	9	0

แหล่งท่องเที่ยวหรือโหนดแต่ละโหนดจะถูกจัดกลุ่มว่าอยู่ในเซตการท่องเที่ยวประเภทใด โดยที่โหนดหนึ่งๆ สามารถถูก

จัดเข้าประเภทการท่องเที่ยวได้หลายประเภทโดยใช้ฟัซซีเซต นั่นคือ ค่าความเป็นสมาชิกของโหนดในแต่ละเซตประเภทการท่องเที่ยวจะมีค่าอยู่ในช่วง 0 ถึง 1 และเอกภพสัมพัทธ์ของฟัซซีเซตนี้ คือ แหล่งท่องเที่ยวที่อยู่ในจังหวัดภูเก็ตจำนวน 80 โหนด ซึ่งเป็นเอกภพสัมพัทธ์แบบไม่ต่อเนื่องและไม่มีลำดับ ค่าความเป็นสมาชิกนี้ได้มาจากการสำรวจความคิดเห็นจากบุคลากรด้านการท่องเที่ยวและคนภูเก็ตท้องถิ่นจำนวน 30 คน โดยให้ผู้ตอบแบบสอบถามกำหนดค่าความเป็นสมาชิกของแต่ละแหล่งท่องเที่ยวว่าอยู่ในการท่องเที่ยวประเภทใด ซึ่งแบบสอบบถามได้ทำการวัดคุณภาพความเชื่อมั่น (reliability) ของแบบสอบถามว่าแต่ละข้อคำถามสอดคล้องกันหรือไม่ โดยการหาค่า Cronbach's alpha (ครอนแบคอัลฟา) หรือค่า "สัมประสิทธิ์ของความเชื่อมั่น" (coefficient of reliability) ซึ่งมีค่าอยู่ที่ 0.7 โดยที่ผลรวมของค่าความเป็นสมาชิกของแต่ละประเภทการท่องเที่ยวในแต่ละแหล่งท่องเที่ยวต้องเท่ากับ 1 จากนั้นนำค่าความเป็นสมาชิกที่ได้มาหาค่าเฉลี่ยเพื่อใช้ในการทดลองต่อไป

ข้อมูลระดับค่าความเป็นสมาชิก เวลาที่ใช้เที่ยวและเวลาเปิด-ปิดทำการของสถานที่นั้นๆ แสดงตัวอย่างข้อมูลดังตารางที่ 3 โดยที่เวลาเปิด-ปิดทำการนี้จะหมายถึงช่วงเวลาที่เหมาะสมในการเที่ยวสถานที่นั้นๆ

ตารางที่ 3 ตัวอย่างข้อมูลแหล่งท่องเที่ยวแบบฟัซซีเซต

โหนด / คุณลักษณะ	01	02	03	04	05
เชิงนิเวศ	0.0	0.0	0.0	0.6	0.0
เชิงวัฒนธรรม	0.7	0.0	0.0	0.2	0.0
ชายหาด	0.0	0.5	0.0	0.0	1.0
จุดชมทิวทัศน์	00	0.5	0.0	0.2	0.0
เชิงศิลปะ	0.0	0.0	1.0	0.0	0.0
เชิงประวัติศาสตร์	0.3	0.0	0.0	0.0	0.0
จำนวนเวลาที่ใช้ (นาที)	40	30	20	45	60
เวลาเปิด (นาฬิกา)	8.00	10.00	6.00	9.00	6.00
เวลาปิด (นาฬิกา)	18.00	20.00	18.00	17.00	19.00

3.2 การค้นหาเส้นทางการท่องเที่ยว

ขั้นตอนการค้นหาเส้นทางการท่องเที่ยวโดยใช้วิธีพาทิเคิลสวอมออปติไมเซชันและฟัซซีเซต สามารถแสดงได้ดังภาพที่ 2 มีรายละเอียดดังนี้

1) การสร้างประชากรเริ่มต้น (Initial particle) โดยใช้การสุ่มค่าให้ X คือ สถานที่ท่องเที่ยว และค่าความเร็วการเคลื่อนที่ให้ V ของอนุภาคแต่ละตัว ดังสมการ (2) และ (3) สุ่มค่าพารามิเตอร์ให้ตัวแปรของ X และ V ตามจำนวนของมิติของปัญหา PSO จะมีการเก็บค่า V ซึ่งเป็นพารามิเตอร์ในการคำนวณความเร็วของการเคลื่อนที่ โดยกำหนดค่า Popular Size (Swarm Size) เท่ากับ 10

X_D = function random value ( ) (2)

V_D-= function random value ( ) (3)

2) การคำนวณค่าความเหมาะสมของอนุภาค (Evaluate fitness value) ฟังก์ชันความเหมาะสมดังสมการ (4) คำนวณโดยส่งค่า X ให้แก่ฟังก์ชันเป้าหมาย FV โดย

Objective function = max(FV) (4)

$$FV = \left(\sum_{i=1}^{n} \sum_{i=1}^{m} \mu_j node_i * h_i \right) \quad (5)$$

เมื่อ n คือ จำนวนโหนด m คือ จำนวนประเภทการท่องเที่ยวที่สนใจ และ μ_j คือค่าความเป็นสมาชิกในเซตประเภทการท่องเที่ยวที่ j ของโหนดลำดับที่ i และ h_i คือค่าความพอใจในประเภทการท่องเที่ยว

3) การเก็บค่าความเหมาะสมที่ดีสุด (Update Pbest) ซึ่งเป็นค่าความเหมาะสมที่ดีสุดในรอบการทำงานปัจจุบัน มีเงื่อนไขการปรับปรุงตามสมการ (6) ถ้าค่าความเหมาะสมของอนุภาคปัจจุบันมีค่าดีกว่า Pbest

IF fitness value > Pbest THEN (6)

Pbest = fitness value ENDIF

4) การเก็บค่าตำแหน่งให้ผลลัพธ์ดีที่สุด (Update Gbest) เป็นตำแหน่งที่ได้ผลลัพธ์ที่ดีสุดพิจารณาจากการวนรอบตั้งแต่รอบแรกถึงปัจจุบัน ซึ่งแตกต่างจาก Pbest ที่พิจารณาการเก็บค่าที่ดีสุดเฉพาะรอบปัจจุบัน Gbest เป็นการเก็บค่าความเหมาะสมที่ดีสุดของการคำนวณทั้งหมด เงื่อนไขการปรับปรุงดังสมการ (7)

IF fitness value > Gbest THEN (7)

$Gbest_D = X_D$ ENDIF

5) การคำนวณความเร็วการเคลื่อนที่ของแต่ละอนุภาค (Update velocity) ดังสมการ (8) โดยกำหนดค่า Inertia Weight Damping Ratio ในงานวิจัยมีค่าเท่ากับ 0.99

$$V_D' = weight * V_D + n_1 * random () * (Pbest - X_D) + n_2 * random () * (Gbest - X_D) \quad (8)$$

6) การปรับปรุงตำแหน่งของแต่ละอนุภาค (Update position) โดยใช้ผลลัพธ์จากการคำนวณในสมการ (8) ทำให้ได้ตำแหน่งใหม่ที่อนุภาคจะเคลื่อนที่ไป จากนั้นสมการ (9) นำผลลัพธ์มาปรับปรุง ซึ่งการปรับปรุงด้วยการบวกค่าเพิ่มจำเป็นต้องตรวจสอบขอบเขตของปัญหาเพื่อป้องกันไม่ให้ตำแหน่งเกินค่าขอบเขตของปัญหา โดยกำหนดค่าสัมประสิทธิการเชื่ออนุภาคตนเอง (Personal Learning Coefficient) เท่ากับ 1.25 และค่าสัมประสิทธิการเชื่อกลุ่ม (Global Learning Coefficient) เท่ากับ 2

$$X_D = X_D + V_D' \quad (9)$$

เงื่อนไขข้อที่หนึ่ง คือ แต่ละโหนดในการซุ่มตำแหน่งและความเร็วของอนุภาคของจะต้องไม่ซ้ำกัน เงื่อนไขข้อที่สอง คือ เวลาที่ใช้ในการท่องเที่ยวนั้นต้องไม่เกินเวลาที่นักท่องเที่ยวมีดังสมการที่ (10) และเงื่อนไขข้อที่สาม คือ แต่ละโหนดของการซุ่มตำแหน่งจะต้องถูกตรวจสอบว่าเวลาที่ไปถึงแต่ละโหนดนั้นเป็นช่วงเวลาที่เหมาะสมสำหรับการท่องเที่ยวที่โหนดนั้นหรือไม่ หากไม่ผ่านจะซุ่มจุดใหม่จนกว่าจะผ่าน

$$\sum_{i=1}^{n} t_i + \sum_{i=1, i \neq j}^{n} t_{ij} \leq Available\ Time \quad (10)$$

7) การตรวจสอบการสิ้นสุดการวนรอบ (Termination Check) เงื่อนไขที่ตรวจสอบให้หยุดการทำงานโดยงานวิจัยนี้กำหนดการหยุดโดยใช้จำนวนรอบการที่ 500 รอบ (Iterations)

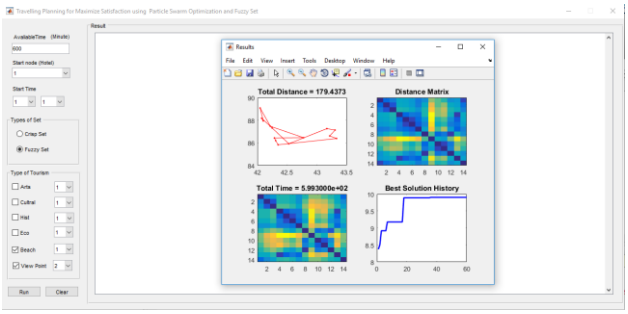
3.4 ผลการดำเนินงาน

ผู้วิจัยได้ทดลองเปรียบเทียบระหว่างการใช้วิธีพาทิเคิลสวอมออปติไมเซชันร่วมกับฟัซซีเซตและการใช้วิธีพาทิเคิลสวอมออปติไมเซชันร่วมกับเซตแบบทวินัยหาเส้นทางการท่องเที่ยวตามประเภทที่สนใจ การทดลองที่ใช้เซตแบบทวินัยนั้นจะต้องมีการปรับค่าความเป็นสมาชิกของฟัซซีเซตให้เป็นเซตแบบทวินัย โดยกำหนดค่าความเป็นสมาชิกของแต่ละโหนดให้เท่ากับ 1 หากค่าความเป็นสมาชิกมีค่ามากกว่า 0 แสดงตัวอย่างดังตารางที่ 4

ตารางที่ 4 ตัวอย่างข้อมูลแหล่งท่องเที่ยวแบบเซตทวินัย

โหนด / คุณลักษณะ	01	02	03	04	05
เชิงนิเวศ	0.0	0.0	0.0	1.0	0.0
เชิงวัฒนธรรม	1.0	0.0	0.0	1.0	0.0
ชายหาด	0.0	1.0	0.0	0.0	1.0
จุดชมทิวทัศน์	00	1.0	0.0	1.0	0.0
เชิงศิลปะ	0.0	0.0	1.0	0.0	0.0
เชิงประวัติศาสตร์	1.0	0.0	0.0	0.0	0.0
จำนวนเวลาที่ใช้ (นาที)	40	30	20	45	60
เวลาเปิด (นาฬิกา)	8.00	10.00	6.00	9.00	6.00
เวลาปิด (นาฬิกา)	18.00	20.00	18.00	17.00	19.00

จากการทดลองให้ผู้ใช้จำนวน 50 คน ซึ่งเป็นนักท่องเที่ยวและผู้เชี่ยวชาญทดลองใช้โปรแกรมต้นแบบและวัดความพึงพอใจว่าผลลัพธ์แหล่งท่องเที่ยวที่ได้ตรงกับความสนใจมากน้อยแค่ไหน โดยใช้แบบสอบถามเป็นเครื่องมือและใช้มาตรวัดลิเคิร์ท (Likert Scale) ให้คะแนน 1 ถึง 5 ซึ่งหมายถึงตรงกับความสนใจน้อยที่สุด น้อย ปานกลาง มาก และมากที่สุดตามลำดับ พร้อมทั้งได้ทำการวัดคุณภาพความเชื่อมั่น (reliability) ของแบบสอบถามว่าแต่ละข้อคำถามสอดคล้องกันหรือไม่ โดยการหาค่า Cronbach's alpha (ครอนแบคอัลฟา) หรือค่า "สัมประสิทธิ์ของความเชื่อมั่น" (coefficient of reliability) ซึ่งมีค่าอยู่ที่ 0.7 ซึ่งผลการทดลองพบว่าการใช้ฟัซซีเซตกำหนดค่าระดับความเป็นสมาชิกนั้นได้คะแนนเฉลี่ยเท่ากับ 4.02 และมีส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.39 ซึ่งมากกว่าการใช้เซตแบบทวินัยที่ได้คะแนนเฉลี่ยเท่ากับ 3.05 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.72 แสดงดังตารางที่ 5



ภาพที่ 2 แสดงส่วนติดต่อกับผู้ใช้ของแบบจำลอง PSO and Fuzzy Set

ทั้งนี้จากการสอบถามผู้ใช้พบว่าผู้ใช้แต่ละคนพบว่ามีความชื่นชอบและพอใจต่อผลคำตอบของแบบจำลองที่สามารถให้คำตอบเป็นสถานที่ท่องเที่ยวที่ตรงต่อความต้องการของนักท่องเที่ยว พร้อมทั้งจัดลำดับสถานที่ท่องเที่ยวว่าจะต้องเดินทางไปสถานที่ท่องเที่ยวใด ก่อน-หลัง ตามลำดับ พร้อมทั้งประมาณเวลาในการเดินทาง ประมาณเวลาในการท่องเที่ยวในแต่ละสถานที่ และเนื่องจากงานวิจัยดังกล่าวนี้ ตั้งอยู่บนการแก้ปัญหาการเดินทางของพนักงานขาย (Traveling Salesman Problem) จึงทำให้นักท่องเที่ยวสามารถกลับมายังโรงแรมได้อย่างถูกต้อง

3.5 สรุป

งานวิจัยนี้นำเสนอแบบจำลองการค้นหาเส้นทางการท่องเที่ยวโดยใช้ฟัซซีเซตกำหนดค่าความเป็นสมาชิกในแต่ละประเภทการท่องเที่ยวของสถานที่ท่องเที่ยวใดๆ และใช้วิธีพาทิเคิลสวอมออปติไมเซชัน ในการค้นหาเส้นทางการท่องเที่ยวเพื่อเสนอแผนการเดินทางท่องเที่ยวที่ตรงกับความสนใจของนักท่องเที่ยวให้มากที่สุดภายใต้เงื่อนไขเรื่องเวลาที่จำกัด จากผลการทดลองเปรียบเทียบระหว่างการใช้ฟัซซีเซตและเซตแบบทวินัยพบว่าการใช้ฟัซซีเซตให้ผลลัพธ์เป็นที่น่าพอใจกว่าการใช้เซตแบบทวินัย เนื่องจากฟัซซีเซตสามารถสะท้อนคุณลักษณะของแหล่งท่องเที่ยวแต่ละแหล่งได้ดีกว่าเซตแบบทวินัย

เอกสารอ้างอิง

[1] P. J. Holland. "*Adaptation in Natural and Artificial Systems"*, Michigan: University of Michigan Press, 1992.

[2] Evolutionary Computation, Internet: https://en.wikipedia.org/wiki/Evolutionary_computation [December 18, 2016].

[3] ธนา สาตรา และคณะ, "การประยุกต์ใช้วิธีอาณานิคมมดกับปัญหาการจัดเส้นทางเดินรถโดยมีข้อจำกัดด้านกรอบเวลาและพิจารณาระดับการบริการ", *การประชุมวิชาการข่ายงานวิศวกรรมอุตสาหการ*, หน้า 16-22, 2555.

[4] กลุ่มงานข้อมูลสารสนเทศและการสื่อสารสำนักงานจังหวัดภูเก็ต, "บรรยายสรุปจังหวัดภูเก็ต", จังหวัดภูเก็ต, 1 กันยายน 2559.[Online].Available: http://www.phuket.go.th/webpk/file_data/intropk/dataPK59.pdf. [Accessed 6 มกราคม 2560].

[5] William Siler and James J. Buckley, "*Fuzzy Expert Systems and Fuzzy Reasoning"*, John Wiley & Sons:New Jersey, 2005.

[6] Jyh-Shing Roger Jang; Chuen-Tsai Sun and Eiji Mizutani, "*Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence"*, Prentice Hall, 1997.

[7] ชนัญฎาไตรวาร, "ระบบค้นหาโฮมสเตย์โดยใช้ฟัซซีลอจิก," *The 5th National Conference on Computing and Information Technology (NCCIT 2009)*, กรุงเทพฯ, หน้า 819-824, 2552.

[8] ศันสนีย์ เอื้อพันธ์วิริยะกุล, "การใช้ฟัซซีลอจิกช่วยในการสร้างระบบการค้นหาโรงแรม", ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเชียงใหม่, 2541.

[9] J. Kennedy and R. Eberhart. *"Particle swarm optimization."* in *Proc. of IEEE International Conference on Neural Networks*, Vol.4, 1995, pp. 1942 – 1948.

[10] สุภกิจ นุตยะสกุล. "*การคำนวณเชิงวิวัฒนาการระหว่างเจเนติกอัลกอริทึม กับพาติคอลสวอมออปติมัยเซชัน.*" *Journal of Information Science and Technology*. [Online]. (2)2, pp. 1322,2011,Available:http://istjournal.mut.ac.th/Journal/JIST_V2N2/Page13-22.pdf [January 8, 2017].

[11] R.C. Eberhart and Yuhui Shi. *"Particle Swarm Optimization: Development, Applications and Resources."* IEEEXplore. [Online] 0-7803-6657-3, 2001,Available:http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=&userType=inst [January 8, 2017].

[12] อนุชา ศรีศิริวัฒน์ และคณะ," การประยุกต์ใช้เทคนิคการหาค่าเหมาะสมที่สุดแบบฝูงอนุภาคสำหรับแก้ปัฐหาทางวิศวกรรม", *Pathumwan Academic Journal*, หน้า 53-60, 2555.

[13] Maria-Iuliana Dascalu, "*Application of particle swarm optimization to formative e-assessment in project management*.", Information Economica, Vol. 15, 2011, pp. 48 – 61.

การทำนายข้อมูลจาก Two-Dimensional Gel Electrophoresis (2D-Gel) ด้วยเทคนิคดาต้าไมน์นิ่งและการคัดเลือกคุณลักษณะ

Prediction on Two-Dimensional Gel Electrophoresis (2D-Gel) using Data Mining and Feature Selection Techniques

เทียนชัย อุ่นน้อย (Theanchai)[1] สำรี มั่นเขตต์กรน์ (Samlee)[2] และกฤษณะ ไวยมัย (Kitsana)[3]

[1]สาขาเทคโนโลยีสารสนเทศ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเกษตรศาสตร์

[2]สาขาวิชารังสีเทคนิค คณะเทคนิคการแพทย์ มหาวิทยาลัยเชียงใหม่

[3]สาขาวิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเกษตรศาสตร์

[1]Theanchai.o@ku.th,[2]samlee@chiangmai.ac.th,[3]fengknw@ku.ac.th

บทคัดย่อ

ข้อมูลการแยกบนเจลด้วยกระแสไฟฟ้าสองมิติ (2D-Gel) เป็นข้อมูลทางโปรตีโอมิกส์ที่ศึกษาการแสดงออกของโปรตีนที่ได้รับความนิยมและใช้อย่างกว้างขวางด้วยคุณสมบัติที่ให้ความทนทานของข้อมูลและให้รายละเอียดของข้อมูลสูง สามารถแสดงความเป็นเอกลักษณ์เฉพาะในแต่ละชนิดของตัวอย่างได้อย่างชัดเจน แต่ก็ส่งผลให้การการตีความและวิเคราะห์ผลข้อมูลการแยกบนเจลด้วยกระแสไฟฟ้าสองมิติ จำเป็นต้องใช้ระยะเวลานานในการวิเคราะห์และที่สำคัญต้องเป็นผู้เชี่ยวชาญที่ผ่านการฝึกอมรบเท่านั้นในการวิเคราะห์ข้อมูลชนิดนี้ เนื่องจากเป็นข้อมูลที่มีจำนวนมิติ ของข้อมูลสูง งานวิจัยนี้มีวัตถุประสงค์ประยุกต์ใช้เทคนิคทางดาต้าไมน์นิ่งและการคัดเลือกคุณลักษณะในการทำนายข้อมูลชนิด การแยกบนเจลด้วยกระแสไฟฟ้าสองมิติ โดยนำเสนอการเตรียมข้อมูลสำหรับข้อมูล การแยกบนเจลด้วยกระแสไฟฟ้าสองมิติ และการทดสอบการคัดเลือกคุณลักษณะรวมไปถึงการสร้างโมเดลในการทำนายเพื่อค้นหาระบบที่ดีที่สุดในการทำนายข้อมูลการแยกบนเจลด้วยกระแสไฟฟ้าสองมิติ จากการทดสอบกับข้อมูลจริงในงานวิจัยนี้พบว่าการทำงานของมอดเดอเรท ที-เทส (Moderated T –Test) ร่วมกับการใช้เทคนิคการแยกประเภทขั้นตอนวิธีการเพื่อนบ้านใกล้ที่สุด(K-NN) มีความถูกต้องในการทำนายมากกว่า 80 เปอร์เซ็นต์ ในทุกสายพันธุ์ใบยาสูบ และเมื่อทำการแก้ปัญหาของความไม่เท่ากันของตัวอย่างในแต่ละคลาสโดยการจำลองชุดข้อมูลใหม่ ที่ใช้พื้นฐานจากข้อมูลเดิมพบว่าผลการจำแนกประเภทในทุกเทคนิคที่เปรียบเทียบมีค่าความถูกต้องในการทำนายมากกว่า 80 เปอร์เซ็นต์ มีเพียงการคัดเลือกคุณลักษณะวิธีการวิเคราะห์องค์ประกอบหลัก (PCA) เท่านั้นที่ให้ค่าความถูกต้องต่ำกว่า 80 เปอร์เซ็นต์ จากผลการทดลองในงานวิจัยนี้น่าจะเป็นประโยชน์ในการเลือกวิธีที่เหมาะสมในการจำแนกประเภทข้อมูล การแยกบนเจลด้วยกระแสไฟฟ้าสองมิติ รวมถึงการเตรียมข้อมูลสามารถนำไปใช้ในการจำแนกประเภทข้อมูล การแยกบนเจลด้วยกระแสไฟฟ้าสองมิติ สาขาอื่นๆ

คำสำคัญ: การแยกบนเจลด้วยกระแสไฟฟ้าสองมิติ ดาต้าไมน์นิ่ง การคัดเลือกคุณลักษณะ

Abstract

Two-Dimensional Gel (2D-Gel) Electrophoresis is one of very powerful and widely used Proteomics tools. The main advantages of this method are its robustness and its unique ability to analyze complete protein at high resolution. However, 2D-Gel contains very high number of dimensions which is difficult to interpret and analyze by the end-users. The objective of this paper is to develop a method for predicting on Two-Dimensional Gel (2D-Gel) Electrophoresis. Four feature selections techniques

which are Principal Component Analysis (PCA), Moderated T –Test, Sparse Partial Least Squares Discriminant Analysis (sPLS-DA), and Recursive Feature Elimination (RFE) are combined, compared and discussed in order to improve accuracy of the prediction model. Performance of three classifiers: Support Vector Machines (SVM), Artificial Neural Network (ANN) and K-Nearest Neighbors (K-NN) is compared and discussed. The proposed method is applied to real-world datasets. The best combination is Moderated T –Test and K-Nearest Neighbors (K-NN) for respectively feature selection and model classification in the case of imbalanced datasets. The experimental results confirm that the proposed method can be applied to various types of 2D-Gel datasets.

Keyword: Two-Dimensional Gel Electrophoresis, 2D-Gel, Data Mining, Feature Selection.

1. บทนำ

โปรตีโอมิกส์(Proteomics)เป็นการศึกษาด้านโปรตีน เช่น การแสดงออกของโปรตีน การวิเคราะห์เชิงปริมาน รวมไปถึงวิเคราะห์โปรตีนที่มีความซับซ้อนจากเซลล์ เนื้อเยื่อ และตัวอย่างทางชีวภาพ ซึ่งโปรตีนถือเป็นส่วนสำคัญของสิ่งมีชีวิตแต่ละชนิด มีหน้าที่การทำงานที่หลากหลายตามชนิดของตัวอย่าง หนึ่งในเทคนิคที่สำคัญทางโปรตีโอมิกส์และใช้กันอย่างกว้างขวางคือ Two-Dimensional Gel Electrophoresis (2D-Gel) ที่เป็นเทคนิคการแยกโปรตีนที่ประกอบไปด้วยการแยกสองขั้นตอน ในขั้นตอนแรกเรียกว่า Isoelectric focusing (IEF) จะทำการแยกโปรตีนโดยใช้ความแตกต่างของค่า Isoelectric point(pl) ขั้นตอนที่สองเรียกว่า Sodium Dodecyl Sulfate Electrophoresis (SDS) จะแยกโปรตีนโดยอาศัยความแตกต่างของขนาดโปรตีน โดยโปรตีนที่มีมวลโมเลกุลมากจะเคลื่อนที่ได้ในระยะทางน้อยกว่าโปรตีนที่มีมวลโมเลกุลน้อย เป็นกระบวนการวิเคราะห์โปรตีนและการถ่ายภาพดิจิตอลในห้องปฏิบัติการ ข้อได้เปรียบหลักของวิธีนี้คือความทนทานของข้อมูลที่ให้รายละเอียดของข้อมูลสูง สามารถแสดงความเป็นเอกลักษณ์เฉพาะในแต่ละชนิดของตัวอย่างได้อย่างชัดเจน เป็นผลให้สามารถจำแนกประเภทของตัวอย่างได้ [1-4] แต่อย่างไรก็ตามการตีความและวิเคราะห์ผลข้อมูล 2D-Gel จำเป็นต้องใช้ระยะเวลานานในการวิเคราะห์และที่สำคัญต้องเป็นผู้เชี่ยวชาญที่ผ่านการฝึกอมรบเท่านั้นในการวิเคราะห์ข้อมูลชนิดนี้ [5]เนื่องจากเป็นข้อมูลที่มีจำนวนมิติ(Dimension)ของข้อมูลสูง

งานวิจัยที่เกี่ยวข้องกับการตีความและการวิเคราะห์ข้อมูล 2D-Gel ได้นำเสนอวิธีที่ง่ายและเพิ่มประสิทธิภาพในการวิเคราะห์ที่สามารถนำไปจำแนกประเภทของตัวอย่างได้อย่างถูกต้อง โดยการประยุกต์ใช้เทคนิคทางด้านดาต้าไมน์นิ่งและวิธีการทางสถิติในการลดจำนวนมิติของข้อมูลและการสร้างโมเดลในการจำแนกประเภทยกตัวอย่างเช่น Cary, Pavisic et al. 2003[4] นำเสนอการจำแนกชนิดของหนูเกิดใหม่กับหนูที่มีอายุ 2 ปี ด้วยการใช้ Neural Network อัลกอริทึม Back-propagation กับข้อมูล 2D-PAGE ได้ค่าความถูกต้องที่มากกว่า 90 เปอร์เซ็นต์ Sampson, Parker et al. 2011 [6] แสดงให้เห็นว่าในการใช้ Partial Least Squares (PLS) และ Support Vector Machines (SVM) มีประสิทธิภาพที่สูงมากในการจำแนกประเภทข้อมูลโปรตีโอมิกส์ Marengo, Robotti et al. 2003 [7] ใช้ Principal Component Analysis (PCA) ร่วมกับ linear discriminant analysis: LDA ในการลดจำนวนมิติของข้อมูลและการทำนายตัวอย่างจากต่อมน้ำเหลืองปกติกับตัวอย่างที่ได้จากมะเร็งต่อมน้ำเหลืองชนิดนอนฮอดจ์กิน โดยปราศจากความผิดพลาด Lee, Lin et al. 2003 [8] ประยุกต์ใช้ Support Vector Machines (SVM) และ Artificial Neural Network (ANN) ได้ค่าความถูกต้องที่ 85 เปอร์เซ็นต์ และ 95.1 เปอร์เซ็นต์ตามลำดับ ด้วยวัตถุประสงค์การประยุกต์ใช้ดาต้าไมน์นิ่งกับข้อมูล 2D–Gel ที่มีจำนวนมิติหรือตัวแปรมาก เทคนิคการลดจำนวนมิติของข้อมูล 2D-Gel ถือเป็นปัจจัยสำคัญในการขจัดปัญหาของมิติของข้อมูล นอกจากจะช่วยลดการใช้ทรัพยากรต่างๆ แล้ว ยังเพิ่มความสามารถในการแยกประเภทข้อมูลได้แม่นยำขึ้นอีกด้วย

การคัดเลือกคุณลักษณะเป็นกระบวนการสำคัญในการทำเหมืองข้อมูลโดยเฉพาะอย่างยิ่งกับข้อมูลที่มีมิติสูงอย่าง 2D-Gel เพื่อเป็นการคัดเลือกคุณลักษณะที่ดีที่สุดประกอบด้วยจำนวนมิติที่สำคัญและส่งผลต่อความแม่นยำในการจำแนก โดยคงสภาพคุณสมบัติข้อมูลเดิมหรือสูญเสียความสำคัญของข้อมูลน้อยที่สุด มีรูปแบบและวิธีการคัดเลือกคุณลักษณะแบ่ง

ออกเป็น 3 ประเภทได้แก่ วิธีการกรอง(Filter Approach) วิธีการควบรวม(Wrapper Approach) และวิธีการฝังตัว (Embedded Approach) งานวิจัยที่เกี่ยวข้องได้นำเสนอการใช้การคัดเลือกคุณลักษณะกับข้อมูลประเภท 2D- Gel อาทิเช่น Artigaud, Gauthier et al. 2013[9] ได้นำเสนอวิธี Moderated T –Test ในการลดจำนวนมิติของข้อมูล และ Silva and Richard. 2016 [10] ได้นำเสนอ Principal Component Analysis (PCA) Moderated T –Test Sparse Partial Least Squares Discriminant Analysis (sPLS-DA) และ Recursive Feature Elimination (RFE) เพื่อหาคุณลักษณะของข้อมูลที่ดีนำไปสู่การวิเคราะห์ความแตกต่างระหว่างตัวอย่างสองชนิดด้วยการค้นหาจุดของโปรตีนหรือฟีเจอร์ ที่สามารถแสดงความแตกต่างของจุดของโปรตีนหรือฟีเจอร์ ในแต่ละชนิดของตัวอย่างจากงานวิจัยที่กล่าว มาพบว่าผลการทำนายเป็นที่สนใจในความถูกต้องของการ จำแนกประเภท และแสดงเหตุผลสำคัญที่ทำให้การวิเคราะห์ข้อมูล 2D-Gel มีประสิทธิภาพเพิ่มขึ้นโดยการลดมิติของข้อมูล อย่างไรก็ตามจากงานวิจัยที่กล่าวมานั้นยังเป็นการทำงานกับชุดข้อมูลที่มีจำนวนตัวอย่างค่อนข้างน้อย ยังเป็นการยากต่อการยอมรับและให้ความน่าเชื่อถือของผลการทดลองที่เกิดขึ้น

วัตถุประสงค์ของงานวิจัยนี้คือการประยุกต์ใช้เทคนิคทางดาต้าไมน์นิ่งและการคัดเลือกคุณลักษณะในการทำนายข้อมูลชนิด 2D-Gel ประกอบไปด้วยการคัดเลือกคุณลักษณะสี่วิธีได้แก่ Moderated T –Test, Principal Component Analysis (PCA) , Sparse Partial Least Squares Discriminant Analysis(sPLS-DA) และ Recursive Feature Elimination (RFE) สำหรับค้นหากลุ่มของคุณลักษณะที่มีศักยภาพในการจำแนกชนิดของตัวอย่าง รวมถึงการเปรียบเทียบการทำงานของการคัดเลือกคุณลักษณะแต่ละวิธี และนำข้อมูลที่ได้จากการคัดเลือกคุณลักษณะมาเป็นอินพุตในการทดสอบประสิทธิภาพของโมเดลในการจำแนกประเภททั้งหมดสามเทคนิคได้แก่ Support Vector Machines (SVM) , Artificial Neural Network (ANN) และ K-Nearest Neighbors (K-NN) เพื่อเปรียบเทียบและวิเคราะห์ผล ในการค้นหาระบบที่ดีที่สุดในการทำนายข้อมูล 2D-Gel โดยงานวิจัยนี้ได้นำเสนอวิธีในการประยุกต์ใช้กับข้อมูลจริง ซึ่งให้ผลการทดลองแสดงให้เห็นว่าการทำงานร่วมกันระหว่างการคัดเลือกคุณลักษณะด้วยวิธี Moderated T –Test กับ เทคนิค K-NN ส่งผลให้ประสิทธิภาพในการจำแนกประเภทค่าความแม่นยำมีค่ามากกว่า 80เปอร์เซ็นต์ในทุกสายพันธุ์ใบยาสูบ ในกรณีที่ข้อมูลมีความไม่สมดุลกันของตัวอย่างแต่ละคลาส และค่าความแม่นยำเพิ่มขึ้นมากกว่า 80 เปอร์เซ็นต์ เมื่อชุดข้อมูลในแต่ละคลาสมีจำนวนเท่ากันมีเพียงการคัดเลือกคุณลักษณะวิธี PCA เท่านั้นที่ให้ค่าความถูกต้องต่ำกว่า 80 เปอร์เซ็นต์

ในหัวข้อต่อไปจะกล่าวถึง ความรู้พื้นฐาน หัวข้อที่ 3 จะแสดงวิธีการวิจัย ผลการวิจัยและอภิปรายผลในหัวข้อที่ 4 และสรุปในหัวข้อที่ 5

2. ความรู้พื้นฐาน

การคัดเลือกคุณลักษณะเป็นกระบวนการสำคัญในการทำเหมืองข้อมูลโดยเฉพาะอย่างยิ่งกับข้อมูลที่มีมิติสูงอย่าง 2D-Gel เพื่อเป็นการคัดเลือกคุณลักษณะที่ดีที่สุดประกอบด้วยจำนวนมิติที่สำคัญและส่งผลต่อความแม่นยำในการจำแนก โดยคงสภาพคุณสมบัติข้อมูลเดิมหรือสูญเสียความสำคัญของข้อมูลน้อยที่สุด และในปัจจุบันเทคนิคที่ได้รับการยอมรับและใช้ในการทำวิจัยกับข้อมูล 2D-Gel ได้แก่

2.1 Moderated T –Test

Moderated T –Test ที่นำเสนอโดย Smyth (2004; modT.Prot function) เป็นสถิติในการหาความแตกต่างระหว่างสองหมวดหมู่ โดยสามารถคำนวณหาความแปรปรวนของแต่ละคลาสเพื่อนำมาหาค่าความแตกต่างที่เป็นตัวบ่งบอกถึง Significant spot หรือจุดที่มีผลเฉพาะและสามารถบ่งบอกถึงหมวดหมู่นั้นๆได้ [9]

2.2 Principal Component Analysis (PCA)

เป็นวิธีที่ได้รับความนิยมในการลดมิติของข้อมูลที่มีความซับซ้อน ซึ่งในการทำงานของ PCA จะทำงานสรุปข้อมูลตั้งต้นให้มีมิติของข้อมูลที่ลดลงโดยไม่ให้เสียสารสนเทศของข้อมูลเดิมและจะสร้างตัวแปรใหม่ซึ่งประกอบขึ้นด้วยความแปรผันหรือความแปรปรวนของตัวแปรเดิมเรียกว่า Principal Components (PCs) จึงนิยมใช้การวิเคราะห์ด้วยเทคนิคนี้เพื่อลดขนาดเมตริกซ์ของตัวแปรให้มีจำนวนน้อยลง[11]

2.3 Sparse Partial Least Squares Discriminant Analysis (sPLS-DA)

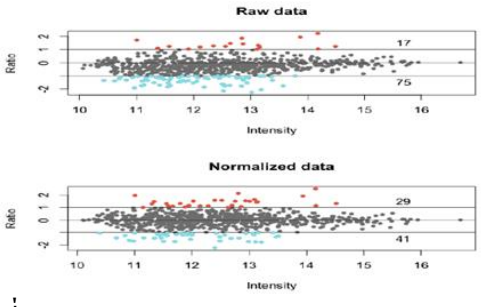
sPLS-DA เป็นเทคนิคที่เหมาะกับการจัดการกับข้อมูลที่มีจำนวนตัวแปรทำนายมากกว่าขนาดตัวอย่าง ซึ่งเป็นปัญหาหลักที่มักเผชิญกับการวิเคราะห์ข้อมูลไมโครอาร์เรย์(microarray) ซึ่ง sPLS-DA เป็นประโยชน์อย่างยิ่งสำหรับข้อมูลที่มีมิติของข้อมูลสูงในการทำ data visualization และ differential analysis ในข้อมูลชนิด gel-based proteomic[10, 12, 13]

2.4 Recursive Feature Elimination (RFE)

Recursive Feature Elimination (RFE) คือการค้นหาชุดของข้อมูลที่มีขนาดเล็กและมีขนาดของฟีเจอร์ที่น้อยที่สุดโดยพิจารณาซ้ำๆ ในขั้นตอนแรกคือการประมาณค่าน้ำหนักในแต่ละฟีเจอร์ในข้อมูลชุดฝึกสอน จากนั้นฟีเจอร์ที่มีค่าน้ำหนักน้อยที่สุดจะถูกตัดออกจากชุดของฟีเจอร์หลัก ขั้นตอนนี้จะถูกทำซ้ำจนได้ฟีเจอร์ที่เหมาะสมสำหรับไปทำนายข้อมูลที่แตกต่างกันได้ ชุดทำนายที่ดีที่สุดจะกำหนดโดยมาตรการวัดประสิทธิภาพยกตัวอย่างเช่น R^2, classification, accuracy, etc [14]

2.5 นอร์มัลไลซ์ (Normalization)

เป็นการปรับช่วงข้อมูลให้อยู่ในช่วงเดียวกันเพื่อขจัดปัญหาความแตกต่างกันของช่วงข้อมูลที่จะทำให้โมเดลทำนายผิดพลาด และในข้อมูล 2D-Gel มีกระบวนการทางห้องปฏิบัติการในการสกัดโปรตีนที่เป็นปัจจัยสำคัญของความแตกต่างของข้อมูลแต่ละตัวอย่างการปรับช่วงข้อมูลที่ได้รับการยอมรับสำหรับข้อมูล 2D-Gel คือ Quantiles Normalization[9] ซึ่งจะปรับค่าข้อมูลในแต่ละคอลัมน์หรือฟีเจอร์ให้อยู่ในระดับเดียวกันรวมถึงปรับการกระจายของข้อมูลในแต่ละคอลัมน์ให้มีการกระจายแบบเดียวกัน โดยในแต่ล่ะค่านั้นจะถูกปรับให้เข้าใกล้ 0 มากที่สุดโดยการใช้ log ratio ทำให้ข้อมูลใหม่ที่มีค่าของปริมาณโปรตีนในระดับที่เท่าๆ กันดังแสดงในภาพที่ 1

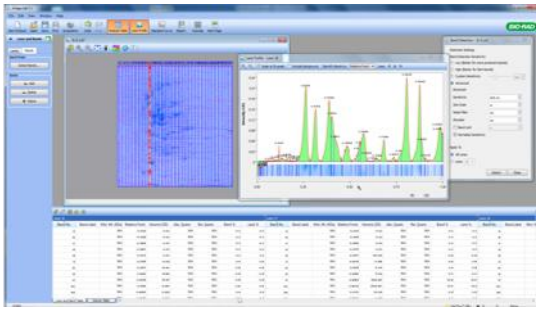


ภาพที่ 1: ตัวอย่างการปรับช่วงข้อมูลด้วย Quantiles Normalization[9]

3. วิธีการวิจัย

3.1 ข้อมูล 2D-Gel

ข้อมูล 2D-Gel ที่ผ่านกระบวนการทางห้องปฏิบัติการและอ่านผลด้วยเครื่อง GS-900 ™ Calibrated Densitometer จากทีมวิจัย รศ.ดร.สำรี มั่นเขตต์กรน์ ภาควิชารังสีเทคนิค มหาวิทยาลัยเชียงใหม่ แสดงความสัมพันธ์ระหว่างมวลโมเลกุลกับค่า Isoelectric point (pI) และค่าปริมาณความเข้มข้นของโปรตีน(Intensity) โดยจะแสดงถึงค่าปริมาณของโปรตีนที่ระดับมวลโมเลกุลต่างๆดังแสดงในภาพที่ 2



ภาพที่ 2: การแสดงออกของโปรตีนในรูปแบบกราฟที่ได้จากเครื่อง GS-900 ™ Calibrated Densitometer

ด้วยข้อมูลที่ใช้ในงานวิจัยนี้เป็นข้อมูลเดียวกับงานวิจัย[15]แต่เป็นการนำข้อมูลในมิติของมวลโมเลกุลกับค่า Isoelectric point (pI) ซึ่งต่างกันที่ใช้มิติของค่า Isoelectric point (pI) และค่า rf มาใช้ในการจำแนกประเภท รวมถึงการเพิ่มประสิทธิภาพในการจำแนกประเภทด้วยการเลือกฟีเจอร์ (Feature Selection)ที่หลากหลายสำหรับข้อมูล 2D-Gel

3.2 การเตรียมข้อมูล (Data preprocessing) 2D-Gel

เนื่องจากข้อมูลจากเครื่อง GS-900 ™ Calibrated Densitometer ให้รายละเอียดของข้อมูลอย่างสูงทำให้ข้อมูลในแต่ละตัวอย่างมีการแสดงค่าความสัมพันธ์ระหว่างมวลโมเลกุลกับค่าIsoelectric point (pI)และค่าปริมาณความเข้มข้นของโปรตีน(Intensity) ที่แตกต่างกันมีช่วงข้อมูลที่กว้างมากทำให้เป็นปัญหาในการประมวลผล ดังนั้นในการประมวลผลของข้อมูลจำเป็นต้องมีการเตรียมข้อมูลที่เหมาะสมถือเป็นส่วนสำคัญการทำเหมืองข้อมูลสำหรับการจำแนกชนิดของข้อมูล 2D-Gel โดยรายละเอียดการเตรียมข้อมูลมีดังนี้

3.2.1 ลดขนาดของข้อมูล (Data reduction) แบ่งช่วงมวลโมเลกุลในแต่ละตัวอย่างให้มีช่วงที่เท่ากันในทุกๆ ตัวอย่าง แบ่งเป็นช่วง ช่วงละ 5 จากนั้นเลือกเฉพาะค่าที่มากที่สุดในแต่ละช่วงเป็นตัวแทนในช่วงนั้นรวมถึงเป็นการลดจำนวนมิติ (Dimension) ลงเพื่อให้เหมาะสมในการประมวลผลข้อมูล

3.2.2 แทนค่าปริมาณโปรตีนที่ขาดหาย ในแต่ละคอลัมน์นั้นๆ มีค่าเท่ากับศูนย์ เพื่อไม่ให้ส่งผลต่อรูปแบบเดิมของข้อมูลตั้งต้น

3.3.3 ขั้นตอนการแปลงรูปแบบข้อมูล (Transforming data) อยู่ในรูปแบบของลำดับจุด (Spot) โดยทำการ Pivot ค่ามวลโมเลกุลรวมกับค่า Isoelectric point (pI) ให้ตัวอย่างข้อมูลในหนึ่งแถวของข้อมูลแทนตัวอย่างหนึ่งตัวอย่าง ทั้งหมด 76 แถวหรือ 76 ตัวอย่างจำนวนฟีเจอร์หลังการ PIVOT 3301 ฟีเจอร์[15] เพื่อให้เหมาะสมและสามารถนำไปประมวลผลและเรียนรู้จดจำโปรแกรม R โดยดาต้าเฟรม(data frame) ของข้อมูลมีการแบ่งเป็นสองตัวอย่างที่ต้องการจำแนกประเภท โดยใช้ตำแหน่ง $X_{j,i}$,j แทนตัวอย่างแต่ละตัวอย่างแสดงเป็นแต่ละคอลัมน์ และ i แทนจุดของโปรตีน(Spot)แสดงเป็นแต่ละแถว ดังแสดงในตารางที่ 1

ตารางที่ 1: รูปแบบข้อมูลนำเข้าสำหรับการประมวลผลด้วยภาษา R

	Sample 1			Sample 2		
	Gel 1	Gel 2	Gel 3	Gel 1	Gel 2	Gel 3
Spot 1	$X_{1,1}$	$X_{2,1}$	$X_{3,1}$	$X_{1',1}$	$X_{2',1}$	$X_{3',1}$
Spot 2	$X_{1,2}$	$X_{2,2}$	$X_{3,2}$	$X_{1',2}$	$X_{2',2}$	$X_{3',2}$
Spot i	$X_{1,i}$	$X_{2,i}$	$X_{3,i}$	$X_{1',i}$	$X_{2',i}$	$X_{3',i}$

3.3 การคัดเลือกคุณลักษณะ

เมื่อข้อมูล 2D-Gel ผ่านกระบวนการเตรียมข้อมูลและแปลงรูปแบบข้อมูลจนสามารถนำมาเป็นข้อมูลอินพุตให้กับการคัดเลือกคุณลักษณะต่อด้วยวิธีต่างๆดังนี้

3.3.1 Moderated T –Test

เป็นการประมวลผลผ่านโปรแกรม R Library Prot2D [9] สำหรับวิเคราะห์ข้อมูล 2D-Gel โดยเฉพาะประกอบไปด้วยนอร์มัลไลซ์เพื่อปรับช่วงข้อมูลที่แตกต่างกันตามการแสดงออกของโปรตีน ในการหาความแตกต่างระหว่างสองหมวดหมู่ และใช้ False Discovery Rate (FDR) ในการกำหนดความถูกต้องของการจำแนก ยกตัวอย่างเช่นในจำนวนข้อมูล 100 ข้อมูลหากกำหนด FDR ที่ 0.2 หมายถึงกำหนดให้ข้อมูล 80 ข้อมูลสามารถจำแนกได้อย่างถูกต้อง ซึ่งจะมีประสิทธิมากกว่า p-value เมื่อนำมาใช้กับข้อมูลที่มีจำนวนมิติมากอย่างข้อมูลในทางชีววิทยา

3.3.2 Principal Component Analysis (PCA)

เป็นวิธีที่ได้รับความนิยมในการลดมิติของข้อมูลที่มีความซับซ้อน กระบวนการสำคัญสำหรับ PCA คือการทำนอร์มัลไลซ์ (Normalization) เพราะข้อมูลที่มีช่วงข้อมูลที่แตกต่างกันเป็นปัญหาสำคัญในการทำ PCA จึงจำเป็นต้องปรับช่วงข้อมูลให้อยู่ในช่วงเดียวกันเพื่อเพิ่มประสิทธิภาพการทำงานของ PCA โดยการทำนอร์มัลไลซ์ใช้เทคนิค Quantiles Normalization จากนั้นเข้าสู่ขั้นตอนการลดตัวแปรโดยกำหนดพารามิเตอร์ ให้ค่าเทคนิคการลดตัวแปรหรือ Dimensionality reduction เป็นแบบ keep variance และเลือกให้ Variance threshold ที่ 90 เปอร์เซ็นต์หมายถึงการกำหนดปริมาณสารเทศใหม่ที่ใช้แทนข้อมูลทั้งหมดที่ 90 เปอร์เซ็นต์

3.3.3 Sparse Partial Least Squares Discriminant Analysis (sPLS-DA)

การคัดเลือกคุณลักษณะด้วย sPLS-DA กระบวนการทำงานเป็นไปตามงานวิจัยของ Tomé S. Silva & Richard, 2016 [10] ซึ่งวิเคราะห์ข้อมูลผ่านโปรแกรม R Library mixOmics จากนั้นนำข้อมูลที่ผ่านการทำการเตรียมข้อมูลและการนอร์มัลไลซ์ (Normalization) มาคัดเลือกคุณลักษณะด้วย sPLS-DA ด้วยคำสั่งดังแสดงในภาพที่ 3

```
require(mixOmics)
results.splsda <- mixOmics::splsda(norm.
data, as.factor(treatment),
  ncomp=2,keepX=c(n.sig.spots,n.sig.spots))
tmp.load <- ((results.splsda$loadings)$X)[,1]
sig.spots.splsda <- names((tmp.load)[abs(tmp.
load) > (.Machine$double.eps)])
```

ภาพที่ 3: คำสั่งการคัดเลือกคุณลักษณะด้วย sPLS-DA

3.3.4 Recursive Feature Elimination (RFE)

ด้วยความสามารถของ Recursive Feature Elimination (RFE) ที่ใช้กับข้อมูลที่มีมิติของข้อมูลสูงและเป็นที่ยอมรับในการทำการคัดเลือกคุณลักษณะในข้อมูลชนิด gel-based proteomic จึงนำมาใช้ในการคัดเลือกคุณลักษณะโดย

กระบวนการทำงานเป็นไปตามงานวิจัยของ Tomé S. Silva & Richard, 2016 [10] นำข้อมูลที่ผ่านการทำการเตรียมข้อมูลและการนอร์มัลไลซ์ (Normalization) มาคัดเลือกคุณลักษณะด้วย RFE ด้วยคำสั่งด้วยคำสั่งดังแสดงในภาพที่ 4

```
require(caret)
require(klaR)
require(e1071)
rfeCtrl <- rfeControl(functions = nbFuncs,
method = "LOOCV")
results.rfe <-rfe(norm.data,factor(treatment),
  sizes=n.sig.spots,rfeControl=rfeCtrl)
sig.spots.nbrfe<-rownames(varImp(results.
rfe))[(1:n.sig.spots)]
```

ภาพที่ 4: คำสั่งการคัดเลือกคุณลักษณะด้วย RFE

3.4 การสร้างโมเดลในการทำนาย

โมเดลในการจำแนกประเภททั้งหมดสามเทคนิคได้แก่ Support Vector Machines (SVM), Artificial Neural Network (ANN) และ K-Nearest Neighbors (K-NN)

4. ผลการวิจัยและการอภิปรายผล

4.1 ข้อมูลที่ใช้ทดสอบระบบ

ข้อมูลใบยาสูบของกระทรวงสรรพสามิตที่ผ่านการบ่มแห้งและเข้าสู่การบวนการแยกโปรตีนด้วยเทคนิค Two-Dimensional Gel Electrophoresis (2D-Gel) และ Densitometer (GS-900 ™ Calibrated Densitometer) ทั้งหมด 76 ตัวอย่างประกอบไปด้วยใบยาสูบสามสายพันธุ์ สายพันธุ์ เวอร์จิเนีย 32 ตัวอย่าง สายพันธุ์เบอร์เลย์ 19 ตัวอย่างและสายพันธุ์เตอร์กิช 25 ตัวอย่าง

4.2 ผลลัพธ์การจำแนกประเภทในแต่ละสายพันธุ์

ในส่วนของประสิทธิภาพของการจำแนกประเภทใบยาสูบของแต่ละโมเดลร่วมกับการคัดเลือกคุณลักษณะได้นำเสนอไว้ในตารางที่ 2

ตารางที่ 2: ประสิทธิภาพการจำแนกประเภทในแต่ละสายพันธุ์

Model Classification	Virginia	Burley	Turkish
Moderated T –Test - K-NN	88.16	85.53	82.89
Moderated T –Test - SVM	77.63	82.89	73.68
Moderated T –Test - Neural Network	63.16	75.24	69.78
sPLS-DA - K-NN	81.43	78.75	63.04
sPLS-DA - SVM	81.07	88.04	73.75
sPLS-DA - Neural Network	91.96	89.64	65.89
PCA - K-NN	68.21	56.93	51.25
PCA - SVM	59.11	73.75	66.96
PCA - Neural Network	64.29	66.07	54.11
RFE - K-NN	59.2	73.39	68.21
RFE - SVM	75	73.57	62.32
RFE - Neural Network	73.68	77.86	59.39

ผลที่ได้จะเห็นได้ว่าโมเดลที่เหมาะสมในการจำแนกสายพันธุ์ใบยาสูบจากข้อมูลที่ได้รับจากห้องปฏิบัติการนี้ คือการทำงานร่วมกันระหว่างการคัดเลือกคุณลักษณะด้วยวิธี Moderated T –Test กับ เทคนิค K-NN ส่งผลให้ประสิทธิภาพในการจำแนกประเภททั้งสามสายพันธุ์ของใบยาสูบให้ค่าความแม่นยำมีค่ามากกว่า 80 เปอร์เซ็นต์ ในทุกสายพันธุ์

4.3 ผลลัพธ์การจำแนกประเภทในแต่ละสายพันธุ์โดยการจำลองข้อมูลขึ้นมาเพื่อทำการวิเคราะห์

เนื่องจาก 2D-Gel ที่ได้จากใบยาสูบที่ส่งขายให้กระทรวงสรรพสามิตทั้งหมดสามชนิดได้แก่ เวอร์จิเนีย เบอร์เลย์ และเตอร์กิช แสดงผลค่าความถูกต้องที่เหมาะสมในการจำแนกประเภทของแต่ล่ะสายพันธุ์ที่เป็นเพียงการทำงานร่วมกันระหว่างการคัดเลือกคุณลักษณะด้วยวิธี Moderated T –Test กับ เทคนิค K-NN ทีให้ค่าความถูกต้องมากกว่า 80 เปอร์เซ็นต์ในทุกสายพันธุ์ แต่อย่างไรก็ตามด้วยการทำงานของวิธีอื่นๆยังไม่สามารถให้ค่าความถูกต้องที่มากพอในการทำนายสายพันธุ์ใบยาสูบได้ อาจเป็นผลมาจากข้อมูลชุดฝึกสอนที่มีจำนวนตัวอย่างของแต่ล่ะสายพันธุ์ที่แตกต่างกันส่งผลให้การคัดเลือกคุณลักษณะที่เลือกใช้กับใบยาสูบทั้งสามสายพันธุ์ ยังไม่สามารถคัดเลือกฟีเจอร์ที่เป็นตัวแทนที่ดีที่สุดออกมาได้ จึงทำให้การคัดเลือกคุณลักษณะและการทดสอบประสิทธิภาพให้ผลการทดลองที่ยังไม่เป็นที่พอใจ จึงทำการจำลองชุดข้อมูลใหม่ ที่ใช้พื้นฐานจากข้อมูลเดิม โดยกำหนดให้มีจำนวนในแต่ละคลาส

เท่ากันที่คลาสละ 40 ตัวอย่างผ่านเทคนิคของ จึงทำการจำลองข้อมูลที่มีจำนวนตัวอย่างที่เท่ากันในแต่ละสายพันธุ์เพื่อสร้างข้อมูลชุดฝึกสอนโดยใช้วิธีการของ Artigaud, Gauthier et al. 2013.ในโปรแกรม R Library Prot2D [9] เป็นการสร้างข้อมูลจากข้อมูลเดิมด้วยคำสั่งดังแสดงในภาพที่ 5

```
Sim.data <- Sim.Prot.2D(data=pecten, nsp=700,
+       nr=10, p0=0.1, s2_0=0.2, d0=3)
```

ภาพที่ 5: คำสั่งการจำลองข้อมูลใน Library Prot2D

โดยรายละเอียดตัวแปรที่สำคัญได้แก่ data คือ ข้อมูลเริ่มต้นที่ต้องการนำไปจำลอง nsp คือ จำนวน spot หรือฟีเจอร์ที่ต้องการแต่ไม่เกินจำนวนฟีเจอร์ของข้อมูลตั้งต้น nr คือ จำนวนตัวอย่างในแต่ล่ะคลาสที่ต้องการ

และเมื่อได้ข้อมูลที่ผ่านการจำลองในแต่ละสายพันธุ์ที่มีจำนวนตัวอย่างเท่ากันในแต่ละคลาสจึงทำให้การคัดเลือกคุณลักษณะในแต่ละวิธีนั้นสามารถเลือกฟีเจอร์ที่ดีและสามารถเป็นตัวแทนของฟีเจอร์ทั้งหมดได้อย่างสมบูรณ์ทำให้การทดสอบประสิทธิภาพด้วยเทคนิคการจำแนกประเภททั้งสามชนิดให้ผลการทดลองที่น่าพอใจ ในส่วนของประสิทธิภาพของการจำแนกประเภทใบยาสูบของแต่ละโมเดลร่วมกับการคัดเลือกคุณลักษณะได้นำเสนอไว้ในตารางที่ 3

ตารางที่ 3: ประสิทธิภาพการจำแนกประเภทในแต่ละสายพันธุ์โดยการจำลองข้อมูลขึ้นมาเพื่อทำการวิเคราะห์

Model Classification	**Virginia**	**Burley**	**Turkish**
Moderated T –Test - K-NN	96.25	98.75	97.5
Moderated T –Test - SVM	100	100	100
Moderated T –Test - Neural Network	100	100	100
sPLS-DA - K-NN	100	98.75	100
sPLS-DA - SVM	100	100	100
sPLS-DA - Neural Network	100	100	100
PCA - K-NN	82.5	95	98.75
PCA - SVM	70	58.75	42.5
PCA - Neural Network	65	82.5	66.25
RFE - K-NN	98.75	98.75	100
RFE - SVM	100	100	100
RFE - Neural Network	100	100	100

ผลที่ได้จะเห็นได้ว่าประสิทธิภาพการจำแนกประเภทดีขึ้นหลังการปรับจำนวนตัวอย่างในแต่ละคลาสให้เท่ากัน มีเพียงการคัดเลือกคุณลักษะวิธี PCA เท่านั้นที่ให้ค่าความถูกต้องต่ำกว่า 80 เปอร์เซ็นต์

5. สรุป

จากการตีความและวิเคราะห์ผลข้อมูล 2D-Gel ที่มีจำนวนมิติ(Dimension) ของข้อมูลสูง ต้องใช้ระยะเวลานานในการวิเคราะห์และที่สำคัญต้องเป็นผู้เชี่ยวชาญที่ผ่านการฝึกอมรบเท่านั้นในการวิเคราะห์ข้อมูลชนิดนี้ วัตถุประสงค์ของงานวิจัยนี้คือการนำเสนอระบบในการทำนายข้อมูลชนิด 2D-Gel เพื่อเพิ่มประสิทธิภาพโมเดลการทำนาย จากการทดลองการจำแนกประเภทข้อมูลใบยาสูบผลการทดลองชี้ให้เห็นว่าการทำงานของ Moderated T –Test ร่วมกับการใช้ Model Classification เทคนิค K-NN มีความถูกต้องในการทำนายมากกว่า 80 เปอร์เซ็นต์ ในทุกสายพันธุ์ใบยาสูบ และเมื่อทำการแก้ปัญหาของความไม่เท่ากันของตัวอย่างในแต่ละคลาสโดยการจำลองชุดข้อมูลใหม่ ที่ใช้พื้นฐานจากข้อมูลเดิมพบว่าผลการจำแนกประเภทในทุกเทคนิคที่เปรียบเทียบมีค่าความถูกต้องในการทำนายมากกว่า 80 เปอร์เซ็นต์ เป็นเพราะการคัดเลือกคุณลักษณะในแต่ละวิธีนั้นสามารถเลือกฟีเจอร์ที่ดีและสามารถเป็นตัวแทนของฟีเจอร์ทั้งหมดได้อย่างสมบูรณ์ และมีเพียงการคัดเลือกคุณลักษณะวิธี PCA เท่านั้นที่ให้ค่าความถูกต้องต่ำกว่า 80 เปอร์เซ็นต์

งานวิจัยนี้แสดงให้เห็นว่าระบบในการทำนายข้อมูลชนิด 2D-Gel มีประสิทธิภาพในการจำแนกประเภทที่มากกว่า 80 เปอร์เซ็นต์ น่าจะเป็นประโยชน์ในการเลือกวิธีที่เหมาะสมในการจำแนกประเภทข้อมูล 2D-Gel รวมถึงการเตรียมข้อมูลที่ใช้ในงานวิจัยนี้สามารถนำไปใช้ในการจำแนกประเภทข้อมูล 2D-Gel สาขาอื่นๆ ทั้งนี้อาจต้องทดสอบกับชุดข้อมูลที่หลากหลายเพื่อศึกษาประสิทธิภาพที่แท้จริง

เอกสารอ้างอิง

1. Mosquera, E., et al., *Two-dimensional gel electrophoresis of Mytilus galloprovincialis: differences in protein expression between intertidal and cultured mussels.* Marine Ecology Progress Series, 2001. **224**: p. 149-156.
2. Rabilloud, T., et al., *Two-dimensional gel electrophoresis in proteomics: Past, present and future.* Journal of Proteomics, 2010. **73**(11): p. 2064-2077.
3. Fernandez-Lozano, C., et al., *Texture Classification of Proteins Using Support Vector Machines and Bio-inspired Metaheuristics*, in *Biomedical Engineering Systems and Technologies: 6th International Joint Conference, BIOSTEC 2013, Barcelona, Spain, February 11-14, 2013, Revised Selected Papers*, M. Fernández-Chimeno, et al., Editors. 2014, Springer Berlin Heidelberg: Berlin, Heidelberg. p. 117-130.
4. Cary, A., D. Pavisic, and R. Vargas, *Feedforward Neural Networks for the Classification of Two-dimensional Polyacrylamide Gel Electrophoresis Images.* Revista Acta Nova, 2003. **1**(3).
5. Malone, J., K. McGarry, and C. Bowerman, *Automated trend analysis of proteomics data using an intelligent data mining architecture.* Expert Syst. Appl., 2006. **30**(1): p. 24-33.
6. Sampson, D.L., et al., *A Comparison of Methods for Classifying Clinical Samples Based on Proteomics Data: A Case Study for Statistical and Machine Learning Approaches.* PLoS ONE, 2011. **6**(9): p. e24973.
7. Silva, T.S., et al., *Data visualization and feature selection methods in gel-based proteomics.* Current Protein and Peptide Science, 2014. **15**(1): p. 4-22.
8. Lee, K.R., et al., *Megavariate data analysis of mass spectrometric proteomics data using latent variable projection method.* PROTEOMICS, 2003. **3**(9): p. 1680-1686.
9. Artigaud, S., O. Gauthier, and V. Pichereau, *Identifying differentially expressed proteins in two-dimensional electrophoresis experiments: inputs from transcriptomics statistical tools.* Bioinformatics, 2013. **29**(21): p. 2729-34.
10. Silva, T.S. and N. Richard, *Visualization and Differential Analysis of Protein Expression Data Using R*, in *Statistical Analysis in Proteomics*, K. Jung, Editor. 2016, Springer New York: New York, NY. p. 105-118.
11. Marengo, E., et al., *New approach based on fuzzy logic and principal component analysis for the classification of two-dimensional maps in health and disease: Application to lymphomas.* Journal of Chromatography A, 2003. **1004**(1–2): p. 13-28.
12. Filzmoser, P., M. Gschwandtner, and V. Todorov, *Review of sparse methods in regression and classification with application to chemometrics.* Journal of Chemometrics, 2012. **26**(3-4): p. 42-51.
13. Chung, D. and S. Keles, *Sparse Partial Least Squares Classification for High Dimensional Data.* Statistical Applications in Genetics and Molecular Biology, 2010. **9**.
14. Kuhn, M., *Variable Selection Using The caret Package.* 2011.
15. สิทธิกร, ส., ระบบจำแนกสายพันธุ์ใบยาสูบโดยใช้เทคนิคดาต้า ไมน์นิ่ง. การศึกษาค้นคว้าอิสระ 2558. วิทยาศาสตร์มหาบัณฑิต สาขาเทคโนโลยีสารสนเทศ ภาควิชาวิศวกรรมคอมพิวเตอร์ มหาวิทยาลัยเกษตรศาสตร์.

การจัดรูปแบบข้อมูลเพื่อสร้างฐานข้อมูลในการวิเคราะห์รายงาน โดยใช้เทคนิคไฮบริดสคีมา

Data customization for ETL used Hybrid schema techniques

สุภาวดี กลมมิตร์ (Supawadee Klommit)[1] และ ศักดิ์ชาย ตั้งวรรณวิทย์ (Sakchai Tangwannawit)[2]
ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
s5807021857555@email.kmutnb.ac.th[1], sakchai.t@it.kmutnb.ac.th[2]

บทคัดย่อ

บทความนี้เป็นงานวิจัยที่ศึกษาและพัฒนารูปแบบข้อมูลโดยเป็นรูปแบบสำหรับพัฒนาระบบรายงานสรุปข้อมูลสินเชื่อเช่าซื้อภายในองค์กร เพื่อช่วยให้ผู้บริหารได้รับข้อมูลจากระบบที่ถูกต้องและรวดเร็ว ช่วยลดการใช้ทรัพยากรบุคลภายในองค์กรในการรวบรวมข้อมูลและจัดทำรายงาน โดยการทำงานในปัจจุบันมักพบข้อผิดพลาดในกระบวนการรวบรวมข้อมูล ซึ่งทำให้ผู้บริหารได้รับข้อมูลรายงานคลาดเคลื่อน ผู้วิจัยจึงพัฒนารูปแบบข้อมูล เพื่อธุรกิจอัจฉริยะขึ้น โดยใช้หลักการของการสร้างฐานข้อมูล การจัดการฐานข้อมูลตามรูปแบบของ Extract, Transform, Load (ETL Process) และการจัดโครงสร้างข้อมูล ในรูปแบบไฮบริดสคีมาสำหรับใช้ในการวิเคราะห์รายงาน

ผลจากการศึกษาพบว่า ระบบรายงานสรุปข้อมูลสินเชื่อเช่าซื้อภายในองค์กรที่ผู้วิจัยพัฒนาขึ้นสามารถใช้งานในองค์กรได้จริงและผลจากการประเมินการยอมรับการใช้งานระบบจากผู้ใช้มีค่าเฉลี่ยในแต่ละปัจจัยดังนี้ ด้านปัจจัยการรับรู้ประโยชน์ได้ค่าเฉลี่ย 4.33 ด้านปัจจัยการรับรู้ความง่ายในการใช้งาน ได้ค่าเฉลี่ย 4.38 ด้านทัศนคติต่อการใช้ ได้ค่าเฉลี่ย 4.53 ด้านพฤติกรรมการยอมรับ ได้ค่าเฉลี่ย 4.59 สรุปได้ว่าระบบรายงานสรุปข้อมูลสินเชื่อเช่าซื้อภายในองค์กรได้รับการยอมรับจากผู้ใช้งานระบบอยู่ในระดับดี

คำสำคัญ: กระบวนการจัดรูปแบบข้อมูล รายงานธุรกิจอัจฉริยะ ไฮบริดสคีมา

Abstract

The purposes of this research was to research and develop report system for Hire Purchase Loans. To help the manager get the accurate, reliable and prompt information from automate system. And to reduce the use of resources for gathering information and producing reports. Nowadays errors are found out in the process of gathering. It causes the management receives inaccurate reports. Therefore, the researcher developed the business intelligence system, using the principles of database. By using Database Management in the form of ETL process and hybrid schema data structure to analyse reports.

The results of the study was found out as followed. The system for reporting which represents data of Hire Purchase Loans developed by the researcher can be used in the organization. In addition of evaluating the user acceptance from users who has been using the system are: Perceived usefulness (PU) average is 4.33. Perceived Ease of Use (PE) average is 4.38. Attitude Toward Using (AU) average is 4.53. Behavioral Intention (BI) average is 4.59. To summarize, data collection system for Hire Purchase Loans has got the excellent feedback from users.

Keyword: ETL Process, Business Intelligence Report, Hybrid Schema.

1. บทนำ

ปัจจุบันการดำเนินธุรกิจมีการแข่งขันที่สูงขึ้นจากในอดีต องค์กรต้องพัฒนาและคิดกลยุทธ์ในการดำเนินธุรกิจมาเพื่อตอบสนองต่อสถานการณ์ที่เกิดขึ้นและเปลี่ยนแปลงตลอดเวลาตามสภาพเศรษฐกิจ การเมือง กฎหมาย และสังคม ยกตัวอย่าง เช่น การเลือกตั้ง การเปลี่ยนแปลงรัฐบาลในการบริหารประเทศ

การเปลี่ยนแปลงกฎหมาย ค่านิยมในการดำเนินชีวิตของประชากรในสังคม เป็นต้น ซึ่งกลยุทธ์เป็นแนวทางในการทำให้องค์กรประสบความสำเร็จประกอบด้วยคำถามที่สำคัญ 3 ข้อ ได้แก่ 1) ปัจจุบันเราอยู่ ณ จุดใด 2) เราต้องการไปถึงจุดใด และ 3) เราจะไปถึงจุดนั้นได้อย่างไร คือวิธีการหรือแนวทางที่องค์กรจะนำเสนอคุณค่า (Value) ให้กับลูกค้า เพื่อให้องค์กรประสบความสำเร็จตามที่ต้องการ โดยการนำแนวคิดการบริหารเชิงกลยุทธ์มาใช้ในองค์กร มีองค์ประกอบหลักที่สำคัญ 3 ประการ ได้แก่ 1) การวิเคราะห์กลยุทธ์ (Strategic Analysis) 2) การจัดทำกลยุทธ์ (Strategic Formulation) และ 3) การปฏิบัติตามกลยุทธ์ (Strategic Implementation)

จากการศึกษาการทำงานระบบเดิม พบว่าองค์กรใช้บุคลากรจำนวนมากเพื่อรวบรวมและวิเคราะห์ข้อมูลจากหลายฝ่ายภายในองค์กร เพื่อจัดทำรายงานในรูปแบบ Excel Files ซึ่งเกิดความล่าช้าและพบข้อผิดพลาดของข้อมูลบ่อยครั้ง ทำให้การแสดงผลรายงานคลาดเคลื่อน ส่งผลต่อการจัดการกลยุทธ์ของผู้บริหาร ดังนั้นปัญหาของผู้บริหารคือการวิเคราะห์ข้อมูลธุรกิจจำนวนมากในระยะเวลาที่จำกัด ถูกต้อง และเชื่อถือได้

จากปัญหาดังกล่าวระบบการวิเคราะห์ข้อมูลภายในองค์กรในรูปแบบธุรกิจอัจฉริยะจึงถูกพัฒนาขึ้น เพื่อช่วยสนับสนุนการตัดสินใจของผู้บริหารให้สามารถวิเคราะห์ข้อมูลได้ถูกต้องและรวดเร็ว ลดการใช้ทรัพยากรบุคลภายในองค์กรในการรวบรวมข้อมูลและจัดทำรายงานและได้รับข้อมูลจากระบบอัตโนมัติที่เชื่อถือได้ ทำให้ผู้บริหารสามารถกำหนดกลยุทธ์ในการดำเนินธุรกิจได้ทันต่อสถานการณ์ที่เกิดขึ้นในปัจจุบัน

2. ทฤษฎีที่เกี่ยวข้อง

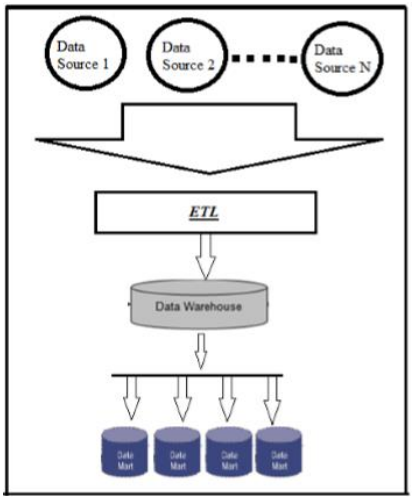
2.1 ETL Process

งานวิจัยหลายๆ งาน [1][2][3] ได้นำกระบวนการจัดรูปแบบข้อมูล (ETL Process) ไปใช้ ซึ่งการทำงานของกระบวนการจัดรูปแบบข้อมูลนั้น ข้อมูลใหม่จะถูกเพิ่มในคลังข้อมูลเป็นระยะโดยทั่วไปอาจจะเป็นรายเดือนรายสัปดาห์ รายวันหรือรายชั่วโมงขึ้นอยู่กับวัตถุประสงค์และประเภทของธุรกิจที่ให้ บริการ ซึ่งกระบวนการทำงานหลักประกอบด้วย

2.1.1 Data Cleansing คือ เมื่อมีข้อมูลมาแล้ว จะต้องมีการตรวจสอบความถูกต้องของข้อมูล แล้วปรับให้ดีขึ้น รวมทั้งกำจัดข้อมูลที่ผิดพลาดไป

2.1.2 Data Transformation คือ ข้อมูลบางอย่างที่ได้มา ซึ่งไม่ได้อยู่ในรูปแบบที่สามารถนำไปวิเคราะห์ได้ ง่ายๆ เช่น มีข้อมูลของคน เช่น วันเดือนปีเกิด ของลูกค้า เมื่อนำไปวิเคราะห์ไม่ได้เฉพาะเจาะจงเป็นวัน จะคำนวนเป็นช่วงอายุ เช่น วัยเด็ก ผู้ใหญ่ ชรา เป็นต้น ทำให้ต้องแปลงก่อนว่าเป็นวัยไหน

2.1.3 Data Loading and Refreshing คือ กำหนด Schedule ได้ว่าจะทำการโหลดข้อมูลทุกๆ กี่วัน รวมทั้งยังสามารถกำหนด Storage ปลายทางได้



ภาพที่ 1:กระบวนการจัดการข้อมูล

2.2 Star Schema

ได้มีการทดลองจัดรูปแบบโครงสร้างข้อมูล [4][5] เพื่อนำเปรียบเทียบความเหมาะสมและความรวดเร็วในการใช้งานในรูปแบบข้อมูลต่างๆ ซึ่งลักษณะของ Star Schema เป็นโครงสร้างแบบดวงดาว คือ มีตารางข้อเท็จจริง ที่เรียกว่า Fact Table อยู่ตรงกลาง ส่วนภายนอกถูกล้อมด้วยตารางมิติ ที่เรียกว่า Dimension Table

2.3 Fact Constellation Schema

จากการเปรียบเทียบความเหมาะสมและความรวดเร็วในการจัดรูปแบบโครงสร้างข้อมูล [4] Fact Constellation Schema เหมาะกับข้อมูลจำนวนมากที่มีการใช้บางตารางร่วมกัน เพื่อลดพื้นที่ในการจัดเก็บข้อมูล ประกอบด้วยตาราง Fact หลายๆ

ตาราง ที่มีการใช้ตาราง Dimensions ร่วมกัน จะได้ภาพที่มองเห็นเป็นการรวมกันของดาว จึงเรียกว่า รูปแบบกาแลคซี่ (Galaxy Schema) หรือกลุ่มดาว Fact (Fact Constellation)

2.4 Hybrid Schema

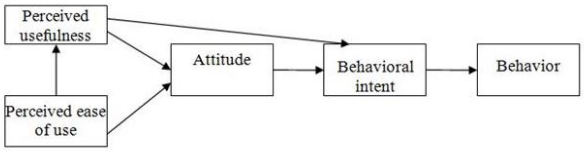
เป็นการนำการจัดรูปแบบโครงสร้างข้อมูลรูปแบบดาว และโครงสร้างข้อมูลรูปแบบกลุ่มดาว มาใช้ในกระบวนการจัดเตรียมข้อมูลรายงาน เนื่องจากจะช่วยให้การอ่านข้อมูลจากฐานข้อมูลมีความรวดเร็ว และประหยัดเนื้อที่ของการจัดเก็บข้อมูลในฐานข้อมูล ที่มีการใช้ข้อมูลบางส่วนร่วมกัน

2.5 Business Intelligence

ธุรกิจอัจริยะ [6][2] เป็นกระบวนการเพิ่มความสามารถในการแข่งขันของธุรกิจ โดยการนำเอาข้อมูลสารสนเทศที่มีอยู่มาก่อให้เกิดประโยชน์สูงสุด ส่วนประกอบของระบบรายงานธุรกิจอัจริยะนั้น เริ่มจากโปรแกรมในกลุ่มของกระบวนการจัดรูปแบบข้อมูล เพื่อนำข้อมูลที่ต้องการเข้าสู่คลังข้อมูล (Data Warehouse) เพื่อเตรียมข้อมูลนั้นให้เหมาะแก่การทำ On-Line Analytical Processing (OLAP)

2.6 Technology Acceptance Model (TAM)

ทฤษฎีแบบจำลองการยอมรับเทคโนโลยี [7] ถูกนำมาใช้เพื่ออธิบายปัจจัยต่างๆ ที่ส่งผลต่อการยอมรับเทคโนโลยี ซึ่งประกอบไปด้วย 4 ปัจจัยหลัก ได้แก่ ปัจจัยการรับรู้ประโยชน์ (PU : Perceived Usefulness) ปัจจัยการรับรู้ความง่ายในการใช้งาน (PE : Perceived Ease of Use) ทัศนคติต่อการใช้ (AU : Attitude Toward Using) พฤติกรรมการยอมรับ (BI : Behavioral Intention)



ภาพที่ 2:ทฤษฎีแบบจำลองการยอมรับเทคโนโลยี

3. วิธีการดำเนินการวิจัย

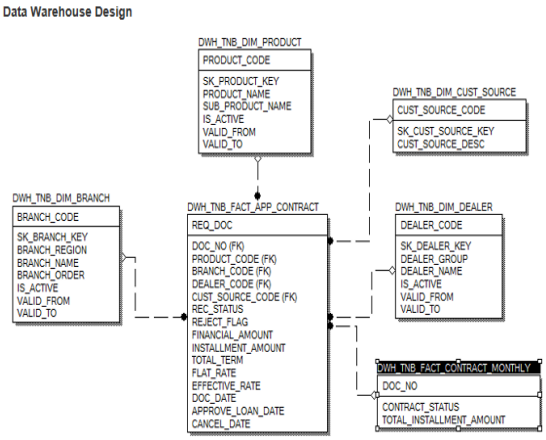
3.1 การทำความเข้าใจกับปัญหา (Planning)

กระบวนการของการศึกษาปัญหาของระบบเกี่ยวกับข้อมูลพฤติกรรมการทำงานของผู้ปฏิบัติงาน และความต้องการของผู้บริหาร ซึ่งศึกษาถึงกระบวนการทำงานทั้งภายในและภายนอกองค์กร นำสภาพแวดล้อมขององค์กรจากการสำรวจมารวบรวมเพื่อวิเคราะห์ถึงจุดบกพร่อง ในกระบวนการทำงานปัจจุบัน รวมถึงการรวบรวมหัวข้อของปัญหาที่พบ โดยให้ความสำคัญกับปัญหาที่มีผลกระทบกับองค์กรมากที่สุดเป็นลำดับแรกเพื่อจัดทำโครงการ ตามความเหมาะสม

3.2 การวิเคราะห์ (Analysis)

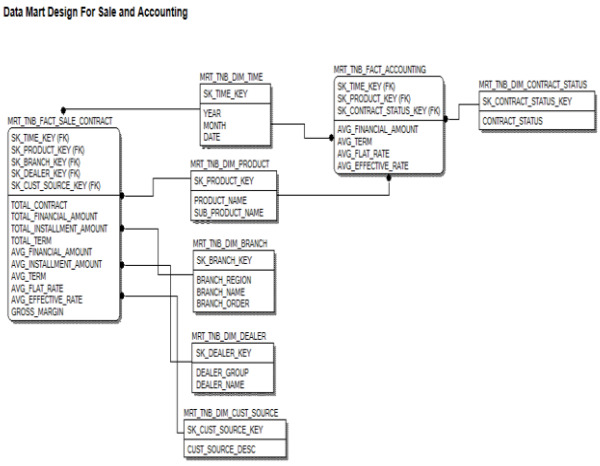
วิเคราะห์ถึงที่มาและความสำคัญของปัญหาในการทำวิจัย รวมถึงศึกษาความเป็นไปได้ในการพัฒนาด้วยเทคนิคการจัดโครงสร้างข้อมูลรูปดาวและการจัดโครงสร้างข้อมูลรูปแบบผสมร่วมกัน ว่าสามารถเข้ามาช่วยแก้ปัญหาดังกล่าวได้

1) การวิเคราะห์โครงสร้างข้อมูล (Data Model) ในส่วนของคลังข้อมูล โดยมีการจัดโครงสร้างความสัมพันธ์ของข้อมูลในรูปแบบโครงสร้างข้อมูลรูปดาว แสดงดัง**ภาพที่ 3**



ภาพที่ 3: โครงสร้างข้อมูลรูปแบบดาว

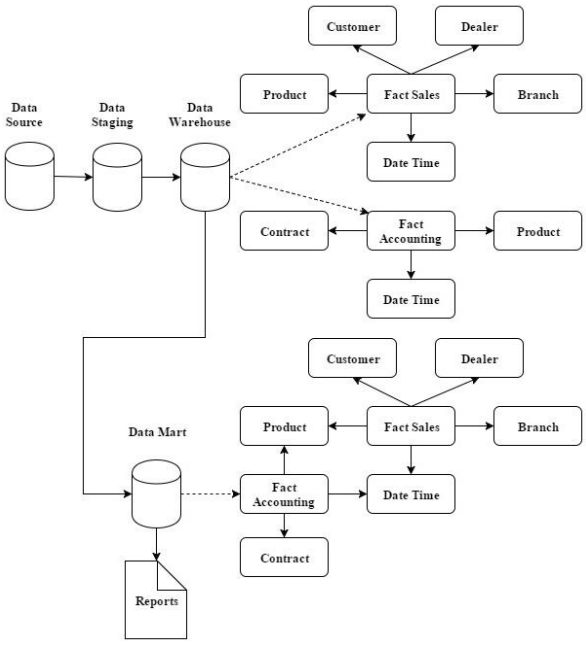
2) การส่งผ่านข้อมูลเข้าสู่กลุ่มข้อมูล (Data Mart) โดยทำการวิเคราะห์โครงสร้างข้อมูล เพื่อส่งผ่านข้อมูลจากคลังข้อมูล เข้าสู่กลุ่มข้อมูล แล้วทำการจัดโครงสร้างความสัมพันธ์ของข้อมูลให้เหมาะสมกับรูปแบบของการออกรายงาน โดยจะใช้โครงสร้างข้อมูลรูปแบบกลุ่มดาว แสดงดัง **ภาพที่ 4**



ภาพที่ 4: โครงสร้างข้อมูลรูปแบบกลุ่มดาว

3.3 การออกแบบ (Design)

สร้างแบบจำลองชุดข้อมูล โดยเริ่มจากการทำกระบวนการจัดรูปแบบข้อมูล เพื่อดึงข้อมูลเข้าสู่ส่วนของการจัดเรียงข้อมูล (Data Staging) และทำการส่งผ่านข้อมูลเข้าสู่คลังข้อมูล แล้วทำการจัดโครงสร้างความสัมพันธ์ของข้อมูลในรูปแบบโครงสร้างรูปดาว จากนั้นทำการส่งผ่านข้อมูลเข้าสู่กลุ่มข้อมูล แล้วทำการจัดโครงสร้างความสัมพันธ์ของข้อมูล ให้เหมาะสมกับรูปแบบของการวิเคราะห์และแสดงผลรายงาน โดยใช้การจัดโครงสร้างรูปแบบกลุ่มดาว โครงสร้างข้อมูล แสดงดัง**ภาพที่ 5**

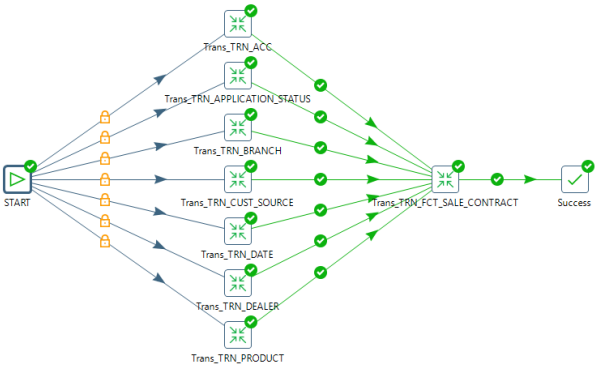


ภาพที่ 5:เทคนิคการจัดโครงสร้างข้อมูลรูปแบบไฮบริดสคีมา

3.4 การพัฒนา (Implementation)

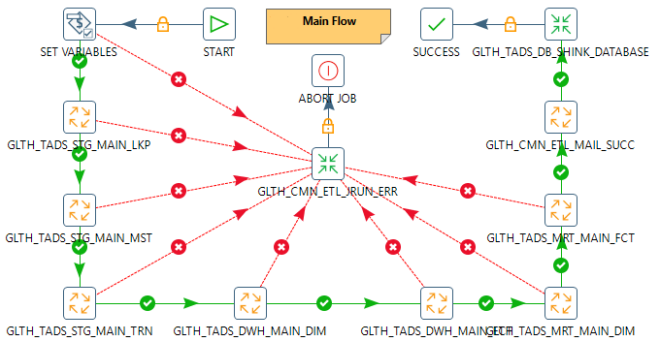
เป็นกระบวนการทำงานเพื่อทำการพัฒนาระบบให้สามารถใช้งานได้ตามวัตถุประสงค์ของการวิจัยและมีรูปแบบของการพัฒนาระบบตามแนวทางในการออกแบบระบบจากขั้นตอนของการออกแบบข้างต้น ซึ่งมีขั้นตอนการพัฒนา ดังนี้

3.4.1 พัฒนาแบบจำลองข้อมูลตามรูปแบบโครงสร้างที่ได้ออกแบบไว้เบื้องต้น ทั้งในส่วนของคลังข้อมูลและกลุ่มข้อมูล โดยการส่งผ่านข้อมูลจากคลังข้อมูลสู่กลุ่มข้อมูล แสดงดัง**ภาพที่ 6**



ภาพที่ 6: การส่งผ่านข้อมูลต้นทางเข้าสู่ตารางในฐานข้อมูล

3.4.2 พัฒนารายงานที่ช่วยสนับสนุนการตัดสินใจของผู้บริหารในด้านการขายและด้านบัญชีของธุรกิจสินเชื่อเช่าซื้อ และทำการตั้งเวลาการทำงานโดยอัตโนมัติให้กับกระบวนการทำงานแต่ละส่วนตามลำดับของการทำงาน แสดงดัง**ภาพที่ 7**



ภาพที่ 7: กระบวนการจัดการข้อมูลอัตโนมัติ

3.5 การทดสอบและประเมินผล (Test & Evaluation)

การทดสอบและประเมินระบบเป็นการทดสอบการยอมรับการใช้งานระบบรายงานสรุปข้อมูลสินเชื่อเช่าซื้อภายในองค์กรจริง ซึ่งเปรียบเทียบจากระบบเดิมใช้เวลาในการจัดเตรียมข้อมูลในการทำงาน 7 วันกับระบบใหม่ใช้เวลาไม่เกิน 1 ชั่วโมงในการจัดเตรียมข้อมูลและมีความผิดพลาดของข้อมูลน้อยกว่า โดยประเมินจากผู้ใช้งานจริง ด้วยวิธีการทำแบบสอบถามตามทฤษฎีแบบจำลองการยอมรับเทคโนโลยี Technology Acceptance Model (TAM) ซึ่งเกณฑ์การยอมรับจะพิจารณาจากคะแนนเฉลี่ยของปัจจัยการยอมรับการใช้งานระบบในแต่ละด้านต้องมีคะแนนเฉลี่ยอยู่ในระดับดี จึงจะยอมรับว่าระบบรายงานสรุปข้อมูลสินเชื่อเช่าซื้อภายในองค์กรที่ผู้วิจัยพัฒนาขึ้นตรงตามความต้องการของผู้ใช้ โดยผู้วิจัยได้กำหนดเกณฑ์ในการทดสอบไว้ 5 ระดับดังนี้ ดัง**ตารางที่ 1**

ตารางที่ 1: เกณฑ์การประเมิน

ระดับเกณฑ์การให้คะแนน		ความหมาย
เชิงคุณภาพ	เชิงปริมาณ	
ดีมาก	4.51 - 5.00	การยอมรับในระดับดีมาก
ดี	3.51 – 4.50	การยอมรับในระดับดี
ปานกลาง	2.51 – 3.50	การยอมรับในระดับปานกลาง
น้อย	1.51 – 2.50	การยอมรับในระดับน้อย
น้อยมาก	1.00 – 1.50	การยอมรับในระดับน้อยมาก

3.6 การบำรุงรักษา (Maintenance)

เป็นกระบวนการแก้ไข ปรับปรุงและดูแลระบบ หลังจากการพัฒนาระบบแล้วเสร็จ โดยผู้วิจัยทำการปรับปรุงหรือแก้ไขข้อมูลในระบบการแสดงรายงาน ในกรณีที่ผู้ใช้งานระบบพบข้อผิดพลาดหรือต้องการปรับเปลี่ยนความต้องการเบื้องต้น

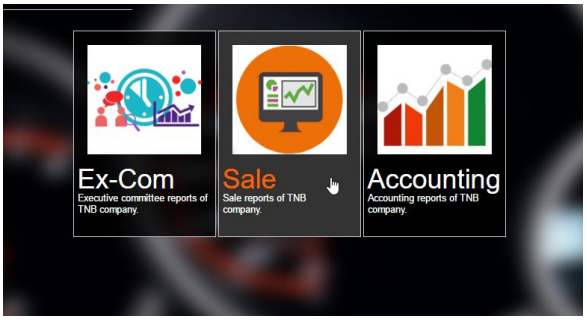
4. ผลการดำเนินงานและทดสอบระบบ

4.1 ผลการดำเนินงาน

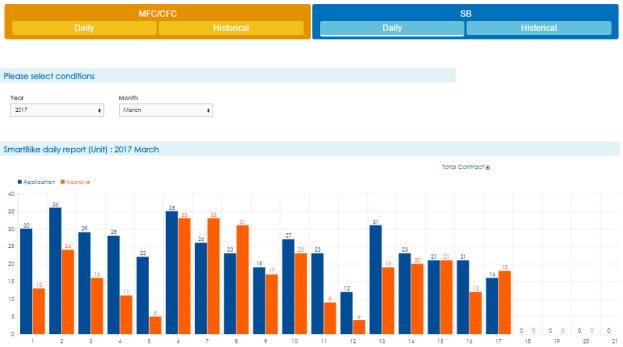
ประกอบไปด้วยส่วนของหน้าจอการใช้งานระบบและการแสดงผลรายงานสินเชื่อเช่าซื้อภายในองค์กร ดัง**ภาพที่ 8-11**



ภาพที่ 8: หน้าจอการเข้าใช้งานระบบ



ภาพที่ 9: หน้าจอการเลือกดูกลุ่มรายงานแยกตามส่วนงาน



ภาพที่ 10: หน้าจอแสดงผลรายงานรูปแบบกราฟ



ภาพที่ 11: หน้าจอแสดงผลรายงานรูปแบบตาราง

4.2 ผลการประเมินการยอมรับการใช้งานระบบ

สรุปผลการประเมินการยอมรับการใช้งานของผู้ใช้ระบบรายงานสรุปข้อมูลสินเชื่อเช่าซื้อภายในองค์กร ทั้ง 4 ด้าน สามารถสรุปผลการประเมินแต่ละด้าน ดังแสดงใน**ตารางที่ 2**

ตารางที่ 2: สรุปผลการประเมิน

รายการประเมิน	$\bar{X}$	SD	ระดับคุณภาพ
1. การรับรู้ประโยชน์	4.33	0.61	ดี
2. การรับรู้ความง่ายในการใช้งาน	4.38	0.61	ดี
3. ทัศนคติต่อการใช้	4.53	0.60	ดีมาก
4. พฤติกรรมการยอมรับ	4.59	0.55	ดีมาก
รวม	4.46	0.59	ดี

5. สรุป

จากการพัฒนาระบบรายงานสรุปข้อมูลสินเชื่อเช่าซื้อภายในองค์กร ด้วยวิธีการจัดรูปแบบข้อมูลเพื่อสร้างฐานข้อมูลในการวิเคราะห์รายงาน โดยใช้เทคนิคไฮบริดสคีมา เมื่อนำระบบที่พัฒนาขึ้นไปใช้งานและทำการประเมินผลการยอมรับการใช้งานของผู้ใช้แต่ละด้านแล้ว สรุปได้ว่าระบบสารมารถทำงานได้ตรงตามความต้องการของผู้ใช้ โดยให้ผลการยอมรับในระดับดี

ปัญหาและอุปสรรคที่พบสำหรับวิจัยนี้หลังจากการทดสอบระบบคือ โปรแกรมที่นำมาช่วยในการพัฒนารายงานมีข้อจำกัดในส่วนของการนำข้อมูลรายงานในรูปแบบของกราฟออกไปใช้งานต่อในรูปแบบ Excel ไม่สามารถทำได้

ข้อเสนอแนะในการพัฒนาต่อ ควรพัฒนาในส่วนของการดูข้อมูลรายงานที่ผู้ใช้งานระบบสามารถออกแบบมุมมองในการวิเคราะห์และสร้างรายงานอย่างง่ายได้ด้วยตนเองสำหรับใช้ในการวิเคราะห์รายงานในลักษณะเฉพาะ ซึ่งจะนำมาใช้ในการประชุมกรณีพิเศษ

เอกสารอ้างอิง

[1] Wiparat Thawitangkul, Oraluck Pattanapratеep, Surasak Leela-Udomlipi, “The use of business intelligence in hospital discharge process,” Journal of the Thai Medical Informatics Association, 1, 20-31, 2016.

[2] Aurelian TITIRISCA, “ETL as a Necessity for Business Architectures,” Database Systems Journal vol. IV, no. 2/2013, 2013.

[3] https://ismnida.wordpress.com/2009/09/26/etl%E0%B8%84%E0%B8%B7%E0%B8%AD%E0%B8%AD%E0%B8%B0%E0%B9%84%E0%B8%A3/

[4] Divya Sharma, “Dimension Modelling Techniques in Business Intelligence,” International Journal of Emerging Trends & Technology in Computer Science (IJETTCS), Web Site: www.ijettcs.org Email: editor@ijettcs.org, Volume 3, Issue 6, November-December 2014.

[5] Khalid Ibrahim Mohammed, “Data warehouse design and implementation based on quality requirements,” International Journal of Advances in Engineering & Technology, Vol. 7, Issue 3, pp. 642-651, July 2014.

[6] Muhammad Sohail, Khalid Khan, Rizwan Iqbal, and Noman Hasany, “Developing Business Intelligence Framework to Automate Data Mapping, Validation, and Data Loading from User Application,” International Journal of Computer Theory and Engineering, Vol. 4, No. 5, October 2012.

[7] F. D. Davis, “Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology,” MIS Quarterly, ABI/INFORM Global, pg. 319, 1989.

[8] Nitin Anand, “ETL and its impact on Business Intelligence,” International Journal of Scientific and Research Publications, Volume 4, Issue 2, February 2014.

[9] ชัชวาลย์ มุ่งแสง, งามนิจ อาจอินทร, สมจิตร อาจอินทร์, “Epidemiological Alert System by Using Business Intelligence," International Computer Science and Engineering Conference: ICSEC 2013 Thai Track Full Papers, 2013.

[10] Antonella Longoa, Sara Giacovellib, Mario A. Bochicchioc, “Fact - Centered ETL: A Proposal For Speeding Business Analytics Up,” Antonella Longo et al. / Procedia Technology 16, pp. 471 – 480, 2014.

[11] Akintola K.G., Adetunmbi A.O. Adeola O.S., “Building data warehousing and data mining from course management system: A case study of FUTA course management information systems,” nformation Technology for People-Centred Development (ITePED) 2011.

[12] Marketa Horakova, Hana Skalska, “Business Intelligence and Implementation in a Small Enterprise,” Journal of Systems Integration, Vol 4, No 2, 2013.

การประยุกต์ใช้การทำเหมืองข้อมูลกฎความสัมพันธ์ร่วมกับตรรกศาสตร์แบบคลุมเครือในบริบทความเข้มแข็งของครอบครัวไทย

Applying Fuzzy Association Rules Mining in the Context of Thai Strength Families

ภิญญาวัฒน์ สุขแก้ว (Phinyawat Sookgaew)[1] และ นพมาศ ปักเข็ม (Noppamas Pukkhem)[2]

สาขาวิชาคอมพิวเตอร์และเทคโนโลยีสารสนเทศ คณะวิทยาศาสตร์ มหาวิทยาลัยทักษิณ

[1]lasamdoria@hotmail.com,[2]noppamas@tsu.ac.th

บทคัดย่อ

บทความนี้นำเสนอกระบวนการสร้างตัวแบบเพื่ออธิบายความสัมพันธ์ระหว่างปัจจัยต่าง ๆ ที่มีผลต่อระดับความเข้มแข็งของครอบครัวไทย โดยการประยุกต์ใช้การทำเหมืองข้อมูลกฎความสัมพันธ์ร่วมกับตรรกศาสตร์แบบคลุมเครือเพื่อแก้ปัญหาการนิยามระดับความเข้มแข็งด้วยตรรกศาสตร์แบบจริงเท็จ โดยมีกระบวนการดังนี้ 1) การเตรียมข้อมูลด้านความเข็มแข็งในครอบครัวไทย 2) การคัดเลือกคุณลักษณะ 3) การออกแบบฟัซซี่เซตและสร้างตัวแบบกฎความสัมพันธ์แบบคลุมเครือ (Fuzzy Association Rules) 4) การประเมินผลตัวแบบ จากผลการวิจัยด้วยข้อมูลความเข้มแข็งของครอบครัวไทยในพื้นที่ อ.ควนขนุน จ.พัทลุง จำนวน 509 ครอบครัว ด้วยระดับ Fuzzy Support เท่ากับ 50% และ Fuzzy Confidence เท่ากับ 90% พบกฎความสัมพันธ์ที่มีค่าความเชื่อมั่นสูงที่สุดจำนวน 10 กฎ โดยกฎที่มีค่าสูงสุดมีค่าความเชื่อมั่นที่ 97% ค่า F-Measure เท่ากับ 85.6% จึงสามารถนำกฎที่ได้จากงานวิจัยไปใช้ในการวิเคราะห์และทราบถึงความสัมพันธ์ระหว่างปัจจัยหลักๆ ที่มีความน่าเชื่อถือเพิ่มขึ้นรวมถึงเป็นข้อมูลสนับสนุนสำหรับการจัดกลุ่มพื้นที่ความเข้มแข็งและการวางนโยบายต่าง ๆ ที่เกี่ยวข้องของประเทศต่อไป

คำสำคัญ: กฎความสัมพันธ์, ตรรกศาสตร์คลุมเครือ, ความเข้มแข็งของครอบครัว, เหมืองข้อมูล

Abstract

This paper proposes the process of associative factors model generation that effect to Thai strength families by using fuzzy association rule mining. This study will be able to solve the problem of Boolean logic explanation. The proposed process consist of 4 main steps: 1) Thai strength data preparation 2) feature selection 3) fuzzy set designing and fuzzy association rule generation 4) model evaluation. In this study, 509 records of Thai Families in Amphoe Khuan Khanun, Phattalung, were used with fuzzy support of 50% fuzzy confidence 90% and F-measure is 85.6%. The 10 fuzzy association rules were discovered and the maximum confidence of all rules is 97%. As the result of rule generation, we able to use the fuzzy association rule to support the area clustering of Thai strength families and able to use in related policy issue in the future.

Keywords: Association Rule, Fuzzy Logic, Strength Families, Data Mining

1. บทนำ

ครอบครัวเป็นสถาบันพื้นฐานที่เป็นหลักสำคัญที่สุดของสังคม ทำหน้าที่หล่อหลอมและขัดเกลาความเป็นมนุษย์ให้แก่สมาชิกของครอบครัว โดยกรมกิจการสตรีและเด็กสถาบันครอบครัวร่วมกับสถาบันวิจัยและให้คำปรึกษาแห่งมหาวิทยาลัยธรรมศาสตร์ [1] ได้จัดเก็บข้อมูลเกี่ยวกับข้อมูลทางด้านความเข้มแข็งของครัวเรือน โดยใช้แบบประเมินความเข้มแข็งของครอบครัวไทย [5] เป็นการเก็บข้อมูลโดยแบ่งเป็น 2 ส่วนหลัก ๆ คือ ส่วนที่ 1 ข้อมูลด้านครอบครัว ซึ่งเป็นข้อมูลพื้นฐานของครอบครัว เช่น เพศ อายุ รายได้เฉลี่ย อาชีพ ศาสนา สถานะภาพสมรส ลักษณะโครงสร้างของครัวเรือน ส่วนที่ 2 เป็นข้อมูลที่ใช้ในการประเมินระดับความเข้มแข็งของ

ครอบครัวไทย โดยกรมกิจการสตรีและเด็กสถาบันครอบครัวได้จัดทำเกณฑ์มาตรฐานกลาง ในการวัดความเข้มแข็งของครอบครัว [1] โดยครอบคลุม 5 ด้าน คือ 1) ด้านสัมพันธภาพ (Family relationship)

2) ด้านบทบาทหน้าที่ (Role)

3) ด้านการพึ่งตนเอง (Self-dependency)

4) ด้านทุนทางสังคม (Social capital)

5) ด้านการหลีกเลี่ยงภาวะเสี่ยง (Avoid risk factors)

แต่ละด้านจะมีความเข้มแข็งเป็น 3 ระดับคือ ระดับต่ำ ระดับปานกลาง และระดับสูง จากข้อมูลดังกล่าวที่เก็บรวบรวมยังขาดการค้นหาความสัมพันธ์ของปัจจัยในด้านต่างๆ ว่าปัจจัยใดบ้างที่เป็นปัจจัยหลักที่ส่งผลต่อความเข้มแข็งของครอบครัวไทย และมีความสัมพันธ์กันหรือไม่และสัมพันธ์กันอย่างไร นอกจากนี้ในการกำหนดช่วงคะแนนของการให้ระดับความเข้มแข็งของครอบครัวไทย ยังใช้วิธีการแบบการจัดระดับความเข้มแข็งตรรกศาสตร์แบบบลูลีนหรือแบบจริงเท็จ ซึ่งเป็นการกำหนดค่าระดับความเป็นสมาชิกในเซต (Degree of Membership) แค่ 0 กับ 1 เท่านั้น ส่งผลให้การวัดระดับความเข้มแข็งในครอบครัวอาจไม่ตรงกับความเป็นจริง เช่น การกำหนดระดับความเข้มแข็งมากที่ระดับคะแนน 80-100 และ กำหนดระดับความเข้มแข็งปานกลางที่ระดับคะแนน 50-79 ในกรณีครอบครัวมีคะแนนความเข้มแข็งที่ 80 คะแนน จะถูกจัดอยู่ในระดับความเข้มแข็งในระดับกับสูง ส่วนครอบครัวที่มีคะแนนความเข้มแข็งที่ 79 คะแนน จะถูกจัดอยู่ในกลุ่มระดับเข้มแข็งระดับปานกลาง ซึ่งจริงๆ แล้วทั้งสองครอบครัวมีคะแนนต่างกันเพียงแค่ 1 คะแนนเท่านั้น ดังนั้นเพื่อให้การอธิบายระดับความเข้มแข็งซึ่งเป็นปัญหาในรูปแบบความคลุมเครืองานวิจัยนี้จึงประยุกต์ใช้ตรรกศาสตร์แบบคลุมเครือมาใช้แก้ปัญหาการนิยามเซตแบบเดิม โดยใช้ร่วมกับการทำเหมืองข้อมูลกฎความสัมพันธ์ เรียกว่า กฎความสัมพันธ์แบบฟัซซี่ (Fuzzy Association Rule) เพื่อหาความสัมพันธ์ระหว่างปัจจัยต่าง ๆของข้อมูลในแบบสอบถาม เพื่อให้ทราบว่าปัจจัยใดบ้างที่ส่งผลหลักต่อความเข้มแข็งในครอบครัวไทย และใช้วิธีการอธิบายแต่ละปัจจัยด้วยวิธีการแบบฟัซซี่

2. ทฤษฎีที่เกี่ยวข้อง

2.1 ดัชนีความเข้มแข็งของครอบครัวไทย [1]

กรมกิจการสตรีและเด็กสถาบันครอบครัวได้จัดทำเกณฑ์มาตรฐานกลางในการวัดความเข้มแข็งของครอบครัวไทยและแบ่งระดับความเข้มแข็งของครอบครัวไทย โดยครอบคลุม 5 ด้าน โดยให้ระดับประเมินความเข้มแข็งของครอบครัวไทยในแต่ละด้านเป็น 3 ระดับ คือ ระดับต่ำ ระดับปานกลาง และระดับสูง ตามช่วงคะแนนที่ได้กำหนดไว้ในตารางที่ 1

ตารางที่ 1 ดัชนีความเข้มแข็งของครอบครัวไทย

ตัวชี้วัด	ช่วงคะแนน		
	ต่ำ	กลาง	สูง
1.ด้านสัมพันธภาพ	0.0-70.6	57.7-70.5	70.6-100
2.ด้านบทบาทหน้าที่	0.0-57.0	57.1-72.0	72.1-100
3.การพึ่งตนเอง	0.0-64.0	64.1-70.0	70.1-100
4.ทุนทางสังคม	0.0-63.0	63.1-75.0	75.1-100
5.หลีกเลี่ยงภาวะเสี่ยง	0.0-73.0	73.1-78.5	78.6-00

2.2 การทำเหมืองกฎความสัมพันธ์

กฎความสัมพันธ์ [2] เป็นการค้นหาความสัมพันธ์ระหว่างเซตรายการ (Itemsets)ในฐานข้อมูลขนาดใหญ่ ซึ่งกฎความสัมพันธ์สามารถสร้างขึ้นจากเซตรายการความถี่ (Frequent Itemsets) ความสำคัญของกฎทำการวัดโดยใช้มาตรวัดสองตัวด้วยกัน คือ ค่าสนับสนุน (Support) [8] ซึ่งเป็นเปอร์เซ็นต์ ของการเกิดขึ้นร่วมกันบ่อยของสิ่งของในเซตข้อมูลขนาดใหญ่ ดังสมการที่ (1) กำหนดให้ D เป็นเซตรายการขนาดใหญ่ ค่าสนับสนุน จาก $A \longrightarrow B$

$$P(A \cap B) = \frac{\text{(จำนวนรายการที่มีทั้ง A และ B)}}{\text{จำนวนรายการทั้งหมด D}} \qquad (1)$$

และข้อมูลตัวที่สองที่นำมาใช้วัดคือ ค่าความเชื่อมั่น (Confidence) [8] ซึ่งเป็นจำนวนของกรณีที่กฎถูกต้องโดยสัมพันธ์กับจำนวนของกรณีที่กฎสามารถนำไปใช้ได้ ดังสมการที่ (2) ค่าความเชื่อมั่น จาก $A \longrightarrow B$

$$P(A \mid B) = \frac{\text{จำนวนรายการที่มีทั้ง A และ B}}{\text{จำนวนรายการที่มี A}} \qquad (2)$$

ในการหากฎความสัมพันธ์นั้นจะมีอัลกอริทึมการหาหลายวิธีด้วยกัน แต่อัลกอริทึมที่เป็นที่รู้จักและใช้อย่างแพร่หลาย คือ อัลกอริทึมเอไพออริ (Apriori)

2.3 อัลกอริทึมเอไพออริ (Apriori)

อัลกอริทึม Apriori [3] เป็นอัลกอริทึมพื้นฐานที่ใช้สำหรับการค้นหากฎความสัมพันธ์ โดยใช้หลักการของการค้นหาเซ็ตของรายการที่ปรากฏบ่อยในฐานข้อมูลรายการด้วยค่าสนับสนุนที่มากกว่าค่าสนับสนุนขั้นต่ำ และสร้างกฎความสัมพันธ์จากเซ็ตของรายการที่ปรากฏบ่อยเหล่านั้นด้วยการใช้ค่าความเชื่อมั่นต่ำสุดเป็นเกณฑ์

2.4 ตรรกศาสตร์คลุมเครือ (Fuzzy Logic Approaches)

ตรรกศาสตร์แบบคลุมครือ [4] เป็นตรรกะที่อยู่บนพื้นฐานความเป็นจริงที่ว่าทุกสิ่งบนโลกแห่งความจริงไม่ใช่มีเฉพาะสิ่งมีความแน่นอนเท่านั้น แต่มีหลายสิ่งหลายเหตุการณ์ ที่เกิดขึ้นอย่างไม่เที่ยงและไม่แน่นอน (Uncertain) อาจเป็นสิ่งที่คลุมเครือ (Fuzzy) ไม่ได้ชัดเจน จากเหตุการณ์ดังกล่าวจึงนำตรรกะศาสตร์แบบคลุมเครือ (Fuzzy Logic) มาเป็นเครื่องมือที่ช่วยในการตัดสินใจภายใต้ความคลุมเครือของข้อมูลโดยยอมให้มีความยืดหยุ่นได้ใช้หลักเหตุผลที่คล้ายการเลียนแบบวิธีความคิดที่ซับซ้อนของมนุษย์ ตรรกศาสตร์แบบคลุมเครือมีลักษณะที่พิเศษกว่าตรรกะแบบจริงเท็จ (Boolean Logic) เป็นแนวคิดที่มีการต่อขยายในส่วนของความจริง (Partial True) โดยค่าความจริงจะอยู่ในช่วงระหว่างจริง (Completely True) กับเท็จ (Completely False) ส่วนตรรกศาสตร์เดิมจะมีค่าเป็นจริงเท็จเท่านั้น รูปแบบของฟังก์ชั่นสมาชิกสำหรับการแก้ปัญหาแบบฟัซซี่มีได้หลายรูป ผู้วิจัยของยกตัวอย่าง 2 แบบ ดังนี้

1. ฟังชั่นสามเหลี่ยม (Tringular Membersihp Function) [10] มีรูปแบบฟังก์ชั่นดังสมการที่ (3)

$$\text{Triangular}(x:a,b,c) = \begin{cases} 0 & x < a \\ (x-a)/(b-a) & a \le x \le b \\ (c-x)/(c-b) & b \le x \le c \\ 0 & x > c \end{cases} \quad (3)$$

2. ฟังก์ชันสี่เหลี่ยมคางหมู (Trapezoidal Membership Function) [10] มีรูปแบบฟังก์ชั่นดังสมการที่ (4)

$$\text{Trapezoidal}(x:a,b,c,d) = \begin{cases} 0 & x < a \\ (x-a)/(b-a) & a \le x < b \\ 1 & b \le x < c \\ (d-x)/(d-c) & c \le x < d \\ 0 & x \ge d \end{cases} \quad (4)$$

2.5 งานวิจัยที่เกี่ยวข้อง

Chiu และคณะ [6] ใช้วิธีการที่เรียกว่าการคลัสเตอร์ตามกฎความสัมพันธ์ (CBAR) ในการหากฎสัมพันธ์ของงานวิจัย วิธีการนี้มีการใช้คลัสเตอร์แบบตาราง เพื่อลดจำนวนในการหาข้อมูลและวิธีการนี้สามารถแยกออกเป็นสองประเด็น ประเด็นที่ 1 คือมีการค้นพบชุดข้อมูลขนาดใหญ่จากรายการฐานข้อมูล ประเด็นที่ 2 คือมีการใช้กราฟความสัมพันธ์ในการหากฎความสัมพันธ์จากฐานข้อมูลลูกค้าและรายการฐานข้อมูลโดยใช้ทฤษฎีฟัซซี่ในการกำหนดค่าความเป็นสมาชิกของลูกค้าในแต่ละกลุ่ม ซึ่งเป็นการแก้ปัญหาในเรื่องความคลุมเครือของการแบ่งกลุ่มลูกค้า

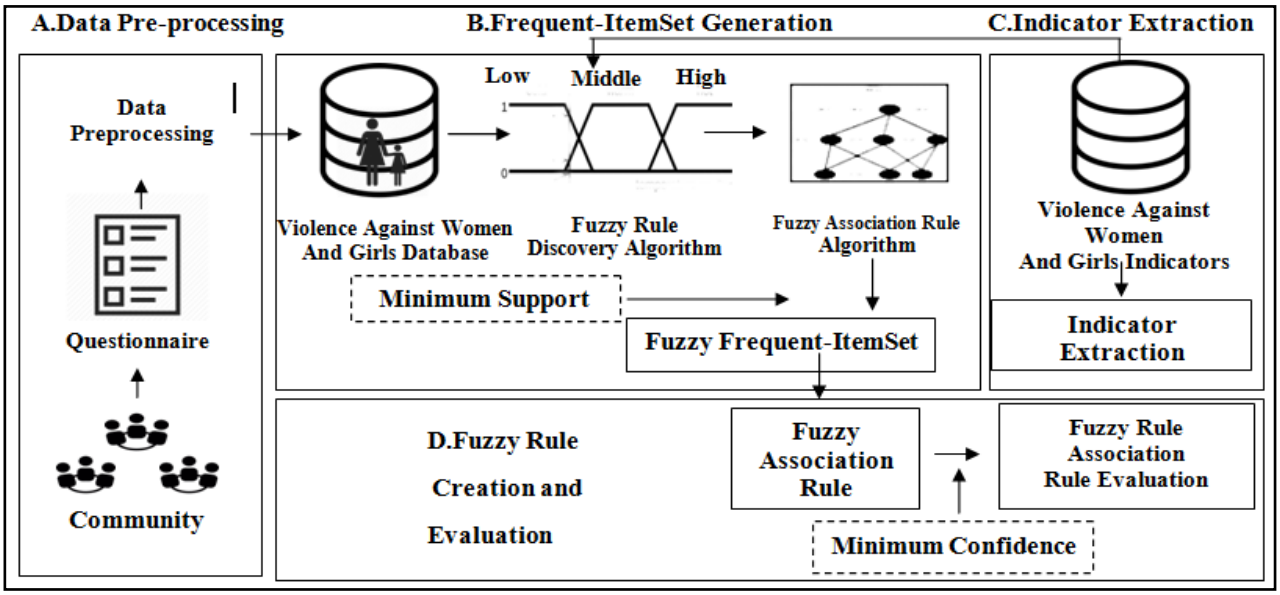
มณเฑียร และพงษ์เทพ [9] ได้ทำการวิจัยการคาดการณ์ผลการเรียนของนักศึกษาโดยใช้กฎความสัมพันธ์ ซึ่งนำผลการเรียนในแต่ละวิชามาหากฎความสัมพันธ์ โดยใช้ค่าสนับสนุนขั้นต่ำ 50% และค่าความเชื่อมั่นขั้นต่ำ 50% จากผลการวิจัยทำให้ได้กฎความสัมพันธ์ของการเรียนในแต่ละวิชา ซึ่งเป็นแนวทางสำหรับนักศึกษาในการเลือกเรียนวิชาที่น่าจะได้เกรดที่ดี

พงษ์เทพ และ อุไรวรรณ์ [7] ได้ใช้เทคนิคกฎความสัมพันธ์เพื่อหาความสัมพันธ์ของผลการเรียนของแต่ละวิชา และใช้เทคนิคฟัซซี่กฎความสัมพันธ์ เพื่อคาดการณ์ผลการเรียนรวมโดยเฉลี่ยว่าอยู่ในระดับต่ำ ปานกลาง หรือสูง เพื่อนำผลไปคัดกรองนักศึกษาที่เรียนวิชาและได้เกรดตามกฎที่ค้นพบเพื่อปรับวิธีการเรียนตั้งแต่ตอนต้น เพื่อที่จะมีผลการเรียนผ่านเกณฑ์ เมื่อเรียนจนครบหลักสูตร

โดยจะเห็นว่างานวิจัยต่าง ๆ ดังที่ได้กล่าวมาข้างต้นมีการประยุกต์ใช้กฎความสัมพันธ์ร่วมกับตรรกศาสตร์แบบคลุมเครือเพื่อแก้ปัญหาในโดเมนต่าง ๆ

3. กรอบแนวคิดการวิจัย

ในงานวิจัยนี้นำเสนอกรอบแนวคิดที่ประกอบไปด้วย 4 ส่วนหลัก ได้แก่ A) การเตรียมข้อมูล B) การสร้าง Fuzzy Frequent-ItemSet C) การสกัดดัชนีชี้วัดความเข็มแข็ง และ D) การสร้างกฎความสัมพันธ์แบบตรรกศาสตร์คลุมเครือและการประเมินค่าความเชื่อมั่นของกฎ โดยภาพกรอบแนวคิดแสดงดังภาพที่ 1



ภาพที่ 1 กรอบแนวคิดของงานวิจัย

กรอบแนวคิดของงานวิจัยมีรายละเอียดดังต่อไปนี้

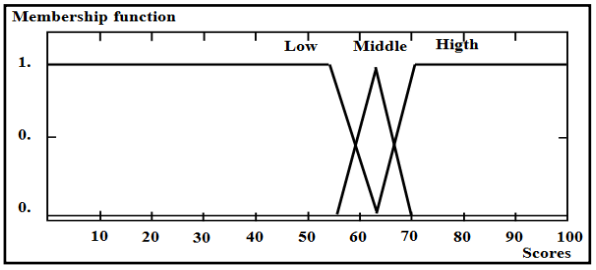
A.การเตรียมข้อมูล (Data pre-processing)

เป็นการบันทึกข้อมูลจากแบบประเมินความเข้มแข็งของครอบครัวไทย [5] ให้อยู่ในรูปแบบ CSV ไฟล์ ข้อมูลแยกเป็น 2 ส่วน คือส่วนที่ 1 เป็นข้อมูลครอบครัว เป็นลักษณะทั่วไปของครอบครัวเช่น เพศ, อายุ, รายได้เฉลี่ย เป็นต้น ส่วนที่ 2 เป็นข้อมูลที่เกี่ยวกับการประเมินความเข้มแข็งของครอบครัวไทยประกอบไปด้วย 5 ด้าน มีจำนวนคุณลักษณะทั้งหมด 49 คุณลักษณะ โดยลักษณะข้อมูลที่บันทึกมีลักษณะข้อมูลที่ไม่สมบูรณ์ (Incomplete Data) ในบางฟิลด์มีข้อมูลที่ขาดหาย (Missing Value) จึงต้องใช้หลักทางสถิติมาคำนวณเช่นค่าเฉลี่ย เพื่อเติมข้อมูลในส่วนที่หายไป หลังจากนั้นได้แปลงข้อมูลแปลง (Data Transformation) ข้อมูลบางคุณลักษณะที่เป็นลักษณะข้อมูลที่ต่อเนื่อง (Continuous) เช่น อายุ, รายได้ ให้เป็นลักษณะข้อมูลแบบไม่ต่อเนื่อง (Discrete) เนื่องจากข้อมูลมีจำนวนหลายคุณลักษณะจึงทำการเลือกคุณลักษณะที่สำคัญและมีผลต่อความเข้มแข็งของครอบครัวไทย (Feature Selection) โดยใช้อัลกอริทึม GainRatioAttributeEval ในโปรแกรม Weka Version 3.7.2 ในการคัดเลือกคุณลักษณะ จากจำนวนทั้งหมด 49 คุณลักษณะ คัดเลือกแล้วมีจำนวนคุณลักษณะที่ใช้ในการค้นหาความสัมพันธ์จำนวน 17 คุณลักษณะ

B. การสร้าง Fuzzy Frequent-ItemSet

การแบ่งโดยใช้ตรรกศาสตร์แบบบลูลีนซึ่งการใช้ตรรกศาสตร์แบบบลูลีนนั้นจะมีค่าคริส์พเซต (crisp set) เป็น 0 หรือ 1 เท่านั้น แต่ฟัซซี่เซตจะมีค่าระหว่าง 0 ถึง 1 แสดงถึงค่า ความเป็นสมาชิกมากน้อยเป็นลำดับ [6] กรณีการกำหนดเกณฑ์การให้คะแนนในด้านที่ 1 ของระดับความเข้มแข็งของครอบครัว คือด้านสัมพันธภาพนั้น จะมีช่วงคะแนนความเข้มแข็งอยู่ที่ 0-100 คะแนน และแบ่งระดับความเข้มแข็งของครอบครัวไทย เป็น 3 ระดับ คือ ความเข้มแข็งของครอบครัวระดับต่ำ ระดับปานกลาง และระดับสูง ถ้าเป็นตรรกศาสตร์แบบบลูลีน จะมีความเข้มแข็งระดับต่ำที่ช่วงคะแนน 0-57.6 ความเข้มแข็งระดับปานกลางที่ช่วงคะแนน 57.7-70.5 และความเข้มแข็งระดับสูงที่ช่วงคะแนน 70.6-100 หากครอบครัวที่ 1 ได้คะแนนความเข้มแข็งของครอบครัว ที่คะแนน 70.6 ซึ่งผลที่ได้จากการให้คะแนนแบบตรรกศาสตร์แบบบลูลีนครอบครัวที่1 จะเป็นครอบครัวที่มีระดับความเข้มแข็งสูง ครอบครัวที่ 2 ได้คะแนนความเข้มแข็งของครอบครัวที่ 70.5 ซึ่งผลที่ได้จากการให้คะแนนแบบตรรกศาสตร์แบบบลูลีนครอบครัวที่ 2 จะเป็นครอบครัวที่มีระดับความเข้มแข็งปานกลาง ซึ่งทั้ง 2 ครัวเรือนได้ระดับคะแนนที่ต่างกันแค่ 0.1 แต่ถูกจัดอยู่ให้อยู่ต่างระดับกัน ส่วนการจัดระดับความเข้มแข็งของครอบครัวโดยใช้ตรรกศาสตร์แบบคลุมเครือนั้นจะแบ่งเป็นเซตระดับ

ความเข้มแข็งต่ำ (L) ปานกลาง (M) และสูง (H) ซึ่งแต่ละระดับความเข้มแข็งจะมีค่าความเป็นสมาชิกในแต่ละเซต แสดงตัวอย่างไว้ในตารางที่ 3 ครอบครัวที่ 2 ได้คะแนนความเข้มแข็งที่ 60 คะแนน มีค่าความเป็นสมาชิก 0.2 0.8 และ 0.0 ในเซตต่ำ เซตปานกลางและเซตสูงตามลำดับ เป็นต้น จากนั้นนำค่าตัวชี้วัดระดับความเข้มแข็งของครอบครัวไทยมาใช้ในการออกแบบ Membership Function เป็นฟังก์ชั่นที่มีการกำหนดระดับความเป็นสมาชิกของตัวแปรที่ต้องการใช้งานโดยเริ่มจากแทนที่กับตัวแทนที่มีความไม่ชัดเจนไม่แน่นอนและคลุมเครือ ซึ่งในงานวิจัยนี้ก็คือระดับความเข้มแข็งของครอบครัวทั้ง 5 ด้าน ซึ่งผู้ออกแบบได้ออกแบบฟังก์ชันความเป็นสมาชิกด้วยฟังก์ชันสามเหลี่ยม และสี่เหลี่ยมคางหมู [6] โดยได้กำหนดค่าพารามิเตอร์ของแต่ละระดับดังต่อไปนี้ Low (0.0 0.0 53.3 62.5), Middle (54.4 62.0 69.8), High (62.7 69.8 100 100) จากการกำหนดค่าดังกล่าวจะได้ดังภาพที่ 2



ภาพที่ 2 Membership Function ความสัมพันธ์ในครอบครัว

เมื่อออกแบบ Membership Function ของทั้ง 5 ด้านจะได้ผลสรุปการออกแบบ Membership Function ของการวัดความเข้มแข็งของครอบครัวไทย 5 ด้าน ดังตารางที่ 2 และตารางที่ 3 เป็นการแสดงตัวอย่างการแทนค่าคะแนนดิบเพื่อให้ได้ค่าสมาชิกแบบฟัซซี่ โดยขอยกตัวอย่างการแทนค่าในด้านที่ 1 ด้านสัมพันธภาพ (Family Relationship) จากทั้งหมด 5 ด้าน

ตารางที่ 2 ตารางสรุปการออกแบบ Membership Function

ประเด็นปัจจัย	Fuzzy Set		
	ระดับต่ำ (L)	ระดับกลาง (M)	ระดับสูง (H)
1. ด้านสัมพันธภาพ	0-62.5	54.4-70.0	62.5-100
2. ด้านบทบาทหน้าที่	0-64.2	55.2-72.2	64.2-100
3.ด้านการพึ่งตนเอง	0-64.2	58.0-72.0	64.2-100
4.ด้านทุนทางสังคม	0-68.6	60.0-75.0	68.6-100
5.ด้านการหลบเลี่ยงภาวะเสี่ยง	0-76.8	70.8-80.6	76.8-100

ตารางที่ 3 การให้คะแนนแบบตรรกศาสตร์แบบคลุมเครือ

คะแนนดิบ	Membership Function			Fuzzy Set
	Low	Middle	High	
58.0	0.5	0.5	0.0	M
60.0	0.2	0.8	0.0	M
65.0	0.0	0.5	0.5	M
68.0	0.0	0.1	0.9	H
70.0	0.0	0.0	1.0	H

จากวิธีการดังกล่าว งานวิจัยนี้ได้ทำการแปลงคะแนนดิบจำนวน 509 ข้อมูล ในด้านด้านสัมพันธภาพ และทำการแปลงคะแนนดิบเพิ่มเติมอีก 4 ด้านที่เหลือ คือ ด้านบทบาทหน้าที่, ด้านการพึ่งตนเอง, ด้านทุนทางสังคมและด้านการหลีกเลี่ยงภาวะเสี่ยง แล้วนำข้อมูลทั้ง 5 ด้านมาหากฎความสัมพันธ์ร่วมกับคุณลักษณะที่ได้ในตอนแรกอีก 12 คุณลักษณะรวมทั้งหมด 17 คุณลักษณะเพื่อสร้างเซตรายการความถี่ ในขั้นตอนนี้เป็นการคัดเลือกชุดความสัมพันธ์ระหว่างเหตุการณ์หรือค่าระหว่างคุณลักษณะต่างๆที่เกิดขึ้นร่วมกันในการคัดเลือก เซตรายการความถี่จะคัดเลือกจากความถี่ของรายการที่มีค่ามากกว่าค่าสนับสนุนขั้นต่ำ (Minimum Support) ที่ได้กำหนดไว้ที่ 50 %

C. การสกัดดัชนีชี้วัดจากฐานข้อมูลความเข้มแข็งของครอบครัวไทย

ขั้นตอนนี้เป็นการนำดัชนีชี้วัดสากลมาสกัดเป็นเกณฑ์สำหรับการกำหนดค่าฟังก์ชันความเป็นสมาชิกให้กับปัจจัยแต่ละด้านของงานวิจัย

D. การสร้างกฎความสัมพันธ์ร่วมกับตรรกศาสตร์คลุมเครือและการประเมินค่าความเชื่อมั่นของกฎ

เป็นขั้นตอนการนำข้อมูลจากขั้นตอน B และ C (ภาพที่ 1) มาประมวลผลด้วยโปรแกรม Weka Version 3.7.2 โดยใช้อัลกอริทึม Association ค้นหากฎตรรกศาสตร์แบบคลุมเครือ (Fuzzy Rule Discovery Algorithm) ผลลัพธ์ได้จากกระบวนการนี้คือตัวแบบของงานวิจัย หรือเรียกว่ากฎความสัมพันธ์ตรรกศาสตร์คลุมเครือ (Fuzzy Association Rule) ซึ่งต่อมาจะถูกคัดเลือกกฎที่น่าสนใจด้วยการนำมาเปรียบเทียบกับค่าความเชื่อมั่น(Confidence) ที่มีค่าน้อยที่สุดโดยได้กำหนดไว้ที่ 90%

4. ผลการดำเนินการวิจัยและการประเมินผล

จากข้อมูลความเข้มแข็งของครอบครัวไทยในปี 2559 ของเทศบาลอำเภอควนขนุน อำเภอควนขนุน จังหวัดพัทลุง โดยใช้เทคนิคฟัซซี่ (Fuzzy Association Rule mining) ค้นหาความความสัมพันธ์ กำหนดค่าสนับสนุน (Support) ที่ 50% และค่าความเชื่อมั่น (Confidence) ที่ 90% และกำหนดจำนวนกฎ 10 กฎ โดยมีการวัดค่าประสิทธิภาพของตัวแบบด้วยค่า F-Measure = 85.6% แสดงดังตารางที่ 4

ตารางที่ 4 กฎความสัมพันธ์แบบฟัซซี่และค่าความเชื่อมั่นของกฎ

กฎความสัมพันธ์	ความเชื่อมั่น
1. role=High 444 ==> self reliance=High 433	97%
2. Religion=Buddhism role=High 430 ==> self reliance=High 419	97%
3. Family structure=Single Extra Family=None 429 ==> self reliance=High 416	97%
4. Religion=Buddhism Family structure=Single 446 ==> self reliance=High 432	97%
5. Religion=Buddhism Extra Family=None 446 ==> self reliance=High 432	97%
6. role=High 444 ==> Religion=Buddhism 430	97%
7. role=High self reliance=High 433 ==> Religion=Buddhism 419	97%
8. Family structure=Single 463 ==> self reliance=High 448	97%
9. Extra Family=None 463 ==> self reliance=High 448	97%
10. Religion=Buddhism 491 ==> self reliance=High 475	97%

จากตารางที่ 4 พบว่ากฎข้อที่ 1 แสดงให้เห็นว่าครอบครัวที่มีระดับความเข้มแข็งด้านการทำบทบาทหน้าที่ของครอบครัวอยู่ในระดับที่ดี จะส่งผลให้การพึ่งพาตนเองของครอบครัวอยู่ในระดับที่ดีด้วย กฎข้อที่ 2 ครอบครัวที่นับถือศาสนาพุทธ และมีการทำบทบาทหน้าที่ครอบครัวอยู่ในระดับที่ดีแล้วจะส่งผลให้การพึ่งพาตนเองของครอบครัวอยู่ในระดับที่ดีด้วย นอกจากนี้การอธิบายระดับคะแนนในแต่ละด้านด้วยนิยามแบบฟัซซี่เซตจะช่วยเพิ่มประสิทธิภาพในเชิงรายละเอียดของปัญหาแบบคลุมเครือได้ดียิ่งขึ้น

5. สรุปและอภิปรายผล

การใช้กฎความสัมพันธ์ร่วมกับฟัซซี่เพื่อค้นหาความสัมพันธ์และคาดการณ์ระดับความเข็มแข็งของครอบครัวไทย โดยเฉพาะคุณลักษณะด้านการพึ่งพาตนเองจะส่งผลในความสัมพันธ์ต่อหลายคุณลักษณะ ซึ่งสังเกตได้จากกฎจำนวน 10 กฎ พบว่า จากการกำหนดค่า Fuzzy Support ที่ 50 เปอร์เซ็นต์ และค่า Fuzzy Confidence ที่ 90 เปอร์เซ็นต์ กำหนดจำนวนกฎ 10 กฎพบว่ามีถึง 9 กฎคิดเป็น 90 เปอร์เซ็นต์ ที่มีคุณลักษณะของด้านการพึ่งพาตนเอง(Self Reliance) ร่วมด้วย ซึ่งสังเกตได้ว่าหากมีสนับสนุนให้ครอบครัวมีความเข้มแข็งในด้านการพึ่งพาตนเองอยู่ในระดับที่ดีแล้วก็จะส่งผลต่อคุณลักษณะอื่นๆ ของครอบครัวตามมาด้วย นอกจากตัวแบบจะสามารถใช้อธิบายความสัมพันธ์ระหว่างปัจจัยที่มีการเกิดขึ้นร่วมกันบ่อยแล้ว ตัวแบบจากการทดลองยังสามารถนำไปประยุกต์ใช้กับพื้นที่อื่น ๆ ของประเทศไทยและสามารถนำไปสู่การจัดกลุ่มพื้นที่ความเข้มแข็งของครอบครัวไทยเพื่อการวางนโยบายที่เกี่ยวข้องได้อย่างเหมาะสมต่อไป

เอกสารอ้างอิง

[1] *รายงานโครงการพัฒนาระบบมาตรฐานครอบครัวเข้มแข็ง* กรมกิจการสตรีและสถาบันครอบครัว กรุงเทพฯ 2557.

[2] Agrawal, R. and Srikant, “Fast algorithms for mining generalized association rules”,*in: Proceeding of the 20th International Conferrence on Very Large Databases VLDB94,*Santiago, Chile, pp.487- 499,1994.

[3] Yu-Shiang, Hung Kuei-Ling B, Chen Chi-Ta and Yang Guang-Feng Deng, “Mining Cluster-based Patterns for Elder Self-care Behavior”, *Innovative DigiTech-Enabled Applications & Services Institute Institute for Information Industry,*Taiwan,2012.

[4] *ระบบฟัซซีและโครงข่ายประสาทเทียม* ศูนย์ผลิตตำราเรียนพระจอมเกล้าพระนครเหนือ กรุงเทพฯ 2555.

[5] แบบประเมินความเข้มแข็งของครอบครัว กรมกิจการสตรีและสถาบันครอบครัว กรุงเทพฯ 2559.

[6] Hung-Pin Chiu, Yi-Tsung Tang and Kun-Lin Hsieh, “ Applying cluster-based fuzzy association rules mining framework into EC environment”,*Applied Soft Computing* 12 ,2012.

[7] พงษ์เทพ รักผกาวงศ์ และ อุไรวรรณ์ รักผกาวงศ์, “ การใช้กฎความสัมพันธ์ ร่วมกับฟัซซี่กฎความสัมพันธ์ เพื่อคาดการณ์ผลการเรียนของนักศึกษา ” , *ว.วิทย มข* 43 2558 หน้า 542-551.

[8] *การทำเหมืองข้อมูล Data Mining* จามจุรีโปรดักส์ จำกัด กรุงเทพฯ 2558.

[9] มณเฑียร อ่อนสำลี และ พงษ์เทพ รักผกาวงศ, “ การคาดการณผลการเรียนของนักศึกษา โดยใชกฎความสัมพันธ์ ”,*การประชุมวิชาการระดับชาติ* ครั้งที่ 7 30-31 มีนาคม 2558 หน้า 1238-1243.

[10] สุวรรณี ธูปจีน และ มณเฑียร รัตนศิริวงศ์วุฒิ , “ การตัดเกรดแบบวิธีการฟัซซีเซต สาหรับการประเมินผลการเรียนรู้ของผู้เรียน ”, *วารสารวิทยาศาสตร์และเทคโนโลยี* 2017.

รูปแบบการตัดสินใจสำหรับระบบแก้ไขปัญหาการใช้งานโดยใช้เทคนิคเหมืองข้อมูล
Decision Model for Helpdesk System using Data Mining Technique

ธันย์ชนก เนาวรัตน์กูล (Thanchanok Naowaratkul)[1]และศักดิ์ชาย ตั้งวรรณวิทย์ (Sakchai Tangwannawit)[2]
ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ
มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
[1]s5807021857083@email.kmutnb.ac.th, [2]sakchai.t@it.kmutnb.ac.th

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนารูปแบบการตัดสินใจสำหรับระบบแก้ไขปัญหาการใช้งาน เพื่อให้สามารถปฏิบัติงานได้อย่างมีประสิทธิภาพ โดยการใช้เทคนิคเหมืองข้อมูลในแบบการจำแนกข้อมูล (Classification) เพื่อเปรียบเทียบอัลกอริทึมที่เหมาะสมกับรูปแบบการตัดสินใจสำหรับระบบแก้ไขปัญหาการใช้งาน โดยนำเอาเทคนิคเหมืองข้อมูล (Data mining) มาใช้ในการเตรียมข้อมูลเพื่อการค้นหากฎการจำแนกข้อมูลที่ให้ค่าความถูกต้องมากที่สุด และอัลกอรึทึมที่ให้ค่าความถูกต้องมากที่สุดคือการจำแนกข้อมูลด้วยต้นไม้ตัดสินใจ (Decision Tree : J48) ซึ่งให้ ค่าความแม่นยำถึง 99.92% และนำเทคนิคการวิเคราะห์ข้อมูลที่ได้มาประยุกต์ใช้ในการแนะนำการส่งต่อผู้รับผิดชอบงานใน แต่ละงานให้มีความถูกต้องและเหมาะสม ซึ่งสามารถนำมาจัดวางแผนการใช้งานและเพื่อป้องกันปัญหาที่จะเกิดขึ้นกับระบบในอนาคต ได้ ผลการประเมินการยอมรับเทคโนโลยีที่ส่งผลต่อความพึงพอใจในการใช้งาน โดยมีผู้ประเมินสองกลุ่มคือ กลุ่มผู้เชี่ยวชาญมีค่าเฉลี่ยเท่ากับ 3.92 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.63 และส่วนของผู้ใช้งานมีค่าเฉลี่ยเท่ากับ 4.13 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.57 สรุปได้ว่า ผู้ใช้งานระบบมีการยอมรับเทคโนโลยีที่ส่งผลต่อความพึงพอใจในการใช้งานอยู่ในระดับดี ดังนั้นระบบที่พัฒนาขึ้นนี้สามารถนำไปใช้งาน วิเคราะห์เหตุการณ์ และช่วยในการส่งต่อปัญหาหรือส่งต่อการซ่อมงานให้กับผู้รับผิดชอบได้

คำสำคัญ: เหมืองข้อมูล การจำแนกประเภทข้อมูล ต้นไม้ช่วยการตัดสินใจ ระบบสนับสนุนการใช้งาน

Abstract

The purpose of this research is to develop the decision model for helpdesk system for effective work performance. The methodology includes data mining classification technique to data preparation for searching the most correct classification rule. The most accurate algorithm is the decision tree (J48), which provides accuracy up to 99.92% and then bring the data analysis technical to the person who responsible appropriately and correctly. This outcome of research is operation plan and preventing the problem which will be happen in the future. After development, the system was evaluated on user technology acceptance form that affect the user satisfaction by two groups of users: experts and users. The average score from experts was 3.92 with standard deviation of 0.63. The average score from users was 4.13 with standard deviation of 0.57. The system can be concluded that users accept. The user's satisfaction is very satisfied. Therefore, the system is able to use for analysis issue and transferring the problem to the responsible team.

Keyword: Data Mining, Classification, Decision Tree, Helpdesk System

1. บทนำ

ระบบสารสนเทศในองค์กรมีขนาดใหญ่มากขึ้น และมีความซับซ้อนทั้งในเรื่องของโครงสร้าง และขั้นตอนการปฏิบัติงาน รูปแบบการตัดสินใจสำหรับระบบแก้ไขปัญหาการใช้งาน (Helpdesk System) จึงเป็นระบบหนึ่งที่หลายองค์กรพัฒนาขึ้น เมื่อระบบสารสนเทศถูกนำมาใช้งานมากขึ้น สิ่งที่

ปรากฏขึ้นหลังจากนั้นได้แก่ ความคาดหวังของผู้ใช้งาน ประสิทธิภาพที่ดีของระบบและการรักษาความปลอดภัย [1]

งานวิจัยนี้ได้ระบุขอบเขตเฉพาะในส่วนของการส่งต่อการแก้ไขปัญหาการใช้งานให้กับผู้รับผิดชอบ ผู้วิจัยจึงนำเสนอความสำคัญในการทำเทคนิคต่างๆ ของการใช้เหมืองข้อมูลเนื่องด้วยในการเปิดเคสต่างๆ ที่เกิดขึ้นในระบบ ManageEngine ServiceDesk Plus System ของบริษัทแห่งหนึ่งพบว่าเมื่อมีเคสเกิดขึ้น คนที่ทำหน้าที่ในการเก็บข้อมูลและส่งต่องานไม่สามารถส่งต่องานให้กับผู้ที่รับผิดชอบในส่วนงานนั้นๆ ได้อย่างมีประสิทธิภาพ เนื่องจากในปัจจุบันการส่งต่อปัญหาหรือส่งต่อการซ่อมงานในบางบริษัทใช้คนในการส่งต่องานให้กับผู้รับผิดชอบ ทำให้พบว่าในบางครั้งอาจเกิดความผิดพลาดในการส่งต่องานให้กับผู้รับผิดชอบได้ ซึ่งการส่งต่องานที่ผิดพลาดทำให้เวลาในการทำงานและแก้ไขงานนั้นล่าช้า เป็นผลทำให้การแก้ไขงานนั้นใช้เวลามากขึ้น และอาจทำให้ส่งผลกระทบกับข้อตกลงระดับบริการ (Service Level Agreement : SLA) ที่ตกลงกับลูกค้าไว้ เพื่อแก้ปัญหาดังกล่าว ผู้วิจัยจึงได้นำเทคนิควิเคราะห์ข้อมูลทางด้านเหมืองข้อมูลมาประยุกต์ใช้ในการสร้างรูปแบบการตัดสินใจเพื่อแนะนำการส่งต่อผู้รับผิดชอบงานในแต่ละงานให้มีความถูกต้อง ซึ่งในงานวิจัยนี้มีการใช้เทคนิคเหมืองข้อมูล เพื่อหาอัลกอริทึมที่เหมาะสมและนำมาพัฒนารูปแบบการตัดสินใจสำหรับระบบแก้ไขปัญหาการใช้งาน เพื่อวิเคราะห์เหตุการณ์และช่วยในการส่งต่อปัญหาหรือส่งต่อการซ่อมงานให้กับผู้รับผิดชอบได้

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 การบริหารจัดการงานด้านเทคโนโลยีสารสนเทศ (IT service Management)

การบริหารจัดการงานด้านเทคโนโลยีสารสนเทศ คือ การใช้เทคโนโลยีสารสนเทศเพื่อสนับสนุนความต้องการและเป้าหมายทางธุรกิจขององค์กร จะเห็นได้ว่าเทคโนโลยีสารสนเทศจะเป็นในลักษณะของการตอบสนองความต้องการของธุรกิจ ซึ่งจะเน้นเรื่องกระบวนการในการให้บริการแก่ผู้ใช้ระบบสารสนเทศและลูกค้าเป็นหลักเพื่อตอบสนองต่อความต้องการของธุรกิจและให้ความสำคัญกับ “คุณภาพในการให้บริการ” เช่นเรื่อง ข้อตกลงระดับบริการ ในสัญญาการให้บริการเทคโนโลยีสารสนเทศเพื่อให้ผู้ใช้ระบบสารสนเทศและลูกค้า เกิดความพึงพอใจสูงสุดในการรับบริการ [2]

2.2 ระบบการแจ้งปัญหาการใช้งาน (ServiceDesk)

ระบบการแจ้งปัญหาการใช้งาน เป็นหัวใจสำคัญสำหรับการร้องขอบริการทั้งหมดจาก Business users ไปถึงหน่วยงาน IT หรือเรียกได้ว่าเป็นจุดบริการเพียงจุดเดียวที่คอยรับการติดต่อจากผู้ใช้งาน (Single point of contact) และ ช่วยผสานงานระหว่างผู้ใช้งานกับหน่วยบริการต่างๆ เพื่อบันทึกและจัดการวัฏจักรของเหตุการณ์ต่างๆ ด้วยการเน้นย้ำให้ความสำคัญกับเรื่องของความเร็วในการทำให้กลับสู่สภาพปกติของบริการด้วยผลกระทบกับธุรกิจน้อยที่สุด [3]

2.3 การทำเหมืองข้อมูล (Data Mining)

การทำเหมืองข้อมูล คือกระบวนการค้นหาสารสนเทศ หรือความรู้ที่อยู่ในฐานข้อมูลขนาดใหญ่ที่ซับซ้อน เพื่อนำความรู้ที่ได้ไปใช้ประโยชน์ในการตัดสินใจ สารสนเทศที่ได้อาจนำมาสร้างการพยากรณ์ หรือสร้างตัวแบบสำหรับการจำแนกหน่วยหรือกลุ่ม หรือแสดงความสัมพันธ์ระหว่างหน่วยต่างๆ หรือให้ข้อสรุปของสาระในฐานข้อมูล [4] โดยการทำเหมืองข้อมูลสามารถแบ่งออกเป็น 5 ขั้นตอนหลักๆ ได้ดังนี้ 1) กำหนดวัตถุประสงค์ 2) จัดการข้อมูลให้อยู่ในรูปแบบมาตรฐานเดียวกัน 3) ขั้นตอนในการแปลงข้อมูล 4) ขั้นตอนการทำเหมืองข้อมูล 5) ขั้นตอนสำหรับวิเคราะห์ข้อมูล [5]

2.4 การสร้างตัวแบบการพยากรณ์ และการจำแนกข้อมูล (Prediction/Classification)

การสร้างตัวแบบการพยากรณ์ และการจำแนกข้อมูล (Prediction/Classification) เป็นการใช้สารสนเทศที่มีอยู่ในฐานข้อมูล มาสร้างตัวแบบเพื่อคาดการณ์สิ่งที่จะเกิดในอนาคต วิธีการสร้างตัวแบบจะใช้วิธีการวิเคราะห์เชิงสถิติเป็นพื้นฐานและเสริมด้วยวิธีปฏิบัติของการทำเหมืองข้อมูล [4] สำหรับงานวิจัยนี้จะใช้การจำแนกข้อมูลโดยใช้เทคนิคต่างๆ ในการแบ่งข้อมูลเพื่อใช้ในการวัดประสิทธิภาพของโมเดลการจำแนกประเภทข้อมูล ดังนี้ 1) การจำแนกข้อมูลด้วยต้นไม้ตัดสินใจใช้โมเดล Trees.J48, 2) การจำแนกข้อมูลด้วยวิธีการเบย์เซียนโดยใช้โมเดล Bayes.NaiveBayes, 3) การจำแนกข้อมูลด้วยการส่งค่าย้อนกลับเป็นอัลกอริทึมเครือข่ายประสาทเทียม ใช้โมเดล Functions.MultilayerPerceptron, 4) การแยกประเภทแบบมีผู้สอน (SVM) ใช้ kernel function ของอัลกอริทึม SMO

โดยใช้โมเดล Function.SMO และ 5) การเรียนรู้แบบเกียจคร้าน ผู้วิจัยเลือกใช้วิธีเพื่อนบ้านใกล้เคียง โดยใช้โมเดล Lazy.IBk [6]

2.5 งานวิจัยที่เกี่ยวข้อง

การพัฒนารูปแบบการตัดสินใจสำหรับระบบแก้ไขปัญหาการใช้งาน ผู้วิจัยได้ศึกษางานวิจัยต่างๆ โดยมีงานวิจัยที่เกี่ยวข้อง ดังนี้

งานวิจัยของเผด็จและพยุง [7] เรื่องโมเดลบนพื้นฐานต้นไม้ตัดสินใจในการมอบหมายงานแบบอัตโนมัติ สำหรับการให้บริการแก้ไขปัญหาด้านเทคโนโลยีสารสนเทศ ได้ทำการศึกษาการใช้เหมืองข้อมูลเพื่อแก้ไขเหตุการณ์ที่เกิดขึ้นภายใต้การดำเนินการของการให้บริการแก้ไขปัญหาด้านเทคโนโลยีสารสนเทศ เนื่องจากปัญหาของการมอบหมายผู้แก้ไขปัญหาที่ไม่เหมาะสมก่อให้เกิดการมอบหมายซ้ำหลายครั้ง โดยใช้ต้นไม้ตัดสินใจแบบ ID3 เพื่อนำไปสู่การมอบหมายที่เหมาะสมในแต่ละเหตุการณ์แบบอัตโนมัติ

งานวิจัยของกรรณิการ์และมาลีรัตน์ [8] เรื่องการพยากรณ์ราคาประเมินที่ดินโดยใช้เทคนิคเหมืองข้อมูล ได้ทำการศึกษาโดยใช้เทคนิคการจัดหมวดหมู่ (Classification) มาทำการสร้างโมเดลเพื่อให้ได้ราคาประเมินที่ดินที่ใกล้เคียงกับราคาตลาดมากที่สุด โดยเปรียบเทียบจากค่าความแม่นยำของ Generalized Linear, Neural Network และ Decision Tree ซึ่งสามารถนำไปใช้ประโยชน์สำหรับการสร้างระบบเพื่อการประเมินราคาที่ดินคราวละมากแปลง (Mass Appraisal) หรือ ระบบแบบจำลองช่วยประเมินราคาคราวละมากแปลง (CAMA: Computer Assisted Mass Appraisal)

งานวิจัยของวรรณวิภา [9] เรื่องดาต้าแวร์เฮาส์และเหมืองข้อมูลสำหรับการบริหาร ได้ทำการศึกษาแนวคิดและการประยุกต์ใช้ดาต้าแวร์เฮาส์และเหมืองข้อมูลสำหรับการบริหาร โดยสาระสำคัญ คือ การใช้เทคนิคเหมืองข้อมูลเป็นเครื่องมือในการหารูปแบบและความสัมพันธ์ของข้อมูลที่ซ่อนอยู่ ในฐานข้อมูลขนาดใหญ่ ซึ่งเป็นวิธีการสำคัญที่ช่วยให้องค์กรสามารถประเมินกลยุทธ์ของหน่วยงานเพื่อสร้างความได้เปรียบในการแข่งขัน หรือที่เรียกว่าธุรกิจอัจฉริยะ

จากการศึกษางานวิจัยข้างต้นพบว่า การทำเหมืองข้อมูลจัดเป็นเทคนิควิธีการในการหารูปแบบที่ซ่อนอยู่ในข้อมูลปริมาณมากเพื่อให้เกิดความรู้ใหม่ที่เป็นประโยชน์ และนำเอาเทคนิคการจำแนกข้อมูล จากข้อมูลขนาดใหญ่ที่มีอยู่มาช่วยวิเคราะห์เหตุการณ์เพื่อช่วยในการส่งต่อปัญหาหรือส่งต่อการซ่อมงานให้กับผู้รับผิดชอบได้ ในงานวิจัยนี้จะใช้กฎการจัดหมวดหมู่ โดยใช้เทคนิคต่างๆ ซึ่งกล่าวไปแล้วในข้อ 2.4 เพื่อนำมาหาค่าความถูกต้องและช่วยวิเคราะห์เหตุการณ์เพื่อช่วยในการส่งต่อปัญหาหรือส่งต่อการซ่อมงานให้กับผู้รับผิดชอบ ซึ่งการใช้เทคนิคต่างๆ มากขึ้นจะทำให้ข้อเปรียบเทียบในเรื่องของความแม่นยำ ความถูกต้องเพิ่มขึ้นและทำให้การพัฒนาระบบมีประสิทธิภาพมากยิ่งขึ้น

3. วิธีการดำเนินงาน

ในการพัฒนารูปแบบการตัดสินใจสำหรับระบบแก้ไขปัญหาการใช้งาน มีวิธีการดำเนินการวิจัยแบ่งออกเป็นขั้นตอนต่างๆ ดังนี้

3.1 ขั้นตอนการเก็บรวบรวมข้อมูล

ในขั้นตอนนี้ข้อมูลที่ผู้วิจัยเลือกใช้ คือ ข้อมูลจากฐานข้อมูล ManageEngine ServiceDesk Plus System ของบริษัทแห่งหนึ่ง ซึ่งเป็นโปรแกรมสำหรับเก็บรวบรวมข้อมูลปัญหาและคำร้องขอต่างๆ ของลูกค้า

ตารางที่ 1: รายการข้อมูลที่ใช้ในการทดลอง

รายการ	รายละเอียด
แหล่งข้อมูล	ManageEngine ServiceDesk Plus System
จำนวนข้อมูล	9,298 ระเบียน
กลุ่มข้อมูล	การตอบปัญหาและข้อสงสัย, การให้บริการตามใบคำร้อง, การแจ้งปัญหาเนื่องจากการขัดข้องของระบบให้กับผู้ใช้งานทราบ

3.2 ขั้นตอนการเตรียมข้อมูลและคัดเลือกข้อมูล

หลังจากผู้วิจัยได้ทำการเก็บรวบรวมข้อมูลในขั้นตอนก่อนหน้าแล้ว ต้องมีการเตรียมข้อมูลและคัดเลือกข้อมูลที่จะนำมาใช้ เนื่องจากข้อมูลอยู่ในรูปแบบที่ยังไม่สามารถนำไปใช้ได้ จึงต้องมีการเตรียมข้อมูลให้เป็นรูปแบบมาตรฐานเดียวกันและคัดเลือกเฉพาะข้อมูลที่สมบูรณ์และรูปแบบถูกต้องครบถ้วน และลดความหลากหลายของข้อมูล เพื่อให้ได้ผลการทดสอบที่มีประสิทธิภาพสูงสุด

โดยมีการคัดเลือกแอทริบิวต์ที่ต้องการนำมาวิเคราะห์ข้อมูลจาก 37 แอทริบิวต์เหลือเพียง 5 แอทริบิวต์ เนื่องจากแอทริบิวต์อื่นๆ ไม่สามารถนำมาวิเคราะห์การหาผู้รับผิดชอบงานได้ ซึ่งมีแอทริบิวต์ที่เลือก ดังต่อไปนี้ หมวดหมู่ (Category), ประเภทย่อย (Subcategory), ประเภทของเคส (Request_Type), เวลาที่เกิด (Created_Time) และรับผิดชอบ (Technical_Group) เป็นคลาสคำตอบ จากนั้นทำการกลั่นกรองข้อมูลโดยนำข้อมูลที่ไม่แน่นอนออกไป และทำการจัดกลุ่มข้อมูล Created_Time เนื่องจากเป็นค่าต่อเนื่องกัน ซึ่งไม่สามารถนำมาวิเคราะห์ข้อมูลได้ จึงทำการจัดกลุ่มช่วงเวลาเกิดเคส แบ่งออกเป็น 4 ช่วงเวลา ดังตารางที่ 2 ซึ่งในแต่ละแอทริบิวต์มีค่าข้อมูลที่เป็นไปได้ ดังนี้

ตารางที่ 2: ค่าของข้อมูลที่เป็นไปได้ที่จะเกิดเคสขึ้นแต่ละเคส

ชื่อแอทริบิวต์	ค่าของข้อมูล
Category	Access, Account, Batch Processing, Storage, Movephase, Card, Brand, etc.
Subcategory	Domain, Outgoing, Trouble, Online, Token, Area, Transaction, Email, etc.
Request_Type	Request (การให้บริการตามใบคำร้อง) Incident (การแจ้งปัญหาการขัดข้อง) Inquiry (การตอบปัญหาและข้อสงสัย)
Created_Time	Duration 1 (00:01- 06:00), Duration 2 (06:01-12:00), Duration 3 (12:01-18:00), Duration 4 (18:01- 00:00)
Technical_Group	Changeteam, Application, Customer Services, System Operator, Business Operator, Security Officer, Infrastructure

ตารางที่ 3: ตัวอย่างข้อมูลที่เกิดขึ้นแต่ละเคส

Category	Sub Category	Request_Type	…	T_G
Account	Email	Request	…	CS
Merchant	Activate	Request	…	BO
Storage	Tape	Incident	…	Infra
…	…	…	…	…

3.3 การทดสอบข้อมูลและเลือกเทคนิคที่เหมาะสม

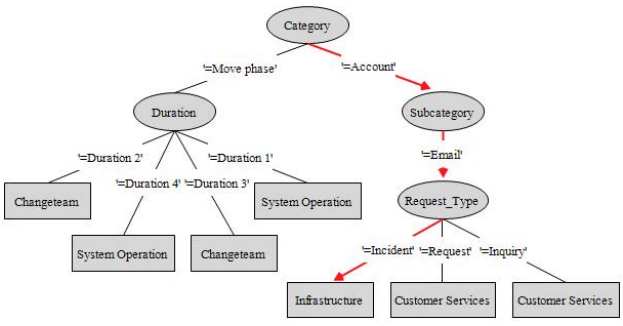
ในขั้นตอนนี้เป็นการนำข้อมูลที่ได้รับการคัดเลือกแล้วจากขั้นตอนที่ผ่านมา แล้วนำข้อมูลดังกล่าว มาทดสอบกับโปรแกรม WEKA ซึ่งต้องเลือกเทคนิคการจำแนกกลุ่มข้อมูลที่เหมาะสม โดยแต่ละเทคนิคจะให้ผลลัพธ์ออกมาในรูปแบบที่แตกต่างกัน ซึ่งในการศึกษาวิจัยครั้งนี้ต้องการนำโมเดลที่เหมาะสม มาใช้ในการวิเคราะห์การส่งต่อเคสที่เกิดขึ้นในแต่ละวัน และนำไปพัฒนารูปแบบการตัดสินใจสำหรับระบบแก้ไขปัญหาการใช้งาน ซึ่งได้มีการทดสอบจำนวน 5 เทคนิค ดังกล่าวไปแล้วในข้อ 2.4 จากนั้นกำหนดตัวชี้วัดโดยใช้วิธีการ Cross-validation เพื่อที่จะประเมินความถูกต้องของตัวจำแนกข้อมูลหรือตัวทำนายข้อมูลได้อย่างน่าเชื่อถือ โดยกำหนดค่า k = 10 ซึ่งจะเรียกว่า 10-fold cross-validation [10]

ผลจากการวิเคราะห์ข้อมูลด้วยโปรแกรม WEKA โดยใช้เทคนิคดังกล่าวข้างต้นสามารถสรุปได้ดังตารางที่ 4 [11] โดยมีการวัดผล ดังนี้ 1) Accuracy เป็นการวัดความถูกต้องของโมเดล โดยพิจารณารวมทุกคลาส 2) Precision เป็นการวัดความแม่นยำของโมเดล โดยพิจารณาแยกทีละคลาส 3) Recall เป็นการวัดความถูกต้องของโมเดล โดยพิจารณาแยกทีละคลาส 4) F-measure เป็นการวัดค่า Precision และ Recall พร้อมกันของโมเดล โดยพิจารณาแยกทีละคลาส และ 5) ROC เป็นค่าแสดงความสัมพันธ์ของข้อมูลที่ทำนายถูกและทำนายผิด [10]

ตารางที่ 4: สรุปผลการวิเคราะห์ข้อมูลด้วยโปรแกรม WEKA

Classifier model	10-fold cross-validation				
	Accuracy	Precision	Recall	F-Measure	ROC
Trees.J48	**99.4300**	**0.994**	**0.994**	**0.994**	**0.999**
NaiveBayes	84.3837	0.856	0.844	0.844	0.981
Function.MLP	94.6333	0.959	0.946	0.948	0.964
Function.SMO	97.9135	0.979	0.979	0.979	0.989
Lazy.IBk	98.1609	0.982	0.982	0.981	0.999

จากตารางแสดงให้เห็นว่าประสิทธิภาพของโมเดล Trees.J48 มีค่าความถูกต้องสูงที่สุดเมื่อเทียบกับโมเดลอื่นๆ ได้จำนวนกฎ 1820 กฎ เมื่อกำหนดค่าความเชื่อมั่นที่ 0.25

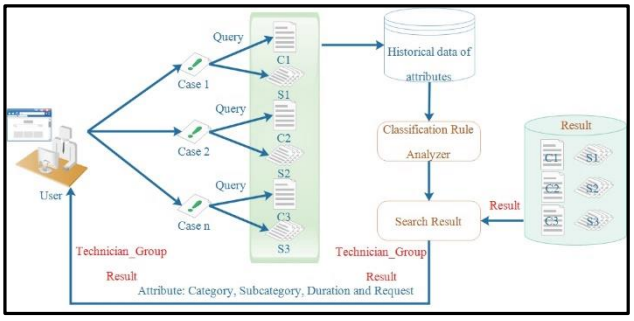


ภาพที่ 2: ตัวอย่างกฎที่ได้จากโปรแกรม WEKA

จากภาพที่ 2 พิจารณาตามทางลูกศร แสดงตัวอย่างกฎหนึ่งที่เกิดขึ้นมีความหมายว่า ถ้า Category เป็น Account และ Sbcategory เป็น Email และ Request_Type เป็น Incident แล้วจะต้องส่งต่องานให้กับ Technical_Group Infrastructure

3.4 พัฒนารูปแบบการตัดสินใจสำหรับระบบการแก้ไขปัญหาการใช้งาน

เป็นการนำกฎที่ได้จากการจำแนกข้อมูลมาเป็นเงื่อนไขในการพัฒนาโปรแกรมให้เป็นเครื่องมือช่วยในการแก้ไขปัญหาการใช้งาน โดยใช้ AngularJS ซึ่งเป็น JavaScript Framework ซึ่งเป็นสูตรสำเร็จของ Library ที่เป็นเครื่องมือในการพัฒนา Web application ในส่วน Front End ทำให้การพัฒนาสะดวกรวดเร็วขึ้น และง่ายต่อการเชื่อมต่อ Restful Service เพราะมีส่วนจัดการ JavaScript Object, การเชื่อมต่อ Data Model ได้สะดวก โดยใช้ MVVM pattern เป็น model ที่ช่วยเหลือในส่วนของการเชื่อมโยงต่อส่วนต่างๆ โดยการสร้าง Controller พร้อม Render data model เข้ากับส่วนแสดงผลหรือ View เข้าด้วยกัน ซึ่งมีขั้นตอนการทำงานของระบบดังภาพที่ 3



ภาพที่ 3: กระบวนการพัฒนาอัลกอริทึมของระบบ

3.5 การทดสอบระบบ

เมื่อพัฒนาระบบแล้วในส่วนของการประเมินผลการใช้งาน ผู้วิจัยใช้แบบสอบถามในการประเมินการยอมรับเทคโนโลยี โดยแบ่งการประเมินผลด้านระบบออกเป็น 5 ปัจจัย คือ 1) การรับรู้ความง่ายต่อการใช้งาน, 2) การรับรู้ประโยชน์, 3) ความน่าเชื่อถือ, 4) ด้านการจัดองค์ประกอบ และ 5) ด้านความพึงพอใจในการให้บริการรูปแบบการตัดสินใจสำหรับระบบแก้ไขปัญหาการใช้งาน โดยผู้วิจัยได้กำหนดเกณฑ์ในการทดสอบไว้ 5 ระดับ ดังตารางที่ 5

ตารางที่ 5: เกณฑ์การประเมินผลการยอมรับเทคโนโลยี

ระดับการยอมรับ	ระดับเกณฑ์การให้คะแนน	
	เชิงปริมาณ	เชิงคุณภาพ
การยอมรับในระดับน้อยมาก	1.00 – 1.50	น้อยมาก
การยอมรับในระดับน้อย	1.51 – 2.50	น้อย
การยอมรับในระดับปานกลาง	2.51 – 3.50	ปานกลาง
การยอมรับในระดับดี	3.51 – 4.50	ดี
การยอมรับในระดับดีมาก	4.51 – 5.00	ดีมาก

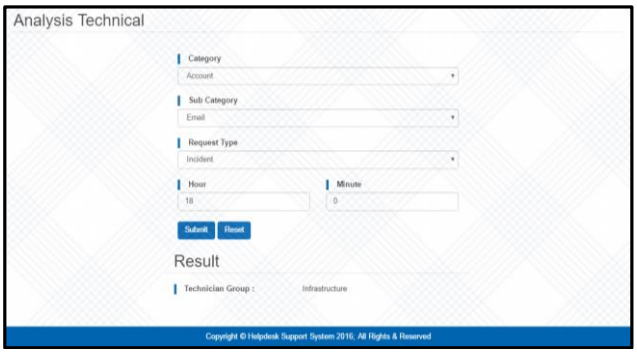
4. ผลการดำเนินงาน

เมื่อพัฒนารูปแบบการตัดสินใจสำหรับระบบแก้ไขปัญหาการใช้งานแล้วผู้ใช้ระบบสามารถทำการส่งต่อปัญหาหรือส่งต่อการซ่อมงานให้กับผู้รับผิดชอบงานได้อย่างถูกต้อง ซึ่งในการพัฒนาระบบผู้วิจัยได้แบ่งการวิเคราะห์ออกเป็น 2 แบบ ดังภาพที่ 4



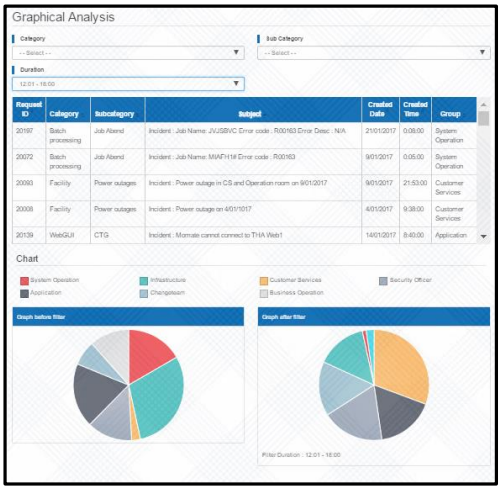
ภาพที่ 4: หน้าการเลือกการวิเคราะห์

แบบที่ 1 การส่งต่องานให้กับผู้รับผิดชอบงาน (Analysis Technical) เช่น เลือกหมวดหมู่ (Category) เป็นAccount และเลือกประเภทย่อย (Subcategory) เป็น Email และเลือกประเภทของเคส (Request_Type) เป็น Incident เมื่อเลือกคำสั่ง Submit จะปรากฏผลลัพธ์ดังภาพที่ 5 ซึ่งสอดคล้องกับตัวอย่างของกฎข้อ 3.3 ดังที่กล่าวมาแล้วข้างต้น



ภาพที่ 5: แสดงผลการวิเคราะห์หน้า Analysis Technical

แบบที่ 2 การวิเคราะห์ข้อมูลย้อนหลังโดยใช้กราฟ (Grapical Analysis) เป็นการนำประสิทธิภาพของโมเดล Trees.J48 ที่มีค่าความถูกต้องสูงที่สุดมาวิเคราะห์ข้อมูลจากรายงานประจำเดือน โดยการกรอง 3 ประเภท คือ การกรองโดยหมวดหมู่, การกรองโดยประเภทย่อย และการกรองช่วงเวลา เช่น เลือกการกรองที่ช่วงเวลา 12:01 ถึง 18:00 น. จะพบว่าในรายงานประจำเดือนฉบับนี้มีปัญหาหรือเคสที่เกิดในช่วงเวลาดังกล่าวกับผู้รับผิดชอบงานกลุ่มใดมากที่สุด ดังภาพที่ 6



ภาพที่ 6: แสดงผลการวิเคราะห์หน้า Graphical Analysis

ผลการประเมินการยอมรับเทคโนโลยีทั้ง 5 ปัจจัย โดยผู้เชี่ยวชาญด้านคอมพิวเตอร์ จำนวน 5 คน ผลที่ได้ คือ มีค่าเฉลี่ยเท่ากับ 3.92 (ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.63) และในส่วนของผู้ใช้งานทั่วไป จำนวน 17 คน ผลที่ได้ คือ มีค่าเฉลี่ยเท่ากับ 4.13 (ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.57)

5. สรุปและอภิปรายผล

งานวิจัยนี้ได้นำเสนอการพัฒนารูปแบบการตัดสินใจสำหรับระบบแก้ไขปัญหาการใช้งาน ในการส่งต่อปัญหาหรือส่งต่อการซ่อมงานให้กับทีมผู้รับผิดชอบงาน ระบบนี้ถือเป็นส่วนหนึ่งที่ช่วยให้ผู้ใช้ระบบทำงานง่ายขึ้นและส่งต่องานได้อย่างถูกต้องและมีประสิทธิภาพ รวมถึงสามารถนำข้อมูลที่ได้จากการวิเคราะห์ข้อมูลรายงานประจำเดือนมาพยากรณ์เหตุการณ์ที่จะเกิดขึ้นในอนาคตได้อีกด้วย ในการพัฒนาระบบได้มีการทดลองใช้งานและทดสอบระบบโดยใช้แบบสอบถามการยอมรับเทคโนโลยี 5 ปัจจัย จากกลุ่มตัวอย่าง 2 กลุ่ม ได้แก่ กลุ่มผู้เชี่ยวชาญทางด้านคอมพิวเตอร์จำนวน 5 คน พบว่ามีค่าเฉลี่ยเท่ากับ 3.92 แสดงให้เห็นว่า ผู้เชี่ยวชาญมีการยอมรับการใช้งานในระดับดี และกลุ่มผู้ใช้ทั่วไปจำนวน 17 คน พบว่ามีค่าเฉลี่ยเท่ากับ 4.13 แสดงให้เห็นว่าผู้ใช้งานระบบมีการยอมรับการใช้งานอยู่ในระดับดี

เอกสารอ้างอิง

[1] C. Suksri, P. Saengsiri and P. James, “Association Rule Discovery from Helpdesk Support System,” *The Tenth National Conference on Computing and Information Technology (NCCIT2014)*, pp. 631-632, 2557.

[2] ปริญญา หอมอเนก, *360 IT Management กลยุทธ์สู่การบริหารอีทีให้ได้มาตรฐานโลก*, เออาร์ไอที: กรุงเทพ, 2554

[3] เบญจพร ฉายาลักษณ์, “ปัจจัยที่มีอิทธิพลต่อความสำเร็จในการนำ ITIL มาใช้ในองค์กร: กรณีศึกษา ธนาคารพาณิชย์แห่งหนึ่ง,” *สารนิพนธ์ วิทยาศาสตรมหาบัณฑิต สาขาวิชาการบริหารเทคโนโลยี วิทยาลัยนวัตกรรม มหาวิทยาลัยธรรมศาสตร์*, 2552.

[4] สุชาดา กีระนันท์, “การเพิ่มศักยภาพระบบการจัดการด้วยเทคโนโลยีสารสนเทศ,” *สัมมนาหนึ่งทศวรรษไอทีจุฬา*, กรุงเทพฯ, 2545.

[5] เรวดี ศักดิ์ดุลยธรรม, “การใช้เทคนิคดาต้าไมน์นิ่งในการสร้างฐานความรู้ เพื่อการทำนายสัมฤทธิ์ผลทางการเรียนของนักศึกษาวิทยาลัยราชพฤกษ์,” *รายงานการวิจัย วิทยาลัยราชพฤกษ์*, 2542.

[6] โกเมศ อัมพวัน, "วิธีการหากฎความสัมพันธ์แบบใหม่โดยต้นไม้แสดงรายการความถี่," *วิทยานิพนธ์ (วท.ม.) จุฬาลงกรณ์มหาวิทยาลัย*, 2548

[7] P. Phomasakha Na Sakolnakorn and P. Meesad, “Tree-Based Model for Automatic Assignment of IT Service Desk Outsourcing,” *The Journal of KMUTNB.*, Vol. 19, No. 1, pp. 19 28, Jan – Apr, 2009

[8] กรรณิการ์ นุชชมภู และ มาลีรัตน์ โสดานิล, “การพยากรณ์ราคาประเมินที่ดินโดยใช้เทคนิคเหมืองข้อมูล,” *ประชุมวิชาการระดับชาติด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศ ครั้งที่ 10*, 2557 หน้า 671 – 676.

[9] วรรณวิภา วงศ์วิไลสกุล, “ดาต้าแวร์เฮาส์และดาต้าไมนิงสำหรับการบริหาร,” *วารสารปัญญาภิวัฒน์*, 2554 หน้า 79-92.

[10] เอกสิทธิ์ พัธรวงศ์ศักดา, *การวิเคราะห์ข้อมูลด้วยเทคนิคดาต้า ไมน์นิง เบื้องต้น*, บริษัท เอเชีย ดิจิตอลการพิมพ์ จำกัด: กรุงเทพฯ, 2557 หน้า 50-58.

[11] ศักดิ์ชาย ตั้งวรรณวิทย์ และ พณณา ตั้งวรรณวิทย์, “การเปรียบเทียบประสิทธิภาพโมเดลการจำแนกและการเลือกคุณลักษณะสำคัญสำหรับคุณภาพการขายผลิตภัณฑ์ทางโทรศัพท์,” *การประชุมวิชาการระดับประเทศด้านเทคโนโลยีสารสนเทศ ครั้งที่ 7*, 2558

การแนะนำประเภทเพลงจากการวิเคราะห์ความสัมพันธ์ ระหว่างคลื่นสมองกับเพลงขณะนอนหลับ

The introduction of music by relation analysis between brainwave and music while human sleeping.

วัฒนชัย วัฒนสิงหะ (Wattanachai Wattanasingha)[1] ชยากร คงทัศน์ (Chayakorn kongtad)[2] และทศพร สายยิ้ม (Tosaporn saiyim)[3]

[1,2,3]สาขาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ ศรีราชา มหาวิทยาลัยเกษตรศาสตร์ วิทยาเขตศรีราชา

wattanachai.ku@gmail.com, chayakorn.k@ku.th, tosaporn.sa@ku.th

บทคัดย่อ

งานวิจัยนี้เป็นการทดสอบสมมุติฐานที่ว่า ขณะที่มนุษย์นอนหลับ เพลงมีอิทธิพลต่อคลื่นสมองหรือไม่ และเพลงประเภทใดที่ทำให้คลื่นสมองมีการตอบสนองของความถี่น้อยที่สุด โดยการหาความสัมพันธ์ระหว่างคลื่นสมองกับกลุ่มทดสอบในช่วงวัยรุ่น ประเภทเพลงที่ใช้ในการทดสอบประกอบด้วย การทดสอบแบบปราศจากเสียงเพลง เพลงประเภทจังหวะเร็ว เพลงประเภทคลาสสิก และเพลงประเภท New Age โดยผ่านการนำค่าความถี่ที่ได้จากการใส่เครื่องตรวจวัดคลื่นสมอง (Electroencephalograph : EEG) มาวิเคราะห์ความสัมพันธ์ ซึ่งกระบวนการวิเคราะห์แบ่งเป็น 2 ขั้นตอน 1) การนำค่าความถี่แสดงผลในรูปแบบกราฟข้อมูลเพื่อหาความสัมพันธ์ระหว่างคลื่นสมองกับเพลง 2) การวิเคราะห์เส้นแบ่งคลื่นวิกฤต (Critical Brainwave : CB) เพื่อบ่งบอกถึงการตอบสนองของคลื่นสมองกับเพลงที่ทดสอบ โดยการทดสอบกับกลุ่มทดสอบพบว่า ค่าการแนะนำประสิทธิภาพของเพลง (Recommend Data Rate : RDR) เพลงประเภท New Age มีค่า RDR = 92.93 % เพลงประเภทคลาสลิก มีค่า RDR = 72.02 % และเพลงประเภทจังหวะเร็ว มีค่า RDR = 35.15 % จึงบอกได้ว่าเพลงประเภท New Age ส่งผลให้คลื่นสมองมีการตอบสนองของความถี่น้อยที่สุด และเป็นประเภทเพลงที่เหมาะสมในการนอนหลับ

คำสำคัญ: ดนตรีบำบัดความเครียด ประเภทดนตรี คลื่นสมอง การตอบสนองของคลื่นสมอง การนอนหลับปกติ

Abstract

This research is the assumptions that would like to know while human sleep, music is influences of stress or not and what kind of music will make the brainwaves have a frequency response in the minimum. By the relationship between the brainwaves of teenager, type of music used in the test composed of test without music, Fast music ,Classical music and New age music through the frequency from the brainwaves monitor (EEG) to analyze. The process consists of two steps 1) the frequency display data showed by chart to determine the relationship between the brainwave and music 2) analysis to critical brainwave to indicate the response of the brainwave with music testing. The tests showed The introduction of music performance (Recommend Data Rate: RDR) RDR of New Age music = 92.93%, RDR of The classic music = 72.02%, RDR of Fast music = 35.15%. The result showed that New Age music is the minimum frequency response and properly to fall asleep.

Keyword: music therapy, music genre, brainwave, brain performance, general sleep.

1. บทนำ

การหลับเป็นพฤติกรรมตามธรรมชาติของมนุษย์ โดยแสดงลักษณะที่มีความรู้สึกตัวแตกต่างกันไป ซึ่งจะเป็นการยับยั้งกล้ามเนื้อที่อยู่ภายใต้อำนาจจิตใจให้มีการตอบสนองต่อสิ่งเร้าลดลง เนื่องด้วยสังคมปัจจุบันเป็นสังคมเร่งรีบในการดำเนินชีวิต จึงส่งผลให้ร่างกายและสมองต้องทำงานหนัก จึงเกิด

อาการล้าหลังจากการทำงาน เมื่อถึงช่วงเวลาพักผ่อนจะทำให้หลับไม่สนิท เกิดการพลิกตัวขณะหลับ อาการเหล่านี้เกิดมาจากคลื่นสมองมีการแปรปรวนจากภาวะปกติ โดยสาเหตุสำคัญคือ ความเครียด เพราะเมื่อมีอาการเครียดคลื่นสมองจะมีการตอบสนองของความถี่มีค่าความถี่พุ่งสูงขึ้น หากเกิดอาการนี้ติดต่อกันในระยะเวลานาน อาจก่อให้เกิดโรคนอนไม่หลับเรื้อรัง ในทางการแพทย์เพลงได้เข้ามามีบทบาทในการบำบัดผู้ป่วยหรือพัฒนาศักยภาพทั้งด้านร่างกาย จิตใจ อารมณ์ ความคิด ด้วยการฟังดนตรีที่มีจังหวะช้าๆ จะช่วยให้การหายใจช้าลง ลดการตึงของกล้ามเนื้อ และส่งผลให้การตอบสนองของความถี่คลื่นสมองอยู่ในภาวะปกติ

การแนะนำเพลงจากการวิเคราะห์ความสัมพันธ์ระหว่างคลื่นสมองกับเพลงขณะนอนหลับ เป็นการวิเคราะห์รูปแบบของคลื่นสมองกับเพลงประเภทต่างๆ โดยตรวจสอบการตอบสนองของคลื่นสมอง กับการวิเคราะห์ความถี่ของเส้นแบ่งคลื่นวิกฤต (CB) แล้วทำการเปรียบเทียบเพลงแต่ละประเภทที่ส่งผลต่อคลื่นสมอง ซึ่งใช้เกณฑ์ของเส้นแบ่งคลื่นวิกฤต (CB) และค่าการแนะนำประสิทธิภาพของเพลง (RDR) เป็นสิ่งที่บอกถึงประสิทธิภาพของเพลงที่มีอิทธิพลคลื่นสมอง ที่ส่งผลให้ความถี่ของคลื่นสมองลดลงจนอยู่ในภาวะปกติ รวมทั้งตลอดระยะเวลาการนอนจะเป็นช่วงเวลาที่พักผ่อนอย่างแท้จริง

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 กลไกของดนตรีต่อบุคคล

การรับเสียงของร่างกาย เกิดขึ้นเมื่อมีการสั่นสะเทือนของวัตถุสองอย่างที่มีความถี่ใกล้เคียงกัน เรียกว่า การสั่นพ้อง (Resonance) การรับเสียงจากหูที่เซลล์ขน (Organ of Corti) จะเปลี่ยนการสั่นสะเทือนของเสียงเป็นคลื่นกระแสประสาทส่งไปยังสมอง โดยกระทบกับน้ำเหลืองทำให้มีการสั่นสะเทือนและเสียงเพลงที่ได้ยินจะส่งผลต่อการเปลี่ยนแปลงโดยอาศัยคลื่นเป็นตัวกลาง

2.2 สัญญาณคลื่นสมอง

สัญญาณที่ตรวจจับได้จากสมองในด้านเทคโนโลยีการสื่อสารระหว่างสมองและคอมพิวเตอร์ (BCI) เรียกว่า สัญญาณคลื่นสมอง (Brainwave) แบ่งคลื่นความถี่ที่สัมพันธ์กับร่างกายมนุษย์เป็น 4 ประเภท ได้แก่

1) คลื่นเดลต้า (Delta) คือ คลื่นที่มีความถี่ต่ำกว่า 4 รอบต่อวินาที (Hz) เป็นคลื่นสมองตอบสนองช้าที่สุด เกิดในขณะที่ร่างกายผ่อนคลายในระดับสูง หรือนอนหลับ

2) คลื่นเธต้า (Theta) คือ คลื่นที่มีความถี่ 4-8 รอบต่อวินาที (Hz) เป็นคลื่นสมองที่ทำงานช้า เกิดขึ้นในขณะที่ร่างกายกำลังจะหลับ

3) คลื่นอัลฟ่า (Alpha) คือ คลื่นที่มีความถี่ 8-12 รอบต่อวินาที (Hz) เป็นคลื่นสมองที่พบบ่อย เกิดขึ้นในขณะที่มีการคิดสิ่งใดสิ่งหนึ่ง

4) คลื่นเบต้า (Beta) คือ คลื่นที่มีความถี่มากกว่า 12 รอบต่อวินาที (Hz) เป็นคลื่นสมองที่มีความถี่สูง มีการตอบสนองสูงที่สุด เกิดขึ้นในขณะวุ่นวายใจ สับสน หรือฟุ้งซ่าน

คลื่นความถี่ที่นำมาใช้ทดสอบ คือ คลื่นเดลต้า และคลื่นเบต้า ซึ่งคลื่นทั้ง 2 ประเภท มีระยะห่างของความถี่แตกต่างกันชัดเจน บ่งบอกถึงการตอบสนองของสมอง ณ ขณะนั้น

2.3 เครื่องตรวจคลื่นไฟฟ้าสมอง

เครื่องตรวจคลื่นไฟฟ้าสมอง (Electroencephalography : EEG) เป็นเครื่องที่ตรวจวัดคลื่นไฟฟ้าในสมอง บริเวณรอบหนังศีรษะ โดยใช้ขั้วไฟฟ้าเป็นตัววัดการไหลของประจุไฟฟ้าภายในเซลล์ประสาทของสมอง ซึ่งอาศัย Neurosky MindWave Mobile MW003 เป็นเครื่องมือที่ใช้ในการวัด โดยส่งค่าคลื่นสมองออกมาทาง Bluetooth เพื่อนำมาประมวลผล

2.4 ประเภทเพลงที่เกี่ยวข้อง

1) เพลงจังหวะเร็ว เป็นดนตรีที่ทำให้ผู้ฟังมีพลังด้านร่างกาย เพิ่มความรู้สึก กระตุ้นอารมณ์ จังหวะเร็ว

2) เพลงคลาสสิก เป็นดนตรีมีรูปแบบชัดเจน สมดุล มีเสียงดนตรีตลอด จังหวะปานกลาง

3) เพลง New Age เป็นดนตรีเพื่อการส่งเสริมสุขภาพ ช่วยลดความเครียด ช่วยให้เกิดสมาธิ จังหวะช้าและต่อเนื่อง

ประเภทของเพลงที่ใช้ในทดสอบการวิเคราะห์หาความสัมพันธ์ระหว่างคลื่นสมองกับเพลงขณะนอนหลับ มีรายละเอียดแสดงผลดังตารางที่ 1

ตารางที่ 1 : ตารางแสดงรายละเอียดเพลงที่ใช้ทดสอบ 3 ประเภท

ประเภทเพลง	ความยาว (นาที)	Tempo (BPM)	เครื่องดนตรี
จังหวะเร็ว	75.27	134.98	เปียโน
คลาสลิก	56.32	107.42	เครื่องสาย, เครื่องลมไม้, เครื่องลมทองเหลือง
New Age	65.16	103.48	เครื่องดนตรีอิเล็กทรอนิกส์, เปียโน

3. วิธีการดำเนินการดำเนินการวิจัย

การดำเนินการวิจัยได้ใช้กลุ่มทดสอบกับช่วงวัยรุ่น ระบบสมการสองตัวแปร เส้นแบ่งคลื่นวิกฤต โดยมีข้อมูลกลุ่มทดสอบ คลื่นสมองและเพลงเป็นตัวแปรสำคัญในการวิเคราะห์หาความสัมพันธ์

3.1 ข้อมูลกลุ่มทดสอบ

การทดสอบการแนะนำเพลงจากวิเคราะห์หาความสัมพันธ์ระหว่างคลื่นสมองกับเพลงขณะนอนหลับ ได้ใช้กลุ่มทดสอบในช่วงวัยรุ่น โดยทดสอบทั้งหมด 6 คน แบ่งเป็นผู้ชาย 3 คน ผู้หญิง 3 คน ซึ่งรายละเอียดของสภาพแวดล้อม ระยะเวลา และตัวแปรต่างๆ แสดงผลดังตารางที่ 2

ตารางที่ 2 : ตารางแสดงรายละเอียดกลุ่มทดสอบ

กลุ่มทดสอบ	เพศ	อายุ	ระยะเวลาในการทดสอบ (21.00-22.00น.)	สภาพแวดล้อม	เครื่องกระจายเสียง
คนที่ 1	ชาย	21	40 นาที	มีเสียงรบกวนไม่มากนัก	ลำโพงขนาด 15 นิ้ว (ห่างจากผู้ทดสอบ 6 เมตร)
คนที่ 2	ชาย	23	41 นาที		
คนที่ 3	หญิง	21	42 นาที		
คนที่ 4	หญิง	22	44 นาที		
คนที่ 5	ชาย	21	43 นาที		
คนที่ 6	หญิง	23	42 นาที		

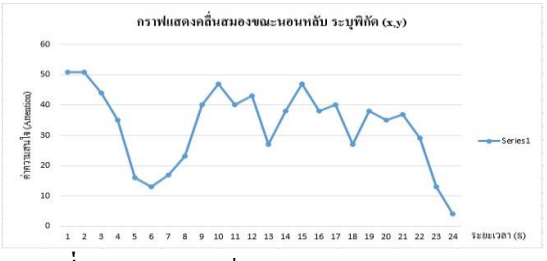
3.2 สมการสองตัวแปร

ระบบสมการสองตัวแปร คือ สมการที่มีตัวแปรสองตัว เลขชี้กำลังของตัวแปรแต่ละตัวเป็นหนึ่ง โดยไม่มีการคูณกันของตัวแปร ซึ่งตัวแปรมีความสัมพันธ์กัน ดังสมการ (1)

$$Ax + By + C = 0 \qquad (1)$$

เมื่อ A,B,C เป็นค่าคงที่ A และ B ไม่เป็นศูนย์พร้อมกัน โดยที่ x และ y เป็นตัวแปร

การแนะนำเพลงจากการวิเคราะห์หาความสัมพันธ์ระหว่างคลื่นสมองกับเพลงขณะนอนหลับ กำหนดให้ตัวแปร y แทนค่าความสนใจ (Attention) ที่ส่งมาจากเครื่องตรวจวัดคลื่นสมอง (EEG) และตัวแปร x แทนระยะเวลาของเพลงที่วัดค่าได้ต่อหน่วยเป็นวินาที ดังภาพที่ 1



ภาพที่ 1 : กราฟแสดงคลื่นสมองขณะนอนหลับ ระบุพิกัด (x,y)

3.3 เส้นแบ่งคลื่นวิกฤต

เส้นแบ่งคลื่นวิกฤต (Critical Brainwave : CB) คือ เส้นที่แบ่งความแตกต่างของคลื่นออกเป็น 2 ประเภท ได้แก่ ประเภทคลื่นที่มีความถี่ต่ำกว่า CB หรือคลื่นที่มีการตอบสนองของคลื่นสมองน้อย และประเภทที่มีความที่สูงกว่า CB หรือคลื่นที่มีการตอบสนองของคลื่นสมองมาก ซึ่งคลื่นนี้จะเป็นสิ่งที่ผู้วิจัยให้ความสนใจในการนำมาวิเคราะห์หาความสัมพันธ์

ค่าเฉลี่ยเลขคณิต (Arithmetic mean) คือ ค่าผลรวมของค่าสังเกตของข้อมูลทั้งหมด หารด้วยจำนวนของข้อมูลทั้งหมดดังสมการที่ (2)

$$\bar{x} = \frac{\sum_i^n x_i}{n} \qquad (2)$$

ค่าของเส้นแบ่งคลื่นวิกฤตได้มาจากการหาค่าเฉลี่ยเลขคณิต (Arithmetic mean) ของกลุ่มทดสอบทั้งหมด 6 คน โดยขณะทดสอบปราศจากเสียงเพลง ซึ่งค่าที่ได้จากการหาค่าเฉลี่ยกลุ่มข้อมูลทดสอบ สามารถระบุเส้นแบ่งจุดวิกฤต (CB) มีค่าความสนใจ (Attention) เท่ากับ 70 จึงได้แทนเส้นแบ่งจุดวิฤตเป็นเส้นสีแดง เพื่อให้การมองเห็นชัดเจน ดังภาพที่ 2



ภาพที่ 2 : กราฟแสดงเส้นแบ่งคลื่นวิกฤต (CB)

4. ผลการดำเนินงาน

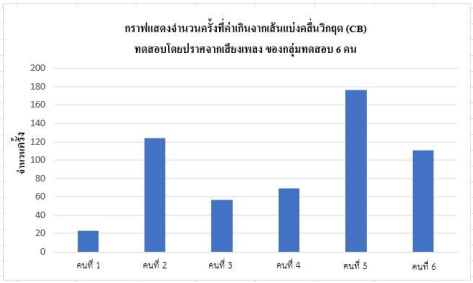
การวิเคราะห์หาความสัมพันธ์ระหว่างคลื่นสมองกับเพลงขณะนอนหลับ ใช้การทดสอบ 4 ประเภท ประกอบด้วย 1) การทดสอบปราศจากเสียงเพลง 2) การทดสอบกับเพลงประเภทจังหวะเร็ว 3) การทดสอบกับเพลงประเภทคลาสสิก 4) การทดสอบกับเพลงประเภท New Age

การเปรียบเทียบความสัมพันธ์ของการทดสอบแบบปราศจากเสียงเพลง และการทดสอบที่ใช้เพลงแต่ละประเภทเป็นตัวแปรในการทดสอบ จะทำให้ทราบถึงการตอบสนองของคลื่นสมองกับการทดสอบแต่ละประเภท หรือบอกความสัมพันธ์ระหว่างคลื่นสมองกับเพลงได้ ซึ่งการวิเคราะห์หาความสัมพันธ์ที่มีเพลงเป็นตัวแปรในการทดสอบ จะใช้เพลง 3 เพลง ในแต่ละประเภทมาใช้ทดสอบเพื่อให้การวิเคราะห์มีความแม่นยำและบอกได้ว่าเพลงประเภทใดที่ส่งผลให้ความถี่คลื่นสมองต่ำกว่าเส้นแบ่งคลื่นวิกฤต (CB) มากที่สุด ซึ่งคำนวณได้จากการหาค่าเฉลี่ยของจำนวนครั้งที่มีค่าเกินจากเส้นแบ่งคลื่นวิกฤต (CB) ของทั้ง 3 เพลงที่ใช้ทดสอบ

สามารถวิเคราะห์ข้อมูลได้จากกราฟแสดงผลคลื่นสมองที่ใช้การทดสอบแบบปราศจากเสียงเพลง และการทดสอบที่ใช้เพลงแต่ละประเภทเป็นตัวแปรในการทดสอบ โดยมีเส้นแบ่งคลื่นวิกฤต (CB) เป็นสิ่งที่บ่งบอกถึงการตอบสนองของคลื่นสมองกับเพลง ซึ่งในการทดลองครั้งนี้ให้ความสนใจช่วงความถี่ที่มีค่าเกินเส้นแบ่งคลื่นวิกฤต (CB)

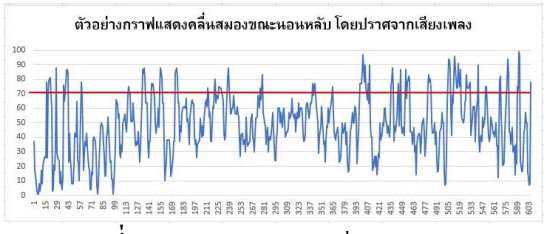
4.1 ข้อมูลการทดลอง

4.1.1 ปราศจากเสียงเพลง ดังภาพที่ 3



ภาพที่ 3 : กราฟแสดงจำนวนครั้งที่ค่าเกินจากเส้นแบ่งคลื่นวิกฤต ทดสอบโดยปราศจากเสียงเพลง ของกลุ่มทดสอบ 6 คน

ดังภาพที่ 3 เป็นการแสดงจำนวนครั้งที่มีค่าเกินจากเส้นแบ่งคลื่นวิกฤต ทดสอบโดยปราศจากเสียงเพลง ของกลุ่มทดสอบ 6 คน จะเห็นได้ว่าคนที่ 1 พบค่าเกินจุด CB 23 ครั้ง คนที่ 2 พบค่าเกินจุด CB 124 ครั้ง คนที่ 3 พบค่าเกินจุด CB 57 ครั้ง คนที่ 4 พบค่าเกินจุด CB 69 ครั้ง คนที่ 5 พบค่าเกินจุด CB 23 ครั้ง และคนที่ 6 พบค่าเกินจุด CB 111 ครั้ง ซึ่งจำนวนครั้งที่ทำการทดสอบโดยปราศจากเสียงเพลง ความถี่มีค่าผันผวนกัน กราฟมีลักษณะขึ้น-ลง สลับกัน หรือเรียกว่าไม่มีความสัมพันธ์ต่อกัน สังเกตได้ดังภาพที่ 4



ภาพที่ 4 : ตัวอย่างกราฟแสดงคลื่นสมองขณะหลับ โดยปราศจากเสียงเพลง

4.1.2 เพลงประเภทจังหวะเร็ว



ภาพที่ 5 : ตัวอย่างกราฟแสดงคลื่นสมอง ทดสอบกับเพลงประเภทจังหวะเร็วขณะนอนหลับ

การทดสอบคลื่นสมองกับเพลงประเภทจังหวะเร็วขณะนอนหลับ ดังภาพที่ 5 ค่าที่เกิดขึ้นมีการเกาะกลุ่มกันมากกว่าการทดสอบแบบปราศจากเสียงเพลง และค่าที่เกิดขึ้นส่วนใหญ่มีค่าความถี่สูงกว่าเส้นแบ่งคลื่นวิกฤต (CB) ซึ่งผลลัพธ์ของจำนวนครั้งที่มีค่าเกินเส้นแบ่งคลื่นวิกฤต (CB) เมื่อทดสอบกับกลุ่มทดสอบ แสดงผลดังตารางที่ 3

ตารางที่ 3 : จำนวนครั้งที่เกิดเส้นแบ่งคลื่นวิกฤต โดยใช้เพลงประเภทจังหวะเร็วทดสอบ

กลุ่มทดสอบ	จำนวนครั้งที่เกินเส้นแบ่งคลื่นวิกฤต (CB) ทดสอบโดยเพลงประเภทจังหวะเร็ว
คนที่ 1	293 ครั้ง
คนที่ 2	242 ครั้ง
คนที่ 3	274 ครั้ง
คนที่ 4	269 ครั้ง
คนที่ 5	288 ครั้ง
คนที่ 6	251 ครั้ง

4.1.3 เพลงประเภทคลาสลิก



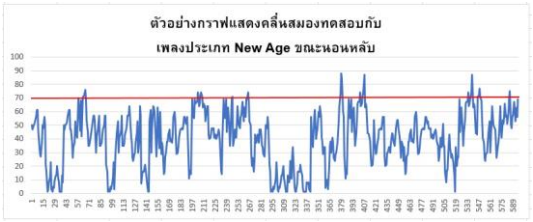
ภาพที่ 6 : ตัวอย่างกราฟแสดงคลื่นสมอง ทดสอบกับเพลงประเภทคลาสลิกขณะนอนหลับ

การทดสอบคลื่นสมองกับเพลงประเภทคลาสลิกขณะนอนหลับ ดังภาพที่ 6 ค่าที่พบมีค่าความถี่ต่ำกว่าเส้นแบ่งคลื่นวิกฤต (CB) บ่อยครั้งขึ้น แต่ยังคงมีความถี่ที่สูงกว่าเส้นแบ่งคลื่นวิกฤต (CB) อยู่บางส่วน ซึ่งผลลัพธ์ของจำนวนครั้งที่มีค่าเกินเส้นแบ่งคลื่นวิกฤต (CB) เมื่อทดสอบกับกลุ่มทดสอบ แสดงผลดังตารางที่ 4

ตารางที่ 4 : จำนวนครั้งที่เกิดเส้นแบ่งคลื่นวิกฤต โดยใช้เพลงประเภทคลาสลิกทดสอบ

กลุ่มทดสอบ	จำนวนครั้งที่เกินเส้นแบ่งคลื่นวิกฤต (CB) ทดสอบโดยเพลงประเภทคลาสลิก
คนที่ 1	114 ครั้ง
คนที่ 2	132 ครั้ง
คนที่ 3	107 ครั้ง
คนที่ 4	120 ครั้ง
คนที่ 5	110 ครั้ง
คนที่ 6	115 ครั้ง

4.1.3 เพลงประเภท New Age



ภาพที่ 7 : ตัวอย่างกราฟแสดงคลื่นสมอง ทดสอบกับเพลงประเภท New Age ขณะนอนหลับ

การทดสอบคลื่นสมองกับเพลงประเภท New Age ขณะนอนหลับ ดังภาพที่ 7 ค่าที่เกิดขึ้นส่วนใหญ่มีค่าเกาะกลุ่มกันและมีความถี่ต่ำกว่าเส้นแบ่งคลื่นวิกฤต (CB) ซึ่งผลลัพธ์ของจำนวนครั้งที่มีค่าเกินเส้นแบ่งคลื่นวิกฤต (CB) เมื่อทดสอบกับกลุ่มทดสอบ แสดงผลดังตารางที่ 5

ตารางที่ 5 : จำนวนครั้งที่เกิดเส้นแบ่งคลื่นวิกฤต โดยใช้เพลงประเภท New Age ทดสอบ

กลุ่มทดสอบ	จำนวนครั้งที่เกินเส้นแบ่งคลื่นวิกฤต (CB) ทดสอบโดยเพลงประเภท New Age
คนที่ 1	29 ครั้ง
คนที่ 2	34 ครั้ง
คนที่ 3	31 ครั้ง
คนที่ 4	25 ครั้ง
คนที่ 5	27 ครั้ง
คนที่ 6	33 ครั้ง

4.2 ผลการทดลอง

จากกราฟแสดงผลคลื่นสมองที่ทำการทดสอบขณะนอนหลับ กับกลุ่มทดสอบ 6 คน โดยมีการทดสอบ 4 ประเภท ประกอบด้วย 1) ทดสอบโดยปราศจากเสียงเพลง 2) ทดสอบกับเพลงประเภทจังหวะเร็ว 3) ทดสอบกับเพลงประเภทคลาสลิก 4) ทดสอบกับเพลงประเภท New Age ซึ่งการทดสอบแต่ละประเภทจะนำจำนวนครั้งที่เกินเส้นแบ่งคลื่นวิกฤต (CB) เป็นการแนะนำเพลงที่ทำให้การนอนหลับเป็นการนอนหลับอย่างแท้จริง

1) ทดสอบโดยปราศจากเสียงเพลง ลักษณะกราฟมีความผันผวนของความถี่ ไม่สม่ำเสมอ ซึ่งแปลได้ว่าไม่มีความถี่ไม่มีความสัมพันธ์ต่อกัน เพราะปราศจากตัวแปรควบคุม

2) เพลงประเภทจังหวะเร็ว ลักษณะกราฟมีค่าความถี่พุ่งสูงขึ้น และค่าจากการทดสอบส่วนใหญ่มีค่าเกินเส้นแบ่งคลื่นวิกฤต (CB) โดยการทดสอบกับกลุ่มทดสอบ พบว่าจำนวนครั้งที่เกินเส้นแบ่งคลื่นวิกฤต (CB) คือ 293, 242, 274, 269, 288 และ 251 ครั้ง ตามลำดับคนทดสอบ

3) เพลงประเภทจังหวะคลาสสิก ลักษณะกราฟมีความถี่ที่ต่ำกว่าเพลงประเภทจังหวะเร็ว ส่งผลให้กราฟส่วนใหญ่อยู่ต่ำกว่าเส้นแบ่งคลื่นวิกฤต (CB) แต่มีบางค่าความถี่ที่เกินเส้นแบ่งคลื่นวิกฤต (CB) โดยการทดสอบกับกลุ่มทดสอบ พบว่าจำนวนครั้งที่เกินเส้นแบ่งคลื่นวิกฤต (CB) 114, 132, 107, 120, 110 และ 115 ครั้งตามลำดับคนทดสอบ

4) เพลงประเภท New Age ลักษณะกราฟความถี่มีแนวโน้มลดต่ำลงกว่าเส้นแบ่งคลื่นวิกฤต (CB) มากที่สุด โดยการทดสอบกับกลุ่มทดสอบ พบว่าจำนวนครั้งที่เกินเส้นแบ่งคลื่นวิกฤต (CB) 29, 34, 31, 25, 27 และ 33 ครั้งตามลำดับคนทดสอบ

การทดสอบ 4 ประเภทข้างต้น พบว่าประเภทของการทดสอบ กราฟของคลื่นสมองมีลักษณะไปในทิศทางเดียวกันกับเพลงแต่ละประเภท หรือเรียกว่าคลื่นสมองกับเพลงขณะนอนหลับ มีความสัมพันธ์กัน ซึ่งในการทดสอบนี้เป็นการแนะนำเพลงที่ส่งผลให้คลื่นสมองมีการตอบสนองต่ำ ทำให้เป็นการนอนหลับอย่างแท้จริง โดยการนำจำนวนครั้งความถี่ที่มีค่าเกินค่าเส้นแบ่งคลื่นวิกฤต (CB) ที่เกิดจากการทดสอบกับกลุ่มทดสอบ มาหาค่าเฉลี่ยเลขคณิต และประเภทของเพลงที่มีค่าเฉลี่ยมีค่าต่ำสุด หรือค่าการแนะนำประสิทธิภาพประเภทของเพลง (Recommend Data Rate : RDR) สูงสุด จะเป็นเพลงที่เหมาะสมต่อการแนะนำ สรุปผลดังตารางที่ 6

ตารางที่ 6 : แสดงความแตกต่างของเพลงแต่ละประเภทที่ทดสอบ

ประเภท	จำนวนครั้งที่เกินค่าเส้นแบ่งคลื่นวิกฤต (CB)						Mean	Recommend Data Rate
	คนที่ 1	คนที่ 2	คนที่ 3	คนที่ 4	คนที่ 5	คนที่ 6		
จังหวะเร็ว	293	242	274	269	288	251	269.5	35.15 %
คลาสสิก	114	132	107	120	110	115	116.3	72.02 %
New Age	29	34	31	25	27	33	29.8	92.83 %

5. สรุป

จากการวิเคราะห์หาความสัมพันธ์ กับกลุ่มทดสอบ โดยมีการทดสอบ 4 ประเภท พบว่าคลื่นสมองกับเพลงมีความสัมพันธ์กัน และเพลงมีอิทธิพลต่อคลื่นสมองขณะนอนหลับ

1) เพลงประเภทจังหวะเร็วมีจำนวนครั้งที่เกินค่าเส้นแบ่งคลื่นวิกฤต (CB) สูงสุด เฉลี่ย 295.5 ครั้ง และมีค่า RDR = 35.15 % ซึ่งอยู่ในระดับต่ำสุดในการแนะนำ 2) เพลงประเภทจังหวะคลาสสิกมีจำนวนครั้งที่เกินค่าเส้นแบ่งคลื่นวิกฤต (CB) เฉลี่ย 116.3 ครั้ง และมีค่า RDR = 72.02 % ซึ่งอยู่ในระดับปานกลางในการแนะนำ 3) เพลงประเภท New Age เห็นข้อแตกต่างได้อย่างชัดเจน โดยตัวแปร CB, RDR เป็นสิ่งที่บ่งบอกถึงประสิทธิภาพที่ดีที่สุด เนื่องจากจำนวนครั้งที่สูงกว่าเส้นแบ่งคลื่นวิกฤต (CB) เฉลี่ย 29.8 ครั้ง ซึ่งมีค่าที่น้อยที่สุด และมีค่า RDR = 92.83 % สูงที่สุด ทำให้สมองได้รับการพักผ่อน มีสมาธิ การตอบสนองความถี่ของคลื่นสมองลดน้อยลง ส่งผลให้ตลอดระยะเวลาการนอนเป็นช่วงเวลาที่พักผ่อนอย่างแท้จริง

เอกสารอ้างอิง

[1] D. R. Chavan, M. S. Kumbhar and R. R. Chavan, "The human stress recognition of brain, using music therapy," *2016 International Conference on Computation of Power, Energy Information and Commuincation (ICCPEIC)*, Chennai, 2016, pp. 200-203.

[2] P. Phasukkit, N. Mahrozeh, M. Kumngern and S. Tungjitkusolmun, "A simple laboratory test of music therapy for insufficient sleep," *2015 8th Biomedical Engineering International Conference (BMEiCON)*, Pattaya, 2015, pp. 1-5.

[3] Y. Li, "Does alpha brainwave music really work?," *2012 5th International Conference on BioMedical Engineering and Informatics*, Chongqing, 2012, pp. 482-485.

[4] Wei Zhao, Xinxi Wang, and Ye Wang. 2010. “Automated sleep quality measurement using EEG signal: first step towards a domain specific music recommendation system*.” In Proceedings of the 18th ACM international conference on Multimedia (MM '10). ACM*, New York, NY, USA, 1079-1082.

[5] D. Jang, Y. J. Park, S. Shin, J. Lee, S. J. Jang and T. B. Lim, "Research about relation of music preference and brain-wave," *2015 International Conference on Information Networking (ICOIN)*, Cambodia, 2015, pp. 466-467.

[6] A. A. Trevisan and L. Jones, "Brain Music System: Standardized Brain Music therapy," *IET Seminar on Assisted Living 2011*, London, 2011, pp. 1-4.

[7] S. A. M. Aris, N. Sulaiman, N. H. A. Hamid and M. N. Taib, "Initial investigation on alpha asymmetry during listening to therapy music," *2010 6th International Colloquium on Signal Processing & its Applications*, Mallaca City, 2010, pp. 1-4.

[8] F. C. Kao, Y. K. Lin, C. C. Chen and C. H. Huang, "Brainwaves Analysis of Relaxation Emotion," *2014 International Symposium on Computer, Consumer and Control, Taichung*, 2014, pp. 308-310.

การจำแนกรุ่นของรถกระบะจากภาพมุมทแยงหน้าโดยใช้ค่าความถี่ของทิศทางตามค่าเกรเดียนท์

Pickup Truck Model Classification from Front Diagonal View using Histogram of Oriented Gradient

ธรานภ สายหรุ่น (Taranop Sairoon)[1] และเสาวลักษณ์ วรรธนาภา (Saowaluk Watanapa)[2]
สาขาวิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยธรรมศาสตร์
t.sairoon@gmail.com[1], scw@cs.tu.ac.th[2]

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อนำเสนอวิธีการจำแนกรุ่นของรถกระบะจากภาพมุมทแยงหน้า โดยใช้ค่าความถี่ของทิศทางตามค่าเกรเดียนท์ (Histogram of Oriented Gradient: HOG) ในการสกัดคุณลักษณะของภาพ จากนั้นใช้การจำแนกประเภทโดยใช้ข้อมูลข้างเคียงใกล้ที่สุด k ตัว (k-Nearest Neighbor: k-NN) และวิธีซัพพอร์ตเวคเตอร์แมชชีน (Support Vector Machines: SVM) โดยชุดข้อมูลที่ใช้ในการทดลองประกอบด้วย รูปภาพมุมทแยงหน้าจำนวน 120 ภาพ จากรถกระบะทั้งหมด 6 รุ่น ซึ่งผลการทดลองแสดงให้เห็นว่า การจำแนกประเภทโดยใช้ข้อมูลข้างเคียงใกล้ที่สุด k ตัว โดยกำหนดให้ k=3 และการจำแนกประเภทด้วยวิธีซัพพอร์ตเวคเตอร์แมชชีนให้ค่าความถูกต้องในการจำแนกร้อยละ 94.2 และ 99.2 ตามลำดับ

คำสำคัญ: การจำแนกรุ่นรถยนต์ ค่าความถี่ของทิศทางตามค่าเกรเดียนท์ ซัพพอร์ตเวคเตอร์แมชชีน

Abstract

This paper presents an algorithm to classify the model of pickup truck from front diagonal view images. The histogram of oriented gradient (HOG) was used in the feature extraction process. Then, the k-nearest neighbor (k-NN) classification technique and the support vector machine (SVM) with linear kernel was applied to classify the pickup truck model. The Experiment dataset consists of 120 images with 6 models of pickup trucks. From the experimental results, the k-nearest neighbor classifier yields 94.2% accuracy, and the linear kernel support vector machine classifier yields 99.2% accuracy.

Keyword: Vehicle Model Recognition, HOG, SVM.

1. บทนำ

ในปัจจุบัน มีการนำระบบวิเคราะห์ที่เกี่ยวข้องกับการจราจรบนท้องถนน เข้ามาใช้งานอย่างหลากหลาย เช่น ระบบตรวจจับการฝ่าไฟแดง ระบบตรวจจับความเร็ว และระบบรู้จำหมายเลขจากแผ่นป้ายทะเบียน เป็นต้น ซึ่งระบบเหล่านี้สามารถวิเคราะห์และตรวจจับการกระทำผิดจากการจราจรบนถนนได้อย่างมีประสิทธิภาพ แต่ในทางปฏิบัติ การตรวจสอบและบังคับใช้กฎหมายจะต้องอ้างอิงรถที่กระทำผิดจากแผ่นป้ายทะเบียน ซึ่งเป็นข้อจำกัดหนึ่งในการใช้งาน เนื่องจากป้ายทะเบียนรถยนต์สามารถถูกปลอมแปลงได้ง่าย ผู้กระทำผิดอาจทำการสวมแผ่นป้ายทะเบียนปลอมหรือแผ่นป้ายทะเบียนจากรถยนต์ที่เกิดอุบัติเหตุ ดังนั้นการใช้ระบบตรวจจับและรู้จำป้ายทะเบียนรถยนต์เพียงอย่างเดียวอาจยังไม่เพียงพอที่จะตรวจสอบการกระทำผิดและบังคับใช้กฎหมายได้

ในส่วนของการตรวจสอบรถยนต์ที่ต้องสงสัยนั้น เบื้องต้นเจ้าหน้าที่จะตรวจสอบจากลักษณะภายนอก ได้แก่ หมายเลขทะเบียน แผ่นป้ายแสดงการชำระภาษีประจำปี รวมถึงยี่ห้อ รุ่น และสีของตัวรถ ในกรณีที่พบรถยนต์ต้องสงสัย เจ้าหน้าที่อาจทำการตรวจสอบเพิ่มเติมโดยการสืบค้นข้อมูลการจดทะเบียนจากกรมการขนส่งทางบก เพื่อความถูกต้องของลักษณะรถยนต์ดังกล่าว

จากการศึกษาสถิติการโจรกรรมรถยนต์ที่ผ่านมาพบว่า เจ้าหน้าที่ตำรวจสามารถจับกุมผู้กระทำความผิดและนำรถยนต์ของกลางมาส่งคืนได้เป็นจำนวนน้อย เมื่อเทียบกับสถิติการโจรกรรมรถยนต์ในแต่ละปี แม้ว่าในปัจจุบันจะมีการนำระบบตรวจจับและรู้จำป้ายทะเบียนรถยนต์มาใช้ในการติดตามรถยนต์ที่สูญหาย หรือต้องสงสัยแล้ว แต่ยังไม่สามารถ

ใช้งานได้จริง เนื่องมาจากอาชญากรมักทำการปลอมแปลงแผ่นป้ายทะเบียน เพื่อให้ยากต่อการตรวจสอบจึงทำให้การติดตามและจับกุมผู้กระทำความผิดเป็นไปได้ยาก

จากข้อมูลยอดจำหน่ายรถยนต์ประจำปี 2557-2559 พบว่าสัดส่วนยอดขายรถกระบะต่อรถยนต์รวมทุกประเภท คิดเป็นร้อยละ 47.74, 48.60 และ 54.27 [1] ตามลำดับ ดังนั้นงานวิจัยนี้จะนำเสนอระบบจำแนกรุ่นของรถกระบะ เนื่องจากรถกระบะเป็นรถยนต์ที่ได้รับความนิยมสูงในประเทศไทย จึงมีสถิติการถูกโจรกรรมสูงตามไปด้วย ซึ่งจะสามารถนำระบบนี้ไปใช้ควบคู่กับระบบการตรวจจับและรู้จำป้ายทะเบียนรถยนต์ และจะช่วยในการตรวจสอบคัดกรองรถยนต์บนท้องถนนได้อย่างมีประสิทธิภาพมากยิ่งขึ้น

2. งานวิจัยที่เกี่ยวข้อง

จากอดีตจนถึงปัจจุบัน มีงานวิจัยที่เกี่ยวข้องกับการวิเคราะห์ภาพยานพาหนะและการสัญจรทางถนน (Traffic Analysis) มากมาย จากการศึกษาเบื้องต้น ผู้วิจัยจึงขอแบ่งงานวิจัยออกเป็นประเภทต่าง ๆ ดังนี้

1. งานวิจัยเกี่ยวกับการจำแนกประเภทของรถยนต์ (Vehicle Type Classification) [2]
2. งานวิจัยเกี่ยวกับการวิเคราะห์ความหนาแน่นของรถยนต์บนท้องถนน [3]
3. งานวิจัยเกี่ยวกับการตรวจจับความเร็ว [4] และตรวจจับการเปลี่ยนเลน [5]
4. งานวิจัยเกี่ยวกับการตรวจจับผู้สัญจรบนท้องถนน (Pedestrian Detection) [6]
5. งานวิจัยเกี่ยวกับการตรวจจับและรู้จำแผ่นป้ายทะเบียนรถยนต์ (License Plate) [7]
6. งานวิจัยเกี่ยวกับการรู้จำรุ่นและยี่ห้อของรถยนต์ (Make and Model Recognition) [8]

จากการศึกษางานวิจัยประเภทต่าง ๆ ผู้วิจัยมีความสนใจเกี่ยวกับการรู้จำรุ่นและยี่ห้อของรถยนต์ เนื่องจากเห็นว่าเป็นงานวิจัยที่กำลังได้รับความนิยมในปัจจุบัน สามารถนำไปใช้ในการป้องกันปราบปรามอาชญากรรมได้จริง และยังไม่พบงานวิจัยในลักษณะเดียวกันนี้ที่ทำการศึกษากับรถกระบะ

งานวิจัยเกี่ยวกับการรู้จำรุ่นและยี่ห้อของรถยนต์นั้น ในอดีตมักจะใช้การเปรียบเทียบกับภาพโมเดลต้นแบบ ทั้งแบบ 2 มิติ [9] และ 3 มิติ [10] แต่วิธีดังกล่าวยังให้ผลลัพธ์ไม่ดีเท่าที่ควรและยังใช้เวลาในการประมวลผลค่อนข้างมาก ในเวลาต่อมาได้มีผู้นำเสนอวิธีการตรวจจับและอธิบายคุณลักษณะเฉพาะ (Feature Detector and Descriptor) ที่สามารถนำมาใช้ในงานรู้จำได้อย่างมีประสิทธิภาพ ได้แก่ Scale-Invariant Feature Transform (SIFT) [11] , Histogram of Oriented Gradient (HOG) [12] และ Speeded Up Robust Features (SURF) [13] โดยอัลกอริทึมดังกล่าว ได้ถูกนำไปประยุกต์กับงานวิจัยต่าง ๆ มากมาย ซึ่งอัลกอริทึมนี้ช่วยเพิ่มประสิทธิภาพให้แก่ระบบจำแนกและรู้จำ และลดระยะเวลาในการประมวลผลได้อย่างมีนัยสำคัญ

การวิเคราะห์คุณลักษณะเฉพาะของรถยนต์นั้น สามารถทำได้หลายมุมมอง โดยมุมมองภาพจากด้านหน้าได้รับความนิยมมากที่สุด เนื่องจากสามารถสังเกตและแยกแยะรุ่นของรถยนต์ได้ง่ายและชัดเจน [14] อีกทั้งยังมีสัญลักษณ์บอกยี่ห้อของรถยนต์ติดอยู่ที่กึ่งกลางด้านหน้าของตัวรถ ซึ่งนำไปใช้ในการรู้จำยี่ห้อของรถยนต์ได้ [15] นอกจากนั้น ยังสามารถเห็นป้ายทะเบียนจากภาพด้านหน้า โดยนำไปใช้ในการตรวจจับตำแหน่งของรถยนต์ในภาพได้ [16] ในขณะเดียวกันยังมีงานวิจัยอื่นที่ใช้ภาพจากด้านท้ายของตัวรถประกอบกับสัญลักษณ์บอกยี่ห้อ และรูปแบบของไฟท้าย ในการรู้จำรุ่นและยี่ห้อของรถยนต์ [17] และยังสามารถใช้ภาพมุมมองด้านข้างในการรู้จำได้อีกด้วย [18]

ในงานวิจัยนี้ ผู้วิจัยได้ทำการทดลองกับภาพมุมทแยงหน้า เนื่องจากในทางปฏิบัติ กล้องที่ใช้สำหรับการตรวจจับรถยนต์จะถูกติดตั้งอยู่ที่ริมถนน ภาพที่ได้จึงเป็นภาพมุมทแยง ซึ่งจากการศึกษางานวิจัยข้างต้น ยังไม่พบว่ามีงานวิจัยใดที่ทำการทดลองกับภาพมุมทแยง

ในงานวิจัยเกี่ยวกับการรู้จำรุ่นและยี่ห้อของรถยนต์นั้น ตัวจำแนก (Classifier) ที่ได้รับความนิยม มีทั้งการใช้ตัวจำแนกแบบพื้นฐาน วัดระยะทางที่สั้นที่สุด โดยใช้ SURF วิเคราะห์สัญลักษณ์บนตัวรถยนต์ [19] หรือใช้ SIFT ในการวิเคราะห์ภาพจากด้านหน้า [14] หรือใช้การจำแนกประเภทโดยใช้ข้อมูลข้างเคียงใกล้ที่สุด k ตัว (k-NN) กับเทคนิคการแปลงแบบ

Fourier, Wavelet และ Curvelet [20] และใช้ SURF ในการหาตำแหน่งของรถยนต์ในภาพ จากนั้นใช้การสกัดคุณลักษณะแบบ HOG ร่วมกับตัวจำแนกแบบซัพพอร์ทเวกเตอร์แมชชีน โดยทำการทดลองกับภาพด้านหน้าของรถยนต์ [21]

นอกจากนี้ยังมีงานวิจัยที่ใช้โมเดลการเรียนรู้แบบเชิงลึก (Deep Learning) ซึ่งกำลังได้รับความนิยมในระบบรู้จำของงานวิจัยสมัยใหม่ ในการวิเคราะห์ภาพด้านหน้าของรถยนต์ ให้ค่าความถูกต้องร้อยละ 88.20 [22] และสามารถนำเทคนิคของ Principal Component Analysis มาประยุกต์ใช้โดยเพิ่มประสิทธิภาพได้ถึงร้อยละ 90.60 [23]

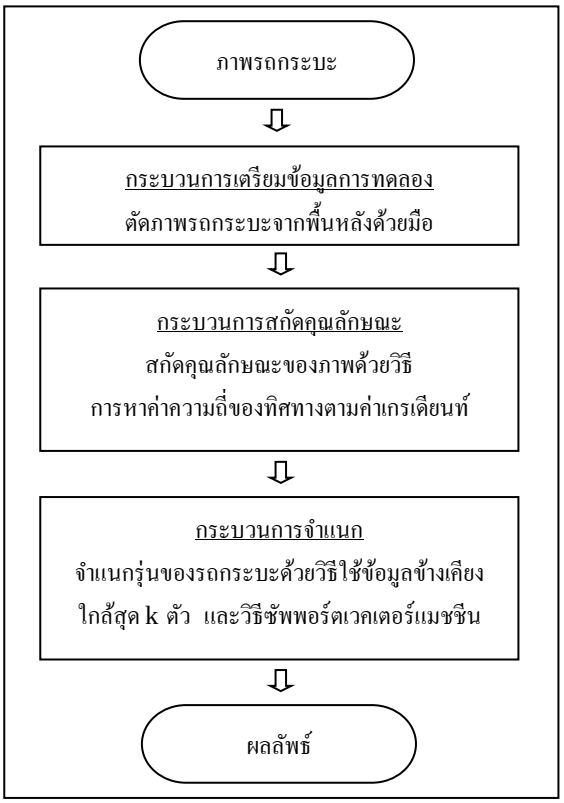
จากการศึกษางานวิจัยที่ผ่านมา พบว่าวิธีการสกัดคุณลักษณะแบบ HOG ได้รับความนิยมจากงานวิจัยเป็นอย่างมาก เนื่องจากการสกัดคุณลักษณะแบบ SIFT และ SURF นั้น จะต้องทำการแปลงจากภาพสีเป็นภาพระดับเทาก่อน ซึ่งทำให้เสียข้อมูลสำคัญของวัตถุไป เนื่องจากความโค้งเว้าของตัวรถอาจขาดหายไปในขั้นตอนการแปลงเป็นภาพระดับเทา ในขณะที่ HOG นั้นทำการวิเคราะห์จากภาพสี ซึ่งสามารถวิเคราะห์รูปร่างได้ชัดเจน และการนำไปใช้กับตัวจำแนกแบบ k-NN และ SVM ให้ผลการจำแนกที่มีประสิทธิภาพ ผู้วิจัยจึงได้ออกแบบการทดลองกับภาพรถกระบะในมุมมองแนวทแยงจากด้านหน้า โดยนำเทคนิคดังกล่าวมาใช้ในการศึกษาและพัฒนาอัลกอริทึม สำหรับงานวิจัยนี้ในบทถัดไป

3. วิธีการดำเนินงาน

วิธีการดำเนินงานวิจัยนั้น สามารถแบ่งออกเป็น 3 ขั้นตอน ดังนี้ การเตรียมข้อมูลการทดลอง (Preprocessing) การสกัดคุณลักษณะ (Feature Extraction) และการจำแนกประเภท (Classification) โดยทำการทดลองกับชุดข้อมูลภาพถ่ายจากกล้องริมถนน ซึ่งเป็นภาพวิดีโอในช่วงกลางวัน ตั้งแต่เวลา 10.00 น. ถึง 13.00 น. จากนั้นทำการคัดเลือกภาพที่เป็นรถกระบะแบบ 2 ประตู จำนวน 120 ภาพ แบ่งเป็นรถกระบะทั้งหมด 6 ประเภท (Class) โดยแสดงขั้นตอนการดำเนินงานดังภาพที่ 1

3.1 การเตรียมข้อมูลการทดลอง

งานวิจัยนี้จะรวบรวมภาพจากกล้องวงจรปิด ซึ่งติดตั้งอยู่ที่ริมถนนมาใช้ในการทดลอง โดยเก็บข้อมูลเฉพาะช่วงเวลากลางวันที่มีแสงสว่างเพียงพอและสภาพอากาศปกติ ตั้งแต่เวลา 10.00 - 13.00 น. จากนั้นนำภาพวิดีโอที่ได้มาทำการแยกเฟรม และเลือกเฉพาะภาพที่เป็นรถกระบะ 2 ประตู ตัดภาพรถกระบะจากพื้นหลังด้วยมือ ซึ่งเป็นภาพรถกระบะเต็มเฟรมที่อาจมีขนาดตั้งแต่ 146×246 พิกเซล ไปจนถึง 191 × 314 พิกเซล และจำแนกประเภทของรถกระบะออกเป็น 6 รุ่น (Class) ดังแสดงในตารางที่ 1 และ 2 โดยแต่ละรุ่นจะมี 20 ภาพ รวมทั้งสิ้นจำนวน 120 ภาพ และลดขนาดของแต่ละภาพให้เหลือ 170 × 270 พิกเซล เพื่อใช้เป็นชุดข้อมูลสำหรับการทดลอง



ภาพที่ 1: ขั้นตอนวิธีการจำแนกรุ่นของรถกระบะ

3.2 ขั้นตอนการสกัดคุณลักษณะ

ในขั้นตอนนี้จะนำชุดภาพที่ได้จากการเตรียมข้อมูลการทดลอง มาทำการสกัดคุณลักษณะโดยอัตโนมัติ ด้วยวิธีการหาค่าความถี่ของทิศทางตามค่าเกรเดียนท์ (Histogram of Oriented Gradient) ดังแสดงในตารางที่ 3 โดยทำการลดทอนขนาดของภาพลง 0.25 เท่า ก่อนนำเข้าสู่กระบวนการสกัดคุณลักษณะ เพื่อเป็นการลดขนาดของเวกเตอร์คุณลักษณะ

(Feature Vector) ให้เหลือ 1,008 คุณลักษณะ (Feature) ซึ่งจะช่วยเพิ่มประสิทธิภาพให้แก่กระบวนการจำแนกในขั้นตอนต่อไป

ตารางที่ 1: รายละเอียดของรถกระบะในแต่ละประเภท

ประเภท	ยี่ห้อ	รุ่น	รุ่นปี
A	อิซูซุ	ดีแมกซ์	2009
B	อิซูซุ	ดีแมกซ์	2012
C	โตโยต้า	ไฮลักซ์ เรโว	2016
D	มิตซูบิชิ	ไทรทัน	2005
E	โตโยต้า	ไฮลักซ์ วีโก้	2008
F	โตโยต้า	ไฮลักซ์ วีโก้ แชมป์	2011

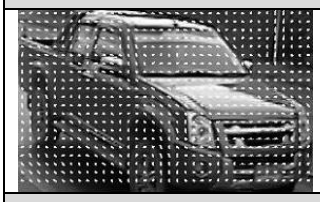
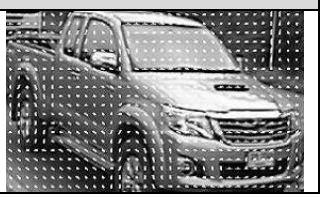
ตารางที่ 2: ภาพของรถกระบะในแต่ละประเภท

ประเภท	ภาพรถกระบะ	ประเภท	ภาพรถกระบะ
A		D	
B		E	
C		F	

3.3 ขั้นตอนการจำแนกประเภท

ในขั้นตอนการจำแนกประเภทของภาพรถกระบะนั้น ผู้วิจัยได้ทำการทดลองโดยใช้ตัวจำแนกประเภทด้วยวิธีใช้ข้อมูลข้างเคียงใกล้สุด k ตัว (k-NN) โดยกำหนดค่า k=3,5,7 จากการทดลองพบว่า ผลการจำแนกที่ k=3 มีความถูกต้องคิดเป็นร้อยละ 94.2 และที่ k=5,7 ได้ผลความถูกต้องเท่ากัน คิดเป็นร้อยละ 93.3 งานวิจัยนี้จึงเลือกใช้ค่า k=3 ซึ่งให้ผลการทดลองที่ดีกว่าจำนวนอื่น และใช้ Euclidean Distance เป็นฟังก์ชันในการวัดระยะห่างระหว่างประเภท (Class) จากนั้นทำการจำแนกประเภทด้วยวิธีซัพพอร์ตเวคเตอร์แมชชีน (Support Vector Machines) โดยใช้เคอร์เนลแบบเชิงเส้น (Linear Kernel) และทำการเปรียบเทียบประสิทธิภาพการจำแนกประเภทของทั้งสองวิธี พร้อมทั้งทำการจับเวลาที่ใช้ในการประมวลผล

ตารางที่ 3: ภาพการสกัดคุณลักษณะด้วยวิธี HOG ในรถกระบะแต่ละประเภท

A	B
	
C	D
	
E	F
	

4. สรุปผลการทดลอง

จากการทดลองพบว่า การจำแนกประเภทด้วยวิธี k-NN ให้ผลลัพธ์ในการจำแนก ร้อยละ 94.2 โดยแสดงผลการทดลองในรูปแบบ Confusion Matrix ในตารางที่ 4 จะพบว่า มิตซูบิชิ ไทรทัน มีความผิดพลาดมากที่สุดอาจเกิดจากรูปแบบของกระจังหน้า หรือการต่อเติมซึ่งแตกต่างไปจากเดิม และการจำแนกประเภทด้วยวิธี SVM มีประสิทธิภาพในการจำแนกสูงถึงร้อยละ 99.2 ดังแสดงผลลัพธ์ในตารางที่ 5 ซึ่งจะพบว่าระบบได้ทำนาย โตโยต้า ไฮลักซ์ วีโก้ แชมป์ ผิดเป็น โตโยต้า ไฮลักซ์ วีโก้ ซึ่งอาจเกิดจากความใกล้เคียงกันทางกายภาพของรถกระบะทั้งสองรุ่น และเมื่อเปรียบเทียบกันทั้งสองวิธีพบว่า SVM ให้ประสิทธิภาพการจำแนกสูงกว่าอย่างมีนัยสำคัญ ดังแสดงในตารางที่ 6 แต่เนื่องจากกระบวนการทำงานของ SVM มีความซับซ้อนกว่า จึงใช้เวลาในการประมวลผลมากกว่า ดังแสดงในตารางที่ 7

ตารางที่ 4: ผลการทดลองการจำแนกรุ่นของรถกระบะด้วยตัวจำแนกแบบ k-Nearest Neighbor , K=3

ประเภทที่ระบบทำนาย

ประเภทที่ถูกต้อง	A	B	C	D	E	F
A	20					
B	1	19				
C			19		1	
D	3			17		
E	1				19	
F					1	19

ตารางที่ 5: ผลการทดลองการจำแนกรุ่นของรถกระบะด้วยตัวจำแนกแบบ SVM

ประเภทที่ระบบทำนาย

ประเภทที่ถูกต้อง	A	B	C	D	E	F
A	20					
B		20				
C			20			
D				20		
E					20	
F					1	19

ตารางที่ 6: ผลการจำแนกรุ่นของรถกระบะด้วยตัวจำแนกแบบ k-Nearest Neighbor และแบบ SVM

ผลการทดลอง	**k-NN (k=3)**	**SVM**
HOG	**94.2 %**	**99.2 %**

ตารางที่ 7: เวลาที่ใช้ในการประมวลผล

การประมวลผล	**เวลารวม (ms)**	**เฉลี่ยต่อภาพ (ms)**
สกัดคุณลักษณะ HOG	524.99	4.37
การจำแนกด้วย k-NN	1,190.80	9.92
การจำแนกด้วย SVM	2,491.70	20.76
เวลารวม k-NN	**1,715.79**	**14.29**
เวลารวม SVM	**3,016.69**	**25.13**

5. ข้อเสนอแนะและแนวทางการพัฒนางานวิจัย

จากการทดลองที่ผ่านมา สรุปได้ว่า การใช้คุณลักษณะแบบ HOG สามารถนำไปใช้ในกระบวนการจำแนกรุ่นของรถกระบะได้อย่างมีประสิทธิภาพ โดยเฉพาะเมื่อใช้ร่วมกับตัวจำแนกแบบ SVM ที่ใช้เคอร์เนลแบบเชิงเส้น

อย่างไรก็ตาม ยังมีข้อผิดพลาดเกิดขึ้น ในการจำแนกรถกระบะบางรุ่นที่มีลักษณะทางกายภาพคล้ายคลึงกันมาก รวมทั้ง ข้อมูลที่ใช้ในการทดลองเป็นภาพที่ได้จากกล้องวงจรปิดในช่วงเวลาที่มีปริมาณความเข้มแสงค่อนข้างมาก แต่ในความเป็นจริง สภาพแสงในบางช่วงเวลาอาจมีการเปลี่ยนแปลงไปตามสภาพแวดล้อมของพื้นที่นั้น ซึ่งในงานวิจัยนี้จะได้ทำการทดลองเพิ่มเติมต่อไป และพัฒนากระบวนการแยกรถกระบะจากพื้นหลังอย่างอัตโนมัติโดยใช้ค่าความแตกต่างของภาพพื้นหลัง (Background Subtraction) รวมถึงการทดลองเพิ่มจำนวนรุ่นและยี่ห้อของรถกระบะ ให้มีความหลากหลายและครอบคลุมมากยิ่งขึ้น เพื่อเพิ่มประสิทธิภาพในการการตรวจจับรถกระบะที่เป็นที่นิยมและพบเจอได้มากบนท้องถนนของประเทศไทย

เอกสารอ้างอิง

[1] บริษัท โตโยต้า มอเตอร์ ประเทศไทย จำกัด. 2557-59. "สรุปยอดจำหน่าย." [ระบบออนไลน์]. แหล่งที่มา http://www.toyota.co.th/sale-volume.php (20 เมษายน 2560).

[2] Zhaojin Zhang, Cunlu Xu and Wei Feng, "Road vehicle detection and classification based on Deep Neural Network," *2016 7th IEEE International Conference on Software Engineering and Service Science (ICSESS)*, Beijing, 2016.

[3] N. V. Hung, L. C. Tran, N. H. Dung, T. M. Hoang and N. T. Dzung, "A traffic monitoring system for a mixed traffic flow via road estimation and analysis," *2016 IEEE Sixth International Conference on Communications and Electronics (ICCE)*, Ha Long, 2016.

[4] J. x. Wang, "Research of vehicle speed detection algorithm in video surveillance," *2016 International Conference on Audio, Language and Image Processing (ICALIP)*, Shanghai, 2016.

[5] I. S. Weon, S. G. Lee and S. C. Moon, "Precise localization of a vehicle within a driving lane by combining the vehicle trajectory with vision information," *2016 16th International Conference on Control, Automation and Systems (ICCAS)*, Gyeongju, 2016.

[6] D. U. Kim, S. H. Park, J. H. Ban, T. M. Lee and Y. Do, "Vision-based autonomous detection of lane and pedestrians," *2016 IEEE International Conference on Signal and Image Processing (ICSIP)*, Beijing, 2016.

[7] M. A. H. Bakhtan, M. Abdullah and A. A. Rahman, "A review on License Plate Recognition system algorithms," *2016 International Conference on Information and Communication Technology (ICICTM)*, Kuala Lumpur, 2016.

[8] Yongguo Ren and Shanzhen Lan, "Vehicle make and model recognition based on convolutional neural networks," *2016 7th IEEE International Conference on Software Engineering and Service Science (ICSESS)*, Beijing, 2016.

[9] J. M. Ferryman, A. D. Worrall, G. D. Sullivan and K. D. Baker, "A generic deformable model for vehicle Recognition," *BMVC '95 Proceedings of the 1995 British conference on Machine vision (Vol. 1)*, Birmingham, UK, 1995.

[10] J. Prokaj and G. Medioni, "3-D model based vehicle recognition," *2009 Workshop on Applications of Computer Vision (WACV)*, Snowbird, UT, 2009.

[11] David G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," *International Journal of Computer Vision Volume 60 Issue 2*, November 2004, Kluwer Academic Publishers Hingham, MA, USA, 2004.

[12] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, San Diego, CA, USA, 2005.

[13] Herbert Bay, Tinne Tuyrelaars and Luc Van Gool, "SURF : Speeded Up Robust Features," Computer Vision and Image Understanding Volume 110 Issue 3, Elsevier Science Inc., NY, USA, 2008.

[14] Y. Shinozuka, R. Miyano, T. Minagawa and H. Saito, "Vehicle make and model recognition by keypoint matching of pseudo frontal view," *2013 IEEE International Conference on Multimedia and Expo Workshops (ICMEW)*, San Jose, CA, 2013.

[15] Mausam Jain and D. Tharun Kumar, "Car make and model recognition," *Mtech. Department of Computer Science and Engineering*, IIT Hyderabad, ODF, Yeddumailaram, 2015.

[16] Huihua Yang, et al., "An Efficient Vehicle Model Recognition Method," *Journal of Software*, Academy Publisher, 2013.

[17] D. F. Llorca, D. Colás, I. G. Daza, I. Parra and M. A. Sotelo, "Vehicle model recognition using geometry and appearance of car emblems from rear view images," *17th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, Qingdao, 2014.

[18] J. Boyle and J. Ferryman, "Vehicle subtype, make and model classification from side profile video," *2015 12th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, Karlsruhe, 2015.

[19] P.Senthilraja and T.Karthikeyan,"Advanced Vehicle Tracking and Model Recognition in Rural Areas using SURF Method," *International Journal of Advanced Research in Computer Science&Technology*, India,2014.

[20] F. M. Kazemi, S. Samadi, H. R. Poorreza and M. R. Akbarzadeh-T, "Vehicle Recognition Based on Fourier, Wavelet and Curvelet Transforms - a Comparative Study," *Information Technology, 2007. ITNG '07. Fourth International Conference on*, Las Vegas, NV, 2007.

[21] Suhan lee, Jeonghwan Gwak and Moongu Jeon,"Vehicle Model Recognition in Video," *International Journal of Signal Processing and Pattern Recognition*, April, 2013.

[22] Yongbin Gao and Hyo Jong Lee,"Deep Learning of Principal Computer for Car Model Recognition," *Int'l Conf. IP, Comp. Vision, and Pattern Recognition*, 2015.

[23] Yongbin Gao and Hyo Jong Lee,"Moving Car Detection and Model Recognition based on Deep Learning," *Advanced Science and Technology Letters*, 2015.

การประเมินประสิทธิภาพของการขจัดหมอกควันในภาพด้วยขั้นตอนวิธีตัวดำเนินการเพิ่มความเข้มภาพแบบฟัซซี่ปรับเกณฑ์สามระดับด้วยวิธีวัดแบบปราศจากภาพอ้างอิง

Evaluation of haze removal of Tuned Tri-threshold Fuzzy Intensification Operators Algorithm with No-Reference Image Quality Assessment

พัฒนพงษ์ โพธิปัสสา (Pattanaphong Pothipassa)[1]

สาขาวิชาเทคโนโลยีสารสนเทศ คณะศิลปศาสตร์และวิทยาศาสตร์ มหาวิทยาลัยราชภัฏศรีสะเกษ

slim4more@hotmail.com

บทคัดย่อ

บทความนี้นำเสนอการประเมินประสิทธิภาพในการขจัดหมอกควันในภาพ (Image Dehaze) ของขั้นตอนวิธีตัวดำเนินการเพิ่มความเข้มของภาพแบบฟัซซี่ปรับเกณฑ์สามระดับ (Tuned Tri-threshold Fuzzy Intensification Operators) โดยในงานวิจัยนี้ ภาพตัวอย่างซึ่งมีสิ่งบดบังแทรกอยู่ภายในจำนวน 10 ภาพ ถูกนำมาทำกระบวนการขจัดหมอกควัน (Dehazing) ด้วยขั้นตอนวิธีตัวดำเนินการเพิ่มความเข้มของภาพแบบฟัซซี่ปรับเกณฑ์สามระดับและขั้นตอนวิธีที่น่าสนใจอีกสองวิธี คือขั้นตอนวิธีการคืนค่าการมองเห็นอย่างเร็ว (Fast Visibility Restoration) และขั้นตอนวิธีลดสีก่อนหน้า (Color Attenuation Prior) โดยใช้ตัวประเมินคุณภาพของภาพแบบไม่ต้องใช้ภาพอ้างอิง (No-Reference Image Quality Assessments) ผลการทดลองพบว่า ขั้นตอนวิธีวิธีตัวดำเนินการเพิ่มความเข้มของภาพแบบฟัซซี่ปรับเกณฑ์สามระดับสามารถขจัดปัจจัยลดทอนในภาพ เช่น หมอก หรือควันได้อย่างมีประสิทธิภาพ โดยได้ภาพที่สามารถให้ข้อสนเทศในภาพแก่ได้มากที่สุด ในขณะที่ใช้เวลาในการประมวลผลน้อยที่สุด นอกจากนี้งานวิจัยนี้ยังได้แสดงให้เห็นว่า Shannon's Entropy เป็นตัวประเมินคุณภาพของภาพที่ผ่านกระบวนการขจัดหมอกควันในภาพมากกว่าตัววัด SNR และ CNR

คำสำคัญ: การประเมินประสิทธิภาพ การขจัดหมอกควัน วิธีประเมินคุณภาพของภาพโดยไม่อาศัยภาพอ้างอิง

Abstract

This paper presents an assessment of the Image Dehaze Algorithm named Tuned Tri-threshold Fuzzy Intensification Operators. In this research, 10 obstructed images were dehazed with Tuned Tri-threshold Fuzzy Intensification Operators and two candidate algorithms (Fast Visibility Restoration and Color Attenuation Prior) and using No-Reference Image Quality Assessments to evaluate the quality of their results. The experimental results show that, the Tuned Tri-threshold Fuzzy Intensification Operators can eliminate image attenuation such as fog or smoke effectively, while output images provided better information than that of the candidate algorithms and required minimum computational time.

Keyword: Performance Evaluation, Dehaze, No-Reference Image Quality Assessment

1. บทนำ

สิ่งบดบังในภาพ เช่น หมอก ควัน หรือสายฝนเป็นปัจจัยสำคัญที่ลดทอนคุณภาพของภาพดิจิตอล และส่งผลต่อประสิทธิภาพในการนำภาพดังกล่าวมาใช้ในทางการประมวลผลภาพ ซึ่งในงานวิจัยชุดก่อนหน้านี้ ผู้วิจัยพบปัญหาการถ่ายภาพจากอากาศยานไร้คนขับ (UAV : Unmanned Ariel Vehicle) ไม่สามารถถ่ายภาพหรือวีดีโอได้ในช่วงเวลาที่มีหมอกปกคลุม หรือมีการเผาเศษขยะในบริเวณดังกล่าว เพื่อแก้ไขข้อจำกัดดังกล่าวของการถ่ายภาพทางอากาศด้วยอากาศ

ยานไร้คนขับ จึงเป็นที่มาของวิจัยเพื่อประเมินประสิทธิภาพของขั้นตอนวิธีขจัดสิ่งบดบังในภาพแบบใหม่ที่ชื่อว่าตัวดำเนินการเพิ่มความเข้มของภาพแบบฟัซซี่ปรับเกณฑ์สามระดับ (Tuned Tri-threshold Fuzzy Intensification Operators) และขั้นตอนวิธีที่น่าสนใจอีก 2 วิธี คือขั้นตอนวิธีคืนค่าการมองเห็นอย่างเร็ว (Fast Visibility Restoration) และขั้นตอนวิธีลดสีก่อนหน้า (Color Attenuation Prior) นอกจากนี้ ในงานวิจัยนี้ยังได้ทำการเปรียบเทียบประสิทธิภาพของตัววัดคุณภาพของภาพแบบไม่อาศัยภาพอ้างอิง (No-Reference Image Quality Assessment) ที่น่าสนใจอีก 3 แบบ เพื่อให้สอดคล้องกับงานถ่ายภาพทางอากาศซึ่งกระทำในสถานการณ์จริงไปพร้อมกัน โดยมีรายละเอียดแสดงในส่วนที่ 2, 3, 4 และ 5 ของบทความวิจัยนี้

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 การประเมินคุณภาพของภาพ

การประเมินคุณภาพของภาพ (Image Quality Assessment) ในทางคอมพิวเตอร์สามารถทำได้ 3 แนวทาง [8] คือการประเมินคุณภาพโดยอาศัยภาพคุณภาพดีมาอ้างอิง (Full-References Image Quality Assessment) การประเมินคุณภาพของภาพโดยอาศัยบางส่วนของภาพคุณภาพดีมาเปรียบเทียบ (Reduced-Reference Image Quality Assessment) [1] และการประเมินคุณภาพของภาพแบบไม่ต้องมีภาพคุณภาพดีมาอ้างอิง (No-Reference Image Quality Assessment) ซึ่งวิธีการแบบอาศัยภาพอ้างอิงที่สำคัญได้แก่ PSNR และ SSIM อย่างไรก็ตาม แม้ว่าทั้งสองวิธีจะได้รับความนิยมในหลายๆงานวิจัย แต่ในทางปฏิบัติหรือในสถานการณ์ที่ไม่สามารถถ่ายภาพดังกล่าวซ้ำได้ ทั้ง PSNR และ SSIM เป็นวิธีประเมินคุณภาพของภาพซึ่งไม่สามารถใช้งานได้จริง

2.2 SNR

SNR หรือ Signal-to-Noise Ratio คือตัววัดคุณภาพแบบไม่ต้องอาศัยภาพคุณภาพดีมาอ้างอิง ซึ่งในงานวิจัย [2] ได้มีการนำเอา SNR มาใช้เพื่อประเมินคุณภาพของภาพ โดยใช้คำนวณจากสมการที่ (1)

$$SNR = 10\log\frac{P_s}{P_n}, \quad (1)$$

โดย SNR คือ ค่า Signal-to-Noise-Ratio ของภาพ

P_s คือ กำลังของสัญญาณในภาพ

P_n คือ กำลังของสิ่งรบกวนหรือ noise ในภาพ [2]

อย่างไรก็ตามจากการทดลองทั้งในงานวิจัยนี้ และงานวิจัยชุดก่อนหน้าของผู้วิจัยพบว่า การใช้ SNR ไม่สามารถสะท้อนคุณภาพของภาพตามที่สายตามนุษย์รับรู้ได้ครอบคลุมทุกกรณี

2.3 CNR

CNR (Contrast-to-Noise Ratio) คือตัววัดคุณภาพของภาพแบบไม่ต้องมีภาพอ้างอิงอีกวิธีการหนึ่ง ซึ่งพัฒนาขึ้นเพื่อใช้แทน PSNR ซึ่งเป็นตัววัดประเภทต้องอาศัยภาพอ้างอิง โดยในงานวิจัย [6] ได้แสดงให้เห็นว่า Contrast-to-Noise-Ratio สามารถใช้ได้กับการวิเคราะห์คุณภาพของภาพในงานสรีระวิทยา อย่างไรก็ตาม แม้ว่า CNR จะง่ายต่อการประยุกต์ใช้ แต่ผลจากการทดลองในงานวิจัยนี้ ทำให้ทราบว่าวิธีการนี้ไม่เหมาะกับการใช้วัดคุณภาพของภาพสีธรรมชาติ (True Color) ที่ผ่านกระบวนการขจัดสิ่งบดบังในภาพ

2.4 Shannon's Entropy

แนวคิดของ Shannon's Entropy ซึ่งใช้ในทางการสื่อสารสามารถนำใช้กับงานวิจัยด้านการประมวลผลภาพได้ โดยวิธีการนี้จะมองว่ารายละเอียดของภาพคือสารสนเทศชนิดหนึ่ง ดังนั้น ภาพที่มีความคมชัดและมีปัจจัยลดทอนน้อยจะมีสารสนเทศที่มากกว่าภาพที่มีปัจจัยลดทอนจำนวนมากแทรกอยู่ สำหรับการคำนวณค่า Shannon's Entropy ของภาพ ในงานวิจัย [7] ได้อธิบายไว้ในสมการที่ (2)

$$H(Y) = -\sum_{j=0}^{2^{\beta}-1} p(a_j)\log(p(a_j)), \quad (2)$$

โดย H คือ ค่า Shannon's Entropy ซึ่งมีค่าเป็น bits/pixels

β คือ ค่า pixel depth ของภาพ

$p(a_j)$ คือ ค่า normalized probability ของแต่ละระดับสีเทาในภาพ

2.5 Tuned Tri-threshold Fuzzy Intensification Operators

ขั้นตอนวิธีตัวดำเนินการเพิ่มความเข้มของภาพแบบฟัซซี่ปรับเกณฑ์สามระดับ (Tuned Tri-threshold Fuzzy Intensification Operators) เป็นวิธีขจัดหมอกควันในภาพซึ่งนำเสนอโดย Zohair Al-Ameen ในงานวิจัย [3] เพื่อจัดการกับ

ภาพที่ถ่ายในสภาพอากาศที่มีฝุ่นควัน จุดเด่นที่น่าสนใจของขั้นตอนวิธีนี้คือเป็นเทคนิคที่ช่วยเผยให้เห็นรายละเอียดของภาพได้ดี โดยที่สีของภาพไม่มีการผิดเพี้ยน

2.6 Fast Visibility Restoration

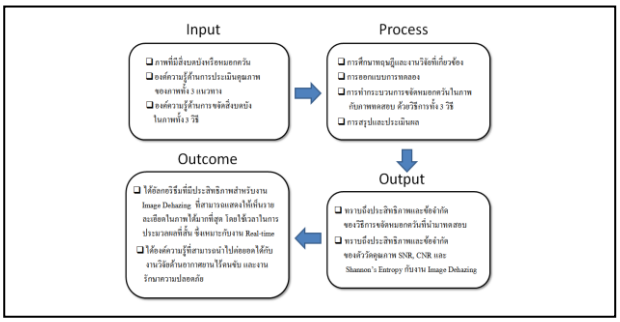
ขั้นตอนวิธีคืนค่าการมองเห็นอย่างเร็ว (Fast Visibility Restoration) เป็นวิธีการขจัดหมอกควันในภาพที่พัฒนาขึ้นโดย Jean-Philippe Tarel และ Nicolas Hautiere [4] จุดเด่นของวิธีการนี้คือผู้ใช้ไม่จำเป็นต้องกำหนดข้อมูลเพิ่มเติมใดๆให้กับกระบวนการขจัดหมอกควันในภาพ ด้วยเหตุที่ขั้นตอนวิธีนี้ทำงานบนพื้นฐานของตัวกรองภาพแบบ Median Filter ดังนั้นจึงใช้เวลาในการประมวลผลค่อนข้างรวดเร็ว เนื่องจากมีการทำงานในแบบ linear function

2.7 Color Attenuation Prior

ขั้นตอนวิธีลดสีก่อนหน้า (Color Attenuation Prior) [5] นำเสนอโดย Qingsong Zhu และคณะฯ ในขั้นตอนวิธีนี้จะทำการลดทอนสีก่อนทำกระบวนการขจัดสิ่งปกคลุมในภาพ โดยการสร้างแบบจำลองเชิงเส้นเพื่อจำลองความลึกของภาพ ภายใต้ prior ใหม่และทำการเรียนรู้พารามิเตอร์ของโมเดลโดยใช้วิธีการเรียนรู้แบบมีผู้สอน ในการนี้ค่า depth information สามารถกู้คืนได้จาก depth map ของภาพที่มีหมอกควันปกคลุม ดังนั้นจึงง่ายต่อการประมาณและฟื้นฟูค่าความกระจายของภาพโดยใช้ atmospheric scattering model

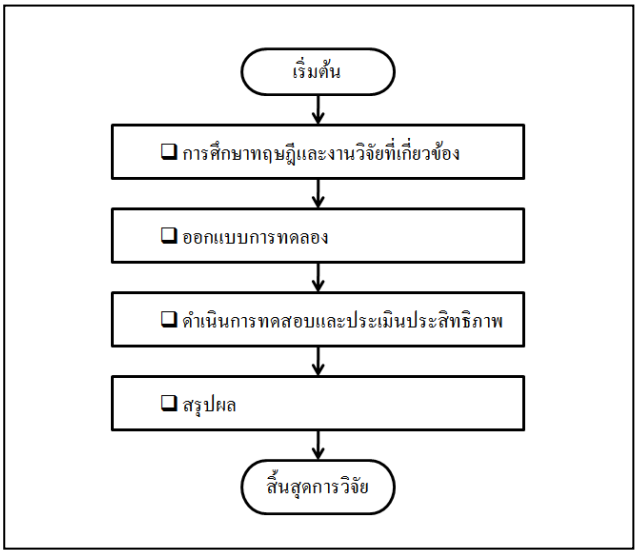
3. วิธีการดำเนินการวิจัย

เพื่อให้ได้ขั้นตอนวิธีขจัดสิ่งบดบังในภาพที่มีประสิทธิภาพ ผู้วิจัยได้ออกแบบกรอบแนวคิดในการวิจัย ดังแสดงในภาพที่ 1



ภาพที่ 1: กรอบแนวคิดในการดำเนินการวิจัย

จากกรอบแนวคิดการทำวิจัยดังกล่าว สามารถแสดงให้อยู่ในรูปของผังงาน เพื่อแจกแจงขั้นตอนในการทำวิจัยได้ ดังแสดงด้วยภาพที่ 2



ภาพที่ 2: ลำดับขั้นตอนในการดำเนินการวิจัย

ทั้งนี้ ขั้นตอนการดำเนินการวิจัยข้างต้นมีรายละเอียดที่สำคัญ ดังนี้

3.1 ศึกษาทฤษฎีและงานวิจัยที่เกี่ยวข้อง

โดยทำการศึกษาอัลกอริธึมขจัดหมอกควันในภาพย้อนหลังตั้งแต่ปี ค.ศ. 2010-2017 จากนั้นเลือกเอาเฉพาะเทคนิคที่น่าสนใจย้อนหลังไม่เกิน 5 ปี จากนั้นทำการศึกษาวิธีการประเมินคุณภาพของภาพทั้งแบบใช้ภาพอ้างอิง (Full-References), แบบใช้ภาพอ้างอิงบางส่วน (Reduced-References) และแบบไม่ใช้ภาพอ้างอิง (No-References)

3.2 การออกแบบการทดลอง

ในขั้นตอนนี้ผู้วิจัยได้ทำการคัดเลือกวิธีการวัดแบบไม่ใช้ภาพอ้างอิงมา 3 วิธี คือ SNR, CNR และ Shannon's Entropy มาใช้ประเมินประสิทธิภาพของอัลกอริธึมขจัดหมอกควันในภาพที่สนใจ โดย SNR และ CNR เป็นตัววัดที่มีแนวคิดมาจาก PSNR ส่วน Shannon's Entropy ผู้วิจัยพบว่าเป็นตัววัดที่ปรากฏในงานวิจัยด้านการประมวลผลภาพหลายฉบับ และยังมีการแสดงวิธีการใช้ในประเทศไทยไม่แพร่หลายนัก

3.3 ขั้นการทดลองและประเมินประสิทธิภาพ

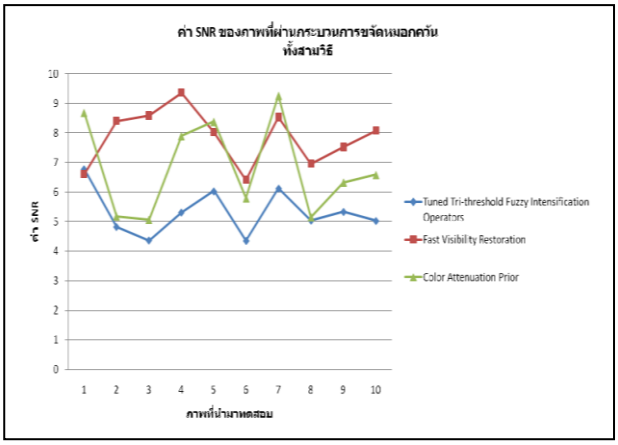
ในขั้นตอนนี้ ภาพที่มีหมอกควันจำนวน 10 ภาพได้ถูกนำมาเข้ากระบวนการขจัดหมอกควันในภาพด้วยอัลกอริธึมทั้ง 3 วิธี จากนั้นดำเนินการประเมินคุณภาพของภาพด้วยตัววัดคุณภาพแบบไม่อาศัยภาพอ้างอิงที่ระบุไว้ในหัวข้อ 3.2

3.4 ขั้นการสรุปผล

ในขั้นตอนนี้ ผลลัพธ์ที่ได้จากกระบวนการขจัดหมอกควันในภาพทั้งหมด จะถูกนำมาพิจารณาในรูปของค่าเฉลี่ย SNR, CNR และ Shannon's Entropy และพิจารณาร่วมกับการยอมรับคุณภาพของภาพโดยมนุษย์

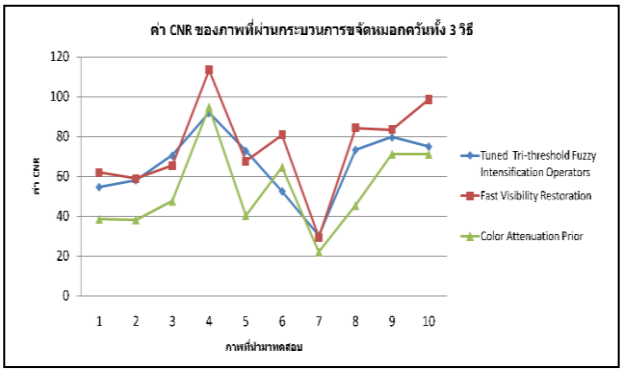
4. ผลการทดลอง

จากการทดลองขจัดหมอกควันในภาพจำนวน 10 ภาพ แบบเดียวกับงานวิจัยเรื่อง Dark Image Prior [2] โดยใช้ขั้นตอนวิธีตัวดำเนินการเพิ่มความเข้มของภาพแบบฟัซซี่ปรับเกณฑ์สามระดับ ขั้นตอนวิธีคืนค่าการมองเห็นอย่างเร็ว และขั้นตอนวิธีลดสีก่อนหน้า ร่วมกับตัวประเมินคุณภาพของภาพแบบไม่อาศัยภาพอ้างอิง ซึ่งประกอบไปด้วย SNR, CNR และ Shannon's Entropy ได้ผลลัพธ์แสดงด้วยภาพที่ 3, 4, 5 และตารางที่ 1 ตามลำดับ



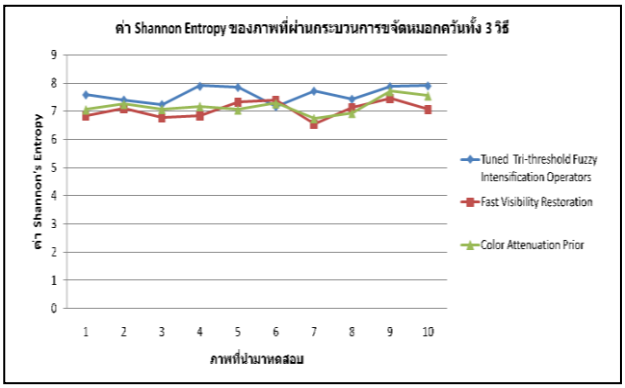
ภาพที่ 3: แสดงค่า SNR ที่ได้จากกระบวนการขจัดหมอกควันในภาพทั้ง 3 วิธี

ผลการทดลองในภาพที่ 3 พบว่า ขั้นตอนวิธีการคืนค่าการมองเห็นอย่างเร็วเป็นวิธีที่ให้ภาพผลลัพธ์ซึ่งมีค่า Signal-to-Noise-Ratio โดยเฉลี่ยสูงที่สุด ตามมาด้วยขั้นตอนวิธีลดสีก่อนหน้าและขั้นตอนวิธีตัวดำเนินการเพิ่มความเข้มของภาพแบบฟัซซี่ปรับเกณฑ์สามระดับ



ภาพที่ 4: แสดงค่า CNR ที่ได้จากกระบวนการขจัดหมอกควันในภาพทั้ง 3 วิธี

ในส่วนของการพิจารณาคุณภาพของภาพด้วยตัววัด CNR นั้น ผลลัพธ์ที่แสดงในภาพที่ 4 ได้แสดงให้เห็นว่าขั้นตอนวิธีการคืนค่าการมองเห็นอย่างเร็วเป็นวิธีการขจัดหมอกควันในภาพที่ให้ค่า Contrast-to-Noise Ratio โดยเฉลี่ยสูงที่สุด ตามมาด้วยขั้นตอนวิธีลดสีก่อนหน้า



ภาพที่ 5: แสดงค่า Shannon's Entropy ที่ได้จากกระบวนการขจัดหมอกควันในภาพทั้ง 3 วิธี

อย่างไรก็ตาม แม้ว่าในการวัดคุณภาพของภาพด้วยตัววัด SNR และ CNR ภาพผลลัพธ์ที่ได้จากขั้นตอนวิธีตัวดำเนินการเพิ่มความเข้มของภาพแบบฟัซซี่ปรับเกณฑ์สามระดับจะให้ค่าโดยเฉลี่ยน้อยที่สุด แต่เมื่อนำมาประเมินในแง่ของการให้สารสนเทศในภาพโดยใช้ Shannon's Entropy พบว่า ขั้นตอน

วิธีตัวดำเนินการเพิ่มความเข้มของภาพแบบฟัซซี่ปรับเกณฑ์สามระดับให้ผลลัพธ์ที่มีค่าเอนโทรปีโดยเฉลี่ยสูงที่สุดและมีความเป็นธรรมชาติมากที่สุด ตามมาด้วยขั้นตอนวิธีลดสีก่อนหน้า ส่วนขั้นตอนวิธีคืนค่าการมองเห็นอย่างเร็วนั้น พบว่าเป็นวิธีการขจัดหมอกควันในภาพที่ให้ค่า Shannon's Entropy โดยเฉลี่ยต่ำที่สุด

นอกจากการประเมินคุณภาพของขั้นตอนวิธีทั้ง 3 วิธีโดยใช้ตัววัดคุณภาพแบบไม่อาศัยภาพอ้างอิงแล้ว ในงานวิจัยนี้ผู้วิจัยยังให้ความสำคัญต่อเวลาที่ใช้ในการประมวลผลของขั้นตอนวิธีทั้ง 3 ด้วย โดยเวลาที่ใช้กับกระบวนการขจัดหมอกควันในภาพทั้งสามวิธีได้แสดงไว้ในตารางที่ 1

ตารางที่ 1: แสดงเวลาที่ใช้ในการประมวลผลของขั้นตอนวิธีขจัดหมอกควันในภาพทั้งสามวิธี (หน่วยเป็นวินาที)

ภาพที่	Tuned Tri-threshold Fuzzy Intensification Operators	Fast Visibility Restoration	Color Attenuation Prior
1	**1.9603**	38.3167	0.0728
2	**1.2985**	29.3217	0.0619
3	**1.1355**	29.9988	0.0611
4	**0.9408**	19.6351	0.0564
5	**3.9111**	212.7322	0.1352
6	**0.7971**	17.0827	0.0508
7	**1.1684**	30.5780	0.0630
8	**2.4223**	68.6955	0.0891
9	**3.8076**	190.7796	0.1281
10	**1.1386**	33.1653	0.0611
Average	**1.8580**	67.0305	0.0780

จากตารางที่ 1 แสดงให้เห็นว่า ขั้นตอนวิธีตัวดำเนินการเพิ่มความเข้มของภาพแบบฟัซซี่ปรับเกณฑ์สามระดับใช้เวลาโดยเฉลี่ยในการขจัดหมอกควันในภาพเพียง 1.8580 วินาที รองลงมาคือขั้นตอนวิธี ลดสีก่อนหน้าและขั้นตอนวิธีการคืนค่าการมองเห็นอย่างเร็วตามลำดับ และเพื่อให้เห็นความแตกต่างของผลลัพธ์แต่ละขั้นตอนวิธีชัดเจนขึ้น ในตอนท้ายของส่วนที่ 4 นี้ แสดงความแตกต่างของทั้ง 3 วิธี ด้วยภาพที่ 6, 7 และ 8



ภาพที่ 6: ผลลัพธ์จากขั้นตอนวิธีตัวดำเนินการเพิ่มความเข้มของภาพแบบฟัซซี่ปรับเกณฑ์สามระดับ

ในภาพที่ 6 จะเห็นได้ว่าขั้นตอนตัวดำเนินการเพิ่มความเข้มของภาพแบบฟัซซี่ปรับเกณฑ์สามระดับสามารถปรับปรุงคุณภาพของภาพถ่ายจากมุมสูงทางอากาศให้มีความชัดเจนขึ้นได้ โดยยังคงไว้ซึ่งความเป็นธรรมชาติของสีในแต่ละส่วนของภาพ

ในขณะเดียวกัน ขั้นตอนวิธีลดสีก่อนหน้าเป็นอีกวิธีหนึ่งที่ทำให้ได้ภาพที่ชัดเจนขึ้น แต่ระดับความสว่างของภาพก็ยังไม่เพียงพอต่อการนำภาพไปใช้งาน หรือทำการประมวลผลต่อด้วยอัลกอริธึมอื่นๆ ดังแสดงด้วยภาพที่ 7



ภาพที่ 7: ผลลัพธ์จากขั้นตอนวิธีลดสีก่อนหน้า

สำหรับขั้นตอนวิธีการคืนค่าการมองเห็นอย่างเร็วนั้น แม้ว่าจะเป็นวิธีขจัดหมอกควันในภาพที่ให้ภาพผลลัพธ์ที่มีค่า SNR และ CNR โดยเฉลี่ยมากที่สุด แต่เมื่อพิจารณาด้วยตัววัด

Shannon's Entropy แล้ว พบว่าเป็นวิธีที่ให้สารสนเทศแก่ผู้ใช้งานภาพน้อยที่สุด เนื่องจากผลลัพธ์ที่ได้มีความผิดเพี้ยนและขาดความเป็นธรรมชาติอย่างเห็นได้ชัด ดังนั้นจึงไม่เหมาะกับการประยุกต์ใช้กับงานปรับปรุงคุณภาพของภาพถ่ายมุมสูงหรือภาพถ่ายทางอากาศ



ภาพที่ 8: ผลลัพธ์จากขั้นตอนวิธีการคืนค่าการมองเห็นอย่างเร็ว

5. สรุปและอภิปรายผล

จากการทดลองพบว่า ขั้นตอนวิธีตัวดำเนินการเพิ่มความเข้มของภาพแบบฟัซซี่ปรับเกณฑ์สามระดับเป็นวิธีการขจัดหมอกควันในภาพที่มีประสิทธิภาพมากที่สุดเมื่อเทียบกับขั้นตอนวิธีการคืนค่าการมองเห็นอย่างเร็วและขั้นตอนวิธีลดสีก่อนหน้า โดยสามารถทำให้ภาพที่ผ่านการปรับปรุงสามารถแสดงรายละเอียดในภาพได้มากที่สุด พิจารณาได้จากค่าแชนนอนเอนโทรปีเฉลี่ยที่สูงที่สุด และเป็นวิธีการที่ใช้เวลาในการประมวลผลโดยเฉลี่ยน้อยที่สุด ดังนั้นจึงเป็นทางเลือกที่น่าสนใจสำหรับการนำไปใช้พัฒนาต่อยอดเป็นอัลกอริธึมปรับปรุงคุณภาพของภาพแบบเวลาจริง นอกจากนี้ผลการทดลองในงานวิจัยนี้ได้แสดงให้เห็นว่าแชนนอนเอนโทรปี เป็นตัววัดคุณภาพของภาพแบบไม่ต้องอาศัยภาพอ้างอิงที่เหมาะกับการวัดคุณภาพของภาพที่ผ่านกระบวนการขจัดหมอกควันในขณะที่ตัววัด SNR และ CNR ไม่สามารถสะท้อนคุณภาพที่แท้จริงของภาพได้ นอกจากนี้ผลการทดลองในงานวิจัยนี้ยังแสดงให้เห็นว่าตัววัดแบบ CNR แม้ว่าจะเป็นที่นิยมใช้กับภาพถ่ายทางการแพทย์ซึ่งเป็นแบบภาพระดับเทา แต่ไม่เหมาะที่จะนำมาใช้ประเมินคุณภาพของภาพสี พร้อมกันนี้ งานวิจัยนี้ยังได้แสดงให้เห็นว่าค่าระดับความแรงของสัญญาณต่อสิ่งรบกวน (SNR) ที่สูงไม่ได้หมายความว่าภาพที่ได้มีคุณภาพดีเสมอไป เนื่องจากระดับของค่า SNR ที่สูงมากๆของภาพ อาจเกิดจากการปรับค่าระดับสีในภาพที่ไม่เหมาะสม จนเป็นสาเหตุทำให้ไม่สามารถแยกความแตกต่างระหว่างส่วนที่เป็นวัตถุออกจากภาพได้

เอกสารอ้างอิง

[1] A. Rehman, and Z. Wang, "Reduced-Reference Image Quality Assessment by Structure Similarity Estimation," *IEEE Trans. on Image Processing*, vol. 21, no. 8, pp. 3378-3389, 2012.

[2] X. Wang, H. He and X. G. Fan, "Signal processing for fast Raman imaging under low SNR condition and its applications in Raman imaging," *Chemmometrics and Intelligent Laboratory Systems*, vol. 163 (2017), pp. 41-48, 2017.

[3] Z. Al-Ameen, "Visibility Enhancement for Images Captured in Dusty Weather via Tuned Tri-threshold Fuzzy Intensification Operators," *I.J. Intelligent Systems and Applications*, pp. 10-17, 2016.

[4] J. P. Tarel and N. Hautiere, "Fast Visibility Restoration from a Single Color or Gray Level Image Images,"

[5] Q. Zhu, J. Mai and L. Shao, "A Fast Single Image Haze Removal Algorithm Using Color Attenuation Prior," *IEEE Trans. on Image Processing*, vol. 24, no. 11, pp. 3522-3533, 2015.

[6] J. Paul, R. W. Bauer, W. Maentele, and T. J. Vogl, "Image fusion in dual energy computed tomography for detection of various anatomic structures – Effect on contrast enhancement, contrast-to-noise ratio, signal-to-noise ration and image quality," *European Journal of Radiology*, vol.80, no. 8, pp. 612-619, 2011.

[7] X. Y. Liu, R. L. Li, H. W. Zhao, T. H. Cheng, G. J. Cui, Q. C. Tan and G. W. Meng, "Quality assessment of speckle patterns for digital image correlation by Shannon entropy," *Optik*, vol.126 (2015), pp. 4206-4211, 2015.

[8] D. Zhang, K. Wnag, Y. Jin, P. Bian and J. Tan, "NR Image Quality Assessment of Satellite Image," *Signal Processing (ICSP), 2016 IEEE 13th International Conference on Signal Processing*, Samrat Ashok Technological Institute, Vidisha, November 7 - 9, 2016.

แอปพลิเคชันวิเคราะห์โรคพืชจากใบบนแอนดรอยด์ โดยใช้ซัพพอร์ตเวกเตอร์แมชชีน
Analysis of Plant Diseases on Leaves Application on Android using Support Vector Machine

ทิพาวรรณ แสวงมี (Thiphawan Sawaengmee)[1] สุพรรษา นาใจตรอง (Supansa Najaitrong)[2]
นิธิศ วังโน (Nitit Wangno)[3] และคำรณ สุนัติ (Khamron Sunat)[4]
ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น
[1]Thiphawan@kkumail.com, [2]Supansa_n@kkumail.com
[3]Nitti@kkumail.com, [4]Khamron_sunat@yahoo.com

บทคัดย่อ

แอปพลิเคชั่นวิเคราะห์โรคพืชจากใบบนแอนดรอยด์ โดยใช้ซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine) เพื่อพัฒนาแอปพลิเคชันระบบแอนดรอยด์ที่สามารถวิเคราะห์ชนิดของโรคพืชจากภาพใบพืชได้ด้วยตนเอง มีความสะดวก รวดเร็วมากยิ่งขึ้น และเพื่อศึกษาการประยุกต์ใช้เทคนิคซัพพอร์ตเวกเตอร์แมชชีนกับการวิเคราะห์โรคพืชบนใบ โดยได้พัฒนาเป็นแอปพลิเคชันบนระบบปฏิบัติการแอนดรอยด์ ร่วมกับโปรแกรม MATLAB 2016 ในขั้นตอนการประมวลผลภาพใช้การแบ่งกลุ่มข้อมูลแบบเคมีนในการแบ่งข้อมูลภาพ เพื่อเพิ่มความแม่นยำในการวิเคราะห์ ในส่วนการวิเคราะห์โรคใช้เทคนิคซัพพอร์ตเวกเตอร์แมชชีนในการจำแนกโรค ได้แก่ โรคใบจุด โรคใบไหม้ และไม่เป็นโรค ภาพที่นำมาวิเคราะห์เป็นภาพดิจิทัล ลักษณะพื้นหลังของภาพใบพืชต้องเป็นสีขาวหรือสีที่ตัดกับสีของโรค พืชที่นำมาวิเคราะห์เป็นพืชเศรษฐกิจของประเทศไทย 4 ชนิด ได้แก่ มะม่วง มันสำปะหลัง พริก และมะนาว จากการวิเคราะห์ความแม่นยำของแอปพลิเคชันเทียบกับผลวิเคราะห์ของผู้เชี่ยวชาญพบว่า ผลการวิเคราะห์ภาพใบพืชที่เป็นโรคใบจุด โรคใบไหม้ และไม่เป็นโรค มีความแม่นยำประมาณ 80.43% 72.31% และ 69.05% ตามลำดับ

คำสำคัญ: โรคพืช ซัพพอร์ตเวกเตอร์แมชชีน การประมวลผลภาพ แอนดรอยด์ การแบ่งกลุ่มข้อมูลแบบเคมีน

Abstract

The Analysis of Plant Diseases on Leaves Application of Vector Machine Support is designed to develop an Application on Android that can analyze the type of disease from plant leaves by themselves are very easy and fast and to study the application of Support Vector Machine techniques The plant diseases from leaves images. This Application is developed on Android operating system together with MATLAB 2016. The image processing process use K-means Clustering in image segmentation for adding the accuracy of the analysis. The diseases analysis process classified diseases that have Leaf Spot, Leaf blight and Healthy Leaf by Support Vector Machine technique. The images are analyzed as digital images that have the background of the leaf must be white or color-contrasted with the color of the disease. The plants were analyzed as four types of economic crops in Thailand: mango, cassava, chilli and lime. The accuracy of the application compared with the expert analysis showed that the Leaf spot, Leaf blight and Healthy Leaf are 80.43% 72.31% and 69.05% respectively.

Keywords: Plant Disease, Support Vector Machine, Processing, Android, K-means Clustering

1. บทนำ

ประเทศไทยมีความเจริญก้าวหน้าด้านการส่งออกผลผลิตทางการเกษตรเป็นอันดับต้นๆของกลุ่มประเทศอาเซียน แต่ในปัจจุบันผลผลิตทางการเกษตรลดลงในขณะที่มีต้นทุนสูงขึ้น สาเหตุหลักเกิดจากปัญหาโรคพืช การเป็นโรคนั้นไม่ได้เกิดเพียงมนุษย์หรือสัตว์เท่านั้น พืชสามารถเป็นโรคได้เช่นกัน หากเกษตรกรระบุชนิดของโรคพืชไม่ได้ จะไม่สามารถแก้ปัญหา

โรคพืชดังกล่าวได้ อย่างไรก็ตามในการสังเกตลักษณะใบพืชบางอาการไม่สามารถแบ่งแยกได้ ต้องให้ผู้เชียวชาญตรวจสอบ ซึ่งต้องใช้เวลาและขั้นตอนมีความซับซ้อน ทางคณะผู้จัดทำจึงพัฒนาแอปพลิเคชั่นวิเคราะห์โรคพืชจากใบบนแอนดรอยด์ โดยใช้เทคนิคซัพพอร์ตเวกเตอร์แมชชีน เพื่อให้เกษตรกรสามารถระบุชนิดของโรคพืชจากใบได้ด้วยตนเองในระยะเวลาอันสั้น ทำให้แก้ไขปัญหาโรคพืชได้โดยไม่ต้องรอผลจากผู้เชี่ยวชาญ มีความสะดวกรวดเร็วมากขึ้น

2. ทฤษฎีที่เกี่ยวข้อง

2.1 โรคพืช

ตามความคิดเห็นของผู้เชี่ยวชาญด้านโรคพืช ให้ความหมายของโรคพืชไว้ว่า เป็นการเปลี่ยนแปลงขบวนการใช้พลังงานในระบบการดำรงชีวิตของ พืชไปจากปกติ ก่อให้เกิด ความเสียหายแก่พืชทำให้มีมูลค่าทางเศรษฐกิจต่ำลง [1], [2]

2.2 อาการของโรคบนใบ

อาการบนใบที่มักพบเสมอ ได้แก่

2.2.1 โรคใบจุด (Leaf Spot)

โรคใบจุดสามารถเกิดขึ้นได้ทุกฤดู โดยเฉพาะฤดูที่มีความชื้นในอากาศสูงลักษณะอาการของโรคใบจุด ส่วนใหญ่จะเกิดขึ้นที่ใบมีลักษณะเป็นแผลจุดเล็ก ๆ รูปร่างไม่แน่นอนหรือค่อนข้างกลมสีน้ำตาลจาง ๆ ตรงกลางขาว มีขอบสีดำหรือม่วงล้อมอยู่โดยรอบ ดังภาพที่ 1

2.2.2 โรคใบไหม้ (Leaf blight)

อาการใบไหม้ แผลแห้งมีขนาดใหญ่ อาจจะแห้งบริเวณเนื้อในหรือจากขอบใบก็ได้ เมื่อเปรียบเทียบกับโรคใบจุดแล้ว แผลที่เกิดจากอาการใบไหม้จะมีขนาดแผลที่ใหญ่กว่า และเป็นบริเวณกว้างกว่า [3] ดังภาพที่ 2



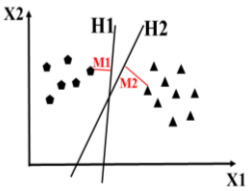
ภาพที่ 1 : อาการใบจุดของมะม่วง



ภาพที่ 2 : อาการใบไหม้ของต้นมะนาว

2.3 ซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machines)

เป็นหนึ่งในวิธีการเรียนรู้แบบมีผู้สอน เมื่อมีข้อมูลฝึกมาให้ และแต่ละข้อมูลถูกจัดอยู่ในประเภทใดประเภทหนึ่งจากสองประเภท ซัพพอร์ตเวกเตอร์แมชชีนจะสร้างแบบจำลองที่สามารถพยากรณ์ได้ว่ากลุ่มนี้จะอยู่ในประเภทใดดังภาพที่ 3



ภาพที่ 3 : SVM

แนวความคิดเกิดจาก การคัดแยกข้อมูลออกเป็น 2 กลุ่ม โดยใช้เส้นแบ่งที่เป็นเส้นตรง จะเห็นว่ามีเส้นตรงจำนวนมากที่สามารถคัดแยกได้ แต่เส้นตรงเส้นไหนที่ดีที่สุด จากภาพที่ 3 นิยาม Margin เป็นผลรวมระยะห่างของเส้นตรงที่แบ่งถึงข้อมูลที่ใกล้ที่สุด (เส้น M1 และ M2) จะเห็นว่า H1 แม้จะสามารถแบ่งข้อมูลทั้งสองกลุ่มออกได้เช่นกัน แต่ระยะในการแบ่งคือเส้น M1 จากเส้นแบ่งไปถึงข้อมูลที่ใกล้ที่สุดนั้นมีขนาดน้อย แต่จากเส้น H2 จะเป็นเส้น M2 ที่แบ่งกลุ่มที่กว้างมากที่สุดของทั้งสองกลุ่มคือให้ค่า Maximum Margin เราเรียกข้อมูลที่อยู่บน Margin นี้ว่า Support Vector

สำหรับรากฐานเดิมของ Support Vector Machine ถูกนำมาใช้กับข้อมูลที่เป็นเชิงเส้น แต่ในความเป็นจริงแล้วข้อมูลที่นำมาใช้ในระบบการสอนให้ระบบเรียนรู้ส่วนใหญ่มักเป็นข้อมูลแบบไม่เป็นเชิงเส้น ซึ่งสามารถแก้ปัญหาดังกล่าวด้วยการนำ Kernal Function มาใช้ เช่น Linear, Polynomial, Radial Basis Function โดยงานวิจัยได้ใช้ Linear Kernel [4]

มีหลายงานวิจัยที่เลือกใช้เทคนิค Support Vector Machine เช่น Automated Plant Disease Analysis (APDA): Performance Comparison of Machine Learning Techniques[5] เป็นงานวิจัยที่นำเทคนิค Feature Extraction ได้แก่ DCT (Discrete cosine transform), DWT (Discrete Wavelet Transform) และ Texture มาใช้ร่วมกับเทคนิค Machine Learning ได้แก่ KNN, Decision Tree, Naïve Bayes, RNN และ SVM คุณสมบัติเหล่านี้ถูกนำมาใช้ในการจำแนกประเภทโรคของใบกุหลาบ ผลการทดลองแสดงให้เห็นว่าเทคนิค DCT+DWT+SVM มี

ประสิทธิภาพ 99.45 % ซึ่งดีกว่าเทคนิคอื่น ๆ สำหรับการจำแนกประเภทของโรค ดังนั้นงานวิจัยนี้จึงเลือกใช้วิธีการจำแนกประเภทโรคด้วยเทคนิค Support Vector Machines

Plant Disease Detection Using Image Processing [6] เป็นงานวิจัยนี้ที่นำเทคนิคการประมวลผลภาพมาใช้ ได้แก่ K-Means Clustering, Otsu Threshold Algorithm และ Color co-occurrence Method ในการแบ่งแยกข้อมูลภาพร่วมกับเทคนิค Machine Learning ในการจำแนกโรค เช่น PCA, SVM และ Neural Network ใช้จำแนกประเภทของโรคใบองุ่น แต่พัฒนาใน MATLAB เท่านั้น ซึ่งไม่สะดวกสำหรับเกษตรกร งานวิจัยนี้จึงนำเอาเทคนิค SVM ใน MATLAB มาใช้ร่วมกับระบบปฏิบัติการแอนดรอยด์ เพื่อให้เกษตรกรหรือบุคคลทั่วไปที่ไม่มีความรู้เรื่องโรคพืชสามารถใช้ได้จริง

2.4 K-means Algorithm [7]

อัลกอลิทึมของ K-meanมีดังนี้

1. กำหนดหรือสุ่มค่าเริ่มต้น จำนวน k ค่า (กลุ่ม) และกำหนดจุดศูนย์กลางเริ่มต้น k จุดหรือที่เรียกว่า Cluster Centers หรือ (Centroid)

2. นำวัตถุทั้งหมดจัดเข้ากลุ่ม โดยทำการหาค่าระยะห่างระหว่างข้อมูลกับจุดศูนย์กลาง หากข้อมูลไหนใกล้ค่าจุดศูนย์กลางตัวไหนที่สุดอยู่กลุ่มนั้น

3. หาค่าเฉลี่ยแต่ละกลุ่ม ให้เป็นค่าจุดศูนย์กลางใหม่

4. ทำซ้ำข้อ 2. จนกระทั่งค่าเฉลี่ยหรือจุดศูนย์กลางในแต่ละกลุ่มจะไม่เปลี่ยนแปลง

2.5 การประมวลผลภาพเชิงดิจิทัล (Digital Image Processing) [8], [9]

เป็นการนำขั้นตอนวิธีการทางคอมพิวเตอร์และการคำนวณทางคณิตศาสตร์มาจัดการกับข้อมูลภาพเชิงดิจิทัล

2.6 การแยกข้อมูลภาพ (Segmentation) [10]

เป็นเทคนิคหรือวิธีการประมวลผลภาพที่มีวัตถุประสงค์ในการแยกวัตถุออกมาจากพื้นหลังของภาพ วิธีการพื้นฐานที่นิยมใช้กันอย่างแพร่หลายมี 2 วิธี คือ การหาขอบภาพ (Edge Detection) และเทคนิคการทำเทรชโฮลด์ (Threshold)

2.7 Otsu Threshold

เป็นการหาค่า Threshold ที่จะนำมาใช้กับทั้งภาพ เพื่อสร้าง Binary image จากภาพสีเทา

2.8 Feature Extraction

เป็นขั้นตอนระบุวัตถุที่อยู่ในภาพ

Color Co-occurrence Method เป็นเมธอดที่ใช้สีและพื้นผิวของภาพมาพิจารณาเพื่อแปลงภาพสีไปเป็น HSI

H : Hue คือ สีบริสุทธิ์ (pure color) เช่น เหลือง ส้ม แดง คำนวณได้จากสมการ[11]

$$H = \begin{cases} theta, & b < g \\ 360 - theta, & b > g \end{cases} \quad (1)$$

S : Saturation คือความบริสุทธิ์ของสีเมื่อผสมกับ แสงสีขาว คำนวณได้จากสมการ

$$S = 1 - \frac{3}{(r+g+b)}[\min(r,g,g)] \quad (2)$$

I : Intensity คือ gray level คำนวณได้จากสมการ

$$I = \frac{1}{3}(r+g+b) \quad (3)$$

หลังจากแปลงภาพเป็น HSI ทำให้ได้เมทริกซ์ภาพสีเทาร่วม จากนั้นนำเมทริกซ์ของ GLCM (Gray Level Co-occurrence Matrices) ที่มีคุณสมบัติในการบอกลักษณะการกระจายเชิงพื้นที่ของค่าโทนสีในภาพ ยกตัวอย่างเช่น หากในเมตริกซ์นั้นมีค่าจำนวนมากอยู่ในช่องแนวทแยงมุมแสดงว่าภาพมีลักษณะหยาบตาม ค่าระหะห่าง d ที่ใช้ในการคำนวณ เป็นต้น มาใช้คำนวณหาค่าสีความคมชัด ดังต่อไปนี้ [10]

$$\text{Contrast} = \sum_{i,j} |i-j|^2 p(i,j)^2 \quad (4)$$

ค่าสีที่ตัดกันวัดความแตกต่างของค่าระดับสีเทา

$$\text{Correlation} = \sum_{i,j} \frac{(i-\mu i)(j-\mu j)p(i,j)}{\sigma_i \sigma_j} \quad (5)$$

ค่าพื้นผิวของภาพ ถ้าค่าติดลบ ภาพนั้นจะมีความหยาบของพื้นผิวมาก

$$\text{Energy} = \sum_{i,j} p(i,j)^2 \quad (6)$$

ผลรวมขององค์ประกอบ (Energy) หรือองค์ประกอบยกกำลังสอง เป็นเปอร์เซ็นต์ของความแปรปรวน

$$\text{Homogeity} = \sum_{i,j} \frac{p(i,j)}{1+|i-j|} \quad (7)$$

ค่ากระจายตัวขององค์ประกอบ (Homogeity) ใช้วัดความใกล้ชิดของการกระจายตัวขององค์ประกอบใน GLCM ในแนวเส้นทแยงมุม หากภาพเป็นเนื้อเดียวกันค่าจะเป็น 1

3.วิธีการดำเนินการวิจัย

3.1 โครงสร้างของแอปพลิเคชัน

ภาพที่ 4 : โครงสร้างของแอปพลิเคชัน

จากภาพที่ 4 นั้น อธิบายได้ดังนี้

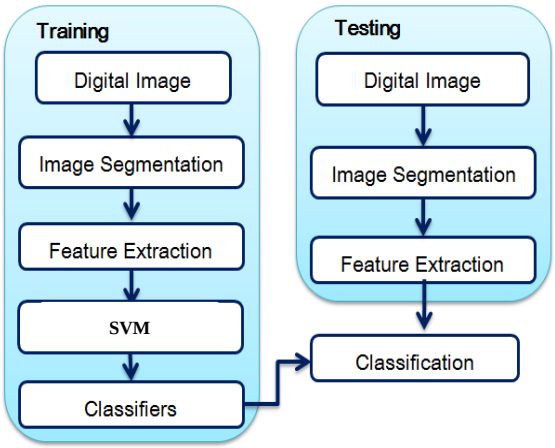
3.1.1 Camera Phone คือหน้า GUI บนระบบแอนดรอยด์ที่ใช้ติดต่อกับผู้ใช้ หากผู้ใช้ถ่ายภาพหรือเลือกภาพใบพืชที่ต้องการวิเคราะห์แล้ว ภาพจะถูกส่งเข้า Server

3.1.2 Server จะเพียงตัวกลางที่ใช้ติดต่อระหว่าง Phone ที่เป็นหน้า GUI ตอบโต้กับผู้ใช้ กับ Matlab

3.1.3 Matlab ทำการวิเคราะห์ภาพใบพืชว่าเป็นโรคใบจุด หรือโรคใบไหม้ หรือไม่เป็นโรค จากนั้นผลวิเคราะห์จะถูกส่งเข้า Database จากนั้น Matlab จะดึงข้อมูลจาก Database เพื่อส่งผ่าน Server ไปยัง Phone

3.1.4 Phone จะดึงผลวิเคราะห์ภาพใบพืชที่ได้จาก Server

3.2 การทำงานภายใน Matlab



ภาพที่ 5 : ขั้นตอนการทำงานใน Matlab

3.2.1 Digital Image การดึงข้อมูลภาพที่นำมาใช้วิเคราะห์ในระบบ จะต้องเป็นภาพสี (RGB)

3.2.2 Image Segmentation

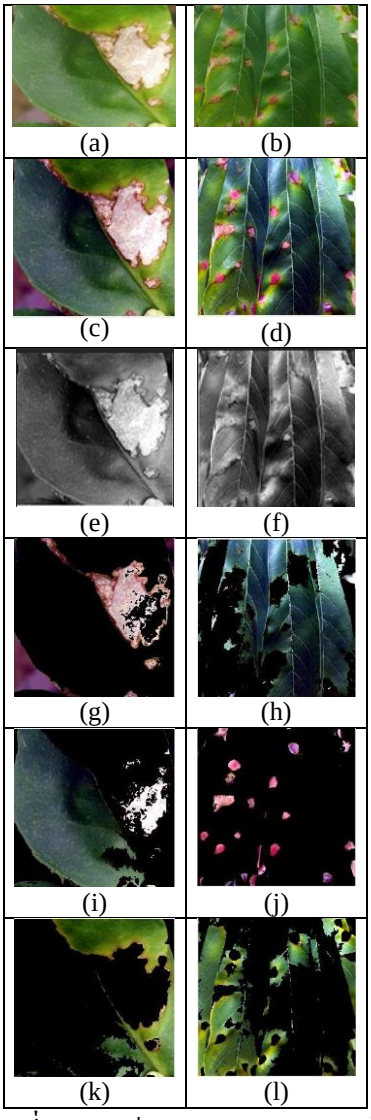
3.2.2.1 ลดสีของภาพที่ไม่จำเป็น เพื่อให้มีความคมชัดเพิ่มขึ้น โดยการแปลง สี RGB เป็นภาพสีเทา

3.2.2.2 Boundary กับ Spot detection Algorithm ใช้ในการหาขอบเขตของภาพ

3.2.2.3 ใช้ Otsu Threshold Algorithm สร้าง Binary image

3.2.3 Feature Extraction ระบุวัตถุที่อยู่ในภาพ

โดยใช้ Gray level Co-occurrence Method จากสมการที่ 4-7 ดังอธิบายไว้ในหัวข้อ 2.8 จากนั้นใช้ อัลกอลิทึม K-means clustering ดังที่อธิบายไว้ในหัวข้อ 2.4 เพื่อแบ่งกลุ่มสีของพื้นที่ใบพืช แบ่งได้ 3 กลุ่ม ให้ผู้ใช้เลือกพื้นที่ส่วนที่เป็นโรค เพื่อเพิ่มความแม่นยำในการวิเคราะห์มากขึ้น



ภาพที่ 6 : การเปลี่ยนแปลงของภาพในแต่ละขั้นตอน

จากภาพที่ 6 นำภาพใบมะนาว (a) และใบมันสำปะหลัง (b) เมื่อนำเข้าระบบแล้วจะต้องทำการปรับความความชัดภาพจึงทำให้ได้ภาพ (c), (d) ก่อนจะทำการแบ่งข้อมูลภาพ (Segment) จะต้องหาค่าความคมชัด (Contrast) ความสัมพันธ์ของพิกเซลที่น่าจะเกิดขึ้น (Correlation) ผลรวมขององค์ประกอบ (Energy) และค่ากระจายตัวขององค์ประกอบ (Homogeity) โดยใช้ Gray level Co-occurrence Method จึงทำให้ได้ภาพ (e), (f) จากนั้นใช้เทคนิค K-mean Clustering ใบมะนาวจะได้ภาพ (g), (i), (k) และใบสำปะหลังได้ (h), (j), (l)

3.2.4 ใช้เทคนิค Support Vector Machine (SVM) ในขั้นตอนการฝึกฝนกับชุดข้อมูลที่มีอยู่ก่อนหน้า เพื่อหาผลลัพธ์ในการวิเคราะห์ โดยนำฟังก์ชัน Kernel มาใช้ในการแปลงข้อมูลให้มีมิติที่สูงขึ้น แปลงข้อมูลให้เป็นข้อมูลเชิงเส้นเพื่อให้แบ่งข้อมูลแบบ Linear ได้ โดยแปลงได้จากสมการที่ 8

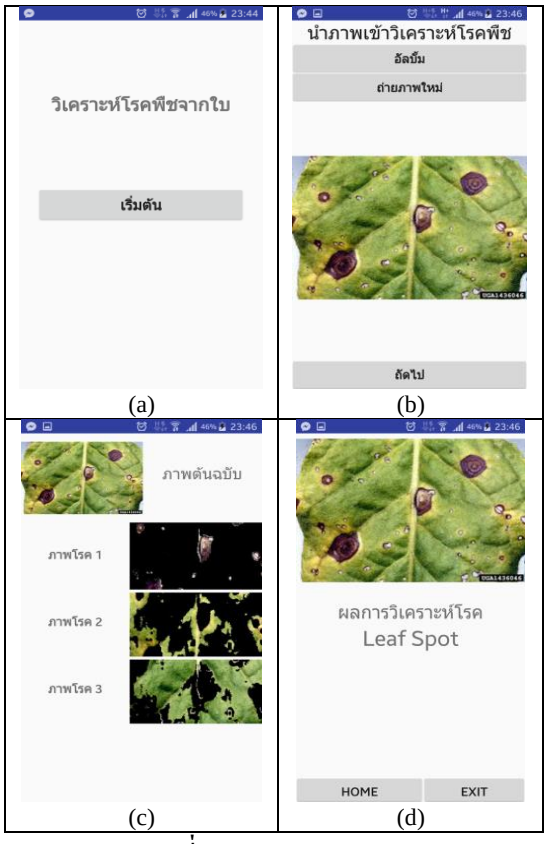
$$c = \sum_i a_i k\left(s_i, x\right) + b_i \qquad (8)$$

โดยที่ a_i เป็นค่าหน้ำหนัก, b เป็นค่าเบี่ยงเบน และ k เป็นฟังก์ชันเคอร์เนล แต่ในกรณีที่เป็นเชิงเส้น ค่า k เป็นจุด ถ้า $C \geq 0$ แล้ว x จะถูกจัดให้อยู่ในกลุ่มแรกมิฉะนั้นจะถูกจัดให้ในกลุ่มที่สอง [12]

3.2.5 Classification ได้ผลวิเคราะห์ภาพใบพืช จำแนกได้ 3 ประเภท ได้แก่ โรคใบจุด โรคใบไหม้ และไม่เป็นโรค

4. ผลการดำเนินงาน

จากการพัฒนาระบบแอนดรอยด์ ได้แอปพลิเคชั่นวิเคราะห์โรคพืชจากใบดังภาพที่ 7



ภาพที่ 7 : ตัวอย่างแอปพลิเคชัน

จากภาพที่ 7 เมื่อเข้าสู่แอปพลิเคชัน จะพบกับหน้าแรก (a) เลือกที่เริ่มต้น จากนั้นเลือกภาพ (b) แอปพลิเคชันจะตัดเฉพาะพื้นที่ส่วนที่เป็นโรคมา 3 ภาพ ให้ผู้ใช้เลือกส่วนที่เป็นโรค (c) แอปพลิเคชันจะแสดงผลวิเคราะห์ (d)

จากการพัฒนาแอปพลิเคชันพบว่าเกษตรกรสามารถวิเคราะห์โรคพืชได้ด้วยตนเอง มีความสะดวก และรวดเร็วในการวิเคราะห์โรคพืชมากขึ้น

จากการวิเคราะห์ภาพใบพืช 4 ชนิด ได้แก่ มะม่วง มันสำปะหลัง พริก และมะนาว คำนวณค่าความแม่นยำ (Accuracy) โดยใช้สมการ

$$Accuracy = (TP + TN)/(TP + TN + FP + FN) \quad (9)$$

เมื่อ True Positive (TP) คือ ผลลัพธ์ที่แอปพลิเคชันทำนายว่าจริง และผู้เชี่ยวชาญบอกว่าผลลัพธ์นั้นจริง

True Negative (TN) คือ ผลลัพธ์ที่แอปพลิเคชันทำนายว่าไม่จริง และผู้เชี่ยวชาญบอกว่าผลลัพธ์นั้นไม่จริง

False Positive (FP) คือ ผลลัพธ์ที่แอปพลิเคชันทำนายว่าจริง แต่ผู้เชี่ยวชาญบอกว่าผลลัพธ์นั้นไม่จริง

False Negative (FN) คือ ผลลัพธ์ที่แอปพลิเคชันทำนายว่าไม่จริง แต่ผู้เชี่ยวชาญบอกว่าผลลัพธ์นั้นจริง

จากจากสมการที่ 9 ความแม่นยำของแอปพลิเคชันเทียบกับผลวิเคราะห์ของผู้เชี่ยวชาญด้านโรคพืชวิทยาพบว่า ผลการวิเคราะห์ภาพใบพืชที่เป็นโรคใบจุด โรคใบไหม้ และไม่เป็นโรค มีความแม่นยำ (Accuracy) เฉลี่ยร้อยละ 80.43 72.31 และ 69.05 ตามลำดับ ดังตารางที่ 1

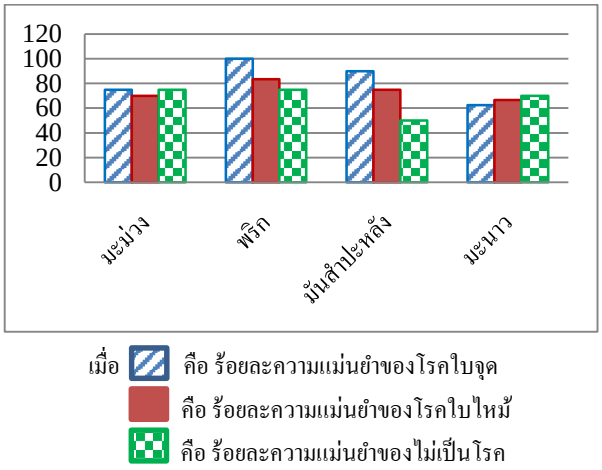
ตารางที่ 1 : ผลการทดสอบข้อมูลแบ่งตามโรค

ประเภทของผลวิเคราะห์	จำนวนข้อมูล (ภาพใบพืช)	ความแม่นยำเฉลี่ย (ร้อยละ)
ใบจุด	46	80.43
ใบไหม้	65	72.31
ไม่เป็นโรค	42	69.05

ตารางที่ 2 : ผลการทดสอบข้อมูลแบ่งตามชนิดพืช

พืช	ร้อยละความแม่นยำ		
	ใบจุด	ใบไหม้	ไม่เป็นโรค
มะม่วง	75	70	75
พริก	100	83.34	75
มันสำปะหลัง	90	75	50
มะนาว	62.50	66.67	70

แผนภูมิที่ 1 : แผนภูมิแสดงผลการทดสอบข้อมูลแบ่งตามชนิดพืช



จากผลการทดสอบในตารางที่ 2 พบว่าผลวิเคราะห์บางส่วนมีความคลาดเคลื่อน เนื่องจากพื้นหลังของภาพมีสีเดียวกับสีของโรค รวมทั้งแสงที่อยู่ในภาพเป็นสาเหตุด้วยเช่นกัน โดยเฉพาะภาพใบที่ไม่เป็นโรค หากมีพื้นหลังหรือแสงมาก ผลวิเคราะห์จะมีความคลาดเคลื่อนสูง

ส่วนโรคใบไหม้พบว่าผลวิเคราะห์ของโรคใบไหม้มีความคลาดเคลื่อนสูงกว่าโรคใบจุด เนื่องจากขนาดของโรคใบไหม้มีขนาดเล็กมาก แอปพลิเคชันจึงวิเคราะห์ออกมาเป็นโรคใบจุด

5. สรุป

แอปพลิเคชั่นวิเคราะห์โรคพืชจากใบบนแอนดรอยด์ โดยใช้ซัพพอร์ตเวกเตอร์แมชชีน มีจุดประสงค์เพื่อพัฒนาประสิทธิภาพความสะดวกรวดเร็วในการวิเคราะห์คุณภาพของใบประยุกต์กับการใช้เทคนิคซัพพอร์ตเวกเตอร์แมชชีน โดยพัฒนาขึ้นจากเทคนิคการประมวลผลภาพบน Matlab ใช้วิธีการหาขอบภาพ การแบ่งข้อมูลแบบเคมีน และการหาเทรชโฮลด์ หลังจากนั้นใช้เทคนิคซัพพอร์ตแมชชีน ในการ

จำแนกประเภทโรค 3 ชนิด ได้แก่ โรคใบจุด โรคใบไหม้ และไม่เป็นโรค จากผลการทดสอบข้อมูลพบว่าเทคนิคซัพพอร์ตเวกเตอร์แมชชีนสามารถจำแนกโรคได้อย่างมีประสิทธิภาพเทียบเท่ากับการวิเคราะห์ของมนุษย์ โดยโรคใบจุดสามารถวิเคราะห์ได้แม่นยำ(Accuracy) ที่สุด 80.43 % รองลงมาเป็นใบไหม้ และไม่เป็นโรคตามลำดับ เกษตรกรสามารถวิเคราะห์โรคพืชได้ด้วยตนเองผ่านแอปพลิเคชันบนระบบแอนดรอยด์ มีความสะดวก และรวดเร็วในการวิเคราะห์โรคพืชจากใบมากขึ้น

เอกสารอ้างอิง

[1] ไพโรจน์ จ๋วงพานิช, "หลักวิชาโรคพืช" มหาวิทยาลัยเกษตรศาสตร์ คณะเกษตร ภาควิชาโรคพืช ปีที่พิมพ์ 2525 พิมพ์ครั้งที่ 2

[2] วรรณวิไล อินทนู, "การวินิจฉัยโรคพืชและการจัดการโรค" คลินิกสุขภาพพืช ภาควิชาโรคพืช มหาวิทยาลัยเกษตรศาสตร์ วิทยาเขตกำแพงแสน, 2547, พิมพ์ครั้งที่ 1

[3] อาการของโรคบนใบ - สารานุกรมไทยสำหรับเยาวชนฯ [Internet]. [cited 2016 Aug 26]. Available from: http://kanchanapisek.or.th/kp6/sub/book/book.php?book=7&chap=4&page=t7-4-infodetail02.html

[4] แนวคิด Support Vector Machine (SVM) [Internet]. เว็บบอร์ด PHP เว็บส่งเสริมการเรียนรู้ Hosting CRM ERP Server Programming ถาม-ตอบปัญหา. [cited 2016 Aug 28]. Available from:http://www.mindphp.com/forums/viewtopic.php?f=29&t=12298

[5] Akhtar A, Khanum A, Khan SA, Shaukat A. Automated Plant Disease Analysis (APDA): Performance Comparison of Machine Learning Techniques. In: 2013 11th International Conference on Frontiers of Information Technology (FIT). 2013. p. 60–5.

[6] Khirade SD, Patil AB. Plant Disease Detection Using Image Processing. In: 2015 International Conference on Computing Communication Control and Automation (ICCUBEA). 2015. p. 768–71.

[7] kmeans [Internet]. [cited 2016 Aug 28]. Available from: https://www.mathworks.com/help/stats/kmeans.html

[8] Najaidee W. วิชุดา นาใจดี: Image Digital [Internet]. วิชุดา นาใจดี. 2556 [cited 2016 Feb 14]. Available from: wichuda111.blogspot.com/2013/11/image-digital.html

[9] รุ่งทิวา เตียวยอดปริยากร จิมขุนทด. โปรแกรมประเมินอัตราการแตกหักของเมล็ดข้าว โดยใช้การประมวลผลภาพดิจิตอล [Internet]. 2556 [cited 2016 Feb 15]. Available from:https://app.enit.kku.ac.th/mis/administrator/doc_upload/20140313090042.pdf.

[10] นวภัค เอื้ออนันต์. Image Segmentation [Internet]. 2545 [cited 2016 Feb 15]. Available from : http://www gear.kku.ac.th/~nawapak/Presentation/Biomed2002.ppt.

[11] สมการ (1)-(7) [Internet]. 2556 [cited 2016 Feb 18]. Available from: https://www.mathworks.com

[12] R. M. Haralick, K. Shanmugam, I. Dinstein. Textural features for Image Classification. IEEE Transactions on Systems, Man and Cybernetics. Vol. SMC-3, 1973. p. 610-621.

การรู้จำโน้ตดนตรีสากลด้วยแสงโดยใช้เทคนิค การเทียบแผ่นแบบ และวิธีการแบ่งส่วนตามความกว้าง

Optical Music Recognition using Template Matching and Amplitude segmentation methods

ปรัชญา ดอกไม้งาม (Pratchaya Dokmaingam)[1] สิริวิมล วรดิษฐ์พัฒนะ (Sirivimon Voraditpatthana)[2]
และทศพร สายยิ้ม(Tosaporn Saiyim)[3]
สาขาวิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ ศรีราชา มหาวิทยาลัยเกษตรศาสตร์
[1]Patchaya.d@ku.th, [2]Sirivimon.v@ku.th, [3]Tosaporn.sa@ku.th

บทคัดย่อ

การรู้จำตัวโน้ตดนตรีด้วยแสง การรู้จำตัวโน้ตดนตรีด้วยแสงปกติแล้วใช้ฮิสโตแกรมในการเทียบสัญลักษณ์ทางดนตรีให้สามารถรับรู้ถึงระดับเสียงของตัวโน้ตดนตรี แต่ด้วยข้อจำกัดดังกล่าวการรู้จำตัวโน้ตได้เพียงแค่ระดับเสียงของตัวโน้ตแต่ไม่ครอบคลุมถึงความยาวของเสียงที่ต้องอ้างอิงจากลักษณะของตัวโน้ต งานวิจัยนี้จึงนำเสนอวิธีการใช้การเทียบแผ่นแบบโดยใช้ภาพกระดาษโน้ตดนตรีที่มีความคมชัดสูงปราศจากสัญญาณรบกวน สัญลักษณ์ต่างๆในเพลงไม่สามารถนำมาใช้งานได้ และวิธีการนี้รองรับเพียงกุญแจประจำหลักที่เป็นกุญแจซอลเท่านั้น จากนั้นนำภาพกระดาษโน้ตดนตรีที่ตรงตามเงื่อนไขดังกล่าวมาเข้าสู่ขั้นตอนการแยกองค์ประกอบของภาพด้วยวิธีการ Bilevel Luminance Thresholding ด้วยขั้นตอนดังกล่าวทำได้ผลลัพธ์เป็นรูปภาพตัวโน้ตดนตรีที่ถูกแบ่งส่วนออกแล้วจึงถูกนำไปหาลักษณะของตัวโน้ตด้วยเทคนิค Template Matching และหาระดับของเสียงด้วยวิธีการ Histogram หลังจากนี้จึงระบุเป็นข้อความในขั้นตอนสุดท้าย งานวิจัยนี้จึงสามารถลดข้อจำกัดของการรู้จำตัวโน้ตดนตรีด้วยแสงแบบเดิมได้

คำสำคัญ: การรู้จำตัวโน้ตดนตรีสากลด้วยแสง การรู้จำอักษรด้วยแสง การแบ่งส่วนตามความกว้าง การเทียบแผ่นแบบ การแบ่งส่วนโน้ตดนตรีสากล กราฟแท่ง

Abstract

Optical Music Recognition (OMR) usually uses histogram to compare musical symbols. To recognize the volume of a musical note. With the limitations of recognizing a music note with general recognition, the note recognizes only the volume of a note but does not cover the length of the note based on the style of the note. Therefore to reduce the limitations as mentioned previously.This article presents the use with Template Matching. This research uses a high-definition noise-canceling picture but music symbol cannot be applied to this method and this method only supports the G-clef. Then, take a music sheet that meets the conditions mentioned above into the segmentation by Amplitude segmentation methods and then use Bilevel Luminance Thresholding. The result is a picture of a divided musical note. The output is then taken to search the characteristic of the note by the template matching technique and find the sound level by using histogram, after that marked as the message in the final state. This research can reduce the limitations of traditional musical note recognition.

Keyword: Optical Music Recognition, Optical Character Recognition, Amplitude Segmentation, Template Matching, Histogram

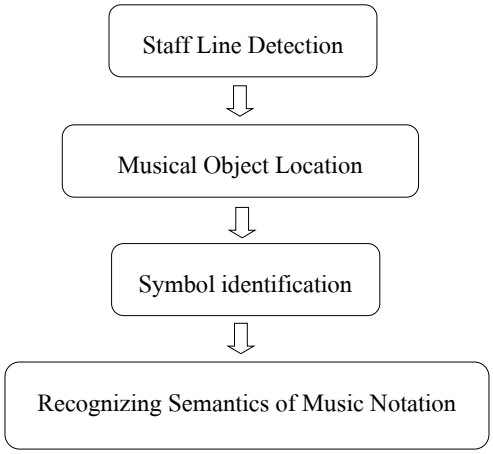
1. บทนำ

ปัจจุบันนี้เทคโนโลยีได้พัฒนาไปอย่างรวดเร็วและไม่มีที่สิ้นสุดเพื่อช่วยอำนวยความสะดวกให้กับมนุษย์ เทคโนโลยีจึงตอบโจทย์การใช้ชีวิตของมนุษย์ในยุคปัจจุบัน และเทคโนโลยีที่กำลังเริ่มได้รับความนิยมเพิ่มขึ้นในยุคปัจจุบันนี้อีกหนึ่งเทคโนโลยีคือ การรู้จำตัวโน้ตดนตรีด้วยแสง เพราะดนตรีเป็นส่วนหนึ่งของวิถีชีวิตมนุษย์ มนุษย์รู้จักดนตรีมาตั้งแต่สมัยโบราณโดยเริ่มจากการนำสิ่งของใกล้ตัวมาทำให้เกิดเสียงและเป็นจังหวะ เช่น ผิวปาก นำกิ่งไม้มาเคาะ ปรบมือ เป็นต้น ซึ่งยุคต่อจากนั้นจึงมีการพัฒนาสร้างเครื่องดนตรีชนิดต่างๆเรื่อยมาจนถึงยุคปัจจุบันเครื่องดนตรีสากลเป็นกลุ่มเครื่องดนตรีที่ได้รับความนิยมในการเล่นมากที่สุด ซึ่งหากนำดนตรีสากลมาผนวกเข้ากับเทคโนโลยี ผลที่ได้จึงเกิดเป็น การรู้จำตัวโน้ตดนตรีด้วยแสง ซึ่งมีพื้นฐานมาจาก การรู้จำอักขระด้วยแสง (Optical character recognition : OCR) [1] ที่เป็นกระบวนการแปลงข้อมูลสื่อสิ่งพิมพ์ต่างๆที่อยู่ในรูปแบบของรูปภาพผ่านกระบวนการประมวลผลภาพ (Image Processing)ให้ได้ผลลัพธ์ออกมาเป็นข้อความเพื่อสามารถนำอักษรที่ได้ไปต่อยอดและพัฒนาต่อไป การรู้จำตัวโน้ตดนตรีด้วยแสงมีหลักการและแนวคิดคล้ายคลึงกับการรู้จำอักขระด้วยแสง แต่เพียงใช้รูปภาพกระดาษโน๊ตดนตรีสากลในการนำมาประมวลผลภาพ ซึ่งการรู้จำตัวโน้ตดนตรีด้วยแสงของงานวิจัยนี้ มีขั้นตอนการทำงานที่เป็นใจความสำคัญ คือการแบ่งตัวโน้ตออกทีละตัวแบบ Amplitude segmentation methods ด้วยวิธีการ Bilevel Luminance Thresholding [2] แล้วจึงหาระดับของเสียงด้วยวิธีการ Histogram ซึ่งเป็นการตรวจสอบและวิเคราะห์ตัวโน้ตและสัญลักษณ์ทางดนตรีที่เกิดขึ้นบนกระดาษดนตรีสากล ซึ่งทั่วไปแล้วการรู้จำตัวโน้ตดนตรีด้วยแสงจะยังไม่สามารถลงรายละเอียดของจังหวะการเล่นตัวโน้ตตัวนั้นๆได้ จะบอกเพียงระดับเสียงของตัวโน้ตที่จับได้ และการตรวจสอบและวิเคราะห์ตัวโน้ตยังมีประสิทธิภาพไม่สูงมากนัก เหตุนี้ที่กล่าวมาทางผู้วิจัยจึงมีแนวคิดที่จะนำนำเอาการเปรียบเทียบแผ่นแบบ (Template Matching) [3] มาเพิ่มเติมในส่วนของการวิเคราะห์และระบุตัวโน้ต โดยภาพโน้ตที่ถูกแบ่งมาแล้ว จะถูกนำมาเปรียบเทียบแม่แบบที่มีอยู่ในชุดข้อมูล เพื่อระบุคุณลักษณะของตัวโน้ตดังกล่าวเพิ่มเติมจากแบบเดิมที่จะระบุได้เพียงระดับเสียง ผลลัพธ์ที่ได้ จะได้อักษรที่ระบุระดับของเสียง และ ความยาวเสียงของตัวโน้ตนั้นๆ เรียงตามลำดับ

2. วรรณกรรมที่เกี่ยวข้อง

2.1 ทฤษฎีการรู้จำตัวโน้ตดนตรีด้วยแสง (Optical music recognition)

การรู้จำตัวโน้ตดนตรีด้วยแสง [4] เรียกโดยย่อว่า *OMR* เป็นกระบวนการที่ทำให้คอมพิวเตอร์รู้จำและแปลงกระดาษโน้ตดนตรีให้แสดงผลลัพธ์ออกมาเป็นเสียง โดยที่คอมพิวเตอร์จะรู้จำและนำมาประมวลผลเปรียบเทียบความเหมือนความต่างของรูปแบบรวมทั้งลักษณะของรูปแบบ เพื่อให้คอมพิวเตอร์วิเคราะห์ความคล้ายคลึงกัน แล้วจึงทำการแสดงผลลัพธ์ออกมา โดย [5]ขั้นตอนการทำงาน *OMR* ที่ดีควรประกอบด้วย *4* ขั้นตอนดังนี้



ภาพที่ 1: แสดงขั้นตอนการทำงานที่ดีของ OMR

ขั้นตอนที่ 1: Staff Line Detection (ตรวจจับบรรทัด 5 เส้น) ขั้นตอนนี้เป็นขั้นตอนที่ตรวจหา บรรทัด 5 เส้น แล้วทำการแยกส่วนออกมาพร้อมกับสัญลักษณ์ทางดนตรี

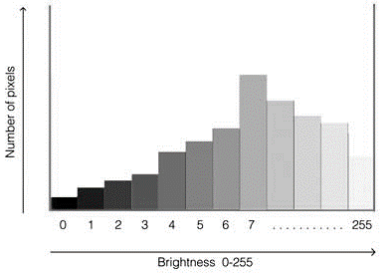
ขั้นตอนที่ 2: Musical Object Location (ตรวจจับสัญลักษณ์ทางดนตรี) ตรวจจับสัญลักษณ์ที่อยู่ในพื้นที่ของบรรทัด 5 เส้น

ขั้นตอนที่ 3: Symbol identification (การระบุสัญลักษณ์) ขั้นตอนนี้สัญลักษณ์จะถูกตรวจจับแล้วแยกออกมาจาก บรรทัด 5 เส้น และถูกระบุตามสัญลักษณ์ เช่น กุญแจประจำหลัก ตัวโน้ต เป็นต้น

ขั้นตอนที่ 4: Recognizing Semantics of Music Notation (ตรวจจับสัญลักษณ์ตามความหมาย) สัญลักษณ์ทางดนตรีจะถูกแบ่งออกมาตามความหมายของแต่ละตัว เช่น ตัวกลม เขบ็ตสองชั้น เขบ็ตหนึ่งชั้น

2.2 Histogram

คือกราฟแท่งที่แสดงให้เห็นถึงปริมาณของพิกเซล [6] ที่มีค่าความสว่างตั้งแต่ค่า 0 ถึง 255 ซึ่งเป็นตำแหน่งที่ขาวที่สุด เห็นได้ว่าแกนทางแนวนอนของกราฟ ภาพที่2 แสดงค่าความสว่าง เป็นภาพเกรสเกล 256 ระดับ ตั้งแต่ 0-255 แนวแกนตั้งแสดงปริมานพิกเซล



ภาพที่ 2: กราฟแสดงค่าความสว่าง

2.3 Image Segmentation

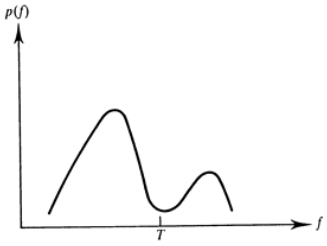
การแบ่งส่วน เป็นวิธีการที่ทำให้สามารถแยกข้อมูลที่เราต้องการออกมา หลักการพื้นฐานของการแบ่งคือการพิจารณา Image amplitude คือการพิจารณาโดยดูจากความสว่างหากเป็นภาพประเภทภาพเกรย์สเกล แต่ถ้าหากภาพเป็นประเภท RGB ให้พิจารณาจากความแตกต่างของสี รวมถึงองค์ประกอบอื่นๆของภาพ เช่น พื้นผิว ขอบของภาพ

2.3.1 Segmentation แบบ Amplitude segmentation methods วิธีการ Bilevel Luminance Thresholding

Segmentation แบบ Amplitude segmentation [7] เป็นการแบ่งส่วนแบบพิจารณาจากความเข้มของแต่ละจุดพิกเซลภายในภาพผลของการแบ่งจะขึ้นอยู่กับวิธีการ Threshold องค์ประกอบ สี และความเข้มของภาพ

วิธีการ Bilevel Luminance Thresholding ใช้สำหรับภาพที่มีลักษณะวัตถุที่สนใจมีความเข้มคงที่เมื่อเทียบกับพื้นหลัง เช่นภาพของตัวอักษร ซึ่งจะทำให้สามารถแยกวัตถุออกจากพื้นหลังได้อย่างชัดเจน

การทำการ Segmentation สามารถทำได้โดยการกำหนดค่า Threshold [8] ซึ่งเป็นค่าความเข้มให้มีค่าที่สามารถแยกความแตกต่างของวัตถุและพื้นหลังได้ตัวอย่างอย่างเช่น ภาพของตัวอักษรที่มีความความเข้มของตัวอักษรเป็น 0 สีดำ และมีความเข้มของพื้นหลังเป็น 255 สีขาว ดังนั้นค่า Threshold จึงควรจะมีค่าเท่ากับ 128 เพื่อที่จะให้สามารถแยกวัตถุออกจากพื้นหลังได้ โดยเลือกค่าจาก histogram การเลือกค่าวัตถุที่จะตัดออกมาควรที่จะเลือกค่า histogram ที่อยู่ที่จุดต่ำสุดที่อยู่ระหว่างจุดสูงสุด (peaks) ดังภาพที่4



ภาพที่ 3: ภาพแสดงจุด peaks บน histogram

2.4 การจับคู่แผ่นแบบ (Template Matching)

คือการตรวจสอบว่ามีส่วนของภาพที่ตำแหน่งใดที่เหมือนกับแม่แบบ [9]โดยหลักการแล้วการจับคู่กับแม่แบบ มีการทำงานโดยมีเลื่อนแม่แบบจากซ้ายไปขวาและจากบนลงล่างทีละพิกเซลจนครบทั้งภาพนับว่ามีพิกเซลที่มีค่าตรงกับแม่แบบกี่พิกเซล ตำแหน่งของการเลื่อนแม่แบบที่ได้จำนวนพิกเซลที่ตรงกับแม่แบบมากพอก็คือตำแหน่งที่มีภาพเหมือนแม่แบบ สามารถคำนวณได้จากสมการ

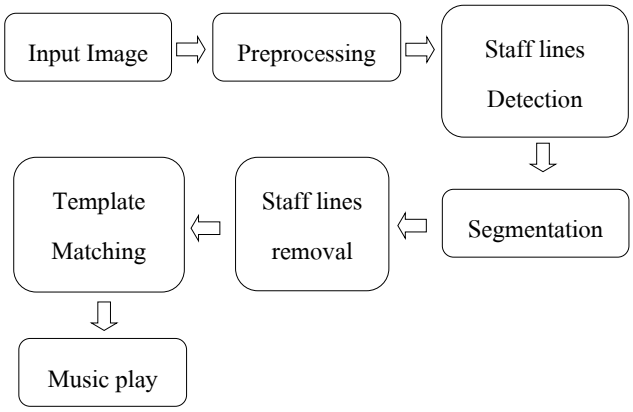
$$R(x,y)^n = \sum_{x',y'}^{n}(\mathrm{T}(\mathrm{x}',\mathrm{y}') - I(x,y))^2 \quad (1)$$

T คือ ภาพแผ่นแบบ
I คือ ภาพที่ต้องการนำไปเปรียบเทียบ
R คือ ระยะทางยูคลิเดียนระหว่างภาพ T และ I

3. วิธีการดำเนินการวิจัย

3.1 ขั้นตอนการดำเนินงาน

เป็นขั้นตอนการแสดงการดำเนินงานของงานวิจัยนี้โดยการทำงานงานของระบบเริ่มจากการรับรูปภาพของกระดาษโน้ตดนตรีสากลเข้ามา ทำการเตรียมรูปภาพ เพื่อเป็นการลดนอยด์และง่ายต่อกระบวนการทำงาน หลังจากนั้นกระทำการตรวจจับบรรทัด 5 เส้น แบ่งส่วนบรรทัด 5 เส้น นำสัญลักษณ์ทางดนตรีออกจากบรรทัด 5 เส้น ลบเส้นของบรรทัด 5 เส้นออก จากนั้นนำสัญลักษณ์ทางดนตรีที่ได้ไปทำการเทียบกับแผ่นแบบที่เก็บไว้ในชุดข้อมูล และทำการแสดงเสียงออกมา

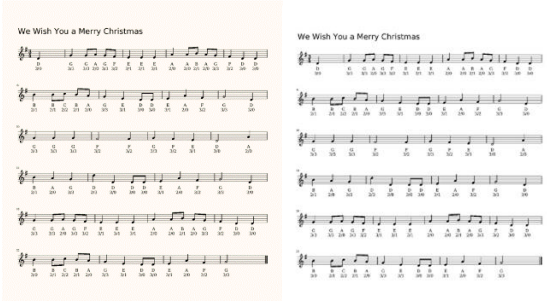


ภาพที่ 4: แสดงขั้นตอนการดำเนินงาน

3.2 ขั้นตอนการเตรียมข้อมูล (Preprocessing)

เป็นขั้นตอนที่ทำให้ภาพสามารถทำการประมวลผลตามระบบได้โดยมีขั้นตอนเตรียม 3 ขั้นตอนด้วยกันคือ ทำภาพให้อยู่ในรูปแบบภาพเกรย์สเกล การทำภาพ Thresholding และ Reverse Image

การทำภาพเกรย์สเกลในส่วนนี้เพื่อให้ภาพมีความคมชัดและง่ายต่อการทำภาพ Thresholding ในขั้นตอนการเตรียมภาพขั้นถัดไป



ภาพที่ 5 : แสดงการแปลงภาพ RGB เป็น ภาพเกรย์สเกล
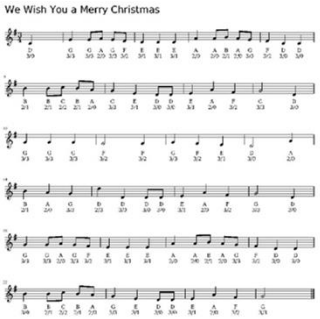
การทำภาพ Thresholding ทำเพื่อแยกวัตถุจากฉากหลัง [8] ให้มีค่าของพิกเซลเพียงแค่ 0,1 โดยที่ 0 คือค่าของพิกเซลสีดำ 1 คือค่าของพิกเซลสีขาว การทำภาพ Thresholding โดยทำการปรับค่าของพิกเซลภายในภาพเกรย์สเกล ที่มีค่าของพิกเซลตั้งแต่ 0-255 ถ้าค่าของพิกเซลภาพมีค่ามากกว่า 128 ให้ค่าของพิกเซลแทนที่ด้วย 1 หรือ สีขาว น้อยกว่า 128 แทนที่ด้วย 0 หรือ สีดำ โดยสามารถเขียนเป็นสมการได้ดังนี้

$$g(x,y) = \begin{cases} 1 & \text{if} \quad f(x,y) > 128 \\ 0 & \text{if} \quad f(x,y) \leq 128 \end{cases} \qquad (2)$$

$g(x,y)$ คือ ภาพ Thresholding ที่เกิดจากค่าพิกเซลของ ภาพ Grayscale

$f(x,y)$ คือ ที่เกิดจากค่าพิกเซลของ ภาพ Grayscale

การทำภาพ Thresholding เป็นวิธีการที่ทำให้ง่ายต่อการแบ่งส่วนภาพเพราะทำให้ค่าของภาพมีแค่สีขาว กับสีดำเท่านั้น และจากการทำภาพ Thresholding จะเห็นได้ว่าภาพมีบางส่วนขาดหายไปเพราะจุดนั้นเป็นจุดที่มีค่าของสีพิกเซลสูงกว่า 128 ทำให้ค่ากลายเป็นสีขาว



ภาพที่ 6: ภาพ Thresholding
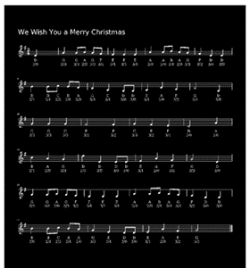
การ Reverse Image ทำเพื่อให้ histogram เน้นจุดสนใจไปที่การจับส่วนสัญลักณ์ทางดนตรี เนื่องจาก histogram จะจับส่วนที่มีค่าความสว่าง ส่วนที่ไม่สำคัญจึงทำให้เป็นสีดำ



ภาพที่ 7: ภาพ Reverse Image

3.3 ตรวจจับบรรทัด 5 เส้น

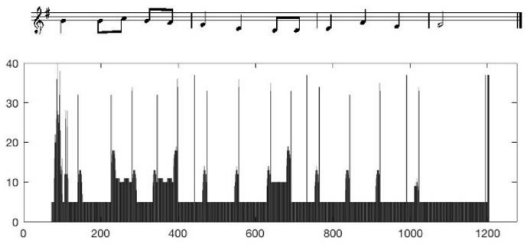
ใช้ histogram หาค่าความสว่างเพื่อทำการตรวจจับบรรทัด 5 เส้น



ภาพที่ 8: แสดงภาพ histogram ของภาพกระดาษโน้ตดนตรี

3.4 การแบ่งส่วน (Segmentation)

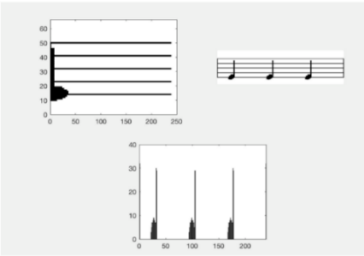
เป็นขั้นตอนการแบ่งกระดาษโน้ตดนตรี โดยอิงตามค่าที่อยู่ระหว่างค่าสูงสุดกับค่าต่ำสุดของ histogram ในแต่ละช่วง หรือการแบ่งส่วนตามแนวนอน เป็นการแบ่งส่วนบรรทัด 5 เส้นออกจากกัน โดยอิงจาก histogram ที่มีค่าสูงสุดแนวนอนถ้ามีความสูง 75% จากค่าสูงสุดให้ทำการแบ่งแนวนอน ของค่าระหว่างความสูงนั้นกับจุดต่ำสุด



ภาพที่ 9 : ภาพแสดงการแบ่งส่วนตามแนวนอน

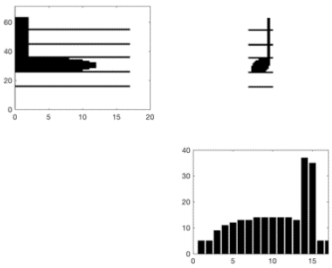
การแบ่งส่วนตามแนวตั้ง

การแบ่งส่วนห้องดนตรี ทำโดยการอ้างอิงจากค่า histogram ที่มีค่าสูงสุด ในแนวนอน ซึ่งคือเส้นกั้นห้องดนตรีแล้วจึงทำการแบ่งส่วนค่าระหว่างจุดสูงสุดกับจุดต่ำสุดออกมา



ภาพที่ 10 : แสดงการแบ่งส่วนห้องดนตรี
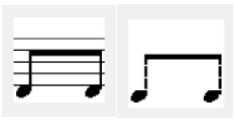
การแบ่งสัญลักษณ์ทางดนตรีออกจากห้องดนตรี เป็นการแบ่งส่วนแนวตั้งที่ทำต่อจากการแบ่งส่วนห้องดนตรี ซึ่งจะทำการแบ่งโดยอ้างอิงจากค่าความถี่ของสัญลักษณ์ทางดนตรีในแนวแกนตั้งระหว่างค่าความถี่สูงสุดและค่าความถี่ที่ต่ำสุด



ภาพที่ 11 : แสดงการแบ่งสัญลักษณ์ทางดนตรี

3.5 ลบบรรทัด 5 เส้น (Staff lines removal)

การลบบรรทัด 5 เส้นดำเนินการลบได้ด้วยการ อ้างอิงจากค่า histogram [10] ที่มีค่าสูงสุดแนวนอน และ ทำให้ส่วนนั้นมีค่าพิกเซลเป็น 1 ซึ่งคือค่าสีขาว



ภาพที่ 12 : แสดงการลบบรรทัด 5 เส้น

3.6 การเทียบแผ่นแบบ (Template Matching)

เป็นการนำตัวโน้ตที่ได้ไปเทียบกับตัวแผ่นแบบที่มีอยู่ในฐานข้อมูลว่าเป็นสัญลักษณ์ทางดนตรีประเภทใดแล้วจึงทำการบันทึก ซึ่งมีการคำนวณค่า [11] correlation coefficient เพื่อให้มีความถูกต้องแม่นยำมากยิ่งขึ้น โดยมีสมการดังนี้

$$r = \frac{\sum_m \sum_n (A_{mn} - \bar{A})(B_{mn} - \bar{B})}{\sqrt{\sum_m \sum_n (A_{mn} - \bar{A}) \sum_m \sum_n (B_{mn} - \bar{B})}} \quad (3)$$

r คือ ค่า correlation coefficient

A คือ ภาพแผ่นแบบ

B คือ ภาพที่นำมาเปรียบเทียบ

$\bar{A}$ คือ ค่าเฉลี่ยของค่าพิกเซลภาพแผ่นแบบ

$\bar{B}$ คือ ค่าเฉลี่ยของค่าพิกเซลภาพที่นำมาเปรียบเทียบ

ซึ่งเลือกจากสัญลักษณ์ทางดนตรีที่มีค่า r สูงที่สุด แปลว่ามีความคล้ายกันมากจึงจะจัดว่าเป็นสัญลักษณ์ทางดนตรีนั้นๆ ถ้าค่า r มีค่าเข้าใกล้ 1 มาก หมายความว่ามีความสัมพันธ์กันมาก หรือมีความคล้ายกันมาก

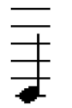
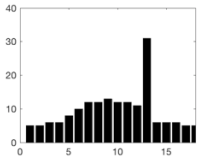
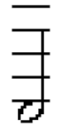
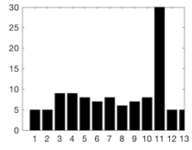
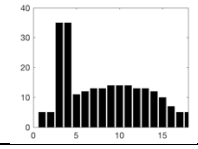
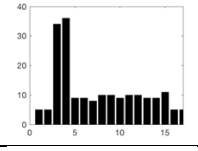
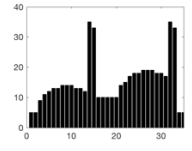
3.7 เล่นเพลงตามสัญลักษณ์ทางดนตรี (Music play)

เป็นขั้นตอนการนำสัญลักณ์ทางดนตรีที่ทำการบันทึกประเภทไว้มาเทียบเสียงเพลงและเล่นเสียงออกมา

4. ผลการดำเนินงาน

การใช้ histogram และ template matching ทำให้ได้ระดับเสียงของตัวโน้ตและความยาวของเสียง เพราะมีการเทียบแผ่นแม่แบบที่เทียบเป็นพิกเซลซึ่งระเอียดกว่าการเทียบแบบ histogram เพียงอย่างเดียว

ตารางที่ 1: แสดง histogram ของตัวโน้ตบางส่วน

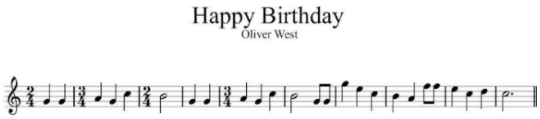
ตัวโน้ต	histogram
	
	
	
	
	

ตารางที่ 2 : แสดงตัวอย่างค่าเฉลี่ย จากค่าcoefficient ที่เกิดระหว่างรูปตัวโน้ตที่ได้กับ ภาพแผ่นแบบ

Template	ค่าเฉลี่ย coefficient
	0.7856
	0.7755
	0.8472
	0.6341

ตารางที่ 2 : แสดงตัวอย่างค่าเฉลี่ย จากค่าcoefficient ที่เกิดระหว่างรูปตัวโน้ตที่ได้กับ ภาพแผ่นแบบ (ต่อ)

	0.6444
	0.7183
	0.6932
	0.6810
	0.5708
	0.6182
	0.6929
	0.5447



ภาพที่ 12 : แสดงรูปภาพกระดาษโน้ตเพลงที่ใช้ทดสอบ

```
g4; g4; a4; g4; c5; b4; g4; g4; a4; g4; c5;
b4; b4; g5; e5; c5; b4; a4; d5; e5; c5; d5;
```

ภาพที่ 13 : ตัวอย่างผลออกมาจาการเทียบแบบ histogram

```
G-clef; time signature: 3/4; quarter g4; quarter g4;
```

ภาพที่ 14 : ตัวอย่างผลออกมาจาการเทียบแบบ histogram และ template matching

5. สรุป

จากผลการดำเนินการวิจัยพบว่า การนำเอาเทคนิคการเปรียบเทียบแผ่นแบบมาทำงานร่วมกับในส่วนของการวิเคราะห์และระบุตัวโน้ตของ OMR ทำให้ได้รายละเอียดในการระบุคุณลักษณะของตัวโน้ตนั้นๆได้อย่างชัดเจน จึงทำให้เมื่อนำไปประยุกต์ใช้งานแล้ว ทำให้สามารถเล่นโน้ตดนตรีที่มีความสั้นและยาวของเสียงตัวโน้ตที่แตกต่างตามคุณลักษณะที่วิเคราะห์ได้จากการเปรียบเทียบแผ่นแบบ แต่ข้อจำกัดของเทคนิคนี้ คือรูปภาพกระดาษโน้ตดนตรีต้องมีความคมชัดระหว่าง 72 dpi ถึง 150 dpi และต้องมีสัญญาณรบกวน(Noise) ที่น้อยหรือไม่มีเลย ถ้าหากไม่เป็นไปตามเงื่อนไขข้างต้นที่กล่าวมา ตัวเทคนิคจะไม่สามารถตรวจจับคุณลักษณะของตัวโน้ตบน

รูปภาพกระดาษโน้ตดนตรีสากลได้ ซึ่งมีวิธีแก้ไข คือการปรับปรุงคุณภาพของภาพก่อนนำไปประมวลผล (Preprocess) โดยการใช้เทคนิคการลดสัญญาณรบกวน การเบลอภาพ และการจัดการจัดการกับภาพด้วยวิธีการอื่นๆ จะช่วยให้สามารถมีประสิทธิภาพเพิ่มขึ้นได้อีกในอนาคต

เอกสารอ้างอิง

[1] M. R. Islam, C. Mondal, M. K. Azam and A. S. M. J. Islam, "Text detection and recognition using enhanced MSER detection and a novel OCR technique," *2016 5th International Conference on Informatics, Electronics and Vision (ICIEV)*, Dhaka, 2016, pp. 15-20.

[2] T. N. Palghamol and S. P. Metkar, "Flexible luminance thresholding for detecting eyes in color images," *Proceedings of The 2014 International Conference on Control, Instrumentation, Energy and Communication (CIEC)*, Calcutta, 2014, pp. 568-572.

[3] A. C. Roy, M. K. Hossen and D. Nag, "License plate detection and character recognition system for commercial vehicles based on morphological approach and template matching," *2016 3rd International Conference on Electrical Engineering and Information Communication Technology (ICEEICT)*, Dhaka, 2016, pp. 1-6.

[4] A. Singh, K. Bacchuwar, A. Choubey, D. Kumar and S. *3rd International Conference on Computer Research and Development*, Shanghai, 2011, pp. 174-178.

[5] Horst Bunke, Patrick Shein-pei Wang, “Handbook analysis”,*Published by World Scientific*, 1997, ISBN 981022270X,9789810222703

[6] D. Kim; C. Kim, "Contrast Enhancement Using Combined 1D and 2D Histogram-Based Techniques," *in IEEE Signal Processing Letters* , vol.PP, no.99, pp.1-1

[7] T. Wangchamhan, S. Chiewchanwattana and K. Sunat, "Multilevel thresholding selection based on chaotic multi-verse optimization for image segmentation," *2016 13th International Joint Conference on Computer Science and Software Engineering (JCSSE)*, Khon Kaen, 2016, pp. 1-6.

[8] P. Pattrapisetwong and W. Chiracharit, "Automatic lung segmentation in chest radiographs using shadow filter and multilevel thresholding," *2016 International Computer Science and Engineering Conference (ICSEC)*, Chiang Mai, 2016, pp. 1-6.

[9] X. Ma and H. Zhang, "Joint geometry and gray-level histogram model for lip-reading," *2016 12th World Congress on Intelligent Control and Automation (WCICA)*, Guilin, 2016, pp. 3083-3087.

[10] J. Cui, H. He and Y. Wang, "An adaptive staff line removal in music score images," *IEEE 10th INTERNATIONAL CONFERENCE ON SIGNAL PROCESSING PROCEEDINGS*, Beijing, 2010, pp. 964-967.

เกมฝึกการหายใจ
BreathingTraining Game

ณัฐธนนท์ หงส์วริทธิ์ธร (Nuttanont Hongwarittorrn)
สาขาวิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์และเทคโนโลยี
มหาวิทยาลัยธรรมศาสตร์
nth@cs.tu.ac.th

บทคัดย่อ

งานวิจัยทางวิทยาศาสตร์สุขภาพชี้ให้เห็นว่าการฝึกหายใจสามารถช่วยลดความดันโลหิต ลดความเครียด และความกังวลได้ และจำนวนผู้คนที่เล่นเกมเพิ่มมากขึ้นในปัจจุบันเพราะเกมแน่นอนย่อมสร้างความสนุกสนาน ความเพลิดเพลินและความท้าทายให้กับผู้เล่นได้ ดังนั้นงานวิจันชิ้นนี้นำเสนอแอปพลิเคชั่นเกมเพื่อฝึกการหายใจ (BreathingTraining) ที่ผู้เล่นเกมได้รับความสนุกสนานเพลิดเพลินไปพร้อมกับการฝึกหัดหายใจเพื่อปรับอัตราการหายใจให้มีความเหมาะสมสอดคล้องกับเกณฑ์การหายใจทางแพทย์และแอปพลิเคชั่นยังแสดงให้เห็นพัฒนาการการฝึกหายใจของผู้ใช้โดยการเก็บประวัติการฝึก และช่วยเพิ่มแรงดึงดูดใจในการฝึกหายใจของผู้ใช้มากยิ่งขึ้น การทำงานของระบบ BreathingTraining จะรับเสียงลมหายใจของผู้ใช้ผ่านไมโครโฟนแล้วนำไปวิเคราะห์อัตราการหายใจเพื่อนำไปกำหนดระดับการฝึกฝนให้ผู้ใช้ในการเล่นเกมในแต่ละระดับระบบจะแสดงกราฟประวัติการฝึกฝนการหายใจด้วย ผลการทดสอบระบบ พบว่าเป็นระบบที่สามารถใช้งานได้ แต่ผลการคำนวณอัตราหายใจยังมีข้อผิดพลาดอยู่โดยเฉพาะการบันทึกการหายใจในสภาพแวดล้อมที่เสียงรบกวน และส่วนต่อประสานกับผู้ใช้ควรได้รับการปรับปรุงเพื่อให้เป็นระบบที่ใช้งานได้ง่ายต่อไป

คำสำคัญ: แอปพลิเคชัน เกม การหายใจ

Abstract

There is health science research indicating that breathing exercise can help ones reduce blood pressure, stress, and anxiety and there are increasing users in Thailand playing games since games are of course fun joyful, and also challenging to users. The current research therefore proposes an application game for users to exercise breathing while they are enjoying playing the game and getting their better breathing rates regarding the medical respiratory rate. The application also provides user's improvement of breathing exercises and persuades users to practice their breathing more. The breatingtraining game allows users to breathe into a microphone and the application will calculate user's respiratory rate so that it will be used to recommend the respiratory rate for users to practice their breathing. The application displays the history of breathing practices. The experimental evaluation of application showed that it was usable even though some respiratory rates calculated from the system and calculated from the users were different particularly in the noisy environment. Additionally, the user interface should be adjusted to make the application easier to use.

Keyword: Application, game, breathing

1. บทนำ

ในการทำกิจกรรมต่างๆของมนุษย์เรานั้น ระบบต่างๆ ในร่างกายจะต้องทำงานสัมพันธ์กัน เช่น ในขณะที่ออกกำลังกาย ร่างกายต้องการใช้พลังงานมากขึ้นจึงต้องการก๊าซออกซิเจนและสารอาหารเพิ่มมากขึ้น ระบบหายใจจึงต้องทำงานหนักขึ้น ถ้าเราหายใจถี่และเร็วเพื่อนำก๊าซออกซิเจนเข้าสู่ร่างกาย และขับก๊าซคาร์บอนไดออกไซด์ออกไป การหมุนเวียนโลหิตในร่างกายก็จะเป็นไปอย่างรวดเร็ว ขณะที่หัวใจ บีบตัวหรือคลายตัวนั้น เลือดจะหมุนเวียนไปทั่วร่างกายพร้อมทั้งมีการลำเลียงสารอาหารและก๊าซออกซิเจนไปยังส่วนต่างๆของร่างกาย เพื่อสร้างพลังงานให้แก่ร่างกาย จากการศึกษาพบว่า การหายใจที่ถูกวิธี คือการหายใจในอัตราที่ช้าและลึกสามารถช่วยลดความดันโลหิตได้ ลดอาการเหนื่อยล้า ลดอาการเจ็บปวด ลดความเครียด ลดความวิตกกังวล หรือส่งผล

ให้มีสุขภาพจิตที่ดีขึ้นแม้ว่าจะไม่ได้เป็นการรักษาโรค หรืออาการเจ็บป่วยโดยตรง

ระบบช่วยฝึกการหายใจ (Breathing Training) เป็นระบบสำหรับช่วยฝึกการหายใจของผู้ใช้ให้มี อัตราการหายใจเหมาะสมตรงตามเกณฑ์การหายใจทางการแพทย์ โดยสามารถใช้งานได้บนโทรศัพท์เคลื่อนที่ อีกทั้งยังสามารถใช้เป็นเครื่องมือทางการแพทย์แบบส่วนตัว หรือนำไปใช้ฝึกการบริหารลมหายใจกับผู้ป่วยได้ โดยมีการเก็บประวัติลมหายใจของผู้ใช้ทั้งก่อนและหลังการฝึก สามารถเรียกดูประวัติการฝึกของผู้ใช้ได้ โดยระบบจะมีรูปแบบการแสดงผลที่เข้าใจง่าย ใช้เวลาเรียนรู้น้อย ใช้งานได้สะดวก และเหมาะสำหรับผู้ใช้ทุกเพศทุกวัย ในทุกช่วงอายุ ตั้งแต่ 12 – 80 ปี (ผู้ใช้มีอายุน้อยหรือมากเกินไปจะต้องมีผู้ดูแลระหว่างการใช้งานระบบเพื่อให้ได้ผลการฝึกที่ตรงตามเป้าหมายที่กำหนดไว้)

2. ทบทวนวรรณกรรม

2.1 ทฤษฎีที่เกี่ยวข้อง

การหายใจ (Respiration) เป็นการนำเอาออกซิเจนเข้าสู่ร่างกาย และนำคาร์บอนไดออกไซด์ออกจากร่างกาย ปัจจัยที่มีผลต่อการเปลี่ยนแปลงการหายใจ การหายใจเป็นการทำงานแบบอัตโนมัติ แต่อย่างไรก็ตาม จังหวะและความลึกของการหายใจ บางขณะก็สามารถควบคุมได้เป็นพักๆ เช่น การเปลี่ยนแปลงของอารมณ์ หรือสภาวะแวดล้อม หรือได้รับอิทธิพลจากความเจ็บป่วย ซึ่งอาจมีผลทำให้หายใจตื้นหรือลึกกว่าปกติได้ ดังนั้นการนับการหายใจ จึงไม่ควรให้ผู้ป่วยรู้สึกตัว โดยมากนิยมนับต่อจากการคลำชีพจร [6]

อัตราเร็วของการหายใจ มีหน่วยเป็นครั้งต่อนาที(bpm) ซึ่งการหายใจ 1 ครั้ง หมายถึงการหายใจเข้าและหายใจออก 1 รอบ อัตราการการหายใจจะแตกต่างกันตามวัยและอายุ โดยเด็กจะหายใจเร็วกว่าผู้ใหญ่ อัตราการหายใจ แรกเกิด(ภายในเดือนแรก) 30-50 ครั้ง/นาที 2 เดือนถึง 2 ปี 30-40 ครั้ง/นาที 3 ปีถึง 12 ปี 20-25 ครั้ง/นาที 12 ปีถึง 18 ปี 18-20 ครั้ง/นาที ผู้ใหญ่ 14-30 ครั้ง/นาที ผู้สูงอายุ 65 ปี 12-28 ครั้ง/นาทีและผู้สูงอายุ 80 ปี 10-30 ครั้ง/นาที [7]

จากการศึกษาของ Bernardi และคณะได้นำเทคนิคการหายใจช้าและลึกไปฝึกในผู้ป่วย chronic heart failure พบว่าการหายใจที่ 6 ครั้ง/นาที จะช่วยเพิ่ม arterial baroreflex sensitivity และ Joseph และคณะได้ทำการฝึกผู้ป่วยที่มีความดันโลหิตสูงด้วยอัตราการหายใจควบคุมที่ 6 ครั้ง/นาที ใช้เวลาฝึก 15 นาทีทุกวันต่อเนื่องเป็นเวลา 8 สัปดาห์ พบว่าผู้ป่วยมีความดัน systolic ลดลงเฉลี่ย 10 mmHg และความดัน diastolic ลดลงเฉลี่ย 4 mmHg จากงานวิจัยทางการแพทย์จะเห็นได้ว่า การฝึกหายใจช้าและลึกสามารถช่วยในการฟื้นฟูหัวใจโดยเฉพาะผู้ป่วยโรคหลอดเลือดหัวใจ ผู้ป่วย Chronic heart failure และผู้ป่วยความดันโลหิตสูง ซึ่งมีผู้ป่วยโรคเหล่านี้มีปริมาณเพิ่มสูงในประเทศไทย

2.2 งานวิจัยที่เกี่ยวข้อง

BreathingZone [1] เป็นแอปพลิเคชันเกี่ยวกับการฝึกหายใจ ที่สามารถช่วยคลายความเครียด ความวิตกกังวลในชีวิตประจำ วัน และยังสามารถช่วยลดอาการความดันโลหิตสูง ระบบมีภาพเคลื่อนไหวสำหรับช่วยให้ผู้ใช้หายใจเข้าและหายใจออกตามจังหวะการเคลื่อนไหวของภาพ และสามารถปรับความเร็วการหายใจได้ตามที่ต้องการฝึก มีเสียงประกอบการหายใจ สามารถวิเคราะห์การหายใจของผู้ใช้ได้เพื่อที่จะฝึกการหายใจต่อไป Pranayama [2] เป็นแอปพลิเคชันเพื่อสอนให้ผู้ใช้หายใจอย่างถูกต้อง เป็นประโยชน์ต่อสุขภาพ สามารถลดความเครียด และความกังวลได้ ช่วยให้นอนหลับได้เร็วขึ้น ช่วยให้การจัดการโรคหอบหืดและโรคปอดอุดกั้นเรื้อรัง ความดันโลหิตลดลง บรรเทาจากอาการปวดหัวและไมเกรน และอารมณ์ดีขึ้น

Breathing Downhill [3] เป็นแอปพลิเคชันเกม ออกแบบมาเพื่อสอนให้เด็กหายใจออกได้ยาวขึ้น โดยขายพร้อมทั้งอุปกรณ์ เสริมในการจัดเก็บลมหายใจ ลักษณะเกมคือ ให้ผู้เล่นทำการหายใจออกทางปากให้ยาวที่สุด ตัวละครในเกมจะทำการขับจักรยานไปตามทาง และแอปพลิเคชันจะบันทึกค่าลมหายใจและคำนวณผลการหายใจเฉลี่ย Breathe2Relax [4] เป็นแอปพลิเคชันหนึ่งที่ช่วยการฝึกการหายใจที่ช่วยลดความเครียด ความกังวล และให้คำแนะนำการฝึกหายใจ สามารถเลือกการฝึกการหายใจเองได้ และมีการให้ผู้ใช้บันทึกความรู้สึกหลังการฝึกหายใจเช่น เครียด ปกติหรือผ่อนคลาย และ BreathCounter [5] เป็นแอปพลิเคชันที่ถูกพัฒนาขึ้นบันทึกเสียงลมหายใจผ่านไมโครโฟน โดยใช้เวลาบันทึก 30 วินาที จากนั้นระบบจะทำการคำนวณจำนวนครั้งที่หายใจเข้าและออกเป็น 1 ครั้งต่อนาที (bpm) และมีการจัดเก็บประวัติการใช้งานให้

จากแอปพลิเคชันที่กล่าวมาเป็นสิ่งที่แสดงให้ได้ว่า การพัฒนาเครื่องมือช่วยฝึกการหายใจเป็นสิ่งที่มีประโยชน์ต่อสุขภาพของผู้ใช้ ซึ่งต่างประเทศได้พัฒนาขึ้นมากมายและนำออกจำหน่ายมากมาย แต่แอปพลิเคชันที่มีอยู่ยังมิได้พัฒนาเป็นลักษณะของเกม

ถึงแม้ BreathingDownhill จะออกแบบเป็นลักษณะเกมแต่ผู้ใช้จำเป็นต้องซื้ออุปกรณ์เสริมซึ่งมีราคาแพงและไม่สะดวกที่ต้องพกพาอุปกรณ์เพิ่มเติม รัฐบาลมีนโยบายที่สนับสนุนให้ประชาชนมีสุขภาพที่ดี การใช้เทคโนโลยีมาช่วยเพื่อให้สุขภาพของคนในประเทศดีขึ้นย่อมเป็นสิ่งที่ควรได้รับการพัฒนาส่งเสริมให้มากยิ่งขึ้น ประเทศไทยมีจำนวนผู้เล่นเกมเป็นจำนวนมาก ดังนั้น ถ้าหากมีเครื่องมือที่เป็นทั้งเกมซึ่งสามารถช่วยการฝึกการหายใจของคนได้และไม่จำเป็นต้องใช้อุปกรณ์พิเศษแล้ว ย่อมจะเป็นประโยชน์สุขภาพได้ในขณะที่ได้รับความสนุกสนานไปด้วย

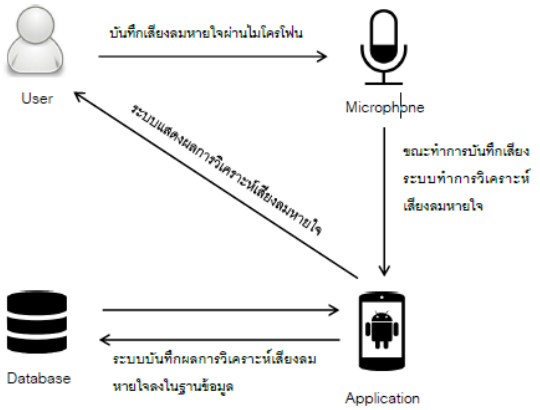
งานวิจัยชิ้นนี้จึงทำการศึกษาแนวการบันทึกการหายใจผ่านไมโครโฟน[8] และนำระดับการหายใจของผู้ใช้ไปใช้ในการออกแบบเกมโดยทั่วไปซึ่งต้องมีลักษณะจูงใจที่ผู้ใช้จะเล่นเพื่อให้ตนเองอยู่ระดับสูงขึ้นหรือเก่งขึ้น จูงใจให้ฝึกการหายใจซึ่งสอดคล้องระดับของการเล่นเกม นั่นคือเพื่อให้อัตราการหายใจดีขึ้นผู้ใช้ก็จะเล่นเกมในระดับที่ท้าทายมากขึ้น

3. ระเบียบการดำเนินงานวิจัย

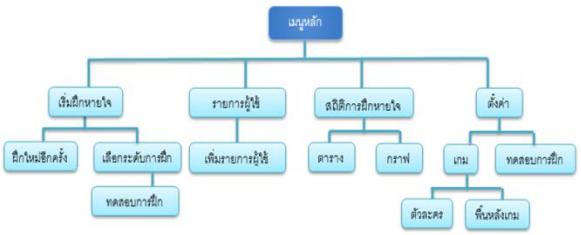
งานวิจัยได้แนวคิดการพัฒนาแอปพลิเคชันการบันทึกเสียงหายใจผ่านทางไมโครโฟน คลื่นเสียงการหายใจที่ผ่านเข้าสู่ไมโครโฟนจะถูกนำไปคำนวณเพื่อหาอัตราการหายใจของผู้ใช้ ข้อมูลส่วนตัวได้แก่อายุ จะถูกนำไปใช้ค้นหาอัตราการหายใจที่เหมาะสมเพื่อนำมาใช้เป็นเกณฑ์เพื่อประเมินอัตราการหายใจของผู้ใช้และนำไปใช้แนะนำระดับในการเล่นเกมเพื่อฝึกให้การหายใจของผู้ใช้ดียิ่งขึ้น งานวิจัยนี้มีสมมติฐานว่า อัตราการหายใจคำนวณได้จากแอปพลิเคชั่นจะไม่แตกต่างจากอัตราการหายใจที่นับจากผู้ใช้

3.1 ภาพรวมและสภาพแวดล้อม (Overall Architecture and Environment

ระบบ BreathingTraining มีสถาปัตยกรรมของระบบดังภาพที่ 1 และภาพรวมของฟังก์ชันการใช้งานดังภาพที่ 2 โดยเริ่มต้นกรอกข้อมูลส่วนตัวได้แก่อายุ และให้ผู้ใช้บันทึกเสียงลมหายใจของผู้ใช้ผ่านไมโครโฟน(ดังภาพที่ 3) แล้วทำการวิเคราะห์คลื่นเสียงลมหายใจ เพื่อแสดงอัตราการหายใจในหน่วย bpm ให้ผู้ใช้ทราบ จากนั้นระบบจะนำข้อมูลที่ได้รับจากผู้ใช้คือ อัตราการหายใจและอายุ มาใช้เป็นเกณฑ์ในการเลือกระดับการฝึกฝน (ดังตัวอย่างจอหน้าในภาพที่ 4) ให้กับผู้ใช้โดยอัตโนมัติในครั้งแรก หากผู้ใช้ฝึกฝนสำเร็จและต้องการฝึกฝนในระดับถัดไป ผู้ใช้จะต้องทำการทดสอบโดยการบันทึกเสียงลมหายใจอีกครั้งเพื่อวัดผลว่าผู้ใช้สามารถรักษาระดับอัตราการหายใจได้คงที่ตามที่ฝึกฝนไว้ และระบบยังสามารถเก็บผลการฝึกฝนของผู้ใช้เพื่อนำมาแสดงการพัฒนา การหายใจของผู้ใช้ในรูปแบบของตารางและกราฟได้อีกด้วย (ดังภาพที่ 7) ระบบ Breathing Training ออกแบบโดยใช้ส่วนต่อประสานกับผู้ใช้ที่สามารถเข้าใจได้ง่าย ทำให้ผู้ใช้ใช้เวลาในการเรียนรู้ระบบน้อยลง อีกทั้งระบบยังสามารถวิเคราะห์คลื่นเสียงลมหายใจของผู้ใช้ได้อย่างทันทีเมื่อรับคลื่นเสียงเข้ามา (Real time) และใช้การออกแบบเกมที่ง่ายๆ เพื่อเป็นการท้าทายให้ผู้ใช้ได้ฝึกฝนหายใจพร้อมกับความสนุกในการเล่นเกมไปด้วย (ดังภาพตัวอย่างที่ 6)



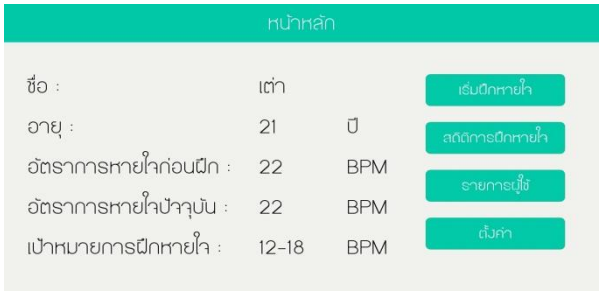
ภาพที่ 1 แสดงสถาปัตยกรรมของระบบ



ภาพที่ 2 แสดงภาพรวมของการใช้งานระบบ



ภาพที่ 3 ตัวอย่างหน้าจอที่จะทำการบันทึกเสียง

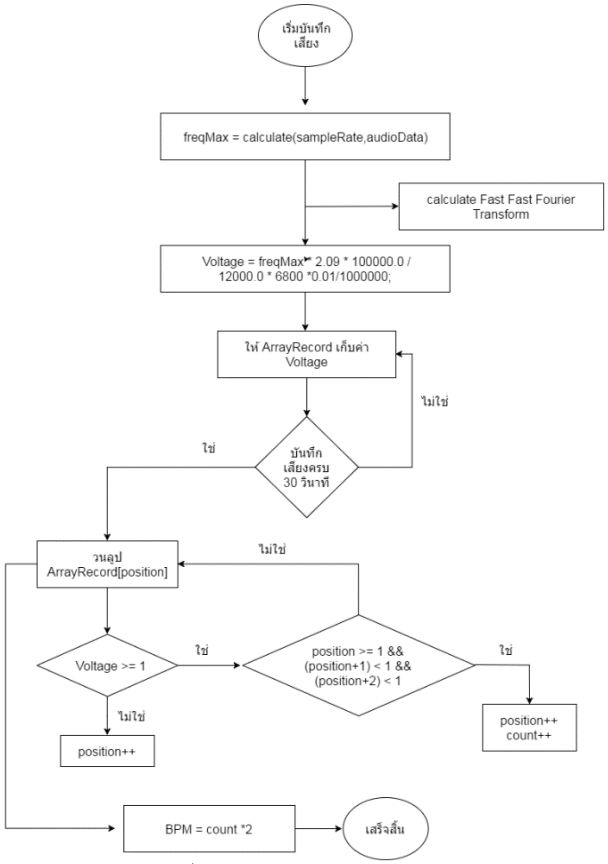


ภาพที่ 4 หน้าจอแสดงผลอัตราการหายใจของผู้ใช้ และเกณฑ์ที่ควรฝึกฝน

วิธีการพัฒนาเกมฝึกหายใจจะใช้ android.graphics.Bitmapในการสร้างหน้าเกม (game panel) ทั้งในส่วนของตัวละคร (character animation) และพื้นหลังเกม (background) แนวทางการออกแบบเกมเน้นที่ให้ผู้ใช้เข้าใจได้ง่าย รูปแบบเกมซึ่งมีลักษณะที่ผู้ใช้อาจจะเคยเล่นเกมหรือคุ้นเคยกับลักษณะของเกมเช่นนี้มาก่อนคือการให้ตัวละครเคลื่อนไปเก็บเหรียญที่เป็นเสมือนรางวัลหรือคะแนน เกมการฝึกการฝึกหายใจในที่นี้ก็จะเป็นไปในลักษณะง่าย ๆ คือการหายใจของผู้ใช้จะเป็นการบังคับการเคลื่อนไหวของตัวละคร การเคลื่อนไหวของตัวละครก็จะสอดคล้องไปกับการหายของใจของผู้ใช้ นั่นคือ เมื่อผู้ใช้หายใจเข้าก็จะเป็นการผลักดันให้ตัวละครเคลื่อนตัวสูงขึ้นไปเก็บเหรียญ ในขณะที่เมื่อผู้ใช้หายใจออก ตัวละครก็จะเคลื่อนกลับลงมาอยู่ในตำแหน่งเดิม ดังนั้นการคำนวณเพื่อให้ตัวละครเคลื่อนที่จะอ้างอิงจากคลื่นเสียงการหายใจของผู้ใช้

ในทุกระดับเกมการฝึกการหายใจ การเคลื่อนไหวของฉากพื้นหลังและเหรียญจะมีความเร็วคงที่ แต่จะแตกต่างกันเพียงความเร็วในการเคลื่อนที่ขึ้นลงของตัวละครและการเคลื่อนที่ของเหรียญสะสมแต้มเท่านั้น โดยจะมีความสัมพันธ์กันกับการหายใจของผู้ใช้ด้วยเงื่อนไข คลื่นเสียงของผู้ใช้ที่คำนวณค่า Voltage ออกมาแล้วมีค่ามากกว่าเท่ากับ 1 (เป็นการสั่งการให้ตัวละครเคลื่อนที่ขึ้น) ในขณะที่ถ้า Voltage มีค่าน้อยกว่า 1 (เป็นการสั่งการให้ตัวละครเคลื่อนที่ลง) ดังลำดับขั้นตอนในการคำนวณในภาพที่ 5

ในการคำนวณการปล่อยตัวของเหรียญในเกมสำหรับเก็บแต้มคะแนนจะแตกต่างกัน ขึ้นอยู่กับระดับในการฝึกหายใจว่ามีค่าเท่าใด ตัวอย่างเช่น ณ ระดับ 20 bpm กล่าวคือ ในเวลา 1 นาทีจะต้องหายใจทั้งหมด 20 ครั้ง แต่เวลาที่ใช้เล่นเกมฝึกหายใจใช้เพียง 15 วินาที แสดงว่าจะต้องหายใจทั้งหมด 20/4 = 5 ครั้งภายใน 15 วินาที เป็นต้น และจะมีข้อจำกัดในเรื่องความต่างของเวลาในเสี้ยววินาทีซึ่งมีความแตกต่างกันน้อยมากทำให้ไม่สามารถเปรียบเทียบความต่างของระดับใกล้ ๆ กันได้อย่างชัดเจน



ภาพที่ 5 ลำดับขั้นตอนการคำนวณหาอัตราการหายใจ

3.2 ขั้นตอนการทดลอง

ในการทดลองเพื่อทดสอบระบบ แอปพลิเคชั่นถูกติดตั้งในมือถือที่มีระบบปฏิบัติการแอนดรอยด์ ใช้ไมโครโฟนเดียวกันดำเนินการทดทองในสถานที่เงียบไม่มีเสียงรบกวนเหมือนกัน จำกัดจำนวนครั้งการทดลองเท่ากัน ใช้สมาร์ทโฟนเครื่องเดียวกันในการทดลอง แอปพลิเคชั่นสามารถติดตั้งและทำงานได้ในสมาร์ทโฟนที่มีระบบปฏิบัติการแอนดรอยด์ แต่อาจจะมีปัญหาในเรื่องของหน้าจออินเตอร์เฟสที่อาจจะมีขนาดแตกต่างกันไปบ้างตามรุ่นและยี่ห้อของสมาร์ทโฟน



ภาพที่ 6 ตัวอย่างหน้าจอเกมการฝึกหายใจ



ภาพที่ 7 ตารางสถิติประวัติการฝึกหายใจ

4. ผลการทดลอง

งานวิจัยนี้มีเป้าหมายเพื่อนำเสนอเกมที่ช่วยการฝึกการหายใจโดยการหายใจผ่านทางไมโครโฟนบนมือถือในระบบแอนดรอยด์ การประเมินผลแอปพลิเคชันประกอบด้วยการเปรียบเทียบความต่างของอัตราการหายใจที่แอปพลิเคชันคำนวณได้เสียงหายใจของผู้ใช้กับการนับอัตราการหายใจด้วยผู้ใช้ การตรวจการทำงานของแต่ละส่วนงานในแอปพลิเคชัน และความพึงพอใจในการทดลองใช้แอปพลิเคชัน ผู้วิจัยจะให้ผู้ใช้ทดลองใช้งานแอปพลิเคชันโดยวัดอัตราความสำเร็จในการใช้งานแต่ละฟังก์ชันในตารางที่ 1 การทดสอบเปรียบเทียบอัตราการหายใจโดยใช้แอปพลิเคชันกับการนับด้วยตนเองจำนวน 20 ครั้งในตารางที่ 2 และผลประเมินความพึงพอใจจากผู้ใช้ที่มีต่อแอปพลิเคชันในตารางที่ 3 ดำเนินการทดสอบกับผู้ใช้จำนวน 20 คนเป็นผู้ชาย 9 คนและผู้หญิง 11 คน ซึ่งมีช่วงอายุ 20-23 ปี ผลการทดสอบมีดังนี้

ตารางที่ 1 แสดงผลการทดสอบการใช้งานแอปพลิเคชัน

งานที่ใช้ทดสอบ	ผิดพลาด (ครั้ง)	สำเร็จ/ไม่สำเร็จ (คน)	อัตราความสำเร็จ
1. เพิ่มบัญชีผู้ใช้	2	สำเร็จ 20	100%
2. ทำการฝึกหายใจ	4	สำเร็จ 20	100%
3. ทดสอบการเลื่อนระดับการหายใจ	-	สำเร็จ 20	100%
4. เลือกดูสถิติการหายใจ แบบตารางและกราฟ	-	สำเร็จ 20	100%
5. เลือกวันที่ ของกราฟสถิติการหายใจ	1	สำเร็จ 20	100%
6. อธิบายความเข้าใจเกี่ยวกับกราฟสถิติการหายใจ	-	สำเร็จ 18 ไม่สำเร็จ 2	90.91%
7. ลบรายการผู้ใช้	3	สำเร็จ 18 ไม่สำเร็จ 2	90.91%

ตารางที่ 2 แสดงผลการเปรียบเทียบอัตราการหายใจโดยใช้แอปพลิเคชันกับการนับด้วยตนเอง

ครั้งที่	อัตราการหายใจโดยใช้แอปพลิเคชัน	อัตราการหายใจโดยการนับด้วยตนเอง	ค่าความต่าง (%)
1	20	22	10.0
2	24	24	0.0
3	8	15	87.5
4	18	18	0.0
5	20	20	0.0
6	18	22	22.2
7	10	18	80.0
8	16	17	6.3
9	14	15	7.1
10	22	20	9.1
11	20	21	5
12	13	16	23.1
13	20	20	0
14	17	18	5.8
15	21	21	0
16	15	15	0
17	16	17	6.25
18	17	20	17.6
19	12	18	50
20	18	19	5.56
ค่าเฉลี่ย	16.95	18.8	16.78

ตาราง 3 แสดงความถี่ระดับความพึงพอใจของผู้ใช้

รายการประเมินผล	ค่าเฉลี่ย	ส่วนเบี่ยงเบนมาตรฐาน
รูปแบบการใช้สีเหมาะ สมกับระบบ	4.10	0.97
รูปแบบไอคอนเหมาะ สมกับระบบ	3.75	1.02
ระบบสามารถช่วยในฝึกหายใจ	3.55	1.00
ระบบมีฟังก์ชั่นการใช้งานที่ครบถ้วน/เหมาะสม	3.25	1.02
ระบบมีการทำงานที่ชัดเจนและเข้าใจได้ง่าย	3.25	0.91
ระบบสามารถใช้งานได้อย่างอิสระ	3.55	0.89
ระบบมีความน่าสนใจในการใช้งาน	3.35	1.09
ระบบเหมาะสมที่จะพัฒนาต่อในอนาคต	4.15	1.04
ภาพรวม	3.62	0.99

ผลการทดสอบฟังก์ชันของระบบในตารางที่ 1 ชี้ให้ เห็นว่า ระบบสามารถใช้งานได้ แต่ยังมีปัญหาเรื่องส่วนต่อประสาน (User Interface) ซึ่งสอดคล้องกับข้อมูลการประเมินความพึงพอใจ

จากผู้ใช้ (ตารางที่ 3) ซึ่งอยู่ในระดับปานกลางในภาพรวม ประเด็นที่ผู้ใช้ให้คะแนนข้างต่ำกว่าข้ออื่นๆ ได้แก่ ระบบควรมีฟังก์ชั่นการใช้งานที่ครบถ้วน และระบบมีการทำงานที่ชัดเจนและเข้าใจได้ง่ายผู้ใช้จึงให้ข้อเสนอแนะอีกหลายประเด็นเช่น การเลือกวันที่ ของข้อมูลสถิติการหายใจ ผู้ใช้ไม่ทราบว่าผลลัพธ์จากการเลือกวันที่จะเป็นอย่างไร ระบบการตรวจวัดเสียงลมหายใจ หากมีการหายใจที่เบา ระบบจะไม่สามารถตรวจจับได้เลย และหน้าอัดเสียงลมหายใจในครั้งแรกหลังการกรอกประวัติ ผู้ใช้ไม่รู้ว่าจะต้องอัดเสียงลมหายใจผ่านไมโครโฟน ผู้ใช้บางคนคิดว่าให้อัดเสียงพูด หรือแค่รออยู่เฉยๆ

การทดสอบสมมติฐานในการเปรียบเทียบอัตราการหายใจจากแอปพลิเคชันกับการนับด้วยตนเองดังแสดงตารางที่ 2 ระบบยังมีความต่างกับการนับการหายใจด้วยตนเองอยู่ แต่ผลการทดสอบการเปรียบเทียบค่าเฉลี่ย [$t(18)=1.19$, $p=0.278$] ไม่พบว่ามีความแตกต่างกันอย่างมีนัยสำคัญ ผลการวิเคราะห์แสดงให้เห็นว่าระบบสามารถคำนวณอัตราการหายใจได้ไม่แตกต่างจากที่ผู้ใช้นับ ทั้งนี้จากการสังเกตส่วนใหญ่ อัตราการหายใจที่บันทึกเข้าสู่ระบบจะมีค่าน้อยกว่าที่ผู้ใช้นับด้วยตนเอง ระบบไม่สามารถตรวจจับเสียงเบาได้โดยเฉพาะเมื่อมีเสียงรบกวน ทำให้ผู้ใช้ต้องหายใจเข้าแรงขึ้นมากกว่าการหายใจออกอย่างชัดเจน จะทำให้อัตราหายใจที่บันทึกในระบบกับที่นับด้วยตนเองจึงแตกต่างกัน ภายหลังการทดสอบในเบื้องต้น ระบบได้รับการปรับปรุงตามที่ผู้ใช้เสนอแนะ

5. สรุปและอภิปรายผล

งานวิจัยนี้เสนอเกมในระบบมือถือช่วยฝึกการหายใจ (Breathing Training) ให้มีอัตราการหายใจเหมาะสมตรงตามเกณฑ์การหายใจทางการแพทย์ โดยมีการเก็บอัตราการหายใจของผู้ใช้ทั้งก่อนและหลังการฝึก สามารถเรียกดูประวัติการฝึกของผู้ใช้ได้ โดยระบบจะมีรูปแบบการแสดงผลที่เข้าใจง่าย ใช้เวลาเรียนรู้น้อย และใช้งานได้สะดวก

ถึงแม้ว่า ผลการทดสอบแอปพลิเคชันบรีธธิง เทรนนิงเบื้องต้นพบว่า เป็นระบบสามารถใช้งานได้ (usable) และผลการศึกษาเป็นไปตามสมมติฐานของการวิจัยที่ตั้งไว้ อย่างไรก็ตาม ระบบควรพัฒนาต่อไป เนื่องจากระบบจะยังคงมีปัญหาเรื่องความถูกต้องในการคำนวณอัตราการหายใจที่บันทึกในสภาพแวดล้อมที่ยังมีเสียงรบกวนและมีการเบี่ยงเบนหรือไม่เสถียรในบางครั้ง การปรับปรุงการคำนวณอัตราหายใจซึ่งอาจต้องใช้คุณสมบัติของเสียงเพิ่มเติมมาเป็นพารามิเตอร์ในการคำนวณหาอัตราการหายใจ เช่น แอมป์พลิจูด(Amplitude), ความถี่(Frequency) และ แรงดันไฟฟ้า (Voltage) เพื่อให้การคำนวณมีความถูกต้องมากขึ้น และประเด็นสำคัญที่ควรได้รับการปรับปรุงคือส่วนต่อประสานกับผู้ใช้ เช่น การเพิ่มการแนะนำการฝึกหายใจก่อนการฝึก และปรับการนำเสนอประวัติการหายใจให้เหมาะสมกับผู้ใช้ในทุกช่วงอายุด้วย

กิตติกรรมประกาศ

ผู้วิจัยขอขอบคุณพิมพกานต์ วรรณศิริกุล และคุณธนพล ขจรกลิ่นมาลา สำหรับการโปรแกรมและขอขอบคุณผู้เข้าร่วมการทดลองทุกท่าน ที่ช่วยให้งานวิจัยสำเร็จลุล่วงไปได้

เอกสารอ้างอิง

[1] "Breathing Zone," 12 October 2013. [Online]. Available: https://play.google.com/store/apps/details?id=com.breathing.zone. [Accessed 15 October 2015].

[2] "Health Through Breath – Pranayama," 2013. [Online]. Available: http://www.saagara.com/apps/breathing/health-through-breath-pranayama. [Accessed 5 October 2015].

[3] Breathing Labs, "Breathing Downhill 3," [Online]. Available: https://itunes.apple.com/us/app/breathing-downhill-3-breathing/id984910302?mt=8. [Accessed 19 May 2015]

[4] The National Center for Telehealth and Technology, "Breathe2Relax," [Online]. Available: https://itunes.apple.com/en/app/breathe2relax/id425720246?mt=8. [Accessed 8 September 2013]

[5] Softrove (Pvt.) Ltd., "BreathCounter," [Online]. Available: https://itunes.apple.com/us/app/breath-counter/id599505398?mt=8. [Accessed 8 March 2012]

[6] ประภัสสร อักษรพันธ์, "การวัดสัญญาณชีพ" [Online]. Available: http://www.publichealth.sru.ac.th/health/images/stories/pdf-word/03.pdf. [Accessed 17 October 2015].

[7] สุขจันทร์ พงษ์ประไพ, "หายใจลึกและช้าส่งผลดีต่อสุขภาพหัวใจ," [Online]. Available: http://www.thaiheart.org/images/Slow_and_deep_brealthing.pdf. [Accessed 10 October 2015].

[8] D. S. Avalur, "Human Breath Detection using Microphone". วิทยานิพนธ์. [Online]. Available: http://www.cs.rug.nl/~aiellom/tesi/avalur [Accessed 10 October 2015].

แอปพลิเคชันอำนวยความสะดวกในการใช้บริการรถโดยสารประจำทางสาธารณะ สำหรับผู้บกพร่องทางการมองเห็น
Bus Facilitation Application for the Blind

เธียรศักดิ์ พลาดิศัยเลิศ (Thearasak Phaladisailoed)[1] ธนวรรธน์ คชเสนี(Tanawat Gajaseni)[2]
และธนิศา นุ่มนนท์ (Thanisa Numnonda)[3]
คณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
[1]ph.thearasak@gmail.com, [2]tanawatgajaseni@gmail.com.com, [3]thanisa@it.kmitl.ac.th

บทคัดย่อ

ปัญหาของผู้บกพร่องทางการมองเห็น ในการใช้รถโดยสารประจำทางสาธารณะในกรุงเทพมหานคร คือไม่สามารถทราบได้ด้วยตนเองว่ารถโดยสารที่ต้องการใช้บริการนั้น จะมาถึงจุดที่ตนเองรออยู่เมื่อใด และเมื่ออยู่บนรถโดยสารก็ไม่สามารถทราบได้ว่าในขณะนั้นตนเองอยู่ที่ไหน และจะถึงป้ายหยุดรถที่ตนเองต้องการลงเมื่อใด

โครงการนี้จึงคิดพัฒนาแอปพลิเคชันบนมือถือ เพื่อให้ผู้บกพร่องทางการมองเห็น รวมไปถึงบุคคลทั่วไป ที่ต้องการใช้บริการรถโดยสารประจำทางสาธารณะ สามารถเดินทางได้สะดวกขึ้น โดยใช้ข้อมูลจาก Traffy API ของ NECTEC ซึ่งมีการใช้ข้อมูลป้ายหยุดรถทั้งหมด 5685 ป้าย และข้อมูลรถโดยสารทั้งหมด 543 เส้นทาง ช่วยในการอำนวยความสะดวกในการใช้บริการรถโดยสารประจำทางสาธารณะ นอกจากนี้ยังได้มีการออกแบบให้สามารถใช้งานผ่านซอฟต์แวร์ TalkBack และ Text-to-Speech ของระบบปฏิบัติการแอนดรอยด์ เพื่อให้ผู้บกพร่องทางการมองเห็นสามารถใช้งานได้สะดวกมากยิ่งขึ้น

คำสำคัญ: ผู้บกพร่องทางการมองเห็น รถโดยสารประจำทางสาธารณะ แอนดรอยด์แอปพลิเคชัน

Abstract

Blind people cannot know by themselves that the bus they are waiting arrives. In addition, they do not know where they are and when to reach the desired bus stop while they are on the bus. These are main problems for them.

Therefore, this project aims to develop mobile application for blind and normal people to use public bus transportation more convenience by using NECTEC's Traffy API. Information of 5,685 bus stops and 543 bus routes are provided to facilitate travelling by bus. TalkBack and Text-to-Speech applications on Android are also included to help blind people to use this mobile application easier.

Keyword: Blind, Public Bus, Android Application

1. บทนำ

ในปัจจุบันมีผู้บกพร่องทางการมองเห็นในกรุงเทพมหานครเป็นจำนวน 3,992 คน [1] ซึ่งผู้บกพร่องทางการมองเห็นส่วนใหญ่ประสบปัญหาในการเดินทาง [2] และจำเป็นต้องอาศัยบริการรถโดยสารสาธารณะในการเดินทาง [3] ปัจจุบันระบบขนส่งมวลชน ที่มีการอำนวยความสะดวกให้แก่บุคคลที่มีความบกพร่องทางการมองเห็นมากที่สุด คือการเดินทางโดยรถไฟฟ้า ซึ่งก็ยังไม่ครอบคลุมพื้นที่ต่าง ๆ ระบบขนส่งที่มีความครอบคลุมมากที่สุดในปัจจุบัน มีเพียงรถโดยสารประจำทางสาธารณะเท่านั้น แต่ก็ไม่เอื้ออำนวยต่อการใช้งานเท่าที่ควร [4] เนื่องจากผู้บกพร่องทางการมองเห็นจะไม่สามารถทราบได้ด้วยตนเองว่า รถโดยสารประจำทางสาธารณะที่ตนเองต้องการใช้บริการจะมาถึงเมื่อใด จึงต้องอาศัยความช่วยเหลือจากบุคคลอื่น และถ้าหากไม่มีคนช่วยเหลือผู้บกพร่องทางการมองเห็น ก็จะไม่สามารถทราบว่า รถที่ตนต้องการโดยสารนั้นมาถึงป้ายหยุดรถเมื่อใด และเมื่ออยู่บนรถโดยสารประจำทางสาธารณะก็ไม่สามารถทราบตำแหน่งปัจจุบันของตนเองได้ จึงเกิดปัญหาความสับสนในการเดินทางตามมา [5]

จากที่กล่าวมาข้างต้น จึงได้มีการพัฒนาระบบอำนวยความสะดวกในการใช้บริการรถโดยสารประจำทางสาธารณะสำหรับ

ผู้บกพร่องทางการมองเห็น เพื่อให้ผู้บกพร่องทางการมองเห็นสามารถใช้บริการรถโดยสารประจำทางสาธารณะได้ด้วยตนเอง ซึ่งแอปพลิเคชันของระบบนั้น จะมีการแจ้งข้อมูลป้ายหยุดรถโดยสารประจำทางสาธารณะที่อยู่ใกล้ผู้ใช้ที่สุด หมายเลขของรถโดยสารประจำทางสาธารณะที่ผ่าน แจ้งเตือนผู้ใช้เมื่อรถโดยสารประจำทางสาธารณะเข้าใกล้ป้ายหยุดรถที่ผู้ใช้อยู่ และแจ้งชื่อของป้ายหยุดรถโดยสารประจำทางสาธารณะถัดไป ที่จะถึง ให้ผู้ใช้ทราบเพื่อให้ผู้ใช้ทราบตำแหน่งปัจจุบันของตน และมีการเตรียมตัวล่วงหน้า เมื่อถึงป้ายหยุดรถโดยสารสาธารณะที่ตนเองต้องการจะลง โดยมีการพัฒนาให้สามารถใช้งานผ่าน TalkBack และมีการใช้ Text-to-Speech และ Voice Recognition เพื่อให้ผู้บกพร่องทางการมองเห็นสามารถใช้งานได้โดยสะดวก [6]

2. วรรณกรรมที่เกี่ยวข้อง

2.1 Java Script Object Notation (JSON)

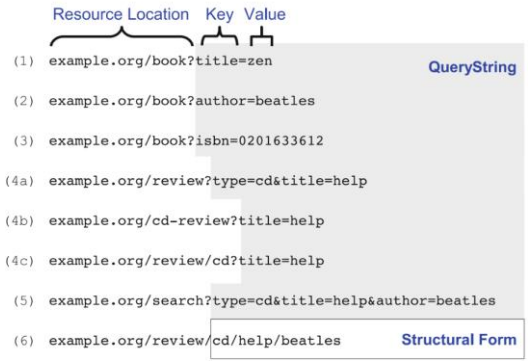
JSON เป็นรูปแบบการแลกเปลี่ยนข้อมูลที่มีโครงสร้างที่เข้าใจง่าย ไม่ซับซ้อน ง่ายต่อการทำความเข้าใจและการสร้าง ทั้งโดยมนุษย์และคอมพิวเตอร์ โดยรูปแบบการเก็บจะมีการเก็บข้อมูลในรูปของ Key และ Value ซึ่งการแลกเปลี่ยนข้อมูลที่อยู่ในรูปแบบ JSON สามารถแลกเปลี่ยนข้อมูลกันได้สะดวก และมีประสิทธิภาพกว่าการแลกเปลี่ยนข้อมูลด้วย XML [7] ซึ่งรูปแบบของ JSON แสดงอยู่ในภาพที่ 1

```
{
        "firstname" : "John",
        "lastname" : "Smith"
}
```

ภาพที่ 1: ตัวอย่างรูปแบบข้อมูลประเภท JSON [7]

2.2 Representational state transfer (REST)

REST เป็นรูปแบบการสร้าง Web service ที่ทำงานและเรียกใช้งานได้ผ่านทาง HTTP Protocol โดยที่ REST นั้นจะใช้การแลกเปลี่ยนข้อมูลกันในรูปของ JSON หรือ XML จึงทำให้มีปริมาณข้อมูลที่รับส่งไม่มากนัก โดยการเรียกใช้งาน Web Service แบบ REST นั้นเราจะเข้าถึง Resource หรือ Service ต่าง ๆ โดยระบุด้วย URI และส่งค่า Parameter ไปให้ผู้ให้บริการ Web Services นั้น ๆ และสามารถใช้ Method ของ HTTP Protocol ในการจัดการกับ Resource หรือ Service ได้ เช่น Create, Read, Update หรือ Delete สามารถดำเนินการผ่านการเรียกใช้งาน HTTP Methods (POST, GET, PUT หรือ DELETE) [8] ในการเรียกดูข้อมูลผ่าน URIs ในลักษณะต่าง ๆ ทั้งแบบกำหนด key value ของข้อมูล และแบบ Structural Form คือระบุแค่ values เท่านั้น [8] ดังแสดงในภาพที่ 2



ภาพที่ 2: ตัวอย่างการใช้งาน REST

2.3 Traffy API

Traffy เป็นส่วนหนึ่งของห้องปฏิบัติการระบบขนส่งและจราจรอัจฉริยะ ซึ่งเป็นห้องวิจัยที่ทำหน้าที่ในการพัฒนาส่งเสริมงานวิจัยในด้านการขนส่ง เดินทาง และจราจร โดย Traffy ได้ให้บริการข้อมูลของรถโดยสารประจำทางสาธารณะในด้านต่าง ๆ ไม่ว่าจะเป็น ต้นสาย-ปลายสาย ป้ายที่รถโดยสารประจำทางสาธารณะสายนั้นผ่าน รวมถึงข้อมูลของป้ายหยุดรถโดยสารประจำทางสาธารณะผ่านทาง Traffy API ซึ่งใน API ก็จะเปิดให้ร้องขอข้อมูลของรถโดยสารประจำทางสาธารณะและป้ายหยุดรถได้ [9]

โดยลักษณะของ Traffy API นั้นจะใช้การแลกเปลี่ยนข้อมูลในรูปแบบ REST API ซึ่งมีการกำหนดวิธีการเรียกดูข้อมูลต่าง ๆ ไว้ในเว็บไซต์ของ Traffy โดยการดูข้อมูลแต่ละส่วนก็มีวิธีการกำหนด Key และ Value ของ URIs ต่างกัน เช่น การดูข้อมูลรถโดยสารประจำทางสาธารณะตามสายที่สนใจ พร้อมข้อมูลป้ายหยุดรถก่อนหน้า และป้ายหยุดรถถัดไป จะเรียกผ่าน http://cloud.traffy.in.th/apis/apitraffy.php?api=get_bus_location&appid=51248f9a&bus_line=73ก โดยที่ค่า bus_line นั้นคือสายรถโดยสารประจำทางสาธารณะ จากตัวอย่างเป็นการร้องขอข้อมูลตำแหน่งของรถโดยสารประจำทางสาธารณะสาย 73ก โดยผู้ใช้จะได้รับข้อมูลตอบกลับในรูปแบบของ JSON ดังแสดงในภาพที่ 3

```
{
    "vehicle_id": "5024130032",
    "bus_line": "73 ก.",
    "speed": 13,
    "heading": 162.9,
    "latitude": 13.749408333333,
    "longitude": 100.54087333333,
    "direction": "inbound",
    "linear_ref": "0.821476171587294",
    "next_stop_id": {⟺},
    "prev_stop_id": {⟺}
},
```


ภาพที่ 3: ตัวอย่างข้อมูล JSON ที่ได้รับจาก Traffy API

2.4 Google Map Direction API

Google Map Direction API คือบริการหนึ่งของ Google ภายใต้ Google Map ที่จะช่วยในการนำทาง การคำนวณระยะเวลา และระยะทางระหว่างจุดต่าง ๆ ในแผนที่ โดยสามารถเลือกรูปแบบการเดินทางได้หลากหลายรูปแบบ และผู้ใช้สามารถที่จะกำหนดค่าต่าง ๆ เพื่อใช้ในการคำนวณการเดินทางได้

ตัวอย่างการใช้งาน Google Map Direction API เพื่อคำนวณการเดินทางจาก Boston ไปยัง Concord โดยผ่าน Charlestown และ Lexington โดยเรียกผ่าน maps.googleapis.com/maps/api/directions/json?origin=Boston,MA&destination=Concord,MA&waypoints= Charlestown,MA|Lexington,MA&key= YOUR_API_KEY ซึ่งเราต้องกำหนดต้นทางไว้ใน origin กำหนดปลายทางไว้ใน destination กำหนดเส้นทางในการเดินทางผ่าน waypoints โดยแต่ละจุดจะคั่นด้วยเครื่องหมาย “ | ” และสุดท้ายจะต้องใส่ Key ไว้ที่ YOUR_API_KEY ซึ่งเราสามารถใช้ Gmail ของเราสร้าง API Key โดยเข้าไปที่ หน้า Google APIs Console [10] โดยตัวอย่างผลลัพธ์จาก API จะแสดงในภาพที่ 4


```
"legs" : [
   {
      "distance" : {
         "text" : "35.9 mi",
         "value" : 57824
      },
      "duration" : {
         "text" : "51 mins",
         "value" : 3062
      },
      "end_address" : "Universal Studios Blvd, Los Angeles, CA 90068, USA",
      "end_location" : {
         "lat" : 34.1330949,
         "lng" : -118.3524442
      },
      "start_address" : "Disneyland (Harbor Blvd.), S Harbor Blvd, Anaheim, CA 92802
      "start_location" : {
         "lat" : 33.8098177,
         "lng" : -117.9154353
      },
```


ภาพที่ 4: ตัวอย่างข้อมูล JSON ที่ได้รับจาก Google Direction API

2.5 งานวิจัยที่เกี่ยวข้อง

MZH Noor และคณะ [11] ได้ใช้เทคโนโลยี RFID ในการระบุข้อมูลรถโดยสารประจำทางสาธารณะ โดยติด Tag RFID ไว้ที่รถโดยสารประจำทางสาธารณะ และในฝั่งของผู้บกพร่องทางการมองเห็นจะมี RFID Reader ซึ่งประกอบด้วย Computer และ เสาอากาศสำหรับการระบุข้อมูลรถโดยสาร ซึ่งเทคโนโลยี RFID นั้น Tag ของ RFID แต่ละอันจะมีหมายเลขไม่เหมือนกัน ทำให้สามารถใส่ข้อมูลแบบเฉพาะเจาะจง สำหรับรถโดยสารประจำทางสาธารณะแต่ละคันได้ แต่ RFID นั้นมีข้อจำกัดที่ระยะในการรับส่งข้อมูลที่น้อยกว่าการใช้ GPS Quoc และคณะ [4] ได้ใช้ Wireless Sensor Network (WSN) ในการระบุหมายเลขรถ และแจ้งข้อมูลให้แก่ผู้บกพร่องทางการมองเห็นโดยติดตั้งอุปกรณ์ WSN ไว้บนรถ และบริเวณป้ายหยุดรถจะถูกแบ่งเป็นพื้นที่ของผู้ใช้บริการทั่วไป และผู้บกพร่องทางการมองเห็น ซึ่งบริเวณพื้นที่ของผู้บกพร่องจะมีการติดตั้งอุปกรณ์เพื่อตรวจสอบว่ามีผู้บกพร่องทางการมองเห็นอยู่หรือไม่ ถ้ามีระบบจะทำการส่งข้อมูลไปยังอุปกรณ์บนรถโดยสารประจำทางสาธารณะ และเมื่อรถโดยสารอยู่ในระยะ 2 เมตรจากป้ายหยุดรถก็จะมีการส่งเสียงประกาศแจ้งหมายเลขรถ ให้ผู้บกพร่องทางการมองเห็นทราบ ซึ่งระบบนี้สามารถแจ้งหมายเลขรถที่หยุดอยู่บริเวณป้ายหยุดรถของผู้บกพร่องทางการมองเห็นได้ จึงช่วยให้ผู้บกพร่องทางการมองเห็น สามารถใช้บริการรถโดยสารประจำทางสาธารณะในสายที่ตนเองต้องการได้ ซึ่งผู้บกพร่องทางการมองเห็นจะต้องทราบเส้นทาง ของรถโดยสารสายนั้นอยู่แล้ว เนื่องจากระบบไม่สามารถแจ้งรายละเอียดอื่นนอกจากหมายเลขรถได้ และการแจ้งหมายเลขรถจะแจ้งเฉพาะ รถที่จอดด้านหน้าป้ายของผู้บกพร่องทางการมองเห็นเท่านั้น Lavanya และคณะ [2] ได้ใช้ ZigBee, WSN, และ Microcontroller โดยติดตั้งทั้งฝั่ง ผู้ใช้ และรถโดยสารประจำทางสาธารณะซึ่งอุปกรณ์ที่อยู่บนรถ จะทำหน้าที่เก็บข้อมูลตำแหน่งของรถโดยสารประจำทางสาธารณะที่ได้จาก GPS เพื่อใช้ในการเปรียบเทียบกับข้อมูลจุดหมายที่ผู้ใช้ต้องการ ในส่วนผู้ใช้จะมีการรับข้อมูลสถานที่ที่ผู้ใช้ต้องการจากไมโครโฟน มีการทำ Voice Recognition เพื่อหาตำแหน่งของสถานที่นั้น และคำนวณหารถโดยสารประจำทางสาธารณะที่ผ่าน เพื่อแจ้งหมายเลขรถ ให้ผู้บกพร่องทางการมองเห็นทราบผ่านทางหูฟัง และส่งข้อมูลไปให้รถโดยสารประจำทางสาธารณะทราบว่า มีผู้บกพร่องทางการมองเห็นต้องการใช้บริการ เมื่อรถมาถึงป้ายที่ผู้ใช้รอก็จะมีการแจ้งเตือนผู้ใช้ด้วยเสียงอีกครั้ง ซึ่งระบบนี้มี

การอำนวยความสะดวกในการใช้งานให้แก่ผู้ใช้ เนื่องจากมีการรับและแจ้งข้อมูลโดยใช้เสียงเป็นหลัก แต่ขนาดของอุปกรณ์ในฝั่งของผู้ใช้หรือผู้บกพร่องทางการมองเห็นนั้น มีขนาดที่ไม่สะดวกต่อการพกพา เมื่อเทียบกับโทรศัพท์มือถือสมาร์ทโฟน Flores และคณะ [5] ได้ใช้วิธีการติดตั้ง Access Point (AP) ไว้ที่รถโดยสารและป้ายหยุดรถ โดยที่ AP ที่ติดตั้งบนรถนั้นจะมีข้อมูลที่เกี่ยวข้องกับรถโดยสารประจำทางสาธารณะคันนั้นอยู่ เช่น ข้อมูลหมายเลขรถ ข้อมูลป้ายหยุดรถทั้งหมด ข้อมูลป้ายหยุดรถที่ผ่านมาแล้ว เพื่อให้ผู้ใช้ทราบว่าถึงจุดหมายที่ต้องตนเองต้องการหรือไม่ โดยการใช้งานของผู้ใช้นั้นจะอาศัยโทรศัพท์มือถือของผู้ใช้เป็นอุปกรณ์ในการรับข้อมูลจาก AP ที่อยู่บนรถ ในการแจ้งเมื่อรถเข้าใกล้ป้ายหยุดรถที่ผู้ใช้อยู่ หรือในกรณีที่ต้องการทราบข้อมูลป้ายหยุดรถที่รถจะเข้าจอด เมื่ออยู่บนรถโดยสารประจำทางสาธารณะแล้ว ซึ่งการใช้งานระบบนี้ไม่จำเป็นที่จะต้องมีการเชื่อมต่อเครือข่ายอินเตอร์เน็ต แต่จำเป็นต้องติดตั้ง AP ไว้ทั้งในตัวรถและที่ป้ายหยุดรถ จึงทำให้มีค่าใช้จ่ายที่สูงกว่าการใช้ RFID [11] และ ZigBee [2] อีกทั้งระยะห่างในการติดต่อสื่อสารระหว่าง AP ที่ติดตั้งไว้ที่รถโดยสารและป้ายหยุดรถจะต้องไม่เกิน 211 ft. หรือ ประมาณ 64.34 เมตร ซึ่งอาจไม่เพียงพอต่อผู้ใช้ ซึ่งเป็นผู้บกพร่องทางการมองเห็นในการเตรียมตัวเพื่อลงจากรถ Korbel และคณะ [12] ได้พัฒนาแอนดรอยด์แอปพลิเคชัน เพื่อช่วยเหลือผู้บกพร่องทางการมองเห็น ในการใช้บริการรถโดยสารสาธารณะ ซึ่งจะมีข้อมูลตำแหน่งของป้ายหยุดรถต่าง ๆ เก็บไว้ในฐานข้อมูลซึ่งเป็น MySQL Server โดยเข้าถึงข้อมูลเหล่านั้นได้จากแอปพลิเคชันบนโทรศัพท์มือถือ และยังมีระบบนำทางไปยังป้ายหยุดรถนั้นด้วย เมื่อไปถึงป้ายหยุดรถจะสามารถจะรับข้อมูลของรถโดยสารผ่าน NFC Tag ที่ติดตั้งตามป้ายหยุดรถต่าง ๆ หากผู้ใช้ต้องการข้อมูลจึงจำเป็นที่จะต้องไปยังป้ายหยุดรถนั้น ๆ ก่อน หรือสามารถรับข้อมูลผ่าน Unstructured Supplementary Service Data (USSD) โดยใช้โทรศัพท์มือถือกดหมายเลขเฉพาะของป้ายหยุดรถ เพื่อให้ระบบส่งข้อมูลตารางเวลาของรถโดยสารมายังโทรศัพท์มือถือได้ ซึ่งผู้ใช้จำเป็นต้องรู้หมายเลขเฉพาะของแต่ละป้าย จึงสามารถดูข้อมูลของป้ายนั้นได้ รวมไปถึงการรับข้อมูลตำแหน่งต่าง ๆ ของป้ายหยุดรถจำเป็นต้อง มีการเชื่อมต่อเครือข่ายอินเตอร์เน็ต เพื่อร้องขอข้อมูลต่าง ๆ จาก Server เมื่อผู้ใช้ไม่ได้เชื่อมต่อกับเครือข่ายอินเตอร์เน็ตก็จะไม่สามารถดูข้อมูลเหล่านั้นได้

3. วิธีการดำเนินการวิจัย

3.1 การวิเคราะห์ความต้องการ

- สามารถแจ้งชื่อป้ายหยุดรถโดยสาร และระยะห่างจากตำแหน่งปัจจุบันของผู้ใช้
- สามารถแจ้งข้อมูลสายรถโดยสารประจำทางสาธารณะที่ผ่านป้ายหยุดรถประจำทางที่ผู้ใช้อยู่
- สามารถแจ้งเตือน เมื่อรถโดยสารประจำทางสาธารณะที่เลือกไว้จะเข้าป้ายที่ผู้ใช้อยู่ได้
- สามารถแจ้งชื่อป้ายหยุดรถโดยสารถัดไป ที่รถโดยสารประจำทางสาธารณะจะเข้าจอดได้
- สามารถแสดงความคิดเห็นต่อการใช้บริการรถโดยสารประจำทางสาธารณะได้
- สามารถแจ้งสภาพจราจรจากเส้นทางของสายรถที่ผู้ใช้เลือกได้
- ค้นหาข้อมูลของป้ายหยุดรถและรถโดยสารประจำทางสาธารณะได้

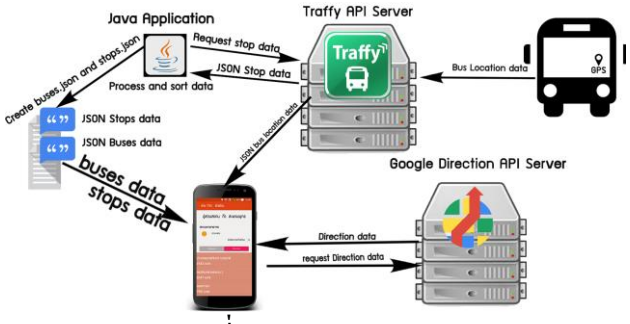
โดยผู้พัฒนาได้ทำการสอบถามความต้องการโดยผู้ใช้จากสมาคมคนตาบอดแห่งประเทศไทย

3.2 การออกแบบและพัฒนาระบบ

ในด้านการออกแบบส่วนติดต่อผู้ใช้ ในปัจจุบันผู้บกพร่องทางการมองเห็นสามารถใช้โทรศัพท์มือถือ ผ่านฟังก์ชันอ่านหน้าจอ โดยในระบบปฏิบัติการแอนดรอยด์จะเรียกฟังก์ชันอ่านหน้าจอว่า TalkBack ซึ่งการออกแบบให้ผู้บกพร่องทางการมองเห็นสามารถใช้งานแอปพลิเคชันได้นั้น จะเน้นไปที่การทำให้ TalkBack สามารถเข้าถึงหรืออ่านส่วนนั้น ๆ ได้ โดยปกติการใช้วิดเจ็ตพื้นฐาน ที่แอนดรอยด์เตรียมไว้ให้ ก็สามารถช่วยให้ TalkBack เข้าถึงได้ [12] รวมถึงการเพิ่มคำอธิบายในส่วนที่เป็นรูปภาพ หรือส่วนอื่นเพื่อขยายความให้ผู้บกพร่องทางการมองเห็นเข้าใจ การใช้งานวิดเจ็ตต่าง ๆ

ในส่วนโครงสร้างของแอปพลิชันมีการพัฒนา โดยแบ่งออกเป็น 2 ส่วนคือ ส่วนการสร้างข้อมูลซึ่งพัฒนาในรูปแบบ

ของ Java Application สำหรับใช้ในการดึงข้อมูลป้ายหยุดรถทั้งหมดจาก Traffy API เพื่อสร้างเป็นไฟล์ข้อมูลป้ายหยุดรถและข้อมูลสายรถในรูปแบบ JSON ไฟล์สำหรับใช้ในแอปพลิเคชัน และส่วนของแอปพลิเคชัน ซึ่งผู้พัฒนาได้ออกแบบให้มีการเรียกดูข้อมูล โดยใช้การเชื่อมต่ออินเตอร์เน็ตน้อยที่สุด จึงพัฒนาให้มีการร้องขอข้อมูลโดยใช้อินเตอร์เน็ต เพียงแค่ข้อมูลตำแหน่งของรถโดยสารประจำทาง และข้อมูลสภาพการจราจรโดยเรียกผ่าน Google Direction API ส่วนข้อมูลอื่น ๆ จะถูกเก็บไว้ในตัวแอปพลิเคชันในรูปแบบ JSON ตาม ภาพที่ 5



ภาพที่ 5: ภาพโครงสร้างระบบ

ในด้านส่วนติดต่อกับผู้ใช้ ได้มีการนำเอาเทคโนโลยีการแปลงเสียงพูดเป็นข้อความโดยใช้ Voice Recognition [13] ในแอนดรอยด์ ช่วยให้ผู้บกพร่องทางการมองเห็น สามารถใช้งานได้โดยสะดวกมากขึ้น [6]

4. ผลการดำเนินงาน

4.1 ผลการทดลอง

จากการสอบถามผลการใช้งาน จากผู้บกพร่องทางการมองเห็นจำนวน 5 คน โดยให้ตอบคะแนนความพึงพอใจตั้งแต่ 1 ถึง 5 ในหัวข้อต่าง ๆ ตามที่ได้ใช้งาน ผลการทดสอบได้เปอร์เซ็นความพึงพอใจในภาพรวมโดยเฉลี่ยอยู่ที่ 72.4 % โดยแสดงรายละเอียดในตารางที่ 1 และพบว่าผู้ใช้มีปัญหา และข้อเสนอแนะต่างๆดังนี้

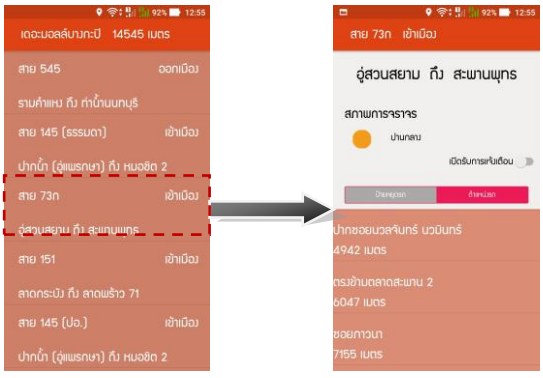
- ข้อมูลตำแหน่งของรถโดยสารสาธารณะยังมีน้อย
- ข้อมูลรถโดยสารสาธารณะไม่ครบถ้วนสมบูรณ์
- ข้อมูลป้ายหยุดรถที่มีไม่อัพเดท
- มีปัญหาความคลาดเคลื่อนในข้อมูลตำแหน่งของผู้ใช้
- การแจ้งเตือนมีความล่าช้า
-

ตารางที่ 1: ระดับความพึงพอใจของผู้ใช้ต่อฟังก์ชันต่าง ๆ

ชื่อฟังก์ชัน	1	2	3	4	5
แจ้งชื่อป้ายหยุดรถโดยสาร และระยะห่างจากตำแหน่งปัจจุบันของผู้ใช้			2	2	1
แจ้งข้อมูลสายรถโดยสารประจำทางสาธารณะที่ผ่านป้ายหยุดรถประจำทางที่ผู้ใช้อยู่			1	3	1
แจ้งเตือนเมื่อ รถโดยสารประจำทางสาธารณะที่เลือกไว้จะเข้าป้ายที่ผู้ใช้อยู่			4		1
แจ้งชื่อป้ายหยุดรถโดยสารถัดไป ที่รถโดยสารประจำทางสาธารณะจะเข้าจอด		1	1		3
แสดงความคิดเห็นต่อการใช้บริการรถโดยสารประจำทางสาธารณะได้			4		1
แจ้งสภาพจราจรจากเส้นทางของสายรถที่ผู้ใช้เลือกได้			3	1	1
ค้นหาข้อมูลของป้ายหยุดรถและรถโดยสารประจำทางสาธารณะได้	1			2	2
ความพึงพอใจต่อการเข้าถึงของแอปพลิเคชัน			2	1	2
ความพึงพอใจโดยรวมต่อแอปพลิเคชัน			3	1	1

4.2 ตัวอย่างหน้าจอแอปพลิเคชัน

ภาพที่ 7 (ก) แสดงข้อมูลป้ายหยุดรถใกล้เคียง โดยแสดงชื่อป้ายและระยะห่างระหว่างป้ายกับผู้ใช้ รวมถึงข้อมูลรถโดยสารประจำทางที่ผ่านป้าย ซึ่งผู้ใช้สามารถกดที่ข้อมูลรถ เพื่อไปยังหน้าแสดงข้อมูลรถโดยสาร เพื่อแสดงข้อมูลรถโดยสาร สภาพการจราจร แสดงข้อมูลป้ายหยุดรถทั้งหมดที่รถสายนั้นผ่าน และแสดงข้อมูลตำแหน่งรถโดยสารหากมีข้อมูล ดังแสดงในภาพที่ 7 (ข)



(ก) (ข)

ภาพที่ 7: หน้าจอป้ายหยุดรถใกล้เคียง และหน้าจอแสดงข้อมูลรถโดยสารประจำทางสาธารณะ

หน้าจอแสดงข้อมูลป้ายหยุดรถถัดไป แสดงข้อมูลป้ายหยุดรถและระยะห่างจากผู้ใช้ โดยสามารถแจ้งเตือนผู้ใช้เมื่อมีการเปิดการแจ้งเตือน



ภาพที่ 8: หน้าจอแจ้งป้ายหยุดรถถัดไป

5. สรุป

ในปัจจุบันผู้บกพร่องทางการมองเห็น ยังไม่ได้รับความสะดวกในการเดินทางด้วยรถโดยสารประจำทางสาธารณะเท่าที่ควร เนื่องจากไม่สามารถทราบข้อมูลที่จำเป็นต่อการใช้บริการได้ จึงเกิดความสับสนในการเดินทางตามมา อีกทั้งในปัจจุบันได้มีการพัฒนา Traffy API เพื่อให้ข้อมูลป้ายหยุดรถและข้อมูลรถโดยสารประจำทางสาธารณะบางส่วน

ดังนั้นผู้พัฒนาจึงใช้ข้อมูลจาก Traffy API พัฒนาแอปพลิเคชันอำนวยความสะดวก ในการใช้บริการรถโดยสารสาธารณะสำหรับผู้บกพร่องทางการมองเห็นขึ้น เพื่อแจ้งข้อมูลที่จำเป็นในการใช้บริการรถโดยสารประจำทางสาธารณะ เช่น ข้อมูลตำแหน่งรถ เพื่อแจ้งเตือนเมื่อรถโดยสารประจำทางสาธารณะเข้าใกล้ป้าย ตำแหน่งป้ายหยุดรถ เพื่อแจ้งข้อมูลป้ายหยุดรถถัดไป และสามารถค้นหาข้อมูลป้ายหรือรถโดยสารประจำทางสาธารณะ

จากการทดลองใช้งานพบว่า ผู้บกพร่องทางการมองเห็นมีเปอร์เซ็นความพึงพอใจในภาพรวมต่อแอปพลิเคชันโดยเฉลี่ยอยู่ที่ 72.4 % ซึ่งปัญหาของแอปพลิเคชันในปัจจุบันคือ การโหลดข้อมูลเมื่อเปิดแอปพลิเคชันใช้เวลานาน และการแจ้งเตือนซึ่งยังมีข้อผิดพลาดอยู่ในบางจุด ในอนาคตจะต้องมีการพัฒนาเพิ่มเติม ในเรื่องการปรับปรุงข้อมูล การพัฒนาแอปพลิเคชันบนแพลตฟอร์มอื่น ตลอดจนการพัฒนาแอปพลิเคชันให้รองรับ API ในอนาคต

เอกสารอ้างอิง

[1] กรมส่งเสริมและคุณภาพชีวิตคนพิการ, “สถิติข้อมูลคนพิการ”, ข้อมูลประมวลผลจากฐานข้อมูลทะเบียนกลางคนพิการ กรมส่งเสริมและพัฒนาคุณภาพชีวิตคนพิการ, 1 เมษายน 2560

[2] Lavanya, G., et al. "Passenger bus alert system for easy navigation of blind." *Circuits, Power and Computing Technologies (ICCPCT), 2013 International Conference on*. IEEE, 2013.

[3] Azenkot, Shiri, et al. "Enhancing independence and safety for blind and deaf-blind public transit riders." *Proceedings of the SIGCHI conference on Human Factors in computing systems*. ACM, 2011.

[4] Quoc, Trung Pham, et al. "Wireless sensor network apply for the blind U-bus system." *International Journal of u- and e-Service, Science and Technology* 3.3 (2010): 13-24. *Knowledge-Based Systems (IJUFKS)*, Vol. 11, No. 4, pp. 445-466, 2003.

[5] Flores, German, et al. "Transit information access for persons with visual or cognitive impairments." *International Conference on Computers for Handicapped Persons*. Springer International Publishing, 2014.

[6] GRUSSENMEYER, WILLIAM, and EELKE FOLMER. "Accessible Touchscreen Technology for People with Visual Impairments: A Survey." (2017).

[7] Nurseitov, Nurzhan, et al. "Comparison of JSON and XML Data Interchange Formats: A Case Study." *Caine* 2009 (2009): 157-162.

[8] Adamczyk, Paul, et al. "Rest and web services: In theory and in practice." *REST: from research to practice*. Springer New York, 2011. 35-57.

[9] ห้องปฏิบัติการขนส่งและจราจรอัจฉริยะ, “โครงการ Traffy Transit”, [ออนไลน์]. แหล่งที่มา: https://its.nectec.or.th /2016/08/10/โครงการ-traffy-transit , เข้าถึงเมื่อ: กุมภาพันธ์ 2017.

[10] Google developer, “Direction API”, [ออนไลน์]. แหล่งที่มา: https://developers.google.com/maps/documentation/directions/start, เข้าถึงเมื่อ: กุมภาพันธ์ 2017.

[11] Noor, Mohd Zikrul Hakim, Ismarani Ismail, and Mohammad Farid Saaid. "Bus detection device for the blind using RFID application." *Signal Processing & Its Applications, 2009. CSPA 2009. 5th International Colloquium on*. IEEE, 2009.

[12] Korbel, Piotr, et al. "Mobile applications aiding the visually impaired in travelling with public transport." *Computer Science and Information Systems (FedCSIS), 2013 Federated Conference on*. IEEE, 2013.

[13] Reddy, B. Raghavendhar, and E. Mahender. "Speech to text conversion using android platform." *International Journal of Engineering Research and Applications (IJERA)* 3.1 (2013): 2538.

การออกแบบและพัฒนาอุปกรณ์เสริมไม้เท้าช่วยเหลือผู้พิการทางสายตา
Additional Function Design and Development of Walking Stick for Supporting a Blind Man

ปัญจรัตน์ หาญพานิช[1] และสุชาติ รมณียารักษ์[2]
[1]สาขาวิชาเทคโนโลยีสารสนเทศ คณะวิทยาศาสตร์และเทคโนโลยี วิทยาลัยเซาธ์อีสท์บางกอก
[2]สาขาวิชาเทคโนโลยีสารสนเทศ คณะวิทยาศาสตร์และเทคโนโลยี วิทยาลัยเซาธ์อีสท์บางกอก
panjarat@hotmail.com, suchart@southeast.ac.th

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์ที่จะออกแบบและพัฒนาอุปกรณ์เสริมไม้เท้าช่วยเหลือผู้พิการทางสายตา ให้สามารถดำรงชีวิตประจำวันเหมือนกับผู้ที่มีสายตาปกติ โดยการนำเอาเซ็นเซอร์อัลตร้าโซนิค โมดูลมาประยุกต์รวมเข้ากับ บอร์ดอาดูโน นำมาติดกับไม้เท้าสำหรับผู้พิการ เพื่อช่วยในการนำทาง เมื่ออุปกรณ์เซ็นเซอร์อัลตร้าโซนิค โมดูลกระทบกับวัตถุสิ่งของต่างๆ ที่อยู่ด้านหน้า และด้านบนศีรษะ ก็จะส่งคลื่นเสียงออกไป เมื่อประมวลผลระยะทางของคลื่นเสียงของวัตถุที่มากระทบ และสามารถแสดงค่าออกมาเป็นเสียง และการสั่น โดยงานวิจัยนี้ได้นำไปให้ผู้พิการทางสายตา ได้ทำการทดสอบกับอุปกรณ์

ผลการวิจัยพบว่า อุปกรณ์เสริมไม้เท้าช่วยเหลือผู้พิการทางสายตา สามารถเพิ่มความมั่นใจให้ผู้พิการทางสายตาได้ในระดับนึง และทำให้ไม่เกิดอันตรายต่อผู้พิการทางสายตา

คำสำคัญ: ไม้เท้าผู้พิการ, อัลตร้าโซนิค , บอร์ดอาดูโน

Abstract

The Objectives of this study are to design and develop the additional function for walking stick to support the blinds which can facilitate them like normal people. By applying an ultrasonic sensor with an Arduino board then sticked with a cane. If the sensor collides with overhead or front object, the system will evaluate data. Then, the response will give on alarm or shake. This study is already tested with the blinds.

The result showed that, the walking stick with the additional function can gain confidence to the blinds and also safe.

Keyword: Walking Stick, Ultrasonic Sensor, Arduino Board

1. บทนำ

ในปัจจุบันเทคโนโลยีได้เข้ามามีบทบาทในการดำรงชีวิตประจำวันมากขึ้น สำหรับผู้ที่มีความบกพร่องทางสายตา การนำเทคโนโลยีมาใช้จึงมีความสำคัญที่จะเข้ามาช่วยให้การดำเนินชีวิตสำหรับผู้พิการทางสายตา ให้สามารถดำรงชีวิตได้อย่างเป็นปกติและมีความปลอดภัยมากยิ่งขึ้น [5]

โดยเฉพาะผู้พิการทางสายตา การใช้ชีวิตประจำวันค่อนข้างยาก อันเนื่องจากการเดินทางไปยังสถานที่ต่างๆ ยังคงมีความยากลำบากอยู่ เนื่องจากสภาพพื้นที่ และพื้นผิวถนน ผู้พิการทางสายตาส่วนใหญ่จะใช้ไม้เท้าในการนำทาง ผู้วิจัยจึงเล็งเห็นความสำคัญในเรื่องนี้จึงได้ศึกษาถึงอุปกรณ์ต่างๆ ที่จะสามารถช่วยเหลือผู้พิการทางสายตาเหล่านี้ ซึ่งผู้วิจัยได้ศึกษาและพบว่ามีการพัฒนาไม้เท้าให้มีล้อหมุน ซึ่งจะมีเซ็นเซอร์เพื่อเตือนในการเดินที่มีพื้นที่ต่างระดับ เช่น บันได รวมถึงการมีอุปกรณ์เซ็นเซอร์ที่สามารถคล้องคอได้ ซึ่งต้องนำติดตัวไว้ตลอด ถ้าหากถอดไว้ในการเดินแต่ละครั้งก็ต้องพกไปคู่กับไม้เท้า ซึ่งผู้วิจัยเห็นว่ายังไม่เกิดความสะดวกต่อผู้พิการเท่าที่ควร ดังนั้นผู้วิจัยจึงได้สอบถามไปยังอาจารย์ที่วิทยาลัยราชสุดาถึงอุปกรณ์เสริมไม้เท้าช่วยเหลือผู้พิการทางสายตา ว่าควรจะให้ออกแบบอย่างไร เพื่อตอบสนองความต้องการของผู้พิการให้มากที่สุด โดยได้ทำการปรับปรุงเพื่อให้เกิดประโยชน์อย่างเห็นได้ชัดเจนและสามารถนำไปใช้งานได้จริง [14]

ทางผู้วิจัย จึงเห็นว่าการใช้ไม้เท้าในการนำทางก็จะมีข้อจำกัดในการตรวจสอบสิ่งกีดขวางแค่ในระยะตรงหน้า และ

เท่าที่ได้สังเกตุการเดินทางของผู้พิการทางสายตา พบว่า ถ้าหากมีอุปสรรคหรือสิ่งกีดขวางที่อยู่ด้านบนศีรษะ เช่น ทางลาดด้านล่างของบันไดสะพานลอย หรือแม้แต่สิ่งกีดขวางที่ห้อยลงมาจากด้านบน อาจจะเป็นป้ายไวนิลที่แขวนเพื่อบังแดด กิ่งไม้ที่ห้อยลงมา ผู้พิการทางสายตาไม่สามารถจะรับรู้ได้ ก็ทำให้สามารถเกิดการชน และอาจได้รับบาดเจ็บได้ [13],[17]

ดังนั้น ผู้วิจัยจึงมีความสนใจที่จะออกแบบและพัฒนาอุปกรณ์เสริมไม้เท้าเพื่อช่วยเหลือผู้พิการทางสายตาให้ได้รับความสะดวกสบายจากการเดินทางด้วยตนเอง และหลีกเลี่ยงการบาดเจ็บจากสิ่งกีดขวางที่อยู่เหนือศีรษะได้

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 ผู้พิการทางสายตา คือ คนพิการเป็นส่วนหนึ่งของทรัพยากรบุคคลที่สมควรส่งเสริมให้มีความรู้ความสามารถและมีศักยภาพในการช่วยเหลือตนเองพึ่งพาตนเองและดำรงชีวิตอยู่ในสังคมอย่างมีความสุขรวมทั้งการช่วยสร้างสรรค์สังคมได้ เช่นเดียวกับคนทั่วไปการให้การศึกษาแก่คนพิการถือเป็นวิถีทางหนึ่งที่จะช่วยให้คนพิการมีชีวิตที่ดีขึ้นได้พัฒนาความสามารถของแต่ละบุคคลได้เต็มศักยภาพมากยิ่งขึ้น [6],[7],[9],[11]

2.2 โมดูลอัลตร้าโซนิค เป็นโมดูลที่ถูกใช้ในงานด้านการตรวจหาวัตถุหลบสิ่งกีดขวางวัดระยะทางหรือเอาไว้ตรวจสอบว่ามีวัตถุเข้ามาในระยะที่กำหนดหรือไม่ โดยมีหลักการเบื้องต้นคือ ใช้การจับเวลาในการเดินทางของคลื่นเซ็นเซอร์จากตัวส่งตกกระทบยังวัตถุ แล้วสะท้อนกลับมายังตัวรับแล้วนำมาคำนวณออกมาเป็นระยะทางโดยใช้ความกว้างของสัญญาณเป็นตัวในการคำนวณ HC-SR04 เป็นโมดูลอัลตร้าโซนิคยอดนิยมโมดูลหนึ่งมีความสามารถวัดระยะได้อย่างแม่นยำ 10 us สักครู่หนึ่งขา Out ก็จะตอบสนองเอาท์พุตลอจิก High ออกมาโดยมีช่วงเวลาที่สัมพันธ์กับระยะทางที่คลื่นสะท้อนวัตถุกลับมาอย่างไรก็ตามหากวัตถุอยู่นอกเหนือขอบเขตที่ตรวจจับได้ ความกว้างสัญญาณที่ส่งกลับจะมีค่า 38 ms HC-SR04 ทำงานที่แรงดันประมาณ + 5V (4.5V ถึง+5.5V) โดยป้อนให้ขา VCC และ GND โมดูลนี้มีขาการเลือกใช้งานโมดูลประเภทนี้มีประเด็นที่สำคัญ เช่น ช่วงระยะห่างของการวัดความกว้างของมุมเมื่อคลื่นเสียงเดินทางออกไปจากตัวส่ง (เรียกว่า Beam Angle) นอกจากนั้นการสะท้อนกลับของคลื่นเสียงที่วัตถุกีดขวางขนาดและรูปทรงของวัตถุและการสะท้อนกลับของเสียงจากหลายทิศทางหรือต่างระยะกันก็มีผลต่อความถูกต้องหรือผิดพลาดในการวัดค่าระยะห่างได้เป็นการวัดระยะทางที่ไม่ได้ใกล้กันมาก [8],[12]

2.3 งานวิจัยที่เกี่ยวข้อง

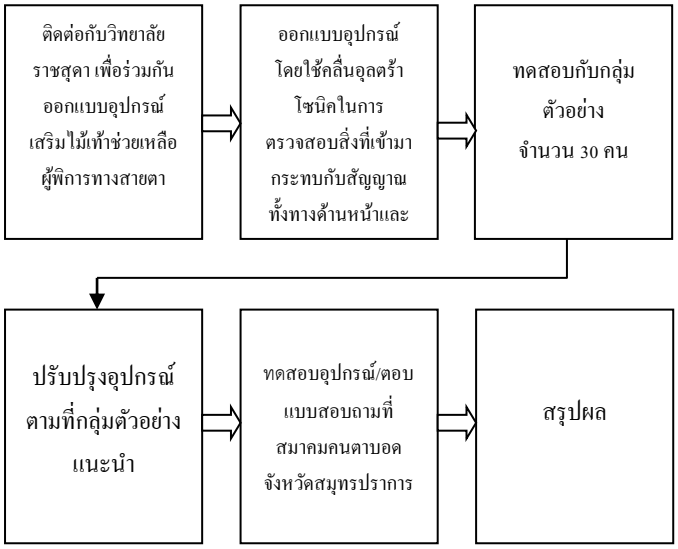
งานวิจัย เรื่อง ไม้เท้าอัจฉริยะสำหรับคนตาบอด ได้กล่าวถึงโครงงานไม้เท้าอัจฉริยะสำหรับคนตาบอด โดยการนำโฟโต้อิเล็กทริคเซ็นเซอร์ มาประยุกต์ใช้ในโครงงานนี้ ซึ่งจัดทำขึ้นมาโดยมีวัตถุประสงค์เพื่อหาหลักการทำงานของไมโครคอนโทรลเลอร์เบอร์ AT89C51 มาใช้ควบคุมการทำงานของไม้เท้าอัจฉริยะสำหรับคนตาบอดที่สามารถนำไปใช้งานเพื่ออำนวยความสะดวกและสร้างความปลอดภัยในการใช้ชีวิตประจำวันได้ ในการจัดทำไม้เท้าอัจฉริยะสำหรับคนตาบอดนี้ถูกออกแบบขึ้นมาโดยใช้แสงอินฟาเรดในการตรวจเช็คสิ่งกีดขวาง ตรวจเช็คบันได ซึ่งสามารถตรวจเช็คได้จากการที่ภาคส่ง ทำการส่งแสงอินฟาเรดไปกระทบสิ่งกีดขวางในระยะ 64 เซนติเมตรและส่งกลับมาที่ภาครับซึ่งก็จะส่งต่อไปที่ตัวประมวลผลซึ่งอยู่ในภาคถัดไปแล้ว จึงทำการประมวลผลออกมา ส่วนการตรวจเช็คบันไดภาคส่งทำการส่งแสงอินฟาเรดไปกระทบกับขอบบันไดขาขึ้นในระยะ 15 เซนติเมตร และขอบบันได ขาลงในระยะ 11 เซนติเมตร แล้วจึงทำการประมวลผลออกมา หลังจากนั้น แล้วก็จะแสดงผลออกมาในรูปการส่งเสียงเตือนและการสั่นเตือนเมื่อตรวจ เช็คเจอบันได ในส่วนของสิ่งกีดขวางจะแสดงผลออกมาในรูปการส่งเสียงเตือน และสามารถสร้างเสียงอัลตร้าโซนิคที่มีความถี่สูงที่ 45 kHz ออกมา เพื่อขับไล่สุนัขไม่ให้มาทำร้ายได้ [18]

งานวิจัย เรื่อง เครื่องบอกเตือนสิ่งกีดขวางสำหรับผู้พิการทางสายตา โดยสำนักวิชาเทคโนโลยีสารสนเทศ มหาวิทยาลัยแม่ฟ้าหลวง แจ้งว่า ได้มีการคิดค้นสิ่งประดิษฐ์ใหม่ คือ "เครื่องบอกเตือนสิ่งกีดขวางสำหรับผู้พิการทางสายตา" ที่มีประสิทธิภาพและราคาถูก โดยมีชื่อว่าไอโซน่าร์ (iSonar) ซึ่งคิดประดิษฐ์ขึ้น โดยนายสุรพล วรภัทราทร อาจารย์ประจำสำนักวิชาเทคโนโลยีสารสนเทศ โดยเครื่องไอโซน่าร์ที่ว่านี้มีขนาดเกือบเท่ากล่องไม้ขีด มีน้ำหนักเบาไม่ถึง 2 กรัม ลักษณะเป็นกล่องสีขาว และมีสายไว้คล้องคอ เพื่อให้เครื่องอยู่ระดับอก และหันส่วนปฏิบัติการออกไปทางด้านหน้า เมื่อผู้คล้อง

อุปกรณ์เปิดเครื่องและเดินตรงไปห่างจากวัตถุกำแพง ฯลฯ ระยะตั้งแต่ 1.30 เมตรลงมา เครื่องก็จะเริ่มสั่น และยิ่งเดินเข้าไปใกล้วัตถุเครื่องก็จะสั่นไหวเพิ่มมากขึ้นเรื่อยๆ ช่วยให้ผู้พิการทางสายตาสามารถทราบการเข้าใกล้วัตถุที่อยู่ด้านหน้าได้ มีรัศมีแนวกว้างทั้งด้านบน ล่าง ซ้าย ขวา ข้างละ 40 ซม.[19]

3. วิธีการดำเนินการวิจัย

การศึกษาอุปกรณ์เสริมไม้เท้าสำหรับผู้พิการทางสายตา เป็นการวิจัยเชิงคุณภาพ โดยมีจุดมุ่งหมายเพื่อประเมินประสิทธิภาพของอุปกรณ์ และจะสามารถนำไปต่อยอดเป็นผลิตภัณฑ์ที่ดี และมีคุณประโยชน์สำหรับผู้พิการทางสายตา โดยผู้วิจัยมีกรอบแนวความคิดในการดำเนินงานวิจัยดังนี้[10]



ภาพที่ 1 ขั้นตอนการดำเนินการวิจัย

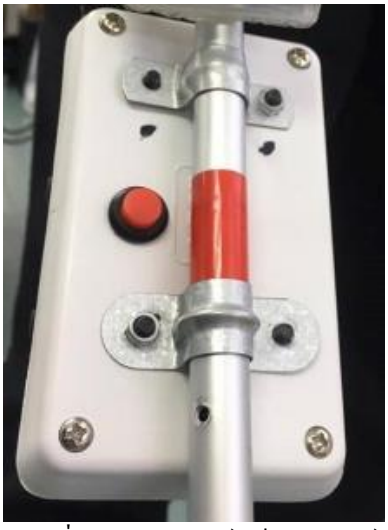
4. ผลการพัฒนาและวิเคราะห์ข้อมูล

อุปกรณ์เสริมไม้เท้าผู้พิการทางสายตาที่ได้พัฒนาขึ้น ประกอบไปด้วยกัน 2 ส่วน ดังนี้



ภาพที่ 1 อุปกรณ์เซ็นเซอร์

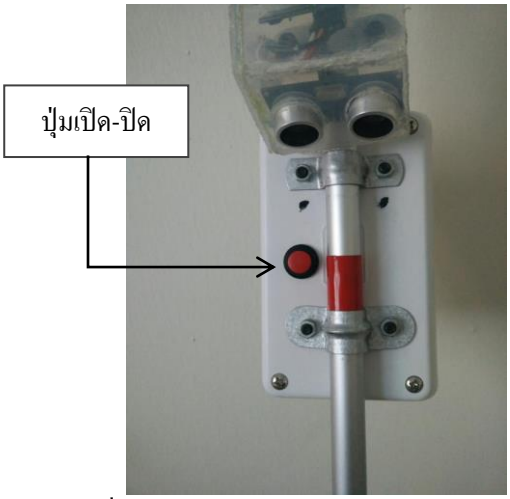
จากภาพที่ 1 เป็นอุปกรณ์เซ็นเซอร์ที่ไว้ตรวจจับสิ่งกีดขวางทั้งด้านบน และด้านหน้า



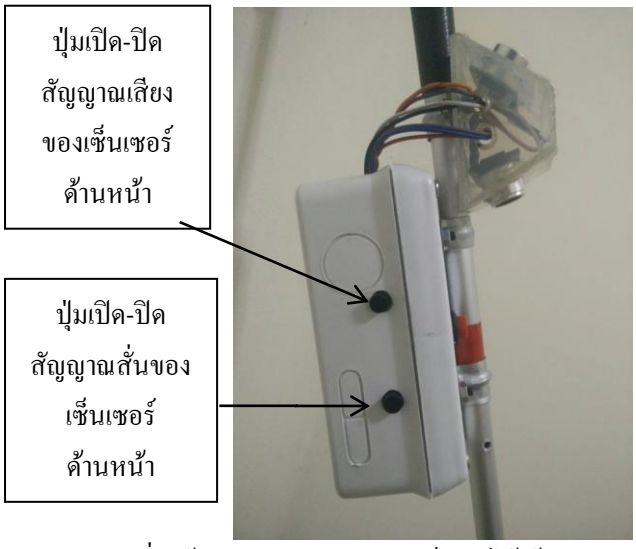
ภาพที่ 2 กล่องอุปกรณ์อิเล็กทรอนิกส์

จากภาพที่ 2 เป็นกล่องอุปกรณ์อิเล็กทรอนิกส์ และมีปุ่มเปิด-ปิดการทำงานของตัวเซ็นเซอร์ในแต่ละส่วน สามารถใช้งาน [15] ได้ดังนี้

1. กดปุ่มสีแดง (Power) ซึ่งอยู่ด้านหน้าของกล่องอุปกรณ์เสริมไม้เท้าผู้พิการทางสายตา เป็นปุ่มเปิด-ปิด เพื่อเป็นการให้อุปกรณ์ต่างๆ เริ่มทำงาน



ภาพที่ 3 แสดงด้านหน้าของกล่องอุปกรณ์เสริมไม้เท้าผู้พิการทางสายตา



ภาพที่ 4 ด้านขวาของกล่องอุปกรณ์เสริมไม้เท้า
ผู้พิการทางสายตา

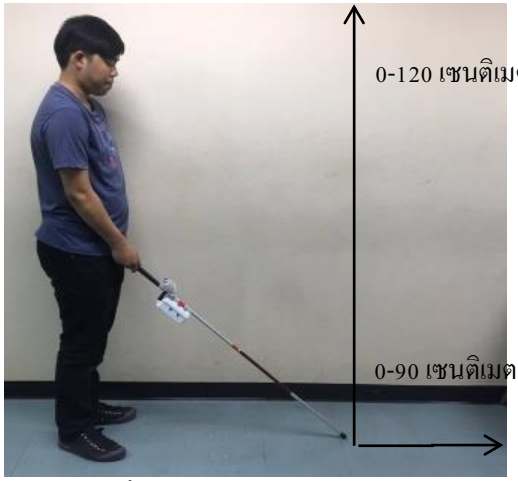
2. ด้านขวาและด้านขวา ประกอบไปด้วย ปุ่มสีดำ ด้านบนและล่างของกล่องอุปกรณ์

2.1 สีดำด้านบน เป็นปุ่มเปิด-ปิดสัญญาณเสียง เมื่อพบสิ่งกีดขวางที่อยู่ด้านหน้าของอุปกรณ์เสริมไม้เท้าผู้พิการทางสายตา

2.2 สีดำด้านล่าง เป็นปุ่มเปิด-ปิดสัญญาณสั่น เมื่อพบสิ่งกีดขวางที่อยู่ด้านหน้าของอุปกรณ์เสริมไม้เท้าผู้พิการทางสายตา

4.1 การทำงานของอุปกรณ์

เมื่อประกอบอุปกรณ์เสริมไม้เท้าผู้พิการทางสายตาเรียบร้อยแล้ว ก็ทำการกดปุ่มเปิด-ปิด (Power) และกดปุ่มการใช้งานเซ็นเซอร์ตามที่ต้องการ (บน-ล่าง, เสียง-สั่น)



ภาพที่ 5 ระยะของการใช้งานอุปกรณ์

จากภาพที่ 5 เมื่อเซ็นเซอร์ด้านหน้าจะทำการตรวจสิ่งกีดขวางในระยะ 0-90 เซนติเมตรจากตัวอุปกรณ์เสริมไม้เท้าสำหรับผู้พิการทางสายตา เมื่อมีสิ่งกีดขวางเข้ามาในระยะ 0-90 เซนติเมตร สัญญาณจะทำการแจ้งเตือนจะเริ่มทำงานไปเรื่อยๆ และถี่ขึ้นเมื่ออยู่ในระยะใกล้ 0 เซนติเมตร และสัญญาณจะหยุดทำงานก็ต่อเมื่อไม่มีสิ่งกีดขวางในอยู่ในระยะที่กำหนด [16]

เมื่อเซ็นเซอร์ด้านบนจะทำการตรวจสิ่งกีดขวางในระยะ 0-120 เซนติเมตรจากตัวอุปกรณ์เสริมไม้เท้าสำหรับผู้พิการทางสายตา เมื่อมีสิ่งกีดขวางเข้ามาในระยะ 0-120 เซนติเมตร สัญญาณจะทำการแจ้งเตือนจะเริ่มทำงานไปเรื่อยๆ และถี่ขึ้นเมื่ออยู่ในระยะใกล้ 0 เซนติเมตร และสัญญาณจะหยุดทำงานก็ต่อเมื่อไม่มีสิ่งกีดขวางในอยู่ในระยะที่กำหนด ดังตัวอย่างการใช้งานอุปกรณ์ในภาพที่ 6



ภาพที่ 6 การใช้งานอุปกรณ์

4.2 ผลการวิเคราะห์ความพึงพอใจของผู้พิการต่ออุปกรณ์เสริมไม้เท้าสำหรับผู้พิการทางสายตา

การเสนอผลการวิเคราะห์ และการแปรผลการวิเคราะห์ข้อมูล ผู้วิจัยได้ใช้โปรแกรมสำเร็จรูปทางสถิติ โดยในการทดสอบสมมุติฐานมีนัยสำคัญทางสถิติที่ระดับ 0.05 ผู้วิจัยได้วิเคราะห์และนำเสนอในรูปแบบของตารางประกอบคำอธิบาย โดยเรียงลำดับหัวข้อการวิเคราะห์ข้อมูลเป็น 4 ส่วนดังนี้

ส่วนที่ 1 ข้อมูลทั่วไปของผู้ตอบแบบสอบถาม มีจำนวนทั้งหมด 15 คน เพศชาย 7 คน และเพศหญิง 8 คน อายุส่วนใหญ่จะอยู่ที่ ระหว่าง 30-39 ปี จำนวน 7 คน 40-49 ปีจำนวน 6 คน และอายุ 50 ปีขึ้นไป จำนวน 2 คน

ส่วนที่ 2 ความพึงพอใจต่ออุปกรณ์ช่วยเหลือผู้พิการทางสายตา

ตารางที่ 4-1 แสดงผลการตอบแบบสอบถามส่วนที่ 2

ความพึงพอใจต่ออุปกรณ์	$\bar{X}$	SD	ระดับความพึงพอใจ
2.1 อุปกรณ์ช่วยเหลือผู้พิการทางสายตามีน้ำหนักที่เหมาะสม	3.80	0.77	มาก
2.2 อุปกรณ์ช่วยเหลือผู้พิการทางสายตาสามารถพกพาได้สะดวก	3.73	0.59	มาก
2.3 อุปกรณ์ช่วยเหลือผู้พิการทางสายตามีเสียงเตือนที่ชัดเจน	4.27	0.70	มาก
2.4 อุปกรณ์ช่วยเหลือผู้พิการทางสายตามีการสั่นสะเทือน	3.80	0.68	มาก
2.5 อุปกรณ์ช่วยเหลือผู้พิการทางสายตามีความไวในการตรวจจับสิ่งกีดขวางที่แม่นยำ	4.40	0.63	มาก
2.6 อุปกรณ์ช่วยเหลือผู้พิการทางสายตามีความคงทนใช้ได้นาน	3.27	0.46	ปานกลาง
2.7 อุปกรณ์ช่วยเหลือผู้พิการทางสายตาเหมาะสมกับทุกเพศ ทุกวัย	3.67	0.62	มาก
2.8 อุปกรณ์ช่วยเหลือผู้พิการทางสายตา รู้สึกมีความปลอดภัยมากขึ้นเมื่อผู้พิการทางสายตานำไปใช้	4.80	0.41	มากที่สุด
รวม	**3.97**	**0.12**	**มาก**

ตารางที่ 4-1 ผลการตอบแบบสอบถามจากผู้พิการทางสายตา มีจำนวนทั้งสิ้น 15 คน แบ่งเป็นชายจำนวน 7 คน และหญิงจำนวน 8 คน พบว่าในภาพรวมผู้พิการทางสายตามีความพึงพอใจต่ออุปกรณ์เสริมไม้เท้าช่วยเหลือผู้พิการทางสายตาอยู่ในระดับมาก คิดเป็นค่าเฉลี่ย 3.97 โดยเมื่อพิจารณารายด้านแล้วผู้พิการทางสายตามีความพอใจในระดับมากที่สุด คือ รู้สึกมีความปลอดภัยมากขึ้นเมื่อผู้พิการทางสายตานำไปใช้ คิดเป็นค่าเฉลี่ย 4.80 รองลงมาคืออุปกรณ์มีความไวในการตรวจจับสิ่งกีดขวางที่แม่นยำ คิดเป็นค่าเฉลี่ย 4.40

ส่วนที่ 3 ความพึงพอใจต่อการใช้งานอุปกรณ์ช่วยเหลือผู้พิการทางสายตา

ตารางที่ 4-2 แสดงผลการตอบแบบสอบถามส่วนที่ 3

ความพึงพอใจต่อการใช้งานอุปกรณ์	$\bar{X}$	SD	ระดับความพึงพอใจ
3.1 อุปกรณ์ช่วยเหลือผู้พิการทางสายตาสามารถเปลี่ยนแบตเตอรี่ได้ง่าย	3.07	0.26	ปานกลาง
3.2 อุปกรณ์ช่วยเหลือผู้พิการทางสายตาสามารถเปลี่ยนไปใช้กับไม้เท้าผู้พิการทางสายตาอันใหม่ได้	4.07	0.46	มาก
3.3 อุปกรณ์ช่วยเหลือผู้พิการทางสายตาควรติดไว้กับไม้เท้า	3.60	0.74	มาก
3.4 ท่านมีความพอใจในอุปกรณ์ช่วยเหลือผู้พิการทางสายตา	4.33	0.62	มาก
3.5 ท่านมีความสนใจที่จะมีอุปกรณ์ช่วยเหลือผู้พิการทางสายตาไว้ใช้	4.47	0.64	มาก
รวม	**3.91**	**0.19**	**มาก**

ตารางที่ 4-2 ผลการตอบแบบสอบถามจากผู้พิการทางสายตา พบว่าในในภาพรวมผู้พิการทางสายตามีความพึงพอใจต่อการใช้งานอุปกรณ์เสริมไม้เท้าช่วยเหลือผู้พิการทางสายตาอยู่ในระดับมาก คิดเป็นค่าเฉลี่ย 3.91 โดยเมื่อพิจารณารายด้านแล้วผู้พิการทางสายตามีความพอใจต่อการใช้งานในระดับมากที่สุด คือ ผู้พิการทางสายตามีสนใจที่จะมีอุปกรณ์ช่วยเหลือผู้พิการทางสายตาไว้ใช้ คิดเป็นค่าเฉลี่ย 4.47 รองลงมาคืออุปกรณ์ช่วยเหลือผู้พิการทางสายตา คิดเป็นค่าเฉลี่ย 4.33

ส่วนที่ 4 ข้อเสนอแนะ ดังนี้

1. ควรมีการล็อกติดกับไม้หรือไม่ก็ควรติดกับไม้อย่างเดียว ควรมีเซ็นเซอร์ติดกับตัวเครื่องก็ได้

2. ควรตั้งระยะเองได้และอยากให้เปลี่ยนแบตเตอรี่เป็นแบบโทรศัพท์

5. สรุป

ผู้วิจัยได้ใช้โปรแกรมสำหรับเขียนชุดคำสั่งต่างๆ ที่ใช้สั่งการไมโครคอนโทรลเลอร์เพื่อให้แสดงผลลัพธ์ออกมาตามความต้องการและในส่วนของ library ต่างๆ ให้เลือกใช้ไม่ว่าจะเป็น library ของตัว Ultra Sonic Sensor และตัวอุปกรณ์อื่นสามารถ download มาเพิ่มและสามารถเรียกใช้ได้ง่ายไม่ยุ่งยาก จึงทำให้การเขียนโปรแกรมหรือการกำหนดค่าต่างๆ ในการควบคุม Micro Controller ทำได้ไม่ยุ่งยากและไม่ซับซ้อนจนเกินไป ซึ่งไม้เท้าที่ผู้วิจัยได้จัดทำขึ้นนี้มีแตกต่างจากไม้เท้าที่มีอยู่ทั่วไป ๆ คือ สามารถถอดหรือเปลี่ยนอุปกรณ์ไปใช้กับไม้เท้าอื่น ๆ ได้อย่างสะดวกและง่ายดาย โดยไม้เท้าทั่ว ๆ ไปไม่สามารถกระทำได้

5.1 อภิปรายผล

จากการได้นำอุปกรณ์ช่วยเหลือผู้พิการทางสายตาไปให้ผู้พิการทางสายตาได้ทำการทดลองการทำงานพบว่า ผู้พิการทางสายตาให้ความสนใจในอุปกรณ์นี้มาก และต้องการให้นำไปใช้งานจริง บางท่านไม่ต้องการให้มีเฉพาะไม้เท้าสำหรับผู้พิการแต่ต้องการแค่ตัวอุปกรณ์ช่วยเหลือผู้พิการทางสายตาเพียงอย่างเดียว เพื่อความคล่องตัวในการช่วยนำทางภายในบ้านของตนเอง

5.2 ข้อเสนอแนะ

จากการสอบถามพูดคุยเชิงลึกกับผู้พิการ พบว่า ตัวอุปกรณ์ช่วยเหลือผู้พิการทางสายตา ต้องการให้มีความยืดหยุ่นในการใช้งานที่มีการถอดที่สะดวก และสามารถพกพาแบบไม่ต้องมีไม้เท้า และแบตเตอรี่ควรจะเป็นแบบโทรศัพท์ คือไม่ต้องคอยถอดเปลี่ยนถ่านบ่อย ๆ และต้องการกำหนดระยะของเซ็นเซอร์เองได้ เพื่อความคล่องตัวในการไปยังพื้นที่ต่างๆ ได้

6. เอกสารอ้างอิง

[1] Bateman, 1964 and Schindale, 1974 อ้างถึงใน Telford and Sawry, 1981

[2] Cartwright, G. P., Cartwright, C. A., and Ward, M. E. (1981). Educating special learners. Belmont, California: Wadsworth Publishing Company.

[3] Charles W Telford; James M Sawrey (1981). The exceptional individual, Englewood Cliffs, N.J. : Prentice-Hall.

[4] Hallahan, D. P., & Kauffman, J. M. (1988). Exceptional children: Introduction to special education (4th ed.). Englewood Cliffs, NJ: Prentice Hall.

[5] กระทรวงศึกษาธิการ. “แผนการศึกษาแห่งชาติ พ.ศ. 2520.” แผนพัฒนาการศึกษาพิเศษ (ด้านคนพิการ) ระยะที่ 8, พ.ศ. 2540 – 2544.

[6] กองการศึกษาเพื่อคนพิการ, กรมสามัญศึกษา. “คู่มือการจัดการเรียนร่วมในระดับมัธยมศึกษา” กรุงเทพฯ : ม.ป.ท., ม.ป.ป.

[7] คณะอนุกรรมการการคัดเลือกและจำแนกความพิการเพื่อการศึกษา, กระทรวงศึกษาธิการ. (2543). “คู่มือการคัดแยกและส่งต่อเพื่อการศึกษา” กรุงเทพฯ : โรงพิมพ์คุรุสภา.คู่มือการคัดแยกและส่งต่อคนพิการเพื่อการศึกษา, 2543.

[8] น.ท.ไพศาล โมลิสกุลมงคล (2550), “คอมพิวเตอร์กราฟิกส์”, กรุงเทพฯ : หจกไทยเจริญ การพิมพ์.

[9] เบญจา ชลธาร์นนท์ (2538), “รวมบทความวิชาการทางการศึกษาพิเศษ”. กรุงเทพฯ : บริษัท สหธรรมิก จำกัด.

[10] ดร.สมพร หวานเสร็จ (2543). “การจัดทำโปรแกรมการศึกษาเฉพาะบุคคล” เอกสารบรรยายในการอบรมเชิงปฏิบัติการเรื่อง “การจัดการศึกษาสำหรับเด็กพิการเรียนร่วม” 18-20 มกราคม. สถาบันราชภัฏสวนดุสิต, กรุงเทพฯ : 2543

[11] ผดุง อารยะวิญญู (2533).“การศึกษาสำหรับเด็กที่มีความต้องการพิเศษ” กรุงเทพฯ : โรงพิมพ์กรุงธนพัฒนา.

[12] มงคล พรหมเทศ (2552).” หนังสืองานไฟฟ้าทั่วไป” กรุงเทพฯ: เอมพันธ์. รัฐธรรมนูญแห่งราชอาณาจักรไทย พ.ศ. 2540.

[13] สว่าง โรจน์รัตนเกียรติ(2524). “สุขภาพจิตของนักเรียนตาบอดในประเทศไทย กรุงเทพมหานคร” : มหาวิทยาลัยธรรมศาสตร์ – IRDC1.

[14] สมทรง พันธุ์สุวรรณ (2529). “การศึกษาสำหรับคนที่บกพร่องทางการเห็น” ภาควิชาการศึกษาพิเศษ คณะครุศาสตร์ วิทยาลัยครูสวนดุสิต.

[15] โอภาส ศิริครรชิตถาวรและวรพจน์ กรแก้ววัฒนกุล และชัยวัฒน์ ลิ้มพรจิตรวิไล (2552), “เรียนรู้ควบคุม อย่างง่ายด้วยโปรแกรมภาษา C กับ Arduino และบอร์ดไมโครคอนโทรลเลอร์”, กรุงเทพฯ : บริษัท อินโนเวตีฟ เอ็กเพอริเมนต์จำกัด.

[16] เอกชัย มะการ (2552). “เรียนรู้เข้าใจใช้งานไมโครคอนโทรลเลอร์ตระกูล AVR ด้วย Arduino”, กรุงเทพฯ : บริษัท อีทีที จำกัด.

[17] เอมอร ตั้งจิตรมณีศักดา (2534), “สภาพการศึกษาปัญหาความต้องการทางการศึกษาและการฝึกอาชีพของนักเรียนตาบอด : ศึกษาเฉพาะกรณีโรงเรียนสอนคนตาบอดภาคเหนือ จังหวัดเชียงใหม่”. กรุงเทพมหานคร : จุฬาลงกรณ์มหาวิทยาลัย.

[18] สมองใสไฮเทค ผลงานนักศึกษา , ISSUE1. Volume20. May-July, 2013.

[19] ASTV ผู้จัดการออนไลน์ , มูลนิธิพัฒนาคนพิการไทย 2 พ.ย. 2556

การพยากรณ์การเป็นสัดในโคนมโดยใช้ตรรกศาสตร์คลุมเครือ
Estrus Prediction in Dairy Cows using Fuzzy Logic

ชนิดา ถาวรรัชชต์ (C. Tawornrach)[1] สายยัญ สายยศ (S. Saiyod) [2]
งามนิจ อาจอินทร์ (N. Arch-Int) [3]จักรชัย โสอินทร์ (C. So-In) [4] และอรัญ จันทร์ลุน (A. Chanlun) [5]
[1,2,3,4]ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น
[5]ภาควิชาอายุรศาสตร์ คณะสัตวแพทยศาสตร์ มหาวิทยาลัยขอนแก่น
[1]Chanida.ta@kkumail.com, [2]Saiyan@kku.ac.th, [3] Ngamnij@kku.ac.th,
[4] Chackso@kku.ac.th, [5]Aran_jan@kku.ac.th

บทคัดย่อ

โคนมเป็นสัตว์เศรษฐกิจชนิดหนึ่งที่ก่อเกิดรายได้ให้กับเกษตรกร ดังนั้นการเพิ่มปริมาณโคจึงเป็นสิ่งสำคัญ แต่โคนมมียังคงปัญหาเรื่องโอกาสความสำเร็จในการผสมเทียม โดยปกติโคนมมีวงรอบการเป็นสัด 18-24 วัน หรือเฉลี่ย 21 วัน ในแต่ละรอบมีความคลาดเคลื่อนในการกลับเข้าสู่วงรอบตรงตามวันเดิม เดิมทีการตรวจจับสัดถูกตรวจโดยการสังเกตของเกษตรกร ในบางครั้งอาจเกิดความคลาดเคลื่อนจากการแสดงออกของพฤติกรรมโคที่ไม่ชัดเจน และพลาดโอกาสในการผสมเทียมครั้งนั้นไป งานวิจัยนี้จึงนำเสนอการใช้เทคนิคฟัซซีลอจิก เพื่อช่วยในการตรวจจับการเป็นสัดในโคนม โดยพิจารณาจากปัจจัยที่เกี่ยวข้องกับการเคลื่อนวงรอบการเป็นสัดในโค จากนั้นสร้างกฎโดยผู้เชี่ยวชาญ จากผลลัพธ์แสดงให้เห็นถึงวงรอบการเป็นสัดของโคที่ได้จากการคำนวณกฎ ผลการทดลองโมเดลแสดงให้เห็นว่าค่าความถูกต้องอยู่ที่ร้อยละ 83.33 ในงานวิจัยพบว่าปัจจัยนำเข้า ฟังก์ชันที่เลือกใช้ หรือกฎ มีผลต่อความแม่นยำของโมเดล

คำสำคัญ: โคนม วงรอบการเป็นสัด ฟัซซีลอจิก การพยากรณ์

Abstract

The dairy cows are the one of economic animals that generates income for the farmer. So that, increasing the amount of cows is very important. Generally, estrus of dairy cow will occur every 18 to 24 days, or average 21 days. In each round, there may be discrepant of estrus cycle. In the past, the estrus of cows is observed by farmer. Sometimes, the behaviors of cows unclearly show therefore the cow miss chance to pregnant. This research presents an algorithm using fuzzy logic to predict the estrus in dairy cows. The factors of estrus cycle shifting are considered as input variables. Then, the fuzzy rules will be created by an expert. The output of fuzzy logic will show the estrus cycle. The experimental results show that the accuracy of model is 83.33%. We found that accuracy of the model dependent on input variables, membership function or fuzzy rules.

Keywords: Dairy Cows, Estrus Cycle, Fuzzy Logic, Prediction

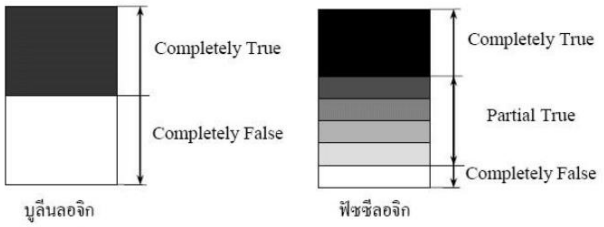
1. บทนำ

โคนมเป็นสัตว์เศรษฐกิจชนิดหนึ่งของประเทศที่มีการเลี้ยงอย่างแพร่หลาย กระจายอยู่ตามภูมิภาคต่าง ๆ ของประเทศไทย [1] และก่อให้เกิดรายได้ให้แก่เกษตรกร ดังนั้นการเพิ่มปริมาณโคให้เพียงพอต่อผู้บริโภคจึงเป็นสิ่งสำคัญ แต่ฟาร์มโคนมส่วนใหญ่ยังคงมีปัญหาเรื่องการผสมพันธุ์อยู่ เนื่องจากปัจจัยหลักที่ทำให้ผสมไม่ติดเกิดจากการจับสัดที่ผิดพลาด

โดยปกติโคนมมีวงรอบการเป็นสัดอยู่ที่ประมาณ 18-24 วัน หรือเฉลี่ย 21 วัน ซึ่งมีความคลาดเคลื่อนในแต่ละรอบ ไม่มาตรงตามวันเดิม ในการเป็นสัดแต่ละครั้งอาจนานเพียง 15-18 ชั่วโมง [2] และควรได้รับการผสมเทียมประมาณ 12 ชั่วโมง หลังการเป็นสัด เดิมใช้วิธีการสังเกตจากพฤติกรรมของโคในขณะที่เป็นสัดโดยเกษตรกร ซึ่งในบางครั้งมีปัจจัยอื่นที่ทำให้โคไม่แสดงพฤติกรรม เช่น ความเครียด ดังนั้นหากมีการจับสัดผิดพลาด อาจทำให้ได้วันและช่วงเวลาในการผสมเทียมที่คลาดเคลื่อนได้ ส่งผลให้พลาดโอกาสในการผสมเทียม และการผลิตลูกโคลดลง

จากปัญหาข้างต้น จึงได้มีหลายงานวิจัยที่มีการใช้เทคนิคต่าง ๆ ในการพยากรณ์การเป็นสัดในโคนมล่วงหน้า เพื่อเพิ่มประสิทธิภาพการตรวจจับสัด ในโค เช่น มีการเพิ่มประสิทธิภาพการตรวจจับสัดในโคนมโดยใช้การตรวจวัดเซนเซอร์ [3] มีการใช้ฟัซซีลอจิก เพื่อหาอัตราความถูกต้องของการจับสัดในโค [4] หรือการตรวจวัดค่าเปลี่ยนแปลงการเคลื่อนไหวของโคขณะที่เป็นสัดโดยใช้ฟัซซีลอจิก [5]

จากงานวิจัยข้างต้นจะเห็นได้ว่ามีการนำฟัซซีลอจิกมาใช้ในการวิจัย เนื่องจากเป็นศาสตร์ด้านการคำนวณที่ให้ค่าได้หลายระดับตั้งแต่ 0 ถึง 1 แตกต่างจากลอจิกบูลีนที่มีเพียง 2 ค่าคือ 0 และ 1 จึงสามารถให้คำตอบได้เหมาะสมกว่า หากในงานวิจัยลอจิกบูลีนอาจให้คำตอบได้เพียงโคเป็นสัดหรือไม่ แต่ฟัซซีลอจิกสามารถแสดงถึงการเป็นสัดในระดับต่าง ๆ ได้มากกว่า ผู้วิจัยจึงได้นำเทคนิคฟัซซีลอจิกมาใช้ในการแก้ไขปัญหาความไม่แน่นอนของวงรอบการเป็นสัดในโค เพื่อช่วยพยากรณ์การเป็นสัดในโคนมล่วงหน้า และรองรับอุปกรณ์ที่ใช้เทคโนโลยีเซนเซอร์ในการตรวจจับสัดให้เกษตรกรสามารถนำไปวางแผนการผสมเทียมได้ดียิ่งขึ้น



ภาพที่ 1: บูลีนลอจิก และ ฟัซซีลอจิก

2.วรรณกรรมและงานวิจัยที่เกี่ยวข้อง

ผู้วิจัยได้ศึกษาวรรณกรรมที่เกี่ยวข้องรวมถึงงานวิจัยต่าง ๆ ที่มีนำเทคนิคต่าง ๆ มาใช้ในการพยากรณ์การเป็นสัดในโคนม โดยในบทความได้แบ่งหัวข้อดังนี้

2.1 วรรณกรรมที่เกี่ยวข้อง

ตรรกศาสตร์คลุมเครือ (Fuzzy Logic) [6] เป็นศาสตร์ด้านการคำนวณ ซึ่งแตกต่างจากบูลีนลอจิก (Boolean Logic) ทั่วไปที่ให้ค่าเพียงใช่หรือไม่ใช่ 0 หรือ 1 แต่ฟัซซีลอจิกเป็นตรรกะหลายระดับ โดยใช้การกำหนดค่าระดับความเป็นสมาชิก (degree of membership) มีค่าตั้งแต่ 0 ถึง 1 ดังภาพที่ 1

ในปัจจุบันได้มีการนำฟัซซีลอจิกมาใช้สร้างเทคนิค หรือระบบด้านคอมพิวเตอร์เพื่อแก้ไขปัญหาต่าง ๆ ที่มีความไม่แน่นอน โดยอาศัยความรู้จากผู้เชี่ยวชาญเฉพาะด้านในการสร้างกฎ ทำให้ระบบสามารถตัดสินใจได้อย่างชาญฉลาดโดยเลียนแบบความคิดมนุษย์

ขั้นตอนการทำงานของฟัซซีลอจิกจะเริ่มจากการแปลงข้อมูลนำเข้าให้อยู่ในรูปแบบค่าตัวแปรฟัซซีเพื่อให้ระบบสามารถนำไปคำนวณความน่าจะเป็นได้ จากนั้นจะประเมินค่าความเป็นสมาชิกจากปัจจัยนำเข้าแต่ละตัว การประเมินค่าจะถูกกระทำด้วยตัวฟัซซีเซต (Fuzzy Set Operation) เช่น AND หรือ OR และตรวจกับว่าตรงกับกฎใดบ้างก่อนนำผลลัพธ์ไปแปลงให้อยู่ในค่าที่สามารถนำไปใช้งานจริงได้

2.2 งานวิจัยที่เกี่ยวข้อง

เดิมการใช้เกษตรกรเป็นผู้สังเกตการเป็นสัดในโค อาจทำให้สูญเสียเวลาจากการเฝ้าสังเกต และได้ประสิทธิภาพไม่ดีเท่าที่ควรหากโคไม่แสดงพฤติกรรม ดังนั้นจึงมีหลายงานวิจัยที่ใช้อุปกรณ์ หรือเทคนิคต่าง ๆ มาช่วยในการตรวจจับให้มีความถูกต้องเพิ่มมากขึ้น เช่น

Leandro dos Anjos Brunassi1 และคณะ[3] ได้ใช้การตรวจวัดเซนเซอร์มาใช้ในการวัดการเคลื่อนไหวของโค ผลผลิตของโค อุณหภูมิต่าง ๆ ในโคแต่ละตัว และนำฟัซซีลอจิกมาคำนวณค่าปัจจัยต่าง ๆ เพื่อหาคำตอบที่มีความแม่นยำมากขึ้น Leandro Ferreira และคณะ[4] ได้มีการใช้ฟัซซีลอจิก เพื่อคำนวณหาอัตราความถูกต้องของการจับสัดในโค โดยใช้ปัจจัย

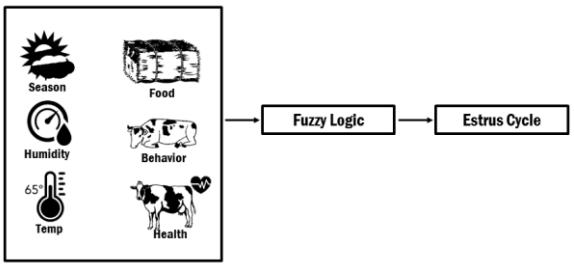
พฤติกรรมการเป็นสัดในโคนม (NBO) ช่วงเวลาที่พยายามขึ้นขี่โคตัวอื่น (APM) และระยะเวลาจำนวนวันตั้งแต่การเป็นสัดครั้งล่าสุด (PSLE) ให้ผลลัพธ์เป็นเปอร์เซ็นต์ของความน่าจะเป็นสัดในโค Abootorabi Zarchi และคณะ [5] ใช้โมเดลฟัซซีวิเคราะห์ว่าโคตัวที่มีค่าเคลื่อนไหวร่างกายตามสมมุติฐานที่ตั้งไว้อยู่ในช่วงเวลาการเป็นสัดจริงหรือไม่

ผลจากการศึกษางานวิจัยที่เกี่ยวข้องพบว่า มีหลายงานวิจัยนำฟัซซีลอจิกมาใช้ในการเพิ่มอัตราความถูกต้องของการจับสัดในโคนม ทำให้ความแม่นยำเพิ่มขึ้นเมื่อเทียบกับวิธีการสังเกตโดยเกษตรกร แต่ละงานวิจัยใช้ปัจจัยที่แตกต่างกันออกไปตามข้อมูลที่มี สภาพแวดล้อมของแต่ละพื้นที่ รวมถึงผู้เชี่ยวชาญในการตรวจสอบ

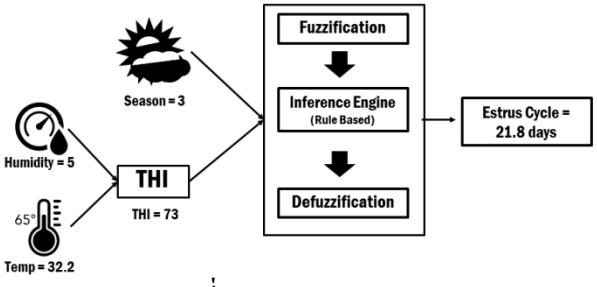
3. วิธีดำเนินงานวิจัย

ในงานวิจัยได้ใช้กระบวนการฟัซซีลอจิก เพื่อทำนายวงรอบการเป็นสัดของโคในรอบถัดไป โดยอาศัยปัจจัยนำเข้ามาจากปัจจัยที่ทำให้โคนมมีวงรอบการเป็นสัดไม่ปกติ หรือไม่มาตรงตามวันเดิม เข้ากระบวนการฟัซซี ผ่านกฎที่สร้างขึ้น และสามารถให้คำตอบได้ว่าโคมีวงรอบการเป็นสัดในรอบถัดไปกี่วัน ดังภาพที่ 2

ปัจจัยที่ทำให้โคนมมีวงรอบการเป็นสัดไม่ตรงตามวันเดิมมีหลากหลายปัจจัย เช่น สภาพภูมิอากาศหรือฤดูกาล ความเครียดของโค อาหาร พฤติกรรม หรือสุขภาพโคในขณะที่ตั้งครรภ์



ภาพที่ 2: ภาพรวมงานวิจัย



ภาพที่ 3: กระบวนการงานวิจัย

ในบางปัจจัยยังคงมีข้อจำกัดในการนำเข้ามาในระบบ เช่น ปัจจัยพฤติกรรมต้องอาศัยการเฝ้าสังเกตจากเกษตรกรหรือปัจจัยทางด้านอาหารที่ไม่สามารถควบคุมให้เหมือนกันได้ในแต่ละพื้นที่ หรือบางปัจจัยยังต้องใช้เครื่องมือทางการเกษตรในการตรวจอยู่ เช่น สุขภาพของโค แต่ปัจจัยบางตัวสามารถใช้อุปกรณ์ในการตรวจวัดค่าเซนเซอร์และใช้เป็นระบบอัตโนมัติได้ เช่น ค่าดัชนีอุณหภูมิและความชื้น (THI) ซึ่งเป็นค่าที่ใช้ในการวัดระดับความเครียดของโค ได้จากอุณหภูมิ และค่าความชื้นของพื้นที่ที่โคอาศัย หรือบางปัจจัยสามารถสังเกตได้โดยง่าย เช่น ฤดูกาล

ดังที่กล่าวมาข้างต้น จะเห็นได้ว่าบางปัจจัยยังคงมีข้อจำกัดการนำมาใช้ในระบบ ดังนั้นในงานวิจัยจึงเลือกมาใช้เพียง 2 ปัจจัย ซึ่งเป็นปัจจัยภายนอกที่สำคัญต่อการเคลื่อนวงรอบการเป็นสัดในโค

ภาพที่ 3 แสดงกรอบงานวิจัย และตัวอย่างค่าข้อมูลต่าง ๆ โดยงานวิจัยนี้ได้ใช้ปัจจัยนำเข้า 2 ปัจจัย คือ ฤดูกาล (Season) เช่น 3 คือช่วงเดือนมีนาคม และค่าดัชนีอุณหภูมิและความชื้น (THI) ซึ่งค่า THI เป็นค่าที่ใช้ในการวัดความเครียดของโค เช่น 73 ได้จากการคำนวณค่าอุณหภูมิ 32.2 °C และค่าความชื้น 5 เปอร์เซ็นต์ เมื่อผ่านการคำนวณกฎของฟัซซีลอจิก สามารถให้ค่าคำตอบวงรอบการเป็นสัดของโคได้ คือ 21.8 วัน ปัจจัยดังกล่าวมีความเกี่ยวข้องกับการเคลื่อนวงรอบการเป็นสัดในโค เช่น หากฤดูร้อนจะส่งผลให้รอบการเป็นสัดยาวขึ้นจาก 21 วัน เป็น 25 วัน หรือ หากโคเกิดความเครียดจะทำให้วงรอบการเป็นสัดยาวขึ้นกว่าปกติ ผู้วิจัยจึงใช้ปัจจัยดังกล่าวเพื่อทำนายวงรอบการเป็นสัดในโคนมโดยผ่านกระบวนการต่าง ๆ ของ

ฟัซซีลอจิก และให้ผลแสดงวงรอบการเป็นสัดของโค โดยมีขั้นตอนดังนี้

3.1 การทำฟัซซี (Fuzzification)

เป็นการนำปัจจัยเข้า และแปลงเป็นค่าระดับความเป็นสมาชิกของแต่ละเซตโดยใช้ฟังก์ชันสมาชิก (Membership Function) ในรูปแบบฟังก์ชันความเป็นสมาชิกของปัจจัยนำเข้า โดยปัจจัยที่นำมาใช้เป็นปัจจัยนำเข้ามีดังนี้

1) ฤดูกาลในประเทศไทย (Season) แบ่งออกเป็น 3 ช่วง ดังนี้ ฤดูร้อนคือช่วงเดือนที่ 3 ถึง 5 ฤดูฝนคือช่วงเดือนที่ 6 ถึง 9 และฤดูหนาวคือช่วงเดือนที่ 10 ถึง 2 และใช้ฟังก์ชันสี่เหลี่ยมคางหมู (Trapmf) เนื่องจากฤดูกาลมีช่วงเวลาที่คงที่ แสดงดังภาพที่ 4(a)

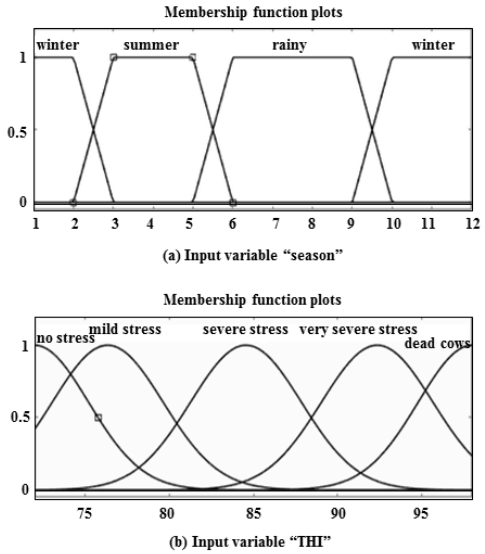
2) ค่าดัชนีอุณหภูมิและความชื้น (THI) เป็นค่าที่ได้จากการคำนวณระหว่างค่าอุณหภูมิ (Temperature) และความชื้น (Humidity) เป็นค่าที่ใช้ในการวัดความเครียดของโค ซึ่งถูกพัฒนาโดย Frank Wiersma ในปี 1990 [7] และได้แบ่งเป็นระดับต่าง ๆ ไว้ดังนี้ ค่า THI น้อยกว่า 72 โคไม่มีความเครียด (No stress) ค่า THI อยู่ระหว่าง 72-78 โคจะมีความเครียดเล็กน้อย (Mild stress) ค่า THI อยู่ระหว่าง 78-89 โคจะเครียดปานกลาง (Severe stress) ค่า THI อยู่ระหว่าง 89-98 โคจะมีความเครียดมาก (Very severe stress) และค่า THI มากกว่า 98 โคจะตายเนื่องจากเครียดจากความร้อน และใช้ฟังก์ชันรูปเกาส์เซียน (gaussmf) เนื่องจากค่า THI มีลักษณะเป็นรูปแบบ nonlinear แสดงดังภาพที่ 4(b)

จากขั้นตอนนี้จะได้ฟังก์ชันความเป็นสมาชิกของปัจจัยนำเข้าแต่ละตัว ซึ่งจะนำไปใช้ในการสร้างกฎเพื่ออนุมานหาคำตอบที่เหมาะสมในขั้นตอนถัดไป

3.2 กฎการอนุมาน (Rules of Inference)

เป็นการสร้างกฎอนุมานเพื่อนำไปหาช่วงคำตอบวงรอบการเป็นสัดในโคที่เหมาะสมตามค่าปัจจัยนำเข้าที่ถูกป้อนเข้าไป ดังตารางที่ 1

เมื่อมีข้อมูลที่ต้องการทดสอบเข้ามา โปรแกรมจะมีการประเมินค่าความเป็นสมาชิกจากปัจจัยนำเข้าแต่ละตัว



ภาพที่ 4: ฟังก์ชันความเป็นสมาชิกของปัจจัยนำเข้า

เพื่อตรวจสอบเงื่อนไขว่าตรงกับกฎใดบ้าง และส่งออกเป็นค่าข้อมูลส่งออก (Output) ซึ่งการประเมินค่าจะถูกกระทำด้วยตัวฟัซซีเซต (Fuzzy Set Operation) เช่น AND ซึ่ง AND จากทฤษฎีเซตจะได้ระดับความเป็นสมาชิกที่ต่ำที่สุดจากปัจจัยแต่ละตัวดังสมการที่ 1

ในงานวิจัยได้สร้างกฎทั้งหมด 18 กฎ โดยใช้ทฤษฎีของผลคูณคาร์ทีเซียน (Cartesian product) โดยปัจจัยที่ 1 มี 4 เซต และปัจจัยที่ 2 มี 5 เซต แต่เซตสุดท้ายของค่า THI มากกว่า 98 จะไม่นำมาคิดเนื่องจากความเครียดสูงเกินไปจนทำให้โคเสียชีวิตได้ ดังนั้นจึงได้กฎทั้งหมดเท่ากับ 4x4 กฎ และเพิ่มกฎเฉพาะ 2 กฎ คือกฎที่ 4 และ 5 ในกรณีที่ฤดูร้อนแต่โคไม่เครียดจะเกิดผลลัพธ์ได้ 3 คำตอบ

$$\mu_{A \cup B}(x)=\max[\mu_A(x),\mu_B(x)] \quad (1)$$

เมื่อ $\mu_A(x)$ คือ ค่าระดับความเป็นสมาชิกของเทอม A ในปัจจัยนำเข้า x

$\mu_B(x)$ คือ ค่าระดับความเป็นสมาชิกของเทอม B ในปัจจัยนำเข้า x

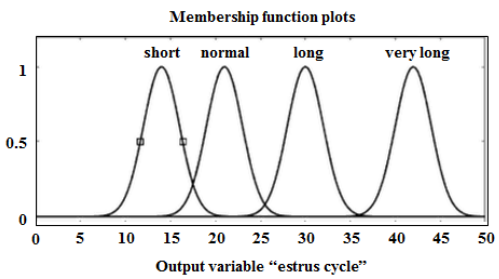
$\mu_{A \cup B}(x)$ คือ ผลรวมค่าระดับความเป็นสมาชิกของเทอม A และ B ในปัจจัยนำเข้า x

ตารางที่ 1: กฎการอนุมานเพื่อหาคำตอบ

Season	THI	Estrus (day)
summer	No Stress	long
summer	Mild Stress	long
summer	severe Stress	very long
summer	very Severe Stress	very long
summer	No Stress	short
summer	No Stress	normal
rainy	No Stress	normal
rainy	Mild Stress	long
rainy	severe Stress	very long
rainy	very Severe Stress	very long
winter	No Stress	short
winter	Mild Stress	normal
winter	severe Stress	long
winter	very Severe Stress	very long
winter	No Stress	short
winter	Mild Stress	normal
winter	severe Stress	long
winter	very Severe Stress	very long

เมื่อแต่ละกฎได้ค่าความเป็นสมาชิกมาแล้ว กฎทั้งหมดที่ผ่านการเข้าเงื่อนไขจะถูกนำมารวมเพื่อหาผลรวมค่าส่งออก (output) โดยใช้ตัวกระทำยูเนียน (Union Operation) ซึ่ง ค่าส่งออกจะมีฟังก์ชันความเป็นสมาชิกเช่นกัน เพื่อแบ่งคำตอบในระดับต่าง ๆ ในงานวิจัยได้ใช้วงรอบการเป็นสัดในโค (Estrus Cycle) ซึ่งเป็นค่าที่ใช้วัดวงรอบการเป็นสัดของโคเพื่อทำนายว่าโคจะเป็นสัดในวันใด แบ่งออกเป็น 4 ช่วง คือ มีการกลับมาเป็นสัดเร็ว (Short) คือระยะเวลาหลังเป็นสัดครั้งสุดท้ายน้อยกว่า 18 วัน ปกติ (Normal) คือระยะเวลา 18-24 วัน [1] ช้า (Long) คือระยะเวลา 28-32 วัน ช้าเกินไป (Very long) คือระยะเวลามากกว่า 42 วัน แสดงดังภาพที่ 5

จากนั้นนำผลลัพธ์ที่ได้เข้าสู่กระบวนการนำไปใช้งานจริงในขั้นตอนถัดไป



ภาพที่ 5: ฟังก์ชันความเป็นสมาชิกของตัวแปรส่งออก

3.3 การทำดีฟัซซี (Defuzzification)

เมื่อได้ค่าผลลัพธ์ของระบบเรียบร้อยแล้ว ขั้นตอนนี้จะเป็นการแปลงผลลัพธ์ที่ได้ให้อยู่ในค่าที่สามารถนำไปใช้งานได้จริง โดยการแปลงค่าผลรวมกฎให้เป็นค่าที่สอดคล้องกับการทำงานของระบบจริง วิธีการทำดีฟัซซีมีหลายรูปแบบ ที่นิยมใช้งานกันอย่างกว้างขวางคือ วิธีหาค่าจุดศูนย์ถ่วง (COG) ดังสมการที่ 2

ระบบจะนำผลรวมค่าส่งออกที่ได้จากขั้นตอนที่ 3.2 คำนวณดังสมการที่ 2 ซึ่งจะเป็นการหาค่าจุดศูนย์ถ่วงของผลรวมค่าส่งออกให้ได้มาเพียงคำตอบเดียว

$$COG = \frac{\sum_{x=a}^{b} \mu_A(x)x}{\sum_{x=a}^{b} \mu_A(x)} \quad (2)$$

เมื่อ COG คือ ค่าส่งออกของการทำดีฟัซซี

$\mu_A(x)$ คือ ค่าระดับความเป็นสมาชิกของ x

4. ผลการดำเนินงาน

ผู้วิจัยได้นำปัจจัยต่าง ๆ ที่เกี่ยวข้องกับการเป็นสัดในโค รวมถึงกฎเพื่อสร้างโมเดลขึ้นโดยใช้เทคนิคฟัซซีลอจิก และได้ทดสอบการใช้งานโมเดล รวมถึงวัดประสิทธิภาพของระบบ ซึ่งมีรายละเอียดดังนี้

4.1 การทดสอบโมเดล

ในโมเดลมีทั้งหมด 18 กฎ แต่เนื่องจากช่วงฤดูหนาวถูกแบ่งเป็น 2 ช่วง ดังนั้นกฎที่ 15-18 จึงเหมือนกับกฎที่ 11-14 ผู้วิจัยจึงสร้างโมเดลจำลองสถานการณ์ต่าง ๆเพียง 12 สถานการณ์ดังตารางที่ 2 เพื่อทดสอบเงื่อนไขของระบบ และสมมุติค่าปัจจัยนำเข้าตามเงื่อนไขในตารางที่ 2 และทดสอบระบบโดยจำลองป้อนค่าพารามิเตอร์ดังตารางที่ 3

4.2 การวัดประสิทธิภาพ

งานวิจัยได้มีการวัดประสิทธิภาพของกฎด้วยการพิจารณาค่าความถูกต้อง (Accuracy) โดยมีสูตรการวัดความถูกต้องดังสมการที่ 3

ตารางที่ 2: สถานการณ์และผลการประเมินจากผู้เชี่ยวชาญ

THI / Season	No stress	Mild Stress	Severe Stress	Very Severe Stress
Summer	long or normal	long	very long	very long
Rainy	normal	long	very long	very long
Winter	short or normal	normal	long	very long

ตารางที่ 3: ผลการทดสอบโมเดล

THI / Season	No stress	Mild Stress	Severe Stress	Very Severe Stress
Summer	21.8	23	35.6	42
Rainy	24.5	27	34.4	42
Winter	16.7	18	24.4	41.8

$$\text{ค่าความถูกต้อง} = \frac{\text{ข้อมูลที่ทำนายถูก}}{\text{ข้อมูลทั้งหมด}} \times 100 \qquad (3)$$

จากการทดลองโมเดลโดยการสมมุติค่าปัจจัยนำเข้าต่าง ๆ ตามเงื่อนไข และวัดประสิทธิภาพตารางที่ 2 และ ตารางที่ 3 โดยใช้การพิจารณาค่าความถูกต้อง พบว่าโมเดลสามารถทำนายวงรอบการเป็นสัดอยู่ในช่วงที่ถูกต้องได้ 10 กฎ จาก 12 กฎ หรือคิดเป็นร้อยละ 83.33

จากการศึกษาวิจัยพบว่าประสิทธิภาพของโมเดลขึ้นอยู่กับปัจจัยนำเข้า ฟังก์ชันที่เลือกใช้ หรือกฎที่แม่นยำ หากมีความครอบคลุมจะทำให้ประสิทธิภาพของโมเดลเพิ่มขึ้น

5. สรุปผลและข้อเสนอแนะ

งานวิจัยนี้ได้นำเสนอการใช้เทคนิคฟัซซีลอจิกในการทำนายวงรอบการเป็นสัดของโค เพื่อช่วยให้การจับสัดของเกษตรกรมีความแม่นยำเพิ่มมากขึ้นกว่าการใช้วิธีสังเกตจากพฤติกรรมโค ซึ่งในบางครั้งไม่แสดงพฤติกรรมออกมา

ในงานวิจัยได้ใช้ปัจจัยนำเข้าที่เกี่ยวข้องกับวงรอบการเป็นสัดในโค ผ่านกระบวนการฟัซซีลอจิก และโมเดลให้ผลการทำนายมีความถูกต้องร้อยละ 83.33

จากการศึกษาพบว่าประสิทธิภาพของโมเดลขึ้นอยู่กับปัจจัยนำเข้า การเลือกใช้ฟังก์ชันสมาชิกที่เหมาะสมกับลักษณะของข้อมูล และความครอบคลุมของกฎที่สร้างขึ้น

ในการวิจัยครั้งนี้ได้ใช้ค่า THI เป็นหนึ่งในปัจจัยนำเข้า ซึ่งได้จากการคำนวณระหว่างค่าอุณหภูมิและค่าความชื้น และต้องใช้อุปกรณ์ในการวัด ดังนั้นในอนาคตผู้วิจัยจะมีการนำค่าอุณหภูมิและค่าความชื้นมาเป็นปัจจัยนำเข้าเพื่อให้โมเดลสามารถคำนวณค่า THI โดยอัตโนมัติจากอุปกรณ์ และสามารถนำไปประยุกต์เป็นแอพลิเคชั่นเพื่อแจ้งเตือนไปยังเกษตรกรได้

เอกสารอ้างอิง

[1] วิโรจน์ ภัทรจิดา, *อาหารและฟาร์มโคนมสมัยใหม่*, โรงพิมพ์มหาวิทยาลัยขอนแก่น, ขอนแก่น, 2546.

[2] สุณีรัตน์ เอี่ยมละมัย, *การสืบพันธุ์ในโค*, หจก.ขอนแก่นการพิมพ์, ขอนแก่น, 2553.

[3] Leandro dos Anjos et al, “Improving detection of dairy cow estrus using fuzzy logic” Scientia Agricola, vol. 67, no. 5, pp. 503-509, 2010.

[4] Leandro Ferreira et al, “Development of Algorithm Using fuzzy Logic to predict Estrus in Dairy Cows: Part I” *Agricultural Engineering International: the CIGR Ejournal*, vol. 9, 2007.

[5] Abootorabi Zarchi et al, “Improving Oestrus Detection in Dairy Cows by Combining Statistical Detection with Fuzzy Logic Classification” *Proceedings Workshop on Advanced Control and Diagnosis*, 2009.

[6] Vilem Novak et al, “Mathematical Principles of Fuzzy Logic” *Kluwer Academic*, 1999.

[7] John Moran, “Appendix1 Temperature Humidity Index” *TROPOCAL DAIRY FARMING*, *Landlinks Press*, *Victoria*, 2005.

[8] วิโรจน์ ภัทรจิดา, *โคนม*, โรงพิมพ์มหาวิทยาลัยขอนแก่น, ขอนแก่น, 2546.

[9] กฤตพล สมมาตย์, *วิทยาศาสตร์และเทคโนโลยีการผลิตโค-กระบือ*, โรงพิมพ์มหาวิทยาลัยขอนแก่น, ขอนแก่น, 2550.

[10] จรัญ จันทลักขณา ประเสริฐ เหรียญแก้ว และ บุญเหลือ เร่งศิริกุล, *การผลิตโคเนื้อ*, สำนักพิมพ์มหาวิทยาลัยเกษตรศาสตร์, กรุงเทพมหานคร, 2515.

การพยากรณ์อุปสงค์เพื่อลดปัญหาการขาดแคลนสินค้าคงคลัง กรณีศึกษาบริษัทเวชภัณฑ์ทางการแพทย์

Forecasting for Inventory Stock-out Reduction Case Study: Pharmaceutical Company

อังสุมา ทองทิว (Angsuma Thongthiew)[1] ฉัฐไชย์ ลีนาวงศ์ (Chartchai Leenawong)[2] และบัณฑิต วรรธนาภา (Bunthit Watanapa)[3]

[1,3]สาขาวิชาระบบสารสนเทศทางธุรกิจ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี

[2]คณะวิทยาศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

[1]angsuma.t@mail.kmutt.ac.th, [2]cleenawong@gmail.com and [3]Bunthit@sit.kmutt.ac.th

บทคัดย่อ

งานวิจัยนี้ได้ศึกษารูปแบบของชุดข้อมูลอุปสงค์ในอดีตจากกรณีศึกษาบริษัทเวชภัณฑ์ทางการแพทย์ โดยมีวัตถุประสงค์เพื่อลดต้นทุนสินค้าคงคลังจากการขาดแคลนสินค้า ที่มีสาเหตุหลักมาจากการพยากรณ์อุปสงค์ที่คลาดเคลื่อนทำให้เกิดการสั่งซื้อสินค้าในปริมาณที่ไม่เหมาะสม ผู้วิจัยพบว่าข้อมูลอุปสงค์ในอดีตของบริษัทเวชภัณฑ์ทางการแพทย์ที่พยากรณ์ด้วยวิธี Holt's Exponential Smooting นั้นมีอิทธิพลของฤดูกาล ผู้วิจัยจึงทำการปรับปรุงค่าพยากรณ์อุปสงค์โดยทำการขจัดอิทธิพลของฤดูกาลออกจากข้อมูล (Deseasonalized Data) ก่อนนำไปพยากรณ์ด้วยวิธี Holt's Exponential Smoothing และทำการพยากรณ์ข้อมูลอุปสงค์ที่มีอิทธิพลของฤดูกาลด้วยวิธี Holt-Winters Exponential Smoothing อีกหนึ่งวิธี งานวิจัยนี้ใช้ค่าความคลาดเคลื่อนในการเปรียบเทียบ ได้แก่ Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE) และ Root Mean Squre Error (RMSE) เพื่อให้ทราบความแตกต่างระหว่างข้อมูลจริงกับข้อมูลที่ได้จากการพยากรณ์ ผลจากการวิจัยพบว่าค่าความคลาดเคลื่อนที่ได้จากการพยากรณ์ด้วยวิธี Holt's Exponential Smoothing บนข้อมูลที่ขจัดอิทธิพลของฤดูกาลออก (Deseasonalized Data) มีค่าต่ำที่สุด กล่าวได้ว่าการพยากรณ์ด้วยวิธีนี้มีความเหมาะสมกับรูปแบบของชุดข้อมูลที่มีอิทธิพลของฤดูกาล ซึ่งจะนำไปสู่การลดปัญหาการขาดแคลนสินค้าเนื่องจากการพยากรณ์ที่ผิดพลาดได้เป็นอย่างดี

คำสำคัญ: การจัดการสินค้าคงคลัง ข้อมูลการพยากรณ์ ค่าความถูกต้องของการพยากรณ์ ต้นทุนสินค้าคงคลัง

Abstract

In this research, the demand data from a pharmaceutical company was examined for finding the best forecasting method with the objective of reducing inventory stock-out costs. One main cause of stock-out remains the inaccurate forecasts leading to inappropriate order quantities. This research attempted to fit the case study's demand data displaying seasonality more accurately by incorporating the seasonal effects into Holt's Exponential Smoothing method from previous work. Both the original Holt's and this modified seasonal Holt's as well as Holt-Winters Exponential Smoothing method are all applied to the same demand data to compare for their accuracies. The results show that the seasonal Holt's Exponential Smoothing on the deseasonlized data gives the lowest MAE, MAPE, and RMSE and thus the best method. If this method is implemented, the inventory costs due to inaccurate forecasts shall be reduced properly.

Keyword: Data Forecasting, Forecasting Accuracy, Inventory Management, Inventory Cost

1. บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

ธุรกิจอุตสาหกรรมนั้นมีการแข่งขันเกิดขึ้นตลอดเวลา ข้อมูลถือเป็นหัวใจสำคัญที่ส่งผลกระทบต่อธุรกิจทั้งทางตรง

และทางอ้อม โดยเฉพาะข้อมูลในการจัดการสินค้าคงคลัง การขาดข้อมูลที่ถูกต้องในการจัดการสินค้าคงคลังอาจทำให้บริษัทสูญเสียรายได้และโอกาสทางการค้า รวมไปถึงการขาดแคลนสินค้าคงคลังที่ก่อให้เกิดความไม่พึงพอใจ ทำให้ลูกค้าขาดความเชื่อมั่นในประสิทธิภาพของบริษัท และอาจเป็นการเปิดโอกาสให้คู่แข่งที่มีความพร้อมมากกว่าในที่สุด

การพยากรณ์ข้อมูลเพื่อใช้ในการจัดการสินค้าคงคลังมีบทบาทสำคัญในการลดปัญหาการขาดแคลนสินค้าคงคลังซึ่งเกิดได้จากหลายสาเหตุดังที่ Joachim and Stölzle [1] ได้ทำการศึกษาและเสนอสาเหตุของการขาดแคลนสินค้าคงคลังว่าเกิดได้จากการทำงานผิดพลาดของพนักงาน เช่น จัดเก็บข้อมูลสินค้าผิดพลาด พนักงานไม่แจ้งเมื่อสินค้าหมด หรือเกิดความผิดพลาดจากตัวแทนจำหน่าย ได้แก่ การจัดส่งสินค้าไม่ถูกต้องหรือล่าช้าไปจากที่ตกลงและสำคัญที่สุดคือเกิดจากการสั่งซื้อสินค้าในปริมาณที่ไม่เหมาะสมเนื่องมาจากการพยากรณ์ที่คลาดเคลื่อนจนต้นทุนเพิ่มขึ้นซึ่งเป็นสาเหตุหลักของปัญหา

1.2 วัตถุประสงค์ของการวิจัย

เพื่อศึกษาและคัดเลือกวิธีการพยากรณ์ที่เหมาะสมกับรูปแบบของชุดข้อมูลอุปสงค์ในอดีตเพื่อลดค่าความคลาดเคลื่อนของการพยากรณ์อันจะนำไปสู่การลดต้นทุนสินค้าคงคลังจากการขาดแคลนสินค้า (Inventory Stock-outs) ที่มีสาเหตุหลักมาจากการพยากรณ์อุปสงค์ที่คลาดเคลื่อนทำให้เกิดการสั่งซื้อในปริมาณที่ไม่เหมาะสม

1.3 ขอบเขตการวิจัย

1.3.1 ข้อมูลที่ใช้ในการวิจัยเป็นข้อมูลอุปสงค์ในอดีตตั้งแต่ พ.ศ. 2554 – 2555 ของบริษัทเวชภัณฑ์ทางการแพทย์รวมระยะเวลา 24 เดือน

1.3.2 งานวิจัยนี้จะทำการปรับปรุงค่าพยากรณ์อุปสงค์เพื่อเปรียบเทียบกับค่าพยากรณ์อุปสงค์จากงานวิจัยอ้างอิงของ Mekel et al [2] โดยทำการพยากรณ์ค่าอุปสงค์ด้วย วิธี Holt's Exponential บนข้อมูลที่มีอิทธิพลของฤดูกาล วิธี Holt's Exponential บนข้อมูลที่ขจัดอิทธิพลของฤดูกาลออก (Deseasonalized Data) และวิธี Holt-Winters Exponential ผ่านโปรแกรม ForecastX

1.3.3 ทดสอบค่าความคลาดเคลื่อนจากการเปรียบเทียบความแตกต่างระหว่างค่าจริงกับค่าพยากรณ์ โดยใช้ MAE, MAPE และ RMSE

2. ทฤษฎีพื้นฐานและงานวิจัยที่เกี่ยวข้อง

2.1 การจัดการสินค้าคงคลัง (Inventory Management)

คำนาย [3] กล่าวว่า การจัดการสินค้าคงคลังสามารถดำเนินการได้หลายวิธี โดยต้องมีการเลือกวิธีการจัดการที่เหมาะสมโดยพิจารณาจากต้นทุนและค่าใช้จ่ายในการจัดการดังนี้

2.1.1 ต้นทุนของสินค้าคงคลัง (Inventory Cost) เป็นมูลค่าของสินค้าคงคลังทั้งหมดหรือผลคูณระหว่างปริมาณสินค้าคงคลังกับต้นทุนต่อหน่วยของสินค้าคงคลัง

2.1.2 ต้นทุนค่าใช้จ่ายในการสั่งซื้อ (Cost of Ordering) ได้แก่ค่าใช้จ่ายในการสั่งซื้อ ค่าขนส่งสินค้านำเข้า ค่าภาษี ฯลฯ เป็นค่าใช้จ่ายคงที่ต่อหนึ่งคำสั่งซื้อ ดังนั้น การสั่งซื้อหรือสั่งผลิตที่มีปริมาณครั้งละมากๆ จะทำให้ประหยัดค่าใช้จ่ายการสั่งซื้อรวมเพราะลดจำนวนครั้งในการสั่งซื้อลง

2.1.3 ต้นทุนค่าใช้จ่ายในการรักษาสินค้าคงคลัง (Cost of Carrying Inventory) แบ่งออกเป็น 3 ประเภท คือ ค่าเก็บรักษา (Storage Cost) ค่าความเสี่ยง (Risk Cost) และค่าของทุน (Capital Cost)

2.1.4 ต้นทุนของสินค้าคงคลังเพื่อความปลอดภัย (Cost of Safety Stock) เป็นค่าใช้จ่ายที่เกิดขึ้นจากการมีสินค้าคงคลังขั้นต่ำไม่เป็นไปตามที่พยากรณ์ไว้ หรือค่าใช้จ่ายเนื่องจากการมีสินค้าคงคลังขั้นต่ำเกินไป

2.1.5 ต้นทุนค่าใช้จ่ายเนื่องจากสินค้าขาดมือ (Stock-out Costs) เป็นค่าใช้จ่ายที่เกิดขึ้นกรณีที่ธุรกิจไม่มีสินค้าส่งมอบให้กับลูกค้าหรือกำไรที่ควรจะได้จากการเสียโอกาสในการขาย เนื่องจากค่าใช้จ่ายประเภทนี้มักมีมูลค่าสูง ดังนั้นธุรกิจมักจะไม่ยอมให้สินค้าคงคลังขาดแคลนโดยใช้วิธีการเก็บรักษาสินค้าคงคลังขั้นต่ำ

จากข้อมูลข้างต้นแสดงให้เห็นถึงความสำคัญของการพยากรณ์ค่าอุปสงค์เพื่อนำไปพิจารณาร่วมกับการกำหนดระดับสินค้าคงคลังปลอดภัย (Safety Stock) และการกำหนดจุดสั่งซื้อใหม่ที่เหมาะสม (Reorder Point) นำไปใช้คำนวณหาปริมาณการสั่งซื้อแบบประหยัด (EOQ) เพื่อลดปัญหาการขาดแคลนสินค้าคงคลังอันเป็นค่าใช้จ่ายที่มีมูลค่าสูง ดังที่ได้ศึกษาจากทฤษฎีและกรณีศึกษาต่อไปนี้

Pibernik [4] จัดการกับสถานการณ์การขาดแคลนสินค้าคงคลังด้วยระบบ Order Fulfillment โดยอาศัยตรรกะในการพิจารณาต่างๆ เช่น จากลำดับการทำสัญญารับการสั่งซื้อจากลูกค้า (Order Promising) การจัดสรรแบบมาก่อนได้ก่อน (First Come First Serve) และการทำ Rank-Based Allocation จัดสรรตามความต้องการและระดับความสำคัญของลูกค้าซึ่งรวมถึงข้อตกลงและความสัมพันธ์ที่มีต่อกัน

การกำหนดระดับสินค้าคงคลังที่ปลอดภัย (Safety Stock Level) จะใช้การคำนวณร่วมกับการกำหนดจุดสั่งซื้อใหม่ที่เหมาะสม (Reorder Point) ในงานวิจัยของ นิรชรและกาญจ์นภา [5] ได้รวบรวมและวิเคราะห์ข้อมูล เพื่อหาวิธีการปรับปรุงการสั่งซื้อถ่านหินของบริษัทกรณีศึกษาและเพื่อเปรียบเทียบค่าใช้จ่ายที่เกิดขึ้นจริง จากนั้นทำการรวบรวมค่าใช้จ่ายในการสั่งซื้อ (Cost of Ordering) และค่าใช้จ่ายในการถือครองสินค้า (Cost of Carrying) เพื่อใช้คำนวณระดับสินค้าคงคลังที่ปลอดภัยของถ่านหินและจุดสั่งซื้อใหม่ที่เหมาะสม

2.2 วิธีการพยากรณ์ (Forecasting Method)

นฤมล [6] ได้อธิบายถึงวิธีการพยากรณ์และสมการพื้นฐานที่ผู้วิจัยเลือกนำมาใช้ในงานวิจัยนี้ดังนี้

2.2.1 Holt's Exponential smoothing method เป็นการพยากรณ์ข้อมูลที่เหมาะสมกับข้อมูลที่มีลักษณะแนวโน้มเชิงเส้น (Trend) ซึ่งจะเป็นเทคนิคที่ให้ความสำคัญกับข้อมูลล่าสุดมากที่สุด ดังนั้นจึงมีค่าคงที่การทำให้เรียบ 2 ค่าคือ α (alpha) และ β (beta) โดยมีสมการดังนี้

$$L_t=\alpha y_t+(1-\alpha)(L_{t-1}+b_{t-1}) \quad (1)$$
$$b_t=\beta(L_t-L_{t-1})+(1-\beta)b_{t-1} \quad (2)$$
$$F_{t+m}=(L_t+b_tm) \quad (3)$$

โดยที่

y = ค่าของข้อมูลที่เวลาต่างๆ

F = ค่าการพยากรณ์ของข้อมูล

t = เวลา

α = ค่าที่ทำให้เรียบระหว่างค่าจริงกับค่าประมาณ

β = ค่าที่ทำให้เรียบระหว่างแนวโน้มจริงกับแนวโน้มประมาณ

α β γ จะมีค่าอยู่ระหว่าง 0 ถึง 1

L = ค่าประมาณของ level ของ series ที่เวลา t

b = ค่าประมาณของความชันของ series ที่เวลา t

m = จำนวนของคาบที่ต้องการพยากรณ์ล่วงหน้า

2.2.2 Holt-Winters Exponential smoothing Method เป็นวิธีที่เหมาะสมกับข้อมูลที่ทั้งแนวโน้ม (Trend) และอิทธิพลของฤดูกาล (Seasonality) มีสมการพื้นฐาน ดังนี้

$$L_t=\alpha\frac{y_t}{s_{t-s}}+(1-\alpha)(L_{t-1}+b_{t-1}) \quad (4)$$
$$b_t=\beta(L_t-L_{t-1})+(1-\beta)b_{t-1} \quad (5)$$
$$S_t=\gamma\frac{y_t}{L_t}+(1-\gamma)S_{t-s} \quad (6)$$
$$F_{t+m}=(L_t+b_tm)S_{t-s+m} \quad (7)$$

โดยที่

y = ค่าของข้อมูลที่เวลาต่างๆ

F = ค่าการพยากรณ์ของข้อมูล

t = เวลา

α = ค่าที่ทำให้เรียบระหว่างค่าจริงกับค่าประมาณ

β = ค่าที่ทำให้เรียบระหว่างแนวโน้มจริงกับแนวโน้มประมาณ

γ = ค่าที่ทำให้เรียบระหว่างค่าฤดูกาลจริงกับค่าประมาณของฤดูกาล

α β และ γ มีค่าอยู่ระหว่าง 0 ถึง 1

L = ค่าประมาณของ level ของ series ที่เวลา t

b = ค่าประมาณของความชันของ series ที่เวลา t

S = ส่วนประกอบของ Seasonality

S = ช่วงระยะเวลาของ Seasonality

2.3 ค่าความคลาดเคลื่อนของการพยากรณ์ (Forecasting Accuracy)

อัจฉรา [7] ได้กล่าวถึงการวัดความผิดพลาดจากการพยากรณ์ว่าจะตรวจสอบได้โดยพิจารณาจากค่าความคลาดเคลื่อน ซึ่งเราสามารถเลือกใช้ตัวตรวจสอบได้ดังนี้

2.3.1 ค่าเฉลี่ยความผิดพลาดสัมบูรณ์ (Mean Absolute Error: MAE) ใช้วัดค่าความแตกต่างระหว่างค่าจริงกับค่าพยากรณ์ หาก MAE มีค่าน้อยแสดงว่าค่าพยากรณ์ใกล้เคียงค่าจริง มีความคลาดเคลื่อนต่ำ สามารถคำนวณได้ดังสมการที่ (8)

$$MAE=\frac{\sum|(A_t-F_t)|}{n} \quad (8)$$

โดยที่

A_t = มูลค่าที่แท้จริงในระยะเวลานั้น

F_t = ค่าพยากรณ์ในระยะเวลานั้น

n = จำนวนระยะเวลา

2.3.2 รากที่สองค่าเฉลี่ยความผิดพลาดกำลังสอง (Root Mean Square Error: RMSE) ใช้วัดค่าความแตกต่างระหว่างค่าจริงกับค่าพยากรณ์ หาก RMSE มีค่าน้อยแสดงว่าค่าพยากรณ์ใกล้เคียงค่าจริง มีความคลาดเคลื่อนต่ำ สามารถคำนวณได้ดังสมการที่ (9)

$$RMSE = \sqrt{\frac{\sum(A_t - F_t)^2}{n}} \quad (9)$$

โดยที่

A_t = มูลค่าที่แท้จริงในระยะเวลานั้น

F_t = ค่าพยากรณ์ในระยะเวลานั้น

n = จำนวนระยะเวลา

2.3.3 ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ (Mean Absolute Percentage Error: MAPE) ใช้วัดความแตกต่างระหว่างค่าจริงกับค่าพยากรณ์ หาก MAPE มีค่าน้อยแสดงว่าค่าพยากรณ์ใกล้เคียงค่าจริง สามารถคำนวณได้ดังสมการที่ (10)

$$MAPE = \frac{\sum|(A_t - F_t)/A_t|}{n} \quad (10)$$

การปรับปรุงการพยากรณ์ (Forecasting)ให้มีค่าความถูกต้องสูงหรือมีค่าความคลาดเคลื่อนต่ำนั้น จะต้องทำก่อนการวิเคราะห์หาปริมาณการสั่งซื้อที่เหมาะสมหรือปริมาณการสั่งซื้อที่ประหยัด (Economic Order Quantity : EOQ) ซึ่งเป็นวิธีที่ช่วยในการกำหนดปริมาณสินค้าที่ต้องการสั่งซื้อในแต่ละครั้งเพื่อให้เกิดต้นทุนหรือค่าใช้จ่ายที่ต่ำที่สุด โดยมีต้นทุนที่สำคัญอยู่ 2 ค่าได้แก่ ต้นทุนการเก็บรักษา (Cost of Carrying) และต้นทุนการสั่งซื้อ (Cost of Ordering) การใช้วิธีนี้จะมีข้อจำกัดหลายประการ เช่น ปัจจัยด้านต้นทุนทั้งหลายต้องคงที่ เหมาะสำหรับระบบปิด ระพีพรรณ [8]

3. วิธีการดำเนินการวิจัย

สำหรับงานวิจัยหัวข้อ การพยากรณ์อุปสงค์เพื่อลดต้นทุนสินค้าคงคลัง กรณีศึกษาบริษัทเวชภัณฑ์ทางการแพทย์ ได้มีกระบวนการดำเนินการประกอบไปด้วยการศึกษารูปแบบของชุดข้อมูลอุปสงค์ในอดีตของกรณีศึกษาบริษัทเวชภัณฑ์ทางการแพทย์ เพื่อเลือกวิธีการพยากรณ์ที่เหมาะสมกับรูปแบบของข้อมูล เพื่อนำไปสู่การสั่งซื้อสินค้าในปริมาณที่เหมาะสมและลดต้นทุนสินค้าคงคลังเนื่องจากการขาดแคลนสินค้าคงคลัง

ซึ่งจากการศึกษาทฤษฎีที่เกี่ยวข้อง พบว่ารูปแบบของชุดข้อมูลที่จะนำมาใช้พยากรณ์แบบอนุกรมเวลานั้นประกอบด้วยค่าต่างๆที่ส่งผลต่อวิธีการพยากรณ์ที่แตกต่างกัน

3.1 ขั้นตอนการดำเนินการวิจัย

3.1.1 ศึกษาข้อมูลอุปสงค์ในอดีตตั้งแต่ พ.ศ. 2554 - 2555 ของบริษัทเวชภัณฑ์ทางการแพทย์เป็นระยะเวลา 24 เดือน

3.1.2 ทำการขจัดอิทธิพลของฤดูกาลออกจากข้อมูล (Deseasonalized Data)

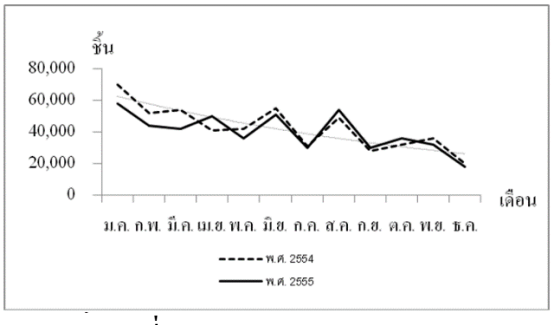
3.1.3 นำข้อมูลหลังการขจัดอิทธิพลของฤดูกาลไปพยากรณ์ด้วยวิธี Holt's Exponential และทำการพยากรณ์ด้วยวิธี Holt-Winters Exponential

3.1.4 หาค่าความคลาดเคลื่อนของข้อมูลที่ได้จากการพยากรณ์ทั้ง 2 วิธี เพื่อเปรียบเทียบกับค่าความคลาดเคลื่อนที่ได้จากการพยากรณ์ของงานวิจัยอ้างอิง

3.1.5 ทำการเปรียบเทียบค่าความคลาดเคลื่อน โดยพิจารณาจาก MAE, RMSE และ MAPE และสรุปผล

3.2 ชุดข้อมูลอุปสงค์ของบริษัทเวชภัณฑ์ทางการแพทย์

งานวิจัยนี้ใช้ข้อมูลอุปสงค์ในอดีตตั้งแต่ พ.ศ. 2554 - 2555 ของบริษัทเวชภัณฑ์ทางการแพทย์เป็นระยะเวลา 24 เดือนซึ่งผู้วิจัยพบว่าชุดข้อมูลอุปสงค์ในอดีตที่นำมาพยากรณ์นั้นมีอิทธิพลของฤดูกาลมาเกี่ยวข้อง ดังกราฟในรูปที่ 1 แสดงให้เห็นว่าข้อมูลอุปสงค์ของ พ.ศ.2554 และ พ.ศ.2555 มีค่าใกล้เคียงกันในช่วงเดือนเดียวกันของทั้งสองปี



ภาพที่ 1: ค่าอุปสงค์ใน พ.ศ. 2554 – 2555 ของบริษัทเวชภัณฑ์ทางการแพทย์

ข้อมูลที่มีการเปลี่ยนแปลงตามฤดูกาลจนมีแบบแผนเดียวกันนั้นสามารถแทนค่าได้ในรูปของดัชนีฤดูกาล (Seasonal Index) ซึ่งผู้วิจัยได้ทำการหาค่าดัชนีฤดูกาลนี้จากการนำค่าอุปสงค์ของแต่ละเดือนมาคูณกับจำนวนช่วงเวลาได้แก่ 12 เดือนและหารด้วยค่าอุปสงค์รวม จะได้ค่าดัชนีฤดูกาลของแต่ละปีจากนั้นจึงนำมาหาค่าเฉลี่ยเพื่อให้ได้ค่าดัชนีฤดูกาลของชุดข้อมูลดังที่แสดงในตารางที่ 1

ตารางที่ 1: ค่าดัชนีฤดูกาลจากข้อมูลอุปสงค์ใน พ.ศ. 2554 - 2555 ของบริษัทเวชภัณฑ์ทางการแพทย์

Period	Seasonal Index
Jan	1.55
Feb	1.16
Mar	1.16
Apr	1.11
May	0.94
Jun	1.28
July	0.74
Aug	1.25
Sep	0.7
Oct	0.83
Nov	0.82
Dec	0.46

4. ผลการวิจัย

4.1 ผลการวิจัย

ในงานวิจัยของ Mekel et al. [2] ได้นำข้อมูลอุปสงค์ในอดีตตั้งแต่ พ.ศ. 2554 - 2555 ของบริษัทเวชภัณฑ์ทางการแพทย์เป็นระยะเวลา 24 เดือนมาทำการพยากรณ์อุปสงค์ด้วยวิธี Holt's Exponential Smoothing แต่วิธี Holt's Exponential Smoothing นั้นเหมาะกับการพยากรณ์ข้อมูลที่ไม่มีอิทธิพลของฤดูกาล ผู้วิจัยจึงนำค่าดัชนีฤดูกาลจากตารางที่ 1 ไปหารออกจากค่าอุปสงค์เดิมทำให้ได้ข้อมูลที่มีการขจัดอิทธิพลของฤดูกาลออกไป (Deseasonalized data) ดังแสดงในตารางที่ 2

ตารางที่ 2: ค่าอุปสงค์ก่อนและหลังขจัดอิทธิพลของฤดูกาล

Period	Actual Demand (Original)		Actual Demand (Deseasonalized)	
	2011	2012	2011	2012
January	70,000	58,000	45,252	37,495
February	52,000	44,000	44,807	37,914
March	54,000	42,000	46,588	36,235
April	41,000	50,000	37,071	45,209
May	42,000	36,000	44,534	38,172
June	55,080	51,000	42,895	39,717
July	31,000	30,000	41,956	40,603
August	49,000	54,000	39,201	43,201
September	28,000	30,000	39,796	42,639
October	32,000	36,000	38,766	43,611
November	36,000	32,000	43,762	38,900
December	20,000	18,000	43,498	39,148

4.1.1 การพยากรณ์ค่าอุปสงค์

เนื่องจาก Mekel et al. [2] ได้พยากรณ์อุปสงค์จากชุดข้อมูลที่มีอิทธิพลของฤดูกาลด้วยวิธี Holt's Exponential Smoothing ผู้วิจัยจึงได้นำชุดข้อมูลอุปสงค์ที่ผ่านการขจัดอิทธิพลของฤดูกาล (Deseasonalized data) จากตารางที่ 2 มาทำการพยากรณ์ด้วยวิธี Holt's Exponential Smoothing และทำการพยากรณ์ชุดข้อมูลที่มีอิทธิพลของฤดูกาลด้วยวิธี Holt-Winters Exponential Smoothing ซึ่งเป็นวิธีที่เหมาะสมกับข้อมูลที่มีทั้งแนวโน้ม (Trend) และอิทธิพลของฤดูกาล (Seasonality) เข้ามาเกี่ยวข้อง นฤมล [8] โดยนำข้อมูลอุปสงค์ตั้งแต่ พ.ศ. 2554 - 2555 เป็นจำนวน 24 เดือนมาพยากรณ์อุปสงค์ 24 เดือนถัดไปด้วยโปรแกรม ForecastX โดยใช้ทั้งสองวิธีและนำผลที่ได้จากการพยากรณ์แต่ละวิธีมาเปรียบเทียบกัน ดังแสดงในตารางที่ 3

ตารางที่ 3: ค่าพยากรณ์อุปสงค์ด้วยวิธีการต่างๆ โดยโปรแกรม ForecastX

Period	Holt's on Original Data	Holt's on Deseasonalized Data	Holt-Winters
	2013	2013	2013
Jan 13	70,000	45,252	61,804
Feb 13	70,000	45,252	46,911
Mar 13	54,880	44,878	46,858
Apr 13	49,821	46,208	47,458
May 13	37,880	38,837	38,970
Jun 13	34,693	41,733	53,517
July 13	46,159	42,187	31,516
Aug 13	32,659	41,750	53,278
Sep 13	41,981	39,310	30,228
Oct 13	29,754	38,808	35,382
Nov 13	27,803	37,979	34,246
Dec 13	31,389	42,033	19,574
Jan 14	20,491	43,848	61,918
Feb 14	47,933	39,447	47,025
Mar 14	49,566	37,570	46,972
Apr 14	47,204	35,492	47,572
May 14	51,730	42,377	39,084
Jun 14	41,365	39,899	53,631
July 14	48,532	39,792	31,630
Aug 14	34,351	40,475	53,392
Sep 14	47,794	42,961	30,342
Oct 14	34,501	43,541	35,496
Nov 14	33,143	44,373	34,360
Dec 14	29,926	40,566	19,688

4.1.2 คำนวณค่าความคลาดเคลื่อนจากการพยากรณ์

หลังจากทำการพยากรณ์ด้วย โปรแกรม ForecastX แล้วนั้น เราสามารถหาค่าความคาดเคลื่อนได้จากการวัดค่าความแตกต่างระหว่างค่าจริงกับค่าที่ได้จากการพยากรณ์ในแต่ละวิธี ดังแสดงในตารางที่ 4

ตารางที่ 4: ค่าความคลาดเคลื่อนของการพยากรณ์อุปสงค์

Method	MAPE	MAE	RMSE
Holt's	31.08%	11,618.90	14,369.54
Holt's (Deseasonlized)	6.71%	2,732.71	3,954.01
Holt-Winters	7.93%	3,135.47	4,006.37

5. สรุปผล

5.1 สรุปผลการศึกษา

ผู้วิจัยได้เลือกใช้วิธี Holt's Exponential Smoothing กับชุดข้อมูลที่ได้มีการขจัดค่าอิทธิพลฤดูกาลออกและวิธี Holt-Winters Exponential Smoothing ซึ่งเหมาะกับข้อมูลที่มีอิทธิพลของฤดูกาลเข้ามาเกี่ยวข้องเพื่อไปเปรียบเทียบกับวิธี Holt's Exponential Smoothing ในงานวิจัยของ Mekel et al. [2] จากนั้นเปรียบเทียบค่าความคลาดเคลื่อนเพื่อหาวิธีพยากรณ์ที่มีประสิทธิภาพดังแสดงในตารางที่ 4 ซึ่งพบว่า วิธี Holt's Exponential Smoothing บนชุดข้อมูลที่ได้มีการขจัดค่าอิทธิพลฤดูกาลออก มีค่า Mean Absolute Error = 2,732.71 Mean Absolute Percentage Error = 6.71% และ Root Mean Square Error = 3,954.01 ซึ่งต่ำกว่าการพยากรณ์ทุกวิธี

ทำการเปรียบเทียบเพื่อเลือกวิธีการที่เหมาะสมด้วยการหาค่าความคลาดเคลื่อนของการพยากรณ์ ศศิธรและกัลยา [9] กล่าวว่าถ้าค่าความคลาดเคลื่อนสูงแสดงว่าวิธีการพยากรณ์นั้นให้ความผิดพลาดในการพยากรณ์มาก ส่วนวิธีการใดมีค่าความคลาดเคลื่อนต่ำ แสดงว่าวิธีการนั้นให้ความผิดพลาดในการพยากรณ์น้อยอันเป็นวิธีที่เหมาะสมและมีความถูกต้องมากกว่า

5.2 อภิปรายผล

ผู้วิจัยได้เลือกวิธี Holt's Exponential Smoothing ร่วมกับการขจัดอิทธิพลฤดูกาลของข้อมูล (Deseasonalized data) มาใช้ในการปรับปรุงค่าพยากรณ์เนื่องจากรูปแบบของข้อมูลนั้นมีอิทธิพลของฤดูกาล แม้ว่าที่จริงแล้วสามารถเลือกพยากรณ์ด้วยวิธี Holt-Winters Exponential Smoothing ได้เนื่องจากถูกสร้างมาสำหรับข้อมูลที่มีสองปัจจัยในการพยากรณ์ ได้แก่ แนวโน้ม (Trend) และ อิทธิพลฤดูกาล แต่สำหรับข้อมูลชุดนี้การพยากรณ์ด้วยวิธี Holt-Winters Exponential Smoothing มีค่าความคลาดเคลื่อนที่สูงกว่า นั่นหมายความว่าหากต้องการเลือกวิธีพยากรณ์ให้มีความเหมาะสมสูงสุด ควรทดลองทำการพยากรณ์ด้วยวิธีที่หลากหลายแล้วนำมาเปรียบเทียบกันเพื่อหาค่าความคลาดเคลื่อนระหว่างค่าจริงกับค่าพยากรณ์

ค่าที่ได้จากปรับปรุงค่าพยากรณ์อุปสงค์นั้นจะนำไปคำนวณหาระดับสินค้าคงคลังปลอดภัย (Safety Stock) และกำหนดจุดสั่งซื้อใหม่ (Reorder Point) ให้ถูกต้องเพื่อลดการขาดแคลนสินค้าคงคลัง และต้องคำนึงถึงปริมาณการสั่งซื้อแบบประหยัดเพื่อให้เกิดความคุ้มค่า ลดต้นทุนการจัดการสินค้าคงคลัง เพิ่มผลกำไรให้กับธุรกิจ

เอกสารอ้างอิง

[1] Joachim, C.F., Ehrenthal & Wolfgang, Stölzle. (2013). An examination of the causes for retail stockouts. *International Journal of Physical Distribution Chair of Logistics Management, 43*(1), 54-69.

[2] Mekel, Christine, Anantadjaya, Samuel & Lahindah, Laura. (2014). Stock out Analysis: An Empirical Study on Forecasting, Re-Order Point and Safety Stock Level at PT Combiphar, Indonesia. *Review of Integrative Business and Economics Research, 3*(1), 52-64.

[3] คำนาย อภิปรัชญาสกุล. (2553). การจัดการสินค้าคงคลัง. กรุงเทพฯ: โฟกัสมีเดีย แอนด์ พับลิชชิ่ง จำกัด

[4] Richard, Pibernik. (2006). Managing stock-outs effectively with order fulfillment systems. *Journal of Manufacturing Technology Management, 17*(6), 721-736

[5] นิรชร วะชุม และ กาญจ์นภา อมรัชกุล. (2559). ตัวแบบปริมาณการสั่งซื้อแบบพลวัตเมื่อระยะเวลานาไม่แน่นอน: กรณีศึกษาถ่านหินอุตสาหกรรมปูนซีเมนต์. *วารสารไทยการวิจัยดำเนินงาน, 4*(2), 32-41.

[6] นฤมล ซิ่งเสถียรตระกูล. (2545). การพยากรณ์โดยเทคนิคการเฉลี่ยน้ำหนักของข้อมูลแต่ละค่าเท่าๆกันและเทคนิคการทำให้เรียบแบบเอกโพแนนเชียล.*วิศวสารลาดกระบัง, 19*(1), 102-107.

[7] อัจฉรา จันทร์ฉาย. (2544). *การพยากรณ์เพื่อการตัดสินใจทางธุรกิจ*. กรุงเทพฯ. สำนักพิมพ์แห่งจุฬาลงกรณ์ มหาวิทยาลัย.

[8] ระพีพรรณ พิริยะกุล. (2550). การจำลองระบบด้วยคอมพิวเตอร์. กรุงเทพฯ : ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยรามคำแหง.

[9] ศศิธร โกฏสืบ และ กัลยา บุญหล้า. (2559). การสร้างตัวแบบเพื่อการพยากรณ์ราคาข้าวหอมมะลิ 105. คณะวิทยาศาสตร์ มหาวิทยาลัยนเรศวร.

การพัฒนาอัลกอริทึมการจัดกลุ่มลูกค้าและการค้นหาความสัมพันธ์ โดยใช้เทคนิค K-Means ร่วมกับ FP-Growth

The Development Customer Segmentation and Associations Algorithm Base on K-Means and FP-Growth

สุพิชฌาย์ แสงชาติ (Supitcha Saengchat)[1] และมณเฑียร รัตนศิริวงศ์วุฒิ (Montean Rattanasiriwongwut)[2]
ภาควิชาการจัดการสารสนเทศ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
[1]S5807021857202@email.kmutnb.ac.th, [2] Montean.r@it.kmutnb.ac.th

บทคัดย่อ

บทความนี้มีวัตถุประสงค์เพื่อพัฒนาอัลกอริทึมการจัดกลุ่มลูกค้าและค้นหาความสัมพันธ์ เพื่อเป็นข้อมูลสนับสนุนกลยุทธ์การขายให้กับองค์กร และตอบสนองความต้องการของลูกค้า สร้างการขายสินค้าต่อเนื่อง [1] และนำความสัมพันธ์ของลูกค้าแต่ละกลุ่มไปจัดโปรโมชั่นส่งเสริมการขายซึ่งผู้วิจัยได้นำข้อมูลการขายสินค้าประเภทเสื้อผ้าของธุรกิจค้าปลีกแห่งหนึ่ง ข้อมูลการขายเดือนกรกฎาคม 2559 ถึง เดือนมกราคม 2560 จำนวน 26,309 รายการ มาวิเคราะห์แนวโน้มที่เกิดขึ้นในอดีตด้วยการทำเหมืองข้อมูลตามกรอบ CRISP-DM [2] โดยใช้โปรแกรม Rapid Minerในการวิเคราะห์ข้อมูล และพัฒนาระบบด้วยภาษา JAVA ผู้วิจัยเลือกใช้อัลกอริทึม K-Means ในการจัดกลุ่มลูกค้า และเทคนิค Elbow Method [3] ในการคำนวณหาจำนวนกลุ่มที่เหมาะสมจากนั้นนำลูกค้าแต่ละกลุ่มมาหาความสัมพันธ์พฤติกรรมการซื้อสินค้าที่ซื้อคู่กันด้วยอัลกอริทึม FP-Growth โดยใช้ค่าสนับสนุนขั้นต่ำ (Minimum Support) 20% ค่าความเชื่อมั่น (Confidence)70% และค่าความสัมพันธ์การซื้อสินค้าร่วมกัน (Lift) เท่ากับ 1 เป็นเกณฑ์ในการวัดประสิทธิภาพของกฎความสัมพันธ์ที่ได้

การประเมินคุณภาพระบบที่พัฒนาขึ้น โดยใช้ผู้เชี่ยวชาญ 5 ท่านผลการประเมินมีค่าเฉลี่ย 4.10 และส่วนเบี่ยงเบนมาตรฐาน 0.49 ซึ่งสรุปได้ว่าระบบที่พัฒนาขึ้นมีคุณภาพอยู่ในเกณฑ์ดีและสามารถนำไปใช้เป็นเครื่องมือในการค้นหาความสัมพันธ์การซื้อสินค้า และการจัดโปรโมชั่นการไปใช้งานได้จริง

คำสำคัญ: เหมืองข้อมูล เคมีน กฎความสัมพันธ์ เอพพรี-กโรธ การจัดกลุ่ม

Abstract

This paper present the development algorithms for finding relationships and grouping customers as database for selling strategies of organization and customer's demand, creating cross selling and bringing the customer relationship to the sales promotion. The researcher uses the data of a retail clothing business which is the 26,309 sales data in July 2016 to January 2017. The data were analyzed to the past trends by Mining CRISP-DM data [2] with Rapid Miner as the analysis tool and developed java. The researcher uses the K-Mean algorithm to group each customer and the Elbow Method [3] technique to determine the appropriate cluster size. Then, each customer group searched for the buying behavior relationship with FP-Growth algorithms by the criteria to measure confidence of the relationship rule with Minimum Support 10%,Confidence 70%,and Lift as 1.

The evaluate performance system by 5 experts. The average result is 4.10 and standard deviation is 0.49 in conclusion, the developed algorithm good performance. It is a tool to find a buying relationship and the actual sale promotion.

Keyword: Data Mining, K-Mean, Association Rule, FP-Growth, Clustering

1. บทนำ

การดำเนินธุรกิจขายปลีกประเภทเสื้อผ้าในห้างสรรพสินค้าปัจจุบันเป็นที่แพร่หลาย ทุกเพศทุกวัยให้ความสนใจ อีกทั้งในธุรกิจค้าปลีกมุ่งเน้นการวางแผนกลยุทธ์ต่างๆ ในการสร้างยอดขาย ซึ่งปัจจุบันมีการแข่งขันทางธุรกิจการค้าสูงเพื่อเป็นผู้นำทางด้านการขายต่อเนื่องและรักษาฐานลูกค้า

เสื้อผ้าจัดอยู่ในกลุ่มปัจจัย 4 เพราะมีความจำเป็นต่อการดำรงชีวิต ปัจจุบันธุรกิจเสื้อผ้ามีการขยายเติบโต ทำให้มีความต้องการสินค้าของกลุ่มผู้บริโภคเพิ่มมากขึ้น ผู้ประ ะกอบการจึงต้องหากลยุทธ์เพื่อสร้างความพึงพอใจให้ลูกค้า จากการสำรวจธุรกิจเสื้อผ้าแห่งหนึ่งพบว่า ปริมาณการซื้อสินค้าประจำวันเพิ่มขึ้นและข้อมูลไม่มีการนำมาวิเคราะห์ความสัมพันธ์การซื้อสินค้า ทำให้ไม่สามารถเข้าใจความต้องการลูกค้าและปริมาณการซื้อสินค้าลดลง

ผู้วิจัยได้ศึกษาค้นคว้าและเปรียบเทียบข้อดีข้อเสียงานวิจัยและนำมาประยุกต์ จึงขอเสนอแนวคิดการทำเหมืองข้อมูลเพื่อให้สอดคล้องกับปัญหางานวิจัยที่ได้กล่าวมาข้างต้น โดยประยุกต์การจัดกลุ่มลูกค้าด้วยอัลกอริทึม K-Means และนำไปหาความสัมพันธ์ร่วมกับอัลกอริทึม FP-Growth เพื่อศึกษาพฤติกรรมการซื้อสินค้าของลูกค้าแต่ละกลุ่ม แล้วนำไปจัดโปรโมชั่นส่งเสริมการขายให้มีความเหมาะสม เพื่อตอบสนองความต้องการลูกค้าให้ได้มากที่สุด

2. ทฤษฎีที่เกี่ยวข้อง

2.1 ทฤษฎีการทำเหมืองข้อมูล

เป็นการค้นหาความสัมพันธ์และรูปแบบทั้งหมด ซึ่งมีอยู่จริงในฐานข้อมูลแต่ได้ถูกซ่อนไว้ภายในข้อมูลฐานข้อมูลขนาดใหญ่ เหมืองข้อมูลจะทำการสำรวจและวิเคราะห์อย่างอัตโนมัติให้อยู่ในรูปแบบที่เต็มไปด้วยความหมายและอยู่ในรูปของกฎ โดยความสัมพันธ์เหล่านี้แสดงให้เห็นถึงความรู้ต่าง ๆ ที่มีประโยชน์ในฐานข้อมูลเพื่อทำนายแนวโน้มและพฤติกรรม

2.2 ทฤษฎีการจัดกลุ่ม

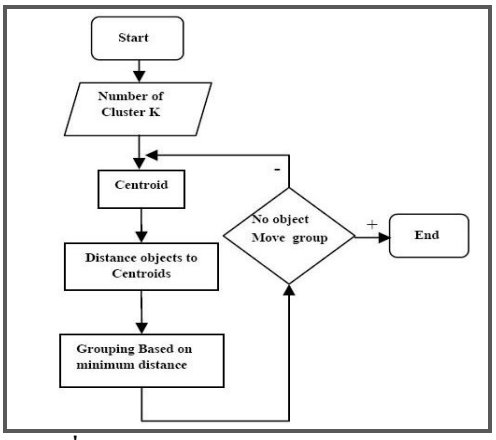
การวิเคราะห์กลุ่มจะแบ่งกลุ่มข้อมูลจากความคล้ายหรือไม่ทราบจำนวนกลุ่มล่วงหน้า เป็นการเรียนรู้แบบไม่มีผู้สอนเพื่อจำแนกกลุ่มข้อมูลที่มีคุณลักษณะคล้ายกันอยู่ในกลุ่มเดียวกัน โดยข้อมูลแต่ละกลุ่มจะถูกเรียกว่าคลัสเตอร์การวิเคราะห์หรือจำแนกกลุ่มข้อมูลสามารถแบ่งออกได้เป็น 2 ประเภท [4] ได้แก่ 1.การจัดกลุ่มแบบจำแนกข้อมูลออกเป็นกลุ่มย่อยๆ ตามจำนวนกลุ่ม (Partitional Clustering) 2.การจัดกลุ่มแบบโครงสร้างลำดับชั้น (Hierarchical Clustering) ซึ่งการจัดกลุ่มจะใช้ทฤษฎีการหาระยะห่างระหว่างของข้อมูล (Distance) โดยวิจัยนี้ได้ใช้การหาระยะห่างด้วยวิธี Euclidean Distance ซึ่งเป็นวิธีการหาระยะห่างเกิดจากรากที่ 2 ของผลต่างระหว่างแอตทริบิวต์ต่างๆ ยกกำลังสอง ดังสมการ

$$\boldsymbol{D}_{Euclidean} = \sqrt{(x1^2 - y1)^2 + (x2^2 - y2)^2 + .. + (xL^2 - yL)^2} \qquad (1)$$

โดยที่ x_1 คือ แอตทริบิวต์ที่ 1 ของข้อมูลจุดที่1 และ y_1 คือ แอตทริบิวต์ที่ 1 ของข้อมูลจุดที่2 โดยข้อมูลทั้ง 2 ตัว(x และ y) มีจำนวนแอตทริบิวต์เท่ากัน

2.3 อัลกอริทึมเคมีน (K-Means Algorithm)

มีการพัฒนาและนำเสนอโดย Mac Queen [8] ซึ่งแสดงให้เห็นถึงการจัดกลุ่มที่สมาชิกภายในกลุ่ม จะมีระยะใกล้จุดศูนย์กลางของกลุ่มโดยแบ่งกลุ่มให้ได้ K กลุ่มโดยค่า K กำหนดโดยผู้วิจัย ซึ่งหลักการทำงานของ K-Means แสดงดังภาพที่ 2



ภาพที่ 1: ขั้นตอนการทำงาน K-Means Algorithm

ขั้นตอนทำงาน K-Means Algorithm

1) ทำการสุ่มค่า K เป็นศูนย์กลางกลุ่มเริ่มต้น

2) คำนวณหาระยะห่างระหว่างข้อมูลแต่ละตัวกับจุดศูนย์กลางของแต่ละคลัสเตอร์

3) จัดกลุ่มข้อมูลใหม่โดยพิจารณาจากระยะห่างของข้อมูลในกลุ่มกับจุดศูนย์กลางว่าข้อมูลนั้นสมควรที่จะอยู่กลุ่มใด

4) ปรับปรุงการจัดกลุ่มโดยย้อนกลับไปทำข้อ 2-3 จนกว่าข้อมูลทุกตัวอยู่ในคลัสเตอร์เดิมหรือไม่มีการเปลี่ยนแปลงแล้ว

2.4 ทฤษฎีกฎความสัมพันธ์ (Associations Rule)

เป็นการหาความสัมพันธ์ของข้อมูล 2 ชุดขึ้นไปภายในข้อมูลที่มีขนาดใหญ่ การหากฎความสัมพันธ์โดยดูจากกฎความสัมพันธ์ของค่าความเชื่อมั่น ค่าสนับสนุน และค่า Lift ที่เกิดขึ้น [3]

2.4.1 นิยามกฎความสัมพันธ์

คำนิยามที่ใช้เรียกเพื่อใช้แทนสิ่งต่างๆในการค้นหากฎความสัมพันธ์ โดยกำหนดให้ I= {i 1, i2,…, in} เป็นเซตของชิ้นข้อมูลหรือรายการสินค้า (items) ที่ประกอบด้วยรายการสินค้าจำนวน n ชิ้น รายการสินค้า และให้ T= {t1, t2,…, tm} คือ เซตข้อมูลรายการซื้อขายสินค้า (Transaction) ในฐานข้อมูล T ที่ประกอบด้วยรายการข้อมูลมีสมาชิกจำนวน m รายการข้อมูล โดยที่แต่ละรายการ ti จะประกอบไปด้วยรายการย่อยของ รายการสินค้าจากกลุ่มข้อมูล I [7]

1) ค่าสนับสนุนขั้นต่ำ หรือเขียนแทนด้วย *minsup* [10] การหาค่าสนับสนุน (Support) จากสมการ (1) คำนวณจากข้อมูลที่ปรากฏร่วมกนของ Xและ Y ต่อสัดส่วนจำนวนรายการ ทั้งหมด (N) ดังสมการ (1)

$$s(X \rightarrow Y) = \sigma \frac{XUY}{N} \quad (1)$$

2) ค่าความเชื่อมั่นขั้นต่ำของกฎหรือเขียนแทนด้วย *minsup* [9] สามารถคำนวณค่า Confidence จากสมการ (2) รายการ Xและ Y ปรากฏร่วมกันต่อสัดส่วนรายการ

$$c(X \rightarrow Y) = \sigma \frac{XUY}{N} \quad (2)$$

จำนวนกฎที่เป็นไปได้จะคำนวณได้จากสมการที่ (3)

$$R = \sum_{k-1}^{d-1} [\binom{d}{k} x \sum_{k-1}^{d-1} \frac{d-k}{j}] \quad (3)$$

เมื่อค่า d คือจำนวนรายการ สมมุติว่ามี 4 รายการ จำนวนกฎที่เป็นไปได้จะเท่ากับ 50 กฎ แต่ถ้ามี 6 รายการจะได้ 602 กฎ : ซึ่งเป็นการเพิ่มขึ้นในลักษณะเอ็กซ์โพเน็นเชียลยิ่งมีรายการสินค้ามากจำนวนกฎจะมากตามไปด้วย

3) Lift คือค่าบ่งบอกการซื้อสินค้าR→ H มีความสัมพันธ์แค่ไหนโดยLift ค่าเกิน1 แสดงว่าสินค้ามีความสัมพันธ์กันมาก

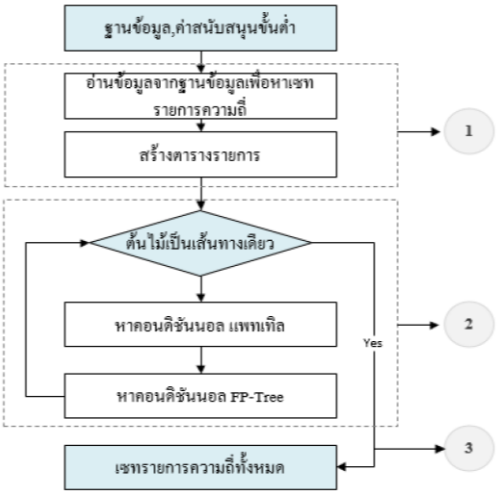
$$Lift(LHS \rightarrow RHS) = \frac{support\ (LHS, RHS)}{Support\ (LHS) x support(RHS)} \quad (4)$$

เมื่อ LHS (Left Hand Side) คือซื้อสินค้าที่อยู่ด้านซ้าย

RHS (Right Hand Side) คือซื้อสินค้าที่อยู่ด้านขวา

2.5 อัลกอริทึม เอฟพี-กโรธ (FP-Growth Algorithm)

Han [9] ได้พัฒนาอัลกอริทึมใหม่ขึ้น เพื่อลดจำนวนการอ่านข้อมูลจากฐานข้อมูล พร้อมเสนอโครงสร้างข้อมูลใหม่ขึ้นมา ชื่อว่า“FP-Tree” เรียกว่าอัลกอริทึม FP-Growth เป็นอัลกอริทึมที่อ่านข้อมูลจากฐานข้อมูลเพียง 2 ครั้ง ขั้นตอนการทำงานแสดงดังภาพที่ 2 [5]



ภาพที่ 2: ขั้นตอนการทำงาน FP-Growth Algorithm

1) อ่านข้อมูลจากฐานข้อมูลเพื่อหาเซตรายการความถี่ระดับที่ 1 ที่ค่าสนับสนุนมากกว่าหรือเท่ากับค่าสนับสนุนขั้นต่ำ

2) อ่านข้อมูลจากฐานข้อมูลอีกครั้งเพื่อตรวจสอบทุกรายการในการเปลี่ยนแปลง หากมีการเปลี่ยนแปลงไม่เป็นสมาชิกของรายการในตารางนั้นให้ลบรายการนั้นออก จากนั้นนำมาเรียงตามลำดับในตารางนำไปสร้างตาราง FP-Tree

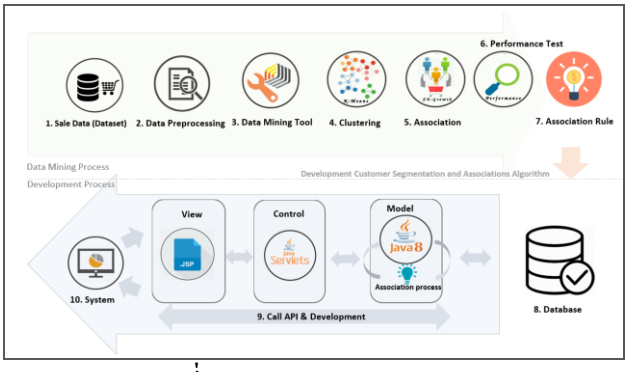
3) หาเซตรายการความถี่ทั้งหมด

พิจารณารายการสุดท้ายก่อน จากนั้นหาทำการหาคอนดิชันนอลแพทเทินเบซในตารางรายการจาก FP-Tree ผลลัพธ์ที่ได้จะเป็นเซตของทรี จากนั้นนำแต่ละคอนดิชันนอลฯ มาสร้าง FP-Tree

ทั้งนี้การพิจารณาเลือกอัลกอริทึม K-Means ในการจัดกลุ่มเพราะเป็นอัลกอริทึมที่เหมาะสมกับชุดข้อมูลที่ไม่ทราบจำนวนกลุ่มที่แน่นอนและเป็นวิธีค้นหาความคล้ายกันของชุดข้อมูลได้เร็วที่สุดเมื่อเทียบกับการจัดกลุ่มแบบลำดับชั้น และการหาความสัมพันธ์โดยใช้อัลกอริทึม FP-Growth เพื่อลดระยะเวลาในการอ่านข้อมูลทำให้ประมวลผลรวดเร็วกว่าอัลกอริทึม Apriori และเหมาะกับชุดข้อมูลที่นำมาใช้ทดสอบเพราะมีจำนวนมาก ข้อเสีย Apriori จะต้องทำการอ่านข้อมูลหลายครั้งและในระหว่างการประมวลผลข้อมูลและ ใช้หน่วยความจำเป็น

จำนวนมากสำหรับการสร้างกลุ่มแข่งขัน ด้วยเหตุผลที่กล่าวมา ผู้วิจัยจึงเลือกใช้ K-Means ร่วมกับ FP-Growth

3. วิธีการดำเนินวิจัย



ภาพที่ 3: กรอบแนวคิดการพัฒนาระบบ

จากภาพที่ 3 ผู้วิจัยได้ดำเนินวิจัยตามกรอบแนวคิดการพัฒนาระบบโดยแบ่งออก เป็น 2 ส่วน มีรายละเอียด ดังนี้

ส่วนที่ 1 ขั้นตอนการทำเหมืองข้อมูล (Data Mining process) โดยมีขั้นตอนการดำเนินการ ดังนี้

3.1 นำข้อมูลการขายจากฐานข้อมูล รวบรวมข้อมูลการขายจากระบบ POS ลักษณะการจัดเก็บเป็นฐานข้อมูลทุติยภูมิตั้งแต่เดือน ก.ค. 2559 - ม.ค. 2560 จำนวน ทั้งสิ้น 26,309 รายการ

3.2 การปรับปรุงข้อมูล(Data Preparation) โดยใช้หลักการของ CRISP-DM[3] เพื่อคัดเลือกและปรับปรุงข้อมูลก่อนให้สมบูรณ์ ซึ่งมีขั้นตอนและรายละเอียด ดังนี้

3.2.1 การคัดเลือก Attribute

คัดกรองเฉพาะคอลัมน์ข้อมูลที่จำเป็นต่อการวิเคราะห์และนำไปสร้างแบบจำลองการจัดกลุ่ม ดังตารางที่ 1

ตารางที่ 1: การคัดเลือก Attribute

ลำดับ	ฟิลด์ข้อมูล	ประเภทข้อมูล	ความหมาย
1	Reciept_No	Polynomial	เลขที่ใบเสร็จ
2	Buy_Day	Integer	วันซื้อสินค้า
3	Time_Buy	Integer	เวลาซื้อสินค้า
4	Item_Des	Integer	ชื่อสินค้า
5	Qty	Integer	จำนวนซื้อ
6	Net Price	Integer	ราคาสุทธิ

3.1.1 การปรับปรุงข้อมูล

จัดการข้อมูลซ้ำ ลดขนาดข้อมูลและทดแทนค่าว่างที่หายไป แสดงดังภาพที่ 4

Receipt_No	Buy_Day	Store	Sex	Item Group	Item_No	Item_Des
P00405843	1/12/2016	CLD	MEN	MO	S16MO01151YE1B	01 - SHIRT
P00405844	1/12/2016	CLD	GIRL	GP	A16GP03022WH1A	03 - T-SHIRT
P00405844	1/12/2016	CLD	GIRL	GP	A16GP03052BK1A	03 - T-SHIRT
P00405845	1/12/2016	CLD	GIRL	GO		
P00405845	1/12/2016	CLD	GIRL	GO	A16GO01062WH1A	01 - SHIRT
P00405845	1/12/2016	CLD	GIRL	GO	A16GO01062BK1A	01 - SHIRT
P00405845	1/12/2016	CLD	GIRL	GO		
P00405845	1/12/2016	CLD	GIRL	GO	A16GO42012GY1B	42 - SWEATE

ภาพที่ 4: การทดแทนค่าว่าง

3.3 นำข้อมูลที่ปรับปรุงแล้วเข้าสู่กระบวนการทำเหมืองข้อมูล (ในบทความนี้ผู้วิจัยใช้ Rapid Miner)

3.4 การจัดกลุ่ม (Clustering) ในการจัดกลุ่มลูกค้าผู้วิจัยเลือกใช้ K-Means ในการจัดกลุ่มและคำนวณหาระยะห่างระหว่างกลุ่มข้อมูลลูกค้าด้วยค่า Euclidean Distance ตามนิยามสมการ ดังนี้

$$\boldsymbol{Receipt}_{Euclidean}=\sqrt{(day1^2-day2)^2+(day2^2-day3)^2+..+(dL^2-dL)^2} \quad (5)$$

โดยหาค่า K ด้วยวิธี Elbow Method [3] ในการจัดกลุ่มเลือกข้อมูลการซื้อสินค้าตามวันเพื่อหาแนวโน้มการซื้อสินค้าลูกค้าแต่ละกลุ่ม

3.4.1 การสร้างแบบจำลองการจัดกลุ่ม

จากชุดข้อมูลทั้งหมด 26,309 นำมาทำ Pivot เพื่อตัดรายการซ้ำ และได้ใบเสร็จทั้งหมด 3,977 รายการจากนั้นนำข้อมูลแปลงเป็นค่า 0-1 โดยไม่สนใจยอดรวมการซื้อสินค้าแต่ละรายการของใบเสร็จ เพื่อนำไปค้นหาวันที่ลูกค้าซื้อสินค้า(จันทร์-อาทิตย์) กรณีมีการซื้อสินค้าจะถูกแทนค่าด้วย “1” และไม่ซื้อสินค้าจะถูกแทนด้วย “0” ผลการจัดกลุ่มลูกค้าตามวัน แสดงดังภาพที่ 5

Attribute	cluster_0	cluster_1	cluster_2	cluster_3
1.0	0.456	1.043	0.004	0.160
2.0	0.377	0.422	0.690	0.159
3.0	0.590	0.620	0.771	0.251
4.0	0.277	0.289	0.322	0.414
5.0	0.241	0.280	0.303	0.384
6.0	1.066	0.007	0.002	0.116
7.0	0.152	0.200	0.087	0.873

ภาพที่ 5: ผลลัพธ์การจัดกลุ่มข้อมูล

จากภาพที่ 6 การจัดกลุ่มทั้งหมด 4 กลุ่มได้แก่

Cluster_0 คือกลุ่มที่นิยมซื้อสินค้าวันศุกร์ 1,011 รายการ

Cluster_1 คือกลุ่มที่นิยมซื้อสินค้าวันอาทิตย์ 1,207 รายการ

Cluster_2 คือคือกลุ่มที่นิยมซื้อสินค้าวันอังคาร1,003 รายการ

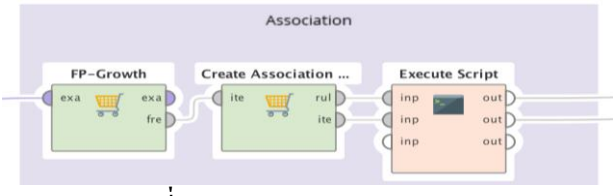
Cluster_3 คือคือกลุ่มที่นิยมซื้อสินค้าวันเสาร์ 756 รายการ

3.5 การหาความสัมพันธ์ (Association) ในการหาความสัมพันธ์ผู้วิจัยได้นำกลุ่มลูกค้าแต่ละกลุ่มมาหาความสัมพันธ์ทีละกลุ่ม เพื่อให้ทราบถึงพฤติกรรมการซื้อสินค้า โดยได้ใช้อัลกอริทึม FP-Growth ที่พัฒนาต่อจาก Appriori เพื่อแก้ไขการอ่านข้อมูลหลายครั้งและการสร้าง Candidate Key หลายครั้ง ซึ่งการคำนวณหาความสัมพันธ์จะใช้หลักการตามภาพที่ 2 เพื่อค้นหา Frequency Item Set หรือรายการสินค้าที่มีความถี่ในการซื้อมาก โดยวิธีการค้นหาเริ่มจาก

1. การหาค่าสนับสนุนการซื้อสินค้า(Item) ที่มีความยาวเท่ากับ 1 ของสินค้าแต่ละประเภท ดังสมการที่ 6

$$Support \quad \frac{\sigma(\text{ชื่อสินค้า})}{\text{จำนวนการซื้อสินค้าทั้งหมด}} \qquad (6)$$

2. เรียงลำดับค่า Support ที่คำนวณได้จากมากไปหาน้อย
3. สร้าง Compact Data Structure หรือ “FP Tree”
4. สร้างรูปแบบที่เกิดขึ้นบ่อยหรือมีการซื้อสินค้าบ่อยครั้งด้วยการกำหนดค่าสนับสนุนขั้นต่ำ(Minimum Support) 20% และค่าความเชื่อมั่น (Confidence) 70% แสดง Process และ Operator การค้นหากฎความสัมพันธ์แสดงดังภาพที่ 6



ภาพที่ 6: Process การค้นหากฎความสัมพันธ์

3.6 การวัดประสิทธิภาพ (Performance)

3.6.1 การจัดกลุ่มลูกค้า มีวิธีการ ดังนี้

ผู้วิจัยได้หาค่า K เริ่มต้นเป็น 4 กลุ่มโดยใช้วิธี Elbow point [2]ในการหาค่าเฉลี่ยกึ่งกลางของข้อมูลเพื่อให้ได้กลุ่มที่เหมาะสมกับชุดข้อมูลซึ่ง Elbow Pointจะทำการหาค่าเฉลี่ยของข้อมูลและ Plot graph ค่า K และค่าเฉลี่ยระยะห่างของข้อมูลเพื่อให้ได้ค่า K ที่เหมาะสม

3.6.2 การหากฎความสัมพันธ์ มีวิธีการ ดังนี้

1. กำหนดค่าสนับสนุนขั้นต่ำการซื้อสินค้าเท่ากับ 20% เพื่อคัดเลือกเฉพาะข้อมูลที่สนใจและความถี่การซื้อ 20%ขึ้นไป

$$Support\ (support \geq 0.2) \qquad [7]$$

2. กำหนดค่าความเชื่อมั่น (Confidence) หรือความน่าจะเป็นในการซื้อสินค้ากำหนดให้เท่ากับ 70% ตามสมมติฐาน

3. คัดเลือกค่าความสัมพันธ์การซื้อสินค้า (Lift) มีค่าเท่ากับหรือมากกว่า 1 ขึ้นไป เพื่อเป็นเงื่อนไขในการจัดโปรโมชั่น

การวัดประสิทธิภาพความสัมพันธ์ที่กล่าวข้างต้นไม่สามารถกำหนดค่าสนับสนุนและค่าความเชื่อมั่นสูง เนื่องจากประเภทสินค้ามีความกระจายตัวของข้อมูลน้อยกว่าสินค้าประเภทอุปโภคบริโภคทำให้โอกาสการซื้อน้อยกว่าทำให้ไม่สามารถกำหนดค่าต่าง ๆ สูงได้ ผู้วิจัยจึงกำหนดค่าที่ยอมรับได้

3.7 กฎความสัมพันธ์(Association Rule) ผลจากการการจัดกลุ่มและเข้าสู่กระบวนการค้นหาความสัมพันธ์ จะได้โมเดลความสัมพันธ์ทั้งหมด 4 กลุ่ม ตามจำนวนของลูกค้า เพื่อนำไปสู่ขั้นตอนการพัฒนาระบบ

<u>ส่วนที่ 2</u> ขั้นตอนการพัฒนาระบบ มีรายละเอียด ดังนี้

3.8 จัดเก็บเข้าสู่ฐานข้อมูล (Database) เป็นขั้นตอนการนำผลลัพธ์จากกระบวนการค้นหาความสัมพันธ์จัดเก็บในฐานข้อมูล (MySQL) เพื่อเรียกใช้งานและนำข้อมูลความสัมพันธ์ไปพัฒนาในการจัดโปรโมชันส่งเสริมการขาย

3.9 ขั้นตอนการติดต่อกับ API ของเครื่องมือในการทำเหมืองข้อมูล และการพัฒนาเว็บไซต์ เพื่อเป็นการประยุกต์อัลกอริทึมให้สามารถใช้งานและตอบสนองความต้องการของลูกค้าได้ โดยในขั้นตอนการพัฒนาระบบ มีรายละเอียดดังนี้

โดยเริ่มจากเชื่อมต่อ API ของ Rapid Miner โดย Process ของK-Means และ FP-Growth ที่สร้างขึ้นซึ่งโปรแกรมได้มี API เพื่อให้ผู้ใช้เชื่อมต่อไปยัง Process โดยใช้ภาษา Java และในการแสดงผลได้ออกแบบแสดงผลผ่านเว็บไซต์ จึงทำให้เลือกใช้ Java Server Pages (JSP) ของภาษา Java เพื่อให้สามารถใช้งาน API ของ Rapid Miner ได้ และเลือกออกแบบ Patterns ให้เป็น MVC Pattern ที่ประกอบด้วย 3 ส่วนได้แก่ 1.Model 2.View 3.Control และส่วนที่ใช้ในการเชื่อมต่อกับ API ของ Rapid Miner คือส่วน Model โดยในModel นั้นจะเรียกใช้ Class ของRapid Miner และใช้ไฟล์ Process ที่ได้สร้างจาก โปรแกรมRapid Miner Package (นามสกุล.rmp)

ในบทความฉบับนี้ได้สร้าง ไฟล์ Process ไว้ 2 ส่วนคือ 1. Process K-Means 2. Process FP-Growth เหตุผลที่ต้องแบ่งเป็น 2 Process เนื่องจากผู้วิจัยต้องการนำข้อมูลการจัดกลุ่มและกฎความสัมพันธ์ทั้ง 4 กลุ่ม จัดเก็บลงฐานข้อมูลเพื่อนำข้อมูล 2 Process เชื่อมต่อกันและแสดงผลหน้าเว็บไซต์ และในการเลือกกฎความสัมพันธ์ที่แสดงการจัดโปรโมชั่นบนเว็บไซต์จะเลือกกฎที่มีความสัมพันธ์ค่า Lift ตั้งแต่ 1 ขึ้นไป เพราะเป็นค่าที่ลูกค้ามีโอกาสซื้อคู่กันสูง ทำให้เป็นที่น่าสนใจในการนำไปวางแผนการจัดโปรโมชั่นส่งเสริมการขายในอนาคตต่อไป

3.10 ระบบที่ได้จากการพัฒนาขึ้นโดยสามารถใช้งานและตอบสนองกับความต้องการของลูกค้าและมีประโยชน์ต่อการจัดโปรโมชั่นส่งเสริมการขาย

4. ผลการดำเนินงาน

4.1 ผลการหาความสัมพันธ์

การวิเคราะห์การจัดกลุ่ม และนำไปค้นหาความสัมพันธ์ สรุปได้เป็น 4 กลุ่ม ดังนี้

กลุ่มที่ 1 กลุ่มลูกค้าที่นิยมซื้อสินค้าวันศุกร์และวันอังคาร กฎความสัมพันธ์การซื้อสินค้าได้ทั้งหมด 66 กฎ

กลุ่มที่ 2 กลุ่มลูกค้าที่นิยมซื้อสินค้าวันอาทิตย์ กฎความสัมพันธ์การซื้อสินค้าได้ทั้งหมด 53 กฎ

กลุ่มที่ 3 กลุ่มลูกค้าที่นิยมซื้อสินค้าวันอังคารและวันพุธ กฎความสัมพันธ์ การซื้อสินค้าได้ทั้งหมด 29 กฎ

กลุ่มที่ 4 กลุ่มลูกค้าที่นิยมซื้อสินค้าวันเสาร์ กฎความสัมพันธ์การซื้อสินค้าได้ทั้งหมด 30 กฎ

ตารางที่ 2: กฎความสัมพันธ์ลูกค้ากลุ่ม 4

xamples, 0 special attributes, 6 regular attributes) Filter (30 / 30 exampl

Conclusion	Premise	Total Supp...	Confidence	Gain	Lift
[BASIC JEANS]	[SHORTS, PANTS, BELT]	0.152	0.906	-0.184	1.515
[BASIC JEANS]	[DRESS/JUMPSUIT, SHORTS, PANTS]	0.179	0.865	-0.234	1.447
[BASIC JEANS]	[DRESS/JUMPSUIT, PANTS, BELT]	0.160	0.852	-0.216	1.425
[PANTS]	[BASIC JEANS, DRESS/JUMPSUIT, BELT]	0.160	0.846	-0.218	1.710
[BASIC JEANS]	[PANTS, SWEATER]	0.164	0.832	-0.230	1.392
[BASIC JEANS]	[SHORTS, HANDBAG, PANTS]	0.149	0.825	-0.213	1.380
[BASIC JEANS]	[SHORTS, PANTS]	0.258	0.812	-0.377	1.359
[BASIC JEANS]	[DRESS/JUMPSUIT, PANTS]	0.250	0.811	-0.366	1.357
[PANTS]	[BASIC JEANS, SHORTS, BELT]	0.152	0.810	-0.224	1.637
[BASIC JEANS]	[PANTS, BELT]	0.213	0.809	-0.313	1.353
[PANTS]	[BASIC JEANS, DRESS/JUMPSUIT, SHORTS]	0.179	0.794	-0.271	1.605
[BASIC JEANS]	[PANTS, T-SHIRT]	0.192	0.792	-0.292	1.325
[PANTS]	[BASIC JEANS, SHORTS, HANDBAG]	0.149	0.774	-0.237	1.565
[BASIC JEANS]	[HANDBAG, PANTS]	0.205	0.752	-0.340	1.258
[DRESS/JUMPSU...	[BASIC JEANS, PANTS, BELT]	0.160	0.752	-0.266	1.346

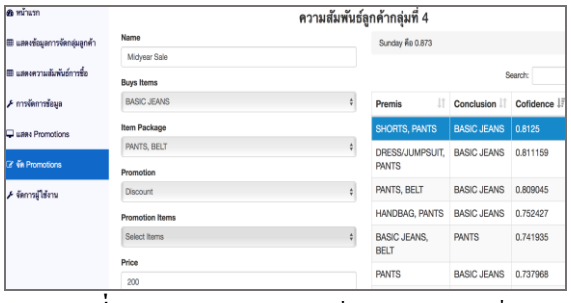
จากตารางที่ 2 ผู้วิจัยขอยกตัวอย่างผลการค้นหากฎความสัมพันธ์ลูกค้ากลุ่ม 4 ซึ่งได้แสดงในรูปแบบตารางผ่าน Operator Execute เพื่อง่ายต่อการวิเคราะห์ จากตารางลูกค้านิยมซื้อสินค้าในวันเสาร์และมักซื้อสินค้าประเภทเสื้อผ้าพร้อมกันบ่อยๆ ซึ่งกฎความสัมพันธ์ที่ได้ทั้งหมด 30 กฎ และซื้อสินค้าคู่กันมากสุด 4 ประเภท สรุปได้ว่าสินค้าประเภทกางเกง (Basic Jeans) เป็นที่นิยมซื้อร่วมกันบ่อย ๆ ด้วยค่าความเชื่อมั่น (Confidence) มากสุด 0.91 หรือ 91% และโอกาสการซื้อคู่กัน 1.51 ซึ่งเมื่อซื้อชุดประเภทยีนส์มีโอกาสซื้อสินค้าคู่กันได้แก่ Short, Pants, Belt, Dress/Jumpsuit และ Sweater แสดงดังตารางที่ 2

4.2 ผลการพัฒนาระบบ

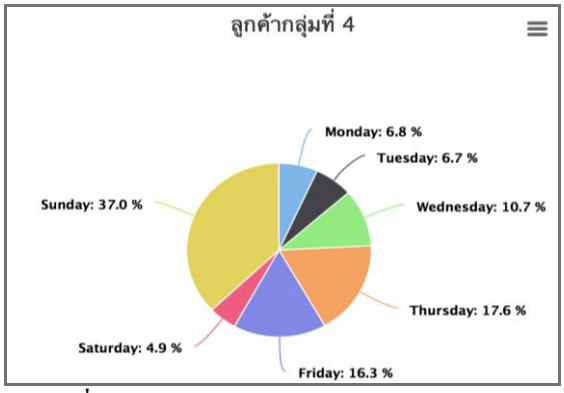
ผู้วิจัยได้นำเสนอการประยุกต์ใช้อัลกอริทึมการจัดกลุ่มและหาความสัมพันธ์ โดยรูปแบบของเว็บแอพพลิเคชั่น ซึ่งระบบนี้จะสนับสนุนการจัดการโปรโมชั่น โดยการจัดโปรโมชั่นตามกฎความสัมพันธ์การซื้อสินค้าแต่ละกลุ่ม แสดงดังภาพที่ 7-9



ภาพที่ 7: แสดงความสัมพันธ์การซื้อสินค้าลูกค้ากลุ่มที่ 4



ภาพที่ 8: หน้าแสดงการจัดโปรโมชั่นสินค้าลูกค้ากลุ่มที่ 4



ภาพที่ 9: กราฟแสดงการซื้อสินค้าลูกค้ากลุ่ม 4 แบ่งตามวัน

จากภาพที่ 7-9 เป็นความสัมพันธ์ลูกค้ากลุ่มที่4 ที่นิยมซื้อสินค้าวันอาทิตย์ดังภาพที่10 โดยเลือกค่าความเชื่อมั่น 80%

จากตัวอย่างนำ Basic Jeans ที่ซื้อคู่กับ Short, Pant มาจัดโปรโมชั่นตามช่วงวันที่ลูกค้านิยมซื้อและสินค้าที่ซื้อคู่กันมากสุด ซึ่งโปรโมชั่นได้แบ่งออกเป็น 3 ประเภทได้แก่ Discount (ส่วนลด),Free (ฟรี), In the price (ส่วนลดเมื่อซื้อชิ้นที่3) แสดงการจัดโปรโมชั่นดังภาพที่ 8

4.3 การประเมินระบบคุณภาพระบบ

ตารางที่ 3: การประเมินคุณภาพระบบ

รายการประเมิน	ระดับคุณภาพ		
	$\overline{X}$	S.D	เชิงคุณภาพ
1.สะดวกต่อการใช้งาน	4.00	0.43	ดี
2.การทำงานของระบบถูกต้อง	4.21	0.44	ดีมาก
3.ประสิทธิภาพในการประมวลผล	3.80	0.56	ดี
4.ความแม่นยำการวิเคราะห์ข้อมูล	4.40	0.52	ดีมาก
สรุปผลการประเมิน	4.10	0.4875	ดี

5. สรุป อภิปรายผล และข้อเสนอแนะ

5.1 สรุป อภิปรายผลการวิจัย

การวิจัยครั้งนี้ สรุปได้ว่าคุณภาพการทำงานของระบบที่พัฒนาขึ้น ค่าเฉลี่ย 4.10 ส่วนเบี่ยงเบนมาตรฐาน 0.49 อยู่ในเกณฑ์ดีโดยสามารถนำไปใช้งานได้เป็นอย่างดี ช่วยอำนวยความสะดวกให้กับผู้ใช้งานและผู้บริหาร เพื่อนำไปวางกลยุทธ์การตลาดเพื่อตอบสนองความต้องการลูกค้าในอนาคตได้

5.2 ข้อเสนอแนะ

ผู้วิจัยมีข้อเสนอแนะเพิ่มเติมสำหรับผู้สนใจที่จะนำระบบนี้ไปพัฒนาต่อให้มีขีดความสามารถเพิ่มขึ้น และจัดการข้อมูลได้หลากหลายมากขึ้น ซึ่งในวิจัยนี้ผู้วิจัย 2 ส่วนดังนี้

5.2.1 ด้านการทำเหมืองข้อมูลสามารถพัฒนาการพยากรณ์ยอดขายสินค้าลูกค้าแต่ละกลุ่มแต่ละเดือนหรือประยุกต์ข้อมูลกนี้ไปพัฒนาปรับใช้กับอัลกอริทึมอื่นๆได้ เนื่องจากวิธีนี้อาจไม่ใช่วิธีที่ดีสุด เพราะผู้วิจัยเลือกอัลกอริทึมที่เหมาะสมกับชุดข้อมูลที่นำมาทดสอบ และสามารถนำไปใช้ในการทำงานจริงได้

5.2.2 ด้านพัฒนาระบบเนื่องจากงานวิจัยนี้ได้กำหนดกลุ่มข้อมูลอยู่ 4 กลุ่มทำให้การ Programming ได้กำหนดการจัดเก็บข้อมูล ไว้แค่ 4 กลุ่ม เพื่อให้เกิดความยืดหยุ่น ของ Programming โดยให้หาจำนวน Output ของกลุ่มข้อมูลจากนั้นให้สร้าง Table ใน Database เพื่อเก็บข้อมูล รวมทั้งส่วนการ Update ไฟล์ Process ใน Server อีกทั้งสามารถ Export ผลลัพธ์ ออกมาเป็นไฟล์.ในรูปแบบ PDF, Excel, Word เป็นต้น

เอกสารอ้างอิง

[1] วชิราภรณ์ สร้อยสน. (2556) “การศึกษาพฤติกรรมผู้บริโภคเพื่อส่งเสริมการขายต่อเนื่อง โดยเทคนิคการทำเหมืองข้อมูล กรณีศึกษาร้านค้าปลีกแห่งหนึ่ง” มหาบัณฑิตมหาวิทยาลัยธรรมศาสตร์.

[2] เอกสิทธิ์ พัชรวงศ์ศักดา. (2558).”การวิเคราะห์ข้อมูลด้วยเทคนิคดาต้าไมนิ่งเบื้องต้น” พิมพ์ครั้งที่ 2. กรุงเทพฯ :บริษัทเอเชีย ดิจิตอล การพิมพ์.

[3] เอกสิทธิ์ พัชรวงศ์ศักดา. (2559).”Practical Data Mining with Rapid Miner Studio7” พิมพ์ครั้งที่ 1. กรุงเทพฯ :บริษัทเอเชีย ดิจิตอลการพิมพ์.

[4] วาทินี นุ้ยเพียร. (2552) “การเปรียบเทียบอัลกอริธึมการจัดกลุ่มข้อมูล Ozone day โดยใช้เทคนิคการทำเหมืองข้อมูล” สารนิพนธ์วิทยาศาสตรมหาบัณฑิต คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยพระจอมเกล้า พระนครเหนือ.

[5] ฉรงค์ศักดิ์ คงทิม. (2554) “การประยุกต์ใช้เอฟพี-กโรธ กับงานแนะแนวการศึกษาต่อในระดับอุดมศึกษา” การประชุมวิชาการทางคอมพิวเตอร์ ฯ (CIT 2011 & Unicom’s 2011), ศาลายา.

[6] ศิริพันธ์ เทพมาก. (2555).”การพัฒนาระบบหาความสัมพันธ์ของการซื้อสินค้าโดยใช้เทคนิคกฎความสัมพันธ์กรณีศึกษาสินค้าประเภทเครื่องสำอาง” ปัญหาพิเศษวิทยาศาสตรมหาบัณฑิตคณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยพระจอมเกล้าพระนครเหนือ.

[7] ประมูล สุขสกาวผ่อง.(2558). ”การทำเหมืองข้อมูลกฎความสัมพันธ์ด้วยวิธีค่าถ่วงน้ำหนักสูงสุด จากกราฟบริบูรณ์แบบสมมาตร”ปัญหาพิเศษวิทยาศาสตรมหาบัณฑิตคณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยพระจอมเกล้าพระนครเหนือ.

[8] Macqueen, (1967) “Some Methods for Classification and Analysis of Multivariate Observation”

[9] Han, J. Pei, J., and Yin, Y. 2000. “Mining Frequent Patterns without Candidate Generation” In proceeding of the 2000 ACM SIGMOD.

[10] R. Agrawal, R. Srikant, “Fast Algorithms for Mining Association Rules in Large Databases,” Proceedings of the 20th International Conference on Very Large Data Bases, 672836: Morgan Kaufmann Publishers Inc.;. pp. 487-99, 1994

ระบบสร้างแผนการท่องเที่ยวจากข้อมูลออนไลน์
Package Tour Generation From Online Data

ณิชมน อุไรศรีพงศ์ (Nichamon Uraisriphong)[1] , วริสรา ธรรมวิรัช (Warisara Thumwirat)[2] ,
บัณฑิต ฐานะโสภณ (Bundit Thanasopon)[3] , พรฤดี เนติโสภากุล (Ponrudee Netisopakul) [4]
คณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง[1,2,3,4]
56070044@kmitl.ac.th[1], 56070117@kmitl.ac.th[2], bundit@it.kmitl.ac.th[3] , ponrudee@it.kmitl.ac.th[4]

บทคัดย่อ

โครงการนี้เกี่ยวข้องกับการพัฒนาระบบสร้างแผนการท่องเที่ยวจากข้อมูลออนไลน์ ทำการสกัดข้อมูลสถานที่ท่องเที่ยวที่เป็นที่นิยมและนำมาใช้ในการวางแผนการท่องเที่ยวด้วยตนเอง ในเบื้องต้นระบบจะมุ่งสกัดข้อมูลท่องเที่ยวในประเทศญี่ปุ่นจากเว็บไซต์พันทิปดอทคอมโดยผู้ใช้งานสามารถกำหนดเงื่อนไขตามต้องการได้ ตัวอย่างเช่น สามารถกำหนดเมือง ฤดูกาล จำนวนวัน และค่าใช้จ่ายเป็นต้น ซึ่งระบบจะทำการสร้างแผนการท่องเที่ยวให้เหมาะสมกับความต้องการของผู้ใช้งาน เทคโนโลยีที่ใช้ในการสกัดข้อความได้แก่ Web Scraping , การสกัดชื่อสถานที่ท่องเที่ยว เทคโนโลยีที่ใช้ในการวิเคราะห์ข้อมูลได้แก่ การนับความถี่ในการพูดถึงสถานที่ท่องเที่ยวและการวิเคราะห์ความคิดเห็นเชิงบวกลบ และเทคโนโลยีที่ใช้ในการพัฒนาระบบได้แก่ Django Framework, ภาษา Python, Mysql, Google Map API , Octoparse ผลการทำงานพบว่าระบบสามารถจัดแผนการท่องเที่ยวรายวันตามความชอบของผู้ใช้งานและให้ข้อมูลที่สำคัญต่างๆ เช่น วันเวลาเปิดปิดสถานที่ท่องเที่ยว อัตราค่าเข้าชม โรงแรมที่พักตามงบประมาณและวิธีการเดินทางโดยรถไฟเบื้องต้น

***คำสำคัญ**: การสกัดความรู้, การสร้างแผนการท่องเที่ยว, การวิเคราะห์อารมณ์, สังคมออนไลน์*

Abstract

This project involves the development of package tour generation system extracts popular tourist attractions and other travel information from online data and uses it to generate a travel plan for a specific traveler. The first version of this work focused on self-traveling in Japan and extracted popular tourist places from the website Pantip.com. Users can set the desired conditions, those are, destination city, season of travelling, number of days, budget, and so on. The system then generates a detailed travel plan for each day, including a recommended hotel, places to go, estimated travel time from place to place and so on. Technologies used in this work are text extraction including web scraping and extraction; and data and text analysis techniques including frequency counting and sentiment analysis. Django Framework, Python, Mysql, Google Map API, and Octoparse are tools used to develop this system. The results showed that the system can arrange daily travel plans according to users' preferences and provide important information such as date, time, location, hotel or hostel within the budget together with travel direction to popular attractions.

Keyword: Knowledge Extraction, Travel Planning Generation, Sentiment Analysis, Social Network.

1. บทนำ

การไปเที่ยวต่างประเทศด้วยตนเองเป็นที่นิยมมากขึ้น เนื่องจาก ข้อมูลส่วนใหญ่สามารถหาได้จากการแบ่งปันหรือแลกเปลี่ยนข้อมูลบนสื่อโซเชียล อย่างไรก็ตาม การหาข้อมูลดังกล่าวมักเสียเวลาค้นหาจากหลาย ๆ แหล่งหรือหลาย ๆ กระทู้ และอาจไม่พบข้อมูลที่สำคัญหรือที่ต้องการ ด้วยเหตุนี้ระบบสร้างแผนการท่องเที่ยวออนไลน์ จึงถูกพัฒนาขึ้นเพื่อช่วยผู้ที่ต้องการเดินทางท่องเที่ยวต่างประเทศด้วยตัวเอง จากผลการสำรวจจำนวนนักท่องเที่ยวไทยเดินทางไปประเทศญี่ปุ่นโดย

JNTO พบว่ามีจำนวนนักท่องเที่ยวเพิ่มขึ้นเป็นจำนวน 6 เท่า ภายในปี 2009 – 2015 [1] ดังนั้น ในเบื้องต้นระบบจะมุ่งสร้างแผนการท่องเที่ยวในประเทศญี่ปุ่น ระบบจะช่วยสร้างให้มีแผนการท่องเที่ยวที่สำเร็จรูปพร้อมรายละเอียดการเดินทางที่ชัดเจน โดยสกัดข้อมูลที่เกี่ยวข้องจากเว็บไซต์เช่น พันทิปดอทคอม แล้วนำข้อมูลเหล่านั้นมาประมวลผลภาษามนุษย์และวิเคราะห์ทางสถิติเพื่อมาสร้างแผนการท่องเที่ยวให้เหมาะสมกับเงื่อนไขที่ผู้เดินทางต้องการ

2. วรรณกรรมที่เกี่ยวข้อง

2.1 Web Scraping

Web Scraping [2] คือ วิธีการหรือเทคนิคการดึงข้อมูลที่เราต้องการจากหน้าเว็บไซต์ Web Scraping จะดึงข้อมูลจากหน้า HTML โดยการเรียนรู้ Markup และรูปร่างภายนอกในการแสดงข้อมูล แล้วจะทำการดึงข้อมูลเนื้อหาในส่วนที่เราต้องการ โดยข้อมูลดังกล่าวสามารถจัดเก็บได้หลายรูปแบบ ตัวอย่างเช่น จัดเก็บเป็นไฟล์ XML XLSX SQL และ TXT เป็นต้น ในงานนี้เราได้นำหลักการของ Web Scraping มาใช้ในการรวบรวมข้อมูลสถานที่ท่องเที่ยวต่าง ๆ จากเว็บไซต์พันทิปดอทคอม ซึ่ง Web Scraping สามารถสร้างขึ้นเองหรือจะใช้ Web Scarping Software ในการทำ Web Scraping ก็ได้เช่นเดียวกัน

2.2 Sentiment Analysis

Sentiment Analysis [3] คือกระบวนการวิเคราะห์อารมณ์ ได้แก่การวิเคราะห์ความคิดเห็น อารมณ์ ความรู้สึก ทัศนคติ ว่าเป็นความคิดเห็นในเชิงบวกหรือเชิงลบ การวิเคราะห์นี้จะทำให้ทราบว่าผู้แสดงความคิดเห็นมีความรู้สึกที่ดี หรือ ความรู้สึกที่ไม่ดี

จากงานวิจัยที่ [4] ตัวอย่างประโยค “The room was fantastic and the service was wonderful. However, we find the restaurant was high-priced” จะพบว่ามีคำในประโยคที่แสดงถึงค่าทาง Sentiment ได้แก่ fantastic, wonderful และ high-priced ซึ่งคำว่า “fantastic” และ “wonderful” นั้นมีค่า Sentiment เป็นบวก (Positive) แต่สำหรับคำว่า “high-priced” นั้นมีค่า Sentiment เป็นลบ (Negative) ซึ่งในงานวิจัยดังกล่าว ได้คำนวณค่าอารมณ์ (Sentiment Score) อย่างง่ายของประโยคข้างต้นจากสูตร

$$Sentiment\ Score = \frac{Positive\ word\ count - Negative\ word\ count}{Positive\ word\ count + Negative\ word\ count} \quad (1)$$

วิธีการคำนวณดังกล่าวนั้นเป็นจุดเริ่มของการออกแบบการวิเคราะห์อารมณ์ ซึ่งสิ่งแรกที่ต้องทำนั้นก็คือการตรวจจับคำที่มีค่า Sentiment เป็นบวก และการตรวจจับคำที่มีค่า Sentiment เป็นลบ หลังจากนั้นทำการใส่ค่าลงในสมการจึงจะได้ค่า Sentiment Score ของทั้งประโยค

2.3 RoutePerfect.com

RoutePerfect [5] เป็นเว็บแอพลิเคชันที่สามารถสร้างแผนการท่องเที่ยวให้กับผู้ใช้งานได้อย่างอัตโนมัติ โดยแผนการท่องเที่ยวนี้ จะเป็นแผนการท่องเที่ยวสำหรับทวีปยุโรปเท่านั้น ผู้ใช้งานสามารถเลือกรูปแบบการท่องเที่ยวและสามารถดูแผนการท่องเที่ยวของตนเองได้บนหน้าเว็บไซต์เท่านั้น แต่ผู้ใช้งานจะไม่สามารถบันทึกและแบ่งปันแผนการท่องเที่ยวได้

เมื่อผู้ใช้งานได้เลือกรูปแบบการท่องเที่ยวเช่น เลือกจุดเริ่มต้น เลือกจุดสิ้นสุด เรียบร้อยแล้ว ระบบจะทำการสร้างแผนการท่องเที่ยว ตามรูปแบบที่ผู้ใช้งานได้เลือกไว้โดยแผนการท่องเที่ยวนั้น จะแสดงเส้นทางการเดินทางบนแผนที่ลำดับสถานที่ต่าง ๆ ที่จะไป และผู้ใช้งานสามารถกำหนดว่าผู้ใช้งานชอบเที่ยวแบบใด เช่น ท่องเที่ยวเชิงธรรมชาติ ท่องเที่ยวเชิงวัฒนธรรม เป็นต้น

2.4 งานวิจัยที่เกี่ยวข้องกับการวางแผนการท่องเที่ยว

งานวิจัยที่เกี่ยวข้องกับการวางแผนการท่องเที่ยวได้แก่ ระบบวางแผนการท่องเที่ยวอัจฉริยะ [6] และระบบวางแผนแพ็กเกจทัวร์โดยใช้ Constraint-based Reasoning [7]

ระบบวางแผนการท่องเที่ยวอัจฉริยะ[6] เป็นการนำเทคโนโลยีซีเมนติกเว็บมาช่วยในการจัดแผนการท่องเที่ยวที่สามารถจัดลำดับการเดินทางได้ด้วยตนเอง ประกอบกับมีการแนะนำสถานที่ท่องเที่ยว ที่เหมาะสมกับรูปแบบของนักท่องเที่ยวแต่ละราย และสร้างแผนการท่องเที่ยวบน Google Map มีการคำนวณระยะทางการเดินทาง ไปยังสถานที่ท่องเที่ยวต่างๆ เพื่อให้ได้การเดินทางไปยังสถานที่ท่องเที่ยวนั้น ๆ ได้ดีที่สุด เทคโนโลยีที่เกี่ยวข้อง ได้แก่

1) Semantic Web ถูกนำมาใช้ในส่วนการแนะนำการท่องเที่ยวให้เหมาะสมกับนักท่องเที่ยวแต่ละคน โดยอาศัยกลไกการอนุมานข้อมูลในองค์ความรู้ ซึ่งงานวิจัยดังกล่าว จะใช้ออนโทโลยีเป็นเครื่องมือในการจัดรูปแบบการท่องเที่ยว และ

แนะนำแหล่งเที่ยวเที่ยว และใช้ Jena API ในการเชื่อมต่อกับออนโทโลยีและใช้ในการอนุมานความรู้

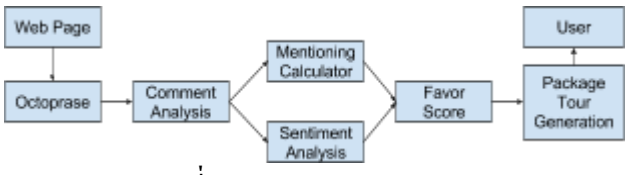
2) Ant Colony Optimization เป็นอัลกอริทึมที่ใช้ในการหาเส้นทางที่สั้นที่สุดสำหรับการไปยังสถานที่ทุกสถานที่ที่กำหนด (Sub Optimal Path)

อย่างไรก็ตาม งานวิจัยดังกล่าว ยังไม่มีการสกัดข้อมูลสถานที่ท่องเที่ยวจากสื่อโซเชียล และจะแนะนำสถานที่ท่องเที่ยวเป็นจุดๆ ก่อน เพื่อให้ผู้ใช้เลือก แล้วจึงสร้างเส้นทางการเดินทางที่สั้นที่สุดให้

อีกงานวิจัยหนึ่งคือ ระบบวางแผนแพ็กเกจทัวร์โดยใช้ Constraint-based Reasoning [7] ซึ่งเป็นการกำหนดปัญหาหรือเงื่อนไขบนตัวแปร เงื่อนไขเหล่านี้จะนำมาใช้ในงานทางด้านปัญญาประดิษฐ์ ซึ่งใช้ Constraint Satisfaction สำหรับการค้นหา Solution ในการแก้ปัญหาด้วยเซตของเงื่อนไข โดยเทคนิคที่จะนำมาใช้ขึ้นอยู่กับชนิดของเงื่อนไขที่ถูกพิจารณา และอัลกอริทึมที่นำมาใช้ในการแก้ปัญหามีด้วยกันหลายวิธี ในงานดังกล่าว เลือกวิธีการแบบ Branch and Bound ซึ่งเป็นอัลกอริทึมพื้นฐานเพื่อใช้ในการหาคำตอบที่ดีที่สุดในการแก้ไขปัญหาแบบไม่ต่อเนื่อง โดยมีเงื่อนไขต่างๆ ในการพิจารณา ได้แก่ ดาวของที่พัก ร้านอาหาร ในบริเวณใกล้เคียง งบประมาณ และรายจ่ายต่างๆ ในการเดินทาง เช่น ค่าน้ำมันรถ ค่าที่พัก ค่าเข้าชม เป็นต้น

งานทั้งสองข้างต้น มุ่งเน้นการท่องเที่ยวในประเทศไทยเท่านั้น ขณะที่ งานวิจัยที่นำเสนอใหม่นี้ มุ่งเน้นการท่องเที่ยวในประเทศญี่ปุ่น

3. วิธีการดำเนินการวิจัย



ภาพที่ 1 :ภาพรวมการทำงานของระบบ

3.1 การรวบรวมและสกัดข้อมูล

ข้อมูลเกี่ยวกับประเทศญี่ปุ่น อาทิ ภูมิภาค จังหวัด ฤดูกาล และข้อมูลสถานที่ท่องเที่ยว จะถูกรวบรวมไว้ก่อนเบื้องต้น จากนั้น ทำการดึงข้อมูลจากหน้าเว็บไซต์พันทิปดอทคอม โดยอันดับแรก เข้าเว็บไซต์ที่ห้องบลูแพลนเน็ต Tag ประเทศญี่ปุ่น จากนั้นคลิกไปเปลี่ยนจากกระทู้ทั้งหมดเป็นกระทู้รีวิวแทน เพื่อจำกัดให้ข้อมูลมีเพียงการรีวิวการท่องเที่ยวประเทศญี่ปุ่นเท่านั้น จากนั้นนำลิงค์แต่ละกระทู้เข้าโปรแกรม Octoparse [8] ซึ่งเป็นโปรแกรม Web Scrapping สำเร็จรูป เพื่อทำการสกัดเฉพาะเนื้อหาที่สนใจในกระทู้ออกมาแล้วจัดเก็บเป็นไฟล์ .txt

3.2 การคำนวณและวิเคราะห์ข้อมูล

3.2.1 การวิเคราะห์ประมวลผลข้อความคิดเห็น (Comment Analysis) ทำการรวบรวมความคิดเห็นของสถานที่ท่องเที่ยวต่าง ๆ ที่ได้จากแต่ละกระทู้ โดยนำประโยคที่แสดงความคิดเห็นเกี่ยวกับสถานที่ท่องเที่ยวนั้น ๆ มาสถานที่ละ 10 ประโยค เพื่อนำมาวิเคราะห์ค่าความคิดเห็นต่อไป

3.2.2 การคำนวณหาจำนวนครั้งที่สถานที่ท่องเที่ยวถูกพูดถึงในกระทู้พันทิป (Mentioning)

สำหรับการคำนวณ Mentioning ของสถานที่ท่องเที่ยวจะใช้ข้อมูลสถานที่ท่องเที่ยวที่ได้จากการรวบรวมในเบื้องต้น แล้วนำมาจัดทำเป็นพจนานุกรมของสถานที่ท่องเที่ยวซึ่งจะประกอบไปด้วยสถานที่ท่องเที่ยวในประเทศญี่ปุ่นซึ่งในเบื้องต้นครอบคลุม 3 จังหวัดคือ โตเกียว โอซาก้า และฮกไกโด และนำไปค้นหาในไฟล์กระทู้พันทิปที่ได้เก็บรวบรวมมาแล้ว

3.2.3 การวิเคราะห์ค่าความคิดเห็นของสถานที่ท่องเที่ยว

การวิเคราะห์ค่าความคิดเห็น ได้นำหลักการการวิเคราะห์อารมณ์มาใช้ โดยทำการรวบรวมความคิดเห็นของสถานที่ท่องเที่ยว แล้วทำการรวบรวมคำที่ปรากฏในประโยคนั้นๆ จากนั้น ทำการแยกประเภทของคำออกเป็น 1) Type of Words ได้แก่Adjective Adverb Noun Verb 2) Intensifier เช่นคำว่า ค่อนข้าง, ไม่ค่อย, เกือบ, แค่, ที่สุด, กว่า, จริง, อีกแล้ว, เกิน, นิดนึง, สุด, โคตร และ 3) Reverse เช่น ได้แก่ ไม่, อย่า, ห้าม, ไม่มี และนำมาให้คะแนนที่บ่งบอกถึงความรู้สึกเชิงบวกและเชิงลบ จากสูตร Phrase Sentiment Score (Phrase-SS) จากสูตรที่ (2) ในกรณีข้อความมีคำประเภท Intensifier และคำ Reverse กับ main word เชิงลบ และสูตรที่ (3) ในกรณีข้อความมีคำประเภท Reverse กับคำ Reverse เชิงบวก ตามลำดับ

$$Phrase - SS = word * (100 + Intensifie - S) / 100 \quad (2)$$

$$Phrase - SS = -1 * main_word \quad (3)$$

ตัวอย่างข้อความที่มี Intensifier เช่น “ค่อนข้างสวย” ซึ่ง “ค่อนข้าง” เป็นคำประเภท Intensifier มีคะแนนอยู่ที่ -30%

และ “สวย” เป็นคำประเภท Adjective มีคะแนนอยู่ที่ 4 คะแนน เมื่อหาคะแนนของคำนี้จะได้เป็น

$$Phrase\text{-}SS = 4 * (100 - 30)/100 = 2.8$$

ซึ่งคะแนนของคำ “ค่อนข้างสวย” จะมีคะแนน 2.8 ซึ่งน้อยกว่าคำว่า “สวย”

ข้อความ “ไม่ลำบาก” ซึ่ง “ไม่” เป็นคำประเภท Reverse มีคะแนนอยู่ที่ -70% และ “ลำบาก” มีคะแนนอยู่ที่ -4.5 คะแนน เมื่อหาคะแนนของคำนี้จะได้เป็น

$$Phrase - SS = -4.5 * (100 - 70)/100 = -1.35$$

หากคำเชิงบวกใดที่มีคำประเภท Reverse นำหน้าอยู่ด้วย กรณีนี้จะทำการคูณค่า -1 ให้กับคำนั้น เพื่อทำให้ความหมายเป็นตรงกันข้าม เช่นคำว่า “สวย” เป็นคำประเภท Adjective มีคะแนนอยู่ที่ 4 ซึ่งเป็นคำที่มีความหมายเชิงบวก เมื่อเติมคำว่า “ไม่” ที่เป็นคำประเภท Reverse เข้าไปด้านหน้า จะกลายเป็น “ไม่สวย” ซึ่งจะกลายเป็นคำที่มีความหมายเชิงลบทันที โดยใช้สูตร ดังนั้น “ไม่สวย” จะมีค่าคะแนนของคำคือ -4 คะแนน

$$Phrase - SS = -1 * 4 = -4$$

จากนั้น ทำการรวมคะแนนคำที่เป็นบวกและลบในแต่ละประโยคเพื่อมาหาค่า Sentence Sentiment Score (Sentence-SS) ดังสูตรต่อไปนี้

$$Sentence - SS = \frac{Positive\ word\ score - Negative\ word\ score}{Positive\ word\ score + Negative\ word\ score} \quad (4)$$

ตัวอย่างประโยคเช่น “ส่วนตัวค่อนข้างชอบบรรยากาศทางเข้าศาลเจ้าแต่ศาลเจ้าอยู่ไม่ใกล้เลยเดินกันขาลาก” จะเห็นได้ว่ามีจำนวนคำที่เป็นเชิงบวกอยู่ 2 คำ ได้แก่คำว่า “ค่อนข้างชอบ” , “บรรยากาศ” และมีจำนวนคำที่เป็นเชิงลบอยู่ 3 คำ ได้แก่คำว่า “ไม่ใกล้” , “เดิน” , “ขาลาก”

ซึ่งคำว่า “ค่อนข้างชอบ” มีคำว่า “ค่อนข้าง” เป็นคำประเภท Intensifier มีคะแนนอยู่ที่ -30% และ “ชอบ” มีคะแนนอยู่ที่ 4 คะแนน เมื่อหาคะแนนของคำนี้จะได้เป็น 2.8 ดังสูตรที่ด้านล่าง

$$Phrase - SS = 4 * (100 - 30)/100 = 2.8$$

และคำว่า “ไม่ใกล้” มีคำว่า “ไม่” เป็นคำประเภท Reverse และ “ใกล้” มีคะแนนอยู่ที่ 4.5 คะแนนซึ่งเป็นคำเชิงบวก ดังนั้นจึงใช้สูตรที่ (3) คำนวณ

$$Phrase - SS = -1 * 4.5 = -4.5$$

ซึ่งเมื่อนำคะแนนแต่ละคำในประโยคมาหาค่า Sentence Sentiment Score จะได้ค่าดังนี้

$$Sentence - SS = \frac{(2.8 + 3) - (4.5 + 2 + 4)}{(2.8 + 3) + (4.5 + 2 + 4)} = \frac{-4.7}{16.3} = -0.28$$

ดังนั้นประโยคนี้จะมีค่า sentiment score คือ -0.28 หรือประโยคนี้มีค่าความคิดเห็นเชิงลบ 28%

3.2.4 การวิเคราะห์ค่าความนิยมของสถานที่ท่องเที่ยว

จากตารางที่ 1 คือกฎการให้ค่าความนิยมของสถานที่ท่องเที่ยว ซึ่งสร้างมาจากผลลัพธ์จากการวิเคราะห์ค่าความคิดเห็น (Sentiment) ซึ่งจะบอกถึงค่าความคิดเห็นต่อสถานที่ท่องเที่ยวนั้น ๆ โดยจะแบ่งเป็นค่าความคิดเห็นเชิงบวก,ค่าความคิดเห็นเชิงลบ,ค่าความคิดเห็นเชิงกลาง และค่าการกล่าวถึงของสถานที่ท่องเที่ยว (Mentioning) จะบอกถึงความนิยมไปสถานที่ท่องเที่ยวนั้นๆ

ตารางที่ 1 : กฎค่าความนิยมของสถานที่ท่องเที่ยว

กฎค่าความนิยมของสถานที่ท่องเที่ยว		Mentioning		
		สูง	ปานกลาง	ต่ำ
Positive	สูง	1	2	3
	ปานกลาง	4	5	6
	ต่ำ	7	8	9
Neutral		10	11	12
Negative	ต่ำ	13	14	15
	ปานกลาง	16	17	18
	สูง	19	20	21

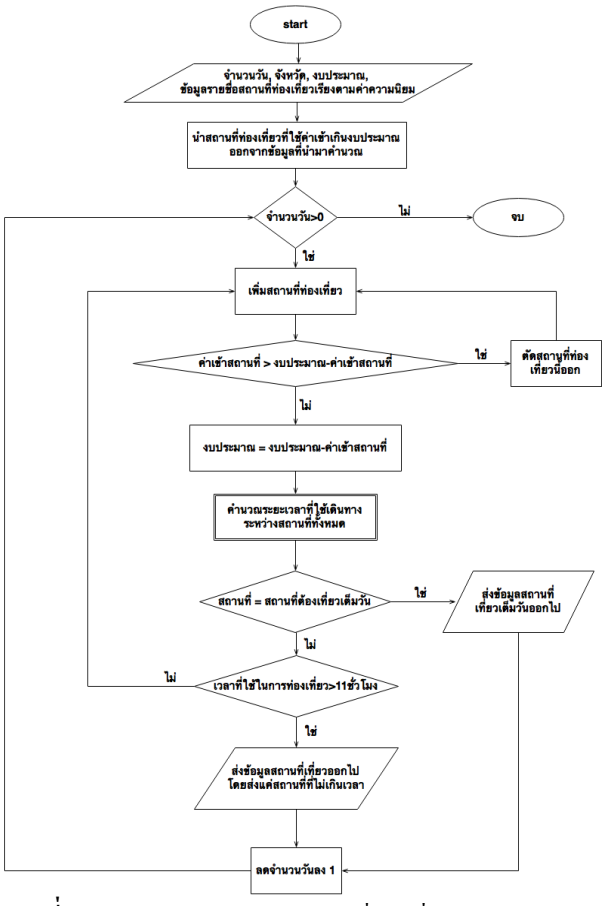
ตัวอย่างเช่น ถ้าสถานที่ท่องเที่ยวนี้มีค่าความคิดเห็นเชิงบวก (Positive) อยู่ในช่วงค่าสูงและมีค่าการกล่าวถึงสถานที่ท่องเที่ยว (Mentioning) อยู่ในช่วงค่าสูง จะมีค่าความนิยมเป็นอันดับที่ 1 ซึ่งหมายถึงสถานที่ท่องเที่ยวนี้จะได้รับการเลือกให้อยู่ในแผนการท่องเที่ยวอันดับแรก และหากสถานที่ท่องเที่ยวนี้มีค่าความคิดเห็นเชิงลบ (Negative) อยู่ในช่วงค่าต่ำและมีค่าการกล่าวถึงสถานที่ท่องเที่ยว (Mentioning) อยู่ในช่วงค่าต่ำ จะมีค่าความนิยมเป็นอันดับที่ 21 ซึ่งหมายถึงสถานที่ท่องเที่ยวนี้จะได้รับการเลือกให้อยู่ในแผนการท่องเที่ยวอันดับสุดท้าย

3.2.5 การสร้างแผนการเดินทางท่องเที่ยว

จะใช้ Google Map API ในการคำนวณหาลำดับการเดินทางที่ดีที่สุดและการค้นหาระยะเวลาที่ใช้ในการเดินทางระหว่างสถานที่ท่องเที่ยวเพื่อนำไปคำนวณหาจำนวนสถานที่ท่องเที่ยวในแต่ละวัน

การวิเคราะห์หาลำดับของสถานที่ท่องเที่ยว จะคำนวณได้จากเส้นทางที่ Google Map API คำนวณออกมาให้ซึ่งเป็นเส้นทางที่สามารถเดินทางได้รวดเร็วที่สุดและดีที่สุด โดยการกำหนดพิกัดของจุดเริ่มต้น พิกัดของจุดปลายทางและพิกัดของ

สถานที่ท่องเที่ยวภายในวันนั้น โดยระบบจะทำการดึงสถานที่ท่องเที่ยวออกมาตามค่าความนิยม



ภาพที่ 2 : อัลกอริทึมการหาจำนวนสถานที่ท่องเที่ยวต่อวัน

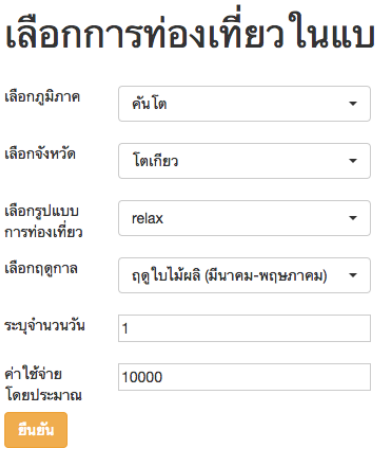
จากภาพที่ 2 การวิเคราะห์หาจำนวนสถานที่ท่องเที่ยวในแต่ละวัน ระบบวิเคราะห์ได้จากระยะเวลาที่ใช้ในการเดินทางระหว่างสถานที่ท่องเที่ยว โดยกำหนดเวลาที่ใช้ในการท่องเที่ยวของแต่ละสถานที่โดยในเบื้องต้นจะกำหนดเวลาเป็นจำนวน 2 ชั่วโมง ซึ่งภายในหนึ่งวันจะรวมเวลาที่ใช้ในการท่องเที่ยวของแต่ละสถานที่ทั้งหมดและรวมระยะเวลาที่ใช้ในการเดินทางระหว่างสถานที่ทั้งหมดจะต้องไม่เกิน 11 ชั่วโมงต่อหนึ่งวันจึงจะได้จำนวนสถานที่ท่องเที่ยวในแต่ละวันออกมาซึ่งค่าเข้าแต่ละสถานที่ท่องเที่ยวจะต้องไม่เกินงบประมาณที่ตั้งไว้และถ้าหากสถานที่ท่องเที่ยวนั้นต้องใช้เวลาเที่ยวเต็มวันก็จะตัดสถานที่ท่องเที่ยวนั้นอยู่ในวันใหม่

สำหรับการคิดค่าใช้จ่ายจะคิดรวมเฉพาะค่าที่พักและค่าเข้าชมสถานที่นั้น ๆ โดยจะไม่รวมค่าเดินทาง

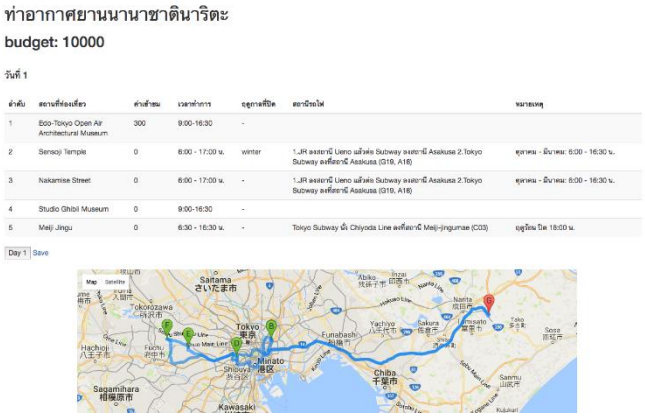
4.ผลการดำเนินงาน

4.1 ผลการดำเนินงานของระบบ

จากภาพที่ 3 ระบบจะให้ผู้ใช้เลือกเงื่อนไขที่ตรงกับความต้องการ ได้แก่ ภูมิภาค จังหวัด รูปแบบการท่องเที่ยว (Relax , Normal , Advance) และฤดูกาล ซึ่งระบบจะเลือกเฉพาะสถานที่ท่องเที่ยวที่เปิดทำการในฤดูกาลนั้นและคำนวณค่าใช้จ่ายโดยประมาณ



ภาพที่ 3 : หน้าสร้างแผนการเดินทาง



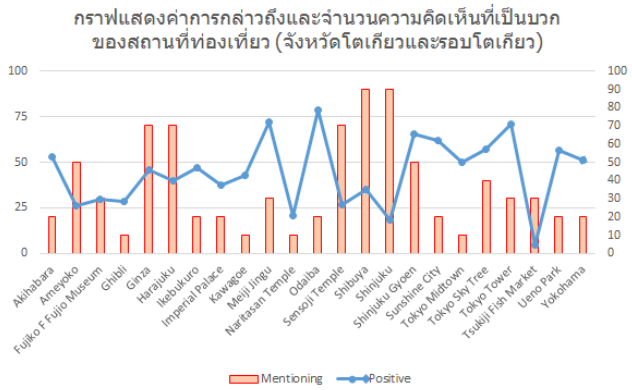
ภาพที่ 4 : หน้าแสดงแผนการเดินทาง

จากภาพที่ 4 ระบบจะแสดงแผนการเดินทางตามเงื่อนไขที่ผู้ใช้ได้เลือกไว้ โดยจะแสดงรายละเอียดดังนี้

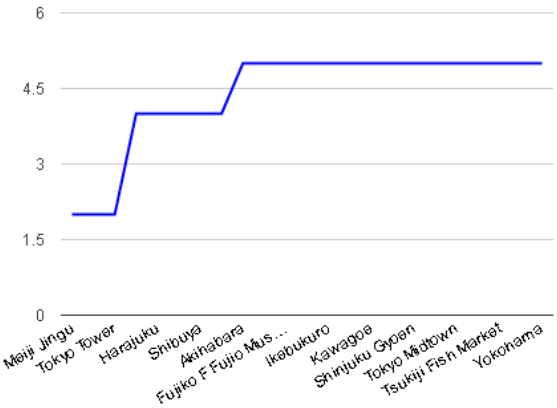
1. แสดงชื่อจุดเริ่มต้นและจุดปลายทาง
2. แสดงงบประมาณที่ผู้ใช้ได้กรอกไว้ โดยระบบจะหักค่าเข้าชมสถานที่และค่าที่พัก
3. แสดงลำดับสถานที่ท่องเที่ยวตามเส้นทางที่ดีที่สุด
4. แสดงรายละเอียดต่าง ๆ ของสถานที่ท่องเที่ยว
5. แสดงแผนที่โดยเรียงการปักหมุดตามลำดับสถานที่ท่องเที่ยวซึ่งเริ่มจากจุดเริ่มต้นซึ่งก็คือที่พักนั่นเอง

4.2 ผลการวิเคราะห์ข้อมูล

จากรูปภาพที่ 5 เป็นกราฟแสดงค่าการกล่าวถึงและจำนวนความคิดเห็นที่เป็นบวกของสถานที่ท่องเที่ยวในจังหวัดโตเกียวและรอบโตเกียว ซึ่งจะเห็นได้ว่าสถานที่ที่มีค่าการกล่าวถึงของสถานที่ท่องเที่ยวในกระทู้พันทิปนั้น ไม่สามารถนำมาใช้วิเคราะห์ค่าความนิยม ของสถานที่ท่องเที่ยวได้เพียงอย่างเดียว แต่ต้องนำจำนวนความคิดเห็นเชิงบวก มาวิเคราะห์ร่วมด้วย เพราะบางสถานที่ มีค่าการกล่าวถึงสูงแต่กลับมีจำนวนความคิดเห็นเชิงบวกน้อยหรือไม่มีเลย



ภาพที่ 5 : กราฟแสดงค่าการกล่าวถึงและจำนวนความคิดเห็นที่เป็นบวกของสถานที่ท่องเที่ยว (จังหวัดโตเกียวและรอบโตเกียว)



ภาพที่ 6 : กราฟแสดงค่าความนิยมของสถานที่ท่องเที่ยว (จังหวัดโตเกียวและรอบโตเกียว)

จากการวิเคราะห์หาค่าความนิยม โดยการใช้ค่า Mentioning และค่า Sentiment Score มาวิเคราะห์ร่วมกัน จะได้ผลลัพธ์ดังภาพที่ 6 ซึ่งจะเห็นได้ว่าสถานที่ที่มีค่านิยมสูงที่สุด ได้แก่ Akihabara, Ameyoko, Fujiko F Fujio Museum, Ghibli, Ikebukuro, Imperial Palace, Kawagoe, Naritasan Temple, Shinjuku Gyoen, Sunshine City, Tokyo Midtown, Tokyo Sky Tree, Tsukiji Fish Market, Ueno Park และ Yokohama ซึ่งสถานที่ที่มีความนิยมสูงที่สุดจะถูกดึงออกไปสร้างแผนการเดินทางก่อนเสมอ

5. บทสรุป

ระบบสร้างแผนการท่องเที่ยวจากข้อมูลออนไลน์จะสร้างแผนการท่องเที่ยวให้ผู้ใช้งานแบบอัตโนมัติ โดยระบบนำข้อมูลที่สกัดมาจากเว็บไซต์พันทิปดอทคอมมาวิเคราะห์ความถี่และวิเคราะห์อารมณ์ ระบบจะนำเงื่อนไขต่าง ๆ ที่ผู้ใช้งานได้กำหนดมา ในการสร้างแผนการท่องเที่ยว ซึ่งแผนการท่องเที่ยวจะบอกลำดับการเดินทาง ค่าเข้าชม เวลาเปิดปิด สถานีใต้ดิน/บนดินและหมายเหตุ โดยการคำนวณลำดับการเดินทางนั้นจะนำสถานที่ที่มีความนิยมสูงที่สุด ในจังหวัดนั้นๆ มาก่อนเสมอ จากนั้นจึงใช้ Google Map API มาคิดคำนวณหาเส้นทางที่ดีที่สุด และแสดงผลแผนการท่องเที่ยวออกมาในรูปแบบของตารางพร้อมแผนที่ให้กับผู้ใช้งาน

6.เอกสารอ้างอิง

[1] Japan National Tourism Organization, “Tourism Statistics” Online: https://www.jnto.go.jp/eng/ttp/sta/, 2560

[2] wannaphong, “ทำ Web Scraping ด้วยภาษา Python” *Online:* https://python3.wannaphong.com/2015/04/ทำ-web-scraping-ด้วยภาษา-python.html, 2558.

[3] Wikipedia, “Sentiment analysis” *Online:* https://en.wikipedia.org/wiki/Sentiment_analysis, 2560.

[4] Mario Muñoz, “Creating your first sentiment analysis application.” *Online:* https://github.com/EuroSentiment/eurosentiment-tutorial-python/wiki/Creating-your-first-sentiment-analysis-application, 2560.

[5] RoutePerfect, “RoutePerfect.” *Online:* https://www.routeperfect.com/, 2560.

[6] สรรค์ เลิศธีรกุล, “ระบบวางแผนการท่องเที่ยวอัจฉริยะ.” วิทยานิพนธ์วิทยาศาสตร์มหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ บัณฑิตวิทยาลัย, สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง, 2553.

[7] สุพิเชฐ สมดี และ จิรนุช โชตนรุ่งเรือง, “ระบบวางแผนแพ็กเกจทัวร์โดยใช้ Constraint-based Reasoning.” ปริญญานิพนธ์วิทยาศาสตรบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศบัณฑิตวิทยาลัย, สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง, 2550.

[8] Octoparse, “Octoparse” *Online:* http://www.octoparse.com/, 2560.

การจำแนกข้อความส่อเสียดในทวิตเตอร์ด้วยการใช้ความน่าจะเป็นของทวีต
Sarcasm Classification in Twitter using Probability of Tweets

กษิดิ์เดช ทาแป้ง (Kasidech Tapang)[1], จันทร์เจ้า มงคลนาวิน (Janjao Mongkolnavin)[2]
[1]คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย
[1]kasidech.tapang@gmail.com
[2]janjao@cbs.chula.ac.th

บทคัดย่อ

ข้อความส่อเสียดเป็นปัญหาหนึ่งในการประมวลผลภาษาธรรมชาติเนื่องจากข้อความส่อเสียดจะกลับขั้วความคิดเห็นของข้อความทำให้การวิเคราะห์ความคิดเห็นของข้อความผิดไปจากความเป็นจริง งานวิจัยนี้ได้เสนอวิธีจำแนกข้อความส่อเสียดออกจากข้อความปกติ โดยประยุกต์ใช้ความน่าจะเป็นของข้อความ และใช้ข้อมูลความคิดเห็นของผู้บริโภคเกี่ยวกับเครือข่ายอินเทอร์เน็ตเครือข่ายหนึ่งในการศึกษา โดยแบ่งการทดลองเป็นสองส่วน ส่วนที่หนึ่งเป็นส่วนการประมวลผลโดยเครื่อง ในส่วนนี้ข้อความแต่ละข้อความจะถูกแบ่งเป็นคำ และแปลงให้อยู่ในโมเดล bigram ซึ่งจะใช้ในการคำนวณความน่าจะเป็นของข้อความโดยใช้วิธีภาวะความควรจะเป็นสูงสุด (Maximum Likelihood Estimation) ในส่วนที่สองกำหนดให้บุคคลจำนวน 5 คนประเมินข้อความแต่ละข้อความว่าข้อความนั้นเป็นข้อความส่อเสียด ข้อความปกติ หรือไม่สามารถระบุได้ แล้วนำคะแนนประเมินมาหาคะแนนเฉลี่ย แล้วนำความน่าจะเป็นของข้อความและคะแนนเฉลี่ยที่ได้จากการประเมินของมนุษย์มาตรวจสอบระดับความสัมพันธ์โดยใช้สหสัมพันธ์ของเพียร์สัน จากผลการทดลองพบว่าค่า P-Value มีค่าเป็น 0.015 ซึ่งสรุปได้ว่าความน่าจะเป็นของข้อความที่คำนวณโดยเครื่องมีความสัมพันธ์ไปในทิศทางเดียวกันกับการจำแนกข้อความส่อเสียดโดยมนุษย์

คำสำคัญ: การจำแนกข้อความส่อเสียด ความน่าจะเป็นของข้อความ การประมวลผลภาษาธรรมชาติ เครือข่ายสังคมออนไลน์ ทวิตเตอร์

Abstract

Sarcasm is one of the issues in Natural Language Processing since it inverts the real sentiment of a phrase; from positive to negative. This study proposes an approach to classify sarcastic phrases, on the topic of one telecommunication service provider in Thailand gathered from Twitter. The experiment consists of two parts, phrase probability estimation by machine and by a group of five people. For the machine part, each phrase segmented into words, which are converted into a bigram model. The phrase probability is calculated from the bigram model using Maximum Likelihood Estimation. For the human part, each person rates each phrase whether it is sarcastic, typical or uncertain, then the average score is computed for each phrase. The relationship of the results from both parts are measured by using Pearson's Correlation Test. The test shows that P-Value is 0.015 which can be concluded that they are correlated in the same direction.

Keyword: Sarcasm Classification, Phrase Probability, Natural Language Processing, Social Networks, Twitter

1. บทนำ

การประมวลผลภาษาธรรมชาตินั้นมีความสำคัญต่อการประมวลผลด้วยคอมพิวเตอร์เป็นอย่างมากเนื่องจากข้อมูลที่มนุษย์ใช้อยู่ในชีวิตประจำวันนั้นมีจำนวนมาก และส่วนใหญ่จะอยู่ในรูปแบบของภาษาธรรมชาติ เช่น ภาษาพูดและภาษาเขียน ดังนั้นถ้าคอมพิวเตอร์สามารถประมวลผลข้อมูลซึ่งอยู่ในรูปแบบของภาษาธรรมชาติได้อย่างถูกต้องและตรงตามความหมายที่แท้จริง จะส่งผลให้เราสามารถใช้ประโยชน์จากข้อมูลจำนวนมหาศาลที่มีอยู่บนโลกนี้ได้อย่างมีประสิทธิภาพ อย่างไรก็ตามคุณลักษณะของภาษาแต่ละภาษาเป็นส่วนสำคัญที่ทำให้การประมวลผลภาษาธรรมชาติแตกต่างกันไปซึ่งส่งผลให้

เทคนิคการประมวลผลภาษาธรรมชาติของภาษาหนึ่งอาจไม่เหมาะสมที่จะนำมาใช้กับอีกภาษาหนึ่ง ตัวอย่างเช่น ภาษาไทยซึ่งถูกจัดอยู่ในภาษาประเภทไม่ตัดคำ และไม่มีตัวอักษรตัวใหญ่และตัวเล็ก [1] มีความแตกต่างจากภาษาอื่น ๆ ที่ตัดคำด้วยช่องว่าง เช่น ภาษาอังกฤษ ดังนั้นงานวิจัยทางด้านการประมวลผลภาษาธรรมชาติในภาษาอื่น อาจไม่สามารถนำมาประยุกต์ใช้กับภาษาไทยได้โดยตรง

นอกจากนี้ บ่อยครั้งที่ข้อมูลอาจไม่ได้อยู่ในรูปแบบที่สามารถตีความได้ตรงตามตัวอักษรซึ่งส่งผลให้การประมวลผลภาษาธรรมชาติมีความคลาดเคลื่อนไปจากความหมายที่แท้จริง เช่น "*ตั้งแต่เกิดมา ไม่เคยใช้สบู่อะไรที่ดีแบบนี้มาก่อนเลย*" ผู้เขียนอาจไม่ได้หมายความว่าสบู่นี้ดีอย่างที่พูด แต่เป็นการเขียนในเชิงส่อเสียดว่าสบู่ก้อนนี้ไม่มีคุณภาพ เป็นต้น ซึ่งหากคอมพิวเตอร์ไม่สามารถจำแนกได้ว่าความคิดเห็นนั้นเป็นความคิดเห็นตามที่ข้อความปรากฏหรือเป็นความคิดเห็นในเชิงส่อเสียด จะส่งผลต่อการสรุปผลความคิดเห็นที่มีต่อประเด็นหรือเรื่องนั้น ๆ โดยเฉพาะอย่างยิ่งในกรณีของความคิดเห็นเชิงส่อเสียดที่มีวัตถุประสงค์ที่จะแสดงความเห็นในเชิงลบอย่างมาก ในขณะที่คอมพิวเตอร์ประมวลผลออกมาเป็นความคิดเห็นเชิงบวกอย่างมาก งานวิจัยนี้มีวัตถุประสงค์เพื่อนำเทคโนโลยีทางด้านการประมวลผลภาษาธรรมชาติและการเรียนรู้ของเครื่องมาใช้ในการจำแนกข้อความส่อเสียด (Sarcasm Classification) ออกจากข้อความปกติในภาษาไทยเพื่อสรุปความหมายที่แท้จริงที่ผู้เขียนข้อความต้องการสื่อซึ่งจะช่วยให้การวิเคราะห์ข้อความโดยเครื่องเป็นไปได้อย่างถูกต้องและแม่นยำมากยิ่งขึ้น

2. วรรณกรรมที่เกี่ยวข้อง

2.1 การประมวลผลภาษาธรรมชาติ

ในปัจจุบันข้อมูลข่าวสารบนโลกอินเทอร์เน็ตล้วนมีความสำคัญอย่างมากสำหรับองค์กร เนื่ององค์กรสามารถนำข้อมูลต่าง ๆ เช่น ผลตอบรับของสินค้าและบริการ (Feedback) มาใช้เพื่อปรับปรุงคุณภาพของสินค้าและบริการให้ดียิ่งขึ้น อย่างไรก็ตามข้อมูลดังกล่าวเป็นข้อมูลที่ไม่มีโครงสร้างที่แน่นอน (Unstructured Data) จึงจำเป็นต้องนำเทคนิคการประมวลผลภาษาธรรมชาติ (Natural Language Processing) มาใช้เพื่อแปลงข้อมูลให้อยู่ในรูปแบบที่คอมพิวเตอร์เข้าใจและสามารถนำไปใช้ในการประมวลผลอื่น ๆ ต่อไปได้

การประมวลผลภาษาธรรมชาติมีการนำไปประยุกต์ในงานด้านต่าง ๆ เช่น การสรุปใจความสำคัญของบทความ การจำแนกชนิดของบทความ [2] การหาประโยคหรือคำที่เกี่ยวข้อง [3] การวิเคราะห์ขั้วและระดับความคิดเห็นของบทวิจารณ์สินค้าโดยผู้บริโภค [4, 5] หรือการแปลภาษาจากภาษาหนึ่งเป็นอีกภาษาหนึ่ง [6] เป็นต้น ศาสตร์ในการประมวลภาษาธรรมชาตินั้นเกี่ยวเนื่องกับปัญญาประดิษฐ์ ซึ่งเป็นศาสตร์ที่กำลังมีการพัฒนาอย่างไม่หยุดยั้ง โดยเชื่อว่าในอนาคตคอมพิวเตอร์จะสามารถเข้าใจและเรียนรู้ข้อมูลในรูปแบบของภาษาธรรมชาติได้ดีมากยิ่งขึ้นตามลำดับ [7]

2.2 งานวิจัยที่เกี่ยวข้องกับการจำแนกข้อความส่อเสียด

ในการจำแนกข้อความส่อเสียดออกจากข้อความปกตินั้น คุณลักษณะของภาษาถือว่าเป็นส่วนสำคัญที่สามารถนำมาใช้ได้ โดยส่วนใหญ่จะดำเนินการโดยสร้างแบบจำลอง N-gram โดยแบ่งข้อความออกเป็นคำ จากนั้นจึงนับจำนวนคำที่เกิดขึ้นในข้อมูลทั้งหมด [8, 9, 10] และใช้จำนวนการปรากฏของคำในแต่ข้อความเป็นองค์ประกอบ (Feature)

นอกจากการใช้ความถี่ในการปรากฏของคำในการจำแนกข้อความส่อเสียดแล้ว เครื่องหมายพิเศษต่าง ๆ เช่น เครื่องหมายอัศเจรีย์ (!) เครื่องหมายปรัศนี (?) หรือสัญลักษณ์พิเศษต่าง ๆ ก็ถูกนับเพิ่มเป็นส่วนหนึ่งขององค์ประกอบในการจำแนกข้อความส่อเสียดออกจากข้อความปกติ [8, 10]

ไอคอนแสดงอารมณ์ก็ถือเป็นอีกองค์ประกอบหนึ่งที่นิยมใช้ในการจำแนกข้อความส่อเสียดออกจากข้อความปกติ [6] เช่น การนับจำนวนครั้งที่ปรากฏขึ้นของไอคอนแสดงอารมณ์ซึ่งมีค่าแสดงอารมณ์ในเชิงบวกและเชิงลบ

Pt´cek [8] ได้ศึกษาการจำแนกข้อความส่อเสียดออกจากข้อความปกติโดยใช้การนับจำนวนของคำที่ใช้ตัวพิมพ์ใหญ่เป็นองค์ประกอบ โดยพบว่าประโยคที่ใช้ตัวอักษรพิมพ์ใหญ่ทั้งหมด เช่น *"THE WEATHER TODAY IS EXTREMELY GOOD."* มีแนวโน้มที่จะเป็นข้อความส่อเสียดมากกว่าประโยค *"The weather today is extremely good."*

2.3 ข้อจำกัดของการศึกษาที่ผ่านมา

จากการศึกษาที่ผ่านมาพบว่าการจำแนกข้อความส่อเสียดนั้นมักจะมีการศึกษากันอย่างแพร่หลายในภาษาต่างประเทศต่าง ๆ เช่น ภาษาอังกฤษ [8, 9, 10] แต่ไม่มีการศึกษาในหัวข้อนี้มากนักในภาษาไทย อีกทั้งผู้วิจัยมีความเห็นว่าการใช้องค์ประกอบอื่น ๆ นอกเหนือจากภาษาในการจำแนกข้อความส่อเสียด เช่น การใช้สัญลักษณ์พิเศษต่าง ๆ ไอคอนแสดงอารมณ์ หรือการใช้ตัวอักษรพิมพ์ใหญ่ติดกัน มักจะขึ้นอยู่กับลักษณะและวัฒนธรรมของภาษาและบุคคลในประเทศนั้น ๆ เช่น ในภาษาไทย จะไม่ปรากฏลักษณะของการใช้ตัวอักษรพิมพ์ใหญ่ ซึ่งจะส่งผลให้เทคนิคการจำแนกข้อความส่อเสียดโดยใช้องค์ประกอบอื่น ๆ ที่ไม่ใช่คำอาจไม่เหมาะสมกับการนำมาประยุกต์กับภาษาไทย

ผู้วิจัยเชื่อว่า การใช้ความน่าจะเป็นของข้อความในการจำแนกข้อความส่อเสียดออกจากข้อความปกติ อาจเป็นอีกวิธีหนึ่งที่น่าสนใจศึกษา โดยผู้วิจัยตั้งสมมุติฐานว่าข้อความปกติควรจะมีความน่าจะเป็นที่ bigram ของคำที่ปรากฏในข้อความจะปรากฏในลำดับดังกล่าวสูงกว่าข้อความส่อเสียดเนื่องจากอัตราส่วนข้อความส่อเสียดจะค่อนข้างน้อยเปรียบเทียบกับจำนวนของข้อความปกติทั่ว ๆ ไป [11] อีกทั้งคำที่ปรากฏร่วมกันในข้อความส่อเสียดนั้นจะมีแนวโน้มว่าเป็นคำที่ไม่เป็นไปตามธรรมชาติ เช่น การใช้คำว่า ความเร็วแสงกับอินเทอร์เน็ต ในประโยค “*อินเทอร์เน็ตยี่ห้อนี้เร็วมากเร็วกว่าความเร็วแสงอีก*” ดังนั้นหากข้อความไหนมีความน่าจะเป็นซึ่งคำนวณจากองค์ประกอบและลำดับของคำที่ใช้ในข้อความในระดับที่ต่ำ นั่นหมายความว่าข้อความนั้นมีโอกาสที่จะเกิดต่ำและมีความเป็นไปได้ที่ข้อความดังกล่าวจะเป็นข้อความส่อเสียด

ผู้วิจัยจึงมีความสนใจที่จะศึกษาประสิทธิภาพของการประยุกต์ใช้ความน่าจะเป็นของข้อความในการจำแนกข้อความส่อเสียดออกจากข้อความปกติรวมถึงศึกษาข้อจำกัดและปัจจัยที่เกี่ยวข้อง

3. การคำนวณความน่าจะเป็นของประโยค

ลำดับ N-gram คือลำดับที่ต่อเนื่องกันของ N รายการ ตัวอย่างของการแปลงข้อความให้อยู่ในลักษณะ N-gram เช่น หากค่า N มีค่าเป็น 1 แล้วข้อความ “I went to school” จะถูกสามารถสร้างเป็น [I, went, to, school] หรือสามารถเรียกแทนได้ว่า Unigram และในกรณีที่ N มีค่าเป็น 2 ข้อความข้างต้นก็จะสามารถจำแนกได้เป็น [I went, went to, to school] ในการศึกษาครั้งนี้ ผู้วิจัยเลือกใช้โครงสร้างข้อมูลแบบ 2-gram หรือ Bigram ในระดับคำ และนำ Bigram ที่ได้ไปสร้างองค์ประกอบทวิภาค (Binary Feature) เพื่อใช้ในการจำแนกข้อความส่อเสียดต่อไป

การคำนวณความน่าจะเป็นที่จะเกิดคำใด ๆ ในข้อความสามารถทำได้โดยพิจารณาจากคำก่อนหน้า k-1 คำ เช่น หากกำหนดค่า N ของ N-gram มีค่าเท่ากับ 2 (Bigram) และประโยค $W = (w_1, w_2, w_3 \ldots w_m)$ ความน่าจะเป็นที่จะเกิดคำ w_2 หลังคำ w_1 นั้นสามารถหาได้จาก $P(w_2|w_1)$ หรือสามารถเขียนให้อยู่ในรูปแบบสมการโดยใช้ความถี่สัมพัทธ์ (Relative Frequency) ซึ่งถือว่าเป็นการประมาณโดยใช้ Maximum Likelihood Estimation (MLE) ได้ตามสมการที่ 1 เนื่องจากอัตราส่วนระหว่างความถี่ที่ลำดับของคำ w_{i-1} จะปรากฏตามด้วยคำ w_i และความถี่ที่คำ w_{i-1} ปรากฏนั้นให้ความน่าจะเป็นสูงสุดที่ลำดับของคำจะเกิดขึ้นในฐานข้อมูลชุดนั้นๆ แล้ว [12]

$$P(w_i \mid w_{i-1}) = \frac{Count(w_{i-1}, w_i)}{Count(w_{i-1})} \qquad \text{สมการ 1}$$

เนื่องจากในหนึ่งข้อความจะประกอบด้วยคำหลายคำ ดังนั้นในการหาความน่าจะเป็นของข้อความนั้นจะใช้ความน่าจะเป็นร่วม (Joint Probability) โดยจะคิดความน่าจะเป็นของคำที่อยู่ติดกันแทน เช่น กำหนดให้ข้อความ W ประกอบด้วยคำ w1, w2, w3, … , wn จะสามารถคำนวณหาความน่าจะเป็นของข้อความ P(w1, w2, w3, … , wn) โดยนำกฎลูกโซ่แห่งความน่าจะเป็น (Chain rule of probability) ซึ่งแสดงในสมการที่ 2 มาใช้

$$\begin{aligned} P(w_1 \ldots w_n) &= P(w_1)P(w_2|w_1)P(w_3|w_2w_1) \ldots P(w_n|w_{n-1}w_{n-2} \ldots w_1) \\ &= P(w_1)\prod_{k=2}^{n} P(w_k|w_{k-1}w_{k-2} \ldots w_1) \end{aligned} \qquad \text{สมการ 2}$$

ซึ่งกฎลูกโซ่ในสมการที่ 2 แสดงถึงการคำนวณหาความน่าจะเป็นร่วมของลำดับ (Joint probability of a sequence) และการคำนวณหาความน่าจะเป็นแบบมีเงื่อนไข (Conditional Probability) ของคำ w โดยกำหนดคำก่อนหน้า w-1 คำ ซึ่งจาก

สมการจะเห็นว่า จะสามารถคำนวณความน่าจะเป็นของข้อความได้โดยนำความน่าจะเป็นแบบมีเงื่อนไขของคำแต่ละคำที่ปรากฏตามลำดับมาคูณกัน และเนื่องจากสมมุติฐานของ Markov [12] ซึ่งระบุว่าเราสามารถทำนายความน่าจะเป็นของหน่วยในอนาคตบางหน่วยได้โดยไม่จำเป็นต้องมองย้อนกลับไปยังอดีตมากเกินไป ดังนั้นจึงสามารถประมาณความน่าจะเป็นแบบมีเงื่อนไขสำหรับคำแต่ละคำในข้อความได้ตามสมการที่ 3

$$P(w_n|w_{n-1}w_{n-2}\dots w_1) \approx P(w_n|w_{n-1}) \qquad \text{สมการ 3}$$

ดังนั้นความน่าจะเป็นของข้อความในสมการที่ 2 จึงสามารถคำนวณได้จาก

$$P(w_1..w_n) \approx P(w_1)\prod_{k=2}^{n} P(w_k|w_{k-1}) \qquad \text{สมการ 4}$$

4. ข้อมูลที่ใช้ในการทดลอง

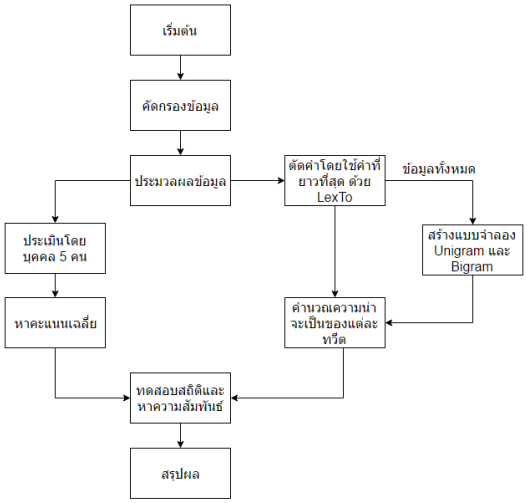
ผู้วิจัยใช้ข้อมูลความคิดเห็นบนเครือข่ายสังคมออนไลน์ทวิตเตอร์เป็นข้อมูลในการทดลอง เนื่องจากข้อความความคิดเห็นต่าง ๆ นั้นมีการเผยแพร่ในเครือข่ายสังคมออนไลน์เป็นส่วนมาก รวมถึงข้อความบนเครือข่ายสังคมออนไลน์ทวิตเตอร์นั้นมีการจำกัดตัวอักษรที่ 140 ตัวอักษร ซึ่งทำให้ความคิดเห็นบนทวิตเตอร์มักเป็นความคิดเห็นเรื่องเดียวและประโยคเดียวต่อ 1 ข้อความซึ่งจะเพิ่มความสะดวกในการวิเคราะห์ นอกจากนี้เครือข่ายสังคมออนไลน์ทวิตเตอร์นั้นไม่มีการจำกัดการเข้าถึงข้อความที่ผู้ใช้เผยแพร่เท่าเครือข่ายสังคมออนไลน์อื่น ๆ

การรวบรวมข้อมูลทำโดยการค้นหาคีย์เวิร์ดผ่านทาง Advanced Search ของเครือข่ายสังคมออนไลน์ทวิตเตอร์โดยค้นหาชื่อผู้ให้บริการอินเทอร์เน็ต X ในประเทศไทย ควบคู่ไปกับการค้นหาโดยใช้แฮชแทกที่เกี่ยวข้อง โดยข้อความที่ใช้ในการค้นหาคือ “X อินเทอร์เน็ต” “Xอินเตอร์เน็ท ” “อินเทอร์เน็ต X” “อินเตอร์เน็ทX” “เน็ตX” “เน็ทX” “XInternet” “InternetX” “X Internet” และ “Internet X”และกำหนดให้ค้นหาข้อความที่เป็นภาษาไทยเท่านั้น โดยเก็บข้อมูลความคิดเห็นของผู้ใช้บริการทั้งหมดระหว่างวันที่ 25 มกราคม 2553 ถึง 9 มิถุนายน 2559 ได้ข้อความทั้งหมด 3,844 ข้อความ จากการจำแนกเบื้องต้นโดยผู้วิจัย พบว่าเป็นข้อความปกติ 3,736 ข้อความและเป็นข้อความส่อเสียด 108 ข้อความ

5. การทดลอง

ผู้วิจัยแบ่งการทดลองออกเป็นสองส่วนซึ่งดำเนินการบนข้อมูลชุดเดียวกัน คือการจำแนกข้อความส่อเสียดโดยมนุษย์และการจำแนกข้อความส่อเสียดโดยใช้ความน่าจะเป็นของข้อความและนำผลมาเปรียบเทียบกัน เพื่อให้ทราบว่าเทคนิคที่ศึกษามีศักยภาพในการนำมาประยุกต์กับการจำแนกข้อความส่อเสียดได้มากน้อยเพียงใด โดยสามารถสรุปขั้นตอนการทดลองได้ดังรูปที่ 1

ผู้วิจัยเริ่มต้นด้วยการกรองข้อมูลที่ได้จากการค้นหาโดยการอ่านและตัดสินใจว่าข้อความข้างต้นเป็นข้อความความคิดเห็นที่เกี่ยวเนื่องกับผู้ให้บริการอินเทอร์เน็ตที่สนใจหรือไม่ ซึ่งหากเป็นข้อมูลที่ไม่เกี่ยวข้องกับผู้ให้บริการอินเทอร์เน็ตที่สนใจจะทำการตัดทวีตดังกล่าวทิ้ง โดยเหลือจำนวนข้อมูลหลังผ่านการกรองแล้วทั้งสิ้นจำนวน 3,844 ข้อความ



รูป 1 ขั้นตอนการทดลอง

ในส่วนของข้อความทั้ง 3,844 ข้อความ จะได้รับการประมวลผลเบื้องต้นดังนี้

1. ตัดข้อความซ้ำซ้อนรวมถึงข้อความที่ทวีตซ้ำ
2. ตัด URL ออกจากข้อความ
3. ตัดแฮชแทกออกจากข้อความ
4. ตัดเครื่องหมาย @ และชื่อบุคคลออกจากข้อความ
5. ตัดตัวอักษรที่ซ้ำกันออกจากข้อความ
6. ตัดอักขระพิเศษ เช่น ! # $ ออกจากข้อความ

จากนั้นในส่วนของการประมวลผลโดยเครื่อง ผู้วิจัยนำข้อความทวีตที่เหลือจากขั้นตอนประมวลผลเบื้องต้นไปตัดคำ

โดยใช้เครื่องมือตัดคำ LexTo[1] ซึ่งเป็นการตัดคำโดยใช้วิธีตัดคำที่ยาวที่สุด (Longest Matching) ดังตัวอย่างในรูปที่ 2 ซึ่งส่งผลให้ข้อความดังกล่าวถูกแปลงเป็นคำในรูปแบบ Unigram จากนั้นตัดคำที่ไม่มีความหมาย เช่น "*สัญาณ*" หรือ "*เปน*" ออกแล้วจึงนำ Unigram ที่เหลือ ในแต่ละข้อความมาสร้างเป็น Bigram เพื่อประกอบการคำนวณความน่าจะเป็นของข้อความตามสมการที่ 4

<s>|ไม่ทัน|ขอบคุณ|เน็ทX|ค่ะ|ที่|แรง|ขนาด|นี้|</s>
<s>|เน็ทX|ตอนกลางคืน|นี่|แบบ|บ|</s>

รูปที่ 2 ตัวอย่างการตัดคำโดยใช้เครื่องมือ LexTo

ในส่วนของการประเมินคะแนนโดยมนุษย์นั้น โดยทั่วไปผู้อ่านข้อความอาจไม่สามารถทราบเจตนาของผู้เขียนได้ถูกต้อง 100% และในทางปฏิบัตินั้นผู้เขียนข้อความก็ไม่จำเป็นที่จะต้องระบุว่าข้อความที่ตนเองเขียนนั้นเป็นข้อความส่อเสียดหรือไม่ ผู้วิจัยจึงเห็นว่าในการศึกษาครั้งนี้ควรกำหนดให้บุคคลอื่นที่มีประสบการณ์การใช้เครือข่ายสังคมออนไลน์ทวิตเตอร์จำนวน 5 คน มาประเมินว่าข้อความที่ผ่านการประมวลผลเบื้องต้นชุดเดียวกับที่ใช้ในการทดลองโดยเครื่องแต่ละข้อความนั้นว่าเป็นข้อความส่อเสียด ไม่เป็นข้อความส่อเสียด หรือไม่แน่ใจว่าเป็นข้อความส่อเสียดหรือไม่ โดยให้คะแนน 0, 1 และ 0.5 ตามลำดับ จากนั้นนำผลที่ได้มาหาค่าเฉลี่ยเป็นค่าความน่าจะเป็นของแต่ละข้อความโดยหากคะแนนมาก จะสรุปว่าข้อความนั้นมีโอกาสเป็นข้อความปกติมาก แต่หากคะแนนน้อย จะสรุปว่าข้อความนั้นมีโอกาสเป็นข้อความปกติน้อย หรือมีโอกาสที่จะเป็นข้อความส่อเสียด โดยผู้วิจัยกำหนดให้บุคคล 5 คนดังกล่าวให้คะแนนข้อความวันละ 100 ข้อความเพื่อลดปัจจัยรบกวนที่เกิดจากความเหนื่อยล้าในการประเมินข้อความจำนวนมากในระยะเวลาที่ต่อเนื่องกัน และดำเนินการเป็นระยะเวลา 40 วันตั้งแต่วันที่ 10 มิถุนายน 2559 จนถึงวันที่ 19 กรกฎาคม 2559 แล้วสรุปรวมคะแนนความส่อเสียดของแต่ละข้อความโดยนำคะแนนที่ได้จากบุคคลที่ทำการประเมินทั้งห้าคนมาเฉลี่ยกัน

ในขั้นตอนสุดท้ายจะเป็นการสรุปผลการทดลองเพื่อสรุปประสิทธิภาพของการจำแนกข้อความส่อเสียดโดยเทคนิคที่ศึกษาโดยเปรียบเทียบกับการจำแนกข้อความส่อเสียดโดยมนุษย์ว่ามีความสอดคล้องกันมากน้อยเพียงใดโดยใช้การทดสอบทางสถิติด้วย Pearson Correlation Test

[1] โปรแกรมตัดคำ LexTo http://www.sansarn.com/lexto/

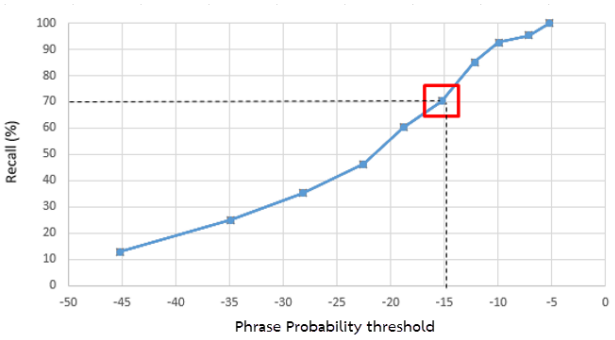
6. สรุปผลการทดลอง

จากข้อความทั้งหมด 3,844 ข้อความ ค่าความน่าจะเป็นของประโยคที่คำนวณได้มีค่าอยู่ระหว่าง 5.77×10^{-84} และ 1.16×10^{-2} เนื่องจากค่าความน่าจะเป็นมีค่าเข้าใกล้ 0 มาก ดังนั้นผู้วิจัยจึงใช้ฟังก์ชันลอการิทึมเพื่อลดความไม่เสถียรทาง (Numerical Instability) โดยค่าความน่าจะเป็นของข้อความหลังใช้ฟังก์ชันลอการิทึมมีค่าอยู่ระหว่าง -83.24 และ -1.94 โดยมีตัวอย่างผลลัพธ์ดังแสดงในตารางที่ 2

ตารางที่ 2 ตัวอย่างการคำนวณความน่าจะเป็นของประโยค

ข้อความ	Probability	Log Probability
<s>|ไม่ทัน|ขอบคุณ|เน็ทX|ค่ะ|ที่|แรง|ขนาด|นี้|</s>	4.41×10^{-13}	-16.98
<s>|เน็ทX|ตอนกลางคืน|นี่|แบบ|บ|</s>	2.02×10^{-8}	-9.46

กราฟที่ 1 แสดงค่าความระลึกที่เส้นแบ่ง (Threshold) ค่าความน่าจะเป็นของข้อความแต่ละช่วงโดยหากความน่าจะเป็นของข้อความมีค่าน้อยกว่าค่าเส้นแบ่ง ข้อความนั้นจะถูกจำแนกเป็นข้อความส่อเสียด ผู้วิจัยพบว่า หากกำหนดเส้นแบ่งค่าความน่าจะเป็นไว้ที่ -15.24 จะสามารถครอบคลุมข้อความส่อเสียดในชุดข้อมูลได้ถึงร้อยละ 70.37



กราฟที่ 1 ความสัมพันธ์ระหว่างค่าความระลึกและเส้นแบ่งในแต่ละช่วง

เพื่อสรุปผลการทดลอง ผู้วิจัยได้ใช้โปรแกรม SPSS ในการทดสอบสมมุติฐานโดยกำหนดค่าความน่าจะเป็นของข้อความที่ได้จากการประเมินโดยมนุษย์และความน่าจะเป็น

ของข้อความที่ได้จากโปรแกรมเป็นตัวแปรที่ใช้ในการคำนวณหาความสัมพันธ์ และกำหนดช่วงความเชื่อมั่นที่ร้อยละ 95 และใช้การทดสอบสมมุติฐานแบบทางเดียว ซึ่งพบว่าค่า p-value มีค่าเท่ากับ 0.006 ซึ่งมีค่าน้อยกว่า 0.05 จึงสรุปได้ว่าจากหลักฐานข้อมูลที่ได้รวบรวมมา มีความเพียงพอที่จะปฏิเสธ H0 ที่ว่า ความน่าจะเป็นของข้อความที่จำแนกโดยเทคนิคที่ศึกษาไม่มีความสัมพันธ์ไปในทางเดียวกันกับความน่าจะเป็นของข้อความที่ได้จากการประเมินโดยบุคคล หรือความน่าจะเป็นของข้อความมีความสัมพันธ์เป็นไปในทิศทางเดียวกันกับค่าความน่าจะเป็นของข้อความที่ได้จากการประเมินโดยบุคคล

ในส่วนของระดับความสัมพันธ์ ผู้วิจัยพบว่าค่า r ซึ่งแสดงถึงระดับความสัมพันธ์ มีค่าเท่ากับ 0.035 เนื่องจากคะแนนเฉลี่ยที่ได้จากการประเมินโดยมนุษย์ไม่ใช่ข้อมูลเชิงปริมาณแท้จริง จากค่าระดับความสัมพันธ์ 0.035 สามารถสรุปได้ว่า ความน่าจะเป็นของข้อความและคะแนนความส่อเสียดที่ได้จากการประเมินของมนุษย์มีความสัมพันธ์กัน

7. ข้อจำกัดของการศึกษา

เนื่องจากงานวิจัยนี้ทำการทดลองภายใต้ข้อมูลเพียงชุดเดียวในหัวข้อความคิดเห็นของผู้บริโภคเกี่ยวกับเครือข่ายอินเตอร์ X เท่านั้น ผู้วิจัยวางแผนที่จะทำการทดลองเพิ่มเติมกับข้อมูลชุดอื่นว่าจะได้ผลเช่นเดียวกันหรือไม่ และในอนาคตอาจมีการนำ k-skip-2-gram มาใช้ในการคำนวณค่าความน่าจะเป็นของ bigram model ซึ่งอาจจะนำไปสู่ข้อสรุปด้านประสิทธิภาพของเทคนิคที่เสนอที่ชัดเจนมากยิ่งขึ้น รวมถึงผู้วิจัยวางแผนที่จะเพิ่มจำนวนบุคคลที่ประเมินคะแนนความส่อเสียดของข้อความให้มากกว่า 5 คนเพื่อจะทำให้มาตราส่วนคะแนนเฉลี่ยที่จะได้มีความละเอียดมากขึ้นซึ่งอาจส่งผลให้ผลสรุปที่ได้จากการศึกษามีความชัดเจนมากยิ่งขึ้นในอนาคต

เอกสารอ้างอิง

[1] Haruechaiyasak, C. (2010). Basic NLP Tools and Text Mining Applications

[2] Yatsko, V. et al Automatic genre recognition and adaptive text summarization. In: Automatic Documentation and Mathematical Linguistics, 2010, Volume 44, Number 3, pp.111-120.

[3] Badrul Sarwar, George Karypis, Joseph Konstan, and John Riedl. 2001. Item-based collaborative filtering recommendation algorithms. In Proceedings of the 10th international conference on World Wide Web (WWW '01). ACM, New York, NY, USA, 285-295.

[4] Thumrongluck, T., Mongkolnavin, J. (2011). A Development of an Automated Thai Customer Review Summarizing System, Chulalongkorn Business Review 33(128): 40-62, 2011

[5] Apisuwankun, P., Mongkolnavin, J. (2013). Opinion Strength Identification in Customer Review Summarizing System Using Association Rule Technique, The International Conference on E-Technologies and Business on the Web (EBW2013), Bangkok, Thailand

[6] Y. Wu, et al. Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation. ArXiv e-prints, 2016.

[7] Chopra, A., Prashar, A., & Sain, C. (2013). Natural Language Processing. International Journal of Technology Enchancements and Emerging Engineering Research, 1(4), 131.

[8] Pt´cek, T. a. s., Habernal, I., & Hong, J. (2014). Sarcasm Detection on Czech and English Twitter. Paper presented at the Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers.

[9] Bamman, D., & Smith, N. A. (2015). Contextualized Sarcasm Detection on Twitter. Paper presented at the Proceedings of the International AAAI Conference on Weblogs and Social Media, Oxford, UK.

[10] Gonzlez-Ibez, R., Muresan, S., & Wacholder, N. (2011). Identifying sarcasm in Twitter: a closer look. Paper presented at the Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies: short papers - Volume 2, Portland, Oregon.

[11] Hsu, V., & Jain, S. (2015). The Lowest Form of Wit: Identifying Sarcasm in Social Media the 14th Annual CS 229 Machine Learning. Arrillaga Center for Sports and Recreation (ACSR)

[12] Jurafsky, D., & Martin, J. H. (2000). Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition: Prentice Hall PTR

การพัฒนาระบบธุรกิจอัจฉริยะ เพื่อสนับสนุนการตัดสินใจ สำหรับผู้บริหารสถาบันการศึกษา
BI System Development to Support Decision of Educational Administrators

วารุณี แต้มคู (Warunee Tamkoo)[1] และกฤษณะ ไวยมัย (Kitsana Waiyamai)[2]
[1]ภาควิชาวิศวกรรมคอมพิวเตอร์, คณะวิศวกรรมศาสตร์, มหาวิทยาลัยเกษตรศาสตร์
[2]ภาควิชาวิศวกรรมคอมพิวเตอร์, คณะวิศวกรรมศาสตร์, มหาวิทยาลัยเกษตรศาสตร์
[1]Warunee.ta@ku.th, [2]Fengknw@ku.ac.th

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาระบบธุรกิจอัจฉริยะสำหรับผู้บริหารสถาบันการศึกษา เพื่อช่วยให้ผู้บริหารมีข้อมูลที่ถูกต้อง สามารถตอบปัญหาเชิงธุรกิจได้อย่างรวดเร็วทันต่อเหตุการณ์ การดำเนินการเริ่มต้นจากการเก็บความต้องการของผู้บริหารในสถาบันการศึกษาจากนั้นทำการวิเคราะห์และสรุปเพื่อนำมาออกแบบคลังข้อมูล และรายงานที่ตอบสนองความต้องการของผู้บริหาร ระบบพัฒนาโดยใช้ Pentaho ซึ่งเป็นเครื่องมือโอเพนซอร์ส BI และใช้เทคนิคการประมวลผลในเชิงวิเคราะห์แบบออนไลน์ (OLAP) ซึ่งเป็นเทคนิคที่ช่วยนำเสนอข้อมูลในหลายมิติ แสดงผลบนเว็บแอพพลิเคชั่น มีการสกัดการเปลี่ยนแปลงและถ่ายโอนข้อมูลไปยังคลังข้อมูล ผลการวิจัยพบว่า ระบบที่พัฒนาขึ้นช่วยให้ผู้บริหารสามารถเรียกดูข้อมูลรายงาน ได้ในหลายมุมมอง และได้รับข้อมูลที่รวดเร็วขึ้น รวมถึงสามารถนำไปใช้ในการวางแผนกลยุทธ์ขององค์กรได้อย่างมีประสิทธิภาพ จากการประเมินคุณภาพของระบบมีค่าเฉลี่ยโดยรวมทุกด้านเท่ากับ 4.46 และค่าเบี่ยงเบนมาตรฐานเท่ากับ 2.19 สรุปผลได้ว่าการประเมินอยู่ในระดับดี

คำสำคัญ: ระบบสนับสนุนการตัดสินใจ ระบบธุรกิจอัจฉริยะ

Abstract

The objective of this research work is to design and develop a set of open-source business intelligence tools for supporting the educational administrators. In the design phase, the requirements are collected from real institutions administrators, an BI-Education data model template and a set of BI-Education report templates are designed from the collected requirements. In the implementation phase, a set of Pentaho open-source business intelligence tools is developed, including the data warehouse, the Extract Transform and Load (ETL), and the reporting tools. In terms of user complacency, average 4.46 and standard deviation 2.19 are obtained as results of the questionnaire on 10 educational administrators.

Keyword: Business Intelligence, Data Warehouse, BI, OLAP

1. บทนำ

ในยุคของเทคโนโลยีสารสนเทศหลายองค์กรมักประสบกับปัญหาการนำข้อมูลที่มีอยู่มหาศาลมาเปลี่ยนเป็นสารสนเทศที่มีคุณค่าต่อองค์กร การนำเทคโนโลยีสารสนเทศที่สมบูรณ์มาช่วยสนับสนุนการตัดสินใจในการดำเนินธุรกิจและการวางแผนงานขององค์กร เนื่องจากธุรกิจการศึกษาเป็นธุรกิจที่ไม่แสวงหาผลกำไรแต่ยังจำเป็นจะต้องใช้ข้อมูลสารสนเทศในการบริหารงานเพื่อให้องค์กรสามารถดำเนินธุรกิจได้อย่างมีประสิทธิภาพ

ปัจจุบันสถาบันการศึกษาส่วนใหญ่มีการจัดเก็บข้อมูลแบบกระจัดกระจายโดยข้อมูลเหล่านี้จะถูกสร้างและจัดเก็บด้วยรูปแบบและโปรแกรมที่แตกต่างกันทำให้ยากต่อการเรียกดูข้อมูลในการนำมาวิเคราะห์และตัดสินใจ เมื่อผู้บริหารต้องการใช้ข้อมูลเจ้าหน้าที่ต้องดำเนินการจัดทำข้อมูลในรูปแบบที่ผู้บริหารต้องการ ทำให้ใช้ระยะเวลาในการดำเนินงานมาก และไม่ทันต่อการใช้งาน รวมถึงยังอาจทำให้เกิดข้อผิดพลาดได้ ดังนั้นจึงเป็นสาเหตุของแนวคิดในการพัฒนาระบบธุรกิจ

อัจฉริยะสำหรับสนับสนุนการตัดสินใจของผู้บริหารในสถาบันการศึกษา

การศึกษาครั้งนี้ทำการเก็บรวบรวมความต้องการของผู้บริหารมาวิเคราะห์ และสรุปเป็นรายงานที่ผู้บริหารจำเป็นต้องใช้ รวมถึงแดชบอร์ดที่เป็นข้อมูลพื้นฐานที่ต้องใช้ในการบริหารงานหลัก [11] ซึ่งทุกสถาบันการศึกษาสามารถนำไปใช้เป็นต้นแบบได้ รวมถึงออกแบบ และพัฒนาโดยใช้ Pentaho ซึ่งเป็นโอเพนซอร์ส BI ที่มีฟังก์ชัน การทำงานครบและได้รับความนิยมจากบริษัทชั้นนำหลายบริษัท [2], [6] เป็นระบบธุรกิจอัจฉริยะที่เป็นเครื่องมือในการรวบรวมข้อมูล สร้างมุมมองของข้อมูล (Cube) สร้างรายงานที่ตอบสนองความต้องการของผู้บริหารในการวางแผนการดำเนินงาน การกำหนดกลยุทธ์ขององค์กรและตอบปัญหาเชิงธุรกิจได้อย่างรวดเร็วและมีประสิทธิภาพ

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 Business Intelligence (BI)

ระบบธุรกิจอัจฉริยะ [4], [10] เป็นกระบวนการรวบรวมข้อมูลขนาดใหญ่ ที่เก็บข้อมูลจากฐานข้อมูลที่ใช้ในระบบงานประจำวันหรือฐานข้อมูลอื่นทั้งที่อยู่ภายในและภายนอกองค์กรนำมาจัดเก็บให้เหมาะสมกับสถานการณ์ สามารถนำข้อมูลเหล่านั้นมาใช้ในการวิเคราะห์และแสดงผลลัพธ์ในรูปแบบของรายงานหรือกราฟ เพื่อให้ผู้บริหารสามารถวิเคราะห์ปัญหา และใช้ในการปรับปรุงประสิทธิภาพการดำเนินงาน [5] รวมถึงช่วยในการสนับสนุนการตัดสินใจได้อย่างรวดเร็ว ถูกต้อง และตอบปัญหาทางธุรกิจในเชิงลึกได้

2.2. คลังข้อมูล (Data Warehouse)

คลังข้อมูลเป็นแหล่งรวบรวมข้อมูลขนาดใหญ่ที่มีความสำคัญ โดยจะนำข้อมูลที่ถูกรวบรวมมาจัดเก็บ วิเคราะห์และสนับสนุนการตัดสินใจในการวางแผน ระบบจะทำการเก็บรวบรวมข้อมูลจากฐานข้อมูลที่ใช้ในระบบงานประจำวัน ซึ่งไม่สามารถตอบสนองความต้องการของผู้บริหารได้ โดยรายการข้อมูลถูกประมวลผลจากฐานข้อมูลที่เป็นรายการในระดับปฏิบัติการ ซึ่งข้อมูลจะเกิดขึ้นซ้ำๆ ทุกวัน (Transaction) โดยเน้นรายละเอียดเป็นหลัก เมื่อผู้บริหารต้องการดูรายงานสรุปเพื่อใช้ในการวางแผนงานในอนาคตอาจทำได้ยาก ดังนั้นคลังข้อมูลจึงนำข้อมูลระดับปฏิบัติงาน (OLTP) มาปรับเปลี่ยนใหม่ให้อยู่ในรูปแบบที่เหมาะสมและสะดวกต่อการนำไปใช้งาน โดยเครื่องมือที่สามารถนำเสนอข้อมูล และวิเคราะห์ข้อมูลดังกล่าว คือ Online Analytical processing (OLAP) ซึ่งเป็นเทคโนโลยีที่ช่วยนำเสนอข้อมูลในหลายมิติ

2.3. การทบทวนวรรณกรรม

Azwa Abdul Aziz [7] นำเสนอสถาปัตยกรรม Education Data Warehouse ที่มีการนำเครื่องมือที่เป็นโอเพนซอร์ส และเครื่องมือที่มีลิขสิทธิ์ มาบูรณาการใช้ร่วมกัน โดยการศึกษาเปรียบเทียบข้อดีข้อเสียของเครื่องมือและนำมาใช้ในแต่ละกระบวนการ ผลที่ได้จากงานวิจัยนี้คือต้นแบบของระบบ ซึ่งสามารถนำมาช่วยสนับสนุนการทำวิจัยในครั้งนี้ได้

Masaaki Ida [8] ได้ทำการพัฒนาโครงสร้างข้อมูลของมหาวิทยาลัย ทำการศึกษาโครงสร้างข้อมูลของมหาวิทยาลัยในญี่ปุ่น โดยวิธีการสำรวจข้อมูลและสร้างระบบธุรกิจอัจฉริยะที่ใช้เทคโนโลยี XBRL Extension ในการสร้างโครงสร้างของข้อมูลรวมถึงการเชื่อมโยงข้อมูลของสถาบันการศึกษา เพื่อให้สามารถวิเคราะห์และได้ความรู้จากโครงสร้างข้อมูลการศึกษา

Manjunath T N [9] ได้ทำการตรวจสอบระบบการจัดส่งข้อมูลสารสนเทศในปัจจุบัน นำเสนอระบบที่ดีและถูกต้องกว่าระบบเดิม อีกทั้งยังสามารถนำมาใช้ในการช่วยในการตัดสินใจขององค์กรการศึกษา โดยทำ การสร้างแบบจำลองข้อมูล แบ่งออกเป็น 3 data mart คือ Employee Mart, Student Mart และ Asset Mart โดยวิธีการรวบรวมข้อมูลใช้วิธี flat file ซึ่งสามารถโหลดข้อมูลเข้าไปยังคลังข้อมูลและคิวรี (Query) ข้อมูลตามที่ต้องการได้ ซึ่งผลของงานวิจัยนี้มุ่งเน้นไปที่คลังข้อมูลสามารถตอบสนองความต้องการของผู้ใช้ และแนะนำให้ใช้แนวคิด OLAP มาช่วยในการเพิ่มประสิทธิภาพในอนาคต

3. วิธีการดำเนินการวิจัย

3.1. การศึกษาและเก็บรวบรวมข้อมูล

งานวิจัยนี้ได้ทำการศึกษาและเก็บรวบรวมข้อมูลความต้องการของผู้บริหารสถาบันการศึกษาระดับอุดมศึกษาในประเทศไทย โดยวิธีการสัมภาษณ์ ศึกษาแนวคิดและกระบวนการการทำระบบธุรกิจอัจฉริยะ การสร้างคลังข้อมูลและขั้นตอนการพัฒนาระบบด้วยเทคนิคการประมวลผลข้อมูล

เชิงวิเคราะห์ (OLAP) การสร้างรายงาน ศึกษาฐานข้อมูลที่นำมาใช้ และชุดเครื่องมือโอเพนซอร์สที่นำมาใช้ในการพัฒนา

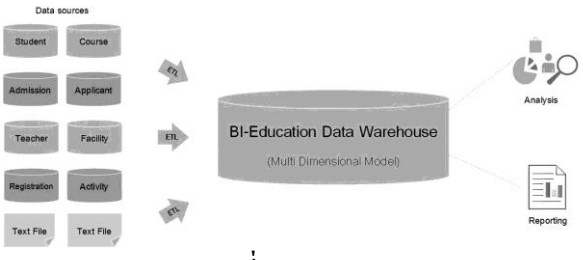
3.2. การวิเคราะห์และการออกแบบระบบ

นำความต้องการที่รวบรวมได้มาทำการวิเคราะห์และสรุปผล จากนั้นทำการออกแบบโครงสร้างของระบบ (BI-Education System Architecture) ออกแบบรายงานแดชบอร์ดและคลังข้อมูลให้ตรงกับความต้องการของผู้บริหาร

ลำดับ	ข้อมูลที่ผู้บริหารต้องการ	มิติมุมมอง (Dimension)
1	ผลงานของอาจารย์	ประเภทผลงาน, คณะ, ภาควิชา, หลักสูตร
2	จำนวนงานวิจัย	สาขางานวิจัย, ระดับของงานวิจัย, ประเภทงานวิจัย, ภาควิชา, นักวิจัย
3	จำนวนหลักสูตรที่เปิดสอน	อาจารย์ประจำหลักสูตร, ประเภทหลักสูตร, ระดับการศึกษา
4	จำนวนผู้เรียน	คณะ, ภาควิชา,หลักสูตร, ระดับการศึกษา, ประเภทหลักสูตร, เพศ, จังหวัดเกิดของ
5	อัตราคงอยู่ของผู้เรียน	ปีการศึกษา
6	จำนวนการลงทะเบียนเรียน	ปีการศึกษา, คณะ, ภาควิชา, หลักสูตร
7	จำนวนการเปิดสอนของอาจารย์	วิชา
8	ผลการเรียน	คณะ, ภาควิชา, หลักสูตร
9	จำนวนการรับเข้า	ปีการศึกษา, ประเภทการรับ
10	จำนวนผู้สำเร็จการศึกษา	คณะ, ภาควิชา, หลักสูตร, ระดับการศึกษา, ประเภทหลักสูตร, เพศ, ปีที่สำเร็จการศึกษา
11	จำนวนบุคลากร	สายงาน, วุฒิการศึกษา, ตำแหน่งทางวิชาการ, สถานะการจ้างงาน
12	จำนวนงบประมาณที่ได้รับ	แหล่งที่มาของงบประมาณ
13	การจัดสรรงบประมาณ	ปีงบประมาณ, หน่วยงานที่ได้รับการจัดสรร
14	การใช้งบประมาณ	ปีงบประมาณ

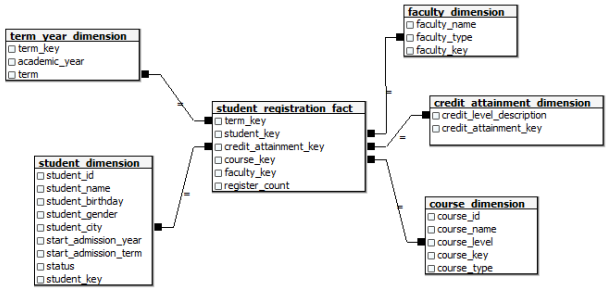
ภาพที่ 1: สรุปความต้องการของผู้บริหาร

จากภาพที่ 1 แสดงรายการข้อมูลที่ผู้บริหารต้องการรวมถึงมิติมุมมองข้อมูลของผู้บริหาร ในการใช้ข้อมูลเพื่อประกอบการตัดสินใจ โดยได้นำความต้องการของผู้บริหารมาวิเคราะห์สรุปและจัดลำดับ



ภาพที่ 2: โครงสร้างระบบ
(BI-Education System Architecture)

จากภาพที่ 2 แสดงโครงสร้างของระบบ มีการนำข้อมูลจากหลายแหล่งมาเข้าสู่กระบวนการสกัด เปลี่ยนแปลง ถ่ายโอนข้อมูลไปยังคลังข้อมูล (ETL) และทำการวิเคราะห์ข้อมูลด้วยเทคนิคการประมวลผลข้อมูลเชิงวิเคราะห์ (OLAP) เพื่อนำเสนอข้อมูลในหลายมุมมอง และแสดงรายงานสรุปแก่ผู้บริหาร

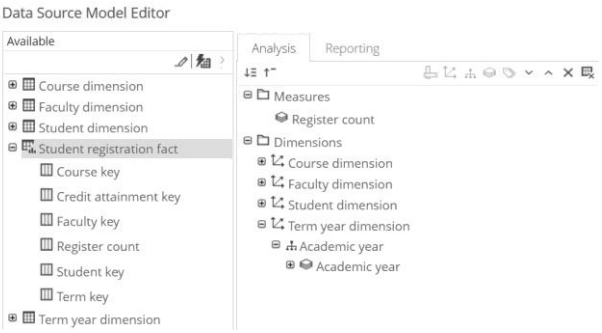


ภาพที่ 3: ตัวอย่างการออกแบบตารางข้อมูลในส่วนการลงทะเบียน

จากภาพที่ 3 แสดงโครงสร้างการออกแบบตารางข้อมูลในคลังข้อมูล โดยลักษณะของฐานข้อมูลแบบหลายมิติ เพื่อรองรับการประมวลผลเชิงวิเคราะห์ (OLAP) ทำให้สามารถแสดงข้อมูลในหลายมุมมอง และรองรับการออกรายงานในมิติต่างๆ

3.3. การพัฒนาระบบ

ทำการติดตั้งโปรแกรมและซอฟต์แวร์ เพื่อเตรียมความพร้อมในการพัฒนาระบบ โดยในส่วนของโปรแกรมจัดการฐานข้อมูลใช้ PostgreSQL และในส่วนของชุดเครื่องมือโอเพนซอร์สใน การสร้างระบบธุรกิจอัจฉริยะใช้ซอฟต์แวร์ Pentaho Community Edition ในส่วนของเครื่องมือในการรวบรวมข้อมูล (Extract, transform and load: ETL) ใช้ Pentaho Data Integration (Kettle) เครื่องมือในการสร้างรายงานแบบ Ad-hoc และ Cube ตัวอย่างดังภาพที่ 4 ใช้ Pentaho BI Server โดยระบบจะนำเสนอข้อมูลผ่านหน้าเว็บไซต์ ใช้การประมวลผลข้อมูลเชิงวิเคราะห์ (OLAP) ช่วยในการวิเคราะห์ข้อมูลแบบหลายมิติ



ภาพที่ 4: ตัวอย่างแสดงหน้าจอการสร้าง Cube

3.4. การทดสอบระบบ

การทดสอบเริ่มจากทดสอบฟังก์ชันการทำงานตามโมดูล (Unit Test) และทดสอบระบบโดยรวม (Integration Test) จากนั้นดำเนินการทดสอบการยอมรับระบบของผู้ใช้ ในขั้นเบต้า (Beta Testing) โดยผู้ใช้งานที่เป็นผู้บริหารจำนวน 10 ท่าน เพื่อประเมินความพึงพอใจในการใช้งานระบบ การทดสอบใช้แบบสอบถามปลายปิด (Closed-ended Form) แบบจัดอันดับความสำคัญ (Rank Order)

3.5. การประเมินผล

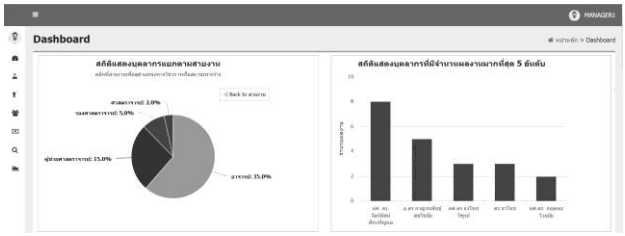
เครื่องมือที่ใช้ในการประเมินผลประสิทธิภาพของระบบ คือ ผลสรุปจากการประเมินความพึงพอใจของผู้ใช้งาน ซึ่งกำหนดเกณฑ์ในการประเมินไว้ 5 ระดับ ดังภาพที่ 5 ซึ่งทำการวิเคราะห์ความพึงพอใจโดยใช้ข้อมูลทางสถิติ ได้แก่ ค่าเฉลี่ยเลขคณิต (Arithmetic mean) และส่วนเบี่ยงเบนมาตรฐาน (Standard deviation: SD)

เกณฑ์	ช่วงคะแนน
ดีมาก	$4.5 \leq X \leq 5.0$
ดี	$4.0 \leq X < 4.5$
ปานกลาง	$3.0 \leq X < 4.0$
พอใช้	$2.0 \leq X < 3.0$
ควรปรับปรุง	$X < 2.0$

ภาพที่ 5: เกณฑ์ในการประเมิน

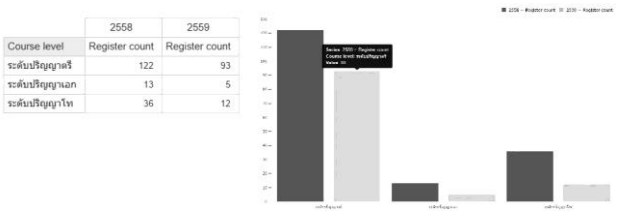
4. ผลการดำเนินงาน

ผลการพัฒนาระบบ BI สำหรับผู้บริหารสถาบันการศึกษา สามารถแสดงข้อมูลจากคลังข้อมูลที่ออกแบบไว้ได้อย่างเหมาะสม สอดคล้องกับความต้องการของผู้ใช้งานโดยอยู่ในรูปแบบของแดชบอร์ด และรายงาน ซึ่งผู้ใช้งานสามารถปรับเปลี่ยนความละเอียดของการพิจารณาข้อมูล จากข้อมูลภาพรวมจนเป็นข้อมูลอย่างละเอียด (Drill Down) และจากส่วนที่เป็นรายละเอียดจนมาเป็นข้อสรุป (Roll up)



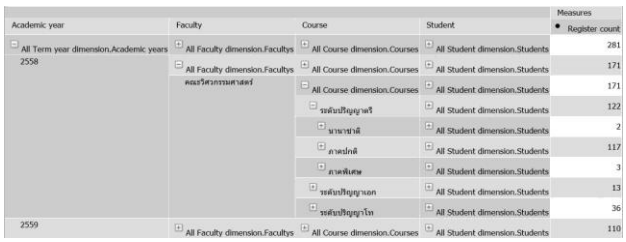
ภาพที่ 6: ตัวอย่างหน้าจอ Dashboard

จากภาพที่ 6 เป็นการแสดงตัวอย่างหน้าจอในส่วนของแดชบอร์ดสำหรับผู้บริหาร (BI-Education Dashboard Template) โดยระบบจะรวบรวมรายงานในรูปแบบของกราฟที่ผู้บริหารจำเป็นต้องใช้ในการติดตามเพื่อสนับสนุนการตัดสินใจ



Course level	2558 Register count	2559 Register count
ระดับปริญญาตรี	122	93
ระดับปริญญาเอก	13	5
ระดับปริญญาโท	36	12

ภาพที่ 7: ตัวอย่างหน้าจอแสดงรายงาน 1 มิติ



Academic year	Faculty	Course	Student	Measures: Register count
All Term year dimension.Academic years	All Faculty dimension.Facultys	All Course dimension.Courses	All Student dimension.Students	281
2558	All Faculty dimension.Facultys	All Course dimension.Courses	All Student dimension.Students	171
	คณะวิศวกรรมศาสตร์	All Course dimension.Courses	All Student dimension.Students	171
		ระดับปริญญาตรี	All Student dimension.Students	122
		นานาชาติ	All Student dimension.Students	2
		ภาคปกติ	All Student dimension.Students	117
		ภาคพิเศษ	All Student dimension.Students	3
		ระดับปริญญาเอก	All Student dimension.Students	13
		ระดับปริญญาโท	All Student dimension.Students	36
2559	All Faculty dimension.Facultys	All Course dimension.Courses	All Student dimension.Students	110

ภาพที่ 8: ตัวอย่างหน้าจอแสดงรายงานหลายมิติ

จากภาพที่ 7 และ 8 แสดงรายงานแบบ 1 มิติและแบบหลายมิติ เป็นรายงานที่ใช้ในการวิเคราะห์ข้อมูลในด้านต่างๆ โดยผู้ใช้งานสามารถปรับเปลี่ยนมุมมองของข้อมูลที่ต้องการวิเคราะห์ได้ และใช้งานผ่านเว็บบราวเซอร์

ลักษณะการใช้งานระบบในด้านต่างๆ	ค่าเฉลี่ย	ระดับความพึงพอใจ
1. การประมวลผล		
มีความถูกต้องในการประมวลผลข้อมูล	4.50	มากที่สุด
ตอบสนองการดูรายงานได้อย่างรวดเร็ว	4.47	มาก
2. การแสดงผล		
มีรายงานครบถ้วน	4.51	มากที่สุด
รายงานมีความถูกต้อง	4.26	มาก
การจัดวางรูปแบบได้เหมาะสมกับการใช้งาน	4.47	มาก
3. การเข้าใช้งานและความปลอดภัย		
ความสะดวกในการเข้าใช้งาน	4.51	มากที่สุด
ความปลอดภัยของข้อมูลภายในระบบ	4.53	มากที่สุด
4. การประยุกต์ใช้		
สามารถประยุกต์ใช้กับงานจริงได้	4.41	มาก
ความพึงพอใจโดยรวม	**4.46**	**มาก**

ภาพที่ 9: ผลการประเมินความพึงพอใจ

จากภาพที่ 9 แสดงผลการประเมินความพึงพอใจในแต่ละด้านและความพึงพอใจโดยรวม ซึ่งคำนวนจากการนำผลรวมของค่าคะแนนจากผลการตอบแบบสอบถามของผู้ใช้งานจำนวน 10 ท่านมาหาค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐาน

5. สรุปผล

งานวิจัยนี้ได้ทำการพัฒนาระบบ BI สำหรับผู้บริหารสถาบันการศึกษา โดยใช้ Pentaho และ PostgreSQL (Open source BI tool) ระบบที่พัฒนาจะแสดงผลในรูปแบบของเว็บแอพพลิเคชั่น ระบบสามารถสร้างรายงานและแดชบอร์ดรวมถึงการเรียกดูข้อมูลรายงานได้ในหลายมุมมอง สามารถดูข้อมูลแบบเจาะลึกลงรายละเอียดได้ (Drill Down) หรือเรียกดูข้อมูลภาพรวมได้ (Roll up) ซึ่งช่วยให้ผู้บริหารสามารถใช้ในการสนับสนุนการตัดสินใจและวางแผนกลยุทธ์ขององค์กรได้อย่างมีประสิทธิภาพทันต่อเหตุการณ์ สำหรับผลการประเมินคุณภาพของระบบจากผู้ใช้งาน 10 ท่าน พบว่าค่าเฉลี่ยโดยรวมทุกด้านเท่ากับ 4.46 และค่าเบี่ยงเบนมาตรฐานเท่ากับ 2.19 ดังนั้นสามารถสรุปผลได้ว่าผลจากการประเมินอยู่ในระดับดี

เอกสารอ้างอิง

[1] ธนาวินท์ รักธรรมานนท์, กฤษณะ ไวยมัย และ ชิดชนก ส่งศิริ, “การใช้เทคนิค data mining เพื่อค้นหาภาควิชาที่เหมาะสมที่สุดให้กับนิสิต,” มหาวิทยาลัยเกษตรศาสตร์, 2555.

[2] Diana, T. Â. R. N. "Pentaho Business Analytics: a Business Intelligence Open Source Alternative." *Database Systems Journal* 3.3 (2012): 23-34.

[3] Gounder, Mohan S., Vani Vasudevan Iyer, and Abdulaziz Al Mazyad. "A survey on business intelligence tools for university dashboard development." *Big Data and Smart City (ICBDSC), 2016 3rd MEC International Conference on.* IEEE, 2016.

[4] Stackowiak, Robert, Joseph Rayman and Rick Greenwald. *Oracle Data Warehousing and Business Intelligence Solutions.* Wiley Publishing, Inc., 2007.

[5] Cui, Zhan, Ernesto Damiani, and Marcello Leida. "Benefits of ontologies in real time data access." *Digital EcoSystems and Technologies Conference, 2007. DEST'07. Inaugural IEEE-IES.* IEEE, 2007.

[6] Liao, Jian-Ming, and Xi-Qiang Zeng. "A framework of call center service management system based on Pentaho." *Wavelet Active Media Technology and Information Processing (ICWAMTIP), 2012 International Conference on.* IEEE, 2012.

[7] Aziz, Azwa Abdul, et al. "A Framework for Educational Data Warehouse (EDW) Architecture Using Business Intelligence (BI) Technologies." *Journal of Theoretical & Applied Information Technology* 69.1 (2014).

[8] Ida, Masaaki. "Structure of university database system and data analysis." *Advanced Communication Technology (ICACT), 2014 16th International Conference on.* IEEE, 2014.

[9] Manjunath, T. N., et al. "Design and Analysis of DWH and BI in Education Domain." *IJCSI* (2011): 544.

[10] Kimball, Ralph, and Margy Ross. *The Data Warehouse Toolkit Third Edition.* John Wiley & Sons, 2013.

[11] วารุณี แต้มคู และกฤษณะ ไวยมัย, “ชุดเครื่องมือโอเพนซอร์ระบบธุรกิจอัจฉริยะสำหรับธุรกิจการศึกษา,” งานประชุมหาดใหญ่วิชาการระดับชาติ และนานาชาติ ครั้งที่ 8, 2560.

การแนะนำการเกษียนหนังสือราชการด้วยวิธีการทำเหมืองข้อความ
Introducing the Document Flow of the Official Document by Text Mining

กันตพิชญ์ อนันตศักดิ์ (Guntapich Anantasak)[1] และณัฐวี อุตกฤษฎ์ (Nattavee Utakrit)[2]
ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
[1]s5807021857024@email.kmutnb.ac.th, [2]nattavee.u@it.kmutnb.ac.th

บทคัดย่อ

การเกษียนหนังสือ เป็นหัวใจสำคัญของการรับ-ส่งหนังสือในหน่วยงานราชการ การนำข้อมูลการเกษียนหนังสือราชการที่มีอยู่มาใช้ให้เกิดประโยชน์ จึงเป็นแนวทางในการช่วยแนะนำการเกษียนหนังสือราชการเพื่อช่วยในการตัดสินใจของบุคลากรในหน่วยงานได้ดียิ่งขึ้น ข้อมูลหนังสือราชการส่วนใหญ่จะอยู่ในรูปแบบข้อความ ดังนั้นผู้วิจัยเลือกใช้วิธีการจัดกลุ่มด้วยการทำเหมืองข้อความ โดยงานวิจัยนี้ได้ทำการเปรียบเทียบประสิทธิภาพอัลกอริทึม อันได้แก่ซัพพอร์ตเวกเตอร์แมชชีน และนาอีฟเบย์เพื่อหาวิธีการจัดกลุ่มที่ดีที่สุดที่จะสามารถนำไปประยุกต์ใช้ในการแนะนำการเกษียนหนังสือราชการได้อย่างมีประสิทธิภาพ ผลการศึกษาพบว่าการจัดกลุ่มข้อมูลหนังสือราชการด้วยอัลกอริทึมซัพพอร์ตเวกเตอร์แมชชีน มีค่าความแม่นยำ (Precision) ค่าความระลึก (Recall) ค่าความถูกต้อง (Accuracy) ดีกว่านาอีฟเบย์

คำสำคัญ: หนังสือราชการ เหมืองข้อความ นาอีฟเบย์ ซัพพอร์ตเวกเตอร์แมชชีน

Abstract

Document flow is the key to receiving and sending documents in government agencies. The use of existing government information on document flow is a good way to guide how the flow of government documents helps the staff make better decisions. This research is a presentation of how to introduce an official document flow. Because the documents are normally in text format, so text mining is applied to group the documents. Support Vector Machine and Naïve Bayes algorithms are used to compare the best result in grouping. It can also be used in the effective introduction of document official flow. From the research results, it was found that grouping with algorithms Support Vector Machine to Precision Recall and Accuracy yielded better results, compared to Naïve Bayes.

Keyword: Official Document, Text Mining, Naïve Bayes, Support Vector Machines

1. บทนำ

ปัจจุบันหน่วยงานราชการ มีการนำระบบงานสารบรรณอิเล็กทรอนิกส์ มาใช้ภายในองค์กร เพื่อรับส่งหนังสือราชการไปยังแต่ละกลุ่ม/ฝ่ายภายในหน่วยงาน เช่น หนังสือภายในหนังสือภายนอก หนังสือตราประทับ หนังสือสั่งการหรือหนังสือประชาสัมพันธ์ เป็นต้น [1] ซึ่งระบบดังกล่าว ช่วยลดการใช้ทรัพยากรกระดาษ และร่นระยะเวลาในการนำเอกสารไปยังแต่ละกลุ่ม/ฝ่ายภายในหน่วยงานได้ดียิ่งขึ้น

ทั้งนี้การส่งหนังสือไปยังกลุ่มฝ่าย โดยเฉพาะหนังสือราชการด้านเทคโนโลยีสารสนเทศ บุคลากรที่รับผิดชอบด้านงานธุรการยังมีความรู้ความเข้าใจในด้านเทคโนโลยีสารสนเทศค่อนข้างน้อย จึงอาจเกิดความผิดพลาดในการส่งหนังสือไปยังกลุ่มฝ่ายต่าง ๆ ผู้วิจัยได้ตระหนักถึงประเด็นดังกล่าว จึงจัดทำการแนะนำการเกษียนหนังสือราชการด้วยวิธีการทำเหมืองข้อความขึ้นเพื่อเป็นเครื่องมือช่วยในการตัดสินใจของบุคลากรในการเกษียนหนังสือราชการด้านเทคโนโลยีสารสนเทศ

2. วรรณกรรมที่เกี่ยวข้อง

2.1 หนังสือราชการ

หนังสือราชการ [1] คือเอกสารที่เป็นหลักฐานในราชการ ได้แก่ หนังสือที่มีไปมาระหว่างส่วนราชการ หนังสือที่มีส่วนราชการมีไปถึงหน่วยงานอื่นใดซึ่งมิใช่ส่วนราชการหรือที่มีไปถึงบุคคลภายนอก หนังสือที่หน่วยงานอื่นใดที่ไม่ใช่ส่วนราชการ หรือบุคคลภายนอกมีมาถึงส่วนราชการ เอกสารที่ทางราชการจัดทำขึ้นเพื่อเป็นหลักฐานในราชการ เอกสารที่ทางราชการจัดทำขึ้นตามกฎหมาย ระเบียบ ข้อบังคับ ข้อมูลข่าวสารหรือหนังสือที่ได้รับเข้าจากระบบสารบรรณอิเล็กทรอนิกส์

ชนิดของหนังสือราชการมีทั้งหมด 6 ชนิดได้แก่ หนังสือภายนอก หนังสือภายใน หนังสือประทับตรา หนังสือสั่งการ หนังสือประชาสัมพันธ์ และหนังสือที่เจ้าหน้าที่ทำขึ้นเพื่อรับไว้เป็นหลักฐานในราชการ

2.2 การทำเหมืองข้อความ (Text Mining)

การทำเหมืองข้อความคือ การสืบค้นและการสกัดความรู้จากข้อความที่ไม่ชัดเจนหรือข้อความที่ไม่มีโครงสร้าง ซึ่งครอบคลุมถึงการนำข้อมูล เช่น เอกสารหรือเว็บไซต์ มาจำแนกข้อความและการจัดกลุ่มเพื่อค้นหารูปแบบ แนวทาง และความสัมพันธ์ที่ซ่อนอยู่ในชุดข้อความนั้น โดยอาศัยหลักสถิติ การรู้จำ การเรียนรู้ของเครื่อง หลักคณิตศาสตร์ หลักการประมวลเอกสาร (Document Processing) หลักการประมวลผลข้อความ (Text Processing) และหลักการประมวลผลภาษาธรรมชาติ (Natural Language Processing) [2] โดยเป้าหมายของการทำเหมืองข้อความแบ่งได้ 3 ประเภท คือ 1.การสรุปเอกสารข้อความ เป็นการลดความซับซ้อนและขนาดของเอกสารข้อความ โดยไม่ทำให้ความหมายหรือสาระสำคัญของข้อมูลเอกสารเสียไป 2.การแบ่งประเภทเอกสารข้อความ (Document Classification) เป็นวิธีการที่ใช้ช่วยในจำแนกประเภทเอกสาร โดยจำเป็นต้องทำการสอนระบบ (Train Model) ให้รู้จำรูปแบบของเอกสารในแต่ละ class ก่อน จากนั้นจึงจำแนกเอกสารตามรูปแบบที่ได้สอนระบบไว้ 3.การแบ่งกลุ่มเอกสารข้อความ (Document Clustering) เป็นการจัดแบ่งเอกสารข้อความออกเป็นกลุ่ม โดยใช้การวัดความคล้ายคลึงและความแตกต่างของคุณลักษณะของเอกสารข้อความ

2.3 ซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machines)

ซัพพอร์ตเวกเตอร์แมชชีน [3] เป็นหนึ่งในอัลกอริทึมการเรียนรู้แบบมีผู้สอน ใช้เพื่อการแบ่งประเภทข้อมูลและการวิเคราะห์การถอดถอย เมื่อมีข้อมูลฝึกมาให้และแต่ละข้อมูลถูกจัดอยู่ในประเภทใดประเภทหนึ่งจากสองประเภท ซัพพอร์ตเวกเตอร์แมชชีนจะสร้างแบบจำลองที่สามารถพยากรณ์ได้ว่าตัวอย่างใหม่นี้จะตกอยู่ในกลุ่มใด

ซัพพอร์ตเวกเตอร์แมชชีน เป็นอัลกอริทึมในการคัดแยกด้านการประมวลผลเป็นภาพ หลักการของ ซัพพอร์ตเวกเตอร์แมชชีนนี้ใช้เพื่อหาระนาบการตัดสินใจในการแบ่งข้อมูลออกเป็นสองส่วน โดยพิจารณาจากจุดของข้อมูลที่อยู่ในกลุ่มที่แตกต่าง จุดเหล่านี้เรียกว่า ซัพพอร์ตเวกเตอร์ ซึ่งจะนำมาหาสมการเส้นแบ่งกลุ่มที่เหมาะสมต่อไป [4]

การเปลี่ยนแปลงมิติของข้อมูลที่เรียกว่า ฟีเจอร์สเปช (Feature Space) ซัพพอร์ตเวกเตอร์แมชชีนจะใช้ฟังก์ชันแม็ปสำหรับย้ายข้อมูลจาก Input Space ไปยัง Feature Space และสร้างฟังก์ชันวัดความคล้ายที่เรียกว่าเคอร์เนลฟังก์ชัน (Kernel Function) บน Feature Space เหมาะสำหรับข้อมูลที่มีมิติของข้อมูลสูง กำหนดให้ $(x_i,y_i),\ldots,(x_n,y_n)$ เป็นตัวอย่างที่ใช้สำหรับการสอน n คือจำนวนข้อมูลตัวอย่าง m คือจำนวนมิติข้อมูลเข้า และ y คือผลลัพธ์ที่มีค่า +1 หรือ -1 ดัง สมการที่ 1

$$(x_i,y_i),\ldots,(x_n,y_n) \text{ เมื่อ } x \in R^m, y \in \{+1,-1\} \quad (1)$$

สำหรับปัญหาเชิงเส้น มิติข้อมูลขนาดสูงได้ถูกแบ่งเป็น 2 กลุ่ม โดยระนาบตัดสินใจ ซึ่งคำนวณได้ดัง สมการที่ 2

$$(w*x)+b=0 \quad (2)$$

เมื่อ w คือ ค่าน้ำหนักและ b คือค่า bias สมการใช้สำหรับจำแนกประเภทของข้อมูลดัง สมการที่ 3 และสมการที่ 4

$$(w*x)+b > 0 \text{ ถ้า } y_i = +1 \quad (3)$$

$$\text{และ } (w*x)+b < 0 \text{ ถ้า } y_i = -1 \quad (4)$$

ซัพพอร์ตเวกเตอร์แมชชีน มีเคอร์เนลฟังก์ชันที่สามารถประยุกต์ใช้งานได้หลายวิธี [5] สำหรับงานวิจัยนี้ได้เลือกใช้วิธี Linear kernel เป็นฟังก์ชันในการทดลอง

2.4 นาอีฟเบย์ (Naïve Bayes)

นาอีฟเบย์ คือการคำนวณความน่าจะเป็น ซึ่งถูกใช้ในการทำนายผล และการแก้ปัญหาที่สามารถคาดการณ์ผลลัพธ์ได้และ

สามารถอธิบายได้ โดยจะทำการวิเคราะห์ความสัมพันธ์ระหว่างตัวแปรเพื่อใช้ในการสร้างเงื่อนไข

กำหนดให้ P(A) คือความน่าจะเป็นที่เหตุการณ์ A เกิดขึ้น และ P(A|B) คือความน่าจะเป็นที่จะเกิดเหตุการณ์ B ขึ้นก่อนและจะมีเหตุการณ์ A ตามมา จากตัวแปรที่กำหนดและแนวคิดของ Bayes' Theorem นั้นสามารถทำนายเหตุการณ์ที่พิจารณาได้จากการเกิดของเหตุการณ์ต่าง ๆ ได้ดัง สมการที่ 5 [6]

$$P(A|B) = [P(B|A) * P(A)]/P(B) \quad (5)$$

นาอีฟเบย์ เป็นวิธีจำแนกประเภทข้อมูลที่มีประสิทธิภาพวิธีหนึ่งโดยที่ใช้งานจัดหมวดหมู่เอกสาร ข้อความ (Text Classification) ได้ดี เป็นวิธีการในการทำงานที่ไม่ซับซ้อน เหมาะกับกรณีของเซตตัวอย่างมีจำนวนมากและคุณสมบัติ (Attribute) ของตัวอย่างไม่ขึ้นต่อกัน โดยกำหนดให้ความน่าจะเป็นของข้อมูลที่จะเป็นดัง สมการที่ 6

$$P(a_1, a_2, \dots, a_n|v_j) = \prod_{i=1}^{n} P(a_i, v_j) \quad (6)$$

กลุ่ม v_j สำหรับข้อมูลที่มีคุณสมบัติ n ตัว X={ $a_1, a_2, \dots, a_n$} หรือ ใช้สัญลักษณ์ว่า $P(a_1, a_2, \dots, a_n|v_j)$ โดยที่ $\prod$ หมายถึงผลคูณของค่า $P(a_i|v_j)$ ทั้งหมด i = 1,2,3,…,n และ j = 1,2,3,…,n ดังนั้น เราจะได้ว่าวิธีการจำแนกประเภทนาอีฟเบย์ดังสมการที่ 7

$$V_{NB} = \underset{v_j \epsilon V}{\operatorname{argmax}} P(v_j) = x \prod_{i=1}^{n} P(a_i|v_j) \quad (7)$$

งานวิจัยนี้ได้เลือกนาอีฟเบย์ (Normal Naïve Bayes) แบบปกติ

2.5 งานวิจัยที่เกี่ยวข้อง

นิเวศ [7] ได้นำเสนอการจัดหมวดหมู่เอกสารภาษาไทยจากหนังสือพิมพ์แบบอัตโนมัติด้วยซัพพอร์ตเวกเตอร์แมชชีน โดยนำวิธีการทำเหมืองข้อความมาช่วยในการตัดคำ เพื่อสกัดข้อความและนำเข้าสู่วิธีการจำลองการจัดหมวดหมู่ตามแนวคิดทางด้านการค้นคืนสารสนเทศ โดยการวัดประสิทธิภาพจัดหมวดหมู่เอกสารภาษาไทยด้วยการวัดค่าความแม่นยำ ค่าความระลึก และค่า F-measure

ศุภเทพ [8] ได้ศึกษาการวัดประสิทธิภาพของวิธีการในการจำแนกประเภทข้อความสั้นบนเว็บไซต์ทวิตเตอร์ โดยนำวิธีการทำเหมืองข้อความมาเข้าสู่วิธีการจำแนกประเภทสำหรับข้อความภาษาไทย ที่เป็นลักษณะข้อความสั้น

กานดา [9] ได้วิเคราะห์ความคิดเห็นบนเครือข่ายสังคมออนไลน์ ได้นำวิธีการในการวิเคราะห์เหมืองข้อความมาประยุกต์ใช้เพื่อวิเคราะห์ความคิดเห็นของผู้บริโภค ซึ่งเรียกว่าการวิเคราะห์เหมืองความคิดเห็น หรือการวิเคราะห์ความรู้สึก โดยการทดสอบผลการวิเคราะห์ความคิดเห็นนั้น ใช้ชุดข้อมูลทดสอบมาจำแนกขั้วความคิดเห็นด้วยแบบจำลองการวิเคราะห์ข้อมูล และให้ผู้เชี่ยวชาญตรวจสอบความถูกต้องอีกครั้ง

ราชวิทย์ [10] ได้นำเสนอการจำแนกกลุ่มคำถามอัตโนมัติบนกระดานสนทนา ซึ่งมีการนำวิธีการทำเหมืองข้อความมาช่วยในการจำแนกคำถามอัตโนมัติซึ่งอยู่ในรูปแบบเว็บแอพพลิเคชั่น ซึ่งได้ทำการศึกษาและเปรียบเทียบประสิทธิภาพการจำแนกข้อความด้วยอัลกอริทึม K-Nearest Neighbors, Decision Tree และ Naïve Bayes ซึ่งผลการจำแนกข้อความภาษาไทยนั้น Naïve Bayes ให้ประสิทธิภาพดีที่สุด มีความถูกต้องที่ ร้อยละ 89.01 โดยการทดสอบผลการจำแนกคำถามอัตโนมัตินั้นจะประเมินจากผลลัพธ์ค่าความถูกต้อง ค่าความแม่นยำ ค่าความระลึก และค่า F-measure

อดิเทพ [11] ได้วิเคราะห์การประมาณการอารมณ์จากความคิดเห็นภาษาไทยโดยใช้อัลกอริทึมซัพพอร์ตเวกเตอร์แมชชีน ซึ่งใช้วิธีการทำเหมืองข้อความในการเก็บรวบรวมข้อความคิดเห็นภาษาไทยมาเพื่อใช้สกัดข้อความ และทำการวิเคราะห์อารมณ์ของข้อความที่มีการแสดงความคิดเห็น โดยประเมินประสิทธิภาพการจำแนกอารมณ์ด้วย การวัดค่าความแม่นยำ ค่าความระลึก

สมศักดิ์ [12] ได้ทำการจัดกลุ่มลูกค้าสินเชื่อจากการปฏิเสธด้วยวิธีการทำเหมืองข้อความมาเพื่อใช้สกัดข้อความ และจัดกลุ่มของข้อความ ด้วยอัลกอริทึม K-Means และ Self-Organizing Maps เพื่อนำผลลัพธ์ที่ได้ไปใช้สนับสนุนการตัดสินใจของผู้ใช้งานต่อไป

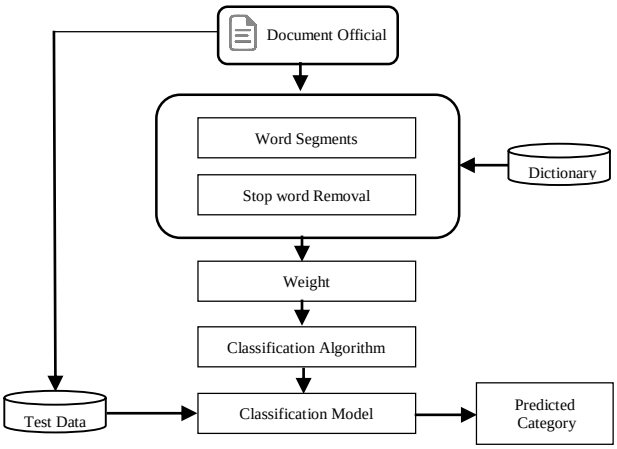
จากการศึกษางานวิจัยข้างต้น การจัดกลุ่มข้อความจะใช้วิธีการทำเหมืองข้อความ ผู้วิจัยจึงได้นำเสนอการแนะนำการเกษียนหนังสือราชการด้วยวิธีการทำเหมืองข้อความ เพื่อเป็นเครื่องมือช่วยตัดสินใจในการส่งหนังสือไปยังกลุ่ม/ฝ่ายให้ใกล้เคียงกับเนื้อหาที่ถูกสร้างขึ้นมากที่สุด และใช้การประเมินด้วยวิธีการวัดค่าความแม่นยำ และค่าความระลึก

ในการวิจัยนี้ ข้อมูลของหนังสือราชการ ได้มีข้อมูลที่ได้ทำการแบ่งกลุ่มไว้อยู่ก่อนแล้ว เพื่อให้การพัฒนาโมเดลสามารถเรียนรู้ได้ง่ายและตรงตามวัตถุประสงค์จึงต้องใช้การเรียนรู้ด้วย Text Classification

จากงานวิจัยที่ศึกษามานั้น งานวิจัยที่มีการเรียนรู้จากการแบ่งกลุ่มของเนื้อหาอยู่ก่อนแล้วจะใช้อัลกอริทึมซัพพอร์ตเวกเตอร์แมชชีน และนาอีฟเบย์ ผู้วิจัยจึงนำอัลกอริทึมดังกล่าวมาเปรียบเทียบเพื่อให้ได้ผลลัพธ์ที่ดีที่สุดเพื่อนำไปประยุกต์ใช้ในการแนะนำการเกษียนหนังสือราชการต่อไป

3. วิธีการดำเนินการวิจัย

งานวิจัยนี้มีวัตถุประสงค์ เพื่อช่วยตัดสินใจในการแนะนำการเกษียนหนังสือราชการ ด้วยวิธีการทำเหมืองข้อความ โดยใช้อัลกอริทึมซัพพอร์ตเวกเตอร์แมชชีน และนาอีฟเบย์ ทดสอบกับรายชื่อหนังสือราชการโดยมีกลุ่มตัวอย่างทั้งหมด 2,000 เอกสาร ดังภาพที่ 1



ภาพที่ 1: แบบจำลองการแนะนำการเกษียนหนังสือราชการ

3.1 การจัดเตรียมข้อมูล

จากกลุ่มตัวอย่างของหนังสือราชการ มีการแบ่งกลุ่มอยู่ก่อนแล้ว เพื่อความถูกต้องในการจัดเตรียมข้อมูล ผู้วิจัยได้อาศัยผู้เชี่ยวชาญเพื่อตรวจสอบความถูกต้องในการกำหนดกลุ่มของหนังสือราชการที่ส่งไปยังกลุ่มฝ่ายต่าง ๆ ภายในหน่วยงาน ดังตารางที่ 1

ตารางที่ 1: การตรวจสอบความถูกต้องของกลุ่มฝ่ายต่าง ๆ

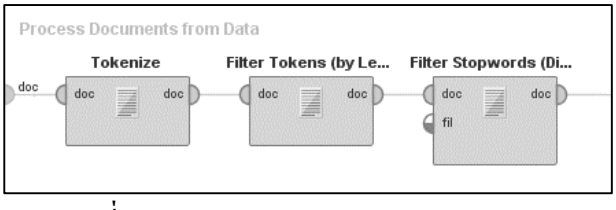
Detail	Class
การเตรียมการจัดทำคำของบประมาณในส่วนการจัดหาระบบคอมพิวเตอร์	Policy
ขอเปลี่ยนทดแทน Monitor	Computer
การบูรณาการฐานข้อมูลประชาชนและการบริการภาครัฐ	Database
ขอรหัสผ่านสำหรับการเข้าใช้งานระบบอินเทอร์เน็ต	Security
ขอรหัสแถบแม่เหล็กของระบบสารบรรณอิเล็กทรอนิกส์	Programmer
ขอความอนุเคราะห์ลงประกาศบนเว็บไซต์	Service

3.2 การแบ่งคำภาษาไทย

ผู้วิจัยได้อาศัยแบบจำลองที่นำเสนอทำการตัดคำภาษาไทยโดยวิธีการตัดคำที่ยาวที่สุด (Longest Matching) ด้วย THSplitLib ร่วมกับการใช้พจนานุกรม ของ Lexitron ดังภาพที่ 2 และทำการกำจัดคำหยุดจากพจนานุกรม ที่กำหนดขึ้นและกำหนดคำที่มีความยาวตั้งแต่ 4 ถึง 25 ตัวอักษร ซึ่งผู้วิจัยได้เพิ่มคำศัพท์ด้านเทคโนโลยีสารสนเทศเข้าไปด้วย โดยใช้โปรแกรม Rapidminer ช่วยในการสกัดคุณลักษณะข้อความของหนังสือราชการดังภาพที่ 3

700	ขอ|รหัสผ่าน|เข้า|ระบบ|อินเตอร์เน็ต|ผอ|ง.	Security
701	ขอ|บัญชี|ผู้ใช้งาน|และ|รหัสผ่าน|ระบบเครือข่าย|อินเต	Security
702	ขอ|รหัสผ่าน|เข้า|ระบบ|อินเตอร์เน็ต|กรมทางหลวง|ผส	Security
703	ขอ|รหัสผ่าน|เข้า|ระบบ|อินเตอร์เน็ต|ผส|บ.	Security
704	ขอ|ผู้ใช้งาน|และ|รหัสผ่าน|ของ|ข้าราชการ|เพื่อ|เข้า|ใ	Security
705	ขอ|ชื่อ|ผู้ใช้งาน|และ|รหัสผ่าน|ผ|สว	Security
706	ขอ|ชื่อ|ผู้ใช้งาน|และ|รหัสผ่าน|ผ|สว	Security
707	ขอ|รหัส|รหัสผ่าน|สำหรับ|เจ้าหน้าที่|เพื่อ|เข้า|ใช้งาน|ร	Security
708	ขอ|ส่ง|แบบ|คำขอ|มี|บัญชีรายชื่อ|ผู้ใช้งาน|และ|รหัสผ	Security
709	การ|ติดตาม|ผลการดำเนินงาน|ตาม|แผนปฏิบัติราชการ|	Policy
710	รายงาน|การ|จัดทำ|แผนแม่บท|เทคโนโลยีสารสนเทศ|	Policy
711	ขอให้|รายงาน|ผลการดำเนินการ|ตาม|แผนปฏิบัติราชก	Policy
712	รายงานการประชุม|คณะกรรมการ|เร่งรัด|ติดตาม|การ|ใ	Policy
713	ขอ|ส่ง|รายละเอียด|โครงการ|จ้าง|ที่ปรึกษา|เพื่อ|พิจาร	Policy
714	ขอ|ส่ง|รายงานผล|การ|ดำเนินโครงการ|ไอซีที|ประจำ|ไ	Policy
715	การรายงานผล|การ|ดำเนินโครงการ|ไอซีที|ประจำ|ปีงบ	Policy
716	การรายงานผล|การ|ดำเนินโครงการ|ไอซีที|ประจำ|ปีงบ	Policy

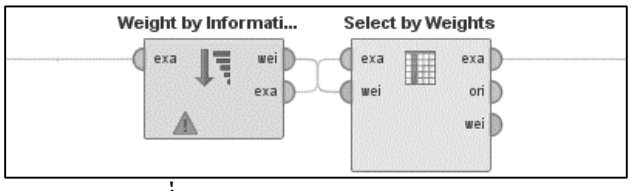
ภาพที่ 2: แสดงผลการตัดคำภาษาไทยด้วยวิธีการตัดคำที่ยาวที่สุด



ภาพที่ 3: การสกัดคุณลักษณะข้อความของหนังสือราชการ

3.3 การลดขนาดคุณลักษณะ

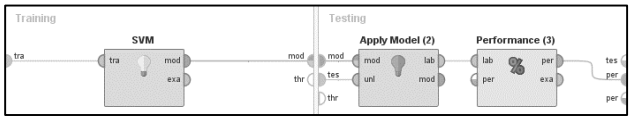
การลดขนาดเป็นการเลือกข้อมูลที่มีประโยชน์ต่อการจัดกลุ่มข้อความ โดยการลดขนาดมิติของเอกสารด้วย Information Gain และ Select Weights [6] โดยวิธีการนี้ช่วยลดคำที่ไม่มีความสำคัญและช่วยให้การประมวลผลมีความเร็วมากขึ้น ดังภาพที่ 4



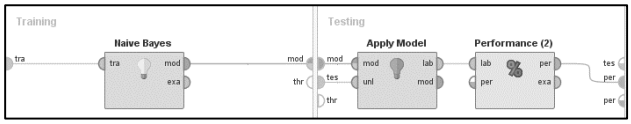
ภาพที่ 4: การลดขนาดมิติของหนังสือราชการ

3.4 การทดสอบวิธีการในการจัดกลุ่ม

การทดสอบวิธีการในการจัดกลุ่มข้อความ โดยนำข้อมูลที่ผ่านการแบ่งคำและการลดขนาดคุณลักษณะแล้ว มาเข้าสู่อัลกอริทึมซัพพอร์ตเวกเตอร์แมชชีน และนาอีฟเบย์ ดังภาพที่ 5 และ 6



ภาพที่ 5: การทดสอบด้วยซัพพอร์ตเวกเตอร์แมชชีน



ภาพที่ 6: การทดสอบด้วยนาอีฟเบย์

ทำการทดสอบเปรียบเทียบประสิทธิภาพด้านความถูกต้องด้วย ค่าความแม่นยำ (Precision) ค่าความระลึก (Recall) ค่าความถูกต้อง (Accuracy) ดัง สมการที่ 8,9 และ 10

$$Precision = \frac{TP}{TP+FP} \quad (8)$$

$$Recall = \frac{TP}{TP+FN} \quad (9)$$

$$Accuracy = \frac{TP+TN}{TP+FP+FN+TN} \quad (10)$$

3.5 แบบจำลองการจัดกลุ่ม

จากการทดสอบวิธีการที่ใช้ในการจัดกลุ่มกำหนดคุณลักษณะแล้ว จะได้แบบจำลองการแนะนำการเกษียนหนังสือราชการที่มีความเหมาะสมมากที่สุด

4. ผลการดำเนินงาน

จากการทดสอบวิธีการจัดกลุ่มหนังสือราชการด้านเทคโนโลยีสารสนเทศ โดยการเปรียบเทียบวิธีการทำเหมืองข้อความจำนวน 2 อัลกอริทึมคือ อัลกอริทึมซัพพอร์ตเวกเตอร์แมชชีนและนาอีฟเบย์ เพื่อวัดประสิทธิภาพความแม่นยำด้วย ค่าความแม่นยำ (Precision) ค่าความระลึก (Recall) ค่าความถูกต้อง (Accuracy) ดังตารางที่ 2, 3 และ 4

ตารางที่ 2: ผลการทดสอบค่าความแม่นยำ (Precision)

Class	Precision	
	Naïve Bayes	**SVM (Linear)**
Computer	86.08%	85.94%
Database	86.61%	97.00%
Security	68.22%	95.18%
Policy	87.72%	92.14%
Programmer	65.00%	95.00%
Service	93.87%	93.20%
Overall Precision	**81.94%**	**93.24%**

ตารางที่ 3: ผลการทดสอบค่าความระลึก (Recall)

Class	Recall	
	Naïve Bayes	**SVM (Linear)**
Computer	78.96%	88.18%
Database	100.00%	100.00%
Security	86.10%	84.49%
Policy	93.17%	90.99%
Programmer	74.29%	81.43%
Service	86.42%	95.83%
Overall Recall	**86.49%**	**90.16%**

ตารางที่ 4: ผลการทดสอบค่าความถูกต้อง (Accuracy)

	Naïve Bayes	**SVM (Linear)**
Accuracy	86.11%	91.75%

จากตารางที่ 2, 3 และ 4 พบว่า ซัพพอร์ตเวกเตอร์แมชชีนให้ค่าความแม่นยำ (Precision) เท่ากับร้อยละ 93.24 ค่าความระลึก (Recall) เท่ากับร้อยละ 90.16 ค่าความถูกต้อง (Accuracy) เท่ากับร้อยละ 91.75 มากกว่า นาอีฟเบย์ ที่ให้ค่าความแม่นยำ (Precision) เท่ากับร้อยละ 81.94 ค่าความระลึก (Recall) เท่ากับร้อยละ 86.49 ค่าความถูกต้อง (Accuracy) เท่ากับร้อยละ 86.11 ซึ่งตารางที่ 2 ค่าความแม่นยำของนาอีฟเบย์ ในกลุ่มของ Programmer และ Security มีค่าน้อยกว่าร้อยละ 70 นั้น เกิดจากความใกล้เคียงกันของเนื้อหาในกลุ่ม Service และ Computer ตามลำดับ

5. สรุปผลและอภิปรายผล

งานวิจัยนี้ได้นำเสนอการแนะนำการเกษียนหนังสือราชการโดยเปรียบเทียบและทดสอบประสิทธิภาพการจัดกลุ่มด้วยอัลกอริทึมซัพพอร์ตเวกเตอร์แมชชีน และอัลกอริทึมนาอีฟเบย์ เพื่อหาวิธีการจัดกลุ่มที่ดีที่สุด โดยผู้วิจัยได้รวบรวมข้อมูลหนังสือราชการ และจัดกลุ่มข้อมูลด้วยวิธีการ Classification โดยปรับค่าพาริเตอร์เคอร์เนลฟังก์ชั่นแบบต่าง ๆ ใช้วิธีการ

ลดคุณลักษณะเพื่อศึกษาประสิทธิภาพในการแนะนำการเกษียนหนังสือราชการ จากการทดลองพบว่า การลดคุณลักษณะด้วยวิธี Information Gain เพื่อคัดเลือกค่าน้ำหนักที่ได้ พร้อมวัดประสิทธิภาพจากค่าความถูกต้อง สามารถสรุปได้ว่าอัลกอริทึมซัพพอร์ตเวกเตอร์แมชชีน ให้ประสิทธิภาพการแนะนำการเกษียนหนังสือราชการโดยเฉลี่ยออกมาดีที่สุดคือร้อยละ 91.75 ขณะที่อัลกอริทึมนาอีฟเบย์ให้ค่าความถูกต้องคือร้อยละ 86.11 ซึ่งสอดคล้องกับงานวิจัยของ นิเวศ ปริญญา และพยุง [7] ที่เปรียบเทียบการจัดหมวดหมู่เอกสารภาษาไทยจากหนังสือพิมพ์ โดยอัลกอริทึมซัพพอร์ตเวกเตอร์แมชชีนให้ผลลัพท์ที่ดีที่สุดเช่นกัน

ทั้งนี้อัลกอริทึมซัพพอร์ตเวกเตอร์แมชชีนมีการจำแนกข้อมูลบนเส้นตรงหลายมิติ จะใช้ส่วนการเลือกที่มีความเหมาะสมที่สุดในการคัดเลือก ซึ่งในการคัดเลือกมาจากข้อมูลที่มีการสอนให้ระบบเรียนรู้ โดยจะใช้เส้นแบ่งที่เป็นเส้นตรงเพื่อหาจุดข้อมูลที่ทำให้ได้สมการเส้นตรงหลายมิติที่ดีที่สุด โดยพิจารณาผลรวมระยะห่างของเส้นตรงที่เป็นเส้นแบ่ง ถึงเส้นตรงที่ผ่านข้อมูลที่ใกล้ที่สุดและขนานกับเส้นแบ่งของแต่ละกลุ่มให้กลุ่มอย่างมีประสิทธิภาพ ต่างจากนาอีฟเบย์ที่ใช้การคำนวณความน่าจะเป็นในการคาดการณ์ผลลัพธ์ที่ได้และ โดยจะทำการวิเคราะห์ความสัมพันธ์ระหว่างตัวแปรเพื่อใช้ในการสร้างเงื่อนไข

6. ข้อเสนอแนะ

6.1 ข้อเสนอแนะทั่วไปจากการวิจัย

ในการจัดกลุ่มของหนังสือราชการ จะต้องใช้ผู้เชี่ยวชาญในการคัดเลือกกลุ่มของหนังสือเพื่อได้ข้อมูลที่ถูกต้อง อีกทั้งควรคำนึงถึงการกำจัดคำหยุดที่จะต้องครอบคลุมคำศัพท์ที่ไม่มีความจำเป็นให้มากที่สุด เนื่องจากจะส่งผลต่อการประมวลผลได้

6.2 ข้อเสนอแนะเพื่อการต่อยอดพัฒนา

งานวิจัยนี้สามารถต่อยอดเพื่อพัฒนาเป็นระบบช่วยแนะนำการเกษียนหนังสือราชการในหน่วยงานที่เกี่ยวข้องเพื่อลดความผิดพลาดในการเกษียนหนังสือราชการของเจ้าหน้าที่ และวิธีการในการวิจัยที่ให้ผลออกมาอย่างมีประสิทธิภาพนี้อาจจะสามารถนำไปประยุกต์เพื่อใช้ในการวิจัยที่เกี่ยวข้องกับการจัดกลุ่มด้วยวิธีการทำเหมืองข้อความ ที่มีข้อมูลรูปแบบเป็นข้อความในลักษณะคล้ายคลึงกันต่อไปได้

เอกสารอ้างอิง

[1] ระเบียบสำนักนายกรัฐมนตรี ว่าด้วยงานสารบรรณ (ฉบับที่ 2) พ.ศ. 2548, (2548, 23 กันยายน). ราชกิจจานุเบกษา. เล่ม 122 ตอนพิเศษ 99ง, หน้า 32.

[2] Kao, A.; Poteet, S. R. “Natural Language Processing and Text Mining”. USA : *Springer Science+Business Media*, 2007.

[3] Cortes, C.; Vapnik, V. “Support-vector networks”. *Machine Learning*, 1995

[4] Joachims. “Text categorization with support vector machines: Learning with many relevant features”. *Proc. of the 10 th European Conf. on Machine Learning*, pp.137–142, 1998.

[5] Press, W. H.; Teukolsky, S. A.; Vetterling, W. T.; Flannery, B. P. "Section 16.5. Support Vector Machines". *Numerical Recipes: The Art of Scientific Computing (3rd edition.).* New York: Cambridge University Press, 2007

[6] บุญเสริม กิจศิริกุล. “อัลกอริทึมการทำเหมืองข้อมูล”. ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย, 2546.

[7] นิเวศ จิระวิชิตชัย ปริญญา สงวนสัตย์ และพยุง มีสัตย์. “การจัดหมวดหมู่เอกสารภาษาไทยแบบอัตโนมัติด้วยซัพพอร์ตเวกเตอร์แมชชีน”. *การประชุมทางวิชาการระดับชาติด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศครั้งที่ 6*, 2553

[8] ศุภเทพ สติมั่น และ ทิพยา จินดาโกวิท. “การวัดประสิทธิภาพของอัลกอริทึมในการจำแนกประเภทข้อความสั้น กรณีศึกษา ข้อความภาษาไทยบนเว็บไซต์ทวิตเตอร์”. *การประชุมทางวิชาการระดับชาติด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศครั้งที่ 7*, 2554

[9] กานดา แผ่วัฒนากุล และดร.ปราโมทย์ ลือนาม. “การวิเคราะห์เหมืองความคิดเห็นบนเครือข่ายสังคมออนไลน์”. *วารสารการจัดการสมัยใหม่*, 2556

[10] ราชวิทย์ ทิพย์เสนา, ฉัตรเกล้า เจริญผล และแกมกาญจน์ สมประเสริฐศรี. “การจำแนกกลุ่มคำถามอัตโนมิติบนกระดานสนทนา โดยใช้เทคนิคเหมืองข้อความ”. *วารสารมหาวิทยาลัยมหาสารคาม*, 2556

[11] อดิเทพ ไชยสาร และ รัฐสิทธิ์ สุขะหุต. “การประมาณอารมณ์จากความคิดเห็นภาษาไทยโดยใช้เทคนิคการเรียนรู้ของเครื่อง”. *การประชุมทางวิชาการระดับชาติด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศครั้งที่ 9*, 2556

[12] สมศักดิ์ วิชัยกิจ และมาลีรัตน์ โสดานิล. “การจัดกลุ่มลูกค้าสินเชื่อจากการปฏิเสธ ด้วยวิธีการทำเหมืองข้อความ”. *การประชุมทางวิชาการระดับชาติด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศครั้งที่ 11*, 2558

การวิเคราะห์ชนิดของคำในประโยคภาษาอีสาน
Isarn Dialect Part of Speech Tagging

ธนกุล ปักกาเวสา (Tanakul Pakkawesa)[1] และพุธษดี ศิริแสงตระกูล (Pusadee Seresangtakul)[2]
ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์
มหาวิทยาลัยขอนแก่น
[1]tanakul@kkumail.com, [2]pusadee@kku.ac.th

บทคัดย่อ

งานวิจัยนี้นำเสนอการวิเคราะห์ชนิดของคำในประโยคภาษาถิ่นอีสานโดยการประยุกต์แบบจำลองฮิดเดนมาร์คอฟ ในวิจัยนี้ผู้วิจัยได้ทำการสร้างแบบจำลองการระบุชนิดของคำ โดยใช้คลังข้อความภาษาถิ่นอีสานจำนวน 20,000 ประโยค (292,441 คำภาษาถิ่นอีสาน) และใช้ประโยคภาษาถิ่นอีสานจำนวน 3,051 ประโยคเพื่อทดสอบความถูกต้องของระบบ ซึ่งผลการทดสอบพบว่าประสิทธิภาพของแบบจำลองที่สร้างด้วยแบบจำลองฮิดเดนมาร์คอฟมีความถูกต้องร้อยละ 93 และ 72 สำหรับการทดสอบด้วยคำที่ปรากฏในชุดข้อมูลฝึกสอน และการทดสอบด้วยคำที่ไม่ปรากฏในชุดฝึกสอน ตามลำดับ

คำสำคัญ: ภาษาถิ่นอีสาน การระบุชนิดของคำ แบบจำลองฮิดเดนมาร์คอฟ

Abstract

This research proposed the analysis of Part of Speech (POS) in Isarn dialect sentences using Hidden Markov Model (HMM). The study used Isarn dialect corpus to construct the POS tagging model. The corpus consists of 20,000 sentences (292,441 words). The model was evaluated by using 3,051 Isarn sentences. The results showed that the performances of the model built with HMM have 93% and 72% accuracy for words existing and not existing in the training model, respectively.

Keyword: Isarn Dialect, Part of Speech Tagging, POS, Hidden Markov Model.

1. บทนำ

การประมวลภาษาธรรมชาติ จัดเป็นหนึ่งในเทคโนโลยีที่มีความสำคัญต่อมนุษย์ เป็นเทคโนโลยีที่ถูกพัฒนาขึ้นเพื่อให้คอมพิวเตอร์สามารถเรียนรู้และเข้าใจภาษาของมนุษย์

ภาษาถิ่นอีสานเป็นภาษาพูดที่ใช้สื่อสารของคนภาคตะวันออกเฉียงเหนือในประเทศไทย เนื่องจากปัจจุบันประเทศไทยมีการใช้ภาษาไทยกลางและอักษรไทย เป็นภาษาราชการ จึงทำให้ภาษาถิ่นอีสานถูกลดบทบาทลง อันนำมาซึ่งความสูญหายของความรู้ วัฒนธรรม และภูมิปัญญาท้องถิ่น ปัจจุบันมีนักวิจัยได้นำเทคโนโลยีการประมวลภาษาธรรมชาติเข้ามาช่วยพื้นฟูภาษาถิ่นอีสาน ทั้งด้านการพัฒนาระบบแปลอักษรอีสานเป็นภาษาไทย [1] การสังเคราะห์เสียงภาษาถิ่นอีสาน [2] ระบบแปลภาษาไทยเป็นภาษาถิ่นอีสาน [3] แต่งานวิจัยเหล่านี้ยังต้องมีการพัฒนาให้มีคุณภาพที่ดีขึ้น จากการศึกษาพบว่างานโครงสร้างประโยคและการระบุชนิดของคำ จะสามารถช่วยเพิ่มประสิทธิภาพงานวิจัยด้านการประมวลภาษาธรรมชาติได้ [4] ช่วยให้คอมพิวเตอร์สามารถวิเคราะห์ประโยคได้ถูกต้องและแม่นยำ เนื่องจากความกำกวมของคำหนึ่งคำอาจมีได้หลายหน้าที่และหลายความหมาย ขึ้นอยู่กับโครงสร้างหรือรูปแบบของประโยค เช่น “หมอมอคิ้ว” (หมอมอ/NCMN คิ้ว/VATT) ให้ความหมาย “หมอดูแม่น” และ “หมอสักคิ้ว” (หมอ/NCMN สัก/VACT คิ้ว/NCMN) ให้ความหมาย “ช่างสักคิ้ว” สำหรับภาษาถิ่นอีสานยังขาดเครื่องมือเพื่อช่วยในการประมวลผลภาษา

ดังนั้นงานวิจัยนี้จึงมุ่งเน้นสร้างแบบจำลองและระบบการระบุชนิดของคำ ผู้วิจัยจึงได้จัดทำเครื่องมือการวิเคราะห์ชนิดของคำในประโยคภาษาอีสาน

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 ภาษาถิ่นอีสาน

ภาษาถิ่นอีสาน [5] เป็นภาษาท้องถิ่นที่ใช้พูดในภาคตะวันออกเฉียงเหนือของประเทศไทย ซึ่งมีสำเนียงการพูดที่เป็นเอกลักษณ์เฉพาะท้องถิ่น ภาษาถิ่นอีสานมีโครงสร้างประโยคที่เหมือนกับภาษาไทย แต่พบว่าภาษาถิ่นอีสานในบางประโยคมีลักษณะสำนวนเป็นคำหรือกลุ่มคำก็ได้

2.2 ชนิดของคำ (Part of Speech: POS)

ชนิดของคำคือตัวบ่งบอกลักษณะหน้าที่คำ ชนิดของคำจะให้ความหมายคำในประโยคได้ชัดเจน โดยการกำหนดมาตรฐานการกำกับชนิดของคำ (Orchid Corpus) [6] แบ่งแยกชนิดของคำเฉพาะเจาะจงออกเป็น 47 หมวดย่อยจากชนิดของคำตามหลักไวยากรณ์ เพื่อลดความกำกวมชนิดของคำที่ใกล้เคียงกันดังตัวอย่างที่แสดงในตารางที่ 1

ตารางที่ 1: ตัวอย่างชนิดของคำใน Orchid Corpus

ชนิดของคำ	คำอธิบาย
NCMN	นามทั่วไป
PDMN	พยางค์สรรพนาม
VACT	กริยาแสดงการกระทำ
XVAE	หลังกริยาช่วย
RPRE	คำบุพบท
EAFF	คำตอบสิ้นสุดประโยค

2.3 การระบุชนิดของคำด้วยแบบจำลองฮิดเดนมาร์คอฟ

ปัญหาการระบุชนิดของคำ [7] เป็นการระบุลำดับของชนิดของคำที่มีโอกาสเกิดขึ้นได้มากที่สุดของลำดับของคำในประโยค โดยทั่วไป ซึ่งสามารถอธิบายได้ดังสมการที่ 1

$$\hat{t}_1^n = \arg\max_{t_1^n} \mathrm{P}\left(t_1^n \mid w_1^n\right) \quad (1)$$

เมื่อ w_1^n คือ ลำดับของคำในประโยค

t_1^n คือ ลำดับของชนิดของคำ

แบบจำลองฮิดเดนมาร์คอฟ [8] สามารถแก้ปัญหาดังกล่าวโดยมีสมมติฐาน 2 ข้อ ดังนี้

1) ความน่าจะเป็นของการเกิดขึ้นของคำจะขึ้นกับชนิดของคำนั้น ๆ เท่านั้น และเป็นอิสระกับคำที่อยู่ข้างเคียง

2) ความน่าจะเป็นของการปรากฏชนิดของคำจะขึ้นกับชนิดของคำที่ปรากฏก่อนหน้า ซึ่งอธิบายได้ดังสมการที่ 2

$$\hat{t}_1^n = \arg\max_{t_1^n} \mathrm{P}\left(t_1^n \mid w_1^n\right) \approx \arg\max_{t_1^n} \prod_{i=1}^{n} \mathrm{P}\left(w_i \mid t_i\right)\mathrm{P}\left(t_i \mid t_{i-1}\right) \quad (2)$$

เมื่อ $\mathrm{P}\left(w_i \mid t_i\right)$ คือ ความน่าจะเป็นของการเกิดคำที่ I เมื่อทราบชนิดของคำที่ i

$\mathrm{P}\left(t_i \mid t_{i-1}\right)$ คือ ความน่าจะเป็นของการเกิดชนิดของคำที่ i เมื่อทราบชนิดของคำของคำก่อนหน้า $i-1$

โดยแบบจำลองฮิดเดนมาร์คอฟ สามารถแทนด้วยออโตมาตาประกอบด้วยเซตของสถานะ ซึ่งหมายถึงชนิดของคำ และเซตของการเปลี่ยนแปลงสถานะที่เกิดขึ้นตามลำดับของค่าสังเกต หรือลำดับของคำที่นำเข้า โดยแบบจำลองฮิดเดนมาร์คอฟ สามารถอธิบายได้ดังสมการที่ 3

$$\lambda = \left(A, B, \pi\right) \quad (3)$$

และประกอบด้วยพารามิเตอร์ต่าง ๆ ดังนี้

N คือ จำนวนของสถานะในแบบจำลอง

$O = \left\{o_1, o_2, ..., o_t\right\}$ คือ ลำดับของค่าสังเกต

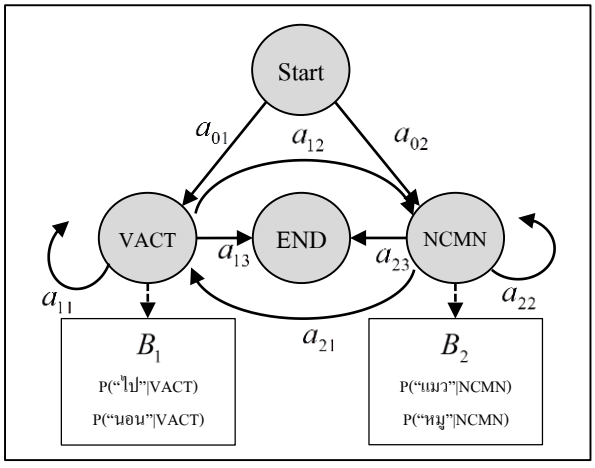
$Q = \left\{q_1, q_2, ..., q_t\right\}$ คือ เซตของสถานะ

$A = \left\{a_{01}, a_{02}, ..., a_{n1}, ..., a_{nn}\right\}$ คือ เซตของความน่าจะเป็นของการเปลี่ยนแปลงสถานะ

โดยที่ a_{ij} แทนความน่าจะเป็นของการเปลี่ยสถานะจากสถานะ i เป็นสถานะ j

$B = b_i\left(O_t\right)$ คือความน่าจะเป็นของการเกิดค่าสังเกต O_t ที่สถานะ i

$\prod = \left\{\pi_1, \pi_2, ..., \pi_n\right\}$ คือ ความน่าจะเป็นที่สถานะ ต่าง ๆ จะเป็นเกิดขึ้นเป็นสถานะแรก



ภาพที่ 1: แบบจำลองฮิดเดนมาร์คอฟสำหรับการระบุชนิดของคำ

จากภาพที่ 1 แสดงแบบจำลองฮิดเดนมาร์คอฟสำหรับการระบุชนิดของคำ ซึ่งจะเห็นได้ว่าในแบบจำลองจะพิจารณาค่าความน่าจะเป็น 2 ส่วน คือ ความน่าจะเป็นของการเปลี่ยนสถานะ A และความน่าจะเป็นของการเกิดค่าสังเกต B ซึ่งแทนสถานะของคำ VACT และ NCMN โดยที่ B_1 และ B_2 จะแทนความน่าจะเป็นของค่าสังเกตหรือคำที่จะเกิดเป็นชนิดของคำ VACT และ NCMN ตามลำดับ

2.2.1 การแก้ปัญหาการระบุชนิดของคำ

การระบุชนิดของคำสามารถทำได้โดยค้นหาลำดับของสถานะที่มีค่าความน่าจะเป็นสูงสุด เมื่อนำเข้าลำดับของค่าสังเกต อัลกอริทึมไวเทอร์บิ (Viterbi) [9] จะถูกใช้ค้นหาลำดับของสถานะที่ ที่มีความน่าจะเป็นมากที่สุด เมื่อกำหนดให้ $Q=\{q_1,q_2,...,q_t\}$ เป็นลำดับการเปลี่ยนแปลงสถานะที่ดีที่สุด และ $O=\{o_1,o_2,...,o_t\}$ เป็นลำดับของการสังเกตต้องการหาค่า

$$\delta_t(i)=\max_{q_1,...,q_t} \mathrm{P}(o_1,...,o_t,q_1,...,q_{t-1},q_t=i\mid\lambda) \quad (4)$$

เมื่อ $\delta_t(i)$ เป็นค่าที่มีความน่าจะเป็นสูงสุดของเส้นทางทั้งหมดในลำดับที่ t ที่สถานะ i และ $\psi_t(i)$ คือเส้นทางที่ให้ความน่าจะเป็นมากสุด ในวิธีของไวเทอร์บิสามารถหาลำดับของสถานะที่มีความน่าจะเป็นสูงสุดได้ตามขั้นตอนดังต่อไปนี้

1) กระบวนการเริ่มต้น

$$\delta_1(i)=\pi_i b_i(o_1), \qquad 1\le i\le N$$
$$\psi_1(i)=0, \qquad 1\le i\le N$$

2) กระบวนการทำซ้ำ

$$\delta_1(i)=\max_i\left[\delta_i(i)a_{ij}\right]o_t, \qquad 1\le i\le N,\ t=2,...,T$$
$$\psi_1(j)=\arg\max_i\left[\delta_i(i)a_{ij}\right], \qquad 1\le i\le N,\ t=2,...,T$$

3) การสิ้นสุด

$$\mathrm{P}(o,q^*\mid\lambda)=\max_i\left[\delta_t(i)\right],$$
$$Q_T=\arg\max_i\left[\delta_t(i)\right]$$

4) เส้นทางย้อนกลับ

$$Q_t=\psi_{t+1}(Q_{t+1})$$

2.4 การวัดประสิทธิภาพด้วยวิธี Cross validation

การทดสอบประสิทธิภาพของแบบจำลองให้มีความน่าเชื่อถือด้วยวิธี Cross validation [10] ทำการแบ่งข้อมูลออกเป็นหลายส่วน (เรียกว่า k-fold cross validation) จะทำการแบ่งข้อมูลออกเป็น k ส่วน โดยที่แต่ละส่วนมีจำนวนข้อมูลเท่ากัน หลังจากนั้นนำข้อมูล k-1 ส่วน Train Model และนำข้อมูลส่วนที่เหลือ มาทำการทดสอบวนทำจนครบ k เพื่อประสิทธิภาพของแบบจำลอง

2.5 งานวิจัยที่เกี่ยวข้อง

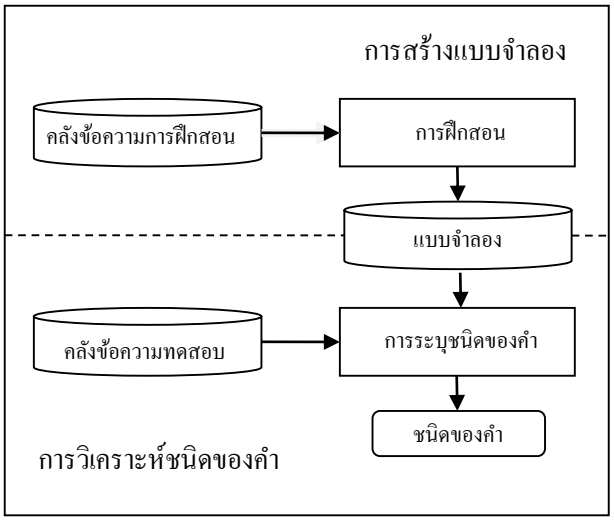
A. P. J, S. P. Mohan and S. K. P. [11] นำเสนอชนิดของคำสำหรับภาษา Malayalam โดยการระบุความคลุมเครือของคำศัพท์ และพัฒนาชนิดของคำ ให้มีประสิทธิภาพและความถูกต้องเพิ่มขึ้น โดยสร้างชนิดของคำขึ้นมาฝึกฝนและทดสอบเพื่อสร้างรูปแบบประโยค จากชนิดของคำจำนวน 29 ชนิด ใช้คลังข้อความขนาด 100 ประโยค (จำนวน 80,000 คำ) พบว่าผลลัพธ์มีประสิทธิภาพและความถูกต้องมากขึ้นเมื่อเปรียบเทียบกับแบบเดิมที่มีคลังข้อความและความคลุมเครืออย่างจำกัด

N. Mishra and A. Mishra [12] นำเสนอการสร้างเครื่องมือระบุชนิดของคำ ที่สามารถอธิบายชนิดของคำภาษาฮินดี ที่เกิดขึ้นในประโยค กับส่วนการระบุชนิดของคำ เช่น นาม สรรพนาม กริยา คำคุณศัพท์ เป็นต้น เป็นการระบุชนิดของคำศัพท์ที่อยู่ในประโยคของบริบทนั้น ๆ ผู้วิจัยได้พัฒนาขึ้นมาสามารถอ่านคลังข้อความ แสดงประโยคและคำศัพท์ของภาษาฮินดีได้ เกิดเป็นเครื่องมือที่ผู้สนใจนำไปใช้ประโยชน์ได้อย่างง่าย

S.K.Sharma และ G.S. Lehal [13] ได้ปรับปรุงการระบุชนิดของคำในภาษา Punjabi โดยการกำหนดชนิดของคำเพื่อแก้ปัญหาความคลุมเครือของคำและความซับซ้อนประโยค ด้วยแบบจำลองฮิดเดนมาร์คอฟ โดยการพิจารณาชนิดของคำในแต่ละคำจากคำก่อนหน้าแบบ (Bi-gram) โดยสร้างคลังข้อความขนาด 20,000 คำ เพื่อใช้ในการฝึกสอนและวัดประสิทธิภาพด้วยคลังข้อความขนาด 26,479 ข้อความ และให้ความถูกต้อง 90.11% เมื่อเทียบกับการระบุชนิดคำด้วยมือ

3. วิธีดำเนินการ

การระบุชนิดของคำในประโยคภาษาถิ่นอีสานได้นำเทคนิควิธีการของแบบจำลองฮิดเดนมาร์คอฟ (Hidden Markov Model: HMM) มาใช้ในการระบุชนิดของคำในประโยคภาษาถิ่นอีสาน ในการดำเนินการวิจัยประกอบด้วย 2 ขั้นตอนหลัก ดังภาพที่ 2



ภาพที่ 2: การวิเคราะห์ชนิดของคำในประโยคภาษาถิ่นอีสาน

3.1 การสร้างแบบจำลอง

การสร้างแบบจำลองการระบุชนิดของคำ เป็นขั้นตอนแรกที่สำคัญในงานวิจัยเพื่อให้แบบจำลองได้เรียนรู้โครงสร้างชนิดของคำในแต่ละคำของประโยค รวมถึงการเชื่อมโยงของโครงสร้างประโยค โดยอาศัยคลังข้อความภาษาถิ่นอีสานเพื่อให้แบบจำลองเรียนรู้คุณลักษณะต่าง ๆ ในขั้นตอนนี้ประกอบด้วย 2 ขั้นตอน ดังนี้

3.1.1 การรวบรวมและเตรียมข้อมูลเพื่อฝึกสอน

ในการวิจัยนี้ได้ใช้ข้อความภาษาถิ่นอีสานจากคลังข้อความภาษาถิ่นอีสานของโครงการแปลภาษาถิ่นอีสาน [3] ซึ่งคลังข้อความดังกล่าวมีการรวบรวม และข้อมูลภาษาถิ่นอีสานที่จัดเก็บในคลังข้อความดังกล่าวจะถูกเขียนด้วยอักษรไทย และในคลังข้อความที่ใช้ประกอบด้วยภาษาถิ่นอีสาน ดังตารางที่ 2 ในการศึกษาวิจัยนี้ผู้วิจัยได้ทำการแบ่งข้อมูลออกเป็น 2 ส่วน โดยส่วนแรกจะถูกใช้ในการเรียนรู้หรือสร้างแบบจำลอง และส่วนที่สองจะถูกใช้เพื่อทดสอบระบบที่พัฒนาขึ้น

ตารางที่ 2: รายละเอียดคลังข้อความภาษาถิ่นอีสาน

รายการ	จำนวนประโยค	จำนวนคำ
การฝึกฝน	20,000	292,441
การทดสอบ	3,051	44,714
รวม	**23,051**	**337,155**

3.1.2 สร้างแบบจำลองฮิดเดนมาร์คอฟ

ในขั้นตอนของการสร้างแบบจำลอง ค่าสถิติภายในแบบจำลองจะถูกคำนวน ซึ่งประกอบด้วยค่าความน่าจะเป็นของสถานะที่จะเกิดเป็นสถานะเริ่มต้น และค่าความน่าจะเป็นของการเปลี่ยนแปลงสถานะ เมื่อกำหนดให้คลังข้อความดังภาพที่ 3 สามารถคำนวนค่าสถิติได้ตามขั้นตอนดังต่อไปนี้

ไป/VACT หม่อง/RPRE นี้/PDMN
ไป/VACT กิน/VACT ข้าว/NCMN
ไป/VACT หา/XVAE นอน/VACT

ภาพที่ 3: ตัวอย่างคลังข้อความฝึกสอน

1) ค่าความน่าจะเป็นของสถานะที่จะเกิดเป็นสถานะเริ่มต้น โดยสามารถคำนวนได้จากจำนวนของการเกิดสถานะนั้นเป็นสถานะเริ่มต้นหารด้วยจำนวนสถานะที่เกิดขึ้นเป็นสถานะเริ่มต้นทั้งหมด จากคลังข้อความข้างต้นสามารถคำนวนค่าความน่าจะเป็นของสถานะที่จะเกิดเป็นสถานะเริ่มต้นได้ดังตารางที่ 3

ตารางที่ 3: ค่าความน่าจะเป็นสถานะเริ่มต้น

	VACT	RPRE	PDMN	NCMN	XVAE
Start	1.0	0	0	0	0

จากตารางที่ 3 มีการเกิดสถานะ VACT เป็นสถานะเริ่มต้นจำนวน 3 ครั้ง และจำนวนครั้งที่เกิดสถานะเริ่มต้นทั้งหมด 3 ครั้ง ดังนั้นที่สถานะ VACT จะมีความน่าจะเป็นที่จะเกิดเป็นสถานะเริ่มต้นเป็น 1.0

2) ค่าความน่าจะเป็นการเปลี่ยนแปลงสถานะ โดยสามารถคำนวนได้โดยนำจำนวนการเปลี่ยนสถานะ i เป็นสถานะ j หารด้วยจำนวนการเปลี่ยนสถานะทั้งหมดของ

สถานะทั้งหมดมายังสถานะ j จากคลังข้อความสามารถคำนวนค่าความน่าจะเป็นของการเปลี่ยนสถานะได้ดังตารางที่ 4

ตารางที่ 4: ค่าความน่าจะเป็นการเปลี่ยนแปลงสถานะ

	VACT	RPRE	PDMN	NCMN	XVAE
VACT	0.25	0.25	0	0.25	0.25
RPRE	0	0	1.0	0	0
PDMN	1.0	0	0	0	0
NCMN	1.0	0	0	0	0
XVAE	1.0	0	0	0	0

จากตารางที่ 4 การเปลี่ยนจากสถานะ VACT ไป VACT พบความถี่ทั้งหมด 1 ครั้ง และการเปลี่ยนแปลงสถานะมายังสถานะ VACT ทั้งหมด 4 ครั้ง ดังนี้ค่าความน่าจะเป็นของการเปลี่ยจากสถานะ VACT เป็นสถานะ VACT P(VACT|VACT) จะเท่ากับ1/4 = 0.25

3) ค่าความน่าจะเป็นของการเกิดค่าสังเกตที่สถานะต่าง ๆ สามารถคำนวนได้จากจำนวนของการเกิดค่าสังเกตในสถานะนั้นหารด้วยจำนวนของการเกิดสถานะนั้นทั้งหมด จากข้อมูลตัวอย่างสามารถคำนวนค่าความน่าจะเป็นของการเกิดค่าสังเกตในสถานะต่าง ๆ ได้ดังตารางที่ 5

ตารางที่ 5: ค่าความน่าจะเป็นของค่าสังเกตคำในสถานะชนิดของคำ

	ไป	หม่อง	นี้	กิน	ข้าว	หา	นอน
VACT	0.6	0	0	0.2	0	0	0.2
RPRE	0	1.0	0	0	0	0	0
PDMN	0	0	1.0	0	0	0	0
NCMN	0	0	0	0	1.0	0	0
XVAE	0	0	0	0	0	1.0	0

จากตารางที่ 5 ยกตัวอย่างความสัมพันธ์ของคำว่า “ไป” กับสถานะของ VACT พบความถี่ทั้งหมด 3 ครั้ง ของการเกิดสถานะชนิดคำ VACT ทั้งหมด 5 ครั้ง ได้ค่าความน่าจะเป็นของ P(“ไป”|VACT) = 3/5 = 0.6

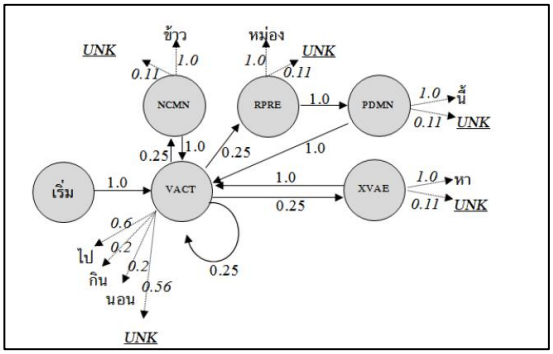
4) การจำลองค่าความน่าจะเป็นค่าสังเกตของคำที่ไม่ปรากฏในข้อมูลฝึกสอน เพื่อเพิ่มประสิทธิภาพแบบจำลองให้สามารถวิเคราะห์ชนิดของคำในประโยคที่ไม่เคยปรากฏในข้อมูลฝึกสอน ดังตารางที่ 6

ตารางที่ 6: ค่าสังเกตของคำที่ไม่เคยปรากฏ

	VACT	RPRE	PDMN	NCMN	XVAE
UNK	0.56	0.11	0.11	0.11	0.11

จากตารางที่ 6 ตัวอย่างการจำลองค่าความน่าจะเป็นของคำที่ไม่เคยปรากฏกับสถานะ VACT จำนวน 5 ครั้งของการเกิดสถานะชนิดคำทั้งหมด ทั้งหมด 9 ครั้ง ได้ค่าความน่าจะเป็นของ P(“UNK”|VACT) = 5/9 = 0.56

ภาพที่ 4 แสดงแบบจำลองฮิดเดนมาร์คอฟของคลังข้อความตัวอย่าง โดยภายในแบบจำลองจะประกอบด้วย 5 สถานะ คือ VACT RPRE PDMN NCMN และ XVAE โดยสถานะ VACT สามารถเกิดเป็นสถานะแรกเท่านั้น



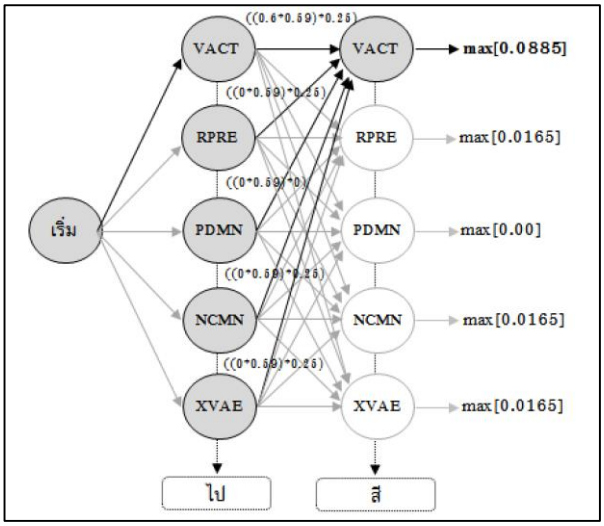
ภาพที่ 4: แสดงแบบจำลองฮิดเดนมาร์คอฟของคลังข้อความตัวอย่าง

3.2 วิเคราะห์ชนิดของคำ

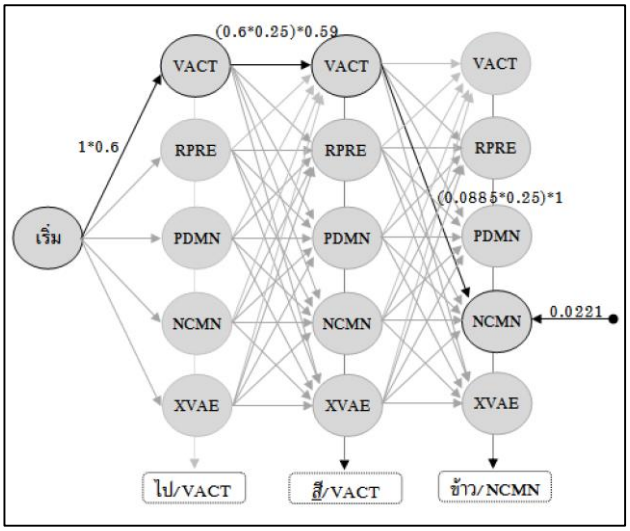
ในส่วนของการวิเคราะห์ชนิดของคำในประโยคภาษาถิ่นอีสาน จะอาศัยวิธีของไวเทอร์บิ เพื่อค้นหาลำดับของสถานะที่มีโอกาสที่จะเกิดขึ้นมากที่สุด เมื่อกำหนดลำดับของคำในประโยคตัวอย่าง “ไปสีข้าว” สามารถค้นหาลำดับของสถานะที่มีความน่าจะเป็นสูงสุดได้ดังต่อไปนี้

1) การหาสถานะเริ่มต้น “ไป” เป็นคำแรกของประโยคซึ่งสามารถเป็น VACT ได้เพียงสถานะเดียวเท่านั้น จะได้ P(“ไป”|VACT) × P(Start | VACT) = 1 × 0.6 = 0.6

2) การหาค่าความน่าจะเป็นลำดับถัดไป “สี” เป็นคำลำดับของสถานะที่สอง (S_2) ซึ่งหาค่าความน่าจะเป็นที่มีค่าสูงสุดจากสถานะก่อนหน้าที่ส่งต่อการเปลี่ยนแปลงสถานะมายังเซ็ตสถานะของ S_2 ดังภาพที่ 5 และแสดงการคำนวนของแต่ละเซ็ตสถานะดังตารางที่ 7



ภาพที่ 5: แสดงตัวอย่างการคำนวนของลำดับถัดไป



ภาพที่ 6: แสดงลำดับเส้นทางที่มีค่าความน่าจะเป็นสูงสุดแต่ละสถานะ

ตารางที่ 7: แสดงการคำนวนหาค่าสถานะถัดไป

S_{i-1}	δ_{i-1}	δ_i	ผลลัพธ์
VACT	0.6	max((0.6*0.25)*0.59)	0.0885
RPRE	0	max((0.6*0.25)*0.11)	0.0165
PDMN	0	max((0.6*0)*0.11)	0
NCMN	0	max((0.6*0.25)*0.11)	0.0165
XVAE	0	max((0.6*0.25)*0.11)	0.0165

3) กระบวนการสิ้นสุด โดยเลือกคำตอบสุดท้ายที่มีความน่าจะเป็นสูงสุดของลำดับสุดท้าย

4) กระบวนการย้อนกลับเพื่อเลือกเส้นทางของลำดับแต่ละสถานะที่มีค่าความน่าจะเป็นสูงสุดในสถานะนั้น ๆ ดังภาพที่ 6

4. ผลการดำเนินงาน

4.1 การวัดประสิทธิภาพแบบจำลองด้วยวิธีการ Cross validation Test กับข้อมูลชุดฝึกสอน โดยจะแบ่งข้อมูลเป็น K ชุด จำนวนเท่า ๆ กัน ในงานวิจัยนี้กำหนดจำนวนชุดข้อมูล (K) ในการทดสอบเป็น 3, 5, 7 และ 10 Fold ในการวัดความถูกต้องของการระบุชนิดของคำจะเปรียบเทียบกับชนิดของคำต้นฉบับที่ระบุไว้ล่วงหน้า จะพิจารณาตามการปรากฏของคำในข้อมูลชุดฝึกสอนให้ความถูกต้องของแบบจำลองดังตารางที่ 8

ตารางที่ 8: ผลการทดสอบประสิทธิภาพของแบบจำลอง

Kfold	คำที่ปรากฏ		คำที่ไม่ปรากฏ	
	จำนวนคำ	ความถูกต้อง	จำนวนคำ	ความถูกต้อง
3 Fold	97,328	91.87	3,644	69.76
5 Fold	58,369	92.83	2,000	70.91
7 Fold	41,854	92.92	1,361	71.43
10 Fold	29,169	92.73	933	70.43

4.2 การวัดประสิทธิภาพของแบบจำลอง โดยใช้ข้อมูลฝึกสอนจำนวนทั้งสิ้น 20,000 ประโยค และทำการทดสอบด้วยข้อมูลจำนวน 3,051 ประโยค ผลการวัดประสิทธิภาพของแบบจำลองที่สร้างขึ้นดังตารางที่ 9

ตารางที่ 9: ประสิทธิภาพของแบบจำลอง

ประเภทชุดทดสอบ	จำนวนคำ	ความถูกต้อง
คำที่ปรากฏ	44,714	93.00%
คำที่ไม่ปรากฏ	1,342	72.00%

จากการทดลองการระบุชนิดของคำโดยใช้แบบจำลองฮิดเดนมาร์คอฟ พบว่าแบบจำลองสามารถให้ประสิทธิภาพความถูกต้องได้ดีเมื่อใช้กับคลังข้อความที่มีคำที่ไม่ปรากฏน้อย และยังพบว่าความผิดพลาดในการระบุชนิดของคำเกิดจากคำที่ไม่เคยปรากฏ เช่น ชื่อคำเฉพาะ

5. สรุปผล

การวิจัยเพื่อสร้างแบบจำลองการระบุชนิดของคำในประโยคภาษาถิ่นอีสานนั้น ผู้วิจัยได้ใช้ข้อมูลภาษาถิ่นอีสานจากคลังข้อความโครงการแปลภาษาอีสานที่ผลการวิจัยพบว่าแบบจำลองชนิดของคำให้ความถูกต้องจากการระบุชนิดของคำมากกว่า 90% โดยมีความถูกต้องที่ 93% และ 72% ของคำที่เคยปรากฏและคำที่ไม่เคยปรากฏ ตามลำดับ ซึ่งระบบที่ได้ทำการพัฒนาขึ้นนี้จะเป็นประโยชน์ในการประยุกต์ใช้เพื่อเพิ่มประสิทธิภาพของงานประมวลด้านภาษาธรรมชาติของภาษาถิ่นอีสาน

เอกสารอ้างอิง

[1] P. Lakkhanawannakun and P. Seresangtakul, "Improved Isarn Dharma Alphabets to Thai Language Translation using Longest Syllable Matching with Named Entities Recognition," *The 5th International Conference on Advanced Computer Control (ICACC)*, Singapore, Dec 14-15, 2013, pp. 635-644.

[2] พงษ์ศธร จันทร์ยอย, "ระบบสังเคราะห์เสียงพูดภาษาอีสาน" *วิทยานิพนธ์ปริญญาวิทยา ศาสตรมหาบัณฑิต สาขาวิทยาการคอมพิวเพตอร์*. บัณฑิตมหาวิทยาลัย. มหาวิทยาลัยขอนแก่น, ขอนแก่น, 2558.

[3] P. Unlee and P. Seresangtakul, " Thai to Isarn dialect Machine Translation using Rule-based and Example-based," *Proceedings of 13th International Joint Conference on Computer Science and Software Engineering*, Khon Kaen, Thailand, July, 2016, pp.13-15.

[4] I. Bordino et al, "Advancing NLP via a distributed-messaging approach," *IEEE International Conference on Big Data (Big Data)*, Washington, DC, 2016, pp. 1561-1568.

[5] เรืองเดช ปันเขื่อนขัติย์, "ภาษาศาสตร์ภาษาไทย" *สถาบันวิจัยภาษาและวัฒนธรรมเพื่อพัฒนาชนบท*. มหาวิทยาลัยมหิดล, กรุงเทพฯ, 2541.

[6] V. Sornlertlamvanich , N. Takahashi and H. Isahara, "Building a Thai part-of-speech tagged corpus (ORCHID)," *Journal of the Acoustical Society*, Japan, 1999, 20(3), pp.189-198.

[7] S.Z. Lee,J. Tsujii and H.C. Rim, " Lexicalized hidden Markov models for part-of-speech tagging," *Proceedings of the 18th conference on Computational linguistics Volume 1*, Stroudsburg, PA, USA, Jul 31 - Aug 04, 2000, pp. 481-487.

[8] Z. Ghahramani, "An Introduction to Hidden Markov Models and Bayesian Networks," *International Journal of Pattern Recognition and Artificial Intelligence 2001*, Feb, 2001, Vol. 15, pp. 9-42.

[9] S. Sun, H. Lin and H. Liu, "A hybrid PSO-Viterbi algorithm for HMMs parameters weighting in Part-of-Speech tagging," *International Conference of Soft Computing and Pattern Recognition (SoCPaR)*, Dalian, 2011, pp. 518-522.

[10] B. Aryoyudanta, T. B. Adji and I. Hidayah, "Semi-supervised learning approach for Indonesian Named Entity Recognition (NER) using co-training algorithm," *2016 International Seminar on Intelligent Technology and Its Applications (ISITIA)*, Lombok, 2016, pp. 7-12.

[11] P.J. Antony, P.M. Santhanu, K.P. Soman, "SVM Based Part of Speech Tagger for Malayalam," *Proceedings of International Conference on Recent Trends in Information, Telecommunication and Computing*, Kochi, Kerala, India, March 12-13, 2010, pp. 339 - 341.

[12] N. Mishra amd A. Mishra, "Part of Speech Tagging for Hindi Corpus," *Proceedings of International Conference on Communication Systems and Network Technologies*, Jammu, India, June 3-5, 2011, pp. 554 - 558.

[13] S.K. Sharma and G.S. Lehal, "Using Hidden Markov Model to improve the accuracy of Punjabi POS tagger," *Proceedings of IEEE International Conference on Computer Science and Automation Engineering*, Shanghai, China, Jun 11-12, 2011, pp. 697-701.

PL-USSR: ตำแหน่งทางกายภาพของจุดเชื่อมต่อแบบไร้สายโดยพิจารณาช่วงความเข้มของสัญญาณที่ยอมรับไม่ได้

PL-USSR: Physical Location of Wireless Access Point with Considering Unacceptable Signal Strength Range

ปรัชญา ใจสุทธิ (Pratchaya Jaisudthi)[1] และสุรเชษฐ์ ยุวรัตน์ (Surachet Yuwarat)[2]

[1,2]สาขาวิชาเทคโนโลยีเครือข่ายคอมพิวเตอร์ คณะวิทยาการคอมพิวเตอร์และเทคโนโลยีสารสนเทศ

มหาวิทยาลัยราชภัฏรำไพพรรณี

pratchaya.j@rbru.ac.th, interceptbto@gmail

บทคัดย่อ

งานวิจัยนี้วิเคราะห์หาตำแหน่งทางกายภาพสำหรับการติดตั้งจุดเชื่อมต่อไร้สายหนึ่งจุดภายในบ้านโดยพิจารณาช่วงความเข้มของสัญญาณที่ยอมรับไม่ได้ ซึ่งใช้แอปพลิเคชันที่ผู้ดูแลระบบเครือข่ายใช้โดยทั่วไป นั่นก็คือ โปรแกรม Ekahau Heatmapper และ inSSIDer โดยจะนำแผนที่และค่าแสดงระดับความความเข้มของสัญญาณที่วัดได้จากทั้งสองแอปพลิเคชันมาช่วยในการวิเคราะห์หาตำแหน่งติดตั้งที่เหมาะสมด้วยวิธีการที่งานวิจัยนี้ได้นำเสนอคือ PL-USSR แล้วมาเปรียบเทียบกับการหาตำแหน่งโดยใช้การคำนวณด้วยค่าเฉลี่ยเลขคณิตถ่วงน้ำหนักของความเข้มของสัญญาณ ซึ่งผลที่ได้วิธี PL-USSR สามารถทำให้สัญญาณครอบคลุมพื้นที่มากยิ่งขึ้น

คำสำคัญ: ตำแหน่งทางกายภาพ, จุดเชื่อมต่อไร้สาย, ความเข้มของสัญญาณ.

Abstract

This research analyzes the physical location for the installation of single wireless access point in the home by considering the unacceptable range of signal strength. We used applications that network administrators commonly use, the Ekahau Heatmapper and inSSIDer. We took the signal strength map and values of both applications to determine the appropriate location of installation with proposed method in this research called PL-USSR. Then we compared our proposed method with positioning by using the weighted arithmetic mean of signal strength. As a result, the PL-USSR method can make the signal cover more area.

Keyword: Physical Location, Wireless Access Point, Signal Strength.

1. บทนำ

ปัจจุบันการใช้เทคโนโลยีเครือข่ายแบบไร้สายถือเป็นปัจจัยสำคัญในชีวิตประจำวัน เพราะผู้ใช้ต้องอาศัยเครือข่ายดังกล่าวในการเชื่อมต่อเพื่อใช้งานอินเทอร์เน็ตในด้านต่าง ๆ ทั้งภายในบ้าน ภายในอาคารและภายนอกอาคาร โดยคนสมัยใหม่ได้ใช้ชีวิตอยู่ภายในอาคารถึงร้อยละ 80 ถึงร้อยละ 90 ของชีวิต [1] แต่ภายในอาคารหรือบ้านมักมีสิ่งกีดขวางเป็นจำนวนมาก ซึ่งจะส่งผลต่อระดับความเข้มของสัญญาณที่น้อยลงจนบางครั้งทำให้ไม่สามารถเชื่อมต่อเพื่อใช้งานอินเทอร์เน็ตได้ [2] ดังนั้นการเลือกจุดสำหรับติดตั้งจุดเชื่อมต่อสัญญาณไร้สายให้ครอบคลุมพื้นที่ใช้งานทั้งหมดจึงเป็นสิ่งที่สำคัญ ที่ผ่านมามีงานวิจัยหลายงานที่พยายามค้นหาวิธีในการหาจุดติดตั้งที่เหมาะสม งานวิจัย [3] เป็นการตำแหน่งติดตั้งจุดเชื่อมต่อสัญญาณโดยใช้โปรแกรมจำลองกา รทำงาน (Simulation) เข้าช่วย นอกจากนี้งานวิจัย [4] ยังได้วิเคราะห์ปัจจัยสำคัญที่ส่งผลต่อความถูกต้องของตำแหน่งที่ได้รับสัญญาณที่ดี แต่ไม่ได้ระบุถึงตำแหน่งของจุดเชื่อมต่อที่เหมาะสม ดังนั้นงานวิจัยนี้จึงมีแนวคิดวิเคราะห์หาตำแหน่งที่เหมาะสมในการติดตั้งโดยแสดงผลออก

โปรแกรมตรวจสอบสัญญาณไร้สายในแต่ละจุด เป็นโปรแกรมที่ได้รับนิยมในการตรวจสอบว่าจุดที่ติดตั้งจุด

เชื่อมต่อสัญญาณครอบคลุมพื้นที่เท่าไหร่ โดยเฉพาะอย่างยิ่งแอปพลิเคชัน inSSIDer [5] ซึ่งเป็นซอฟต์แวร์ฟรี สามารถตรวจสอบ`สัญญาณไวไฟและตรวจหาคุณสมบัติของแต่ละ SSID ได้ เพื่อใช้ในการปรับปรุงการกระจายสัญญาณไวไฟ นอกจากนี้งานวิจัย [6] ได้นำ inSSIDer ไปใช้ในการหาความเข้มของสัญญาณสำหรับช่วงความถี่ 2.4 กิกะเฮิร์ต หรือ 5 กิกะเฮิร์ต ดังนั้นงานวิจัยนี้จึงนำ inSSIDer มาช่วยในหาตำแหน่งที่เหมาะสมในการติดตั้งจุดเชื่อมต่อสัญญาณไร้สาย ซึ่งรองรับสัญญาณความถี่ในช่วงดังกล่าว

แอปพลิเคชันสำหรับสำรวจสถานที่ตั้ง (Site Survey) เป็นโปรแกรมที่เป็นประโยชน์มากในการออกแบบระบบเครือข่ายไร้สาย แต่อย่างไรก็ตามโปแกรมดังกล่าวมีราคาค่อนข้างสูง จึงมีการนำแผนที่ระดับความเข้มของสัญญาณ (Heatmapper) มาช่วยเป็นเครื่องมือในการออกแบบแทน เช่น แอปพลิเคชัน Ekahau Heatmapper [7] ซึ่งเป็นแอปพลิเคชันที่มีให้ดาวน์โหลดใช้ฟรีและนำมาประยุกต์ใช้กับการติดตั้งเครือข่ายและตรวจสอบความแรงของสัญญาณ งานวิจัย [8] ได้นำ Ekahau Heatmapper มาใช้วัดความแรงของสัญญาณ แล้วนำเปรียบเทียบกับโปรแกรมที่พัฒนาขึ้น ซึ่งทำให้ช่วยประหยัดเวลาโดยผู้ใช้ไม่ต้องไปวัดสัญญาณในแต่ละจุด ดังนั้นในงานวิจัยนี้จึงได้นำ Ekahau Heatmapper มาช่วยวิเคราะห์หาตำแหน่งที่เหมาะสมในการติดตั้งจุดเชื่อมต่อสัญญาณไร้สายร่วมกับโปรแกรม inSSIDer

จากที่กล่าวมาทั้งหมดผู้วิจัยจึงนำเสนอวิธีการคำนวณตำแหน่งที่เหมาะสมของจุดเชื่อมต่อไร้สายหนึ่งจุด โดยเรียกวิธีนี้ว่า ตำแหน่งทางกายภาพของจุดเชื่อมต่อแบบไร้สายโดยพิจารณาช่วงความเข้มของสัญญาณที่ยอมรับไม่ได้ (PL-USSR : Physical Location of Wireless Access Point with Considering Unacceptable Signal Strength Range) โดยอาศัยแอปพลิเคชัน inSSIDer และ Ekahau Heatmapper เข้ามาช่วยในการวิเคราะห์หาจุดติดตั้งที่เหมาะสม

2. สิ่งที่เสนอ

ในงานวิจัยนี้ผู้วิจัยได้เสนอวิธี PL-USSR ซึ่งมีขั้นตอนคือนำค่าที่ได้จากการสำรวจโปรแกรม inSSIDer และ Ekahau Heatmapper มาใช้ในการพิจารณาจุดที่เหมาะสมในการกำหนดตำแหน่งของจุดเชื่อมต่อไร้สาย โดยหาค่าเฉลี่ยของสัญญาณของค่าที่ได้จากสำรวจด้วยโปรแกรม inSSIDer ตามสมการที่ 1

$$\bar{I} = \frac{\sum_{i=1}^{N} I_i}{N} \quad (1)$$

โดย

I_{i} แทนความเข้มของสัญญาณจุดที่ i

N คือจำนวนจุดวัดระดับสัญญาณทั้งหมด

$\bar{I}$ คือค่าเฉลี่ยของความเข้มของสัญญาณที่วัดได้ทั้งหมด

หลังจากนั้นหาช่วงสัญญาณที่ยอมรับได้โดยการบวกลบด้วยค่าเบี่ยงเบนมาตรฐาน ซึ่งหมายความว่าจุดใด ๆ ที่ยังอยู่ในช่วงสัญญาณดังกล่าว ผู้ใช้งานสามารถยอมรับได้ โดยแสดงตามสมการที่ 2

$$A = \{x | \bar{I} - aS \leq x \leq \bar{I} + aS\} \quad (2)$$

โดย

A แทนเซตของความเข้มของสัญญาณที่ยอมรับได้

a แทนค่าที่ไม่เจาะจงซึ่งผู้ใช้สามารถกำหนดเองได้

S แทนค่าเบี่ยงเบนมาตรฐานของความเข้มของสัญญาณที่วัดได้จากโปรแกรม inSSIDer

หลังจากนั้นแบ่งกลุ่มของสัญญาณตามระดับความเข้มของสัญญาณหรือตามสีที่ได้จากการสำรวจด้วยโปรแกรม Ekahau Heatmapper ทำการเลือกกลุ่มของสัญญาณหรือแถบสีที่มีส่วนเบี่ยงเบนมาตรฐานจากค่าเฉลี่ยที่วัดได้จากโปรแกรม inSSIDer ที่ต่ำที่สุด เรียกว่า E แล้วเลือกสัญญาณที่ไม่อยู่ในช่วงสัญญาณที่ยอมรับได้แต่ยังอยู่ในกลุ่ม E ว่า G_e ตามสมการที่ 3

$$G_e = \{x | x \notin A \ \wedge x \in E\} \quad (3)$$

แล้วคำนวณจุดที่เหมาะสมโดยใช้ค่าเฉลี่ยแบบถ่วงน้ำหนักตามสมการ (4) และ (5)

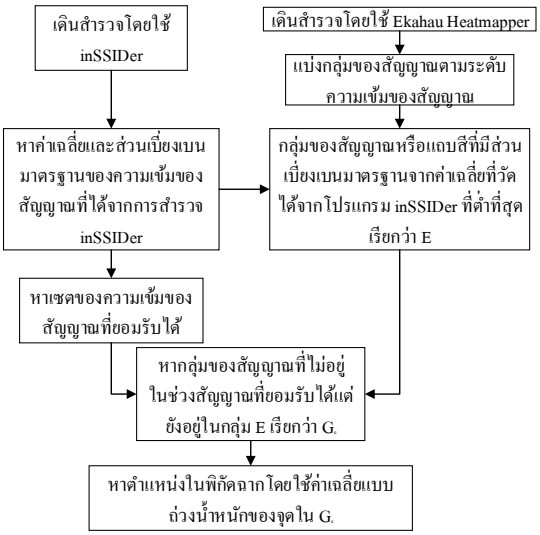
$$X = \frac{\sum_{i=1}^{N} I_i X_i}{N}, \ I_i \in G_e \quad (4)$$

$$Y = \frac{\sum_{i=1}^{N} I_i Y_i}{N}, \ I_i \in G_e \quad (5)$$

โดย

X_{i} และ Y_{i} คือตำแหน่งในพิกัดฉากของสัญญาณที่ i

จากนั้นทำการเปลี่ยนจุดเชื่อมต่อสัญญาณไร้สายแล้วหาระดับสัญญาณใหม่อีกครั้งเพื่อตรวจสอบระดับสัญญาณที่ได้เทียบกับจุดเดิมและจุดที่เป็นค่าเฉลี่ยของสัญญาณทั้งหมด ซึ่งขั้นตอนทั้งหมดสามารถเขียนเป็นแผนผังได้ดังภาพที่ 1



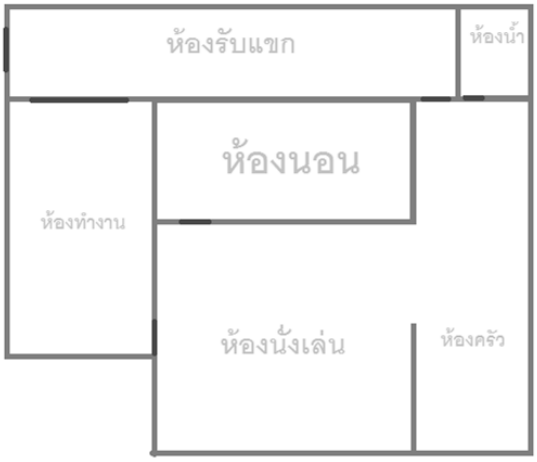
ภาพที่ **1**: แผนผังของการทำงานของวิธีที่นำเสนอ

3. วิธีการดำเนินงาน

การกำหนดตำแหน่งทางกายภาพของจุดเชื่อมต่อแบบไร้สายโดยใช้แอปพลิเคชันสร้างแผนที่ระดับความเข้มของสัญญาณมีขั้นตอนการดำเนินงานดังนี้

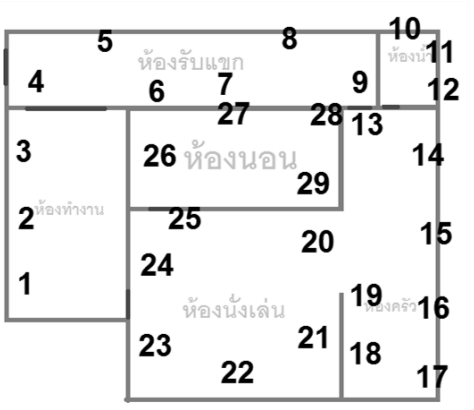
3.1 กำหนดตำแหน่งบริเวณที่ต้องการสำรวจ

ในขั้นตอนนี้ต้องทำแผนผังบริเวณที่ต้องการจะสำรวจโดยจำลองใช้พื้นที่บ้านพักแห่งหนึ่งดังภาพที่ 2 ซึ่งมีจุดเชื่อมต่อแบบไร้สายหนึ่งจุด



ภาพที่ **2**: แผนผังแสดงบริเวณที่ต้องการจะสำรวจ

หลังจากนั้นทำการกำหนดตำแหน่งที่จะเดินสำรวจสัญญาณทั้งหมด 29 ตำแหน่ง ดังภาพที่ 3



ภาพที่ **3**: ตำแหน่งบริเวณที่ต้องการสำรวจ

3.2 สำรวจสัญญาณโดยใช้โปรแกรม inSSIDer

ขั้นตอนนี้ต้องทำการสำรวจสัญญาณโดยใช้โปรแกรม inSSIDer จะได้ค่ามาในรูปแบบของ dBm

3.3 สำรวจสัญญาณโดยใช้โปรแกรม ekahau

ขั้นตอนนี้ต้องทำการสำรวจสัญญาณโดยใช้โปรแกรม Ekahau Heatmapper จะออกมาในรูปแบบของระดับความเข้มของสัญญาณเป็นสี โดยบริเวณที่มีสีเข้มจะมีระดับความเข้มของสัญญาณสูงกว่าบริเวณอื่น ๆ กล่าวคือของบริเวณที่ใกล้กับจุดเชื่อมต่อ (Access Point) จะมีระดับความเข้มของสัญญาณสูงซึ่งเป็นสีเขียวเข้ม และจะลดลงเมื่ออยู่ห่างจากบริเวณจุดเชื่อมต่อ

3.4 หาตำแหน่งที่เหมาะสม

โดยใช้วิธีการคำนวณเลขคณิตถ่วงน้ำหนัก และวิธี LUSSR แล้วนำภาพการกระจายสัญญาณของจุดเชื่อมต่อไร้สายของทั้งสองวิธีมาเปรียบเทียบกัน โดยอาศัยการวัดค่าการกระจายสัญญาณจากโปรแกรม Ekahau Heatmapper ทั้งนี้กำหนดค่าที่ไม่เจาะจง (a) ไว้ที่ 0.25 เพื่อต้องการหาบริเวณที่มีสัญญาณอยู่ระดับใกล้ค่าเฉลี่ยมากที่สุด

4. ผลการดำเนินงาน

จากวิธีดำเนินการจะได้ผลการทดลองดังนี้

4.1 การกระจายสัญญาณเมื่อกำหนดตำแหน่งจุดเชื่อมต่อไร้สายอยู่ที่จุด (0, 0) หรือตำแหน่งที่ 1

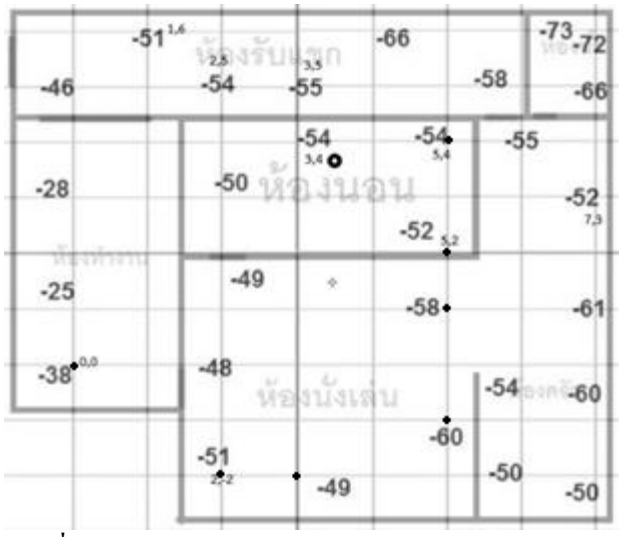
จากการสำรวจหาประสิทธิภาพความเข้มของสัญญาณเครือข่าย โดยใช้ค่า dBm ในโปรแกรม inSSIDer และได้บริเวณ

ความเข้มของสัญญาณจากโปรแกรม Ekahau Heatmapper เมื่อนำภาพทั้งสองโปรแกรมมาซ้อนทับกันดังแสดงในภาพที่ 4 ซึ่งจะพบว่าความเข้มของสัญญาณนั้นยังกระจายไปไม่ทั่วพื้นที่



ภาพที่ 4: การกระจายสัญญาณเมื่อกำหนดตำแหน่งจุดเชื่อมต่อไร้สายอยู่ที่จุด (0, 0) หรือตำแหน่งที่ 1

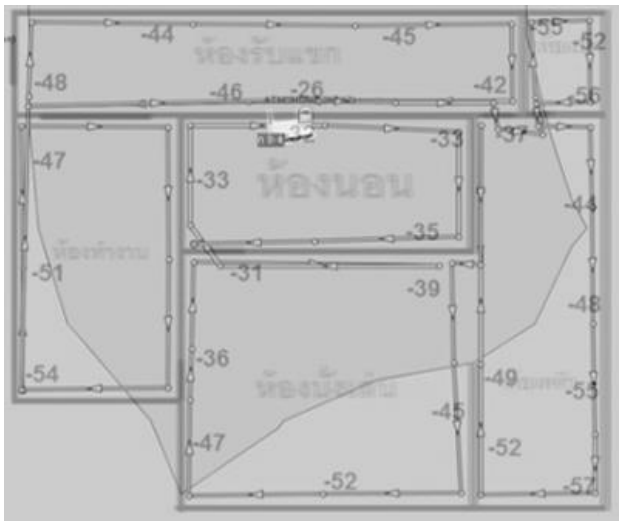
4.2 การกระจายสัญญาณเมื่อกำหนดตำแหน่งจุดเชื่อมต่อไร้สายอยู่ที่ตำแหน่งที่เกิดจากการคำนวณด้วยค่าเฉลี่ยเลขคณิตถ่วงน้ำหนักของความเข้มของสัญญาณ



ภาพที่ 5: ตำแหน่งที่เหมาะสมในการติดตั้งจุดเชื่อมต่อไร้สายโดยคำนวณจากค่าเฉลี่ยเลขคณิตถ่วงน้ำหนัก

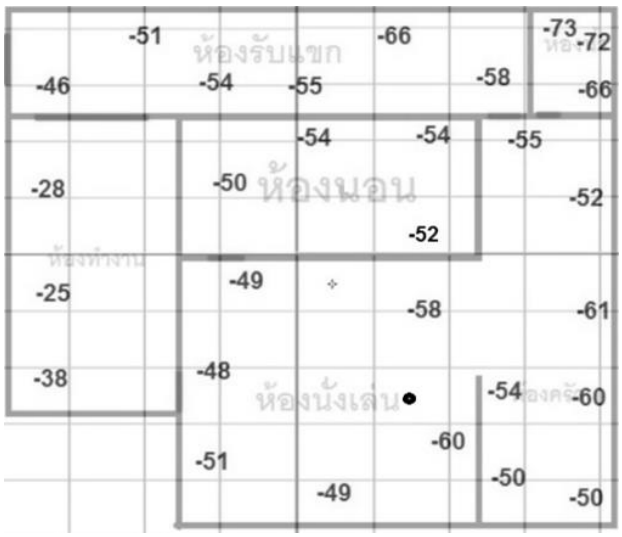
จากผลที่ได้ข้อ 4.1 สัญญาณยังกระจายไม่ดีมากนัก จึงใช้วิธีการคำนวณตำแหน่งที่เหมาะสมโดยอาศัยการคำนวณด้วยค่าเฉลี่ยเลขคณิตถ่วงน้ำหนักของความเข้มของสัญญาณ ซึ่งจะได้ผลดังแสดงในรูปที่ 5 ซึ่งก็คือจุด (3, 4)

และเมื่อนำโปรแกรม Ekahau Heatmapper สำรวจพื้นที่ครอบคลุมของสัญญาณ จะได้ผลดังแสดงในภาพที่ 6 ซึ่งสัญญาณครอบคลุมพื้นที่ส่วนใหญ่มากยิ่งขึ้น ซึ่งดีขึ้นกว่าในข้อ 4.1



ภาพที่ 6: บริเวณความเข้มของสัญญาณเมื่อกำหนดตำแหน่งจุดเชื่อมต่อไร้สายจากค่าเฉลี่ยเลขคณิตถ่วงน้ำหนัก

4.3 การกระจายสัญญาณเมื่อกำหนดตำแหน่งจุดเชื่อมต่อไร้สายอยู่ที่ตำแหน่งที่เกิดจากคำนวณด้วยวิธี PL-USSR

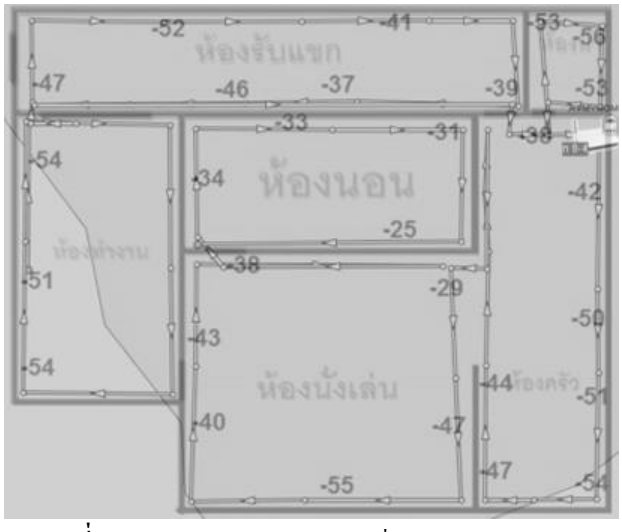


ภาพที่ 7: ตำแหน่งที่เหมาะสมในการติดตั้งจุดเชื่อมต่อไร้สายโดยคำนวณจากวิธี PL-USSR

จากผลที่ได้ข้อ 4.2 สัญญาณยังกระจายไม่ทั่ว ซึ่งพบว่ายังมีจุดที่ยังไม่ครอบคลุมซึ่งเป็นปัญหาอยู่ ดังนั้นจึงใช้วิธีการ

คำนวณตำแหน่งที่เหมาะสมโดยอาศัยการคำนวณด้วยค่าเฉลี่ยเลขคณิตถ่วงน้ำหนักของความเข้มของสัญญาณ ซึ่งจะได้ผลดังแสดงในภาพที่ 7 ซึ่งก็คือจุด (4.5, -0.6)

และเมื่อนำโปรแกรม Ekahau Heatmapper สำรวจพื้นที่ครอบคลุมของสัญญาณ จะได้ผลดังแสดงในภาพที่ 8 ซึ่งสัญญาณครอบคลุมพื้นที่ส่วนใหญ่ของทั้งหมด ซึ่งดีขึ้นกว่าในข้อ 4.1 และข้อ 4.2



ภาพที่ 8: บริเวณความเข้มของสัญญาณเมื่อกำหนดตำแหน่งของจุดเชื่อมต่อไร้สายจากวิธี PL-USSR

5. สรุปผล

จากการทดลองการกำหนดตำแหน่งทางกายภาพของจุดเชื่อมต่อแบบไร้สายโดยใช้แอปพลิเคชันสร้างแผนที่ระดับความเข้มของสัญญาณนั้น ผลลัพธ์ที่ได้ออกมาสามารถสรุปได้คือ สมการที่ใช้หาค่าเฉลี่ย กับส่วนเบี่ยงเบนมาตรฐานในกรณีแรกนั้นสามารถหาตำแหน่งของการวางตัว Access Point ได้ส่วน Signal Strength ยังอยู่ในระดับที่ยังไม่น่าพอใจ แต่สมการในกรณีที่สองนั้น สามารถหาตำแหน่งของการวางตัว Access Point ได้ดีกว่าสมการในกรณีที่หนึ่ง เพราะตัว Signal Strength ที่สำรวจได้จากการวาง Access Point ในตำแหน่งที่หาได้จากกรณีที่สองนั้น มีค่าเฉลี่ยที่ดีกว่า กรณีแรก และครอบคลุมพื้นที่ภายในบริเวณอาคารที่สำรวจได้มากกว่า

เอกสารอ้างอิง

[1] X. Zhu, Y. Feng., "*RSSI-based Algorithm for Indoor Localization*," Communications and Network, vol. 5, no. 2, pp. 37–42, 2013.

[2] Y. H. Kho, N. S. Chong, G. A. Ellis and R. C. Kizilirma., "*Exploiting RF Signal Attenuation for Passive Indoor Location Tracking of an Object*," IEEE 2015 International Conference on Computer, Communication, and Control Technology (I4CT 2015), pp. 152-156.

[3] K. Farkas, Á. Huszák and G. Gódor, "*Optimization of Wi-Fi Access Point Placement for Indoor Localization,"* Journal IIT (Informatics & IT Today), pp. 28–33, ISSN 1339--147X, 2013.

[4] H. Miao, Z. Wang, J. Wang, L. Zhang and Z. Liu., "*A Novel Access Point Selection Strategy for Indoor Location with Wi-Fi*," 26th Chinese Control and Decision Conference (CCDC), pp. 5260-5265, 2014.

[5] http://www.metageek.com/products/inssider/

[6] F. Idris, N. Hashim, "*Deployment and Evaluation of a Multi-Hop Wireless Network,"* International Journal of Computer Trends and Technology (IJCTT), Vol. 5, ISSN 2231-2803, pp. 80-83, 2013.

[7] https://www.ekahau.com/products/heatmapper/overview/

[8] ศังกร เวชพิทักษ์, ธนพล เจนสุทธิเวชกุล และเกียรติศักดิ์ โยชะนัง, The 6TH National Conference on Computing and Information Technology, NCCIT2010, pp. 285-290, 2010.

ปัจจัยที่มีผลต่อการเลือกใช้ Cloud ERP ของบริษัทและโรงงานอุตสาหกรรมในประเทศไทย
Factors effecting Cloud ERP Adoption in Manufacturing Firms in Thailand

พิสุทธิ์ หน่อแก้ว (Pisut Norkaew)[1] วัชพล สุขสกุล (Watchapon Suksakul)[2] และ บัณฑิต ฐานะโสภณ (Bundit Thanasopon)[3]
[1,2,3]ภาควิชาสารสนเทศ คณะสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
Nutpsnk@gmail.com[1], Iambank.ws@gmail.com[2], bundit@it.kmitl.ac.th[3]

บทคัดย่อ

ในปัจจุบันเทคโนโลยีมีบทบาทที่สำคัญในการเข้ามาแก้ไขปัญหาต่างๆ ที่เกิดขึ้นภายในองค์กร และเทคโนโลยี Cloud ERP ก็เป็นหนึ่งในเทคโนโลยีที่ถูกยอมรับและได้รับความนิยมในกลุ่มผู้ประกอบการ ในการเข้ามาช่วยบริหารจัดการส่วนต่างๆ ภายในองค์กร ทำให้ทางผู้วิจัยเกิดความสนใจที่จะศึกษาค้นคว้าปัจจัยที่มีผลต่อการยอมรับการใช้งานเทคโนโลยี Cloud ERP จากบริษัทและโรงงานอุตสาหกรรม โดยการเก็บข้อมูลแบบสอบถามจากกลุ่มตัวอย่างบริษัทและโรงงานอุตสาหกรรมที่ประกอบกิจการผู้ผลิตสินค้าอุปโภคและบริโภคในประเทศไทย งานวิจัยนี้ใช้การวิเคราะห์การถดถอยโลจิสติก (Logistic Regression) ในการวิเคราะห์ข้อมูลเพื่อบ่งชี้ปัจจัยสำคัญที่มีผลต่อการยอมรับ Cloud ERP ในประเทศไทย

คำสำคัญ: การยอมรับนวัตกรรม, ทฤษฎีการเผยแพร่นวัตกรรม, การประมวลผลแบบกลุ่มเมฆ, ระบบการวางแผนทรัพยากรองค์กร

Abstract

Nowadays, technology plays a key role for solving many organizational problems. Cloud ERP is one of the technologies that have been widely accepted and popular among entrepreneurs to help manage the different parts of the organization. Therefore, this research studies factors effecting the adoption of Cloud ERP in manufacturing firms. We collected data from manufacturing firms in consumer goods industry in Thailand. By using logistic regression technique, we are be able to identify important factors influencing Cloud ERP adoption in Thailand.

Keyword: Innovation adoption, Diffusion of innovation, (DOI) theory, Cloud computing, Enterprise Resource Planning systems

1. บทนำ

ในปัจจุบันความเจริญก้าวหน้าทางเทคโนโลยีสารสนเทศได้ทำให้เกิดการเปลี่ยนแปลงในทุกๆกระบวนการทำงานในส่วนต่างๆขององค์กร การนำเอาเทคโนโลยีสารสนเทศมาปรับใช้นั้นสำคัญอย่างยิ่งสำหรับการทำให้ธุรกิจสามารถอยู่รอดและเติบโตได้ หนึ่งในเทคโนโลยีที่สำคัญนั้นคือ ระบบการวางแผนทรัพยากรองค์กร (Enterprise Resource Planning: ERP) ซึ่งเป็นระบบที่ใช้ในการจัดการและวางแผนทรัพยากรต่างๆขององค์กร สามารถรวมงานหลักต่างๆขององค์กรเข้าด้วยกัน อีกทั้ง ปัจจุบันมีการพัฒนาอย่างรวดเร็วของเทคโนโลยีเข้าสู่ยุคการประมวลผลแบบกลุ่มเมฆ (Cloud Computing) ซึ่งเป็นการให้บริการบนอินเตอร์เน็ต โดยผู้ให้บริการจะมีการจัดสรรทรัพยากรคอมพิวเตอร์ต่างๆให้กับผู้ใช้บริการ โดยที่ผู้ใช้บริการไม่จำเป็นต้องลงทุนด้านฮาร์ดแวร์และซอฟต์แวร์ เพียงแค่มีเครือข่ายอินเตอร์เน็ตก็สามารถเข้าถึงบริการได้

จากการสำรวจมูลค่าตลาดซอฟต์แวร์และบริการซอฟต์แวร์ในไทย โดย SIPA และ IMC Institution [1] แสดงให้เห็นว่าในปี 2557 ตลาดมีการเติบโตร้อยละ 9.4 และมีการเติบโตในปี 2558 ร้อยละ 1.2 อีกทั้งได้มีการคาดการณ์ การเติบโตของตลาดในปี 2559 และ 2560 อยู่ร้อยละ 4.4 และ 4.3 ตามลำดับ ซึ่งแสดงให้เห็นว่ามีแนวโน้มที่ผู้บริโภคซอฟต์แวร์จะเปลี่ยนจากการใช้ซอฟต์แวร์สำเร็จรูปบนเครื่องคอมพิวเตอร์เป็นการใช้บริการซอฟต์แวร์บนเทคโนโลยีคลาวด์มากขึ้น

ผู้ให้บริการจึงได้มีการทำการประยุกต์ใช้ Cloud Computing เข้ากับระบบ ERP หรือเรียกว่า Cloud ERP เพื่อเข้ามาช่วยแก้ปัญหาต่างๆ และทำให้ระบบมีประสิทธิภาพที่ดีขึ้น ซึ่งในการใช้งาน Cloud ERP ยังมีข้อจำกัดคือจำเป็นต้องเชื่อมต่อกับอินเตอร์เน็ตทุกครั้ง แต่อย่างไรก็ตาม ตลาดของ Cloud Computing มีการเติบโตและเข้ามามีบทบาทต่อตลาดไอทีในประเทศไทยมากขึ้น เพื่อตอบสนองความต้องการของผู้ใช้งาน Cloud ในประเทศไทยที่มีแนวโน้มจะเพิ่มสูงขึ้นอย่างต่อเนื่อง

ดังนั้น การศึกษาเรื่องปัจจัยที่มีผลต่อการเลือกใช้ Cloud ERP ของบริษัทและโรงงานอุตสาหกรรมในประเทศไทย จะทำให้สามารถนำผลที่ได้จากการวิจัยไปใช้เป็นแนวทางในการตัดสินใจเลือกใช้ Cloud ERP ขององค์กร และผู้ให้บริการ Cloud ERP และนำข้อมูลที่ได้ไปพัฒนาให้เกิดการยอมรับการใช้งานให้เพิ่มขึ้นในอนาคตได้อีกด้วย

2. ทฤษฎีที่เกี่ยวข้องและทบทวนวรรณกรรม

2.1 การประมวลผลแบบกลุ่มเมฆ (Cloud Computing)

การประมวลผลแบบกลุ่มเมฆ [2] (Cloud Computing) เป็นการย้ายข้อมูลและระบบประมวลผลจากคอมพิวเตอร์ตั้งโต๊ะและพกพา ไปยังศูนย์กลางข้อมูลขนาดใหญ่และมุ่งเน้นการใช้งานบนระบบอินเตอร์เน็ตซึ่งมีความยืดหยุ่นสูง เชื่อมต่อการทำงานระยะทางไกลได้อย่างสะดวก โดยไม่จำเป็นต้องจัดเตรียมอุปกรณ์ต่างๆ และไม่จำเป็นต้องมีความรู้เชิงลึกในการทำงานก็สามารถใช้งานได้ ในกลุ่มผู้สำรวจอุตสาหกรรมกล่าวว่า Cloud computing มีการทำการตลาดที่มีศักยภาพ โดย IDC ได้คาดการณ์ไว้ว่าเทคโนโลยีนี้ จะเติบโตมากขึ้น

2.2 ระบบบริหารวางแผนทรัพยากรองค์กรบนกลุ่มเมฆ (Cloud ERP)

Cloud ERP คือ ระบบบริหารวางแผนทรัพยากรองค์กร ซึ่งใช้งานบนระบบของกลุ่มเมฆ ที่สามารถจัดการธุรกิจให้เกิดความยืดหยุ่นมากขึ้น สามารถลดค่าใช้จ่ายที่จำเป็นในการซื้ออุปกรณ์ฮาร์ดแวร์ เซิร์ฟเวอร์ ซอฟต์แวร์ การบำรุงรักษาตลอดการใช้งาน และถือเป็นพัฒนาการในทางไอทีที่ทำให้เกิดคุณสมบัติใหม่ๆ เกิดขึ้น เพื่อให้ตอบสนองต่อการทำธุรกิจและเกิดประโยชน์สูงสุดกับองค์กร

2.3 แนวคิดและทฤษฎีเกี่ยวกับการยอมรับ

ในการศึกษาวิจัยครั้งนี้ ได้สร้างกรอบแนวคิดโดยนำทฤษฎีการแพร่กระจายของนวัตกรรม (Diffusion of Innovation Theory) ของ Rogers เพื่อนำมาอธิบายการยอมรับนวัตกรรมการเลือกใช้ Cloud ERP ของบริษัทและโรงงานอุตสาหกรรมในประเทศไทย โดยสามารถอธิบายการยอมรับได้ดังนี้ [3]

1. คุณลักษณะประโยชน์เชิงเปรียบเทียบ (Relative Advantage) หมายถึง การที่บุคคลหรือองค์กรรับรู้ว่านวัตกรรมนั้นเป็นประโยชน์มากกว่าสิ่งเดิมที่ได้ใช้อยู่ ยิ่งนวัตกรรมนั้นประโยชน์มากเท่าไหร่ก็จะส่งผลให้เกิดการยอมรับมากขึ้น

2. คุณลักษณะความเข้ากันได้ (Compatibility) หมายถึง ความสอดคล้องกับการทำงานแบบเดิม ประสบการณ์ที่มีอยู่ในตัวบุคคล ซึ่งถ้ามีความสอดคล้องกับที่มีอยู่เดิมมากเท่าไหร่ ก็จะทำให้เกิดการยอมรับได้ง่ายขึ้น

3. คุณลักษณะความซับซ้อน (Complexity) หมายถึง นวัตกรรมที่ได้รับการยอมรับและเกิดความเชื่อ โดยนวัตกรรมที่มีความซับซ้อนน้อย ใช้งานง่าย จะได้รับการยอมรับที่มากกว่า นวัตกรรมที่ใช้งานยาก

4. คุณลักษณะความสามารถนำไปทดลองใช้ (Trialability) หมายถึง นวัตกรรมที่สามารถนำมาทดลองใช้ได้เพื่อให้เห็นผลได้จริง ทำให้เกิดอัตราการรับนวัตกรรมที่สูง

5. คุณลักษณะสามารถสังเกตได้ (Observability) หมายถึง นวัตกรรมนั้นสามารถสังเกตเห็นผลได้จากผู้อื่น ยิ่งถ้าสังเกตเห็นได้ชัดเจน จะทำให้เกิดการยอมรับนวัตกรรมนั้น

2.4 สมมติฐานของงานวิจัย

1. ปัจจัยด้านค่าใช้จ่าย เป็นปัจจัยที่มีความสำคัญต่อการตัดสินใจเลือกใช้ Cloud ERP ซึ่งระบบ On-premise ERP ต้องใช้เงินลงทุนทางด้านการเงิน และด้านทรัพยากรมนุษย์อย่างมากในการติดตั้ง ดูแลรักษา และพัฒนาระบบ ดังนั้น Cloud ERP จึงถูกนำมาใช้ในการแก้ไขปัญหาต่างๆ [4] โดยการประหยัดค่าใช้จ่ายด้านทรัพยากรต่างๆ จะเป็นแรงผลักดันในการนำเทคโนโลยี Cloud ERP มาใช้งาน [8] และยังสามารถลดค่าใช้จ่ายที่ไม่จำเป็นในการก่อตั้งบริษัทขนาดเล็ก ที่เพิ่งเริ่มต้นทำกำไร และนั่นทำให้บริษัทเกิดผลกำไร ที่ดีตั้งแต่การก่อตั้งบริษัทในช่วงต้น [13]

H_1 : ปัจจัยด้านค่าใช้จ่าย มีอิทธิพลเชิงบวกต่อการตัดสินใจเลือกใช้ Cloud ERP ของบริษัทในประเทศไทย

2. ปัจจัยด้านโครงข่ายอินเตอร์เน็ต ส่งผลต่อการเลือกใช้ Cloud ERP เนื่องจากการทำงานต่างๆ ของระบบจำเป็นต้องอาศัยการทำงานผ่านโครงข่ายอินเตอร์เน็ตที่รวดเร็ว คล่องตัว มีประสิทธิภาพ และจำเป็นต้องเชื่อมต่ออินเตอร์เน็ตตลอดเวลา แม้ว่าปัจจุบันประเทศไทยมีการพัฒนาระบบอินเตอร์เน็ตมากขึ้นแล้ว แต่บางพื้นที่ต่างจังหวัดก็ยังคงขาดความเสถียรของอินเตอร์เน็ตอยู่ ทำให้เกิดความกังวลต่อการตัดสินใจของผู้ใช้งานได้ [6] และประสิทธิภาพของงานที่ต้องการการคำนวณมากๆ ของ Cloud ERP มีการคาดหวังว่าจะดีกว่าระบบ On-premise ERP [4] แต่อย่างไรก็ตาม ผู้เขียนมีความเห็นว่าปัญหาด้านโครงข่ายอินเตอร์เน็ตในประเทศไทยมีความเป็นไปได้ว่าจะส่งผลให้ Cloud ERP ไม่ได้รับการยอมรับเท่าที่ควร

H_2 : ปัจจัยด้านโครงข่ายอินเตอร์เน็ต มีอิทธิพลเชิงลบต่อการตัดสินใจเลือกใช้ Cloud ERP ของบริษัทในประเทศไทย

3. ความเข้ากันได้ แสดงถึง การที่ระบบการทำงานรูปแบบเดิมที่สามารถทำงานร่วมกับระบบใหม่ๆได้ ยิ่งถ้าระบบใหม่สามารถเข้ากันกับระบบเดิมได้มาก ก็ยิ่งเกิดการยอมรับการทำงานนั้นง่ายขึ้น ซึ่งการที่องค์กรเลือกใช้ ERP package ในการเปลี่ยนแปลงองค์กรให้ดีขึ้นนั้น นวัตกรรมที่องค์กรนำมาใช้ ต้องมีความสอดคล้องกับรูปแบบการทำงานที่เป็นอยู่จะช่วยให้เกิดการยอมรับที่ง่ายขึ้น [7]

H_3 : ปัจจัยด้านความเข้ากันได้ มีอิทธิพลเชิงบวกต่อการตัดสินใจเลือกใช้ Cloud ERP ของบริษัทในประเทศไทย

4. ค่านิยมขององค์กร จากการกดดันจากคู่แข่งทางการตลาดซึ่งมีการแข่งขันที่สูงมาก ทำให้บริษัทต้องมีการพัฒนาตัวเองอยู่เสมอ ทำให้เกิดแรงกดดันในการนำเทคโนโลยีเข้ามาใช้งานเพื่อให้เกิดการพัฒนาและก้าวขึ้นเป็นผู้นำ [8] ไม่เพียงแต่ความกดดันจากคู่แข่งเท่านั้น ความกดดันจากคู่ค้า ก็เป็นปัจจัย ที่ช่วยกระตุ้นการยอมรับการใช้งานเช่นกัน เพื่อให้องค์กรเกิดการตื่นตัวและพัฒนาอยู่เสมอ [14]

H_4 : ปัจจัยด้านค่านิยม มีอิทธิพลเชิงบวกต่อการตัดสินใจเลือกใช้ Cloud ERP ของบริษัทในประเทศไทย

5. การใช้งาน จะต้องไม่ซับซ้อน และเกิดความเข้าใจง่ายให้มากที่สุด เพื่อให้ง่ายต่อการยอมรับเทคโนโลยี ซึ่งการใช้ Cloud ERP สามารถที่จะเข้าถึงการใช้งานได้ทุกที่ ผ่านทาง Web browsers และโทรศัพท์มือถือ ซึ่งแตกต่างจาก On-premise ERP โดยการเข้าถึงที่ดีขึ้นทำให้ผู้บริหารสามารถเข้ามาตรวจสอบและอัพเดทข้อมูลสำคัญได้อย่างทันที [4] และ Cloud ERP ช่วยให้การสื่อสารระหว่างองค์กรกับลูกค้า พนักงาน และผู้ส่งสินค้ามีความคล่องตัวมากขึ้น [9]

H_5 : ปัจจัยด้านการใช้งาน มีอิทธิพลเชิงบวกต่อการตัดสินใจเลือกใช้ Cloud ERP ของบริษัทในประเทศไทย

6. การขาดทักษะความรู้ในการเลือกใช้งาน Cloud ERP เนื่องจาก Cloud ERP ยังเป็นระบบที่ค่อนข้างใหม่ ผู้ใช้งานควรมีทักษะความรู้ ความเข้าใจ พื้นฐานทางเทคโนโลยีที่ดีที่เพียงพอต่อการใช้งาน โดยจากการสำรวจบริษัท SMEs 1 ใน 3 พบว่าบุคคลากรในบริษัทยังขาดทักษะความรู้ ความเข้าใจในความหมาย และประโยชน์จากการใช้งาน Cloud ERP ที่ถูกต้อง ที่ทำให้เกิดความเข้าใจที่คลาดเคลื่อนต่อเทคโนโลยี [5]

H_6 : ปัจจัยด้านการขาดทักษะความรู้ มีอิทธิพลเชิงลบต่อการตัดสินใจเลือกใช้ Cloud ERP ของบริษัทในประเทศไทย

7. การทดลองใช้ เป็นสิ่งสำคัญที่ช่วยให้องค์กรมีความสนใจที่จะเลือกใช้เทคโนโลยี Cloud ERP ตัดสินใจได้ง่ายขึ้น ซึ่งกลยุทธ์ของผู้ให้บริการ Cloud ERP มักจะมุ่งเน้นไปที่ การนำเสนอ Feature ใหม่ๆ และการเพิ่มการสนับสนุนของระบบจากผู้ให้บริการสู่ผู้ใช้บริการ [10] และ การนำมาทดลองใช้กับสภาพแวดล้อมการทำงาน ก่อนที่องค์กรจะนำมาใช้จริง [7]

H_7 : ปัจจัยด้านการทดลองใช้ มีอิทธิพลเชิงบวกต่อการตัดสินใจเลือกใช้ Cloud ERP ของบริษัทในประเทศไทย

8. ความปลอดภัยเป็นสิ่งที่หลายบริษัทให้ความสำคัญ โดยความกังวลของบริษัท ในเรื่องของสถานที่ที่ใช้ในการเก็บข้อมูลจากผู้ให้บริการ และความเป็นส่วนตัวของข้อมูล เนื่องจาก ผู้ให้บริการอาจมีการเปลี่ยนแปลงสถานที่เก็บข้อมูลและกฎหมายการป้องกันข้อมูลแต่ละประเทศที่แตกต่างกัน [4] และ ความกังวลเกี่ยวกับความปลอดภัยต่อผู้ให้บริการ ซึ่งบริษัทมองว่าเป็นสิ่งที่นอกเหนือการควบคุม ทำให้ยากต่อการยอมรับเทคโนโลยีเข้ามาใช้งาน [8]

H_8 : ปัจจัยด้านความปลอดภัย มีอิทธิพลเชิงลบต่อการตัดสินใจเลือกใช้ Cloud ERP ของบริษัทในประเทศไทย

9. ผู้บริหารระดับสูง เป็นบุคคลที่สำคัญในการขับเคลื่อนและตัดสินใจขององค์กร โดยการสร้างวิสัยทัศน์ที่ดีต่อการนำเทคโนโลยีเข้ามาใช้ เพื่อให้พนักงานเกิดทัศนคติที่ดี และถ่ายทอดส่งต่อไปยังระดับต่างๆ ภายในองค์กรได้ [11]

H_9 : ปัจจัยด้านผู้บริหารระดับสูง มีอิทธิพลเชิงบวกต่อการตัดสินใจเลือกใช้ Cloud ERP ของบริษัทในประเทศไทย

10. การตลาดเป็นการกระตุ้นให้ลูกค้าเกิดความสนใจและอยากใช้สินค้านั้น โดยผู้ให้บริการพยายามที่จะสร้างแรงจูงใจโดยการโฆษณาสินค้าบนอินเตอร์เน็ตและนิตยสาร ซึ่งทำให้ผู้บริโภคเกิดความสนใจและอยากใช้ในสินค้ามากขึ้น [12]

H_{10} : ปัจจัยด้านการตลาด มีอิทธิพลเชิงบวกต่อการตัดสินใจเลือกใช้ Cloud ERP ของบริษัทในประเทศไทย

3. วิธีดำเนินการวิจัย

วิจัยนี้เป็นการวิจัยเชิงสำรวจ โดยใช้แบบสอบถามแบบปลายปิด เป็นเครื่องมือในการเก็บรวบรวมข้อมูล ซึ่งกลุ่มตัวอย่างคือ บริษัทและโรงงานอุตสาหกรรมที่เกี่ยวข้องกับสินค้าอุปโภคหรือบริโภคที่ตั้งอยู่ในประเทศไทย เนื่องจากผู้วิจัยเห็นว่ากลุ่มผู้ประกอบการเกี่ยวกับสินค้าอุปโภคและบริโภคนั้นมีความหลากหลายและมีจำนวนมากทั่วประเทศไทย ซึ่งมีอัตราการตอบรับ ดังนี้ 1.นิคมอุตสาหกรรมลาดกระบัง จำนวน 27 ชุด ตอบกลับจำนวน 21 ชุด 2.นิคมอุตสาหกรรมบางพลี จำนวน 20 ชุด ตอบกลับจำนวน 14 ชุด 3.นิคมอุตสาหกรรมบางชัน จำนวน 15 ชุด ตอบกลับจำนวน 10 ชุด และ 4. พื้นที่อื่นๆในประเทศไทย จำนวน 150 ชุด ตอบกลับจำนวน 25 ชุด โดยหลังจากการเก็บรวบรวมข้อมูลมีแบบสอบถามที่ใช้ในการวิเคราะห์ได้ทั้งหมดจำนวน 69 ชุด จากนั้นนำมาประมวลผล โดยใช้โปรแกรมสำเร็จรูป SPSS

4. การวิเคราะห์ข้อมูล

4.1 การวิเคราะห์ความเชื่อมั่นของเครื่องมือ

จากการทดสอบความเชื่อมั่นของเครื่องมือที่ใช้ในการวัดค่าตัวแปร (Reliability of Instruments) พบว่าค่า Cronbach's Alpha ของแต่ละตัวแปรที่นำมาวิเคราะห์นั้น ไม่มีค่าความเชื่อมั่นของตัวแปรใดที่มีค่าต่ำกว่า 0.7 ซึ่งแสดงว่าตัวแปรของแบบสอบถามสอดคล้องกัน และมีความเชื่อมั่นที่ยอมรับได้ สามารถนำไปใช้เป็นเครื่องมือในงานวิจัยต่อไปได้

4.2 การทดสอบสมมติฐานเบื้องต้นของการวิเคราะห์การถดถอยโลจิสติก

การวิเคราะห์ค่าสัมประสิทธิ์สหสัมพันธ์ (Correlation) ของปัจจัยทั้ง 10 ตัวมีค่าอยู่ระหว่าง -0.028 ถึง 0.729 ซึ่งไม่มีคู่ตัวแปรใดที่มีค่า Pearson's Correlation เกิน 0.8 อีกทั้งไม่มีปัจจัยด้านใดที่มีค่า VIF เกิน 5 แสดงให้เห็นว่าตัวแปรอิสระไม่มีความสัมพันธ์กันเอง จึงสามารถสรุปได้ว่าไม่พบปัญหาความสัมพันธ์ระหว่างตัวแปรอิสระ (Multicollinearity)

นอกจากนั้นเรายังได้ทำการการตรวจสอบเพื่อหากรณีที่ผิดปกติ (Outliers or Influential Cases) พบว่ามีชุดข้อมูลที่น่าจะมีปัญหาดังกล่าว 1 ชุด เนื่องจากการตรวจสอบค่า Mahalanobis Distance และ Stanrdized DFFit พบว่ามีค่าเกินจากที่กำหนดไปมาก และไม่เกาะกลุ่มกับแบบสอบถามชุดอื่น ดังนั้น จึงเหลือข้อมูลที่นำไปวิเคราะห์ทั้งสิ้น 68 ชุด

4.3 การวิเคราะห์การถดถอยโลจิสติกแบบไบนารี

ผลการวิเคราะห์ข้อมูลพบว่าค่า -2 Log likelihood (-2LL) มีค่าเท่ากับ 34.86 และ สมการที่สร้างขึ้นมีค่า Nagelkerke R Square เท่ากับ 0.776 จึงสามารถอธิบายความผันแปรของตัวแปรต้นได้ 77.6%

จากตารางที่ 1 ผลการวิเคราะห์ความถดถอยโลจิสติก พบว่ามี 3 ปัจจัยที่ส่งผลตรงตามสมมติฐานที่ตั้งไว้ และอีก 1 ปัจจัยที่ส่งผลกระทบในทิศทางที่ตรงกันข้ามกับสมมติฐานที่ตั้งไว้ โดยค่า Sig. มีค่าน้อยกว่า 0.05 โดยสามารถอธิบายรายละเอียดของแต่ละปัจจัยที่มีนัยสำคัญได้ดังนี้

สำหรับสมมติฐานที่ 2 (โครงข่ายอินเตอร์เน็ต) จากการวิเคราะห์ พบว่าค่า β = -3.125 และค่า Sig. = 0.049 น้อยกว่า 0.05 ดังนั้นจึงยอมรับสมมติฐานที่ 2 แสดงว่า ปัจจัยด้านโครงข่ายอินเตอร์เน็ตมีความสัมพันธ์ต่อการเลือกใช้ Cloud ERP และสมมติฐานที่ 5 (การใช้งาน) จากการวิเคราะห์ พบว่าค่า β = 3.137 และค่า Sig. = 0.014 น้อยกว่า 0.05 ดังนั้นจึงยอมรับสมมติฐานที่ 5 แสดงว่า ปัจจัยด้านการใช้งานมีความสัมพันธ์เชิงบวกต่อการเลือกใช้ Cloud ERP อีกทั้งสมมติฐานที่ 6 (การขาดทักษะความรู้) จากการวิเคราะห์ พบว่าค่า β = 3.371 และค่า Sig. = 0.011 น้อยกว่า 0.05 ดังนั้น

ถึงแม้ว่าปัจจัยการขาดทักษะความรู้จะมีผลต่อการยอมรับ Cloud ERP อย่างไรก็ตามทิศทางของผลกระทบนั้นเป็นไปในทิศทางตรงกันข้ามกับสมมติฐานที่ 6 จึงปฏิเสธสมมติฐานที่ 6 และสำหรับสมมติฐานที่ 9 (ผู้บริหารระดับสูง) จากการวิเคราะห์ พบว่า ค่า β = 3.954 และค่า Sig. = 0.008 น้อยกว่า 0.05 ดังนั้นจึงยอมรับสมมติฐานที่ 9 แสดงว่าปัจจัยด้านผู้บริหารระดับสูงมีความสำคัญต่อการเลือกใช้ Cloud ERP ของบริษัทและโรงงานอุตสาหกรรมในประเทศไทย

ตารางที่ 1: แสดงผลการวิเคราะห์การถดถอยโลจิสติกแบบไบนารี

	B	Sig.	Exp(B)
Expenses	1.053	.370	2.867
Network infrastructure	-3.125	.049*	.044
Compatability	2.330	.124	10.278
Values	1.566	.219	4.789
Usage	3.137	.014*	23.040
Lack of Knowledge	3.371	.011*	29.102
Trialability	-2.376	.234	.093
Security	.816	.491	2.262
Top Management	3.954	.008*	52.125
Marketing	-2.565	.128	.077
Constant	-27.953	.017	.000

* มีนัยสำคัญทางสถิติที่ระดับ 0.05

5. สรุปผลการวิจัย และข้อเสนอแนะ

5.1 สรุปผลการวิจัย

การวิจัยครั้งนี้ ผู้จัดทำได้นำข้อมูลที่เก็บได้จากกลุ่มตัวอย่างบริษัทและโรงงานอุตสาหกรรมที่เกี่ยวข้องกับสินค้าอุปโภคหรือบริโภคที่ตั้งอยู่ในประเทศไทย จำนวน 68 ชุด โดยผลการวิเคราะห์ข้อมูลด้วยเทคนิคการถดถอยโลจิสติกแบบไบนารี สรุปได้ว่า มี 3 ปัจจัยซึ่งส่งผลตรงตามสมมติฐานที่ตั้งไว้ ได้แก่ ปัจจัยด้านโครงข่ายอินเตอร์เน็ต ปัจจัยด้านการใช้งาน และปัจจัยด้านผู้บริหารระดับสูง และอีก 1 ปัจจัยที่พบว่ามีความสำคัญแต่ส่งผลกระทบในทิศทางที่ตรงกันข้ามกับสมมติฐานที่ตั้งไว้ ได้แก่ ปัจจัยการขาดทักษะความรู้ ซึ่งส่งผลต่อการเลือกใช้ Cloud ERP ของบริษัทและโรงงานอุตสาหกรรมในประเทศไทย

5.2 อภิปรายผลการทดลอง

งานวิจัยนี้พบว่า ปัจจัยด้านโครงข่ายอินเตอร์เน็ตมีอิทธิพลต่อการเลือกใช้ Cloud ERP อย่างไรก็ตามเราพบว่าปัจจัยนี้ส่งผลกระทบในเชิงลบต่อการยอมรับ Cloud ERP ของบริษัทและโรงงานอุตสาหกรรมในประเทศไทย โดยเหตุผลที่พบผลกระทบเชิงลบนั้นอาจเป็นเพราะ ปัจจุบันการใช้งานอินเตอร์เน็ตบางพื้นที่ในประเทศไทยนั้น ยังคงขาดความเสถียรในการใช้งาน แม้ว่าจะถูกพัฒนาอย่างต่อเนื่องก็ตาม แต่ก็ยังคงไม่สามารถทำให้ผู้บริโภคเชื่อถือได้มากพอ [6] ซึ่งสอดคล้องกับงานวิจัยของ Guo Chao และ Alex Peng (2014) [4] กล่าวว่า การตอบสนองของระบบ Cloud ERP ถูกคาดหวังว่าจะมีการทำงานดีที่ขึ้นกว่าระบบ On-premise ERP ซึ่งในประเทศไทยนั้น Cloud ERP อาจยังไม่สามารถสร้างการยอมรับได้ตามที่คาดหวังไว้

เรายังพบว่าปัจจัยด้านการใช้งาน มีอิทธิพลเชิงบวกต่อการเลือกใช้ Cloud ERP ของบริษัทและโรงงานอุตสาหกรรมในประเทศไทย โดยอาจเนื่องมากจาก การที่จะนำนวัตกรรมมาประยุกต์ใช้งาน จะต้องมีประสิทธิภาพที่ดีไม่ซับซ้อน และเกิดความเข้าใจง่ายให้มากที่สุด ซึ่งระบบ On-premise ERP มีข้อจำกัดในเรื่องของความซับซ้อน นอกจากนั้นการใช้ Cloud ERP สามารถเข้าถึงการใช้งานได้ทุกที่ ซึ่งแตกต่างจาก On-premise ERP ที่ต้องใช้อุปกรณ์ในการเข้าถึงเมื่ออยู่ต่างสถานที่ ซึ่งทำให้ผู้บริหารสามารถตรวจสอบและอัพเดทข้อมูลสำคัญได้ทุกที่ทุกเวลา [4] และยังสามารถปรับปรุงเรื่องการโต้ตอบสื่อสารระหว่างลูกค้า พนักงาน และ Suppliers ได้ง่ายและคล่องตัวมากขึ้น [15]

เรายังพบอีกว่าการสนับสนุนของผู้บริหารระดับสูงมีอิทธิพลอย่างมากต่อการเลือกใช้ Cloud ERP ของบริษัทในประเทศไทย ซึ่งสอดคล้องกับ Hervé Allart (2014) [10] ได้แนะนำไว้ว่าบริษัทจะต้องมีผู้บริหารระดับสูงที่คอยเตรียมการอนุญาต ให้ความช่วยเหลือ และ สนับสนุนให้กับผู้ใช้งาน และบทบาทของผู้บริหารระดับสูงในการสร้างวิสัยทัศน์ที่ดีต่อการนำเทคโนโลยีเข้ามาใช้ เพื่อให้พนักงานเกิดทัศนคติที่ดี และถ่ายทอดส่งต่อไปยังระดับต่างๆ ภายในองค์กรอีกด้วย [9]

ปัจจัยด้านการขาดทักษะความรู้นั้นได้ผลที่ตรงข้ามกับสมมติฐานที่ตั้งไว้ โดยการขาดทักษะความรู้เกี่ยวกับ Could

ERP นั้นส่งผลกระทบในเชิงบวกต่อการยอมรับ Cloud ERP ซึ่งคือ ยิ่งผู้ตัดสินใจมีความรู้เกี่ยวกับเทคโนโลยีน้อยก็จะยิ่งมีแนวโน้มที่จะเลือกใช้ ซึ่งตรงกันข้ามกับคำแนะนำในงานศึกษาอื่นๆ โดยอาจเป็นเพราะ บุคคลากรในบริษัทที่ยังขาดทักษะความรู้ ความเข้าใจในความหมาย และประโยชน์จากการใช้งาน Cloud ERP ที่ถูกต้อง จะทำให้เกิดความเข้าใจที่คลาดเคลื่อนต่อการใช้เทคโนโลยี [5] ซึ่งผู้จัดทำมีความเห็นว่าการที่ผู้ใช้งานไม่มีความรู้เชิงลึกเกี่ยวกับนวัตกรรมนั้น ก็อาจจะทำให้ง่ายต่อการโน้มน้าวให้ยอมรับนวัตกรรมจากคำโฆษณาของผู้ให้บริการ หรือว่าผู้ที่นำเสนอได้มากกว่าผู้ใช้งานที่มีความรู้เชิงลึก ซึ่งทำให้ได้ผลดังที่เราพบในการศึกษานี้

5.3 ข้อเสนอแนะสำหรับงานวิจัยในอนาคต

1. งานวิจัยในอนาคตสามารถศึกษาปัจจัยอื่นๆที่เกี่ยวข้องเพิ่มเติม เช่น ปัจจัยด้านประเภทขององค์กร ปัจจัยด้านเศรษฐกิจ เป็นต้น เนื่องจาก แต่ละช่วงเวลาของสภาพเศรษฐกิจมีความแตกต่างกัน ทำให้แต่ละองค์กรมีสภาพแวดล้อมที่แตกต่างกัน

2. สามารถศึกษากลุ่มตัวอย่าง หรือกลุ่มประชากรอื่นๆ ที่อยู่ในภูมิภาคอื่น หรือประเทศอื่น เพื่อที่จะนำผลการศึกษามาเปรียบเทียบกันว่าแต่ละบริบทนั้นส่งผลให้ปัจจัยสำคัญมีความเหมือนหรือต่างมากน้อยเพียงใด

เอกสารอ้างอิง

[1] “SIPA เผยผลสำรวจตลาดซอฟต์แวร์ไทย ปี 2558 มูลค่าการผลิตเติบโตร้อยละ1.2”[Online].Avaliable: http://www.aripfan.com /sipa-software-market-thai/

[2] ชุตินันท์ พึ่งเกตุ และ พีริยา จารุเศรษฐการ. “เทคโนโลยีของการประมวลผลแบบกลุ่มเมฆ”*วารสารวิชาการปทุมวัน*. 2557, ฉบับที่ 4. เดือนพฤษภาคม - สิงหาคม 2557

[3] วัชรพล คงเจริญ. “ปัจจัยการยอมรับนวัตกรรมที่มีอิทธิพลต่อความพึงพอใจของผู้ใช้บริการธนาคารผ่านโทรศัพท์เคลื่อนที่ในเขตกรุงเทพมหานคร.” *วิทยานิพนธ์บริหารธุรกิจมหาบัณฑิต*, มหาวิทยาลัยกรุงเทพ. 2558

[4] Guo Chao Alex Peng, Chirag Gala. “Cloud ERP: a New Dilemma to Modern Organisations?.”*Information School, University of Sheffield*. 2014

[5] Björn Johansson, Amar Alajbegovic, Vasileios Alexopoulos, Achilles Desalermos. “Cloud ERP Adoption Opportunities and Concerns: A Comparison between SMEs and Large Companies.” *Department of Informatics, School of Economics and Management*, Lund University, Sweden. 2014

[6] Hervé Allart. “Adoption Factors and Implementation Strategies of on-Premise and Cloud Based ERP Systems by SMEs in Thailand.” *MBA International program, Stamford International University*. June-December 2014

[7] Ben Light & Anastasia Papazafeiropoulou. “Reasona Behind ERP Package Adoption: A Diffusion of Innovations Perspective.” *University of Salford, Brunel University*. 2004

[8] Tiago Oliveira, Manoj Thomas, Mariana Espadanal. Information & Management. “Assessing the determinants of cloud computing adoption: An analysis of the manufacturing and services sectors”. 2014, Volume 51. July 2014

[9] Chinyao Low and Yahsueh Chen, Mingchang Wu. “Understanding the determinants of cloud computing adoption.” *Department of Industrial Engineering and Management, National Yunlin University of Science & Technology*. 2011

[10] Ezer Osei Yeboah-Boateng, Kofi Asare Essandoh. Emerging Trends in Computing and Information Sciences. “The Level of Awareness Amongst Small & Medium-sized Enterprises (SMEs) in Developing Economies”. 2014

[11] ภานุพงศ์ เสกทวีลาภ “ปัจจัยที่ส่งผลต่อความตั้งใจด้านพฤติกรรมการใช้ Cloud Storage ใน ระดับ Software-as-a-Service (SaaS) ของพนักงานองค์กรเอกชน ในเขตพื้นที่เศรษฐกิจของกรุงเทพมหานคร.” *วิทยานิพนธ์บริหารธุรกิจมหาบัณฑิต บัณฑิตวิทยาลัย,* มหาวิทยาลัยกรุงเทพ. 2558

[12] ณกฤช เศวตนันทน์ “ความเร็วอินเตอร์เน็ต.” [Online]. Available: http://lawyerthailand.com/ความเร็วอินเตอร์เน็ต/. 2013

[13] Maricela-Georgiana Avram. INTER-ENG. “Advantages and challenges of adopting cloud computing from an enterprise perspective”. 2014

[14] Byeong-Yun Chang, Pham Hoang Hai, Dong-Won Seo, Jong-Hun Lee, Seung Hyun Yoon. “Thedeterminant of adoption in cloud computing in Vietnam.” *Department of Business Administration, Ajou University*. 2013

[15] Ezer Osei Yeboah-Boateng, Kofi Asare Essandoh. Emerging Trends in Computing and Information Sciences. “The Level of Awareness Amongst Small & Medium-sized Enterprises (SMEs) in Developing Economies”. 2014

การกรองแพตเก็ตแบบปรับได้ด้วยการวิเคราะห์สถิติความน่าจะเป็นในระบบป้องกันการบุกรุก

Adaptive Packet Filter with Probability Statistic Analysis in Intrusion Prevention System

วีรยุทธ ศักดิ์คำดวง (Weerayuth Sukkhamduang)[1] และสายยัญ สายยศ (Saiyan Saiyod)[2]
ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น
[1]weerayuth.s@kkumail.com, [2]saiyan@kku.ac.th

บทคัดย่อ

จากปัญหาการพยายามบุกรุกเครือข่ายที่ผ่านมา ซึ่งเกิดจากผู้ไม่ประสงค์ดีนั้นต้องการที่จะทำให้เกิดความเสียหายต่อระบบเครือข่าย ด้วยเหตุนี้เราจึงต้องป้องกันเหตุการณ์ที่จะเกิดขึ้น ดังนั้นในงานวิจัยนี้จึงได้นำเสนอระบบป้องกันการบุกรุกโดยใช้ทฤษฎีความน่าจะเป็นทำงานร่วมกับระบบตรวจจับการบุกรุก ในการทดลองนั้นใช้ชุดข้อมูล (datasets) ที่สร้างขึ้นมาประกอบด้วยแพคเกตปกติ และแพคเกตที่ไม่ปกติส่งข้อมูลผ่านระบบและตัวระบบนั้นสามารถปรับระดับการกรองแพคเกตได้ โดยมีค่าน้ำหนักในการทดสอบในระดับที่แตกต่างกัน ผลการทดลองพบว่าระบบที่สร้างขึ้นนั้นมีประสิทธิภาพในการป้องกันความพยายามในการบุกรุกหรือโจมตีเครือข่ายจากชุดข้อมูลที่สร้างขึ้นมามีความถูกต้องที่สุดที่ค่าน้ำหนักเท่ากับ 5 จากผลการทดลองจึงสามารถสรุปได้ว่าอัลกอริทึมของระบบที่นำเสนอมีประสิทธิภาพ สามารถนำไปใช้งานได้จริงในสถานการณ์อื่นๆ ได้ในอนาคต

คำสำคัญ : ระบบป้องกันการบุกรุก การกรองแพตเก็ต ความน่าจะเป็น จำนวนสามเหลี่ยม

Abstract

The past problem of trying to intrusion network because intrusion person want to cause damage to network. For this reason, we must prevent against any attempted intrusion. In this research, we have introduced intrusion prevention system using intrusion detection system works with probability theory. In the experiment, use the datasets generated to test the system and the system can adjust the level of packet filtering. Contains normal package and abnormal package send data through system by weight of the test at different levels. The results showed that the system were effective in preventing intusion attempts or attack the network from the generated datasets. The best accuracy was obtained at the weight of 5. From the results, it can be concluded that the algorithms of the proposed system are effective. The proposed system can be used in other situations in the future.

Keyword : Intrusion Prevention System, Packet Filter, Probability, Triangular Number.

1. บทนำ

ระบบเครือข่ายคอมพิวเตอร์มีความสำคัญในการติดต่อสื่อสารในชีวิตประจำวันเป็นอย่างมาก ซึ่งเห็นได้จากการได้รับความนิยมไม่ว่าจะเป็นทั้งในองค์กรภาครัฐและเอกชน เช่น การให้บริการเว็บไซต์เพื่อใช้ประชาสัมพันธ์ บริการอีเมล์สำหรับการรับส่งข้อมูลภายในและภายนอกองค์กร ทำให้เกิดปัญหาที่ตามมาคือ เกิดการบุกรุก หรือการพยายามโจมตีเพื่อสร้างความเสียหายให้แก่ระบบ ซึ่งเกิดจากผู้ไม่ประสงค์ดีนั้นต้องการที่จะทำให้เกิดความเสียหายต่อระบบทำให้ข้อมูลที่เป็นความลับขององค์กรถูกเปิดเผยหรือเสียหาย

ด้วยเหตุนี้ สิ่งแรกที่เครือข่ายคอมพิวเตอร์จะต้องมี คือ ระบบรักษาความปลอดภัย เช่น ไฟร์วอลล์ (Firewall) เพื่อป้องกันไม่ให้มีผู้บุกรุกส่งชุดคำสั่งหรือแพคเกต (package) ที่ไม่ได้รับอนุญาตเข้ามาในเครือข่ายได้ แต่ในการป้องกันของไฟร์วอลล์นั้นจะทำการป้องกันการบุกรุกเครือข่ายจากชุดของกฎที่กำหนดไว้แล้ว

โดยการปิดกั้นข้อมูลจาก IP ต้นทางทั้งหมด ซึ่งในบางกรณี IP นั้นเป็นของหน่วยงาน อาจจะมีทั้งผู้บุกรุกและผู้ใช้งานปกติในเวลาเดียวกัน เมื่อเราทำการปิดกั้น IP ไปแล้ว จะส่งผลทำให้ผู้ใช้งานปกติได้รับผลกระทบไปพร้อมกัน

จากงานวิจัยที่ผ่านมานั้นจะมุ่งเน้นไปยังการพัฒนาในด้านการตรวจจับการบุกรุก และพัฒนาในการป้องกันการบุกรุกเครือข่ายเป็นหลัก [1], [2], [3] ซึ่งมีประสิทธิภาพในการตรวจจับและป้องกันการบุกรุกเป็นอย่างดี พร้อมทั้งสามารถลดภาระงานให้แก่ระบบ จึงทำให้ลดการใช้ทรัพยากรเครื่องโดยไม่จำเป็นได้ แต่ในอีกด้านหนึ่ง ยังไม่สามารถแก้ปัญหาผลกระทบที่มีต่อผู้ใช้งานปกติได้

ดังนั้น งานวิจัยนี้จึงนำเสนอวิธีการในการป้องกันการบุกรุกเครือข่ายด้วยวิธีการกรองแพคเกตที่ผ่านเข้ามาในเครือข่าย (Packet filter) โดยใช้หลักการทางสถิติ คือทฤษฎีสมการความน่าจะเป็น (Probability) ผสมผสานกับสมการจำนวนสามเหลี่ยม (Triangular number) เพื่อนำมาทำงานร่วมกับระบบตรวจจับการบุกรุก (Intrusion Detection System)

วัตถุประสงค์เพื่อสร้างระบบป้องกันการบุกรุก (Intrusion Prevention System) ที่มีประสิทธิภาพ โดยที่ผู้ใช้งานปกติได้รับผลกระทบน้อย และยังสามารถปรับ (Adaptive) ใช้ได้ในสภาพแวดล้อมเครือข่ายจริงได้ในอนาคต

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

จากการศึกษาวิเคราะห์ทฤษฎีและงานวิจัยที่เกี่ยวข้องกับระบบป้องกันการบุกรุกและหลักการทางสถิตินั้น สามารถแยกย่อยได้ดังนี้

2.1 ทฤษฎีที่เกี่ยวข้อง

ทฤษฎีที่เกี่ยวข้องนั้นประกอบไปด้วย

2.1.1 ระบบตรวจจับการบุกรุกการ (Intrusion Detection System) หรือ (IDS)

เป็นขั้นตอน หรือกระบวนการตรวจสอบเหตุการณ์ที่เกิดขึ้นบนระบบเครือข่ายคอมพิวเตอร์โดยการวิเคราะห์ข้อมูลจากแพคเก็ตที่ส่งไปมาภายในเครือข่าย แล้วนำมาตรวจสอบว่าเข้าข่ายว่าเป็นการบุกรุกหรือไม่ และแจ้งเตือนว่าเป็นรูปแบบการโจมตีประเภทใด สามารถแบ่งออกได้ดังนี้ [4]

2.1.2 ระบบป้องกันการการบุกรุก (Intrusion Prevent System)

Intrusion Prevention System (IPS) เป็นระบบที่มีความสามารถมากกว่า Intrusion Detection System (IDS) โดยมีความสามารถในการหยุดยั้งการบุกรุกหรือตอบโต้การโจมตีได้โดยอัตโนมัติทันทีที่ตรวจพบ [5]

2.1.3 สถิติ (Statistic)

สถิติ (Statistics) คือ การนำเอาข้อมูลในอดีตและปัจจุบัน มาช่วยในการคาดการณ์อนาคต จะมีข้อกำหนดที่เป็นหลักวิทยาศาสตร์ แบ่งลักษณะการนำไปประยุกต์ใช้งานได้เป็นสองลักษณะใหญ่ๆ ของแต่ละกลุ่มได้ดังต่อไปนี้ [6]

a. สถิติเชิงพรรณนา (Descriptive Statistics)

เป็นการใช้สถิติเพื่ออธิบายข้อมูลที่มีอยู่ โดยไม่ได้นำไปพยากรณ์ประชากร แต่อย่างใด

b. สถิติเชิงอนุมาน (Inferential Statistics)

เป็นการศึกษาข้อมูลของกลุ่มตัวอย่าง แล้วนำข้อมูลนั้นไปประมาณหรือคาดการณ์ประชากร หรือเหตุการณ์ที่จะเกิดขึ้นในอนาคต

2.1.4 ความน่าจะเป็น (Probability)

ความน่าจะเป็น [6] หมายถึง จำนวนที่แสดงให้ทราบว่าเหตุการณ์ใดเหตุการณ์หนึ่งมีโอกาสเกิดขึ้นมากน้อยเพียงใด สามารถเขียนเป็นสมการได้ดังนี้ (1)

$$P(E) = \frac{n}{N} \quad (1)$$

เมื่อ

E คือ เหตุการณ์ที่เราสนใจ

P(E) คือโอกาสความน่าจะเป็นของเหตุการณ์

n คือ จำนวนเหตุการณ์ที่เกิดขึ้น

N คือจำนวนเหตุการณ์ที่อาจจะเกิดขึ้นได้ทั้งหมด

ดังนั้น เมื่อเราทราบจำนวนเหตุการณ์ที่เกิดขึ้นมาแล้วย่อมสามารถหาโอกาสความน่าจะเป็นที่จะเกิดขึ้นต่อไปได้

2.1.4 จำนวนสามเหลี่ยม (Triangular number)

จำนวนสามเหลี่ยม [7] หมายถึง ผลบวกบางส่วนของอนุกรม $1 + 2 + 3 + 4 + 5 + \cdots$ คือ 1, 3, 6, 10, 15 ฯลฯ ผลบวกที่ *n* หาได้จากสมการ (2)

$$Tn = \sum_{k=1}^{n} k = \frac{n(n+1)}{2} \quad (2)$$

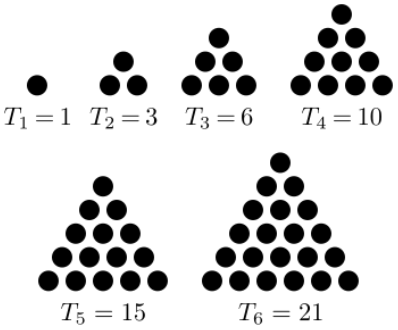
เมื่อ

Tn คือ จำนวนสามเหลี่ยม

n คือ จำนวนเต็มทั้งหมดที่จะต้องนำมาบวก

k คือ ตัวเลขแรกของอนุกรม

เป็นทฤษฎีที่คิดค้นได้ในช่วงศตวรรษที่ 6 ก่อนคริสตกาลโดยพีทาโกรัสและศิษย์ (Pythagorean) เกิดจากการสร้างสามเหลี่ยมด้านเท่า ดังที่แสดงใน ภาพที่ 1



ภาพที่ 1: จำนวนสามเหลี่ยมหกจำนวนแรก

2.2 งานวิจัยที่เกี่ยวข้อง

งานวิจัยที่เกี่ยวข้องนั้นแบ่งออกได้เป็น ด้าน Network Intrusion Detection และ Network Intrusion Prevention

2.2.1 งานวิจัยด้าน Network Intrusion Detection

Kim และ Park (2003) [1] ได้เสนอการใช้ Support Vector Machine IDS เพื่อใช้ตรวจจับการบุกรุก ซึ่งเป้าหมายของงานวิจัยนี้คือ ใช้ SVM ในแบบ Multi Class มาจำแนกประเภทได้และให้ประสิทธิภาพในส่วนของการตรวจจับ Anomaly แบบใหม่ Ren Hui Gong และคณะ [2] (2005) เสนอการนำวิธีทางพันธุกรรมในการตรวจสอบความผิดปกติของเครือข่าย (Misuse Detection) ได้ใช้ชุดข้อมูล DARPA datasets จาก MIT Lincoln Laboratory ผลการทดลองแสดงให้เห็นว่าวิธีทางพันธุกรรมเป็นวิธีที่มีประสิทธิภาพและมีความยืดหยุ่นในการตรวจสอบการบุกรุกเครือข่ายได้อย่างแม่นยำ

2.2.2 งานวิจัยด้าน Network Intrusion Prevention

Yuxin Meng และ Lam-For Kwok [3]ได้นำเสนอวิธีการในการป้องกันการบุกรุกโดย สามารถลดจำนวนแพคเกตไม่ปกติที่เข้ามาในระบบ ประกอบด้วยการทดลองจากภายในสถานที่ทำการวิจัย การใช้งาน Honeypod ดักจับข้อมูลที่มีการเชื่อมต่อเข้ามายังสถานที่ทำวิจัย และการใช้ DARPAR dataset ในการทดสอบ จากนั้นจึงใช้วิธีการวัดประสิทธิภาพโดยเปรียบเทียบระหว่าง การใช้ระบบ และการไม่ใช้ระบบ ซึ่งแสดงให้เห็นว่า สามารถลดภาระงานของระบบตรวจจับการบุกรุก รวมไปถึงการลดการใช้ทรัพยากรของเครื่องได้เป็นอย่างดี

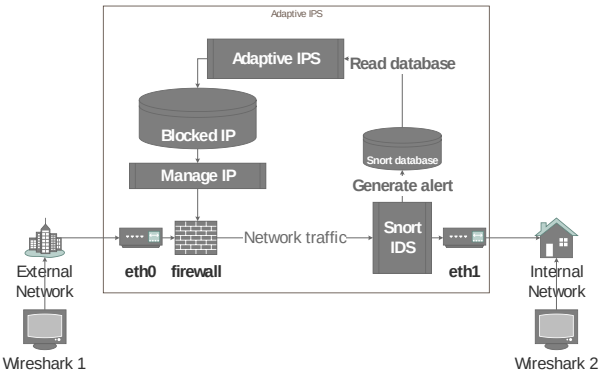
3. วิธีดำเนินการวิจัย

เนื่องจากระบบตรวจจับการบุกรุกนั้น มีความสามารถเพียงการแจ้งเตือนเมื่อตรวจพบสิ่งผิดปกติที่เข้ามาในระบบ แต่ไม่สามารถทำการป้องกันได้ ผู้วิจัยจึงต้องวิเคราะห์ว่าจะสามารถป้องกันได้อย่างไร

ซึ่งคำตอบที่ได้นั้นคือ ไฟร์วอลของระบบ แต่ไฟร์วอลนั้นยังไม่มีประสิทธิภาพเพียงพอที่จะใช้ในเป็นระบบป้องกันการบุกรุก เพราะไฟร์วอลนั้นจะทำงานได้ตามกฎที่กำหนดไว้เท่านั้น

ในทางกลับกัน เมื่อระบบทำการปิดกั้นการเชื่อมต่อโดยใช้ไฟร์วอลเพื่อความปลอดภัยจากผู้บุกรุก จะส่งผลต่อผู้ใช้งานปกติไปพร้อมกัน เช่น หากระบบนั้นใช้ไฟร์วอลปิดกั้น IP จากบริษัท ก เพื่อป้องกันจากผู้บุกรุกจำนวนน้อย แต่ผู้ใช้งานปกติอีกจำนวนมากนั้นจะต้องได้รับผลกระทบ เป็นต้น จึงทำให้ผู้วิจัยต้องวิเคราะห์ว่า จะใช้กระบวนการแบบใดจึงจะเหมาะสมในการป้องกันการบุกรุกอย่างมีประสิทธิภาพ และให้ผู้ใช้งานปกติได้รับผลกระทบน้อย

ดังนั้น งานวิจัยนี้ ผู้วิจัยต้องคำนึงถึงเป้าหมายหลักทั้งสองอย่าง ได้แก่ ความสามารถในการป้องกันการบุกรุกเครือข่ายอย่างมีประสิทธิภาพ และต้องให้ผู้ใช้งานปกติได้รับผลกระทบน้อย ผู้วิจัยจึงได้ออกแบบและสร้างระบบกรองแพคเกตขึ้นมา โดยระบบนี้จะทำงานร่วมกับระบบตรวจจับการบุกรุกและไฟร์วอลที่มีอยู่แล้ว เพื่อสร้างเป็นระบบป้องกันการบุกรุกที่มีประสิทธิภาพมากกว่าการใช้ไฟร์วอล หรือระบบตรวจจับการบุกรุกเพียงอย่างเดียว



ภาพที่ 2: ภาพรวมของระบบ

รายละเอียดแต่ละขั้นตอนของภาพที่ 2 มีดังนี้

3.1 Snort IDS จะทำการอ่านแพคเก็ตที่ผ่านเข้ามา ไม่ว่าจะเป็นแพคเก็ตปกติ หรือไม่ปกติ จากทุกๆ IP และทำการบันทึกข้อมูลของแพตเก็ตลงฐานข้อมูลลงใน Snort database

3.2 ระบบ Adaptive IPS จะต้องหาความน่าเชื่อถือของ IP ที่เชื่อมต่อเข้ามาในระบบ ด้วยวิธีการเรียกดูข้อมูลจากฐานข้อมูล Snort Database ได้แก่ IP ที่มีการเชื่อมต่อเข้ามา จำนวนแพคเกตปกติ จำนวนแพคเกตไม่ปกติของแต่ละ IP ในช่วงเวลา 1 นาทีที่ผ่านมา

ในส่วนของการตัดสินความน่าเชื่อถือของ IP นั้น หากระบบตรวจพบแพคเกตไม่ปกติเข้ามาในระบบ แล้วทำการปิดกั้นไฟร์วอลทันที จะส่งผลต่อผู้ใช้ปกติไปพร้อมกัน ผู้วิจัยจึงต้องทำการหาวิธีตัดสินว่าเครือข่ายที่มีการเชื่อมต่อเข้ามานั้นมีเจตนามุ่งร้ายหรือไม่ จากข้อมูลที่ได้รับจากระบบตรวจจับการบุกรุก ซึ่งประกอบด้วย หมายเลข IP ต้นทาง จำนวนของแพคเกตปกติและไม่ปกติของ IP ต้นทางนั้น ก่อนที่จะตัดสิน

ตัวอย่าง หากมีการส่งแพคเกตเกิดขึ้นทั้งหมด 3 ครั้ง มีแพคเกตปกติทั้งหมด 2 ครั้ง และแพคเกตไม่ปกติทั้งหมด 1 ครั้ง 1 เมื่อเลือกใช้ทฤษฎีความน่าจะเป็น (Probability) จำนวนเหตุการณ์ที่เราสนใจในเหตุการณ์นี้คือ จำนวนแพคเกตไม่ปกติ 5 และเมื่อนำมาเปรียบเทียบกับจำนวนแพคเกตปกติที่มีอยู่จะได้ความน่าจะเป็นออกมาเท่ากับ 1/2 = 0.5

หากมีการส่งแพคเกตเกิดขึ้นทั้งหมด 30 ครั้ง มีแพคเกตปกติทั้งหมด 20 ครั้ง และแพคเกตไม่ปกติทั้งหมด 10 ครั้ง (1+1+1+...+1) หากยังคงเลือกใช้ทฤษฎีความน่าจะเป็นเพียงอย่างเดียว จำนวนเหตุการณ์ที่เราสนใจในเหตุการณ์นี้คือจำนวนแพคเกตที่ไม่ปกติคือ 10 และเมื่อนำมาเปรียบเทียบกับจำนวนแพคเกตปกติ คือ 20 จากนั้นนำมาคำนวณจะได้ความน่าจะเป็นเท่ากับ (1+1+1+...+1)/20 = 0.5 นั้นแสดงให้เห็นว่าไม่มีความแตกต่างกันแต่อย่างใด จึงยังไม่ตอบโจทย์ว่ามีความพยายามหรือรูปแบบพฤติกรรมในการบุกรุก

ดังนั้นผู้วิจัยจึงต้องให้ระบบใช้วิธิการเพิ่มระดับการระวังให้มากขึ้นในทุกๆ ครั้งที่มีแพคเกตไม่ปกติเข้ามา โดยในวิธีการนี้จะทำการเพิ่มการเฝ้าระวังความเสี่ยงในทุกๆ ครั้งที่มีแพคเกตไม่ปกติเข้ามา และนำไปเปรียบเทียบกับแพคเกตปกติ โดยใช้สมการจำนวนสามเหลี่ยม (Triangular number) เข้ามาช่วยทำงานร่วมกันกับสมการความน่าจะเป็น จะได้ว่า (1+2+3+4+5)/10 = 1.5 ซึ่งค่าที่ได้จะแตกต่างจากการนำสมการความน่าจะเป็นมาใช้เพียงอย่างเดียว ทำให้ระบบทำงานได้ใกล้เคียงการวิเคราะห์ของผู้วิจัยได้มากขึ้น

ในบางสถานการณ์หรือสภาพแวดล้อมของระบบเครือข่ายแต่ละแห่งนั้นแตกต่างกัน จึงต้องเพิ่มความยืดหยุ่นให้แก่สมการของระบบด้วยการปรับ (Adaptive) เพิ่มหรือลดตัวแปรค่าน้ำหนักไปตามสภาพของระบบเครือข่าย เนื่องจากความแตกต่างกันในหลากหลายสถานการณ์ บางสถานการณ์นั้นต้องการความสะดวกจึงต้องมีการให้ค่าน้ำหนักที่น้อย เพื่อลดจำนวนครั้งของการปิดกั้นเครือข่าย บางสถาการณ์นั้นต้องการความปลอดภัยมากจึงใช้ค่าน้ำหนักที่สูง เพื่อเพิ่มจำนวนครั้งในการปิดกั้นเครือข่าย จากนั้นจึงเขียนออกมาในรูปของสมการ (3) ดังนี้

$$IP\ Confidence = \frac{(\frac{K(K+1)}{2})\times W}{I} \qquad (3)$$

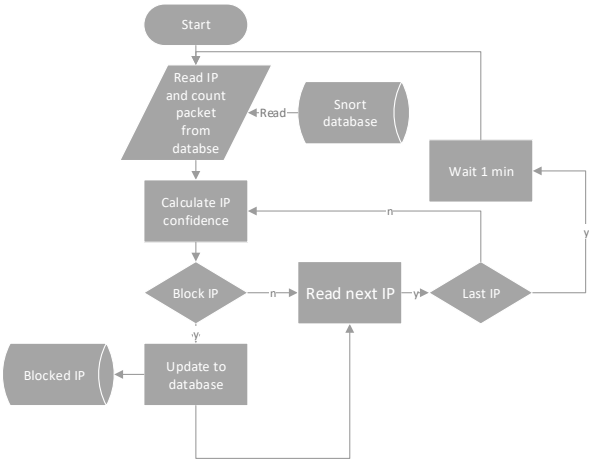
เมื่อ

I คือ จำนวนแพคเกตปกติ

K คือ จำนวนของแพตเกตไม่ปกติ

W คือค่าน้ำหนักที่ใช้ในการคำนวณ

โดยค่าที่ได้นั้นมาจากการนับจำนวนแพคเกตปกติ และไม่ปกติที่ระบบตรวจจับการบุกรุกได้บันทึกไว้ใน Snort Database และค่า W นั้นมาจากการที่ผู้วิจัยได้กำหนดขึ้น เพื่อทดสอบว่าค่าน้ำหนักเท่าใดเหมาะสมในสถานการณ์นี้



ภาพที่ 3: กระบวนการทำงานของ Adaptive IPS

3.3 หากค่า IP Confidence มากกว่า 1 จะทำการปิดกั้น IP นั้น โดยบันทึกค่าไปยัง Blocked IP และทำซ้ำจนครบทุกๆ IP

3.4 Manage IP ทำการควบคุม Firewall ให้ยกเลิกการปิดกั้น IP ที่ปิดกั้นไว้ก่อนหน้านี้ และ update IP ที่จะปิดกั้นใหม่ลงไปที่ Firewall แทน

3.5 ครบรอบเวลา 1 นาที จึงทำการเริ่มกระบวนการใหม่อีกครั้ง

3.6 ขั้นตอนสุดท้ายเป็นการวัดความแม่นยำของระบบการกรองแพคเกตที่สร้างขึ้นมา

3.6.1 สร้าง datasets ขึ้นมานั้นใช้โปรแกรม Hping ในการส่งข้อมูลจาก external network ไปยัง internal network จากนั้นจึงใช้โปรแกรม wireshark ในการบันทึกข้อมูลที่ส่ง ประกอบด้วยแพคเกตปกติจำนวน 5,000 แพตเกต และแพคเกตไม่ปกติจำนวน 5,000 แพคเกต กระจายส่ง โดยใช้เวลาส่งทั้งหมด 1 ชั่วโมง และบันทึกออกมาเป็น datasets

3.6.2 ทดสอบระบบด้วยวิธีการใช้ datasets ขึ้นมา ซึ่งจะส่งแพคเกตผ่านระบบกรองแพคเกตที่สร้างขึ้นมา

3.6.3 เมื่อส่งแพคเกตครบ ในเวลา 1 ชั่วโมง จากนั้นเราจะได้จำนวนแพคเกตปกติ และไม่ปกติที่ผ่านระบบมาได้ จากนั้นจึงนำมาวัดความแม่นยำต่อไป

4. ผลการดำเนินงาน

4.1 ในการทดสอบ เราทดลองส่งแพคเก็ตทั้งหมดให้ระบบทำการคัดกรอง โดยทดสอบเปรียบเทียบค่า *W* ระหว่าง 1 ถึง 10

4.2 วัดประสิทธิภาพของระบบป้องกันการบุกรุกโดยการส่งแพคเกตข้อมูลทั้งหมด จะได้ผลลัพธ์ทั้งหมด 4 แบบ ได้แก่

True Positive (*TP*) คือแพตเก็ตปกติที่ผ่านไฟร์วอล

True Negative (*TN*) คือแพตเกตปกติที่ไม่ผ่านไฟร์วอล

False Positive(*FP*) คือแพตเก็ตไม่ปกติที่ผ่านไฟร์วอล

False Negative(*FN*) คือแพตเก็ตไม่ปกติที่ไม่ผ่านไฟร์วอล

ตารางที่ 1: แสดงผลการทดสอบในแต่ละค่าน้ำหนัก

ค่า *W*	*TP*	*TN*	*FP*	*FN*
1	4,876	124	4,645	355
2	4,344	656	4,555	445
3	3,928	1,072	4,001	999
4	3,635	1,365	3,827	1,173

ตารางที่ 1: แสดงผลการทดสอบในแต่ละค่าน้ำหนัก (ต่อ)

ค่า *W*	*TP*	*TN*	*FP*	*FN*
5	3,444	1,556	3,210	1,790
6	2,999	2,001	2,928	2,072
7	2,544	2,456	2,726	2,274
8	2,100	2,900	2,345	2,655
9	1,958	3,042	2,017	2,983
10	1,588	3,412	1,999	3,001

จากตารางที่ 1 อธิบายได้ว่า ค่าน้ำหนัก (W) ต่ำ จะทำให้ค่า True Positive และ False Positive สูง ซึ่งเป็นผลมาจากการที่ระบบมีการปิดกั้นน้อยครั้งจึงทำให้แพคเกตนั้นผ่านเข้ามาเป็นจำนวนมาก แต่ถ้าหากค่าน้ำหนักสูง จะส่งผลตรงข้าม คือทำให้ค่า True Negative และ False Negative สูงขึ้น ซึ่งเป็นผลมาจากการที่ไฟร์วอลนั้นมีการปิดกั้นบ่อยครั้งมากกว่าเมื่อมีค่าน้ำหนักที่สูงจึงส่งผลให้จำนวนแพคเกตนั้นผ่านเข้ามาได้น้อย

4.3 ทดสอบประสิทธิภาพของระบบป้องกันการบุกรุก โดยวิธีการวัดประสิทธิภาพดังนี้ (3)

$$\text{Accuracy} = \frac{(TP+FN)}{(TP+TN+FP+FN)} \qquad (3)$$

เมื่อ

Accuracy คือค่าที่บ่งบอกความแม่นยำของระบบ

ตารางที่ 2: แสดงข้อมูลร้อยละของความถูกต้อง

ค่า W	*Accuracy (%)*
1	52.31
2	47.89
3	49.27
4	48.08
5	52.34
6	50.71
7	48.18
8	47.55
9	49.41
10	45.89

จากผลการทดลองในตารางที่ 1 แสดงให้เห็นว่าหากต้องการให้ระบบมีความสามารถในการป้องกันการบุกรุกอย่างมีประสิทธิภาพสามารถใช้ค่า W ที่สูง เพื่อทำให้ระบบนั้นปิดกั้นด้วยไฟร์วอลบ่อยครั้งสำหรับป้องกันการโจมตี แต่หากต้องการให้ผู้ใช้งานปกติใช้งานได้อย่างมีประสิทธิภาพ ได้รับผลกระทบน้อย ระบบจะต้องใช้ค่า W ที่ต่ำ เพื่อให้ไฟร์วอลนั้นไม่ปิดกั้นการเชื่อมต่อบ่อยเกินไป

แต่เนื่องจากวัตถุประสงค์งานวิจัยนี้ต้องการสร้างระบบป้องกันการบุกรุกที่มีประสิทธิภาพ โดยที่ให้ผู้ใช้งานปกตินั้นได้รับผลกระทบน้อย จึงมี 2 ปัจจัยที่ต้องนำมาวิเคราะห์ด้วยกัน ได้แก่ TP หมายถึงจำนวนของผู้ใช้งานปกติ และ FN หมายถึงจำนวนในการป้องกันการบุกรุก จากข้อมูลของตารางที่ 1 เมื่อนำมาคำนวณในสมการที่ 3 แล้ว ผลลัพท์ดังแสดงในตารางที่ 2 จะแสดงให้เห็นว่าค่า W เท่ากับ 5 นั้นให้ค่าความถูกต้องที่สามารถป้องกันการบุกรุก และส่งผลกระทบต่อผู้ใช้งานน้อยที่ค่าความแม่นยำ 52.34%

5. สรุป

หากต้องการป้องกันการบุกรุกได้ดีที่สุด คือการปิดกั้นการเชื่อมต่อตลอดเวลา ซึ่งจะทำให้ผู้ใช้งานปกตินั้นได้รับผลกระทบ แต่หากต้องการให้ผู้ใช้งานปกติได้รับความสะดวกเราจำเป็นต้องลดระดับความปลอดภัยลง

ดังนั้น ผลการทดสอบระบบกรองแพคเกตที่สร้างขึ้นมาเพื่อป้องกันการบุกรุกเครือข่าย จากการทดลองใช้ datasets ที่สร้างขึ้นมา มีความแม่นยำที่ดีที่สุดที่ค่า W เท่ากับ 5 แต่หากในสถานการณ์อื่นๆ ก็จะให้ผลที่แตกต่างกันออกไป

ผู้พัฒนาสามารถนำระบบที่สร้างขึ้นนี้เป็นแนวทางในการพัฒนาระบบป้องกันการบุกรุกให้มีประสิทธิภาพมากยิ่งขึ้น โดยสามารถนำไปใช้งานในสถานการณ์จริงได้ในอนาคต

เอกสารอ้างอิง

[1] Kim, D. S., & Park, J. S. “Network-based intrusion detection with support vector machines”. In Information Networking (pp. 747-756). Springer Berlin Heidelberg. 2003

[2] Gong, R. H., Zulkernine, M., & Abolmaesumi, P. “A software implementation of a genetic algorithm based approach to network intrusion detection”. SNPD/SAWN . 6th International Conference, Ontario, Canada. (pp. 246-253). IEEE. 2005

[3] Meng Y., Kwok L.F. “Adaptive blacklist-based packet filter with a statistic-based approach in network intrusion detection”, Journal of Network and Computer Applications. 2013

[4] Bace, R., & Mell, P. *NIST special publication on intrusion detection systems*. BOOZ-ALLEN AND HAMILTON INC MCLEAN VA.2001.

[5] Scarfone, K., & Mell, P. *Guide to intrusion detection and prevention systems (idps)*. NIST special publication, 800, 94.2007.

[6] สุธิชัย โง้วศิริ. *หลักสถิติ Principles of Statistics*, ภาควิชาสถิติและคอมพิวเตอร์. คณะวิทยาศาสตร์, สำนักพิมพ์มหาวิทยาลัยรามคำแหง, 2538

[7] Pengelley, David J. "The bridge between the continuous and the discrete via original sources", in Otto Bekken et al, Study the Masters: The Abel-Fauvel Conference, National Center for Mathematics Education, University of Gothenburg, Sweden, p. 3. 2002

เจทอล์ค : แอปพลิเคชันบนไอโอเอสเพื่อการสื่อสารในภาษาญี่ปุ่น
J-Talk: An iOS Application for Communication in Japanese

ดุลยารัตน์ ขาววิเศษ (Dunyarat kawviset)[1] และวรสิทธิ์ ชูชัยวัฒนา (Worasit Choochaiwattana)[2]
สาขาวิชาวิศวกรรมเว็บและการพัฒนาแอปพลิเคชันบนอุปกรณ์พกพา
วิทยาลัยครีเอทีฟดีไซน์ แอนด์ เอ็นเตอร์เทนเมนต์เทคโนโลยี มหาวิทยาลัยธุรกิจบัณฑิตย์
[1]dunyarat.kaw@dpu.ac.th, [2]worasit.cha@dpu.ac.th

บทคัดย่อ

จากข้อมูลการท่องเที่ยวประเทศญี่ปุ่นพบว่า จำนวนนักท่องเที่ยวชาวไทยเพิ่มขึ้นสูงมาก เป็นผลมาจากการยกเลิกการใช้วีซ่าท่องเที่ยวสำหรับคนไทย ดังนั้นงานวิจัยนี้มีวัตถุประสงค์ที่จะทำการพัฒนาไอโอเอสแอปพลิเคชัน สำหรับการให้ข้อมูลเกี่ยวกับบทสนทนาในชีวิตประจำวันในภาษาญี่ปุ่น เพื่ออำนวยความสะดวกให้นักท่องเที่ยวชาวไทย แอปพลิเคชันดังกล่าวมีชื่อว่า "เจ ทอล์ค" ในการประเมินผลการใช้งานแอปพลิเคชัน เจ ทอล์ค ใช้วิธีการสำรวจด้วยแบบสอบถาม โดยให้ผู้ตอบแบบสอบถามจำนวน 60 คน ทำการทดลองใช้แอปพลิเคชัน แล้วทำการประเมินความพึงพอใจในด้านต่าง ๆ

คำสำคัญ: แอปพลิเคชันบนอุปกรณ์พกพา แอปพลิเคชันเพื่อการสื่อสารในภาษาญี่ปุ่น การพัฒนาไอโอเอสแอปพลิเคชัน

Abstract

The information from Japan national tourism organization stated that the number of Thai tourists has been increased due to the exemption of tourist visa for Thai people. Thus, this paper aims at developing an iOS application for providing everyday-used conversation in Japanese language for Thai tourists. The application was named "J-Talk". To evaluate the effectiveness of J-talk, 60 participants have been recruited. They were asked to use the application and then answer the questions in an evaluation questionnaire. The results showed that the participants were satisfied with the application.

Keywords: Mobile Application, Application for Japanese Communication, iOS Application Development

1. บทนำ

คำว่า การสื่อสาร (Communication) หมายถึงกระบวนการที่ผู้ส่งสารต้องการถ่ายทอดข้อมูลต่างๆ ไปยังผู้รับสารให้เกิดความเข้าใจตรงกันระหว่างผู้สื่อสารและผู้รับสาร โดยผ่านการถ่ายทอดการสื่อสารในรูปแบบต่างๆ เช่น การพูด การเขียน อารมณ์ ความคิด ความรู้สึก หรือการแสดงลักษณะอื่นๆ ซึ่งการสื่อสารนั้นจะมีรูปแบบการสื่อสารที่แตกต่างกันไป โดยขึ้นอยู่กับสถานการณ์นั้นๆ [1]

ในปัจจุบันโทรศัพท์มือถือ (Smartphone) ถือเป็นปัจจัยสำคัญอย่างหนึ่งในการติดต่อสื่อสารในชีวิตประจำวันของคนในยุคปัจจุบันไม่ว่าจะเป็นการเข้าอินเตอร์เน็ต ดูหนัง ฟังเพลง หรือการส่งอีเมล์ (E- mail) เพื่ออำนวยความสะดวกและรวดเร็วในการติดต่อสื่อสาร ดังนั้นมือถือไม่ได้เป็นแค่อุปกรณ์ที่ใช้ในการติดต่อสื่อสารไร้สายเท่านั้น แต่ทว่าโทรศัพท์มือถือได้รับการพัฒนารูปแบบการทำงานให้มีประสิทธิภาพสูงสุด และมีความสามารถในการติดต่อสื่อสารข้อมูลให้เกิดประโยชน์สูงสุดกับผู้ใช้งาน ก็จะก่อให้เกิดความสะดวกและรวดเร็วในการติดต่อสอบถามข้อมูลจริงได้ตลอดเวลา

ในปี 2557 สถิติขององค์การส่งเสริมการท่องเที่ยวแห่งประเทศญี่ปุ่นในเดือน มกราคม - กุมภาพันธ์ 2557 มีนักท่องเที่ยวเดินทางไปประเทศญี่ปุ่น 61,500 คน เพิ่มขึ้นร้อยละ 70.9 โดยมีนักท่องเที่ยวเข้าไปท่องเที่ยวในประเทศญี่ปุ่นมากอย่างต่อเนื่อง [2]

ในปี 2558 มีคนไทยเดินทางไปท่องเที่ยวญี่ปุ่นจำนวนประมาณ 822,000 –855,000 คน เพิ่มขึ้นประมาณร้อยละ 25.0–30.0 ซึ่งยังคงเป็นอัตราสูงต่อเนื่องจากปี 2557 ที่เติบโตร้อยละ 45.0 (จำนวนนักท่องเที่ยว 657,600 คน) ขณะที่ภาพรวมการใช้จ่ายด้านต่างๆในญี่ปุ่นของคนไทย มีแนวโน้มคิดเป็นมูลค่า

ประมาณ 26,300–27,500 ล้านบาท เพิ่มขึ้นร้อยละ 20.0–25.0 จากที่เติบโตร้อยละ 45.3 ในปี 2557 [3] ตามภาพที่ 1

ตลาดนักท่องเที่ยวคนไทยที่เดินทางไปท่องเที่ยวญี่ปุ่นในปี 2558

	2557	2558
จำนวนนักท่องเที่ยวคนไทยที่เดินทางไปญี่ปุ่น (คน)	657,600	822,000 – 855,000
อัตราการขยายตัว (ร้อยละ)	45.0	25.0 – 30.0
มูลค่าการใช้จ่ายของนักท่องเที่ยวคนไทยในญี่ปุ่น (ล้านบาท)	22,000	26,300 – 27,500
อัตราการขยายตัว (ร้อยละ)	45.3	20.0 – 25.0

ภาพที่ 1: แผนภาพแสดงตลาดนักท่องเที่ยวคนไทยที่เดินทางไปท่องเที่ยวญี่ปุ่นในปี 2558

ดังนั้นผู้วิจัยได้เล็งเห็นว่าการพูด การสื่อสารของคนไทยที่มีต่อภาษาญี่ปุ่นนั้น ยังพบว่ายังคงเป็นปัญหาในเรื่องของการออกสำเนียงหรือน้ำเสียงตามบริบทต่างๆ แต่หากผู้รับสารได้รับการสื่อสารที่ถูกต้องก็จะทำให้เกิดความเข้าใจตรงกันระหว่างผู้ส่งสารและผู้รับสาร และผู้วิจัยซึ่งมีความสนใจด้านภาษาญี่ปุ่นและมีความสนใจในการพัฒนาแอปพลิเคชันบนโทรศัพท์มือถือสมาร์ทโฟน จึงเห็นความสำคัญสำหรับการติดต่อสื่อสารในการท่องเที่ยวที่ประเทศญี่ปุ่น เพื่ออำนวยความสะดวกและรวดเร็วให้กับผู้ใช้งานโทรศัพท์สำหรับกลุ่มเป้าหมายที่นิยมท่องเที่ยวประเทศญี่ปุ่นซึ่งเหมาะสำหรับผู้ที่ไม่มีความรู้ด้านภาษาญี่ปุ่นหรือมีความรู้ด้านภาษาญี่ปุ่นในระดับพื้นฐานสามารถติดต่อสื่อสารได้โดยอัตโนมัติและสอบถามแลกเปลี่ยนข้อมูลต่าง ๆ ระหว่างการท่องเที่ยวได้อย่างสะดวกและรวดเร็ว

2. งานวิจัยและแอปพลิเคชันที่เกี่ยวข้อง

2.1 งานวิจัยที่เกี่ยวข้อง

จากการจัดทำวิจัยครั้งนี้มีวัตถุประสงค์เพื่อการพัฒนาแอปพลิเคชัน เจ ทอล์ค แอปพลิเคชันบนไอโอเอสเพื่อการสื่อสารในภาษาญี่ปุ่น ซึ่งมีงานวิจัยที่เกี่ยวข้องกับแอปพลิเคชัน เจ ทอล์ค

ในขณะที่งานวิจัยเรื่อง **"ปัจจัยที่มีผลต่อการใช้โมบายแอปพลิเคชันบนสมาร์ทโฟน และแท็บเล็ต"** โดย รศ. วงหทัย ตันชีวะวงศ์ [4] วิจัยว่าเทคโนโลยีการสื่อสารไร้สายผ่านเครื่องมือสื่อสารแบบ พกพา (Wireless communication through portable devices)ได้แก่สมาร์ทโฟนและสมาร์ทแท็บเล็ตมีพัฒนาการอย่างรวดเร็วและได้รับความนิยมมากเพราะ เป็นการสื่อสารแบบจอสัมผัส (Touch-screen media) ทำให้พกพาสะดวกใช้งานง่ายผู้ใช้ไม่ต้องแบกน้ำหนักของเครื่องมือสื่อสารเหมือนเดิมเมื่อมีเทคโนโลยีสื่อสารไร้สายความเร็วสูงรองรับก็ยิ่งทำให้ผู้ใช้สามารถสื่อสารได้ทุกที่ทุกเวลาและในหลายๆมากยิ่งขึ้นเท่าไหร่ ยิ่งทำมให้ผู้ใช้งานสามารถทำการสื่อสารไร้สายรูปแบบต่างๆได้คล่องตัวมากยิ่งขึ้น ซึ่งตรงกับความต้องการของผู้ศึกษาทำการวิจัยแอปพลิเคชัน เจ ทอล์ค ในครั้งนี้ที่ต้องการให้การใช้งานแอปพลิเคชันเกิดความสะดวกและรวดเร็วต่อการใช้งาน

งานวิจัยเรื่อง " ปัจจัยที่มีผลต่อการดาวน์โหลดแอพพลิเคชั่น (Application) ของผู้ใช้โทรศัพท์มือถือไอโฟน (iPhone)" โดย กนกวรรณ ครินชัย [5] ซึ่งเป็นวิจัยที่ผู้วิจัยแอปพลิเคชัน ศึกษาปัจจัยที่มีผลต่อการดาวน์โหลดแอปพลิเคชันและการศึกษาครั้งนี้ซึ้งตรงกับความต้องการของผู้วิจัยแอปพลิเคชัน เจ ทอล์ค ที่ต้องการพัฒนาให้แอปพลิเคชันสามารถตอบสนองความต้องการและทำให้เกิดประโยชน์สูงสุดรวมทั้งจะช่วยส่งผลต่อการดาวน์โหลดแอปพลิเคชัน เจ ทอล์ค ได้อีกด้วย

2.2 แอปพลิเคชันที่เกี่ยวข้อง

ตารางที่ 1: แสดงชื่อแอปพลิเคชันที่เกี่ยวข้องเพื่ออำนวยความสะดวกในการติดต่อสื่อสารพร้อมด้วยคุณสมบัติและฟังก์ชันการทำงานของระบบ

คุณสมบัติฟังก์ชัน	รายชื่อแอปพลิเคชัน			
	Bravolol	JP Visit Travel Lite	Jicon	J-Talk โดยผู้วิจัย
ประโยคสื่อสารแบ่งเป็นประเภทหมวดหมู่	/	/	/	/
ประโยคความหมายภาษาไทย				/
ประโยคคำอ่านแบบอักษรโรมันจิ	/	/	/	/
เสียงพูดภาษาญี่ปุ่น	/	/		/

จากตารางที่ 1 แอปพลิเคชันที่เกี่ยวข้องในการวิจัยในครั้งนี้เป็นการเปรียบเทียบความแตกต่างของคุณสมบัติและฟังก์ชันการทำงานของระบบที่นำมาทำการวิจัย โดยผู้วิจัยได้เพิ่มส่วน

แตกต่างที่แอปพลิเคชันตัวอื่นไม่มี เช่น ความหมายที่เป็นภาษาไทย และเสียงพูดภาษาญี่ปุ่น เป็นต้น

3. ขั้นตอนวิธีการดำเนินการวิจัย

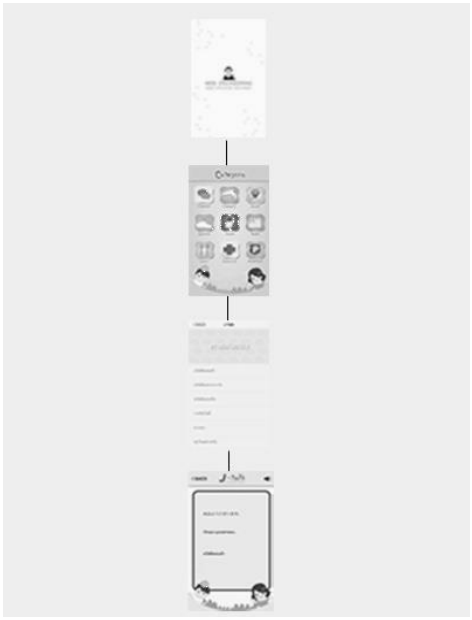
งานวิจัยนี้ผู้วิจัยได้พัฒนาแอปพลิเคชันบนไอโฟน โดยมีการกำหนดเนื้อหาที่เกี่ยวข้องกับ การสื่อสารสนทนา ณ ประเทศญี่ปุ่น และเพื่ออำนวยความสะดวกให้กับนักท่องเที่ยวชาวไทยที่ไปท่องเที่ยวไว้สำหรับติดต่อสอบถามข้อมูลต่างๆ ในแต่ละสถานการณ์เมื่อท่องเที่ยว ณ ประเทศญี่ปุ่น โดยได้นำเนื้อหา บทสนทนาภาษาญี่ปุ่นมาจาก [6], [7], [8], [9]

3.1 การวิเคราะห์ข้อมูลปัญหาและศึกษาค้นคว้าข้อมูล

การวิเคราะห์ปัญหาและศึกษาค้นคว้าข้อมูลเพื่อใช้ในการพัฒนาได้ศึกษาปัญหาและข้อจำกัดในด้านต่างๆที่มีผลกระทบต่อการพัฒนาระบบตลอดจนทฤษฎีและแนวคิดที่เกี่ยวของในการดำเนิน โดยใช้เครื่องมือที่ใช้โปรแกรม Xcode ร่วมกันภาษา Objective-C ในการพัฒนาแอปพลิเคชัน, SQLite Database ใช้สำหรับการสร้างฐานข้อมูล, โปรแกรม Photoshop และ Illustrator ใช้การตกแต่งภาพ รวมทั้ง Audacity สำหรับการตัดต่อเสียง

3.2 การวิเคราะห์และการออกแบบระบบ

เริ่มต้นจากการออกแบบ โครงสร้างของแอปพลิเคชัน J-Talk ซึ่งประกอบด้วยหน้าจอการทำงานต่างๆของแอปพลิเคชันตามภาพที่ 2



ภาพที่ 2: แสดงถึงการออกแบบสถาปัตยกรรมของระบบ J-Talk

3.3 การประเมินผลการใช้งานระบบ

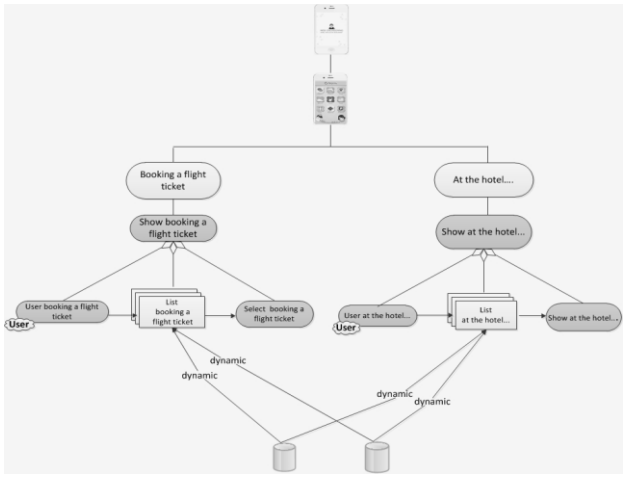
ผู้วิจัยได้สร้างแบบประเมินความพึงพอใจโดยการสร้างแบบสอบถามซึ่งจะใช้คำถามแบบมาตราส่วนประมาณค่า (Rating Scale) และคำถามปลายเปิด (Open-Ended Question) เพื่อให้ผู้ใช้แสดงคิดเห็นและให้ข้อเสนอแนะ กลุ่มผู้ได้รับการประเมินจะแบ่งออกเป็น 2 กลุ่ม คือกลุ่มผู้ที่มีความรู้ด้านภาษาญี่ปุ่นและกลุ่มผู้ที่ไม่มีความรู้ด้านภาษาญี่ปุ่น กลุ่มละ 30 คน

4. ผลการดำเนินการวิจัย

หลังจากผู้วิจัยได้ทำการศึกษาวิเคราะห์และการออกแบบระบบการพัฒนาแอปพลิเคชันนั้น ผู้วิจัยได้ทำการออกแบบขั้นตอนกระบวนการทำงานของระบบ เพื่อจัดวางโครงสร้างและกำหนดรูปแบบของแอปพลิเคชันให้ทำงานได้อย่างมีระเบียบแบบแผน เช่น การออกแบบกระบวนการทำงานของระบบ ระบบฐานข้อมูล หน้าจอหลัก หน้าจอแสดงประโยคของแต่ละหมวดหมู่ หน้าจอแสดงรายละเอียดของประโยคพร้อมเสียงของคำศัพท์ และผลของการทดสอบระบบโดยผู้ใช้งานจะทำการทดลองใช้งาน เพื่อสรุปผลการประเมินการใช้งานของระบบให้ได้ระบบที่ทำงานถูกต้องตามฟังก์ชันงาน

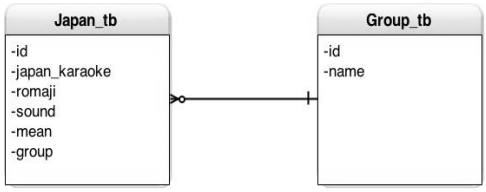
4.1 ผลการออกแบบและการพัฒนาระบบ

แบบการจำลองการทำงานของระบบ Information Design : Blueprint การออกแบบโครงสร้างของระบบ J-Talk Application มีการออกแบบโครงสร้างการทำงานสามารถนำเสนอในรูปแบบการเข้าถึงข้อมูลของระบบ ตามภาพที่ 3



ภาพที่ 3: Information Design : Blueprint ของระบบ J-Talk Application

การออกแบบระบบฐานข้อมูล J-Talk Application เป็นออกแบบการเก็บฐานข้อมูลที่ใช้ในการงานระบบฐานข้อมูล ซึ่งมีการเก็บข้อมูล 2 ตารางความสัมพันธ์ ได้แก่ฐานข้อมูล Japam_tb และ Group_tb ตามภาพที่ 4



ภาพที่ 4: Database Schema Design ของระบบ J-Talk Application

หน้าจอการทำงานของระบบ J-Talk จะประกอบไปด้วยหน้าจอแสดงหมวดหมู่ เมื่อผู้ใช้งานเข้ามาสู่จะพบกับหมวดหมู่ของแต่ละประเภทที่แยกไว้เพื่อให้ผู้ใช้ได้เลือกประเภทของประโยคสื่อสารตามความต้องการ ตามภาพที่ 5



ภาพที่ 5: หน้าจอแสดงหมวดหมู่

หน้าจอที่แสดงรายการคำสื่อสารสนทนา เมื่อผู้ใช้งานกดเลือกประโยคสนทนาที่ต้องการ ประโยคสนทนาที่เลือกนั้นจะโชว์รายละเอียดของคำอ่านที่เป็นภาษาญี่ปุ่นและเสียงพูดที่เป็นภาษาญี่ปุ่นของประโยคสนทนานั้น ตามภาพที่ 6



ภาพที่ 6: หน้าจอแสดงรายการประโยคของหมวดหมู่การสนทนา

เมื่อผู้ใช้งานกดเลือกประโยคสนทนาจากหน้าจอแสดงรายการ ก็จะเข้ามาหน้าจอแสดงรายละเอียดของประโยคที่ผู้ใช้เลือก ประกอบไปด้วย ประโยคสื่อสารแบบภาษาญี่ปุ่น,แบบภาษาโรมันจิและเสียงพูดที่เป็นภาษาญี่ปุ่นของประโยคนั้นๆ ตามภาพที่ 7



ภาพที่ 7: หน้าจอแสดงรายละเอียดของประโยคหมวดหมู่สนทนา

4.2 ผลการประเมินความพึงพอใจ

จากข้อมูลลักษณะกลุ่มตัวอย่างจากการประเมินความเหมาะสม/ความพึงพอใจของผู้ทดลองใช้งานแอปพลิเคชัน เจทอล์ค: แอปพลิเคชันบนไอโอเอสเพื่อการสื่อสารในภาษาญี่ปุ่น โดยแบ่งเป็นกลุ่มผู้ที่มีความรู้ด้านภาษาญี่ปุ่นและกลุ่มผู้ที่ไม่มีความรู้ด้านภาษาญี่ปุ่น กลุ่มละ 30 คน

ตารางที่ 2: ประสิทธิภาพ/ความง่ายต่อการใช้งานของระบบ ของผู้ที่มีความรู้ด้านภาษาญี่ปุ่นและผู้ที่ไม่มีความรู้ด้านภาษาญี่ปุ่น

ประสิทธิภาพ/ความง่ายต่อการใช้งานของระบบ	ของผู้ที่มีความรู้ด้านภาษาญี่ปุ่น		ของผู้ที่ไม่มีความรู้ด้านภาษาญี่ปุ่น	
	X	S.D.	X	S.D.
ความรวดเร็วในการแสดงผลหน้าจอ	4.83	0.38	4.53	0.63
ความง่าย (User Friendly)ของการใช้งาน	4.87	0.35	4.63	0.61
มีความสะดวกในการใช้งานการค้นหาข้อมูล	4.63	0.56	4.33	0.76
เฉลี่ยรวม	4.78	0.43	4.50	0.67

จากตารางที่ 2 ประสิทธิภาพ/ความง่ายต่อการใช้งานของระบบ จำนวน 3 คำถามของผู้ที่มีความรู้ด้านภาษาญี่ปุ่น ในด้านความคิดของผู้ตอบแบบสอบถาม พบว่ามีความพึงพอใจโดยภาพรวมในระดับมากที่สุด โดยมีค่าเฉลี่ยอยู่ที่ 4.78 และผู้ที่ไม่มีความรู้ด้านภาษาญี่ปุ่นในด้านความคิดของผู้ตอบแบบสอบถาม พบว่ามีความพึงพอใจโดยภาพรวมในระดับมาก โดยมีค่าเฉลี่ยอยู่ที่ 4.50

ตารางที่ 3: ความเหมาะสม/ความพึงพอใจด้านการทำงานได้ตามหน้าที่ของระบบของผู้ที่มีความรู้ด้านภาษาญี่ปุ่นและผู้ที่ไม่มีความรู้ด้านภาษาญี่ปุ่น

การทำงานได้ตามหน้าที่ของระบบ ความถูกต้องของข้อมูลที่ได้จากการค้นหา	ของผู้ที่มีความรู้ด้านภาษาญี่ปุ่น		ของผู้ที่ไม่มีความรู้ด้านภาษาญี่ปุ่น	
	X	S.D.	X	S.D.
ความถูกต้องของประโยคภาษาญี่ปุ่น	4.50	0.68	4.47	0.78
ความถูกต้องของสำเนียงภาษาญี่ปุ่น	4.70	0.47	4.57	0.63
ความถูกต้องของหมวดหมู่ที่ใช้งาน	4.67	0.55	4.70	0.47
ความน่าเชื่อถือของข้อมูล	4.67	0.48	4.73	0.45
การทำงานได้ตามหน้าที่ของระบบ	4.50	0.73	4.60	0.56
เฉลี่ยรวม	4.61	0.58	4.61	0.58

จากตารางที่ 3 ความเหมาะสม/ความพึงพอใจด้านการทำงานได้ตามหน้าที่ของระบบ จำนวน 5 ข้อคำถามของผู้ที่มีความรู้ด้านภาษาญี่ปุ่น ในด้านความคิดของผู้ตอบแบบสอบถาม พบว่ามีความพึงพอใจโดยภาพรวมในระดับมากที่สุด โดยมีค่าเฉลี่ยอยู่ที่ 4.61 และที่ไม่มีความรู้ด้านภาษาญี่ปุ่น ในด้านความคิดของผู้ตอบแบบสอบถาม พบว่ามีความพึงพอใจโดยภาพรวมในระดับมากที่สุด โดยมีค่าเฉลี่ยอยู่ที่ 4.61

ตารางที่ 4: ความเหมาะสม/ความพึงพอใจด้านการออกแบบ/การจัดรูปแบบแอปพลิเคชันของผู้ที่มีความรู้ด้านภาษาญี่ปุ่นและผู้ที่ไม่มีความรู้ด้านภาษาญี่ปุ่น

ด้านการออกแบบ/การจัดรูปแบบแอปพลิเคชัน	ของผู้ที่มีความรู้ด้านภาษาญี่ปุ่น		ของผู้ที่ไม่มีความรู้ด้านภาษาญี่ปุ่น	
	X	S.D.	X	S.D.
ความน่าสนใจของข้อมูล	4.57	0.63	4.53	0.63
ความหลากหลายของข้อมูล	4.20	0.76	4.20	0.71
ความสวยงามของหน้าจอแสดงผล	4.60	0.67	4.67	0.66
ตัวอักษรง่ายต่อการอ่านและการใช้งาน	4.47	0.68	4.63	0.49
ภาพกับเนื้อหามีความสอดคล้องกัน	4.53	0.68	4.67	0.55
เฉลี่ยรวม	4.47	0.68	4.54	0.61

จากตารางที่ 4 ความเหมาะสม/ความพึงพอใจด้านการออกแบบ/การจัดรูปแบบแอปพลิเคชัน จำนวน 5 ข้อคำถามของผู้ที่มีความรู้ด้านภาษาญี่ปุ่น ในด้านความคิดของผู้ตอบแบบสอบถาม พบว่ามีความพึงพอใจโดยภาพรวมในระดับมากที่สุด โดยมีค่าเฉลี่ยอยู่ที่ 4.47 และผู้ไม่มีความรู้ด้านภาษาญี่ปุ่น ในด้านความคิดของผู้ตอบแบบสอบถาม มีความพึงพอใจโดยภาพรวมในระดับมากที่สุด โดยมีค่าเฉลี่ยอยู่ที่ 4.54

5. สรุป

จากการพัฒนาแอปพลิเคชัน เจ ทอล์ค : แอปพลิเคชันบนไอโอเอสเพื่อการสื่อสารในภาษาญี่ปุ่น มีวัตถุประสงค์เพื่อพัฒนาระบบให้ตรงกับความต้องการของคนสังคมในปัจจุบันที่นิยมไปท่องเที่ยว ณ ประเทศญี่ปุ่น ผู้พัฒนาระบบได้วางแผนเลือกการทำแอปพลิเคชันเป็นภาษาญี่ปุ่นและพัฒนาระบบอย่างเป็นขั้นตอน มีการกำหนดปัญหาและความต้องการ ซึ่งได้แก่ ประโยคที่นักท่องเที่ยวชาวไทยนิยมใช้บ่อยเพื่อพัฒนาแอปพลิเคชันโทรศัพท์พกพา (I-Phone) ให้เหมาะกับการสื่อสารที่ประเทศญี่ปุ่น เพื่ออำนวยความสะดวกให้กับนักท่องเที่ยวคนไทยที่ไปท่องเที่ยวไว้สำหรับติดต่อสอบถามข้อมูลต่างๆ ในแต่ละสถานที่ของประเทศญี่ปุ่นเพื่อต้องการให้แอปพลิเคชันโทรศัพท์พกพาได้เกิดประโยชน์สูงสุดในชีวิตประจำวัน เพื่อพัฒนาแอปพลิชันให้เหมาะสมกับผู้ที่มีความรู้ด้านภาษาญี่ปุ่นและผู้ที่ไม่มีความรู้ด้านภาษาญี่ปุ่น

หากนำแอปพลิเคชันบนไอโอเอสเพื่อการสื่อสารภาษาญี่ปุ่นนี้อยู่ใน App Store ผู้ตอบแบบสอบถามที่มีความรู้ด้านภาษาญี่ปุ่นทั้งหมด และผู้ที่ไม่มีความรู้ด้านภาษาญี่ปุ่นแสดงความคิดเห็นว่าหากนำแอปพลิเคชัน เจ ทอล์ค : แอปพลิเคชันบนไอโอเอสเพื่อการสื่อสารภาษาญี่ปุ่น อยู่ใน App Store มีผู้ Download จำนวนอย่างละทั้งหมด 30 คน หรือร้อยละ 100

ดังนั้นผลจาการประเมินแอปพลิเคชัน เจ ทอล์ค : แอปพลิเคชันบนไอโอเอสเพื่อการสื่อสารในภาษาญี่ปุ่น มีการทำงานได้ตรงกับวัตถุประสงค์ที่ตั้งไว้ ซึ่งเหมาะกับคนไทยที่มีความรู้ด้านภาษาญี่ปุ่นและไม่มีความรู้ด้านภาษาญี่ปุ่นที่ต้องการไปท่องเที่ยวประเทศญี่ปุ่น

เอกสารอ้างอิง

[1] ภาษาไทยเพื่อการสื่อสาร, “ความหมายของการสื่อสาร” สืบค้น 14 พฤศจิกายน 2557, จาก http://www.ipesp.ac.th/learning/thai/chapter1-1.html

[2] สถานเอกอัครราชทูต (ณ กรุงโตเกียว), สืบค้น 14 พฤศจิกายน 2557, จาก http://www.thaiembassy.jp

[3] สำนักข่าวหุ้นอินไซด์, "คนไทยเที่ยวญี่ปุ่นปี 2558 ฯ" สืบค้น 3 เมษายน 2558, จาก http://www.hooninside.com/news

[4] รศ. วงหทัย ตันชีวะวงศ์, “ปัจจัยที่มีผลต่อการใช้โมบายแอปพลิเคชันบนสมาร์ทโฟน และแท็บเล็ต” สืบค้น 3 เมษายน 2558, จาก,

http://beyond.library.tu.ac.th/cdm/ref/collection/research/id/233

[5] กนกวรรณ ครินชัย (2554) งานวิจัยเรื่อง "ปัจจัยที่มีผลต่อการดาวน์โหลดแอพพลิเคชั่น (Application) ของผู้ใช้โทรศัพท์มือถือไอโฟน (iPhone)" สืบค้น 3 เมษายน 2558, จาก http://digi.library.tu.ac.th/thesis/it/0786/title- biography.pdf

[6] โลกออนไลน์เพื่อการเรียนภาษาญี่ปุ่น, สืบค้น 14 พฤศจิกายน 2557, จาก http://www.j-campus.com/kanji

[7] *คู่มือท่องเที่ยวตามใจชอบญี่ปุ่น*, สำนักงานพิมพ์ ทิบ ไทย อินเตอร์ บุ๊ค, คันไซ, (พิมพ์ครั้งที่1), (2556).

[8] สมาคมส่งเสริมเทคโนโลยี (ไทย-ญี่ปุ่น), *มินะ โนะ นิฮงโกะ 1*, สำนักพิมพ์ภาษาและวัฒนธรรม สมาคมส่งเสริมเทคโนโลยี (ไทย-ญี่ปุ่น), กรุงเทพฯ,(2546).

[9] สมาคมส่งเสริมเทคโนโลยี (ไทย-ญี่ปุ่น), *มินะ โนะ นิฮงโกะ 2*, สำนักพิมพ์ภาษาและวัฒนธรรม สมาคมส่งเสริมเทคโนโลยี (ไทย-ญี่ปุ่น), กรุงเทพฯ, (2546).

การพัฒนาระบบแนะนำสมุนไพรไทยเพื่อสุขภาพบนโทรศัพท์มือถือระบบแอนดรอย์
Development of Thai Herb Recommendation System for Healthcare on Android Phone

ณัฐวดี หงษ์บุญมี (Nattavadee Hongboonmee)[1] และฉัตรสุดา นิลสกุล (Chatsuda Ninsakul)[2]
ภาควิชาวิทยาการคอมพิวเตอร์และเทคโนโลยีสารสนเทศ คณะวิทยาศาสตร์ มหาวิทยาลัยนเรศวร
[1]nattavadeeho@nu.ac.th

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาระบบแนะนำสมุนไพรไทยสำหรับอำนวยความสะดวกแก่ผู้ใช้ในการดูแลสุขภาพภายในระบบประกอบด้วยการตรวจสอบอาการป่วยเบื้องต้นและแนะนำการรักษาด้วยสมุนไพรไทยโดยใช้เทคนิคตารางตัดสินใจ ซึ่งสร้างจากปัจจัยในการเลือกสมุนไพรจำนวน 4 ปัจจัย นอกจากนี้ระบบยังสามารถค้นหาข้อมูลสมุนไพร ข้อมูลร้านยาสมุนไพรโดยออกแบบระบบในรูปแบบแอปพลิเคชันบนโทรศัพท์มือถือระบบแอนดรอย์ เครื่องมือในการพัฒนาใช้โปรแกรม Android Studio และภาษา JAVA โปรแกรม MySQL สำหรับการจัดการฐานข้อมูล การประเมินประสิทธิภาพระบบใช้การวัดระดับความพึงพอใจของผู้ใช้ 2 กลุ่มคือผู้เชี่ยวชาญจำนวน 3 คนและผู้ใช้ทั่วไปจำนวน 35 คน ผลการประเมินจากผู้เชี่ยวชาญพบว่ามีค่าเฉลี่ยรวม 4.40 ส่วนเบี่ยงเบนมาตรฐาน 0.38 อยู่ในระดับดี ส่วนผู้ใช้ทั่วไปมีความพึงพอใจต่อประสิทธิภาพของระบบโดยรวมอยู่ในระดับดีด้วยค่าเฉลี่ยรวมเท่ากับ 4.11 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.62 ดังนั้นสามารถสรุปได้ว่าระบบนี้มีประสิทธิภาพสามารถนำไปใช้งานได้

คำสำคัญ: ระบบแนะนำ สมุนไพรไทย สุขภาพ แอปพลิเคชันแอนดรอย์

Abstract

This research aims to develop Thai herb recommendation system for facilitating users in health care. This includes initial disease screening and the system will recommend Thai herbal treatments by using decision table technique. Which was built from the factors in choosing herbs were 4 factors. Furthermore, the system can also search herb information and herb shop information. By designing the system as application on Android phone. Development tools use Android Studio and JAVA programming, MySQL for database management. The performance evaluation system used to measure the satisfaction level of 2 set users. There are 3 experts and 35 general users. The results of the experts' evaluation showed that the average was 4.40 and the standard deviation 0.38, which is in a good level. The general users were satisfied with the performance of the overall system is at a good level. Overall average equal to 4.11,standard deviation equal to 0.62. So it can be concluded that this system is effective and usable.

Keyword: Recommendation System, Thai Herb, Health, Application, Android.

1. บทนำ

การมีสุขภาพที่ดีนอกจากการดูแลตนเองแล้วปัจจัยที่สำคัญอีกอย่างหนึ่งได้แก่การรักษาสุขภาพ ซึ่งวิธีที่คนไทยรู้จักและคุ้นเคยเป็นอย่างดีก็คือการรักษาสุขภาพแบบแพทย์แผนไทย โดยวิธีนี้เป็นการนำสมุนไพรไทยมาใช้ในการรักษาโรคต่างๆ เนื่องจากสมุนไพรไทยมีสรรพคุณที่หลากหลายสามารถรักษาได้ทุกโรค ตัวยาจากสมุนไพรมักไม่มีผลกระทบข้างเคียงเหมือนตัวยาแผนปัจจุบัน ทั้งนี้ปัจจุบันกระแสนิยมในการดูแลสุขภาพด้วยการแพทย์แผนไทยเป็นที่ยอมรับกันเพิ่มมากขึ้น โดยเฉพาะการใช้สมุนไพรดูแลสุขภาพก่อนเกิดอาการ องค์การอนามัยโลก[1] ได้แนะนำให้ใช้การแพทย์ทั้งการแพทย์แผนปัจจุบันและการแพทย์แผนโบราณคู่ขนานกันไป

ด้วยเหตุผลดังกล่าวในงานวิจัยนี้จึงนำเสนอวิธีการรักษาสุขภาพแบบแพทย์แผนไทยในรูปแบบของการพัฒนาระบบ

แนะนำสมุนไพรไทยเพื่อดูแลสุขภาพบนโทรศัพท์มือถือระบบปฏิบัติการแอนดรอยด์ โดยประยุกต์ใช้เทคนิคตารางตัดสินใจ ระบบนี้จะช่วยให้สามารถดูแลสุขภาพและตรวจสอบอาการป่วยของตนเองได้ทันทีเพื่อที่จะรู้อาการและวิธีการรักษาเบื้องต้นโดยใช้ยาสมุนไพรไทยทำให้ไม่ต้องเสียเวลาในการเดินทางไปยังสถานพยาบาลต่างๆ และเป็นการอำนวยความสะดวกให้กับบุคคลในสังคมปัจจุบันที่สนใจวิธีการรักษาแบบแพทย์แผนไทย นอกจากนั้นยังสามารถค้นหาร้านขายยาสมุนไพรไทยและบอกเส้นทางในการเดินทางไปยังร้านขายยาสมุนไพรเพื่อความสะดวกและรวดเร็วในการเดินทางเพิ่มมากขึ้น

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 ระบบแนะนำ

ระบบแนะนำ (Recommendation System) เป็นระบบสำหรับการแนะนำข้อมูลสินค้าและบริการที่เหมาะสมหรือน่าสนใจให้กับผู้ใช้ระบบ [2] ระบบให้คำแนะนำประกอบไปด้วย 3 ส่วนหลักได้แก่ 1) ส่วนข้อมูลที่จำเป็นต้องใช้ในการประมวลผลเช่น ข้อมูลของผู้ใช้แต่ละคน 2) ส่วนของการประมวลผลเป็นส่วนสำคัญที่สุดเพื่อให้การแนะนำข้อมูลตรงความต้องการของผู้ใช้มากที่สุดและ 3) ส่วนของการนำเสนอคำแนะนำสินค้าหรือบริการให้กับผู้ใช้ ซึ่งในปัจจุบันมีการนำระบบแนะนำมาประยุกต์ใช้แนะนำข้อมูลแทบทุกด้านเช่น การแนะนำร้านอาหาร การแนะนำท่องเที่ยว การแนะนำเลือกซื้อคอมพิวเตอร์ การแนะนำเลือกรายวิชาเรียน เป็นต้น

2.2 สมุนไพรไทย

สมุนไพรหมายถึงพืชที่มีสรรพคุณในการรักษาโรคหรืออาการเจ็บป่วยต่างๆ [3] การนำสมุนไพรตั้งแต่สองชนิดขึ้นไปมาผสมรวมกันเพื่อใช้ในการรักษาโรคหรืออาการเจ็บป่วยเรียกว่ายา ปัจจุบันยาสมุนไพรได้รับการคัดเลือกเข้าบัญชียาหลักแห่งชาติจำนวน 71 รายการเพื่อให้ประชาชนมีโอกาสเลือกใช้ยาแผนไทยมากขึ้น

ยาสมุนไพรถ้าเป็นยาไทยหรือยาประกอบไม่ใช่ยาเดี่ยว จะมีองค์ประกอบของยาสมุนไพรหลายอย่าง การรักษาจะเป็นการรักษาให้เกิดความสมดุลไม่ได้รักษาเฉพาะส่วนเหมือนยาแผนปัจจุบัน ข้อดีของการใช้ยาสมุนไพรคือ 1) เป็นการดูแลโดยองค์รวม 2) เป็นการรักษาที่เสียค่าใช้จ่ายน้อยประหยัดงบประมาณ 3) สมุนไพรสามารถหาได้ในท้องถิ่น

2.3 ระบบปฏิบัติการแอนดรอย์

แอนดรอยด์ (Android) เป็นระบบปฏิบัติการสำหรับอุปกรณ์พกพาเช่นโทรศัพท์มือถือ แท็บเล็ตคอมพิวเตอร์ [4] พัฒนาขึ้นโดยบริษัทแอนดรอยด์ ภายหลังกูเกิลได้ซื้อบริษัทแอนดรอยด์และนำระบบปฏิบัติการแอนดรอยด์ไปพัฒนาต่อโดยกูเกิลได้อนุญาตให้นักพัฒนาสามารถแก้ไขโค้ดต่างๆ ด้วยภาษาจาวาและควบคุมอุปกรณ์ผ่านทางชุด Java libraries ลิขสิทธิ์ของโค้ดแอนดรอยด์สามารถใช้ในลักษณะของซอฟต์แวร์เสรี

2.4 เทคนิคตารางการตัดสินใจ

ตารางการตัดสินใจ (Decision Table) เป็นเครื่องมือที่ใช้แสดงเงื่อนไขการตัดสินใจ [5] และการเลือกทำงานภายใต้เงื่อนไขที่ระบุในรูปแบบของตาราง เหมาะสำหรับการตัดสินใจที่มีความซับซ้อนและเป็นการเลือกสิ่งที่ดีที่สุดจากตัวเลือกที่มีหลากหลาย โดยทั่วไปตารางการตัดสินใจมีส่วนประกอบที่สำคัญ 4 ส่วน ดังตารางที่ 1 และการสร้างกฎหรือการระบุกฎที่เป็นไปได้ภายใต้เงื่อนไขที่ระบุ ดังตัวอย่างในตารางที่ 2

ตารางที่ 1: แสดงส่วนประกอบของตารางการตัดสินใจ

เงื่อนไข (Condition)	ทางเลือกของเงื่อนไข
การกระทำ (Action)	กฎของการกระทำ

ตารางที่ 2: แสดงขั้นตอนการระบุกฎของการกระทำ

เงื่อนไข (Condition)	**ทางเลือกของเงื่อนไข**			
เงื่อนไข 1	Y	Y	N	Y
เงื่อนไข 2	Y	N	Y	Y
เงื่อนไข 3	N	Y	Y	Y
การกระทำ (Action)	**กฎของการกระทำ**			
การกระทำ 1			X	
การกระทำ 2	X			
การกระทำ 3		X		

2.5 งานวิจัยที่เกี่ยวข้อง

ประภาวดีและเกียรติศักดิ์ [6] นำเสนองานวิจัยระบบแนะนำโปรแกรมตรวจสุขภาพกรณีศึกษาโรงพยาบาลเอกชนในเขตกรุงเทพมหานคร โดยใช้เทคนิค Collaborative Filtering ซึ่ง

เป็นระบบสำหรับแนะนำโปรแกรมตรวจสุขภาพที่เหมาะสมกับความต้องการของผู้ใช้มากที่สุด โดยออกแบบระบบในรูปแบบเว็บแอปพลิเคชัน ผลการประเมินจากผู้ใช้พบว่ามีความพึงพอใจอยู่ในระดับดี

รุจิราและทองพูล [7] นำเสนองานวิจัยเรื่องระบบสนับสนุนการตัดสินใจเพื่อช่วยค้นหาโรงแรมในอำเภอเมืองหัวหินโดยใช้ตารางการตัดสินใจ เนื่องจากตารางตัดสินใจเป็นเทคนิคที่เหมาะกับกรณีที่เงื่อนไขของการตัดสินใจมีหลายเงื่อนไข เงื่อนไขที่ผู้ใช้เลือกจะเป็นตัวกำหนดผลการค้นหาโรงแรม ผลการวัดประสิทธิภาพในการค้นหามีค่าเฉลี่ยเกินร้อยละ 80

สมุนไพรไทยสรรพคุณเพื่อสุขภาพ [8] แอปพลิเคชันนี้รวบรวมข้อมูลเกี่ยวกับสมุนไพรไทยแสดงรายละเอียดสมุนไพรเช่น ชื่อวิทยาศาสตร์ ชื่อสามัญ สรรพคุณและวิธีการใช้งาน

แอปพลิเคชันสมุนไพรไทย-220+ ชนิด [9] แอปพลิเคชั่นที่รวบรวมข้อมูลพืชสมุนไพร 226 ชนิดพร้อมทั้งรายละเอียดข้อมูลทางด้านวิทยาศาสตร์ รูปภาพ สรรพคุณ สูตรสมุนไพร โดยสามารถแยกการแสดงผลเรียงตามพยัญชนะ

จากการศึกษาแอปพลิเคชันทั้ง 2 พบว่าแอปพลิเคชันทั้ง 2 นี้ยังขาดในส่วนของตรวจสอบอาการป่วยและแสดงโรคที่เกี่ยวข้องกับอาการป่วยพร้อมทั้งการแนะนำวิธีรักษาโรคด้วยสมุนไพร

3. วิธีการดำเนินการวิจัย

การสร้างระบบแนะนำสมุนไพรไทยเพื่อดูแลสุขภาพบนโทรศัพท์มือถือระบบแอนดรอย์ มีขั้นตอนการดำเนินการดังนี้

3.1 การศึกษาข้อมูลที่เกี่ยวข้องกับการดูแลสุขภาพด้วยสมุนไพรไทย

การศึกษาและกำหนดขอบเขตการดำเนินงาน การรวบรวมข้อมูลจากเอกสารต่างๆ รวบถึงการสอบถามจากบุคลากรด้านแพทย์แผนไทย สามารถกำหนดขอบเขตการทำงานของระบบได้ดังนี้ 1) การวิเคราะห์ข้อมูลและแนะนำสมุนไพรไทยสำหรับรักษาอาการป่วยเบื้องต้น 2) ข้อมูลและรายละเอียดสมุนไพรไทย 3) ข้อมูลร้านยาสมุนไพรไทย

3.2 การวิเคราะห์สมุนไพรไทยที่เหมาะสมโดยใช้ตารางการตัดสินใจ

ในงานวิจัยนี้เลือกใช้เทคนิคตารางการตัดสินใจในการวิเคราะห์เพื่อเลือกสมุนไพรไทยที่เหมาะสมจากการตอบคำถามของผู้ใช้ระบบ เนื่องจากเหมาะสำหรับการเลือกในกรณีตัวเลือกมีความสัมพันธ์กันและตัวเลือกมีความหลากหลาย ซึ่งจากการศึกษาหลักการแนะนำการรักษาด้วยสมุนไพรจากเอกสาร [10] พบว่ามีปัจจัยที่สำคัญจำนวน 4 ปัจจัยคือ เพศ อายุ บริเวณร่างกายและอาการป่วย โดยสรุปเงื่อนไขในการสร้างตารางการตัดสินใจได้ดังตารางที่ 3

ตารางที่ 3: การสร้างตารางตัดสินใจ

เงื่อนไข	กฎสำหรับการตัดสินใจ
เพศ	1. ชาย 2. หญิง
อายุ	1. 0-29 2. 30-44 3. 45-60 4. มากกว่า 60
บริเวณร่างกาย	1. ศีรษะ 2. ลำตัวส่วนกลาง 3. ลำตัวส่วนล่าง 4.ทั่วไป
อาการ	1. ปวดศีรษะ 2. เครียด 3. ตาพร่า ตามัว 4. เจ็บคอ 5. ปวดหลัง 6. ปวดท้อง 7. นอนไม่หลับ 8. คลื่นไส้ อาเจียน ฯลฯ

จากตารางที่ 3 สามารถสรุปปัจจัยในการแนะนำสมุนไพรไทยได้ทั้งหมด 4 ปัจจัย และสามารถสร้างเป็นตารางการตัดสินใจได้ดังตารางที่ 4 ซึ่งใช้ในส่วนแนะนำสมุนไพรไทยของระบบที่พัฒนาขึ้น โดยในตารางได้ยกตัวอย่างเพียง 3 ตัวอย่าง

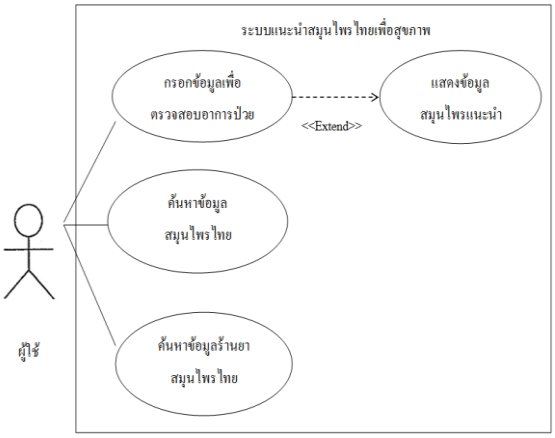
ตารางที่ 4: ตัวอย่างตารางตัดสินใจของระบบที่พัฒนาขึ้น

เงื่อนไข	สมุนไพรสูตร1	สมุนไพรสูตร2	สมุนไพรสูตร3
เพศ	หญิง	ชาย	หญิง
อายุ	45 - 60	30 - 44	30 - 44
บริเวณร่างกาย	ทั่วไป	ลำตัวส่วนกลาง	ศีรษะ
อาการ	นอนไม่หลับ	ปวดท้อง	ปวดศีรษะ

3.3 การออกแบบระบบ

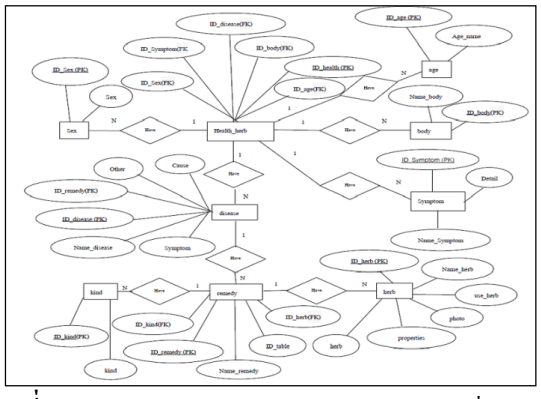
ใช้การออกแบบเชิงวัตถุด้วยแผนภาพ UML (Unified Modeling Language) ประกอบไปด้วย 1) Use Case Diagram และ 2) ER Diagram

จากภาพที่ 1 ภายในระบบถูกแบ่งออกเป็น 3 ส่วนหลักคือ 1) การตรวจสอบอาการป่วยโดยผู้ใช้กรอกข้อมูลระบบจะวิเคราะห์ข้อมูลแล้วแนะนำตัวยาสมุนไพรไทยสำหรับการรักษา 2) การค้นหาข้อมูลสมุนไพรไทย 3) การค้นหาร้านยาสมุนไพรและแผนที่



ภาพที่ 1: Use Case Diagram ของระบบแนะนำสมุนไพรไทยเพื่อสุขภาพ

ER Diagram (Entity Relationship Diagram) แบบจำลองที่ใช้อธิบายโครงสร้างของฐานข้อมูล ซึ่งในระบบนี้ทำการออกแบบข้อมูลทั้งหมด 9 ตาราง คือ age sex body symptom health_herb disease herb remedy kind ดังภาพที่ 2



ภาพที่ 2: ER Diagram ของระบบแนะนำสมุนไพรไทยเพื่อสุขภาพ

3.4 การพัฒนาระบบ

พัฒนาระบบจากแนวคิดที่ได้ศึกษาและออกแบบ ซึ่งระบบงานนี้พัฒนาในรูปแบบแอปพลิเคชันบนโทรศัพท์มือถือระบบแอนดรอย์เพื่อช่วยอำนวยความสะดวกให้กับผู้ใช้ โดยใช้โปรแกรม Andriod Studio ภาษา JAVA เป็นภาษาหลักและโปรแกรม MySQL สำหรับการจัดการฐานข้อมูล

3.5 เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ในการวิจัยประกอบด้วย 1) แอปพลิเคชันแนะนำสมุนไพรไทยเพื่อดูแลสุขภาพที่พัฒนาเรียบร้อยแล้ว 2) แบบประเมินประสิทธิภาพแอปพลิเคชันโดยผู้เชี่ยวชาญและ 3) แบบประเมินความพึงพอใจโดยผู้ใช้ทั่วไป

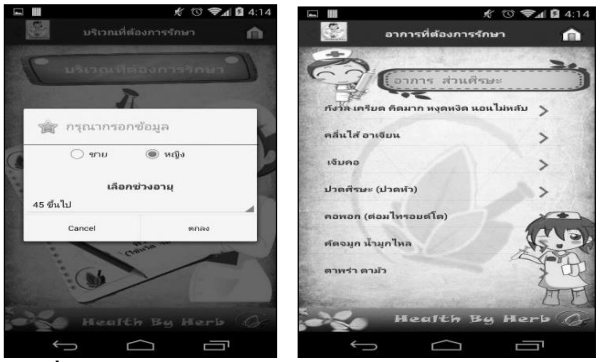
3.6 การประเมินประสิทธิภาพของระบบ

ระบบที่พัฒนาเสร็จเรียบร้อยแล้วจะนำไปทดสอบและประเมินประสิทธิภาพด้วยการทดสอบการทำงานของระบบและใช้แบบสอบถามเพื่อประเมินประสิทธิภาพจากผู้ใช้ 2 กลุ่ม คือ กลุ่มผู้เชี่ยวชาญด้านแพทย์แผนไทยจำนวน 3 คน และกลุ่มผู้ใช้ทั่วไปจำนวน 35 คน การให้คะแนนมีระดับคะแนนตั้งแต่ 1-5 (คะแนน 5 คือดีมาก – คะแนน 1 คือน้อยมาก) ซึ่งเกณฑ์การให้คะแนนสำหรับการสรุปผลการประเมินประสิทธิภาพของระบบ มีดังนี้ 4.51-5.00 ความพึงพอใจของผู้ใช้ที่มีต่อระบบอยู่ในระดับดีมาก 3.51-4.50 ความพึงพอใจของผู้ใช้ที่มีต่อระบบอยู่ในระดับดี 2.51-3.50 ความพึงพอใจของผู้ใช้ที่มีต่อระบบอยู่ในระดับปานกลาง 1.51-2.50 ความพึงพอใจของผู้ใช้ที่มีต่อระบบอยู่ในระดับน้อย 1.00-1.50 ความพึงพอใจของผู้ใช้ที่มีต่อระบบอยู่ในระดับน้อยมาก

4. ผลการดำเนินงาน

4.1 ผลการพัฒนาระบบ

ผลจากการพัฒนาระบบแนะนำสมุนไพรเพื่อสุขภาพสามารถแสดงหน้าจอบางส่วนได้ดังภาพที่ 3



ภาพที่ 3: ผู้ใช้ใส่ข้อมูลเพื่อรับคำแนะนำสมุนไพรสำหรับดูแลสุขภาพ

จากภาพที่ 3 เป็นการทำงานของระบบส่วนการแนะนำสมุนไพรด้วยเทคนิคตารางการตัดสินใจ โดยผู้ใช้จะต้องทำการเลือกปัจจัยสำหรับการวิเคราะห์ข้อมูลคือเพศ อายุ บริเวณร่างกายและอาการป่วย ระบบจะแสดงผลการวิเคราะห์โรคที่เกี่ยวข้องและแนะนำวิธีการรักษาด้วยสมุนไพรไทยเพื่อบรรเทาอาการป่วยเบื้องต้น ดังภาพที่ 4 และภาพที่ 5



ภาพที่ 4: แสดงโรคที่เกี่ยวข้องและวิธีการรักษาเบื้องต้น



ภาพที่ 5: แสดงสมุนไพรที่ใช้รักษาโรคและรายละเอียด

ตัวอย่างหน้าจอแสดงการค้นหาข้อมูลสมุนไพรไทย โดยผู้ใช้สามารถเลือกดูรายละเอียดของสมุนไพรไทยแต่ละชนิดได้ และหน้าจอการเลือกร้านยาสมุนไพร ระบบจะแสดงรายละเอียดของข้อมูลร้านสมุนไพรไทยและแผนที่ ดังภาพที่ 6



ภาพที่ 6: การค้นหาสมุนไพรและข้อมูลร้านขายสมุนไพรพร้อมแผนที่

4.2 ผลการประเมินประสิทธิภาพ

ผลการประเมินประสิทธิภาพของระบบแนะนำสมุนไพรไทยเพื่อสุขภาพบนโทรศัพท์มือถือระบบแอนดรอย์ ได้ใช้การประเมินประสิทธิภาพโดยทดสอบการทำงานของระบบและตอบแบบสอบถาม ซึ่งทำการทดสอบกับ 2 กลุ่มคือกลุ่มผู้เชี่ยวชาญได้แก่บุคลากรจากจัตุรัสมังกร ศูนย์ข้อมูลสมุนไพร คณะเภสัชศาสตร์ มหาวิทยาลัยนเรศวร จำนวน 3 คนและกลุ่มผู้ใช้งานทั่วไปจำนวน 35 คน ดังตารางที่ 5 และตารางที่ 6

ตารางที่ 5: สรุปผลการประเมินประสิทธิภาพโดยผู้เชี่ยวชาญด้านสมุนไพรไทย

หัวข้อที่ประเมิน	ประสิทธิภาพ		
	$\bar{X}$	S.D.	ความหมาย
1. ด้านความสามารถของระบบตรงกับความต้องการ	4.00	0.00	ดี
2. ด้านความสะดวกและง่ายต่อการใช้งานระบบ	4.67	0.47	ดีมาก
3. ด้านความเร็วในการใช้งานระบบ	4.33	0.47	ดี
4.ด้านความถูกต้องของระบบ	4.33	0.47	ดี
5.ด้านประโยชน์ของระบบ	4.67	0.47	ดีมาก
รวม	**4.40**	**0.38**	**ดี**

จากตารางที่ 5 พบว่าความพึงพอใจในประสิทธิภาพระบบจากผู้เชี่ยวชาญในภาพรวมมีค่าเฉลี่ยเท่ากับ 4.40 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.38 อยู่ในระดับดี

ตารางที่ 6: สรุปผลการประเมินประสิทธิภาพโดยผู้ใช้ทั่วไป

หัวข้อที่ประเมิน	ประสิทธิภาพ		
	$\bar{X}$	S.D.	ความหมาย
1. ด้านความสามารถของระบบตรงกับความต้องการ	3.95	0.59	ดี
2. ด้านความสะดวกและง่ายต่อการใช้งานระบบ	4.20	0.60	ดี
3. ด้านความเร็วในการใช้งานระบบ	4.05	0.74	ดี
4.ด้านความถูกต้องของระบบ	4.10	0.54	ดี
5.ด้านประโยชน์ของระบบ	4.25	0.62	ดี
รวม	**4.11**	**0.62**	**ดี**

จากตารางที่ 6 พบว่าความพึงพอใจในประสิทธิภาพระบบจากผู้ใช้ทั่วไปในภาพรวมมีค่าเฉลี่ยเท่ากับ 4.11 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.62 อยู่ในระดับดี ซึ่งผลการประเมินจากผู้เชี่ยวชาญและผู้ใช้ทั่วไป พบว่าหัวข้อประเมินที่ได้คะแนนมากที่สุดเหมือนกันคือด้านความสะดวกและง่ายต่อการใช้งานและด้านประโยชน์ของระบบ แสดงว่าระบบนี้ช่วยอำนวยความสะดวกให้แก่ผู้ใช้และมีประสิทธิภาพ ดังนั้นจึงสามารถนำระบบที่พัฒนาไปใช้งานได้จริง

5. สรุปและข้อเสนอแนะ

การศึกษาวิจัยนี้นำเสนอระบบแนะนำสมุนไพรไทยเพื่อดูแลสุขภาพบนโทรศัพท์มือถือระบบแอนดรอย์ โดยประยุกต์ใช้เทคนิคตารางการตัดสินใจเพื่อให้การแนะนำสมุนไพรไทยที่ใกล้เคียงและเหมาะสมกับผู้ใช้มากที่สุด ซึ่งการพัฒนาระบบแนะนำสมุนไพรไทยเพื่อสุขภาพในรูปแบบแอปพลิเคชันบนโทรศัพท์มือถือให้มีรูปแบบที่ใช้งานง่ายและช่วยอำนวยความสะดวกในการใช้งานเป็นการตอบโจทย์ความต้องการของกลุ่มผู้ใช้ที่ต้องการดูแลสุขภาพและยังช่วยอนุรักษ์ภูมิปัญญาการแพทย์แผนไทยอีกด้วย ผลจากการสำรวจความพึงพอใจ โดยทำการประเมินความพึงพอใจของผู้ใช้งานต่อระบบที่พัฒนาขึ้น ผลที่ได้คือผู้ใช้งานมีความพึงพอใจอยู่ในระดับดีและการประเมินประสิทธิภาพระบบโดยผู้เชี่ยวชาญได้ผลออกมาว่าอยู่ในระดับดีเช่นกัน ดังนั้นจึงสามารถนำระบบแนะนำสมุนไพรไทยนี้ไปใช้งานได้อย่างมีประสิทธิภาพ

ในส่วนของข้อเสนอแนะสำหรับการพัฒนาต่อไปในอนาคตนั้นเนื่องจากงานวิจัยนี้มีข้อจำกัดเรื่อง ข้อมูลปัจจัยในการวิเคราะห์มีจำนวนน้อยส่งผลให้การแนะนำสมุนไพรยังไม่ครอบคลุมกับอาการเจ็บป่วยทั้งหมด ในอนาคตถ้ามีการเพิ่มจำนวนข้อมูลปัจจัยจะทำให้ได้การแนะนำที่มีความสอดคล้องตามความเป็นจริงและสามารถเพิ่มประสิทธิภาพในการแนะนำสมุนไพรไทยให้เหมาะสมกับผู้ใช้มากยิ่งขึ้นอีกด้วย

เอกสารอ้างอิง

[1] สมาคมแพทย์แผนไทยแห่งประเทศไทย. (2559). “การแพทย์แผนไทย”. [ออนไลน์]. สืบค้นเมื่อ 3 มีนาคม 2559. จาก: http://www.utts.or.th

[2] ขวัญชัย สุขแสน, ณัฐวุฒิ มูลศาสตร์และมนัสนันท์ จันโท, “ระบบมัคคุเทศก์พกพาสำหรับการท่องเที่ยว”, *The 10th National Conference on Computing and Information Technology (NCCIT2014)*, ภูเก็ต, ประเทศไทย, 8-9 พฤษภาคม 2557.

[3] พาณี ศิริสะอาด (2559). “ประเด็นของการรักษาด้วยแผนปัจจุบันและแผนไทย”. [ออนไลน์]. สืบค้นเมื่อ 15 มีนาคม 2559. จาก: http://www.pharmacy.cmu.ac.th/web2553/n45.php

[4] ณัฐณิชา วีระมงคลเลิศ, พงษ์ศธร จันทร์ยอยและจักรชัย โสอินทร์, *Android App Development ฉบับสมบุรณ์*, ไอดีซี พรีเมียร์: กรุงเทพฯ, 2555.

[5] สุธีรา พึ่งสวัสดิ์, พงศกร บำรุงไทยและนฤมล มีเอี่ยม, “ระบบแนะนำการเลือกพันธุ์ยางพาราสำหรับพื้นที่ภาคตะวันออกของประเทศไทยโดยใช้ตารางตัดสินใจ”, *The 11th National Conference on Computing and Information Technology (NCCIT2015)*, กรุงเทพฯ, ประเทศไทย, 2-3 กรกฎาคม 2558.

[6] ประภาวดี ศรีสอนและเกียรติศักดิ์ โยชะนัง, “ระบบแนะนำโปรแกรมตรวจสุขภาพกรณีศึกษาโรงพยาบาลเอกชนในเขตกรุงเทพมหานคร”, *The 9th National Conference on Computing and Information Technology (NCCIT2012)*, กรุงเทพฯ, ประเทศไทย, 9-10 พฤษภาคม 2556.

[7] รุจิรา ธรรมสมบัติและทองพูล หีบไธสง, “ระบบสนับสนุนการตัดสินใจเพื่อช่วยค้นหาโรงแรมในอำเภอหัวหินโดยใช้ตารางการตัดสินใจ”, *The 5th National Conference on Computing and Information Technology (NCCIT2009)*, กรุงเทพฯ, ประเทศไทย, 22-23 พฤษภาคม 2552.

[8] iTumApp. (2559).“สมุนไพรไทย สรรพคุณเพื่อสุขภาพ”. [ออนไลน์]. สืบค้นเมื่อ 10 มีนาคม 2559.จาก: https://play.google.com/store/apps/details?id=com.herb.androidherb

[9] DooDroid.com. (2559). “แอปพลิเคชันสมุนไพรไทย-220+ ชนิด”. [ออนไลน์]. สืบค้นเมื่อ 11 มีนาคม 2559. จาก: https://play.google.com/store/apps/details?id=com.doodroid.herb

[10] ธรรมนิตย์, *คัมภีร์ยาสมุนไพรไทย ตำรับหมอพร กรมหลวงชุมพรเขตรอุดมศักดิ์*, ไทยควอลิตี้บุ๊คส์: กรุงเทพฯ, 2552.

ระบบคำนวณมาตรวัดซอฟต์แวร์เพื่อการทดสอบด้วยระบบธุรกิจอัจฉริยะ
Software Metrics Calculation System for Testing with Business Intelligence

ศิริวรรณ มูลเม้า (Siriwan Moolmao)[1] และณัฐวี อุตกฤษฎ์ (Nattavee Utakrit)[2]
ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้า พระนครเหนือ
[1]s5807021857172@email.kmutnb.ac.th, [2]nattavee.u@it.kmutnb.ac.th

บทคัดย่อ

ปัจจุบันการทดสอบระบบของซอฟต์แวร์เป็นส่วนงานที่สำคัญในวงจรการพัฒนาระบบ จึงต้องมีมาตรวัดคุณภาพซอฟต์แวร์ที่มีความชัดเจนและเป็นที่ยอมรับ และจากการวิจัยพบว่ามักใช้การคำนวณค่ามาตรวัดด้วยโปรแกรมไมโครซอฟต์เอกเซล ซึ่งใช้เวลาและมีความเสี่ยงในความแม่นยำ ดังนั้นผู้วิจัยจึงพัฒนาและออกแบบระบบคำนวณมาตรวัดซอฟต์แวร์เพื่อการทดสอบด้วยระบบธุรกิจอัจฉริยะ เพื่อใช้เป็นตัวประเมินและชี้วัดคุณภาพได้อย่างอัตโนมัติ นำไปซึ่งการพัฒนาหรือปรับปรุงซอฟต์แวร์ให้ตอบโจทย์ทางธุรกิจให้มากที่สุด โดยมีการทดสอบความแม่นยำโดยการทดสอบแบบกล่องดำด้วยการเปรียบเทียบผลลัพธ์จากการคำนวณด้วยมือกับผลลัพธ์จากระบบ ได้ค่าความแม่นยำร้อยละ 90.90 และทำการประเมินความพึงพอใจทั้ง 4 ด้านคือ 1) ด้านความสามารถของระบบตรงกับความต้องการ มีค่าเฉลี่ย 4.24 2) ด้านความน่าเชื่อถือใช้งานของระบบ มีค่าเฉลี่ย 4.20 3) ด้านความสะดวกและการใช้งานของระบบ มีค่าเฉลี่ย 4.15 4) ด้านความปลอดภัยในการใช้งานของระบบมีค่าเฉลี่ย 4.50 ซึ่งอยู่ในเกณฑ์ดี

คำสำคัญ: การทดสอบซอฟต์แวร์ มาตรวัดซอฟต์แวร์ ระบบธุรกิจอัจฉริยะ

Abstract

Today, software testing is an important part of the software development lifecycle. Therefore, there must be clear and acceptable software quality metrics. Form the majority of researches, it was found that most were conducted with the metrics from calculation with Microsoft Excel, which takes time and there is a risk with accuracy. Therefore, the researcher designed and developed software metrics calculation system for testing with business intelligence to be automatically used as the quality assessor and indicator. This led to the software development or update to meet the business as much as possible with the accuracy testing done by black box testing by comparing the manual calculated results with the results from the system. The test has an accuracy of 90.90%. The evaluation was also done on the satisfaction of the 4 areas: 1) the system ability to match the needs with the average was at 4.24 2) the system reliability with the average was at 4.20 3) the system convenience and usability with the average at 4.15 4) safety in use of the system and have an average was at 4.50 which was in a good level.

Keyword: Software Testing, Software Metrics, Business Intelligence

1. บทนำ

ปัจจุบันการทดสอบระบบของซอฟต์แวร์นับเป็นส่วนงานที่สำคัญในวงจรพัฒนาระบบ เพื่อให้แน่ใจว่าซอฟต์แวร์มีคุณภาพจึงต้องมีการจัดการทดสอบระบบซอฟต์แวร์ [1] ซึ่งเป็นการตรวจสอบทุกกิจกรรมของการทดสอบ ทำให้มักเจอปัญหาต่างๆ ในระหว่างการทดสอบ และในบางครั้งก็เป็นปัญหาที่พบได้บ่อย [2] โดยเป้าหมายหลักของทีมทดสอบคือการประเมินและตรวจสอบคุณภาพซอฟต์แวร์ ดังนั้นจึงต้องมีมาตรวัดคุณภาพซอฟต์แวร์ที่มีความชัดเจนและเป็นที่ยอมรับ เพื่อใช้ในการสนับสนุนการตัดสินใจในการนำไปวางแผนหรือแก้ไขปัญหาได้อย่างทันท่วงที และจากการวิจัยส่วนใหญ่พบว่ามักใช้การคำนวณค่ามาตรวัดด้วยโปรแกรมไมโครซอฟต์เอกเซล (Microsoft Excel) ซึ่งใช้เวลาและมีความเสี่ยงในความแม่นยำ ดังนั้นผู้วิจัยจึงพัฒนาและออกแบบระบบคำนวณมาตรวัดซอฟต์แวร์เพื่อการทดสอบด้วยระบบธุรกิจอัจฉริยะ เพื่อใช้เป็นตัวประเมินและชี้วัดคุณภาพได้อย่างอัตโนมัติ และเห็นภาพมากกว่าด้วยการแสดงผลออกเป็นรายงาน นำไปซึ่งการพัฒนาหรือปรับปรุงซอฟต์แวร์ให้ตอบโจทย์ทางธุรกิจให้มากที่สุด

จากงานวิจัยของ Kreedy [3] ในการสำรวจเครื่องมือมาตรวัดซอฟต์แวร์ต่างๆ พบว่าโดยส่วนใหญ่แล้วมักใช้คำนวณมาตรวัดจากโปรแกรม Excel ถึงร้อยละ 43 เช่นเดียวกับงานวิจัยของ Gargและ Lakra [4] ที่ใช้โปรแกรม Excel ในการคำนวณค่ามาตรวัดในเรื่องสมรรถภาพของเครือข่ายซึ่งใช้เวลาและมีความเสี่ยงในความแม่นยำ ต่างจากงานวิจัยของ Lincke, Lundberg และ Lowe [5] ที่ใช้เครื่องมือคำนวณอัตโนมัติ แต่ใช้โปรแกรม Excel ในการรวบรวมเก็บผลลัพธ์

ด้วยเหตุดังกล่าวผู้วิจัยจึงทำการพัฒนาระบบคำนวณมาตรวัดซอฟต์แวร์เพื่อการทดสอบด้วยระบบธุรกิจอัจฉริยะ เพื่อช่วยแสดงมาตรวัดมาตรฐานได้อย่างอัตโนมัติ ในการนำไปปรับปรุงคุณภาพการผลิตซอฟต์แวร์ในปัจจุบัน โดยสามารถแยกแยะได้ว่าต้องจัดการหรือเจาะลึกที่ส่วนใด ทำให้บรรลุความพึงพอใจของลูกค้าได้อย่างถูกจุดยิ่งขึ้น

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 การทดสอบซอฟต์แวร์

การทดสอบซอฟต์แวร์ เป็นกระบวนการที่จะช่วยให้ซอฟต์แวร์ที่พัฒนาขึ้นมีความถูกต้อง ความสมบูรณ์ มีความปลอดภัย และมีคุณภาพที่ดี [6] และให้คำนิยามว่า การทดสอบซอฟต์แวร์ คือขั้นตอนหลักของการพัฒนาซอฟต์แวร์อีกด้วย

2.2 เครื่องมือจัดการทดสอบ

เครื่องมือจัดการทดสอบ คือเครื่องมือที่ใช้ในการเก็บข้อมูลเกี่ยวกับการทดสอบ ตั้งแต่วางแผนกิจกรรมการทดสอบ จนไปถึงการออกรายงานสถานะของกิจกรรมการทดสอบ ทำให้มีโอกาสในการเข้าถึงและวิเคราะห์ข้อมูลได้อย่างรวดเร็วและง่ายในการติดต่อสื่อสารกันของหน่วยงานแต่ละโครงการ

2.3 มาตรวัดซอฟต์แวร์เพื่อการทดสอบ

มาตรวัดซอฟต์แวร์ (Software Metric) นั้นเป็นหน่วยวัดมาตรฐานของการประเมินผล มาตรวัดใช้สำหรับการประเมินกระบวนการซอฟต์แวร์ หรือแม้แต่ผลิตภัณฑ์หรือบริการ

ในกระบวณการพัฒนาซอฟต์แวร์ได้มีสูตรมาตรวัดต่างๆ ประกอบไปด้วย มาตรวัดสำหรับรูปแบบการวิเคราะห์ มาตรวัดสำหรับรูปแบบการออกแบบ มาตรวัดสำหรับซอสโค้ด มาตรวัดสำหรับการทดสอบ และมาตรวัดสำหรับการบำรุงรักษา [7]

ประเภทของมาตรวัดคุณภาพซอฟต์แวร์หรือมาตรวัดสำหรับการทดสอบแบ่งออกเป็น 4 ประเภท คือ มาตรวัดการทดสอบด้วยตนเอง มาตรวัดการทดสอบสมรรถภาพ มาตรวัดการทดสอบแบบอัตโนมัติ และมาตรวัดการทดสอบแบบร่วมกัน ซึ่งในการวิจัยนี้มีขอบเขตเฉพาะมาตรวัดการทดสอบด้วยตนเองเท่านั้น

จากงานวิจัยของ Leitner, Ciupa, Meyer และ Howard [8] ได้กล่าวว่า ในการทดสอบด้วยตนเองมีข้อดีในการจับภาพกรณีทดสอบที่ลึกหรือสามารถคาดเดาได้ยาก ซึ่งการทดสอบแบบอัตโนมัติไม่สามารถคาดเดาได้ แต่ในทางกลับกันการทดสอบด้วยตนเองไม่สามารถให้ความครอบคลุมกรณีทดสอบที่เยอะมากได้ ดังนั้นหากมองในแง่ของการทดแทนจะพบว่าการทดสอบแบบอัตโนมัติไม่สามารถคาดเดากรณีที่ลึกหรือคาดเดาได้ยากได้ แต่ในทางกลับกันในเรื่องของความกว้างของกรณีทดสอบ การทดสอบด้วยตนเองสามารถทดแทนได้โดยการเพิ่มเวลาในการทดสอบ หรือเพิ่มจำนวนผู้ทดสอบนั่นเอง

จากบทความของ Gulechha [9] กล่าวว่ามาตรวัดการทดสอบด้วยตนเอง มีทั้งหมด 8 มาตรวัด ประกอบด้วย

2.3.1 Test Case Productivity (TCP) เป็นมาตรวัดที่ช่วยในการเปรียบเทียบแนวโน้มของค่าผลผลิตกรณีทดสอบได้

2.3.2 Test Execution Summary (TES) มาตรวัดนี้จะช่วยในการจัดหมวดหมู่ของกรณีทดสอบที่เกี่ยวสถานะต่างๆ

2.3.3 Defect Acceptance (DA) มาตรวัดนี้เป็นตัวแสดงจำนวนข้อบกพร่องที่ผ่านการตรวจสอบแล้วว่าถูกต้อง ซึ่งระบุโดยทีมงานทดสอบในระหว่างที่มีการปฏิบัติการทดสอบ

2.3.4 Defect Rejection (DR) มาตรวัดนี้เป็นตัวแสดงจำนวนของข้อบกพร่องที่ถูกตีกลับระหว่างที่มีการปฏิบัติการทดสอบว่าทำการระบุไม่ถูกต้อง

2.3.5 Bad Fix Defect (B) มาตรวัดนี้เป็นตัวตรวจสอบประสิทธิภาพของขั้นตอนการแก้ไขข้อบกพร่อง

2.3.6 Test Execution Productivity (TEP) มาตรวัดนี้จะช่วยแสดงแนวโน้มการปฏิบัติการทดสอบ ตั้งแต่ครั้งแรกจนถึงการทดสอบครั้งสุดท้ายที่ผลการทดสอบผ่าน

2.3.7 Test Efficiency (TE) มาตรวัดนี้ตรวจสอบประสิทธิภาพของทีมงานทดสอบในการระบุข้อบกพร่อง ระหว่างการทดสอบ จนถึงเปิดตัวใช้งานโปรแกรม

2.3.8 Defect Severity Index (DSI) มาตรวัดนี้ใช้ตรวจสอบคุณภาพของซอฟต์แวร์ เพื่อใช้การตัดสินใจสำหรับการเปิดตัวโปรแกรม

2.4 ระบบธุรกิจอัจฉริยะและเครื่องมือที่ใช้ในการพัฒนา

ระบบธุรกิจอัจฉริยะ (Business Intelligence : BI) คือซอฟต์แวร์ที่นำข้อมูลที่มีอยู่มาจัดทำรายงานในรูปแบบต่างๆ ที่เหมาะสมกับมุมมองในการวิเคราะห์ และตรงกับความต้องการของผู้ใช้งาน [10]

สำหรับงานวิจัยนี้ ผู้วิจัยเลือกใช้ซอฟต์แวร์ Microsoft SQL Server 2008 R2 เนื่องจากเป็นซอฟต์แวร์ DBMS ที่มีเครื่องมือสร้างรายงานเพื่อใช้ในการตัดสินใจ นั่นคือโปรแกรม SQL Server Reporting Services (SSRS)

3. วิธีการดำเนินงาน

3.1 การศึกษาและรวบรวมข้อมูล

3.1.1 ศึกษาการใช้งานระบบจัดการทดสอบ โดยเลือกใช้โปรแกรม QA Book Desktop โดยศึกษาการใช้งานระบบทุกมอดูล ตั้งแต่การกำหนดค่าเริ่มต้น การจัดการความต้องการ การจัดการกรณีทดสอบ การจัดการการปฏิบัติการทดสอบ การจัดการข้อบกพร่อง และการจัดการสิ่งแวดล้อมในการทดสอบ

3.1.2 ศึกษาและทำความเข้าใจการคำนวณค่ามาตรวัดเพื่อการทดสอบ เพื่อให้สามารถคำนวณค่าเพื่อแสดงออกทางรายงานได้อย่างถูกต้อง ประกอบด้วย

3.1.2.1 Test Case Productivity เป็นมาตรวัดที่ช่วยในเรื่องของการผลิตกรณีทดสอบ โดยมีสูตรการคำนวณคือ

$$Test\ Case\ Productivity = \left[\frac{Total\ Raw\ Test\ Steps}{Efforts\ (hours)}\right] Step(s)/hour$$

วิธีการคำนวณคือทำการนับจำนวนทั้งหมดของขั้นตอนดิบมาหารด้วยชั่วโมงการทำงานต่อวันคูณด้วยวันทำงานจริง ตามตัวอย่างในตารางที่ 1

ตารางที่ 1: ตัวอย่างการนับจำนวนทั้งหมดของขั้นตอนดิบ

ชื่อกรณีทดสอบ	จำนวนขั้นตอนดิบ
ABC_1	30
ABC_2	32
ABC_3	40
ABC_4	36
ABC_5	45
จำนวนทั้งหมดของขั้นตอนดิบ	183

จากตัวอย่างในตารางที่ 1 จะเห็นว่าใช้เวลาเขียนกรณีทดสอบทั้งหมด 183 ขั้นตอน ชั่วโมงการทำงานต่อ 1 วันคือ 8 ชั่วโมง เช่น คำนวณต่อสัปดาห์ มี 5 วัน ดังนั้น TCP = 183/(8x5) = 4.575 หรือผลผลิตกรณีทดสอบประมาณ 5 ขั้นตอน/ชั่วโมง

3.1.2.2 Test Execution Summary มาตรวัดนี้ใช้การคิดสัดส่วนของแต่ละสถานะของกรณีทดสอบเป็นค่าร้อยละ (%)

3.1.2.3 Defect Acceptance มาตรวัดนี้เป็นตัวกำหนดจำนวนของข้อบกพร่อง โดยทีมทดสอบในระหว่างที่มีการปฏิบัติการทดสอบ โดยมีสูตรการคำนวณคือ

$$Defect\ Acceptance = \left[\frac{Number\ of\ Valid\ Defects}{Total\ Number\ of\ Defects} * 100\right]\%$$

วิธีการคำนวณคือทำการนับจำนวนสถานะของข้อผิดพลาดที่ไม่ใช่ “Not a Bug” ซึ่งบ่งบอกว่าระบุข้อผิดพลาดถูกต้อง นำมาหารด้วยจำนวนข้อผิดพลาดทั้งหมด คูณด้วย 100

3.1.2.4 Defect Rejection มาตรวัดนี้เป็นตัวกำหนดจำนวนของข้อบก พร่องที่ถูกตีกลับระหว่างที่มีการปฏิบัติการทดสอบ โดยมีสูตรการคำนวณคือ

$$Defect\ Rejection = \left[\frac{Number\ of\ Defect(s)\ Rejected}{Total\ Number\ of\ Defects} * 100\right]\%$$

วิธีการคำนวณคือทำการนับจำนวนสถานะของข้อผิดพลาดที่มีสถานะ “Not a Bug” ซึ่งบ่งบอกว่าทำการระบุข้อผิดพลาดไม่ถูกต้อง หารด้วยจำนวนข้อผิดพลาดทั้งหมด แล้วคูณด้วย 100

3.1.2.5 Bad Fix Defect มาตรวัดนี้เป็นตัวตรวจสอบประสิทธิภาพการแก้ไขข้อบกพร่อง โดยมีสูตรการคำนวณคือ

$$Bad\ Fix\ Defect = \left[\frac{Number\ of\ of\ Bad\ Fix\ Defect(s)}{Total\ Number\ of\ Valid\ Defects} * 100\right]\%$$

วิธีการคำนวณคือนับจำนวนข้อผิดพลาดที่มีสถานะ “Re-opened” ซึ่งบ่งบอกว่ายังมีข้อผิดพลาดหลังแก้ไขแล้ว จากนั้นหารด้วยจำนวนข้อผิดพลาดที่ระบุถูกต้อง ซึ่งมาจากมาตรวัด DA แล้วคูณด้วย 100

3.1.2.6 Test Execution Productivity มาตรวัดนี้จะบ่งบอกถึงจำนวนการผลิตการปฏิบัติการทดสอบต่อวัน โดยมีเงื่อนไขและการคำนวณ คือ

$$Test\ Execution\ Productivity = \left[\frac{Total\ No.\ of\ TC\ executed\ (Te)}{Execution\ Efforts\ (hours)} * 8\right] Execution(s)/Day$$

ในตารางที่ 2 จะเป็นการแบ่งแยกข้อมูลออกเป็น 4 ส่วนคือ 1) Base Test Case = จำนวนของกรณีทดสอบที่ดำเนินการอย่างน้อยหนึ่งครั้ง 2) T (1) = จำนวนกรณีทดสอบ ที่ต้องทดสอบใหม่ คิดเป็นร้อยละ 71 ถึงร้อยละ 100 ของจำนวนรวมกรณีทดสอบ 3) T (0.66) = จำนวนกรณีทดสอบ ที่ต้องทดสอบใหม่ คิดเป็น 41% ถึง 70% ของจำนวนรวมกรณีทดสอบ 4) T (0.33) = จำนวนกรณี

ทดสอบ ที่ต้องทดสอบใหม่ คิดเป็นร้อยละ 1 ถึง ร้อยละ 40 ของจำนวนรวมกรณีทดสอบ

ตารางที่ 2: ตัวอย่างการจำแนกข้อมูลการปฏิบัติการทดสอบ

TC Name	Base Run Effort (hr)	Re-Run1 Status	Re-Run1 Effort (hr)	Re-Run2 Status	Re-Run2 Effort (hr)	Re-Run3 Status	Re-Run3 Effort (hr)
XYZ_1	2	T(0.66)	1	T(0.66)	0.45	T(1)	2
XYZ_2	1.3	T(0.33)	0.3	T(1)	2		
XYZ_3	2.3	T(1)	1.2				
XYZ_4	2	T(1)	2				
XYZ_5	2.15						

ดังนั้น Te = 5+((1*4)+(2*0.66)+(1*0.33))) = 10.65 จะได้ค่า TEP = (10.65/19.7)*8 = 4.3 การปฏิบัติการ/วัน

3.1.2.7 Test Efficiency มาตรวัดนี้ตรวจสอบประสิทธิภาพของทีมทดสอบในการระบุข้อบกพร่อง ระหว่างการทดสอบ จนเปิดตัวของโปรแกรม โดยมีเงื่อนไขและการคำนวณคือ

$$Test\ Efficiency = \left[\frac{DT}{DT + DU} * 100 \right] \%$$

DT = จำนวนข้อบกพร่องที่ถูกระบุในระหว่างการทดสอบ

DU = จำนวนข้อบกพร่องที่ถูกระบุโดยผู้ใช้หลังจากเปิดตัวโปรแกรม

3.1.2.8 Defect Severity Index มาตรวัดนี้ใช้ตรวจสอบดัชนีความรุนแรงของข้อผิดพลาด เพื่อใช้ตัดสินใจเรื่องเปิดตัวโปรแกรม แบ่งเป็น 2 ส่วน และมีการคำนวณคือ

$$Defect\ Severity\ Index = \left[\frac{\sum (Severity\ Index * No.\ of\ Valid\ Defect(s)\ for\ this\ severity}{Total\ Number\ of\ Valid\ Defects} \right]$$

ส่วนที่ 1 คือ DSI for All Status defect ซึ่งอยู่ภายใต้ขั้นตอนการปฏิบัติการทดสอบ และส่วนที่ 2 คือ DSI for Open Status defect ค่านี้อยู่ในขณะที่เปิดตัวใช้งานระบบจริง

3.1.3 ศึกษาเครื่องมือในการพัฒนาระบบธุรกิจอัจฉริยะ โดยใช้เครื่องมือที่ชื่อ SQL Server Reporting Services (SSRS) และใช้โปรแกรมจัดการฐานข้อมูล คือ Microsoft SQL Server 2012

3.1.4 ศึกษาข้อจำกัดและความเป็นไปได้ของระบบเป็นการศึกษาเกี่ยวกับระบบที่พัฒนาขึ้น เพื่อให้องค์กรที่นำระบบไปใช้งาน สามารถดำเนินการได้สะดวกยิ่งขึ้น

3.2 การวิเคราะห์การทำงานและออกแบบหน้ารายงาน

ทำการวิเคราะห์ว่าเกี่ยวข้องกับขั้นตอนและข้อมูลใดบ้าง เนื่องจากเป็นการนำข้อมูลจากระบบจัดการทดสอบ ของโปรแกรม QA Book Desktop มาใช้ในการคำนวณ จึงไม่จำเป็นต้องทำการออกแบบฐานข้อมูลอีก แต่จำเป็นต้องทำความเข้าใจและวิเคราะห์ว่าในฐานข้อมูลนั้นมีข้อมูลใดบ้างที่มีความเกี่ยวข้อง ซึ่งประกอบด้วยข้อมูลจากเทเบิ้ลต่างๆ ทั้งหมด 27 เทเบิ้ล คือข้อมูลเกี่ยวกับข้อมูลโครงการ ข้อมูลทีมงาน ข้อมูลความต้องการของโครงการ ข้อมูลกรณีทดสอบ ข้อมูลการปฏิบัติการทดสอบ และข้อมูลข้อบกพร่องของระบบ รวมถึงออกแบบรูปแบบการแสดงรายงาน และเงื่อนไขการเรียกดูรายงาน เพื่อให้ตอบโจทย์การใช้งานและง่ายต่อความเข้าใจ เช่นมาตรวัด Test Case Productivity (TCP) จะออกแบบให้เรียกเงื่อนไขเป็นปีที่ต้องการค้นหา และออกแบบให้ผู้ใช้งานสามารถ Drill-Down ข้อมูลดูรายละเอียดย่อยต่างๆ คือดูแบบรายละเอียดแต่ละเดือน หรือรายละเอียดแต่ละโครงการ หรือรายละเอียดแต่ละความต้องการได้

3.3 การสร้างและพัฒนาระบบ

ในขั้นตอนการสร้างและพัฒนาระบบการคำนวณมาตรวัดซอฟต์แวร์เพื่อการทดสอบด้วยระบบธุรกิจอัจฉริยะนั้น เลือกใช้ SSRS โดยใช้เครื่องมือที่ชื่อว่า Business Intelligence Development Studio (BIDS) ซึ่งเป็นเครื่องมือที่นักพัฒนาคุ้นเคยอินเตอร์เฟซกันดีอยู่แล้ว เพราะเป็นอินเตอร์เฟซของ Microsoft Visual Studio เนื่องจากเป็น IDE ตัวเดียวกัน และในส่วนของการจัดการฐานข้อมูลใช้ Microsoft SQL Server 2012

ในขั้นตอนของการพัฒนารายงานแต่ละค่ามาตรวัด จะสร้าง Data Source จากงานวิจัยจะเลือก Embedded Connection แบบ Microsoft SQL Server และระบุ Connection String เป็น DataSource = TK\SQLEXPRESS2 0 0 8 ;Initial Catalog=Vienna ซึ่งเป็นการระบุแหล่งที่มาหรือสร้างกลุ่มข้อมูลที่สามารถนำมาใช้ร่วมกันได้

จากนั้นสร้างความสัมพันธ์ของข้อมูลภายใต้ Data Source จากนั้นทำการสร้าง Data Set โดยกำหนดชื่อ และเลือก Data Source ที่ต้องการเรียกใช้ข้อมูล แล้วกำหนด Query แบบ Text เพื่อทำการกำหนดค่าเป็นคำสั่ง Query ที่ต้องการได้

จากนั้นทำการสร้างรูปแบบรายงานตามที่ออกแบบไว้ ทั้งการแสดงข้อมูลรายงานแบบต่างๆ คือแบบ Tabular หรือแบบ Matrix รวมถึงการแสดงข้อมูลรูปแบบกราฟ ซึ่งสามารถนำข้อมูลฟิลด์ต่างๆ ที่อยู่ใน Data Set นำไปแสดงได้ ซึ่งในแสดงข้อมูลต่างๆ ที่กล่าวถึงตอนต้นนั้น จำเป็นต้องสร้างนิพจน์ (Expression) เพื่อทำการคำนวณค่าจากข้อมูลตามสูตรคำนวณของแต่ละมาตรวัดด้วย เมื่อทำการสร้างรายงานเสร็จแล้ว สามารถสร้างค่าพารามี

เตอร์ (Parameters) ตามที่ออกแบบไว้ เพื่อใช้เป็นเงื่อนไขในการเรียกดูรายงานของผู้ใช้งาน

3.4 การทดสอบความถูกต้องของระบบ

งานวิจัยนี้ทดสอบกับข้อมูลการทดสอบในโครงการต่างๆ ของบริษัทโลคัส เทเลคอมมูนิเคชั่น อิงค์ ประกอบด้วยข้อมูลความต้องการ ข้อมูลกรณีทดสอบ ข้อมูลการปฏิบัติการทดสอบ และข้อมูลข้อบกพร่องของระบบ เพื่อนำมาเป็นข้อมูลดิบในการทำการคำนวณเพื่อหามาตรวัดซอฟต์แวร์ต่างๆ

การประเมินความถูกต้องของระบบ ใช้เทคนิคการทดสอบแบบกล่องดำ โดยเปรียบเทียบระหว่างค่าการคำนวณด้วยมือ กับค่าพยากรณ์ใหม่ที่ได้จากระบบ ตามกรณีทดสอบที่สร้างขึ้น

ผู้วิจัยได้ทำกรณีทดสอบเพื่อเปรียบเทียบการคำนวณระหว่างการคำนวณจากระบบที่พัฒนาขึ้น กับการคำนวณจากโปรแกรม Excel ในแง่ของความรวดเร็วในการคำนวณ

และทำการประเมินความพึงพอใจของผู้ใช้ต่อระบบที่พัฒนาขึ้น โดยใช้มาตราส่วนประมาณค่า (Rating Scale) 5 ระดับ ตามแบบของไลเคิร์ท (Likert Scale) ใช้ผู้ใช้งานระบบทั่วไป 15 คน ประกอบด้วยผู้ใช้งานที่เกี่ยวข้องระบบ คือตำแหน่งผู้จัดการโครงการ (Project Manager), ผู้จัดการงานทดสอบซอฟต์แวร์ (Test Manager), นักพัฒนาระบบ (Programmer) และนักทดสอบซอฟต์แวร์ (Software Tester) ทำการทดสอบ 4 ด้าน คือ 1) ด้านความสามารถของระบบตรงกับความต้องการ 2) ด้านความน่าเชื่อถือใช้งานของระบบ 3) ด้านความสะดวกและการใช้งานของระบบ 4) ด้านความปลอดภัยในการใช้งานของระบบ

หลังจากที่เก็บรวบรวมผลการประเมินที่ได้มาสรุปผลการประเมินความพึงพอใจของระบบ เพื่อใช้แปลความหมายของข้อมูลและนำมาสรุปผลการประเมินความพึงพอใจ เกณฑ์การให้คะแนนของการประเมินระบบงานดังตารางที่ 3

ตารางที่ 3: เกณฑ์การให้คะแนนของการประเมินระบบงาน

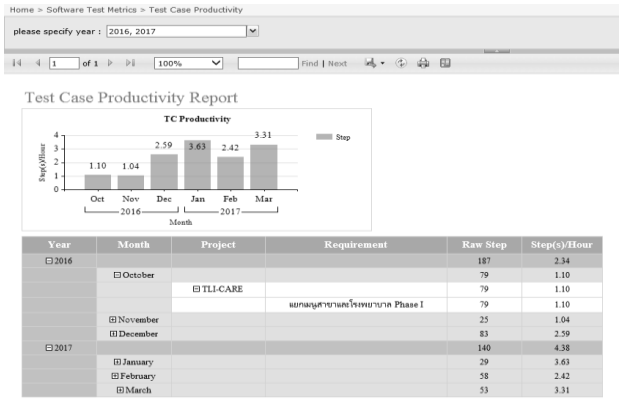
ระดับเกณฑ์การแปลความหมาย		ระดับเกณฑ์การแปลความหมาย	ความหมาย
เชิงคุณภาพ	เชิงปริมาณ		
ดีมาก	5	4.51 – 5.00	มีความพึงพอใจในระดับดีมาก
ดี	4	3.51 – 4.50	มีความพึงพอใจในระดับดี
ปานกลาง	3	2.51 – 3.50	มีความพึงพอใจในระดับปานกลาง
น้อย	2	1.51 – 2.50	มีความพึงพอใจในระดับน้อย
น้อยมาก	1	1.00 – 1.50	มีความพึงพอใจในระดับน้อยมาก

มีการนำวิธีทางสถิติ มาช่วยเพื่อสรุปผลการประเมินความพึงพอใจของระบบ โดยการคำนวณหาค่าเฉลี่ย (Mean) และค่าส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation)

4. ผลการดำเนินงานและการทดสอบระบบ

4.1 ผลการดำเนินงาน

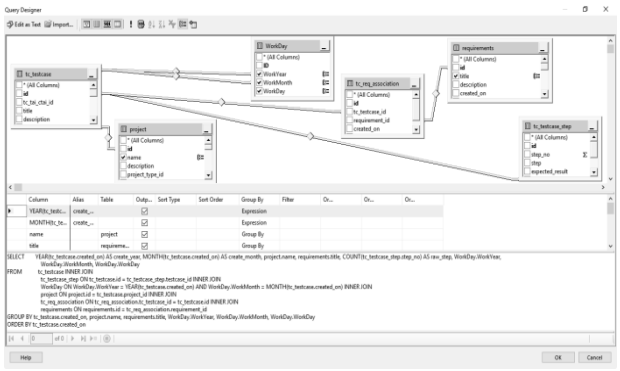
ผลการดำเนินงานจากการพัฒนาระบบคำนวณมาตรวัดซอฟต์แวร์เพื่อการทดสอบ ออกมาในรูปแบบรายงานต่างๆ ตามที่ทำการออกแบบ โดยสามารถเลือกเงื่อนไขการเรียกดูรายงานได้ตามต้องการ และสามารถดูข้อมูลแบบเจาะลึก (Drill-Down) ได้อีกด้วย โดยสามารถดูตัวอย่างได้จากภาพที่ 1 ซึ่งเป็นหน้ารายงานของมาตรวัด Test Case Productivity



Year	Month	Project	Requirement	Raw Step	Step(s)/Hour
⊟2016				187	2.34
	⊟October			79	1.10
		⊟TLI-CARE		79	1.10
			แยกเกณฑ์สาขาและโรงพยาบาล Phase I	79	1.10
	⊞November			25	1.04
	⊞December			83	2.59
⊟2017				140	4.38
	⊞January			29	3.63
	⊞February			58	2.42
	⊞March			53	3.31

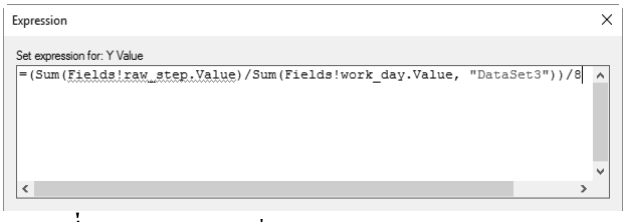
ภาพที่ 1: ตัวอย่างรายงาน Test Case Productivity

โดยการสร้างรายงาน TCP นี้ได้สร้าง DataSet ขึ้นมา 1 รายการ ประกอบด้วยตาราง ข้อมูลเก็บกรณีทดสอบ ข้อมูลเก็บโครงการ ข้อมูลเก็บความต้องการ ข้อมูลเก็บขั้นตอนการทดสอบ จากนั้นนำตารางข้อมูลมาเชื่อมความสัมพันธ์กัน จากภาพที่ 2 ด้านล่างเป็นตัวอย่างการสร้าง Data Set ของรายงานมาตรวัด Test Case Productivity



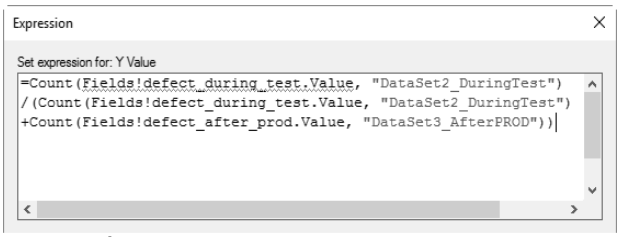
ภาพที่ 2: หน้าจอการสร้าง Data Set ของรายงาน Test Case Productivity

ในส่วนของการคำนวณที่แสดงออกมาในรูปแบบกราฟนั้น จะเลือกข้อมูล Detail เป็นข้อมูลขั้นตอนดิบ และจัดกลุ่มข้อมูลด้วยเดือนที่สร้างขั้นตอนทดสอบ จากนั้นใส่นิพจน์เพื่อกำหนดสูตรคำนวณ จากภาพที่ 3 เป็นตัวอย่างการสร้างนิพจน์ของรายงานมาตรวัด Test Case Productivity

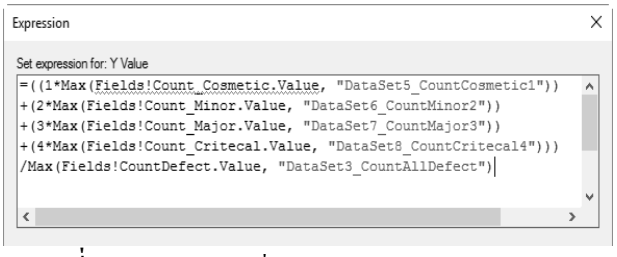


ภาพที่ 3: การสร้างนิพจน์เพื่อคำนวณรายงาน Test Case Productivity

ส่วนการทำรายงานอื่นจะมีเงื่อนไขในการสร้าง Data Set แตกต่างกัน แต่จะมีการสร้างนิพจน์เกิดขึ้นเพื่อทำการคำนวณ จากภาพที่ 4 และ 5 เป็นตัวอย่างนิพจน์ในมาตรวัดอื่นๆ



ภาพที่ 4: การสร้างนิพจน์เพื่อคำนวณรายงาน Test Efficiency



ภาพที่ 5: การสร้างนิพจน์เพื่อคำนวณรายงาน Defect Severity Index

4.2 ผลการทดสอบระบบ

จากการทดสอบความแม่นยำด้วยการทดสอบแบบ Black Box นั้น ผู้วิจัยได้ทำการคำนวณด้วยมือเพื่อเปรียบเทียบผลที่ได้กับในระบบ โดยทดสอบกรณีทดสอบละ 2 ครั้ง ในช่วงเวลาที่ต่างกัน ทำให้จำนวนข้อมูลที่มีแตกต่างกัน ดังตารางที่ 4 และ 5 พบว่าการทดสอบทั้งสองครั้งมีความถูกต้องร้อยละ 90.90 เท่ากัน โดยวัดจากสูตร

$$\frac{\text{[Total No. of Status Passed x 100]}}{\text{Total No. of All Test Case}}$$

ตารางที่ 4: ตัวอย่างกรณีทดสอบเพื่อวัดความแม่นยำของระบบ ครั้งที่ 1

กรณีทดสอบ	ค่าการคำนวณด้วยมือ	ค่าการคำนวณด้วยระบบ	ผลการทดสอบ	หมายเหตุ
ค่า TCP ตามปีที่ต้องการ	0.117	0.117	ผ่าน	-
ค่า TCP ตามปี,เดือนที่ต้องการ	0.065	0.065	ผ่าน	-
ค่า TCP ปี,เดือน,โปรเจคที่ต้องการ	0.065	0.065	ผ่าน	-
ค่า TES ตามโปรเจคที่ต้องการ	2,3,4,18	2,3,4,18	ผ่าน	-
ค่า DA ตามปี,เดือนที่ต้องการ	87.50	87.50	ผ่าน	-
ค่า DR ตามปี,เดือนที่ต้องการ	12.50	12.50	ผ่าน	-
ค่า B ตามปี,เดือนที่ต้องการ	25	25	ผ่าน	-
ค่า TEP ตามปี,เดือนที่ต้องการ	-	-	ไม่ผ่าน	ระบบไม่รองรับ
ค่า TE ตามปี,เดือนที่ต้องการ	88	88	ผ่าน	-
ค่า DSI for All Status ตามปีเดือนที่ต้องการ	2.5	2.5	ผ่าน	-
ค่า DSI for Open Status ตามปีเดือนที่ต้องการ	2.5	2.5	ผ่าน	-

ตารางที่ 5: ตัวอย่างกรณีทดสอบเพื่อวัดความแม่นยำของระบบ ครั้งที่ 2

กรณีทดสอบ	ค่าการคำนวณด้วยมือ	ค่าการคำนวณด้วยระบบ	ผลการทดสอบ	หมายเหตุ
ค่า TCP ตามปีที่ต้องการ	2.34	2.34	ผ่าน	-
ค่า TCP ตามปี,เดือนที่ต้องการ	1.10	1.10	ผ่าน	-
ค่า TCP ปี,เดือน,โปรเจคที่ต้องการ	1.10	1.10	ผ่าน	-
ค่า TES ตามโปรเจคที่ต้องการ	59,2,3,4	59,2,3,4	ผ่าน	-
ค่า DA ตามปี,เดือนที่ต้องการ	87.50	87.50	ผ่าน	-
ค่า DR ตามปี,เดือนที่ต้องการ	12.50	12.50	ผ่าน	-
ค่า B ตามปี,เดือนที่ต้องการ	14.29	14.29	ผ่าน	-
ค่า TEP ตามปี,เดือนที่ต้องการ	-	-	ไม่ผ่าน	ระบบไม่รองรับ
ค่า TE ตามปี,เดือนที่ต้องการ	92.31	92.31	ผ่าน	-
ค่า DSI for All Status ตามปีเดือนที่ต้องการ	2.5	2.5	ผ่าน	-
ค่า DSI for Open Status ตามปีเดือนที่ต้องการ	2.5	2.5	ผ่าน	-

จากตารางที่ 6 ด้านล่าง เป็นการเปรียบเทียบความเร็วในการคำนวณจากระบบที่ทำการพัฒนา กับการคำนวณจาก Excel พบว่าระบบที่ทำการพัฒนาสามารถให้ผลลัพธ์ที่ไวกว่า ถึงแม้ว่าการคำนวณจาก Excel จะสามารถใส่สูตรคำนวณไว้ล่วงหน้าได้ แต่ยังต้องเสียเวลาในการกรอกข้อมูลดิบที่จะนำมาคำนวณ และมีความเสี่ยงหากผู้ใช้งานกรอกข้อมูลเหล่านี้ผิดอีกด้วย

ตารางที่ 6: ตัวอย่างกรณีทดสอบเพื่อเปรียบเทียบกับโปรแกรม Excel

กรณีทดสอบ	เวลาคำนวณจากระบบ (วินาที)	เวลาคำนวณจาก Excel (วินาที)	ผลการทดสอบ
การหาค่า TCP	2	300	ระบบคำนวณไวกว่า
การหาค่า TES	2	180	ระบบคำนวณไวกว่า
การหาค่า DA	2	480	ระบบคำนวณไวกว่า
การหาค่า DR	2	480	ระบบคำนวณไวกว่า
การหาค่า B	2	300	ระบบคำนวณไวกว่า
การหาค่า TE	2	300	ระบบคำนวณไวกว่า
การหาค่า DSI for All Status	2	600	ระบบคำนวณไวกว่า
การหาค่า DSI for Open Status	2	600	ระบบคำนวณไวกว่า

หลังจากให้ผู้ใช้งานระบบทั่วไป 15 คน ใช้งานระบบแล้วพบว่าประเมินความพึงพอใจทั้ง 4 ด้าน คือ 1) ด้านความสามารถของระบบตรงกับความต้องการ มีค่าเฉลี่ย 4.34 และส่วนเบี่ยงเบนมาตรฐาน 0.65 2) ด้านความน่าเชื่อถือใช้งานของระบบ มีค่าเฉลี่ย 4.20 และส่วนเบี่ยงเบนมาตรฐาน 0.54 3) ด้านความสะดวกและการใช้งานของระบบ มีค่าเฉลี่ย 4.23 และส่วนเบี่ยงเบนมาตรฐาน 0.64 4) ด้านความปลอดภัยในการใช้งานของระบบมีค่าเฉลี่ย 4.62 และส่วนเบี่ยงเบนมาตรฐาน 0.61เมื่อนำค่าเฉลี่ยที่ได้มาทำการแปลงความหมายข้อมูลจากเกณฑ์การแปลงความหมายข้อมูลและพิจารณาจากค่าเฉลี่ย ในตารางที่ 3 พบว่ามีความพึงพอใจต่อระบบในระดับดี

5. สรุปผล

หลังใช้ระบบ ทำให้สามารถจัดการคุณภาพของซอฟต์แวร์ได้ดีขึ้น เนื่องจากสามารถมองเห็นแนวโน้มในด้านต่างๆ จากตัวรายงาน และยังคำนวณได้รวดเร็วกว่าการใช้โปรแกรม Excel เนื่องจากไม่ต้องเสียเวลากรอกข้อมูลดิบ โดยในส่วนฟังก์ชั่นการใช้งานของโปรแกรม QA Book Desktop นั้นค่อนข้างสนับสนุนการคำนวณมาตรวัดซอฟต์แวร์เพื่อการทดสอบ แต่พบว่ามีบางมาตรวัดที่ระบบยังไม่สนับสนุน นั่นคือมาตรวัด TEP เนื่องจากระบบยังไม่มีการเก็บข้อมูลเกี่ยวกับความคืบหน้าของการปฏิบัติการทดสอบ กรณีทดสอบหลังพบข้อผิดพลาดในระบบ

6. ข้อเสนอแนะ

ในกระบวณการพัฒนาซอฟต์แวร์ สามารถชี้วัดค่าต่างๆ ได้หลายประเภท จากการวิจัยได้ทำการชี้วัดเพียงตัวประเภทเดียวคือสำหรับการทดสอบ ดังนั้นหากมีการพัฒนาเพิ่มในส่วนอื่น จะช่วยให้การชี้วัดคุณภาพซอฟต์แวร์เป็นไปได้อย่างมีประสิทธิภาพและครอบคลุมยิ่งขึ้น และผลจากการศึกษาวิจัยในครั้งนี้ยังสามารถนำไปต่อยอดหรือประยุกต์ใช้ในคำนวณหาค่ามาตรวัดเพื่อการทดสอบ แบบอัตโนมัติได้ เนื่องจากโปรแกรม QA Book Desktop สามารถสนับสนุนการใช้งานในเรื่องนี้ได้

เอกสารอ้างอิง

[1] Sheena Kukreja, Abhishek Singhal and Abhay Bansal, “A Critical Survey on Test Management in IT Projects,” *International Conference on Computing, Communication and Automation (ICCCA2015)*, Department of CSE, ASET Amity University, pp. 791 ,2015.

[2] Buppha Wilaiphan, What–Why–Who–How Software Testing, Retrieved from http://wcm.gable.com/portal/page/portal/gable/thai/it_talks/Y2013/it_talks_V34_01/G-Magz_V34_1.pdf

[3] Aika Ibraheem Kreedy, “Metrics Tools For Reengineering Software Projects,” *International Journal of Scientific & Technology Research*, vol.3, pp. 361-363, 2014.

[4] Garg and Lakra, "A Software Metric for the Activeness of a Local Area Network," *International Journal of Computer Application*, vol.156, pp.11, 2016.

[5] Rüdiger Lincke, Jonas Lundberg and Welf Löwe, Comparing Software Metrics Tools, Retrieved from http://arisa.se/files/LLL-08.pdf

[6] อุไร ทองหัวไผ่, "การทดสอบซอฟต์แวร์." วารสารเกษมบัณฑิต, ฉบับที่ 16, 2558.

[7] Dinesh Thakur, Classification of Software Metrics in Software Engineering, Retrieved from http://ecomputernotes.com/software-engineering/classification-of-software-metrics

[8] Leitner, A., Ciupa, I., Meyer, B., Howard, M., "Reconciling Manual and Automated Testing: the AutoTest Experience," in Proceedings of the 40th Hawaii International Conference on System Sciences (Waikoloa, Big Island, Hawaii), January, 2007.

[9] Lokesh Gulechha, Software Testing Metrics, Retrieved from http://artemisa.unicauca.edu.co/~cardila/CS__Software_Testing_Metrics.pdf

[10] มลิวัลย์ แซ่อึ้ง, "ตอนที่ 1 : มารู้จัก SSRS กันเถอะ," บล็อกของชุมชนเครือข่ายผู้ปฏิบัติงานไอที มหาวิทยาลัยสงขลานครินทร์

การพัฒนาระบบสารสนเทศสำหรับให้บริการการบำรุงรักษารถยนต์ โดยใช้เทคนิคการรักษาลูกค้าสัมพันธ์

The Development of Information System for Fleet Maintenance Using Customer Retention Technique

ภัชรีญา มีทา (Phatchareeya Meetha)[1] และศักดิ์ชาย ตั้งวรรณวิทย์ (Sakchai Tangwannawit)[2]
ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ
มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
[1]s5807021857121@email.kmutnb.ac.th, [2]sakchai.t@it.kmutnb.ac.th

บทคัดย่อ

งานวิจัยชิ้นนี้มีวัตถุประสงค์เพื่อพัฒนาระบบสารสนเทศสำหรับให้บริการการบำรุงรักษารถยนต์ โดยใช้เทคนิคการรักษาลูกค้าสัมพันธ์ การพัฒนางานวิจัยนี้ใช้ทฤษฎีการบริหารลูกค้าสัมพันธ์ในด้านกลยุทธ์การรักษาลูกค้าและความต้องการของลูกค้า นำมาพัฒนาเป็นระบบสารสนเทศสำหรับให้บริการการบำรุงรักษารถยนต์ ระบบถูกพัฒนาขึ้นในลักษณะ Web Application ซึ่งภายในระบบสารสนเทศสำหรับให้บริการการบำรุงรักษารถยนต์นี้ประกอบไปด้วย 7 โมดูลได้แก่ โมดูลการลงทะเบียน โมดูลการลงทะเบียนสมาชิก โมดูลจัดการอะไหล่ โมดูลจัดการเคมีภัณฑ์ โมดูลสอบถามค่าใช้จ่ายด้านผลิตภัณฑ์ โมดูลรายงานการเข้าบำรุงรักษา และโมดูลรายงานค่าใช้จ่าย ผลที่ได้จากการประเมินการยอมรับและการใช้เทคโนโลยีในการพัฒนาระบบสารสนเทศ จากผู้เชี่ยวชาญและผู้ใช้งานทั่วไป พบว่า ให้การยอมรับการใช้งานและระบบสามารถใช้งานได้ดีและมีประโยชน์ ซึ่งผลการประเมินการยอมรับและการใช้เทคโนโลยีในการพัฒนาระบบสารสนเทศเพื่อการบำรุงรักษายานพาหนะโดยใช้เทคนิคการรักษาลูกค้า ใน 4 ด้าน ได้แก่ 1. ปัจจัยการรับรู้ผลประโยชน์ ค่าเฉลี่ยเท่ากับ 4.43 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.24 2. ปัจจัยการรับรู้ความยากง่ายในการใช้งาน ค่าเฉลี่ยเท่ากับ 4.41 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.30 3. ทัศนคติต่อการใช้ ค่าเฉลี่ยเท่ากับ 4.52 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.26 และ 4. พฤติกรรมความตั้งใจ ค่าเฉลี่ยเท่ากับ 4.55 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.22 สรุปได้ว่า ระบบที่พัฒนาขึ้นมีความพึงพอใจอยู่ในระดับดี

คำสำคัญ: ลูกค้า การบำรุงรักษารถยนต์ กลยุทธ์การรักษาลูกค้า

Abstract

This purpose of this research was to the development of information system for fleet maintenance using customer retention technique. The development of this research by customer relationship management theory, customer retention strategy and customer requirements for the development of fleet maintenance system. Web application development is the process and practice of developing web applications. The fleet maintenance system was composed of seven modules, register module, member module, spare parts module, consumable part (chemical) module, expense inquiry module, maintenance report module, and cost report module. The results of system from experts and general users found that. Showed that users were acceptance with the system, which is very usable and useful. The evaluation results of acceptance and use of technology of to the development of information system for fleet maintenance using customer retention technique. In 4 areas: 1.Perceived usefulness $\bar{x}$ = 4.43 , S.D. = 0.24 2. Perceived Ease of Use $\bar{x}$ = 4.41 , S.D. = 0.30 3. Attitude Toward Using $\bar{x}$ =4.52 , S.D. = 0.26 and 4. Behavioral Intention $\bar{x}$ =4.55 , S.D. = 0.22 In overall, satisfactions of the developed system appeared at a good level.

Keyword: Customer, Fleet Maintenance, Customer Retention Strategy

1. บทนำ

บริษัท โตโยต้า มอเตอร์ ประเทศไทย จำกัด ได้ตระหนักถึงความสำคัญในทุกการติดต่อของลูกค้า รวมทั้งรับฟังเสียงของลูกค้าเสมือนหนึ่งการให้บริการ จากการสอบถามลูกค้าการจากเข้ารับการตรวจเช็คระยะและการเข้าศูนย์บริการเพื่อทำการดูแลรักษา ณ ศูนย์บริการลูกค้า TOYOTA Petra ซึ่งเป็นการสำรวจโดยการพูดคุยกับลูกค้าที่มาใช้บริการภายในศูนย์ พบปัญหาที่เกิดจากความต้องการของลูกค้าในเรื่องการเข้าถึงข้อมูลค่าใช้จ่ายที่เกี่ยวข้องกับผลิตภัณฑ์และค่าบริการด้านงานซ่อมบำรุง ซึ่งเป็นในส่วนของการบริการหลังการขายผ่านระบบงานออนไลน์ นอกเหนือจากการสอบถามข้อมูลผ่านทาง Call Center ซึ่งลูกค้าสามารถเข้าถึงและตรวจสอบข้อมูลได้ง่ายและสามารถตอบสนองต่อความต้องการของลูกค้าที่อยู่ในช่วงการสื่อสารในยุคดิจิทัล

จากปัญหาที่พบจากการให้บริการขั้นต้น ซึ่งเกิดจากช่องโหว่ในการให้บริการลูกค้าที่ไม่ครอบคลุมในทุก ๆ ด้าน และมีที่มาจากระบบการทำงานในรูปแบบเดิม ที่ส่งผลให้ไม่เอื้ออำนวยต่อความต้องการของลูกค้า ที่ต้องการการทำงานในรูปแบบออนไลน์ จึงเป็นที่มาของงานวิจัยนี้ เพื่อพัฒนาให้ระบบมีความน่าสนใจและตอบสนองต่อลูกค้าออนไลน์ได้อย่างแท้จริง จึงเกิดแนวคิดในการการพัฒนาระบบสารสนเทศสำหรับให้บริการการบำรุงรักษารถยนต์ โดยใช้เทคนิคการรักษาลูกค้าสัมพันธ์ ให้เกิดความประทับใจและเป็นการรักษาลูกค้าในระยะยาว

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 การบริหารลูกค้าสัมพันธ์

(Customer Relationship Management : CRM) คือ การสร้างความสัมพันธ์อันดีแก่ลูกค้า โดยการสร้างมูลค่าที่เกิดขึ้นจากการสร้างความสัมพันธ์ที่ดีเหนือกว่าคู่แข่ง จากการวิเคราะห์ในมุมมองของการตลาดสามารถจำแนกสาเหตุที่ทำให้เกิด CRM ได้ดังต่อไปนี้ 1. ลูกค้าพร้อมตีจากแม้ว่าจะพอใจ 2. สถานการณ์การแข่งขันที่รุนแรง 3. ลูกค้ามีทางเลือกไม่จำกัด 4. ลูกค้าภักดีต่อหลายตรายี่ห้อในเวลาเดียวกัน (Multi-brand Loyalty) 5. ร้อยละ 80 ของกำไรทั้งหมดมาจากลูกค้าเพียงร้อยละ 20 (Pareto's Rule of 80/20) 6. ปริมาณข้อมูลจำนวนมากในการดำเนินธุรกิจ แต่มีการนำมาใช้ประโยชน์น้อยมาก [1]

2.2 กลยุทธ์การรักษาลูกค้า (Customer retention strategy)

กลยุทธ์การรักษาลูกค้า (Customer retention strategy) เป็นเทคนิคหนึ่งของงานด้าน CRM โดยใช้วิธีการประเมินผลการบริหารลูกค้าสัมพันธ์และการรักษาความเติบโตของความสัมพันธ์ของลูกค้า โดยมีวัตถุประสงค์เพื่อสร้างความประทับในในตราสินค้าและสร้างสัมพันธภาพที่ดีกับลูกค้าปัจจุบัน [2]

ซึ่งวิธีการรักษาลูกค้าจะแสดงให้เห็นถึงความสัมพันธ์ระหว่างการหาลูกค้ารายใหม่กับการรักษาลูกค้าเก่า (Customer acquisition and customer retention link) โดยมีขั้นตอนการรักษาลูกค้าที่เกี่ยวข้องกับ 7 ประเด็นดังต่อไปนี้ 1. การสร้างความเข้าใจในเรื่องความคาดหวังของลูกค้ากับคุณภาพที่ส่งมอบให้กับลูกค้า (Customer expectation and delivered quality) 2. คุณค่า (Value) 3. เอกลักษณ์ผลิตภัณฑ์กับความเหมาะสม (Product uniqueness and suitability) 4. การสร้างกลไกเพื่อสร้างความภักดีต่อตัวผลิตภัณฑ์ (Loyalty mechanisms) 5. ความสะดวกสบายในการซื้อหรือรับบริการ (Ease of purchase) 6. การบริการลูกค้า (Customer service) 7. การสร้างอุปสรรคไม่ให้ลูกค้าเลิกใช้ผลิตภัณฑ์ (Exit barrier) ในปัจจุบันธุรกิจต่าง ๆ ให้การยอมรับในเรื่องการสร้างความพอใจให้กับลูกค้าและการรักษาลูกค้าที่มีอยู่มากขึ้น เพราะการรักษาซึ่งลูกค้าไว้ด้วยการทำให้ลูกค้ามีความพึงพอใจและมั่นใจในด้านการให้บริการสร้างความสัมพันธ์อันดีที่ได้รับจากธุรกิจ จะส่งผลให้ลูกค้ามีแนวโน้มที่จะซื้อบริการหรือผลิตภัณฑ์จากธุรกิจเพิ่มมากขึ้น [3]

กมลวรรณ [4] ได้พัฒนาระบบลูกค้าสัมพันธ์ สำหรับเกรย์มาร์เก็ต กรณีศึกษา บริษัทเทดดี้ ออโต้ เซลส์ จำกัด ซึ่งพบปัญหาจากการทำงานของฝ่าย CRM และฝ่ายขาย ซึ่งเมื่อทำการพัฒนาระบบขึ้นพบว่าข้อมูลซึ่งเดิมที่ต้องใช้ในการทำงานร่วมกันแต่ไม่สัมพันธ์กันเมื่อพัฒนาให้มีความพันธ์ร่วมกันแล้วการทำงานได้เกิดความรู้ใหม่ ๆ เกิดขึ้น สามารถเรียกใช้ข้อมูล

ที่ใช้ร่วมกันเป็นชุดเดียวกัน ส่งผลให้มีประสิทธิ์ภาพในการทำงานเพิ่มมากขึ้น

ลินดา [5] ได้พัฒนาระบบการจัดการองค์ความรู้ เรื่อง ระบบสารสนเทศภายในองค์กร โดยจัดทำระบบแสดงการค้นหาปัญหาและแนวทางการแก้ไขปัญหา เพื่อค้นหาข้อมูลแนวทางการแก้ไขปัญหา โดยนำมาช่วยในการปฏิบัติงานของพนักงานในองค์กร โดยได้มีการรวบรวมองค์ความรู้ที่มีอยู่อย่างกระจัดกระจาย มาพัฒนาให้เป็นระบบ เพื่อให้ทุกคนในองค์กรสามารถเข้าถึงความรู้และ นำความรู้ที่ได้ไปประยุกต์ใช้ในการปฏิบัติงานให้เกิดประสิทธิภาพ

พัชรพล [6] ได้พัฒนาระบบการผสมผสานการวิเคราะห์แบบหลายมิติ และวิถีความเป็นเลิศแบบโตโยต้า กับระบบรายงานผลการตรวจสอบคุณภาพงานซ่อมรถยนต์ โดยได้เลือกประยุกต์ระบบสารสนเทศเพื่อการจัดการ (Management Information Systems) ควบคู่กับการประมวลผลในเชิงวิเคราะห์แบบออนไลน์ (On-Line-Analytic Processing : OLAP) และวิถีความเป็นเลิศแบบโตโยต้าในการจัดทำระบบรายงานผลการตรวจสอบคุณภาพงานซ่อมรถยนต์

O.Khadeer [7] ประยุกต์ธุรกิจอัจฉริยะและระบบการบริหารลูกค้าสัมพันธ์เข้าด้วยกัน และการทำเหมืองข้อมูลถูกนำมาใช้ในการเก็บรวบรวมข้อมูลจากระบบการบริการลูกค้าสัมพันธ์ซึ่งมีวิธีการคือ การแสดงข้อมูลเพื่อที่จะสามารถจับความสัมพันธ์ของลูกค้ากับธุรกิจขององค์กรและทำการกำหนดความสัมพันธ์ของลูกค้ากับองค์กรโดยใช้ข้อมูลจากฐานความรู้เดิมที่มีอยู่ภายในองค์กรและนำมากำหนดกลยุทธ์กับลูกค้ารายอื่นที่มีพฤติกรรมคล้ายเคียงกัน

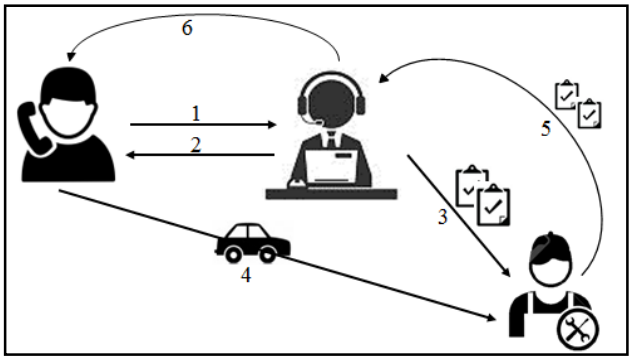
P. Ing [8] ได้พัฒนาระบบการจัดการความสัมพันธ์กับลูกค้าโดยเน้นความเรียบง่ายบนความปลอดภัย โดยใช้สถาปัตยกรรมระบบกระจาย โดยที่ผ่านมาใช้วิธีการแบ่งกลุ่มผู้ซื้อ และการวัดความพึงพอใจของลูกค้า โดยที่องค์กรไม่ได้มุ่งเน้นลูกค้าขององค์กร ซึ่งต้องศึกษาพฤติกรรมของลูกค้าและการแสวงหาโอกาสที่จะมีอิทธิพลต่อพฤติกรรมดังกล่าวผ่านรูปแบบต่าง ๆ ของการสื่อสาร ซึ่งจะสามารถดึงดูดลูกค้าใหม่ ๆ และสามารถรักษาลูกค้ารายเดิม ส่งผลให้เกิดประโยชน์ที่กับองค์กรและสร้างความภักดีของลูกค้าต่อองค์กรได้

3. วิธีการดำเนินงาน

วิธีการในการดำเนินงานวิจัยของระบบสารสนเทศสำหรับให้บริการการบำรุงรักษารถยนต์ โดยใช้เทคนิคการรักษาลูกค้าสัมพันธ์ มีขั้นตอนการดำเนินงานตามหลักวงจรการพัฒนาระบบ (System Development Life Cycle: SDLC) ซึ่งมีขั้นตอนดังนี้

3.1 การศึกษาและรวบรวมข้อมูลระบบ

ศึกษารายละเอียดเกี่ยวกับธุรกิจการให้บริการด้านการบำรุงรักษารถยนต์ โดยเฉพาะการเข้ารับบริการการบำรุงรักษารถยนต์หลังการขาย ดังภาพที่ 1



ภาพที่ 1: การทำงานของระบบงานเดิม

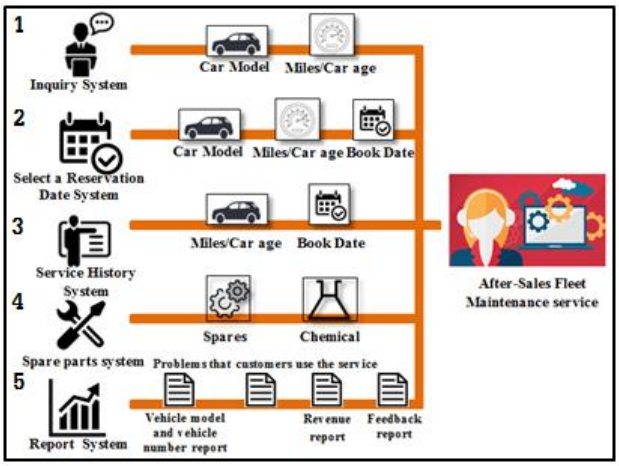
จากภาพการทำงานของระบบงานเดิมสามารถอธิบายรูปแบบการทำงานของระบบงานเดิมได้พอสังเขปคือ 1. เมื่อลูกค้าต้องการนำรถเข้ามาตรวจเช็คสภาพตามระยะ/รอบบริการ ลูกค้าต้องทำการติดต่อเข้ามาเพื่อสอบถามข้อมูลค่าใช้จ่ายเบื้องต้นแก่พนักงานของศูนย์บริการและแจ้งเวลาที่ลูกค้าต้องการนำรถเข้าศูนย์บริการ 2. พนักงานจะทำการตรวจสอบข้อมูลที่ลูกค้าต้องการนำข้อมูลการขอเข้ารับบริการจัดส่งเป็นเอกสารให้แก่แผนกบำรุงรักษา 3. เมื่อแผนกบำรุงรักษาได้รับเอกสาร 4. รับรถของลูกค้ามาและทำการประเมินค่าใช้จ่ายที่แท้จริง 5. ช่างซ่อมบำรุงจะทำการส่งมอบเอกสารผ่านทางพนักงาน 6. พนักงานจะนำเอกสารค่าใช้จ่ายส่งมอบให้แก่ลูกค้า ซึ่งจากระบบงานเดิมพบว่าการทำงานที่เป็นอยู่ในปัจจุบันมีความซ้ำซ้อนและหลายขั้นตอน อีกทั้งลูกค้าไม่สามารถเรียกดูรายงานการซ่อมบำรุงผ่านระบบออนไลน์ได้

ซึ่งปัญหาดังกล่าวสามารถนำมากำหนดขอบเขตของปัญหาที่เกิดขึ้นและกำหนดความต้องการ (Requirements) ของลูกค้าร่วมกับผู้ปฏิบัติงาน โดยข้อมูลดังกล่าวได้ทำการสำรวจข้อมูล

จากการสัมภาษณ์พนักงานและความต้องการที่เกิดจากลูกค้าที่เข้ามาใช้บริการ จากนั้นสำรวจการทำงานด้านการบำรุงรักษาภายในศูนย์บริการ แล้วจึงทำการสรุปปัญหาออกมาเป็นข้อกำหนด (Requirement Specification)

3.2 วิเคราะห์และออกแบบระบบ

การวิเคราะห์และออกแบบระบบสารสนเทศเพื่อให้บริการในด้านการสอบถามข้อมูลค่าใช้จ่ายงานบำรุงรักษารถยนต์ โดยใช้วิธีการบริหารงานลูกค้าสัมพันธ์ เป็นขั้นตอนที่จะทำให้การพัฒนาระบบมีประสิทธิภาพ การวิเคราะห์การทำงานต่าง ๆ ตามโมดูลการบำรุงรักษารถยนต์ ดังภาพที่ 2



ภาพที่ 2: โมดูลการบำรุงรักษารถยนต์

จากภาพที่ 2 โมดูลการบำรุงรักษารถยนต์สามารถอธิบายการทำงานได้ตามลำดับหมายเลข 1-5 ได้ดังต่อไปนี้

หมายเลข 1 โมดูลสอบถามค่าใช้จ่ายด้านผลิตภัณฑ์ เป็นโมดูลที่นำข้อมูลรุ่นรถยนต์, ข้อมูลด้านผลิตภัณฑ์, ข้อมูลค่าใช้จ่ายมาเชื่อมโยงความสัมพันธ์ และใช้เงื่อนด้านอายุรถยนต์และเลขระยะทางเพื่อแสดงผลให้ที่ถูกต้องให้แก่ผู้ใช้งาน

หมายเลข 2 โมดูลตารางปฏิทินและการจองคิวเข้าใช้บริการเป็นโมดูลที่นำโมดูลสอบถามค่าใช้จ่ายด้านผลิตภัณฑ์มารวมกับโมดูลตารางปฏิทินเพื่อทำการจองคิว ซึ่งลูกค้าจะต้องเป็นสมาชิกของระบบและทำการเพิ่มรุ่นรถยนต์จากนั้นจึงจะสามารถใช้ตารางปฏิทินได้

หมายเลข 3 โมดูลรายงานการเข้าบำรุงรักษา เป็นโมดูลที่เก็บข้อมูลรถ, ข้อมูลรายงานค่าใช้จ่าย และข้อมูลวันที่เข้ารับบริการของลูกค้าไว้ภายในระบบ

หมายเลข 4 โมดูลจัดการอะไหล่เป็นโมดูลที่เชื่อมโยงความสำพันธ์ระหว่างโมดูลจัดการเคมีภัณฑ์ และค่าบริการรุ่นรถยนต์แต่ละประเภทมาทำการผูกเงื่อนไขให้ตรงตามข้อมูลรุ่นรถยนต์แต่ละรุ่นรถ

หมายเลข 5 โมดูลรายงานการเข้าบำรุงรักษาเป็นโมดูลที่เก็บข้อมูล 4 ด้านคือ 1. รุ่นรถและจำนวนรถที่เข้ามาใช้บริการ 2. ปัญหาที่ลูกค้าเข้ามาใช้บริการ 3. รายได้ที่เกิดขึ้นหลังจากลูกค้าเข้าศูนย์บริการ และ 4. ผลตอบรับหลังจากการติดตามหลังการเข้าใช้บริการของลูกค้า ซึ่งเป็นระบบที่รวบรวมการทำงานที่รวบรวมการทำงานทั้งหมดหลังจากได้ให้บริการแก่ลูกค้าเก็บไว้ยังฐานข้อมูล

3.3 การพัฒนาระบบ

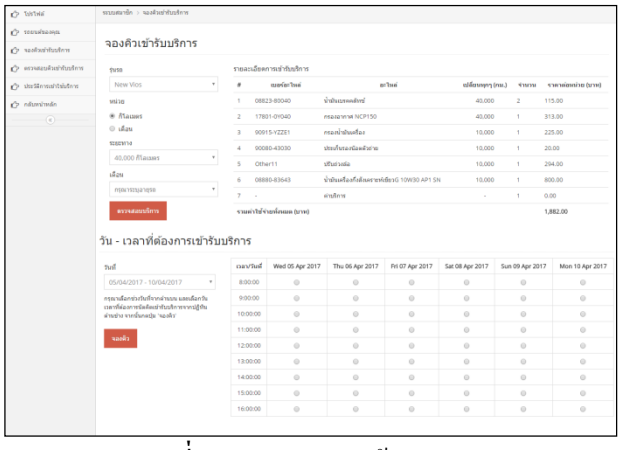
การพัฒนาระบบสารสนเทศเพื่อให้บริการในด้านการสอบถามข้อมูลค่าใช้จ่ายงานบำรุงรักษารถยนต์ โดยใช้วิธีการบริหารงานลูกค้าสัมพันธ์ใช้ MySQL และ PHP เป็นเครื่องมือในการพัฒนาให้อยู่ในรูปแบบเว็บแอพพลิเคชั่น ตัวอย่างหน้าระบบที่ได้ถูกพัฒนาขึ้นมีดังนี้

3.3.1 เมนูตรวจสอบค่าบริการตามระยะการใช้งาน ประกอบไปด้วย เมนูสอบถามข้อมูลค่าใช้จ่ายโดยที่ลูกค้าสามารถตรวจสอบได้จากการเลือกรุ่นรถยนต์ โดยสามารถใช้เงื่อนไขจากเลขไมล์ปัจจุบันหรือระยะเวลาที่ได้ใช้งานในปัจจุบันดังภาพที่ 3



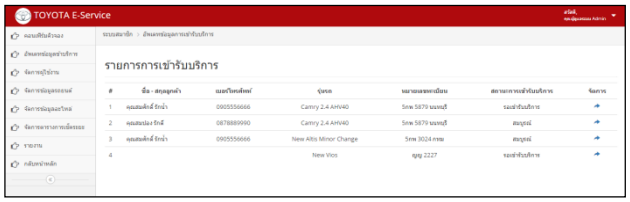
ภาพที่ 3: เมนูสอบถามข้อมูลค่าใช้จ่าย

3.3.2 เมนูจองเวลาเพื่อขอเข้ารับบริการ ประกอบไปด้วย ข้อมูลรถยนต์ของลูกค้าและวันที่ที่ลูกค้าต้องการนำรถยนต์เข้ามาใช้บริการภายในศูนย์บริการ โดยระบบจะแสดงค่าใช้จ่ายเบื้องต้นให้แก่ลูกค้าดังภาพที่ 4



ภาพที่ 4: เมนูจองคิวขอเข้ารับบริการ

3.3.1 เมนูยืนยันการขอเข้ารับบริการ เมนูนี้จะแสดงข้อมูลที่สำคัญของลูกค้าประกอบด้วย ชื่อ – นามสกุล, เบอร์โทรศัพท์, รุ่นรถ, หมายเลขทะเบียน, และแสดงสถานะการเข้ารับบริการ ดังภาพที่ 5



ภาพที่ 5: เมนูยืนยันการขอเข้ารับบริการ

3.4 การทดสอบประเมินผลของระบบ

เมื่อพัฒนาระบบและทำการทดสอบระบบให้สามารถทำงานได้ตรงตามวัตถุประสงค์ที่กำหนดไว้เป็นที่เรียบร้อยแล้ว จะทำการประเมินความพึงพอใจของระบบจากการทดสอบความสมบูรณ์ของระบบโดยผู้พัฒนา ซึ่งต้องใช้ข้อมูลสมมติในการทดสอบ และทำการประเมินความพึงพอใจระบบ โดยผู้ใช้งานทั่วไปและลูกค้าที่เข้ามาใช้บริการภายในศูนย์บริการทั้งหมดจำนวน 15 ท่าน เพื่อวัดการยอมรับการใช้งานที่มีต่อระบบ และทำการวัดความพึงพอใจที่มีต่อระบบ โดยผู้เชี่ยวชาญจำนวน 5 ท่าน โดยใช้แบบประเมินในการเก็บรวบรวมข้อมูล

4. ผลการดำเนินงาน

การพัฒนาระบบสารสนเทศสำหรับให้บริการการบำรุงรักษารถยนต์ โดยใช้เทคนิคการรักษาลูกค้าสัมพันธ์ สามารถแสดงผลข้อมูลที่ออกแบบและพัฒนาได้เหมาะตามความต้องการของลูกค้า ซึ่งลูกค้าและผู้ดูและระบบสามารถใช้งานโมดูลการลงทะเบียน, โมดูลการลงทะเบียนสมาชิก, โมดูลจัดการอะไหล่, โมดูลจัดการเคมีภัณฑ์, โมดูลสอบถามค่าใช้จ่ายด้านผลิตภัณฑ์, โมดูลรายงานการเข้าบำรุงรักษา และโมดูลรายงานค่าใช้จ่ายได้ โดยที่ลูกค้าไม่ต้องทำการสอบถามกับพนักงานของศูนย์บริการ ซึ่งถือได้ว่าเป็นการให้บริการที่สามารถตอบสนองต่อความต้องการของลูกค้าได้อย่างแท้จริง

โดยมีผลการประเมินการยอมรับการใช้งานระบบสารสนเทศสำหรับให้บริการการบำรุงรักษารถยนต์ โดยใช้เทคนิคการรักษาลูกค้า ใน 4 ด้านได้แก่ 1. ปัจจัยการรับรู้ผลประโยชน์ 2. ปัจจัยการรับรู้ความยากง่ายในการใช้งาน 3. ทัศนคติต่อการใช้ และ 4. พฤติกรรมความสามารถสรุปผลการประเมินความพึงพอใจระบบ โดยผู้ใช้งานทั่วไปและลูกค้าที่เข้ามาใช้บริการภายในศูนย์บริการทั้งหมดจำนวน 15 ท่าน เพื่อวัดการยอมรับการใช้งานที่มีต่อระบบ และทำการวัดความพึงพอใจที่มีต่อระบบโดยผู้เชี่ยวชาญจำนวน 5 โดยใช้แบบประเมินในการเก็บรวบรวมข้อมูล

4.1 สรุปผลการประเมินการยอมรับการใช้งานระบบ

จากการการยอมรับการใช้งานระบบสารสนเทศสำหรับให้บริการการบำรุงรักษารถยนต์ โดยใช้เทคนิคการรักษาลูกค้า สรุปได้ดังตารางที่ 1 และการวัดความพึงพอใจที่มีต่อระบบโดยผู้เชี่ยวชาญ สรุปได้ดังตารางที่ 2

ตารางที่ 1 : สรุปผลการประเมินการยอมรับการใช้งานระบบ

หัวข้อการประเมินคุณภาพ	$\bar{x}$	S.D	ระดับคุณภาพ
1. ปัจจัยการรับรู้ผลประโยชน์	4.43	0.24	ดี
2. ปัจจัยการรับรู้ความยากง่ายในการใช้งาน	4.41	0.30	ดี
3. ทัศนคติต่อการใช้	4.52	0.26	ดี
4. พฤติกรรมความตั้งใจ	4.55	0.22	ดี
ค่าเฉลี่ยโดยรวมทุกด้าน	4.48	0.04	ดี

ตารางที่ 2 : สรุปการประเมินความพึงพอใจระบบโดยผู้เชี่ยวชาญ

หัวข้อการประเมินความพึงพอใจ	$\bar{x}$	S.D	ระดับคุณภาพ
1. Functional Requirement Test	4.57	0.10	ดี

หัวข้อการประเมินความพึงพอใจ	$\bar{x}$	S.D	ระดับคุณภาพ
2. Functional Test	4.53	0.08	ดี
3. Performance Test	4.55	0.05	ดี
4. Usability Test	4.53	0.06	ดี
5. Security Test	4.36	0.06	ดี
ค่าเฉลี่ยโดยรวมทุกด้าน	4.51	0.02	ดี

เมื่อพิจารณาผลสรุปการประเมินการยอมรับการใช้งานระบบจากผู้ใช้งานทั่วไปและลูกค้าจำนวน 15 คน พบว่า ผลการประเมินการยอมรับการใช้งานระบบอยู่ในระดับดี ซึ่งค่าเฉลี่ยโดยรวมเท่ากับ 4.48 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.04 ดังนั้นความพึงพอใจในการยอมรับการใช้งานระบบอยู่ในระดับดี และการประเมินความพึงพอใจของระบบโดยผู้เชี่ยวชาญจำนวน 5 ท่าน พบว่า ผลการประเมินความพึงพอใจของระบบอยู่ในระดับดี ซึ่งมีค่าเฉลี่ยโดยรวมเท่ากับ 4.51 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.02 สรุปได้ว่าระบบสารสนเทศสำหรับให้บริการการบำรุงรักษารถยนต์ โดยใช้เทคนิคการรักษาลูกค้ามีความพึงพอใจอยู่ในระดับดี

5. สรุปผลและข้อเสนอแนะ

5.1 สรุปผล

ระบบสารสนเทศสำหรับให้บริการการบำรุงรักษารถยนต์โดยใช้เทคนิคการรักษาลูกค้า โดยใช้ MySQL และภาษา PHP เป็นเครื่องมือในการพัฒนาให้อยู่ในรูปแบบเว็บแอพพลิเคชั่น ระบบสามารถนำไปใช้งานได้ในทางปฏิบัติงานได้อย่างมีคุณภาพและมีประสิทธิภาพ มีความถูกต้องในการประมวลผลและสามารถตอบสนองได้ตรงตามความต้องการของผู้ใช้อยู่ในระดับดี

5.2 ข้อเสนอแนะ

ระบบควรพัฒนาให้สามารถแจกแจงข้อมูลให้มีรายละเอียดมากขึ้น เช่น รายงานรูปแบบกราฟ เพื่อนสนับสนุนการทำงานของผู้บริหารและตอบสนองต่อความต้องการของลูกค้าได้ดียิ่งขึ้น

ระบบควรพัฒนาให้สามารถใช้งานบนโทรศัพท์เคลื่อนที่ได้ในรูปแบบของแอพพลิเคชัน เพื่อขยายความสามารถในการให้บริการได้อย่างแท้จริง

เอกสารอ้างอิง

[1] Cong Li and Li Ma, Operation Design of Customer Relationship Management System, Sichuan Normal University, Chengdu, China, 2015.

[2] อำภา สิงห์คำพุด. การบริหารลูกค้าสัมพันธ์ แรงจูงใจที่มีอิทธิพลต่อพฤษติกรรมในการตัดสินใจเลือกซื้อประกันภัยรถยนต์ของผู้บริโภคในกรุงเทพมหานคร กับบริษัทเอกชนแห่งหนึ่ง. กรุงเทพมหานคร. มหาวิทยาลัยศรีนครินทรวิโรฒ , 2554.

[3] ฤทัยชนก ไผ่ผาด. การศึกษากลยุทธ์การรักษาลูกค้าของโรงแรมห้าดาวในจังหวัดเชียงใหม่. มหาวิทยาลัยนเรศวร, 2555.

[4] กมลวรรณ บุญทวี. ระบบจัดการลูกค้าสัมพันธ์ สำหรับเกรย์มาร์เก็ต กรณีศึกษา บริษัทเทดดี้ ออโต้ เซลส์ จำกัด. กรุงเทพมหานคร. สถาบันเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2555.

[5] ลินดา โสบิน. ระบบการจัดการองค์ความรู้ เรื่องระบบสารสนเทศภายในองค์กร โดยใช้เทคนิคการให้เหตุผลตามกรณีเป็นหลัก. กรุงเทพมหานคร. สถาบันเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2555.

[6] พัชรพล นรสาร. การผสมผสานการวิเคราะห์แบบหลายมิติ และวิถีความเป็นเลิศแบบโตโยต้า กับระบบรายงานผลการตรวจสอบคุณภาพงานซ่อมรถยนต์. กรุงเทพมหานคร. สถาบันเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2558.

[7] Omar Khadeer Hussain, Elizabeth Chang, Vish Ramakonar, Tharam S Dillon, A Customer Relationship Management Ecosystem that utilizes Multiple Sources and types of Information Conjointly, School of Information Systems, Curtin University, Perth Australia, Alsys MSC Sdn Bhd, Kuala Lumpur Malaysia ,Digital Ecosystems and Business Intelligence Institute Pty Ltd, Perth Australia, 2013.

[8] Prof. Ing and Ing.Ján Hurtuk, Customer Relationship Management System proposal with emphasis on implicity and security using distributed system architecture, Dept. of Electronics of Computers and Informatics, FEI TU of Košice, Slovak Republic, 2014.

การพัฒนาคลังข้อมูลการประเมินความเสี่ยงสำหรับการจัดการด้านภัยพิบัติ
A Development of Data warehouse in Risk Assessment for Disaster management

ธนาภรณ์ ปานรังศรี (Tanaporn Panrungsri)[1] และ เอสเธอร์ เสงี่ยมกุล (Esther Sangiamkul)[2]
คณะเทคโนโลยีและสิ่งแวดล้อม มหาวิทยาลัยสงขลานครินทร์ วิทยาเขตภูเก็ต
[1]tanapornpanrungsri@gmail.com, [2]esther.j@psu.ac.th

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาระบบคลังข้อมูล (Data Warehouse) โดยการประยุกต์ใช้ข้อมูลทั้งจาก Geographic Information Systems (GIS) และข้อมูลจากหน่วยงานที่มีส่วนเกี่ยวข้องในการจัดการภัยพิบัติ ข้อมูลที่นำมาใช้เป็นข้อมูลที่เกี่ยวกับระดับความรุนแรงทางด้านอุทกภัยและดินถล่ม จำนวนประชากรและครัวเรือนที่ได้ผลกระทบจากภัยแล้ง ข้อมูลทั้งหมดได้ถูกนำมารวบรวมด้วยกระบวนการ Extract-Transform-Load (ETL) และวิเคราะห์โดยใช้ Online Analytical Processing (OLAP) ซึ่งผลลัพธ์จากการวิเคราะห์ได้รับการนำเสนอบนแดชบอร์ด (Dashboard) โดยประสิทธิภาพของคลังข้อมูลจะถูกนำไปตรวจสอบ ด้วยการประเมินการใช้งานจริงและการซักถามความคิดเห็นจากผู้เชี่ยวชาญทางด้านภัยพิบัติ ผลลัพธ์ของงานวิจัยมีส่วนช่วยในการตัดสินใจและวางแผนในขั้นตอนการเตรียมการป้องกันภัยพิบัติให้กับเจ้าหน้าที่ของกรมป้องกันและบรรเทาสาธารณภัยจังหวัดภูเก็ต

คำสำคัญ: คลังข้อมูล ธุรกิจอัจฉริยะ ระบบสนับสนุนการตัดสินใจ ระบบสารสนเทศทางภูมิศาสตร์ การจัดการด้านภัยพิบัติ

Abstract

This research aims to develop a disaster data warehouse by integrating Geographic Information Systems (GIS) data and information from disaster mitigating organizations. Damage assessment from flood and landslide, the population, and the number of households affected by drought are considered. Extract-Transform-Load (ETL) process is employed to prepare the data for the Multi-Dimensional Modeling. The model is analyzed using Online Analytical Processing (OLAP) method. The result of the analysis is presented as a dashboard for users. The data warehouse is evaluated through trial uses and opinions from experts in the field of disaster. The result can facilitate planning and preparing for disasters for the Department of Disaster Prevention and Mitigation (DDPM) in Phuket.

Keyword: Data Warehousing, Business Intelligence, Decision Support System, GIS, Disaster Management.

1. บทนำ

ภัยพิบัติทางธรรมชาติทำให้เกิดความเสียหายทั้งชีวิตและทรัพย์สิน ส่งผลกระทบทางเศรษฐกิจอย่างมากมาย จึงต้องมีการจัดการด้านภัยพิบัติเพื่อลดความสูญเสียที่เกิดขึ้น [1] กระบวนการจัดการด้านภัยพิบัติแบ่งเป็น 4 ขั้นตอน ดังนี้ (1) การเตรียมการ (Preparedness) (2) การตอบสนอง (Response) (3) การฟื้นฟู (Recovery) และ (4) การบรรเทาทุกข์ (Mitigation) โดยมีหน่วยงานที่เกี่ยวข้องทั้งภาครัฐและเอกชน เช่น กรมอุตุนิยมวิทยา กรมทรัพยากรน้ำ กรมป้องกันและบรรเทาสา-ธารณภัย ซึ่งแต่ละหน่วยงานมีข้อมูลที่จัดเก็บตามกฎระเบียบภายในองค์กร อาทิ ข้อมูลการตัดสินใจเพื่อลดความเสียหาย ข้อมูลการวิเคราะห์ความเสี่ยง ข้อมูลการจัดอบรม ข้อมูลการแจ้งเตือน ข้อมูลการประเมินความเสียหาย [2][3] และข้อมูลจากสังคมออนไลน์ เช่น เฟซบุ๊ก (Facebook) ทวิตเตอร์ (Twitter) และ คราวด์ซอร์สซิ่ง (Crowdsourcing) [3] รวมทั้งข้อมูล GIS ที่สามารถแสดงข้อมูลแผนที่ [4]

จังหวัดภูเก็ตเป็นจังหวัดที่ประกอบด้วย 3 อำเภอ (อำเภอเมือง อำเภอถลาง และอำเภอกะทู้) แบ่งเป็น 17 ตำบล และ 103 หมู่บ้าน มีความโดดเด่นด้านการท่องเที่ยวทางทะเลอีกหนึ่งจังหวัดของภาคใต้ชายฝั่งอันดามันและเป็นที่นิยมของชาวต่างชาติในปี พ.ศ. 2548 ได้รับผลกระทบจากเหตุการณ์

ธรณีพิบัติทำให้นักท่องเที่ยวลดลงถึงร้อยละ 40-50 จากเหตุการณ์คลื่นยักษ์สึนามิ (Tsunami) มีผลต่อการท่องเที่ยวคิดเป็นร้อยละ 0.1 ของผลิตภัณฑ์รวมของประเทศ [5] นอกจากนี้จังหวัดภูเก็ตมีพื้นที่ทางทรัพยากรน้ำค่อนข้างมากสามารถรับน้ำได้โดยประมาณ 1,244 ตารางกิโลเมตร อีกทั้งมีแนวโน้มความต้องการน้ำเพิ่มมากขึ้นเฉลี่ยร้อยละ 0.2 ต่อปี ทำให้ปริมาณน้ำไม่เพียงพอต่อความต้องการทั้งหมด [6] จึงเกิดภาวะภัยแล้ง ผู้วิจัยได้สัมภาษณ์ผู้เชี่ยวชาญด้านภัยพิบัติพบว่าการเจริญเติบโตของเมืองเพิ่มขึ้นและมีสิ่งก่อสร้างกีดขวางทางน้ำไหลให้เกิดอุทกภัยและดินถล่มบ่อยครั้งในพื้นที่จังหวัดภูเก็ต

ด้วยเหตุนี้จึงมีความจำเป็นต้องศึกษาพื้นที่เสี่ยงอุทกภัย ดินถล่ม และภัยแล้ง เพื่อบรรเทาความรุนแรงของภัยพิบัติ รวมถึงลดการสูญเสียชีวิตและทรัพย์สิน อีกทั้งเพื่อเป็นแนวทางการเฝ้าระวังให้หน่วยงาน ซึ่งมีหน่วยงานที่ดูแลและเกี่ยวข้องด้านการจัดการภัยพิบัติ เช่น ศูนย์ป้องกันและบรรเทาสาธารณภัย เขต 18 ภูเก็ต (ปภ.เขต18 ภูเก็ต) ป้องกันและบรรเทาสาธารณภัยจังหวัดภูเก็ต (ปภ.ภูเก็ต) และการประปาส่วนภูมิภาคสาขาภูเก็ต ปัญหาของงานวิจัยนี้คือ การจัดการข้อมูลและบูรณาการเป็นคลังข้อมูลที่สามารถใช้ประเมินผลกระทบและความเสี่ยงอุทกภัย ดินถล่ม และภัยแล้งในอนาคต ผู้วิจัยได้ตั้งคำถามวิจัยว่า กระบวนการและเทคโนโลยีใดที่สามารถจัดการข้อมูลและเพิ่มประสิทธิภาพในการประเมินความเสี่ยงและผลกระทบ

การบูรณาการข้อมูล (Data Integration) จึงเป็นกระบวนการทางเทคโนโลยีและเครื่องมือที่ใช้แก้ปัญหาความแตกต่างของข้อมูลที่ทั้งโครงสร้างและเครื่องมือให้อยู่ในรูปแบบเดียวกันและจัดเก็บในคลังข้อมูล (Data warehouse) ที่มีการออกแบบโครงสร้างข้อมูลแบบหลายมิติ (Multi-dimensional modeling) [7] โดยมีปัจจัยเป็นตัวชี้วัดความความเสี่ยงได้แก่ ลักษณะภัย ลักษณะภูมิศาสตร์ จำนวนประชากร จำนวนครัวเรือน พื้นที่ด้านการเกษตร โครงสร้างพื้นฐาน ระยะเวลาเกิดภัย การอบรมเจ้าหน้าที่ อุปกรณ์เตือนภัย อุปกรณ์กู้ภัย และข้อมูลย้อนหลังการเกิดภัย และประเภทมิติข้อมูลได้แก่ เวลา พื้นที่ ประเภทภัย โดยมีกระบวนการสำคัญดังนี้ (1) การสกัด การเปลี่ยนแปลงและถ่ายโอนข้อมูล (ETL) จากแหล่งข้อมูล (2) การออกแบบ Multi-dimensional modeling (3) การประมวลผลแบบ OLAP (4) การนำเสนอข้อมูลในรูปแบบ เช่น แดชบอร์ด และรายงาน การพัฒนาคลังข้อมูลสามารถจัดการข้อมูลในโครงสร้างหลายมิติที่แตกต่างและมีปริมาณมากเพื่อใช้สนับสนุนการประเมินความเสี่ยงและผลกระทบที่มีประสิทธิภาพและรวดเร็วโดยการใช้ข้อมูลภายในองค์กร [7][8]

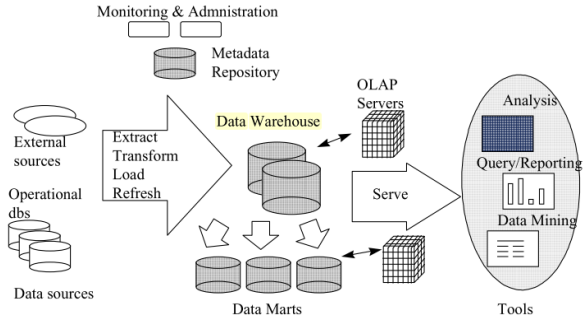
ความท้าทายของงานวิจัยคือการบูรณาการข้อมูลโดยประยุกต์ข้อมูล GIS กับข้อมูลด้านภัยพิบัติเพื่อพัฒนาคลังข้อมูลที่สามารถวิเคราะห์ระดับความเสี่ยงในจังหวัดภูเก็ต เพื่อทำให้เกิดความรวดเร็ว ความยืดหยุ่น และแสดงถึงความร่วมมือระหว่างองค์กรที่ดูแลด้านภัยพิบัติ ในงานวิจัยได้แบ่งการประเมินความเสี่ยงเป็นสองด้านดังนี้ (1) ระดับความรุนแรงของพื้นที่เสี่ยงอุทกภัยและดินถล่ม (2) จำนวนครัวเรือนและผู้ได้รับผลกระทบจากภัยแล้งพื้นที่จังหวัดภูเก็ต รวมทั้งได้ประเมินคลังข้อมูลและการใช้งานโดยผู้เชี่ยวชาญด้านภัยพิบัติ งานวิจัยนี้ประกอบด้วย ส่วนที่ 2 รวบรวมทฤษฎีและงานวิจัยที่เกี่ยวข้อง ส่วนที่ 3 กรอบการทำงานและการบูรณาการข้อมูล ส่วนที่ 4 ผลลัพธ์ของงานวิจัย ส่วนที่ 5 การประเมินคลังข้อมูลและการใช้งาน ส่วนที่ 6 สรุปผลและงานในอนาคต

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 คลังข้อมูล

คลังข้อมูล (Data Warehouse) [7] คือ ระบบฐานข้อมูลที่จัดเก็บข้อมูลรูปแบบหลายมิติ Multidimensional Modeling คลังข้อมูลมีวัตถุประสงค์เพื่อช่วยในการตัดสินใจและการวางแผนตามความต้องการผู้ใช้งานแบ่งเป็น 4 ด้าน คือ (1) ความต้องการด้านธุรกิจ (Business Requirements) (2) ความต้องการด้านข้อมูล (Data Requirements) (3) ความต้องการด้านฟังก์ชัน (Functional Requirements) (4) ความต้องการทางด้านเทคนิค (Technical Requirements) เพื่อช่วยในการวิเคราะห์ได้อย่างมีประสิทธิภาพและรวดเร็ว นอกจากนี้การจัดเก็บคลังข้อมูลมีลักษณะเด่นแบ่งเป็น 4 ลักษณะคือ (1) Subject-Oriented เป็นลักษณะโครงสร้างคลังข้อมูลที่จัดเก็บเป็นหมวดหมู่ตามเนื้อหา (2) Integrated เป็นคุณสมบัติที่สามารถรวบรวมข้อมูลจากแหล่งข้อมูลต่างๆ ทั้งภายในองค์กรและภายนอกองค์กร (3) Time-Variant เป็นคุณลักษณะสามารถวิเคราะห์ข้อมูลตามช่วงเวลา ซึ่งสามารถเก็บข้อมูลที่เกิดขึ้นในอดีต เพื่อช่วยวิเคราะห์แนวโน้มและคาดการณ์ข้อมูล

ที่จะเกิดขึ้นในอนาคต (4) Non-Volatile เป็นลักษณะข้อมูลที่นำมาจัดเก็บที่ไม่มีการเปลี่ยนแปลงรูปแบบข้อมูล โดยการนำข้อมูลเพื่อมาประมวลผลแบบ OLAP เพื่อใช้วิเคราะห์ข้อมูลได้หลายมิติ ทั้งการแสดงผลแบบโดยรวม (Roll up) แบบละเอียด (Drill down) แบบแยกสิ่งที่สนใจ (Slice) แบบพลิกแกน (Dice) และลูกบาศก์ (Cube) ทำให้สามารถแสดงข้อมูลได้หลายแกนช่วยตัดสินใจได้เร็วขึ้นในระยะเวลาที่เหมาะสม [7] [9] OLAP มีความแตกต่างจากฐานข้อมูลที่ในรูปแบบการประมวลผลธุรกรรมออนไลน์



ภาพที่ 1: สถาปัตยกรรมคลังข้อมูล

(Online transaction processing: OLTP) เนื่องจาก OLTP เป็นการทำงานโดยอัตโนมัติสำหรับงานประมวลผลข้อมูลธุรกรรม เช่น รายการสั่งซื้อและการทำธุรกรรมด้านการธนาคาร และธุรกรรมต้องมีรายละเอียดการอัพเดทข้อมูลแล้วซึ่งมีโครงสร้างและความซ้ำซ้อนโดยมีการกำหนดผ่านคีย์หลัก (Primary key) ซึ่งข้อมูลมีความสัมพันธ์ภายในตารางและยากต่อการวิเคราะห์ในเชิงลึก [9] นอกจากนี้มีตัวอย่างงานวิจัยที่ประยุกต์ใช้ธุรกิจอัจฉริยะในด้านภัยพิบัติ เช่น Asghar และคณะ [7] ได้ออกแบบ Business Intelligence Model เพื่อเป็นแนวทางการจัดการข้อมูลและออกแบบคลังข้อมูล โดยใช้ Oracle Warehouse Builder (OWB) ในการสร้างรายงานเพื่อสนับสนุนการตัดสินใจด้านภัยพิบัติในองค์กรของประเทศปากีสถานในงานวิจัยนี้มีการใช้แหล่งข้อมูลจากหน่วยงานเดียว

2.2 ระบบสารสนเทศภูมิศาสตร์

GIS เป็นเครื่องมือและกระบวนการทำงานของข้อมูลเชิงพื้นที่ด้วยระบบคอมพิวเตอร์โดยกำหนดข้อมูลสารสนเทศที่มีความสัมพันธ์เชิงพื้นที่ (Spatial Data) และข้อมูลคุณลักษณะประจำหรือข้อมูลอรรถาธิบาย (Attribute) เพื่อใช้บริหารทรัพยากรธรรมชาติและสิ่งแวดล้อมที่มีประสิทธิภาพ [4] [6]

Eckle และคณะ [10] ได้นำระบบสารสนเทศภูมิศาสตร์ (GIS) มาประยุกต์ใช้กับการจัดการด้านภัยพิบัติเพื่อแก้ปัญหาและประเมินความเสี่ยงด้านอุทกภัย ได้ใช้ OpenStreetMap (OSM) ซึ่งเป็นฐานข้อมูลด้านแผนที่แบบออนไลน์โดยออกแบบและนำเสนอต้นแบบ OpenFloodRiskMap (OFRM) เพื่อช่วยในการตัดสินใจเชิงพื้นที่

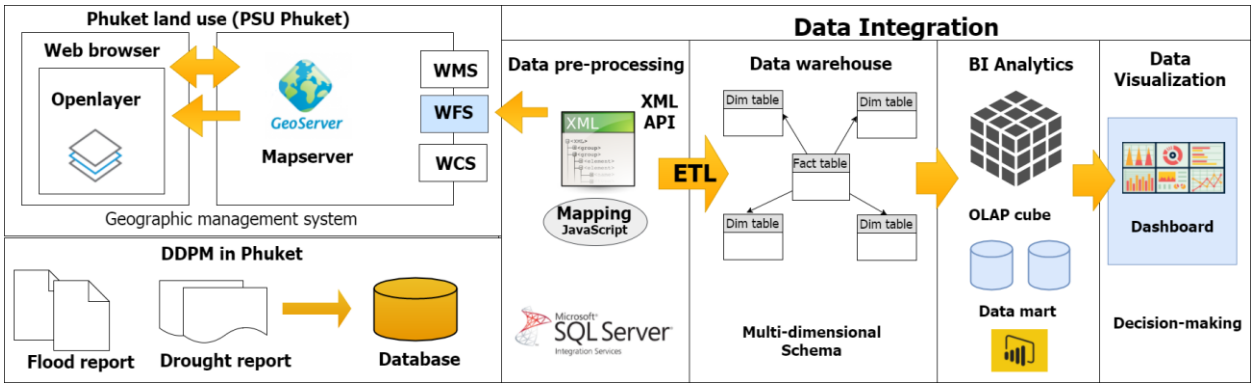
David Bitner และคณะ [4] ได้ประยุกต์ข้อมูล GIS ในระบบจัดการด้านภัยพิบัติแบบ Sahana (Sahana Disaster Management System) ซึ่ง GIS เป็นระบบฐานข้อมูลเฉพาะใช้วิเคราะห์และแสดงรูปแบบแผนที่และตารางเพื่อใช้ในการตัดสินใจโดยงานวิจัยชิ้นนี้ได้อธิบาย GIS prototype ในระบบจัดการด้านภัยพิบัติแบบ Sahana ซึ่งมีการใช้ GIS platforms เช่น GRASS ArcMap Google Earth และ GeoServer [4]

3. ขั้นตอนการวิจัย

ผู้วิจัยออกแบบกรอบการทำงานของธุรกิจอัจฉริยะ จากการศึกษาการจัดการด้านภัยพิบัติและรวบรวมเทคโนโลยีสารสนเทศเพื่อประยุกต์ข้อมูลด้านภัยพิบัติกับข้อมูล GIS โดยมีขอบเขตการวิเคราะห์ดังนี้ (1) การวิเคราะห์ระดับความรุนแรงของพื้นที่เสี่ยงด้านอุทกภัยและดินถล่ม (2) การวิเคราะห์จำนวนประชากรและครัวเรือนที่ได้ผลกระทบจากภัยแล้ง จากการสัมภาษณ์เจ้าหน้าที่ ปภ.จังหวัดภูเก็ต พบว่าอุทกภัย ดินถล่ม และภัยแล้ง เป็นภัยพิบัติทางธรรมชาติมักเกิดขึ้นในพื้นที่ ผู้วิจัยได้ประยุกต์ข้อมูลการใช้ประโยชน์ที่ดิน [11] และข้อมูลรายงานอุทกภัย ดินถล่ม และภัยแล้ง โดยแสดงกระบวนการทำงานทั้งหมด 3 ขั้นตอนได้แก่ (1) การรวบรวมข้อมูล (2) การเตรียมข้อมูล (3) การออกแบบคลังข้อมูล

3.1 การรวบรวมข้อมูล

ผู้วิจัยได้รวบรวมข้อมูลแบบทุติยภูมิโดยมีรายละเอียดดังนี้ (1) จำนวนประชากร จำนวนครัวเรือน ลักษณะภัย ระยะเวลาเกิดภัย พื้นที่ด้านการเกษตร โครงสร้างพื้นฐาน อุปกรณ์แจ้งเตือน การจัดอบรมเจ้าหน้าที่ การสำรองน้ำเพื่อการเกษตรและบริโภค ข้อมูลจุดจ่ายน้ำสำรองจากหน่วยงานปภ.จังหวัดภูเก็ต (2) ข้อมูลพื้นที่ระดับหมู่บ้าน พิกัดลองจิจูดและละติจูด ลักษณะทางภูมิศาสตร์ การใช้ประโยชน์ที่ดิน จากกลุ่มวิเคราะห์ข้อมูลและวิจัยด้านท่องเที่ยว พลังงานและสิ่งแวดล้อม มอ.ภูเก็ต



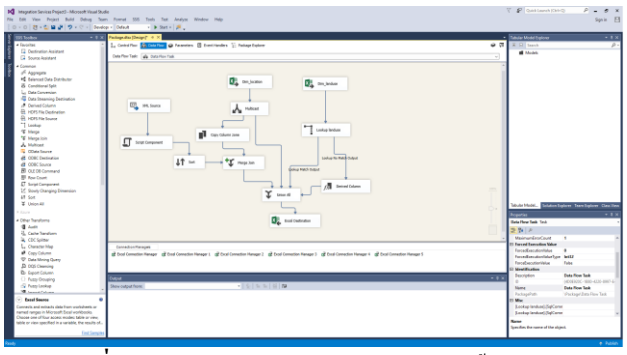
ภาพที่ 2: ขั้นตอนการวิจัย

3.2 การเตรียมข้อมูล

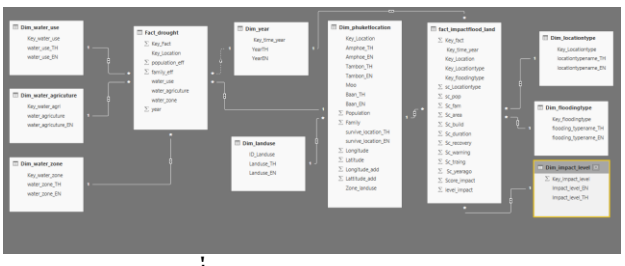
การรวบรวมข้อมูลผู้วิจัยได้แบ่งข้อมูลเป็นสองส่วน (1) ข้อมูลรายงานอุทกภัย ดินถล่มและภัยแล้ง โดยสกัดข้อมูลและจัดเก็บในฐานข้อมูล (2) ข้อมูลการใช้ประโยชน์ที่ดิน มีผู้วิจัยได้ใช้ช่องทาง API บนบริการ WFS ของ GEOServer ซึ่งสามารถดึงประเภทการใช้ประโยชน์ที่ดินในจังหวัดภูเก็ตใช้พิกัด Latitude กับ Longitude ในตารางพื้นที่จังหวัดภูเก็ต เชื่อมโยง (Mapping) กับพิกัดใน Shape file นำข้อมูลออกในรูปแบบไฟล์ XML ตามภาพที่ 2 นอกจากนี้ได้กำจัดข้อมูลที่ไม่เกี่ยวข้องและเปลี่ยนรูปแบบจัดเก็บในคลังข้อมูลโดยใช้กระบวนการ ETL ผ่านบริการ SQL Server Integration Services (SSIS) โปรแกรม Microsoft Visual Studio 2015 บนฐานข้อมูล Windows SQL server 2014ผู้วิจัยได้เชื่อมโยง (mapping) ข้อมูลที่ตรงกันกับ Attributeในตารางมิติพื้นที่จังหวัดภูเก็ตตามภาพที่ 3

3.3 การออกแบบคลังข้อมูล

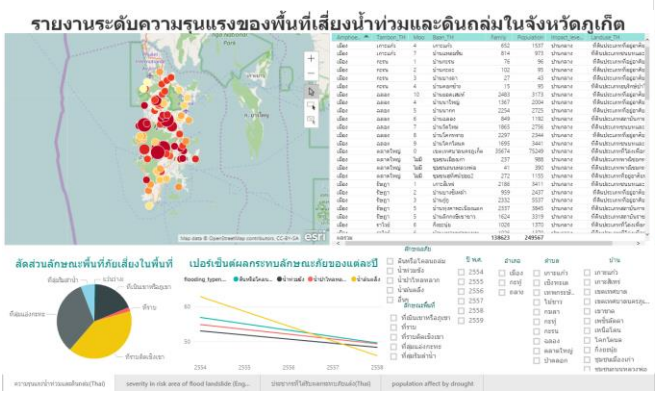
ผู้วิจัยพัฒนาคลังข้อมูลในรูปแบบ Multi-dimensional modeling เพื่อเก็บข้อมูลด้านอุทกภัย ดินถล่ม และภัยแล้งของจังหวัดภูเก็ตในการวิเคราะห์ระดับความรุนแรงของพื้นที่เสี่ยงของอุทกภัยและดินถล่ม โดยมีตัวชี้วัดคือคะแนนระดับความรุนแรง ส่วนการวิเคราะห์จำนวนประชากรและครัวเรือนที่ได้รับผลกระทบจากภัยแล้ง มีผลรวมของจำนวนประชากร และผลรวมของจำนวนครัวเรือนเป็นตัวชี้วัด โครงสร้าง Multi-dimensional modeling ประกอบด้วย ตารางข้อเท็จจริง 2 ตาราง ดังนี้ (1) ตารางวิเคราะห์รุนแรงพื้นที่เสี่ยงของอุทกภัยและดินถล่มโดยคำนวณตามปัจจัย การวิเคราะห์ซึ่งประยุกต์จากการกำหนดค่าน้ำหนักและคะแนนตามของรายงานระดับความรุนแรงของพื้นที่เสี่ยงอุทกภัยและดินถล่มของหน่วยงาน ปภ. ภูเก็ต โดยคะแนนเต็ม100 คะแนนแบ่งเป็น 3 ระดับ ดังนี้ คะแนนมากกว่า60 มีความรุนแรงของความเสี่ยงระดับสูง คะแนนระหว่าง 30-60 มีความรุนแรงของความเสี่ยงระดับปานกลาง น้อยกว่า 30 มีความรุนแรงของความเสี่ยงระดับต่ำ (2) ตารางวิเคราะห์ภัยแล้งโดยคำนวณค่าจากผลรวมของจำนวนประชากรและครัวเรือนในพื้นที่จังหวัดภูเก็ต และมีตารางมิติจำนวน 8 ตารางดังนี้ (1) ตารางมิติลักษณะพื้นที่ (2) ตารางมิติทางภูมิศาสตร์ (3) ตารางมิติเวลา (4) ตารางมิติประเภทภัย (5) ตารางมิติสำรองน้ำเพื่อการเกษตร (6) ตารางมิติสำรองน้ำเพื่อบริโภค (7) ตารางแหล่งจ่ายน้ำสำรองดังภาพที่ 4



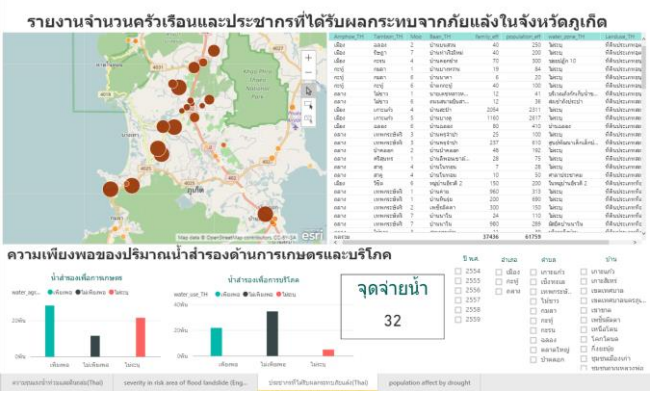
ภาพที่ 3: กระบวนการ ETL ในตารางมิติลักษณะพื้นที่



ภาพที่ 4: Multi-dimensional modeling



ภาพที่ 5: แดชบอร์ดรายงานระดับความรุนแรงในพื้นที่เสี่ยงอุทกภัยและดินถล่ม



ภาพที่ 6: แดชบอร์ดรายงานจำนวนประชากรและครัวเรือนที่ได้รับผลกระทบจากภัยแล้ง

4. ผลการดำเนินงาน

ผู้วิจัยได้เลือกใช้การวิเคราะห์โดยใช้การประมวลผลแบบ OLAP เพื่อวิเคราะห์ระดับความรุนแรงของพื้นที่เสี่ยงตามมิติของข้อมูลประเภทภัย แบ่งการวิเคราะห์เป็น 2 รายงาน คือ (1) ประเมินระดับความรุนแรงพื้นที่เสี่ยงด้านอุทกภัยและดินถล่มในจังหวัดภูเก็ตตามภาพที่ 5 (2) วิเคราะห์จำนวนครัวเรือนและประชากรที่ได้รับผลกระทบจากภัยแล้งของจังหวัดภูเก็ตตามภาพที่ 6 ตัวอย่าง เช่น การจำลองตารางมิติพื้นที่ภูเก็ต ตารางมิติลักษณะพื้นที่ ตารางมิติเวลา และตารางมิติประเภทภัย ในรูปแบบข้อมูลเจาะลึกและรูปแบบลูกบาศก์ภายในแต่ละมิติของข้อมูลโดยใช้ Microsoft Power BI เป็นเครื่องมือในการสร้างและนำเสนอผลการวิเคราะห์ข้อมูลระดับความรุนแรงของพื้นที่เสี่ยงและในรูปแบบเช่น แดชบอร์ด กราฟ และตาราง ตามหน่วยย่อยของมิติข้อมูลเพื่อสามารถวิเคราะห์ข้อมูลทั้งในระดับเชิงลึกและเชิงตื้น เช่น อำเภอถลาง=>ตำบลเทพกษัตรี=>หมู่ที่ 4=>บ้านดอน ในปี พ.ศ.2558 มีความรุนแรงในระดับปานกลาง ลักษณะเป็นน้ำท่วมขัง เนื่องจากมีพื้นที่เป็นแอ่งกระทะ เป็นที่ดินประเภทที่อยู่อาศัยหนาแน่นน้อย

5. การประเมินผลและการอภิปราย

ผลจากงานวิจัยนี้ ผู้วิจัยได้ให้ผู้เชี่ยวชาญจากกรมป้องกันและบรรเทาสาธารณภัย จำนวน 8 ท่าน โดยแบ่งการประเมินเป็น 2 ด้าน คือ (1) การประเมินด้านคลังข้อมูล (2) การประเมินด้านการใช้งาน โดยแบ่งคะแนนการประเมินเป็น 5 ระดับ คือ ระดับดีมาก ระดับดี ระดับปานกลาง ระดับน้อย และระดับควรปรับปรุง ผลจากการประเมินด้านคลังข้อมูลพบว่า จำนวนจุดจ่ายน้ำการสำรองน้ำเพื่อการเกษตร จำนวนประชากรและจำนวนครัวเรือน อยู่ในระดับดี มิติข้อมูลด้านภูมิศาสตร์และการใช้ประโยชน์ที่ดิน มิติข้อมูลหน่วยย่อยพื้นที่ มิติด้านภัยพิบัติ อยู่ในระดับดีมากมิติข้อมูลการสำรองน้ำเพื่อการบริโภคและมิติข้อมูลด้านเวลาในระดับปานกลาง โดยมิติข้อมูลด้านเวลามีข้อจำกัดในด้านการเก็บข้อมูลจากภาครัฐซึ่งข้อมูลที่หน่วยงานจัดเก็บในระดับปี โดยรวมการประเมินคลังข้อมูลอยู่ในเกณฑ์ระดับดีประเมินด้านการใช้งาน ขนาดหน้าจอเหมาะสมต่อการใช้งานในระดับดีมาก การแสดงทั้งสองภาษา ขนาดกราฟรูปแบบกราฟในระดับดี ระบบข้อมูลเป็นหมวดหมู่อยู่ในระดับดีและสามารถนำไปใช้ประโยชน์ได้จริงอยู่ในเกณฑ์ระดับดีมาก ด้านสนับสนุนการตัดสินใจอยู่ในระดับดี โดยรวมการใช้งานระบบอยู่ในเกณฑ์ระดับดี นอกจากนี้ผู้เชี่ยวชาญได้แนะนำการจัดการข้อมูล เนื่องจากหน่วยงานที่เกี่ยวข้องมีการจัดเก็บข้อมูลในรูปแบบรายงานทำให้ยากต่อการรวบรวมข้อมูลจึงได้แนะนำให้ออกแบบคลังข้อมูลที่วิเคราะห์เฉพาะเรื่องเพื่อง่ายต่อการออกแบบ ส่วนการวิเคราะห์หรือการตัดสินใจควรแบ่งระดับผู้ใช้งานเนื่องจากบางข้อมูลมีความสำคัญต่อกลุ่มใดกลุ่มหนึ่ง

งานวิจัยนี้พบว่าการบูรณาการข้อมูลและพัฒนาคลังข้อมูลเป็นเทคโนโลยีที่ช่วยจัดการข้อมูลเพื่อประเมินผลกระทบและความเสี่ยงด้านอุทกภัย ดินถล่มและภัยแล้ง โดยใช้ปัจจัยที่มีการกำหนดไว้ ทำให้สามารถประเมินผลกระทบและความเสี่ยงในมุมมองมิติของ พื้นที่ ลักษณะทางภูมิศาสตร์ เวลา ประเภทภัย สามารถสนับสนุนการตัดสินใจและเป็นแนวทางการวางแผนด้านจัดการภัยพิบัติในจังหวัดภูเก็ต

6. สรุปผลและแนวทางงานวิจัยในอนาคต

งานวิจัยนี้เป็นการบูรณาการข้อมูลด้านภัยพิบัติ โดยประยุกต์นำข้อมูล GIS กับข้อมูลด้านภัยพิบัติ และพัฒนาคลังข้อมูลที่มีการจัดเก็บหลายมิติ Multi-dimensional modeling โดยผ่านกระบวนการเตรียมข้อมูล ETL และประมวลผลแบบ OLAP เพื่อวิเคราะห์ตามมิติข้อมูล สามารถวิเคราะห์เชิงภาพรวมและแบบเจาะลึกนำเสนอในรูปแบบแดชบอร์ด โดยนำเสนอรายงานเพื่อประเมินระดับความรุนแรงของความเสี่ยงด้านอุทกภัยและดินถล่ม และวิเคราะห์จำนวนประชากรและครัวเรือนที่ได้รับผลกระทบจากภัยแล้ง เพื่อสนับสนุนการตัดสินใจ และเป็นประโยชน์ต่อเจ้าหน้าที่ด้านการจัดการภัยพิบัติจังหวัดภูเก็ต โดยในอนาคตสามารถเชื่อมโยงและประยุกต์ข้อมูลจากแหล่งข้อมูลอื่นได้ เช่นข้อมูลแบบจำลองทางน้ำไหล เพื่อสามารถวิเคราะห์พื้นที่เสี่ยงและประมาณการความเสียหายที่เกิดขึ้น เพื่อใช้ข้อมูลในการวางแผนเตรียมตัวและป้องกัน ได้อย่างแม่นยำและทันต่อสถานการณ์

7. กิตติกรรมประกาศ

ขอบคุณ ปภ.จังหวัดภูเก็ตสำหรับข้อมูลประเภทรายงานและข้อเสนอแนะด้านการจัดการด้านภัยพิบัติในจังหวัดภูเก็ต

ขอบคุณ ข้อมูลการใช้ประโยชน์ที่ดินจังหวัดภูเก็ต กลุ่มวิเคราะห์ข้อมูลและวิจัยด้านท่องเที่ยว พลังงาน และสิ่งแวดล้อม (Group for data analysis and research on tourism Energy and Environment; GDEN) คณะเทคโนโลยีและสิ่งแวดล้อม มหาวิทยาลัยสงขลานครินทร์วิทยาเขตภูเก็ต

เอกสารอ้างอิง

[1] นิตยา โชคทวีพาณิชย์ และ สุวิณี วิวัฒน์วานิช, "การศึกษาการจัดการพยาบาลในภาวะภัยพิบัติสำหรับผู้สูงอายุ" *วารสาร พยาบาลตำรวจ (Journal of the Police)*, 7.1, 183-196, 2015

[2] Pu, C., and M. Kitsuregawa, "Big data and disaster management: a report from the jst/nsf joint workshop." *Georgia Institute of Technology, CERCS* , 2013.

[3] Flachberger, C., & Gringinger, E., "Decision Support for Networked Crisis & Disaster Management–a Comparison with the Air Traffic Management Domain.", 2016

[4] Careem, M., Bitner, D., & Silva, R., "Gis integration in the sahana disaster management system." *Proceedings the International Conference on Information Systems for Crisis Response and Management, Delft, The Netherlands*, 2007.

[5] รอดภิญโญ ศศิพล, “การสื่อสาร การตลาดของการท่องเที่ยวแห่งประเทศไทย (ททท.) และการตัดสินใจท่องเที่ยว 6 จังหวัด ชายฝั่งทะเลอันดามัน หลังเหตุธรณีพิบัติ คลื่นยักษ์สึนามิของนักท่องเที่ยวชาวไทยโดย ศศิ พล และ รอด ภิญโญ” *Diss. The University of the Thai Chamber of Commerce.*

[6] พวงแก้ว นัฐพงษ์ สุวรรณประสิทธิ์ ชนิดา และ ศรีชัย นัยนา, "การศึกษาปริมาณความชื้นในดินโดยประยุกต์ใช้ระบบสารสนเทศภูมิศาสตร์ในเขต จังหวัดภูเก็ต." *วารสาร วิทยาศาสตร์ บูรพา* 21.2, 61-71, 2016.

[7] Asghar, S., Fong, S., & Hussain, T., "Business intelligence modeling: A case study of disaster management organization in pakistan." *Computer Sciences and Convergence Information Technology, 2009, ICCIT'09, Fourth International Conference on*. IEEE, 2009.

[8] Liyang, T., Zhiwei, N., Zhangjun, W., & Li, W., "A conceptual framework for business intelligence as a service (saas bi)." *Intelligent Computation Technology and Automation (ICICTA), 20 International Conference on* Vol. 2, IEEE, 2011.

[9] Chaudhuri, Surajit, and Umeshwar Dayal., "An overview of data warehousing and OLAP technology." *ACM Sigmod record*, 26.1, 65-74, 1997.

[10] Eckle, M., Albuquerque, J. P. D., Herfort, B., Wolff, R., Zipf, A., Leiner, R., & Jacobs, C., "Leveraging OpenStreetMap to support flood risk management in municipalities: A prototype decision support system.", 2016

[11] Ratchaniphont A., Keson J., Wongsai S., Kokkaew E., Wongsai N., Senrit D. and Onthong S., “Land-use change in the Andaman Triangle provinces, Thailand.”, *Journal of Remote Sensing and GIS Association of Thailand.*, 17(1), 2016

ระบบจัดตารางสอนแบบอัตโนมัติโดยการประยุกต์ใช้ออนโทโลยีและกฎ
Automatic Course Scheduling System Applying Ontology and Rule-based

บุรัสกร อยู่สุข (Burasakorn Yoosooka)[1] ปองพล นิลพฤกษ์ (Pongpon Nilaphruek)[2] และ คงเทพ บุญมี (Khongthep Boonme)[3]
ภาควิชาคณิตศาสตร์และวิทยาการคอมพิวเตอร์ สาขาวิชาวิทยาการคอมพิวเตอร์
คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี
[1]y.burasakorn@rmutt.ac.th, [2] pongpon_n@rmutt.ac.th,[3] khongthep@rmutt.ac.th

บทคัดย่อ

งานวิจัยนี้มีจุดประสงค์เพื่อศึกษาปัญหาการจัดตารางสอนในรูปแบบปัญหาเอ็นพีบริบูรณ์ซึ่งเป็นปัญหาที่ไม่สามารถหาคำตอบดีที่สุดได้ และพัฒนาระบบจัดตารางสอนแบบอัตโนมัติให้กับสาขาวิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี โดยได้มีการประยุกต์ใช้ออนโทโลยีเพื่อเป็นตัวแทนสารสนเทศที่เกี่ยวข้องกับระบบ และนำมาใช้ในการออกแบบขั้นตอนวิธีเพื่อแก้ปัญหาของงานวิจัย ผู้วิจัยได้ทำการสร้างออนโทโลยีโดยอ้างอิงตามรูปแบบความต้องการของกรณีศึกษา โดยประยุกต์ใช้กฎ และนำมาใช้ในการพัฒนาระบบจัดตารางสอนแบบอัตโนมัติ ผลการศึกษาแสดงความสามารถของการนำเอาออนโทโลยีมาประยุกต์ใช้กับกฎสามารถเพิ่มกระบวนการในการหาข้อสรุปได้ดีขึ้น โดยเปรียบเทียบการทำงานระหว่างการใช้ระบบกับการจัดตารางสอนด้วยมือ ซึ่งผู้ใช้งานหลักมีระดับความพึงพอใจในการใช้งานระบบอยู่ในระดับมาก คิดค่าเฉลี่ยโดยรวมได้เท่ากับ 4.04 และมีส่วนเบี่ยงเบนมาตรฐานเฉลี่ยเท่ากับ 0.62 แต่อย่างไรก็ตามคำตอบที่ได้จากระบบยังไม่เป็นคำตอบที่ดีที่สุดจากความต้องการทั้งหมด ในอนาคตสามารถประยุกต์ใช้เครื่องมือและขั้นตอนวิธีอื่น ๆ เพื่อแก้ปัญหาต่อไป

คำสำคัญ: การจัดตารางสอนแบบอัตโนมัติ ออนโทโลยี กฎ

Abstract

This research aims to study the course scheduling in NP-Complete problem, which cannot find the best solution as well as to develop the course scheduling system for Computer Science Division, Faculty of Science and Technology, Rajamangala University of Technology Thanyaburi applying the ontology to represent the related information in the system. Also, proposed ontology was used in the proposed algorithm to solve the above-mentioned problem. The proposed ontology was constructed by referencing from user requirements in the case study and applied to the rules for developing the automatic course scheduling. The result of this research explains that the abilities of ontology combining with rule-based can improve the procedure to find the consensus by comparing between using the proposed system and manual management. The satisfaction in the proposed system from main users is in good level; the average score is 4.04, and SD is 0.62. However, the consensus from the system is not the best solution from the requirements. In future works, other tools and algorithms can be applied to solve the problem in addition.

Keyword: Automatic Course Scheduling, Ontology, Rule-based.

1. บทนำ

การจัดตารางสอนนั้นนับเป็นงานที่ท้าทาย เนื่องจากมีปัจจัยสภาพแวดล้อมที่เกี่ยวข้องกันหลากหลายมิติ อาทิเช่น ความชอบส่วนตัวของอาจารย์ ชนิดห้องเรียน ความจุของห้องเรียน จำนวนนักศึกษา ชนิดของรายวิชา และอุปกรณ์เฉพาะของรายวิชา เป็นต้น นอกจากนี้การจัดตารางสอนด้วยกระบวนการแบบทำมือ ซึ่งจัดการด้วยมนุษย์นั้น ใช้เวลาและทรัพยากรในการจัดการเป็นอย่างมาก หากต้องการจัดให้ตรงกับ

ความต้องการของผู้ใช้งานให้ได้มากที่สุด ปัญหานี้จึงจัดได้ว่ามีความสำคัญทั้งในเชิงวิชาการและในเชิงการนำไปประยุกต์ใช้

โดยงานวิจัยนี้ได้ใช้กรณีศึกษาจากสาขาวิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี ซึ่งผู้วิจัยได้เล็งเห็นการนำปัญหาการจัดตารางสอนมาสร้างเป็นระบบแบบอัตโนมัติ และถ้ามีปัจจัยบังคับที่หลากหลายและไม่จำกัดจำนวน อาจส่งผลให้ปัญหานี้จัดอยู่ในรูปแบบปัญหาเอ็นพีบริบูรณ์ (NP-Complete) [1] ซึ่งเป็นปัญหาที่ไม่สามารถหาคำตอบดีที่สุดได้ในช่วงเวลาจำกัด อย่างไรก็ตามปัญหานี้สามารถลดรูปให้อยู่ในรูปแบบที่สามารถคำนวณได้ในทางปฏิบัติ แต่คำตอบที่ได้จะขึ้นอยู่กับการลดรูปหรือจำนวนข้อบังคับตามปัจจัยที่เกี่ยวข้อง

ดังนั้นผู้วิจัยจึงเสนอการพัฒนาระบบจัดตารางสอนแบบอัตโนมัติ โดยการประยุกต์ใช้ออนโทโลยีและกฎ ซึ่งได้รับการยอมรับในกระบวนการพัฒนาเว็บเชิงความหมาย [2] เพื่อเพิ่มประสิทธิภาพในกระบวนการหาข้อสรุป และยังได้เสนอปัจจัยที่เกี่ยวข้องกับการจัดตารางสอนแบบอัตโนมัติให้อยู่ในปัจจัยพื้นฐานที่ยอมรับได้จากผู้ที่เกี่ยวข้อง รวมถึงการเปรียบเทียบระดับความพึงพอใจในการใช้งานระบบกับกระบวนการจัดตารางสอนด้วยมือ

2. วรรณกรรมที่เกี่ยวข้อง

ออนโทโลยี (Ontology) [3] [4] ในเชิงเทคโนโลยีสารสนเทศ คือการนิยามหรือการกำหนดรูปแบบโครงสร้างของสิ่งที่เราสนใจ (domain) ให้มีความหมายตามขอบเขตขององค์ความรู้ หรือกล่าวอีกนัยหนึ่งคือตัวแทนการแสดงองค์ความรู้ ซึ่งออนโทโลยีได้ถูกกล่าวถึงในหลายแขนง โดยเฉพาะระบบปัญญาประดิษฐ์ ซึ่งปัจจุบันออนโทโลยีได้ถูกนำใช้ในเว็บเชิงความหมายมากขึ้น เนื่องจากสามารถแบ่งปันข้อมูลกันได้ง่ายผ่านโปรโตคอล HTTP [5] โดยมาตรฐานของการจำลองข้อมูลของออนโทโลยี นิยมใช้กรอบมาตรฐานที่เรียกว่า Resource Description Framework (RDF) และ Ontology Web Language (OWL) หรือเขียนโดยย่อคือ RDF/OWL [6]

กฎ (Rule-based) [7] คือเงื่อนไขการทำงานเชิงตรรกศาสตร์ ซึ่งสามารถสรุปข้อมูลจากข้อเท็จจริง (fact) เพื่อสร้างข้อมูลใหม่ได้ ในที่นี้ข้อเท็จจริงอาจอยู่ในรูปแบบขององค์ความรู้หรือสารสนเทศก็ได้ โดยกฎนั้นถูกผนวกและนำมาประยุกต์ใช้กับออนโทโลยีตามแนวคิดของเว็บเชิงความหมาย โดยใช้หลักการทางตรรกศาสตร์ของเงื่อนไขต่าง ๆ เช่น สมมุติกำหนดกฎเป็นภาษาคำพูดต่อไปนี้

- (ถ้า x เป็นพ่อ y) และ(y เป็นพ่อ z) จะได้ (x เป็นปู่ z)

ถ้าในกรณีนี้ มีข้อมูลที่เป็นข้อเท็จจริงคือ

- fact 1 => นาย ก เป็นพ่อ นาย ข
- fact 2 => นาย ข เป็นพ่อ นาย ค

จากกรณีดังกล่าว ถ้าประยุกต์ใช้กฎด้านบน จะทำให้เกิดข้อเท็จจริงขึ้นมาใหม่อีก 1 ข้อมูลคือ {face 3 => นาย ก เป็นปู่ นาย ค} ซึ่งจะสังเกตได้ว่า ระบบทำการเรียนรู้จากกฎที่สร้างขึ้นเพื่อหาข้อเท็จจริงใหม่ที่เกิดขึ้น

จากประโยชน์ของออนโทโลยีและกฎที่ได้กล่าวไว้ ในทางปฏิบัติ ผู้วิจัยได้เปรียบเทียบออนโทโลยีเสมือนข้อเท็จจริงของสารสนเทศหรือองค์ความรู้ เช่น ข้อมูลอาจารย์ผู้สอน ข้อมูลรายวิชา และข้อมูลห้องเรียน เป็นต้น รวมถึงกฎยังเปรียบเสมือนกระบวนการหาข้อสรุปในการจัดตารางสอน ซึ่งผู้วิจัยได้ใช้รูปแบบกฎในมาตรฐานภาษากฎของเว็บเชิงความหมาย (Semantic Web Rule Language: SWRL) [8]

นอกจากนี้ผู้วิจัยยังได้ทบทวนงานวิจัยที่เกี่ยวข้องหลายงาน โดยสรุปมีดังต่อไปนี้

[9] งานวิจัยที่เกี่ยวข้องนี้ได้ทำการรวบรวมแนวทางการหาคำตอบของการแก้ปัญหาการจัดตารางสอนอัตโนมัติด้วยขั้นตอนวิธีเชิงพันธุกรรม (Genetic algorithm) ขั้นตอนวิธีการลอกแบบ (Memetic algorithm) การค้นหาแบบทาบู (Tabu search) การจำลองการอบเหนียว (Simulated annealing) กระบวนการแบ่งกลุ่ม (Clustering methods) และกระบวนการเชิงเส้น (Linear method) ซึ่งผลลัพธ์สามารถแก้ปัญหาได้ระดับหนึ่งถ้าไม่มีเงื่อนไขมากจนเกินไป

[10] งานวิจัยที่เกี่ยวข้องนี้ได้นำเสนอรูปแบบลำดับของการเลือกตารางเวลาสอนของผู้สอนและตารางเวลาของผู้เรียนให้สอดคล้องกัน โดยขึ้นอยู่กับเงื่อนไขหลากหลายรูปแบบ ซึ่งการจัดตารางสอนเน้นไปที่ความชอบของผู้สอนเป็นหลักด้วยหลักการของการทำเหมืองแร่ข้อมูล (Data Mining) โดยผลลัพธ์ที่ได้เป็นที่น่าพอใจถ้าเปรียบเทียบกับการจัดตารางสอนแบบเก่าซึ่งใช้วิธีการทำด้วยบุคคลากรโดยตรงแบบดั้งเดิม งานวิจัยนี้ได้

ใช้กรณีศึกษาของสาขาวิชาวิทยาการคอมพิวเตอร์และสาขาอื่น ๆ ที่เกี่ยวข้องใน Jordan University of Science and Technology ประเทศจอร์แดน แต่อย่างไรก็ตาม งานวิจัยนี้ได้นำเสนอโดยใช้แค่ความชอบของผู้สอนเท่านั้น โดยไม่ได้อ้างอิงถึงความชอบของนักศึกษาและประเด็นเรื่องห้องเรียน ชนิดของรายวิชา และข้อมูลอื่นที่เกี่ยวข้อง

[11] งานวิจัยที่เกี่ยวข้องนี้ได้มุ่งเน้นการสร้างตารางสำหรับระบบการศึกษา โดยทำการศึกษาไปที่ลักษณะของปัญหาทางวิศวกรรมซอฟแวร์ โดยมีการกำหนดปัญหาเงื่อนไขความพึงพอใจซึ่งถูกแสดงในศาสตร์ของวิศวกรรมซอฟแวร์ เพื่อให้ได้มาซึ่งตารางกำหนดการต่าง ๆ ที่สามารถถูกสร้างขึ้นมาได้จากพื้นฐานการกระตุ้นคิดได้ด้วยตนเองของสารสนเทศในระบบ (heuristic information) งานวิจัยนี้ได้พัฒนาระบบเป็นลักษณะของโปรแกรมแบบติดตั้งใช้ในเครื่องอย่างเดียว (standalone) โดยผู้ใช้สามารถกำหนดเงื่อนไขต่าง ๆ เข้าไปในระบบ หลังจากนั้นระบบจะทำการสร้างตารางกำหนดการสอนให้แบบอัตโนมัติ โดยสรุปแล้วงานวิจัยชิ้นนี้สามารถช่วยลดภาระการทำงานที่เกี่ยวข้องกับระบบจัดตารางสอนที่ต้องทำด้วยบุคคลแบบเก่าได้มาก แต่อย่างไรก็ตามการแสดงผลของโปรแกรมยังไม่เป็นที่น่าพอใจนัก ประกอบกับโปรแกรมมีลักษณะเป็นการติดตั้งใช้ในเครื่องได้อย่างเดียว จึงทำให้มีข้อจำกัดในการใช้งานร่วมกัน

3. วิธีการดำเนินการโครงการวิจัย

การดำเนินงานและจัดการวิจัยนี้ เป็นการออกแบบออนโทโลยีสำหรับสารสนเทศที่เกี่ยวข้องพร้อมกับกฎ รวมถึงการออกแบบขั้นตอนวิธีสำหรับเลือกรายวิชา ห้องเรียน และอาจารย์ผู้สอน ให้มีความสอดคล้องกัน เพื่อนำมาประยุกต์ใช้ในการสร้างระบบจัดตารางสอนแบบอัตโนมัติของกรณีศึกษา ซึ่งสามารถสรุปขั้นตอนการดำเนินการวิจัยได้ดังนี้

3.1 การรวบรวมข้อมูลการจัดตารางสอนภายในสาขาวิชาวิทยาการคอมพิวเตอร์

ผู้วิจัยได้รวบรวมข้อมูลการจัดตารางสอนภายในสาขาวิชาโดยให้อยู่บนพื้นฐานความต้องการที่ยอมรับได้ของอาจารย์ที่เกี่ยวข้อง ซึ่งตารางสอนของสาขาวิชาจะแบ่งออกเป็นกรณีต่างๆ ดังต่อไปนี้

- ภาคปกติ มีวิชาทฤษฎี ซึ่งสอนทั้งหมด 3 ชั่วโมง ส่วนมากไม่ใช้ห้องปฏิบัติการ ส่วนวิชาปฏิบัติสอนทั้งหมด 4 ชั่วโมง ต้องใช้ห้องปฏิบัติการเท่านั้นในการเรียนการสอน และควรจะจัดให้อยู่ในช่วงเวลาราชการ (ตั้งแต่วันจันทร์ถึงศุกร์ เวลา 8.00 น. ถึง 17.00 น.) ถ้าเป็นไปได้
- ภาคพิเศษ จะมีลักษณะรายวิชาเหมือนภาคปกติทุกประการ ยกเว้นเวลาในการสอนควรจะจัดให้อยู่ในช่วงเวลานอกราชการ (ตั้งแต่วันจันทร์ถึงศุกร์ เวลา 17.00 น. ถึง 21.00 น. หรือ วันเสาร์ถึงวันอาทิตย์ เวลา 8.00 น. ถึง 17.00 น.)
- กรณีวิชาศึกษาทั่วไปลำดับความสำคัญมีมากกว่าวิชาในสาขา ดังนั้นวิชาไหนเป็นวิชาศึกษาทั่วไปจะถูกจัดตารางก่อน เพื่อความสะดวกในการทำงาน
- การจัดตารางสอนให้เทียบความสำคัญของรายวิชาและอาจารย์ผู้สอนตามลำดับ
- ห้องปฏิบัติการ โดยปกติแบ่งห้องออกเป็นสองขนาดคือ สามารถรับผู้เรียนได้ 30 คนและ 60 คน
- ห้องทฤษฎี แบ่งห้องออกเป็นสองขนาดเช่นเดียวกัน คือ สามารถรับผู้เรียนได้ 30 คนและ 60 คน

3.2 การออกแบบออนโทโลยี

การออกแบบออนโทโลยีสำหรับใช้เป็นข้อมูลข้อเท็จจริง (fact data) ให้กับระบบนั้น ผู้วิจัยได้เลือกมาตรฐาน RDF/OWL เป็นตัวแทนของข้อมูลชุดนี้ โดยผู้วิจัยได้แบ่ง Concept ของออนโทโลยีดังต่อไปนี้

- Prof ใช้สำหรับนิยามอาจารย์ผู้สอน
- Room ใช้สำหรับนิยามข้อมูลห้องเรียน ณ เวลาหนึ่ง
- Subject ใช้สำหรับนิยามรายวิชา
- Section ใช้สำหรับนิยามการเรียนการสอนที่จะเกิดขึ้น ในที่นี้ 1 section เท่ากับการเรียนการสอนที่เกิดขึ้น 1 ห้องเรียน

โดยรายละเอียดของ Concept แต่ละแบบ มีดังนี้

3.2.1 Prof ประกอบไปด้วย Data Properties คือ

- name ใช้สำหรับเก็บชื่อผู้สอน เป็น string
- priority ใช้สำหรับเก็บระดับความสำคัญของผู้สอน โดยเป็น set ของ decimal ระหว่าง 0.0-1.0 ค่ายิ่งมากยิ่งสำคัญ
- preferenceDaySet ใช้สำหรับเก็บ set ของวันที่ผู้สอนชอบ โดยเป็น set ของลำดับวันเรียงจากวันที่ชอบมากที่สุดไปถึงวันที่ชอบน้อยสุด เช่น {“Tue”, ”Wed”, ”Thu”, ”Mon”, ”Fri”, ”Sat”, ”Sun”} จากตัวอย่างนี้ผู้สอนมีลำดับความชอบ คือ วันอังคาร พุธ พฤหัส จันทร์ ศุกร์ เสาร์ และอาทิตย์ ตามลำดับ
- preferenceTimeSet ใช้สำหรับเก็บ set ของเวลาที่ผู้สอนชอบ โดยเป็น set ของเวลา 3 ค่า ดังต่อไปนี้ {“8.00-12.00”, “13.00-17.00”, “17.00-21.00”} ผู้สอนเลือกเวลาที่ชอบมากที่สุดเป็นอันดับ 1 และเรียงลำดับความชอบต่อไปตามลำดับ

3.2.2 Room ประกอบไปด้วย Data Properties คือ

- seat ใช้สำหรับเก็บจำนวนที่นั่ง เป็น integer ตั้งแต่ 0-*
- roomNumber ใช้เก็บเลขของห้องเรียนเป็น string
- isLab ใช้สำหรับเก็บข้อมูลว่าห้องนี้มี Lab หรือไม่ เป็น Boolean โดย เก็บค่าเป็นtrue หรือ false
- avaliableTime ใช้สำหรับเก็บช่วงเวลาใด ๆ ไว้ โดยเลือกจาก set ของเวลา 3 ค่า ดังต่อไปนี้ {“8.00-12.00”, “13.00-17.00”, “17.00-21.00”, “false”} หมายเหตุ ห้องที่มีช่วงเวลาต่างกันจะแยกเป็นอิสระต่อกัน และนำ day + timeslot เช่น “MON 8.00-12.00” เป็นต้น

3.2.3 Subject ประกอบไปด้วย Data Properties คือ

- subjectID ใช้สำหรับเก็บรหัสวิชา เป็น string
- subjectName ใช้สำหรับเก็บชื่อรายวิชา เป็น string
- subjectPriority ใช้สำหรับเก็บระดับความสำคัญของรายวิชาเป็น decimal set ระหว่าง 0.0-1.0 ค่ายิ่งมากยิ่งสำคัญ
- labRequired ใช้สำหรับเก็บข้อมูลว่าวิชานี้ต้องการ Lab หรือไม่ เป็น Boolean โดย เก็บค่าเป็น true หรือ false
- duration ใช้สำหรับเก็บจำนวนชั่วโมงการสอนต่อหนึ่งครั้ง โดยเป็น set ของ integer ให้เลือกเอาระหว่าง {1, 2, 3, 4} โดยหน่วยเป็น จำนวนชั่วโมง

นอกจากนี้ยังมี Object Properties คือ

- hasProf อ้างอิงถึง Class Prof

3.2.4 Section ประกอบไปด้วย Data Properties คือ

- sectionID ใช้สำหรับเก็บรหัส section เป็น string
- studentAmount ใช้สำหรับเก็บจำนวนของผู้เรียน เป็น integer ตั้งแต่ 0-*

และมี Object Properties คือ

- hasSubject อ้างอิงถึง Class Subject เพื่อบอกให้ทราบว่า section นี้จัดการเรียนการสอนวิชาอะไร
- hasRoom อ้างอิงถึง Class Room เพื่อบอกให้ทราบว่า section นี้เรียนห้องไหน

3.3 การออกแบบกฎ

กฎส่วนของตรรกศาสตร์ที่ถูกหยิบยกขึ้นมา เพื่อทำให้ข้อมูลที่ได้ออกแบบไว้แล้วในหัวข้อที่ 3.2 เกิดข้อเท็จจริงใหม่ ๆ

ในกรณีนี้ ผู้วิจัยได้นิยามกฎตามหลักการของ SWRL ที่ได้กล่าวไว้แล้วในหัวข้อที่ 2 เพื่อกำหนดให้ระบบเรียนรู้การกำหนดตารางสอนให้แก่ Section ต่าง ๆ ว่ากำหนดไปแล้วหรือยัง โดยถ้ากำหนดไปแล้วให้แสดงผลลัพธ์ของ Section ดังกล่าวให้เทียบเท่ากับ BookedSection หรือถ้า Section ใด ๆ ยังไม่ถูกกำหนดลงในตารางสอน ระบบต้องสามารถตรวจได้ว่า Section ดังกล่าวนั้นจะเทียบเท่ากับ UncompletedSection ในทำนองเดียวกัน ระบบต้องสามารถตรวจสอบสถานะของห้องเรียนว่าว่างหรือไม่ โดยผู้วิจัยได้กำหนดกฎขึ้นมาดังต่อไปนี้

Rule 1:

(?sec type_of Section) ^ (?s type_of Subject) ^
(?p type_of Prof) ^ (?r type_of Room) ^
(?sec hasSubject ?s) ^ (?sec hasRoom ?r) ^
(?s hasProf ?p) -> (?sec type_of BookedSection)

Rule 2:

(?sec type_of Section) ^ not(?sec type_of BookedSection) -> (?sec type_of UncompletedSection)

Rule 3:
(?sec type_of Section) ^ (?r type_of Room) ^
(?sec type_of BookedSection) ^ (?sec hasRoom ?r) ^
(?r avaliableTime false)->(?r type_of BookedRoom)

Rule 4:
(?r type_of Room) ^ not(?r type_of
?BookedRoom)^ -> (?r type_of AvaliableRoom)

จากกฎดังกล่าว สามารถหาได้ว่า Section ไหนบ้างที่ถูกนำไปลงในตารางสอนแล้ว และ Room ไหนบ้างที่ว่าง ซึ่งจะมีประโยชน์สำหรับประยุกต์ใช้กับขั้นตอนวิธีในหัวข้อถัดไป

3.4 การออกแบบขั้นตอนวิธีสำหรับระบบ

จากขั้นตอนที่ผ่านมา ผู้วิจัยได้ออกแบบออนโทโลยีและกฎสำหรับเป็นข้อมูลในการจัดตารางสอน ขั้นตอนในหัวข้อนี้จะเป็นการออกแบบขั้นตอนวิธีสำหรับการจัดลำดับ Section ของการเรียนการสอน ว่า Section ใด ๆ ควรจะถูกนำไปใส่ไว้ในตารางสอน ณ ตำแหน่งใด เป็นต้น

สมมุติมี set ตัวอย่างข้อมูลเป็นดังต่อไปนี้

- :t1 :type_of :Prof
- :t1 :name “ปองพล นิลพฤกษ์”
- :t1 :priority 0.3
- : t1 : preferenceDay {“Tue”, ”Wed”, ”Thu”, ”Mon”, ”Fri”, ”Sat”, ”Sun”}
- :t1 :preferenceTime {“8.00-12.00”, “13.00-17.00”, “17.00-21.00”}
- :t2 :type_of :Prof
- :t2 :name “บุรัสกร อยู่สุข”
- :t2 :priority 0.7
- : t2 : preferenceDay {”Mon”, “Tue”, ”Wed”, ”Thu”, ”Fri”, ”Sat”, ”Sun”}
- :t2 :preferenceTime {“8.00-12.00”, “13.00-17.00”, “17.00-21.00”}

ขั้นที่ 1 นำข้อมูลเท็จจริงที่มีไปเข้ากฎที่ออกแบบไว้

เริ่มต้นจากการนำข้อมูลเท็จจริงที่มีไปเข้ากฎที่ออกแบบไว้ และหาว่ามี ข้อมูลที่เป็น type_of UncompletedSection หรือไม่ ถ้ามี :=> ให้ไปทำขั้นตอนที่ 2 ถ้าไม่มี :=> ให้จบการทำงานของโปรแกรม จากขั้นตอนนี้ ได้ว่า sec1 และ sec2 เป็น type_of UncompletedSection ทั้งสอง ดังนั้นจึงไปทำขั้นที่ 2 ต่อ

ขั้นที่ 2 เรียงลำดับความสำคัญ (priority) ของ Section จากน้อยไปมากโดยใช้ priority ของ Prof และ Subject เพื่อเลือกตัวแทนข้อมูล ตามลำดับ ซึ่งมีการกำหนดค่าน้ำหนัก (weight) ของ priority ดังตารางที่ 1

ตารางที่ 1: การกำหนด factor และ weight เพื่อเลือกตัวแทนข้อมูลขั้นที่ 2

$factor_i$	**data from**	**Weight**
$factor_1$	Prof.priority	$w_1 = 0.7$
$factor_2$	Subject.subjectPriority	$w_2 = 0.5$

หลังจากนั้นให้ทำการหาอาจารย์และรายวิชาที่เกี่ยวข้องกัน ถ้าไม่มี :=> ให้ออกจากโปรแกรม ถ้ามี :=> ทำต่อ

ถ้ามีข้อมูลให้เราทำการหาความสัมพันธ์ของข้อมูลต่อ จากตัวอย่างข้อมูล เรามีอาจารย์อยู่ 2 คน คือ t1 “ปองพล” และ t2 “บุรัสกร” โดยวิชาที่เกี่ยวข้องคือ

sec1 เรียนวิชา “Internet Technology in Daily Uses” มีผู้สอนคือ “ปองพล”

sec2 เรียนวิชา “Data Structure and Algorithm” มีผู้สอนคือ “บุรัสกร”

เพราะฉะนั้นเราจะเรียงลำดับ section จากสมการดังนี้

(factor1 x w1) + (factor2 x w2) จะได้

f(sec1) = (**t1.priority x** 0.7) + (**s1.subjectPriority x** 0.5)
= (0.3 **x** 0.7) + (0.8 **x** 0.5) = 0.61

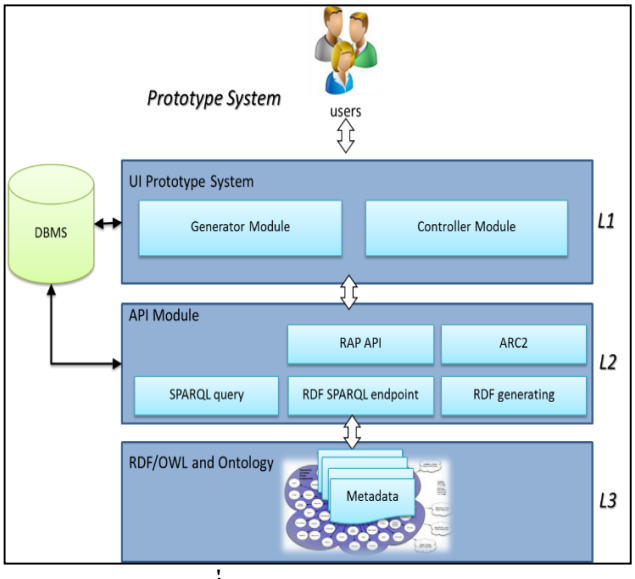
f(sec2) = (**t2.priority x** 0.7) + (**s2. subjectPriority x** 0.5)
= (0.7 **x** 0.7) + (0.3 **x** 0.5) = 0.64

ดังนั้น จะได้ลำดับของ section คือ sec2 > sec1 ตามลำดับ จากนั้นไปทำขั้นตอนที่ 3

ขั้นที่ 3 หาห้องให้เหมาะสมสำหรับ section และอาจารย์ผู้สอน โดยเรียงตามลำดับจากขั้นตอนที่ผ่านมา จากขั้นตอนที่แล้วเราได้ sect2 > sect1 ตามลำดับ ขั้นตอนต่อไปนี้คือการหาห้องที่เหมาะสมให้กับ section และอาจารย์ผู้สอน โดยใช้คุณสมบัติของออนโทโลยีที่ได้ออกแบบไว้

4. สถาปัตยกรรมของระบบ

ในการจัดการข้อมูลมีการเชื่อมโยงกันในแต่ละ layer โดยมีการติดต่อกับผู้ใช้ในส่วนของ layer 1 และนำคำสั่งที่ได้รับมาติดต่อกับส่วนฐานข้อมูล และ layer 2 ซึ่งเป็นส่วนของโมดูล API จากนั้นจะมีการติดต่อกับ layer 3 ซึ่งเป็นส่วนขององค์ความรู้หรือออนโทโลยีที่ได้ออกแบบ โดยดูได้จากภาพที่ 1



ภาพที่ 1: สถาปัตยกรรมของระบบ

ระบบต้นแบบพัฒนาโดยใช้ RAP API [12] เป็น API หลักในการดึงฐานความรู้ออกมาใช้งาน และใช้ SPARQL [13] เพื่อทำการเรียกใช้งานฐานความรู้และกฎที่ได้ออกแบบไว้ ในการทำงานผู้ใช้งานเพียงสร้างฐานความรู้ให้สอดคล้องกับความต้องการและเข้าสู่ระบบเพื่อสร้างตารางสอนแบบอัตโนมัติ จากนั้นระบบจะทำการแสดงข้อมูลตารางสอนของผู้สอนแต่ละท่านที่อยู่ในระบบแบบอัตโนมัติ

5. ผลการทดลองและผลประเมิน

ระบบจัดตารางสอนแบบอัตโนมัติของกรณีศึกษา ได้นำออนโทโลยีสำหรับสารสนเทศที่เกี่ยวข้องพร้อมกับกฎ รวมถึงการออกแบบขั้นตอนวิธีสำหรับเลือกรายวิชา ห้องเรียน และอาจารย์ผู้สอนให้สอดคล้องกัน และเรียงลำดับความสำคัญที่มีความใกล้เคียงที่สุดโดยอาศัยการคำนวณทางคณิตศาสตร์ เพื่อช่วยให้ระบบทำการลงตารางสอนได้สะดวกรวดเร็ว โดยเมื่อเปรียบเทียบกับลักษณะการทำงานแบบดั้งเดิมพบว่า การใช้ออนโทโลยีสำหรับจัดเก็บองค์ความรู้ เพื่อทำงานแบบเว็บเชิงความหมายซึ่งมีคุณสมบัติในการแบ่งปันข้อมูล การจัดเก็บข้อมูลให้อยู่ศูนย์กลางโดยไม่มีการทำสำเนาข้อมูลแบบซ้ำซ้อน รวมถึงการสนับสนุนความคิดเชิงตรรกศาสตร์ ส่งผลให้ระบบใหม่ทำงานได้เร็วกว่าระบบแบบเดิมที่ทำงานด้วยมนุษย์

จากการสำรวจระดับความพึงพอใจในการใช้งานระบบจัดตารางสอนแบบอัตโนมัติ โดยเมื่อเปรียบเทียบการทำงานแบบเดิมและการใช้ระบบในส่วนของผู้ใช้งานหลักพบว่าผู้ใช้งานหลักมีระดับความพึงพอใจในการใช้งานระบบอยู่ในระดับมาก คิดค่าเฉลี่ยโดยรวมได้เท่ากับ 4.04 และมีส่วนเบี่ยงเบนมาตรฐานเฉลี่ยเท่ากับ 0.62

แต่ถึงกระนั้น คำตอบที่ได้จากระบบยังไม่เป็นคำตอบที่ดีที่สุดจากความต้องการทั้งหมด ข้อเสนอแนะคือ ถ้าต้องการให้มีความแม่นยำสูงควรเพิ่มปัจจัยแวดล้อม เครื่องมือและขั้นตอนวิธีอื่น ๆ ในการคำนวณ แต่อย่างไรก็ตามการใส่เงื่อนไขให้กับระบบมากเกินไปอาจทำให้ระบบช้าตามรูปแบบปัญหาของเอ็นพีบริบูรณ์ ซึ่งอาจเป็นผลเสียมากกว่าผลดี

เอกสารอ้างอิง

[1] G. Kendall, P. J. Andrew and S. Kristian, "A Survey of NP-Complete Puzzles," *ICGA Journal*, vol. 31.1, pp. 13-34, 2008.

[2] P. Hitzler, M. Krotzsch and S. Rudolph, Foundations of Semantic Web Technologies, CRC Press, 2009.

[3] S. Steffen and R. Studer, Handbook on ontologies, Springer Science & Business Media, 2013.

[4] F. Arvidsson and A. Flycht-Eriksson, "Ontologies I," 2008.

[5] W3C, "Hypertext Transfer Protocol," 2014. [Online]. Available: https://www.w3.org/Protocols/.

[6] W3schools, "Introduction to OWL," 2011. [Online]. Available: http://www.w3schools.com/rdf/rdf_owl.asp.

[7] J. F. Sowa and ed, Principles of semantic networks: Explorations in the representation of knowledge, Morgan Kaufmann, 2014.

[8] W3C Member, "A Semantic Web Rule Language Combining OWL and RuleML," 2004.

[9] S. Mirhassani and Habibi, "Solution approaches to the course timetabling problem," *Artificial Intelligence Review*, 2011.

[10] I. Aljarah, A. Salhieh and H. Faris, "An Automatic Course Scheduling Approach Using Instructors' Preferences," Jordan University of Science & Technology, Amman, 2014.

[11] S. Gupta and S. Senthilkumar, "Timetable Generation Using AI Algorithms," 2008.

[12] D. Westphal and C. Bizer, "Introduction to RAP," 2004.

[13] W3C SPARQL Group, "SPARQL," 2013. [Online]. Available: http://www.w3.org/TR/rdf-sparql-query/.

แอปพลิเคชันช่วยเหลือผู้บกพร่องทางการได้ยินในการเดินทางด้วยระบบขนส่งรถไฟฟ้า
Airport Rail Link Transportation Application for the Hearing Impaired (NextStop)

อานนท์ ชลวุฒิ (Arnon Chonrawut)[1] และตรีเทพ รัตนพิภพ (Trithep Ratanapipop)[2]
คณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
[1]anonchollawut@gmail.com, [2]the9uide@outlook.com

บทคัดย่อ

ปัจจุบันการเดินทางด้วยรถไฟฟ้าเป็นสิ่งที่คนไทยเริ่มให้ความสนใจ และหันมาใช้กันมากขึ้น ในการเดินทางด้วยรถไฟฟ้าจำเป็นที่จะต้องรับฟังประกาศต่าง ๆ ในขบวนรถไฟฟ้า เพื่อให้การเดินทางเป็นไปได้อย่างราบรื่น เช่น การประกาศสถานีถัดไปที่รถไฟฟ้าจะจอด และสถานีปลายทางที่รถไฟฟ้าจะไปถึง ซึ่งจะมีผลต่อผู้โดยสารที่เป็นผู้บกพร่องทางการได้ยิน ในช่วงเวลาที่มีคนใช้บริการรถไฟฟ้าจำนวนมาก ทำให้โดนเบียดไปอยู่ในตำแหน่งที่ไม่สามารถมองเห็นป้ายประกาศได้ชัดเจนดังนั้นโครงการนี้จึงได้พัฒนาแอนดรอยด์แอปพลิเคชัน ที่ช่วยแจ้งข้อมูลที่สำคัญให้กับผู้โดยสารที่เป็นผู้บกพร่องทางการได้ยิน เพื่อให้การเดินทางด้วยรถไฟฟ้าของผู้บกพร่องทางการได้ยินสะดวกมากยิ่งขึ้น จึงจำเป็นต้องมีการแจ้งข้อมูลที่จำเป็นให้กับผู้บกพร่องทางการได้ยิน เพื่อป้องกันปัญหาการนั่งเลยสถานีที่ตัวเองต้องการจะลงหรือการลงสถานีที่ผิดพลาดเพราะได้รับข้อมูลไม่เพียงพอ

คำสำคัญ: ผู้พิการทางการได้ยิน รถไฟฟ้า แอนดรอยด์ แอปพลิเคชัน

Abstract

Nowadays, skytrain transportation interests Thai people and increase number of using. Traveling by the skytrain, Passenger should pay attention to the announcement carefully. (e.g. Announcement for the next stop, the terminal station etc.) To be unable to get an information by announcement could be significant obstacle for hearing impaired passengers. Therefore, we developed android application to inform and notify necessary information to hearing impaired passengers while traveling. This information can help hearing impaired passengers to travel more comfortable with skytrain and prevent them from missing their destination.

Keyword: Hearing Impaired, Skytrain, Android Application

1. บทนำ

ปัจจุบันการเดินทางด้วยรถไฟฟ้าเป็นสิ่งที่คนเริ่มให้ความสนใจ และหันมาใช้กันมากขึ้น ในการเดินทางด้วยรถไฟฟ้าจำเป็นต้องใช้การรับฟังประกาศต่าง ๆ เพื่อให้การเดินทางเป็นได้อย่างราบรื่น เช่น การประกาศสถานีถัดไปที่รถไฟฟ้าจะหยุด และสถานีปลายทางที่รถไฟฟ้าจะไปถึง จากสถิติจำนวนผู้พิการในกรุงเทพมหานคร มีจำนวนผู้พิการทางการได้ยินทั้งหมด 12,055 คน [1] ที่มีบัตรประจำตัวผู้พิการ การเดินทางด้วยรถไฟฟ้าที่ในช่วงเวลาที่มีคนใช้บริการเป็นจำนวนมาก จึงส่งผลให้ผู้โดยสารที่เป็นผู้บกพร่องทางการได้ยิน เดินทางได้ไม่สะดวกสบายเท่าที่ควร ถึงแม้ว่าจะมีการติดตั้งจอแสดงผลไว้ให้บริการในขบวนรถไฟฟ้าแล้วก็ตาม แต่ก็ยังทำให้การแจ้งข่าวสารเป็นไปได้ลำบากเพราะผู้โดยสารที่บกพร่องทางการได้ยินจะต้องคอยดูที่หน้าจอการแสดงผลตลอดเวลาที่รถไฟฟ้ามีการชะลอตัว และหากรถไฟฟ้าเที่ยวนั้นมีผู้ใช้บริการเป็นจำนวนมาก ก็จะส่งผลทำให้ผู้โดยสารที่บกพร่องทางการได้ยินอาจโดนเบียดเข้าไปในจุดที่ทำให้มองจอแสดงผลได้ลำบากหรือไม่สามารถมองเห็นได้เลย [2]

เพื่อให้การเดินทางด้วยรถไฟฟ้าของผู้บกพร่องทางการได้ยินสะดวกมากยิ่งขึ้น จึงจำเป็นต้องมีการแจ้งข้อมูลที่จำเป็น

ให้กับผู้บกพร่องทางการได้ยิน เพื่อป้องกันปัญหาการนั่งเลยสถานีที่ต้องการจะลงหรือการลงสถานีที่ผิดพลาดเพราะข้อมูลที่ไม่เพียงพอ เราจึงได้พัฒนาแอนดรอยด์แอปพลิเคชันที่ช่วยแจ้งข้อมูลที่สำคัญให้กับผู้โดยสารที่เป็นผู้บกพร่องทางการได้ยิน โดย แอปพลิเคชันจะแจ้งข้อมูลว่ารถไฟฟ้าที่กำลังโดยสารอยู่นั้นผ่านสถานีอะไรมาแล้ว และสถานีถัดไปคือสถานีอะไรผ่านระบบ GPS และนำข้อมูลตำแหน่งที่ได้มาไปคำนวณระยะเวลาที่ใช้ในการเดินทางไปยังสถานีปลายทางที่ต้องการจะลงผ่าน Google map API พร้อมกับแจ้งเตือนเมื่อใกล้จะถึงสถานีปลายทางที่ผู้ใช้ต้องการจะลง และยังบอกข้อมูลเกี่ยวกับสถานที่ที่เป็นที่นิยมของแต่ละสถานี เช่น ร้านอาหาร ร้านเครื่องดื่ม เป็นต้นผ่าน Foursquare API เพื่อให้ข้อมูลกับผู้ใช้เป็นตัวช่วยในการตัดสินใจเพื่อที่จะแวะลงสถานีนั้น ๆ ในการเดินทางครั้งต่อไป

2. เทคโนโลยีที่ใช้ในการพัฒนาแอปพลิเคชัน

2.1 Global Positioning System (GPS)

เป็นระบบสำหรับระบุตำแหน่งบนพื้นโลก โดยใช้สัญญาณจากดาวเทียมอย่างน้อย 3 ดวง ในการใช้ระบุตำแหน่งของ GPS ซึ่งจะเปลี่ยนสัญญาณเป็นตำแหน่ง (X, Y, Z) มีประโยชน์ในการใช้ระบุตำแหน่งในแผนที่ หรือสามารถมาประยุกต์ใช้ในการนำทางได้ [3]

2.2 Foursquare API

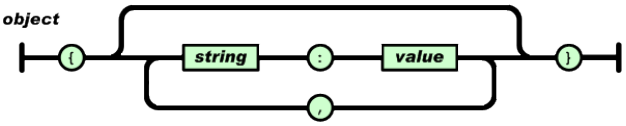
เป็นบริการของ Foursquare ในการให้บริการข้อมูลของสถานที่ต่าง ๆ ผ่านอินเทอร์เน็ต โดยผู้พัฒนาจะต้องทำการลงทะเบียนเพื่อขอสิทธิ์ในการใช้งาน ตัวระบบมีการใช้ RESTful และรองรับ XML, JSON, JSONP ในการส่งข้อมูล ซึ่งมีความสามารถในการค้นหาสถานที่ต่าง ๆ อย่างเป็นหมวดหมู่ และสามารถระบุบริเวณรอบพิกัดที่จะต้องการค้นหาได้ [4]

2.3 Google Map Direction API

เป็นบริการที่ Google เปิดให้ใช้บริการ ซึ่งเป็นบริการย่อยของ Google Map ที่ช่วยในการคำนวณการเดินทางต่าง ๆ โดยสามารถใช้คำนวณระยะทางจากต้นทางถึงปลายทาง คำนวณระยะเวลาที่ใช้ในการเดินทาง และสามารถเลือกรูปแบบการเดินทางเป็น รถยนต์ รถไฟ เดิน และจักรยานได้ [5]

2.4 JavaScript Object Notation (JSON)

เป็นรูปแบบการเก็บข้อมูลประเภทหนึ่งที่นิยมใช้มากในปัจจุบัน ซึ่งเป็นภาษาที่คนสามารถอ่านและเขียนได้ง่าย และคอมพิวเตอร์สามารถนำข้อมูลไปใช้ได้ง่าย โดยนิยมนำมาใช้ในการเก็บข้อมูลไว้สำหรับใช้ในแอปพลิเคชัน หรือใช้ในการส่งข้อมูลผ่าน API ต่างๆ ซึ่งมีความสามารถในการเก็บข้อมูลเป็นรูปแบบของ Object, Array, Value, String, Number สำหรับเตรียมเพื่อที่สามารถนำไปใช้ในโปรแกรมได้ง่าย โดยหลายๆ ภาษานั้นจะมีเครื่องมือที่ช่วยในการแปลงให้เข้ากับแต่ละภาษาที่ใช้ได้ ซึ่งมีเครื่องมือให้สามารถใช้ได้ในภาษา C, C++, C#, Java, JavaScript, Perl, Python และอื่นๆ [6]



ภาพที่ 1: การเก็บข้อมูลรูปแบบ Object

2.5 งานวิจัยที่เกี่ยวข้อง

ข้อมูลในการเดินทางที่ดีนั้นเป็นส่วนหนึ่งที่สำคัญในการบริการของระบบขนส่งสาธารณะ Nina Waara [7] ได้ทำวิจัยเกี่ยวกับความต้องการของข้อมูลในการเดินทางของผู้สูงอายุและผู้พิการพบว่า ผู้พิการยังคงมีความต้องการในการเข้าถึงข้อมูลของการเดินทาง ในระหว่างที่กำลังเดินทางอยู่บนขบวนรถไฟด้วย เช่น สถานีปลายทาง สถานีถัดไปที่รถไฟจะจอด หรือข้อมูลที่บอกว่าผู้โดยสารควรลงฝั่งไหนของขบวนรถไฟ

Chui Yin Wong [8] ได้ทำการวิจัยเกี่ยวกับการออกแบบส่วนติดต่อกับผู้ใช้สำหรับผู้พิการทางการได้ยิน พบว่าภาษาเขียนที่ใช้ในการสื่อสารนั้นยังมีผลต่อความเข้าใจของตัวผู้พิการทางการได้ยินอีกด้วย เพราะพวกเขาไม่สามารถเข้าใจภาษาพูดได้อย่างพวกคำอุทานได้ หรือแม้กระทั่งตัวย่อของคำต่าง ๆ แม้ว่าผู้พิการทางการได้ยินจะชอบสัญลักษณ์ต่าง ๆ แต่บางครั้งสัญลักษณ์ก็ทำให้พวกเขาเข้าใจได้ยาก หากไม่มีคำอธิบายประกอบ Gary R. Bettger และ Timothy J. Pearson [9] ได้ทำวิจัยเกี่ยวกับการช่วยเหลือผู้พิการทางการได้ยินในระบบขนส่งมวลชนสาธารณะในรัฐแมสซาชูเซ็ท ซึ่งได้แนะนำวิธีการในการช่วยเหลือด้วยเทคโนโลยีที่มีอยู่ในปัจจุบันนี้ คือ Amplifier/Receiver ทำหน้าที่แปลงประกาศสาธารณะต่าง ๆ เป็นสัญญาณ FM และส่งไปยังตัวรับที่มีลักษณะคล้ายกับหูฟัง

แต่ยังไม่สามารถใช้ได้เพราะบางสถานีไม่อนุญาต Induction loop เป็นหลักการที่ใช้การสร้างสนามแม่เหล็กไฟฟ้า ที่ส่งผลกับตัว Telecoil ในการช่วยเหลือการได้ยิน แต่ก็มีปัญหากับขบวนรถที่ทำมาจากเหล็กเพราะทำให้อาจเกิดข้อผิดพลาดกับสัญญาณที่ส่งไปยังตัว Telecoil ได้ การใช้จอโทรทัศน์ทั่ว ๆ ไป ในการแสดงข้อมูลของรถไฟก็ไม่ค่อยเหมาะสม เพราะว่ามีข้อมูลแสดงผลบนจอเยอะมากจนเกินไป Electronic readerboard เป็นป้าย LED ที่จะมีข้อความวิ่ง แม้ว่าจะทำงานได้ดีในตัวขบวนรถไฟฟ้า แต่ก็มีปัญหามองเห็นได้ไม่ชัดเจนในช่วงเวลาที่มีคนใช้บริการเป็นจำนวนมาก Telecommunication Device for the Deaf (TDDs) จะต้องใช้ TDD ในการส่งและการรับเท่านั้น ทำให้มีปัญหาการติดต่อสื่อสารไม่ได้ครอบคลุม [9] Action On Hearing Loss [10] ก็ได้เสนอนโยบายของการให้บริการกับผู้พิการทางการได้ยินในการใช้บริการรถไฟไว้คือ ต้องมีเจ้าหน้าที่ที่ถูกฝึกในการสื่อสารกับผู้พิการทางการได้ยิน ต้องมีจอแสดงผลข้อมูลในแต่ละขบวนรถไฟฟ้า และต้องเตือนผู้โดยสารในเรื่องของสถานีถัดไป ต้องมีสิ่งที่คอยแจ้งเตือนผู้โดยสารผ่านการมองอย่างเช่นไฟบนประตูในกรณีที่ประตูเปิดปิด หรือไฟกระพริบในกรณีที่เกิดเหตุฉุกเฉิน ในแต่ละสถานีต้องมีจอแสดงข้อมูลแบบ real-time และต้องมีพื้นที่ที่ช่วยเหลือด้านข้อมูลที่ใช้งานได้ง่ายสำหรับผู้พิการทางการได้ยินโดยจะมีจอแสดงผลและแป้นพิมพ์ และภาษาที่ใช้ในการเขียนขั้นตอนต่าง ๆ บนสถานีควรเขียนให้ชัดเจน ทาง Papworth Trust [11] ได้บอกว่าวิธีหนึ่งในการสื่อสารข้อมูลให้กับผู้โดยสารให้มีประสิทธิภาพ คือ การทำแอปพลิเคชันบน Smartphone ผู้โดยสารจะรับรู้ข้อมูลได้ดีขึ้นเกี่ยวกับการเดินทางของพวกเขา ถ้าหากมีการเปิดข้อมูลเหล่านั้นออกสู่สาธารณะมากพอ มีแอปพลิเคชันหลายตัวที่พร้อมให้บริการแล้ว แต่ก็ยังคงขาดข้อมูลต่าง ๆ ที่มาจากผู้ให้บริการของสถานีรถไฟเอง

3. วิธีการดำเนินการวิจัย

3.1 การวิเคราะห์ความต้องการ

จากการสอบถามและเก็บข้อมูลจากสมาคมคนหูหนวกแห่งประเทศไทย ทำให้เราวิเคราะห์ความต้องการได้ดังนี้

- สามารถแสดงเส้นทางการเดินรถของระบบรถไฟฟ้า Airport Rail Link ได้
- สามารถแสดงตำแหน่งสถานีปัจจุบันของผู้ใช้ได้
- สามารถดูรายละเอียดของการเดินทางได้
- สามารถแจ้งเตือนเมื่อใกล้ถึงสถานีที่ต้องการจะลงหรือแจ้งเตือนเมื่อเลยสถานีแล้วได้
- สามารถยกเลิกการแจ้งเตือนได้
- สามารถดูรายละเอียดของแต่ละสถานีได้
- สามารถดูรายละเอียดของสถานที่บริเวณใกล้รถไฟฟ้า
- สามารถแสดงสถานีที่ผู้ใช้ต้องการจะลงได้

โดยผู้พัฒนาได้ทำการสอบถามและเก็บความต้องการจากสมาคมคนหูหนวกแห่งประเทศไทย

3.2 การออกแบบและพัฒนาระบบ

ในการออกแบบส่วนติดต่อกับผู้ใช้นั้นหลังจากที่ได้ทำการเก็บข้อมูลจากสมคมคนหูหนวกแห่งประเทศไทยและการศึกษางานวิจัยต่าง ๆ พบว่าผู้ที่มีการบกพร่องทางการได้ยินนั้นจะมีปัญหาเรื่องการอ่านหนังสือในส่วนที่เขียนเป็นภาษาพูดทำให้ไม่เข้าใจในการสื่อสาร ต้องใช้เป็นภาษาเขียนจึงจะสามารถเข้าใจได้ง่าย และจะเข้าใจได้ดียิ่งขึ้นหากมีการใช้สัญลักษณ์ เข้ามาช่วยอธิบาย ๆ ดังนั้นในการออกแบบส่วนติดต่อกับผู้ใช้ จึงจะใช้คำภาษาเขียนที่เข้าใจได้ง่าย และมีการนำสัญลักษณ์เข้ามาช่วยอธิบายในส่วนต่าง ๆ ด้วย

ในส่วนของการออกแบบการทำงานแอปพลิเคชั้นนั้น จะมีการเก็บข้อมูลของสถานีรถไฟฟ้าโดยใช้รูปแบบไฟล์ JSON และในการคำนวณเวลาที่ผู้ใช้จะถึงสถานีปลายทาง จะทำโดยการอ้างอิงตำแหน่งของผู้ใช้จาก GPS และส่งข้อมูลผ่านอินเตอร์เน็ตไปยัง Google Map Direction API และรับข้อมูลกลับมาเป็น JSON ในส่วนของสถานที่โดยรอบของสถานีจะทำการส่งข้อมูลตำแหน่งของแต่ละสถานีไปยัง Foursquare API เพื่อรับข้อมูลของสถานที่โดยรอบกลับมาเป็น JSON

4 ผลการดำเนินงาน

4.1 ผลการทดลอง

เนื่องจากการติดตั้งป้ายแสดงผลบนขบวนรถไฟฟ้า Airport Rail Link นั้น มีไม่เพียงพอต่อการแจ้งข้อมูลให้กับผู้โดยสารที่

เป็นผู้พิการทางการได้ยิน หรือบางครั้งอาจโดนเบียดไปอยู่ในตำแหน่งที่ไม่สามารถมองเห็นจอแสดงผลได้อย่างชัดเจน เราจึงนำข้อมูลเหล่านั้น เช่น เวลาในการเดิน สถานีถัดไปที่รถไฟฟ้าจะจอด สถานีปลายทางที่ผู้ใช้ต้องการจะลง และการแจ้งเตือน เป็นต้น มาแสดงผลในรูปแบบของ Mobile Application ซึ่งเป็นสิ่งที่ผู้พิการทางการได้ยินสามารถเข้าถึงได้ง่าย

ในอนาคตทางผู้พัฒนาตั้งใจจะรวมระบบนี้เข้ากับระบบของทางสมาคมผู้พิการทางการได้ยิน เพื่อนำบัตรประจำตัวผู้พิการและการยืนยันตัวบุคคลเข้ามาช่วยในเรื่องของการออกตั๋วโดยสาร ทำให้การออกตั๋วโดยสารระหว่างเจ้าหน้าที่และผู้พิการทางการได้ยินมีความสะดวกมากยิ่งขึ้น และจะพัฒนาระบบที่รองรับการแจ้งเตือนไปยังตัวแอปพลิเคชัน ในกรณีที่รถไฟฟ้าเกิดเหตุการณ์ขัดข้อง

4.2 ตัวอย่างหน้าจอแอปพลิเคชัน

เมื่อผู้ใช้เข้าแอปพลิเคชันครั้งแรก ก็จะพบกับหน้ายืนยันตำแหน่งเริ่มต้นของสถานีที่ผู้ใช้อยู่ ผ่านการตรวจจับด้วยระบบ GPS ดังภาพที่ 2



ภาพที่ 2: หน้าจอถามสถานีผู้ใช้งาน

ผู้ใช้สามารถเลือกสถานีปลายทางเองได้จากหน้าจอแสดงรายชื่อสถานีของรถไฟฟ้าสำหรับเลือกสถานีปลายทางที่จะต้องการจะแจ้งเตือนดังภาพที่ 3



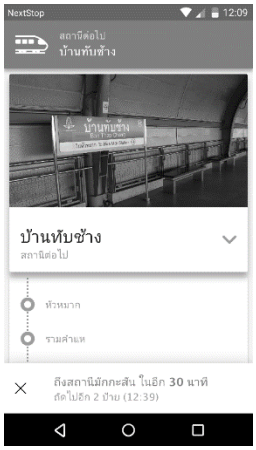
ภาพที่ 3: หน้าจอเลือกสถานี

หลังจากเลือกสถานีปลายทางแล้วแอปพลิเคชันจะแสดงข้อมูลโดยรวมในการเดินทาง ดังภาพที่ 4



ภาพที่ 4: หน้าจอแสดงข้อมูลการเดินทาง

เมื่อกดปุ่ม “ เริ่ม” แอปพลิเคชันจะแสดงรายละเอียดของการเดินทางโดยจะบอกว่าใช้เวลาอีกกี่นาทีในการเดินทาง เหลืออีกกี่สถานีและสถานีต่อไปเป็นสถานีอะไร ดังภาพที่ 5



ภาพที่ 5: หน้าจอแสดงรายละเอียดระหว่างการเดินทาง

ผู้ใช้สามารถกดดูรายละเอียดของแต่ละสถานีได้ ดังภาพที่ 6



ภาพที่ 6: หน้าจอแสดงรายระเอียดของแต่ละสถานี

เมื่อกดดูรายละเอียดแล้ว ผู้ใช้สามารถปัดหน้าจอขึ้นเพื่อดูสถานที่ที่อยู่บริเวณรอบสถานีนั้น ๆ ได้ ดังภาพที่ 7



ภาพที่ 7: หน้าจอแสดงสถานที่บริเวณรอบสถานี

เมื่อถึงสถานีปลายทางแล้วแอปพลิเคชันจะทำการแจ้งเตือน โดยจะมีการแจ้งเตือน 2 ครั้ง ในครั้งแรกก่อนถึงสถานีปลายทาง 1 ป้ายและ ครั้งที่สองเมื่อถึงสถานีปลายทาง ดังภาพที่ 8



ภาพที่ 8: หน้าจอการแจ้งเตือนเมื่อถึงสถานีปลายทาง

ผู้ใช้สามารถกดกากบาทที่มุมด้านซ้ายล่าง เพื่อยกเลิกการแจ้งเตือนสถานีปลายทางที่เลือกไปแล้วได้ ดังภาพที่ 9



ภาพที่ 9: หน้าจอแสดงการยกเลิกการแจ้งเตือน

5. สรุป

การทำงานของระบบนี้จะมีการทำงานหลักคือการแจ้งเตือนและให้ข้อมูลการเดินทางให้กับผู้ใช้งานแอปพลิเคชันเพื่อป้องกันการนั่งเลยสถานีปลายทางที่ต้องการจะลง โดยเมื่อเริ่มใช้แอปพลิเคชัน ระบบจะทำการค้นหาตำแหน่งปัจจุบันและหาสถานีที่ใกล้ที่สุดของผู้ใช้ให้โดยอัตโนมัติ ผู้ใช้ทำการกำหนดสถานีปลายทางที่ต้องการจะลง ในระหว่างเดินทางจะสามารถดูรายละเอียดของการเดินทาง และจะมีการบอกเวลาที่ใช้ในการ

เดินทางโดยจะคำนวณจากตำแหน่งปัจจุบันของผู้ใช้ และรายละเอียดของสถานที่ในบริเวณรอบ ๆ สถานีที่ดึงข้อมูลจาก Foursquare

ตัวระบบจะทำการตรวจจับสถานีเริ่มต้นของผู้ใช้ให้เองอัตโนมัติ ผู้ใช้แค่ทำการเลือกสถานีที่ต้องการจะลงเพียงแค่นั้น ตัวระบบก็จะแจ้งข้อมูลที่สำคัญ ๆ ให้กับผู้ใช้ พร้อมกับแจ้งเตือนก่อนถึงสถานีปลายทาง 1 ป้าย และแจ้งเตือนซ้ำอีกครั้งเมื่อถึงสถานีปลายทางที่ผู้ใช้เลือกไว้ โดยการแจ้งเตือนจะใช้การสั่นและกระพริบแสงแฟลชของตัว Smart phone ตัวแอปพลิเคชันสามารถทำงานแบบ Background process ได้ ผู้ใช้สามารถกดปิดหน้าจอหลังจากที่เลือกสถานีปลายทางไปแล้วได้ ตัวระบบก็ยังคงทำงานอยู่ และจะแจ้งเตือนเมื่อถึงสถานีปลายทางที่ผู้ใช้เลือกไว้

ตัวแอปพลิเคชันมีปัญหาในเรื่องของการคำนวณเวลาที่ยังไม่แม่นยำมากพอ หากตัวขบวนรถไฟฟ้าเกิดอาการผิดปกติหรือหยุดกะทันหัน และไม่สามารถคำนวณเวลาที่รถไฟฟ้าจะเข้ามาเทียบท่าที่สถานีเริ่มต้นของผู้ใช้ได้ และในอนาคตจะมีการนำแอปพลิเคชันออกมาทดสอบกับผู้พิการทางการได้ยินเพื่อประเมินคุณภาพและความพึงพอใจของผู้ใช้

เอกสารอ้างอิง

[1] กระทรวงการพัฒนาสังคมและความมั่นคงของมนุษย์, “รายงานข้อมูลสถานการณ์ด้านคนพิการในประเทศไทย,” .2016 11 10]ออนไลน์ .[Available: http://dep.go.th/th/page/รายงานสถานการณ์คนพิการ.

[2] พ.เสือเขียว, “โลกของคนหูหนวก,” เว็บท่าสำหรับคนพิการ, .[ออนไลน์]Available: http://www.pwdsthai.com/index.php?option=com_content&view=article&id=-05-07-11-02-527:2014 17&catid=:96cat-handic-hear&Itemid=] .507 ที่เข้าถึง 6 Feburary .[2017

[3] Wikipedia, “Global Positioning System,” .[ออนไลน์] Available: https://en.wikipedia.org/wiki/Global_Positioning_System. [ที่เข้าถึง.[2017 2 12

[4] Foursquare, “Foursquare for Developers,” .2016 11 10]ออนไลน์ .[Available: https://developer.foursquare.com.

[5] Google, “Google Maps Directions API | Google Developers,” Google, .[ออนไลน์] Available: https://developers.google.com/maps/documentation/directions/. [ที่เข้าถึง.[2017 2 13

[6] E. International, “JSON,” .[ออนไลน์] Available: http://www.json.org. [ที่เข้าถึง.[2017 1 12

[7] N. Waara, OLDER AND DISABLED PEOPLE’S NEED AND VALUATION, Sweden : Lund University, .2009

[8] C. Y. Wong และ C. W. Khong, Designing a Mobile User Interface for the Deaf: A Malaysian Context, Malaysia: Multimedia University, .2015

[9] Gary และ Timothy, Accomodating deaf and hard of hearing persons on public transportation system in Massachusetts, Massachusetts: Harvard University, .1988

[10] A. O. H. Loss, Access to rail travel for people with hearing loss Policy statement, United Kingdom, .2014

[11] P. Trush, Improving rail travel for disabled passengers in Britain, United Knigdom: Cambridge, .2012

การประยุกต์ใช้ระบบระบุตำแหน่งภายในอาคารเพื่อจัดการองค์ความรู้ประจำจุด

An Applied of Indoor Positioning System for Local-based Knowledge Management

ชิษณุพงศ์ สุขสว่าง (Chisanupong Sooksawang)[1] ธนกฤช ตันติธนกร (Thanakit Tantithanakorn)2

บุญประเสริฐ สุรักษ์รัตนสกุล (Boonprasert Surakratanasakul)3

คณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

[1]b3nziie@hotmail.com, [2]billthebigest@gmail.com, [3]boonprasert@it.kmitl.ac.th

บทคัดย่อ

การจัดการองค์ความรู้ประจำจุด คือ การเข้าถึงองค์ความรู้ที่มีความเกี่ยวเนื่องกับตำแหน่งพิกัดของวัตถุในองค์กร เช่น องค์ความรู้ที่เกี่ยวข้องกับอุปกรณ์หรือเครื่องจักร งานวิจัยนี้จึงได้นำเสนอเครื่องมือที่ช่วยจัดการองค์ความรู้ประจำจุดโดยใช้วิธีการระบุตำแหน่งพิกัดภายในองค์กร ซึ่งอยู่ในรูปแบบของแอปพลิเคชันบนมือถือ เพื่อให้ผู้ใช้สามารถเรียกดูองค์ความรู้ได้ทุก ๆ ตำแหน่งพิกัดที่ประจำอยู่ภายในองค์กร ประโยชน์ที่ได้รับจากการพัฒนาระบบนี้ ทำให้การจัดการองค์ความรู้ภายในองค์กรมีความเป็นรูปธรรม และช่วยส่งเสริมให้เกิดกิจกรรมที่เกี่ยวข้องกับการจัดการความรู้ รวมถึงการพัฒนาต่อยอดองค์ความรู้ใหม่

คำสำคัญ: การระบุตำแหน่งภายในอาคาร องค์ความรู้ประจำจุด การจัดการความรู้

Abstract

An important of location-based knowledge management is an accessibility of organizational knowledge that embedded in physical of organization, for example, knowledge in machinery at industrial factory. For these, this research purposes an approach to manage knowledge using the location-based within an organization. This system developed based on mobile platform for user accessible and utilizable knowledge on reference location. The benefits of this approach are tangible of knowledge and encourage knowledge management activities for sustainable.

Keyword: Indoor Positioning System, Location-based Knowledge, Knowledge Management.

1. บทนำ

การระบุตำแหน่งภายในอาคารด้วยเทคโนโลยีแบบ Global Positioning System (GPS) นั้น ไม่สามารถทำการระบุตำแหน่งได้ เนื่องจากมีสนามแม่เหล็กจากวัสดุภายในอาคารก่อกวนสัญญาณดาวเทียมที่ใช้ทำการระบุตำแหน่ง ด้วยเหตุนี้จึงเกิดเทคโนโลยีในการระบุตำแหน่งภายในอาคาร Indoor Positioning System (IPS) ซึ่งเป็นการระบุตำแหน่งภายในอาคารโดยใช้สัญญาณต่าง ๆ เช่น สัญญาณWi–Fi สัญญาณบลูทูธ รวมถึงคลื่นสนามแม่เหล็ก ประกอบกับการจัดการความรู้ในปัจจุบัน เป็นกลไกหนึ่งที่ใช้ในการขับเคลื่อนและเป็นตัวชี้วัดความสำเร็จขององค์กร โดยทั่วไปแล้ว องค์ความรู้ที่มีอยู่ในองค์กรมีหลากหลายรูปแบบ บางรูปแบบขึ้นอยู่กับภูมิทัศน์ในแต่ละองค์กร ตัวอย่างเช่น องค์ความรู้ประจำจุด ซึ่งจะพบได้ค่อนข้างมากในองค์กรที่ใช้เครื่องมือหรือเครื่องจักรเป็นหลักในการดำเนินงาน ซึ่งผู้ใช้งานประจำมักจะมีประสบการณ์หรือความรู้เฉพาะบุคคลที่เกิดจากการใช้งานเครื่องมือนั้น ๆ

งานวิจัยนี้ จึงนำเสนอเทคโนโลยีในการระบุตำแหน่งภายในอาคารด้วยคลื่นสนามแม่เหล็ก เพื่อประยุกต์ใช้ในการต่อยอดการแบ่งปันองค์ความรู้ในองค์กรโดยเฉพาะองค์ความรู้ประจำจุด โดยนำเสนอ ดังนี้ ส่วนที่ 2 กล่าวถึงเทคโนโลยีที่เกี่ยวข้องกับการระบุตำแหน่งภายในอาคารและหลักการจัดการความรู้ในองค์กร ส่วนที่ 3 เป็นบทวิเคราะห์และออกแบบระบบ ส่วนที่ 4 เป็นพัฒนาระบบและการทดลอง และส่วนสุดท้าย เป็นบทสรุปและแนวทางการพัฒนาต่อยอดในอนาคต

2. ทฤษฎีและเทคโนโลยีที่เกี่ยวข้อง

2.1 การระบุตำแหน่งภายในอาคาร

ระบบระบุตำแหน่งภายในอาคาร (Indoor Positioning System: IPS) [1] ทำหน้าที่คล้ายคลึงกับระบบทางภูมิศาสตร์ (Global Positioning System: GPS) เพื่อใช้ในการระบุพิกัดตำแหน่งของวัตถุต่าง ๆ แต่ระบบ GPS มีข้อจำกัดเมื่อนำมาใช้กับการระบุตำแหน่งภายในอาคารหรือภายใต้สิ่งปลูกสร้าง เนื่องจากสิ่งปลูกสร้างเหล่านี้ลดทอนสัญญาณจากดาวเทียม จึงทำให้เกิดความคลาดเคลื่อนสูง [2] ดังนั้น จึงมีการพัฒนาเทคโนโลยีในการระบุตำแหน่งเมื่ออยู่ภายในอาคารที่เรียกว่า IPS ซึ่งใช้งานบนพื้นที่จำกัดและมีความซับซ้อนมากกว่า ปัจจุบันระบบ IPS มีวิธีการจัดการอยู่หลายรูปแบบ ดังนี้

1. การระบุโดยใช้สัญญาณบลูทูธ

การระบุตำแหน่งในสภาพแวดล้อมภายในอาคารโดยใช้สัญญาณบลูทูธที่นิยม คือ Beacons ซึ่งเป็นเทคโนโลยีที่ใช้วัดระยะความใกล้ ซึ่งหากหลุดจากรัศมีของสัญญาณก็จะทำให้อุปกรณ์ไม่สามารถรับข้อมูลจาก Beacons ได้ ซึ่งการทำงานของ Beacons จะมีหน้าที่ในการส่งสัญญาณออกมาอย่างเดียว ส่วนตัวอุปกรณ์พกพาหรือโทรศัพท์มือถือเมื่อได้รับสัญญาณจาก Beacons จะนำข้อมูลที่ได้รับมาประมวลผล เพื่อหาระยะห่างระหว่างอุปกรณ์โดยการใช้ APIs เช่น Estimote ซึ่งสามารถนำระยะห่างที่ได้นั้นมาแบ่งเป็นระยะใกล้หรือไกล ทำให้รู้ว่าผู้ใช้เดินเข้าใกล้ตัว Beacons นี้ในระยะแค่ไหนแล้วให้แสดงข้อมูลอะไร [3]

2. การระบุโดยใช้ความเข้มของสัญญาณWi-Fi

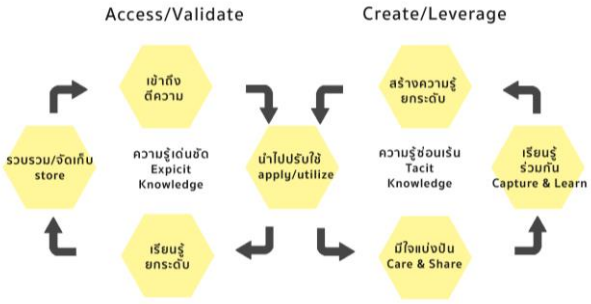
สัญญาณ Wi–Fi สามารถระบุตำแหน่งได้แม่นยำกว่าการใช้สัญญาณ 3G/4G เพราะบริเวณรัศมีของสัญญาณ Wi–Fi นั้นแคบกว่า ทำให้ระบุตำแหน่งได้ใกล้เคียงกว่า โดยมีหลักการทำงาน คือ การคำนวณระยะจากความเข้มของสัญญาณที่อุปกรณ์จับสัญญาณอยู่กับตัว Access Point ที่ระบุ [4]

3. การระบุโดยใช้สัญญาณคลื่นสนามแม่เหล็ก

เป็นเทคโนโลยีที่เปลี่ยนการรบกวนจากสนามแม่เหล็กในอาคารให้เกิดประโยชน์ โดยใช้การรบกวนของสนามแม่เหล็กเป็นตัวนำทาง ซึ่งก่อนทำการระบุ ผู้ใช้ต้องทำการเก็บค่าสนามแม่เหล็กภายในอาคารก่อน [5] โดยมีวิธีในการเก็บค่าดังนี้ ทำการอัพโหลดแผนผังพื้นอาคารเก็บไว้เป็นพื้นฐาน จากนั้นก็ทำการเดินไปตามทางเดินภายในตัวอาคาร โดยถือโทรศัพท์สมาร์ทโฟนที่จะทำการบันทึกค่าสนามแม่เหล็กไปตลอดเส้นทางเดิน ซึ่งแผนที่สนามแม่เหล็กนี้จะถูกบันทึกเก็บไว้ในฐานข้อมูลของ IndoorAtlas ทันทีที่แผนที่ถูกสร้างขึ้นผู้ใช้ก็สามารถใช้งานระบบการระบุตำแหน่งภายในอาคารได้เลย และยังสามารถพัฒนาแอปพลิเคชันโดยการใช้ APIs เพื่อดึงค่าตำแหน่งจากฐานข้อมูล

2.2 การจัดการความรู้

การจัดการความรู้ คือ การรวบรวมองค์ความรู้ที่มีอยู่อย่างกระจัดกระจายทั้งที่อยู่ในตัวบุคคลหรือรูปแบบที่จับต้องได้ มาพัฒนาให้เป็นระบบ เพื่อให้ทุกคนภายในองค์กรสามารถเข้าถึงความรู้ รวมทั้งสามารถนำไปประยุกต์ใช้ปฏิบัติงาน โดยมีวัตถุประสงค์ 3 ประการ คือ บรรลุเป้าหมายของงาน บรรลุเป้าหมายในการพัฒนาคน และบรรลุเป้าหมายการพัฒนาองค์กร ซึ่งการทำให้บรรลุเป้าหมายนั้น จำเป็นต้องเกี่ยวข้องกับงาน 2 ส่วน คือ ส่วนของการเข้าถึงความรู้ที่ใช่และถูกต้อง และส่วนของการสร้างและต่อยอดองค์ความรู้ โดยส่วนของการเข้าถึงจะมุ่งเน้นไปที่ความรู้แบบชัดแจ้ง ในขณะที่ส่วนของการสร้างและต่อยอดองค์ความรู้จะมุ่งเน้นที่องค์ความรู้แบบซ่อนเร้น โดยมีส่วนการนำไปใช้และประยุกต์ เป็นจุดเชื่อมต่อระหว่างทั้งสองงาน ดังแสดงในภาพที่ 1



ภาพที่ 1: วงจรการจัดการองค์ความรู้แบบชัดแจ้งและองค์ความรู้แบบซ่อนเร้น

กลไกการแลกเปลี่ยนเรียนรู้ (Knowledge conversion) [6] ก็เป็นพื้นฐานหนึ่งที่สำคัญในการจัดการความรู้ โดยการแลกเปลี่ยนนั้นอยู่บนพื้นฐานของกลไก 4 แบบ ดังนี้

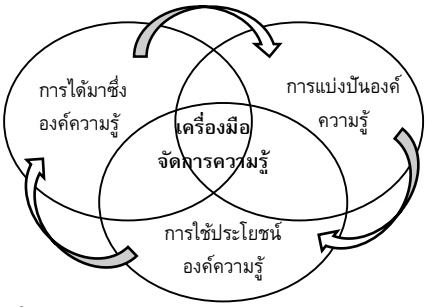
- กลไก Socialization เป็นการแลกเปลี่ยนจากองค์ความรู้แบบซ่อนเร้น สู่ องค์ความรู้แบบซ่อนเร้น ตัวอย่างเช่น การแลกเปลี่ยนหรือแบ่งปันประสบการณ์กันระหว่างบุคคล

- กลไก Externalization เป็นการแลกเปลี่ยนจากองค์ความรู้แบบซ่อนเร้น สู่ องค์ความรู้แบบชัดแจ้ง ตัวอย่างเช่น การบันทึกความรู้จากตัวบุคคลลงสื่อต่างๆ เช่น เขียนบทความจากประสบการณ์ลงหนังสือ

- กลไก Combination เป็นการแลกเปลี่ยนจากองค์ความรู้แบบชัดแจ้ง สู่ องค์ความรู้แบบชัดแจ้ง ตัวอย่างเช่น การรวบรวม การสรุป หรือการสำรวจจากสื่อต่าง ๆ

- กลไก Internalization เป็นการแลกเปลี่ยนจากองค์ความรู้แบบชัดแจ้ง สู่ องค์ความรู้แบบซ่อนเร้น ตัวอย่างเช่น การจัดอบรมถ่ายทอดจากผู้ที่มีประสบการณ์

ซึ่งกลไกข้างต้น จำเป็นต้องมีการดำเนินงานอย่างต่อเนื่องและอย่างเป็นระบบ โดยมีเป้าหมายให้บุคลากรในองค์กรสามารถเข้าถึงความรู้และพัฒนา บนพื้นฐานองค์ประกอบสำคัญของการจัดการความรู้ คือ คน เทคโนโลยี และกระบวนการนำมาใช้ [7] เพื่อก่อให้เกิดวงจรการจัดการองค์ความรู้ ดังภาพที่ 2 ซึ่งเป็นแนวคิดพื้นฐานสำคัญของการพัฒนาเครื่องมือการจัดการองค์ความรู้ [8]



ภาพที่ 2: พื้นฐานของการพัฒนาเครื่องมือการจัดการองค์ความรู้

3. การเปรียบเทียบและออกแบบระบบ

จากงานวิจัยที่ได้ทำการศึกษามานั้น ในเรื่องของการจัดการองค์ความรู้จะเป็นในรูปแบบของ Web-Base Application เพื่อใช้สำหรับการจัดการองค์ความรู้บนเว็บไซต์ [9] ซึ่งไม่ได้มีการพูดถึงตำแหน่งขององค์ความรู้ เราจึงได้ทำการประยุกต์ใช้ในการจัดการองค์ความรู้ซึ่งมีข้อดี คือทำให้การจัดการองค์ความรู้ภายในองค์กรมีความเป็นรูปธรรม และช่วยให้สามารถเข้าถึงข้อมูลองค์ความรู้ในสถานที่จริงได้ เพื่อแก้ไขเหตุการณ์เฉพาะหน้า

ในส่วนนี้จึงได้ทำการเปรียบเทียบข้อดี-ข้อเสียของระบบระบุตำแหน่งภายในอาคารรูปแบบต่าง ๆ เพื่อเป็นข้อมูลสำหรับการวิเคราะห์ออกแบบระบบและพัฒนาแอปพลิเคชัน

3.1 การเปรียบเทียบระบบระบุตำแหน่งภายในอาคาร

ในส่วนนี้ได้ทำการเปรียบเทียบระบบระบุตำแหน่งภายในอาคารประเภทต่าง ๆ เพื่อวิเคราะห์หาระบบที่เหมาะสำหรับการประยุกต์ใช้กับแอปพลิเคชันที่จะพัฒนา โดยทำการเปรียบเทียบสื่อสัญญาณที่ใช้ในการระบุพิกัด ซึ่งประกอบด้วย สัญญาณบลูทูท (Beacon) สัญญาณWi-Fi (Locata) และสัญญาณคลื่นสนามแม่เหล็ก (IndoorAtlas) [10] โดยแสดงผลการเปรียบในตารางที่ 1

ตารางที่ 1: ตารางเปรียบเทียบระหว่างระบบระบุพิกัดที่ใช้สัญญาณ Wi-Fi บลูทูธ และคลื่นสนามแม่เหล็ก

	Wi-Fi (Locata)	**Bluetooth (Beacon)**	**สนามแม่เหล็ก (IndoorAtlas)**
พื้นที่ครอบคลุมเฉลี่ย	บริเวณรอบ ๆ Access point 6-10 เมตร	บริเวณรอบ ๆ Beacon 1-2 เมตร	ไม่จำกัด
การใช้อุปกรณ์เสริมในการติดตั้ง	ใช้ตัวกระจายสัญญาณ	ใช้ตัวกระจายสัญญาณ	ไม่จำเป็น
ระบบปฏิบัติการที่รองรับ	Android	iOS/Android	iOS/Android
อัตราการตอบสนองระบุตำแหน่งแบบเวลาจริง (Real-time)	มีการตอบสนองทันที	มีความหน่วง (Delay) จากการตอบสนอง	มีการตอบสนองทันที
จัดการโดยใช้เพียงซอฟต์แวร์เท่านั้น	ต้องใช้อุปกรณ์เพิ่ม	ต้องใช้อุปกรณ์เพิ่ม	ใช้เพียงซอฟต์แวร์ในสมาร์ทโฟน
มีบริการสำหรับนักพัฒนา	ไม่มี	ไม่มี	มี SDK ให้ในการพัฒนา

จากตารางที่ 1 ระบบระบุพิกัดภายในอาคารโดยใช้สัญญาณบลูทูธ ต้องอาศัยตัวส่งสัญญาณที่เรียกว่า Beacon ซึ่งครอบคลุมพื้นที่สัญญาณประมาณ 2-3 ตารางเมตร ซึ่งอุปกรณ์พกพาส่วนใหญ่ มักมีส่วนติดต่อรับสัญญาณบลูทูธอยู่แล้ว จึงสามารถตรวจจับสัญญาณจาก Beacon เพื่อระบุตำแหน่งพิกัดได้ แต่มีข้อเสีย ระยะทำการที่ค่อนข้างแคบ

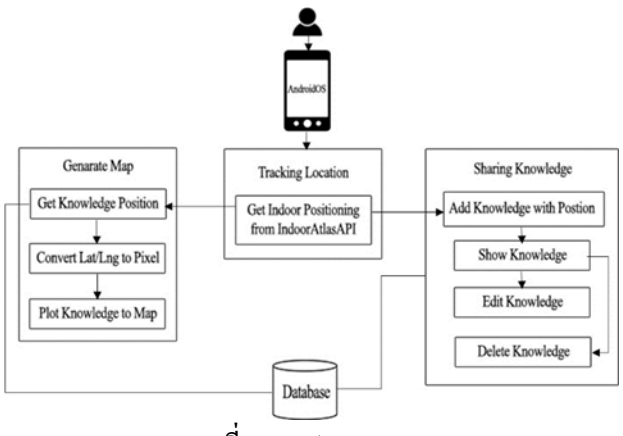
ส่วนระบบที่ใช้สัญญาณ Wi-Fi ต้องใช้งานร่วมกับแผนที่ติดตั้งอุปกรณ์จุดเชื่อมต่อไร้สาย (Access point) ซึ่งอุปกรณ์

พกพาส่วนใหญ่ สามารถรับสัญญาณ Wi-Fi ได้อยู่แล้ว จึงสามารถระบุตำแหน่งพิกัดได้ โดยข้อดีของการใช้ระบบที่เป็นสัญญาณ Wi-Fi อยู่ที่ค่าใช้จ่ายต่ำ แต่ก็ยังมีความแม่นยำไม่สูงมาก โดยเฉลี่ยจะอยู่ในระดับกลาง คือ 6-10 เมตร (กรณีที่มีความแม่นยำสูงจะอยู่ในช่วง 1-5 เมตร และความแม่นยำต่ำอยู่ที่มากกว่า 11 เมตร)

ระบบสุดท้าย เป็นระบบระบุพิกัดโดยใช้คลื่นสนามแม่เหล็กที่เรียกว่า IndoorAtlas เป็นระบบที่ใช้สนามแม่เหล็กภายในตัวอาคารในการระบุตำแหน่งและติดตามตำแหน่งของบุคคลในตัวอาคารหรือวัตถุต่าง ๆ การทำงานจะควบคู่กับอุปกรณ์พกพา ซึ่งข้อมูลจะถูกนำมาใช้ในการติดตามพิกัดตำแหน่งโดยเปรียบเสมือนกับการติดตามโดย GPS ซึ่งแสดงผลแบบตามเวลาจริง (Real time) โดยที่ไม่ต้องพึ่งพาอุปกรณ์เพิ่มเติม นอกจากนี้ระบบดังกล่าว ยังมี Geomagnetic technology ซึ่งให้ความแม่นยำสูงในระดับ 1-2 เมตร และที่สำคัญ ผู้พัฒนาสามารถจัดการข้อมูลและประยุกต์การพัฒนาได้ผ่านชุดของซอฟต์แวร์ (Software Development Kit: SDK) ที่เตรียมไว้ให้ ซึ่งผู้วิจัยได้นำระบบระบุพิกัดนี้มาพัฒนาแอปพลิเคชัน

3.2 สถาปัตยกรรมระบบ

สถาปัตยกรรมของระบบจะแบ่งออกเป็น 2 ส่วนหลัก คือ ส่วนของหน้างาน(การใช้งานผ่านแอปพลิเคชัน) และส่วนของการเก็บองค์ความรู้ ซึ่งทั้ง 2 ส่วนจะทำงานสอดคล้องกันโดยผ่านการระบุพิกัดตำแหน่งภายในอาคาร (IndoorAtlas API) ซึ่งมีการแบ่งออกเป็น 5 ส่วนย่อย ดังแสดงในภาพที่ 3



ภาพที่ 3: สถาปัตยกรรมของระบบ

1. ส่วนติดตามตำแหน่งผู้ใช้ (Tracking Location)

เป็นการดึงข้อมูลตำแหน่งผู้ใช้ผ่าน IndoorAtlas APIs โดยก่อนจะดึงค่านั้น ต้องมีการเดินเก็บสำรวจเพื่อเก็บค่า ละติจูดและลองจิจูดภายในอาคารก่อน เพื่อระบุตำแหน่งผู้ใช้ว่าอยู่ตำแหน่งใดภายในอาคาร

2. สร้างแผนผังภายในอาคาร (Generate Map)

เป็นการสร้างแผนผังภายในอาคาร พร้อมแสดงแผนผังตำแหน่งองค์ความรู้ ระบุพิกัดตำแหน่งของผู้ใช้ว่าอยู่ตำแหน่งใดของแผนผังภายในอาคาร โดยหลังจากที่มีการดึงค่าละติจูดลองจิจูดมาแล้ว จะทำการแปลงพิกัดเป็นพิกเซลเพื่อแสดงผลบนแผนผัง

3. การแบ่งปันองค์ความรู้ (Knowledge Sharing)

เป็นส่วนของการจัดการองค์ความรู้ด้วย 4 ขั้นตอนพื้นฐาน ผู้ใช้สามารถเพิ่มองค์ความรู้โดยระบุพิกัดตำแหน่งที่อยู่ขององค์ความรู้นั้นผ่านแอปพลิเคชัน มีการแสดงองค์ความรู้ต่าง ๆ ที่มีอยู่แล้วบนแผนผังภายในอาคาร ผู้ใช้สามารถแก้ไของค์ความรู้ที่ต้องการ และสามารถลบความองค์ความรู้ที่ไม่ต้องการได้

4. ส่วนการแสดงผลข้อมูล (User Interface)

เป็นส่วนติดต่อกับผู้ใช้ โดยจัดทำในรูปแบบ Android Application ซึ่งพัฒนาโดย Android Studio และ Java ซึ่งเป็นโอเพนซอร์ส ซึ่งปัจจุบันมีผู้ใช้งานระบบ Android ค่อนข้างมาก จึงทำให้ง่ายต่อการเรียนรู้และใช้งาน

5. ส่วนฐานข้อมูล (Database)

เป็นส่วนที่นำข้อมูลที่ต้องการบันทึกทั้งหมดมาเก็บไว้ เพื่อนำมาวิเคราะห์ทั้งจากที่ผู้ใช้งานป้อนเข้ามา หรือส่วนของผลลัพธ์หลังจากการประมวลผลเสร็จเรียบร้อย

3.3 การทดสอบและวัดประสิทธิภาพ

การทำงานของระบบระบุตำแหน่งภายในอาคารโดยคลื่นสนามแม่เหล็ก IndoorAtlas เริ่มจากการวาดแผนผังภายในอาคารแล้วทำการอัพโหลดแผนผังภายในอาคารเข้าสู่ระบบของ IndoorAtlas หลังจากนั้นจึงทำการเก็บค่าสนามแม่เหล็กภายในอาคารโดยใช้แอปพลิเคชัน MapCreator2 ที่มากับ IndoorAtlas โดยเริ่มจากการกำหนดจุดต้นทางและปลายทางเพื่อคำนวณระยะเส้นทางเปรียบเทียบบนแผนผังกับระยะทางจริง จากนั้นทำการเดินตามเส้นทางที่กำหนดเพื่อเก็บค่าสนามแม่เหล็ก ซึ่งเมื่อทำการเก็บค่าตามเส้นทางเสร็จ Generate Map จะทำการสร้าง Heat Map เพื่อให้ผู้พัฒนานำไปใช้งาน ดังภาพที่ 4 จะเป็นภาพ Heat map ของแผนผังภายในอาคารที่เคยทำการเก็บค่าสนามแม่เหล็กไว้แล้ว ซึ่งจะเห็นได้ว่าหากเป็นพื้นที่สีเขียวนั้นแสดงว่ามีการเก็บค่าที่ดี และแม่นยำ สีเหลืองก็

คือปานกลาง และสีแดงคือไม่มีความแม่นยำ และมีโอกาสคลาดเคลื่อนสูง



ภาพที่ 4: ภาพแผนผังที่มีการนำ Heat map มาผสม

ซึ่งจากการเก็บค่าทางตรง การกำหนดจุดต้นทางและปลายทางที่เป็นทางยาวจะได้ค่าที่แม่นยำกว่า ส่วนการเก็บค่าในที่โล่งและทางโค้ง ซึ่งต้องกำหนดจุดเป็นระยะสั้น ๆ เพื่อที่จะให้ได้ค่าสนามแม่เหล็กที่มีความแม่นยำสูงที่สุด

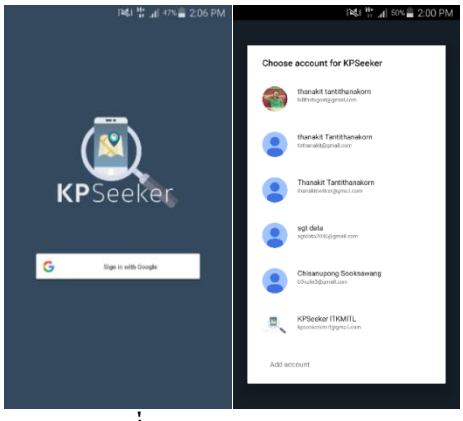


ภาพที่ 5: ภาพแสดงการทดลองพิกัดตำแหน่งจริง

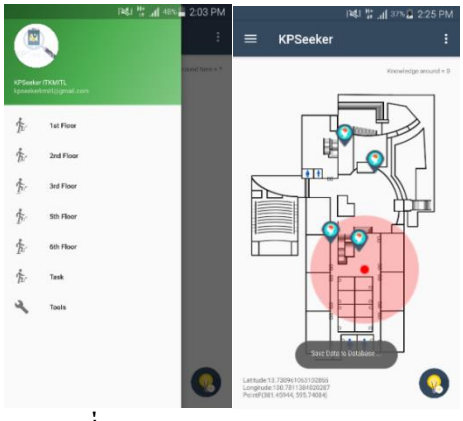
จากการทดลองการระบุพิกัดตำแหน่งภายในอาคารจากสถานที่จริง ดังแสดงในภาพที่ 5 เมื่อผู้ใช้งานยืนอยู่ที่ตำแหน่งหน้าประตู แอปพลิเคชันก็สามารถแสดงพิกัดตำแหน่งที่แม่นยำ และเมื่อมีการเคลื่อนที่ จุดพิกัดก็มีการขยับตาม แต่อย่างไรก็ตาม หากอินเทอร์เน็ตที่สมาร์ทโฟนมีการเชื่อมต่อมีความเร็วต่ำ การตอบสนองของแอปพลิเคชันในแสดงพิกัดตำแหน่งก็มีความล่าช้าไปด้วย

4. ผลการดำเนินงาน

ในส่วนนี้ แสดงถึงแอปพลิเคชันที่ผู้วิจัยได้พัฒนาขึ้น ชื่อว่า KPSeeker ซึ่งผู้ใช้งานจำเป็นต้องมีระบุตัวตนก่อน โดยการระบุผ่าน Google Account ดังภาพที่ 6 หลังจากเข้าสู่ระบบแอปพลิเคชันจะแสดงแผนผังพร้อมกับพิกัดตำแหน่งของผู้ใช้ตำแหน่งขององค์ความรู้ต่าง ๆ เนื่องจากแผนผังภายในอาคารเป็นรูปแบบ 2 มิติ หากมีองค์ความรู้ที่อยู่ในพิกัดเดียวกันในแนวดิ่งจะมีไดอะล็อกให้เลือกว่าผู้ใช้ต้องการที่จะดูองค์ความรู้ใด ที่ถูกแบ่งปันไว้ ในทางด้านมุมซ้ายบน โดยผู้ใช้สามารถกดเมนูเพื่อเลือกชั้น ในส่วนด้านขวาล่างมีปุ่มเพิ่มองค์ความรู้ไว้สำหรับ การเพิ่มองค์ความรู้ใหม่ ดังภาพที่ 7

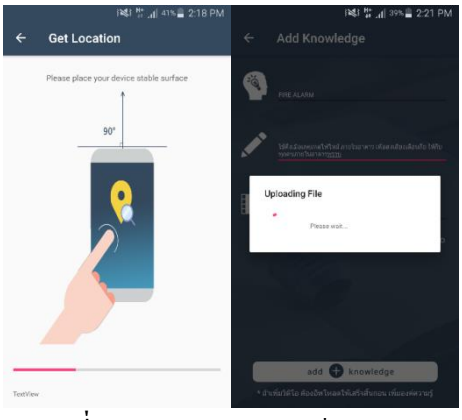


ภาพที่ 6 ส่วนติดต่อการเข้าสู่ระบบ

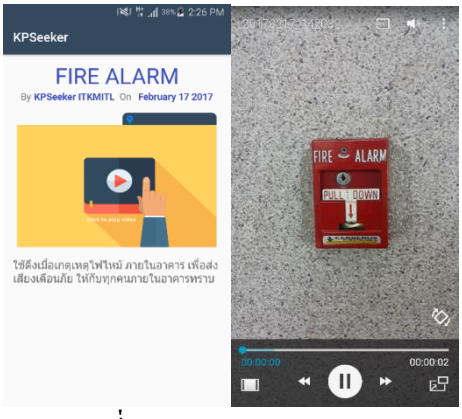


ภาพที่ 7: ส่วนแสดงแผนผังพิกัดของผู้ใช้และตำแหน่งขององค์ความรู้ที่มีอยู่

ในส่วนของภาพที่ 8 และ 9 เป็นแสดงขั้นตอนการเพิ่มองค์ความรู้เข้าสู่ระบบ ซึ่งผู้ใช้สามารถทำได้โดย นำสมาร์ทโฟนวางไว้ในตำแหน่งที่ต้องการเพิ่มองค์ความรู้ ทำการกดปุ่มเพิ่มองค์ความรู้ แอปพลิเคชันจะทำการดึงค่าละติจูดและลองจิจูดมาเก็บไว้ในระบบและทำการส่งค่าไปที่ส่วนให้ผู้ใช้กรอกข้อมูล โดยผู้ใช้สามารถกรอกชื่อองค์ความรู้ รายละเอียดคำอธิบาย และสามารถอัพโหลดรูปภาพหรือวีดีทัศน์ที่เกี่ยวข้องกับองค์ความรู้นั้น ๆ เข้าสู่ระบบฐานข้อมูลได้



ภาพที่ 8: ส่วนติดต่อผู้ใช้ในการเพิ่มองค์ความรู้



ภาพที่ 9: ส่วนแสดงรายละเอียดองค์ความรู้

เพื่อให้การเข้าถึงและแบ่งปันองค์ความรู้กระทำได้สะดวก แอปพลิเคชันได้ออกแบบการใช้งานแบบ Zooming Interface คือ Zoom in และ Zoom out โดย Zoom out คือ การแสดงเพียงหัวข้อในมุมมองกว้าง ๆ หรืออยู่นอกอาณาเขตที่ให้ความสนใจ (Area of interest) แต่ในถ้าผู้ใช้มีการเลือกองค์ความรู้ที่ต้องการเข้าถึง ระบบจะทำการ Zoom in เพื่อแสดงรายละเอียดขององค์ความรู้ที่มีอยู่แบบเต็ม เพื่อให้ง่ายต่อการใช้งานในพื้นที่ที่มีแหล่งขององค์ความรู้จำนวนมาก และช่วยลดความซ้ำซ้อนของพื้นที่ที่แสดงผล

5. บทสรุป

การประยุกต์ใช้ระบบระบุพิกัดตำแหน่งภายในอาคารเพื่อจัดการองค์ความรู้ประจำจุด สามารถตอบสนองความต้องการในการใช้งานได้ แต่อย่างไรก็ตาม การใช้งานเริ่มต้นยังมีความยุ่งยากอยู่บ้าง เนื่องจากผู้ใช้งานต้องทำการสร้างแผนที่คลื่นสนามแม่เหล็กก่อน สำหรับการพัฒนาต่อยอดในอนาคต ผู้วิจัยจะเพิ่มความแม่นยำของระบบระบุพิกัดและการใช้งานที่ง่ายขึ้น เพื่อการแบ่งปันองค์ความรู้ในองค์กรให้มีความน่าใช้ แปลกใหม่และเกิดประโยชน์สูงสุด

เอกสารอ้างอิง

[1] Senion.com, “What is Indoor Positioning System?,” [Online] Available:https://senion.com/indoor-positioning-system/.

[2] ONLINE, “GPS คืออะไร”, [Online] Available: http://www.onelink.co.th/onelink/gps.php.

[3] Komkid, “ไขความลับ Beacons เทคโนโลยี Internet of Things ที่ฮาร์ดแวร์สตาร์ทอัพชอบใช้”, [Online] Available: http://startitup.in.th/beacon-internet-of-thing-hardware-startup.

[4] BECKI WOOD, “การนำเทคโนโลยี Wi - Fi หรือ beacons มาใช้กับการตลาดในการซื้อขายสินค้า”, [Online] Available: http://www.adslthailand.com/post/การนำเทคโนโลยี-Wi - Fi หรือ-beacons-มาใช้กับการตลาดในการซื้อขายสินค้า

[5] Clay Dillow, “การทำแผนที่ภายในอาคาร ด้วยวิธีตรวจติดตามสนามแม่เหล็ก”, [Online] Available: https://arthitonline.me/2012/07/11/indoor-mapping-magnetic-field/

[6] Nonaka I. and Takeuchi H., The Knowledge Creating Company, Oxford University Press, 1995.

[7] Herschel, R.T., Nemati, H. and Steiger, D., “Tacit to Explicit Knowledge Conversion: Knowledge Exchange Protocols”, Journal of Knowledge Management, vol.35, pp.113-125.

[8] Amrit Tiwana, The Knowledge Management Toolkit, Prentice Hall, 2002.

[9] Mark Ginsburg and Ajit Kambil, “A Web-based Knowledge Management Support System for Document Collections”, Hawaii International Conference, 1999

[10] IndoorAtlas, “What is indooratlas?”, [Online] Available: https://www.indooratlas.com/how-it-works/

[11] Davenport T.H., Prusak L., Working Knowledge: How Organizations Manage What They Know, Harvard Business School Press, Boston, 1997.

[12] Guus S., Hans A. and et al., Knowledge Engineering and Management: The CommonKADS Methodology, A Bradford Book, The MIT Press, 1999.

[13] Chaiyo hosting, “Android application”, [Online] Available:http://www.chaiyohosting.com/android-application.

[14] Boonprasert S., Kazuhiko H., “Knowledge Memory System Design on Virtual Space”, Proceeding of ITE Tokai University, vol.5, no.1, pp.13-19, 2012

วิธีการค้นหาร่องรอยที่ไม่ดีประเภทลองเมธอดโดยใช้แบบจำลอง
Model-based Approach for Long Method Bad Smell Detection

ธีรภัทร์ เสถียรพงษ์ (Theraphat Satheanphong)[1] และทรงศักดิ์ รองวิริยะพานิช (Songsak Rongviriyapanich)[2]
สาขาวิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยธรรมศาสตร์
[1]theraphat.satheanphong@gmail.com, [2]rongviri@cs.tu.ac.th

บทคัดย่อ

งานวิจัยนี้นำเสนอแบบจำลองมาตรวัดสำหรับค้นหาร่องรอยที่ไม่ดีประเภทลองเมธอด โดยแบบจำลองนี้จะเป็นแบบจำลองที่รองรับภาษาโปรแกรมเชิงวัตถุ วิธีการคือแปลงซอร์สโค้ดโปรแกรมที่เป็นภาษาเชิงวัตถุให้อยู่ในรูปของแบบจำลอง GAST จากนั้นนำมาเข้าสู่กระบวนการแปลงแบบจำลองด้วยเมตาโมเดล PDG ที่ได้ออกแบบขึ้นหลังจากนั้นแปลงแบบจำลองให้เป็นภาษาจาวาเพื่อหามาตรวัดที่ต้องการ จากนั้นทำการทดสอบโดยการเปรียบเทียบกับกรณีตัวอย่าง พบว่าเมตาโมเดลที่ได้ออกแบบนั้นให้ผลตรงกับผลการทดลองจากกรณีตัวอย่าง

คำสำคัญ: แบบจำลอง มาตรวัด ร่องรอยที่ไม่ดี ลองเมธอด คอนโทรลโฟลว ดาต้าโฟลว

Abstract

This research proposes Model-based approach for detecting long method bad smells. The model supports any object-oriented programming language. The method consists of three steps: 1) Extract source code in to model, 2) Design meta-model and Transform with model from first step, 3) Calculate metrics for detect long method bad smell. We demonstrate our approach by giving an example of identify refactoring with Replace temp with query technic. Our approach proves to be able to detect long method bad smells.

Keyword: Model Transformation, Metrics, Program Dependency Graph, Bad Smell Detection, Control Flow Graph, Data Flow Graph

1. บทนำ

ร่องรอยที่ไม่ดีประเภทลองเมธอด (Long Method Bad Smell) [14] คือลักษณะโครงสร้างเมธอดของโปรแกรมเชิงวัตถุที่มีขนาดมากกว่า 12 บรรทัด [16] หรือมีความซับซ้อนสูง ยากที่จะเข้าใจ สามารถเป็นตัวชี้วัดได้ว่าโปรแกรมหรือการออกแบบนั้นมีปัญหาและอาจสร้างความลำบากในการบำรุงรักษาได้

การค้นหาร่องรอยที่ไม่ดีในซอร์สโค้ดโปรแกรมเชิงวัตถุ (Bad Smell Detection) นั้นมีวิธีการหนึ่งคือการใช้มาตรวัดซอฟต์แวร์ (Software Metrics) มาตรวัดซอฟต์แวร์จะใช้ในการคำนวณคุณสมบัติต่างๆเพื่อค้นหาร่องรอยที่ไม่ดีประเภทลองเมธอดที่มีอยู่ในซอร์สโค้ด

ทุกวันนี้ได้มีภาษาโปรแกรมเชิงวัตถุ (Object Oriented Programming) เกิดขึ้นมากมาย เช่น Java, C/C++ เป็นต้น ซึ่งนักวิจัยต่างก็มุ่งเน้นไปที่การค้นหาร่องรอยที่ไม่ดีของซอร์สโค้ดโปรแกรมเชิงวัตถุไปที่ภาษาใดภาษาหนึ่ง ทำให้เกิดข้อจำกัดในการพัฒนาต่อยอดเป็นเครื่องมือได้ โดยที่ในปัจจุบันนั้นได้มีเทคนิคการพัฒนาซอฟต์แวร์แบบเอ็มดีเอซึ่งเป็นเทคนิคการพัฒนาซอฟต์แวร์ที่ไม่อิงต่อภาษาโปรแกรมใดๆ

เอ็มดีเอ (Model Driven Architecture : MDA) [15] เป็นการพัฒนาซอฟต์แวร์ด้วยการใช้แบบจำลอง (Model) ซึ่งแบบจำลองนี้จะเป็นแบบจำลองที่อิสระจากแพลตฟอร์ม (Platform Independent Model) ไว้ใช้อธิบายถึงการทำงานของระบบงานซึ่งยังไม่มีการกำหนดแพลตฟอร์ม โดยทั่วไปแล้วแบบจำลองลักษณะนี้จะถูกอธิบายโดยใช้ภาษา UML (Unified Modeling Language)

จากปัญหาดังกล่าว งานวิจัยนี้จึงนำเสนอวิธีการค้นหาร่องรอยที่ไม่ดีประเภทลองเมธอดด้วยการใช้แบบจำลอง โดยแบบจำลองเป็นแบบจำลองของภาษาโปรแกรมเชิงวัตถุและเป็นอิสระจากแพลตฟอร์ม (Platform Independent) ซึ่งแบบจำลองนี้สามารถใช้ในการคำนวณมาตรวัดซอฟต์แวร์เพื่อค้นหาร่องรอยที่ไม่ดีประเภทลองเมธอด

เนื้อหาในบทความนี้ประกอบด้วยหัวข้อที่ 2 กล่าวถึงงานวิจัยที่เกี่ยวข้อง ในหัวข้อที่ 3 จะอธิบายถึงภาพรวมของกระบวนการสร้างและค้นหาร่องรอยที่ไม่ดีประเภทลองเมธอดโดยใช้แบบจำลอง หัวข้อที่ 4 จะเป็นการทดลองโดยใช้กรณีศึกษาและหัวข้อที่ 5 จะเป็นการสรุปผล

2. งานวิจัยที่เกี่ยวข้อง

ในหัวข้อนี้จะกล่าวถึงงานวิจัยที่เกี่ยวข้องซึ่งสามารถแบ่งได้เป็น 2 หัวข้อ คือ สำรวจและรวบรวมมาตรวัดสำหรับร่องรอยที่ไม่ดีประเภทลองเมธอด และสำรวจแนวทางการแปลงซอร์สโค้ดโปรแกรมเชิงวัตถุให้อยู่ในรูปของแบบจำลอง

2.1 สำรวจและรวบรวมมาตรวัดร่องรอยที่ไม่ดีประเภทลองเมธอด

Fontana และ Braione [4] ได้สำรวจเครื่องมือและรวบรวมมาตรวัดสำหรับการค้นหาร่องรอยที่ไม่ดีประเภทต่างๆ สำหรับร่องรอยที่ไม่ดีประเภทลองเมธอดที่ได้รวบมานั้น ประกอบด้วย Number Line Of Code (NLOC) , Cyclomatic Complexity (VG) และ มาตรวัด Halstead Complexity โดยเครื่องมือที่ใช้ทดลองคือ JDeodorant , PMD และ Checkstyle ซึ่งผลการทดลองเห็นได้ชัดว่า JDeodorant ให้ผลดีที่สุด

Charalampidou และ Ampatzoglou [2] ได้สำรวจและรวบรวมมาตรวัดที่ใช้ทำนายการเกิดร่องรอยที่ไม่ดีประเภทลองเมธอดซึ่งส่วนใหญ่จะเกี่ยวข้องกับขนาด (Size) และกลุ่มก้อน (Cohesion) ซึ่งรวบรวมได้ทั้งหมด 13 มาตรวัดได้แก่ Uncommend Line Of Code (LOC), Lack Of CohesionMetric1,2,3,4,5(LCOM1),(LCOM2),(LCOM3), (LCOM4),(LCOM5), Cohesion Metric (Coh), Tight Class Cohesion (TCC),Loose Class Cohesion (LCC), Degree Of Cohesion Direct (DC_D) ,Degree Of Cohesion Indirect (DC_I), Class Cohesion (CC), Class Cohesion Metric (SCOM) และ Low-Level Design Similarity Based Class Cohesion (LSCC) จากนั้นได้ทำการสร้างกรณีศึกษาขึ้นมาพบว่ามาตรวัดที่ให้ค่าที่โดดเด่นที่สุด ได้แก่ LCOM1, LCOM2 และ COH โดยมาตรวัดที่ใช้งานง่ายที่สุดจะเป็น LCOM1 และ LCOM2

Sukhdeep และ Raman [5] ได้ทำการสำรวจและรวบรวมมาตรวัดสำหรับค้นหาร่องรอยที่ไม่ดีประเภทต่างๆโดยมาตรวัดสำหรับลองเมธอดที่เลือกคือ Number Of Line Of Code (NLOC), Cyclomatic Complexity และมาตรวัด Halstead

Panita, Songsakdi และ Taweesup [1] ได้เสนอมาตรวัดสำหรับรีแฟคทอริงร่องรอยที่ไม่ดีประเภทลองเมธอด ซึ่งมาตรวัดประกอบไปด้วย Cyclomatic Complexity (MCX), Line Of Code In Method (LOC) และ Lack Of Cohesion Of Method (LCOM)

Bryton , Fernando และ Miguel [6] ได้นำเสนอวิธีการค้นหาร่องรอยที่ไม่ดีประเภทลองเมธอด โดยใช้โมเดลที่ชื่อว่า Binary Logistic Regression (BLR) ซึ่งเป็นโมเดลที่ใช้กลุ่มมาตรวัดในการคำนวณหาร่องรอยที่ไม่ดีประเภทลองเมธอด โดยกลุ่มมาตรวัดประกอบด้วย Method Line Of Code (MLOC),Nested Block Depth(NBD), Cyclomatic Complexity และ Number Of Parameter (PAR) จากการทดลองโดยใช้กรณีศึกษาพบว่าโมเดล BLR นั้นสามารถค้นหาร่องรอยที่ไม่ดีประเภทลองเมธอดได้

จากการสำรวจมาตรวัดที่ใช้สำหรับค้นหาร่องรอยที่ไม่ดีประเภทลองเมธอดจากงานวิจัยที่ผ่านมา แสดงดังตารางที่ 1

ตารางที่ 1: สรุปมาตรวัดสำหรับร่องรอยที่ไม่ดีประเภทลองเมธอดจากงานวิจัยที่ผ่านมา

มาตรวัด	งานวิจัย
NLOC	[4],[5]
Cyclomatic Complexity	[4],[5],[1], [6]
Halstead Complexity	[4],[5]
Coh	[2]
LOC	[1][6]
LCOM	[1]
LSCC	[2]
NBD	[6]
PAR	[6]
TCC	[2]
LCOM1,LCOM2,LCOM3,LCOM4,LCOM5	[2]
LCC	[2]

มาตรวัด	งานวิจัย
DC_D	[2]
DC_I	[2]
CC	[2]
SCOM	[2]

2.2 การแปลงซอร์สโค้ดโปรแกรมเชิงวัตถุให้อยู่ในรูปของแบบจำลอง

Yania และคณะ [7] ได้เสนอวิธีการใช้มาตรวัดซอฟต์แวร์โปรแกรมเชิงวัตถุที่เป็นอิสระจากแพลตฟอร์ม ซึ่งใช้กระบวนการพัฒนาซอฟต์แวร์แบบเอ็มดีเอ โดยได้ขยายเมตาโมเดลชื่อ Visitor ซึ่งเป็นเมตาโมเดลที่ใช้อธิบายภาษาโปรแกรมเชิงวัตถุจากนั้นได้ทดลองกับกรณีศึกษา ผลการทดลองพบว่าเมตาโมเดลที่ถูกขยายนั้นสามารถรองรับการคำนวณโดยใช้มาตรวัดได้

Tae Woong Kim และคณะ [8] ได้เสนอวิธีการค้นหาร่องรอยที่ไม่ดีโดยใช้ภาษา OCL วิธีการคือแปลงซอร์สโค้ดภาษาจาวาให้อยู่ในรูปแบบภาษา XMI โดยมีเมตาโมเดล JavaEAST ใช้สำหรับอธิบาย XMI Element หลังจากนั้นใช้ OCL ค้นหาร้องรอยที่ไม่ดี จากนั้นได้ทดลองกับร่องรอยที่ไม่ดีประเภท Refused Bequest พบว่าสามารถค้นหาได้อย่างถูกต้อง

งานวิจัยนี้ได้เสนอวิธีการค้นหาร่องรอยที่ไม่ดีประเภทลองเมธอดของภาษาโปรแกรมเชิงวัตถุโดยใช้แบบจำลอง โดยวิธีการจะแบ่งเป็น การแปลงซอร์สโค้ดให้อยู่ในรูปของแบบจำลอง GAST . การออกแบบเมตาโมเดลและการสร้างกฎการแปลงและการค้นหาร่องรอยที่ไม่ดีประเภทลองเมธอดโดยใช้มาตรวัด

3. ระเบียบวิธีวิจัย

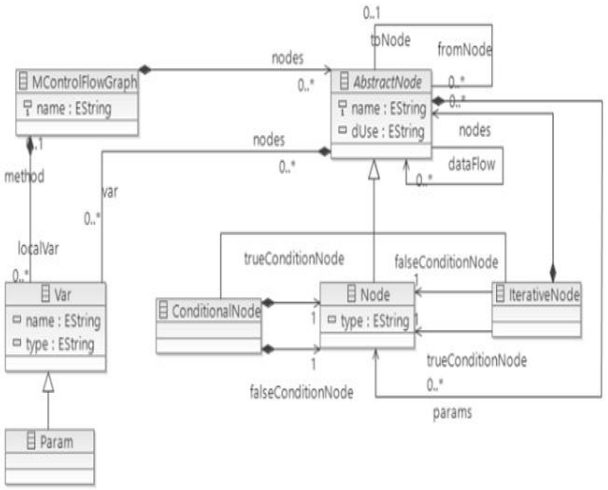
งานวิจัยนี้ได้นำเทคนิคการพัฒนาซอฟต์แวร์โดยใช้แบบจำลอง(MDA)โดยมีขั้นตอนที่สำคัญคือการแปลงแบบจำลอง (Model Transformation) ซึ่งต้องอาศัยแบบจำลองต้นทางและแบบจำลองปลายทาง หลังจากนั้นจะเป็นการคำนวณหาค่ามาตรวัดจากแบบจำลองปลายทางเพื่อนำไปใช้สำหรับค้นหาร่องรอยที่ไม่ดีประเภทลองเมธอด

3.1 แบบจำลองสำหรับโปรแกรมภาษาเชิงวัตถุ (GAST)

เมตาโมเดล GAST (General Abstract Syntax Tree Meta-model)[17] ใช้สำหรับอธิบายโครงสร้างของภาษาโปรแกรมเชิงวัตถุ ที่มีอยู่ในปัจจุบัน (JAVA,C,C++) ประกอบไปด้วยเอนทิตี้ต่างๆ เช่น Class, Packages , Type, Function, Statement ,Variables เป็นต้นซึ่ง การแปลงซอร์สโค้ดของโปรแกรมภาษาเชิงวัตถุให้อยู่ในรูปแบบจำลอง GAST นั้นสามารถทำด้วยเครื่องมือ SISSy [9] ซึ่งเป็นส่วนเสริมของโปรแกรมอีคลิป (Eclipse Plugin)โดยแบบจำลอง GAST จะเป็นแบบจำลองต้นทางเพื่อใช้สำหรับกระบวนการแปลงแบบจำลองของงานวิจัยนี้

3.2 เมตาโมเดล Program Dependency Graph (PDG)

เมตาโมเดล PDG คือเมตาโมเดลที่ใช้อธิบายลักษณะโครงสร้างการทำงานของเมธอดรวมไปถึงลำดับการไหล (Control-flow) ของ Statement และลำดับการไหลของข้อมูล (Data-flow) ซึ่งจะไหลจาก Statement หนึ่งไปสู่อีก Statement โดย PDG นั้นได้ออกแบบตามแนวคิดของงานวิจัย [12] ซึ่งได้ใช้ Control-Flow Graph และ Data-Flow Graph วิเคราะห์โครงสร้างของเมธอด



ภาพที่ 2: เมตาโมเดล PDG ที่สร้างด้วย Ecore

ผู้วิจัยได้ใช้ Ecore ซึ่งเป็นส่วนหนึ่งของ Eclipse Modeling Framework (EMF) [10] สร้างเมตาโมเดล PDG ดังภาพที่ 2 เอนทิตี้ MControlflowGraph ใช้สำหรับแทนเมธอดที่อยู่ในคลาสซึ่งในเมธอดจะประกอบด้วย Statement ต่างๆที่เรียงลำดับการไหล (Control-flow) ตั้งแต่ Statement แรกไปจนถึง Statement สุดท้ายโดย Statement นั้นจะแทนด้วย

AbstractNode ซึ่งมี fromNode ใช้สำหรับบอกลำดับการไหลของ Node ก่อนหน้าและ toNode ใช้สำหรับบอก Node ลำดับถัดไปสำหรับลำดับการไหลของข้อมูล จะใช้ dataFlow เพื่อบอก Node ถัดไปของข้อมูล ,เอนทิตี้ Node นั้นใช้สำหรับแทน Statement ทั่วไป เช่น double totalPrice = 3.5; เป็นต้น, เอนทิตี้ ConditionNode ใช้สำหรับแทน Statement ที่เกี่ยวข้องกับเงื่อนไข เช่น if , else เป็นต้น เอนทิตี้ IterativeNode นั้นจะใช้สำหรับแทน Statement ที่เกี่ยวข้องกับ Loop , เอนทิตี้ Param แทนพารามิเตอร์ของเมธอด

3.3 สร้างกฎการแปลงแบบจำลองต้นทางไปปลายทาง

การสร้างกฎการแปลงนั้นจะต้องอาศัยการจับคู่ของแบบจำลองต้นทางกับแบบจำลองปลายทาง เนื่องจากข้อจำกัดในเนื้อที่ของบทความที่จะนำเสนอ ผู้วิจัยจึงขอยกตัวอย่างการจับคู่แบบจำลองการแปลง SimpleStatement ของเมตาโมเดล GAST ไปสู่ Node ของเมตาโมเดล PDG ดังตารางที่ 2

ตารางที่ 2: ตัวอย่างการจับคู่แบบจำลอง

GAST	PDG	Rule
SimpleStatement	Node	SimpleStatement2Node

การสร้างกฎการแปลงนั้นงานวิจัยนี้ได้เลือกใช้ภาษา ATL [11] ซึ่งเป็นภาษาสำหรับแปลงแบบจำลอง โดยกฎการแปลงแปบจำลองที่ได้จากการจับคู่นั้นแสดงดังภาพที่ 3

```
1  rule SimpleStatement extends Statement2Node {
2     from
3        s : GAST!SimpleStatement
4     to
5        t : PDG!Node (
6           toNode <- if s.oclIsUndefined() then
7                        s.reflmmediateComposite()
8                        .reflmmediateComposite()
9                     else
10                       s.next
11                    endif
12       )
13 }
```

ภาพที่ 3: ตัวอย่างกฎ SimpleStatement2Node

จากภาพที่ 3 คือการแปลง SimpleStatement ของเมตาโมเดล GAST ไปสู่เอนทิตี้ Node ของเมตาโมเดล PDG โดยมีแอตทริบิวท์ที่สำคัญคือ toNode ซึ่งเป็นการทำ Control-Flow ของ Statement บรรทัดที่ 6 “s.OclIsUndefined()” เป็นการตรวจสอบว่า Statement ปัจจุบันนั้นได้ถูกกำหนดไว้หรือไม่ ถ้าไม่ถูกกำหนดไว้ แสดงว่า Statement ปัจจุบันนั้นคือ Statement แรกซึ่งก็คือบรรทัดที่ 7 แต่ถ้าถูกกำหนดไว้จะเป็นการเรียก Statement ต่อไปแทนในบรรทัดที่ 9 ซึ่ง “s.next” นั้นเป็นฟังก์ชั่น Helper ของภาษา ATL ซึ่ง Helper นี่ถูกสร้างขึ้นมาสำหรับสืบทอด Control-Flow ของ Statement ถัดไปดังภาพที่ 4

```
1  helper context GAST!BlockStatement def: getMap() :
2  Map(GAST!Statement, GAST!Statement) =
3  self.statements->iterate(e; acc: TupleType(prev: GAST!Statement,map:
4  Map(GAST!Statement, GAST!Statement)) = Tuple{ prev = Sequence {}->
5  first(), map = Map{}} | Tuple{prev = e, map =
6  if acc.prev.oclIsUndefined() then
7           acc.map
8              else
9           acc.map.including(acc.prev, e)
10             endif
11    }).map;
12 helper def: next : Map(GAST!Statement, GAST!Statement) =
13    GAST!BlockStatement.allInstances()->asSequence()->
14    iterate(e; acc : Map(GAST!Statement, GAST!Statement) = Map {} |
15          acc.union(e.getMap())
16    );
17 helper context GAST!Statement def: next :
18 GAST!Statement = thisModule.next.get(self);
```

ภาพที่ 4: Helper สำหรับ Control-Flow

จากภาพที่ 4 next() คือ Statement ปัจจุบันที่จะหา Statement ต่อไปโดยจะเรียกใช้ next() ที่มีชนิดการคืนค่าเป็น Map ที่รับ Parameter เป็น Statement โดยจะคืนค่า Statement ลำดับต่อไปจนกระทั้งถึง Statement สุดท้าย

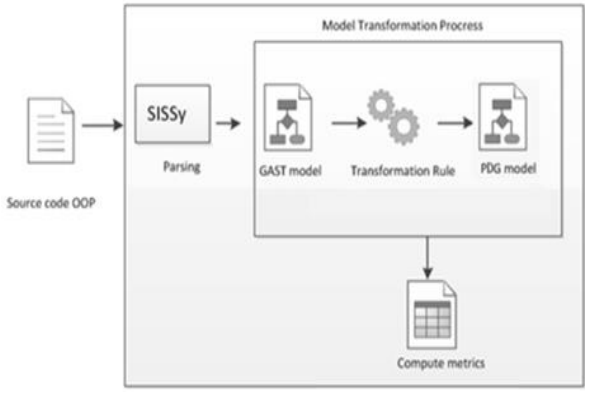
3.4 การคำนวณมาตรวัดเพื่อหาร่องรอยที่ไม่ดี

การคำนวณมาตรวัดสามารถทำได้หลังจากแปลงแบบจำลอง ซึ่งมาตรวัดที่แบบจำลองสามารถคำนวณค่าได้นั้นจะต้องเป็นมาตรวัดในระดับเมธอด ซึ่งจากการสำรวจมาตรวัดจากงานวิจัยที่ผ่านมา พบว่ามาตรวัดระดับเมธอดบางชนิดนั้นไม่สามารถคำนวณได้จากเมตาโมเดล PDG เช่น กลุ่มของมาตรวัด Halstead Complexity ซึ่งจำเป็นจะต้องใช้ Operator มาช่วยคำนวณ เป็นต้น ซึ่งเมตาโมเดล PDG ไม่ได้รองรับในส่วนนี้ ดังนั้นมาตรวัดที่เมตาโมเดล PDG คำนวณได้ ดังตารางที่ 3

ตารางที่ 3: สรุปมาตรวัดงานวิจัย

มาตรวัด	วิธีการ
VG	Edge - node+2
LOC	Number Of Statement In Method
PAR	Number Of Parameter
Coh	Coh=1-(1-1/n)LCOM5 ,n : Number Line Of Code
LCOM	(m a/v)/(m-1) ,m : the number of methods in the class, a : the number of methods in class that access an instance
LCOM1	LCOM1 = P, where P is the number of pairs of lines that do not share variables.
LCOM2	LCOM2 = P – Q, if P − Q ≥ 0 / otherwise LCOM2 = 0,where P is the number of pairs of lines that do not share variables, and Q is the number of pairs of lines that share variables.
LCOM3	Number of connected components in a graph, where each node represents a line of code and each edge the common use of at least one variable
LCOM4	Similar to LCOM3. Method calls are treated as edges
LCOM5	LCOM5 = (a - nl) / (l - nl) where n is the number of lines, a is the number of variables used in a line, and l is the total number of variables.
DC_D	DCD =|ED| / [n * (n – 1) / 2] where ED the number of edges in a graph connecting directly related lines of code
DC_I	DCI =|EI| / [n * (n – 1) / 2] where EI the number of edges in a graph connecting indirectly related lines of code
LCC	LCC = (NDC + NIC) / NP where NDC and NP as defined above, and NIC the number of indirectly connected pairs of lines
TCC	TCC = NDC / NP where NDC the number of directly connected pairs of lines

จากระเบียบขั้นตอนของงานวิจัยนี้เริ่มจากแปลงซอร์สโค้ดภาษาโปรแกรมเชิงวัตถุใดใดให้อยู่ในรูปของแบบจำลอง GAST หลังจากนั้นแปลงไปสู่แบบจำลอง PDG โดยใช้กฎการแปลงที่ได้เสนอไว้ จากแบบจำลอง PDG สามารถนำไปใช้คำนวณมาตรวัดหาร่องรอยที่ไม่ดีประเภทลองเมธอดจากมาตรวัดที่ได้ทำการสำรวจได้ ซึ่งภาพรวมของงานวิจัยนี้สามารถอธิบายดังภาพที่ 4



ภาพที่ 4: ภาพรวมของงานวิจัย

4. การทดลอง

เพื่อที่จะทดสอบว่าแบบจำลองที่สร้างขึ้นนั้นสามารถค้นหาร่องรอยที่ไม่ดีประเภทลองเมธอดได้ ผู้วิจัยจะทดสอบกับผลการทดลองของงานวิจัย [12] ซึ่งได้เสนอเทคนิค Refactoring Filter Condition (RFC) เพื่อช่วยระบุวิธีการรีแฟคทอริงร่องรอยที่ไม่ดีประเภทลองเมธอด ซึ่งประกอบด้วย 3 รีแฟคทอริงได้แก้ 1) Replace Temp With Query (RTWQ) 2) Preserve Whole Object (PWO) 3) Introduce Parameter Object (IPO) ซึ่ งงานวิจัยดังกล่าวนั้นผู้วิจัยได้เลือกเทคนิคการระบุรีแฟคทอริงประเภท RTWQ ซึ่งมีวิธีการคือ

Rule 1: $\exists$ v1 $\in DU_{mc1}$· size($PATH_{v1}$) > 0 |

$\forall path \in PATH_{v1}$| $(du(v1)_{path} = 1) \wedge (cu(v1)_{path} > 0 \vee pu(v1)_{path} > 0)$

โดยที่

$du(v)_{path}$ คือ assignment ของ v ในเส้นทาง

$cu(v)_{path}$ คือ computation ของ v ในเส้นทาง

$pu(v)_{path}$ คือ predication ของ v ในเส้นทาง

นักวิจัยดังกล่าวได้นำตัวอย่างซอร์สโค้ดจาก [14]มาทดสอบโดยได้เลือกเมธอด getPrice() มาแปลงให้อยู่ในรูปของ Data Flow Graph (DFG) เพื่อหาเส้นทางการไหลของตัวแปร basePrice และ discountFactor ซึ่งได้ผลตามตารางที่ 4

ตารางที่ 4: ผลการทดลอง RTWQ ของงานวิจัย [12]

V1	$PATH_{v1}$	du(v1)	cu(v1)	pu(v1)
basePrice	Line 1,3,7	1	1	1
	Line 1,7	1	1	0
discountFactor	Line 2,4,7	1	1	0
	Line 2,6,7	1	1	0

จากตารางที่ 4 เมื่อนำเส้นทางของตัวแปรเข้าสู่ Rule 1 พบว่าตัวแปรทั้งคู่ให้ค่าตรงกับสมการ ดังกล่าว ดังนั้น เมธอด getPrice() ควรถูกรีแฟคทอริงด้วย Replace Temp With Query

จากการทดลองของงานวิจัย [12] ผู้วิจัยจึงได้ประยุกต์ใช้กับงานวิจัยนี้โดยได้นำตัวอย่างเมธอด getPrice() จาก[14]ผ่านกระบวนการแปลงแบบจำลอง จากนั้น แปลงแบบจำลองให้เป็นภาษาจาวาโดยใช้ EMF Generator Model [10] เพื่อค้นหาหาเส้นทางของตัวแปร basePrice และ discountFactor ตามเทคนิคของงานวิจัยดังกล่าว ซึ่งผลลัพธ์ของการค้นหาเส้นทางของตัวแปรทั้งคู่ นั้นแสดงดังตารางที่ 5

ตารางที่ 5: ผลการค้นหา Node จากเมธอด getPrice()

Var	Node (du)	ConditionNode (pu)	Node (cu)
basePrice	1	1	1
	1	1	0
discountFactor	1	1	0
	1	1	0

จากตารางที่ 5 เมื่อนำผลของการหาเส้นทางการของตัวแปร basePrice และ discountFactor ไปคำนวณกับ Rule 1 ของงานวิจัย [12] พบว่าค่าที่ได้นั้นจะเท่ากันกับผลการทดลองของงานวิจัยดังกล่าว แสดงให้เห็นว่าเมตาโมเดล PDG สามารถระบุรีแฟคทอริงโดยใช้ Replace Temp With Query ได้ นอกจากนี้ผู้วิจัยได้ทำการคำนวณเพิ่มเติมกับมาตรวัดที่ได้สรุปไว้ในตารางที่ 3 เนื่องจากข้อจำกัดในเนื้อที่ของบทความที่จะนำเสนอ ผู้วิจัยจึงขอยกตัวอย่างเฉพาะมาตรวัดที่เกี่ยวข้องกับเมธอดดังนั้นมาตรวัดที่เกี่ยวกับ Cohesion จึงไม่สามารถหาได้จากตรงนี้ ซึ่งจากผลการทดลองพบว่าเมตาโมเดล PDG นั้นสามารถคำนวณหามาตรวัดดังกล่าวได้ซึ่งสรุปผลไว้ดังตารางที่ 6

ตารางที่ 6: สรุปมาตรวัดที่เมตาโมเดล PDG สามารถหาได้

มาตรวัด	ผลลัพธ์
VG	3
LOC	6
PAR	0

5. สรุป

งานวิจัยนี้เสนอวิธีการค้นหาร่องรอยที่ไม่ดีประเภทลองเมธอดโดยใช้แบบจำลองซึ่งสามารถรองรับกับภาษาโปรแกรมเชิงวัตถุ โดยการสร้างเมตาโมเดลสำหรับหามาตรวัดที่ใช้สำหรับคำนวณหาร่องรอยที่ไม่ดีประเภทลองเมธอดหลังจากนั้นนำเมตาโมเดลมาทดลองกับกรณีตัวอย่างพบว่าเมตาโมเดลที่ได้ทำการสร้างขึ้นมานั้นให้ผลการทดลองเป็นที่น่าพอใจ โดยในอนาคตนั้นงานวิจัยนี้สามารถพัฒนาต่อยอดไปเป็นเครื่องมือสำหรับวิเคราะห์โปรแกรม หรือหากมีผู้สนใจ สามารถนำไปประยุกต์กับการหาร่องรอยที่ไม่ดีประเภทอื่นๆได้

เอกสารอ้างอิง

[1] Meananeatra, P., S. Rongviriyapanish, and T. Apiwattanapong. *Using software metrics to select refactoring for long method bad smell*. in *The 8th Electrical Engineering/ Electronics, Computer, Telecommunications and Information Technology (ECTI) Association of Thailand - Conference 2011*. 2011.

[2] Charalampidou, S., A. Ampatzoglou, and P. Avgeriou, *Size and cohesion metrics as indicators of the long method bad smell: An empirical study*, in *Proceedings of the 11th International Conference on Predictive Models and Data Analytics in Software Engineering*. 2015, ACM: Beijing, China. p. 1-10.

[3] Shyam R. Chidamber, C.F.K., *A Metrics suite for object oriented design.* 1993.

[4] Fontana, F.A., P. Braione, and M. Zanoni, *Automatic detection of bad smells in code: An experimental assessment.* Journal of Object Technology, 2012. **11**(2): p. 5: 1-38.

[5] Kaur, S. and D.R.M. , *Analysis of Various Software Metrics Used To Detect Bad Smells.* 2016.

[6] Bryton, S., F.B.e. Abreu, and M. Monteiro. *Reducing Subjectivity in Code Smells Detection: Experimenting with the Long Method.* in *2010 Seventh International Conference on the Quality of Information and Communications Technology*. 2010.

[7] Crespo, Y., et al., *Language Independent Metric Supporttowards Refactoring Inference.* 2005.

[8] Kim, T.-W., T.-G.K. , and a.J.-H. Seu, *Specification and Automated Detection of Code Smells using OCL.* 2013.

[9] "SISSy – Structural Investigation of Software Systems." [Online]. Avialable : http://www.sqools.org/sissy/

[10] The Eclipse Foundation, Eclipse Modelling Project. Eclipse IDE website. 2011. Available at: http://www.eclipse.org/modeling/. Accessed at August 13, 2011.

[11] F. Jouault, I. Kurtev, "Transforming Models with ATL" LNCS 3844, pp. 128-138, 2006.

[12] Meananeatra, P., S. Rongviriyapanish, and T. Apiwattanapong. *Identifying Refactoring Through Formal Model Based on Data Flow Graph*. Software Engineering (MySEC),*2011 5th Malaysian Conference in. IEEE, 2011.*

[14] Martin Fowler.Refactoring: Improving the Design of Existing code. Addison-Wesley, 1999.

[15] J.Miller and J. Mukerji, “MDA Guild Version 1.0” OMG, 2003.

[16] Cunha, Jácome, et al. "Towards a catalog of spreadsheet smells." *Computational Science and Its Applications–ICCSA 2012* (2012): 202-216.

[17] Trifu, Mircea, and Peter Szulman. "Language independent abstract metamodel for quality analysis and improvement of oo systems." *Proceedings of the 7th German Workshop on Software-Reengineering (WSR 2005), Bad Honnef, Germany*. Vol. 25. 2005.

วงจรการบริหารจัดการระบบสารสนเทศด้วยสื่อสังคมออนไลน์
Information System Life Cycle Management using Social Network

ชาติสยาม พิลาม (Chatsayam Philam)[1] นวพร วิสิฐพงศ์พันธ์ (Nawaporn Wisitpongphan)[2]
และนลินภัสร์ ปรวัฒน์ปรียกร (Nalinpat Porrawatpreyakorn)[3]
[1,3]ภาควิชาเทคโนโลยีสารสนเทศ [2]ภาควิชาการสื่อสารข้อมูลและเครือข่าย คณะเทคโนโลยีสารสนเทศ
ศูนย์วิจัยเฉพาะทางเทคโนโลยีสารสนเทศและการสื่อสาร มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
[1]s5707011859530@email.kmutnb.ac.th, [2]nawaporn.w@it.kmutnb.ac.th, [3]nalinpat.p@it.kmutnb.ac.th

บทคัดย่อ

ปัจจุบันระบบสารสนเทศที่พัฒนาตามกระบวนการพัฒนาระบบทั่วไปของภาครัฐหลายระบบไม่ถูกนำไปใช้งานให้เกิดประโยชน์อย่างคุ้มค่า เนื่องจากกระบวนการพัฒนาระบบที่ใช้ส่วนใหญ่นั้นล้วนมุ่งเน้นการพัฒนาเพื่อให้สำเร็จตามความต้องการของผู้ใช้งานเพียงอย่างเดียว แต่มิได้คำนึงถึงการดูแลระบบแบบต่อเนื่องจนกระทั่งระบบถูกใช้งานอย่างแท้จริง ดังนั้น งานวิจัยนี้จึงได้นำเสนอการนำวงจรการบริหารจัดการระบบสารสนเทศ ที่พัฒนามาจากแนวปฏิบัติที่ดีจากงานวิจัยของ Budde [3], Marchetta [4], และ Conello[5] มาใช้ร่วมกับการพัฒนาระบบบริหารจัดการสินทรัพย์ของกรมชลประทาน

ผลจากการพัฒนาตามแนวทางที่นำเสนอแสดงให้เห็นว่ากระบวนการบริหารจัดการระบบหลังการส่งมอบนั้นเป็นสิ่งที่สำคัญอย่างยิ่ง และพบว่าการสร้างปฏิสัมพันธ์ระหว่างทีมผู้พัฒนาและผู้ใช้งานโดยตรงและรวดเร็วผ่านแอพลิเคชั่นไลน์เป็นปัจจัยหลักอันก่อให้เกิดการมีส่วนร่วมและความกระตือรือร้นในการช่วยปรับปรุงระบบให้สามารถใช้งานให้เกิดประโยชน์กับองค์กรได้อย่างแท้จริง

คำสำคัญ: วงจรการบริหารจัดการระบบ การส่งมอบระบบ การมีส่วนร่วม สังคมออนไลน์

Abstract

Nowadays, there are many information systems in public sectors that have never been used despite the fact that they were developed according to the users' requirements and follow a standard software development process. In this paper, we applied best software development practice proposed by Budde [3], Marchetta [4], and Conello [5] on developing an asset management system for the Royal Irrigation Department.

Results obtained from using our approach suggested that information system management is very important even after the product delivery phase. Furthermore, prompt service and direct interaction between development team and users provided by, in this case, Line application is the key factor which encourages users' participation and build enthusiasm which resulted in collaborative product improvement and benefit to the organization.

Keyword: Information System Life Cycle Management, Product Delivery, Users' Participation, Social Network.

1. บทนำ

ในปัจจุบันการพัฒนาระบบสารสนเทศ โดยทั่วไปจะมุ่งเน้นการพัฒนาระบบให้สำเร็จและตรงกับความต้องการของผู้ใช้ในขณะนั้น แต่ผู้ใช้งานระบบสารสนเทศภายในองค์กรมักจะเกิดความต้องการใหม่ ๆ เพิ่มขึ้นภายหลัง จากการใช้งานระบบสารสนเทศจริง ในช่วงระยะเวลาหนึ่ง ซึ่งพอที่จะรู้ว่าระบบสารสนเทศนั้นไม่สามารถตอบสนองความต้องการ ที่เพิ่มมากขึ้น เป็นเหตุผลที่ทำให้องค์กรล้มเหลวในการนำระบบสารสนเทศเข้าไปใช้งาน [1]

หากเราพิจารณาให้ดี ปัญหาที่ทำให้องค์กรล้มเหลวเมื่อนำระบบสารสนเทศเข้าไปใช้งานนั้น ส่วนใหญ่

ไม่ได้เกิดจากขั้นตอนในการพัฒนา แต่ปัญหากับเกิดหลังจากระบบมีการส่งมอบแล้ว [2] หากหลังการส่งมอบระบบ มีขั้นตอนในการติดตาม ขับเคลื่อน ปรับปรุงและบำรุงรักษาให้ระบบตอบสนองต่อความต้องการของผู้ใช้อย่างต่อเนื่องเป็นการบริหารจัดการระบบสารสนเทศอย่างครบวงจร ก็จะทำให้ระบบเกิดการใช้งานอย่างคุ้มค่า และองค์กรประสบความสำเร็จในการนำระบบสารสนเทศเข้าไปใช้งาน

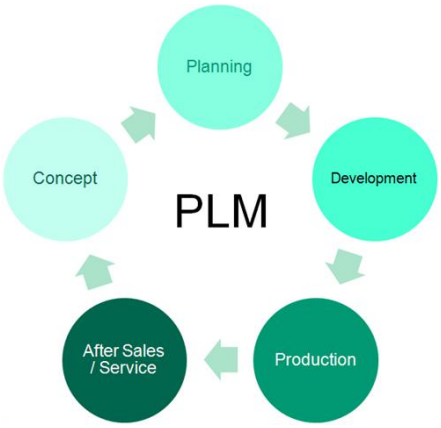
จึงมีนักวิจัยหลายท่านให้ความสำคัญ และนำเสนอทฤษฎีที่เกี่ยวกับวงจร กรอบแนวทางในการบริหารจัดการระบบสารสนเทศมาแก้ปัญหาดังกล่าว โดย Budde มองว่า หากมีการส่งมอบระบบสารสนเทศ ให้ผู้ใช้แล้ว ควรมีแผนการบำรุงรักษาแบบบูรณาการโดยการยึดวิสัยทัศน์ขององค์กร และสังเกตกลุ่มผู้ใช้ เพื่อโน้มน้าวผู้ใช้ให้มีความต้องการใช้งานระบบสารสนเทศอย่างต่อเนื่อง แต่ถ้าหากองค์กรนั้นมีการปรับเปลี่ยนของผู้ใช้ที่เกี่ยวข้องกับผลิตภัณฑ์ที่ส่งมอบนั้นก็จะทำให้ต้องเริ่มกำกับพฤติกรรมของผู้ใช้กลุ่มนั้นใหม่ ทำให้ใช้ระยะเวลานาน เพื่อให้เกิดความคุ้มค่า [3] Marchetta และคณะนำเสนอแนวทางในการปรับทัศนคติของผู้ใช้ ที่มีต่อระบบสารสนเทศให้มีความกระตือรือร้นที่ต้องการใช้งานเป็นหลัก แล้วทำการปรับเปลี่ยนระบบสารสนเทศให้สอดคล้องความต้องการของผู้ใช้ในขณะนั้น โดยองค์กรที่จะประสบความสำเร็จนั้น จำเป็นต้องมีงบประมาณสนับสนุนสูง [4] Conello และคณะแนะนำให้ยึดการเปลี่ยนแปลงทางเทคโนโลยีมาเป็นตัวกระตุ้นให้ผู้ใช้มีความต้องการอยากใช้งาน โน้มนาวให้ผู้ใช้มีมุมมองที่ดีต่อเทคโนโลยีและทำการปรับปรุงระบบสารสนเทศให้ทันเทคโนโลยีในปัจจุบัน ซึ่งองค์กรจำเป็นจะต้องมีบุคลากรด้านการพัฒนาระบบโดยเฉพาะ จึงจะประสบความสำเร็จ [5] จากแนวทางข้างต้น ที่เป็นวิธีการที่มีการบริหารจัดการระบบสารสนเทศภายหลังการส่งมอบ ต่อจากกระบวนการพัฒนาระบบ เพื่อให้ระบบมีการใช้งานอย่างคุ้มค่าต่อองค์กร แต่ก็ยังมีข้อเสียในเรื่องที่ต้องใช้ระยะเวลาในการควบคุมดูแลที่ยาวนาน ใช้เงินทุนสนับสนุนสูง และจำเป็นต้องมีบุคลากรเฉพาะด้าน ประจำภายในองค์กร

ดังนั้น งานวิจัยนี้ จึงได้นำเสนอวงจรการบริหารจัดการระบบสารสนเทศที่นำแอพลิเคชั่นไลน์ซึ่งเป็นแอพลิเคชั่นสื่อสังคมออนไลน์ที่เป็นที่นิยมในประเทศไทยมาใช้เป็นช่องทางสื่อสารเพื่อสร้างความตื่นตัวและการมีส่วนร่วมของผู้ใช้งาน โดยผู้วิจัยได้ทำการทดสอบแนวทางที่นำเสนอกับการพัฒนาระบบบริหารจัดการสินทรัพย์ของกรมชลประทาน

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

งานวิจัยนี้นำเสนอผลการศึกษา จากการนำวงจรในการบริหารจัดการระบบสารสนเทศให้เกิดความคุ้มค่าในการใช้งานและการลงทุนทางด้านระบบสารสนเทศ ทั้งนี้เพื่อเป็นการทบทวนงานวิจัยและบทความ ที่มีส่วนเกี่ยวข้องต่อวงจรการบริหารจัดการระบบสารสนเทศโดยมุ่งเน้นไปที่ กรอบแนวทางวงจรที่มีส่วนเกี่ยวข้อง สามารถสรุปประเด็นและข้อเสนอแนะจากงานวิจัยและบทความต่าง ๆ ได้ดังนี้

วงจรการบริหารจัดการผลิตภัณฑ์ คือ กระบวนการจัดการผลิตภัณฑ์โดยรวมของการผลิต ผลิตภัณฑ์ ตั้งแต่ขั้นแนวคิด ขั้นการออกแบบ ขั้นการผลิต ไปจนถึงการใช้งานผลิตภัณฑ์ การให้บริการหลังการขายจนกระทั่งการทำลายทิ้ง [6]



ภาพที่ 1: วงจรการบริหารจัดการผลิตภัณฑ์
ที่มา: https://virtualdutchman.com/2010/04/04/plm [Accessed May 01 2017]

Budde มองว่า หากมีการส่งมอบระบบให้ผู้ใช้แล้ว ควรมีแผนการบำรุงรักษา แบบบูรณาการ โดยการยึดวิสัยทัศน์ขององค์กร พร้อมทั้งใช้วิธีการสังเกตกลุ่มผู้ใช้ระบบต่าง ๆ แล้วยึดกลุ่มที่มีการแสดงออกในทางที่ ต่อต้านการใช้งานระบบสารสนเทศเป็นสิ่งสำคัญ และใช้วิธีการแก้ปัญหาด้วยการ โน้มน้าวผู้ใช้ด้วยการ ให้ข้อมูลที่เป็นประโยชน์สูงสุดต่อผู้ใช้ แนะนำผู้ใช้ในทางที่ไม่เกิดผลเสียต่อผู้ใช้เมื่อจำเป็นต้องใช้งานระบบ

เพื่อให้ผู้ใช้ระบบเกิดความต้องการใช้งาน แล้วจึงทำการประเมิน ผู้ใช้งานระบบต่าง ๆ อย่างต่อเนื่อง ใส่ใจและให้ความสำคัญต่อการให้บริการหลังการขายเทียบเท่ากระบวนการในการพัฒนาระบบ แต่ถ้าหากองค์กรนั้นมีการปรับเปลี่ยนของผู้ใช้ที่เกี่ยวข้องกับระบบที่ส่งมอบนั้น ก็จะทำให้ต้องเริ่มกำกับพฤติกรรมของผู้ใช้กลุ่มนั้นใหม่ ทำให้ใช้ระยะเวลานานและเงินทุนสนับสนุนสูง เพื่อให้เกิดการใช้งานอย่างคุ้มค่า [3]

Marchetta และคณะนำเสนอแนวทางด้วยการวางกรอบที่ใช้ในการปรับทัศนคติ ของผู้ใช้ ที่มีต่อเทคโนโลยีให้มีความกระตือรือร้นที่ต้องการใช้งานระบบสารสนเทศ มากกว่าการทำให้ผู้ใช้เกิดการต่อต้านต่อการเปลี่ยนแปลงขององค์กรในการนำระบบสารสนเทศเข้ามาใช้งานร่วมกับพนักงานในองค์กร แล้วจึงทำการปรับเปลี่ยนระบบที่ใช้ในการทำงานร่วมกับผู้ใช้ให้สอดคล้องความต้องการของผู้ใช้ในขณะนั้น เพื่อให้เกิดความสมดุลขององค์กร พร้อมทั้งเฝ้าติดตามให้ความช่วยเหลือแก่ผู้ใช้ และประเมินการทำงานของพนักงานที่ทำงานร่วมกับระบบสารสนเทศอย่างสม่ำเสมอ แต่ยังมีผลเสีย อยู่ที่องค์กรจะประสบความสำเร็จนั้น จะมีต้นทุนแฝงเกิดขึ้นจำเป็นต้องมีงบประมาณสนับสนุนสูง [4]

Conello และคณะแนะนำให้ยึดการเปลี่ยนแปลงทางเทคโนโลยี มาเป็นตัวกระตุ้นให้ผู้ใช้มีความต้องการอยากใช้งาน ทำการโน้มนาวให้ผู้ใช้ มีมุมมองที่ดีต่อเทคโนโลยีและทำการปรับปรุงระบบสารสนเทศ ให้ทันกับเทคโนโลยี ที่มีอยู่ในปัจจุบัน เพื่อให้เกิดความสะดวกต่อการใช้งาน โดยคำนึงถึงความยากง่ายในการใช้งานระบบเป็นปัจจัยสำคัญ โดยเฉพาะในส่วนที่มีการปฏิสัมพันธ์กับผู้ใช้ ซึ่งองค์กรจำเป็นจะต้องมีบุคลากรที่มีความชำนาญด้านการพัฒนาระบบสารสนเทศ โดยเฉพาะจึงจะสามารถนำพาองค์กรไปสู่ความสำเร็จ ในการนำระบบเข้าไปใช้งานภายในองค์กร [5]

3. วิธีการดำเนินการวิจัย

จากการศึกษาผลงานวิจัยและทบทวนบทความต่างๆ ที่เกี่ยวข้องกับวงจรและกรอบแนวทางในการบริหารจัดการระบบสารสนเทศ ผู้วิจัยได้สังเคราะห์และเลือกปัจจัยที่เหมาะสมเพื่อนำไปปรับใช้ในวงจรการบริหารจัดการระบบสารสนเทศให้เกิดการใช้งานระบบสารสนเทศอย่างคุ้มค่า ดังนี้

3.1 การคัดเลือกระบบสารสนเทศตัวอย่าง

ระบบสารสนเทศที่ถูกนำมาทดลองใช้กับวงจรการบริหารจัดการระบบสารสนเทศที่พัฒนาขึ้นนี้เป็นระบบที่ผู้วิจัยสามารถเข้าถึงได้ทั้งกลุ่มผู้ใช้และผู้พัฒนาระบบ ระบบดังกล่าวคือ ระบบสารสนเทศเพื่อการบริหารจัดการสินทรัพย์ของกรมชลประทาน ซึ่งเป็นหน่วยงานภาครัฐภายใต้กระทรวงเกษตรและสหกรณ์

ระบบสารสนเทศเพื่อการบริหารจัดการสินทรัพย์กรมชลประทาน พัฒนาขึ้นในรูปแบบ Web Application ที่สามารถใช้บริการผ่านอินเตอร์เน็ตได้ โดยระบบให้บริการด้านการบริหารทะเบียนสินทรัพย์ ตั้งแต่การขึ้นทะเบียนสินทรัพย์ การโอนยืมสินทรัพย์ การซ่อมบำรุงสินทรัพย์ การจำหน่ายสินทรัพย์ การจัดทำรายงานสินทรัพย์ในรูปแบบต่างๆ และมีผู้ใช้บริการทั่วประเทศ ซึ่งสามารถแบ่งตามบริบทการทำงานได้ทั้งหมด 4 ส่วน ดังนี้

ส่วนที่ 1 กลุ่มผู้ใช้งานทั่วไป ผู้ใช้งานระบบด้านการจัดการทะเบียนสินทรัพย์ด้านต่างๆ กระจายอยู่ตามหน่วยงานชลประทานทั่วประเทศ จำนวน 400 users

ส่วนที่ 2 กลุ่มผู้ใช้ระดับหัวหน้าสำนัก ผู้ใช้งานระบบด้านการติดตาม ตรวจสอบ การทำงานของกลุ่มผู้ใช้งานทั่วไป ซึ่งอยู่ตามหน่วยงานใหญ่ในภูมิภาคต่าง ๆ จำนวน 50 users

ส่วนที่ 3 กลุ่มผู้ใช้ระดับผู้ติดตาม ผู้ใช้งานระบบด้านการติดตามการดำเนินงานของหน่วยงานชลประทาน และรายงานข้อมูลต่อจากกลุ่มผู้ใช้ระดับหัวหน้าสำนัก ปฏิบัติงานภายในส่วนกลาง กรมชลประทาน จำนวน 5 users

ส่วนที่ 4 กลุ่มผู้ใช้ระดับผู้ดูแลระบบ ผู้ใช้งานระบบด้านการดูแลระบบและจัดการข้อมูลพื้นฐานของระบบ จำนวน 2 users

3.2 การคัดเลือกประชากรและกลุ่มตัวอย่าง

จากระบบสารสนเทศตัวอย่างที่มีกลุ่มผู้ใช้งานระบบทั้งสิ้น 4 ส่วน มีผู้ใช้งานระบบจำนวน 457 users ผู้วิจัยได้ใช้ผู้ใช้งานระบบทั้งหมดเป็นกลุ่มประชากร และทำการกำหนดกลุ่มตัวอย่างในครั้งนี้ด้วยวิธีการสุ่มตัวอย่างแบบ Convenience Sampling โดยเลือกผู้ใช้งานระบบที่ผู้วิจัยสามารถติดต่อเพื่อเก็บข้อมูลได้สะดวก

3.3 การสังเคราะห์แนวทางการบริหารจัดการระบบสารสนเทศ

การคัดเลือกแนวทางเพื่อนำมาใช้ในการดำเนินวงจรการบริหารจัดการระบบสารสนเทศเพื่อให้เกิดการใช้งานระบบสารสนเทศอย่างคุ้มค่านั้น ผู้วิจัยได้วิเคราะห์วงจรและกรอบแนวทางที่เกี่ยวข้อง 3 วิธี มาเปรียบเทียบและสังเคราะห์วิธีการที่ใช้ว่ามีข้อดีและข้อเสียต่างกันอย่างไร และวิธีไหนเหมาะสมสำหรับการนำมาใช้เพื่อติดตามกำกับดูแลระบบสารสนเทศหลังจากการส่งมอบระบบ สำหรับงานวิจัยนี้ผู้วิจัยได้เลือกแนวทางที่เป็นข้อดีมาใช้ร่วมกัน ซึ่งแสดงรายละเอียด ดังตารางที่ 1

ตารางที่ 1: แนวทางวิธีการบริหารจัดการระบบสารสนเทศ

วิธีการ / แนวทาง	Budde	Marchetta	Conello	แนวทางที่เลือก
ควบคุมพฤติกรรมผู้ใช้	×	✓	✓	×
ยึดการเปลี่ยนแปลงทางเทคโนโลยี	✓	×	✓	✓
ยึดวิสัยทัศน์องค์กร	✓	×	×	✓
ติดตามให้ความช่วยเหลือผู้ใช้	×	✓	×	✓
ให้การสนับสนุนทางเงินทุน	✓	×	×	×
เพิ่มบุคลากรด้านการพัฒนาระบบ	×	×	✓	×
ปรับปรุงระบบให้ตอบสนองความต้องการผู้ใช้ตลอดเวลา	×	✓	✓	✓
คำนึงถึงความยากง่ายในการใช้ระบบ	×	×	✓	✓

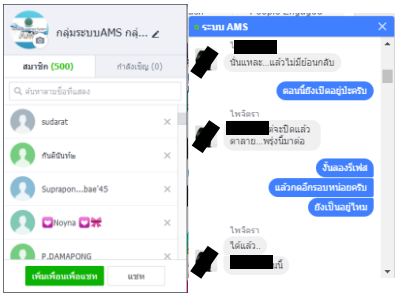
3.4 การออกแบบวงจรการบริหารจัดการระบบสารสนเทศ

วงจรการบริหารจัดการระบบสารสนเทศในงานวิจัยนี้หมายถึง กระบวนการในการจัดการระบบตั้งแต่ขั้นตอนการออกแบบพัฒนาระบบจนถึงขั้นตอนการส่งมอบ และดูแลระบบหลังการส่งมอบ ให้มีการใช้งานระบบอย่างต่อเนื่องและคุ้มค่าต่อการลงทุน ที่ผ่านมาจะเห็นได้ว่า ในการพัฒนาระบบโดยทั่วไปจะให้ความสำคัญต่อกระบวนการในการพัฒนา แต่ยังขาดแนวทางในการบริหารจัดการระบบ หลังจากระบบถูกส่งมอบเรียบร้อยแล้ว ในงานวิจัยนี้จึงได้สังเคราะห์แนวทางในการบริหารจัดการระบบเพื่อนำมาใช้ร่วมกับระบบสารสนเทศตัวอย่าง ซึ่งมีแนวทางดังต่อไปนี้ 1) ติดตามให้ความช่วยเหลือผู้ใช้ผ่านแอพลิเคชั่นสื่อสังคมออนไลน์ เป็นแนวทางในการให้ความช่วยผู้ใช้งานผู้ใช้งานระบบเมื่อเกิดปัญหาต่างๆ ขึ้น 2) ปรับปรุงระบบให้ตอบสนองความต้องการผู้ใช้ตลอดเวลา เป็นแนวทางที่ต้องปรับปรุงระบบสารสนเทศให้ตรงกับความต้องการของผู้ใช้ที่เพิ่มขึ้นในส่วนที่อยู่ภายใต้ข้อตกลงและสัญญาของการพัฒนาระบบ 3) คำนึงถึงความยากง่ายในการใช้งาน เป็นแนวทางในการสังเกตผลการใช้งานของผู้ใช้งานระบบ ซึ่งวิเคราะห์จากปัญหาที่เกิดขึ้นระหว่างการใช้งานที่แจ้งผ่านทางสื่อสังคมออนไลน์ โดยปัญหาบางส่วนจะได้รับการปรับปรุงให้อำนวยความสะดวกต่อการใช้งานของผู้ใช้แต่ละส่วน 4) ยึดการเปลี่ยนแปลงทางเทคโนโลยี เป็นการปรับปรุงระบบสารสนเทศโดยนำเทคโนโลยีใหม่ ๆ ที่เกิดขึ้นเข้ามาปรับปรุงให้ระบบสามารถปฏิสัมพันธ์กับผู้ใช้ให้ดีขึ้น 5) ยึดวิสัยทัศน์องค์กร เป็นแนวทางในการติดตามควบคุมผู้ใช้ ให้ปฏิบัติตามสิ่งที่ผู้บริหารมุ่งหวังให้เกิดขึ้นเกี่ยวกับการนำระบบสารสนเทศมาใช้งานในองค์กร

3.5 การทดลองวงจรและติดตามผล

จากการเลือกใช้แนวทางเพื่อนำมาใช้ภายในวงจร ผู้วิจัยได้มุ่งเน้นการติดตามผลภายหลังที่กระบวนการพัฒนาเสร็จสิ้นในส่วนการส่งมอบและหลังการส่งมอบระบบแล้ว โดยให้ความสำคัญในแนวทางการติดตามให้ความช่วยเหลือผู้ใช้งานระบบสารสนเทศ การปรับปรุงระบบสารสนเทศให้ตอบสนองความต้องการของผู้ใช้ และคำนึงถึงความยากง่ายในการใช้งานเป็นแนวทางหลักในการทดลองใช้งานวงจร เฉกเช่นเดียวกับวิธีการให้บริการหลังการขาย

ผู้วิจัยได้ให้ผู้พัฒนาระบบสารสนเทศเป็นผู้กำกับดูแล คอยติดตามให้ความช่วยเหลือ ปรับปรุงระบบให้ตรงตามความต้องการผู้ใช้ พร้อมทั้งคำนึงถึงความยากง่ายในการใช้งานระบบสารสนเทศและประเมินผลลัพธ์จากจำนวนผู้ใช้งาน ระยะเวลาการปรับปรุงระบบหลังการส่งมอบ โดยการให้ความช่วยเหลือผู้ใช้จะต้องรวดเร็วและทันต่อความต้องการของผู้ใช้ พร้อมทั้งรับผลการตอบกลับจากผู้ใช้ เพื่อการพัฒนาระบบให้ดียิ่งขึ้นนั้น ผู้วิจัยจึงใช้ช่องทางสื่อสังคมออนไลน์ติดต่อสื่อสารกับผู้ใช้ผ่านแอพลิเคชั่นที่เป็นที่นิยม เช่น ไลน์ เฟรชบุ๊ค อีเมล เป็นต้น ซึ่งปัจจุบันสามารถใช้สื่อสารผ่านรูปภาพและวิดีโอ ได้โดยง่ายเพื่อความเข้าใจระหว่างผู้พัฒนาระบบและผู้ใช้



ภาพที่ 2: การติดต่อสื่อสารระหว่างผู้พัฒนาและผู้ใช้งานผ่านสื่อสังคมออนไลน์

4. ผลการดำเนินการวิจัย

จากการใช้งานวงจรการบริหารจัดการระบบสารสนเทศที่พัฒนาขึ้นนี้ร่วมกับระบบสารสนเทศตัวอย่าง ด้วยการเพิ่มแนวทางเพื่อจัดการบริการหลังการขาย ปรากฏว่าการให้ความช่วยเหลือผู้ใช้งานผ่านช่องทางสื่อสังคมออนไลน์ ทำให้สามารถเข้าถึงและติดต่อกับผู้ใช้ได้โดยง่ายและสามารถให้ความช่วยเหลือได้ทันเหตุการณ์ นอกจากนี้การใช้ทีมผู้พัฒนาระบบเป็นผู้กำกับดูแล คอยให้ความช่วยเหลือ และปรับปรุงระบบให้ตรงตามความต้องการของผู้ใช้โดยตรง ยังทำให้สามารถใช้เวลาในการแก้ปัญหาต่าง ๆ ได้อย่างรวดเร็ว และสามารถปรับปรุงผลิตภัณฑ์ให้ปรับเปลี่ยนไปตามเทคโนโลยีในขณะนั้น เพื่อตอบสนองการใช้งานที่ง่ายและสะดวกขึ้นได้อย่างง่าย ซึ่งสามารถวัดได้จาก การทำงานของระบบที่มีความชัดเจน การกรอกข้อมูลไม่ซับซ้อน เมนูและข้อความในระบบมีความชัดเจนเข้าใจได้ง่ายขึ้นกว่าระบบที่ส่งมอบ

นอกจากการปรับปรุงระบบแล้ว ปัจจัยหลักที่ใช้ขับเคลื่อนระบบ คือ ความสอดคล้องกับวิสัยทัศน์องค์กร ซึ่งสำหรับกรมชลประทานแล้วมีวิสัยทัศน์ที่ชัดเจน คือการนำระบบสารสนเทศเข้ามาช่วยแบ่งเบาภาระในการทำงาน ซึ่งในกรณีศึกษานี้ระบบจะไม่สามารถใช้งานให้เกิดประโยชน์ได้เลย หากผู้บริหารไม่ให้ความสำคัญกับระบบและไม่เปลี่ยนแปลงกฎระเบียบ และนโยบายภายในให้เอื้อต่อการเปลี่ยนแปลงกระบวนการภายในมาสู่การใช้ระบบ



ภาพที่ 3: การใช้สื่อสังคมออนไลน์กำกับดูแลให้ความช่วยเหลือผู้ใช้



ภาพที่ 4: วิวัฒนาการการปรับปรุงผลิตภัณฑ์ระบบสารสนเทศ

ความสำเร็จของการพัฒนาระบบซึ่งยกมาเป็นกรณีศึกษานี้ สามารถสะท้อนให้เห็นได้ในรายละเอียดการปรับแก้ ที่รวบรวมได้จากผู้ใช้แต่ละครั้ง ดังแสดงในตารางที่ 2 ซึ่งล้วนมีเนื้อหามุ่งเน้นให้ระบบสามารถทำงานได้ง่ายขึ้น

ตารางที่ 2: ตัวอย่างรายละเอียดการปรับปรุงระบบแต่ละเวอร์ชั่น

เวอร์ชั่น	รายละเอียดการปรับปรุง	เหตุผลการปรับปรุง
1.0	เวอร์ชั่นระบบส่งมอบระบบแล้ว	-
1.1	1. การ upload ไฟล์ซึ่งมีขนาดใหญ่บางครั้งต้องใช้เวลานาน ดังนั้น ควรมีการบอกสถานะการ upload ไฟล์ 2. อายุการรับประกันของครุภัณฑ์บางประเภทมีหลายระยะ ควรให้สามารถกรอกโดยการระบุรายละเอียดอายุการรับประกันอื่น ๆ ในวงเล็บ	1. แก้ไขให้ใช้งานง่าย 2. แก้ไขให้ใช้งานง่าย
1.2	1. การโอน/การยืม ให้ค้นหาด้วยหมายเลขครุภัณฑ์ 2. เวลาตอบกลับให้แก้ไขคำขอ ให้มีกล่องไว้เรียกดูรายละเอียดที่ตอบกลับ	1. แก้ไขให้ใช้งานง่าย 2. แก้ไขให้ใช้งานง่าย
2.0	1. เพิ่มการค้นหารายการในการจัดการรายงานประจำปี 2. เพิ่มชื่อผู้ขึ้นทะเบียนครุภัณฑ์	1. แก้ไขให้ใช้งานง่าย 2. แก้ไขให้ใช้งานง่าย
2.1	1. เพิ่มหน้าค้นหาชื่อครุภัณฑ์ 2. เพิ่มสถานะครุภัณฑ์ ระบุเป็นไม่ใช่ของหน่วยงานพร้อมรายงาน 3. เปลี่ยนการเลือกลักษณะครุภัณฑ์เป็นการป้อนข้อมูล 4. เพิ่มการลบข้อมูลในรายงานครุภัณฑ์ที่ไม่ใช่ของหน่วยงาน 5. เพิ่มหน้า จัดการส่วนงาน	1. แก้ไขให้ใช้งานง่าย 2. แก้ไขให้ใช้งานง่าย 3. แก้ไขให้ใช้งานง่าย 4. เพิ่มเติมให้เข้ากับบริบทการทำงาน 5. เพิ่มเติมให้เข้ากับบริบทการทำงาน
3.0	1. เอาลายเซ็นใน พด.44 พด.45 ออก 2. เพิ่มช่องรายการค่าเสื่อมสุทธิในรายงานครุภัณฑ์	1. ปรับเปลี่ยนตามเทคโนโลยี 2. แก้ไขให้ใช้งานง่าย

5. สรุป

จากผลการดำเนินการวิจัยทั้งหมด พบว่าการบริหารจัดการระบบสารสนเทศอย่างครบวงจร ด้วยการเพิ่มแนวทางในการบริหารจัดการบริการหลังการขายแก่ผู้ใช้งาน เพิ่มเติมจากกระบวนการพัฒนาระบบ ทำให้ระบบถูกนำไปใช้งานให้เกิดประโยชน์อย่างคุ้มค่า ด้วยการประยุกต์ใช้สื่อสังคมออนไลน์เป็นช่องทางในการติดต่อสื่อสาร กำกับดูแล และติดตามผู้ใช้ ทำให้ช่องว่างการให้ความช่วยเหลือแก่ผู้ใช้นั้นหมดไป

ผลการทดลองใช้วงจรที่นำเสนอสะท้อนให้เห็นว่าแนวทางการยึดการเปลี่ยนแปลงทางเทคโนโลยี วิสัยทัศน์องค์กร การติดตามให้ความช่วยเหลือผู้ใช้ การปรับปรุงระบบให้ตอบสนองความต้องการผู้ใช้ตลอดเวลา และการคำนึงถึงความยากง่ายในการใช้งาน ล้วนเป็นแนวทางในการให้บริการหลังการขายภายใต้วงจรการบริหารจัดการระบบสารสนเทศที่มีความสำคัญไม่ยิ่งหย่อนไปกว่ากัน โดยปัจจัยหลักที่ส่งผลให้ระบบสารสนเทศถูกนำไปใช้งานอย่างคุ้มค่าและเป็นพื้นฐานอันก่อให้เกิดหลายปัจจัยดังที่กล่าวมาข้างต้น คือ ช่องทางการติดต่อสื่อสารผ่านสื่อสังคมออนไลน์

เอกสารอ้างอิง

[1] Jenner, S. (2009). Realising Benefits from Government ICT Investment – A fool's errand. Academic Publishing International Ltd. Reading. UK.

[2] Chaffey, D. and Smith, P. (2013). E Marketing Excellence: Planning and Optimizing Your Digital Marketing. New York: Routledge.

[3] Budde, O. (2010). Holistic PLM-model – deduction of a holistic PLM-model from the general dimensions of an integrated management, International Conference on Product Lifecycle Management Jan 2010.

[4] Marchetta, M. Mayer, F. and Forradellas, R. (2011). A reference framework following a proactive-product approach for Product Lifecycle Management, Computers in Industry 62 (2011) 672–683.

[5] Corallo, A. Lazoi, M. Margarito, A. and Pinna, D. (2013). Developing a PLM Framework: a case study application in an energy company, The IFIP WG5.1 10th International Conference on Product Lifecycle Management – PLM13 July 2013.

[6] Fatallah, A. Cardinal, J. Ermine, J. and Bocquet, J. (2009). Enterprise modelling: building a product lifecycle (PLM) model as a component of the integrated vision of the enterprise. Research in Interactive Design, Springer Verlag, 2009, vol. 3.

การพัฒนาออนโทโลยีเพื่อการสืบค้นสื่อความหมายสำหรับการเพาะปลูกผักอินทรีย์
Ontology Development for Semantic Search of Organic Vegetable Cultivation

วรพจน์ สุวรรณภิภพ (Worapoj Suwanpipob)[1] งามนิจ อาจอินทร์ (Ngamnij Arch-Int)[2]
สมจิตร อาจอินทร์ (Somjit Arch-Int)[3] ปวีณา วันชัย (Paweena Wanchai)[4]
ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น
[1]worapoj_suwan@kkumail.com, [2]ngamnij@kku.ac.th, [3]somjit@kku.ac.th, [4]wpaweena@kku.ac.th

บทคัดย่อ

ปัจจุบันมีการพัฒนาเว็บแอพพลิเคชันสำหรับให้ข้อมูลทางด้านเกษตรอินทรีย์จำนวนมาก แต่เว็บไซต์ส่วนใหญ่ยังสืบค้นโดยใช้คำสำคัญ และผู้ใช้งานไม่สามารถสืบค้นข้อมูลในลักษณะถามตอบ หรือสืบค้นแบบสื่อความหมายได้ ทำให้งานวิจัยนี้มีวัตถุประสงค์ในการพัฒนาออนโทโลยีสำหรับการสืบค้นแบบสื่อความหมายของการเพาะปลูกผักอินทรีย์ ซึ่งงานวิจัยนี้ได้ใช้คุณสมบัติ OWL เข้ามาแก้ไขปัญหาคำพ้องความหมายที่สนับสนุนการสืบค้นแบบสื่อความหมาย อีกทั้งยังมีส่วนต่อประสานกับผู้ใช้สำหรับสืบค้นข้อมูลในรูปแบบของ ทริปเปิล โดยใช้เทคนิคออโต้คอมพลีต ซึ่งผลลัพธ์การสืบค้น และตรงกับคำตอบของผู้เชี่ยวชาญ พบว่ามีค่าความแม่นยำอยู่ที่98 เปอร์เซ็นต์

คำสำคัญ: ออนโทโลยี เว็บสื่อความหมาย เกษตรอินทรีย์

Abstract

Although many web applications are developed for providing agriculture information, most applications support only keyword-based searching. Users unable to search data in the form of question answering or semantic search. This research proposes the ontology development for semantic search of organic vegetable cultivation. The research aims to provide the OWL properties for resolving synonym conflicts and to support semantic search. This research also provides query interface to enable users posting their queries in a triple form by using autocomplete technique. The system was validated with the precision, and recall measurements and evaluated by expert which has 9 8 percent of F-measure rate.

Keyword: Ontology, Semantic Web, Organic Farming

1. บทนำ

การเผยแพร่ความรู้ทางด้านเกษตรอินทรีย์จะอยู่ในรูปแบบที่หลากหลาย เช่น ในรูปแบบของหนังสือ หรือเว็บไซต์ ซึ่งเว็บไซต์ส่วนใหญ่จะแสดงข้อมูลอยู่ในรูปแบบของบทความที่สามารถสืบค้นความรู้เกษตรอินทรีย์ในลักษณะคำสำคัญ (Keywords) ซึ่งไม่สามารถทำให้คอมพิวเตอร์เข้าใจความหมายของข้อมูลมากพอ อีกทั้งผู้ใช้ทั่วไปต้องค้นหาข้อมูลเกษตรอินทรีย์จากหลากหลายเว็บไซต์ เช่น เว็บไซต์ศัตรูพืช เว็บไซต์เพาะปลูก และเว็บไซต์สูตรปุ๋ย ซึ่งทำให้เสียเวลาในการค้นหา อีกทั้งอาจได้ข้อมูลที่ไม่ตรงกับความต้องการของผู้ใช้งาน นอกจากนี้เว็บไซต์ความรู้ทางด้านการเกษตรทั่วไปไม่สามารถสืบค้นในลักษณะแบบถามตอบ (Question Answering: QA) ทำให้ในปัจจุบันหลายวิจัย [1] [2] [3] ได้จัดทำระบบสืบค้น โดยใช้ออนโทโลยี (Ontology) [4] เข้ามาเป็นแหล่งเก็บองค์ความรู้และพัฒนาระบบสืบค้นสื่อความหมายทางด้านการเกษตร อย่างไรก็ตามก็ยังมีปัญหาทางด้านการสืบค้น ดังต่อไปนี้

(1) ไม่สามารถสืบค้นในลักษณะคำพ้องความหมาย (Synonym) หรือคำที่เขียนต่างกันแต่มีความหมายเหมือนกัน เช่น อีตู่ไทยที่เป็นภาษาอีสานจะพ้องความหมายกับกะเพราที่เป็นภาษากลาง หรือเต่าทองที่พ้องความหมายกับด้วงเต่าลาย

(2) ไม่มีการใช้คุณสมบัติของ OWL (Web Ontology Language) [5] เพื่อช่วยให้คอมพิวเตอร์สามารถสืบค้นข้อมูลได้อย่างสื่อความหมาย ทำให้การสืบค้นได้ข้อมูลที่ไม่ถูกต้องสมบูรณ์

(3) มีการสืบค้นในลักษณะของคำสำคัญ แต่ไม่สามารถสืบค้นในลักษณะแบบถามตอบได้

จากปัญหาดังกล่าวงานวิจัยนี้จึงได้จัดทำการพัฒนาออนโทโลยีเพื่อการสืบค้นแบบสื่อความหมายสำหรับการเพาะปลูกผักอินทรีย์ เพื่อแก้ไขปัญหาแต่ละข้อ ดังนี้ (1) ออกแบบออนโทโลยีเพื่อ

รองรับการสืบค้นในลักษณะคำพ้องความหมาย (2) ออกแบบการใช้คุณสมบัติของ OWL เช่น owl:equivalentProperty, owl:InverseOf และ owl:hasValue เป็นต้น เพื่อช่วยให้คอมพิวเตอร์สามารถอนุมาน (infer) ความสัมพันธ์อื่น ๆ ของข้อมูลเพิ่มเติม (3) ออกแบบ และพัฒนาระบบสืบค้นในลักษณะถามตอบในรูปแบบทริปเปิล (Triple) ที่ใช้เทคนิคออโต้คอมพลีต (Autocomplete) โดยดึงข้อมูลระดับอินสแตนด์ (Instant) ที่ถูกอนุมาน โดยเครื่องมืออนุมาน (Inference Engine) ให้แสดงออกในรูปแบบของทริปเปิลทั้งหมด เพื่อช่วยให้ผู้ใช้งานสามารถสืบค้นข้อมูลทริปเปิลได้อย่างสะดวก และตรงกับความต้องการอย่างแท้จริง

2. ทฤษฎีพื้นฐานและงานวิจัยที่เกี่ยวข้อง

2.1 ทฤษฎีพื้นฐาน

เทคโนโลยีเว็บเชิงความหมาย (Semantic Web Technologies) [6] เป็นเทคโนโลยีที่ถูกใช้ในการบูรณาการข้อมูลเชิงความหมาย โดยสร้างความสัมพันธ์ให้กับข้อมูล และจัดเก็บองค์ความรู้ลงในออนโทโลยี โดยการอธิบายถึงสิ่งต่างๆ เพื่อให้สามารถเข้าใจความหมายของสิ่งนั้นได้อย่างชัดเจน และสร้างความสัมพันธ์ระหว่างกันที่อยู่ภายใต้ขอบเขตที่สนใจ ทั้งนี้จะมีภาษาที่ถูกพัฒนาเพื่อใช้แทนออนโทโลยี เช่น ภาษา OWL สำหรับอธิบายความหมายของข้อมูลในออนโทโลยี พร้อมกำหนดความสัมพันธ์และคุณสมบัติของข้อมูล เพื่อให้เครื่องคอมพิวเตอร์สามารถเข้าใจความหมายได้มากขึ้น

2.2 งานวิจัยที่เกี่ยวข้อง

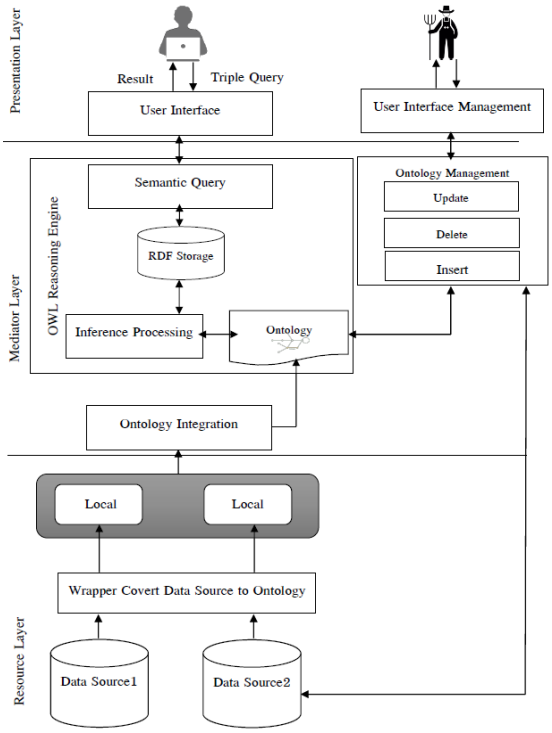
งานวิจัยที่เกี่ยวข้องกับการสืบค้นทางด้านเกษตรได้รับความสนใจเป็นอย่างมาก เนื่องจากเป็นการเพิ่มช่องทางในการสืบค้นข้อมูลแบบสื่อความหมาย และตรงกับความต้องการของผู้ใช้งาน ซึ่งในงานวิจัย Ontology based System for Pests and Disease Management of Grapes in India [2] ได้นำเสนอออนโทโลยีในการรวบรวมความรู้ของศัตรูพืช และโรคขององุ่นจากหลากหลายเว็บไซต์ และสกัดคำต่าง ๆ ให้อยู่ในรูปแบบของคำสำคัญ นอกจากนี้ยังมีการพยากรณ์การเกิดศัตรูพืช โดยใช้ตรรกศาสตร์คลุมเครือ (fuzzy logic) และพัฒนาแอพพลิเคชั่น (Application) สำหรับเผยแพร่ความรู้ทางด้านจัดการศัตรูพืช ทั้งนี้งานวิจัยนี้มีข้อดีคือสามารถพยากรณ์การเกิดศัตรูพืช ข้อเสียคือสามารถสืบค้นแบบคำสำคัญ โดยใช้ตำแหน่งของผู้ใช้งานเพียงอย่างเดียวเท่านั้น, ไม่แสดงรายละเอียดของคุณสมบัติ OWL และไม่มีการแก้ไขปัญหาคำพ้องความหมายที่มีอยู่ในข้อมูล ต่อมาในงานวิจัยการออกแบบออนโทโลยีข้าว-ไทย [3] ได้นำเสนอการสืบค้นด้วยคำสำคัญจากออนโทโลยี โดยใช้คุณสมบัติ OWL ได้แก่ owl:InverseOf และ owl:TransitiveProperty ซึ่งในงานวิจัยนี้มีข้อเสียคือไม่มีผู้เชี่ยวชาญทางด้านข้าวมายืนยันผลลัพธ์ของข้อมูล ทำให้คำตอบของการสืบค้นไม่น่าเชื่อถือ, ไม่แสดงตัวอย่างการใช้คุณสมบัติ OWL และไม่มีการสืบค้นในลักษณะคำพ้องความหมาย สุดท้ายเป็นงานวิจัย การพัฒนาออนโทโลยีเพื่อจัดการความรู้ด้านแปรรูปข้าว [4] ได้นำเสนอออนโทโลยีสำหรับสืบค้นในลักษณะคำสำคัญ โดยใช้คุณสมบัติ OWL เพื่อให้เข้าใจความหมายของข้อมูลได้หลากหลาย ซึ่งในงานวิจัยนี้มีข้อเสียคือไม่แสดงคุณสมบัติของ OWL และไม่มีการสืบค้นในลักษณะคำพ้องความหมาย

จากการศึกษาปัญหาของงานวิจัยที่เกี่ยวข้องงานวิจัยนี้จึงได้แก้ปัญหาการอนุมานคำตอบโดยภาษา OWL, สืบค้นข้อมูลแบบทริปเปิลในลักษณะถามตอบ โดยใช้เทคนิคออโต้คอมพลีตเข้ามาช่วยเหลือในการสืบค้น โดยดึงข้อมูลระดับอินสแตนด์ภายในออนโทโลยีให้แสดงออกในรูปแบบของ ทริปเปิลทั้งหมด เพื่อช่วยในการสืบค้นของผู้ใช้งาน อีกทั้งงานวิจัยนี้ได้แก้ไขปัญหาคำพ้องความหมายจากการสืบค้นชื่อผักของภาษาท้องถิ่นเพื่อให้สามารถสืบค้นได้หลากหลาย

3. วิธีดำเนินงานวิจัย

3.1 ออกแบบสถาปัตยกรรมของระบบ

งานวิจัยนี้ได้ออกแบบสถาปัตยกรรมระบบของฐานความรู้แบบสื่อความหมายสำหรับเพาะปลูกผักอินทรีย์ ดังภาพที่ 1

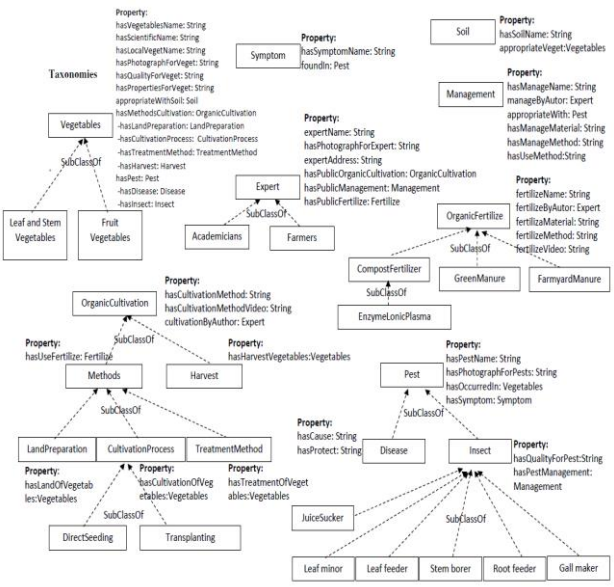


ภาพที่ 1: ระบบสถาปัตยกรรมระบบเพื่อการสืบค้นเชิงความหมายสำหรับเพาะปลูกผักอินทรีย์

จากภาพที่ 1 แสดงสถาปัตยกรรมของระบบที่มีสามระดับชั้น ได้แก่ ชั้นที่หนึ่งระดับนำเสนอ (Presentation Layer) เป็นส่วนที่ติดต่อกับผู้ใช้งาน โดยแบ่งผู้ใช้งานออกเป็น 2 กลุ่ม ได้แก่ กลุ่มผู้ใช้งานที่สามารถสืบค้นคำถามจากระบบ และกลุ่มผู้ดูแล หรือผู้เชี่ยวชาญที่สามารถเพิ่ม ลบ และแก้ไขข้อมูล ชั้นที่สองระดับตัวกลาง (Mediator Layer) ทำหน้าที่บูรณาการออนโทโลยีที่มาจากระดับแหล่งข้อมูล และนำข้อมูลจากระดับนำเสนอเข้ามาผ่านกระบวนต่าง ๆ ประกอบด้วยสามส่วนได้แก่ส่วนที่หนึ่ง คือ ส่วนเครื่องมือเชิงเหตุผลแบบ OWL (OWL Reasoning Engine) ประกอบด้วยส่วนการสืบค้นแบบสื่อ-ความหมาย (Semantic Query) ในลักษณะออโต้คอมพลีต, ส่วนแหล่งจัดเก็บข้อมูล RDF (RDF Storage) เป็นแหล่งเก็บข้อมูลที่ได้จากกระบวนการอนุมานข้อมูลของเครื่องมือเชิงเหตุผลแบบ OWL ซึ่งข้อมูลจะถูกเก็บในลักษณะของทริปเปิล, ส่วนการประมวลผลเพื่อการอนุมานข้อมูล (Inferencing Processing) ทำหน้าที่เรียนรู้ หรือวิเคราะห์ความรู้จากออนโทโลยีเพื่อแปลงเป็นทริปเปิล และเก็บไว้ในแหล่งข้อมูล RDF และส่วนที่เป็นออนโทโลยี เป็นแหล่งเก็บความรู้เกี่ยวกับผักอินทรีย์ ส่วนที่สองคือ ส่วนการจัดการออนโทโลยี (Ontology Management) ทำหน้าที่จัดการข้อมูลในออนโทโลยี เช่น การเพิ่ม ลบ หรือแก้ไขข้อมูลในออนโทโลยีเป็นต้น และส่วนที่สามคือ ส่วนการบูรณาการออนโทโลยี (Ontology Integration) ทำหน้าที่ในการบูรณาการออนโทโลยีเข้าด้วยกัน โดยใช้วิธีการเชื่อมโยงความสัมพันธ์เชิงความหมาย และชั้นที่สามระดับแหล่งข้อมูล (Resource Layer) ประกอบด้วย แหล่งข้อมูล (Data Source) ทำหน้าที่เก็บข้อมูลการเพาะปลูกผักอินทรีย์ โดยแบ่งออกเป็น 2 แหล่งข้อมูลคือ แหล่งข้อมูลที่ 1 เป็นแหล่งเก็บข้อมูลที่นำเข้าจากแหล่งอื่น ซึ่งอยู่ในรูปแบบของฐานข้อมูลที่ประกอบด้วย ตารางผัก, ตารางเพาะปลูก, ตารางศัตรูพืช, ตารางดิน, ตารางลักษณะอาการ, ตารางผู้เชี่ยวชาญ, ตารางวิธีการจัดการศัตรูพืช และตารางปุ๋ย และแหล่งข้อมูลที่ 2 เป็นแหล่งเก็บข้อมูลการเพาะปลูกผักอินทรีย์ของระบบ โดยจัดเก็บอยู่ในรูปแบบฐานข้อมูล ซึ่งในส่วนระดับชั้นแหล่งข้อมูลนี้จะมี แรพเพอะ (Wrapper) ทำหน้าที่แปลงข้อมูลจากฐานข้อมูลให้อยู่ในรูปแบบของออนโทโลยีท้องถิ่น (Local Ontology) เพื่อเตรียมส่งให้กับส่วนการบูรณาการออนโทโลยีในระดับตัวกลางต่อไป

3.2 ออกแบบโครงสร้างออนโทโลยี

งานวิจัยนี้ได้ออกแบบโครงสร้างออนโทโลยีการเพาะปลูกผักอินทรีย์ โดยได้นำเสนอคลาส (Class) และพรอปเพอร์ตี (Property) ในรูปแบบแทกโซโนมี (Taxonomies) ดังภาพที่ 2



ภาพที่ 2: โครงสร้างออนโทโลยีลักษณะแทกโซโนมี

จากภาพที่ 2 ประกอบด้วยคลาสต่าง ๆ เช่น คลาสผัก (Vegetables Class) คลาสเพาะปลูกอินทรีย์ (OrganicCultivation Class) คลาสผู้เชี่ยวชาญ (Expert Class) คลาสศัตรูพืช (Pest Class) คลาสลักษณะอาการ (Symptom Class) คลาสวิธีจัดการ (Management Class) คลาสดิน (Soil Class) และคลาสปุ๋ยอินทรีย์ (OrganicFertilize Class) ทั้งนี้แต่ละคลาสมีอ็อบเจกต์พรอปเพอร์ตี (Object Property) และดาต้า-ไทป์พรอปเพอร์ตี (DataType Property) ตัวอย่างดาต้าไทป์- พรอปเพอร์ตีของคลาสผัก ได้แก่ มีชื่อสามัญ (hasVegetablesName) มีชื่อวิทยาศาสตร์ (hasScientificName) มีชื่อท้องถิ่น (hasLocalVegetName) มีรูปภาพ (hasPhotographForVeget) มีลักษณะ (hasQualityForVeget) มีสรรพคุณ (hasPropertiesForVeget) และตัวอย่างอ็อบเจกต์พรอปเพอร์ตีของคลาสผัก ได้แก่ เหมาะสมกับดิน (appropriateWithSoil) มีขั้นตอนเพาะปลูก (hasMethodsOCultivation) และมีศัตรูพืช (hasPest) เป็นต้น

3.3 คุณสมบัติของออนโทโลยี

ในส่วนนี้แสดงให้เห็นถึงคุณสมบัติภาษา OWL ที่ใช้ในงานวิจัยนี้ เช่น owl:equivalentProperty, owl:InverseOf และ owl:hasValue เพื่ออนุมานข้อมูลจากความสัมพันธ์ให้เรียนรู้คำตอบได้มากขึ้น ดังนี้

owl:equivalentProperty เป็นการเชื่อมพรอปเพอร์ตีที่มีลักษณะที่สอดคล้องกัน เช่น ดาต้าไทป์พรอปเพอร์ตี “hasScientificName” กับดาต้าไทป์พรอปเพอร์ตี “มีชื่อวิทยาศาสตร์” เป็นดาต้าไทป์พรอปเพอร์ตีที่สอดคล้องกัน ซึ่งกำหนดให้พรอปเพอร์ตีทั้งสองเป็น owl:equivalentProperty กัน ทำให้เครื่องมืออนุมานเข้าใจว่าพรอปเพอร์ตีทั้งสองเป็นพรอปเพอร์ตีเดียวกัน ดังภาพที่ 3

```
<owl:DatatypeProperty
rdf:about="&or;hasScientificName">
      <rdfs:domain rdf:resource="&or;Vegetables"/>
      <owl:equivalentProperty rdf:resource="&or;มีชื่อวิทยาศาสตร์"/>
      <rdfs:range rdf:resource="&xsd;string"/>
   </owl:DatatypeProperty>
```

ภาพที่ 3: ตัวอย่างการใช้ owl:equivalentProperty ระดับสคีม่า

```
Query:
prefix or: <http://202.28.94.53/organic#>
SELECT ?ScientificName
where {
        or:V001 or:มีชื่อวิทยาศาสตร์ ?ScientificName.
}
Output Query Inferred:
Solanum lycopersicum L
```

ภาพที่ 4: ตัวอย่างสปาเคิล และผลสืบค้นของ owl:equivalentProperty

จากออนโทโลยีในภาพที่ 2 มีการกำหนดพรอปเพอร์ตี “hasScientificName” ในคลาสผัก แต่ยังมีการใช้คุณสมบัติ owl:equivalentProperty กับพรอปเพอร์ตี “มีชื่อวิทยาศาสตร์” ทำให้สามารถสืบค้นข้อมูลในออนโทโลยี โดยใช้พรอปเพอร์ตี “มีชื่อวิทยาศาสตร์” ได้ด้วยเช่นกัน เช่น สืบค้น “V001” หรือมะเขือเทศมีชื่อวิทยาศาตร์อะไร ทำให้ระบบแสดงคำตอบการสืบค้นแบบอนุมานคือ Solanum lycopersicum L ดังภาพที่ 4

- owl:InverseOf ในออนโทโลยีผักอินทรีย์ได้กำหนดอ็อบเจกต์พรอปเพอร์ตี เช่น “appropriateVeget” ให้มีคุณสมบัติเป็น owl:InverseOf กับ “appropriateWithSoil” ซึ่งเครื่องมืออนุมานจะทำการสลับค่าโดเมนกับเรนจ์ของ “appropriateVeget” เพื่อกำหนดให้กับ “appropriateWithSoil” แทน ดังภาพที่ 5

```
<owl:ObjectProperty rdf:about="&or;appropriateVeget">
      <rdfs:domain rdf:resource="&or;Soil"/>
      <rdfs:range rdf:resource="&or;Vegetables"/>
      <owl:inverseOf
rdf:resource="&or;appropriateWithSoil"/>
      <owl:equivalentProperty rdf:resource="&or;เหมาะสมกับผัก"/>
   </owl:ObjectProperty>
```

ภาพที่ 5: ตัวอย่างการใช้ owl:InverseOf ระดับสคีม่า

```
Query:
prefix or: <http://202.28.94.53/organic#>
SELECT ?Soil
where {
        or:V001 or:appropriateWithSoil ?Soil.
}
Output Query Inferred:
Soil003 (ดินร่วนปนทราย)
```

ภาพที่ 6: ตัวอย่างสปาเคิล และผลสืบค้นของ owl:InverseOf

การใช้คุณสมบัติ owl:inverseOf จะทำให้สามารถสืบค้นความสัมพันธ์ของข้อมูลผ่านพรอปเพอร์ตี “appropriateWithSoil” ได้ ถึงแม้จะไม่มีการกำหนดความสัมพันธ์ของพรอปเพอร์ตีนี้อยู่ในออนโทโลยี เช่น สืบค้น “V001” หรือมะเขือเทศมีความเหมาะสมกับดินชนิดใด โดยสืบค้นผ่านพรอปเพอร์ตีที่มีชื่อว่า “appropriateWithSoil” ทำให้ระบบสามารถแสดงคำตอบอนุมานคือ Soil003 (ดินร่วนปนทราย) ดังภาพที่ 6

owl:hasValue เป็นคุณสมบัติช่วยจำกัดขอบเขตข้อมูลที่เป็นค่าของคลาส ซึ่งค่าข้อมูลของคุณสมบัตินี้มีค่าที่ไม่เหมือนกันเมื่อถูกนำไปใช้กับคลาสที่แตกต่างกัน โดยเป็นการกำหนดค่าข้อมูลที่อยู่ในระดับอินสแตนด์ (instance-level) เท่านั้น เช่น คลาสแมลงจำพวกกัดกินใบ (Leaf feeder Class) มีการกำหนดค่าข้อมูลผ่านพรอปเพอร์ตี “hasSymptom” ให้มีค่าข้อมูลเป็น “S004” หรือกัดกินใบ ดังภาพที่ 7

```
<owl:Class rdf:about="&or;LeafFeeder">
    <rdfs:subClassOf rdf:resource="&or;Insect"/>
    <rdfs:subClassOf>
      <owl:Restriction>
        <owl:onProperty
rdf:resource="&or;hasSymptom"/>
        <owl:hasValue rdf:resource="&or;S004"/>
      </owl:Restriction>
    </rdfs:subClassOf>
  </owl:Class>
```

ภาพที่ 7: ตัวอย่างการใช้ owl:hasValue ระดับสคีม่า

Query:

```
prefix or: <http://202.28.94.53/organic#>
SELECT ?Symptom
where {
        or:P007 or:hasSymptom ?Symptom.
}
```

Output Query Inferred:

S004 (กัดกินใบ)

ภาพที่ 8: ตัวอย่างสปาเคิล และผลสืบค้นของ owl:hasValue

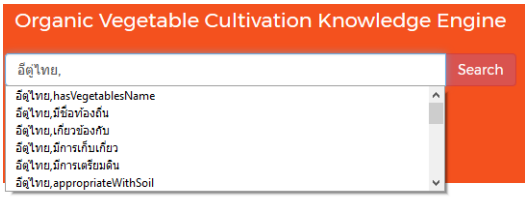
เป็นการใช้คุณสมบัติ owl:hasValue ในการกำหนดค่าข้อมูลให้กับสมาชิกทุกตัวของคลาสแมลงจำพวกกัดกินใบ โดยแสดงพรอปเพอร์ตี “hasSymptom” ดังนั้นจากภาพที่ 8 เป็นการสืบค้นข้อมูล “P007” หรือหนอนกระทู้ผัก ซึ่งอยู่ในคลาสแมลงจำพวกกัดกินใบว่ามีลักษณะอาการ (hasSymptom) อย่างไร ทำให้ระบบสามารถแสดงคำตอบอนุมานคือ S004 (กัดกินใบ)

4. ผลการดำเนินงาน

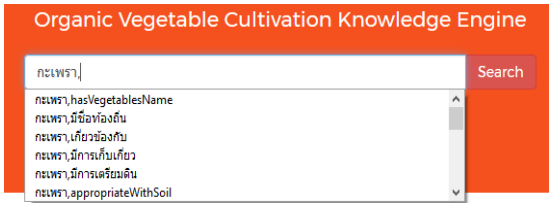
สำหรับผลการทดลองของงานวิจัยนี้ได้ข้อมูลมาจากผู้เชี่ยวชาญทางด้านเกษตรอินทรีย์ และคู่มือการเพาะปลูกผักอินทรีย์ [7] นอกจากนี้ได้พัฒนาออนโทโลยีเพาะปลูกผักอินทรีย์ด้วยโปรแกรมโปรทีเจ (Protégé) เวอร์ชัน 4.3.1 [8] และพัฒนาเว็บแอพพลิเคชั่นในการสืบค้นข้อมูลเพาะปลูกผักอินทรีย์แบบสื่อความหมายด้วยเน็ตบีนส์ (NetBeans) เวอร์ชั่น 8.2 และอาปาเช่จีน่า (Apache Jena) ภายใต้อาปาเช่ทอมแคท (Apache Tomcat) เวอร์ชั่น 7.0.73 เพื่อเป็นเว็บเซิร์ฟเวอร์ (Web Server) สำหรับเจเอสพี (Java Server Pages: JSP) และนำเสนอตัวอย่างผลสืบค้นสื่อความหมายแบบทริปเปิล โดยใช้ออโต้-คอมพลีต และแสดงผลการแก้ไขปัญหาคำพ้องความหมายของชื่อผัก ดังนี้

4.1 ผลสืบค้นสื่อความหมายแบบทริปเปิล

จากการเพิ่มเติมคุณสมบัติของ OWL ให้กับออนโทโลยีทำให้เครื่องมืออนุมานสามารถอนุมานความสัมพันธ์ต่าง ๆ เพิ่มเติมจากที่กำหนดไว้ในออนโทโลยี ซึ่งงานวิจัยนี้ได้พัฒนาเว็บแอปพลิเคชันสำหรับผู้ใช้งานเพื่อช่วยในการสืบค้นข้อมูลการเพาะปลูกผักอินทรีย์แบบออโต้คอมพลีต โดยใช้เครื่องมืออนุมานดึงเอาข้อมูลในลักษณะทริปเปิลออกมาทั้งหมดที่อนุมานได้จากออนโทโลยีเก็บไว้ในแหล่งจัดเก็บข้อมูล RDF เพื่อรองรับการสืบค้นแบบสื่อความหมายในรูปแบบออโต้คอมพลีต ดังภาพที่ 9 และ 10



ภาพที่ 9: ส่วนต่อประสานกับผู้ใช้ของการสืบค้นแบบทริปเปิล โดยใช้ออโต้คอมพลีต กรณีสืบค้นอีตู่ไทย



ภาพที่ 10: ส่วนต่อประสานกับผู้ใช้ของการสืบค้นแบบทริปเปิลโดยใช้ออโต้คอมพลีต กรณีสืบค้นกะเพรา

จากภาพที่ 9 และ 10 มีการเก็บข้อมูลเฉพาะกะเพราลงในออนโทโลยี แต่สามารถสืบค้นคำพ้องความหมายของกะเพราคืออีตู่ไทยได้ ซึ่งทำให้ผู้ใช้งานสามารถสืบค้นทริปเปิลได้ทั้งกะเพรา และอีตู่ไทย



ภาพที่ 11: ส่วนต่อประสานกับผู้ใช้ในลักษณะถามตอบ

จากภาพที่ 11 เป็นตัวอย่างแสดงผลลัพธ์ในลักษณะถามตอบจากการสืบค้นแบบทริปเปิล โดยใช้ออโต้คอมพลีต ซึ่งมีคำถามว่ากะเพรามีชื่อท้องถิ่นอะไรบ้าง หรือแสดงในรูปแบบของทริปเปิลคือ “กะเพรา,มีชื่อท้องถิ่น” โดยระบบแสดงคำตอบของคำถามคือ อีตู่ไทย, กอมก้อ, อีตู่ข้า, กะเพราขาว, กะเพราขน, กะเพราแดง และกอมก้อดง

4.2 ผลประเมินประสิทธิภาพของการสืบค้น

การวัดประสิทธิภาพด้วยวิธีสืบค้นข้อมูลเชิงความหมายโดยใช้ค่าความถูกต้อง (Precision) ดังสมการที่ 1 ค่าความครบถ้วน (Recall) ดังสมการที่ 2 และค่าความแม่นยำ (F-Measure) ดังสมการที่ 3 สามารถแสดงสมการได้ [9] ดังนี้

สมการที่ 1 $$\text{Precistions} = \frac{A}{A+B}$$

สมการที่ 2 $$\text{Recall} = \frac{A}{A+C}$$

สมการที่ 3 $$\text{F-Measure} = \frac{2(\text{Precistion} \times \text{Recall})}{\text{Precistion}+\text{Recall}}$$

จากสมการทั้ง 3 มีรายละเอียดของตัวแปร ดังนี้ A คือ จำนวนข้อมูลที่สืบค้นได้ และถูกต้องตามความสนใจ, B คือ จำนวนข้อมูลที่ถูกต้องตามความสนใจ แต่ไม่ถูกสืบค้น และ C คือ จำนวนข้อมูลที่สืบค้นได้แต่ไม่ถูกต้องตามความสนใจ

จากการทดลองสืบค้นจำนวน 50 คำถาม และนำไปเทียบกับคำตอบของผู้เชี่ยวชาญทางด้านเกษตร ปรากฏว่ามีค่าความถูกต้องเท่ากับ 97 เปอร์เซ็นต์ ค่าความครบถ้วนข้อมูลเท่ากับ 100 เปอร์เซ็นต์ และค่าความแม่นยำเท่ากับ 98 เปอร์เซ็นต์ ดังนั้นการพัฒนาออนโทโลยีเพื่อการสืบค้นสื่อความหมายสำหรับการเพาะปลูกผักอินทรีย์มีประสิทธิภาพอยู่ในระดับที่ดีมาก ถึงแม้ว่าค่าความถูกต้องมีประสิทธิภาพที่สูง แต่ก็ยังน้อยกว่า 100 เปอร์เซ็นต์ แสดงว่าระบบยังมีข้อผิดพลาดในการดึงคำตอบจากระบบออกมาแสดง ซึ่งข้อมูลที่ไม่ถูกแสดงอาจจะเกิดมาจากการเชื่อมโยงข้อมูลระหว่างพรอปเพอร์ตี

5. สรุปและข้อเสนอแนะ

งานวิจัยนี้ได้นำเสนอการแก้ไขปัญหาการสืบค้นแบบสื่อ-ความหมายแบบทริปเปิล โดยใช้เทคนิคออโต้คอมพลีต และสืบค้นเชิงภาษาท้องถิ่นเพื่อแก้ไขปัญหาของคำพ้องความหมาย โดยการพัฒนาออนโทโลยีเพาะปลูกผักอินทรีย์ในประเทศไทย ซึ่งมีผลการวัดค่าความแม่นยำอยู่ที่ 98 เปอร์เซ็นต์

อย่างไรก็ตามงานวิจัยนี้รองรับการสืบค้นแบบทริปเปิลในลักษณะออโต้คอมพลีต แต่ยังไม่สามารถทำการสืบค้นในลักษณะของภาษาธรรมชาติ (Natural Language) ของมนุษย์ได้ ดังนั้นในอนาคตผู้วิจัยอาจพัฒนาการสืบค้นในลักษณะของภาษาธรรมชาติ เพื่อรองรับการสืบค้นเชิงความหมายของการเพาะปลูกผักอินทรีย์ในประเทศไทย

เอกสารอ้างอิง

[1] A. Chougule, V. Kumar Jha and D. Mukhopadhay. “Ontology Based System for Pests and Disease Management of Grapes in India” *IEEE 6th International Conference on Advanced Computing (IACC 2016)*, Bhimavaram, India, February 27-28, 2016.

[2] ขนิษญา กุลนาวิน, ธนิดา วงศ์กาฬสินธุ์ และ นพรัตน์ ยติกร, “การออกแบบออนโทโลยีข้าวไทย” *วารสารนเรศวรพะเยา* ปีที่ 7 ฉบับที่ 1 ม.ค.-เม.ย. 2557 หน้า 76-84.

[3] โรสริน อัคนิจ, ธนิต พุทธพงษ์ศิริพร, น้ำฝน ลำดับวงศ์ และ อารีย์ ธัญกิจจานุกิจ, “การพัฒนาออนโทโลยี เพื่อการจัดการความรู้ด้านการแปรรูปข้าว” *วารสารเกษตร* ปีที่ 27 ฉบับที่ 3 ต.ค 2554 หน้า 267-274.

[4] Jens Lehmann and Johanna Völker. (2014). *An Introduction to Ontology Learning*. Berlin: AKA Heidelberg.

[5] Jeff Heflin. *An Introduction to the OWL web Ontology language*. Retrieved August 10, 2016, from www.cse.lehigh.edu/~heflin/IntroToOWL.pdf.

[6] World Wide Web Consortium (W3C), *Semantic Web*. Retrieved August 10, 2016, from https://www.w3.org/standards/semanticweb.

[7] อธิชาติ ศรีสะอาด และ เกรียงไกร ยอดชมภู. (2555). *คู่มือการเพาะปลูกผักอินทรีย์*. กรุงเทพฯ : นาคา อินเตอร์มีเดีย.

[8] Stanford University, *Protégé*. Retrieved October 5, 2016, from http://protege.stanford.edu.

[9] Creighton University Health Sciences Library. *Measuring Search Effectiveness*. Retrieved November 3, 2016, from. https://www.creighton.edu/fileadmin/.../Searching_-_Recall_Precision.pdf.

ปัจจัยแห่งความสำเร็จของการพัฒนาซอฟต์แวร์แบบต่างสถานที่
A Key Success Factors for Distributed Software Development

สุดารัตน์ วรเวทวิชัย (Sudarat Worawetwichai)[1] เทวา คำปาเชื้อ (Tawa Khampachua)[2]
และนวพร วิสิฐพงศ์พันธ์ (Nawaporn Wisitpongphan)[3]
[1,2]ภาควิชาคอมพิวเตอร์ศึกษา คณะครุศาสตร์อุตสาหกรรม [3]ภาควิชาการสื่อสารข้อมูลและเครือข่าย คณะเทคโนโลยีสารสนเทศ
มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
[1]Oilsudarat.w@gmail.com, [2]Tawa.k@fte.kmutnb.ac.th, [3]Nawaporn.w@it.kmutnb.ac.th

บทคัดย่อ

ในการพัฒนาซอฟต์แวร์ที่ทีมงานโครงการเดียวกันไม่เคยรู้จักกันมาก่อนมาทำงานร่วมกันแบบแยกกันทำงานต่างสถานที่เป็นประเด็นที่นักวิจัยทางด้านวิศวกรรมซอฟต์แวร์ต่างให้ความสำคัญ เพื่อให้สามารถทำงานร่วมกันได้เร็วที่สุดเพื่อให้งานสำเร็จลุล่วง ซอฟต์แวร์มีคุณภาพ ดำเนินงานตรงระยะเวลาที่กำหนด และใช้งบประมาณไม่เกินที่กำหนดต้องอาศัยปัจจัยหลายอย่างด้วยกัน ดังนั้นการวิจัยในครั้งนี้จึงได้ทำการศึกษาโครงการพัฒนาซอฟต์แวร์ที่มีการทำงานร่วมกันแบบต่างสถานที่เพื่อวิเคราะห์หาปัจจัยแห่งความสำเร็จของโครงการ โดยได้มีการดำเนินการเก็บรวบรวม ศึกษาข้อมูลจำนวน 6 โครงการ สัมภาษณ์ผู้ที่เกี่ยวข้อง และวิเคราะห์ข้อมูลที่ได้รับ เพื่อจำแนกว่าปัจจัยอะไรที่ก่อให้เกิดความสำเร็จในการทำงานร่วมกัน และปัจจัยอะไรที่ทำให้โครงการล้มเหลว เพื่อนำเอาผลที่ได้จากการศึกษาไปใช้ในการพัฒนากรอบแนวทางการทำงานร่วมกันของการพัฒนาซอฟต์แวร์แบบต่างสถานที่ต่อไป

คำสำคัญ: การพัฒนาซอฟต์แวร์แบบต่างสถานที่ การพัฒนาซอฟต์แวร์ร่วมกันแบบกระจาย วิศวกรรมซอฟต์แวร์ ปัจจัยแห่งความสำเร็จ การทำงานร่วมกัน

Abstract

Distributed software development project, which involves team members that work in different locations and have never before met or known one another, has long been an issue that researchers in software engineering try to solve. The team has to find the best way to collaborate in order to deliver the product efficiently in a timely manner and stay under the budget. In this work, we analyzed different collaboration approaches used by multiple distributed software development teams working in different locations on six different projects. Comprehensive interview and work output from each project were analyzed in order to identify key factors which cause the project to fail. Lesson learned from this study will later be used to design collaborative framework for distributed software development.

Keyword: Distributed Software Development, DSD, Software Engineering, Key Success Factors, Collaboration.

1. บทนำ

1.1 หลักการและเหตุผล

ในกระบวนการการพัฒนาซอฟต์แวร์ ต้องใช้คนที่มีความชำนาญหลายด้าน ซอฟต์แวร์สำหรับงานบางอย่าง มีผู้สร้างไว้จำหน่ายแล้ว แต่ถ้าซอฟต์แวร์ที่สั่งให้คอมพิวเตอร์ทำงานตามที่ต้องการไม่มีจำหน่าย ก็ต้องจ้างคนมาพัฒนาซอฟต์แวร์ให้ แต่ไม่ใช่จ้างเฉพาะคนเขียนโปรแกรมเท่านั้น ต้องจ้างผู้เชี่ยวชาญในการวิเคราะห์ ออกแบบ และทดสอบ มาร่วมมือกัน อีกทั้งต้องเลือกกระบวนการการพัฒนา ที่เหมาะสมกับลักษณะของซอฟต์แวร์ด้วย จึงจะได้ซอฟต์แวร์ที่ดี เสร็จตรงตามเวลาที่กำหนด และราคาไม่แพงเกินไป [1]

การใช้คำว่า "การพัฒนาซอฟต์แวร์" มีความหมายมากกว่า "การเขียนโปรแกรม" เพราะมีความซับซ้อนมากกว่า ในการ

พัฒนาซอฟต์แวร์ นักวิเคราะห์ที่ไม่ใช่โปรแกรมเมอร์ มักจะมาร่วมศึกษาวิเคราะห์ความซับซ้อนของโจทย์ปัญหา โดยจะมีเอกสารประกอบการดำเนินการ และมีการประสานงาน เพื่อให้คนหลายคนสามารถทำงานร่วมกันได้ ซึ่งกระบวนการพัฒนาซอฟต์แวร์ส่วนใหญ่ในปัจจุบันเป็นการพัฒนาแบบร่วมกลุ่มกันของบุคคล ทีมงานที่พัฒนาจะต้องอยู่ในสถานที่เดียวกัน เช่น สำนักงาน แต่ในทศวรรษที่ผ่านมา มีการพัฒนาซอฟต์แวร์แบบที่เรียกว่า Offshore ซึ่งเป็นการพัฒนาซอฟต์แวร์แบบทีมงานกระจายกันอยู่ เช่น การพัฒนาที่ประเทศอื่น การทำงานที่บ้านไม่ต้องไปหน้างาน การทำงานร่วมกันของบุคคลที่อยู่ในโซนเวลาที่แตกต่างกัน เป็นต้น สำหรับในบริบทของประเทศไทย มีการพัฒนาซอฟต์แวร์ที่ไม่ต้องทำงานในสถานที่เดียวกัน เนื่องจาก บุคคลในทีมทำงานอยู่คนละจังหวัด หรือมีปัญหาในการเดินทาง บริษัทพัฒนาซอฟต์แวร์ส่วนใหญ่เป็นบริษัทเล็กๆ มีโครงการที่ทำไม่ต่อเนื่อง โอกาสอยู่รอดยาก หลายบริษัทจึงลดค่าใช้จ่ายโดยการไม่เช่าสำนักงาน ไม่จ้างพนักงานประจำ จ้างฟรีแลนซ์ในการทำงาน ซึ่งการทำงานในรูปแบบนี้มีข้อเสียที่อาจจะกระทบต่องาน ไม่ว่าจะเป็น การสื่อสารที่ผิดพลาด เช่น การสื่อสารระหว่างนักวิเคราะห์ระบบกับนักพัฒนาระบบ การพูดคุยแสดงความคิดเห็นที่ไม่ตรงกัน ไม่ตรงกับความต้องการของลูกค้า เป็นต้น นอกจากนี้ยังมีเรื่องของการติดตามงานที่เป็นไปได้ยาก เช่น ฝั่งของผู้บริหารจัดการโครงการไม่สามารถติดตามความก้าวหน้าของทีมพัฒนาระบบได้ เนื่องจากไม่เห็นงานจริง ไม่เห็นการทำงานของอีกฝ่าย ทำให้เกิดความล่าช้าและความเสียหายกับโครงการได้

1.2 วัตถุประสงค์

วัตถุประสงค์เพื่อศึกษากรณีศึกษาของโครงการที่ดำเนินการพัฒนาซอฟต์แวร์ร่วมกันแบบต่างสถานที่ และนำเสนอปัจจัยแห่งความสำเร็จในการพัฒนาซอฟต์แวร์แบบต่างสถานที่

2. วรรณกรรมที่เกี่ยวข้อง

2.1 วงจรการพัฒนาระบบ

กระบวนการพัฒนาซอฟต์แวร์ (Software Development Process) เมื่อทำหน้าที่เป็นกรอบในการสร้างผลิตภัณฑ์ใดๆ แล้ว จะสามารถเรียก “กระบวนการ” ได้อีกอย่างหนึ่งว่า “วงจรชีวิตหรือวัฏจักรชีวิต” (Life Cycle) ของผลิตภัณฑ์นั้น กระบวนการพัฒนาซอฟต์แวร์จึงเรียกได้อีกอย่างหนึ่งว่า วงจรการพัฒนาระบบ (System Development Life Cycle) [2] ซึ่งวงจรการพัฒนาระบบ หรือกรอบแนวทางการพัฒนาซอฟต์แวร์นั้น ในปัจจุบันมีหลายแนวทาง ซึ่งแต่ละแนวทางก็มีข้อดีข้อเสียที่แตกต่างกันไป ตัวอย่างเช่น Waterfall Model เป็นแบบจำลองพื้นฐานที่ประกอบด้วยขั้นตอนการดำเนินงานที่เรียงต่อเนื่องกันเป็นลำดับ เหมาะกับการพัฒนาซอฟแวร์ขนาดใหญ่ ต้องรอให้การออกแบบทั้งหมดเสร็จสมบูรณ์ก่อน จึงจะเริ่มเขียนโปรแกรมได้ และไม่สามารถรองรับความต้องการของลูกค้าที่เปลี่ยนแปลงอย่างรวดเร็วได้ [3] หรือ แบบจำลองกระบวนการพัฒนาซอฟต์แวร์แบบ Rational Unified Process (RUP) แบ่งการทำงานออกเป็นเฟส แสดงกิจกรรมที่ต้องดำเนินงาน และหลักในการดำเนินการ ซึ่งเป็นแบบจำลองที่มีเป้าหมายในการทำให้กระบวนการในขั้นตอนนั้นๆ สมบูรณ์ที่สุด มีการวนและทำซ้ำเรื่อยๆ จนขั้นตอนสมบูรณ์ โดยจะมีเป้าหมายในแต่ละขั้นตอนว่าควรจะได้ผลอะไรออกมาบ้าง [4]

2.2 การพัฒนาซอฟต์แวร์แบบต่างสถานที่

การพัฒนาซอฟต์แวร์แบบต่างสถานที่ (Distributed Software Development: DSD) เป็นการดำเนินการพัฒนาซอฟต์แวร์ร่วมกันของบุคคลที่ทำงานกระจายกันต่างสถานที่ การพัฒนาซอฟต์แวร์แบบดังเดิม ที่ทีมงานนั่งทำงานในที่เดียวกัน สังกัดองค์กรเดียวกัน อยู่ในประเทศเดียวกัน นั้นเริ่มจะมีต้นทุนและค่าใช้จ่ายที่สูงขึ้นและศักยภาพในการแข่งขันน้อยลง [5] ความก้าวหน้าทางเศรษฐกิจ สื่อที่สนับสนุน และความกดดันจากค่าใช้จ่ายที่สูงขึ้น ทำให้เกิดการลงทุนในการพัฒนาซอฟต์แวร์ร่วมกันแบบต่างสถานที่ ตามที่ Brooks [6] ได้บอกไว้ ซอฟต์แวร์มีคุณสมบัติหลายอย่างที่ทำให้มันเหมาะสำหรับวิธีการพัฒนาซอฟต์แวร์ร่วมกันแบบต่างสถานที่ ซอฟต์แวร์สามารถที่จะทำซ้ำ ส่งถ่าย ปรับปรุงแก้ไข และนำมาใช้ได้ไม่จำกัด เหมือนกับไม่มีค่าใช้จ่าย แต่ก็ยังเป็นเรื่องของการเปลี่ยนแปลง สิ่งที่แตะต้องไม่ได้ และอาจกลายเป็นความซับซ้อนที่อาจจะเกิดขึ้นได้ จากในบางคำพูดของเขา Herbsleb [7] กล่าวว่า มันไม่ได้เป็นเรื่องแปลกที่การพัฒนาซอฟต์แวร์จะมีทีมงานกระจายอยู่คนละสถานที่ บางครั้งก็คนละทวีป การเติบโตที่เพิ่มขึ้น ทำให้แข่งขันสูงขึ้น บริษัทจึงได้หันมาใช้วิธีการพัฒนาซอฟต์แวร์แบบต่างสถานที่ ซึ่งถือเป็นการที่จะทำให้ต้นทุนต่ำ

และบริษัทได้ขยายไปในตลาดโลก การใช้วิธีการพัฒนาซอฟต์แวร์ร่วมกันแบบต่างสถานที่นี้ทำให้เกิดผลกระทบในตลาดไม่ว่าจะเป็นการจัดจำหน่าย การผลิต (การเขียนโปรแกรม) การทดสอบ และการส่งมอบซอฟต์แวร์ให้กับลูกค้า

ดังนั้น การพัฒนาซอฟต์แวร์แบบต่างสถานที่ ที่ได้เห็นในช่วงที่ผ่านมา จึงเป็นทางเลือกในการพัฒนาซอฟต์แวร์ ซึ่งเป็นที่นิยมในทศวรรษที่ผ่านมา โดยมีลักษณะของการทำงานร่วมกัน การให้ความร่วมมือกันระหว่างหน่วยงานและองค์กร เป็นการสร้างกลุ่มของนักพัฒนาที่ทำงานคนละเมือง คนละประเทศต่างสถานที่กัน [8] ซึ่งบริษัทที่ทำงานในสภาพแวดล้อมที่มีการพัฒนาแบบต่างสถานที่ ไม่ว่าจะเป็นด้านเทคนิค คน องค์กร ปัญหาการสื่อสาร การจัดการ ความสัมพันธ์ระหว่างสิ่งต่างๆ นั้นถือว่าการพัฒนาซอฟแวร์เป็นกระบวนการที่ซับซ้อน และ การพัฒนาซอฟต์แวร์ร่วมกันแบบต่างสถานที่ ช่วยให้เกิดความท้าทายไม่ว่าจะเป็นทางด้านกายภาพ การกระจายตัวทำงานต่างสถานที่ชั่วคราว และวัฒนธรรมที่แตกต่างกัน เมื่อโครงการมีการพัฒนาที่เกี่ยวข้องกับประเทศอื่นๆ ทำให้เป็นที่รู้จักกันในการพัฒนาซอฟต์แวร์ระดับโลก การพัฒนาซอฟต์แวร์แบบต่างสถานที่ก็เกิดบนพื้นฐานการพัฒนาซอฟต์แวร์แบบดังเดิม ได้แก่ การจัดทำความต้องการของระบบ วิเคราะห์ระบบ ออกแบบ พัฒนา และทดสอบระบบ บางอย่างอาจจะต้องทำ ณ สถานที่ทำงานของลูกค้า (On-Site) หรือนอกสถานที่ทำงานของลูกค้า (Off Shore) หรือทั้งสองแบบ (Multi-Site) [9]

3. วิธีการดำเนินการวิจัย

การศึกษาใช้กลุ่มตัวอย่างของโครงการที่มีการพัฒนาซอฟต์แวร์ร่วมกันแบบต่างสถานที่ 6 โครงการ ซึ่งประกอบด้วยโครงการต่างๆ ดังตารางที่ 1

ตารางที่ 1: รายชื่อโครงการที่นำมาเป็นกรณีศึกษา

ลำดับ	โครงการ	ผลลัพธ์
C1	Sourcing Control	สำเร็จ
C2	JV Online	สำเร็จ
C3	PDT Temper Glass	สำเร็จ แต่ไม่ได้ Go live
C4	Pawn Academy	สำเร็จ
C5	Safety Incident	ถูก terminate
C6	E-dairy	เลยกำหนดการส่งมอบงาน

การศึกษาได้ใช้กรณีศึกษาที่เป็นโครงการที่มีมูลค่าต่ำกว่า 10 ล้านบาท และเป็นโครงการที่ใช้ระยะเวลาในการพัฒนาไม่เกิน 1 ปี ใช้วิธีการติดตามโครงการและสัมภาษณ์ผู้มีส่วนเกี่ยวข้อง โดยให้เล่ารายละเอียดของโครงการ ซึ่งใช้รายการหัวข้อคำถามดังตารางที่ 2

ตารางที่ 2: รายการหัวข้อคำถาม

ลำดับ	คำถาม
1	งบประมาณและระยะเวลาในการดำเนินงาน
2	เครื่องมือที่ใช้ในการดำเนินงาน (ติดตามงาน)
3	เครื่องมือที่ใช้ในการติดต่อสื่อสารระหว่างบุคคล
4	เอกสารประกอบการดำเนินงาน
5	ภาษาที่ใช้ในเอกสารและการติดต่อสื่อสาร
6	จำนวนทีมงาน
7	บทบาทหน้าที่ของผู้มีส่วนเกี่ยวข้อง
8	ปัจจัยที่คิดว่าทำให้โครงการสำเร็จ/ล้มเหลว

หลังจากเก็บรวบรวมข้อมูลในแต่ละโครงการ ก็ได้นำมาวิเคราะห์หาปัจจัยแห่งความสำเร็จ ในการทำงานของการพัฒนาซอฟต์แวร์ร่วมกันแบบต่างสถานที่ เพื่อหาปัจจัยแห่งความสำเร็จในการทำงานร่วมกันของการพัฒนาซอฟต์แวร์แบบต่างสถานที่

4. ผลการดำเนินงาน

การศึกษาใช้กลุ่มตัวอย่างของโครงการที่มีการพัฒนาซอฟต์แวร์ร่วมกันแบบต่างสถานที่ มีการกระจายงานกันทำ ปัจจัยด้านเครื่องมือนำเสนอเป็นเครื่องมือที่ใช้ดำเนินการและเครื่องมือที่ใช้ในการติดต่อสื่อสาร รายละเอียดเครื่องมือและงบประมาณของแต่ละโครงการดังตารางที่ 3

ตารางที่ 3: รายละเอียดเครื่องมือและงบที่เกี่ยวข้องของแต่ละโครงการ

โครงการ	งบประมาณ	เครื่องมือที่ใช้	การสื่อสาร
C1	300,000 บาท	Trello	Line E-mail การโทรศัพท์ การประชุม On-Site

ตารางที่ 3: รายละเอียดเครื่องมือและงบที่เกี่ยวข้องของแต่ละโครงการ

โครงการ	งบประมาณ	เครื่องมือที่ใช้	การสื่อสาร
C2	1,800,000 บาท	Weekly Report	Line E-mail การโทรศัพท์ การประชุม On-Site
C3	250,000 บาท	Trello (ไม่สำเร็จ)	Social Media E-mail การประชุม
C4	990,000 บาท	Google Sheet Trello (ไม่สำเร็จ)	Facebook การประชุม
C5	2,200,000 บาท	Dropbox Google Drive	Line E-mail การโทรศัพท์ การประชุม On-Site ทุกสัปดาห์
C6	1,790,000 บาท	Teamwork Google Sheet	Line E-mail การโทรศัพท์ การประชุม On-Site

ในโครงการ C1 มีการใช้ Trello ในการติดตามงานเพื่ออัพเดทความก้าวหน้า เช่นเดียวกับ โครงการ C3 และ C4 แต่โครงการ C3 และ C4 ไม่ได้รับประโยชน์จากเครื่องมือนี้ เนื่องจากทีมงานไม่ให้ความร่วมมือในการเข้าไปอัพเดทใน Trello ส่วนในโครงการ C6 มีการใช้ Teamwork แต่ทีมงานไม่ได้เข้าไปอัพเดทเช่นกัน จึงเปลี่ยนมาใช้ Google Sheet ในการอัพเดทความก้าวหน้าของงาน

การใช้เอกสารประกอบการดำเนินงานและภาษาที่ใช้ในการดำเนินงานก็มีความสำคัญด้วย เพราะโครงการที่จะประสบความสำเร็จได้บุคคลในทีมจะต้องมีความสามารถในการสื่อสารเข้าใจอย่างรวดเร็ว การใช้เอกสารของแต่ละโครงการแสดงในตารางที่ 4

ตารางที่ 4: รายละเอียดเอกสารและภาษาที่ใช้ของแต่ละโครงการ

โครงการ	เอกสารที่ส่งมอบให้ลูกค้า	ภาษาที่ใช้
C1	FD TD UAT Manual	ภาษาไทย
C2	FD TD SIT Unit Test UAT	ภาษาอังกฤษ
C3	SDD Manual	ภาษาไทย
C4	UAT Manual	ภาษาไทย
C5	SOW Weekly report Meeting	ภาษาอังกฤษ
C6	UAT	ภาษาไทย

ในโครงการที่ใช้ภาษาอังกฤษแต่ทีมงานต้องใช้ระยะเวลาในการทำความเข้าใจภาษา และใช้ระยะเวลาในการจัดทำเอกสารเป็นภาษาอังกฤษส่งผลให้งานล่าช้า และไม่ประสบความสำเร็จ แต่ในโครงการที่ใช้ภาษาอังกฤษและทีมงานสามารถใช้ภาษาอังกฤษได้ จะไม่ส่งผลต่อการดำเนินงาน

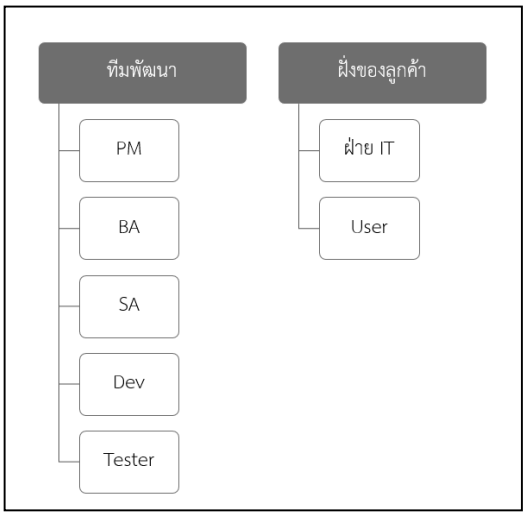
สำหรับจำนวนทีมงานพัฒนาในแต่ละโครงการมีจำนวนอยู่ระหว่าง 4 – 7 คน ผู้ที่มีส่วนเกี่ยวข้องในทีมที่มี 4 คน ได้แก่ ผู้บริหารจัดการโครงการ นักวิเคราะห์และออกแบบระบบ โปรแกรมเมอร์ และนักทดสอบระบบ ซึ่งเป็นทีมที่มีจำนวนน้อยที่สุด ส่วนที่เพิ่มมา อาจเป็นจำนวนโปรแกรมเมอร์หรือนักวิเคราะห์ที่เพิ่มขึ้นมา

ปัจจัยแห่งความสำเร็จในการพัฒนานอกเหนือจากสิ่งที่ได้กล่าวไปแล้ว สิ่งสำคัญก็คือบุคคลในทีมต้องมีความสามารถและความรับผิดชอบในการทำงาน ถ้าเป็นทีมงานที่เป็นคนเก่งทำงานรวมกัน ไม่ว่าจะทำงานแยกกันไกลแค่ไหนก็สามารถประสบความสำเร็จได้เนื่องจากเครื่องมือสำหรับสนับสนุนการทำงานในปัจจุบันมีมากมาย

5. อภิปรายผล

จากกระบวนการพัฒนาซอฟต์แวร์พื้นฐานโดยทั่วไปจะแบ่งส่วนงานหลักๆ ออกเป็น การวางแผน (Planning) การระบุความต้องการสำหรับการพัฒนาระบบ (Requirement Development) การวิเคราะห์และออกแบบระบบ (System Analysis and Design) การเขียนโปรแกรม (Implement) การทดสอบระบบ (Test) และการนำซอฟต์แวร์ที่ได้ขึ้นเซิร์ฟเวอร์เพื่อนำไปใช้งาน (Deploy)

บุคคลที่เกี่ยวข้องสำหรับการพัฒนาซอฟต์แวร์โดยทั่วไปจะประกอบด้วย ผู้บริหารโครงการ (Project Manager :PM) ผู้วิเคราะห์ธุรกิจ (Business Analysis : BA) ผู้วิเคราะห์และออกแบบระบบ (System Analysis and Design : SA) ผู้พัฒนาระบบหรือโปรแกรมเมอร์ (Developer : Dev) ผู้ทดสอบระบบ (Tester) และบุคคลที่อยู่ในฝั่งขององค์กรที่เป็นผู้ต้องการระบบ ได้แก่ ฝ่ายสารสนเทศขององค์กร (Information Technology : IT) และผู้ใช้งานซอฟต์แวร์ที่จะพัฒนาขึ้น (User) ดังภาพที่ 1



ภาพที่ 1: ผู้เกี่ยวข้องในการดำเนินการพัฒนาซอฟต์แวร์

สำหรับกระบวนการพัฒนาซอฟต์แวร์ร่วมกันแบบต่างสถานที่นั้นจะมีลักษณะที่เหมือนกับกระบวนการพัฒนาซอฟต์แวร์พื้นฐานโดยทั่วไป เพียงแต่ลดขั้นตอนของการนั่งทำงานร่วมกัน เผชิญหน้ากัน โดยนำเครื่องมือและเทคโนโลยีต่างๆ เข้ามาช่วยสนับสนุนการติดต่อสื่อสารแทน เพื่อเป็นการลดค่าใช้จ่ายหรือลดต้นทุนในการผลิตซอฟต์แวร์ และยังคงไว้ซึ่งคุณภาพของซอฟต์แวร์ซึ่งซอฟต์แวร์ที่ผลิตได้จะต้องตรงต่อความต้องการของผู้ใช้ ตรงตามระยะเวลาที่ได้ตั้งไว้ และตรงตามงบประมาณที่ได้ตั้งไว้ด้วย

ในขั้นตอนการ Planning หรือการวางแผนการพัฒนาซอฟต์แวร์ เป็นความรับผิดชอบของ PM ในการจัดทำ Project Plan เริ่มต้นจากเอกสาร TOR ซึ่งเป็นเอกสารที่จะระบุขอบเขตของซอฟต์แวร์ที่ต้องผลิต โดยเอกสาร TOR จะได้มาจากความต้องการของผู้ใช้ และฝ่าย IT ของลูกค้าได้ระบุขอบเขตมา เพื่อเป็นข้อกำหนดให้ PM นำมาพิจารณาวางแผน เพื่อทำ Project-Plan หลังจากที่ PM ได้จัดทำแผนการทำงานแล้ว จะต้องส่งมอบให้ลูกค้าหรือฝ่าย IT ของลูกค้า ไว้สำหรับติดตามงานให้เป็นไปตามแผนที่ได้วางไว้ และยังต้องส่งแผนที่ได้ให้กับทีมงานพัฒนาของตนเองด้วย อันได้แก่ BA, SA, Dev และ Tester ซึ่งแผนที่จะส่งให้กับภายในทีมจะต้องกำหนดเส้นตายของงานให้เร็วขึ้นกว่าแผนจริงที่ส่งให้ลูกค้า เพื่อเป็นการลดความเสี่ยงที่อาจเกิดขึ้นได้ หากมีฝ่ายใดฝ่ายหนึ่งติดปัญหา ทำงานไม่เสร็จหรืองานไม่เดิน เนื่องจากไม่ได้เผชิญหน้ากันจึงไม่สามารถรู้ได้ว่างานเป็นตามแผนหรือไม่ จึงต้องมีการกำหนดแผนให้เร็วกว่าเพื่อเผื่อเวลาสำหรับการตรวจสอบและแก้ไข

ในขั้นตอนการเก็บข้อมูลความต้องการของระบบ เป็นหน้าที่ของ BA ที่จะต้องเข้าไปคุยกับฝั่งของลูกค้าเพื่อหาข้อสรุปของความต้องการทั้งหมดออกเป็นข้อๆ ให้ครบถ้วน เพื่อที่จะเป็นข้อมูลสำหรับการวิเคราะห์และออกแบบระบบต่อไป โดยฝ่าย IT หรือ ผู้ใช้งาน จะต้องเป็นผู้ให้ข้อมูล และที่สำคัญจะต้องเป็นข้อมูลที่ถูกต้องครบถ้วนด้วย เพราะถ้าหากเกิดการเปลี่ยนแปลงก็จะส่งผลต่อระยะเวลาและงบประมาณในการพัฒนา เมื่อ BA เรียบเรียงข้อมูลความต้องการได้ทั้งหมดแล้วจะต้องส่งให้กับฝั่งของลูกค้ายืนยันความถูกต้อง จากนั้นจึงดำเนินการวิเคราะห์และออกแบบต่อไป ซึ่งผลลัพธ์ที่ได้จะถูกนำส่งเป็นข้อมูลให้กับ SA และ Dev เพื่อให้เป็นข้อมูลประกอบการทำงาน สำหรับขั้นตอนนี้จำเป็นจะต้องมีการพบกันระหว่างฝั่งลูกค้ากับฝั่งทีมพัฒนาอย่างน้อย 1 ครั้งเพื่อหาข้อสรุปของความต้องการระบบได้อย่างมีประสิทธิภาพ

ในขั้นตอนการวิเคราะห์และออกแบบระบบให้ได้มาซึ่งเอกสาร TD (Technical Design) และเอกสาร FD (Function Design) ที่เป็นรูปธรรม เพื่อให้การทำงานมีความเข้าใจที่ตรงกัน โดยเป็นหน้าที่ของ SA ในการจัดทำ โดยที่เอกสาร TD จะแสดงรายละเอียดเกี่ยวกับระบบ ว่ากระบวนการทำงานจริงเป็นอย่างไร มีกระบวนการในบ้างที่ต้องมีระบบมาเกี่ยวข้อง

แล้วระบบต้องมีขั้นตอนการทำงานอย่างไรบ้าง รวมถึงการวิเคราะห์ออกแบบโมดูลต่างๆของระบบ สิทธิ์การเข้าถึงของผู้ใช้งานแต่ละระดับ และการออกแบบฐานข้อมูลพร้อมคำอธิบายต่างๆ ส่วนเอกสาร FD จะแสดงหน้าจอระบบที่ได้ออกแบบไว้ซึ่งเป็นส่วนของ View (V) ในการเขียนโปรแกรมรูปแบบ MVC เพื่อให้รายละเอียดแบบเป็นภาพที่เห็นได้จริง พร้อมทั้งระบุคำอธิบายหน้าจอด้วย เมื่อได้เอกสารทั้งสองตัวจากการวิเคราะห์และออกแบบแล้ว จะต้องส่งให้ฝั่งของลูกค้ายืนยันความถูกต้องและยอมรับสิ่งที่ได้ออกแบบ เมื่อลูกค้าตกลงยืนยันเรียบร้อยเอกสารจะถูกส่งให้กับ Dev และ Tester เพื่อใช้ในการดำเนินงานต่อไป

ในขั้นตอนการ Implement หรือขั้นตอนการเขียนโปรแกรมเป็นหน้าที่ของโปรแกรมเมอร์ในการเขียนโปรแกรมตามที่ได้มีการออกแบบไว้ ซึ่งข้อมูลของระบบที่โปรแกรมเมอร์จะต้องพัฒนานั้นทราบมาตั้งแต่แผนการทำงาน ข้อมูลความต้องการและเอกสารการออกแบบ TD/FD แล้ว โปรแกรมเมอร์มีหน้าที่พัฒนาตามแผนที่กำหนดไว้ จากนั้นส่งให้ Tester ทดสอบเพื่อตรวจสอบความถูกต้องของระบบ ในขั้นตอนนี้อาจเป็นขั้นตอนที่ใช้ระยะเวลายาวนานที่สุด PM จะต้องควบคุมให้งานเป็นไปตามแผนที่วางไว้ด้วย เมื่อพัฒนาโมดูลใดโมดูลหนึ่งเสร็จเรียบร้อยแล้วก็จะต้องส่งให้ Tester ได้ตรวจสอบต่อไป

ในขั้นตอนการทดสอบ เป็นหน้าที่ของ Tester ที่จะต้องทดสอบความถูกต้องของระบบตามความต้องการของลูกค้าและตามสิ่งที่ได้ออกแบบไว้ เมื่อ Tester ได้รับข้อมูลของซอฟต์แวร์แล้ว Tester จะต้องเตรียมการสำหรับการทดสอบระบบ โดยอาจจะมีการพูดคุยสื่อสารระหว่าง โปรแกรมเมอร์และ Tester ในการทำงาน เพื่อแจ้งปัญหาข้อผิดพลาด และการแก้ไข นำส่งให้ทดสอบใหม่ เมื่อทำการทดสอบก็จะต้องมีการบันทึกข้อมูล หากเห็นว่าถูกต้องครบถ้วน ในส่วนของขั้นตอนการทดสอบนี้จะต้องมีการทำเอกสารสำหรับการทดสอบร่วมกับผู้ใช้งานจริง เพื่อให้ลูกค้ายอมรับระบบว่าเป็นระบบที่สมบูรณ์สามารถใช้งานได้จริง

และในขั้นตอนสุดท้าย ขั้นตอนการ Deploy หรือขั้นตอนการส่งมอบระบบให้กับลูกค้า จะเป็นการนำระบบที่เสร็จสมบูรณ์แล้วติดตั้งขึ้นบนเซิร์ฟเวอร์ของลูกค้า โดย PM มีหน้าที่ประสานงานกับฝั่ง IT ของลูกค้า แล้วให้ Dev ทำการส่งมอบระบบที่สมบูรณ์ที่ผ่านการทดสอบและยอมรับจากทั้งสองฝ่าย

6. สรุป

จากการศึกษากรณีศึกษาทั้งหกโครงการ ซึ่งนักพัฒนาอยู่ต่างสถานที่กัน และเป็นโครงการขนาดกลางทำให้ได้ทราบถึงปัจจัยแห่งความสำเร็จในการดำเนินโครงการ การจะพัฒนาซอฟต์แวร์ต่างสถานที่ให้ประสบผลสำเร็จ ขึ้นอยู่กับปัจจัยหลายอย่าง ไม่ว่าจะเป็นการบริหารจัดการ มีความต้องการของผู้ใช้ที่ชัดเจน มีการทำงานอย่างมีความรับผิดชอบ การประสานงานและสื่อสารอย่างมีประสิทธิภาพ บุคคลหรือทีมงานที่ร่วมกันพัฒนาซอฟต์แวร์เป็นปัจจัยสำคัญในการทำงานให้ออกมาบรรลุเป้าหมาย PM ต้องมีความสามารถในการบริหาร BA และ SA จะต้องวิเคราะห์ออกแบบได้อย่างถูกต้อง นักพัฒนาต้องเก่งมีความรู้ความสามารถ และ Tester ต้องมีความละเอียดรอบคอบทำให้ซอฟต์แวร์ที่พัฒนาเป็นที่ยอมรับของลูกค้าให้ได้ ภาษาที่ใช้ควรเป็นภาษาที่บุคคลในทีมสามารถสื่อสารได้อย่างเข้าใจ ทุกคนต้องมีส่วนร่วมและให้ความร่วมมือ วัฒนธรรมการทำงาน ภาษา และวิธีการ ต้องคล้ายคลึงกันเพื่อที่จะทำงานได้สำเร็จ ได้มาซึ่งซอฟต์แวร์ที่มีคุณภาพ ตรงตามความต้องการของลูกค้า ตรงตามระยะเวลาที่กำหนดในสัญญา และไม่เกินงบประมาณที่ตั้งไว้

เอกสารอ้างอิง

[1] ชนม์ชนก วีรวรรณ. "การพัฒนาซอฟต์แวร์" *สารานุกรมไทยสำหรับเยาวชนฯ* เล่มที่ 25 หน้า 123-145.

[2] กิตติ ภักดีวัฒนะกุล. "*วิศวกรรมซอฟต์แวร์ (SOFTWARE ENGINEERING)*". กรุงเทพฯ : เคทีพี คอมพ์ แอนด์ คอนซัลท์, [2550]

[3] นิตยา ปลีกะมล. (สืบค้นวันที่ 11 พฤศจิกายน 2559) โมเดลการพัฒนาซอฟต์แวร์ [ออนไลน์]. เข้าถึงได้ที่ : https://sites.google.com/site/nityaplikaml/home/model-kar-phathna-sxftwaer

[4] AuSawinz. (สืบค้นวันที่ 11 พฤศจิกายน 2559) Rational Unified process (RUP) คืออะไร [ออนไลน์]. เข้าถึงได้ที่ : https://ispying.blogspot.com/2013/10/rational-unified-process-rup.html

[5] Prikladnicki, R., Audy, J. "Comparing Offshore Outsourcing and the Internal Offshoring of Software Development: A Qualitative Study," *Proceedings of the Americas Conference on Information Systems (AMCIS)*, San Francisco, USA. 2009.

[6] Brooks, F. P. The Mythical Man-Month: Essays on Software. 1st Addison-Wesley Longman Publishing Co.,Inc. 1978.

[7] Herbsleb, J. D. Global Software Engineering: The Future of Socio-technical Coordination. *IEEE Computer Science*. P188-198. 2007.

[8] Meyer, B. The unspoken revolution in software engineering. *IEEE Computer. 2006.*

[9] Rodrigo G. C. Rocha, Catarina Costa, Cleyton Rodrigues, Ryan Ribeiro de Azevedo, Ivaldir H. Junior, Silvio Meira, Rafael Prikladnicki. "Collaboration Models in Distributed Software Development: a Systematic Review" *CLEI ELECTRONIC JOURNAL*, VOLUME 14, NUMBER 2, PAPER 1, AUGUST 2011

ระบบขอใช้และติดตามรถยนต์โดยใช้ Google Map กรณีศึกษา บริษัท ไปรษณีย์ไทย จำกัด

Car Reservation and Tracking System Using Google Map Case Study Thailand Post Co., Ltd

อัญชลี รุ่งเทียน (Unchalee Rungthian)[1] มาลีรัตน์ โสดานิล (Maleerat Sodanil)[2]
ภาควิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
[1]joy.enjoys@gmail.com, [2]maleerat.s@it.kmutnb.ac.th

บทคัดย่อ

จากการศึกษาปัญหาระบบการขอใช้บริการรถยนต์ บริษัทไปรษณีย์ไทย พบว่ามีการเก็บข้อมูลที่เป็นเอกสารและใช้เวลาการดำเนินงานแต่ละขั้นตอนนาน ไม่สามารถติดตามสถานะการใช้งานได้แบบเรียลไทม์ งานวิจัยนี้มีวัตถุประสงค์เพื่อแก้ปัญหาดังกล่าว โดยพัฒนาระบบการขอใช้บริการและติดตามรถยนต์ระหว่างปฏิบัติงาน พัฒนาเป็น 2 ส่วน คือ 1) เว็บแอปพลิเคชันผ่านเครือข่ายอินเทอร์เน็ตสำหรับการขอใช้บริการรถยนต์ 2) โมบายแอปพลิเคชันทำงานบนระบบปฏิบัติการแอนดรอยด์ สำหรับพนักงานขับรถเพื่อใช้ติดตามรถยนต์ในขณะปฏิบัติงาน ทั้งนี้ได้ประยุกต์ใช้จีพีเอสกับกูเกิลแมพเอพีไอ ในการค้นหาสถานที่ คำนวณระยะเวลา และเส้นทางจากบริษัทไปรษณีย์ไทย ไปยังปลายทาง ระบบสามารถระบุตำแหน่งรถยนต์ และแจ้งเตือนหากพบว่าใช้เวลาในการเดินทางมากกว่าปกติ นอกจากนี้ยังมีการนำเสนอรายงานสรุปต่างๆ เพื่อใช้ประกอบการประเมินจำนวนการใช้รถยนต์ ในการประเมินความพึงพอใจของระบบ โดยผู้ใช้งานที่เกี่ยวข้องกับระบบ ได้ค่าเฉลี่ย 4.20 ค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.43 ตามลำดับ สรุปได้ว่าระบบที่พัฒนาขึ้นมีประสิทธิภาพอยู่ในระดับดี สามารถนำไปใช้เพื่อลดเวลาการดำเนินงานในแต่ละขั้นตอนติดตามสถานะการปฏิบัติงานของพนักงานขับรถและรถยนต์ได้อย่างมีประสิทธิภาพ

คำสำคัญ: ระบบติดตาม จีพีเอส กูเกิลแมพเอพีไอ

Abstract

According to the current problem of car reservation process that is not going paperless and no tracking in real time, this research aims to develop a system of car reservation and tracking system using Google Map in order to solve this problem. The system is divided into 2 parts 1) Web Application for requester and approve manager 2) Mobile Application for drivers. Google Map is applied to use in order to identify the car location which we can calculate the distance between destination location and Thailand Post including the notification when the operation time is over from the initial calculation. The experimental result is given good satisfaction which average score 4.20, the standard deviation score was 0.43 In conclusion the developed system had effectively good level, could be reduce operation times each of procedures, tracking operating status of drivers and car effectively.

Keywords: Tracking System, GPS, Google Map API.

1. บทนำ

บริษัท ไปรษณีย์ไทย จำกัด เป็นหน่วยงานรัฐวิสาหกิจทำธุรกิจบริการไปรษณีย์มีหน้าที่จัดส่งสินค้าและข่าวสารต่างๆ ทางไปรษณีย์ทั้งในประเทศและนอกประเทศ โดยบริษัทฯ มีฝ่ายรถยนต์ขนส่ง มีหน้าที่ให้บริการรถยนต์ เฉพาะเจ้าหน้าที่บริษัทเฉพาะสำนักงานใหญ่ เพื่อเดินทางไปปฏิบัติภารกิจต่างๆ ปัจจุบันการขอใช้บริการรถยนต์จะต้องดำเนินการผ่านเอกสารแบบฟอร์มซึ่งเกิดปัญหาต่างๆ เช่น ใช้เวลานานแต่ละขั้นตอน การรายงานผลการขอรถล่าช้า และพนักงานขับรถบางครั้งไม่รู้

เส้นทาง ทำให้สิ้นเปลืองน้ำมันเชื้อเพลิง ใช้เวลาขับรถไม่ตรงตามเวลาที่กำหนด ผู้วิจัยจึงพัฒนาออกเป็น ระบบการขอใช้รถยนต์เป็นเว็บแอปพลิเคชัน และโมบายแอปพลิเคชันเพื่อติดตามรถยนต์ ระบบปฏิบัติการแอนดรอยด์ และใช้ Google MAP API ค้นหาสถานที่ เลือกเส้นทาง คำนวณระยะเวลา และแสดงตำแหน่งของรถยนต์

2. ทฤษฎีและวรรณกรรมที่เกี่ยวข้อง

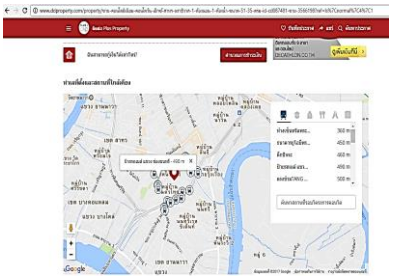
บริษัท ไปรษณีย์ไทย มีรถให้บริการทั้งหมด 7 ประเภท ได้แก่รถยนต์นั่ง (เก๋ง) รถยนต์ไมโครบัส รถยนต์ไมโครบัส (VIP) รถยนต์บรรทุกขนาดเล็ก (ปิกอัพ ตอนครึ่ง-CAB) รถยนต์ 18 ที่นั่ง รถยนต์ 40 ที่นั่ง และรถยนต์ 45 ที่นั่ง ตามลำดับ ปัจจุบันพบว่า การขอใช้บริการรถยนต์จะต้องดำเนินการผ่านเอกสารแบบฟอร์ม ซึ่งใช้เวลากระบวนการแต่ละขั้นตอนนาน สิ้นเปลืองกระดาษ การตรวจสอบตารางสถานะพนักงานขับรถและรถยนต์ผ่านเอกสาร ซึ่งสาเหตุล้วนมาจาก การจัดเก็บเป็นเอกสาร และไม่สามารถติดตามรถยนต์ขณะปฏิบัติงานได้ ผู้วิจัยได้ศึกษาค้นคว้าข้อมูลงานวิจัยที่เกี่ยวข้อง มีผู้พัฒนาเป็นเว็บแอปพลิเคชัน เพื่อเพิ่มประสิทธิภาพการบริหารจัดการรถยนต์และตอบสนองความต้องการของผู้ใช้ได้อย่างรวดเร็ว [1] ผ่านทั้งอินทราเน็ตและอินเทอร์เน็ต สามารถประมวลผลสถานะรถยนต์และพนักงานขับรถ และพิมพ์รายงานสรุปสถิติการขอใช้รถ [2] ซึ่งงานวิจัยที่กล่าวมาแล้วนั้นมีข้อจำกัดคือ ขาดระบบติดตามรถยนต์ [3] ผู้วิจัยจึงนำแนวคิดจุดเด่นของงานวิจัยต่างๆ มาพัฒนาเป็นเว็บแอปพลิเคชันสำหรับการขอใช้รถยนต์ ซึ่งระบบประมวลผลสถานะรถยนต์หรือพนักงานขับรถที่พร้อมปฏิบัติงานเพื่ออำนวยความสะดวกให้กับผู้ใช้งาน และใช้เทคโนโลยี GPS ติดตามรถยนต์ บันทึกตำแหน่งพิกัดตั้งแต่เริ่มปฏิบัติงานจนสิ้นสุดการปฏิบัติงาน และใช้เทคโนโลยี Google Map API ค้นหาสถานที่ เลือกเส้นทางที่เหมาะสม [4] คำนวณระยะเวลา [5] จากบริษัทฯ ไปยังปลายทาง [6] , [7] หากเกินเวลาเดินทางจาก Google MAP API คำนวณ จะแจ้งเตือน [8] ให้ผู้อนุมัติทราบว่ามีพนักงานขับรถเกินเวลาที่กำหนด

2.1 GPS

GPS ย่อมาจาก Global Positioning System เป็นระบบกำหนดตำแหน่งบนพื้นโลก ซึ่งเป็นเครื่องรับสัญญาณจากดาวเทียม NAVSTAR GPS (Navigation System with Time and Ranging Global Positioning System) เป็นระบบสำรวจหาตำแหน่งบนพื้นโลกด้วยดาวเทียม โดยจะค้นหาตำแหน่งและนำทางด้วยดาวเทียม จะประกอบด้วยดาวเทียมจำนวน 24 ดวง โคจรรอบโลกวันละ 2 รอบ ตัวเครื่องรับสัญญาณจะต้องมองเห็นไม่น้อยกว่า 4 ดวง จะทำให้สามารถนำข้อมูลไปคำนวณหาตำแหน่งได้ทุกแห่งบนพื้นโลก การทำงานของ GPS จะคำนวณระยะทางระหว่างดาวเทียมกับเครื่อง GPS ซึ่งจะต้องใช้ระยะทางจากดาวเทียมอย่างน้อย 3 ดวงขึ้นไป โดยใช้สูตรคำนวณทางฟิสิกส์ คือ ระยะทางเท่ากับความเร็ว * เวลา โดยดาวเทียมทั้ง 3 ดวงจะส่งสัญญาณคลื่นวิทยุ โดยเคลื่อนที่ด้วยความเร็วคงที่มายังเครื่อง GPS โดยความเร็ว 186,000 ไมล์ต่อวินาที เครื่อง GPS นั้นจะมีความถูกต้องแม่นยำ ขึ้นอยู่กับจำนวนดาวเทียมที่สามารถรับสัญญาณได้ในขณะนั้น [9]

2.2 Google Maps API (Application Programming Interface)

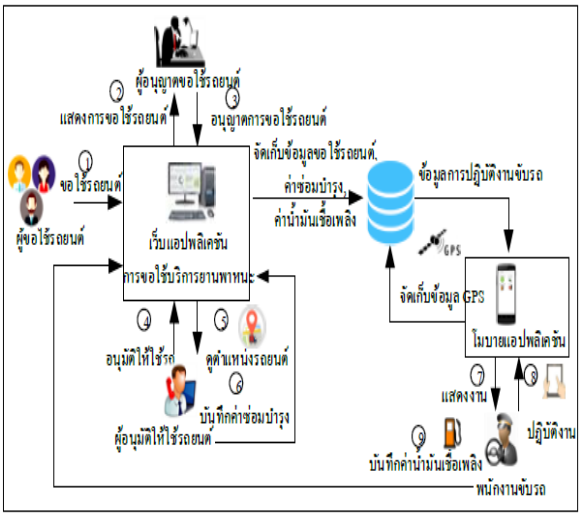
API (Application Programming Interface) เป็นลักษณะการเขียนโปรแกรมเพื่อเชื่อมต่อกับผู้ให้บริการ ซึ่งผู้ให้บริการจะมีไลบรารีหรือคลาสเพื่อให้บริการ โดยการเรียกใช้ผ่านอินเตอร์เฟสที่จัดไว้ให้ Google Maps API เป็นชุด API เพื่อเชื่อมต่อกับ Google เพื่อเรียกอ่านข้อมูลทางภูมิศาสตร์ ในการเรียกดูแผนที่แต่ละครั้งต้องมีการระบุ Google Map API Key ให้ทำการประมวลผล จากนั้น Google จะส่งแผนที่กลับมาผู้ที่ร้องขอผ่านทางเว็บเพจ [10] การเลือกเส้นทางของ Google Map ใช้ไลบรารี่ชื่อ Google Direction ซึ่งระยะทางจะเลือกแบบ Manhattan Distance หรือเรียกอีกชื่อว่าระยะทางแบบ city block ใช้ฟังก์ชันวัดระยะห่างซึ่งพิจารณาความแตกต่างของมิติค่าเฉลี่ย การประยุกต์การใช้งาน สามารถประยุกต์ได้หลากหลาย เช่น แผนที่ท่องเที่ยว ธุรกิจอสังหาริมทรัพย์ เป็นต้น [11], [12]



ภาพที่ 1: ตัวอย่างการใช้งาน Google Map API

3. วิธีดำเนินงานวิจัย

วิธีการดำเนินงานวิจัยนี้ มีกรอบแนวคิดในการพัฒนาระบบ ดังภาพที่ 2



ภาพที่ 2: กรอบแนวความคิดของงานวิจัย

จากภาพที่ 2 สามารถอธิบายกรอบแนวความคิดของระบบออกเป็นออกเป็น 2 ส่วน คือ 1) เว็บแอปพลิเคชันการขอใช้บริการรถยนต์ 2) โมบายแอปพลิเคชันสำหรับการติดตามรถยนต์ โดยได้แยกตามประเภทของผู้ใช้งานดังนี้

ผู้ขอใช้รถยนต์ กรอกข้อมูลเพื่อขอใช้รถยนต์และสามารถตรวจสอบสถานการณ์ขอใช้รถยนต์

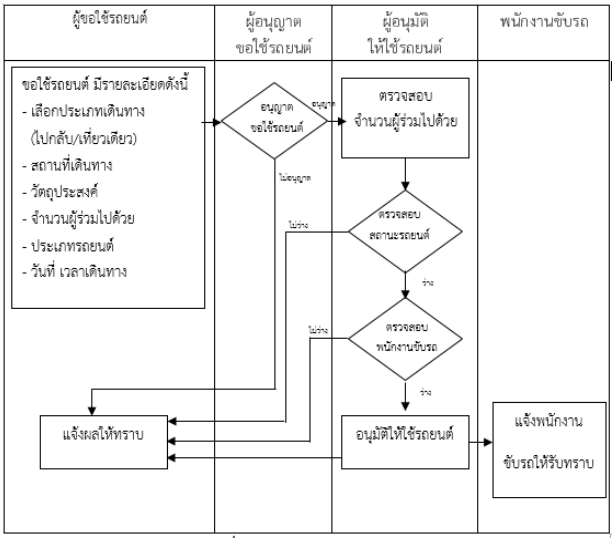
ผู้อนุญาตใช้รถยนต์ เป็นผู้อนุญาตการขอใช้รถยนต์

ผู้อนุมัติให้ใช้รถยนต์ เป็นผู้อนุมัติรถยนต์พร้อมพนักงานขับรถ และสามารถดูตำแหน่งรถยนต์ที่ปฏิบัติงาน กรณีมีค่าซ่อมแซม จะเป็นผู้กรอกค่าซ่อมแซม และสามารถดูรายงานการขอใช้รถยนต์ได้

พนักงานขับรถ จะดูงานที่ได้รับมอบหมาย ขณะปฏิบัติงานระบบจะบันทึกตำแหน่งจีพีเอสผ่านโมบายแอปพลิเคชัน ลงฐานข้อมูล กรณีมีค่าน้ำมันเชื้อเพลิง จะเป็นผู้กรอกค่าน้ำมันพร้อมแนบใบเสร็จ

3.1 ผังงานของระบบ (System Flow Diagram)

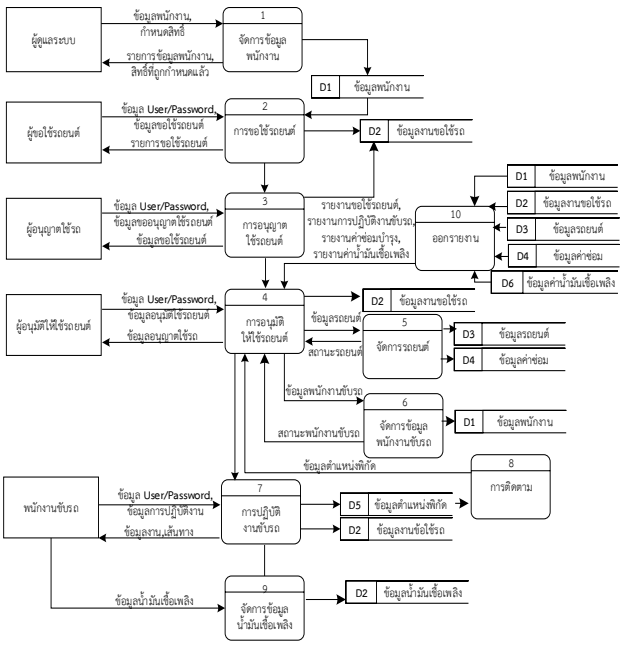
ขั้นตอนการขอใช้บริการรถยนต์ ขั้นตอนแรกผู้ขอใช้รถยนต์จะดำเนินการเลือก ประเภทเดินทาง (ไปกลับ/เที่ยวเดียว) สถานที่เดินทาง จำนวนผู้ร่วมเดินทาง และประเภทรถยนต์ ต่อมาผู้อนุญาตจะอนุญาตการขอใช้รถยนต์ จากนั้นผู้อนุมัติให้ใช้รถยนต์จะพิจารณาจำนวนผู้ร่วมเดินทางเพื่อเลือกประเภทรถยนต์ที่เหมาะสมและดูพนักงานขับรถที่มีคิวว่างในช่วงเวลานั้น ถ้าพบว่ารถยนต์และพนักงานขับรถว่าง จะแจ้งผลให้ผู้ขอใช้รถยนต์ ดังภาพที่ 3



ภาพที่ 3: ขั้นตอนการขอใช้รถยนต์

3.2 การออกแบบและพัฒนาระบบ

การออกแบบกระบวนการทำงานของระบบ แสดงด้วยแผน ภาพการไหลของข้อมูล ระดับ 0 ดังภาพที่ 4



ภาพที่ 4: แผนการไหลของข้อมูลระดับ 0

การพัฒนาระบบแบ่งเป็น 2 ส่วน คือ เว็บแอปพลิเคชันใช้เครื่องมือ Adobe Dreamweaver CS3 โมบายแอปพลิเคชันใช้เครื่องมือ Android Studio 2.2 ส่วนฐานข้อมูลใช้ MariaDB

และใช้ Google MAP API เป็นเครื่องมือค้นหาสถานที่ แสดงตำแหน่ง คำนวณเวลา

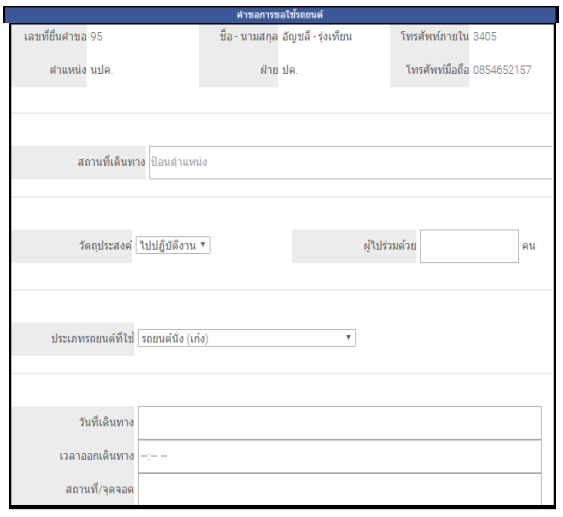
3.3 การทดสอบระบบ

การทดสอบเพื่อตรวจหาข้อผิดพลาดจากระบบมี 2 ขั้นตอน คือ ขั้นตอนที่ 1 ทดสอบระบบในขั้นแอลฟา (Alpha Stage) ผู้พัฒนาทำการทดสอบหาข้อผิดพลาดของระบบ เป็นการทดสอบโปรแกรม แต่ละฟังก์ชัน และดูความถูกต้องของระบบ ขั้นตอนที่ 2 ทดสอบในขั้นเบต้า (Beta Stage) ทำแบบประเมินคุณภาพของระบบ โดยผู้เชี่ยวชาญด้านเทคโนโลยีสารสนเทศของบริษัทฯ จำนวน 3 คน ใช้แบบสอบถามประเมินประสิทธิภาพของระบบ และใช้แบบสอบถามประเมินความพึงพอใจของผู้ใช้ที่เกี่ยวข้องกับระบบโดยตรง 5 คน พิจารณา 5 ด้านคือ ด้านการทำงานตรงตามความต้องการของผู้ใช้, ด้านความถูกต้องของการทำงานตรงตามฟังก์ชันต่างๆ ที่มีอยู่ในระบบ, ด้านความง่ายต่อการใช้งานระบบ ด้านการรักษาความปลอดภัยของข้อมูล และด้านภาพรวมของระบบ โดยได้กำหนดเกณฑ์การให้คะแนนเชิงคุณภาพและเชิงปริมาณ แบ่งเป็น 5 ระดับ (ดีมาก ดี ปานกลาง น้อย และน้อยมาก) และเชิงปริมาณ (1-5) จากนั้นสรุปผลโดยใช้ค่าเฉลี่ย (Mean) และค่าเบี่ยงเบนมาตรฐาน (S.D.)

4. ผลการดำเนินงาน

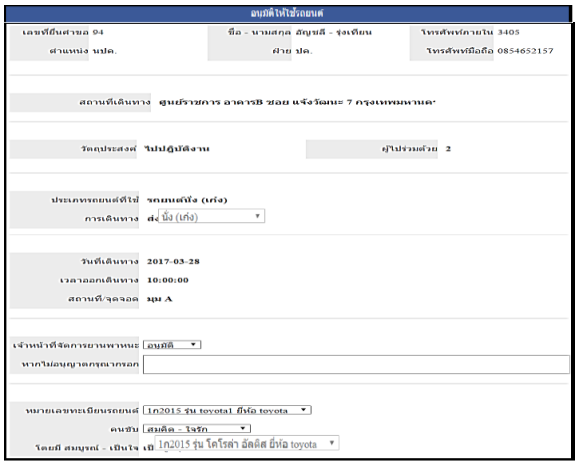
4.1 ผลการพัฒนาระบบ

ระบบการขอใช้บริการและติดตามรถยนต์ ผู้วิจัยได้พัฒนาขึ้น โดยแยกสิทธิการใช้งานตามประเภทผู้ใช้



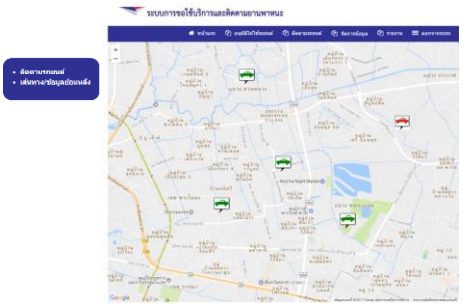
ภาพที่ 5: หน้าจอการขอใช้รถยนต์

จากภาพที่ 5 ส่วนของหน้าจอการขอใช้รถยนต์ ผู้ใช้งานสามารถค้นหาสถานที่เดินทางโดยใช้ Google Map API จากนั้นเลือกประเภทรถยนต์ วันที่และเวลาเดินทาง ต่อมาผู้อนุญาตจะดำเนินการอนุญาตการขอใช้รถ หลังจากนั้นผู้อนุมัติจะเป็นอนุมัติให้ใช้รถยนต์พร้อมพนักงานขับรถ โดยระบบจะประมวลผลสถานะรถยนต์และพนักงานขับที่พร้อมปฏิบัติงานดังภาพที่ 6



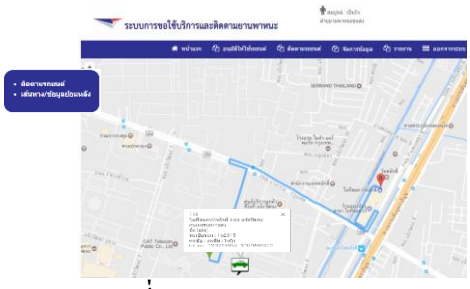
ภาพที่ 6: หน้าจอการอนุมัติให้ใช้รถยนต์

ส่วนของการติดตามรถยนต์ โดยระบบจะแสดงพิกัดตำแหน่งของผู้ใช้งานบนแผนที่ Google MAP API ดังภาพที่ 7 และแสดงเส้นทางการปฏิบัติงานได้ ดังภาพที่ 8



ภาพที่ 7: หน้าจอแสดงตำแหน่งผู้ใช้บนแผนที่ Google Map

โดยแสดงเป็นรูปภาพรถ (ทั้งนี้จะแสดงรูปรถสีเขียว หมายถึง ขับรถอยู่ในเวลาที่กำหนด ส่วนรูปรถสีแดง หมายถึง ขับรถเกินเวลาที่กำหนด)



ภาพที่ 8: หน้าจอแสดงเส้นทางปฏิบัติงาน

สำหรับโมบายแอปพลิเคชัน เมื่อพนักงานขับรถ ล็อกอินเข้าสู่ระบบ จะแสดงหน้าหลักดังภาพ ที่ 9 จากนั้นกดรับทราบงาน จะแสดงรายละเอียดของงานดังภาพที่ 10 และให้ Google MAP API คำนวณเส้นทางที่เหมาะสมดังภาพที่ 11



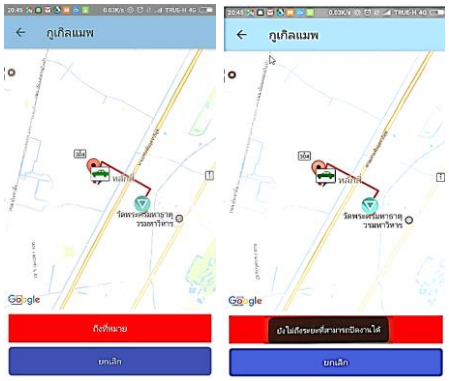
ภาพที่ 9: หน้าจอแสดงหน้าหลัก



ภาพที่ 10: หน้าจอแสดงรายละเอียดของงาน

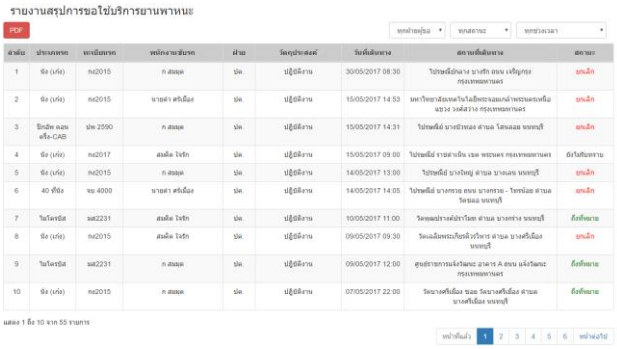
หากไม่ถึงปลายทางภายในรัศมี 1 กิโลเมตร ผู้ใช้งานไม่สามารถกดถึงที่หมายได้ ระบบจะแจ้งเตือน ดังภาพที่ 11 (ข)

นอกจากนี้ ระบบสามารถออกรายงานซึ่งประกอบด้วย รายงานการขอใช้รถยนต์ รายงานสรุปงานพนักงานขับรถ รายงานค่าซ่อมบำรุงรักษารถยนต์และรายงานค่าน้ำมันเชื้อเพลิง ดังภาพที่ 12



(ก) (ข)

ภาพที่ 11: (ก) หน้าจอแสดงเส้นทางปฏิบัติงาน
(ข) หน้าจอแจ้งเตือนถ้าไม่ถึงปลายทาง



ภาพที่ 12: หน้าจอแสดงรายงานการขอใช้รถยนต์

4.2 ผลการประเมินคุณภาพของระบบ

ผลการประเมินประสิทธิภาพของระบบโดยผู้เชี่ยวชาญด้านเทคโนโลยีสารสนเทศของบริษัท ไปรษณีย์ไทย อยู่ในระดับดี ค่าเฉลี่ยเท่ากับ 4.24 และค่าเบี่ยงเบนมาตรฐานเท่ากับ 0.43 ดังแสดงในตารางที่ 1

ตารางที่ 1: ผลการประเมินประสิทธิภาพของระบบ

ด้านที่ประเมิน	ประสิทธิภาพ		
	Mean	S.D.	แปล
ด้านการทำงานตามความต้องการของผู้ใช้	4.38	0.49	ดี
ด้านความถูกต้องของการทำงานตรงตามฟังก์ชันต่างๆ ที่มีอยู่ในระบบ	4.11	0.33	ดี
ด้านความง่ายต่อการใช้งานระบบ	4.41	0.47	ดี
ด้านการรักษาความปลอดภัยของข้อมูล	4.08	0.20	ดี
ด้านภาพรวม	4.16	0.41	ดี
สรุปผล	4.24	0.43	ดี

ผลการประเมินความพึงพอใจของผู้ใช้งานที่เกี่ยวข้องกับระบบโดยตรง อยู่ในระดับดี ค่าเฉลี่ยเท่ากับ 4.20 และค่าเบี่ยงเบนมาตรฐานเท่ากับ 0.43 ดังแสดงในตารางที่ 2

ตารางที่ 2: ผลการประเมินความพึงพอใจที่มีต่อระบบ

ด้านที่ประเมิน	ความพึงพอใจ		
	Mean	S.D.	แปล
ด้านการทำงานตรงตามความต้องการของผู้ใช้	4.27	0.46	ดี
ด้านความถูกต้องของการทำงานตรงตามฟังก์ชันต่างๆ ที่มีอยู่ในระบบ	4.3	0.48	ดี
ด้านความง่ายต่อการใช้งานระบบ	4.24	0.46	ดี
ด้านการรักษาความปลอดภัยของข้อมูล	4.2	0.42	ดี
ด้านภาพรวม	4	0	ดี
สรุปผล	4.20	0.43	ดี

5. สรุปผล

จากการพัฒนาระบบการขอใช้บริการและติดตามรถยนต์โดยการประยุกต์ใช้ Google Map API กรณีศึกษา บริษัท ไปรษณีย์ไทย จำกัด มีวัตถุประสงค์เพื่อให้การขอใช้บริการรถยนต์ได้ง่าย สะดวกรวดเร็ว และสามารถติดตามรถยนต์ได้ ซึ่งผู้วิจัยได้พัฒนาขึ้นออกเป็น 2 ส่วนคือ 1) เว็บแอปพลิเคชันสำหรับการขอใช้บริการรถยนต์ผ่านเครือข่ายอินเทอร์เน็ต ซึ่งได้กำหนดสิทธิการใช้งานตามประเภทผู้ใช้ คือ ผู้ขอใช้รถยนต์ ผู้อนุญาตขอใช้รถยนต์ ผู้อนุมัติให้ใช้รถยนต์ และพนักงานขับรถ 2) โมบายแอปพลิเคชัน สำหรับพนักงานขับรถใช้เพื่อติดตามรถยนต์ ในการทดสอบคุณภาพของระบบ ผลการประเมินความพึงพอของผู้มีส่วนเกี่ยวข้องกับระบบ อยู่ในระดับดี ทั้งนี้การคำนวณเส้นทาง และเวลาเดินทาง ระบบจะบันทึกในตอนเริ่มต้นเท่านั้น อย่างไรก็ตามการพัฒนาระบบการขอใช้บริการและติดตามรถยนต์ โดยการประยุกต์ใช้ Google Map API กรณีศึกษา บริษัท ไปรษณีย์ไทย จำกัด ที่พัฒนาขึ้นสามารถนำมาใช้งานได้จริง

เอกสารอ้างอิง

[1] ยุทธนา สงนรินทร์. ระบบฐานข้อมูลเพื่อการจัดการรถยนต์ : กรณีศึกษาระบบจัดการรถยนต์มหาวิทยาลัยทักษิณ. สารนิพนธ์วิทยาศาสตร์มหาบัณฑิต สาขาวิชาการจัดการเทคโนโลยีสารสนเทศ มหาวิทยาลัยสงขลานครินทร์ วิทยาเขตหาดใหญ่, 2550

[2] ณัฏฐกิตติ์ กิตติธำรงกุล. การพัฒนาระบบขอใช้รถยนต์ราชการของสำนักงานปลัดกระทรวงศึกษาธิการ. ปัญหาพิเศษ วิทยาศาสตร์มหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2551.

[3] ฐิติกานต์ เฉยบำรุง. แอปพลิเคชั่นติดตามรถยนต์ด้วยเครื่อง GPS. ปริญญานิพนธ์ สาขา วิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเกษตรศาสตร์ ,2555.

[4] C. Fu, Y. Wang, Y. Xu and Q. Li, "The Logistics Network System Based on the Google Maps API", International Conference on Logistics Systems and Intelligent Management 2010 , pp. 1486-1489, 2010.

[5] R. Hofstede, and T. Fioreze, "SURFmap: A network monitoring tool based on theGoogle Maps API", International Symposium on Integrated Network Management 2009, pp. 679-690, 2009.

[6] โสภาวดี โชติกลา และคนอื่นๆ . "เว็บแมพเซอร์วิสสสำหรับการท่องเที่ยวชุมชนในพื้นที่ชายฝั่งทะเลภาคตะวันออก : ชลบุรี ระยอง จันทบุรี ตราด" วารสารวิจัยราชภัฏ พระนคร. ปีที่ 10 ฉบับที่ 1 2558 : 54-64.

[7] สุพาพร เรืองแจ่ม. ระบบวิเคราะห์การใช้น้ำมันเชื้อเพลิงของรถจากการเดินทาง โดย Google Map API. การประชุมวิชาการ NCCIT2012. 2555 : 510-517.

[8] นิโรธ เกียรติวุฒิกาญจน์ ระบบติดตามเด็กเล็กเทคโนโลยีจีพีเอสบนโทรศัพท์มือถือระบบปฏิบัติการแอนดรอยด์. ปริญญานิพนธ์ สาขาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเกษตรศาสตร์, 2555.

[9] อรพิมพ์ มงคลเคหา. "Global Positioning Systeme (GPS) สู่โลกกว้างนวัตกรรมในยุค IT." ก้าวทันโลกวิทยาศาสตร์. ปีที่ 7 ฉบับที่ 2 (2550) :6-18

[10] ปฐมพงษ์ ฉับพลัน และฐิมาพร เพรชแก้ว. การประยุกต์ใช้ Google Maps API ในการเฝ้าระวังปัญหายาเสพติด. การประชุมวิชาการ NCCIT2010. 2553 : 196-201

[11] กฤษณา ชูลิตะพันธ์พงศ์ และวรพันธ์ แก้วพิทยาภรณ์. รายงานการวิจัยการประยุกต์ใช้ Google Map API สำหรับฐานข้อมูลสหกิจศึกษา. กรุงเทพฯ: มหาวิทยาลัยเทคโนโลยีราชมงคลกรุงเทพ, 2555.

[12] วรพันธ์ แก้วพิทยาภรณ์ และกฤษณา ชูลิตะพันธ์พงศ์. รายงานการวิจัยการประยุกต์ใช้ Google Map API การประยุกต์ใช้งาน Google Map APIสำหรับการประชาสัมพันธ์พิพิธภัณฑ์ท้องถิ่นของกรุงเทพฯ. กรุงเทพฯ : มหาวิทยาลัยเทคโนโลยีราชมงคลกรุงเทพ, 2555.

การเปรียบเทียบช่องโหว่ในระบบจัดการเนื้อหา: จุมลา เวิร์ดเพรส และดรูพอล
Comparison of Vulnerabilities in Content Management Systems: Joomla, Wordpress and Drupal

ปรัชญา ใจสุทธิ (P. Jaisudthi)[1] กิตติ มาลา (K. Mala)[2] และสัญชัย บุญเหลือ (S. Boonluea)[3]
สาขาวิชาเทคโนโลยีเครือข่ายคอมพิวเตอร์ คณะวิทยาการคอมพิวเตอร์และเทคโนโลยีสารสนเทศ มหาวิทยาลัยราชภัฏรำไพพรรณี
[1]pratchaya.j@rbru.ac.th, [2]kitti607@gmail.com, [3]teesanchai1@gmail.com

บทคัดย่อ

การเลือกระบบจัดการเนื้อหาเพื่อพัฒนาเว็บไซต์ของตนเองหรือบริษัทต้องคำนึงถึงความปลอดภัยของข้อมูลบนเว็บไซต์เป็นสำคัญ ดังนั้นบทความนี้ศึกษาช่องโหว่ของระบบจัดการเนื้อหา ซึ่งได้แก่ จุมลา เวิร์ดเพรส และดรูพอล โดยใช้เครื่องมือสแกนช่องโหว่จำนวนสามตัวได้แก่ อคูเนทิกซ์ เนสซุส และเวกา เพื่อเป็นข้อมูลเบื้องต้นในการเลือกระบบจัดการเนื้อหาที่เป็นที่นิยมในปัจจุบัน ซึ่งผลการวิจัยพบว่าดรูพอลมีช่องโหว่น้อยที่สุด แต่อย่างไรก็ตามผู้ใช้สามารถนำจุมลาและเวิร์ดเพรสไปใช้ได้โดยควรปิดช่องโหว่เหล่านั้นก่อน

คำสำคัญ: การเปรียบเทียบ ช่องโหว่ ระบบจัดการเนื้อหา

Abstract

Choosing a content management system for developing own or company website, the developer needs to consider the security of the information on website. Hence, this article studied the vulnerabilities of the content management system including the Joomla, Wordpress and Drupal. And we also used the three vulnerability scanners those were Acunetix, Nessus and Vega which were the most popular content management system. From the results, we found that Drupal had the least vulnerabilities. However, if users would like to use Joomla and Wordpress, they should to elclose those vulnerabilities first.

Keyword: Comparison, Vulnerability, Content Management System.

1. บทนำ

ปัจจุบันระบบจัดการเนื้อหา (Content Management System: CMS) เป็นระบบซอฟต์แวร์ที่นิยมใช้กันอย่างแพร่หลาย เพราะระบบนี้มีความซับซ้อนน้อย ทำให้ลดขั้นตอนในการพัฒนาและบริหารเว็บไซต์ ซึ่งระบบต้องใช้ควบคู่กับเครื่องบริการ (Server) และเครื่องบริการฐานข้อมูล (Database Server) ตัวอย่างระบบจัดการเนื้อหา [1] เช่น เวิร์ดเพรส (Wordpress) ซึ่งใช้งานได้ดีกับเว็บไซต์ขนาดเล็กไปจนถึงขนาดกลางรวมทั้งบล็อก จุมลา (Joomla) ซึ่งเป็นระบบที่เหมาะสมกับเว็บไซต์ธุรกิจออนไลน์ แต่ต้องอาศัยทักษะการเขียนโค้ดในบางส่วน ดรูพอล (Drupal) เป็นระบบที่พัฒนาขึ้นซึ่งช่วยให้ผู้ใช้ออกแบบเว็บไซต์ได้เสมือนจริง และง่ายต่อการใช้งานมากกว่าระบบอื่น โดยจากสถิติการใช้งานในเดือนมกราคมปี พ.ศ. 2560 [2] พบว่าเวิร์ดเพรสครองส่วนแบ่งทางการตลาดมากที่สุด รองลงมาคือ จุมลา และดรูพอล ตามลำดับ

ถึงแม้ว่าหลายหน่วยงานเริ่มหันมาใช้ระบบจัดการเนื้อหาเพื่อพัฒนาเว็บไซต์ของหน่วยงานเพื่อลดต้นทุนหรือทำให้ง่ายต่อการพัฒนา แต่ระบบซอฟต์แวร์ดังกล่าวยังคงมีระบบรักษาความปลอดภัยที่ไม่เสถียร โดยเฉพาะอย่างยิ่งช่องโหว่ที่เป็นจุดอ่อน ซึ่งผู้โจมตี (Attacker) มักใช้ช่องทางดังกล่าวโจมตีเว็บไซต์หรือเครื่องบริการเว็บ ปัญหาที่พบมากคือ ผู้โจมตีระบบเว็บจะทำให้เครื่องบริการเว็บกินทรัพยากรปริมาณมากซึ่งเป็นสาเหตุของการล่มของเว็บไซต์ [3]

ดังนั้นจึงมีการพัฒนาซอฟต์แวร์ต่าง ๆ เพื่อช่วยหาช่องโหว่ของเว็บไซต์ เรียกว่า ตัวสแกนช่องโหว่ของเว็บ (Web Vulnerability Scanner) ซึ่งเจ้าของเว็บไซต์จะใช้ซอฟต์แวร์ดังกล่าวเพื่อตรวจหาจุดอ่อนหรือช่องโหว่แล้วทำการแก้ไข

จุดอ่อนเหล่านั้นก่อนที่จะมีผู้ไม่ประสงค์ดีทำการเจาะหรือทำให้เกิดความเสียหายแก่เว็บไซต์หรือระบบฐานข้อมูลได้ ตัวอย่างตัวสแกนช่องโหว่ของเว็บ เช่น [4] อคูเนทิกซ์ (Acunetix) ซึ่งถูกพัฒนาขึ้นเพื่อหาช่องโหว่ของเว็บแอปพลิเคชัน โดยมีวิธีการค้นหาช่องโหว่ที่สำคัญคือ ซึ่งประกอบด้วยเอสคิวแอล อินเจคชัน (SQL Injection) และครอสไซท์สคริปปิ้ง (Cross Site Scripting) มีทั้งเวอร์ชันเสียเงินและทดลองใช้ฟรี ตัวอย่าง เช่น เนสซุส (Nessus) เวกา (Vega) เป็นต้น

สำหรับงานวิจัยนี้ได้นำเสนอการเปรียบเทียบระบบจัดการเนื้อหาโดยเปรียบเทียบจำนวนช่องโหว่ของเวิร์ดเพรส จุมลา และดรูพอล โดยใช้ตัวสแกนช่องโหว่เนสซุส เวกา และอคูเนทิกซ์

2. วรรณกรรมที่เกี่ยวข้อง

งานวิจัย [5] ได้ศึกษาการเปรียบเทียบระบบจัดการเนื้อหาบนเว็บไซต์ ได้แก่ Alfresco, Typo3, Dotnetnuke, Drupal, Joomla, Wordpress และ Plone โดยศึกษาความต้องการของระบบ ประสิทธิภาพ ความปลอดภัย คุณสมบัติ การสนับสนุนและการจัดการ ซึ่งสรุปได้ว่าจุมลามีระบบที่มียืดหยุ่นมากกว่าขณะที่ความปลอดภัยในส่วนของจุมลา ดรูพอลและเวิร์ดเพรสดีกว่าระบบอื่น ๆ Typo3 ให้ประสิทธิภาพที่ดีที่สุดและมีประโยชน์ต่อธุรกิจมาก DotNetNuke มีปลั๊กอินมากที่สุด จุมลามีการจัดการทรัพย์สิน โฆษณา และคลิปบอร์ดได้ดีที่สุด

เวิร์ดเพรสมีคุณสมบัติในการจัดการน้อยที่สุด Alfresco ใช้ง่ายที่สุด งานวิจัยดังกล่าวสามารถเปรียบเทียบในหลายๆ ด้านรวมทั้งทางด้านความปลอดภัย แต่อย่างไรก็ตามผู้วิจัยยังไม่สามารถบอกได้แน่ชัดว่าระบบจัดการเนื้อหาบนเว็บไซต์ระหว่างจุมลา เวิร์ดเพรส และดรูพอล ตัวใดดีที่สุด เนื่องจากเป็นเพียงการพิจารณาว่ามีฟังก์ชันเกี่ยวกับความปลอดภัยหรือไม่เท่านั้น ดังนั้นจึงจำเป็นต้องใช้เครื่องมือในการตรวจหาช่องโหว่เพิ่มเติม

งานวิจัย [6] ได้ศึกษาเครื่องมือที่ใช้วิเคราะห์ช่องโหว่ของเว็บภายในองค์กร โดยใช้ Nessus, Acunetix และ ZAP โดยได้ให้แนวทางในการสแกนหาช่องโหว่ แต่เป็นการตรวจสอบเว็บที่สร้างขึ้นมาภายในองค์กรเท่านั้นซึ่งผู้วิจัยไม่ได้ใช้ระบบจัดการเนื้อหา

งานวิจัย [7] ได้ศึกษาประสิทธิภาพของตัวสแกนเพื่อหาช่องโหว่ของบริการเว็บ ซึ่งได้แก่ Netsparker และ Acunetix ซึ่งผลที่ได้ทั้งสองตัว พบว่า ทางด้าน Cross Site Scripting (XSS) และ SQL Injection ซอฟต์แวร์ Acunetix สามารถตรวจสอบได้ทุกด้าน แต่อย่างไรก็ตามการตรวจสอบยังต้องได้รับการปรับปรุง

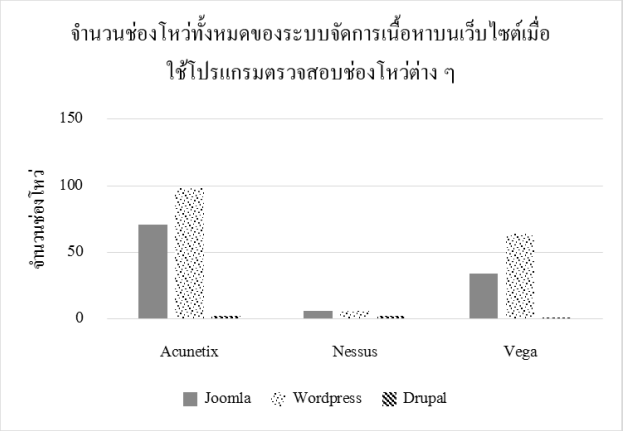
งานวิจัย [8] ได้ศึกษาประสิทธิภาพของความปลอดภัยใน Joomla, Drupal และ WordPress ในสภาวะเงื่อนไขเดียวกัน โดยใช้วิธีการเจาะระบบเข้าไปซึ่งผู้วิจัยได้สรุปว่า เวิร์ดเพรสมีความปลอดภัยที่ดีกว่าอีก 2 ตัว โดยงานวิจัยดังกล่าวได้ทำการเจาะระบบโดยไม่ใช้เครื่องมือสแกน

3. วิธีการดำเนินการวิจัย

บทความนี้นี้จะนำเสนอเครื่องมือที่สามารถบอกถึงช่องโหว่ของระบบจัดการเนื้อหา สำหรับการตรวจสอบช่องโหว่นั้นจะใช้เครื่องมือ Vulnerability Scanner ซึ่งเป็นโอเพนซอร์สตรวจสอบช่องโหว่ของเว็บไซต์ แล้วแสดงระดับความอันตรายของช่องโหว่ ในงานวิจัยนี้ได้เลือกเว็บไซต์ซึ่งเป็นบริการระบบจัดการเนื้อหามาสามตัว ได้แก่ Joomla เวอร์ชัน 3.6.5, Wordpress เวอร์ชัน 3.7.1 และ Drupal เวอร์ชัน 7.54 ในส่วนของซอฟต์แวร์ Vulnerability Scanner ได้เลือกมาสามตัว ได้แก่ Acunetix เวอร์ชัน 11, Nessus เวอร์ชันทดลองใช้ และ Vega เวอร์ชันทดลองใช้ โดยใช้หน้าเว็บไซต์เริ่มต้น (Default Website Page) ของซอฟต์แวร์ CMS แต่ละประเภทและใช้ตัวสแกนหาช่องโหว่โดยเว็บไซต์ 1 เว็บ จะต้องถูกสแกนจากตัวสแกนหาช่องโหว่ทั้งสามตัว แล้วนำผลที่ได้มาพล็อตเป็นกราฟแท่งแล้วทำการวิเคราะห์ต่อไป

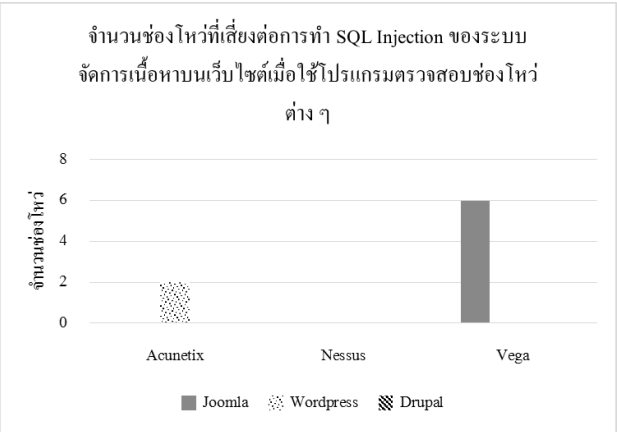
4. ผลการดำเนินงาน

งานวิจัยนี้ได้ดำเนินการสร้างเว็บไซต์โดยเครื่องมือ CMS ขึ้นมา 3 เว็บไซต์ เพื่อทำการทดสอบประสิทธิผลของเว็บไซต์ CMS ว่ามีช่องโหว่มากน้อยแค่ไหน โดยทำการตรวจสอบช่องโหว่โดยซอฟต์แวร์ Vulnerability Scanner ผู้วิจัยจึงทดสอบความปลอดภัยของเว็บไซต์ CMS โดยประเมินผลจากการตรวจสอบช่องโหว่ของแต่ละเว็บไซต์ ด้วยการสแกนหาช่องโหว่ด้วยการใช้ซอฟต์แวร์ Vulnerability Scanner ซึ่งมีผลดังนี้



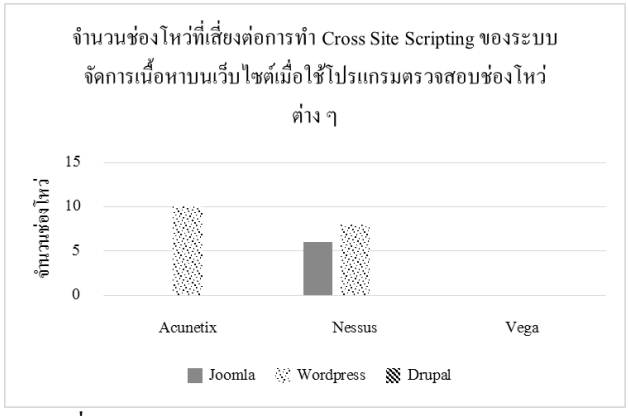
ภาพที่ 1: จำนวนช่องโหว่ทั้งหมดของระบบจัดการเนื้อหาบนเว็บไซต์เมื่อใช้โปรแกรมตรวจสอบช่องโหว่ต่าง ๆ

จากภาพที่ 1 แสดงจำนวนช่องโหว่ทั้งหมดของระบบจัดการเนื้อหาบนเว็บไซต์จูมลา เวิร์ดเพรส และดรูพอล โดยใช้โปรแกรมตรวจสอบช่องโหว่ได้แก่ อะคิวเนติกซ์ เนสซุส และเวกา พบว่ากราฟมีแนวโน้มในทิศทางเดียวกันคือ ระบบจัดการเนื้อหาบนเว็บไซต์ของเวิร์ดเพรส จูมลา และเวกามีปริมาณช่องโหว่เรียงจากมากไปน้อยตามลำดับ โดยช่องโหว่ที่นำมาใช้พิจารณาได้แก่ ช่องโหว่ที่มีระดับความเสี่ยงสูงสุด (High Vulnerabilities) และช่องโหว่ที่มีระดับความเสี่ยงปานกลาง (Medium Vulnerabilities) รวมกัน นอกจากนี้ทั้งสามโปรแกรมพบจำนวนช่องโหว่ทั้งหมดไม่เท่ากัน แต่จะพบว่าทั้งสามโปรแกรมตรวจพบจำนวนช่องโหว่ของเวอร์ดเพรสมากกว่าระบบจัดการเนื้อหาบนเว็บไซต์ตัวอื่น ๆ



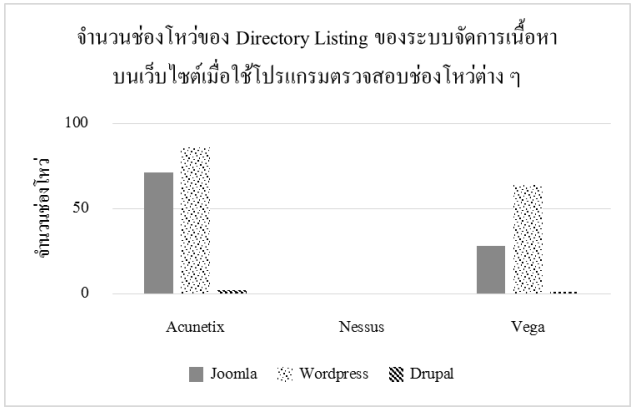
ภาพที่ 2: จำนวนช่องโหว่ที่เสี่ยงต่อการทำ SQL Injection ของระบบจัดการเนื้อหาบนเว็บไซต์เมื่อใช้โปรแกรมตรวจสอบช่องโหว่

เมื่อพิจารณาช่องโหว่ที่เสี่ยงต่อการทำ SQL Injection ตามภาพที่ 2 พบว่าดรูพอลไม่มีความเสี่ยงทางด้านนี้ เนื่องจากดรูพอลได้มีการแก้ไขระบบเพื่อปิดช่องโหว่ในส่วนนี้แล้ว ส่วนระบบจัดการเนื้อหาบนเว็บไซต์อื่นๆ ยังมีส่วนที่เสี่ยงต่อการทำ SQL Injection อยู่



ภาพที่ 3: จำนวนช่องโหว่ที่เสี่ยงต่อการทำ Cross Site Scripting ของระบบจัดการเนื้อหาบนเว็บไซต์เมื่อใช้โปรแกรมตรวจสอบช่องโหว่ต่าง ๆ

ส่วนเรื่องของจำนวนช่วงโหว่ที่เสี่ยงต่อการทำ Cross Site Scripting ตามที่แสดงในภาพที่ 3 พบว่าเวิร์ดเพรสยังเป็นระบบจัดการเนื้อหาบนเว็บไซต์ที่มีช่องโหว่ทางด้านนี้มากที่สุด ส่วนดรูพอลไม่พบช่องโหว่ดังกล่าวแต่อย่างไร



ภาพที่ 4: จำนวนช่องโหว่ของ Directory Listing ของระบบจัดการเนื้อหาบนเว็บไซต์เมื่อใช้โปรแกรมตรวจสอบช่องโหว่ต่าง ๆ

นอกจากนี้เมื่อพิจาณาภาพที่ 4 แสดงจำนวนช่องโหว่เนื่องจาก Directory Listing ซึ่งพบว่าในเวิร์ดเพรสและจูมลามีปริมาณช่องโหว่มากกว่าดรูพอล ทั้งนี้เนื่องจากทั้งเวิร์ดเพรสและจูมลาไม่ได้มีการปิดฟังก์ชัน Directory Listing ทำให้

ผู้ไม่หวังดีสามารถเข้าถึงพาธที่เป็นโฟลเดอร์ได้ ซึ่งจะทำให้ผู้ไม่หวังดีเข้าถึงไฟล์ในเครื่องบริการโดยไม่ได้รับอนุญาตได้ง่ายกว่าดรูพอล

5. สรุป

บทความนี้ได้อธิบายถึงความเสี่ยงด้านความปลอดภัยของระบบจัดการเนื้อหาเว็บไซต์ โดยใช้โปรแกรม Acunetix, Nessus และ Vega สแกนหาช่องโหว่ของ Joomla, Wordpress และ Drupal สำหรับการประเมินความปลอดภัย พบว่าโดยรวมหรือแม้กระทั่งช่องโหว่ที่เสียงต่อการโจมตีที่นิยมไม่ว่าจะเป็น SQL Injection และ Cross-Site Script เมื่อใช้ตัวสแกนช่องโหว่ทั้งสามพบว่าดรูพอลเป็นระบบจัดการเนื้อหาที่มีช่องโหว่น้อยที่สุด

เอกสารอ้างอิง

[1] S. Gilani, A. Majeed , M. Muzammal, A.U. Rehman, H.Zaheer5 and Z. Jan, "*A Navigational Evaluation Model for Content Management Systems*", The Nucleus 53, No. 2, pp. 82-88, 2016.

[2] C. A. Contu, E. C. Popovici, O. Fratu and M. G. Berceanu, "*Security Issues in Most Popular Content Management Systems*", The 11th International Conference on COMMUNICATIONS - COMM 2016, At Bucharest, ROMANIA, 2016.

[3] P. Taral and B. Gite, "*CMS Website Security Threat Protection Oriented Analyzer System*", International Journal of Computer Applications Technology and Research. Vol. 3, pp. 809 - 811, 2014.

[4] L. Dukes, X. Yuan, F. Akowuah, "*A Case Study on Web Application Security Testing with Tools and Manual Testing*", Southeastcon, 2013 Proceedings of IEEE, 2013.

[5] A. Mirdha, A. Jain and K. Shah, "*Comparative Analysis of Open Source Content Management Systems*", 2014 IEEE International Conference on Computational Intelligence and Computing Research, 2014.

[6] N. I. Daud, K. A. Abu Bakar and M. S. Md Hasan, "*A Case Study on Web Application Vulnerability Scanning Tools*", Science and Information Conference 2014, pp. 595-600, 2014.

[7] C. Joshi and U. K. Singh, "*Performance Evaluation of Web Application Security Scanners for More Effective Defense*", International Journal of Scientific and Research Publications, Vol. 6, June 2016.

[8] S. K. Patel, V. R. Rajkot and J. B. Prajapati, "*Comparative Analysis Of Web Security In Open Source Content Management System*", 2013 International Conference on Intelligent Systems and Signal Processing (ISSP), 2013.

การประยุกต์ใช้กูเกิลแมปของโรคพิษสุนัขบ้าในมนุษย์และสัตว์
Applied Google Map on Rabies in Humans and Animals

ปารวี ภัทรกวิน (Paravee Pattarakavin)[1] และ เกริก ภิรมย์โสภา (Krerk Piromsopa)[2]
ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
[1]Paravee.secu@gmail.com,[2] Krerk.P@chula.ac.th

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาระบบการทำแผนที่ระบาดวิทยาสำหรับข้อมูลจากโรคพิษสุนัขบ้าที่เกี่ยวข้องกับข้อมูลอื่น ๆ เช่น โรคมนุษย์ โรคสัตว์ อุณหภูมิ และ ปริมาณน้ำฝน ปัจจัยเหล่านี้มีความสัมพันธ์กัน แต่ก็ไม่สามารถเข้าถึงได้ง่ายในรูปแบบที่เป็นอันหนึ่งอันเดียวกันและมีความหมาย งานวิจัยนี้นำข้อมูลที่จะแสดงบนแผนที่ของกูเกิลในรูปแบบที่สะดวกในการเข้าถึงข้อมูล นอกจากนั้นระบบนี้จะเป็นเครื่องมือสำหรับการเรียนรู้เกี่ยวกับการผสานรวมข้อมูลประเภทต่างๆ ในการประเมินระบบที่เสนอนั้นแบบสอบถามจะถูกแจกจ่ายให้กับบุคคลทั่วไปจำนวน 32 คน ผลการสำรวจพบว่า 25% เห็นว่า ระบบมีประโยชน์มากที่สุด 25% มีประโยชน์มาก 28% มีประโยชน์ ซึ่งผู้ที่เห็นว่ามีประโยชน์รวมทั้งสิ้น 78% ยิ่งไปกว่านั้นยังพบว่ามีผู้ใช้ 38% ที่สนใจข้อมูลการแพร่ระบาด แต่ไม่ทราบวิธีเข้าถึงข้อมูลสาธารณะ ซึ่งผู้ใช้กลุ่มนี้จะได้รับประโยชน์สูงสุดจากการใช้ระบบที่เรานำเสนอ

คำสำคัญ: ระบบสารสนเทศภูมิศาสตร์ โรคระบาด การทำจินตทัศน์ กูเกิลแมป

Abstract

This research aims to develop an epidemiological mapping system for Rabies data which is related to other data such as human disease, animal disease, temperature and rain fall. These factors are related but they are not easily access in a unified and meaningful way. This research brings information to be displayed on a Google Map in a convenient form to access the data. Moreover, this system will be a tool for learning about integration of different type of data. To evaluate the proposed system, the questionnaire is distributed to 32 general users. The result of the survey shows that 25% agree that the system is most useful, 25% very useful, 28% useful and 78% overall useful. Moreover, there are 38% of users who are interested in outbreak information but do not know how to access the public data. This group of users will gain most benefit from using our proposed system.

Keyword: *GIS, Epidemiology, Data Visualization, Google Map*

1. ความเป็นมาหรือความสำคัญ

การได้มาซึ่งข้อมูลนั้น จำเป็นต้องใช้กำลังคน ค่าใช้จ่ายและเวลา จึงเป็นสิ่งมีค่า ด้วยความหลากหลายของรูปแบบข้อมูลเพื่อการแสดงผล จึงเป็นแนวทางเพื่อนำมาผ่านกระบวนการให้อยู่ในรูปแบบที่ง่ายขึ้น เพื่อให้นักวิจัยหรือผู้สนใจได้สามารถเข้าถึงและนำไปใช้ประโยชน์

ข้อมูลสถิติด้านวิทยาการระบาด (Epidemiology) ทั้งของคนและสัตว์มีความสัมพันธ์กันในทางระบาดวิทยา โดยเฉพาะโรคติดต่อซึ่งเป็นชนิดที่มีเชื้อ (Infectious Diseases) หลายชนิดที่มีความสัมพันธ์กันกับอุบัติการณ์ของโรค (Incidence) ระหว่างคนกับสัตว์ ทั้งรูปแบบที่ติดต่อระหว่างคนและสัตว์โดยตรงหรือรูปแบบของพาหะ ในการศึกษาการระบาดนั้นยังมีปัจจัยด้านสิ่งแวดล้อมที่เข้ามามีผลต่อการระบาด เช่น อุณหภูมิ ปริมาณน้ำฝน ซึ่งโรคระบาดแบ่งเป็นโรคระบาดในช่วงฤดูฝนและฤดูร้อน โดยนักระบาดวิทยาใช้เวลาและสถานที่เป็นเครื่องมือวัดการกระจายของโรค

การทำจินตทัศน์ (Visualization) ถือเป็นเครื่องมือหนึ่งเพื่อแสดงผลลัพธ์ความสัมพันธ์ของข้อมูลออกมา จึงนับว่าเป็นกุญแจเพื่อให้ผู้ใช้เข้าถึงและเข้าใจความสัมพันธ์ของข้อมูล ซึ่ง

เป็นประโยชน์ต่อการตีความและการนำผลลัพธ์นั้นไปใช้งาน [1] โดยงานวิจัยนี้ได้รับการสนับสนุนข้อมูลจากสำนักระบาดวิทยาสำนักควบคุมป้องกันและบำบัดโรคสัตว์ และ กรมอุตุนิยมวิทยา มาทำจินตทัศน์ในการแสดงผลลัพธ์ รูปแบบและความสัมพันธ์ของข้อมูล เพื่อเป็นประโยชน์ต่อการตีความและการนำผลลัพธ์นั้นไปใช้รวมถึงการใช้ประโยชน์จากบุคคลทั่วไปที่สนใจ ซึ่งโรคที่ใช้เป็นตัวอย่างในการศึกษาครั้งนี้คือโรคพิษสุนัขบ้า (Rabies)

ผู้วิจัยได้ตั้งสมมติฐานว่า ระบบที่พัฒนาขึ้นจะมีประโยชน์ต่อผู้ที่สนใจและใช้งาน ดังนี้

H_0 คือ การนำข้อมูลโรคระบาดของคนและสัตว์ รวมถึงอุณหภูมิ และปริมาณน้ำฝน มาซ้อนทับ บนแผนที่เดียวมีประโยชน์ต่อผู้สนใจข้อมูลด้านโรคระบาด

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ผู้วิจัยได้ศึกษาทฤษฎีและงานวิจัยที่เกี่ยวข้องในด้านระบาดวิทยาและการทำแผนที่บนกูเกิลแมป

2.1 ระบบสารสนเทศภูมิศาสตร์ (GIS) [2]

ระบบสารสนเทศภูมิศาสตร์ (Geographic Information System: GIS) คือ กระบวนการทำงานเกี่ยวกับข้อมูลในเชิงพื้นที่ด้วยระบบคอมพิวเตอร์ ที่ใช้กำหนดข้อมูลและสารสนเทศที่มีความสัมพันธ์กับตำแหน่งในเชิงพื้นที่ ที่มีความสัมพันธ์กับตำแหน่งในแผนที่ ตำแหน่งของเส้นรุ้ง (Latitude) เส้นแวง (Longitude)

2.2 ฐานข้อมูลที่มีส่วนสัมพันธ์กับข้อมูลเชิงพื้นที่ (Spatial Data) [2]

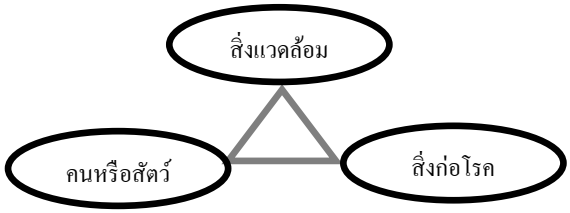
ฐานข้อมูลที่มีส่วนสัมพันธ์กับข้อมูลเชิงพื้นที่ ซึ่งจัดอยู่ในรูปแบบและความสัมพันธ์ ของข้อมูลเชิงพื้นที่ทั้งหลาย ซึ่งจะสามารถนำมาวิเคราะห์ด้วยระบบสารสนเทศภูมิศาสตร์และทำให้สื่อความหมายเรื่องการเปลี่ยนแปลงที่สัมพันธ์กับเวลาได้ เช่น การแพร่ขยายของโรคระบาด ข้อมูลเหล่านี้ เมื่อปรากฏบนแผนที่ทำให้สามารถแปล สื่อความหมายและสะดวกในการทำความเข้าใจรวมถึงการนำไปใช้งาน

2.3 วิทยาการระบาด (Epidemiology)

ระบาดวิทยา[3] คือ การศึกษาปัจจัยที่มีผลต่อสุขภาพ ความเจ็บป่วยของประชากร ความสนใจในสาธารณสุข และเวชศาสตร์การป้องกันระบาดวิทยาเป็นพื้นฐานสำคัญของงานวิจัยด้านสาธารณสุข เวชศาสตร์อิงหลักฐาน (Evidence-Based Medicine) ในการหาปัจจัยเสี่ยงของโรคและประเมินวิธีการรักษาที่เหมาะสม

ระบาดวิทยาเป็นศาสตร์ที่ศึกษาทั้ง โรคที่ติดเชื้อ (Infectious Disease) และโรคไร้เชื้อ (Non-infectious Disease) เพื่อหาสาเหตุของโรค ธรรมชาติของโรค สืบสวนการระบาดของโรค เฝ้าระวังโรค วางมาตรการป้องกันและควบคุม วิทยาการระบาดมีรากฐานมาจากความจริงที่พบว่า โรคหรือความเจ็บป่วย มีโอกาสเกิดในแต่ละบุคคลด้วยอัตราเสี่ยงที่ไม่เท่ากันขึ้นอยู่กับปัจจัยอื่น เช่น ตัวบุคคล สิ่งที่ก่อให้เกิด หรือ สภาพแวดล้อม เป็นต้น

จากภาพที่ 1 แสดงสามเหลี่ยมการเกิดโรคทางระบาดวิทยา (Epidemiologic Triangle) ประกอบด้วย คน หรือสัตว์ (Host) สิ่งก่อโรค (Agent) และ สิ่งแวดล้อม (Environment) สภาวะอากาศ อุณหภูมิ ซึ่งหากมาบรรจบกันอย่างเหมาะสม จะทำให้เกิดโรคขึ้นได้ ซึ่งหลักการนี้สามารถประยุกต์ใช้กับการเกิดโรคภัยไข้เจ็บได้ทุกโรค ทั้งคนและสัตว์โดยวิธีการทางระบาดวิทยา



ภาพที่ 1 แผนภาพแสดงปัจจัยสามทางที่ทำให้เกิดโรค [3]

2.4 โรคพิษสุนัขบ้าหรือโรคกลัวน้ำ (Rabies) [4][5]

โรคพิษเกิดจากเชื้อไวรัส (Rabies) ซึ่งเป็น อาร์ เอน เอ ไวรัส (RNA virus) หรือโรคพิษสุนัขบ้า ซึ่งแพร่เชื้อโดยถูกสัตว์ที่มีเชื้อกัด เชื้อที่อยู่ในน้ำลายจะเข้าสู่บาดแผลและผ่านเข้าสู่เส้นประสาทส่วนปลาย ไขสันหลัง ระยะฟักตัวของเชื้อไวรัสเรบิส ในคนนั้นพบว่า 87% มีระยะฟักตัวของโรค 3 เดือน แต่ทุกรายมีระยะฟักตัวของโรคไม่เกิน 1 ปี ทั้งนี้ขึ้นกับปัจจัยอื่น โรคพิษสุนัขบ้า เป็นหนึ่งในโรคในฤดูร้อน ที่ป้องกันได้ แต่ก็มีผู้เสียชีวิตจากโรคนี้ทุกปี โดยมีพาหะหลักจากสุนัข แมว ที่แพร่เชื้อไวรัสจากการกัด ข่วน หรือเลียผิวหนังคนที่มีแผล ซึ่งปัจจุบันยังไม่มียารักษาโรคพิษสุนัขบ้า ดังนั้น ไม่ว่ามนุษย์หรือสัตว์ที่ติดเชื้อนี้เมื่อปรากฏอาการของโรค จะเสียชีวิตทุกราย

2.5 การทำจินตทัศน์ (Visualization)

จินตทัศน์ [2] [6] คือการแสดงผลลัพธ์รูปแบบความสัมพันธ์ของข้อมูลออกมาให้ผู้ใช้สามารถมองเห็น รับรู้ และเข้าถึงความหมายของข้อมูลที่ต้องการนำเสนอ ซึ่งการทำจินตทัศน์มีหลายรูปแบบ เช่น กราฟ แผนที่ มายด์แมป เป็นต้น [7] โดยมีนักวิจัยจำนวนมากที่ให้ความสนใจจินตทัศน์และนำมาประยุกต์ใช้กับการแสดงรูปแบบความสัมพันธ์ในด้านการกระจายตัวเชิงพื้นที่ เช่น นักวิจัยชาวจีน[8]ได้นำจินตทัศน์ในรูปแบบแผนที่บนกูเกิลแมปแสดงการใช้ประโยชน์การใช้ที่ดิน นอกจากนี้ยังมีการประยุกต์ใช้การแสดงผลแบบแผนที่กับโรคระบาด เช่น นักวิจัยชาวคาซัคสถาน [9] ได้ทำการวิจัยเรื่องโรคแท้งติดต่อหรือโรคบรูเซลลิส (Brucellosis) โดยใช้การแสดงผลแบบออฟไลน์ด้วยควอนตัมจีไอเอส ซึ่งแสดงความสัมพันธ์ของระดับแท้งติดต่อของสัตว์เลี้ยงในฟาร์ม โดยใช้สีเขียวและสีชมพูแทน แพะและแกะ ในงานวิจัยดังกล่าว เป็นการแสดงความสัมพันธ์ข้อมูลเพียงแค่โรคที่เกิดขึ้น ไม่ได้มีการแสดงถึงปัจจัยที่เกี่ยวข้องอื่นๆ ยิ่งไปกว่านั้น ได้มีนักวิจัยจากหน่วยระบาดวิทยาด้านสุขภาพและอนามัยของกรมอนามัยและบริการมนุษย์ (DHHS) ในรัฐแทสเมเนีย[10] พัฒนาเครื่องมือที่ใช้บน กูเกิลแมป ชื่อ ‘เว็บ อิพิ’ เพื่อแทนที่ซอฟต์แวร์แผนที่แบบดั้งเดิมที่มีราคาสูง ในการทำแผนที่สำหรับตรวจสอบข้อมูลทางระบาดวิทยาใช้ดำเนินการติดตามผู้ป่วยและเสียชีวิตด้วยโรคมะเร็ง

มีผู้ทำการวิจัยในการประยุกต์ใช้ระบบสารสนเทศภูมิศาสตร์และ การสร้างแบบจำลอง ทางภูมิศาสตร์ เชิงพื้นที่ของระบาดวิทยาสิ่งแวดล้อม[11] ผู้วิจัยได้นำการประยุกต์ดังกล่าวมาเป็นแนวทางในโครงงานวิจัยนี้ เนื่องจากเป็นงานวิจัยที่คล้ายคลึงกัน

จากการศึกษาค้นคว้า จินตทัศน์ที่ได้รับความนิยมในการแสดงรูปแบบความสัมพันธ์ในการกระจายตัวเชิงพื้นที่และโรค คือ แผนที่ ซึ่งงานวิจัยนี้ได้ทำแบบสอบถามเพื่อให้ผู้ทำการประเมินเลือก รูปแบบการแสดงผล โดยเปรียบเทียบระหว่างแผนที่ กราฟเส้น และตารางข้อมูลซึ่งเป็นที่นิยมใช้ในปัจจุบัน

2.6 การแบ่งฤดูกาลในประเทศไทย [12]

ฤดูกาลในประเทศไทยแบ่งออกเป็น 3 ฤดู ได้แก่ ร้อน ฝน หนาว

2.6.1 ฤดูร้อน ประมาณกลางเดือนกุมภาพันธ์ - กลางเดือนพฤษภาคม แบ่งช่วงอากาศร้อนเป็น 2 ช่วง คือ ร้อน 35.0 -39.9 องศาเซลเซียส และ ร้อนจัด คือ 40.0 องศาเซลเซียส ขึ้นไป

2.6.2 ฤดูฝน ประมาณกลางเดือน พฤษภาคม - กลางเดือนตุลาคม การแบ่งช่วงของปริมาณน้ำฝนแบ่งเป็น 5 ช่วง คือ ฝนวัดจำนวนไม่ได้ ฝนเล็กน้อย ฝนปานกลาง ฝนหนัก และ ฝนหนักมาก

2.6.3 ฤดูหนาว ประมาณกลางเดือน ตุลาคม - กลางเดือนกุมภาพันธ์

การแบ่งช่วงของอากาศในช่วงหน้าหนาว แบ่งเป็น 3 ช่วง คือ อากาศหนาวจัด อากาศหนาว อากาศเย็น

2.7 กูเกิลแมป (Google Map) [13] [14]

กูเกิลแมปเป็นบริการแผนที่ออนไลน์ โดยกูเกิลได้เปิด เอพีไอให้กับนักพัฒนาระบบสามารถโหลดมาใช้งานโดยไม่เสียค่าใช้จ่าย

2.8 เจสัน (JavaScript Object Notation: JSON) และ เอ็กเอ็มแอล (Extensible Markup Language: XML) [16][17]

เจสันและเอ็กเอ็มแอลเป็นรูปแบบการแลกเปลี่ยนข้อมูลที่ใช้ทรัพยากรน้อย ลักษณะข้อมูลสามารถอ่านหรือแก้ไขได้ง่ายผ่านจาวาสคลิปต์ โดยสามารถใช้ได้หลากหลาย เมื่อเปรียบเทียบเอ็กเอ็มแอล กับ เจสัน ประสิทธิภาพในการผ่านข้อมูลจากคำสั่งรูปแบบการเก็บข้อมูลนั้นเจสันสามารถเข้าถึงข้อมูลได้ง่ายและรวดเร็วกกว่าเอ็กเอ็มแอล ผู้วิจัยจึงได้เลือกใช้เจสันในงานวิจัยนี้

3. ขั้นตอนวิธีการดำเนินการวิจัย

งานวิจัยนี้ได้แบ่งขั้นตอนการวิจัยออกเป็น 6 ขั้นตอน คือ การเก็บรวบรวมข้อมูล การจัดการข้อมูล การวิเคราะห์ข้อมูล การเลือกรูปแบบการแสดงผล การพัฒนาระบบ และการประเมินระบบ

3.1 การเก็บรวบรวมข้อมูล ประกอบด้วย 3 แหล่งข้อมูลดังนี้

3.1.1 ข้อมูลสถิติการเกิดโรคพิษสุนัขบ้าในคน

สำนักระบาดได้เก็บข้อมูลสถิติผู้ป่วยและผู้เสียชีวิต ในฐานข้อมูลระบบรายงานการเฝ้าระวังโรค 506 ซึ่งจัดรวบรวมโดยสำนักงานสาธารณสุขจังหวัด โรงพยาบาล และสถานีอนามัยทุกแห่ง จากโรงพยาบาลรัฐทุกแห่ง โดยยังไม่ครอบคลุม

ทั้งหมดของโรงพยาบาลเอกชน ซึ่งสามารถนำออกมา ในรูปแบบไฟล์อาร์ทีเอฟ

3.1.2 ข้อมูลสถิติการเกิดโรคพิษสุนัขบ้าในสัตว์

ข้อมูลสถิติการเกิดโรคในสัตว์นั้นเป็นผู้รวบรวมและจัดเก็บข้อมูล โดยรวบรวมข้อมูลและจัดเก็บในฐานข้อมูลของสำนักควบคุม ป้องกัน และบำบัดโรคสัตว์ กรมปศุสัตว์ และ สามารถเข้าถึงได้จากเว็บไซต์ขององค์การเฝ้าระวังโรคระบาดสัตว์ระหว่างประเทศ

3.1.3 ข้อมูลสถิติด้านปริมาณน้ำฝนและอุณหภูมิ

3.2 การจัดการข้อมูล โดยนำเข้าข้อมูลสถิติโรคคน โรคสัตว์ อุณหภูมิ และ ปริมาณน้ำฝนที่ได้มาแปลงให้อยู่ในรูปแบบไฟล์เอ็กเซล

3.3 วิเคราะห์ข้อมูล โดยวิเคราะห์ข้อมูลแต่ละส่วน เพื่อหาความสัมพันธ์ในเบื้องต้น โดยจัดแบ่งกลุ่มข้อมูลเป็น 2 แบบ คือ แบบแรกเพื่อใช้สำหรับสร้างชุดข้อมูลในการเปรียบเทียบแต่ละปี ในส่วนนี้จะแบ่งกลุ่มข้อมูล จาก ปี 2005 – 2015 (2548 – 2558) แบบที่สองจะแบ่งข้อมูลออกเป็นช่วงตามฤดูกาลเพื่อใช้สร้างการแสดงผลที่ ฤดูกาล อุณหภูมิ และ ปริมาณน้ำฝนมาเกี่ยวข้อง

3.4 เลือกรูปแบบการแสดงผล

ขั้นตอนนี้จะทำการวิเคราะห์ รูปแบบ การแสดงข้อมูลหรือการทำจินตทัศน์ จะให้ผู้ประเมินทำการประเมิน จากแบบสอบถาม โดยให้เลือกการแสดงผลข้อมูลโรคระบาด ที่จัดไว้ในรูปแบบตารางข้อมูล กราฟ และ แผนที่ ผลการทดลองส่วนนี้ 63 % เลือกแผนที่ 16% เลือก กราฟ และ 22% เลือกตารางแสดงข้อมูล จากผลการทดลองในส่วนนี้ ผู้วิจัยจึงเลือกแสดงข้อมูลบนแผนที่ ซึ่งการใช้ข้อมูลเชิงพื้นที่ ร่วมกับข้อมูลสถิติ โดยเก็บข้อมูลทั้ง 2 ส่วน ในรูปเจสันไฟล์ ดังภาพที่ 2

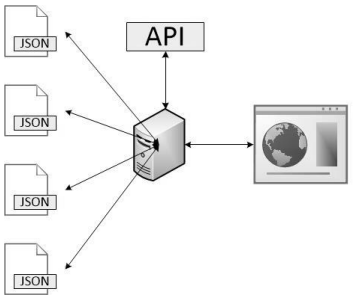

```
data_callback
(
{"type":"FeatureCollection","metadata":
{"generated":1395197681000},"features":
[
{ "type": "Feature", "properties": { "incident": 3 }, "geometry": {
"type": "MultiPolygon", "coordinates": [ [ [ [ 100.03302845043332,
13.458753372896366 ], [ 100.03280465676387,
13.458814906583894 ], ],…....
……. , [ 100.03302845043332, 13.458753372896366 ] ] ] ] } },
```


ภาพที่ 2 แสดงรูปแบบการเก็บข้อในรูปแบบไฟล์เจสัน

3.5 การพัฒนาระบบ

3.5.1 การออกแบบระบบ

เพื่อให้บุคคลทั่วไปสามารถเข้าถึงข้อมูลได้ จึงได้ออกแบบให้แสดงผลแผนที่ผ่านเว็บ โดยการเก็บข้อมูล ดังแผนภาพที่ 4 ซึ่งแสดงสถาปัตยกรรมการเชื่อมต่อระหว่างไฟล์เจสันที่เก็บข้อมูลทางภูมิศาสตร์และสถิติของการพบโรค อุณหภูมิ และปริมาณน้ำฝน โดยแสดงถึงการมีปฏิสัมพันธ์ของส่วนต่าง ๆ เรียกใช้เอพีไอ (Application Programing Interface: API) และแสดงผลข้อมูลที่ซ้อนทับบนบราว์เซอร์ดังภาพที่ 4



ภาพที่ 3 แผนภาพแสดงสถาปัตยกรรมการเชื่อมต่อข้อมูล

ในการทำงานไฟล์ไมโครซอฟต์เอ็กเซลล์ถูกประมวลจากผู้วิจัยก่อนจะแปลงเป็นไฟล์นามสกุลเจสัน ซึ่งจะถูกนำเข้าไปแสดงเป็นภาพบนแผนที่ หลังจากนั้นจะผ่านกระบวนการจินตทัศน์เพื่อให้มนุษย์ได้รับรู้ผ่านทางสายตาบุคคลทั่วไป

3.5.2 การออกแบบการแสดงผล

ในงานวิจัยนี้เป็นการซ้อนข้อมูล ที่แตกต่างกัน ซ้อนทับอยู่บนแผนที่เดียวกัน การออกแบบการแสดงผลในแต่ละชั้น จึงออกแบบให้แสดงผลในลักษณะ รูปหลายเหลี่ยม เพื่อเป็นสัญลักษณ์ ของข้อมูลแต่ละชนิดดังตารางที่ 2 โดยกำหนดการแสดงผลข้อมูลสภาพอากาศและปริมาณน้ำฝนด้วยสีดังตารางที่ 3

ตารางที่ 1 แสดงสัญลักษณ์ที่ใช้แทนข้อมูลที่ซ้อนบนแผนที่

ข้อมูล	สัญลักษณ์/การแสดงผล
โรคคน	
โรคสัตว์	
อุณหภูมิ	
ปริมาณน้ำฝน	

ตารางที่ 2 แสดงการแบ่งช่วงฤดูในประเทศไทย

อากาศ/ฝน	ค่า	สี
อากาศร้อน	35.0 - 39.9 องศาเซลเซียส	แดง
อากาศร้อนจัด	40.0 องศาเซลเซียส ขึ้นไป	ส้ม
อากาศหนาวจัด	ต่ำกว่า 8.0 องศาเซลเซียส	ม่วง
อากาศหนาว	8.0 -15.9 องศาเซลเซียส	น้ำเงินเข้ม
อากาศเย็น	16.0 -22.9 องศาเซลเซียส	น้ำเงิน
ฝนวัดจำนวนไม่ได้	0.1 มิลลิเมตร	เขียวเข้ม
ฝนเล็กน้อย	0.1 -10.0 มิลลิเมตร	เขียวอ่อน
ฝนปานกลาง	10.1 -35.0 มิลลิเมตร	เหลือง
ฝนหนัก	35.1 -90.0 มิลลิเมตร	ฟ้า
ฝนหนักมาก	90.1 มิลลิเมตรขึ้นไป	ฟ้าเข้ม

3.6 การประเมินระบบ

3.6.1 แบ่งกลุ่มการทดสอบ

ในการทดสอบจะใช้วิธีการสุ่มโดยไม่ได้จำกัดว่าจะต้องเป็นผู้ที่เกี่ยวข้องหรือผู้เชี่ยวชาญหรือไม่

3.6.2 อุปกรณ์ในการทดสอบและประเมิน

แบบสอบถามการใช้งานแบบออนไลน์ และระบบแผนที่ตัวอย่างที่พัฒนาแล้ว โดยแบบสอบถามแบ่งออกสองส่วน ส่วนแรกสอบถามความเกี่ยวข้อง ความรู้ของผู้ตอบแบบสอบถามและการเลือกรูปแบบการแสดงผลข้อมูลโรคระบาด ส่วนที่สองเป็นการทดลองใช้งานระบบเบื้องต้นจากข้อมูลตัวอย่าง

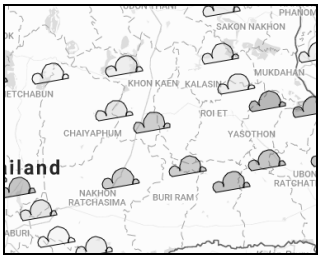
4. ผลการดำเนินการวิจัย

4.1 การจัดทำแผนที่

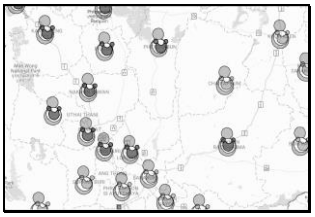
การทำจินตทัศน์ของระบบนี้เป็นการซ้อนข้อมูลในแต่ละชั้นเริ่มจาก ข้อมูลอุณหภูมิ ที่แสดงเป็นชั้นที่ 1 ดังภาพที่ 4 โดยซ้อนบนกูเกิลแมป ซ้อนข้อมูล ปริมาณน้ำฝนดังภาพที่ 5 เมื่อซ้อนทับข้อมูลโรคและสัตว์จะปรากฏดังภาพที่ 6 การแทนข้อมูลด้วยสัญลักษณ์ซ้อนทับข้อมูลบนแผนที่เดียวกันนั้นทำให้ข้อมูล อยู่ในรูปแบบที่เข้าใจง่ายสำหรับผู้ศึกษาและผู้สนใจด้านระบาดวิทยา



ภาพที่ 4 ภาพการซ้อนทับข้อมูลอุณหภูมิบนแผนที่



ภาพที่ 5 ภาพการซ้อนทับข้อมูลปริมาณน้ำฝนบนแผนที่



ภาพที่ 6 ภาพการซ้อนทับข้อมูลจำนวนอุบัติการณ์ (Incident) ของโรคพิษสุนัขบ้าในคนและสัตว์ บนแผนที่

ตารางที่ 3 แสดงผลการประเมินจากแบบสอบถาม

รายการคำถาม	คำตอบ	จำนวน	%
1. ท่านเป็นผู้ใช้ข้อมูล ศึกษา หรือทำงานด้านโรคคน และ/หรือ โรคสัตว์หรือไม่	ใช่	8	25
	ไม่ใช่	23	72
2. ท่านเป็นผู้สนใจข้อมูลการระบาดของโรคหรือไม่	ใช่/สนใจ	21	66
	ไม่ใช่/ไม่สนใจ	11	34
3. ท่านทราบหรือไม่ว่าแหล่งข้อมูลด้านโรคระบาดสามารถค้นหาได้จากที่ใด	ทราบ	15	47
	ไม่ทราบ	17	53
4. ท่านทราบหรือไม่ว่าอุณหภูมิมีผลต่อการระบาดของโรค	ทราบ	23	72
	ไม่ทราบ	8	25
5. ท่านคิดว่าระบบที่ได้พัฒนาขึ้นนี้มีประโยชน์กับท่านมากน้อยเพียงใด	มีประโยชน์มากที่สุด	8	25
	มีประโยชน์มาก	8	25
	มีประโยชน์	9	28
	มีประโยชน์น้อย	1	3
	ไม่มีประโยชน์	6	19

จากตารางที่ 5 ผลที่ได้จากการทดลองนั้นในส่วนของการเข้าถึงข้อมูล พบว่า 29% ของผู้ตอบแบบสอบถามเป็นผู้ใช้ หรือ

ศึกษาหรือทำงานด้านโรค แต่มีถึง 63% ของผู้ตอบแบบสอบถามทั้งหมดสนใจข้อมูลด้านการระบาดของโรค โดยมีเพียง 47% ของทั้งหมดเท่านั้นที่ทราบว่าจะได้ข้อมูลหรือแหล่งข้อมูลมาจากที่ใด

จากผลการทดลอง ผู้ทำการประเมิน เห็นว่าระบบที่นำเสนอนั้นมีประโยชน์มาก – มีประโยชน์มากที่สุด รวมทั้งสิ้น 78% ซึ่งมากกว่า มีประโยชน์น้อย หรือ ไม่มีประโยชน์ จึงสามารถสรุปผลได้ว่าการวิจัยได้ว่า ยอมรับสมมติฐาน H_0 คือ การนำข้อมูลโรคระบาดของคนและสัตว์ รวมถึงอุณหภูมิ และปริมาณน้ำฝนมาซ้อนทับ บนแผนที่เดียวมีประโยชน์ต่อผู้สนใจข้อมูลด้านโรคระบาด

5. สรุป

การเก็บข้อมูลด้านโรคระบาดในคนและสัตว์ ตลอดจนข้อมูลด้านอุณหภูมิและปริมาณน้ำฝนที่มีผลต่อการเกิดและการระบาดของโรคถูกจัดเก็บมีหน่วยงานที่รับผิดชอบแตกต่างกัน ไม่ได้มีการจัดทำข้อมูลแบบศูนย์ข้อมูล จึงจัดเก็บอยู่ในรูปแบบที่แตกต่างกัน จากเหตุผลที่กล่าวมาข้างต้น ทำให้ผู้ที่ใช้ข้อมูลยากต่อการเข้าถึงและยากต่อการเข้าใจ ดังนั้นงานวิจัยนี้จึงได้รวบรวมข้อมูลดังกล่าวนำมาคัดกรองและจัดเก็บให้อยู่ในรูปแบบที่สามารถเข้าถึงและเข้าใจได้ง่าย เพื่อเป็นประโยชน์ต่อผู้ที่สนใจหรือ เพื่อหาแนวทางป้องกัน หลีกเลี่ยงจากพื้นที่ที่เสี่ยงต่อการเกิดโรคระบาด

ปัญหาที่พบ คือ ข้อมูลที่รวบรวมมีความหลากหลาย ทั้งข้อมูลที่นำมาใช้ประโยชน์ได้และใช้ประโยชน์ไม่ได้ ทั้งความผิดพลาดของข้อมูลเช่น การสะกดชื่อจังหวัดไม่ต้องกัน ความละเอียดของข้อมูล จึงทำให้ยากต่อการแยกและรวบรวมข้อมูล เช่น การเก็บสถิติของโรคแท้งติดต่อ มีการเก็บสถิติในคน แต่ไม่มีการเก็บสถิติในสัตว์ทำให้ไม่สามารถนำมาแสดงความสัมพันธ์ได้

นอกจากนี้ผู้ทำการประเมินให้ข้อเสนอแนะว่าควรนำโรคระบาดอื่น ๆ มาแสดงด้วยซึ่งจะจัดทำในลำดับต่อไป

6. กิตติกรรมประกาศ

ขอขอบคุณ สำนักระบาดวิทยา สำนักควบคุม ป้องกันและบำบัดโรคสัตว์ กรมปศุสัตว์ กระทรวงเกษตรและสหกรณ์ และกรมอุตุนิยมวิทยา ที่ให้ความอนุเคราะห์ข้อมูลสำหรับงานวิจัยฉบับนี้

เอกสารอ้างอิง

[1] ญาใจ ลิ่มปิยะกรณ์, *การทำเหมืองข้อมูล*, ภาควิชาวิศวกรรมคอมพิวเตอร์ ตำราวิศว ฯ 100 ปี: กทม., พ.ศ. 2456 – 2556.

[2] สมบัติ อยู่เมือง, “ความหมายของคำว่า "ระบบสารสนเทศภูมิศาสตร์", 2542. [ออนไลน์]. Available: http://www.gisthai.org/about-gis/gis.html.

[3] ไพบูลย์ โล่ห์สุนทร, *ระบาดวิทยา*, กรุงเทพมหานคร: พิมพ์ครั้งที่ 7, โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย, พ.ศ. 2552.

[4] เจษฎา ทองเหม, “โรคพิษสุนัขบ้าในสัตว์ (Rabies)” *วารสารสัตวแพทย์ผู้ประกอบการบำบัดโรคสัตว์* ปีที่ 1 ฉบับที่ 3 ประจำเดือน กรกฎาคม – กันยายน 2544

[5] *Lecture notes of Zoonoses of Medical Importance (SIMI 603)* คณะแพทยศาสตร์ศิริราชพยาบาล, กทม. 1982

[6] S. Bresciani and M. J. Eppler, “ The Benefits of Synchronous Collaborative Information Visualization: Evidence from an Experimental Evaluation,” *IEEE Transactions on Visualization and Computer Graphics*, vol 15, no.16, pp. 1073 - 1080, Oct 23, 2009.

[7] W. S. Cleveland and R. McGill, “Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods,” *Journal of the American Statistical Association*, vol. 79, no. 387, pp. 531-554 , 1984.

[8] H. Zhang , M. Li, C. Chen, Z. Bao, Q. Huang and D. Cai, “Land use information release system based on Google Maps API and XML,” *2010 18th International Conference on Geoinformatics*, Beijing, 18-20 June 2010.

[9] A. N. Kuznetsov and M. S. Syzdykov, “Evaluation of spatial brucellosis distribution using the Geographic Information System: Towards building a high performance spatial epidemiology system for supervision on zoonotic infections,” *2014 IEEE 8th International Conference on Application of Information and Communication Technologies (AICT)*, Astana, 2014.

[10] H. Shi, Y. Zhang, J. Zhang, P. Wan and K. Shaw, “Development of Web-Based Epidemiological Reporting System for Tasmania Utilizing a Google Maps Add-On,” *9th Biennial Conference of the Australian Pattern Recognition Society on Digital Image Computing Techniques and Applications (DICTA 2007)*, Glenelg, Australia, , 2007, pp. 118-123.

[11] S. E. W. a. S. L. R. John E. Vena, " Applications of geographic information systems and geo-spatial modeling in environmental epidemiology," *Proceedings of the ITI 2013 3 5th International Conference on Information Technology Interfaces (ITI)*, Cavtat, 2013.

[12] กรมอุตุนิยมวิทยา หนังสืออุตุนิยมวิทยา "ฤดูกาลในประเทศไทย", [ออนไลน์]. Available: 2560 https://www.tmd.go.th/

[13] Y. Wu, Z. Liang and L. Liu, "Design and Implementation of Tourism Information System Based on Google Maps API," *2013 21st International Conference on Geoinformatics*, Kaifeng, pp. 1-4, 2013.

[14] Google , "Google Maps APIs", 2017. [online]. Available:https://enterprise.google.com/intl/en/maps/products/mapsapi.html

[15] Wikipedia, "Google Maps", 2017 https://en.wikipedia.org/wiki/Google_Maps

[16] Palo Alto, "Stop Comparing JSON and XML", November, 2017 [online]. Available: http://www.yegor256.com/2015/11/16/json-vs-xml.html

[17] IBM, "Introduction to XML" [online]. Available: https://www.ibm.com/developerworks/xml/tutorials/xmlintro/xmlintro.html

การเปรียบเทียบประสิทธิภาพการจำแนกข้อมูลโดยใช้เทคนิคการรวมกลุ่มจำแนกข้อมูล เพื่อวิเคราะห์ปัจจัยที่ส่งผลต่อการสอบวัดผลการทดสอบทางการศึกษาระดับชาติ ทางด้านอาชีวศึกษา

The Comparison of Data Classification Efficiency by using ENSEMBLE Technique for Analysis of Factors Affecting on the National Educational Test Results of Vocational Education (V-NET)

พุทธิดา สกุลวิริยกิจกุล (Puttida Sakulviriyakitkul)[1] จรัญ แสนราช (Charun Sanrch)[2] และ ธันว์รัชต์ สินธนะกุล (Thanrat Sintanakul)[3]

ภาควิชาคอมพิวเตอร์ศึกษา คณะครุศาสตร์อุตสาหกรรม มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

[1]puttidasak@aksorn.ac.th, [2]jsr@kmutnb.ac.th, [3]thanrat.s@fte.kmutnb.ac.th

บทคัดย่อ

บทความนี้นำเสนอผลการเปรียบเทียบประสิทธิภาพการจำแนกข้อมูลโดยใช้เทคนิคการรวมกลุ่มจำแนกข้อมูล (ENSEMBLE) ทั้ง 3 เทคนิคคือ Vote Ensemble, Bootstrap Aggregating และ Random Forest และนำอัลกอลิทึ่มที่มีประสิทธิภาพสูงสุดนำมาใช้วิเคราะห์ค้นหาปัจจัยที่ส่งผลต่อการสอบวัดผลการทดสอบทางการศึกษาระดับชาติ ทางด้านอาชีวศึกษาโดยการลดการนำเข้าทีละตัวแปร ข้อมูลวิจัยที่ใช้ในการวิจัยเป็นข้อมูลนักศึกษาจากวิทยาลัยเทคโนโลยีอักษรพัทยา ระหว่างปีการศึกษา 2556-2558 จำนวน 2,034 ชุด ข้อมูล 6 แอทริบิวต์ ผลการเปรียบเทียบพบว่า Random Forest ให้ค่าประสิทธิภาพความถูกต้องสูงสุด 100% ซึ่งมากกว่า Vote ENSEMBLE และ Bootstrap Aggregating ซึ่งมีค่าเท่ากับ 99.95% และ ผลการค้นหาปัจจัยที่ส่งผลต่อการสอบวัดผลการทดสอบทางการศึกษาระดับชาติ ทางด้านอาชีวศึกษา พบตัวแปรที่มีความสำคัญจากสำคัญมากไปหาน้อยได้ผลดังนี้ สาขาที่สังกัด ระดับชั้น และ เพศ

คำสำคัญ: เทคนิคการรวมกลุ่มจำแนกข้อมูล(ENSEMBLE) เทคนิคเหมืองข้อมูล การจำแนกข้อมูล เปรียบเทียบประสิทธิภาพ

Abstract

This article presented the comparative results of data classification using 3 techniques of ENSEMBLE : the Vote ENSEMBLE, Bootstrap Aggregating and Random Forest. The most efficient algorithms were used to analyze the factors affecting the National Educational Test of Vocational Education by reducing imports by one variable. The data used in the research were students from Aksorn Pattaya Technological College during the 2013-2016 academic year for 2,034 data set of 6 attributes. The comparative results showed that Random Forest had the highest accuracy of 100%, which was more than the Vote ENSEMBLE and Bootstrap Aggregating with the accuracy of 99.95%. The finding results of factors affecting the National Educational Test of Vocational Education revealed that the important variables were as follows: department, study level and gender

Keyword: ENSEMBLE Technique, Data Mining, Classification, Algorithm, Efficiency

1. บทนำ

ปัจจุบันในการจัดการศึกษาได้ให้ความสำคัญกับรูปแบบและกระบวนการที่ทำให้นักเรียนประสบความสำเร็จในการเรียนสูงสุด ไม่ว่าจะเป็นการนำรูปแบบการจัดการเรียนการสอนโดยใช้เทคนิคการสอนเข้ามาใช้ในชั้นเรียน รวมไปถึงนำเทคโนโลยีที่พัฒนามากขึ้นเข้ามาใช้ร่วมกัน กับการเรียนการสอน ทั้งหมดนั้นก็เพื่อประโยชน์ในการเรียนของนักเรียน เพื่อให้ได้ผลสัมฤทธิ์ทางการเรียนที่ดีขึ้น และสามารถประยุกต์ใช้รายวิชาที่หลากหลาย รวมไปถึงการใช้ชีวิตประจำวัน ดังนั้นเพื่อเป็นการวัดผลการเรียนรู้ของนักเรียน-

นักศึกษาจึงมีหน่วยงานมากมายจัดทดสอบวัดความรู้ในด้านต่างๆ ให้กับนักเรียน-นักศึกษา ไม่ว่าจะเป็นการทดสอบระดับสถานศึกษา สูงไปจน ระดับประเทศ การจัดการเรียนการสอนในอาชีวะศึกษาในปัจจุบันนั้น มีความก้าวหน้าและพัฒนาอย่างต่อเนื่องจะเห็นได้จากผลงานของนักศึกษาที่เลือกเรียนในสายวิชาชีพ เนื่องมาจากพัฒนาการจัดการเรียนการสอนที่นำกระบวนการเรียนรู้เข้ามาใช้ในชั้นเรียน ซึ่งเป็นการพัฒนาผลสัมฤทธิ์ทางการเรียนของนักเรียน-นักศึกษา โดยสามารถวัดผลได้หลายรูปแบบ ไม่ว่าจะเป็น การนำผลความคืบหน้าทางการเรียนมาทำการเปรียบเทียบกัน หรือ การวัดผลโดยการทดสอบต่างๆ โดยสถาบันที่จัดสอบ นอกจากนี้นักเรียนที่สำเร็จการศึกษาในขั้นประกาศนียบัตรวิชาชีพและประกาศนียบัตรวิชาชีพชั้นสูงนั้นจะต้องผ่านการทดสอบ โดยสถาบันการทดสอบทางการศึกษาแห่งชาติ (สทศ.) V-NET ย่อมาจาก Vocational National Education Testหมายถึง การทดสอบทางการศึกษาระดับชาติ ทางด้านอาชีวศึกษา เป็นการทดสอบวัดความรู้มาตรฐานทางวิชาการและวิชาชีพ ตามหลักสูตรประกาศนียบัตรวิชาชีพ (ปวช.) สำหรับนักศึกษาอาชีวศึกษาทีจะจบระดับชั้น ปวช.3 ผลการทดสอบ V-NET สามารถนำไปใช้ในการปรับปรุง พัฒนาการเรียนการสอนของครูอันจะส่งผลให้ผลสัมฤทธิ์ทางการเรียนของนักศึกษาสูงขึ้นได้ รวมทั้งการนำผลการสอบ V-NET ไปใช้ในการประกันคุณภาพการศึกษาด้วย [1]

วิทยาลัยเทคโนโลยีอักษรพัทยา สังกัดสำนักงานคณะกรรมการการอาชีวศึกษา กระทรวงศึกษาธิการ นั้นจัดการเรียนการสอนในระดับประกาศนียบัตรวิชาชีพ (ปวช.) และประกาศนียบัตรวิชาชีพชั้นสูง (ปวส.) ได้มีการจัดทดสอบทางการศึกษาระดับชาติ ทางด้านอาชีวศึกษา เป็นการทดสอบวัดความรู้มาตรฐานทางวิชาการและวิชาชีพ โดยสาขาที่เข้าร่วมการทดสอบประกอบไปด้วย สาขาการบัญชี สาขาการตลาด สาขาการโรงแรม สาขาคอมพิวเตอร์ธุรกิจ สาขาช่างไฟฟ้าและอิเล็กทรอนิกส์ และสาขาช่างยนต์ โดยใน 3 ปี ที่ผ่านมาพบว่ามีนักเรียนจำนวนไม่น้อยที่สามารถผ่านการทดสอบ โดยถ้าย้อนกลับไปดู ก็สามารถพบว่านักเรียนที่ผ่านการทดสอบนั้นเป็นนักเรียนที่มีผลการเรียนในระดับดี เมื่อตรวจสอบผลการทดสอบในปีการศึกษา 2558 พบว่า นักเรียนกว่าร้อยละ 40 ที่ผ่านการทดสอบนั้น มาจากนักเรียนที่มีผลการเรียนในระดับปานกลาง ถึง ต่ำ ดังนั้นจึงไม่สามารถบอกได้ทีเดียวว่า นักเรียนที่มีผลการเรียนดีนั้นจะสามารถผ่านการทดสอบ และการทดสอบดังกล่าวนั้นยังส่งผลต่อคุณภาพการจัดการศึกษาในภาพรวมซึ่งทางวิทยาลัยนั้นจำเป็นจะต้องมียอดนักเรียน-นักศึกษาผ่านการทดสอบไม่น้อยกว่าร้อยละ 50 [2]

การพยากรณ์ ในอนาคตนั้นเข้ามามีบทบาทอย่างมาก ไม่ว่าจะเป็น เรื่องของธุรกิจ การดำเนินชีวิต รวมไปถึงการศึกษา การพยากรณ์ในทางการศึกษานั้นมีข้อดีคือ เพื่อเป็นการวางแผนการจัดการเรียนการสอนในอนาคต นอกจากนี้ยังเป็นการวิเคราะห์ข้อมูลในอดีต และที่กำลังจะส่งผลถึงปัจจุบันอีกด้วย ซึ่งเทคโนโลยีต่างที่ใช้ในการ พยากรณ์หรือวิเคราะห์นั้นมีอยู่มากมาย ซึ่งหนึ่งในนั้นคือ เทคนิคการรวมกลุ่มจำแนกข้อมูล (ENSEMBLE) เทคนิคนี้ได้เป็นเทคนิคที่มีประสิทธิภาพสูง ดังจะเห็นได้จากงานวิจัยของ จักรกฤช ,ชลวิช และธนารักษ์ เรื่อง Boosting-based Ensemble learning with penalty profiles for automatic Thai unknown word recognition ที่ใช้เทคนิคการรวมกลุ่มจำแนกข้อมูล ในการวิเคราะห์เพื่อรูปแบบที่สามารถจำแนกหมวดหมู่ โดยเทคนิคการรวมกลุ่มจำแนกข้อมูลนั้นสามารถแบ่งออกได้ 3 ประเภท คือ Vote Ensemble, Bootstrap Aggregating (Bagging) และ Random Forest [3]

งานวิจัยชิ้นนี้มีวัตถุประสงค์เพื่อเปรียบเทียบเทคนิคที่จะนำมาใช้พยากรณ์ผลการทดสอบทางการศึกษาระดับชาติ ทางด้านอาชีวศึกษา วิทยาลัยเทคโนโลยีอักษรพัทยา โดยใช้เทคนิคการรวมกลุ่มจำแนกข้อมูล (Ensemble) เนื่องจากเป็นเทคนิคที่ใช้โมเดล classification หลายๆ โมเดลมาช่วยในการพยากรณ์และคาดคะเนผลการทดสอบ และ เพื่อศึกษาปัจจัยที่ส่งผลต่อผลการทดสอบทางการศึกษาระดับชาติ ทางด้านอาชีวศึกษา โดยทำการจัดลำดับความสำคัญของปัจจัยที่ใช้ในการพยากรณ์ ประกอบไปด้วย สาขาที่สังกัด เกรดเฉลี่ย และชั้นปี เป็นกรณีศึกษา

2. วรรณกรรมที่เกี่ยวข้อง

งานวิจัยชิ้นนี้ใช้เทคนิคการรวมกลุ่มจำแนกข้อมูล เพื่อการหาประสิทธิ์ภาพในการพยากรณ์และการจำแนกข้อมูล โดยใช้โปรแกรม Rapid Miner Studio ในการทดสอบเปรียบเทียบ

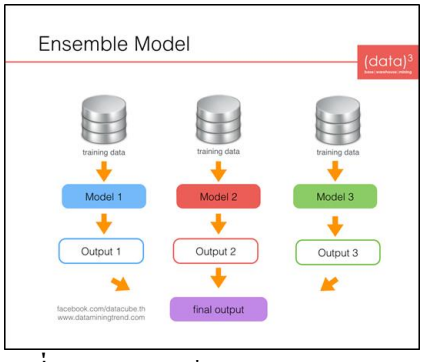
ประสิทธิภาพและค้นหาปัจจัยที่ส่งผลต่อการสอบวัดผลการทดสอบทางการศึกษาระดับชาติ ทางด้านอาชีวศึกษา

2.1 เหมืองข้อมูล (Data Mining)

การทำเหมืองข้อมูลเป็นกระบวนการเพื่อค้นหารูปแบบและความสัมพันธ์จากข้อมูลเป็นจำนวนมากที่ซ่อนอยู่ในชุดข้อมูล โดยอาศัยหลักสถิติ การรู้จำ รวมถึงการเรียนรู้ของเครื่อง ซึ่งเทคนิคที่ใช้ในการทำเหมืองข้อมูลแบ่งเป็น [4] กฎความสัมพันธ์ (Association Rule) การจำแนกข้อมูล (Data Classification) และการจัดกลุ่มข้อมูล (Data Clustering)

2.2 การจำแนกข้อมูล (Classification)

การจำแนกข้อมูล [4] เป็นการจำแนกประเภทของข้อมูล โดยเป็นวิธีที่ใช้ในการทำเหมืองข้อมูลเนื่องจากได้ตัวจำแนกที่มีองค์ความรู้เกี่ยวกับคุณสมบัติของข้อมูลแต่ละประเภท โดยเทคนิคที่ใช้ในการวิจัยในครั้งนี้ คือ เทคนิคการรวมกลุ่มจำแนกข้อมูล แบ่งออกได้ 3 ประเภทดังนี้ [5]



ภาพที่ 1 : การนำเทรนนิ่ง ดาต้า มาสร้างโมเดลต่างๆ

จากภาพที่ 1 เป็นการนำเทรนนิ่ง ดาต้า มาสร้างโมเดลต่างๆ โดยข้อมูลเทรนนิ่งดาต้าเหล่านี้จะเป็นข้อมูลชุดเดียวกันก็ตัวอย่างเช่น วิธีการ Vote Ensemble หรือจะเป็นข้อมูลที่ต่างกันก็ ตัวอย่างเช่น วิธี Bagging และ Random Forest หลังจากได้โมเดลมาชุดหนึ่งแล้วจึงนำไปทำนายข้อมูลที่ยังไม่รู้คำตอบ สำหรับการทำนายด้วยเทคนิคการรวมกลุ่มจำแนกข้อมูล ซึ่งมีหลายๆ โมเดลแต่ละโมเดลก็จะให้คำตอบออกมา ในขั้นตอนสุดท้ายจะต้องนำคำตอบมารวมกันเพื่อตรวจสอบคำตอบที่เหมาะสมที่สุด โดยใช้วิธีการ vote เลือกคำตอบที่ตอบตรงกันมากที่สุด หลักการสร้างโมเดลการรวมกลุ่มจำแนกข้อมูล คือโมเดลที่สร้างควรจะมีความหลากหลายเพื่อให้ทำนายข้อมูลแบบต่างๆ กันได้มาก การสร้างโมเดลที่หลากหลายนี้ทำได้โดยการใช้เทคนิค classification หลายๆ ประเภท หรือการสร้างเทรนนิ่ง ดาต้า ที่มีลักษณะต่างกัน จึงสรุปวิธีการสร้างโมเดลการรวมกลุ่มจำแนกข้อมูล ดังนี้

- Vote Ensemble เป็นการใช้เทรนนิ่ง ดาต้า ชุดเดียวกันแต่สร้างโมเดลด้วยเทคนิคต่างกัน เพื่อให้โมเดลมีความหลากหลายมากขึ้น การสร้างโมเดลจะใช้ 3 เทคนิคที่ต่างกัน
- Bootstrap Aggregating (Bagging) เป็นการสุ่มเทรนนิ่งดาต้าให้เป็นหลายชุด แต่สร้างโมเดลด้วยเทคนิคเดียวกันทั้งหมด โดยใช้การสร้างโมเดลด้วยเทคนิค classification เดียวกัน เช่น ใช้เทคนิค Decision Tree หรือ เทคนิค Neural Networks ทั้งหมด แม้ว่าจะเป็นเทคนิค Decision Tree เหมือนกันแต่ข้อมูลที่ใช้ในการสร้างโมเดลต่างกันก็ทำให้โมเดลที่สร้างขึ้นมาได้มีลักษณะที่ต่างกัน
- Random Forest เป็นเทคนิคที่คล้ายๆ กับ Bagging แต่แทนที่จะสุ่มข้อมูลอย่างเดียวก็ทำการสุ่มเลือกแอตทริบิวต์ (ฟีเจอร์) ต่างๆ ออกมาเป็นหลายๆ ชุดด้วย และสร้างโมเดลด้วยเทคนิค Decision Tree หลายต้น โมเดลด้วยการสุ่มแอตทริบิวต์ด้วยแทนที่จะเป็นการสุ่มเฉพาะข้อมูลตัวอย่างเพียงอย่างเดียว

2.3 งานวิจัยที่เกี่ยวข้อง

งานวิจัยครั้งนี้ได้ศึกษาเรื่องการเปรียบเทียบประสิทธิภาพการจำแนกข้อมูลโดยใช้เทคนิคการรวมกลุ่มจำแนกข้อมูล เพื่อวิเคราะห์ปัจจัยที่ส่งผลต่อการสอบวัลผลการทดสอบทางการศึกษาระดับชาติ ทางด้านอาชีวศึกษา ได้ศึกษางานวิจัยที่เกี่ยวข้องดังนี้

งานวิจัยของวิรชา และธวัชชัย[6]ได้เปรียบเทียบประสิทธิภาพระหว่างเทคนิค Decision Tree, C4.5, K-NN และ Naïve Bayes เพื่อใช้พยากรณ์ผลสอบมาตรฐานด้านเทคโนโลยีสารสนเทศ มหาวิทยาลัยราชภัฏนครปฐม พบว่าเทคนิค Naïve Bayes ให้ประสิทธิภาพสูงที่สุดเท่ากับ 94.35%

งานวิจัยของชลนิศา [7] ได้นำเทคนิคDecision Tree มาใช้สำหรับจำแนกสถานภาพการสำเร็จการศึกษาของนักศึกษา โดยแบ่งกลุ่มสถานภาพการสำเร็จการศึกษาออกเป็น 3 กลุ่มคือ

สำเร็จการศึกษา ไม่สำเร็จการศึกษาและกำลังศึกษา โดยงานวิจัยดังกล่าวจะทำการเปรียบเทียบค่า Precision และ Recall ที่ได้กับตัวแบบของเทคนิคโครงข่ายประสาทเทียม

งานวิจัยของชิดชนก ธนาวินท์และกฤษณะ [8] ได้นำเสนอการใช้เทคนิคDecision Tree เพื่อช่วยสำหรับการแก้ปัญหาการเลือกสาขาวิชาเรียนของนิสิตคณะวิศวกรรมศาสตร์ มหาวิทยาลัยเกษตรศาสตร์ และแบ่งกลุ่มความเหมาะสมของนิสิตเป็น 2 กลุ่มคือ Good และ Bad มีค่าความถูกต้องร้อยละ 80

3. วิธีการดำเนินการวิจัย

งานวิจัยครั้งนี้ได้ใช้เครื่องมือในการทำเหมืองข้อมูลคือโปรแกรม Rapid Miner Studio สำหรับการดำเนินการวิจัยและมีวิธีการวิจัยมีขั้นตอนดังนี้

3.1 การจัดเตรียมข้อมูล

งานวิจัยครั้งนี้ได้มีการจัดเตรียมข้อมูลโดยใช้ข้อมูลของนักเรียน-นักศึกษา ระดับประกาศนียบัตรวิชาชีพชั้นปีที่ 3 และระดับประกาศนียบัตรวิชาชีพชั้นสูงชั้นปีที่ 2 วิทยาลัยเทคโนโลยีอักษรพัทยา โดยข้อมูลนักเรียน-นักศึกษา เป็นข้อมูลปีการศึกษา 2556-2558 จำนวนทั้งสิ้น 2,394 ข้อมูลดังตารางที่ 1

ตารางที่ 1 : ตัวอย่างข้อมูลนักเรียน-นักศึกษา ปีการศึกษา 2556-2558

ระดับชั้น	แผนก	เพศ	เกรด	ปีการศึกษา	ผลสอบ V-Net
ปวช	บัญชี	หญิง	3.23	2556	ผ่าน
ปวส	โรงแรม	ชาย	2.9	2557	ไม่ผ่าน
…..	…..	…..	…..	…..	…..

จากนั้น นำข้อมูลไปปรับปรุงในส่วนของค่าข้อมูลที่ขาดหายไปโดยใช้วิธีการตัดเรคคอร์ดที่ข้อมูลที่ไม่สมบูรณ์ทิ้งทั้งหมด การเลือกแอตทริบิ้วต์เฉพาะที่เกี่ยวข้องและมีการปรับเปลี่ยนรูปแบบของข้อมูลบางแอตทริบิ้วต์เพื่อลดขนาดของข้อมูลที่ใช้ในการประมวลผล จากการปรับปรุงข้อมูลเรียบร้อยพบว่าข้อมูลที่สมบูรณ์พร้อมสำหรับการนำไปวิเคราะห์ในงานวิจัยครั้งนี้ประกอบด้วย 6 แอตทริบิ้วต์ และมีจำนวนเรคคอร์ดข้อมูลจำนวน 2,034 เรคคอร์ด สามารถแสดงตัวอย่างที่สมบูรณ์ได้ดังตารางที่ 2

ตารางที่ 2 : ตัวอย่างข้อมูลที่ผ่านกระบวนการจัดเตรียมอย่างสมบูรณ์

Sex	Major	Dep.	Grade	V-NET	Results
M	V	HT	3.23	Fail	Medium
F	HV	AC	2.9	Pass	Best
…..	…..	…..	…..	…..	…..

จากตารางที่ 2 แสดงตารางข้อมูลนักเรียน-นักศึกษา ระดับประกาศนียบัตรวิชาชีพชั้นปีที่ 3 และ ระดับประกาศนียบัตรวิชาชีพชั้นสูงชั้นปีที่ 2 วิทยาลัยเทคโนโลยีอักษรพัทยา โดยข้อมูลนักเรียน-นักศึกษา เป็นข้อมูลปีการศึกษา 2556-2558 ที่ผ่านการปรับปรุงและสมบูรณ์แล้ว ประกอบไปด้วย แอตทริบิ้วต์ Sex (เพศ) Major (ปวช,ปวส) Dep. (สาขาที่สังกัด) Grade (เกรดเฉลี่ยรวม) V-NET (ผลการทดสอบทางการศึกษาระดับชาติ ทางด้านอาชีวศึกษา) และ Results (ระดับผลการเรียน)

3.2 การออกแบบโมเดลเพื่อทำการทดลอง

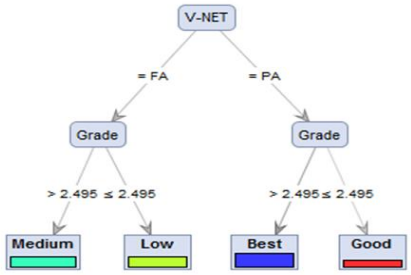
ผู้วิจัยได้ทำการออกแบบโมเดลตามรูปแบบของเทคนิคการรวมกลุ่มจำแนกข้อมูล และใช้โปรแกรม Rapid Miner Studio แบ่งออกเป็น 3 โมเดลดังนี้

- Vote Ensemble งานวิจัยนี้ได้เลือกใช้อัลกอลิทึม Decision Tree, เทคนิค Naive Bayes และเทคนิค Rule Induction ในการหาประสิทธิภาพของ เทคนิค Vote Ensemble เพราะข้อมูลนักเรียน-นักศึกษา ที่เก็บรวบรวมได้มาอยู่ในรูปแบบประเภทข้อมูลที่ไม่ใช่ตัวเลข (Nominal)
- Bootstrap Aggregating (Bagging) งานวิจัยครั้งนี้เลือกใช้ เทคนิค Rule Induction เนื่องจากเป็นอัลกอลิทึ่มที่ให้ประสิทธิภาพที่ดีที่สุดจากอัลกอลิทึ่มที่ทำการ Vote ในโมเดล Vote Ensemble ยังทำงานกับแอตทริบิวต์ที่เป็นรูปแบบประเภทข้อมูลที่ไม่ใช่ตัวเลข (Nominal) ซึ่งจะเห็นได้จากค่าความแม่นยำ (Accuracy) ดังตารางที่ 3

ตารางที่ 3 : ผลการหาประสิทธิภาพของโมเดลที่ใช้ในโมเดล Vote

Model	Accuracy	Recall	Precision	MAE
Decision Tree	75.41%	75.00%	62.66%	0.249
Naive Bayes	94.98%	94.33%	94.02%	0.107
Rule Induction	99.95%	99.95%	99.95%	0.001

- Random Forest เป็นอัลกอลิทึมที่ใช้โมเดลเทคนิค Decision Tree โดยจะสุ่มข้อมูลและสุ่มเลือกแอตทริบิวต์ต่างๆ จำนวน 10 ชุด และสร้างโมเดลด้วยเทคนิค Decision Tree ตามจำนวนที่ได้ตั้งไว้ใน Number of Trees ซึ่งผู้วิจัยทำการตั้งค่าไว้ที่ 10 ต้น ดังภาพที่ 2



ภาพที่ 2: ตัวอย่าง Decision Tree สร้างด้วย Random Forest

3.3 การประเมินผล

- การประเมินผลโมเดล จะได้จากการประเมินประสิทธิภาพซึ่งตรวจสอบจากค่าความถูกต้อง (Accuracy) ค่าความแม่นยำ (Precision) ค่าความระลึก (Recall) และ ค่าความคลาดเคลื่อน (MAE) ทุกกระบวนการประเมินผลโมเดลจะใช้โปรแกรม Rapid Miner Studio เป็นเครื่องมือในการหาประสิทธิภาพ
- การค้นหาปัจจัยโดยนำอัลกอลิทึมที่สร้างแบบจำลองที่ดีที่สุดจากการเปรียบเทียบ มาค้นหาปัจจัยที่ส่งผลต่อการสอบวัดผลการทดสอบทางการศึกษาระดับชาติ ทางด้านอาชีวศึกษาโดยการลดการนำเข้าทีละแอทริบิวต์ จากนั้นดูผลประสิทธิภาพ และสตรวจสอบค่าความแม่นยำ (Accuracy)

4. ผลการดำเนินงาน

งานวิจัยครั้งนี้สามารถสรุปผลการดำเนินงานได้โดยแบ่งเป็น 2 ส่วนดังนี้

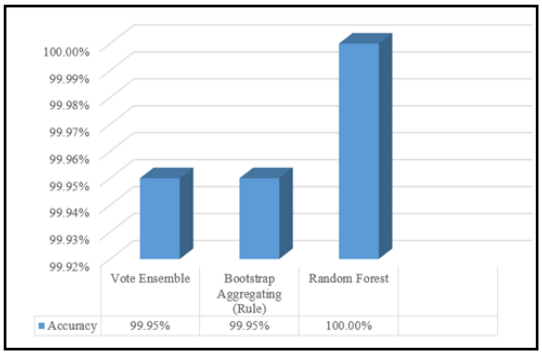
4.1 ผลการเปรียบเทียบประสิทธิภาพ

งานวิจัยครั้งนี้ได้ออกแบบโมเดลตามรูปแบบของเทคนิคการรวมกลุ่มจำแนกข้อมูล และใช้โปรแกรม Rapid Miner Studio ในการทดสอบและหาประสิทธิภาพของโมเดล สามารถแสดงการเปรียบเทียบได้ ดังตารางที่ 4

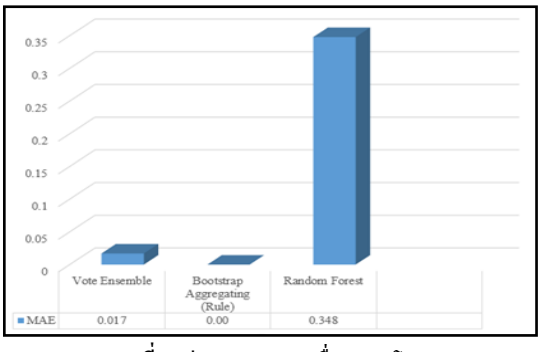
ตารางที่ 4 : ผลการหาประสิทธิภาพของโมเดลที่ใช้เทคนิคการรวมกลุ่มจำแนกข้อมูล

Algorithms	Accuracy	Recall	Precision	MAE
Vote Ensemble	99.95%	99.95%	99.97%,	0.017
Bootstrap Aggregating (Rule)	99.95%	99.95%	99.95%	0.001
Random Forest	100.00%	100.00%	100.00%	0.000

จากตารางที่ 4 แสดงให้เห็นผลการเปรียบเทียบค่าประสิทธิภาพของโมเดลที่ใช้เทคนิคการรวมกลุ่มจำแนกข้อมูลที่นำ 3 อัลกอลิทึม ซึ่งพบว่า Random Forest ให้ค่าประสิทธิภาพสูงสุด ที่ 100.00% และ ค่าความคลาดเคลื่อน (MAE) ที่ 0.000 และลำดับสุดท้ายพบว่า Vote Ensemble ให้ค่าประสิทธิภาพสูงสุด ที่ 99.95% และ ค่าความคลาดเคลื่อน (MAE) ที่ 0.017 ดังภาพที่ 3 และภาพที่ 4



ภาพที่ 3: ค่าประสิทธิภาพสูงสุดของโมเดล



ภาพที่ 4: ค่าความคลาดเคลื่อนของโมเดล

4.2 ผลการค้นหาปัจจัย

จากผลการหาประสิทธิภาพโมเดล ผู้วิจัยจึงนำอัลกอลิทึมที่มีผลประสิทธิภาพสูงสุด นำมาใช้ในการค้นปัจจัยที่มีผลต่อการสอบวัดผลการทดสอบทางการศึกษาระดับชาติ ทางด้านอาชีวศึกษาโดยการลดการนำเข้าทีละแอทริบิวต์ แสดงแอทริบิวต์ที่มีความสำคัญตามลำดับดังนี้

ตารางที่ 4 : ผลการค้นหาปัจจัยที่มีผลต่อการสอบวัดผลการทดสอบทางการศึกษาระดับชาติ ทางด้านอาชีวศึกษา

Factors	Accuracy	No.
SEX	81.80%	3
Major	61.98%	2
Dep.	39.20%	1

จากตารางที่ 4 แสดงให้เห็นผลการค้นปัจจัยที่มีผลต่อการสอบวัดผลการทดสอบทางการศึกษาระดับชาติ ทางด้านอาชีวศึกษาโดยใช้เทคนิคการรวมกลุ่มจำแนกข้อมูลที่มีประสิทธิภาพสูงสุด ซึ่งอัลกอลิทึมที่ใช้ในการค้นหาคือ Random Forest ซึ่งสามารถเรียงตามลำดับความสำคัญของปัจจัยได้ดังนี้ Dep. (สาขาที่สังกัด), Major (ปวช, ปวส) และ Sex (เพศ)

5. อภิปรายและสรุปผลที่ได้จากการวิจัย

การวิจัยครั้งนี้เป็นการนำเสนอผลการทดสอบประสิทธิภาพของ นำเสนอผลการเปรียบเทียบประสิทธิภาพการจำแนกข้อมูลโดยใช้เทคนิคการรวมกลุ่มจำแนกข้อมูล ทั้ง 3 คือ Vote Ensemble, Bootstrap Aggregating ,Random Forest และนำอัลกอลิทึ่มที่มีประสิทธิภาพสูงสุดนำมาใช้วิเคราะห์ค้นหาปัจจัยที่ส่งผลต่อการสอบวัดผลการทดสอบทางการศึกษาระดับชาติ ทางด้านอาชีวศึกษาโดยการลดการนำเข้าทีละตัวแปร ผลการวิจัยสรุปได้ว่า อัลกอลิทึม Random Forest ให้ค่าประสิทธิภาพสูงสุด ที่ 100.00% และ ค่าความคลาดเคลื่อน (MAE) ที่ 0.000 ซึ่งมีประสิทธิภาพในการพยากรณ์แม่นยำมากที่สุด และใช้ในการวิเคราะห์เพื่อค้นปัจจัยที่มีผลต่อการสอบวัดผลการทดสอบทางการศึกษาระดับชาติ ทางด้านอาชีวศึกษาโดยการลดการนำเข้าทีละแอทริบิวต์ ซึ่งสามารถเรียงตามลำดับความสำคัญของปัจจัยได้ดังนี้ Dep. (สาขาที่สังกัด), Major (ปวช,ปวส) และ Sex (เพศ) จากผลการวิจัยในครั้งนี้จะสามารถพัฒนาต้นแบบกระบวนการวิเคราะห์ปัจจัยที่ส่งผลต่อผลสัมฤทธิ์ทางการเรียนด้วยกระบวนการพยากรณ์ รวมไปถึงการประยุกต์ใช้เทคนิคเหมืองข้อมูลกับการวิเคราะห์พฤติกรรมที่ส่งผลต่อผลสัมฤทธิ์ทางการเรียนอีกด้วย

เอกสารอ้างอิง

[1] สถาบันการทดสอบทางการศึกษาแห่งชาติ, “การสอบวัดผลการทดสอบทางการศึกษาระดับชาติ ทางด้านอาชีวศึกษา” เข้าถึงได้จาก http://www.niets.or.th/th/catalog/view/242 สืบค้นเมื่อ ม.ค. 2560

[2] วิทยาลัยเทคโนโลยีอักษรพัทยา, “ผลการเรียนชั้นปีสุดท้าย” เข้าถึงได้จาก http://a-tech.aksorn.ac.th/ สืบค้นเมื่อ ม.ค. 2560

[3] Jakkrit TeCho,Cholwich Nattee and Thanaruk Theeramunkong, “Boosting-based Ensemble learning with penalty profiles for automatic Thai unknown word recognition,” *International Journal of ELSEVIER*, Vol. 63, No. 6, pp. 1117-1134, 2012.

[4] Jiawei Han and Micheline Kamber, “*Data Mining Concepts and Technique*” San Francisco, Morgan Kqufmann Publishers, 2006.

[5] Data Mining Trend, “การสร้างโมเดล ENSEMBLE” เข้าถึงได้จาก https://goo.gl/JDiYLO สืบค้นเมื่อ ม.ค. 2560

[6] วิรชา วิรัชกุล และ ธวัชชัย งามสันติวงศ์, “*เปรียบเทียบประสิทธิภาพระหว่างเทคนิค Decision Tree C4.5, K-NN และ Naïve Bayes เพื่อใช้พยากรณ์ผลสอบมาตรฐานด้านเทคโนโลยีสารสนเทศ มหาวิทยาลัยราชภัฏนครปฐม*” การประชุมวิชาการระดับชาติ มหาวิทยาลัยราชภัฏภูเก็ต ครั้งที่ 4 พ.ศ. 2558

[7] ชลนิศา สาระ, “การจำแนกกลุ่มสถานภาพการสำเร็จการศึกษาโดยแบบจำลองDecision Tree”, ภาควิชาวิทยาการคอมพิวเตอร์และสารสนเทศ บัณฑิตวิทยาลัย มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2550.

[8] ชิดชนก ส่งศิริ, ธนาวินท์ รักธรรมานนท์ และกฤษณะ ไวยมัย, “*การใช้เทคนิค Data Mining เพื่อค้นหาภาควิชาที่เหมาะสมที่สุดให้กับนิสิต*” การประชุมทางวิชาการของมหาวิทยาลยั เกษตรศาสตร์ ครั้งที่ 39. พ.ศ. 2544 หน้า ที่ 43-50.

การศึกษาเปรียบเทียบ NETPIE กับ แพลตฟอร์ม Internet of Things อื่น
Comparison Study of NETPIE with other IoT Platforms

ขวัญชนก ศรีมูล (Khwanchanok Srimool)[1] ฐาปนี ฉายากุล (Thapanee Chayakul)[2] เอมอัชนา นิรันตสุขรัตน์(Aimaschana Niruntasukrat)[3] พนิตา พงษ์ไพบูลย์ (Panita Pongpaibool)[4] สุขุมาล กิติสิน(Sukumal Kitisin)[5]

[1,2,5]วิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยเกษตรศาสตร์

[3,4]ห้องปฏิบัติการวิจัยนวัตกรรมอินเทอร์เน็ต ศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ

khwanchanok.sm@ku.th, thapanee.ck@ku.th, aimaschana.Niruntasukrat@nectec.or.th, panita.Pongpaibool@nectec.or.th,sukumal.i@ku.th

บทคัดย่อ

Internet of Things เป็นเทคโนโลยีที่เข้ามามีบทบาทสำคัญและมีคุณค่าทางเศรษฐกิจ NECTEC ได้พัฒนา NETPIE เพื่อเป็นแพลตฟอร์มบนคลาวด์ที่ให้บริการการเชื่อมโยงอุปกรณ์ IoT ต่างๆเข้าด้วยกันแบบไร้รอยต่อ งานวิจัยนี้เป็นการศึกษาพัฒนาเพื่อทดลองเปรียบเทียบการทำงานและประสิทธิภาพของแพลตฟอร์ม NETPIE กับแพลตฟอร์ม Internet of Things อื่นๆ ได้แก่ Anto, AWS IoT(Amazon), Azure IoT Hub(Microsoft), Blynk, Firebase Realtime Database(Google) และ IBM Watson Internet of Things (IBM Bluemix) เพื่อให้เห็นความแตกต่างของแต่ละแพลตฟอร์มและสามารถเลือกใช้ได้อย่างเหมาะสม

คำสำคัญ: Internet of Things NETPIE IoT แพลตฟอร์ม

Abstract

This paper describe an experimental study on Internet of Things platform by comparing NETPIE, an IoT platform developed by NECTEC, with other internet of things platforms, including Anto, AWS IoT(Amazon), Azure IoT Hub(Microsoft), Blynk, Firebase Realtime Database(Google) and IBM Watson Internet of Things (IBM Bluemix). The objective of the study is to compare them in various aspects so that the developers can choose an appropriate platform.

Keyword: Internet of Things, NETPIE, IoT platforms

1. บทนำ

เทคโนโลยี Internet of Things (IoT) ได้รับความสนใจอย่างกว้างขวาง เนื่องจากเทคโนโลยี IoT จะเปลี่ยนรูปแบบและกระบวนการผลิตในภาคอุตสาหกรรมไปสู่ยุคใหม่ มีนักพัฒนาที่สร้างสรรค์ชิ้นงานเพื่อสร้างธุรกิจบนเทคโนโลยี IoT เกิดขึ้นมากมาย ผู้ให้บริการแพลตฟอร์มสำหรับการสื่อสารเชื่อมโยงอุปกรณ์ IoT เข้าด้วยกันและให้บริการ API เพื่ออำนวยความสะดวกต่อนักพัฒนาก็มีหลากหลายเช่นกัน แต่ละบริการก็มีจุดเด่น และบริการที่แตกต่างกันไป เพื่อที่นักพัฒนาจะสามารถเลือกใช้บริการได้เหมาะสมกับงานและความต้องการของตน การศึกษาจุดเด่นและจุดด้อยของแต่ละบริการของแพลตฟอร์มต่างๆ จึงมีประโยชน์อย่างยิ่ง

2. IoT แพลตฟอร์ม

แพลตฟอร์ม IoT ที่ให้บริการในปัจจุบันมีทั้งสัญชาติไทยดังเช่น Network Platform for Internet of Everything (NETPIE) ซึ่งพัฒนาโดย NECTEC และ Anto ของบริษัทเอกชน และแพลตฟอร์มของต่างประเทศ อาทิเช่น AWS IoT(Amazon) Azure IoT Hub(Microsoft) Blynk Firebase Realtime Database(Google) IBM Watson Internet of Things (IBM Bluemix) แพลตฟอร์มต่างๆมีการรองรับภาษาและบริการที่แตกต่างกัน ดังภาพที่ 1 และภาพที่ 2 ตามลำดับ

	MICROSOFT AZURE IOT HUB		AMAZON AWS IOT	Bluemix Watson Internet of Things		Firebase	Anto	Blynk	NETPIE
	Device	Service		Device	Application				
C++	✓	x	✓ (Arduino Yún)	✓	x	x	✓ (ESP8266)	✓	✓ (Arduino Mega, ESP8266)
Node.js	✓	✓	✓	✓	✓	x	x	x	✓
JavaScript	x	x	x	x	x	✓	x	x	✓
Java	✓	✓	✓	✓	✓	x	x	x	✓
.NET	✓	✓	x	x	x	x	x	x	x
Python	✓	x	✓	✓	✓	x	x	x	✓
Ruby	x	x	x	x	✓	x	x	x	x
PHP	x	x	x	x	✓	x	x	x	x
Go	x	x	x	x	✓	x	x	x	x
Android	x	x	✓	x	x	✓	x	x	✓
Swift	x	x	✓	x	x	✓	x	x	x
REST API	✓	x	x	x	x	✓	✓	x	✓

ภาพที่ 1:แสดง Programming Languages ที่ IoT Platforms รองรับ

	MICROSOFT AZURE IOT HUB	AMAZON AWS IOT	Bluemix Watson Internet of Things	Firebase	Anto	Blynk	NETPIE
Data Storage	✓	✓	✓	✓ (Realtime Database)	x	✓ (Blynk Local Server)	✓ (NETPIE Feed)
Visualization	✓	✓	✓	✓	✓	✓	✓
Encryption Connection	✓	✓	✓	✓	✓	✓	✓
Analytics	✓ (Azure Stream Analytics)	✓ (Rule Engine, Amazon Kinesis, AWS Lamda)	✓ (IBM IoT Real-time Insights)	✓ (Firebase Analytics)	x	x	x

ภาพที่ 2:**แสดง** Feature **ต่างๆ ที่** IoT Platforms **รองรับ**

3. การเปรียบเทียบ NETPIE กับแพลตฟอร์ม IoT อื่น

3.1 การวิเคราะห์และออกแบบการทดลอง

ผู้วิจัยได้ออกแบบการวิเคราะห์เปรียบเทียบทั้งในเชิงปริมาณและเชิงคุณภาพโดยประเด็นในการวิเคราะห์เปรียบเทียบเชิงคุณภาพมีดังนี้ การเริ่มต้นใช้งานสำหรับผู้เริ่มต้นว่ายากหรือง่ายเพียงใด การออกแบบหน้าเว็บไซต์ใช้งานสะดวกหรือไม่ ตัวอย่างที่มีในเว็บไซต์เพียงพอหรือไม่ การเขียนโปรแกรมโดยใช้บริการนั้นๆ ยากหรือง่ายเพียงใด และบริการนั้นๆ รองรับภาษาและแพลตฟอร์มแบบใดบ้าง

ประเด็นการวิเคราะห์เปรียบเทียบเชิงปริมาณมีดังนี้ ทดลองวัดเวลาไป-กลับในการส่งข้อความ (Round-trip time) ทดลองวัดปริมาณข้อมูลที่ Receiver ได้รับจาก Sender ในช่วงเวลาหนึ่งๆ (Throughput)

จากประเด็นที่ต้องการศึกษา ประกอบกับคุณสมบัติและบริการของแต่ละแพลตฟอร์มที่สรุปได้ดังภาพที่ 1 และภาพที่ 2 จึงได้มีการออกแบบการทดลองโดยการจัดกลุ่มการทดลองตามแพลตฟอร์มหรือบอร์ดที่บริการต่างๆรองรับ โดยมีจุดประสงค์เพื่อวัดระยะเวลาที่ใช้ในการส่งข้อความจากอุปกรณ์หนึ่งไปยังอีกอุปกรณ์หนึ่ง ในที่นี้จะเรียกว่า Sender และ Receiver โดยมี Sender เป็นผู้ลงบันทึกเวลาตั้งแต่เริ่มส่งข้อความออกไปจนกระทั่งได้รับข้อความตอบกลับมาจาก Receiver ส่วน Receiver มีหน้าที่รอรับข้อความจาก Sender เมื่อได้รับแล้วทำการปิด-เปิดไฟที่เชื่อมต่อกับบอร์ด แล้วส่งข้อความกลับไปยัง Sender ทั้งนี้ Receiver จะเป็น NodeMCU บอร์ดเสมอ ในขณะที่ Sender อาจจะเปลี่ยนแปลงตามความเหมาะสมกับการทดลองนั้นๆ

การศึกษาเปรียบเทียบแบ่งออกเป็น 2 แบบ แบบแรกคือการวัดเวลาไป-กลับในการส่งข้อความ (Round-trip time) แบบที่สองเป็นการทดลองวัดปริมาณข้อมูลที่ Receiver ได้รับจาก Sender ในช่วงเวลาหนึ่งๆ (Throughput) โดยแบ่งออกเป็น 4 กลุ่มการทดลองดังนี้

การทดลองกลุ่มที่ 1 NodeMCU และ NodeMCU เป็นการส่งข้อความระหว่าง NodeMCU สองตัว ตัวหนึ่งทำหน้าที่เป็น Sender อีกตัวเป็น Receiver โดยใช้ภาษา C++ ในการพัฒนา โดยแพลตฟอร์มที่ใช้ในกลุ่มการทดลองนี้ได้แก่ Anto และ NETPIE

การทดลองกลุ่มที่ 2 Node.js และ NodeMCU เป็นการทดลองโดยมีโปรแกรมที่พัฒนาด้วย Node.js เป็น Sender ส่งคำสั่งปิด-เปิดไฟไปยัง NodeMCU ซึ่งทำหน้าที่เป็น Receiver พัฒนาด้วยภาษา C++ แพลตฟอร์มที่ใช้ในกลุ่มการทดลองนี้ ได้แก่ Bluemix Watson IoT Azure IoT Hub และ NETPIE การทดลองนี้จัดทำขึ้นเนื่องจาก Bluemix Watson IoT และ Azure IoT Hub ไม่สนับสนุนการส่งข้อความระหว่างอุปกรณ์สองตัวซึ่งพัฒนาด้วย C++ ทั้งคู่ ทั้งสองแพลตฟอร์มมีการแบ่งฟังก์ชันการทำงานระหว่างตัวส่งและตัวรับข้อมูลอย่างชัดเจน โดยตัวรับจะไม่สามารถสื่อสารกับตัวรับด้วยกันได้ ตัวส่งก็เช่นกัน ทำให้ทั้งสองแพลตฟอร์มนี้ไม่สามารถทำการทดลองตามที่ออกแบบไว้ในการทดลองที่ 1 ได้

การทดลองกลุ่มที่ 3 Python และ Python บน Raspberry Pi เป็นการทดลองโดยมีโปรแกรมที่พัฒนาด้วยภาษา Python เป็น Sender ส่งคำสั่งปิด-เปิดไฟไปยัง Raspberry Pi ซึ่งทำหน้าที่เป็น Receiver ซึ่งพัฒนาด้วยภาษา Python เช่นเดียวกัน แพลตฟอร์มที่ใช้ในกลุ่มการทดลองนี้ได้แก่ Bluemix Watson IoT, AWS IoT และ NETPIE

การทดลองกลุ่มที่ 4 REST API ระหว่าง Python และ Python บน Raspberry Pi เป็นการทดลองโดยมีโปรแกรมที่พัฒนาด้วย REST API บนภาษา Python เป็น Sender ส่งคำสั่งปิด-เปิดไฟไปยัง Raspberry Pi ซึ่งทำหน้าที่เป็น Receiver พัฒนาด้วย REST API บนภาษา Python เช่นกัน แพลตฟอร์มที่อยู่ในกลุ่มการทดลองนี้ได้แก่ Firebase Anto และ NETPIE

3.2 การวัดผลการทดลอง

แบ่งเป็นสองส่วน คือ การทดลองวัด Round-trip time ทำการวัดเวลาทุกๆครึ่งชั่วโมงโดยจับเวลาทั้ง connection time และ sending time ตั้งแต่วันที่ 20 ธันวาคม พ.ศ.2559 เวลา 18:00น. ถึงวันที่ 23 ธันวาคม พ.ศ.2559 เวลา 12:00น. และการทดลองวัด Throughput โดยการส่งข้อความทุก 3000 มิลลิวินาที ในเวลา 30 นาที แล้วเพิ่มความถี่ในการส่งข้อความเป็นทุก 1000 มิลลิวินาที และ 500 มิลลิวินาที ตามลำดับ

สมมติฐานการทดลอง

1. การส่งข้อติดต่อกันหลายๆครั้งและเร็วขึ้นจะทำให้เกิดการสูญหายของข้อมูล

2. การส่งข้อความไป-กลับของแต่ละแพลตฟอร์มมีเวลาที่แตกต่างกัน

4. การเปรียบเทียบ NETPIE กับแพลตฟอร์ม IoT อื่น

จากการศึกษาคุณสมบัติของแต่ละแพลตฟอร์มได้ตารางสรุปการแบ่งกลุ่มการทดลอง ดังภาพที่ 3 ผู้วิจัยเริ่มจากการใช้ NodeMCU การทดลองนี้ เป็นการสั่งงานปิด-เปิดไฟบนบอร์ดผ่านหน้าเว็บไซต์ด้วยภาษา HTML และ JavaScript แต่เนื่องจากมีบริการที่รองรับการเขียนโดยใช้ JavaScript บนแพลตฟอร์มเพียง Firebase และ NETPIE เท่านั้น ทำให้ไม่สามารถทำการทดลองสำเร็จได้ทั้งหมด

	MICROSOFT AZURE IOT HUB	AMAZON AWS IOT	Bluemix Watson Internet of Things	Firebase	Anto	Blynk	NETPIE
NodeMCU	✓	x	✓	✓	✓	✓	✓
Raspberry Pi	✓	✓	✓	✓	x	✓	✓

ภาพที่ 3:แพลตฟอร์มต่างๆที่ NodeMCU **และ** RaspberryPi **รองรับ**

จากปัญหาที่พบ จึงได้มีการออกแบบการทดลองใหม่ (ช่วงที่2) ผลการทดลอง (ช่วงที่ 2) จากการทดลองสั่งงานผ่านเว็บที่เขียน โดย JavaScript ที่มีการออกแบบการทดลองใหม่ โดยให้อุปกรณ์สองตัวสื่อสารกันโดยตรง โดยตัวหนึ่งทำหน้าที่เป็นตัวส่งคำสั่ง และอีกตัวเป็นตัวรับคำสั่งและปิด-เปิดไฟ ซึ่งจะตัดปัญหาเรื่องความแตกต่างของภาษา และแพลตฟอร์มออกไปได้ โดยแบ่งกลุ่มการทดลองตามข้อมูลจากภาพที่ 1 ได้ดังนี้

กลุ่มที่ 1 NodeMCU และ NodeMCU

การทดลองเปรียบเทียบเวลาส่งข้อความไป-กลับระหว่าง NodeMCU 2 ตัว โดยใช้ IoT แพลตฟอร์ม Firebase Anto และ NETPIE

กลุ่มที่ 2 Node.js และ NodeMCU(ESP8266)

การทดลองเปรียบเทียบเวลาส่งข้อความไป-กลับระหว่าง Node.js และ NodeMCU(ESP8266) โดยเขียนบน C++ ในการเปิด-ปิดไฟ LED บน NodeMCU(ESP8266) จากภาพที่ 1 IoT แพลตฟอร์มที่รองรับ Node.js ได้แก่ Microsoft Azure IoT Hub, Bluemix Watson Internet of Things และ NETPIE

กลุ่มที่ 3 โปรแกรม Python และ โปรแกรม Python บน RaspberryPi

การทดลองเปรียบเทียบเวลาส่งข้อความไป-กลับระหว่าง Python และ Python บน RaspberryPi ในการเปิด-ปิดไฟ จากภาพที่ 1 IoT Platform ที่รองรับภาษา Python มี Amazon AWS IoT, Bluemix Watson Internet of Things, Firebase และ NETPIE

จากการทดลองทั้ง 3 กลุ่มในครั้งนี้ ซึ่งได้แก่ กลุ่ม NodeMCU และ NodeMCU, กลุ่มที่ 2 Node.js และ NodeMCU (ESP8266) และ กลุ่มที่ 3 โปรแกรม Python และโปรแกรม Python บน RaspberryPi มีการวัดผล 3 แบบ คือ

1. การวัดเป็น 2 รอบละ 20 ครั้ง แบ่งเป็นช่วงเวลาประกอบไปด้วย ช่วงสาย 9:00-11:0น. และช่วงบ่าย 14:00-16:00น.
2. การวัดเป็น 2 รอบละ 50 ครั้ง แบ่งเป็นช่วงเวลาประกอบไปด้วย ช่วงสาย 9:00-11:00น. และช่วงบ่าย 14:00-16:00น.
3. การวัดเป็น 2 รอบละ 100 ครั้ง แบ่งเป็นช่วงเวลาประกอบไปด้วย ช่วงสาย 9:00-11:00น. และช่วงบ่าย 14:00-16:00น.

จากการทดลองพบว่า Firebase ใช้เวลานาน และผลที่ได้มีแปรปรวนสูง ทำให้การทดลองไม่สามารถสรุปผลออกมาได้ จึงได้มีการออกแบบการทดลองใหม่ (ช่วงที่ 3)

ผลการทดลอง (ช่วงที่ 3)

4.1 ผลการทดลองวัด Round-trip time

จากผลการทดลองช่วงที่ 2 พบว่า Library ของ Firebase นั้น เป็น REST API ทั้งภาษา Python และ C++ จึงได้มีการออกแบบการทดลองใหม่ โดยเพิ่มกลุ่มที่เป็น REST API เป็นกลุ่มที่ 4 ในการทดลอง มีรายละเอียดดังนี้

กลุ่มที่ 1 NodeMCU และ NodeMCU

IoT Platform ที่จะนำมาเปรียบเทียบ ได้แก่ Anto และ NETPIE โดย NodeMCU ที่เป็น Sender จะต่อกับ Module SD Card เพื่อทำการบันทึกเวลา และ NodeMCU ที่เป็น Receiver จะทำการส่งข้อความตอบกลับมายัง Sender และทำการเปิด-ปิดไฟ

ตารางที่ 1 กลุ่มที่ 1 NodeMCU และ NodeMCU

แพลตฟอร์ม	Connection time (ms)		Sending time (ms)	
	Avg.	S.D.	Avg.	S.D.
Anto	194.52	121.54	313.29	149.86
NETPIE	2386.92	165.32	196.94	63.91

กลุ่มที่ 2 Node.js และ NodeMCU(ESP8266)

ใช้กลุ่มการทดลองเดิม ตามการทดลองช่วงที่ 2

ตารางที่ 2 กลุ่มที่ 2 Node.js และ NodeMCU(ESP8266)

แพลตฟอร์ม	Connection time (ms)		Sending time (ms)	
	Avg.	S.D.	Avg.	S.D.
Azure	1036.77	282.88	770.32	242.63
Bluemix	3024.56	630.63	1006.56	204.26
NETPIE	3726.62	793.40	313.10	161.07

กลุ่มที่ 3 Python และ Python บน RaspberryPi

IoT Platform ที่จะนำมาเปรียบเทียบ ได้แก่ Amazon AWS IoT, Bluemix Watson Internet of Things และ NETPIE

ตารางที่ 3 กลุ่มที่ 3 Python และ Python บน RaspberryPi

แพลตฟอร์ม	Connection time (ms)		Sending time (ms)	
	Avg.	S.D.	Avg.	S.D.
AWS IoT	3184.90	233.40	2273.90	793.40
Bluemix	1969.58	447.84	3244.72	464.66
NETPIE	15714.62	2038.74	340.68	202.46

กลุ่มที่ 4 REST API ภาษา Python และ REST API ภาษา Python บน RaspberryPi

Microsoft Azure IoT Hub จะแบ่งการรับ-ส่งข้อความ โดยมองเป็น Device กับ Service ซึ่ง Device สามารถส่งข้อความไปยัง Cloud และ Service รับข้อความจาก Cloud มายัง Service จากการทดลองเป็นการทดลองระหว่าง REST API ภาษา Python บน Raspberry 2 ตัว เนื่องจากความสามารถของ Azure ใน REST API นั้นจะเป็นแค่ Device ที่ส่งข้อความไปยัง Cloud เท่านั้น ไม่สามารถเขียนให้ Device รับข้อความจาก Cloud ได้ จึงไม่สามารถทำให้ REST API ภาษา Python บน Raspberry 2 ตัวส่งข้อความหากันได้ และเนื่องจากกลุ่มนี้เป็น REST API จึงไม่สามารถวัด Connection time ได้ IoT Platform ที่นำมาเปรียบเทียบ ได้แก่ Firebase, Anto และ NETPIE

ตารางที่ 4 กลุ่มที่ 4 REST API ภาษา Python และ REST API ภาษา Python บน RaspberryPi

แพลตฟอร์ม	Sending time (ms)	
	Avg.	S.D.
Firebase	6560.44	1191.77
Anto	1884.57	516.07
NETPIE	5422.88	1075.26

4.2 ผลทดลองวัด Throughput

จากการทดลองได้ผลในแต่ละกลุ่มดังต่อไปนี้

กลุ่มที่ 1 NodeMCU และ NodeMCU

Platform IoT ที่จะนำมาเปรียบเทียบ ได้แก่ Anto และ NETPIE ได้ผลการทดลอง ดังตารางที่ 5

ตารางที่ 5 Throughput กลุ่ม NodeMCU และ NodeMCU

แพลตฟอร์ม	ความถี่ในการส่งข้อความ					
	3000 ms		1000 ms		500 ms	
	ส่ง	รับ	ส่ง	รับ	ส่ง	รับ
Anto	600	450	1800	1647	3600	3287
NETPIE	600	530	1800	1459	3600	2094

กลุ่มที่ 2 Node.js และ NodeMCU(ESP8266)

Platform IoT ที่จะนำมาเปรียบเทียบ ได้แก่ Azure IoT HUB, Bluemix Watson Internet of Things และ NETPIE ได้ผลการทดลอง ดังตารางที่ 6

ตารางที่ 6 Throughput กลุ่ม Node.js และ NodeMCU

แพลตฟอร์ม	ความถี่ในการส่งข้อความ					
	3000 ms		1000 ms		500 ms	
	ส่ง	รับ	ส่ง	รับ	ส่ง	รับ
Azure IoT Hub	600	600	1800	1800	3600	3303
Bluemix Watson IoT	600	600	1800	1800	3600	3600
NETPIE	600	600	1800	1800	3600	3600

กลุ่มที่ 3 Python และ Python บน RaspberryPi

Platform IoT ที่จะนำมาเปรียบเทียบ ได้แก่ Amazon AWS IoT, Bluemix Watson IoT และ NETPIE ได้ผลการทดลอง ดังตารางที่ 7

ตารางที่ 7 Throughput กลุ่ม Python และ Python บน Raspberry Pi

แพลตฟอร์ม	ความถี่ในการส่งข้อความ					
	3000 ms		1000 ms		500 ms	
	ส่ง	รับ	ส่ง	รับ	ส่ง	รับ
AWS IoT	600	600	1800	1800	3600	3600
Bluemix	600	600	1800	1800	3600	3600
NETPIE	600	599	1800	1800	3600	3600

กลุ่มที่ 4 REST API ภาษา Python และ REST API ภาษา Python บน RaspberryPi

IoT Platform ที่จะนำมาเปรียบเทียบได้แก่ Firebase Anto และ NETPIE ได้ผลการทดลอง ดังตารางที่ 8

ตารางที่ 8 Throughput กลุ่ม REST API ภาษา Python และ REST API ภาษา Python บน RaspberryPi

แพลตฟอร์ม	ความถี่ในการส่งข้อความ					
	3000 ms		1000 ms		500 ms	
	ส่ง	รับ	ส่ง	รับ	ส่ง	รับ
Firebase	600	427	1800	774	3600	824
Anto	600	556	1800	1419	3600	2413
NETPIE	600	484	1800	1048	3600	1034

4.3 ความซับซ้อนในการเขียนโปรแกรมใช้งานแพลตฟอร์ม

จากการทดลองผู้วิจัยสรุปจำนวนบรรทัดของโค้ดที่ต้องเขียนเพิ่มเติมจากตัวอย่างโค้ดที่แต่ละแพลตฟอร์มให้ไว้ในเว็บไซต์ โดยนับเป็นบรรทัดได้ดังนี้

ตารางที่ 9 ตารางแสดงจำนวนบรรทัดของโค้ดของกลุ่มการทดลองที่ 1 ได้แก่ Firebase Anto และ NETPIE

	Sender	Receiver
Firebase	22	10
Anto	11	6
NETPIE	16	6

จากตารางที่ 9 สรุปได้ว่า Anto เขียนโค้ดเพิ่มน้อยที่สุด รองลงมาคือ NETPIE และสุดท้าย คือ Firebase

ตารางที่ 10 ตารางแสดงจำนวนบรรทัดของโค้ดของกลุ่มการทดลองที่ 2 ได้แก่ Azure, Bluemix และ NETPIE

	Sender	Receiver
Azure	14	5
Bluemix	20	12
NETPIE	13	6

จากตารางที่ 10 สรุปได้ว่า Azure กับ NETPIE เขียนโค้ดเพิ่มไม่ต่างกันมากนัก ส่วน Bluemix เขียนโค้ดเพิ่มมากกว่า

ตารางที่ 11 ตารางแสดงจำนวนบรรทัดของโค้ดของกลุ่มการทดลองที่ 3 ได้แก่ AWS IoT, Bluemix, Firebase และ NETPIE

	Sender	Receiver
AWS IoT	15	14
Bluemix	16	14
NETPIE	18	13
Firebase	22	22

จากตารางที่ 11 สรุปได้ว่า AWS IoT Bluemix และ NETPIE เขียนโค้ดเพิ่มแทบไม่ต่างกัน ส่วน Firebase เขียนโค้ดเพิ่มมากที่สุด

5. สรุปผลการทดลอง

จากผลการทดลองพบว่า Anto มีความเร็วในการรับส่งข้อความมากที่สุด NETPIE และ Azure มีความเร็วรองลงมา Bluemix, AWS IoT และ Firebase ช้าลงมาตามลำดับ ความช้าเร็ว และความเหวี่ยงของเวลาขึ้นอยู่กับความเร็ว และความเสถียรของอินเตอร์เน็ตขณะนั้นด้วย อาจจะส่งผลให้การวัดความเร็วในแต่ละแพลตฟอร์มไม่เที่ยงตรงมาก

นัก แต่ก็สามารถอนุมานจนได้ข้อสรุปว่า ในการนำไปใช้งานในด้านการเก็บข้อมูลเพื่อวิเคราะห์ Firebase, Bluemix และ Azure มีความเหมาะสมมาก Bluemix และ Azure นั้นอาจจะมีข้อจำกัดอยู่บ้าง เพราะจะแบ่งการทำงานเป็น 2 ส่วน คือ Device กับ Application หรือ Service ทำให้ต้องเลือกใช้อย่างจำกัดว่าจะเลือกทำส่วนไหน Bluemix จะค่อนข้างช้าเพราะ Server ตั้งอยู่อเมริกา ทำให้ค่อนข้างช้า เมื่อเทียบกับแพลตฟอร์มอื่นๆแล้ว ส่วนใหญ่ตั้งอยู่ที่สิงคโปร์ ในส่วนของ Azure มีความยากตรงการนำตัวอย่างมาใช้ทำให้การศึกษาในช่วงเริ่มต้นนั้นช้ามาก แต่ในส่วนการส่งข้อความไปกลับนั้นเร็วเทียบเท่ากับ NETPIE และยังสามารถเก็บข้อมูลได้อีกด้วย ส่วน Firebase แม้จะมีความเร็วช้า แต่ก็มีความแม่นยำ และการจัดการการเก็บข้อมูลที่ดี เมื่อไฟดับหรืออินเทอร์เน็ตขาดการเชื่อมต่อ Firebase สามารถจัดการปัญหา และมีการอัพเดทค่าต่อได้ ซึ่งถือว่าเป็นข้อดีอย่างมาก แต่เนื่องจาก Firebase ไม่มีไลบรารีที่รองรับ Python หรือ C++ ออกมาอย่างเป็นทางการ และใน README.md ใน Github ก็ระบุไว้ชัดเจนว่า “The Arduino library is under heavy development, experimental, unversioned and its API is not stable.” ณ . ขณะนี้จึงยากที่จะนำ Firebase มาใช้ในการพัฒนา IoT แอปพลิเคชัน

ในเรื่องความปลอดภัย AWS IoT มีประสิทธิภาพมากกว่าตัวอื่นๆ เนื่องจากต้องมีการยืนยัน certificate ทุกครั้งที่อุปกรณ์ต้องการเชื่อมต่อไปยัง AWS IoT แต่เพราะเหตุผลนี้เองที่ทำให้ AWS IoT ไม่เหมาะกับอุปกรณ์ที่มีความจำน้อย และไม่เหมาะกับผู้เริ่มต้นศึกษา IoT เพราะขั้นตอนการสร้างและนำไปใช้ของ certificate และ Policy ค่อนข้างซับซ้อนพอสมควร

Anto มีความเร็วในหลักสิบมิลลิวินาที ถือว่ามีความเร็วมาก แต่มีข้อจำกัดคือสามารถทำงานได้กับ ESP8266 เท่านั้น ถ้าต้องการนำไปประยุกต์ใช้งานแบบอื่น อาจจะต้องใช้ REST API ทำให้การทำงานอาจจะมีความสามารถไม่เทียบเท่าไลบรารีภาษานั้นๆโดยตรง และการวัดความเร็วในการส่ง-รับข้อความจาก NodeMCU ที่ได้ทำการทดลองไปนั้น ต้องมีเว็บเป็นตัวกลางเพราะไม่สามารถทำให้ NodeMCU ส่ง-รับข้อความกันได้ตรงๆ จึงจำเป็นต้องสร้าง Toggle switch ไว้รองรับการกดปิด-เปิดไฟ โดยเขียนให้ Sender ส่งข้อความไปเปลี่ยน Toggle switch อีกอัน ผ่าน Receiver สลับการทำงานไปมา เหมาะแก่การพัฒนา IoT ในระดับเริ่มต้นมากกว่า เพราะจำนวนโค้ดที่ต้องเขียนเพิ่มมีจำนวนน้อยที่สุด ใช้งานง่าย

NETPIE มีความเร็วในระดับดี ถ้าเทียบกับแพลตฟอร์มอื่นๆ ในเรื่องการจัดการ Database นั้น NETPIE และ Anto เทียบ Firebase, Bluemix, AWS IoT และ Azure ไม่ได้ แต่ในด้านความเร็ว NETPIE มีความเร็วเป็นอันดับสองรองจาก Anto เมื่อเทียบกับแพลตฟอร์มอื่นๆ ในเรื่อง SDK/Language ถือได้ว่า NETPIE ได้เปรียบ เพราะมีการรองรับภาษาที่หลากหลาย มีทั้งไลบรารีและ REST API ในการเรียนรู้ และเข้าใจในการเขียนโค้ดเพิ่มก็อยู่ในระดับปานกลางถึงค่อนข้างง่าย เหมาะต่อการนำไปใช้เรียนรู้ และพัฒนาเทคโนโลยีที่เกี่ยวกับ IoT

จึงสรุปได้ว่าหากจะนำ IoT แพลตฟอร์มไปใช้งานจริงในการพัฒนานั้น ควรเลือกจากคุณสมบัติให้เหมาะสมตามการใช้งานเป็นหลัก

6. กิตติกรรมประกาศ

ขอขอบคุณ คุณชาวีร์ อิสริยภัทร์ ที่ให้คำแนะนำและความช่วยเหลือตลอดระยะเวลาการทำวิจัย และขอขอบคุณศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ ห้องปฏิบัติการวิจัยนวัตกรรมอินเทอร์เน็ต(INO) และภาควิชาวิทยาการคอมพิวเตอร์มหาวิทยาลัยเกษตรศาสตร์ ที่สนับสนุนการศึกษาค้นคว้าและทำวิจัยครั้งนี้

เอกสารอ้างอิง

[1] KKDEVELOPER. BASIC ARDUINO ตอนที่ 1: ARDUINO คืออะไร. [สืบค้นเมื่อ 31 ตุลาคม 2559] http://www.sat2you.com/site/?p=2702/

[2] Pinmanas As. ภาษาซีพลัสพลัส. [ออนไลน์] [สืบค้นเมื่อ 29 พฤศจิกายน 2559] http://pinmanass.blogspot.com/2013/09/blog-post.html

[3] Suratin Pattanawongthai. การติดตั้ง Node.js และ NPM. [ออนไลน์] [สืบค้นเมื่อ 29 พฤศจิกายน 2559] http://www.helloho.me/how-to-install-npm-and-nodejs/

[4] ThaiEasyElec. บทความการพัฒนาโปรแกรมบน Raspberry Pi ด้วย Qt. [ออนไลน์] [สืบค้นเมื่อ 31 ตุลาคม 2559] http://www.thaieasyelec.com/article-wiki/embedded-electronics-application/บทความการพัฒนาโปรแกรมบน-raspberry-pi-ด้วย-qt.html/

กรอบการประเมินสมรรถนะด้านเทคโนโลยีสารสนเทศระหว่างการพัฒนาซอฟต์แวร์แบบอไจล์ในองค์กรภาครัฐ

Information Technology Competency Assessment Framework during Agile Software Development in Public Sector

เทวา คำปาเชื้อ (Tawa Khampachua)[1] และสมคิด แซ่หลี (Somkid Saelee)[2]

ภาควิชาคอมพิวเตอร์ศึกษา คณะครุศาสตร์อุตสาหกรรม มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

[1]tawa.k@fte.kmutnb.ac.th, [2]somkid.s@fte.kmutnb.ac.th

บทคัดย่อ

หลักการพัฒนาระบบเทคโนโลยีสารสนเทศแบบอไจล์ที่สำคัญคือการเน้นการมีส่วนร่วมของผู้ใช้งานเพื่อให้ระบบที่พัฒนาขึ้นมานั้นสามารถนำไปใช้งานได้จริงและมีฟังก์ชั่นที่ตอบสนองความต้องการใช้งานขององค์กรมากที่สุด โดยของการพัฒนาแบบอไจล์จะแบ่งวงรอบการพัฒนาที่เรียกว่า Sprint ออกเป็นหลายวงรอบตามจำนวนโมดูลหลักและโมดูลย่อยของระบบสารสนเทศ เพื่อเปิดโอกาสให้ผู้พัฒนาสามารถส่งมอบและทดสอบโมดูลเหล่านั้นกับผู้ใช้งานได้บ่อยขึ้น วิธีการพัฒนาซอฟต์แวร์เช่นนี้จะเอื้อต่อการประเมินสมรรถนะของผู้ใช้งานระหว่างการพัฒนาแทนการประเมินสมรรถนะก่อนหรือหลังการพัฒนา ซึ่งผลที่ได้รับจากการประเมินนี้จะถูกนำมาใช้เพื่อออกแบบหลักสูตรการอบรม ออกแบบคู่มือการใช้งาน และยังสามารถนำผลป้อนกลับที่ได้ในแต่ละวงรอบการพัฒนามาใช้เพื่อปรับปรุงระบบสารสนเทศให้ใช้งานง่ายขึ้นและสามารถสนับสนุนการปฏิบัติงานได้อย่างแท้จริง

เนื่องจากผู้วิจัยได้รับมอบหมายให้ดำเนินโครงการพัฒนาระบบสารสนเทศเพื่อการบริหารจัดการสินทรัพย์ของกรมชลประทาน จึงได้ทดลองใช้กรอบแนวทางนี้กับกลุ่มตัวอย่างซึ่งเป็นเจ้าหน้าที่พัสดุจากสำนักงานชลประทานทั่วประเทศจำนวนทั้งสิ้น 613 คน โดยวงรอบการพัฒนามีจำนวนทั้งสิ้น 19 รอบ ผลการทดสอบพบว่ากรอบแนวทางที่นำเสนอสามารถช่วยให้ผู้พัฒนาวางแผนการปรับปรุงและอบรมการใช้งานระบบได้อย่างมีประสิทธิภาพ โดยผู้ใช้งานระบบสามารถขึ้นทะเบียนครุภัณฑ์ได้ภายใน 20 นาที ซึ่งใช้เวลาน้อยกว่าการทดสอบใช้ระบบในวงรอบแรกถึง 3 เท่า

คำสำคัญ : *กรอบแนวทางการประเมินสมรรถนะ วงจรชีวิตการพัฒนาซอฟต์แวร์แบบอไจล์*

Abstract

The key principle of Agile software development method is the user's participation which is viable by means of continuous delivery through the multiple sprints. The number of sprints vary with the number of modules/sub-modules in the system. With this method, developers will get to deliver and test the software module with real users more frequently. User's IT competency assessment could passively be conducted by exploiting such mechanism throughout the development process as oppose to before or after the software development. The IT literacy results evaluated from each sprint were used as feedback to either improve the software training material as well as curriculum or improve the software quality to better serve the current work flow.

Being given an opportunity to develop an Asset Management System (AMS) for the Royal Irrigation Department (RID), we had applied our framework with a sample group of 613 people from procurement departments coming from every RID office around the country. Based on the results obtained from the 19 sprints of applying our practice, users were able to register the asset within 20 minutes which is 3 times less than the time taken on the first sprint.

Keyword: *Competency Assessment Framework, Agile Software Development Life Cycle,*

1. บทนำ

ปัจจุบันระบบเทคโนโลยีสารสนเทศและการสื่อสาร หรือ “ไอซีที” ได้ถูกนำมาเป็นเครื่องมือประยุกต์ใช้ในการบริหารจัดการองค์กรและการให้บริการลูกค้าหรือประชาชนมากขึ้น แต่ถึงกระนั้น องค์กรต่างๆ เหล่านั้น โดยเฉพาะในกลุ่มผู้บริหารมักประสบปัญหา หรือมีปัญหาหนักใจเกี่ยวกับการพิจารณาความคุ้มค่าด้านการลงทุนด้านไอซีที ว่าควรลงทุนด้านไอซีทีแค่ไหนเพียงใดจึงจะเหมาะสม และยิ่งกว่านั้นจะมั่นใจได้อย่างไรว่างบประมาณที่ลงทุนไปนั้น จะมีประโยชน์และมีความคุ้มค่ากับทุนและทรัพยากรที่ลงไป โดยเฉพาะการลงทุนด้านไอซีทีในหน่วยงานภาครัฐหรือหน่วยงานภาคบริการประชาชนซึ่งเป้าหมายการวัดผลความคุ้มค่ามิได้วัดเฉพาะผลกำไรขององค์กร แต่ประกอบด้วยปัจจัยหลายด้าน ซึ่งด้านหนึ่งที่สำคัญคือการถ่ายทอดความรู้และประโยชน์ในการใช้งานระบบให้กับผู้ใช้งาน [1]

ในอดีตโครงการพัฒนาระบบสารสนเทศส่วนใหญ่เมื่อดำเนินมาถึงขั้นตอนการส่งมอบงานก็มักจะมีการจัดอบรมและจัดทำคู่มือให้กับผู้ใช้งาน จากการติดตามและสังเกตพฤติกรรมของบุคลากรหน่วยงานภาครัฐ พบว่าวิธีการอบรมมักจะได้ผลกับกลุ่มคนที่มีประสบการณ์ในการใช้งานคอมพิวเตอร์มากกว่ากลุ่มคนที่ขาดทักษะในการใช้งานคอมพิวเตอร์ และผู้ที่มีอายุมาก คนเหล่านี้มักสามารถจดจำเนื้อหาการอบรมได้ในขณะที่อยู่ในห้องเรียนเท่านั้น แต่หลังจากการอบรมเสร็จสิ้นแล้วก็มักจะลืมเนื้อหาไป และหากระบบสารสนเทศมีขั้นตอนทำงานที่ซับซ้อน จะส่งผลให้เกิดมุมมองหรือทัศนคติที่ไม่ดีต่อระบบสารสนเทศดังกล่าว โดยระบบอาจถูกมองว่าเป็นภาระ และอ้างเป็นเหตุผลให้ไม่ใช้งานระบบในที่สุด

ดังนั้น งานวิจัยนี้มีวัตถุประสงค์เพื่อนำเสนอกรอบการประเมินสมรรถนะด้านเทคโนโลยีสารของบุคลากรระหว่างการพัฒนาระบบแบบอไจล์ โดยกรอบการประเมินสมรรถนะนี้ถูกนำไปทดลองใช้ในโครงการพัฒนาระบบสารสนเทศเพื่อการบริหารจัดการสินทรัพย์ของกรมชลประทานกับกลุ่มตัวอย่างซึ่งเป็นเจ้าหน้าที่พัสดุจากสำนักงานชลประทานทั่วประเทศจำนวนทั้งสิ้น 613 คน โดยวงรอบการพัฒนาจะใช้เวลาประมาณรอบละ 2-3 สัปดาห์ มีจำนวนทั้งสิ้น 19 รอบ โดยวัดสมรรถนะจากระยะเวลาในการใช้งาน 7 โมดูลหลักของระบบในทุกวงรอบการส่งงาน

2. สมมติฐานของการวิจัย

กรอบแนวทางที่นำเสนอสามารถนำมาใช้เป็นเครื่องมือเพื่อวางแผนการพัฒนาบุคลากร และช่วยลดระยะเวลาในการใช้งานระบบสารสนเทศในการปฏิบัติงานได้อย่างน้อยร้อยละ 50

3. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

3.1 การประเมินสมรรถนะด้านเทคโนโลยีสารสนเทศ

สมรรถนะด้านเทคโนโลยีสารสนเทศนั้นสามารถจำแนกได้หลายด้าน ขึ้นอยู่กับมุมมองของผู้ใช้งาน อาทิ Calvani และคณะ [2] ได้จำแนกสมรรถนะด้านดิจิทัลออกเป็น 3 ด้าน ประกอบด้วย

- ด้านเทคโนโลยี คือความสามารถในการใช้และเรียนรู้เทคโนโลยีใหม่
- ด้านการรับรู้ คือความสามารถในการเข้าถึงและประมวลผลสารสนเทศ
- ด้านจริยะธรรม คือความสามารถในการใช้ ICT ได้ตามหลักการที่เหมาะสม

Martin และ Grudziecki [3] แบ่งระดับของสมรรถนะด้านดิจิทัลออกเป็น 3 ระดับ ประกอบด้วย Digital Competence ซึ่งถือว่าเป็นระดับแรกที่ต้องเข้าใจเทคโนโลยีในหลักการ ระดับที่ 2 คือ Digital Usage หรือความสามารถในการใช้งานเทคโนโลยี ส่วนระดับที่ 3 คือ Digital Transformation หรือระดับที่สามารถสร้างนวัตกรรมได้ ส่วน Mason et. al. [4] และ Leahy et. al. [5] มีมุมมองที่ต่างออกไป กล่าวคือ Mason แบ่งสมรรถนะด้านเทคโนโลยีสารสนเทศออกเป็น 2 ด้านคือ ความสามารถในการใช้งาน และความตระหนักในการใช้งาน แต่ Leahy แบ่งสมรรถนะออกเป็น 2 ด้านเช่นกัน แต่แบ่งตามวัตถุประสงค์ในการใช้งาน คือ เพื่อการดำรงชีวิตอยู่ในสังคม และเพื่อการเรียนและการทำงาน

ปัจจุบันเทคโนโลยีเข้ามามีบทบาทในการเรียนการสอนและแม้แต่การทำงานมากขึ้น ดังนั้น การเตรียมความพร้อมด้านเทคโนโลยีสารสนเทศของนักเรียน นักศึกษา หรือบุคลากรในองค์กร จึงเป็นสิ่งที่สถาบันศึกษาและองค์กรต่างๆทั้งภาครัฐและเอกชนให้ความสนใจ ดังจะเห็นได้จากการจัดสอบ

มาตรฐานหรือใบรับรองต่าง ๆ ที่เกี่ยวข้องกับเทคโนโลยีสารสนเทศมีมากขึ้นเรื่อย ๆ อาทิ หลักสูตร Digital Literacy ของ Microsoft, IC3 Digital Literacy Certificate, ICDL (International Computer Driving License) เป็นต้น โดยสมรรถนะที่ใช้ประเมินส่วนใหญ่คือ การใช้งานคอมพิวเตอร์พื้นฐาน การจัดการไฟล์บนเครื่องคอมพิวเตอร์ การใช้งานเว็บบราวเซอร์ ความเสี่ยงและความมั่นคงปลอดภัยคอมพิวเตอร์ การใช้งานแอพลิเคชั่นเกี่ยวกับภาพ เสียง และวิดีโอ และการใช้งานแอพลิเคชั่นพื้นฐานเช่น Word Processing, Spreadsheets, Presentation เป็นต้น [5]

3.2 การพัฒนาซอฟต์แวร์แบบอไจล์

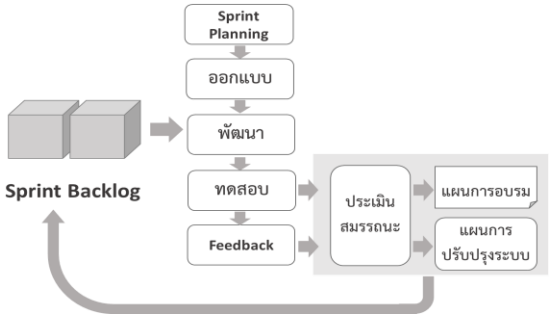
การพัฒนาซอฟต์แวร์แบบอไจล์นั้นเป็นการพัฒนาแบบวนรอบที่มุ่งเน้นไปที่การส่งมอบงานทีละนิดเพื่อเปิดโอกาสให้ผู้ใช้งานระบบได้มีโอกาสให้คำแนะนำในการพัฒนาเป็นระยะ ซึ่งแตกต่างจากวิธีการพัฒนาแบบดั้งเดิมหลายด้าน [6,7] ดังแสดงในตารางที่ 1

ตารางที่ 1 ความแตกต่างระหว่างแนวทางการพัฒนาแบบดั้งเดิมและการพัฒนาแบบอไจล์

มิติการพัฒนา	แบบดั้งเดิม	แบบอไจล์
การควบคุม	เน้นกระบวนการ	เน้นตัวบุคคล
สไตล์การจัดการ	ควบคุมโดยการสั่งการ	มีผู้นำแต่เน้นการทำงานร่วมกัน
การกำหนดบทบาท	กำหนดตามความถนัด	ทำงานเป็นทีม สามารถสับเปลี่ยนบทบาทได้
การสื่อสาร	สื่อสารเมื่อจำเป็นอย่างเป็นทางการ	สื่อสารแบบไม่เป็นทางการตลอดเวลา
การมีส่วนร่วมของลูกค้า	มีส่วนร่วมเฉพาะช่วงที่ต้องการการวิเคราะห์เพื่อตัดสินใจ	ตลอดเวลาเนื่องจากเป็นสิ่งสำคัญมาก
เทคโนโลยี	ไม่มีข้อจำกัด	นิยมใช้เทคโนโลยีแบบ object-oriented
สถานที่ทำงานของทีม	ส่วนใหญ่ทำงานคนละแห่ง	ส่วนใหญ่ทำงานที่เดียวกัน
ขนาดของทีม	ส่วนใหญ่มากกว่า 10 คน	ส่วนใหญ่น้อยกว่า 10 คน
การมีส่วนร่วมของคนในทีม	ไม่จำเป็น	จำเป็น
การวางแผน	ทำล่วงหน้า	วางแผนตลอดเวลา
Feedback	ไม่ค่อยมีกลไกที่เอื้อต่อการรับ feedback	มีหลายกลไกที่เอื้อต่อการรับ feedback
เอกสาร	เน้นการส่งเอกสารทุกขั้นตอน	ไม่เน้นการส่งเอกสาร

ความท้าทายในการพัฒนาระบบแบบอไจล์ในองค์กรภาครัฐที่หลายงานวิจัยกล่าวถึงนั้น คือ ปัญหาการมีส่วนร่วมของบุคลากรภาครัฐที่เป็นผู้ใช้งานระบบ กระบวนการเอกสารของทางราชการที่เมื่อเป็นไปตามหลักการอไจล์ และกระบวนการจัดซื้อจัดจ้างที่ไม่รองรับการกำหนดขอบเขตงานที่ยืดหยุ่นแบบอไจล์ [8-11] อย่างไรก็ตามการพัฒนาซอฟต์แวร์แบบอไจล์นั้นมิได้มีรายละเอียดเกี่ยวกับการผลักดันให้เกิดการใช้งานระบบแต่อย่างใดเนื่องจากหลักการพัฒนาระบบแบบอไจล์เน้นการพัฒนาเพื่อให้ระบบใช้งานได้ตามความต้องการของลูกค้าเป็นหลัก แต่วิธีการปรับแต่งระบบและปฏิสัมพันธ์กับลูกค้าเพื่อให้เกิดการยอมรับนั้นยังไม่ได้กล่าวถึงในรายละเอียด ดังนั้นงานวิจัยนี้จะได้กล่าวถึงวิธีการประเมินสมรรถนะของลูกค้า (ผู้ใช้งานระบบ) ระหว่างการพัฒนาแบบอไจล์เพื่อให้สามารถนำผลการประเมินมาใช้เป็นกรอบแนวทางปรับปรุงระบบและแนวทางการจัดอบรมเพื่อให้ลูกค้าสามารถใช้งานระบบได้อย่างแท้จริง

4. กรอบการประเมินสมรรถนะระหว่างการพัฒนาซอฟต์แวร์แบบอไจล์



ภาพที่ 1 กรอบการประเมินสมรรถนะระหว่างการพัฒนาซอฟต์แวร์แบบอไจล์

ภาพที่ 1 แสดงวงรอบการพัฒนาระบบ 1 วงรอบ ซึ่งประกอบด้วยการนำ Sprint Backlog มาวางแผนการพัฒนา โดย

เริ่มตั้งแต่การออกแบบ พัฒนา ทดสอบ และส่งมอบให้ลูกค้าหรือผู้ใช้งานระบบได้ทดลองใช้และให้ feedback กลับคืน โดยระหว่างการทดสอบในแต่ละวงรอบนั้นจะมีการประเมินสมรรถนะของผู้ใช้งานระบบ ซึ่งวิธีการประเมินจะประกอบด้วย 4 ขั้นตอน คือ

1) สังเคราะห์กิจกรรมที่จำเป็นในการใช้งานระบบที่ส่งมอบในวงรอบนั้น ๆ เช่น การกรอกข้อมูล การแนบไฟล์ การดาวน์โหลดไฟล์ การอัพโหลดไฟล์ การค้นหา การใช้งานเมนู การแก้ไขข้อความ การแก้ไขรูปภาพ เป็นต้น
2) ประเมินสมรรถนะในรูปแบบของระยะเวลาที่ใช้ในการทำกิจกรรมต่าง ๆ ที่สังเคราะห์ได้ในขั้นตอนแรก หากคะแนนประเมินต่ำกว่าเกณฑ์ที่กำหนด ทีมจะต้องพิจารณาหาสาเหตุของปัญหาว่าเป็นเพราะผู้ใช้งานขาดสมรรถนะที่จำเป็นในการใช้งานระบบ หรือเป็นเพราะการออกแบบ UX ที่ไม่เอื้อต่อการใช้งาน
3) หากเป็นเพราะสมรรถนะของผู้ใช้งานระบบ ทีมจะต้องออกแบบวิธีการนำเสนอและอบรมผู้ใช้งานใหม่
4) หากเป็นเพราะ UX ทีมจะต้องร่วมหารือกับผู้ใช้งานเพื่อเร่งปรับปรุงการออกแบบระบบเพื่อให้สามารถใช้งานได้ง่ายขึ้น ตามหลักการ Lean UX [12-13]

ขั้นตอนการปรับปรุง UX นั้น หากเป็นการพัฒนาให้กับหน่วยงานภาครัฐซึ่งมีการจัดซื้อจัดจ้างที่กำหนดระยะเวลาและขอบเขตแบบตายตัวแล้ว จะต้องมีการบันทึกเงื่อนไขการเปลี่ยนแปลงทุกวงรอบการพัฒนา เพื่อนำข้อมูลดังกล่าวมาอ้างอิงหากไม่สามารถดำเนินการได้แล้วเสร็จตามระยะเวลาที่กำหนด

5. กรณีศึกษาระบบบริหารจัดการสินทรัพย์

5.1 ระบบบริหารจัดการสินทรัพย์ กรมชลประทาน

ผู้วิจัยได้มีโอกาสพัฒนาระบบบริหารจัดการสินทรัพย์ให้กับกรมชลประทาน โดยระบบดังกล่าวนี้เมื่อพัฒนาแล้วเสร็จจะถูกนำมาใช้ทดแทนกระบวนการจัดการครุภัณฑ์ตลอดช่วงอายุ คือ ตั้งแต่การขึ้นทะเบียน การโอน การยืม การจ่ายยืม การจัดทำครุภัณฑ์ประจำปี การบันทึกประวัติการซ่อมครุภัณฑ์ ตลอดจนถึงการแทงจำหน่ายครุภัณฑ์ ซึ่งวิธีการบริหารจัดการครุภัณฑ์ของกรมฯ เดิมใช้วิธีการสำเนาเอกสารสำคัญ ลงนามกำกับ และส่งเอกสารทางไปรษณีย์ ซึ่งการขึ้นทะเบียนในแต่ละครั้งใช้ระยะเวลานานถึง 2-6 เดือน การจัดเก็บข้อมูลเป็นเอกสารหรือไฟล์ excel จึงยากต่อการตรวจเช็ค ดังนั้น ความท้าทายของการพัฒนาระบบบริหารจัดการสินทรัพย์ คือการบริหารการเปลี่ยนแปลงกระบวนการระหว่างการพัฒนาด้วย

5.2 สมรรถนะพื้นฐานสำหรับแต่ละโมดูล

ระบบบริหารจัดการสินทรัพย์ แบ่งออกเป็น 7 โมดูลหลัก โดยผู้ใช้งานในแต่ละโมดูลจำเป็นต้องมีสมรรถนะพื้นฐาน ดังแสดงในตารางที่ 2

ตารางที่ 2 สมรรถนะพื้นฐานที่จำเป็นในการใช้งานระบบบริหารจัดการสินทรัพย์

โมดูล	สมรรถนะพื้นฐาน
การขึ้นทะเบียนครุภัณฑ์	การแนบไฟล์ การ Navigate ระหว่างขั้นตอน การสืบค้นผู้ขาย การบันทึกรายละเอียดการจัดซื้อ* การแก้ไขข้อมูล*
การยืมครุภัณฑ์ การโอนครุภัณฑ์	การเข้าถึงข้อมูลเมื่อมีการแจ้งเตือน การสืบค้นครุภัณฑ์ การอนุมัติผ่านระบบ* การเรียกดูสถานะของครุภัณฑ์
การจ่ายยืม บันทึกประวัติการซ่อม	การสืบค้นครุภัณฑ์ การบันทึกประวัติการซ่อม* การบันทึกประวัติการจ่ายยืม
การจำหน่ายครุภัณฑ์	การเปลี่ยนสถานะครุภัณฑ์ การเรียกดูรายงานการจำหน่าย
การจัดทำรายงานประจำปี	การเรียกดูรายงานโดยใช้ตัวกรอง

* สมรรถนะที่ต้องอาศัยความเข้าใจกระบวนการทางปฏิบัติงาน

สมรรถนะพื้นฐานดังแสดงในตารางที่ 2 สามารถนำมาจำแนกได้เป็น 3 กลุ่มสมรรถนะ คือ

1) สมรรถนะด้านการใช้งานคอมพิวเตอร์เบื้องต้น เช่น การ Navigate ระหว่างหน้า การใช้เมนูแบบ drop-down การแนบไฟล์ เป็นต้น

2) สมรรถนะด้านการเข้าถึงและประมวลผลสารสนเทศ เช่น การสืบค้น การใช้ตัวกรองเพื่อเรียกดูรายงาน เป็นต้น

3) สมรรถนะการใช้คอมพิวเตอร์เพื่อการปฏิบัติงาน คือ สมรรถนะที่ต้องอาศัยความเข้าใจในขั้นตอนการปฏิบัติงานประกอบการความสามารถในการใช้งานคอมพิวเตอร์เบื้องต้น เช่นการบันทึกข้อมูลต่าง ๆ ลงในระบบ จำเป็นต้องเข้าใจความหมายของข้อมูลจึงจะสามารถนำเข้าข้อมูลได้ถูกต้อง เป็นต้น

5.3 กลุ่มประชากรและวงรอบการทดสอบระบบ

กลุ่มผู้ใช้งานระบบบริหารจัดการสินทรัพย์ครุภัณฑ์ที่จะมาเป็นกลุ่มตัวอย่างในการทดลองนี้ คือเป็นเจ้าหน้าที่พัสดุของกรมชลประทานจากทั่วประเทศจำนวน 613 คน จากสำนักงานและโครงการชลประทานทั่วประเทศ โดยผู้วิจัยแบ่งวงรอบการพัฒนาออกเป็น 19 วงรอบ วงรอบละ 2-3 อาทิตย์

6. ผลการวิจัย

ผู้วิจัยได้ทดลองประเมินประสิทธิภาพของกรอบแนวทางการประเมินสมรรถนะด้านเทคโนโลยีสารสนเทศระหว่างการพัฒนาซอฟต์แวร์แบบอไจล์กับกลุ่มตัวอย่างซึ่งเป็นเจ้าหน้าที่พัสดุจากสำนักงานชลประทานทั่วประเทศ จำนวนทั้งสิ้น 613 คน ร้อยละ 82 เป็นผู้หญิง และร้อยละ 27 มีอายุอยู่ในช่วง 46-60 ปี ซึ่งเป็นกลุ่มที่มีสมรรถนะทางด้านเทคโนโลยีสารสนเทศในระดับที่ต่ำมาก กล่าวคือ ขาดสมรรถนะด้านการใช้งานระบบคอมพิวเตอร์เบื้องต้น แต่กลับเป็นกลุ่มที่มีความรู้ความเข้าใจในกระบวนการปฏิบัติงานในระดับดีถึงดีมาก เนื่องจากมีประสบการณ์ทำงานมาก

เมื่อพิจารณาด้านอายุราชการ ร้อยละ 38 มีประสบการณ์การทำงานไม่ถึง 6 ปี ดังนั้น จึงไม่มีความชำนาญในเรื่องการบริหารจัดการสินทรัพย์ เนื่องจากมีประสบการณ์กับการจัดการสินทรัพย์บางประเภทเท่านั้น ดังนั้นผู้ใช้งานในกลุ่มนี้จะขาดสมรรถนะด้านการใช้คอมพิวเตอร์เพื่อการปฏิบัติงาน

จำนวนวงรอบในการทดลองเพื่อประเมินประสิทธิภาพกรอบแนวทางเท่ากับวงรอบการพัฒนาคือ 19 วงรอบ ระยะเวลาของแต่ละรอบประมาณ 2-3 สัปดาห์ โดยแต่ละวงรอบมีจำนวนโมดูลที่ส่งมอบไม่เท่ากัน กล่าวคือ 2 วงรอบแรกนั้นเป็นการส่งมอบโมดูลการขึ้นทะเบียนครุภัณฑ์ วงรอบที่ 3 มีโมดูลการยืมและโอนเพิ่มเติม ส่วนวงรอบที่ 4 นั้นมีโมดูลการจ่ายยืมและการบันทึกประวัติการซ่อม ส่วนวงรอบที่ 5-19 นั้น มีทุกโมดูล โดยกลุ่มผู้ใช้งานในแต่ละวงรอบนั้นต่างกัน ผลการประเมินประสิทธิภาพใช้วิธีการวัดระยะเวลาที่ผู้ใช้งานทำงานในแต่ละกระบวนการ โดยมีผลลัพธ์ ดังแสดงในตารางที่ 3

ตารางที่ 3 เปรียบเทียบเวลาของการทำงานรอบแรกและรอบสุดท้าย

โมดูลของระบบ	เวลารอบแรก (นาที)	เวลารอบสุดท้าย (นาที)
การขึ้นทะเบียนครุภัณฑ์	50-70	15-20
การยืมครุภัณฑ์	30-40	10-15
การโอนครุภัณฑ์	30-40	10-15
การจ่ายยืม	10-20	5-10
บันทึกประวัติการซ่อม	5-10	2-3
การจำหน่ายครุภัณฑ์	10-20	5-10
การจัดทำรายงานประจำปี	10-20	5-10

ผลจากการทดสอบในวงรอบแรกของการอบรมเพื่อทดสอบระบบนั้น พบว่าสมรรถนะพื้นฐานของแต่ละกลุ่มผู้ใช้งานที่ทดสอบระบบนั้นค่อนข้างต่างกัน ดังนั้น แม้ว่าจะมีการปรับระบบให้สอดคล้องกับวิธีการปฏิบัติงานมากเท่าใด แต่หากผู้ใช้งานขาดสมรรถนะด้านการใช้คอมพิวเตอร์พื้นฐานก็จะไม่สามารถใช้งานระบบได้เลย ดังนั้น ในวงรอบถัด ๆ มา ผู้วิจัยจึงปรับแนวทางการอบรม โดยนอกจากจะเน้นการอธิบายวิธีการใช้งานคอมพิวเตอร์พื้นฐานแล้วยังต้องจัดเจ้าหน้าที่เพิ่มเติมเพื่อคอยช่วยเหลือกลุ่มผู้ใช้งานเหล่านี้เพื่อให้สามารถใช้งานระบบได้เมื่อสิ้นสุดการอบรมเพื่อการทดสอบระบบ อย่างไรก็ตาม ผู้วิจัยพบว่ามีบุคลากรบางส่วนในจำนวนนี้ (ประมาณร้อยละ 3%) ที่ปิดกั้นการรับรู้และมีทัศคติที่ไม่เปิดรับเทคโนโลยีใหม่ โดยมักเป็นผู้ใช้งานที่มีอายุใกล้เกษียณ ดังนั้นในกรณีนี้จึงไม่สามารถปรับหลักสูตรเพื่อให้เอื้อต่อการปรับทัศนคติของบุคลากรกลุ่มนี้ได้

7. สรุปและอภิปรายผลการวิจัย

การพัฒนากรอบแนวทางการประเมินสมรรถนะด้านเทคโนโลยีสารสนเทศระหว่างการพัฒนาซอฟต์แวร์แบบอไจล์ที่นำเสนอนี้ สามารถนำมาใช้เพื่อเป็นเครื่องมือสำหรับสนับสนุนการวางแผนพัฒนาบุคลากรด้านเทคโนโลยีสารสนเทศขององค์กร เนื่องจากในแต่ละวงรอบการพัฒนาจะมีการประเมินสมรรถนะที่จำเป็นสำหรับการใช้งานโมดูลนั้นๆ ของผู้ใช้งาน ซึ่งผลที่ได้จากการประเมินนี้จะถูกนำมาใช้เพื่อออกแบบหลักสูตรการอบรม ออกแบบคู่มือการใช้งานระบบสารสนเทศ และนำผลป้อนกลับมาใช้เพื่อปรับปรุงให้ระบบสามารถใช้งานได้ง่ายขึ้นและสนับสนุนการปฏิบัติงานได้อย่างแท้จริง

จากการทดลองใช้กรอบแนวคิดนี้กับกลุ่มตัวอย่าง ผลการทดสอบโมดูลการขึ้นทะเบียนครุภัณฑ์ซึ่งเป็นโมดูลที่ยากและใช้ระยะเวลานานที่สุด พบว่าระยะเวลาเฉลี่ยที่ใช้ในการขึ้นทะเบียนครุภัณฑ์ของวงรอบการพัฒนาแรกอยู่ที่ประมาณ 50 – 70 นาที และลดลงเหลือเพียง 15-20 นาทีในวงรอบการพัฒนาสุดท้าย โดยผู้วิจัยได้มีการปรับหลักสูตรการอบรมให้ผู้ใช้งานระบบสามารถเรียนรู้การใช้งานและลดระยะเวลาการปฏิบัติงานได้มากกว่าร้อยละ 50 ดังสมมติฐานที่ตั้งไว้

เอกสารอ้างอิง

[1] Khampachua, Tawa, and Nawaporn Wisitpongphan. "Implementing Successful IT Projects in Thailand Public Sectors: A Case Study." In *Proceedings of The 15th European Conference on eGovernment ECEG 2015 University of Portsmouth*, p. 403. 2015.

[2] Calvani, A., et.al., Assessing Digital Competence in Secondary Education – Issues, Models, and Instruments. In *Issues in Information and Media Literacy: Education, practice*, and Pedagogy, 2009, pp. 153 – 172

[3] Martin, A., Grudziecki, J., “DigEuLit: Concepts and Tools for Digital Literacy Development”, University of Glasgow, Scotland (2006)

[4] Mason, John, and Robert Mc Morrow. "YACLD: yet another computer literacy definition." *Journal of Computing Sciences in Colleges* 21.5 (2006): 94-100.

[5] Leahy, Denise, and Dudley Dolan. "Digital literacy: A vital competence for 2010?." *In Key competencies in the knowledge society*, pp. 210-221. Springer Berlin Heidelberg, 2010.

[6] Nerur, Sridhar, RadhaKanta Mahapatra, and George Mangalaraj. "Challenges of migrating to agile methodologies." *Communications of the ACM* 48, no. 5 (2005): 72-78.

[7] Schuh, Peter. *Integrating agile development in the real world*. Charles River Media, Inc., 2004.

[8] Wisitpongphan, Nawaporn, and Tawa Khampachua. "Agile in public sector: Case study of dairy farm management projects." *In 13th IEEE International Joint Conference on Computer Science and Software Engineering (JCSSE)*, 2016, pp. 1-5

[9] U.S. Government Accountability Office, ”Software Development: Effective Practices and Federal Challenges in Applying Agile Methods,” GAO- 12-681, July 27, 2012

[10] A. Kaczorowska, “Traditional and Agile Project Management in Public Sector and ICT,” *Proc. of the Federated Conference on Computer Science and Information Systems*, vol. 5, 2015, pp. 1521–1531

[11] J. Little and A. Karaj, “Transforming a Public Sector Organization: From Stone Age to Agile,” *Agile Conference*, 2013, pp. 74–81

[12] Gosper, Julian, Jean-Luc Agathos, Richard Rutter, and Terry Coatta. "UX design and agile: a natural fit?." *Communications of the ACM* 54, no. 1 (2011): 54-60.

[13] Gothelf, Jeff, and Josh Seiden. Lean UX: Applying lean principles to improve user experience. " O'Reilly Media, Inc.", 2013.

ระบบค้นหาโรงพยาบาลบริเวณใกล้เคียงในกรุงเทพมหานคร บนระบบปฏิบัติการแอนดรอยด์
Find Nearby Hospitals in Bangkok System on The Android Operating System

จารุนันท์ บุญวิทยาลิขิต (Jarunan Boonwittayalikit)[1] และมณเฑียร รัตนศิริวงศ์วุฒิ (Montean Rattanasiriwongwut)[2]
ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
[1]s5807021857032@email.kmutnb.ac.th, [2]montean.r@it.kmutnb.ac.th

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาระบบค้นหาโรงพยาบาลบริเวณใกล้เคียงในกรุงเทพมหานครบนระบบปฏิบัติการแอนดรอยด์ โดยประยุกต์ใช้กูเกิลแมพและระบบจีพีเอส ในการนำทางไปโรงพยาบาล ซึ่งค้นหาโรงพยาบาลที่อยู่ใกล้ที่สุด อีกทั้งยังมีข้อมูลการปฐมพยาบาลเบื้องต้น และเกร็ดความรู้เพื่อแนะนำและให้ความรู้เกี่ยวกับสุขภาพ โดยพัฒนาบนระบบปฏิบัติการแอนดรอยด์เพื่อความสะดวกและง่ายต่อการใช้งาน

การประเมินคุณภาพของผู้เชี่ยวชาญที่มีต่อระบบที่พัฒนาขึ้นได้ค่าเฉลี่ยเท่ากับ 4.13 และค่าเบี่ยงเบนมาตรฐานเท่ากับ 0.45 อยู่ในระดับดี และผลการความพึงพอใจของผู้ใช้งานระบบที่พัฒนาได้ค่าเฉลี่ยเท่ากับ 3.96 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.58 อยู่ในระดับดี

คำสำคัญ: กูเกิลแมพ จีพีเอส แอนดรอยด์

Abstract

The purpose of this research was to develop a Find Nearby Hospital in Bangkok System by applying google map and GPS system to navigate to the nearby hospitals. Which searches the nearest hospital. The first aid information and tips to introduce and educate about health. Developed on the Android operating system to facilitate and easy to use.

The evaluation from experts discovered an average quality in the satisfaction level was at 4.1 3 with a standard deviation of 0.45, the results showed that system has a good quality level the evaluation from users towards discovered an average satisfaction level was at 3.96 with a standard deviation of 0.58 the results showed that system has a good quality level.

Keywords: Google Map, GPS, Android

1. บทนำ

ในแต่ละวันอุบัติเหตุเกิดขึ้นได้ตลอดเวลา โดยเกิดได้จากการเดินทาง การทำงาน การดำเนินชีวิตประจำวัน ทำให้เกิดการบาดเจ็บ สูญเสียทั้งชีวิตและทรัพย์สิน หากเกิดอุบัติเหตุขึ้นแล้วและมีการบาดเจ็บ จำเป็นจะต้องนำผู้บาดเจ็บส่งโรงพยาบาลเพื่อทำการรักษาให้ทันท่วงที ในการเดินทางเพื่อไปโรงพยาบาลหากไม่ชำนาญเส้นทางก็จะทำให้เสียเวลา ซึ่งอาจส่งผลให้เกิดความสูญเสียตามมา และด้วยในปัจจุบันผู้คนหันมาดูแลให้ความสำคัญกับการดูแลสุขภาพ รวมทั้งตระหนักในการป้องกันโรคภัยไข้เจ็บต่างๆ มากขึ้น เมื่อเกิดอาการหรือเจ็บป่วยก็จะพบแพทย์เพื่อรักษาแต่เนิ่นๆ และยังมีการใช้บริการทางการแพทย์อื่นที่ไม่ใช่การรักษาโดยตรง เช่น การเสริมความงาม เป็นต้น ซึ่งในการค้นหาโรงพยาบาลนั้น หากไม่ใช่คนในพื้นที่ อาจไม่ทราบว่าโรงพยาบาลหรือสถานพยาบาลอยู่ที่ไหนบ้าง และเนื่องจาก Google Maps มีการแสดงข้อมูลบนแผนที่อย่างกว้างขวาง ไม่เจาะจง ทำให้ต้องใช้เวลาในการค้นหาโรงพยาบาลที่ต้องการ

เทคโนโลยี มีความสำคัญกับชีวิตประจำวันของผู้คนเป็นอย่างมาก ทั้งทางตรงและทางอ้อมซึ่งอุปกรณ์สื่อสารที่เรียกว่าโทรศัพท์มือถือ เป็นสิ่งที่ขาดไม่ได้ในการดำเนินชีวิตประจำวัน

ของทุกคนและโทรศัพท์มือถือในปัจจุบันก็มีการพัฒนากันอย่างต่อเนื่อง มีการพัฒนาแอปพลิเคชันมากมาย ทำให้โทรศัพท์มือถือในปัจจุบัน (Smart Phone) มีประสิทธิภาพเพิ่มมากขึ้น

ดังนั้น ผู้วิจัยจึงมีแนวคิดในการพัฒนาระบบค้นหาโรงพยาบาลบริเวณใกล้เคียงในกรุงเทพมหานครบนระบบปฏิบัติการแอนดรอยด์ เพื่อช่วยในการค้นหาโรงพยาบาลได้อย่างรวดเร็ว โดยแสดงโรงพยาบาลที่อยู่ใกล้ที่สุด สามารถนำทางไปโรงพยาบาลได้ทันที หรือสามารถดูข้อมูลโรงพยาบาล และโทรออกเพื่อแจ้งเหตุจากแอพพลิเคชันได้ทันที อีกทั้งยังสามารถเลือกได้ว่าต้องการค้นหาโรงพยาบาลรัฐบาล โรงพยาบาลเอกชน หรือศูนย์บริการสาธารณสุข เป็นการเพิ่มทางเลือกในการเข้ารับการรักษา และยังสามารถดูข้อมูลการปฐมพยาบาลเบื้องต้นและเกร็ดความรู้ได้อีกด้วย ซึ่งระบบที่พัฒนาขึ้นนี้เป็นการพัฒนาในรูปแบบของแอพพลิเคชันบนระบบปฏิบัติการแอนดรอยด์ที่แสดงผลข้อมูลโรงพยาบาลบริเวณใกล้เคียงในกรุงเทพมหานคร

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 ระบบสารสนเทศทางภูมิศาสตร์ (Geographic Information System : GIS)

ระบบสารสนเทศภูมิศาสตร์ หรือ Geographic Information System : GIS คือกระบวนการทำงานเกี่ยวกับข้อมูลในเชิงพื้นที่ด้วยระบบคอมพิวเตอร์ ที่ใช้กำหนดข้อมูลและสารสนเทศ ที่มีความสัมพันธ์กับตำแหน่งในเชิงพื้นที่ เช่น ที่อยู่ บ้านเลขที่ สัมพันธ์กับตำแหน่งในแผนที่ ตำแหน่ง เส้นรุ้ง เส้นแวง ข้อมูลและแผนที่ใน GIS เป็นระบบข้อมูลสารสนเทศที่อยู่ในรูปของตารางข้อมูล และฐานข้อมูลที่มีส่วนสัมพันธ์กับข้อมูลเชิงพื้นที่ (Spatial Data) ซึ่งรูปแบบและความสัมพันธ์ของข้อมูลเชิงพื้นที่ทั้งหลาย จะสามารถนำมาวิเคราะห์ด้วย GIS และทำให้สื่อความหมายในเรื่องการเปลี่ยนแปลงที่สัมพันธ์กับเวลาได้ เช่น การแพร่ขยายของโรคระบาด การเคลื่อนย้ายถิ่นฐาน การบุกรุกทำลาย การเปลี่ยนแปลงของการใช้พื้นที่ ฯลฯ ข้อมูลเหล่านี้ เมื่อปรากฏบนแผนที่ทำให้สามารถแปลและสื่อความหมาย ใช้งานได้ [1]

2.2 ระบบปฏิบัติการแอนดรอยด์ (Android)

แอนดรอยด์ (Android) เป็นซอฟต์แวร์ที่มีโครงสร้างแบบเรียงทับซ้อนหรือแบบสแต็ก (Stack) ซึ่งรวมเอาระบบปฏิบัติการ (Operating System) มิดเดิลแวร์ (Middleware) และแอพพลิเคชันที่สำคัญเข้าไว้ด้วยกัน เพื่อใช้สำหรับทำงานบนอุปกรณ์พกพาเคลื่อนที่ (Mobile Devices) เช่น โทรศัพท์มือถือ เป็นต้น การทำงานของแอนดรอยด์ มีพื้นฐานอยู่บนระบบลีนุกซ์ เคอร์เนล (Linux Kernel) ซึ่งใช้ Android SDK (Software Development Kit) เป็นเครื่องมือสำหรับการพัฒนาแอพพลิเคชันบนระบบปฏิบัติการแอนดรอยด์ และใช้ภาษา Java ในการพัฒนา [2]

2.3 ระบบหาพิกัดบนพื้นโลก (GPS)

จีพีเอส (GPS) เป็นระบบหาพิกัดบนพื้นโลกโดยการอ้างอิงจากดาวเทียมที่มีความแม่นยำสูงสามารถใช้หาพิกัดใดๆ บนพื้นโลกได้ทุกเวลา ทุกสภาพอากาศ ซึ่งในปัจจุบันเทคโนโลยีที่เริ่มเข้ามามีบทบาทในชีวิตประจำวันมากขึ้น มีการนำไปใช้ในกิจกรรมต่างๆ มากมาย เช่น การช่วยวางแผนเดินทางด้วยแผนที่ระบบติดตามตำแหน่งบนพื้นผิวโลก เป็นต้น [3]

2.4 เอพีไอ (Application Programming Interface)

API ย่อมาจาก Application Programming Interface คือ ช่องทางการเชื่อมต่อระหว่างเว็บไซต์หนึ่งไปยังอีกเว็บไซต์หนึ่ง หรือเป็นการเชื่อมต่อระหว่างผู้ใช้งานกับ Server หรือจาก Server เชื่อมต่อไปหา Server ซึ่ง API นี้เปรียบได้เป็นภาษาคอมพิวเตอร์ที่ทำให้คอมพิวเตอร์สามารถสื่อสารและแลกเปลี่ยนข้อมูลกันได้อย่างอิสระ เช่น Google Maps API คือ บริการของ Google อีกรูปแบบหนึ่งที่เราสามารถนำข้อมูลของ Google Maps ที่ทาง Google โดย API ถือเป็นกลุ่มของฟังก์ชัน ขั้นตอน หรือคลาส (Class) ที่ระบบปฏิบัติการ (OS) หรือผู้ให้บริการสร้างขึ้นมาเพื่อรองรับการเรียกขอข้อมูลจากโปรแกรมอื่นๆ ทั้งนี้ API สามารถใช้งานได้กับภาษาในการเขียนโปรแกรมที่รองรับเท่านั้น ซึ่งมันจะถูกจัดทำให้อยู่ในรูปแบบ Syntax หรือ element ที่สามารถนำไปใช้ได้อย่างสะดวกสบาย [4]

2.5 งานวิจัยที่เกี่ยวข้อง

ปัจจุบันมีงานวิจัยที่ใช้ Google Maps พิกัดจีพีเอส และระบบปฏิบัติการแอนดรอยด์ มาประยุกต์ใช้ในงานวิจัยอย่างหลากหลาย ซึ่งงานวิจัยที่ได้ศึกษา ดังนี้

ชไมพร (2557) ทำการการวิจัยเรื่อง ระบบแนะนำสถานที่ท่องเที่ยวอุทยานแห่งชาติทางภาคใต้ของประเทศไทยด้วย

Google Maps API บนมือถือสมาร์ทโฟน เป็นการพัฒนาระบบแนะนำ สถานที่ท่องเที่ยวอุทยานแห่งชาติทางภาคใต้ของประเทศไทยโดยใช้ Google Maps และใช้เทคนิค Collaborative Filtering ร่วมกับ K-Mean มาช่วยในการแนะนำข้อมูลจากชิ้นข้อมูลเรตติ้ง (Rating) ในระบบเทียบกับเรตติ้งของผู้ใช้งานและทำนายสถานที่ที่คาดว่าตรงกับความต้องการของผู้ใช้โดยใช้อัลกอรึทึม K-Mean ในการหาค่าที่มีความใกล้เคียงของข้อมูล และประเมินความพึงพอใจในการใช้ระบบโดยการหาค่าเฉลี่ย (Mean) และค่าส่วนเบี่ยงเบนมาตรฐาน (S.D) [5]

มยุรี (2557) ทำการการวิจัยเรื่อง ระบบติดตามสถานการณ์สถานศึกษาที่ประสบอุทกภัย โดยใช้สมาร์ทโฟนแอนดรอยด์ ซึ่งระบบนี้สามารถรองรับการรายงานข้อมูลอุทกภัยแบบเรียลไทม์ โดยสถาปัตยกรรมของการพัฒนาระบบมี 2 ส่วน ส่วนที่ 1 สมาร์ทโฟนแอนดรอยด์ ใช้ในการรายงานพิกัดจีพีเอส (Global Positioning System: GPS) ของสถานศึกษา ซึ่งสามารถแสดงผลผ่าน Google Map API ส่วนที่ 2 คือ Web-Based Application สามารถรองรับการนำเข้าข้อมูล [6]

สุพาพร (2555) ทำการการวิจัยเรื่อง ระบบวิเคราะห์การใช้น้ำมันเชื้อเพลิงของรถจากการเดินทาง โดย Google Map API โดยระบบสามารถช่วยในการวิเคราะห์การเดินทางและคำนวณระยะทาง โดยใช้เทคนิคการวิเคราะห์ในการจำแนกข้อมูล วิธีการถดถอยเชิงเส้นพหุคูณเพื่อประมาณค่าน้ำมันเชื้อเพลิงในการเดินทางเบื้องต้นให้กับผู้ใช้ ซึ่งจะช่วยในการตัดสินใจในการเดินทางให้กับผู้ใช้ได้ในระดับหนึ่ง ข้อมูลจะแสดงในรูปแบบของเว็บแอปพลิเคชั่น [7]

วิษณุ (2555) ทำการการวิจัยเรื่อง ระบบสารสนเทศภูมิศาสตร์เพื่อสืบค้นข้อมูลโรงงานโดยใช้ Google Map ได้นำทฤษฎีของฟัซซี่เซตเข้ามาช่วยในการค้นหาข้อมูล ทำให้ระบบมีความสามารถในการค้นหาข้อมูลโรงงานได้ดียิ่งขึ้น และแสดงผลในรูปแบบของแผนที่ โดยใช้ Google Map [8]

กศวายุธ (2555) ทำการการวิจัยเรื่อง ระบบสารสนเทศภูมิศาสตร์อาคารชลประทานโดยใช้ Google Maps นำเสนอระบบสารสนเทศภูมิศาสตร์อาคารชลประทาน ซึ่งเป็นการ Mashup ข้อมูลจาก 3 แหล่ง คือ 1) ข้อมูลสารสนเทศของอาคารชลประทาน 2) ข้อมูลพิกัดตำแหน่งอาคารชลประทาน ข้อมูลเส้นแบ่งขอบเขตตำบล ขอบเขตอำเภอ ขอบเขตจังหวัด ถนน ทางน้ำ และ3) ข้อมูล Google Maps เข้าด้วยกันและแสดงผลลัพธ์ออกมาในรูปแบบเชิงพื้นที่โดยแสดงตำแหน่งพิกัดทางภูมิศาสตร์ของอาคารชลประทานบนแผนที่แบบโต้ตอบทันทีของ Google Maps [9]

3. วิธีดำเนินการวิจัย

ระบบค้นหาโรงพยาบาลบริเวณใกล้เคียงในกรุงเทพมหานครบนระบบปฏิบัติการแอนดรอยด์ ผู้วิจัยได้แบ่งการทำงานออกเป็นขั้นตอนซึ่งประกอบไปด้วย 5 ขั้นตอนดังนี้

3.1 การศึกษาและรวบรวมข้อมูลของระบบ

ในการศึกษาข้อมูลและเอกสารที่เกี่ยวข้องเพื่อระบบค้นหาโรงพยาบาลบริเวณใกล้เคียงในกรุงเทพมหานครบนระบบปฏิบัติการแอนดรอยด์ ผู้วิจัยได้แบ่งหัวข้อในการศึกษาข้อมูลดังนี้

3.1.1 ศึกษาเครื่องมือที่ใช้ในการพัฒนาระบบ โดยวิธีการต่าง ๆ ที่จะนำไปใช้ในการพัฒนาระบบนั้น โปรแกรมสำหรับพัฒนา Mobile Application โดยภาษาที่ใช้ในการเขียนโปรแกรมคือ ภาษา Java PHP และ MySQL

3.1.2 ศึกษาเกี่ยวกับโรงพยาบาลในเขตกรุงเทพมหานคร โดยศึกษาข้อมูล และนำข้อมูลเกี่ยวกับโรงพยาบาล ได้แก่โรงพยาบาลรัฐบาล โรงพยาบาลเอกชน และศูนย์บริการสาธารณสุข มาประกอบการวิจัย และได้ศึกษาจากแหล่งข้อมูลอื่นๆ เช่น เอกสาร หนังสือต่าง ๆ ข้อมูลทางอินเตอร์เน็ต เพื่อให้ได้ข้อมูลอย่างครบถ้วน

3.1.3 ศึกษาการพัฒนาโปรแกรมประยุกต์บนโทรศัพท์มือถือระบบปฏิบัติการแอนดรอยด์ รวมถึงเครื่องมือที่ใช้ในการพัฒนาระบบ ภาษาที่ใช้พัฒนา วิธีการ เทคนิค ขั้นตอนต่างๆ ในการพัฒนาระบบ

3.2 การวิเคราะห์ระบบ

จากการศึกษาข้อมูลจากแหล่งข้อมูลต่าง ๆ และการวิเคราะห์ข้อมูล ทำให้ทราบขั้นตอนและกระบวนการทำงานการพัฒนาระบบค้นหาโรงพยาบาลบริเวณใกล้เคียงในกรุงเทพมหานคร บนระบบปฏิบัติการแอนดรอยด์ โดยประยุกต์ใช้ระบบสารสนเทศทางภูมิศาสตร์ และ Google Maps ซึ่งพัฒนาในรูปแบบแอปพลิเคชัน (โปรแกรมประยุกต์

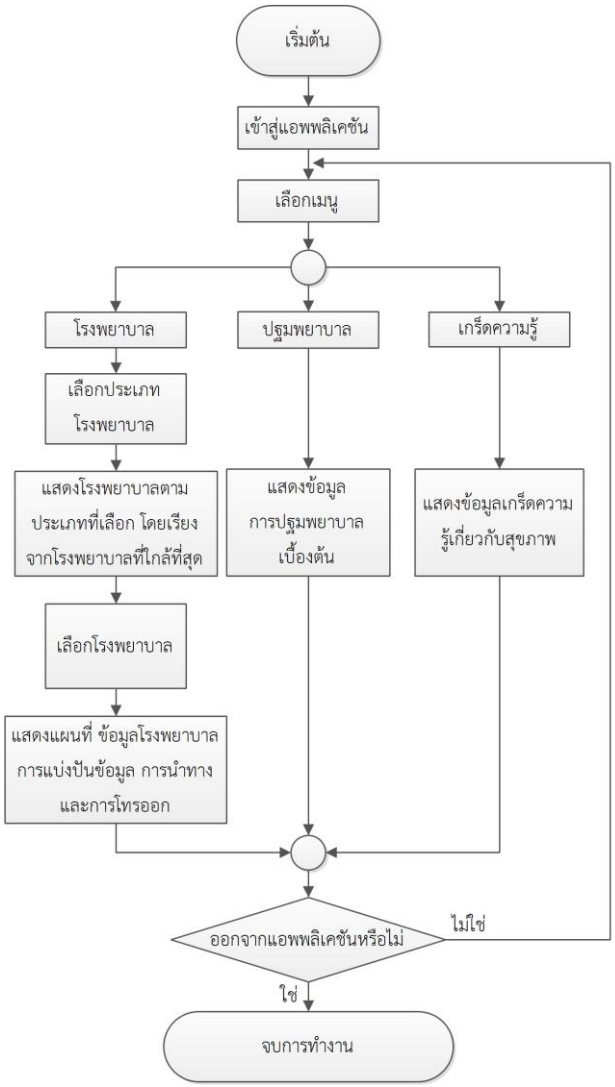
บนระบบปฏิบัติการแอนดรอยด์) การทำงานแบ่งเป็น 3 ส่วน ได้แก่

3.2.1 ส่วนโรงพยาบาลของการแสดงผลโรงพยาบาล เป็นส่วนที่ค้นหาโรงพยาบาลตามประเภทที่กำหนด แสดงแผนที่ข้อมูลโรงพยาบาล เช่น การให้บริการ ที่อยู่ เบอร์โทรศัพท์ เป็นต้น การแบ่งปันข้อมูล การนำทางจะใช้เทคนิคของระบบสารสนเทศทางภูมิศาสตร์ พิกัดจีพีเอส ในการค้นหาตำแหน่งร่วมกับ Google Maps API และการโทรออกจากแอพพลิเคชัน

3.2.2 ส่วนปฐมพยาบาล เป็นส่วนที่แสดงข้อมูลการปฐมพยาบาลเบื้องต้น

3.2.3 ส่วนเกร็ดความรู้ เป็นส่วนที่แสดงข้อมูลเกร็ดความรู้เกี่ยวกับสุขภาพ

โดยภาพรวมของระบบแสดงไว้ในภาพที่ 1 ดังนี้



ภาพที่ 1: ภาพรวมของระบบ

3.3 การพัฒนาระบบ

การพัฒนาระบบค้นหาโรงพยาบาลบริเวณใกล้เคียงในกรุงเทพมหานคร บนระบบปฏิบัติการแอนดรอยด์ จากการศึกษา วิเคราะห์ และออกแบบระบบ ทำให้ทราบขั้นตอนการดำเนินงาน ซึ่งผู้วิจัยได้พัฒนาระบบในรูปแบบ Mobile Application เพื่อความสะดวก รวดเร็ว และใช้งานง่าย โดยใช้ภาษา PHP และ Java พัฒนาระบบ การพัฒนาระบบแบ่งเป็น 2 ส่วน ดังนี้

3.3.1 การออกแบบส่วนติดต่อกับผู้ใช้งาน (User Interface) เป็นการออกแบบหน้าจอ เพื่อติดต่อกับผู้ใช้งาน เพื่อความสะดวก รวดเร็ว และใช้งานง่าย

3.3.2 การพัฒนาระบบในรูปแบบ Mobile Application โดยใช้โปรแกรมและเครื่องมือ ดังนี้

3.3.2.1 Androids Studio, Androids SDK

3.3.2.2 ระบบจัดการฐานข้อมูลที่ใช้ คือ My SQL

3.3.2.3 Google Maps API

3.4 การทดสอบระบบ

การทดสอบระบบ เป็นการทดสอบและตรวจสอบข้อผิดพลาดที่เกิดจากระบบ เพื่อให้สอดคล้องกับความต้องการ ระบบทำงานได้รวดเร็วและถูกต้อง โดยใช้กระบวนการ Black-Box Testing เป็นการทดสอบการทำงานในการหาข้อบกพร่องและข้อผิดพลาด ซึ่งเป็นการทดสอบระบบแบบไม่สนใจกระบวนการทำงานภายในของโปรแกรม แต่เน้นไปที่ข้อมูลเข้าและผลลัพธ์ที่ได้ว่าการทำงานต่างๆ ถูกต้องตามวัตถุประสงค์ที่กำหนดไว้หรือไม่ แบ่งออกเป็น 2 ขั้นตอน ดังนี้

3.4.1 การทดสอบระบบขั้นอัลฟ่า (Alpha Testing) เป็นการทดสอบการทำงานของระบบ โดยผู้วิจัยทำการทดสอบกระบวนการทำงานว่าถูกต้องตามวัตถุประสงค์ที่ได้กำหนดไว้หรือไม่ เมื่อทำการทดสอบแล้วพบปัญหาซึ่งอาจเกิดจากข้อผิดพลาดของกระบวนการทำงานที่ไม่ถูกต้องหรือไม่สามารถทำงานได้ตรงตามวัตถุประสงค์ที่กำหนดไว้จะทำการแก้ไขและปรับปรุงให้ดีขึ้น

3.4.2 การทดสอบระบบขั้นเบต้า (Beta Testing) เป็นการทดสอบประสิทธิภาพของระบบ และนำผลการประเมินจากการทดสอบมาสรุปผลเพื่อปรับปรุงแก้ไขให้ระบบมีประสิทธิภาพมากขึ้น

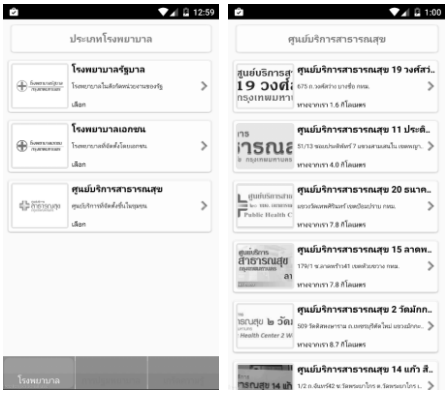
4. ผลการดำเนินงาน

ผลการดำเนินงานระบบค้นหาโรงพยาบาลบริเวณใกล้เคียงในกรุงเทพมหานคร บนระบบปฏิบัติการแอนดรอยด์ แบ่งเป็น 2 ส่วน ดังนี้

4.1 ผลการพัฒนาระบบ

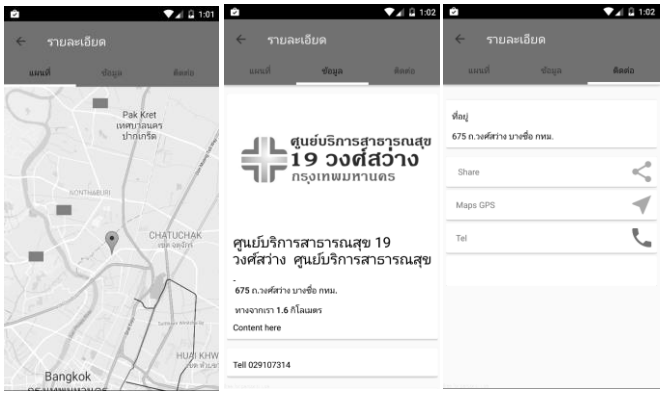
ในการพัฒนาระบบค้นหาโรงพยาบาลบริเวณใกล้เคียงในกรุงเทพมหานคร บนระบบปฏิบัติการแอนดรอยด์ แบ่งการทำงาน ดังนี้

4.1.1 ส่วนแสดงผลโรงพยาบาล เป็นส่วนแสดงประเภทโรงพยาบาล ในที่นี้กำหนดไว้ 3 ประเภท คือ โรงพยาบาลรัฐบาล โรงพยาบาลเอกชน และศูนย์บริการสาธารณสุข เมื่อเลือกประเภทโรงพยาบาลแล้ว จะแสดงรายการโรงพยาบาลโดยเรียงลำดับตามระยะทางที่ใกล้ที่สุด ดังภาพที่ 2



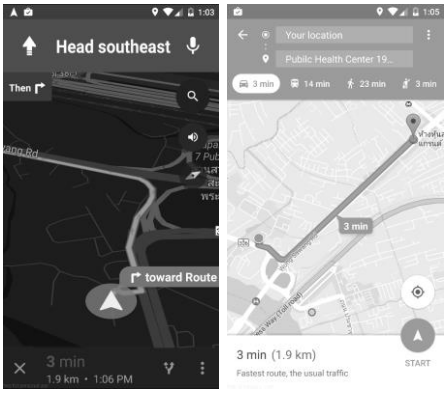
ภาพที่ 2: ส่วนแสดงผลโรงพยาบาล

4.1.2 ส่วนแสดงข้อมูลโรงพยาบาล เป็นส่วนที่แสดงข้อมูลของโรงพยาบาล ประกอบด้วย 3 ส่วน คือ แผนที่ ข้อมูล และติดต่อ ดังภาพที่ 3



ภาพที่ 3: ส่วนแสดงข้อมูลโรงพยาบาล

4.1.3 ส่วนการนำทางไปยังโรงพยาบาลที่เลือก เป็นการนำทาง จากจุดที่ผู้ใช้งานอยู่ ไปยังโรงพยาบาลจุดหมาย ดังภาพที่ 4



ภาพที่ 4: ส่วนการนำทางไปยังโรงพยาบาลที่เลือก

4.1.4 ส่วนแสดงผลการปฐมพยาบาล เป็นส่วนแสดงข้อมูลการปฐมพยาบาลเบื้องต้น เพื่อแนะนำแก่ผู้ใช้งาน ดังภาพที่ 5



ภาพที่ 5: ส่วนแสดงผลการปฐมพยาบาล

4.1.5 ส่วนแสดงผลเกร็ดความรู้ เป็นส่วนแสดงความรู้ในการดูแลสุขภาพแก่ผู้ใช้งาน ดังภาพที่ 6



ภาพที่ 6: ส่วนแสดงผลเกร็ดความรู้

4.2 ผลการประเมิน

4.2.1 เมื่อนำระบบที่ได้พัฒนาไปทดสอบเพื่อประเมินคุณภาพของระบบโดยผู้เชี่ยวชาญ สามารถสรุปผลการประเมินได้ ดังนี้

ตาราง 1: ผลการประเมินคุณภาพโดยผู้เชี่ยวชาญ

รายการประเมิน	ระดับคุณภาพ		
	$\bar{x}$	S.D.	ระดับ
ด้านความสามารถในการทำงานของระบบ	4.28	0.43	ดี
ด้านความถูกต้องของระบบ	4.28	0.53	ดี
ด้านการประเมินการใช้งานของระบบ	4.03	0.44	ดี
ด้านประสิทธิภาพโดยรวมของระบบ	3.95	0.41	ดี
สรุปผลการประเมิน	4.13	0.45	ดี

จากการประเมินคุณภาพของผู้เชี่ยวชาญที่มีต่อระบบที่พัฒนาขึ้นได้ค่าเฉลี่ยเท่ากับ 4.13 และค่าเบี่ยงเบนมาตรฐานเท่ากับ 0.45 อยู่ในระดับดี

4.2.2 เมื่อนำระบบที่ได้พัฒนาไปทดสอบเพื่อประเมินความพึงพอใจโดยผู้ใช้งานระบบ สามารถสรุปผลการประเมินได้ ดังนี้

ตาราง 2: ผลการประเมินความพึงพอใจโดยผู้ใช้งานระบบ

รายการประเมิน	ระดับความพึงพอใจ		
	$\bar{x}$	S.D.	ระดับ
ด้านรูปแบบการนำเสนอระบบ	4.04	0.57	ดี
ด้านการใช้งานระบบ	3.93	0.64	ดี
ด้านความรวดเร็วในการประมวลผล	3.80	0.55	ดี
ด้านประโยชน์ของระบบ	4.08	0.55	ดี
สรุปผลการประเมิน	3.96	0.58	ดี

จากการประเมินความพึงพอใจของผู้ใช้งานระบบที่พัฒนาได้ค่าเฉลี่ยเท่ากับ 3.96 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.58 อยู่ในระดับดี

5. สรุปผล

การพัฒนาระบบค้นหาโรงพยาบาลบริเวณใกล้เคียงในกรุงเทพมหานคร บนระบบปฏิบัติการแอนดรอยด์ ทำให้สามารถค้นหาโรงพยาบาลได้สะดวกขึ้น โดยแบ่งประเภทโรงพยาบาล เพื่อช่วยในการค้นหาโรงพยาบาลที่เฉพาะเจาะจงมากขึ้นตามความต้องการของผู้ใช้งาน ซึ่งในกรณีที่มีเหตุการณ์ฉุกเฉิน ผู้ใช้งานสามารถค้นหาโรงพยาบาล ข้อมูลการปฐมพยาบาลได้ทันที นอกจากนี้ผู้ใช้งานสามารถดูข้อมูลรายละเอียดของโรงพยาบาล แผนที่โรงพยาบาล แบ่งปันข้อมูลโรงพยาบาล ให้ระบบนำทางไปโรงพยาบาล และโทรศัพท์ออกจากแอพพลิเคชันได้โดยไม่ต้องเสียเวลาในการค้นหาเบอร์โทรศัพท์ นอกจากนี้ยังรวบรวมข้อมูลการปฐมพยาบาลเบื้องต้นในด้านต่างๆ รวมถึงการดูแลรักษาสุขภาพหากมีอาการด้านสุขภาพ และยังมีเกร็ดความรู้ เพื่อแนะนำผู้ใช้งานในการรักษาสุขภาพอีกด้วย

จากการพัฒนาระบบ ผู้วิจัยได้มีการทดลองใช้งานและทดสอบคุณภาพของระบบโดยใช้แบบประเมินคุณภาพจากกลุ่มตัวอย่าง 2 กลุ่ม คือ การประเมินคุณภาพโดยผู้เชี่ยวชาญจำนวน 5 คน และการประเมินความพึงพอใจโดยกลุ่มผู้ใช้งานทั่วไป จำนวน 30 คน ซึ่งผลการประเมินอยู่ในระดับดี

ข้อเสนอแนะและแนวทางในการพัฒนา ระบบค้นหาโรงพยาบาลบริเวณใกล้เคียงในอนาคต ควรมีข้อมูลในส่วนของโรงพยาบาล และคลินิกครอบคลุมทั้งประเทศ และมีการค้นหาเส้นทางอื่นเมื่อมีการจราจรหนาแน่น อีกทั้งควรพัฒนาระบบให้รองรับระบบปฏิบัติการที่หลากหลาย เพื่อรองรับการใช้งานของผู้ใช้งานให้สามารถใช้งานได้อย่างทั่วถึง

เอกสารอ้างอิง

[1] ระบบสารสนเทศภูมิศาสตร์ (Geographic Information System) GIS. [ออนไลน์]. [สืบค้นวันที่ 1 พฤศจิกายน 2559]. จาก http://www.gisthai.org/about-gis/gis.html

[2] กูเกิลแอนดรอยด์ (Google Android). [ออนไลน์]. [สืบค้นวันที่ 3 พฤศจิกายน 2559]. จาก http://know.9choo.in.th/2013/07/google-android.html

[3] ระบบหาพิกัดบนพื้นโลก GPS. [ออนไลน์]. [สืบค้นวันที่ 10 พฤศจิกายน 2559]. จาก http://nectec.or.th/rd/electronics/be205-45/be205-45.php

[4] API คืออะไร ทำหน้าที่อะไร ประโยชน์ของ API มีอะไรบ้าง. [ออนไลน์]. [สืบค้นวันที่ 12 พฤศจิกายน 2559]. จาก http://www.เกร็ดความรู้.net/api/

[5] ชไมพร ทองขาว และ มาลีรัตน์ โสดานิล, “ระบบแนะนำสถานที่ท่องเที่ยวอุทยานแห่งชาติทางภาคใต้ของประเทศไทยด้วยกูเกิลแมพเอพีไอ บนมือถือสมาร์ท” *The Eighth National Conference on Computing and Information Technology* พ.ศ.2557

[6] มยุรี จี้ระมาตย์ และ ศักดิ์ชาย ตั้งวรรณวิทย์, “ระบบติดตามสถานการณ์สถานศึกษาที่ประสบอุทกภัย โดยใช้สมาร์ทโฟนแอนดรอยด์” *The Eighth National Conference on Computing and Information Technology* พ.ศ.2557

[7] สุพาพร เรืองแจ่ม, “ระบบวิเคราะห์การใช้น้ำมันเชื้อเพลิงของรถจากการเดินทาง โดย Google Map API” *The Eighth National Conference on Computing and Information Technology* พ.ศ.2555

[8] วิษณุ อิ่มอารมณ์ และ สุพจน์นิตย์สุวัฒน์, “ระบบสารสนเทศภูมิศาสตร์เพื่อสืบค้นข้อมูลโรงงานโดยใช้ Google Map” *The Eighth National Conference on Computing and Information Technology* พ.ศ.2555

[9] กศวายุธ รักษ์กุล และ ศิริรัตน์ วณิชโยบล, “ระบบสารสนเทศภูมิศาสตร์อาคารชลประทานโดยใช้ Google Maps” *The Eighth National Conference on Computing and Information Technology* พ.ศ.2555

การพัฒนาต้นแบบโปรแกรมระบบแอนดรอยด์เพื่อให้การสนับสนุนการพัฒนาพฤติกรรมการดูแลสุขภาพในผู้ป่วยเบาหวาน
Android Application Prototype Development for Supporting Self-care Behavior in Diabetic Patient

ดร.รังสรรค์ สุวรรณหงส์ (Dr. Rungsan Suwannahong)[1] และผศ.ดร.สุรีรัตน์ อินทร์หม้อ (Asst. Prof. Dr. Sureerut Inmor)[2]

สาขาวิชาระบบสารสนเทศ คณะบริหารธุรกิจ มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี

[1]rungsan_s@rmutt.ac.th, [2]sureerut_i@rmutt.ac.th

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อ ศึกษาพฤติกรรมการส่งเสริมสุขภาพ และความต้องการข้อมูลด้านการส่งเสริมสุขภาพของผู้ป่วยเบาหวานในประเทศไทย ตลอดจนนำข้อมูลที่ได้ไปออกแบบโปรแกรมต้นแบบเพื่อสนับสนุนการใช้งานด้านการส่งเสริมสุขภาพของผู้ป่วยเบาหวาน โดยมีการตั้งสมมุติฐานว่า ลักษณะทางประชากรศาสตร์ที่แตกต่างกันของผู้ป่วยเบาหวาน (เพศ, อายุ, สถานภาพการสมรส, ระดับการศึกษา, อาชีพ, รายได้ และระยะเวลาที่เป็นโรคเบาหวาน) ส่งผลต่อพฤติกรรมการส่งเสริมสุขภาพที่แตกต่างกัน ประชากรที่ใช้ในการสำรวจมาจากผู้ป่วยเบาหวานในประเทศไทยที่ไม่ต้องใช้อินซูลิน โดยต้องเป็นผู้ป่วยที่มีการใช้งานโทรศัพท์สมาร์ทโฟน และเคยใช้งานแอพพลิเคชั่นในสมาร์ทโฟนอย่างน้อยหนึ่งประเภท

คำสำคัญ: ปัญหาสุขภาพ เบาหวาน สมาร์ทโฟน การดูแลสุขภาพ

Abstract

Research is aim to promote healthy behavior of people with diabetes in Thailand. The data obtained from the research will be used for design application prototype to support those tasks of health care. The hypothesis states that different demographic characteristics of patients (gender, age, marital status, education, occupation, income and duration of diabetes) affect different health promoting behaviors. The population survey of diabetes patients in Thailand that does not require insulin with the use of mobile smart phones and has experience at least one application on smart phone.

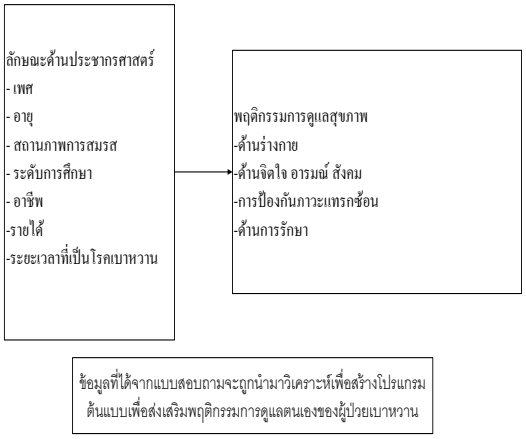
Keyword: Health Problem, Diabetic, Smart Phone, Health-care

1. บทนำ

การส่งเสริมสุขภาพของประชากรจัดเป็นมาตรการที่มีความสำคัญ ที่เกี่ยวข้องกับการป้องกันและรักษาสุขภาพ เพราะการที่ประชากรมีสุขภาพแข็งแรงย่อมหมายถึงงบประมาณในการดูแลผู้ป่วยที่ลดลง และยังเป็นการป้องกันภาวะแทรกซ้อนในผู้ป่วยได้อีก

โรคเบาหวานเป็นโรคที่มีการตรวจพบผู้ป่วยจำนวนมากในประเทศไทย และเป็นโรคที่ก่อให้เกิดภาวะแทรกซ้อนได้อีกหลายด้าน เช่น ทางตา ไต ระบบประสาท หลอดเลือด ความดันโลหิตสูง และอาการชาตามปลายมือและปลายเท้าเป็นต้น ดังนั้นการให้ความรู้ด้านการส่งเสริมสุขภาพแก่ผู้ป่วยเบาหวาน จึงเป็นสิ่งสำคัญต่อการลดภาวะแทรกซ้อนดังกล่าว

ในปัจจุบันโทรศัพท์มือถือได้ถูกนำมาใช้เป็นช่องทางในการสื่อสารข้อมูลทางการแพทย์และสุขภาพมากยิ่งขึ้น โดยได้มีการนำมาใช้งานด้านด้านการส่งเสริมการออกกำลังกาย การควบคุมอาหาร และการเฝ้าระวังต่างๆ และยังมีการใช้งานพื้นฐานอื่นๆ เช่น ตารางการนัดหมายของแพทย์ ตลอดจนนำเสนอข้อมูลด้านสุขภาพอันเป็นประโยชน์ งานวิจัยนี้จัดทำขึ้นโดยมีวัตถุประสงค์ที่จะใช้ประโยชน์จากแอพพลิเคชั่นแอนดรอยด์บนสมาร์ทโฟน โดยสำรวจความต้องการใช้งานแอพพลิเคชั่นของผู้ป่วยเบาหวาน ในด้านที่เกี่ยวข้องกับโปรแกรมส่งเสริมสุขภาพ แล้วนำมาจัดสร้างระบบต้นแบบเพื่อนำไปใช้ให้เป็นประโยชน์ต่อผู้ป่วยดังกล่าว



ภาพที่ 1: กรอบการวิจัย

2. วรรณกรรมที่เกี่ยวข้อง

2.1 การบริโภคอาหารสำหรับผู้ป่วยเบาหวาน

คำแนะนำในการบริโภคอาหารสำหรับผู้ป่วยเบาหวานมีดังนี้คือ

ปฏิบัติตามคำแนะนำทั่วไปของการบริโภคอาหารให้ครบ 5 หมู่ โดยมีสัดส่วนที่เหมาะสมตามคำแนะนำคือแป้ง (55-60%) โปรตีน (15-20%) และไขมัน (25%) โดยรับประทานในปริมาณที่สม่ำเสมอทุกมื้ออาหาร

ถ้าน้ำหนักผู้ป่วยมาก ควรคุมอาหารให้เหมาะสมกับการลดน้ำหนัก งดอาหารทอด รวมทั้งรับประทานอาหารตามเวลา ไม่ทานของขบเคี้ยวระหว่างมื้ออาหาร

เพื่อทำให้สุขภาพด้านการขับถ่ายดีขึ้นควรเพิ่มอาหารที่มีกากใยให้มากขึ้น

ถ้ามีอาการแทรกซ้อนด้านความดันโลหิตสูงหรือโรคไต ควรควบคุมอาหารที่มีรสเค็มเป็นกรณีพิเศษ

สำหรับอาหารที่ไม่ควรรับประทาน ได้แก่ น้ำตาลทุกชนิดรวมทั้งน้ำผึ้ง ขนมรวมถึงผลไม้ที่มีรสรสหวาน (โรคเบาหวานฯ, 2558)

2.2 การป้องกันโรคเบาหวาน

เนื่องจากโรคเบาหวานเป็นโรคที่ก่อให้เกิดโรคแทรกซ้อนต่างๆ มักจะมีอาการรุนแรงจนทำให้คุณภาพชีวิตและสุขภาพของผู้ป่วยแย่ลง ถึงแม้จะมียาที่ช่วยควบคุมอาการของผู้ป่วยเบาหวานได้ แต่ยาเหล่านี้ก็ไม่สามารถส่งผลถึงการลดอาการแทรกซ้อนได้ ดังนั้นการป้องกันตนเองให้ห่างไกลจากอาการของโรคเบาหวานจึงเป็นสิ่งที่ควรจะกระทำ คำแนะนำในการป้องกันมีดังนี้

การปรับเปลี่ยนพฤติกรรมใน 3 ด้านคือ การลดน้ำหนัก การออกกำลังกาย การรับประทานอาหารที่มีเส้นใยเพื่อช่วยระบบขับถ่าย และรับประทานอาหารและผลไม้ที่มีรสหวานให้น้อยที่สุด

อาจมีการใช้ยาในการป้องกันโรคเบาหวาน โดยต้องมีการปรึกษาแพทย์ผู้เชี่ยวชาญ และจะมีการใช้ในกลุ่มที่ผ่านการคัดกรองแล้วว่ามีความเสี่ยงสูง (ความรู้ทั่วไปเรื่องเบาหวาน, 2558)

2.3 อาการโรคเบาหวาน

โรคเบาหวานเกิดจาก การทำงานผิดปกติของฮอร์โมนอินซูลิน ทำให้การเผาผลาญน้ำตาลเพื่อใช้เป็นพลังงานไม่สามารถทำได้ดี ทำให้เกิดการสะสมของน้ำตาลในเลือดสูง และจะถูกกรองออกมาทางปัสสาวะ บุคคลทั่วไปควรสังเกตตัวเองว่ามีอาการบ่งชี้ที่จะนำไปสู่การเป็นผู้ป่วยโรคเบาหวานหรือไม่ เพื่อจะได้ดำเนินการรักษาหรือป้องกันได้ทัน โดยอาการบ่งชี้ต่างๆ มีดังต่อไปนี้คือ

อาการหิวน้ำบ่อยและปัสสาวะบ่อยครั้งขึ้น

น้ำหนักลดผิดปกติ โดยประมาณ 5-10 กิโลกรัม ภายในเวลา 2-3 เดือน

อาการหิวบ่อยและอยากกินของขบเคี้ยวแบบไม่มีสาเหตุ ทั้งที่เพิ่งกินอาหารมื้อหลัก

อาการผิวหนังแห้ง อันเนื่องมาจากอาการแพ้อินซูลินในร่างกาย บางรายอาจมีอาการคันบนผิวหนัง

อาการบาดแผลหายช้า ถ้าหากหายช้ากว่าที่ควรจะเป็นก็อาจเป็นอาการของโรคเบาหวานได้

อาการติดเชื้อราง่าย เพราะเมื่อขาดน้ำตาลจะทำให้ภูมิคุ้มกันอ่อนแอลง การต้านทานเชื้อราก็ทำได้น้อย

อารมณ์แปรปรวนง่าย และมีอาการอ่อนเพลีย

อาการตาพร่ามัวมองไม่ชัด เห็นแสงลอยไปมา

อาการชาตามมือและเท้า มักจะเกิดกับผู้ที่เป็นเบาหวานติดต่อกันเป็นเวลานาน

ระดับน้ำตาลในเลือดสูง โดยต้องได้รับการตรวจคัดกรองจากผู้เชี่ยวชาญและมีการวินิจฉัยที่เหมาะสมจึงจะสามารถทราบได้

ผู้ที่มีอาการดังกล่าวข้างต้น ควรหมั่นสังเกตตัวเอง หากพบความผิดปกติต้องรีบปรึกษาแพทย์ผู้เชี่ยวชาญเพื่อการตรวจสอบ และวินิจฉัยโรคที่ละเอียดแม่นยำต่อไป (อาการโรคเบาหวาน, 2558)

2.4 แนวทางในการออกกำลังกายในผู้ป่วยเบาหวาน

(แพทย์หญิงเนติมา คูนีย์, 2555) เนื่องจากผู้ป่วยเบาหวาน มีความเสี่ยงต่อการได้รับผลกระทบจากการออกกำลังกาย (Exercise) ที่เกินจากร่างกายต้องการ จึงได้มีการกำหนดแบบแผนในการออกกำลังกาย โดยใช้ชื่อเรียกว่า กิจกรรมทางกาย (Physical Activity) ทางองค์การอนามัยโลก ได้มีการกำหนดความแตกต่างของทั้งสองคำดังนี้คือ

- การออกกำลังกาย มีการดำเนินการในรูปแบบซ้ำๆ กัน เพื่อเสริม หรือรักษาระดับของสมรรถภาพทางกาย

- กิจกรรมทางกาย เป็นกิจกรรมการเคลื่อนไหวส่วนต่างๆ ของร่างกาย เพื่อทำให้เกิดพลังงาน เช่น การทำงานบ้าน การประกอบอาชีพ การเดินทาง และงานอดิเรก

ประโยชน์ของกิจกรรมทางกายมีดังต่อไปนี้คือ

ด้านร่างกาย ช่วยลดอัตราการตาย ลดระดับน้ำตาลในเลือดและความดันโลหิต ทำให้ร่างกายมีความไวต่ออินซูลินเพิ่มขึ้น ไตรกรีเซอไรด์ลดลง และยังช่วยควบคุมน้ำหนักอีกด้วย

ด้านจิตใจ ส่งผลให้ผ่อนคลาย ลดความเครียด และลดอาการซึมเศร้าสำหรับผู้ป่วยที่มีอาการเรื้อรังได้

ด้านสังคม เนื่องจากถ้าหากมีการออกกำลังกายเป็นกลุ่มร่วมกันในชุมชน จะมีส่วนช่วยสร้างเครือข่ายชุมชนได้ มีส่วนลดความรู้สึกโดดเดี่ยวของผู้ป่วยลงได้

ด้านเศรษฐกิจ ทำให้ช่วยลดค่าใช้จ่ายลง เนื่องจากทำให้สุขภาพโดยรวมของผู้ป่วยดีขึ้น

2.5 ความเสี่ยงจากการออกกำลังกายของผู้ป่วยเบาหวาน

(แพทย์หญิงเนติมา คูนีย์, 2555) ผู้ป่วยเบาหวานอาจมีความเสี่ยง ที่เกิดขึ้นจากการออกกำลังกายที่ไม่เหมาะสม หรือไม่ได้มีการเตรียมความพร้อมทางด้านร่างกาย โดยความเสี่ยงต่างๆ ที่อาจเกิดขึ้นจากการออกกำลังกายได้แก่

การเสียชีวิตแบบเฉียบพลัน โดยเฉพาะในกรณีที่ผู้ป่วยมีโรคหัวใจและมีการออกกำลังกายในระดับหนัก ดังนั้นก่อนการออกกำลังกาย ควรต้องมีการตรวจสอบและวินิจฉัยอาการของผู้ป่วยโรคหัวใจก่อน

โรคกล้ามเนื้อหัวใจตาย มักจะเกิดขึ้นกับกรณีที่ผู้ป่วยอยู่ในกลุ่มที่ออกกำลังกายไม่สม่ำเสมอ แล้วมีการออกกำลังกายในระดับหนัก

การบาดเจ็บของระบบกระดูกและกล้ามเนื้อ เป็นอาการที่มีโอกาสเกิดขึ้นบ่อยที่สุด เกิดจากระดับความแรงและลักษณะของแรงกระแทกจากการออกกำลังกาย และการบาดเจ็บโดยทางอ้อม เช่น ข้ออักเสบ และการปวดหลัง

ความเสี่ยงในด้านอื่นๆ เช่น ภาวะน้ำตาลในเลือดผิดปกติ และความดันโลหิตสูงหลังการออกกำลังกาย

2.6 งานวิจัยที่เกี่ยวข้อง

กฤษณา คำลอยฟ้า (2011) ได้ทำงานวิจัยเรื่อง พฤติกรรมการดูแลสุขภาพตนเองของผู้ป่วยโรคเบาหวาน

ในคลินิกโรคเบาหวาน โรงพยาบาลแก้งสนามนาง จังหวัดนครราชสีมา โดยผลของการวิจัยพบว่าพฤติกรรมการดูแลสุขภาพตนเองของผู้ป่วยเบาหวานโดยรวมอยู่ในระดับมาก ซึ่งแยกเป็นด้านต่างๆ คือ ร่างกาย จิตใจ อารมณ์ และสังคม ส่วนคะแนนเฉลี่ยด้านพฤติกรรมการดูแลตนเองในด้านการป้องกันภาวะแทรกซ้อนอยู่ในระดับที่ต่ำ จึงมีข้อเสนอแนะให้สนับสนุนพัฒนาพฤติกรรมการป้องกัน ภาวะแทรกซ้อน เพื่อให้ผู้ป่วยเบาหวานมีภาวะสุขภาพที่ดีขึ้น

Vladimir Villarreal, Jesus Fontecha, Ramon Hervas, and Jose Bravo (2014) ได้ทำการวิจัยเรื่อง Mobile and ubiquitous architecture for the medical control of chronic diseases through the use of intelligent devices: Using the architecture for patients with diabetes โดยทีมวิจัยได้ทำการพัฒนารูปแบบโปรแกรมบนโทรศัพท์มือถือ เพื่อใช้ในการติดตามและควบคุมอาการ โดยได้กำหนดเป็นรูปแบบกึ่งอัตโนมัติ ให้สามารถปรับใช้กับโรคชนิดอื่นๆ ได้ และได้เน้นไปที่การใช้เป็นโปรแกรมในการควบคุมอาการต่างๆ ของผู้ป่วย เพื่อมุ่งเน้นให้มีคุณภาพชีวิตที่ดีขึ้น และเสนอรูปแบบมาตรฐานในการพัฒนาโปรแกรม ซึ่งประกอบด้วยคุณสมบัติด้านความง่ายในการใช้งาน ความสามารถในการเพิ่มเติมข้อมูลและฟังก์ชันการทำงาน ความสามารถในการเชื่อมโยงและประสานข้อมูล รวมถึงความสามารถในการปรับแต่งระบบให้สอดคล้องกับความต้องการ ของผู้ใช้งาน

Predrag Klasnja and Wanda Pratt (2011) ได้ทำการวิจัยเรื่อง Healthcare in the pocket: Mapping the space of mobile-phone health interventions ศึกษาภาพรวมของการใช้งานโทรศัพท์มือถือ ในการรับส่งข้อมูลที่เกี่ยวข้องทางการแพทย์ มีการศึกษารูปแบบการทำงานของโทรศัพท์มือถือที่เหมาะสมต่อการใช้งาน 5 ประเภท ได้แก่ ติดตามข้อมูลทางด้านสุขภาพ การติดต่อทีมผู้ให้การรักษา การใช้งานด้านเครือข่ายสังคม การเข้าถึงข้อมูลทางด้านสุขภาพและการใช้งานเพื่อวัตถุประสงค์ในการผ่อนคลาย นอกจากนี้ยังได้ศึกษาถึงแนวโน้มของการใช้งานโทรศัพท์มือถือ ในงานด้านการแพทย์

Kevin Patrick, William G. Griswold, Fred Raab, and Stephen S. Intille (2008) กล่าวถึงเทคโนโลยีโทรศัพท์มือถือกับการใช้เพื่อข้อมูลทางการแพทย์ ไว้ในงานวิจัยเรื่อง Health and the Mobile Phone ว่าเทคโนโลยีของโทรศัพท์มือถือในปัจจุบัน มีประสิทธิภาพสูงขึ้นแต่ราคาถูกลง ทำให้ถูกนำมาใช้ในด้านการให้บริการด้านสุขภาพ รวมถึงสนับสนุนด้านการรักษาสุขภาพส่วนบุคคล แต่ปัญหาที่พบคือ ยังไม่สามารถเข้าถึงประชากรส่วนใหญ่ได้ และยิ่งทำให้ผู้ที่ไม่สามารถเข้าถึงข้อมูลสุขภาพในช่องทางอื่นๆได้ ยิ่งห่างไกลจากแหล่งข้อมูลมากขึ้น

3. วิธีการดำเนินการวิจัย

ขอบเขตด้านประชากรเป็นผู้ป่วยเบาหวานในประเทศไทยที่ไม่ต้องใช้อินซูลิน เนื่องจากยังอยู่ในกลุ่มที่มีสุขภาพแข็งแรงเพียงพอที่จะเข้าสู่โปรแกรมการดูแลสุขภาพตนเอง โดยต้องเป็นผู้ป่วยที่มีการใช้งานโทรศัพท์สมาร์ทโฟน และเคยใช้งานแอพพลิเคชั่นในสมาร์ทโฟนอย่างน้อยหนึ่งประเภท ไม่จำกัดอายุ การสุ่มตัวอย่างใช้วิธีการสุ่มตามความสะดวก ในช่วง มกราคม ถึง กุมภาพันธ์ พ.ศ.2558 ใช้สูตรของ Cochran ในการคำนวณประชากรกลุ่มตัวอย่าง เนื่องจากไม่ทราบจำนวนประชากรที่แน่นอน ใช้ค่าความความเชื่อมั่น 95% และมีค่าความผิดพลาด ±5% ได้ค่าประชากรกลุ่มตัวอย่าง 385 คน นำข้อมูลมาวิเคราะห์เพื่อหาพฤติกรรมและความต้องการด้านการส่งเสริมสุขภาพ อันนำไปสู่การออกแบบระบบต้นแบบต่อไป สถิติที่ใช้ในการวิเคราะห์ เป็นสถิติเชิงพรรณนา และเชิงอนุมาน ได้แก่ Independent t-test และ One-way ANOVA

4. ผลการดำเนินงาน

4.1 การวิเคราะห์ข้อมูลประชากรศาสตร์

ตารางที่ 1 : แสดงจำนวน (ความถี่) และค่าร้อยละทางประชากรศาสตร์

	ความถี่	ร้อยละ
ชาย	165	42.9
หญิง	220	57.1
อายุต่ำกว่า 35 ปี	105	27.3
อายุ 35 ปีขึ้นไป – 55 ปี	136	35.3
อายุ 55 ปีขึ้นไป – 75 ปี	113	29.4
อายุ 75 ปีขึ้นไป	31	8.1
เป็นโรคเบาหวานต่ำกว่า 5 ปี	200	52.2
เป็นโรคเบาหวาน 5 ปีขึ้นไป–10 ปี	134	34.8
เป็นโรคเบาหวาน 10 ปี ขึ้นไป–20 ปี	40	10.4
เป็นโรคเบาหวาน 20 ปีขึ้นไป	10	2.6

4.2 การวิเคราะห์ข้อมูลด้านพฤติกรรมการดูแลสุขภาพ

ตารางที่ 2: แสดงพฤติกรรมการดูแลสุขภาพของผู้ป่วยเบาหวาน

	$\bar{x}$	S.D.	อันดับ
ด้านร่างกาย			
1. ออกกำลังกายที่เหมาะสม	3.02	1.132	1
2. ออกกำลังกายทุกวันครั้งละ 10-20 นาที	3.02	1.132	2
3. ออกกำลังจนมีอาการหอบน้อยๆและเหงื่อซึม	3.02	1.132	3
ด้านจิตใจ อารมณ์ สังคม			
1. คิดว่าเป็นไปไม่ได้ที่จะเป็นโรคเบาหวาน	3.01	1.076	3
2. มีความเครียดกับการควบคุมอาการไม่ได้	3.03	1.053	2
3. เกิดความกังวลต่ออาการแทรกซ้อน	3.19	1.145	1
ด้านการป้องกันภาวะแทรกซ้อน			
1. หลีกเลี่ยงอาหารที่ได้รับคำแนะนำให้เลี่ยง เช่น น้ำตาล ไขมัน	3.77	1.006	2
2. ดูแลเท้าตามคำแนะนำเรื่องการบริหารเท้าสำหรับผู้ป่วยเบาหวาน	3.78	1.040	1
ด้านการรักษา			
1. ยาก่อนอาหาร-ทานอาหารหลังยา 30 นาที	3.90	1.082	1
2. ยาพร้อมอาหาร-กินยาพร้อมอาหารคำแรก	3.43	1.291	3
3. ยาหลังอาหาร-ทานหลังอาหารทันที	3.83	1.136	2

ตารางที่ 3: แสดงพฤติกรรมการดูแลสุขภาพของผู้ป่วยเบาหวาน

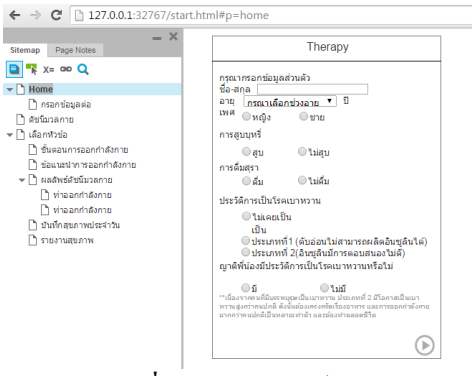
พฤติกรรมการดูแลสุขภาพของผู้ป่วยเบาหวาน	$\bar{x}$	S.D.	อันดับ
1. ด้านร่างกาย	3.16	0.919	3
2. ด้านจิตใจ อารมณ์ สังคม	3.08	0.909	4
3. ด้านการป้องกันภาวะแทรกซ้อน	3.78	0.915	1
4. ด้านการรักษา	3.72	0.931	2

ตารางที่ 4: การทดสอบสมมติฐาน

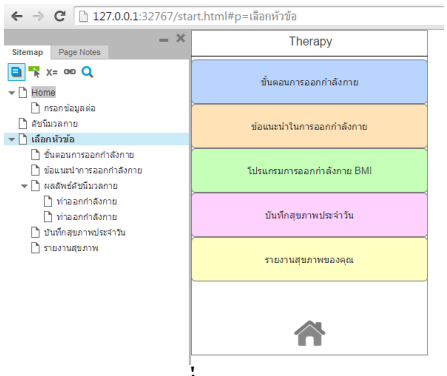
สมมติฐานที่ 1 ลักษณะด้านประชากรศาสตร์ที่แตกต่างกันส่งผลต่อพฤติกรรมการดูแลสุขภาพของผู้ป่วยเบาหวานที่แตกต่างกัน

ลักษณะประชากรศาสตร์	พฤติกรรมการดูแลสุขภาพของผู้ป่วยเบาหวาน			
	ด้านร่างกาย	ด้านจิตใจอารมณ์สังคม	ด้านการป้องกันภาวะแทรกซ้อน	ด้านการรักษา
สมมติฐานที่ 1.1 เพศ	✓	-	✓	-
สมมติฐานที่ 1.2 อายุ	✓	✓	-	-
สมมติฐานที่ 1.3 ระยะเวลาที่ป่วยเป็นโรคเบาหวาน	✓	✓	-	-

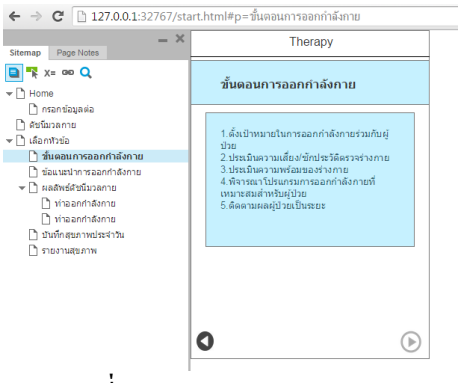
4.3 โปรแกรมต้นแบบการออกกำลังกาย



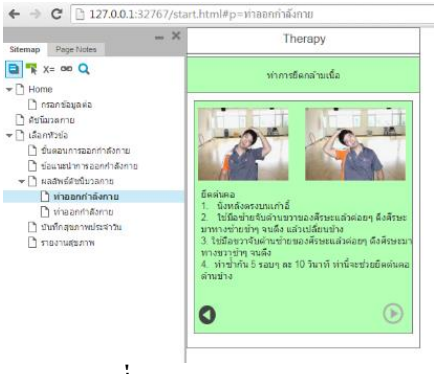
ภาพที่ 2: ข้อมูลส่วนตัวเบื้องต้น



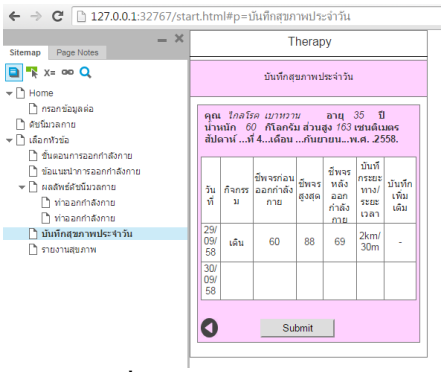
ภาพที่ 3: หัวข้อเมนู



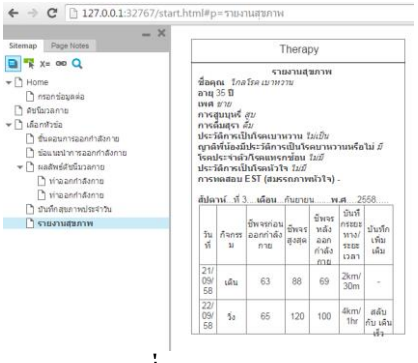
ภาพที่ 4: ข้อแนะนำในการออกกำลังกาย



ภาพที่ 5: ท่าทางการออกกำลังกาย



ภาพที่ 6: บันทึกสุขภาพประจำวัน



ภาพที่ 7: รายงานสุขภาพ

5. สรุป

1. ผู้ตอบแบบสอบถามส่วนใหญ่ให้ความสำคัญกับพฤติกรรมการดูแลสุขภาพด้านการป้องกันภาวะแทรกซ้อนและด้านการรักษาเป็นอย่างมาก เนื่องจากเป็นเรื่องที่ต้องอาศัยความรู้ในการปฏิบัติตนจากแพทย์ผู้เชี่ยวชาญ จึงต้องการช่องทางการรับข้อมูลข่าวสารที่ถูกต้อง เพื่อมั่นใจว่าตนเองสามารถเฝ้าระวังต่างๆ ที่จะเกิดภาวะแทรกซ้อนตามมาได้

2. ผู้ป่วยโรคเบาหวานอายุน้อยและที่มีระยะเวลาป่วยเป็นโรคเบาหวานต่ำกว่า 5 ปี มีการดูแลสุขภาพมากกว่าผู้ที่อายุมากและป่วยมาแล้วหลายปี ในด้านจิตใจ อารมณ์ สังคม สรุปได้ว่าการเจ็บป่วยเป็นระยะเวลานานๆ อาจทำให้ผู้ป่วยหมดกำลังใจในการดูแลตนเอง ดังนั้นควรส่งเสริมพฤติกรรมด้านจิตใจให้ผู้ป่วยมีสุขภาพจิตดีและคุณภาพชีวิตที่ดีขึ้น รวมถึงไม่เป็นการวิตกกังวลกับโรคภัยจนเกินไป

การศึกษาครั้งต่อไปควรศึกษาถึงผลกระทบที่ส่งผลต่อการดูแลสุขภาพของผู้ป่วยเบาหวาน เพื่อให้ทราบถึงปัจจัยที่ทำให้ผู้ป่วยเบาหวานไม่สามารถดูแลตนเองได้ถูกต้อง และนำข้อมูลด้านนี้ไปวิเคราะห์ในการพัฒนาฟังก์ชันการทำงานเพิ่มเติมในโปรแกรม เช่น การนำเอาข้อมูลในรายงานสุขภาพ และ อาการของผู้ใช้งานที่ได้รับ มาวิเคราะห์เพิ่มเติมให้เหมาะสมกับแต่ละบุคคล

ควรศึกษาถึงประสิทธิภาพในการใช้งานของโปรแกรมที่ใช้งานบนสมาร์ทโฟน(Application) การดูแลสุขภาพของผู้ป่วยเบาหวาน เพื่อนำมาพัฒนาให้สอดคล้องกับผู้ป่วยในแต่ละช่วงอายุ และระยะเวลาการป่วย ซึ่งจะทำให้ผู้ใช้งานได้รับข้อมูลอย่างถูกต้องและเหมาะสม

เอกสารอ้างอิง

[1] โรคเบาหวาน สาเหตุเบาหวาน อาการเบาหวาน อาหารเบาหวาน. เข้าถึงเมื่อวันที่ 21 กันยายน 2558 http://thaidiabetes.blogspot.com/

[2] ความรู้ทั่วไปเรื่องเบาหวาน. Siamhealth.net. เข้าถึงเมื่อวันที่ 21 กันยายน 2558 http://www.siamhealth.net/public_html/Disease/endocrine/DM/intro.htm#.VgKB-tLtlBd

[3] อาการโรคเบาหวาน 10 สัญญาณควรระวัง เช็กได้ง่ายนิดเดียว. กระปุกดอทคอม. เข้าถึงเมื่อวันที่ 21 กันยายน 2558 http://health.kapook.com/view121479.html

[4] แพทย์หญิงเนติมา คูนีย์ (2555), แนวทางเวชปฏิบัติการออกกำลังกายในผู้ป่วยเบาหวานและความดันโลหิตสูง (Exercise in Patients with Diabetes and Hypertension), สถาบันวิจัยและประเมินเทคโนโลยีทางการแพทย์ กรมการแพทย์

[5] กฤษณา คำลอยฟ้า (2011), พฤติกรรมการดูแลสุขภาพตนเองของผู้ป่วยโรคเบาหวานในคลินิกโรคเบาหวาน โรงพยาบาลแก้งสนามนาง จังหวัดนครราชสีมา ,The Journal of Baramarajonani College of Nursing, Nakhnratchasima, Vol. 17 No. 1 January-June 2011

[6] Vladimir Villarreal, Jesus Fontecha, Ramon Hervas, Jose Bravo (2014), Mobile and ubiquitous architecture for the medical control of chronic diseases through the use of intelligent devices: Using the architecture for patients with diabetes, Future Generation Computer Systems, Volume 34, May 2014, Pages 161-175. Available at http://www.sciencedirect.com/science/article/pii/S0167739X1300277X

[7] Predrag Klasnja and Wanda Pratt (2011), Healthcare in the pocket: Mapping the space of mobile-phone health intervention, Journal of Biomedical Informatics, Volume 45, Issue 1, February 2012, Pages 184–198. Available at http://www.sciencedirect.com/science/article/pii/S1532046411001444

[8] Kevin Patrick, William G. Griswold, Fred Raab, Stephen S. Intille (2008), Health and the Mobile Phone, American Journal of Preventive Medicine, Volume 35, Issue 2, August 2008, Pages 177-181. Available at http://www.sciencedirect.com/science/article/pii/S0749379708004078

ระบบแนะนำสารอาหารส่วนบุคคลสำหรับผู้ป่วยเบาหวาน โดยใช้ออนโทโลยีและวิธีการอนุมาน

Personalized Nutrient Recommendation System for Diabetic Patient using Ontology and Rule of Inference Approach

พัทนีย์ อาบุญงาม (Pattanee Arboonngam)[1] งามนิจ อาจอินทร์ (Ngamnij Arch-int)[2]
สมจิตร อาจอินทร์ (Somjit Arch-int)[3] และสุกัญญา ขยันทำ (Sukanya Khayantham)[4]
[1,2,3]สาขาวิชาเทคโนโลยีสารสนเทศ ภาควิชาวิทยาการคอมพิวเตอร์ มหาวิทยาลัยขอนแก่น

บทคัดย่อ

การควบคุมอาหารเป็นสิ่งจำเป็นสำหรับผู้ป่วยเบาหวาน โดยผู้ป่วยต้องรู้จักเลือกรับประทานชนิดและปริมาณอาหารที่เหมาะสมกับตนเอง เพื่อให้ร่างกายสามารถควบคุมระดับน้ำตาลและไขมันในเลือด รวมถึงป้องกันหรือลดภาวะแทรกซ้อนของโรคได้ งานวิจัยนี้จึงนำเสนอระบบแนะนำสารอาหารส่วนบุคคลสำหรับผู้ป่วยเบาหวาน โดยระบบจะคำนวณและแนะนำสารอาหารที่เหมาะสมให้กับผู้ป่วยแต่ละรายโดยใช้หลักการอาหารแลกเปลี่ยน โดยคำนึงถึงปัจจัยต่าง ๆ ของผู้ป่วยทั้งข้อมูลส่วนตัวและโรคแทรกซ้อน งานวิจัยนี้ได้พัฒนาออนโทโลยีสำหรับข้อมูลผู้ป่วยเบาหวาน โดยใช้ภาษา OWL และสร้างกฎการอนุมานโดยใช้ภาษา SWRL เพื่อคำนวณส่วนของอาหารและแนะนำสารอาหารที่เหมาะสมกับผู้ป่วยได้เสมือนผู้เชี่ยวชาญแนะนำ ซึ่งระบบงานนี้จะถูกประเมินความถูกต้องโดยผู้เชี่ยวชาญและมีการวัดความถูกต้องในการให้คำแนะนำของระบบโดยใช้ค่า F-measure เฉลี่ย 0.94

คำสำคัญ: ออนโทโลยี ผู้ป่วยเบาหวาน สารอาหารแนะนำ กฎการอนุมาน

Abstract

Diet balancing is essential for a patient with diabetes. They need to comprehend how to deliberately select each meal, in either types or quantities that is beneficial to their body, helps maintaining the levels of glucose and lipids in the blood, as well as helps preventing or reducing the risk of complications. Being concerned to assist people with diabetes, this research proposes the personalized nutrition recommendation system for a patient with diabetes. The system can do the calculations and provide proper advisory information for individual patients according to food exchange program. Each patient's personal information and complication condition are also taken into account. In developing this system, OWL is used for representing the patient ontology and SWRL is used for calculation and providing advisory information regarding the servings and nutrition facts as appropriately as an expert. This system is evaluated by expert and measured the precision of system in recommendation with average F-measure of 0.94

Keyword: Ontology, Diabetic Patient, Nutrition Recommendation, Rules of Inference.

1. บทนำ

สถานการณ์โรคเบาหวานประเทศไทยในปัจจุบันเป็นโรคเรื้อรังที่เป็นปัญหาสาเหตุการเสียชีวิตของคนไทย และมีแนวโน้มที่จะเพิ่มขึ้นเรื่อย ๆ การปรับเปลี่ยนวิถีการดำรงชีวิตประจำวันในการควบคุมอาหารจึงจำเป็นสำหรับผู้ป่วยเบาหวาน เพื่อช่วยให้ผู้ป่วยมีสุขภาพที่แข็งแรงสามารถควบคุมระดับน้ำตาลและไขมันในเลือด รวมถึงช่วยป้องกันลดภาวะแทรกซ้อนของโรคได้ [1] เบาหวานควบคุมได้ด้วยการควบคุมชนิดและปริมาณอาหารที่รับประทานแต่ละมื้อ ปัจจุบันผู้ป่วยเบาหวานสามารถเลือกรับประทานอาหารต่างๆ ได้

เช่นเดียวกับคนปกติ แต่จำเป็นต้องเรียนรู้ว่ารับประทานได้มากน้อยเพียงใดจึงจะไม่ทำให้น้ำตาลในเลือดสูง

องค์กรเกี่ยวกับสุขภาพมีการเผยแพร่เอกสารเกี่ยวกับการป้องกันและแนะนำการดูแลตนเองของโรคเบาหวาน แต่เป็นเพียงเอกสารพื้นฐานและเข้าใจได้ยากไม่มีความแน่นอนเกี่ยวกับเงื่อนไขต่าง ๆ ของผู้ป่วย ทำให้ปัจจุบันมีนักวิจัยเล็งเห็นความสำคัญในการสร้างระบบการแนะนำการดูแลตนเองส่วนบุคคลสำหรับผู้ป่วยเบาหวาน ทั้งแนะนำการปฏิบัติตนพื้นฐาน การรับประทานอาหารให้กับผู้ป่วยเบาหาน โดยอาศัยประเภทผู้ป่วยเบาหวาน โรคแทรกซ้อน ผลตรวจทางร่างกาย ผลตรวจห้องปฏิบัติการ แต่ยังขาดในส่วนการคำนวณปริมาณสารอาหารและคำนวณส่วนอาหารที่ผู้ป่วยควรได้รับที่เป็นกุญแจสำคัญที่ใช้ในการควบคุมระดับน้ำตาลในเลือด [2] บางงานวิจัยได้มีการพัฒนาระบบฐานความรู้อัจฉริยะด้านเบาหวานซึ่งเป็นระบบเกี่ยวกับการวางแผนการรับประทานอาหารเพื่อสุขภาพอัตโนมัติโดยใช้พีระมิดอาหารเบาหวาน (Diabetic Food Pyramid) ในการจัดกลุ่มอาหาร และใช้ข้อมูลส่วนตัว น้ำหนัก ส่วนสูง อายุและกิจกรรมทางกาย รวมถึงโรคอื่น ๆ ของผู้ป่วยมาคำนวณพลังงานสารอาหารที่ผู้ป่วยรับประทานได้และใช้กฎในแบ่งส่วนอาหาร แล้วเลือกอาหารในหมวดหมู่นั้นๆออกมาแนะนำ แต่ยังขาดผลตรวจด้านสุขภาพที่อาจมีภาวะแทรกซ้อนที่จำเป็นต่อการแนะนำอาหาร รวมถึงขาดกฎที่ใช้ในการแนะนำประเภทอาหาร ทั้งอาหารแนะนำที่ควรรับประทาน อาหารที่รับประทานได้แต่จำกัด และอาหารที่ควรหลีกเลี่ยง ที่เป็นจุดสำคัญของอาหารผู้ป่วยเบาหวาน เพื่อให้น้ำตาลในเลือดลดลงสู่เกณฑ์ปกติ

ผู้วิจัยจึงมีแนวคิดในการศึกษาและออกแบบพัฒนาระบบแนะนำสารอาหารส่วนบุคคลสำหรับผู้ป่วยเบาหวานที่จะเน้นการแนะนำสารอาหารที่เหมาะสมสำหรับผู้ป่วยเบาหวานที่อาจมีภาวะแทรกซ้อน โดยจะใช้ข้อมูลของผู้ป่วยเบาหวานในการสร้างเงื่อนไขในการคำนวณสารอาหารตามความต้องการพลังงานต่อวันสอดคล้องกับภาวะสุขภาพของผู้ป่วยเบาหวานตามหลักการอาหารแลกเปลี่ยน (Food Exchange) [3] แล้วเลือกอาหารแนะนำที่ควรรับประทาน อาหารที่รับประทานได้แต่จำกัด และอาหารที่ควรหลีกเลี่ยง พร้อมทั้งแสดงปริมาณอาหารที่สามารถรับประทานได้ ปริมาณสารอาหารต่าง ๆ รวมถึงปริมาณเครื่องปรุงรส โดยผู้วิจัยได้นำหลักการเว็บเชิงความหมาย (Semantic Web) ที่จะทำให้ข้อมูลนั้นมีความหมายที่ชัดเจนขึ้น โดยจะช่วยในการนำเสนอข้อมูลแบบมีโครงสร้าง ทำให้ช่วยในการวิเคราะห์หรือจำแนกข้อมูลที่มีความสัมพันธ์กับข้อมูลอื่นในแต่ละระดับ การพัฒนาระบบจะเป็นการสร้างฐานความรู้ออนโทโลยี (Ontology) โดยใช้ภาษา OWL (Web Ontology Language) และมีการสร้างกฎการอนุมานโดยใช้ภาษา SWRL (Semantic Web Rule Language) เพื่อจะควบคุมและแนะนำอาหารให้ผู้ป่วยได้อย่างเป็นอัตโนมัติถูกต้องเหมาะสมสำหรับเงื่อนไขของผู้ป่วยที่แตกต่างกันมากยิ่งขึ้น

2. วรรณกรรมที่เกี่ยวข้อง

2.1 ออนโทโลยี

ออนโทโลยี (Ontology) [4] เป็นการอธิบายถึงขอบเขตเรื่องใดเรื่องหนึ่งหรือสิ่งที่เราสนใจ (Domain) อย่างมีโครงสร้าง โดยนำเสนอความหมาย คุณลักษณะ และความสัมพันธ์ ขยายคำศัพท์ที่ใช้เป็นตัวแทนแนวคิด (Concept) ให้สามารถเข้าใจร่วมกันได้ รวมถึงสนับสนุนการแลกเปลี่ยน แบ่งปันความรู้และสามารถนำมาใช้ร่วมกันใหม่ได้

2.2 ภาษาเชิงความหมาย

OWL (Web Ontology Language) [5] เป็นภาษาที่ใช้ในการพัฒนาออนโทโลยีมีโครงสร้างพื้นฐานมากจาก RDF (Resource Description Framework) แต่มีการสื่อความหมายมากขึ้น โดยจะมีการอธิบายความหมายของความสัมพันธ์ (Relationships) ระหว่างคลาส (Class) และคุณสมบัติ (Properties) เพิ่มส่วนของการกำหนดชนิดและขนาดข้อมูล พร้อมทั้งอธิบายเชิงตรรกะเข้าไป ทำให้คอมพิวเตอร์สามารถอ่านค่าและเข้าใจความหมายของขอบเขตต่าง ๆ มากขึ้น

2.3 กฎเชิงความหมาย

SWRL (Semantic Web Rule Language) [6] เป็นภาษาที่มีพื้นฐานอยู่บนภาษา OWL Lite และ OWL DL ซึ่งเป็นภาษาย่อยของ OWL โดยใช้โครงสร้างตาม RuleML XML Syntax สามารถสร้างกฎเพิ่มเติมในการนำมาใช้งานกับออนโทโลยีภาษา OWL ได้โดยกฎที่สร้างขึ้นจะประกอบด้วย 2 ส่วน คือ ส่วนของด้านซ้ายมือจะเป็นเงื่อนไข (Antecedent) มีลูกศรชี้ไป

ด้านขวามือจะเป็นส่วนของผลลัพธ์ (Consequent) และจะมีเครื่องหมาย ? นำหน้าตัวแปรหรือตัวอย่างข้อมูล

2.4 งานวิจัยที่เกี่ยวข้อง

จากการศึกษางานวิจัยมีส่วนสำคัญต่อการพัฒนาระบบแนะนำสารอาหารส่วนบุคคลสำหรับผู้ป่วยเบาหวาน ดังนี้

ในปี ค.ศ.2008 Chang-Shing Lee และคณะผู้วิจัย ได้พัฒนาเอเจนต์อัจฉริยะสำหรับการแนะนำอาหารผู้ป่วยโรคเบาหวาน [7] พัฒนาโดย แบ่งออนโทโลยีเป็น 2 ส่วน คือ 1) ออนโทลีอาหารไต้หวัน (Taiwanese Food Ontology) ที่จะแบ่งอาหารออกเป็น 6 หมู่ ตามหลักการอาหารแลกเปลี่ยน 2) ออนโทโลยีเฉพาะบุคคล (Personal Food Ontology) ที่จะใช้ข้อมูลประวัติ (Profile) ส่วนเป้าหมายการบริโภค (Diet Goal) และส่วนของอาหารที่ชื่นชอบ (Favorite) ใช้กลไกอนุมานฟัซซี่ (Fuzzy Inference Mechanism) แนะนำรายการอาหารสำหรับมื้อเย็นแยกเป็น 6 หมู่ และบอกสัดส่วนที่สามารถรับประทานได้ แต่ยังขาดในส่วนของการนำประวัติด้านสุขภาพ ผลตรวจต่างๆ ที่จำเป็นต่อการแนะนำอาหารเข้ามาพิจารณา

ในปี ค.ศ.2015 Benjamas Hempo และคณะผู้วิจัย ได้มีการพัฒนาระบบการแนะนำการดูแลตนเองส่วนบุคคลสำหรับผู้ป่วยเบาหวาน [8] โดยการสร้างฐานความรู้ผู้ป่วยเบาหวาน (Diabetes knowledge-based Ontology หรือ DKOs) ซึ่งจะเก็บข้อมูลส่วนตัวผู้ป่วย ผลตรวจทางร่างกาย ผลตรวจทางห้องปฏิบัติการ โรคแทรกซ้อนและความผิดปกติ แล้วใช้ OWL Property และ SWRL Rule เพื่อการแนะนำการปฏิบัติตนพื้นฐานและการแนะนำอาหาร แต่ยังขาดการคำนวณสารอาหารตามหลักการอาหารแลกเปลี่ยน

งานวิจัยของ Ibrahin M.Ahmed และคณะผู้วิจัย ได้มีการพัฒนาระบบฐานความรู้อัจฉริยะด้านเบาหวาน [9] เพื่อวางแผนการรับประทานอาหารเพื่อสุขภาพโดยอัตโนมัติให้กับผู้ป่วยเบาหวานชนิดที่ 2 ให้คล้ายกับคำแนะนำของผู้เชี่ยวชาญของมนุษย์ โดยสร้างออนโทโลยีอาหารตามพีระมิดอาหารเบาหวาน ซึ่งแบ่งกลุ่มอาหารออกเป็น 6 หมู่ แนะนำอาหารโดยใช้กฎการอนุมาน IF-Then Rule ซึ่งพิจารณาจากเพศ น้ำหนัก ส่วนสูง กิจกรรมทางกาย เพื่อคำนวณพลังงานสารอาหารที่ผู้ป่วยรับประทานได้และใช้อายุรวมถึงโรคอื่นๆที่ผู้ป่วยเป็นแบ่งส่วนอาหารให้กับผู้ป่วยตามเงื่อนไขที่สร้างไว้ แต่ยังขาดในส่วนของผลตรวจห้องปฏิบัติการ ภาวะแทรกซ้อนโรคเบาหวานที่จะใช้สร้างกฎเพื่อแนะนำชนิดอาหารที่เหมาะสมกับผู้ป่วย คืออาหารแนะนำที่ควรรับประทาน อาหารที่รับประทานได้แต่จำกัด และอาหารที่ควรหลีกเลี่ยง รวมถึงคำแนะนำที่เสมือนผู้เชี่ยวชาญแนะนำ

3. วิธีการดำเนินงาน

จากการศึกษาเอกสารทฤษฎีและงานวิจัยที่เกี่ยวข้อง ผู้วิจัยได้ออกแบบสถาปัตยกรรมของระบบ การออกแบบออนโทโลยีและกฎต่างๆ ดังนี้

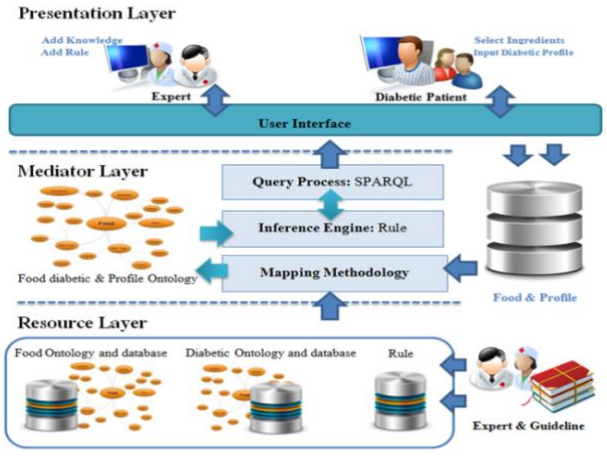
3.1 การออกแบบสถาปัตยกรรมของระบบ

สถาปัตยกรรมของระบบจะประกอบด้วยชั้นต่าง ๆ ดังแสดงในภาพที่ 1 ที่มีรายละเอียด ดังต่อไปนี้

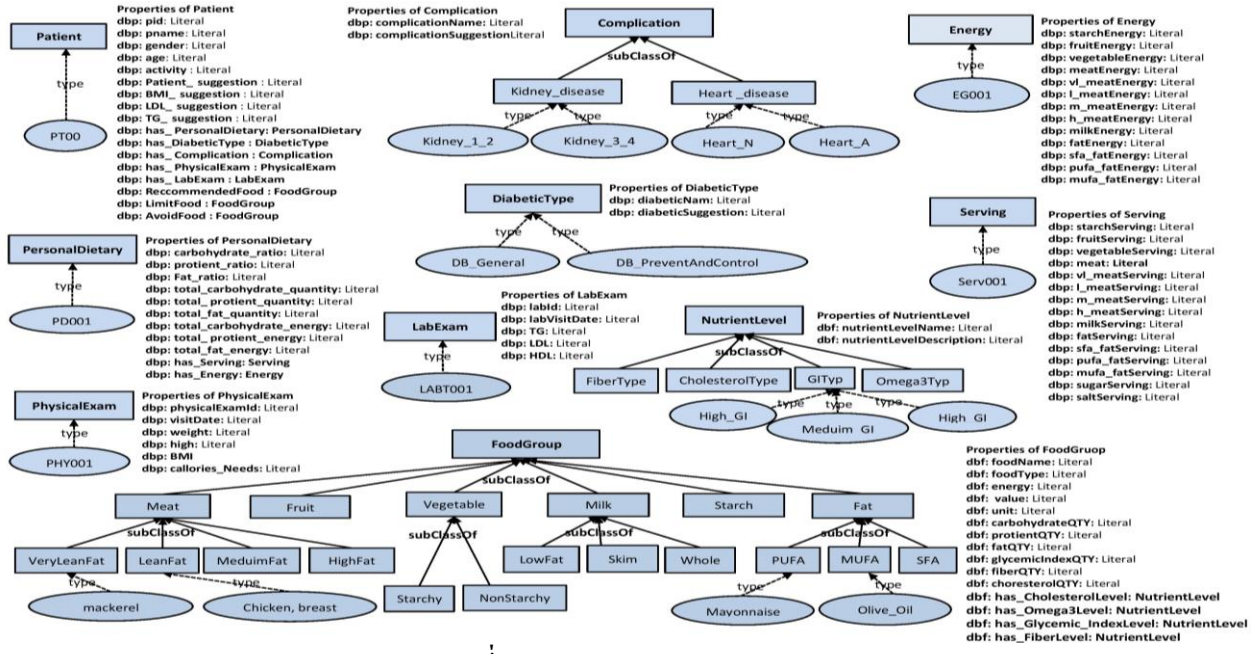
1) ชั้นนำเสนอ (Presentation Layer) เป็นส่วนติดต่อประสานงานกับผู้ใช้งาน 2 กลุ่ม คือ 1) ผู้ป่วย ที่สามารถกรอกข้อมูลส่วนตัว ประวัติสุขภาพ ภาวะแทรกซ้อนโรคเบาหวาน และเลือกวัตถุดิบที่ต้องการรับประทานได้ 2) ผู้เชี่ยวชาญ ที่สามารถเพิ่มความรู้เกี่ยวกับสารอาหารและกฎต่าง ๆ โดยส่วนนี้จะรับข้อมูลเพื่อนำไปจัดเก็บในฐานข้อมูลและแสดงผลลัพธ์ให้กับผู้ใช้งาน

2) ชั้นกลาง (Mediator Layer) เป็นส่วนที่เชื่อมต่อและทำการอนุมาน เพื่อแนะนำอาหารที่เหมาะสมให้แก่ผู้ป่วย

3) ชั้นแหล่งข้อมูล (Resource Layer) เป็นส่วนที่จะเก็บข้อมูลความรู้เกี่ยวกับโรคเบาหวานและคุณค่าโภชนาการของอาหาร รวมถึงกฎที่ใช้ในการอนุมานให้กับระบบ



ภาพที่ 1: สถาปัตยกรรมระบบแนะนำสารอาหารส่วนบุคคลสำหรับผู้ป่วยเบาหวาน



ภาพที่ 2: ออนโทโลยีอาหารผู้ป่วยเบาหวาน

3.2 การออกแบบออนโทโลยีอาหารผู้ป่วยเบาหวาน

การออกแบบออนโทโลยีสร้างโดยโปรแกรม Protégé และใช้ภาษา OWL โดยโครงสร้างออนโทโลยีจะแสดงในภาพที่ 2 ซึ่งจะแบ่งเป็น 2 ส่วน ดังนี้

3.2.1 ออนโทโลยีผู้ป่วยเบาหวาน

ในส่วนนี้จะเป็นออนโทโลยีที่จัดเก็บข้อมูลพื้นฐานของผู้ป่วยและส่วนของการคำนวณสารอาหารที่เหมาะสมสำหรับผู้ป่วยเบาหวาน ดังนี้

1) คลาสผู้ป่วย (Patient) คลาสเก็บข้อมูลส่วนตัวของผู้ป่วยทั้งชื่อ อายุ เพศ กิจกรรมทางกาย และคำแนะนำเกี่ยวกับอาหารสำหรับผู้ป่วยเบาหวาน ประเภทผู้ป่วยเบาหวาน (has_DiabeticType), มีโรคแทรกซ้อน (has_Complication), มีผลตรวจทางร่างกาย (has_PhysicalExam), มีผลตรวจทางห้องปฏิบัติการ (has_LabExam) มีการบริโภคอาหารส่วนบุคคลของผู้ป่วย (has_PersonalDietary), อาหารแนะนำที่ควรรับประทาน (RecommendedFood), อาหารที่รับประทานได้แต่จำกัด (LimitFood) และสุดท้ายอาหารที่ควรหลีกเลี่ยง (AvoidFood)

2) คลาสประเภทผู้ป่วยเบาหวาน (DiabeticType) คลาสเก็บชื่อ คำแนะนำของประเภทผู้ป่วยเบาหวาน

3) คลาสโรคแทรกซ้อน (Complication) คลาสเก็บชื่อและคำแนะนำของโรคแทรกซ้อนเบาหวาน

4) คลาสผลตรวจทางร่างกาย (PhysicalExam) คลาสเก็บข้อมูลเกี่ยวกับร่างกาย เช่น ส่วนสูง น้ำหนัก ความต้องการพลังงานต่อวัน ค่าดัชนีมวลกาย และความดันโลหิต

5) คลาสผลตรวจทางห้องปฏิบัติการ (LabExam) คลาสเก็บข้อมูลผลตรวจระดับไขมันไตรกลีเซอร์ไรด์ในเลือด (TG) ระดับไขมันดี (HDL-C) และระดับไขมันไม่ดี (LDL-C) ในร่างกายของผู้ป่วยเบาหวาน

6) คลาสการบริโภคอาหารส่วนบุคคล (Personal Dietary) คลาสเก็บข้อมูลสัดส่วนสารอาหารหลักที่ควรรับประทานต่อวัน (คาร์โบไฮเดรต: โปรตีน: ไขมัน) พลังงานของสารอาหารที่ควรได้รับและปริมาณของสารอาหารที่ควรได้รับ โดยมีความสัมพันธ์กับคลาสอื่น ๆ คือ มีส่วนอาหาร (has_Serving) และมีพลังงานจากอาหาร (has_Energy)

7) คลาสส่วนอาหาร (Serving) คลาสเก็บข้อมูลส่วนของอาหารในกลุ่มต่าง ๆ ที่สามารถรับประทานได้ต่อวัน

8) คลาสพลังงานจากอาหาร (Energy) คลาสเก็บข้อมูลพลังงานจากอาหารแต่ละกลุ่มที่ผู้ป่วยรับประทานได้

3.2.2 ออนโทโลยีอาหาร

ออนโทโลยีอาหารประกอบด้วย 2 คลาสหลัก คือ คลาสกลุ่มอาหาร (FoodGroup) ซึ่งมีความสัมพันธ์กับคลาสระดับสารอาหาร (NutrientLevel) ดังนี้

1) คลาสกลุ่มอาหาร (FoodGroup) คลาสเก็บข้อมูลเกี่ยวกับอาหาร ทั้งปริมาณ, พลังงาน, ขนาด และสารอาหารต่างๆ คือ คาร์โบไฮเดรต โปรตีน ไขมัน ไฟเบอร์ ค่าดัชนีน้ำตาล คอเลสเตอรอล และโอเมก้า-3 แบ่งหมวดหมู่ตามหลักการอาหารแลกเปลี่ยนโดยแบ่งอาหารเป็น 6 คลาสย่อยคือ แป้ง (Starch), เนื้อสัตว์ (Meat), ผัก (Vegetable), ผลไม้ (Fruit), ไขมัน (Fat) และผลิตภัณฑ์จากนม (Milk)

2) คลาสระดับสารอาหาร (Nutrient Level) เป็นคลาสเก็บชื่อและคำอธิบายของระดับสารอาหารไว้ โดยระดับสารอาหารแบ่งเป็น 4 คลาส คือ ระดับค่าดัชนีน้ำตาล ระดับคอเลสเตอรอล ระดับใยอาหารและระดับโอเมก้า-3

3.3 การออกแบบกฎ

การออกแบบกฎที่ใช้ในการคำนวณและแนะนำสารอาหารจะใช้ภาษา SWRL ในการเขียนกฎอนุมานเงื่อนไข ผู้วิจัยได้ยกตัวอย่างข้อมูลผู้ป่วย ดังภาพที่ 3 และนำเสนอกฎเป็น 2 ส่วน ดังนี้

กรณีตัวอย่าง ผู้ป่วยเบาหวานโดยรวมไม่มีโรคแทรกซ้อน (General Diabetic Patient) เพศชาย (male) มีกิจกรรมทางกายในระดับปานกลาง (moderately active) ส่วนสูง (high) 160 cm น้ำหนัก (weight) 100 kg ความดันโลหิตปกติ (SBP < 140 mmHg และ DBP < 90 mmHg) มีไขมันไม่ดีในเลือดปกติ (LDL-C < 100 mg/dL^3) แต่มีผลตรวจไตรกลีเซอไรด์สูง (TG > 150 mg/dL)

ภาพที่ 3: ตัวอย่างข้อมูลผู้ป่วยที่จำเป็นกับการแนะนำอาหาร

1) กฎการคำนวณสารอาหารที่เหมาะสม

กฎที่ 1 การคำนวณค่าดัชนีมวลกาย (BMI) และความต้องการพลังงานต่อวัน (Calories Needs)

- BMI = weight (kg) / (high (m))^2
- Calories Needs for Male, BMI>25 and moderately active
  = (high -100) * 25

โดยการนำข้อมูลผู้ป่วยใน ภาพที่ 3 และสูตรข้างต้นมาใช้เขียนกฎตาม ภาพที่ 4

dbp:Patient(?p) ∧ dbp:activity(?p, "moderately active") ∧ dbp:gender(?p, "male") ∧ dbp:has_PhysicalExam(?p, ?phy) ∧ dbp:has_High(?phy, ?h) ∧ dbp:has_Weight(?phy, ?w) ∧ swrlb:multiply(?hd, ?h, 0.01) ∧ swrlb:pow(?hd2, ?hd, 2) ∧ swrlb:divide(?bmi, ?w, ?hd2) ∧ swrlb:subtract(?ideal, ?h, 100.0) ∧ swrlb:greaterThanOrEqual(?bmi, 25.0) ∧ swrlb:multiply(?cal, ?ideal, 25.0) → dbp:has_BMI(?phy, ?bmi) ∧ dbp:has_CalloriesNeed(?phy, ?cal)

ภาพที่ 4: ตัวอย่างกฎคำนวณความต้องการพลังงานต่อวัน

กฎที่ 2 การให้สัดส่วนสารอาหารหลักตามข้อมูลผู้ป่วยตาม ภาพที่ 3 ให้สัดส่วน คาร์โบไฮเดรต:โปรตีน:ไขมัน = 60 : 15 : 25 ตาม ภาพที่ 5

dbp:Patient(?p) ∧ dbp:has_DiabeticType(?p, DiabeticType_General) ∧ dbp:has_PhysicalExam(?p, ?phy) ∧ dbp:has_BMI(?phy, ?bmi) ∧ swrlb:greaterThan(?bmi, 25.0) ∧ dbp:has_LabExam(?p, ?l) ∧ dbp:TG(?l, ?tg) ∧ swrlb:greaterThan(?tg, 150) ∧ dbp:has_PersonalDietary (?p, ?pd) → dbp:carbohydrate_ratio(?pd, 60) ∧ dbp:protient_ratio(?pd, 15) ∧ dbp:fat_ratio(?pd, 25)

ภาพที่ 5: ตัวอย่างกฎการให้สัดส่วนสารอาหารหลัก

กฎที่ 3 การคำนวณหาพลังงานและปริมาณของสารอาหารหลัก (energy and quantity) โดยการนำตัวแปรความต้องการพลังงานต่อวัน (?cal) จาก ***กฎที่ 1*** มาคำนวณกับตัวแปรสัดส่วนสารอาหารหลักทั้งคาร์โบไฮเดรต, โปรตีนและไขมัน (ตัวแปร ?cr, ?pr, ?fr,) จาก ***กฎที่ 2*** ยกตัวอย่างคำนวณพลังงานและปริมาณคาร์โบไฮเดรตที่ผู้ป่วยควรได้รับ

- total_carbohydrate_energy = Calories Needs *Carbohydrate_ratio/100
- total_carbohydrate_quantity = total_carbohydrate_energy/4

ใช้ตัวแปรและสูตรข้างต้นสร้างกฎตาม ภาพที่ 6

dbp:Patient(?p) ∧ dbp:has_PhysicalExam(?p, ?phy) ∧ dbp:has_CalloriesNeed(?phy, ?cal) ∧ dbp:has_ PersonalDietary(?p, ?pd) ∧ dbp:fat_ratio(?pd, ?fr) ∧ dbp: carbohydrate_ratio(?pd, ?cr) ∧ dbp:protient_ratio(?pd, ?pr) ∧ swrlb:multiply(?tc, ?cr, 0.01, ?cal) ∧ swrlb:multiply(?tp, ?pr, 0.01, ?cal) ∧ swrlb:multiply(?tf, ?fr, 0.01, ?cal) ∧ swrlb:divide(?tcq, ?tc, 4.0) ∧ swrlb:divide(?tpq, ?tp, 4.0) ∧ swrlb:divide(?tfq, ?tf, 9.0) → dbp:total__fat_energy(?pd, ?tf) ∧ dbp:total__protient_energy(?pd, ?tp) ∧ dbp:total_carbohydrate_energy(?pd, ?tc) ∧ dbp:total__fat_quantity(?pd, ?tfq) ∧ dbp:total_protient_quantity(?pd, ?tpq) ∧ dbp:total_carbohydrate_quantity(?pd, ?tcq)

ภาพที่ 6: ตัวอย่างคำนวณพลังงานและปริมาณสารอาหารหลัก

กฎที่ 4 ให้ส่วนอาหาร (serving) สามกลุ่มแรกตามเงื่อนไขของข้อมูลผู้ป่วยใน ภาพที่ 3 ดังนี้ ผัก 3 ส่วน ผลไม้ 5 ส่วน ผู้ป่วยที่มีภาวะอ้วนเกินมาตรฐาน นม 0 ส่วน (ควรงดนม) สร้างกฎออกมาตาม ภาพที่ 7

dbp:Patient(?p) ∧ dbp:has_LabExam(?p, ?l) ∧ dbp:TG(?l, ?tg) ∧ swrlb:greaterThanOrEqual(?tg, 150) ∧ dbp:has_PhysicalExam (?p, ?phy) ∧ dbp:has_BMI(?phy, ?bmi) ∧ swrlb:greaterThan (?bmi, 25.0) ∧ dbp:has_PersonalDietary (?p, ?pd) ∧ dbp:has_Serving(?pd, ?s) ∧ dbf:Milk(?milk) → dbp:vegetable_serving(?s, 3) ∧ dbp:fruit _serving(?s, 5) ∧ dbp:milk_serving(?s, 0) ∧ dbp:AvoidFood(?p, ?milk)

ภาพที่ 7: ตัวอย่างการให้ส่วนอาหารสามกลุ่มแรก

กฎที่ 5 การหาปริมาณคาร์โบไฮเดรตที่เหลือจากส่วนอาหารสามกลุ่มแรก ผัก ผลไม้และนม โดยนำตัวแปรที่เป็น

ส่วนอาหาร (?vs, ?fs, ?ms) จาก ***กฎที่ 4*** มาใช้ในคำนวณกับตัวแปรปริมาณสารอาหารหลักคาร์โบไฮเดรต (?tcq) จาก ***กฎที่ 3*** ด้วยวิธีการคำนวณของหลักการอาหารแลกเปลี่ยน ดังนี้

- carbohydrate in vegetable (c1) = vegetable _serving*5
- carbohydrate in fruit (c2) = fruit _serving*15
- carbohydrate in milk (c3) = milk _serving*12
- total carbohydrate in 3 food group (ac1) = c1+c2+c3
- redue1_carbohydrate_quantity (rcg1) = tcq – ac1

(rcg1= ปริมาณคาร์โบไฮเดรตที่เหลือ)

ใช้ตัวแปรและสูตรข้างต้นสร้างกฎออกมาตาม ภาพที่ 8

dbp:Patient(?p) ∧ dbp:has_PersonalDietary(?p, ?pd) ∧ dbp:has_Serving(?pd, ?s) ∧ dbp:total__fat_quantity(?pd, ?tfg) ∧ dbp:total_protient_quantity(?pd, ?tpg) ∧ dbp:total_carbohydrate_quantity(?pd, ?tcg) ∧ dbp:vegetable_serving(?s, ?vs) ∧ dbp:milk_serving(?s, ?ms) ∧ dbp:fruit_serving(?s, ?fs) ∧ dbp:has_Energy(?pd, ?e) ∧ swrlb:multiply(?ve, ?vs, 25) ∧ swrlb:multiply(?me, ?ms, 90) ∧ swrlb:multiply(?fe, ?fs, 60) ∧ swrlb:multiply(?c1, ?fs, 15) ∧ swrlb:multiply(?c2, ?vs, 5) ∧ swrlb:multiply(?c3, ?ms, 12) ∧ swrlb:multiply(?p1, ?fs, 0) ∧ swrlb:multiply(?p2, ?vs, 2) ∧ swrlb:multiply(?p3, ?ms, 8) ∧ swrlb:multiply(?f1, ?fs, 0) ∧ swrlb:multiply(?f2, ?vs, 0) ∧ swrlb:multiply(?f3, ?ms, 5) ∧ swrlb:add(?ac1, ?c1, ?c2, ?c3) ∧ swrlb:add(?ap1, ?p1, ?p2, ?p3) ∧ swrlb:add(?af1, ?f1, ?f2, ?f3) ∧ swrlb:subtract(?rpg1, ?tpg, ?ap1) ∧ swrlb:subtract(?rcg1, ?tcg, ?ac1) ∧ swrlb:subtract(?rfg1, ?tfg, ?af1) → dbp:fruit_energy(?e, ?fe) ∧ dbp:milk_energy(?e, ?me) ∧ dbp:vegetable_energy(?e, ?ve) ∧ dbp:redue1_carbohydrate_quantity(?pd, ?rcg1) ∧ dbp:redue1_fat_quantity(?pd, ?rfg1) ∧ dbp:redue1_protient_quantity(?pd, ?rpg1)

ภาพที่ 8: ตัวอย่างกฎคำนวณคาร์โบไฮเดรตที่เหลือ

กฎที่ 6 การหาส่วนอาหารของกลุ่มข้าว โดยการนำปริมาณคาร์โบไฮเดรตที่เหลือคำนวณได้จาก ***กฎที่ 5*** นำมาหารด้วย 18 จะได้ส่วนของอาหารกลุ่มข้าว แล้วจะใช้ **swrlb:round** ทำให้ส่วนที่คำนวณออกมาได้เป็นทศนิยมสามารถปัดเศษให้เป็นจำนวนเต็ม เพื่อให้ง่ายต่อการเลือกส่วนอาหารที่จะรับประทานของผู้ป่วย แล้วนำส่วนของข้าวที่เป็นจำนวนเต็มคูณด้วย 80 เพื่อเก็บค่าพลังงานจากอาหารกลุ่มข้าวไว้อีกด้วยโดยสร้างกฎตาม ภาพที่ 9

dbp:Patient(?p) ∧ dbp:has_PersonalDietary(?p, ?pd) ∧ dbp:has_Serving(?pd, ?s) ∧ dbp:has_Energy(?pd, ?e) ∧ dbp:redue1_carbohydrate _quantity(?pd, ?rcg1) ∧ swrlb:divide(?ss1, ?rcg1, 18) **∧ swrlb:round(?ss, ?ss1)** ∧ swrlb:multiply(?se, ?ss, 80) → dbp:starch_serving(?s, ?ss) ∧ dbp:starch_energy(?e, ?se)

ภาพที่ 9: ตัวอย่างกฎคำนวณหาส่วนของอาหารกลุ่มข้าว

กฎที่ 7 การแนะปริมาณเครื่องปรุงรส ยกตัวอย่างการแนะนำโซเดียมให้ผู้ป่วยตามเงื่อนไขใน *ภาพที่* 3 ซึ่งเป็นผู้ป่วยเบาหวานโดยรวมไม่มีโรคแทรกซ้อนและมีความดันปกติ คือ *SBP น้อยกว่า 140 mm/Hg และ DBP น้อยกว่า 90 mm/Hg* สามารถรับประทานโซเดียมได้ 3000 mg/day ตามภาพที่ 10

dbp:Patient(?p) ∧ dbp:has_PhysicalExam(?p, ?phy) ∧ dbp:has_Serving(?pd, ?s) ∧ dbp:has_PersonalDietary(?p, ?pd) ∧ dbp:has_DiabeticType(?p, DiabeticType_General) ∧ dbp:SBP(?phy, ?sbp) ∧ dbp:DBP(?phy, ?dbp) ∧ swrlb:lessThan(?sbp, 140)∧ swrlb:lessThan(?dbp, 90) → dbp:salt_serving(?s, 3000)

ภาพที่10: ตัวอย่างกฎให้ส่วนของโซเดียมต่อวัน

2) กฎในการแนะนำอาหาร

กฎที่ 8 การแนะนำประเภทอาหารกลุ่มข้าวที่ควรรับประทานให้แก่ผู้ป่วยตามเงื่อนไขใน *ภาพที่ 3* คือ*ผู้ป่วย*เบาหวาน BMI มากกว่า 25 และ LDL-C น้อยกว่า 100 mg/dL ให้แนะนำอาหารที่มีค่าดัชนีน้ำตาลต่ำ รวมถึงต้องเป็นอาหารกลุ่มข้าวประเภทคาร์โบไฮเดรตเชิงซ้อน แสดงกฎตามภาพที่ 11

dbp:Patient(?p) ∧ dbp:has_PhysicalExam(?p, ?phy) ∧ dbp:has_BMI(?phy, ?bmi) ∧ swrlb:greaterThanOrEqual(?bmi, 25) ∧ dbp:has_LabExam(?p, ?l) ∧ dbp:LDL-C(?l, ?ldl) ∧swrlb:lessThan(?ldl, 100) ∧ dbf:Starch(?starchy) ∧ dbf:has_Glycemic_Index_Level(?starchy, GIType_Low) ∧ dbf:food_type(?starchy, "complex carbohydrate") → dbp:ReccommendedFood(?p, ?starchy)

ภาพที่ 11: กฎแนะนำอาหารที่ควรรับประทานกลุ่มข้าว

กฎที่ 9 การให้คำแนะนำเสมือนผู้เชี่ยวชาญ โดยคำแนะนำเกี่ยวกับการรับประทานอาหารต่าง ๆ จะเก็บไว้ในคำแนะนำตามเงื่อนไขของผู้ป่วย คือ diabeticType_Suggestion, complication_Suggestion, BMI_Suggestion, TG_Suggestion , LDL_ Suggestion การเขียนใช้ **swrlb:stringConcat** เพื่อรวบรวมคำแนะนำจากเงื่อนไขต่าง ๆ ไว้ในที่เดียวเก็บไว้ใน Patient_Suggestion ตามภาพที่ 12

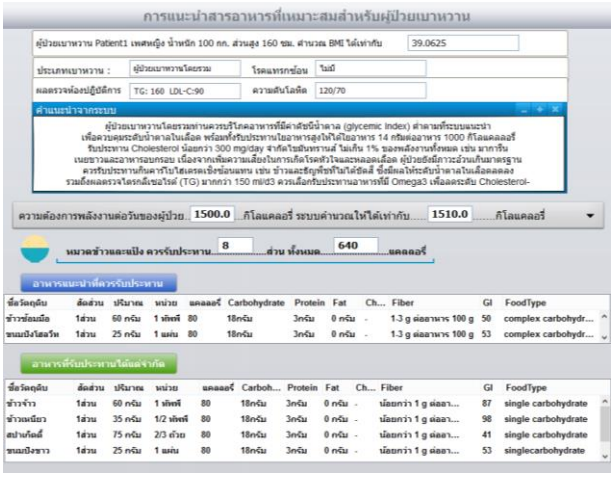
dbp:Patient(?p) ∧ dbp:has_DiabeticType(?p, DiabeticType_General) ∧ dbp:BMI_Suggestion(?p, ?sugbmi) ∧ dbp:TG_Suggestion(?p, ?tgsug) ∧ dbp:LDL_Suggestion(?p, ?ldlsug) ∧ dbp: diabeticType_Suggestion(DiabeticType_General, ?dbsug) **∧ swrlb:stringConcat**(?ptsug, ?dbname, ?dbsug, ?sugbmi, ?tgsug, ?ldlsug) → dbp:Patient_Suggestion(?p, ?ptsug)

ภาพที่ 12: กฎแสดงคำแนะนำเสมือนผู้เชี่ยวชาญ

4. ผลการดำเนินงาน

ผู้วิจัยได้ออกแบบหน้าจอของระบบเพื่อแสดงให้เห็นผลของการแนะนำสารอาหารที่เหมาะสมสำหรับผู้ป่วยเบาหวานตามกรณีตัวอย่างใน ***ภาพที่3*** โดยระบบจะสามารถแสดงข้อมูลพื้นฐานของผู้ป่วย แสดงคำแนะนำเสมือนผู้เชี่ยวชาญเพื่อให้ผู้ป่วยสามารถเข้าใจวิธีการการเลือกรับประทานอาหารเบื้องต้น ต่อมาจะแสดงความต้องการพลังงานต่อวันของผู้ป่วยเทียบกับ

ความต้องการพลังงานต่อวันตามส่วนอาหารที่ระบบคำนวณได้ แสดงส่วนอาหารและพลังงานที่รับประทานได้ตามกลุ่มอาหาร (กลุ่มข้าว) พร้อมทั้งแสดงรายการอาหาร ปริมาณอาหารและสารอาหารต่างๆแบ่งตามประเภทอาหาร (เช่น อาหารรับประทานได้แต่จำกัด) เพื่อให้ผู้ป่วยสามารถเลือกรับประทานอาหารให้ดีต่อสุขภาพของตนเองมากที่สุดตามภาพที่ 13



ภาพที่ 13: ตัวอย่างออกแบบหน้าจอของระบบ

วัดประสิทธิภาพของการออกแบบออนโทโลยีและกฎด้วย F-measure ค่าที่ใช้วัดความสัมพันธ์ของ ค่าความแม่นยำ (Precision) และค่าเรียกคืน (Recall) โดยทดสอบระบบจากตัวอย่างข้อมูลผู้ป่วยเบาหวานที่แตกต่างกันไป 40 ราย แล้วเปรียบเทียบการแนะนำอาหารของผู้เชี่ยวชาญกับระบบเพื่อตรวจสอบความถูกต้องตามสมการ (1), (2) และ (3) ดังนี้

$$Precision = \frac{\text{จำนวนคำตอบของระบบที่ถูกต้องตรงกับผู้เชี่ยวชาญแนะนำ}}{\text{จำนวนคำตอบทั้งหมดของระบบ}} \quad (1)$$

$$Recall = \frac{\text{จำนวนคำตอบของระบบที่ถูกต้องตรงกับผู้เชี่ยวชาญแนะนำ}}{\text{จำนวนคำตอบที่ถูกต้องทั้งที่ถูกแนะนำและไม่ถูกแนะนำ}} \quad (2)$$

$$F\text{-}measure = 2 \times \left\{ \frac{Precision \times Recall}{Precision + Recall} \right\} \quad (3)$$

จากการประเมินผลประสิทธิภาพของระบบพบว่า มีค่าความแม่นยำ *(Precision)* เท่ากับ 0.92 มีค่าเรียกคืน *(Recall)* เท่ากับ 0.95 โดยค่าความสัมพันธ์เฉลี่ยที่ดีที่สุด *(F-measure)* ของทั้งสองค่า เท่ากับ 0.94 ซึ่งผลที่ออกมาแสดงให้เห็นว่าระบบสามารถแนะนำสารอาหารให้กับผู้ป่วยเบาหวานได้เป็นอย่างดี

5. สรุป

ระบบสามารถแนะนำสารอาหารสำหรับผู้ป่วยเบาหวานได้อย่างเหมาะสมตามเงื่อนไขของผู้ป่วยเบาหวาน ซึ่งมีความใกล้เคียงและมีความถูกต้องเสมือนผู้เชี่ยวชาญหรือนักโภชนาการแนะนำ ข้อเสนอแนะระบบยังขาดในส่วนของการพัฒนาเว็บแอพพลิเคชั่นที่ให้ผู้ป่วยสามารถเข้าใช้งานได้ง่ายและเลือกชนิดนิดอาหารได้ด้วยตนเอง เพื่อที่จะได้ดึงอาหารมาคำนวณโดยตรง ซึ่งผู้วิจัยจะดำเนินการพัฒนาต่อไป

เอกสารอ้างอิง

[1] พัทธนันท์ ศรีม่วงธุ์, “อาหารเพื่อสุขภาพและโภชนบำบัด”, กรุงเทพฯ: ห้างหุ้นส่วนจำกัด เอ็ม แอนด์ เอ็ม เลเซอร์พริ้นต์, 2555.

[2] สมาคมโรคเบาหวานแห่งประเทศไทย, “แนวทางเวชปฏิบัติสำหรับโรคเบาหวาน พ.ศ. 2557”, กรุงเทพฯ: ศรีเมืองการพิมพ์, 2557.

[3] กลุ่มโภชนาการประยุกต์ สำนักโภชนาการ กรมอนามัย, “อิ่มอร่อยได้สุขภาพสไตล์เบาหวาน สำหรับผู้ที่เป็นเบาหวา ที่ยังไม่มีภาวะของโรคแทรกซ้อน”, กรุงเทพฯ: ศูนย์สื่อและสิ่งพิมพ์แก้วเจ้าจอม, 2555.

[4] K.K. Breitman, M.A. Casanova and W. Truszkowski, *“*Semantic Web: Concepts, Technologies and Applications*”*, Springer-Verlag:London, 2007.

[5] World Wide Web Consortium: W3C, *“*OWL Web Ontology Language Overview*”*, Retrieved Jan 9, 2016, from http://www.w3.org/TR/owl-features

[6] World Wide Web Consortium: W3C, *“*SWRL: A Semantic Web Rule Language Combining OWL and RuleML.*”*, Retrieved Dec 27, 2015, from http://www.w3.org/Submission/ SWRL

[7] Chang-Shing Lee, Mei-Hui Wang, Huan-Chung Li and Wen-Hui Chen,“Intelligent Ontological Agent for Diabetic Food Recommendation” *IEEE World Congress on Computational Intelligence, 2008.*

[8] B. Hempo, N. Arch-int, S. Arch-in, C. Pattarapongsin, “Personalized Care Recommendation Approach Diabetes patients Using Ontology and SWRL” *Springer-Verlag Berlin Heidelberg 2015. Information Science and Applications, vol. 51, pp. 959-966, 2015.*

[9] Ibrahim M.Ahmed, Marco AlFonse, Abeer M.Mahmoud, and Abdel-Badeem M.Salem, and Wen-Hui Chen. “Knowledge Acquisition for Developing Knowledge-Base of Diabetic Expert System” *ICIT 2015 The 7th Inter national Conference on Information Technology, 2015.*

การย่อความข่าวสุขภาพโดยการประยุกต์ใช้ออนโทโลยี
Ontology-based Health News Summarization

พรรณภา นวลปลอด (Pannapa Nuanplord)[1] และมาลีรัตน์ โสดานิล (Maleerat Sodanil)[2]
ภาควิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ
มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
[1]pannapa.np@gmail.com, [2]maleerat.s@it.kmutnb.ac.th

บทคัดย่อ

ข่าวออนไลน์เป็นสื่อรูปแบบหนึ่งที่สำนักพิมพ์และสำนักข่าวต่างๆ ได้นำเทคโนโลยีมาช่วยในการกระจายข่าวผ่านทางอินเทอร์เน็ต เพื่อให้ผู้อ่านสามารถติดตามข่าวสารประจำวันที่เกิดขึ้นได้ทุกที่ทุกเวลา และในปัจจุบันพบว่าข่าวที่ปรากฏอยู่บนอินเทอร์เน็ตมีเป็นจำนวนมาก ส่งผลให้การอ่านและทำความเข้าใจข่าวสารที่ตรงกับความต้องการนั้นทำได้ยาก จึงได้มีงานวิจัยที่เกี่ยวข้องกับการย่อความเอกสาร เพื่อแก้ไขปัญหาให้ผู้อ่านสามารถทำความเข้าใจเนื้อหาข่าวและคัดกรองข้อมูลที่ตรงกับความต้องการได้สะดวกมากยิ่งขึ้น แต่จากการศึกษางานวิจัยพบว่าผลลัพธ์ที่ได้จากการย่อความยังมีองค์ประกอบของข่าวที่ไม่ครบถ้วน ดังนั้นงานวิจัยนี้จึงได้นำเสนอแนวทางในการย่อความข่าวสุขภาพทางด้านอาหาร โดยการประยุกต์ใช้ออนโทโลยีในการสกัดคำสำคัญของข่าวสุขภาพตามหลัก 5W1H เพื่อให้การย่อความข่าวสุขภาพได้ใจความที่ครบถ้วนสมบูรณ์ตามหลักการรายงานข่าวและการสรุปสาระสำคัญของข่าว ซึ่งผลการทดลองพบว่า ผลลัพธ์ที่ได้จากการย่อความด้วยวิธีการประยุกต์ใช้ออนโทโลยี สามารถย่อความข่าวได้ดี

คำสำคัญ: การย่อความ ออนโทโลยี ข่าวสุขภาพ

Abstract

Publishers and News agencies use internet technology for publishing online news. The reader can read daily news during the whole day. In the present, there is a lot of online news on the internet, it causes trouble reading and understanding. Therefore, the researches and the concepts related to Text Summarization were studied to solve this problem. However, the results of Text Summarization lacked news elements. For this reason, this paper present Ontology-based Health News Summarization with 5W1H elements. The ontology was designed for keyword extraction. The experiment showed that results of Text Summarization based on 5W1H elements work well.

Keyword: Text Summarization, Ontology, Health News.

1. บทนำ

ในปัจจุบันเทคโนโลยีได้เข้ามามีบทบาทต่อชีวิตประจำวันของมนุษย์ โดยการนำเทคโนโลยีมาช่วยอำนวยความสะดวกในการดำเนินชีวิตเพื่อให้เกิดความสะดวกสบาย รวมถึงช่วยในการเผยแพร่ข่าวสารผ่านทางเว็บไซต์หรือสื่อต่างๆ ส่งผลให้การกระจายข้อมูลและข่าวสารสามารถทำได้อย่างรวดเร็ว ได้รับข่าวสารที่ทันสมัย ตรงกับความต้องการ สามารถประหยัดเวลาและค่าใช้จ่ายได้ แต่เนื่องจากปริมาณของข่าวที่เผยแพร่อยู่บนอินเทอร์เน็ตมีเป็นจำนวนมาก การอ่านและทำความเข้าใจข่าวทั้งหมดที่ต้องการจึงทำได้ยากและใช้ระยะเวลานาน ประกอบกับในปัจจุบันคนไทยอ่านหนังสือน้อยลง แต่ติดตามข่าวออนไลน์เพิ่มมากขึ้น การย่อความจึงเป็นวิธีการหนึ่งที่ช่วยลดระยะเวลาและลดปริมาณของเนื้อหาโดยคงไว้ซึ่งสาระสำคัญของข่าว ซึ่งประกอบด้วย ใคร (Who) อะไร (What) ที่ไหน (Where) เมื่อไหร่ (When) ทำไม (Why) และอย่างไร (How)

หรือ 5W1H เพื่อให้ผู้อ่านสามารถทำความเข้าใจเนื้อหาและสาระสำคัญของข่าวได้สะดวกมากยิ่งขึ้น

ดังนั้นผู้วิจัยจึงมีแนวคิดที่จะนำออนโทโลยีมาใช้ในการสกัดคำสำคัญของข่าวสุขภาพตามหลัก 5W1H เพื่อนำไปประยุกต์ใช้กับการย่อความข่าวสุขภาพทางด้านอาหารให้มีความครบถ้วนตามหลักการรายงานข่าวและการสรุปสาระสำคัญของข่าว [1]

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 ข่าวและการสรุปสาระสำคัญของข่าว

ข่าว คือ การรายงานเหตุการณ์หรือเรื่องราวที่มีความน่าสนใจ มีความสำคัญ และเป็นข้อเท็จจริงที่มีความทันสมัยสำหรับผู้รับสาร โดยเนื้อหาของข่าวมุ่งเน้นรายงานข้อเท็จจริงที่เกิดขึ้นตามหลักการรายงานข่าว 5W1H [1] ซึ่งสามารถช่วยในการพิจารณาข้อมูลที่เป็นสาระสำคัญของข่าวและนำไปใช้ในการสรุปสาระสำคัญ ซึ่งหลักการ 5W1H ประกอบด้วย

- who คือ บุคคลใดที่มีความเกี่ยวข้องกับเหตุการณ์ที่เกิดขึ้น
- what คือ เกิดเหตุการณ์อะไรขึ้นภายในเนื้อหาข่าว
- where คือ สถานที่ที่เกิดเหตุการณ์
- when คือ ระยะเวลาที่เกิดเหตุการณ์ภายในข่าวนั้น
- why คือ เหตุของการเกิดเหตุการณ์ภายในข่าว
- how คือ เหตุการณ์นั้นเกิดขึ้นอย่างไร

2.2 ออนโทโลยี

ออนโทโลยี เป็นโครงสร้างของความรู้ของสิ่งที่สนใจในเรื่องใดเรื่องหนึ่งในเชิงของแนวความคิด [2], [3] ที่สามารถแสดงความสัมพันธ์และคุณสมบัติของสิ่งต่างๆ โดยใช้โหนดในการแสดงแนวความคิด และใช้เส้นเชื่อมเพื่อแสดงความสัมพันธ์ระหว่างแนวความคิด ซึ่งรูปแบบทั่วไปของออนโทโลยีจะมีลักษณะเป็นลำดับชั้น [4] ออนโทโลยีถูกนำไปประยุกต์ใช้ในงานด้านต่างๆ เช่น การอ้างอิงหรือใช้งานซ้ำ การจัดหมวดหมู่คำศัพท์ การค้นหาข้อมูลโดยพิจารณาความหมายที่เกี่ยวข้องกับคำค้น [5], [6] รวมถึงการนำไปประยุกต์ใช้กับการย่อความในด้านต่างๆ [7], [8], [9] โดยองค์ประกอบของออนโทโลยีจะประกอบด้วย แนวความคิด (Concept), สล็อท (Slot), ฟาเซ็ท (Facets), และอินสแตนซ์ (Instances)

แนวทางในการพัฒนาออนโทโลยี สามารถพัฒนาได้ 3 แนวทาง ดังนี้

2.2.1 การพัฒนาแบบบนลงล่าง (Top-Down) เป็นวิธีการพัฒนาที่เริ่มจากการกำหนดแนวความคิดหลักของโดเมนก่อน แล้วจึงทำการกำหนดแนวความคิดที่เฉพาะเจาะจงในภายหลัง

2.2.2 การพัฒนาแบบล่างขึ้นบน (Bottom-Up) เป็นวิธีการพัฒนาที่เริ่มจากการกำหนดคลาสย่อยหรือแนวความคิดที่เฉพาะเจาะจงก่อน แล้วนำแนวความคิดเหล่านั้นมาจัดกลุ่มเป็นแนวความคิดหลักของโดเมน

2.2.3 การพัฒนาแบบผสม (Combination) เป็นวิธีการพัฒนาที่ผสมผสานระหว่างการพัฒนาแบบบนลงล่าง และแบบล่างขึ้นบน โดยจะทำการกำหนดแนวความคิดที่เฉพาะเจาะจงก่อน แล้วจึงกำหนดแนวความคิดหลักของโดเมนบางส่วน จากนั้นเชื่อมโยงแนวความคิดทั้งสองกับแนวคิดระดับกลาง

2.3 กระบวนการพัฒนาออนโทโลยี

กระบวนการพัฒนาออนโทโลยีตามแนวคิดของ Noy และ McGuinness [10] สามารถแบ่งออกเป็น 7 ขั้นตอน ดังนี้

2.3.1 ขั้นตอนการกำหนดขอบเขต (Determine Scope) เป็นขั้นตอนของการกำหนดขอบเขตและวัตถุประสงค์ของการพัฒนา ซึ่งส่งผลต่อการออกแบบ การประเมินผล และการนำออนโทโลยีกลับมาใช้ใหม่

2.3.2 ขั้นตอนการพิจารณาการนำออนโทโลยีกลับมาใช้ใหม่ (Consider Reuse) เป็นขั้นตอนการพิจารณาถึงการนำออนโทโลยีที่มีอยู่กลับมาใช้ใหม่เพื่อลดขั้นตอนในการพัฒนา

2.3.3 ขั้นตอนการระบุรายละเอียดและเงื่อนไขที่เกี่ยวข้อง (Enumerate Terms) เป็นขั้นตอนการระบุรายละเอียดที่เกี่ยวข้องกับเทอมทั้งหมด รวมถึงคุณสมบัติต่างๆ ภายใต้โดเมนที่สนใจ โดยไม่คำนึงถึงการซ้อนทับกันระหว่างแนวความคิดกับคุณสมบัติ

2.3.4 ขั้นตอนการกำหนดคลาส (Define Classes) เป็นขั้นตอนการกำหนดคลาสตามโดเมนที่ได้กำหนดไว้ ซึ่งเป็น

ขั้นตอนที่มีความสำคัญในการออกแบบ โดยคลาสที่ได้กำหนดนั้นมาจากขั้นตอนระบุรายละเอียดและเงื่อนไขที่เกี่ยวข้อง

2.3.5 ขั้นตอนการกำหนดคุณสมบัติของคลาส (Define Properties) เป็นขั้นตอนการกำหนดประเภทให้กับคุณสมบัติของคลาสแบบพื้นฐาน เช่น มีค่าเป็น String หรือ Number และคุณสมบัติแบบซับซ้อน เช่น วัตถุต่างๆ ที่เป็นอินสแตนซ์

2.3.6 ขั้นตอนการกำหนดเงื่อนไขให้กับคุณสมบัติ (Define Constraints) เป็นขั้นตอนในการกำหนดคุณสมบัติให้กับสล็อท ซึ่งสล็อทจะมีฟาเซ็ทที่แตกต่างกัน คือ มีการระบุถึงจำนวนของสมาชิก เพื่ออธิบายถึงจำนวนสูงสุดและต่ำสุดของสล็อทที่อยู่ภายในคลาส

2.3.7 ขั้นตอนการสร้างอินสแตนซ์ของคลาส (Create Instance) เป็นขั้นตอนในการกำหนดอินสแตนซ์ให้กับคลาส หรือการสร้างตัวแทนข้อมูล โดยการสร้างอินสแตนซ์ของคลาสและใส่ค่าลงในสล็อท

3. ขั้นตอนการดำเนินงาน

การวิจัยนี้มีเป้าหมายนำเสนอแนวทางในการย่อความข่าวสุขภาพโดยการประยุกต์ใช้ออนโทโลยี ซึ่งในขั้นตอนการดำเนินงานผู้วิจัยได้ แบ่งขั้นตอนการดำเนินงานออกเป็น 2 ขั้นตอน ดังนี้

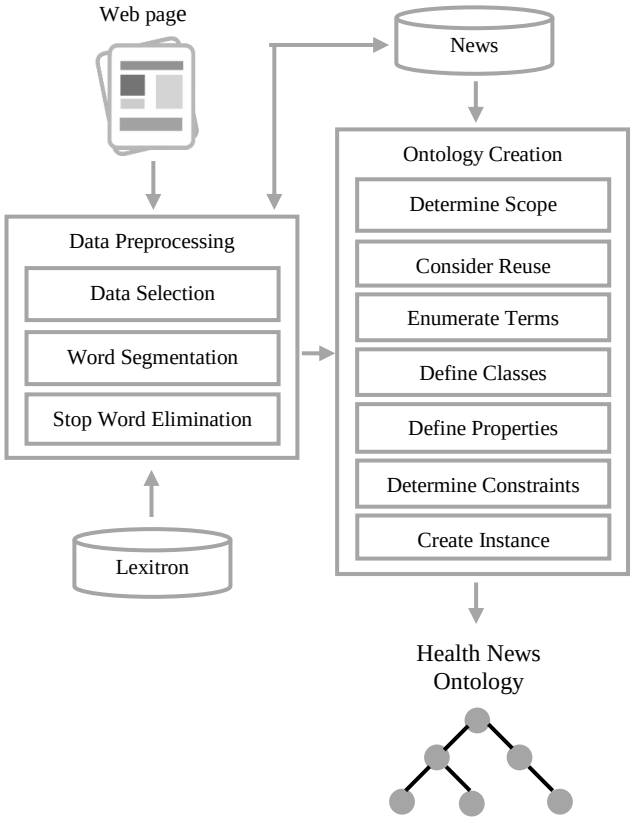
3.1 เก็บรวบรวมข้อมูล

ทำการเก็บรวบรวมข่าวสุขภาพทางด้านอาหารจำนวน 500 ข่าว ระหว่างช่วงปีพุทธศักราช 2552 ถึง 2559 จากเว็บไซต์ที่ทำการเผยแพร่ข่าวสุขภาพ www.health.kapook.com และ www.thaihealth.or.th มาทำการคัดเลือกข้อมูล ตัดคำ และกำจัดคำหยุด จากนั้นจัดเก็บลงในฐานข้อมูล เพื่อศึกษาลักษณะของข่าวสุขภาพ

3.2 ออกแบบและพัฒนาออนโทโลยี

ผู้วิจัยได้ทำการออกแบบและพัฒนาออนโทโลยีโดยใช้โปรแกรม Protégé โดยมีขั้นตอนการออกแบบและพัฒนาออนโทโลยีตามแนวคิดของ Noy และ McGuinness [11] ซึ่งมีขั้นตอนดังต่อไปนี้

3.2.1 ขั้นตอนการกำหนดขอบเขต ออนโทโลยีข่าวสุขภาพทางด้านอาหารมีวัตถุประสงค์ในการพัฒนาขึ้นเพื่อนำไปประยุกต์ใช้กับขั้นตอนวิธีในการย่อความเอกสารและเพื่อช่วยในการสกัดคำสำคัญของข่าวในโดเมนข่าวสุขภาพทางด้านอาหารที่ครอบคลุมถึงบุคคล การกระทำ สถานที่ เวลา สาเหตุ ผลกระทบและการป้องกันที่ปรากฏอยู่ในข่าว ซึ่งการนำออนโทโลยีมาช่วยในการสกัดคำสำคัญจะทำให้ผู้อ่านข่าวสามารถเห็นภาพรวมของข่าวสุขภาพได้อย่างชัดเจนมากยิ่งขึ้นและลดระยะเวลาในการอ่านและทำความเข้าใจข่าวสุขภาพ



ภาพที่ 1: กรอบแนวคิดการวิจัย

3.2.2 ขั้นตอนการพิจารณาการนำออนโทโลยีกลับมาใช้ใหม่ ผู้วิจัยได้ทำการพัฒนาออนโทโลยีข่าวสุขภาพขึ้นมาใหม่ทั้งหมดเพื่อนำไปประยุกต์ใช้กับขั้นตอนวิธีในการย่อความข่าวสุขภาพเนื่องจากออนโทโลยีที่มีการเผยแพร่อยู่ในปัจจุบัน ยังไม่ครอบคลุมในเรื่องสุขภาพ

3.2.3 ขั้นตอนการระบุรายละเอียดและเงื่อนไขที่เกี่ยวข้อง ในขั้นตอนนี้ ได้ทำการพิจารณาคำศัพท์ที่มีความเกี่ยวข้องกับข่าวสุขภาพทางด้านอาหาร ที่ได้จากการศึกษาเนื้อหาข่าวสุขภาพจากเว็บไซต์ www.health.kapook.com และ www.thaihealth.or.th โดยพิจารณาจากคำศัพท์ที่เกี่ยวข้องกับ บุคคล การกระทำ วันที่เกิดเหตุการณ์ สถานที่เกิดเหตุ สาเหตุของเหตุการณ์ ผลกระทบที่เกิดขึ้น และวิธีการป้องกัน เพื่อนำไปใช้สกัดคำสำคัญของข่าวสุขภาพทางด้านอาหาร ซึ่งคำศัพท์ที่เกี่ยวข้องกับขอบเขตดังกล่าว ยกตัวอย่างเช่น เนื้อ หมู ไก่ ปลา ผัก ผลไม้ ส้ม องุ่น เป็นต้น

3.2.4 ขั้นตอนการกำหนดคลาส ในขั้นตอนของการกำหนดคลาสของออนโทโลยีข่าวสุขภาพทางด้านอาหารผู้วิจัยได้ทำการพัฒนาออนโทโลยีแบบบนลงล่าง (Top-Down) เนื่องจากไม่มีออนโทโลยีที่สามารถนำกลับมาใช้งานใหม่ได้ โดยผู้วิจัยได้รวบรวมคำสำคัญในข่าวสุขภาพ ซึ่งครอบคลุมหลักการในการสรุปสาระสำคัญของข่าว 5W1H และทำการแบ่งคลาสออกเป็น 7 คลาสหลัก ได้แก่ คลาสบุคคล (Person) คลาสการกระทำ (Action) คลาสสถานที่ (Place) คลาสเวลา (Time) คลาสผลกระทบ (Effect) คลาสสาเหตุ (Cause) และคลาสการป้องกัน (Protect)

3.2.5 ขั้นตอนการกำหนดคุณสมบัติของคลาส เพื่อช่วยให้การอธิบายข้อมูลมีความละเอียดมากยิ่งขึ้น ผู้วิจัยได้ทำการกำหนดคุณสมบัติของคลาสออกเป็น 2 รูปแบบ คือ Object property ซึ่งเป็นการกำหนดคุณสมบัติของคลาสที่เกี่ยวข้องกับความสัมพันธ์ระหว่างคลาส เช่น hasAction แทนความสัมพันธ์ระหว่างคลาสบุคคลกับคลาสการกระทำ และ Data property เป็นการกำหนดคุณสมบัติของข้อมูลที่อยู่ภายในแต่ละคลาส เช่น คลาสบุคคล จะประกอบด้วยชื่อบุคคล นามสกุล และตำแหน่งงาน มีการกำหนดชนิดข้อมูลเป็น String เป็นต้น

3.2.6 ขั้นตอนการกำหนดเงื่อนไขให้กับคุณสมบัติ หลังจากที่ได้ทำการกำหนดคุณสมบัติของคลาสของออนโทโลยีข่าวสุขภาพทางด้านอาหารแล้ว ในขั้นตอนนี้จะเป็นการกำหนดเงื่อนไขให้กับคุณสมบัติของคลาสต่างๆ ซึ่งในขั้นตอนนี้ ผู้วิจัยไม่ได้ทำการกำหนดเงื่อนไขให้กับคุณสมบัติของคลาส

3.2.7 ขั้นตอนการสร้างอินสแตนซ์ของคลาส เป็นขั้นตอนในการสร้างอินสแตนซ์ของคลาสหรือตัวแทนของข้อมูล โดยทำการกำหนดข้อมูลลงในอินสแตนซ์ของแต่ละคลาส เช่น นายกฤษดา ศิรามพุช ผู้อำนวยการศูนย์เวชศาสตร์นานาชาติ เป็นอินสแตนซ์ของคลาสบุคคล เป็นต้น

4. ผลการดำเนินงาน

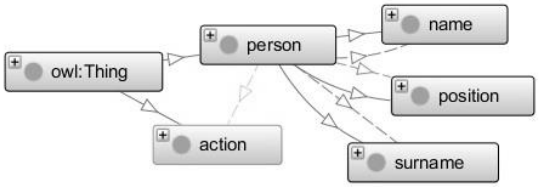
ผลลัพธ์การออกแบบและพัฒนาออนโทโลยีข่าวสุขภาพทางด้านอาหาร และผลการทดลองสามารถแสดงได้ดังนี้

4.1 ผลลัพธ์การออกแบบและพัฒนาออนโทโลยี

ผลลัพธ์การออกแบบและพัฒนาออนโทโลยีข่าวสุขภาพทางด้านอาหาร ประกอบด้วย 7 คลาสหลัก 17 คลาสย่อย และ มีตัวแทนข้อมูลทั้งหมด 1,468 อินสแตนซ์ สามารถแสดงผลลัพธ์ออกเป็นคลาสหลักๆ ได้ดังนี้

4.1.1 คลาสบุคคล (person) และคลาสการกระทำ (action)

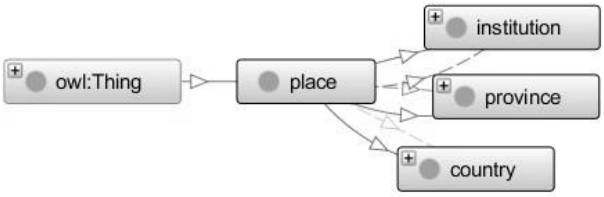
คลาสบุคคลเป็นคลาสที่แทนข้อมูลบุคคลมีความเกี่ยวข้องกับคลาสการกระทำ ดังภาพที่ 2



ภาพที่ 2: คลาสบุคคลและคลาสการกระทำ

4.1.2 คลาสสถานที่ (place)

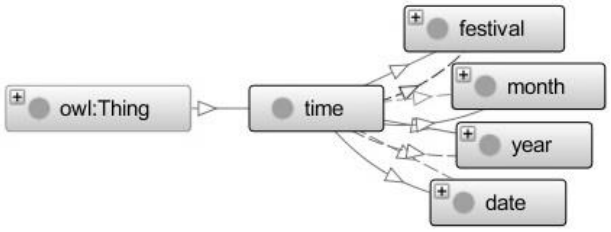
เป็นคลาสแทนข้อมูลสถานที่ แบ่งคลาสย่อยออกเป็น คลาสชื่อหน่วยงาน (institution) คลาสชื่อประเทศ (country) คลาสชื่อจังหวัด (province) ดังภาพที่ 3



ภาพที่ 3: คลาสสถานที่

4.1.3 คลาสเวลา (time)

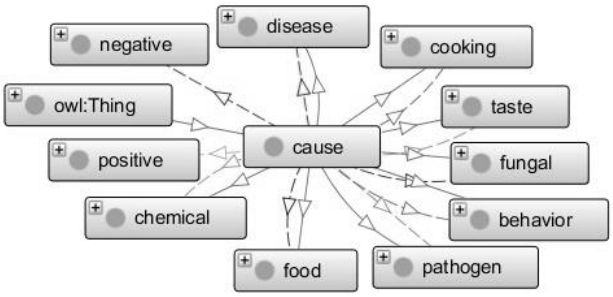
เป็นคลาสแทนข้อมูลเวลา แบ่งคลาสย่อยออกเป็น 4 คลาส คลาสวันที่ (date) คลาสเดือน (month) คลาสปี (year) และคลาสเทศกาล (festival) ดังภาพที่ 4



ภาพที่ 4: คลาสเวลา

4.1.4 คลาสสาเหตุ (cause) และ คลาสผลกระทบ (effect)

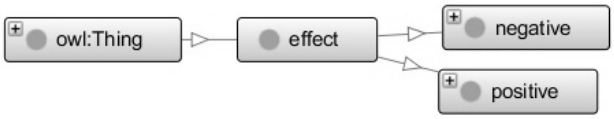
เป็นคลาสแทนข้อมูลสาเหตุแบ่งคลาสย่อยออกเป็น 8 คลาสย่อย ได้แก่ คลาสโรค (disease) คลาสรสชาติ (taste) คลาสอาหาร (food) คลาสเชื้อรา (fungal) คลาสเชื้อโรค (pathogen) คลาสพฤติกรรมการรับประทานอาหาร (behavior) คลาสสารเคมี (chemical) และคลาสการทำอาหาร (cooking) มีความเกี่ยวข้องกับคลาสย่อยของคลาสผลกระทบ คือ คลาสผลกระทบทางบวก (positive) และคลาสผลกระทบทางลบ (negative) ดังภาพที่ 5



ภาพที่ 5: คลาสสาเหตุ และคลาสผลกระทบ

4.1.5 คลาสผลกระทบ (effect)

เป็นคลาสแทนผลกระทบที่เกิดขึ้นภายในข่าวสุขภาพ แบ่งออกเป็น 2 คลาสย่อย ได้แก่ คลาสผลกระทบทางบวก (positive) และคลาสผลกระทบทางลบ (negative) ดังภาพที่ 6



ภาพที่ 6: คลาสสาเหตุ

4.1.6 คลาสการป้องกัน (protect)

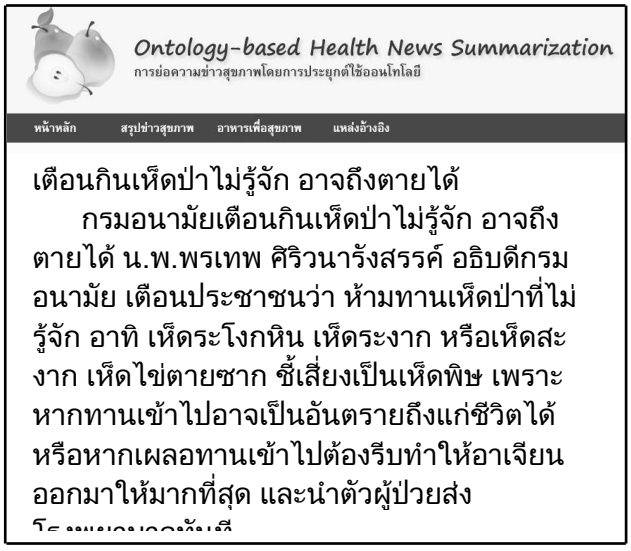
เป็นคลาสแทนวิธีการป้องกันที่ปรากฏอยู่ภายในข่าว มีความเกี่ยวข้องกับคลาสผลกระทบทางลบ ดังภาพที่ 7



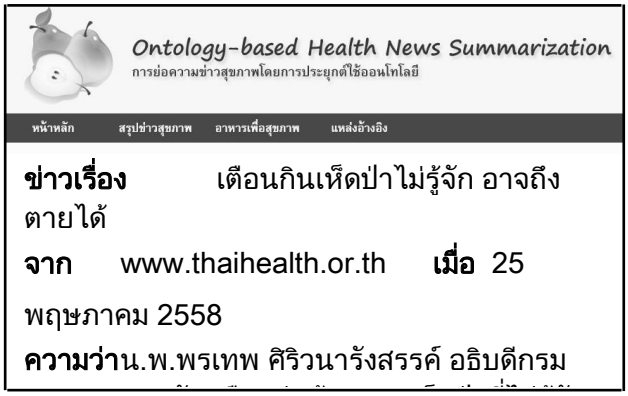
ภาพที่ 7: คลาสการป้องกัน

4.2 ผลการทดลอง

การทดสอบออนโทโลยีข่าวสุขภาพทางด้านอาหารของงานวิจัยนี้ ผู้วิจัยได้ทำการทดลองโดยการย่อความข่าวสุขภาพจำนวน 30 ข่าว และได้ทำการประเมินผลโดยกำหนดเกณฑ์การประเมินผล ออกเป็น 5 ระดับ คือ ดีมาก ดี ปานกลาง พอใช้ และปรับปรุง ซึ่งข่าวต้นฉบับและผลการย่อความข่าวสุขภาพสามารถแสดงได้ดังภาพที่ 8 และภาพที่ 9 ตามลำดับ และผลการประเมินการย่อความข่าว สามารถแสดงตัวอย่างได้ดังตารางที่ 1



ภาพที่ 8: ข่าวต้นฉบับ



ภาพที่ 9: ผลการย่อความข่าว

ตารางที่ 1: ผลการทดลองการย่อความข่าวสุขภาพ

ข่าวที่	ชื่อเรื่อง	ผลการประเมิน
1	เตือนผู้บริโภคระวังปลาดิบ	ดี
2	อากาศร้อนจัด ระวังอาหารเป็นพิษ	ดี
3	ลด ละ ภาชนะโฟม เสี่ยงอันตราย	ดีมาก
4	ลดโซเดียม เปลี่ยนพฤติกรรมการกิน	ดี
5	แนะกินผักผลไม้ ช่วยร่างกายคลายร้อน	ปานกลาง
6	คุมเมนูเด็กช่วงปิดเทอม เลี่ยงภาวะอ้วน	ดี
7	เพิ่มคุณภาพอาหารริมทาง ให้ถูกหลัก	ดีมาก
8	ล้างผักถูกวิธี ลดสารเคมีตกค้าง	ดี
9	อาหารส่งผลต่อความง่วง	ดี
10	เตือนกินเห็ดป่าไม่รู้จัก อาจถึงตายได้	ดี

5. สรุปผลการดำเนินงานและข้อเสนอแนะ

งานวิจัยนี้ได้นำเสนอแนวทางการย่อความข่าวสุขภาพโดยการประยุกต์ใช้ออนโทโลยีในการสกัดคำสำคัญของข่าวตามหลักการรายงานข่าวและการสรุปสาระสำคัญของข่าว 5W1H และได้ทำการทดลองโดยการทดสอบการย่อความข่าวสุขภาพจำนวน 30 ข่าว และให้ผู้ใช้ทำการประเมินผลของการย่อความ ซึ่งผลจากการประเมิน พบว่าผลลัพธ์ที่ได้จากการย่อความที่เกิดจากการประยุกต์ใช้ออนโทโลยีอยู่ในระดับดี สำหรับแนวทางการดำเนินงานวิจัยในอนาคต เพื่อให้การย่อความข่าวสุขภาพมีความถูกต้องมากยิ่งขึ้น ควรมีการนำกฎและไวยากรณ์ทางภาษามาช่วยในการวิเคราะห์และประมวลผล

เอกสารอ้างอิง

[1] จุไรรัตน์ ลักษณะศิริ, *ภาษาไทยเพื่อการสื่อสาร*, นครปฐม : โรงพิมพ์มหาวิทยาลัยศิลปากร, 2558.

[2] ผุสดี บุญรอด, "การศึกษาปัจจัยที่มีผลต่อการย่อความภาษาไทยและการพัฒนาเทคนิคการย่อความภาษาไทยโดยใช้การประมวลผลภาษาธรรมชาติร่วมกับฐานความรู้ออนโทโลยี", วิทยานิพนธ์ปรัชญาดุษฎีบัณฑิต สาขาเทคโนโลยีสารสนเทศ ภาควิชาเทคโนโลยีสารสนเทศ บัณฑิตวิทยาลัย มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2551.

[3] วิชุดา โชติรัตน์. "ระบบวิเคราะห์ข่าวออนไลน์โดยใช้ฐานความรู้ออนโทโลยี และการทำเหมืองข้อความ กรณีศึกษาข่าวในพื้นที่ 5 จังหวัดชายแดนภาคใต้ ประเทศไทย", วิทยานิพนธ์วิทยาศาสตรมหาบัณฑิต สาขาเทคโนโลยีสารสนเทศ ภาควิชาเทคโนโลยีสารสนเทศ บัณฑิตวิทยาลัย มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2554.

[4] กิตย์ศิริ ช่อเจี้ยง และนวลวรรณ สุนทรภิษัช, "ขั้นตอนวิธีสรุปข้อคิดเห็นภาษาไทยโดยใช้ออนโทโลยี," *วารสารเทคโนโลยีสารสนเทศ* ฉบับที่ 7 มกราคม-มิถุนายน พ.ศ. 2554 หน้า 13-18.

[5] สมชาย ปราการเจริญ, "ออนโทโลยี: ทางเลือกของการพัฒนาฐานความรู้ในรูปแบบเชิงเนื้อหา, " *NCCIT05*. 2548 หน้า 92-99.

[6] นิศา สถาพรวจนา, "การใช้การทำเหมืองข้อความและออนโทโลจีในการจำแนกความสนใจของนักท่องเที่ยวต่อการท่องเที่ยวในประเทศไทย", สารนิพนธ์วิทยาศาสตรมหาบัณฑิต สาขาวิทยาการคอมพิวเตอร์ ภาควิชาวิทยาการคอมพิวเตอร์และสารสนเทศ บัณฑิตวิทยาลัย มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2553.

[7] Wu, Chia-Wei. And Liu, Chao-Lin. "Ontology-based Text Summarization for Business News Articles." *Computers and Their Applications*. 2003, pp.389-392.

[8] Hennig, L., Umbrath, W. And Wetzker, R. "An Ontology-Based Approach to Text Summarization." *Web Intelligence and Inte lligent Agent Technology*. 2008, pp.291-294

[9] Plaza, L. Díaz, A. And Gervás, P. "Concept-graph based Biomedical Automatic Summarization using Ontologies." *Graph-Based Algorithms in Natural Language Processing*. 2008, pp.53-56.

[10] Noy, F. And McGuiness, L. "Ontology Development 101 : A Guide to Creating Your First Ontology." *Knowledge Systems Laboratory*. 2001, pp.1-25.

รูปแบบเว็บไซต์เพื่อการจัดการความรู้ที่เหมาะสมกับวิธีการทำงานของบุคลากรฝ่ายพัสดุและจัดหาในประเทศ การไฟฟ้าฝ่ายผลิตแห่งประเทศไทย

A Suitable Knowledge Management Website's Model for Officer in Domestic Supply and Procurement Division, Electricity Generating Authority of Thailand

ภริษา ภัทรธนาวัชร์ (Parisa Pattarathanawat)[1] และณัฐวี อุตกฤษฎ์ (Nattavee Utakrit)[2]

[1,2] ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

[1]s5807021857237@email.kmutnb.ac.th, [2]nattavee.u@it.kmutnb.ac.th

บทคัดย่อ

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อศึกษารูปแบบเว็บไซต์เพื่อการจัดการความรู้ที่เหมาะสมสำหรับหน่วยงานรัฐวิสาหกิจ โดยใช้ฝ่ายพัสดุและจัดหาในประเทศ การไฟฟ้าฝ่ายผลิตแห่งประเทศไทยเป็นกรณีศึกษา ผู้วิจัยได้ทำการศึกษา สังเคราะห์ข้อมูลจากหลักการและทฤษฎี ตลอดจนงานวิจัยที่เกี่ยวข้อง รวมถึงการใช้แบบสอบถามและการสัมภาษณ์จากกลุ่มตัวอย่างเพื่อพัฒนารูปแบบเว็บไซต์เพื่อการจัดการความรู้ที่เหมาะสมขึ้น ผลการตรวจสอบความเหมาะสมโดยผู้เชี่ยวชาญด้านการจัดการความรู้ของฝ่ายฯ ต่างเห็นด้วยกับรูปแบบดังกล่าวนี้ ค่าเฉลี่ยเท่ากับ 4.39 และส่วนเบี่ยงเบนมาตรฐาน 0.24 จากนั้นจึงได้พัฒนาเป็นเว็บไซต์เพื่อการจัดการความรู้ของฝ่ายขึ้นภายใต้รูปแบบดังกล่าว ทดลองใช้งานจริงและประเมินความพึงพอใจโดยผู้ใช้ที่เป็นบุคลากรของฝ่าย ค่าเฉลี่ยรวมทุกด้านเท่ากับ 4.24 และส่วนเบี่ยงเบนมาตรฐาน 0.39

คำสำคัญ: การจัดการองค์ความรู้ เว็บไซต์เพื่อการจัดการความรู้ รูปแบบเว็บไซต์เพื่อการจัดการความรู้

Abstract

This research aims to find a suitable knowledge management website's model for state enterprise. The domestic supply and procurement division of the Electricity Generating Authority of Thailand was used for testing in this case study. The data to develop the model was collected from relevant literature and gathered through interview with staff users. The model was checked by the knowledge management experts who approved. The evaluation results from experts showed the mean of 4.39 with standard deviation at 0.24. Then, development of the knowledge management website based on this model was undertaken. In conclusion, the website has been used in the domestic supply and procurement division and evaluation results from staff users have shown their satisfactions with this particular system to be at an acceptable level. The satisfaction results from users showed the mean of 4.24 with standard deviation at 0.39.

Keywords: Knowledge Management (KM), Knowledge Management Website, Knowledge Management Website's Model

1. บทนำ

ปัจจุบันโลกได้เข้าสู่ยุคเศรษฐกิจฐานความรู้ (Knowledge-based Economy - KBE) ความรู้คือทรัพยากรที่มีค่าอย่างหนึ่ง งานต่าง ๆ จำเป็นต้องใช้ความรู้มาสร้างผลผลิตให้เกิดมูลค่าเพิ่มมากยิ่งขึ้น [1] การจัดการความรู้จึงเป็นสิ่งที่จำเป็นอย่างยิ่งต่อการพัฒนาองค์กรให้ขับเคลื่อนไปตามเป้าหมายที่กำหนดไว้อย่างมีประสิทธิภาพ ปัจจุบันหน่วยงานต่าง ๆ เริ่มเล็งเห็นถึงความสำคัญในเรื่องดังกล่าว จึงได้มีการจัดทำโครงการเพื่อบริหารจัดการความรู้ภายในหน่วยงานของตนเอง ทั้งรูปแบบการรวบรวมเอกสาร การจัดกิจกรรม หรือแม้กระทั่งการจัดทำระบบจัดการองค์ความรู้ผ่านระบบออนไลน์ในรูปแบบของเว็บไซต์ เป็นต้น

การไฟฟ้าฝ่ายผลิตแห่งประเทศไทย (กฟผ.)ได้ดำเนินการวางระบบเทคโนโลยีสารสนเทศเพื่อการจัดการความรู้ในรูปแบบเว็บไซต์ (KM Portal & Corner) [2] หากแต่มีสถิติการแบ่งปันความรู้ลดลง จากการวิเคราะห์ถึงสาเหตุ พบว่าสาเหตุที่สำคัญอย่างหนึ่งคือเว็บไซต์ถูกออกแบบโดยที่ไม่คำนึงถึงความต้องการของผู้ใช้งาน ทำให้ผู้ใช้ไม่อยากเข้าใช้งานเว็บไซต์ ส่งผลให้การแบ่งปันความรู้ลดลง ผู้วิจัยได้นำแนวคิดระบบจัดการความรู้ยุคใหม่ หรือ KM 3.0 มาปรับใช้กับงานวิจัย ซึ่งมีลักษณะสำคัญคือ “อยู่ในวิถี” หมายถึง อยู่ในวิธีการทำงาน เนียนอยู่ในเนื้องาน หรือเนียนอยู่ในขั้นตอนการทำงานจนเป็นอัตโนมัติ

ฝ่ายพัสดุและจัดหาในประเทศเป็นหน่วยงานหนึ่งใน กฟผ. ที่มีนโยบายจะจัดทำเว็บไซต์ที่ตอบสนองความต้องการของผู้ใช้งานเพื่อเป็นช่องทางในการแลกเปลี่ยนองค์ความรู้ภายในหน่วยงาน ดังนั้นงานวิจัยนี้ผู้วิจัยจึงได้เลือกหน่วยงานดังกล่าวเป็นกรณีศึกษาเพื่อออกแบบและพัฒนารูปแบบเว็บไซต์เพื่อการจัดการความรู้ที่เหมาะสม ซึ่งผลจากการวิจัยนี้จะสามารถใช้เป็นแนวทางในการต่อยอดพัฒนาไปสู่หน่วยงานอื่นของ กฟผ. หรือองค์กรอื่นที่มีสภาพแวดล้อมที่คล้ายกันได้ในอนาคต

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 องค์ประกอบหลักของการแลกเปลี่ยนเรียนรู้

องค์ประกอบหลักของการแลกเปลี่ยนเรียนรู้ (Knowledge Sharing) มี 3 องค์ประกอบ [3] ได้แก่

คน (People) ถือว่าเป็นองค์ประกอบที่สำคัญที่สุด เพราะเป็นแหล่งศูนย์รวมของความรู้ที่สมควรนำออกมาแบ่งปันเป็นอย่างยิ่ง โดยที่ควรจะเป็นคนที่มีความรู้จากการปฏิบัติจริง และอยากจะมาแบ่งปันและแลกเปลี่ยนความรู้นั้นด้วยความเต็มใจ

สถานที่และบรรยากาศ (Place) เป็นองค์ประกอบที่สำคัญอีกประการหนึ่งที่จะทำให้การแลกเปลี่ยนเรียนรู้มีชีวิตชีวาและน่าสนใจ เพราะสถานที่และบรรยากาศที่ดี (สบาย ๆ ผ่อนคลาย) มีความเหมาะสมกับแต่ละกลุ่มคน จะทำให้คนเหล่านั้นมาเจอกัน พูดคุย ปรึกษา วิเคราะห์ปัญหา แบ่งปัน และแลกเปลี่ยนเรียนรู้ซึ่งกันและกันอย่างสบายใจ

สิ่งอำนวยความสะดวกต่างๆ (Infrastructure) เป็นองค์ประกอบที่สำคัญที่ช่วยให้การแบ่งปันและแลกเปลี่ยนเรียนรู้เกิดได้ง่ายและสะดวกขึ้น เช่น กระดานสำหรับเขียน คอมพิวเตอร์สำหรับการสรุปและจัดเก็บความรู้ รวมถึงการแบ่งปัน (Share) หรือการส่งต่อข้อมูล

ดังนั้นการพัฒนารูปแบบเว็บไซต์เพื่อการจัดการความรู้จึงมีความจำเป็นที่จะต้องพัฒนาให้สอดคล้องกับความต้องการของผู้ใช้ ซึ่งถือว่าเป็นองค์ประกอบที่สำคัญที่สุดของการแลกเปลี่ยนเรียนรู้

2.2 ทฤษฎีระบบจัดการความรู้ยุคใหม่ หรือ KM 3.0 [4]

เป็นการจัดการความรู้ยุคใหม่เพื่อการบรรลุเป้าหมายหลักขององค์กร ที่เน้นใช้ Human KM เปลี่ยนกระบวนทัศน์ ผ่านเทคโนโลยีใหม่ในยุคดิจิตอล สร้างเครื่องมือและระบบที่มีประสิทธิภาพ เพื่อให้โอกาสคนในสังคมทุกภาคส่วนและทุกระดับได้เข้าถึงข้อมูลความรู้ เพื่อสร้างโอกาส ความเท่าเทียม สามารถใช้ในชีวิตประจำวันหรืองานประจำ เพื่อบรรลุเป้าหมายสูงสุด คือ เกิดพฤติกรรมการเรียนรู้สู่การเปลี่ยนแปลง (Transformative Learning) สู่สุขภาวะของคนในสังคมไทย ซึ่งประกอบด้วยลักษณะสำคัญ ดังนี้ 1) อยู่ในวิถี หมายถึง อยู่ในวิถีการทำงาน หรือวิถีชีวิตจนเป็นอัตโนมัติ KM 3.0 ต้องเป็นกระบวนการจัดการความรู้ที่ผู้ใช้สามารถใช้งานได้ดี 2) มีเป้าหมาย หมายถึง ต้องมุ่งเน้นเป้าหมายหลักขององค์กร ระบุความรู้ที่มีความสำคัญอย่างยิ่งยวดไว้ให้ชัดเจน 3) ใช้เครื่องมือไอที ซึ่งระบบไอทีจะเป็นเครื่องมือสำคัญของการที่จะทำให้เกิดกระบวนการจัดการความรู้ที่ผู้ใช้สามารถใช้งานได้ดี 4) มีพลังจัดการเป้าหมาย และมีการจัดการความรู้จากภายนอกนำมาใช้ในการพัฒนางานของตนด้วยวิธีการต่าง ๆ

การศึกษาแนวทางการจัดการความรู้ยุคใหม่ หรือ KM 3.0 เพื่อเป็นแนวทางในการออกแบบขั้นตอนการใช้งานเว็บไซต์เพื่อการจัดการความรู้ที่จะพัฒนาขึ้น ให้มีขั้นตอนที่ทำให้ผู้ใช้พึงพอใจ สามารถเข้าถึงแต่ละเมนูที่ต้องการได้ง่าย ไม่ซับซ้อน

2.3 ทฤษฎีการออกแบบอินเตอร์เฟส

การคำนึงถึงความต้องการของผู้ใช้งาน ด้วยศาสตร์การออกแบบตามประสบการณ์ของผู้ใช้งาน (User Experience Design) เป็นความพยายามอย่างต่อเนื่องที่จะเรียนรู้เกี่ยวกับผู้ใช้งาน สังเกตพฤติกรรมการใช้งาน แล้วนำไปพัฒนาสินค้าและบริการ การประสบความสำเร็จด้าน User Experience Design สิ่งที่สำคัญ คือ ไม่ใช่แค่ให้ใช้งานได้เท่านั้น (Usability) แต่ต้องรวมถึงการใช้ทรัพยากรที่เหมาะสมในการออกแบบ (Resource) ความสามารถในการเรียนรู้ (Learnability) การตอบสนองด้าน

อารมณ์และพฤติกรรมต่อสินค้าและบริการ (Emotional and Behavioral Response) ดังนั้นการวัดความสำเร็จของ User Experience Design จึงไม่สามารถพึ่งพาเพียงแค่การดูค่าตัวเลขยอดการดาวน์โหลด หรือจำนวนคนเข้าชมเว็บไซต์ต่อวัน (Page Views) แต่ต้องใช้วิธีการเฝ้าสังเกตหรือทดสอบผู้ใช้งานให้ใช้งานเว็บไซต์จริง ๆ (User Testing)

ผู้วิจัยยังได้ศึกษาเว็บไซต์เกี่ยวกับการจัดการองค์ความรู้และเครือข่ายสังคมออนไลน์ ที่มีผู้นิยมใช้งานในปัจจุบัน อาทิ GotoKnow.org Pantip.com SoccerSuck.com เป็นต้น ซึ่งผู้วิจัยได้วิเคราะห์จุดเด่นต่าง ๆ และจุดด้อยของแต่ละเว็บไซต์เพื่อนำมาสังเคราะห์เป็นรูปแบบเว็บไซต์เพื่อการจัดการความรู้ในครั้งนี้

2.4 ทฤษฎีการเก็บรวบรวมข้อมูลในการวิจัย

การเก็บรวบรวมข้อมูล คือ การที่ผู้วิจัยพยายามรวบรวมหลักฐานต่าง ๆ นำมาพิจารณา วิเคราะห์ วิจารณ์ แล้วสรุปผลมาเป็นคำตอบปัญหาการวิจัยว่าเป็นไปตามที่ได้ตั้งสมมติฐานไว้หรือไม่ โดยใช้เครื่องมือการสัมภาษณ์ การสำรวจ การสังเกต การกรอกแบบสอบถาม รายงาน และเอกสารต่าง งานวิจัยครั้งนี้ได้ใช้วิธีเก็บรวบรวมข้อมูลความต้องการของผู้ใช้ด้วยแบบสอบถาม และการสัมภาษณ์ เพื่อที่จะได้พัฒนารูปแบบเว็บไซต์เพื่อการจัดการความรู้ให้สามารถตอบสนองความต้องการในการใช้งานของผู้ใช้ได้อย่างดีที่สุด

2.5 งานวิจัยที่เกี่ยวข้องอื่น ๆ

งานวิจัยของยุทธพงศ์ [5] ได้นำเสนอเรื่อง การออกแบบอินเตอร์เฟสเพื่อการซื้อขายออนไลน์ มีวัตถุประสงค์เพื่อหารูปแบบที่ใช้งานง่าย เหมาะสมกับพฤติกรรมการใช้งานของคนไทย โดยทำการศึกษาเว็บไซต์จำนวน 12 เว็บไซต์ ทำการวิจัยแบบ Focus Group และทำแบบประเมินความพึงพอใจ ผลที่ได้คือ รูปแบบการจัดวางอินเตอร์เฟสในเว็บไซต์ รูปแบบการจัดวางสินค้า รูปแบบของการกระทำบนเมนูและเว็บไซต์ รูปแบบการป้อนข้อมูล รูปแบบการใช้ Social Media และองค์ประกอบในการออกแบบ และเบญนภา [6] ได้นำเสนอเรื่อง การพัฒนายูสเซอร์อินเตอร์เฟสเชิงปฏิสัมพันธ์เพื่อการรับรู้ และการใช้ประโยชน์ของกลุ่มวัยทำงานในช่วงอายุ 45-65 ปี เพื่อศึกษาเกี่ยวกับการออกแบบยูสเซอร์อินเตอร์เฟสที่เหมาะสม สรุปได้ว่าควรออกแบบโดยคำนึงถึงความต้องการในการใช้งาน ข้อจำกัดทางด้านการรับรู้ สมรรถภาพที่เสื่อมถอยลง ซึ่งควรออกแบบให้ง่ายต่อการใช้งาน ง่ายต่อการมองเห็น มีความชัดเจน เพื่อให้คนในวัยนี้มีโอกาสใช้งานเทคโนโลยีที่เหมาะสมกับช่วงวัยและได้รับประโยชน์สูงสุดในการเข้าถึงเทคโนโลยีสมัยใหม่ ที่มีการพัฒนาต่อเนื่องอย่างไม่มีที่สิ้นสุด

3. วิธีการดำเนินการวิจัย

3.1 การศึกษาข้อมูล

3.1.1 ศึกษาข้อมูลรูปแบบการบริหารจัดการความรู้ที่มีอยู่ในปัจจุบัน รวมทั้งที่มีอยู่เดิมของฝ่าย

3.1.2 กำหนดวัตถุประสงค์และขอบเขตของการวิจัย

3.2 การรวบรวมข้อมูล

3.2.1 ศึกษาและรวบรวมข้อมูลจากแหล่งข้อมูลเอกสาร บทความ งานวิจัยที่เกี่ยวข้อง แหล่งข้อมูลเว็บไซต์ และหนังสือดังที่กล่าวมาแล้วในส่วนของทฤษฎีและงานวิจัยที่เกี่ยวข้อง

3.2.2 จัดทำแบบสอบถามความต้องการใช้งานเว็บไซต์เพื่อการจัดการความรู้ ซึ่งแต่ละข้อคำถามได้มาจากการวิเคราะห์และสังเคราะห์ทฤษฎีและงานวิจัยที่เกี่ยวข้อง นำแบบสอบถามไปให้ผู้ทรงคุณวุฒิเพื่อประเมินความสอดคล้องระหว่างข้อคำถามกับวัตถุประสงค์ (IOC) จากนั้นจึงนำแบบสอบถามที่ได้ปรับปรุงแล้วจำนวน 30 ชุดสำรวจความคิดเห็นและสัมภาษณ์เพื่อเก็บข้อมูลจากกลุ่มตัวอย่างที่เป็นบุคลากรภายในฝ่าย

3.3 การสังเคราะห์ข้อมูลและพัฒนารูปแบบเว็บไซต์

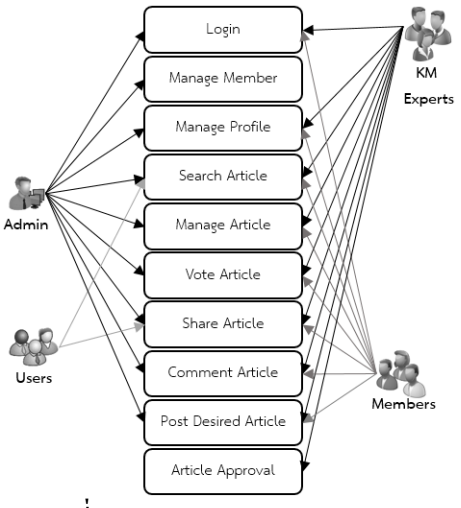
3.3.1 สังเคราะห์ข้อมูลทั้งหมดที่ได้ทำการศึกษาจากแบบสอบถาม และการสัมภาษณ์ถึงความต้องการของผู้ใช้ แล้วจึงพัฒนารูปแบบเว็บไซต์เพื่อการจัดการความรู้ขึ้น

3.3.2 ตรวจสอบความเหมาะสมของรูปแบบเว็บไซต์เพื่อการจัดการความรู้ โดยผู้เชี่ยวชาญด้านการจัดการความรู้ด้วยแบบสอบถามและการสัมภาษณ์

3.4 การพัฒนาเว็บไซต์

3.4.1 พัฒนาเว็บไซต์เพื่อการจัดการความรู้ของฝ่ายขึ้นภายใต้รูปแบบที่ได้ออกแบบ

3.4.2 ผู้วิจัยใช้ยูเอ็มแอล (UML) มาใช้ในการวิเคราะห์ระบบ ซึ่งสามารถอธิบายด้วยยูสเคสไดอะแกรมดังภาพที่ 1 เพื่อแสดงความสัมพันธ์ของการโต้ตอบระหว่างกลุ่มผู้ใช้กับเว็บไซต์เพื่อการจัดการความรู้



ภาพที่ 1: Use Case Diagram ของระบบโดยรวม

จากภาพดังกล่าวระบบโดยรวมจะแบ่งผู้ใช้ออกเป็น 4 กลุ่ม ดังนี้ กลุ่มแรก คือ ผู้ดูแลระบบ (Admin) ซึ่งจะสามารถเข้าใช้งานระบบได้ทุกส่วน ยกเว้นระบบตรวจสอบความถูกต้องของแต่ละบทความ กลุ่มที่สอง คือ กลุ่มสมาชิก (Members) โดยผู้ดูแลระบบจะให้สิทธิ์การเป็นสมาชิกเฉพาะพนักงานฝ่ายพัสดุและจัดหาในประเทศเท่านั้น ซึ่งสมาชิกสามารถใช้รหัสประจำตัวพนักงาน และรหัสผ่านเดียวกันกับที่ใช้เข้าอีเมลขององค์กรเพื่อเข้าสู่เว็บไซต์ สมาชิกสามารถเข้าใช้งานในส่วนของระบบจัดการข้อมูลสมาชิก การค้นหาบทความ การจัดการบทความ การโหวตบทความ การแบ่งปันบทความ การแสดงความคิดเห็น และช่องทางสำหรับแจ้งบทความที่อยู่ในความสนใจ กลุ่มที่สาม คือ กลุ่มผู้เชี่ยวชาญ ซึ่งจะสามารถเข้าใช้งานระบบต่าง ๆ ได้เช่นเดียวกันกับกลุ่มสมาชิก หากแต่ผู้เชี่ยวชาญจะสามารถเข้าใช้ระบบตรวจสอบความถูกต้องของแต่ละบทความในหมวดหมู่ที่ตนเองได้รับสิทธิ์ได้ เพื่อเป็นเครื่องหมายยืนยันว่าบทความดังกล่าวมีความถูกต้อง และกลุ่มที่สี่ คือ ผู้ใช้ทั่วไป ซึ่งจะสามารถค้นหาและอ่านบทความบนเว็บไซต์ได้

3.5 การประเมิน

งานวิจัยนี้แบ่งการประเมินออกเป็น 2 ส่วน คือ 1) ประเมินความเหมาะสมของรูปแบบเว็บไซต์เพื่อการจัดการความรู้โดยผู้เชี่ยวชาญด้านการจัดการความรู้จำนวน 3 ท่าน 2) ประเมินความพึงพอใจของกลุ่มตัวอย่างต่อการใช้งานเว็บไซต์เพื่อการจัดการความรู้ที่ได้พัฒนาขึ้น ดังภาพที่ 2 โดยกลุ่มตัวอย่างคือบุคลากรของฝ่ายจำนวน 15 ท่านที่มีประสบการณ์ในการใช้งานเว็บไซต์เพื่อการจัดการความรู้ ผู้วิจัยได้ใช้การเลือกกลุ่มตัวอย่างแบบเจาะจงกำหนดขนาดของกลุ่มตัวอย่างโดยพิจารณาจากขนาดของประชากรเป้าหมาย และเลือกใช้เกณฑ์ค่าเฉลี่ย (Mean) และค่าส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation หรือ S.D.) ในการประเมินความเหมาะสมและการประเมินความพึงพอใจ

4. ผลการดำเนินวิจัย

4.1 รูปแบบเว็บไซต์เพื่อการจัดการความรู้ที่เหมาะสม

ผลการศึกษาและสังเคราะห์ความเหมาะสมของรูปแบบเว็บไซต์เพื่อการจัดการความรู้ ซึ่งผ่านการประเมินความเหมาะสมโดยผู้เชี่ยวชาญ มีรายละเอียดดังภาพที่ 2

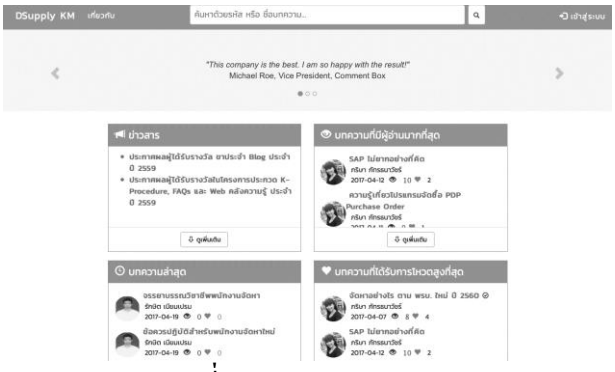


ภาพที่ 2: รูปแบบของเว็บไซต์เพื่อการจัดการความรู้

4.2 ผลการพัฒนาเว็บไซต์เพื่อการจัดการความรู้ของฝ่ายพัสดุและจัดหาในประเทศ

เว็บไซต์เพื่อการจัดการความรู้ที่พัฒนาขึ้นภายใต้รูปแบบที่ได้ออกแบบมีลักษณะหน้าจอดังภาพที่ 3 โดยเป็นการเรียงลำดับเมนูตามที่ผู้ใช้งานให้ความสนใจ และถึงแม้ว่าจากการศึกษาและสังเคราะห์ความเหมาะสมของรูปแบบเว็บไซต์เพื่อการจัดการความรู้ พบว่าผู้ใช้งานส่วนใหญ่ใช้คอมพิวเตอร์เพื่อเข้าถึงเว็บไซต์เพื่อการจัดการความรู้ แต่จากการศึกษาทิศทางใน

การออกแบบเว็บไซต์ปี ค.ศ. 2017 พบว่า เว็บไซต์ควรต้องมีการปรับเปลี่ยนรูปแบบการแสดงผลตามขนาดหน้าจอของอุปกรณ์ต่าง ๆ เพื่อที่จะแสดงผลได้ดีที่สุดในทุก ๆ ขนาดหน้าจอแสดงผลด้วย [7] ผู้วิจัยจึงได้พัฒนาเว็บไซต์ให้สามารถปรับเปลี่ยนรูปแบบการแสดงผลตามขนาดหน้าจอของอุปกรณ์ต่าง ๆ โดยใช้เทคนิค Mobile-First Responsive Site [8] ซึ่งเป็นการสร้างเว็บไซต์ขึ้นมาใหม่โดยออกแบบให้รองรับกับหน้าจอขนาดเล็กก่อน จากนั้นจึงพัฒนาให้เหมาะกับหน้าจอขนาดใหญ่



ภาพที่ 3: หน้าจอหลักของเว็บไซต์

ภาพที่ 4 เป็นการแสดงเนื้อหาบทความ โดยจะแสดงชื่อรูปภาพของผู้นำเสนอ จำนวนผู้อ่าน และจำนวนโหวต ในส่วนของเนื้อหาสามารถแทรกรูปภาพ และวีดิโอได้



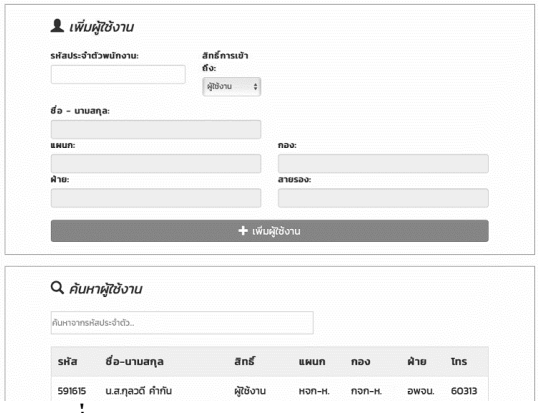
ภาพที่ 4: หน้าจอแสดงเนื้อหาบทความ

ภาพที่ 5 แสดงรายละเอียดข้อมูลของผู้ใช้ บทความที่ผู้ใช้ได้เขียนเรียงตามลำดับวัน เวลาที่นำเสนอบนเว็บไซต์ โดยผู้ใช้สามารถปรับเปลี่ยนรูปภาพประจำตัว แก้ไขบทความ และตั้งค่าการแจ้งเตือนได้ที่หน้าจอนี้



ภาพที่ 5: หน้าจอการจัดการข้อมูลและบทความของสมาชิก

ผู้ดูแลระบบสามารถจัดการข้อมูลต่าง ๆ ในส่วนที่ตนเองได้รับสิทธิ์ได้ ดังภาพที่ 6



ภาพที่ 6: หน้าจอบริหารจัดการข้อมูลสำหรับผู้ดูแลระบบ

4.3 ผลการประเมิน

ผลการประเมินความเหมาะสมของรูปแบบเว็บไซต์โดยผู้เชี่ยวชาญ ดังตารางที่ 1 และผลการประเมินความพึงพอใจทุกด้านต่อเว็บไซต์เพื่อการจัดการความรู้ที่พัฒนาขึ้นโดยผู้ใช้งานจริง สรุปได้ดังตารางที่ 2

ตารางที่ 1: สรุปผลการประเมินความเหมาะสมจากผู้เชี่ยวชาญ

ผู้ประเมิน	$\bar{x}$	S.D.
ผู้เชี่ยวชาญด้านการจัดการความรู้	4.39	0.24

ตารางที่ 2: สรุปผลการประเมินความพึงพอใจจากผู้ใช้งานทั่วไป

หัวข้อการประเมิน	$\bar{x}$	S.D.
รูปแบบของการเชื่อมโยงข้อมูล	4.07	0.56
รูปแบบของการวาง Layout	4.34	0.43
รูปแบบของการจัดการบทความ	4.18	0.59
รูปแบบของ Action	4.32	0.45
รูปแบบการแสดงข้อมูล	4.36	0.42

ตารางที่ 2: สรุปผลการประเมินความพึงพอใจจากผู้ใช้งานทั่วไป (ต่อ)

หัวข้อการประเมิน	$\bar{x}$	S.D.
รูปแบบการป้อนข้อมูล	4.45	0.47
องค์ประกอบในการออกแบบ	4.04	0.68
ความพึงพอใจโดยรวมต่อเว็บไซต์	4.20	0.41
สรุปภาพรวมในทุกด้าน	4.24	0.39

5. สรุปผลและอภิปรายผล

งานวิจัยนี้ได้นำเสนอการพัฒนารูปแบบเว็บไซต์เพื่อการจัดการความรู้ที่เหมาะสมสำหรับหน่วยงานรัฐวิสาหกิจ โดยใช้ฝ่ายพัสดุและจัดหาในประเทศ การไฟฟ้าฝ่ายผลิตแห่งประเทศไทยเป็นกรณีศึกษา ผลการศึกษาครั้งนี้นอกจากจะทำให้ทราบถึงรูปแบบที่จะใช้เป็นแนวทางในการพัฒนาเว็บไซต์เพื่อการจัดการความรู้ของหน่วยงานดังกล่าวแล้ว ผลการประเมินยังแสดงให้เห็นว่าเว็บไซต์เพื่อการจัดการความรู้ที่พัฒนาขึ้นนั้นมีรูปแบบการทำงานบนเว็บไซต์สอดคล้องกับวิธีการทำงานของบุคลากรภายในฝ่าย สามารถตอบสนองความพึงพอใจของกลุ่มตัวอย่างได้เป็นอย่างดี อย่างไรก็ตามเมื่อให้กลุ่มตัวอย่างได้ทดลองใช้งานเว็บไซต์ ต่างมีความเห็นตรงกันว่าเมื่อเลือกบทความที่ต้องการอ่านควรจะแสดงผลที่หน้าจอเดิม ซึ่งประเด็นนี้แตกต่างจากรูปแบบที่ผู้วิจัยได้ออกแบบไว้ คือ เมื่อเลือกบทความควรจะแสดงผลที่หน้าจอใหม่ นอกจากนี้ผู้ใช้ต่างมีความประสงค์ที่จะให้มีการเชื่อมโยงความรู้กับหน่วยงานอื่นด้วย

6. ข้อเสนอแนะ

6.1 ข้อเสนอแนะทั่วไปจากการวิจัย

6.1.1 ภายหลังการเก็บความต้องการโดยใช้แบบสอบถาม ควรพัฒนาเว็บไซต์ต้นแบบ นำไปให้ผู้ใช้งานได้ทดลองใช้ และใช้วิธีการสัมภาษณ์เพื่อเก็บความต้องการของผู้ใช้อีกครั้ง แล้วจึงนำข้อมูลที่ได้มาสังเคราะห์เพื่อออกแบบรูปแบบเว็บไซต์ ทั้งนี้เพื่อให้ได้ข้อมูลที่ตรงกับความต้องการของผู้ใช้มากที่สุด

6.2 ข้อเสนอแนะเพื่อการต่อยอดพัฒนา

6.2.1 ผลจากการวิจัยนี้ทำให้ทราบถึงรูปแบบเว็บไซต์เพื่อการจัดการความรู้ที่เหมาะสมกับวิธีการทำงานของบุคลากรภายในฝ่าย ซึ่งอาจนำมาปรับใช้เป็นแนวทางในการต่อยอดพัฒนาไปสู่หน่วยงานอื่นของ กฟผ. หรือองค์กรอื่นที่มีสภาพแวดล้อมที่คล้ายกันได้ในอนาคต แต่อย่างไรก็ตามควรมีการเชื่อมโยงข้อมูลกับระบบจัดการความรู้อื่น ๆ หรือฐานข้อมูลอื่น ๆ ที่เกี่ยวข้อง เพื่อจัดทำเป็นฐานข้อมูลกลางของหน่วยงานต่อไปด้วย

เอกสารอ้างอิง

[1] ศิริรัตน์ จำปีเรือง และ อมรรัตน์ วัฒนาธร, “ความรู้ที่จำเป็นของคนยุคเศรษฐกิจฐานความรู้” *วารสารศึกษาศาสตร์มหาวิทยาลัยนเรศวร* ปีที่ 12 ฉบับที่ 1 ม.ค.-เม.ย. 2553 หน้า 165-171.

[2] ภาวนา อังคณานุวัฒน์. 2557. “การสัมมนาการแบ่งปันประสบการณ์ดำเนินการตามระบบ SEPA ในองค์กร SEPA Conference 2014: Driving for Excellence”. (ออนไลน์) แหล่งที่มา:http://www.irdp.org/2015/news_files/249/attachment1.pdf. 1 ตุลาคม 2559.

[3] *การแลกเปลี่ยนเรียนรู้ (Knowledge Sharing)*. [ออนไลน์] 2552. [สืบค้นวันที่ 1 ตุลาคม 2559]. จาก https://kminbusiness.wordpress.com/2009/12/26/การแลกเปลี่ยนเรียนรู้-knowledge/

[4] สถาบันส่งเสริมการจัดการความรู้เพื่อสังคม (สคส.). *ขอบฟ้าใหม่ในการจัดการความรู้*. [ออนไลน์] 2559. [สืบค้นวันที่ 22 กรกฎาคม 2559]. จาก http://kmi.or.th/wp-content/uploads/2015/11/new-sky-of-km.pdf

[5] ยุทธพงศ์ ญาณโยธิน. (2555). *การออกแบบอินเตอร์เฟสเพื่อการซื้อขายออนไลน์*. ปริญญาศิลปศาสตรมหาบัณฑิต สาขาวิชาการออกแบบนิเทศศิลป์ ภาควิชาการออกแบบนิเทศศิลป์ บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร.

[6] เบญนภา ชาติเชื้อ. (2556). *การพัฒนายูสเซอร์อินเตอร์เฟสเชิงปฏิสัมพันธ์เพื่อการรับรู้และการใช้ประโยชน์ของกลุ่มวัยทำงานในช่วงอายุ 45-65 ปี*. ปริญญาวิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสื่อสารมวลชน คณะเทคโนโลยีสื่อสารมวลชน มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี.

[7] *สรุป 7 Web Design Trend ที่จะมาแรงในปี 2017*. [ออนไลน์] 2560. [สืบค้นวันที่ 1 กุมภาพันธ์ 2560]. จาก https://www.designil.com/trend-web-design-2017.html

[8] *4 วิธีการทำ Responsive จากเว็บไซต์ระดับโลก*. [ออนไลน์] 2559. [สืบค้นวันที่ 1 กุมภาพันธ์ 2560]. จาก https://www.designil.com/responsive-web-design-4-ways.html

การประยุกต์ใช้เทคนิคการบริหารความเสี่ยงร่วมกับระบบติดตามและควบคุมโครงการด้านเทคโนโลยีสารสนเทศ

An Application of Risk Management Techniques with IT Project Monitoring and Controlling System

พงศกร นาคลมัย (Pongsakorn Naklamai)[1] และณัฐวี อุตกฤษฎ์ (Nattavee Uttakrit)[2]
ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยพระจอมเกล้าพระนครเหนือ
[1]S5807021857113@email.kmutnb.ac.th, [2]nattavee.u@it.kmutnb.ac.th

บทคัดย่อ

ปัจจุบันโครงการด้านเทคโนโลยีสารสนเทศมีจำนวนเพิ่มขึ้นอย่างต่อเนื่อง ดังนั้นการบริหารโครงการเหล่านั้นให้ประสบผลสำเร็จจึงต้องอาศัยกระบวนการบริหารโครงการที่มีประสิทธิภาพ กระบวนการติดตามและควบคุมเป็น 1 ใน 5 ของกลุ่มกระบวนการบริหารโครงการตามหลักการของ PMBOK 5th ที่แนะนำให้ผู้บริหารโครงการควรปฏิบัติ ควบคู่กับความรู้ต่างๆ ที่สำคัญ ปัจจุบันมีการนำระบบสารสนเทศเข้ามาช่วยในกระบวนการติดตามและควบคุมเพื่อช่วยอำนวยความสะดวกให้ผู้บริหารโครงการสามารถดำเนินงานตามกระบวนการดังกล่าวได้ งานวิจัยนี้ได้นำเสนอการประยุกต์เทคนิคการบริหารความเสี่ยงซึ่งเป็นองค์ความรู้หนึ่งที่สำคัญในการบริหารโครงการมาประยุกต์ใช้กับระบบติดตามและควบคุมโครงการด้านเทคโนโลยีสารสนเทศ จากผลการทดลองกับกลุ่มตัวอย่างที่เป็นผู้บริหารโครงการพบว่าสามารถช่วยให้การติดตามและควบคุมการดำเนินงานในโครงการดำเนินงานมีประสิทธิภาพถึงร้อยละ 95 อีกทั้งกลุ่มตัวอย่างยังมีความพึงพอใจในการใช้งานอยู่ในระดับดีอีกด้วย

คำสำคัญ: การบริหารโครงการ การบริหารความเสี่ยง
ระบบติดตามและควบคุมโครงการ

Abstract

Currently, Information technology projects has increased steadily. Successful for project management requires effective project management process. Monitoring and controlling process is one-fifth of the PMBOK 5th project management group that advises project managers to follow. Along with the important knowledge, information systems are now being used to assist in monitoring and controlling process to facilitate project managers to work on the process. This research presents an application of risk management techniques, one of the most important aspects of project management, to IT project monitoring and controlling system. The results showed that monitoring and controlling the operation of project to operate effectively 95%, and the user satisfaction of the system users are at a good level.

Keyword: Project Management, Risk Management,
Project Monitoring and Controlling System.

1. บทนำ

การดำเนินโครงการในปัจจุบันผู้บริหารโครงการจะต้องทำงานในหลายด้านพร้อมกัน จึงอาจทำให้การบริหารจัดการอาจจะไม่มีประสิทธิภาพ ดังนั้นเพื่อหลีกเลี่ยงป้องกันปัญหาที่อาจจะเกิดขึ้น จำเป็นต้องใช้เครื่องมือเข้ามาช่วยในการติดตามควบคุมโครงการ จากการสืบค้นข้อมูลงานวิจัยในปัจจุบันพบว่ามีนักวิจัยหลายท่านได้พัฒนาระบบขึ้นมาเพื่อช่วยในการติดตามความคืบหน้าของโครงการทั่วไป ซึ่งส่วนใหญ่พัฒนาเพื่อใช้งานผ่านเว็บไซต์ [1], [2], [3], [4] แต่ผู้วิจัยได้เล็งเห็นว่าการบริหารความเสี่ยงเป็นอีกองค์ความรู้หนึ่งที่สำคัญตามหลักของ PMBOK 5th ในการบริหารโครงการกระบวนติดตามและควบคุมโครงการ [5] ซึ่งที่ผ่านมายังไม่มีการศึกษาและประยุกต์ใช้กับระบบบริหารโครงการด้านเทคโนโลยีสารสนเทศในปัจจุบัน งานวิจัยนี้จึงได้ดำเนินการค้นคว้าแนวทางการประยุกต์ใช้เทคนิคการบริหารความเสี่ยงร่วมกับระบบติดตามและควบคุมเพื่อทำให้ระบบมีการทำงานได้ครอบคลุมและมีความสมบูรณ์ยิ่งขึ้น ซึ่งช่วยให้โครงการสำเร็จตาม

เป้าหมายได้อย่างมีประสิทธิภาพ ลดความไม่แน่นอนและความสูญเสียในการปฏิบัติงานให้น้อยที่สุด นอกจากระบบจะช่วยในการติดตามควบคุมเพื่อบริหารโครงการแล้ว ยังมีส่วนช่วยในการตัดสินใจการบริหารจัดการความเสี่ยงที่เกิดขึ้นในโครงการ ระบุความเสี่ยงและปัญหาที่อาจเกิดขึ้น รวมทั้งสามารถให้คำแนะนำในการบริหารจัดการความเสี่ยงได้

2. ทฤษฎีที่เกี่ยวข้อง

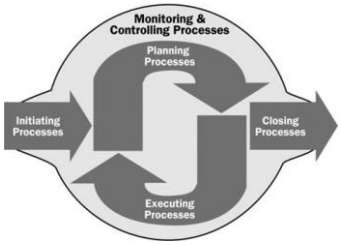
2.1 ทฤษฎีการบริหารโครงการ

การบริหารโครงการเป็นการประยุกต์ใช้องค์ความรู้ ทักษะ เครื่องมือ และเทคนิคต่างๆกับกิจกรรมของโครงการเพื่อตอบสนองต่อความต้องการของโครงการ โดยการประยุกต์และผสมผสานกลุ่มกระบวนการบริหารโครงการดังนี้

กลุ่มกระบวนการเริ่มต้นโครงการ (Initiating Process Group) เป็นการดำเนินการเพื่อกำหนดโครงการใหม่โดยการทำให้โครงการได้รับการอนุมัติเพื่อเริ่มต้น กลุ่มกระบวนการวางแผน (Planning Process Group) เป็นการกำหนดขอบเขตของโครงการ กลั่นกรองวัตถุประสงค์และระบุแนวทางในการปฏิบัติงาน กลุ่มกระบวนการในการดำเนินงาน (Executing Process Group) เป็นการทำให้งานในโครงการเสร็จสมบูรณ์ตามที่กำหนดไว้ตามแผนการบริหาร กลุ่มกระบวนการติดตามและควบคุม (Monitoring and Controlling Process Group) เป็นกระบวนการติดตาม ประเมิน กำหนดวิธีการตรวจสอบความก้าวหน้า และผลการดำเนินงาน กลุ่มกระบวนการปิดโครงการ (Closing Process Group) กระบวนการนี้เป็นการดำเนินการเพื่อทำให้กิจกรรมของทุกกลุ่มกระบวนการเสร็จสมบูรณ์ คณะทำงานบริหารโครงการจะต้องให้ความสนใจความสัมพันธ์ระหว่างกลุ่มกระบวนการข้างต้น ในลักษณะที่ว่าถ้ามีการเปลี่ยนแปลงของปัจจัยหนึ่งจะมีผลทำให้ปัจจัยอื่นเปลี่ยนตามไปได้อยู่เสมอ [5], [6]

2.2 ทฤษฎีการติดตามและควบคุมโครงการ

จากการค้นคว้าข้อมูลผู้วิจัยพบว่า กระบวนการติดตามและควบคุมโครงการเป็นส่วนสำคัญที่จะทำให้โครงการสำเร็จ โดยกระบวนการนี้ประกอบด้วย การติดตามและการควบคุมความก้าวหน้า และประสิทธิภาพการทำงานของโครงการโดยเกี่ยวข้องกับการจัดการการเปลี่ยนแปลง การให้คำแนะนำ มาตรการการแก้ไข หรือการป้องกันปัญหาที่อาจจะเกิดขึ้น การติดตามกิจกรรมของโครงการที่ดำเนินงาน โดยเปรียบเทียบกับแผนการบริหารโครงการและบรรทัดฐานสำหรับการวัดผล เพื่อนำผลการดำเนินงานมาปรับปรุงเพื่อพัฒนาประสิทธิภาพ รวมถึงควบคุมปัจจัยที่อาจหลุดรอดจากกระบวนการควบคุมเปลี่ยนแปลง หรือระบบการบริหารจัดการด้านการทำงานโครงการเพื่อให้แน่ใจว่าการเปลี่ยนแปลงที่ได้รับการอนุมัติเท่านั้นจะสามารถถูกดำเนินการได้ โดยกระบวนการติดตามและควบคุมโครงการนี้จะอยู่ในกระบวนการตั้งแต่เริ่มต้นโครงการจนกระทั่งส่งมอบโครงการ [5] แสดงดังภาพที่ 1



ภาพที่ 1: กระบวนการติดตามและควบคุมโครงการ [5]

กระบวนการติดตามและควบคุมโครงการมีความเกี่ยวข้องกับการติดตามและควบคุมงานภายในโครงการ การจัดการการเปลี่ยนแปลงในโครงการ การระบุและกำหนดขอบเขตของโครงการที่ชัดเจน การควบคุมตารางเวลาของโครงการ การควบคุมค่าใช้จ่าย ต้นทุนภายในโครงการ การดูแลคุณภาพของโครงการ การควบคุมการสื่อสารภายในโครงการ การความคุมความเสี่ยงภายในโครงการ ควบคุมการจัดซื้อจัดจ้าง รวมถึงความพึงพอใจของผู้มีส่วนได้ส่วนเสียภายในโครงการ [5], [7]

2.3 ทฤษฎีการบริหารจัดการความเสี่ยงของโครงการ

ตามหลักของ PMBOK 5th ซึ่งเป็นแนวทางเกี่ยวกับองค์ความรู้ในการบริหารโครงการ [5] ได้ระบุไว้ว่า การบริหารจัดการความเสี่ยงประกอบไปด้วยกระบวนการวางแผนบริหารความเสี่ยง การระบุความเสี่ยง การวิเคราะห์ความเสี่ยง การวางแผนตอบสนองความเสี่ยง และการควบคุมความเสี่ยงของโครงการ โดยมีวัตถุประสงค์เพื่อเพิ่มความเป็นไปได้และผลกระทบของเหตุการณ์ในเชิงบวก ลดความเป็นไปได้และผลกระทบของเหตุการณ์ในเชิงลบต่อโครงการซึ่ง ประกอบด้วย กระบวนการระบุความเสี่ยง (Identify Risks) เป็นการค้นหาความเสี่ยงที่อาจส่งผลกระทบต่อโครงการ และการบันทึกลักษณะของความเสี่ยงว่าความเสี่ยงดังกว่ามีลักษณะรูปแบบใด กระบวนการประเมินความเสี่ยง (Risk Assessment/Evaluation) เป็นการวิเคราะห์ถึงปัจจัยหรือ

สถานการณ์ต่างๆ ที่เป็นสาเหตุทำให้โครงการเกิดปัญหา โดยประเมินจากโอกาสและผลกระทบที่อาจจะเกิดขึ้น กระบวนการตอบสนองความเสี่ยง (Risk Responses) เป็นการพัฒนาทางเลือกและดำเนินการเพื่อเพิ่มโอกาสและลดอุปสรรคที่ส่งผลกระทบต่อวัตถุประสงค์ของโครงการ โดยสร้างกลยุทธ์เพื่อจัดการ และกระบวนการควบคุมความเสี่ยง (Control Risks) เป็นการนำแผนตอบสนองความเสี่ยงมาใช้ การติดตามความเสี่ยงที่ถูกระบุ การระบุความเสี่ยงใหม่ๆที่เกิดขึ้น และประมาณการประสิทธิภาพของกระบวนการบริหารความเสี่ยงตลอดทั้งโครงการ โดยใช้การสร้างกิจกรรมเพื่อบริหารความเสี่ยง จากการศึกษาข้อมูลของผู้วิจัยพบว่ายังไม่มีระบบที่นำการจัดการความเสี่ยงมาประยุกต์ใช้ จึงเล็งเห็นความสำคัญในส่วนนี้ เพื่อนำมาพัฒนาประสิทธิภาพของระบบติดตามและควบคุมโครงการ [8]

2.4 เทคนิคบูทสแตรป

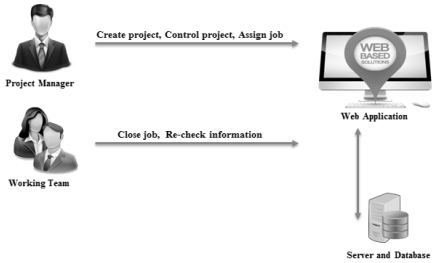
ปัจจุบันเทคโนโลยีมีส่วนสำคัญในการทำงานรวมถึงการบริหารโครงการ เทคนิคบูทสแตรป (Bootstrap) ทำให้ใช้งานแอพพลิเคชั่นทำงานผ่านโทรศัพท์มือถือได้ ง่ายต่อการเข้าถึง [9] ซึ่งผู้วิจัยเล็งเห็นว่า เทคนิคนี้สามารถช่วยสนับสนุนให้ผู้บริหารโครงการสามารถใช้โทรศัพท์มือถือทำงานได้อย่างดี เนื่องจากสามารถติดตามและควบคุมโครงการได้ผ่านทางโทรศัพท์ ซึ่งจะช่วยให้การดำเนินโครงการมีประสิทธิภาพมากยิ่งขึ้นและยังช่วยลดระยะเวลาในการดำเนินโครงการอีกด้วย

3. วิธีดำเนินงาน

วิธีการดำเนินงานการพัฒนาการประยุกต์ใช้เทคนิคการบริหารความเสี่ยงร่วมกับระบบติดตามและควบคุมโครงการด้านเทคโนโลยีสารสนเทศ ผู้วิจัยได้ศึกษา ออกแบบ และพัฒนาสร้างระบบโดยใช้วิธีการดำเนินงานตามรูปแบบการสร้างโมเดลต้นแบบเร็ว ซึ่งมีขั้นตอนดังนี้

3.1 วิเคราะห์ปัญหาและศึกษาแนวทางแก้ไข

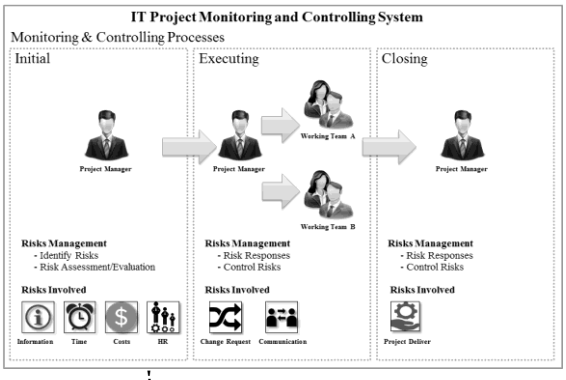
ผู้วิจัยได้แนวคิดจากปัญหาที่ผู้บริหารโครงการจะต้องทำงานในหลายด้าน ซึ่งต้องมีการดูแลจัดการตลอดระยะเวลาในการดำเนินโครงการ ตั้งแต่กระบวนการเริ่มต้นโครงการจนถึงการส่งมอบโครงการ อาจทำให้ประสิทธิภาพในการบริหารโครงการไม่บรรลุผลสำเร็จตามเป้าหมาย และจากการศึกษางานวิจัยในปัจจุบันพบว่า ระบบติดตามและควบคุมโครงการด้านเทคโนโลยีสารสนเทศในปัจจุบัน เป็นระบบที่ทำงานโดยรวบรวมข้อมูลในโครงการและดำเนินการขั้นตอนในโครงการตามขั้นตอน ซึ่งยังไม่มีการบริหารจัดการความเสี่ยงจึงไม่มีประสิทธิภาพในการทำงานที่มากพอ แสดงดังภาพที่ 2



ภาพที่ 2: ระบบติดตามและควบคุมโครงการในปัจจุบัน

3.2 การออกแบบระบบ

จากปัญหาที่กล่าวมาไว้ในข้างต้น ผู้วิจัยได้ทำการออกแบบการประยุกต์ใช้เทคนิคการบริหารความเสี่ยงร่วมกับระบบติดตามและควบคุมโครงการด้านเทคโนโลยีสารสนเทศ โดยใช้เทคนิคการบริหารจัดการโครงการจาก PMBOK 5th [5] เพื่อให้ผู้บริหารโครงการใช้งานดำเนินโครงการอย่างมีประสิทธิภาพและมีการจัดการความเสี่ยงทำให้โครงการดำเนินการตามเป้าหมาย การประยุกต์ใช้เทคนิคการบริหารความเสี่ยงร่วมกับระบบติดตามและควบคุมโครงการด้านเทคโนโลยีสารสนเทศ จึงเป็นสิ่งที่ผู้วิจัย ศึกษาพัฒนาเพื่อนำมาใช้บริหารจัดการโครงการ ช่วยให้โครงการสำเร็จตามเป้าหมายและจัดการความเสี่ยงได้อย่างมีประสิทธิภาพ โดยมีภาพรวมแสดงกระบวนการทำงาน แสดงดังภาพที่ 3



ภาพที่ 3: ภาพรวมแสดงกระบวนการทำงาน

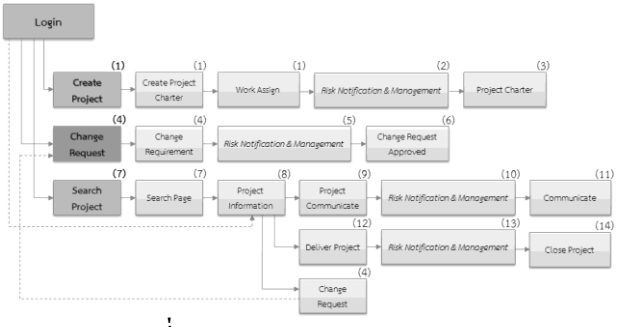
จากภาพที่ 3 แสดงถึงขั้นตอนการติดตามและควบคุมโครงการ ในขั้นเริ่มต้นและวางแผนโครงการจะมีการจัดการความเสี่ยง การระบุความเสี่ยงและการประเมินความเสี่ยงขั้นตอนที่เกี่ยวข้องจะเป็นขั้นตอนการสร้างโครงการโดยมี

ความเสี่ยงที่เกี่ยวข้องคือ ความเสี่ยงจากความถูกต้องของข้อมูล ความเสี่ยงจากต้นทุนโครงการ ความเสี่ยงจากระยะเวลาดำเนินโครงการและความเสี่ยงจากปริมาณทรัพยากรบุคคล ถัดมาเป็นขั้นตอนการดำเนินงานโครงการ ขั้นตอนนี้จะมีการตอบสนองและควบคุมความเสี่ยงที่เกิดขึ้น ความเสี่ยงที่เกี่ยวข้องจะมีความเสี่ยงด้านการสื่อสาร และความเสี่ยงจากการขอเปลี่ยนแปลงโครงการ สุดท้ายขั้นตอนการส่งมอบโครงการ ในขั้นตอนนี้จะมีการตอบสนองและควบคุมความเสี่ยงที่เกิดขึ้นกับโครงการ โดยความเสี่ยงที่เกี่ยวข้องจะเป็นความเสี่ยงจากการส่งมอบโครงการ ซึ่งในแต่ละส่วนระบบจะมีการแจ้งเตือนช่วยในการจัดการความเสี่ยงที่อาจจะเกิดขึ้นในโครงการ มีการให้คำแนะนำสำหรับจัดการความเสี่ยงรวมถึงการประเมินระดับความเสี่ยงที่เกิดขึ้น

3.3 การพัฒนาระบบ

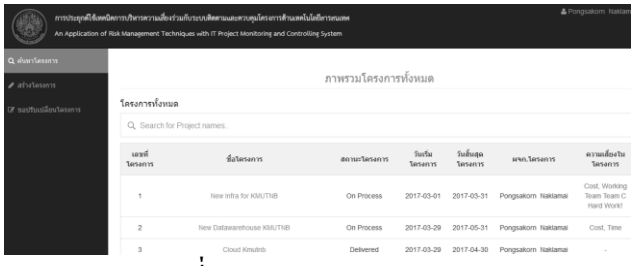
การประยุกต์ใช้เทคนิคการบริหารความเสี่ยงร่วมกับระบบติดตามและควบคุมโครงการด้านเทคโนโลยีสารสนเทศ ผู้วิจัยดำเนินการพัฒนาระบบตรวจสอบและควบคุมโครงการสามารถใช้งานผ่านโทรศัพท์สมาร์ทโฟน เพื่ออำนวยความสะดวกให้กับผู้บริหารโครงการ รวมถึงการประยุกต์ใช้เทคนิคการบริหารความเสี่ยงเข้ากับระบบติดตามและควบคุมโครงการ ซึ่งจะช่วยให้ผู้บริหารโครงการทำการบริหารโครงการได้อย่างมีประสิทธิภาพ โดยมีระบบนี้ช่วยในการดำเนินโครงการให้สำเร็จลุล่วงตามหลักของการบริหารโครงการ

กระบวนการทำงานมีดังนี้ โดยผู้บริหารโครงการเข้าสู่ระบบ (1) สามารถสร้างโครงการ กฎบัตรโครงการเพื่อใช้เป็นข้อมูลดำเนินโครงการ แล้วทำการจัดสรรงานให้ทีมเพื่อดำเนินการ เมื่อสร้างโครงการแล้วระบบจะวิเคราะห์ทำการตรวจสอบ (2) แจ้งเตือนว่ามีความเสี่ยงใดที่อาจจะเกิดขึ้นภายในโครงการ ตามเทคนิคการบริหารจัดการความเสี่ยงจาก PMBOK 5th [5] ความเสี่ยงที่เกี่ยวข้องในขั้นตอนนี้จะเป็นความเสี่ยงจากความถูกต้องของข้อมูล ความเสี่ยงจากต้นทุนโครงการ ความเสี่ยงของระยะเวลาในการดำเนินโครงการ และความเสี่ยงจากทรัพยากรบุคคล จากนั้นผู้บริหารโครงการจะต้องเป็นผู้บริหารจัดการความเสี่ยงที่เกิดขึ้น (3) จึงได้กฎบัตรของโครงการมา (4) ระบบสามารถขอเปลี่ยนแปลงโครงการได้ โดยผู้บริหารโครงการจะต้องแจ้งข้อมูลความต้องการที่ต้องการเปลี่ยนแปลงภายในโครงการนั้นๆ (5) ระบบจะวิเคราะห์ทำการตรวจสอบและแจ้งเตือนว่ามีความเสี่ยงที่เกี่ยวข้องกับการเปลี่ยนแปลงโครงการ (6) แล้วจึงอนุมัติการร้องขอเปลี่ยนแปลงโครงการ (7) ระบบค้นหาของโครงการซึ่งในในการค้นหาตรวจสอบโครงการ (8) จะแสดงรายละเอียดของโครงการต่างๆ เพื่อถึงสามารถจัดการโครงการผ่านระบบได้ (9) ระบบสามารถสื่อสารภายในโครงการ (10) และมีการแจ้งเตือนความเสี่ยงด้านการสื่อสาร (11) เพื่อให้ทีมดำเนินการสื่อสารภายในโครงการได้อย่างราบรื่น (12) เมื่อดำเนินโครงการเสร็จสิ้นจะมีการส่งมอบโครงการ (13) ระบบจะมีการตรวจสอบข้อมูลสารสนเทศที่เกี่ยวข้องกับโครงการหากไม่ครบถ้วนระบบก็จะมีการแจ้งเตือนความเสี่ยงที่อาจจะเกิดขึ้นได้ (14) เมื่อดำเนินการเรียบร้อยจะสามารถส่งมอบโครงการได้โดยสมบูรณ์ แสดงดังภาพที่ 4



ภาพที่ 4: แสดงกระบวนการทำงานของระบบ

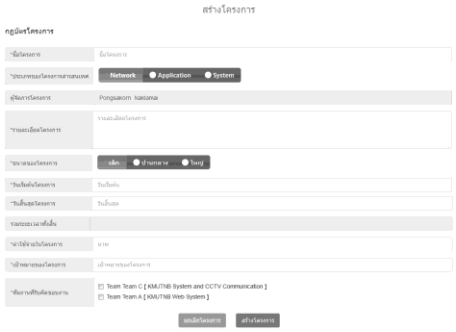
เมื่อผู้บริหารโครงการเข้าสู่ระบบจะมีหน้าแสดงภาพรวมของโครงการ แสดงข้อมูลพื้นฐาน ชื่อโครงการ สถานะโครงการ วันเริ่ม วันสิ้นสุดโครงการ ผู้จัดการโครงการและความเสี่ยงที่มีอยู่ในโครงการ แสดงดังภาพที่ 5



ภาพที่ 5: แสดงหน้าภาพรวมของระบบ

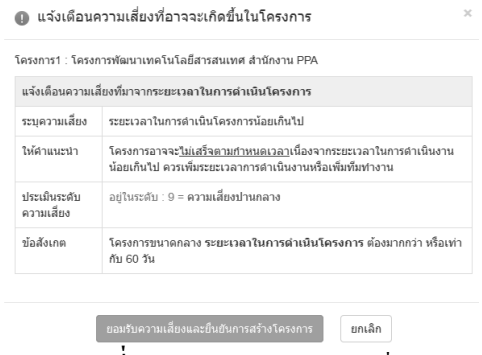
ระบบสามารถสร้างโครงการโดยการสร้างกฎบัตรโครงการเพื่อใช้เพิ่มประสิทธิภาพในการบริหารโครงการ รวมถึงเพื่อ

จัดเก็บข้อมูลสำคัญในโครงการและนำไปประมวลผลเพื่อจัดการความเสี่ยง แสดงดังภาพที่ 6



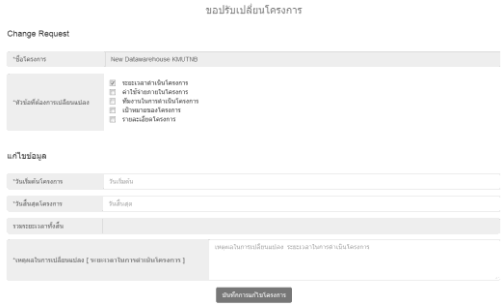
ภาพที่ 6: แสดงหน้าสำหรับสร้างกฎบัตรโครงการ

เมื่อระบบตรวจสอบพบว่ามีความเสี่ยงที่อาจส่งผลกระทบต่อโครงการ ระบบจะมีการแจ้งเตือนให้ผู้บริหารโครงการรับทราบและทำการจัดการความเสี่ยงที่เกิดขึ้น แสดงดังภาพที่ 7



ภาพที่ 7: แสดงหน้าแจ้งเตือนความเสี่ยง

ระบบสามารถร้องขอปรับเปลี่ยนโครงการได้ ซึ่งจะมีการประเมินความเสี่ยงอีกครั้งหลังจากที่มีการขอเปลี่ยนแปลงโครงการและแจ้งให้ทีมงานรับทราบ แสดงดังภาพที่ 8



ภาพที่ 8: แสดงหน้าขอปรับเปลี่ยนโครงการ

การประยุกต์ใช้เทคนิคการบริหารความเสี่ยงร่วมกับระบบติดตามและควบคุมโครงการด้านเทคโนโลยีสารสนเทศ ทำให้ผู้บริหารโครงการสามารถจัดการโครงการได้อย่างมีประสิทธิภาพเพิ่มขึ้น ลดปัญหาที่อาจจะเกิดขึ้นในอนาคต รวมถึงสามารถรับมือกับปัญหาที่อาจจะเกิดได้

4. วิธีการทดสอบ

4.1 การประเมินประสิทธิภาพของระบบ

เป็นการประเมินความถูกต้องของระบบ ที่มีการแจ้งเตือนความเสี่ยงซึ่งสอดคล้องตรงตามทฤษฎีบริหารโครงการและการบริหารจัดการความเสี่ยงของ PMBOK 5th โดยให้ผู้เชี่ยวชาญที่มีประสบการณ์ด้านการบริหารโครงการ นำโครงการที่เคยดำเนินการในอดีตจำนวน 10 โครงการมาทดสอบใช้งานกับระบบ แล้วตรวจสอบจากระบบ ว่าการทำงานของระบบนั้นมีความถูกต้องตามเหตุการณ์ที่เกิดขึ้นจริงในโครงการ

ตารางที่ 1: แนวทางการประเมินความถูกต้องของการทำงานของระบบ

แนวทางประเมินประสิทธิภาพของระบบ	ความถูกต้อง (ร้อยละ)
ความครบถ้วนถูกต้องของข้อมูลแสดงผล	100
ความเหมาะสมของการจัดวางเมนูบนหน้าจอ	90
มีความรวดเร็วในการแสดงผล	100
การแจ้งเตือนความเสี่ยงมีความถูกต้อง	100
การให้คำแนะนำการจัดการความเสี่ยงถูกต้อง	90
การสร้างโครงการ มีความถูกต้องในการสร้าง	100
การขอเปลี่ยนแปลงโครงการมีความถูกต้อง	100
แสดงข้อมูลการสื่อสารมีความถูกต้อง	90
การค้นหาเรียกดูข้อมูลโครงการมีความถูกต้อง	90
ระบบมอบหมายงานทำงานได้ถูกต้อง	90
รวม	95

4.2 การประเมินความพึงพอใจต่อการใช้งาน

จากนั้นได้ทำการสอบถามความพึงพอใจต่อการใช้งานระบบ ซึ่งผู้วิจัยได้กำหนดเกณฑ์การให้คะแนนของแบบประเมินความพึงพอใจต่อการใช้งานโดยเกณฑ์การให้คะแนนแบ่งออกเป็น 5 ระดับ ซึ่งใช้เกณฑ์ตามแบบกำหนดเกณฑ์ตามวิธีของไลเกิร์ต โดยใช้แบบสอบถามที่ผ่านการตรวจสอบค่าความสอดคล้องของข้อคำถามกับวัตถุประสงค์ในการวัด (IOC) กับกลุ่มตัวอย่างซึ่งเป็นผู้เชี่ยวชาญด้านการบริหารโครงการ

ตารางที่ 2: ผลสรุปการประเมินความพึงพอใจต่อการใช้งานระบบ

รายการประเมิน	ระดับความพึงพอใจ		
	$\bar{x}$	S.D.	ผลประเมิน
ด้านการแสดงผล	4.27	0.57	ดี
ด้านการใช้งาน	4.55	0.25	ดีมาก
ด้านประโยชน์	4.47	0.43	ดี
รวม	4.43	0.42	ดี

5. สรุปและอภิปรายผล

งานวิจัยนี้ได้ประยุกต์ใช้เทคนิคการบริหารความเสี่ยงร่วมกับระบบติดตามและควบคุมโครงการด้านเทคโนโลยีสารสนเทศ สามารถช่วยในการตัดสินใจบริหารโครงการและทำให้การบริหารโครงการมีประสิทธิภาพ ซึ่งผลจากการนำระบบไปให้ผู้เชี่ยวชาญที่มีประสบการณ์ด้านการบริหารโครงการทดสอบใช้งานจริงกับข้อมูลโครงการในอดีตทั้งหมด 10 โครงการ พบว่าระบบติดตามและควบคุมทำงานแนะนำในการบริหารจัดการความเสี่ยงได้ถูกต้องถึงร้อยละ 95 อีกทั้งการใช้เทคนิคบูทสแตรปในการออกแบบทำให้ระบบสามารถใช้งานรองรับการทำงานรูปแบบของหน้าจอการแสดงผลทั้งเครื่องคอมพิวเตอร์ส่วนบุคคล หรือโทรศัพท์มือถือสามาร์ทโฟนได้ ซึ่งทำให้ผลการประเมินความพึงพอใจในการใช้งานระบบอยู่ในระดับดี มีโดยค่าเฉลี่ยความพึงพอใจรวมอยู่ที่ 4.43 และส่วนเบี่ยงเบนมาตรฐาน (SD) เท่ากับ 0.42

6. ประโยชน์ที่ได้รับจากการวิจัย

การประยุกต์ใช้เทคนิคการบริหารความเสี่ยงร่วมกับระบบติดตามและควบคุมโครงการสามารถช่วยผู้บริหารโครงการในการควบคุม และเตรียมการปรับปรุงเพื่อป้องกัน แก้ไขปัญหาที่อาจจะเกิดขึ้น ซึ่งจะช่วยในการพัฒนาการดำเนินโครงการให้ดีขึ้นลดโอกาศเกิดปัญหา ลดระยะเวลาในการแก้ไขปัญหา ส่งผลให้สามารถลดต้นทุนของโครงการ ส่งมอบโครงการได้ในระยะเวลาอันรวดเร็ว ทำให้การดำเนินโครงการอย่างมีประสิทธิภาพ นอกจากนี้ยังสามารถนำการประยุกต์ใช้เทคนิคการบริหารความเสี่ยงร่วมกับระบบติดตามและควบคุมโครงการไปใช้ในองค์กร หน่วยงานต่างๆ โดยการนำไปปรับใช้กับโครงการในด้านอื่นๆ รวมถึงศึกษาข้อมูลเพิ่มเติมเพื่อนำไปปรับใช้ให้เหมาะสมกับโครงการ

7. ข้อเสนอแนะจากงานวิจัย

การประยุกต์ใช้เทคนิคการบริหารความเสี่ยงร่วมกับระบบติดตามและควบคุมโครงการด้านเทคโนโลยีสารสนเทศ ได้ทดสอบกับโครงการที่ผ่านมาในอดีต ผลการทดสอบพบว่าทำงานได้อย่างถูกต้อง ช่วยทำให้การบริหารโครงการมีประสิทธิภาพ อย่างไรก็ตาม การบริหารโครงการควรคำนึงถึงเทคนิคการบริหารโครงการด้านอื่นด้วย ดังนั้นในการต่อยอดจากงานวิจัยนี้ อาจศึกษาเทคนิค และองค์ความรู้สำคัญในการบริหารโครงการมาพิจารณาปรับใช้ให้เกิดประโยชน์สูงสุด เช่น การจัดการผู้มีส่วนได้ส่วนเสียภายในโครงการ การจัดการแผนการสื่อสาร เป็นต้น รวมถึงสามารถนำข้อมูลความเสี่ยงที่ได้รับมาจากระบบไปพัฒนาต่อยอดหาแนวทางป้องกันแก้ไขปัญหาที่อาจจะเกิดขึ้นในอนาคตได้

เอกสารอ้างอิง

[1] Mona Taghavi and Ahmed Patel, “Web Base Project Management System for Development of ICT Project Outsourced by Iranian Government,” ICOS, 2011.

[2] Antonie Jetter, “Project management in product development: Toward a framework for targeted flexibility,” PICMET, 2015.

[3] Jeffrey M. Kurcz, “Improving Communication and Project Management through an Adaptive Collaborative Learning System,” ICALT, 2015

[4] ธาดา จันตะคุณ และ ณัฐวี อุตกฤษฎ์, “การพัฒนาระบบสารสนเทศ สำหรับการบริหารโครงการพัฒนาซอฟต์แวร์ผ่านอินเทอร์เน็ต,” คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ , 2556.

[5] PMI, “A Guide to the Project Management Body of Knowledge(5th ed.),” PMI global standards continually and accurately reflect the evolving profession, 2013.

[6] Ramakrishna P. Saloni, “Enterprise Compliance Risk Management,” John Wiley & Sons Inc (US), 2015.

[7] Gregory M. Horine, “Project Management Absolute Beginner’s Guide (3rd Edition),” QUE Publishing, 2012.

[8] Paul Hopkin, “Fundamentals of Risk Management : Understanding, Evaluating and Implementing Effective Risk Management,” Kogan Page Ltd, 2017.

[9] สถิตย์ เรียนพิศ, “Bootstrap and JQuery,” กรุงเทพฯ Developmate, 2557.

การพัฒนาระบบสนับสนุนการตัดสินใจเพื่อช่วยมอบหมายงานให้เจ้าหน้าที่ฝ่ายเทคโนโลยีสารสนเทศ

A Development of Decision Support System to Assign Jobs to the IT Technical Support

อำนาจ กองสุข (Amnat Gongsuk)[1] และ ณัฐวี อุตกฤษฎ์ (Nattavee Utakrit)[2]

ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

[1]S5807021857261@email.kmutnb.ac.th,[2]nattavee.u@it.kmutnb.ac.th

บทคัดย่อ

งานวิจัยนี้เป็นการศึกษาและพัฒนาระบบสนับสนุนการตัดสินใจเพื่อช่วยมอบหมายงานให้เจ้าหน้าที่ฝ่ายเทคโนโลยีสารสนเทศได้อย่างเหมาะสม โดยการประยุกต์ใช้เทคนิคต้นไม้ตัดสินใจ ระบบนี้สามารถทำงานผ่านเว็บบราวเซอร์ (Web Browser) โทรศัพท์มือถือ และแท็บเล็ต วิธีการศึกษาและออกแบบระบบได้ใช้เทคนิคการออกแบบแบบ UML (Unified Modeling Language) ใช้ภาษา PHP (Hypertext Preprocessor) ในการพัฒนาโปรแกรม และใช้ MySQL เพื่อช่วยในการจัดการฐานข้อมูล ประเมินระบบโดยใช้วิธีการประเมินแบบกล่องดำ (Blackbox Testing Technique) ประเมินผลโดยกลุ่มตัวอย่างจำนวน 3 กลุ่มได้แก่ ผู้จัดการฝ่ายเทคโนโลยีสารสนเทศ เจ้าหน้าที่ฝ่ายเทคโนโลยีสารสนเทศ และเจ้าหน้าที่ฝ่ายอื่นๆ ผลการประเมินในภาพรวมอยู่ในระดับดี โดยมีค่าเฉลี่ย และส่วนเบี่ยงเบนมาตราฐานตามลำดับดังนี้ 4.06 (SD 0.54), 4.20 (SD 0.53) และ 4.15 (SD 0.60)

คำสำคัญ: ระบบสนับสนุนการตัดสินใจ การจำแนกประเภทด้วยต้นไม้ตัดสินใจ

Abstract

This research was conducted to study and develop A development of decision support system to assign jobs to the IT technical Support by applying the decision tree technique. This system is able to function through web browser, mobile phones and tablets. The system was studied and designed using Unified Modeling Language (UML) and Hypertext Preprocessor (PHP) to develop the program MySQL was also used to manage the databases. Blackbox Testing Technique was used to evaluate the system by 3 groups: information technology manager, information technology staff. and other staffs. The overall assesment result was at a good level with the means and standara deviations as follows: 4.06 (SD 0.54), 4.20 (SD 0.53), and 4.15 (SD 0.60), respectively.

Keyword: Decision Support System, Decision Tree Classification.

1. บทนำ

ปัจจุบันเทคโนโลยีสารสนเทศเข้ามามีบทบาทสำคัญต่อการดำเนินธุรกิจของหน่วยงานองค์กรทั้งภาครัฐและภาคเอกชน จากการสำรวจการใช้เทคโนโลยีสารสนเทศ จากข้อมูลในปี 2559 ที่ผ่านมาพบว่าสถานประกอบการที่มีคนทำงานตั้งแต่ 16 คนขึ้นไปจำนวน 49,895 แห่ง โดยมีสัดส่วนของสถานประกอบการที่ใช้คอมพิวเตอร์ร้อยละ 87.5 มีความเสี่ยงที่จะเกิดปัญหาและสร้างความเสียหายอย่างรุนแรงต่อระบบสารสนเทศ [1] นอกจากนี้ในอดีตยังเคยเกิดเหตุการณ์ร้ายแรงองค์กรต่างๆ เช่น ในปี 2004 มีไวรัส Mydoom ซึ่งเป็นไวรัสที่ดึงข้อมูลผู้ใช้และเผยแพร่ข้อมูลส่วนบุคคลไปทางอีเมล Outlook สร้างความเสียหายเป็นจำนวน 38,000,000,000 ดอลล่า [2] จากปัญหาที่เกิดขึ้นในหลายองค์กรจึงได้จัดตั้งฝ่ายเทคโนโลยีสารสนเทศเพื่อทำหน้าที่ป้องกัน ปรับปรุง และสนับสนุนทางด้านโครงสร้างพื้นฐานของระบบเทคโนโลยีสารสนเทศขององค์กร

จากการสัมภาษณ์ ผู้จัดการฝ่ายเทคโนโลยีสารสนเทศ บริษัทฟาสโก้ มอเตอร์ (ประเทศไทย) จำกัด [3] มีกระบวนการในการร้องของานเทคโนโลยีสารสนเทศ โดยผู้ที่ร้องของานสารสนเทศ (User) สามารถร้องของานผ่านทางโทรศัพท์ อีเมล และแบบคำร้องส่งมาให้กับผู้จัดการฝ่ายเทคโนโลยีสารสนเทศ (IT Manager) ซึ่ง IT Manager จะรวบรวมและนำรายการที่ได้ร้องขอมาแยกประเภทของปัญหาจัดลำดับความสำคัญของงาน แล้วจึงจ่ายงานให้กับเจ้าหน้าที่ฝ่ายเทคโนโลยีสารสนเทศ (IT Support) แต่เนื่องจากแผนกเทคโนโลยีสารสนเทศต้องดูแล แก้ไขปัญหาด้านระบบสารสนเทศให้กับบุคลากรทั้งองค์กร ซึ่งบางครั้งอาจทำให้เกิดความล่าช้าในการเข้าไปแก้ไขปัญหา IT Manager ต้องทำการบันทึกปัญหาจำนวนของงานที่เข้ามาในแต่ละวัน อีกทั้งยังต้องวิเคราะห์ประเด็นปัญหา (Issue) เพื่อที่จะได้แจกจ่ายงานให้กับ IT Support ได้อย่างเหมาะสม จากปัญหาดังกล่าวผู้วิจัยจึงเห็นว่ามีความจำเป็นที่จะต้องพัฒนาระบบสารสนเทศที่ช่วยให้ผู้จัดการฝ่ายเทคโนโลยีสารสนเทศสามารถรับงานได้เร็ว ช่วยตัดสินใจแจกจ่ายงานได้ แสดงสถานะของการร้องขอและเพื่อให้เกิดความสะดวกและรวดเร็วจึงจำเป็นที่ต้องพัฒนาระบบที่สามารถทำงานผ่านเว็บบราวเซอร์โทรศัพท์มือถือ และแท็บเล็ตได้

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

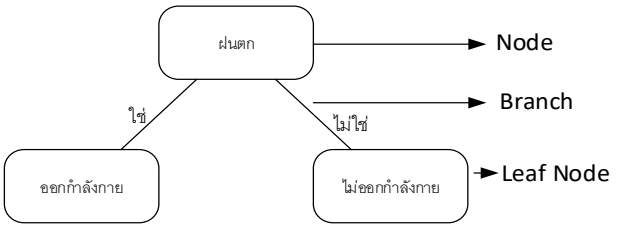
ในการจัดทำงานวิจัยครั้งนี้ผู้จัดทำได้ศึกษาเอกสารและงานวิจัยที่เกี่ยวข้อง เพื่อที่จะนำมาประยุกต์ และพัฒนาใช้กับระบบที่พัฒนาขึ้นโดยมีรายละเอียดเทคนิค และงานวิจัยที่เกี่ยวข้องดังต่อไปนี้

2.1 ระบบการสนับสนุนการตัดสินใจ

เพื่อให้มั่นใจว่าระบบที่ออกแบบสามารถทำงานโดยสนับสนุนการตัดสินใจของผู้จัดการฝ่ายเทคโนโลยีสารสนเทศ ก่อนการออกแบบระบบจึงจำเป็นที่จะต้องศึกษาระบบสนับสนุนการตัดสินใจ (Decision Support System) [4] เพื่อนำมาประยุกต์ใช้กับระบบ ซึ่งระบบการสนับสนุนการตัดสินใจเป็นเทคนิคที่รวบรวมข้อมูลเพื่อนำมาวิเคราะห์และสร้างตัวแบบที่ซับซ้อนภายในซอฟต์แวร์เดียวกัน เพื่อแก้ปัญหาที่มีความยุ่งยากซับซ้อน เป็นการประสานการทำงานระหว่างบุคลากรกับเทคโนโลยี โดยการโต้ตอบกันเพื่อแก้ปัญหาแบบมีโครงสร้าง ไม่มีโครงสร้าง และแบบกึ่งโครงสร้าง ซึ่งอยู่ภายใต้การควบคุมของผู้ใช้ตั้งแต่เริ่มต้นถึงสิ้นสุดขั้นตอน

2.2 แบบจำลองต้นไม้

ระบบที่จัดทำขึ้นเป็นระบบที่สนับสนุนการตัดสินใจ เพื่อให้ระบบสามารถวิเคราะห์ตัดสินใจในการแนะนำผู้จัดการฝ่ายเทคโนโลยีสารสนเทศในการเลือกจัดสรรงานให้กับเจ้าหน้าที่ฝ่ายเทคโนโลยีสารสนเทศ เทคนิคที่สามารถช่วยวิเคราะห์ได้คือเทคนิคต้นไม้ตัดสินใจ (Decision Tree) [5] ซึ่งเป็นเทคนิคในการสร้างต้นไม้ขึ้นโดยต้นไม้ตัดสินใจจะประกอบไปด้วย Noed, Branch, Leaf Node โดยมีการแนบข้อมูลอยู่ด้วย โดยแต่ละกิ่งก้านจะประกอบไปด้วยเงื่อนไข Yes หรือ No เปรียบเทียบจากโหนดบนสุดไปยังโหนดล่างสุดจนถึงโหนดสุดท้ายจะทำให้ทราบถึงข้อมูลที่ทำการพิจารณาหรือวิเคราะห์การทำงานของการจำแนกประเภทด้วยต้นไม่ตัดสินใจ



ภาพที่1: ตัวอย่างแบบจำลองต้นไม้ตัดสินใจ

โหนด (Node) เป็นส่วนของเงื่อนไขในการตัดสินใจ ซึ่งสามารถใช้ในการทำนายพยากรณ์

Branch เป็นส่วนของการเชื่อมต่อระหว่างโหนดแสดงคุณลักษณะที่เป็นไปได้ทั้งหมดของแต่ละโหนด

Leaf Node เป็นส่วนที่แสดงกลุ่มที่ได้จากเงื่อนไขการตัดสินใจ

2.3 การพัฒนาเว็บแอปพลิเคชั่น

ด้วยเทคโนโลยีระบบเครือข่ายในปัจจุบันมีประสิทธิภาพสูง ทำให้เป็นที่นิยมที่จะพัฒนาระบบแอปพลิเคชั่นที่ทำงานผ่านเว็บ เพราะเข้าถึงได้รวดเร็วไม่สิ้นเปลืองทรัพยากรของเครื่องลูกข่าย (Client) ไม่จำเป็นต้องลงทุนสูง ผู้วิจัยจึงได้เลือกใช้วิธีการพัฒนาแอปพลิเคชั่นที่สามารถทำงานบนเว็บ (Web Browser) [6] โดยขั้นตอนการพัฒนาได้ศึกษาขั้นตอนการทำงานของแอปพลิเคชั่นที่ทำงานบนเว็บและภาษาที่ใช้ในการเขียนโปรแกรมที่สามารถทำงานบนเว็บได้ เว็บแอปพลิเคชั่นคือ

โปรแกรมประยุกต์ที่เข้าถึงด้วยการค้นดูผ่านเว็บ เครือข่ายคอมพิวเตอร์ อินเทอร์เน็ต โดยไม่จำเป็นต้องติดตั้งแอปพลิเคชั่นบนเครื่องลูกข่าย โดยการติดต่อระหว่างเครื่องลูกข่าย (Client) และเครื่องแม่ข่าย (Server) ผ่านเว็บโปรโตคอล (http) เครื่องแม่ข่ายทำหน้าที่ให้บริการแก่เครื่องลูกข่ายโดยโปรแกรมที่ทำหน้าที่ให้บริการเว็บ (Web Services) เช่น อาปาเช่เว็บเซิร์ฟเวอร์ ไมโครซอฟท์ไอไอเอส [4]

2.4 การพัฒนาเว็บแอปพลิเคชั่นด้วยบูทสแตรป

ปัญหาสำคัญในการพัฒนาเว็บแอปพลิเคชั่นคือ เนื่องจากในปัจจุบันมีอุปกรณ์ที่สามารถรองรับการทำงานของเว็บเบราว์เซอร์ได้หลากหลาย เช่น คอมพิวเตอร์ แล็บท็อป โทรศัพท์มือถือและอื่นๆ และด้วยขนาดของหน้าจอของอุปกรณ์ที่มีหลายขนาดเช่น โทรศัพท์ iPhone7 มีขนาดหน้าจอ 4.7 นิ้ว แต่ iPhone7 Plus จะมีขนาดหน้าจอ 5.5 นิ้ว ทำให้การพัฒนาเว็บแอปพลิเคชั่นจำเป็นต้องใช้เทคนิคบูทสแตรป (Bootstrap) [7] ในการพัฒนา บูทสแตรปเป็นเครื่องมือที่ช่วยให้พัฒนาเว็บแอปพลิเคชั่นได้อย่างรวดเร็ว และดูสวยงาม สามารถทำให้เว็บแอปพลิเคชั่นทำงานบนอุปกรณ์มือถือ รองรับบริการเว็บเบราว์เซอร์ได้หลากหลาย มีองค์ประกอบให้ใช้งานหลากหลายไม่ว่าจะเป็นเทมเพลต เมนูบาร์ ตาราง ไอคอน หน้าต่าง และอีกหลายๆรายการที่สามารถเลือกใช้งานซึ่งส่วนประกอบสำคัญด้วยกันอยู่ คือ CSS, Fonts, และ JS

2.5 งานวิจัยที่เกี่ยวข้อง

การค้นคว้างานวิจัยที่เกี่ยวข้องมีวัตถุประสงค์เพื่อนำความรู้หลักการเทคนิคต่างๆ มาเสริมการทำงานวิจัยให้มีประสิทธิภาพมากยิ่งขึ้นเช่น

พูลศักดิ์ [8] ได้นำเสนอเกี่ยวกับการพัฒนาระบบสารสนเทศเพื่อการจัดการงานบริการของสายงานเทคโนโลยี และสารสนเทศภายในบริษัท: กรณีศึกษา บริษัทประกันภัยแห่งหนึ่งในกรุงเทพมหานคร โดยปัญหาของงานวิจัยนี้ คือ กระบวนการแจ้งปัญหาสารสนเทศที่ไม่เป็นระเบียบ ไม่สามารถตรวจสอบรายงานการแก้ปัญหาย้อนหลังได้ สาเหตุเกิดจากช่องทางการแจ้งปัญหาที่มีหลายช่องทาง การจัดเก็บข้อมูลที่ไม่เป็นระเบียบ โดยเสนอวิธีการแก้ไขปัญหาโดยวิเคราะห์สาเหตุของปัญหางานเดิมและนำมาพัฒนาเป็นระบบขอใช้บริการสายงานเทคโนโลยีสารสนเทศ ผลที่ได้ สามารถแก้ปัญหาการแจ้งเตือนเจ้าหน้าที่ได้ทันที ลดระยะเวลาการแจ้งข้อใช้บริการ สามารถสืบค้นข้อมูลย้อนหลังได้

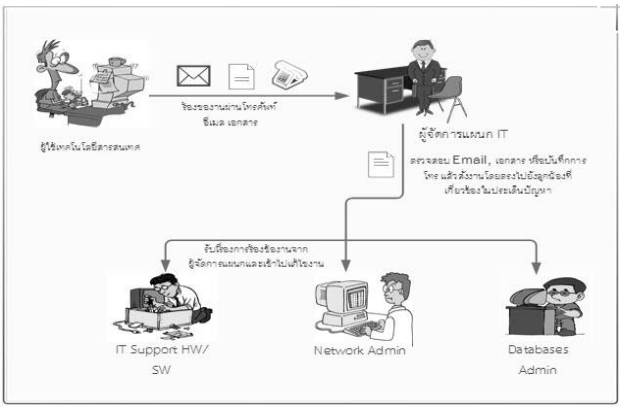
กำธร [9] ได้นำเสนอเกี่ยวกับระบบสนับสนุนการตัดสินใจสำหรับผู้ใช้บริการของการท่องเที่ยวโดยปัญหาของงานวิจัยนี้คือการตัดสินใจเลือกใช้บริการของการท่องเที่ยว จุดหมายที่พักและสถานที่ท่องเที่ยวซึ่งยากต่อการตัดสินใจโดยสาเหตุเกิดจากจำนวนสถานที่ท่องเที่ยว การเลือกจุดหมาย งบประมาณโดยเสนอวิธีแก้ปัญหาสร้างระบบบนเว็บเซอร์วิสที่ช่วยตัดสินใจเลือกจุดหมายปลายทางโดยใช้หลักการจำแนกประเภทด้วยต้นไม้ตัดสินใจ ผลที่ได้ทำให้ผู้ใช้งานสามารถช่วยเลือกจุดหมายปลายทางในสถานที่ท่องเที่ยวได้

3. วิธีการดำเนินการวิจัย

วิธีการดำเนินงานวิจัยประกอบด้วยการดำเนินงานดังต่อไปนี้

3.1 การศึกษาและวิเคราะห์ปัญหา

เพื่อให้ทราบถึงปัญหาและสาเหตุที่ทำให้กระบวนการร้องของงานจากฝ่ายเทคโนโลยีสารสนเทศมีความล่าช้าและไม่ถูกต้อง ผู้วิจัยจึงได้เข้าไปศึกษาและวิเคราะห์เอกสารการดำเนินงานของฝ่ายเทคโนโลยีสารสนเทศทั้งผังงาน โครงสร้างของฝ่าย คู่มือการปฏิบัติงาน และแบบฟอร์มต่างๆที่ใช้ในการปฏิบัติงานทำให้ทราบถึงขั้นตอนการปฏิบัติงาน การสังเกตการณ์ในกระบวนการทำงานเพื่อทราบถึงข้อมูลเชิงลึกในการทำงานการร้องของงานจากทางฝ่ายเทคโนโลยีสารสนเทศจะต้องผ่านผู้จัดการฝ่ายเทคโนโลยีสารสนเทศ ผู้จัดการฝ่ายจะเป็นผู้ตัดสินใจในการแจกจ่ายงานให้กับเจ้าหน้าที่ฝ่ายเทคโนโลยีสารสนเทศโดยการจดบันทึกและตัดสินใจ

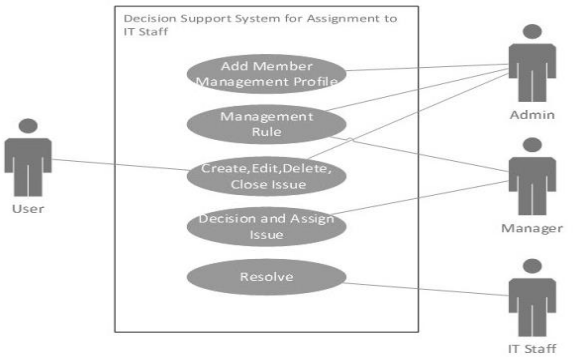


ภาพที่2: แสดงการทำงานของระบบงานในปัจจุบัน

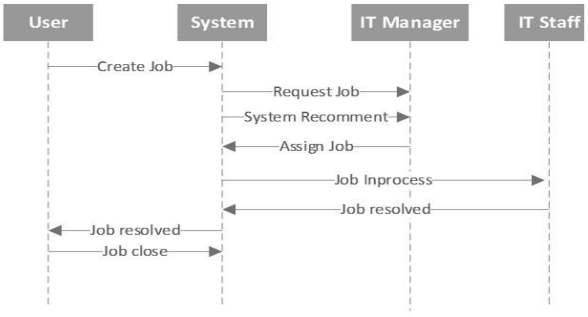
3.2 การออกแบบและพัฒนาระบบ

หลังจากการศึกษาและวิเคราะห์ปัญหาทำให้เห็นถึงกรอบแนวความคิดการสนับสนุนงานทางด้านความช่วยเหลือระบบสารสนเทศแบบเดิม และผู้วิจัยได้ออกแบบกรอบแนวความคิดแบบใหม่เพื่อปรับปรุงกระบวนการสนับสนุนงานทางด้านสารสนเทศให้มีความเหมาะสมโดยในขั้นตอนการออกแบบได้ออกแบบดังนี้

3.2.1 ออกแบบจำลองการทำงานด้วย UML (Unified Modeling Language) [8] เพื่อให้เห็นถึงการทำงานของระบบได้อย่างชัดเจนจึงได้ใช้ UML ไดอะแกรมในการออกแบบและพัฒนาระบบ UML ไดอะแกรมจะแสดงมุมมองเชิงพฤติกรรมการนำเข้าข้อมูลการส่งออกข้อมูลโดยสามารถแสดงแบบจำลองของระบบ ซึ่งจากภาพที่ 3 จะประกอบไปด้วยผู้ที่เกี่ยวข้องในระบบ (Actor) และหน้าที่การทำงาน (Use Case) ซึ่ง Actor ของ User ทำหน้าที่ใช้ Use Case สร้าง แก้ไข และปิดประเด็นการร้องของงาน Actor ของ Admin ทำหน้าที่ใช้ Use Case สร้างผู้ใช้งาน และชนิดการร้องขอปัญหา Actor ของ IT Manager ทำหน้าที่จ่ายงานให้ IT Support และ Actor ของ IT Support ทำหน้าที่ใช้ Use Case เพื่อแก้ไขปัญหางาน

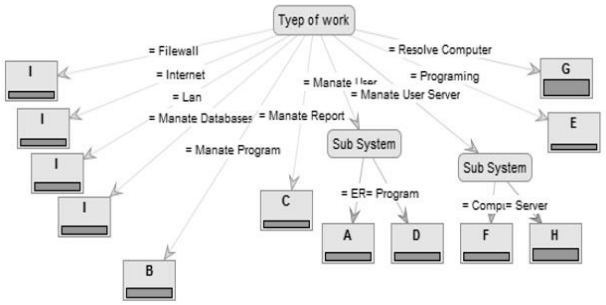


ภาพที่3: Use Case Diagram ของระบบแสดงการทำงาน



ภาพที่4: ขั้นตอนการทำงานของระบบ

3.2.2 ออกแบบต้นไม้ตัดสินใจ (Decision Tree) [5] ในการออกแบบโมเดลต้นไม้ตัดสินใจได้ใช้โปรแกรม RapidMiner Studio 7.4 ในการออกแบบและหาประสิทธิภาพของโมเดลที่ออกแบบโดยใช้ Gini Index และหาความถูกต้องของโมเดลโดยใช้ Nominal Cross-Validation ได้ค่า accuracy ที่ 96.67% ที่ความคลาดเคลื่อน +/- 10 % ซึ่งผลลัพธ์ที่ได้จากการสร้างโมเดลแสดงดังภาพที่ 5



ภาพที่ 5: แสดงภาพต้นไม้ตัดสินใจของระบบ

3.2.3 ออกแบบและสร้างฐานข้อมูลโดยใช้โปรแกรม Visio ในการออกแบบฐานข้อมูลและใช้ Phpmyadmin ในการสร้างฐานข้อมูล

3.2.4 พัฒนาเว็บแอปพลิเคชั่นโดยใช้เทคนิคบูทสแตรปมาช่วยในการพัฒนาเว็บโดยใช้ภาษา PHP, HTML, Bootstrap [9],[5] โดยทำการพัฒนาตามขั้นตอนที่ได้ออกแบบ

3.3 การประเมินประสิทธิภาพของระบบ

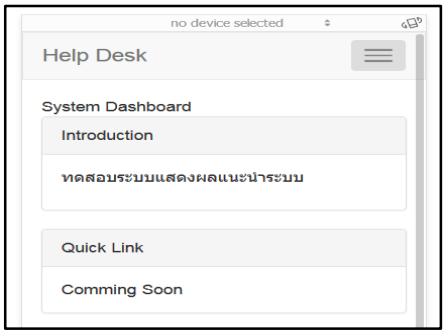
งานวิจัยนี้ผู้วิจัยได้ใช้การประเมินประสิทธิภาพของระบบโดยการทำแบบประเมินประสิทธิภาพ และส่งให้กลุ่มตัวอย่างทำการประเมินจำนวน 3 กลุ่ม แบ่งออกเป็นกลุ่มละ 5 คนได้แก่ ผู้จัดการฝ่ายเทคโนโลยีสารสนเทศ เจ้าหน้าที่ฝ่ายเทคโนโลยีสารสนเทศ และเจ้าหน้าที่ฝ่ายอื่นๆ ซึ่งแบบประเมินจะประกอบไปด้วย 3 ตอน ตอนที่ 1 ข้อมูลส่วนบุคคล ตอนที่ 2 การให้คะแนนต่อประสิทธิภาพของระบบ 5 ระดับ และตอนที่ 3 ความคิดเห็นเพิ่มเติม

4. ผลการดำเนินการวิจัย

4.1 ผลการพัฒนาระบบ

จากการพัฒนาแอปพลิเคชั่น ระบบสนับสนุนการตัดสินใจเพื่อช่วยมอบหมายงานให้เจ้าหน้าที่ฝ่ายเทคโนโลยีสารสนเทศสามารถแยกการทำงานโดยแบ่งเป็นส่วนต่างๆดังนี้

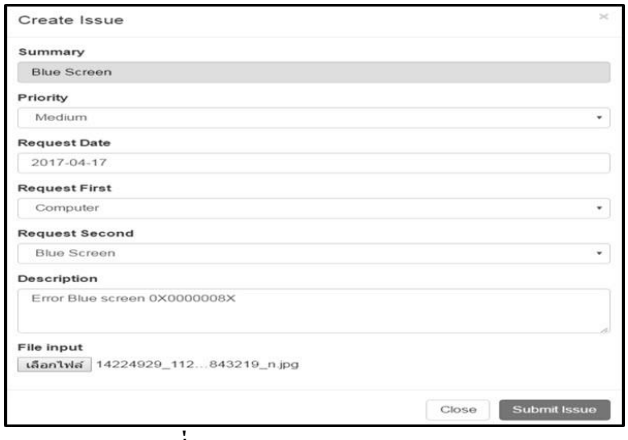
หน้าหลักของโปรแกรมเป็นหน้าแรกหลังจากผ่านการล็อกอินเข้าใช้งาน ซึ่งโปรแกรมแสดงให้เห็นถึงการแสดงผลได้อย่างเหมาะสมบนอุปกรณ์อื่นๆ เช่น คอมพิวเตอร์ โน๊ตบุ๊ค และโทรศัพท์มือถือซึ่งเป็นคุณสมบัติของการใช้ Bootstrap [5] แสดงดังภาพที่ 6



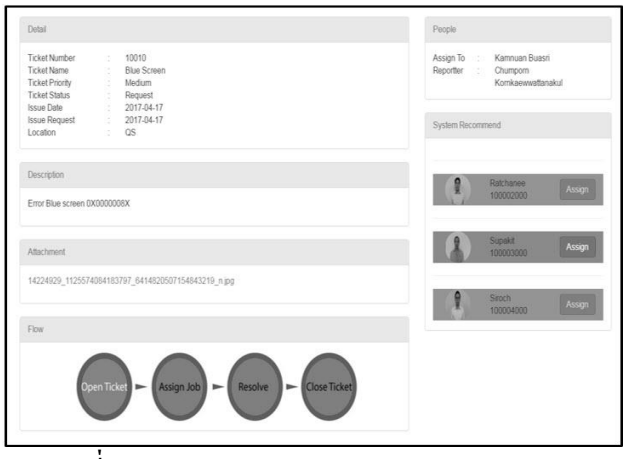
ภาพที่ 6: แสดงหน้าต่าง Home บนอุปกรณ์มือถือ

ในภาพที่ 7-8 แสดงให้เห็นถึงการใช้หลักการของต้นไม้ตัดสินใจ[4] โดยในภาพที่ 7 แสดงให้เห็นถึงขั้นตอนการสร้างประเด็นที่ใช้เพื่อการร้องของานโดยใช้วิธีการเลือกจากหัวข้อแรก Request First จากนั้นระบบจะดึงขอมูลในส่วนที่เกี่ยวกับหัวข้อแรกขึ้นมาแสดง หลังจากนั้นก็เลือกในหัวข้อถัดไป ซึ่งจะช่วยให้ผู้ที่ร้องของานสามารถร้องของานได้ตรงประเด็นปัญหาและรวดเร็ว ซึ่งจะส่งผลต่อการเลือกผู้ที่จะเข้ามาแก้ปัญหาในประเด็นการร้องของาน

ภาพที่ 8 จะแสดงให้เห็นถึง ประเด็นปัญหาที่ได้มีการรองขอเพื่อให้เจ้าหน้าที่ฝ่ายเทคโนโลยีสารสนเทศเข้าไปแก้ไขในประเด็นปัญหา ซึ่งจะประกอบด้วยรายละเอียดที่ทางผู้ร้องของานได้กำหนดตรงขั้นตอนสร้างการร้องของานสารสนเทศ (Create Issue)สถานะของงานจะเป็นตัวกำหนดว่าอยู่ในขั้นตอนใด ระบบจะแสดงให้เห็นเป็นรูปภาพชัดเจนโดยต้นไม้ตัดสินใจจะแยกแยะผู้ที่จะเข้าไปดำเนินการแก้ไข ตั้งแต่ขั้นตอนการ Create Issue แล้ว เมื่อแยกแยะประเด็นปัญหาได้แล้ว ต้นไม้ตัดสินใจจะแยกแยะประเด็นปัญหา และส่งไปยังฉลาก (Label) Label จะทำหน้าที่เป็นตัวแทนของเจ้าหน้าที่ฝ่ายเทคโนโลยีสารสนเทศอยู่ ซึ่งถูกกำหนดโดยผู้จัดการฝ่ายเทคโนโลยีสารสนเทศ ซึ่งสามารถกำหนดเลือกจำนวนผู้รับผิดชอบในประเด็นปัญหาของแต่ละหัวข้อหากมีเจ้าหน้าที่เทคโนโลยีสารสนเทศท่านใดสถานะว่างอยู่ ระบบจะแสดงเป็นสีเขียวและสามารถกำหนดงานให้เจ้าหน้าที่ท่านนั้นได้ แสดงดังภาพที่ 8



ภาพที่ 7: แสดงหน้าต่างการ Create Issue



ภาพที่ 8: แสดงหน้าต่าง Issue และ System Recommend

เมื่อผู้จัดการฝ่ายเทคโนโลยีสารสนเทศได้ Assign งานให้กับเจ้าหน้าที่ฝ่ายแล้วทางผู้ได้รับมอบหมายงานจะเข้าไปแก้ไขปัญหาแล้วเปลี่ยนสถานะในระบบเป็น Resolve หลังจากนั้น Issue จะถูกส่งกลับไปยังผู้ร้องของานเพื่อทำการปิดงาน

4.2 ผลการประเมินประสิทธิภาพของระบบ

งานวิจัยนี้ผู้วิจัยได้ใช้การทำการประเมินผลโดยกลุ่มตัวอย่างจำนวน 3 กลุ่มได้แก่ ผู้จัดการฝ่ายเทคโนโลยีสารสนเทศ เจ้าหน้าที่ฝ่ายเทคโนโลยีสารสนเทศ และเจ้าหน้าที่ฝ่ายอื่นๆ ซึ่งผลการประเมินในภาพรวามอยู่ในระดับดีมีค่าเฉลี่ย และส่วนเบี่ยงเบนมาตราฐานได้ตามลำดับดังนี้ 4.06 (SD 0.54), 4.20 (SD 0.53) และ 4.15 (SD 0.60) ซึ่งแสดงให้เห็นดังตารางที่ 1

ตารางที่ 1: ผลการประเมินประสิทธิภาพโดยผู้เชี่ยวชาญ

รายการประเมิน	ระดับประสิทธิภาพ		
	ค่าเฉลี่ย	ส่วนเบี่ยงเบนมาตรฐาน	ผลการประเมิน
สำหรับเจ้าหน้าที่ผู้ร้องของาน (User)			
1.ด้านฟังก์ชั่นการใช้งาน	4.13	0.79	ดี
2.ด้านการประมวลผล	4.15	0.58	ดี
3.ด้านความยากง่าย	4.00	0.61	ดี
4.ด้านความปลอดภัย	4.33	0.43	ดี
สำหรับผู้จัดการฝ่ายไอที (IT Manager)			
1.ด้านฟังก์ชั่นการใช้งาน	4.00	0.57	ดี
2.ด้านการประมวลผล	4.00	0.67	ดี
3.ด้านความยากง่าย	4.10	0.38	ดี
4.ด้านความปลอดภัย	4.15	0.53	ดี
สำหรับเจ้าหน้าที่ฝ่ายไอที (IT Technical Staff)			
1.ด้านฟังก์ชั่นการใช้งาน	4.13	0.76	ดี
2.ด้านการประมวลผล	4.10	0.31	ดี
3.ด้านความยากง่าย	4.15	0.49	ดี
4.ด้านความปลอดภัย	4.40	0.54	ดี

5. สรุป

จากการศึกษาและวิจัยในครั้งนี้ทำให้ได้ระบบสนับสนุนการตัดสินใจเพื่อช่วยมอบหมายงานให้เจ้าหน้าที่ฝ่ายเทคโนโลยีสารสนเทศ สามารถเลือกบุคลากรมารับงานได้อย่างเหมาะสม ทำให้เจ้าหน้าที่ฝ่ายเทคโนโลยีสารสนเทศเข้าไปแก้ไขงานได้อย่างมีประสิทธิภาพ และผู้ร้องของานได้รับการแก้ไขปัญหาจากการร้องขอได้อย่างมีประสิทธิภาพ ผลการประเมินจากกลุ่มตัวอย่างทั้ง3กลุ่มได้ผลอยู่ในระดับดีในทุกประเด็นซึ่งระบบนี้สามารถนำมาประยุกต์ใช้งานจริงได้

สำหรับข้อเสนอแนะจากการทำวิจัยในครั้งนี้ผู้วิจัยได้พบอุปสรรคและปัญหาในการออกแบบระบบ การเก็บข้อมูลที่ยังไม่ครอบคลุม ถ้าต้องการจะพัฒนาระบบให้ดีขึ้นผู้วิจัยมีความเห็นว่า การทำงานหรือเก็บข้อมูลในแต่ละขั้นตอนควรได้รับความคิดเห็นจากผู้เชี่ยวชาญระบบ ในขั้นตอนการออกแบบ ระบบควรให้เจ้าหน้าฝ่ายเทคโนโลยีสารสนเทศเข้ามามีส่วนร่วมในการออกแบบ และสำหรับการต่อยอดงานวิจัยนี้ผู้วิจัยมีข้อเสนอแนะว่าควรมีการศึกษาเพิ่มเติมในเรื่องการกำหนดระยะเวลาในการปฏิบัติงานของเจ้าหน้าที่ฝ่ายเทคโนโลยีสารสนเทศ และให้คำแนะนำในการแก้ไขปัญหาในเบื้องต้นแก่ผู้ร้องของงานและให้คำแนะนำในการแก้ไขปัญหากับเจ้าหน้าที่ฝ่ายเทคโนโลยีสารสนเทศ

เอกสารอ้างอิง

[1] สำนักงานสถิติแห่งชาติ, กระทรวงดิจิทัลเพื่อเศรษฐกิจและสังคม. (2560). สำรวจการมีการใช้เทคโนโลยีสารสนเทศและการสื่อสารในสถานประกอบการ พ.ศ. 2559. กทม : บริษัท ธนาเพรส จำกัด.

[2] 5 อันดับไวรัสคอมพิวเตอร์ที่สร้างความเสียหายรายแรง. (2556). [ออนไลน์]. [สืบค้นวันที่ 21 กันยายน 2559]. จาก http://www.techxcite.com/topic/15187.html

[3] คำนวน บัวศรี. (วันที่ 21 กันยายน 2559). ผู้จัดการฝ่ายเทคโนโลยีสารสนเทศ บริษัท ฟาสโก้ มอเตอร์ (ประเทศไทย) จำกัด สัมภาษณ์.

[4] ระบบสนับสนุนการตัดสินใจ(DecisionSupportSystem). (2552). [ออนไลน์]. [สืบค้นวันที่ 1 ตุลาคม 2559]. จาก http://armka2518 .exteen.com/20090209/decision-support-system-dds

[5] ขั้นตอนการสร้างโมเดลต้นไม้ตัดสินใจ (Decision Tree). (2014). [ออนไลน์]. [สืบค้นวันที่ 1 ตุลาคม 2559] จาก http://dataminingtrend.com/2014/decision-tree-model/

[6] Web Services Architecture. (2014). [online]. [cited 2014]. Available from : http://www.w3.org/TR/Ws-arch.

[7] Bootstrap Framework for development responsive. (2011). [online]. [cited 2011]. Available from : http://getbootstrap.com.

[8] พูลสักดิ์ หลาบสีดา. (2559). การพัฒนาระบบสารสนเทศเพื่อการจัดการงานบริการของสายงานเทคโนโลยีและสารสนเทศภายในบริษัท: กรณีศึกษา บริษัทประกันภัยแห่งหนึ่งในกรุงเทพมหานคร. คณะวิศวกรรมศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ

[9] กำธร ศรีอุดม. (2554). ระบบสนับสนุนการตัดสินใจสำหรับผู้ใช้บริการเว็บเซอร์วิสการท่องเที่ยวโดยใช้ต้นไม้การตัดสินใจ. ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น

[10] โปรแกรมเชิงวัตถุ และยูเอ็มแอล (UML). (2560). [ออนไลน์]. [สืบค้นวันที่ 20 มีนาคม 2560]. จาก http://www.thaiall.com/uml/indexo.html

[11] Hypertext Preprocessor (PHP), (1999). [online]. [cited 1999]. Availble from: https://www.w3schools.com/php/default.asp

แอปพลิเคชันคัดพยัญชนะอาเซียนบนอุปกรณ์เคลื่อนที่
Asean Alphabet Handwriting Application on Mobile Device

จิรัฐติกาล กันตีมูล (Jiruttigarn Gunteemoon)[1] ศุภลักษณ์ มงคลรัตน์ (Supaluck Mongkhonrat)[2]
เพชร อิ่มทองคำ (Phet Imtongkhum)[3] และจักรชัย โสอินทร์ (Chakchai So-In)[4]
ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น
[1]jiruttigarn@gmail.com, [2]Supaluck.mongkhonrat@gmail.com, [3]phet@kkumail.com, [4]chakso@kku.ac.th

บทคัดย่อ

เนื่องจากประเทศไทยได้ก้าวเข้าสู่การเป็นสมาชิกประชาคมอาเซียนในปีที่ผ่านมา เมื่อเปรียบเทียบกับสถิติการศึกษาของประเทศไทยกับกลุ่มประเทศอาเซียนนั้น พบว่าระดับการศึกษาอยู่ในอันดับรั้งท้าย จึงอาจทำให้ได้รับผลกระทบในด้านต่าง ๆ จากการเข้าร่วมเป็นสมาชิกนี้ ดังนั้นทางผู้จัดทำจึงมีแนวคิดที่จะผนวกเทคโนโลยีมาช่วยพัฒนาด้านการศึกษา โดยพัฒนาแอปพลิเคชันคัดพยัญชนะอาเซียนบนอุปกรณ์เคลื่อนที่ ซึ่งประกอบด้วย 4 ฟังก์ชันหลัก ได้แก่ ฝึกคัด แบบฝึกหัด เกม และ AR Activity Mode โดยจะมุ่งเน้นให้เด็กนักเรียนในวัยปฐมวัยได้รู้จักกับพยัญชนะอาเซียน เพื่อนำไปเป็นพื้นฐานในการเขียนและสามารถประยุกต์ใช้ในการเรียนรู้ภาษาอาเซียนต่อไป โดยผลการทดสอบปรากฏผลผ่านในทุกฟังก์ชัน และจากผู้ใช้งานที่โรงเรียนอนุบาลสาธิต มหาวิทยาลัยขอนแก่น และอนุบาลขอนแก่น แสดงถึงความพึงพอใจเฉลี่ยอยู่ระดับดี

คำสำคัญ: ประชาคมอาเซียน พยัญชนะ แบบฝึกคัด เทคโนโลยีภาพเสมือนจริง

Abstract

Since Thailand has entered the ASEAN Community in the past year, statistically compared to the education of Thailand and ASEAN countries, it was found that the level of education is in the rear. It may be affected by the various aspects of joining this membership. Thus, the authors have ideas to bring technology to help develop the education by building an ASEAN alphabetic application on a mobile device. It consists of 4 main functions: practice, exercises, game, and AR Activity Mode. This application focuses on early childhood learners with the ASEAN alphabets to put it as a basis for writing and can be further applied to learn the language of ASEAN. The functionality evaluation yields pass for all functions and from users of the demo application at kindergarten, Khon Kaen University, it appears that the average satisfaction is good.

Keywords: ASEAN, Alphabet, Handwriting Exercises, Augmented Reality

1. บทนำ

เนื่องจากประเทศไทยได้ก้าวเข้าสู่การเป็นประชาคมอาเซียนในวันที่ 31 ธันวาคม 2558 โดยประกอบด้วยประเทศสมาชิกทั้งหมด 10 ประเทศ ซึ่งได้แก่ ไทย, พม่า, ลาว, เวียดนาม, มาเลเซีย, สิงคโปร์, อินโดนีเซีย, ฟิลิปปินส์, กัมพูชา และบรูไน ซึ่งการเข้าเป็นสมาชิกประชาคมอาเซียนของประเทศไทยนั้นถือว่ามีผลดีหลายประการอย่างไรก็ตามเมื่อพิจารณาในรายละเอียดก็จะปรากฏถึงผลกระทบบางประเด็นโดยเฉพาะ ด้านคุณภาพการศึกษา เมื่อเปรียบเทียบสถิติการศึกษาของประเทศไทยกับกลุ่มประเทศอาเซียนแล้วพบว่า คุณภาพการศึกษา ที่เป็นพื้นฐานสำคัญในการพัฒนาประเทศไทย ในรอบ 3 ปี ย้อนหลังอยู่ในอันดับรั้งท้าย ซึ่งผลที่ตามมาอาจทำให้ประเทศไทยเสียโอกาส หรือได้รับผลกระทบในด้านต่าง ๆ ตามมา [1]

ในปัจจุบันเทคโนโลยีได้เข้ามามีบทบาทมากในชีวิตประจำวันและเข้าถึงได้ทุกเพศทุกวัย ทำให้การเลี้ยงดูเด็ก ๆ ของผู้ปกครองส่วนใหญ่เปลี่ยนแปลงไป โดยผู้ปกครองมักจะจัดหาสมาร์ตโฟน แท็บเล็ต หรืออุปกรณ์ทางคอมพิวเตอร์ให้เด็กได้ฝึกการเรียนรู้ผ่านทางแอปพลิเคชันต่าง ๆ เช่น แอปพลิเคชันที่เกี่ยวกับเรียนรู้ด้านภาษา ด้านการคำนวณ หรือด้านศิลปะ

ซึ่งก็มักจะมีรูปแบบที่แตกต่างกัน อีกทั้งยังมีภาพ เสียงหรือเกม ทำให้ดึงดูดความสนใจของเด็กได้เป็นอย่างมาก และยังได้รับความสนุกสนานอีกด้วย

ดังนั้นทางผู้จัดทำจึงได้มีแนวคิดที่จะสร้างเครื่องมือหรือสื่อการเรียนการสอนเพื่อช่วยส่งเสริมการเรียนรู้ให้กับการศึกษาของเด็กเพื่อที่จะปรับตัวเข้าสู่ประชาคมอาเซียน โดยพัฒนาแอปพลิเคชันคัดพยัญชนะอาเซียนบนอุปกรณ์เคลื่อนที่ซึ่งประกอบด้วย 4 ฟังก์ชันหลักคือ ฝึกคัด แบบฝึกหัด เกม และ AR Activity Mode ที่ช่วยกระตุ้นให้เด็กมีความสนใจในการเรียนรู้เพิ่มมากขึ้น อีกทั้งยังเป็นพื้นฐานในการฝึกเขียน และการเรียนรู้ภาษาอาเซียนต่อไปในอนาคต

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 ภาษาอาเซียน

ภาษาราชการของประเทศสมาชิกอาเซียน ได้แก่ ประเทศอินโดนีเซียใช้ภาษาอินโดนีเซีย, ประเทศมาเลเซีย และประเทศบรูไนใช้ภาษามาเลย์, ประเทศฟิลิปปินส์และสิงคโปร์ใช้ภาษาอังกฤษ, ประเทศเวียดนามใช้ภาษาเวียดนาม, ประเทศลาวใช้ภาษาลาว, ประเทศพม่าใช้ภาษาพม่า, ประเทศกัมพูชาใช้ภาษาเขมร และประเทศไทยใช้ภาษาไทย [2]

2.2 การเรียนรู้ (Learning Ecology)

- ฮิลการ์ด และ เบาเวอร์ [3] “การเรียนรู้ เป็นกระบวนการเปลี่ยนแปลงพฤติกรรม อันเป็นผลมาจากประสบการณ์และการฝึก ทั้งนี้ไม่รวมถึงการเปลี่ยนแปลงของพฤติกรรมที่เกิดจากการตอบสนองตามสัญชาตญาณ”

- พจนานุกรมของเว็บสเตอร์ [4] “การเรียนรู้ คือ กระบวนการเพิ่มพูน และปรุงแต่งระบบความรู้ ทักษะ นิสัย หรือการแสดงออกต่าง ๆ อันมีผลมาจากสิ่งกระตุ้นอินทรีย์โดยผ่านประสบการณ์ การปฏิบัติ หรือการฝึกฝน”

- ประดินันท์ อุปรมัย [5] “การเรียนรู้คือการเปลี่ยนแปลงของบุคคลอันมีผลเนื่องมาจากการได้รับประสบการณ์ โดยการเปลี่ยนแปลงนั้นเป็นเหตุทำให้บุคคลเผชิญสถานการณ์เดิมแตกต่างไปจากเดิม ประสบการณ์ที่ก่อให้เกิดการเปลี่ยนแปลงพฤติกรรมหมายถึงทั้งประสบการณ์ทางตรงและประสบการณ์ทางอ้อม”

2.3 แอนดรอยด์ (Android)

เป็นระบบปฏิบัติการสำหรับอุปกรณ์พกพา เช่น โทรศัพท์มือถือ แท็บเล็ต และคอมพิวเตอร์ ซึ่งเป็นระบบที่มีพื้นฐานการทำงานบนลินุกซ์ เคอร์เนล ปัจจุบันเป็นของบริษัทกูเกิลและได้พัฒนาในนามของ Open Handset Alliance ทางกูเกิลได้เปิดให้นักพัฒนาสามารถแก้ไขโค้ดต่างๆ ด้วยภาษา Java และควบคุมอุปกรณ์ผ่านทางชุด Java libraries ที่กูเกิลพัฒนาขึ้น แอนดรอยด์ จึงเป็นซอฟต์แวร์ Open Source ที่ได้รับความนิยมเป็นอย่างมาก และยังมีฟังก์ชันที่หลากหลาย [6]

2.4 Augmented Reality (AR)

เป็นเทคโนโลยีที่พัฒนาในรูปแบบ Human-Machine Interface ที่อาศัยเทคโนโลยีคอมพิวเตอร์ และระบบเสมือนจริง โดยที่วัตถุเสมือนนั้น ๆ จะถูกสร้างมาผสมกับสภาพในโลกจริงในรูป 3 มิติ และแสดงผลแบบ Real Time โดยเทคโนโลยีนี้จะต้องประกอบด้วย 3 ระบบ คือ ระบบ Tracking ระบบแสดงภาพ และระบบประมวลผลเพื่อสร้างวัตถุ 3 มิติ โดยระบบ Tracking (กล้อง) จะรับข้อมูลรูปภาพ เช่น รูปแบบ ตำแหน่ง และทิศทาง จากนั้นระบบประมวลผลก็จะนำไปแปลความหมาย และแสดงภาพ 3 มิติ ออกมาในตำแหน่งและทิศทางเดียวกันกับภาพที่กล้องจับได้ [7]

2.5 งานวิจัยที่เกี่ยวข้อง

2.5.1 แอปพลิเคชันการเรียนรู้ภาษากัมพูชาบนระบบปฏิบัติ การแอนดรอยด์ เป็นเว็บแอปพลิเคชันพัฒนาสอนภาษากัมพูชาบนสมาร์ตโฟนระบบปฏิบัติการแอนดรอยด์มีการสอน 3 หัวข้อหลัก คือ พยัญชนะ คำศัพท์ และสนทนา โดยมีประโยคสนทนาพร้อมกับมีเสียง สะดวกต่อการใช้งาน และสามารถเรียกใช้ตามหมวดหมู่ได้ แต่จะเป็นการเน้นการสอนพยัญชนะ คำศัพท์ และการสนทนา เท่านั้น และจำกัดภาษาเพียงกัมพูชา [8]

2.5.2 แอปพลิเคชันการเขียนอักษรภาษาไทย สำหรับเด็กในวัยเริ่มต้นชุด ก-ฮ โดยเป็นแอปพลิเคชันการเขียนหรือการคัดอักษรภาษาไทย ซึ่งมีการประยุกต์ใช้เทคโนโลยีแฟลชที่ทำงานบนระบบปฏิบัติการแอนดรอยด์ โดยมีข้อดีคือ การแสดงภาพแบบแอนิเมชันให้ผู้เรียนสามารถเข้าใจได้ง่าย แต่จะเน้นที่การสอนพยัญชนะ การผสมคำ และคำศัพท์ และจำกัดเพียงภาษาไทย [9]

2.5.3 แอปพลิเคชันเก่งไทย เป็นแอปพลิเคชันบนสมาร์ตโฟน ซึ่งประกอบด้วยเนื้อหาภาษาไทยพื้นฐาน ตั้งแต่ พยัญชนะ สระวรรณยุกต์ ตัวเลข โดยมีลักษณะการทำงานเป็น

Game-based Learning ที่ช่วยแก้ปัญหาการอ่านหนังสือไม่ออกและเขียนหนังสือไม่ได้ตั้งแต่ระดับอนุบาลหรือประถมศึกษา โดยมีการแสดงลำดับและทิศทางในการเขียน มีภาพการ์ตูนแอนิเมชันประกอบตัวอักษรหลังจากเขียนเสร็จ และมีรูปแบบการคัดที่ครบถ้วน ได้แก่ การคัดพยัญชนะ สระ และตัวเลข แต่ในส่วนของเกมจะจำกัดที่ 1 เกม หรือมีค่าใช้จ่ายที่มากกว่านั้น [10]

2.5.4 แอปพลิเคชันคัดลายมือ Thai Handwriting เป็นแอปพลิเคชันคัดลายมือ ซึ่งมีการคัดพยัญชนะ ก – ฮ, A – Z, a – z และตัวเลข 0 – 9, ๐ – ๙ โดยสามารถเขียนซ้ำได้ เปลี่ยนขนาดตัวอักษร และมีแบบการฝึกคัดพยัญชนะภาษาไทยและพยัญชนะภาษาอังกฤษ อย่างไรก็ตามจะไม่มีส่วนของการตรวจสอบความถูกต้องของการเขียน [11]

จากแอปพลิเคชันที่ได้กล่าวมา เมื่อเปรียนเทียบกับแอปพลิเคชันคัดพยัญชนะอาเซียนบนอุปกรณ์เคลื่อนที่ จะพบว่าแอปพลิเคชันที่พัฒนาขึ้นมีความโดดเด่นที่แตกต่าง คือ มีความหลากหลายของภาษา ซึ่งประกอบไปด้วยภาษาในประเทศอาเซียนจำนวน 8 ภาษา อีกทั้งยังทั้งในส่วนของ ฝึกคัดและแบบฝึกหัด รวมไปถึงมี เกม และผนวกฟังก์ชันของ AR Activity Mode ที่มาเพิ่มความน่าสนใจในการใช้งานอีกด้วย

3. วิธีดำเนินงาน

3.1 วิเคราะห์ปัญหาและความต้องการ

การวิเคราะห์และหาความต้องการจากระบบการทำงานเดิมของแอปพลิเคชันที่มีความเกี่ยวข้องกับการฝึกทักษะการคัดพยัญชนะในภาษาต่าง ๆ

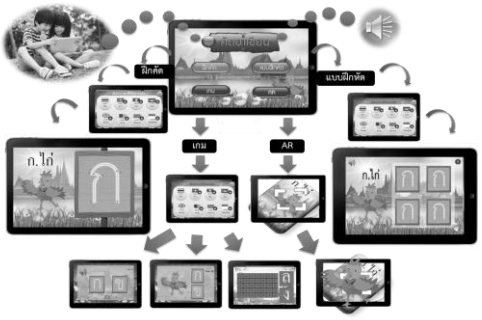
3.2 ค้นคว้า ศึกษาทฤษฎี และงานวิจัยที่เกี่ยวข้อง

ผู้จัดทำได้ทำการศึกษาและรวบรวมข้อมูลจากหลายแหล่ง โดยแบ่งออกเป็น 3 ส่วน คือ

- ศึกษาแอปพลิเคชันที่มีความเกี่ยวข้องจากเว็บไซต์ต่าง ๆ
- ศึกษาและรวบรวมข้อมูลเกี่ยวกับพยัญชนะในประเทศอาเซียน จากศูนย์ภาษาอาเซียนมหาวิทยาลัยขอนแก่น และจากเว็บไซต์ต่าง ๆ
- ศึกษาข้อมูลและวิธีใช้งานเครื่องมือต่าง ๆ เช่น Android studio, Autodesk 3ds Max และ Unity

3.3 การออกแบบระบบ และวิเคราะห์ระบบ

การออกแบบระบบของแอปพลิเคชันประกอบด้วย 4 ฟังก์ชันหลัก คือ ฝึกคัด แบบฝึกหัด เกม และ AR Activity Mode (ดังภาพที่ 1)



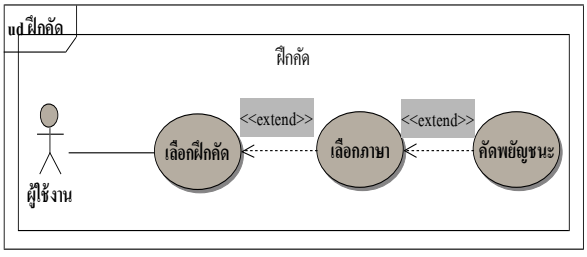
ภาพที่ 1: ภาพรวมของแอปพลิเคชัน

3.4 Use Case Diagram (แผนภาพแบบจำลอง)

Use Case Diagram แบ่งเป็น 4 ฟังก์ชันหลักดังนี้คือ

3.4.1 Use Case Diagram ฝึกคัด

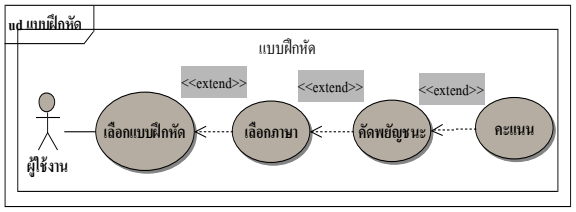
ผู้ใช้งานเลือกฟังก์ชันฝึกคัด เลือกภาษา และเลือกพยัญชนะที่ต้องการคัด แล้วแอปพลิเคชันแสดงหน้าจอการฝึดคัด โดยมีทั้งหมด 8 ภาษา คือ ภาษาไทย ภาษาอังกฤษ ภาษามาเลย์ ภาษาพม่า ภาษาเขมร ภาษาลาว ภาษาอินโดนีเซีย และภาษาเวียดนาม



ภาพที่ 2: Use Case Diagram ฝึกคัด

3.4.2 Use Case Diagram แบบฝึกหัด

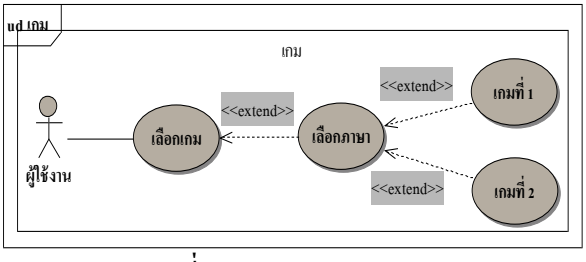
ผู้ใช้งานเลือกฟังก์ชันแบบฝึกหัด เลือกภาษา และเลือกพยัญชนะที่ต้องการคัด โดยมีทั้งหมด 8 ภาษา คือ ภาษาไทย ภาษาอังกฤษ ภาษามาเลย์ ภาษาพม่า ภาษาเขมร ภาษาลาว ภาษาอินโดนีเซีย และภาษาเวียดนาม แอปพลิเคชันแสดงหน้าจอแบบฝึกหัด แล้วเมื่อคัดพยัญชนะแล้วจะมีการให้คะแนนเป็นดาว (คะแนนเต็ม 4 ดาว คิดจากเปอร์เซ็นต์ความครอบคลุมของพื้นที่)



ภาพที่ 3: Use Case Diagram แบบฝึกหัด

3.4.3 Use Case Diagram เกม

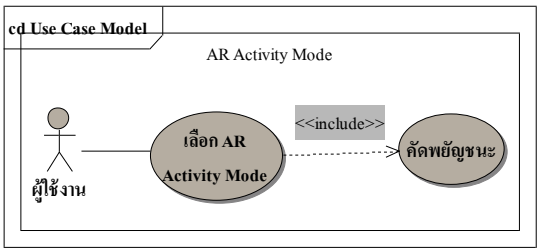
ผู้ใช้งานเลือกฟังก์ชันเกม เลือกภาษาโดยมีทั้งหมด 8 ภาษา คือ ภาษาไทย ภาษาอังกฤษ ภาษามาเลย์ ภาษาพม่า ภาษาเขมร ภาษาลาว ภาษาอินโดนีเซีย และภาษาเวียดนาม จากนั้นเลือกเกม มี 2 เกม คือ เกมคัดพยัญชนะจากภาพที่ปรากฏ และเกมคัดพยัญชนะจากเสียงที่ได้ยิน โดยความถูกต้องที่ได้เปรียบเทียบกับเปอร์เซ็นต์ความครอบคลุมของพื้นที่



ภาพที่ 4: Use Case Diagram เกม

3.4.4 Use Case Diagram AR Activity Mode

ผู้ใช้งานเลือกฟังก์ชัน AR และส่องไปที่การ์ด (ในตัวอย่างนี้มี 2 ชุดภาษา คือ ภาษาไทยและอังกฤษ) ก็จะแสดงภาพ 3 มิติขึ้นมา เมื่อคลิกที่ภาพก็จะปรากฏแอปพลิเคชันในส่วนแบบฝึกหัด



ภาพที่ 5: Use Case Diagram AR Activity Mode

3.4.5 เทคนิคที่ใช้ในการพัฒนา

- ฟังก์ชันฝึกคัด แบบฝึกหัด และเกม ใช้เทคนิค Drawing ในการเขียน และใช้ Grid Coordinates ในการตรวจ สอบความถูกต้องของการเขียน
- ฟังก์ชัน AR Activity Mode ใช้ เทคนิค AR ร่วมกับ Unity และ Vuforia ในการแสดงภาพ เสมือนจริง

4. ผลการดำเนินงาน

4.1 ตัวอย่างการพัฒนาแอปพลิเคชัน



ภาพที่ 6: หน้าแรกของแอปพลิเคชันประกอบไปด้วย 4 ฟังก์ชัน คือ ฝึกคัด, แบบฝึกหัด, เกม และ AR Activity Mode



ภาพที่ 7: กรณีที่เลือกฟังก์ชันฝึกคัด แบบฝึกหัด และเกม จะแสดงหน้าเลือกภาษาที่ต้องการคัดประกอบไปด้วย 8 ภาษา คือ ภาษาไทย ภาษาอังกฤษ ภาษาลาว ภาษาพม่า ภาษามาเลย์ ภาษาเขมร ภาษาเวียดนาม และภาษาอินโดนีเซีย





ภาพที่ 8: เลือกภาษาไทย หรือเลือกภาษาอังกฤษจะแสดงหน้าพยัญชนะของภาษาไทย ผู้ใช้งานเลือกคัดพยัญชนะตัวที่ต้องการ หากเลือกภาษาอื่นก็จะมีรูปแบบเดียวกัน แต่พยัญชนะที่แสดงจะแตกต่างกัน





ภาพที่ 9: เลือกคัด ก (ภาษาไทย) หรือ A (ภาษาอังกฤษ) จะแสดงภาพพยัญชนะ 4 ตัว และเสียงให้ ผู้ใช้งานได้ฝึกคัด หากเลือกคัดพยัญชนะตัวอื่น ๆ ก็จะมีรูปแบบเดียวกันแต่แตกต่างกันที่ภาพพยัญชนะและเสียง




ภาพที่ 10: กรณีที่คัดพยัญชนะไม่ถูกต้องและไม่ครบจะแสดงดังภาพด้านซ้ายมือพร้อมมีเสียง ให้ลองคัดใหม่อีกครั้ง และเมื่อคัดพยัญชนะถูกต้องและคัดครบถ้วนจะแสดงดังภาพด้านขาวมือพร้อมเสียงพยัญชนะตัวนั้น ๆ ส่วนภาษาอื่นมีรูปแบบเดียวกันแต่แตกต่างกันที่ภาพพยัญชนะและเสียง





ภาพที่ 11: เมื่อเลือกฟังก์ชันแบบฝึกหัด จะแสดงหน้าดังภาพที่ 7 และ 8 จากนั้นถ้าผู้ใช้งานเลือกคัด ก (ภาษาไทย) หรือเลือกคัด A (ภาษาอังกฤษ) จะแสดงภาพพยัญชนะ และเสียงให้ผู้ใช้งานคัด หากเลือกคัดพยัญชนะตัวอื่น ๆ ก็จะมีรูปแบบเดียวกันแต่แตกต่างกันที่ภาพพยัญชนะและเสียง





ภาพที่ 12: เมื่อผู้ใช้งานคัดพยัญชนะจากภาพที่ 11 เรียบร้อยแล้วจะมีการให้คะแนนความถูกต้อง โดยให้คะแนนเป็นดาว ซึ่งคะแนนเต็มจะอยู่ที่ 4 ดาว และลดลงมาตามลำดับความถูกต้อง



ภาพที่ 13: เมื่อเลือกฟังก์ชันเกมจะแสดงหน้าดังภาพที่ 7 และ 8 จากนั้นให้ผู้ใช้งานเลือกเกม ซึ่งจะมี 2 เกม ส่วนภาษาอื่น ๆ ก็จะมีรูปแบบเดียวกัน





ภาพที่ 14: เลือกเกมที่ 1 เป็นเกมที่ให้ผู้เล่นเลือกคัดพยัญชนะจากภาพที่แสดงให้ถูกต้อง เมื่อคัดเสร็จระบบจะสุ่มภาพและพยัญชนะจนกว่าจะหมดเวลาที่กำหนด (30 วินาที) และแสดงคะแนนทั้งหมดที่คัดถูกต้อง





ภาพที่ 15: เลือกเกมที่ 2 เป็นเกมที่ให้ผู้เล่นเลือกคัดพยัญชนะจากเสียงที่ได้ยินให้ถูกต้อง เมื่อคัดเสร็จระบบจะสุ่มภาพและพยัญชนะไปเรื่อย ๆ จนกว่าจะหมดเวลาที่กำหนด (30 วินาที) และแสดงคะแนนทั้งหมดที่คัดถูกต้อง




ภาพที่ 16: เมื่อเลือกฟังก์ชัน AR จะเข้าสู่โหมด AR Activity ผู้เล่นจะต้องเลือกการ์ดที่ต้องการ แล้วทำการส่องไปที่การ์ด มีจะภาพ 3 มิติแสดงขึ้นมา และถ้ากดที่ภาพ 3 มิติ จะเข้าสู่หน้าคัดพยัญชนะตัวนั้น ดังภาพที่ 11

4.2 การทดสอบฟังก์ชันหลักในแอปพลิเคชัน

เป็นการทดสอบ 4 ฟังก์ชันหลักตาม Use Case Diagram โดยทดสอบจำนวน 10 ครั้ง (ผ่าน/ไม่ผ่าน) ดังนี้

4.2.1. ทดสอบฟังก์ชันการทำงาน

- **การเข้าฟังก์ชันฝึกคัด** เลือกฟังก์ชันฝึกคัด เลือกภาษา และเลือกตัวอักษร แอปพลิเคชันสามารถแสดงหน้าคัดพยัญชนะได้ และแสดงตัวอักษรที่ต้องการคัด สรุปว่าผ่าน ทั้ง 10 ครั้ง
- **การเข้าฟังก์ชันแบบฝึกหัด** เลือกฟังก์ชันแบบฝึกหัด เลือกภาษา และเลือกตัวอักษรที่ต้องการ แอปพลิเคชันสามารถแสดงหน้าคัดพยัญชนะได้ และแสดงตัวอักษรที่ต้องการคัด สรุปว่าผ่าน ทั้ง 10 ครั้ง
- **การเข้าฟังก์ชันเกม** เลือกฟังก์ชันเกม เลือกภาษา และเลือกเกมที่ 1 แอปพลิเคชันสามารถแสดงหน้าเกมที่ 1 ได้ สรุปว่าผ่าน ทั้ง 10 ครั้ง และเลือกฟังก์ชันเกมที่ 2 ได้ สรุปว่าผ่าน ทั้ง 10 ครั้ง
- **การเข้าฟังก์ชัน AR Activity Mode** เลือกฟังก์ชัน AR Activity Mode เข้าสู่โหมดกล้อง และส่องไปที่การ์ด แอปพลิเคชันแสดงภาพ 3 มิติ และเข้าสู่หน้าแบบฝึกหัด สรุปว่าผ่าน ทั้ง 10 ครั้ง

4.2.2 ทดสอบความแม่นยำของการคัด

- ฟังก์ชันฝึกคัด จะแสดงพยัญชนะจำนวน 4 ตัว กรณีที่ผู้ใช้งานคัดพยัญชนะ 1-2 ตัว จะแสดงภาพและเสียงให้ลองใหม่อีกครั้ง และหากผู้ใช้งานคัดพยัญชนะ 3-4 ตัว จำนวน 10 ครั้ง คัดถูก 5 ครั้ง ผลปรากฏถูกทั้ง 5 ครั้ง และคัดผิด 5 ครั้ง ผลปรากฏผิดทั้ง 5 ครั้ง (ถูกต้อง 100%)
- ฟังก์ชันแบบฝึกหัด จะแสดงพยัญชนะจำนวน 1 ตัวเมื่อผู้ใช้งานคัดพยัญชนะเสร็จ จะมีการให้คะแนนซึ่งแบ่งเป็น 4 ระดับ โดยจะให้คะแนนเป็นดาว จำนวน 8 ครั้ง (1-4 ดาว ครั้งละ 2 ตรงและไม่ตรง) คัดตรงทั้ง 4 ครั้ง ผลปรากฏตรงทั้ง 4 ครั้ง และคัดไม่ตรง 4 ครั้ง ผลปรากฏไม่ตรงทั้ง 4 ครั้ง (ถูกต้อง 100%)
- ฟังก์ชันเกม ประกอบด้วย 2 เกม เกมที่ 1 เลือกคัดพยัญชนะตามภาพที่แสดง และเกมที่ 2 เลือกคัดพยัญชนะตามเสียงที่ได้ยิน จำนวน 10 ครั้ง คัดถูก 5 ครั้ง ผลปรากฏถูกทั้ง 5 ครั้ง และคัดผิด 5 ครั้ง ผลปรากฏผิดทั้ง 5 ครั้ง (ถูกต้อง 100%)

4.3 แบบสอบถาม

4.3.1 แบบสอบถาม (สำหรับนักเรียน)

ผลจากการทดสอบของผู้ใช้งานหรือนักเรียนชั้นอนุบาล 2-3 โรงเรียนอนุบาลสาธิต มหาวิทยาลัยขอนแก่น และอนุบาลขอนแก่น จำนวน 10 คน พบว่า ระดับความพึงพอใจอยู่ในระดับมาก คิดเป็นร้อยละ 80 ระดับความพึงพอใจอยู่ในระดับปานกลาง คิดเป็นร้อยละ 20 และระดับความพึงพอใจอยู่ในระดับน้อยร้อยละ 0 (ดังตารางที่ 4)

ตารางที่ 4: ผลจากการตอบแบบสอบถามของเด็กผู้ใช้งาน

รายการสอบถามการใช้งาน แอปพลิเคชันคัดพยัญชนะอาเซียน	ระดับความพึงพอใจ		
	มาก	ปานกลาง	น้อย
ผู้ใช้งานมีความพึงพอใจในระดับมากน้อยเพียงใด	8	2	0

4.3.2 แบบสอบถาม (สำหรับผู้ปกครอง)

ผลสรุปจากการตอบแบบสอบถามของผู้ปกครอง โรงเรียนอนุบาลสาธิต มหาวิทยาลัยขอนแก่น และอนุบาลขอนแก่น จำนวน 5 คน พบว่า กระบวนการติดตั้งคิดเป็นค่าเฉลี่ยเท่ากับ 8.00 รูปแบบและภาพคิดเป็นค่าเฉลี่ยเท่ากับ 13.8 การใช้งานคิดเป็นค่าเฉลี่ยเท่ากับ 16.4 และภาพรวมของแอปพลิเคชันคิดเป็นค่าเฉลี่ยเท่ากับ 9.00 (ดังตารางที่ 5)

ตารางที่ 5: ผลจากการตอบแบบสอบถามของผู้ปกครอง ระดับประมาณค่าประสิทธิภาพและคุณภาพ 5 ระดับดังนี้ 5 = ระดับดีมาก, 4 = ระดับดี, 3 = ระดับปานกลาง, 2 = ระดับน้อย, 1 = ระดับปรับปรุง

รายการสอบถามการใช้งาน แอปพลิเคชันคัดพยัญชนะอาเซียนบนอุปกรณ์เคลื่อนที่		ค่าเฉลี่ย
1. กระบวนการติดตั้ง		
1.1	การติดตั้งแอปพลิเคชันง่ายและเหมาะสม	4
1.2	ผู้ใช้งานสามารถเข้าใจ เรียนรู้ และใช้งานแอปพลิเคชันได้อย่างรวดเร็ว	4
2. รูปแบบและภาพ		
2.1	สีสันของภาพ สีตัวอักษร มีความเหมาะสมมากน้อยเพียงใด	4.8
2.2	ตำแหน่งการจัดวางปุ่ม มีความเหมาะสมมากน้อยเพียงใด	4
2.3	ขนาด และรูปแบบของตัวอักษร มีความเหมาะสมมากน้อยเพียงใด	5
3. การใช้งาน		
3.1	แอปพลิเคชันสามารถทำงาน และแสดงผลได้ถูกต้อง	4
3.2	ความเร็วของในการตอบสนองของแอปพลิเคชันมากน้อยเพียงใด	3.4
3.3	หลักการคัดของแต่ละภาษาตรงตามหลักสูตร	4
3.4	แอปพลิเคชันมีประโยชน์ต่อผู้ใช้งานมากน้อยเพียงใด	5
4. ภาพรวมของแอปพลิเคชัน		
4.1	ความน่าสนใจของแอปพลิเคชัน	4.6
4.2	ความทันสมัยของรูปแบบแอปพลิเคชัน ผู้ใช้งานมีความพึงพอใจในระดับใด	4.4

5. สรุปผลและข้อเสนอแนะ

แอปพลิเคชันคัดพยัญชนะอาเซียนบนอุปกรณ์เคลื่อนที่ทำงานบนระบบปฏิบัติการแอนดรอยด์รุ่น 4.0 ขึ้นไป ซึ่งรูปแบบเหมาะสำหรับเด็กที่อยู่ในวัยปฐมวัย (ช่วงอายุ 5-7 ปี) โดยประกอบด้วยพยัญชนะในประเทศอาเซียน (10 ประเทศ 8 ภาษา) ซึ่งแบ่งออกเป็น 4 ฟังก์ชันหลัก คือฟังก์ชันฝึกคัด และฟังก์ชันแบบฝึกหัดมีพยัญชนะ 8 ภาษา พร้อมภาพ และเสียง ฟังก์ชันเกมมี 2 เกม โดยมีรูปแบบเหมือนกันทั้ง 8 ภาษา แต่จะมี ภาพ และเสียงของพยัญชนะที่แตกต่างกัน และ AR Activity

Mode จะมีจำนวนการ์ด 2 ชุด คือ ก – ฮ และ A – Z ซึ่งการให้คะแนนมีเป็น 4 ระดับ โดยจะให้คะแนนเป็นดาว

จากการทดสอบการฟังก์ชันหลักทั้ง 4 ฟังก์ชัน แอปพลิเคชันสามารถทำงานได้ตามที่ผู้ใช้งานต้องการ เช่น ผู้ใช้งานสามารถคัดพยัญชนะ ทำแบบฝึกหัด เล่นเกม และมีความสนใจใน AR Activity Mode นอกจากนี้ในส่วนของการตอบแบบสอบถาม โดยมีผลความพึงพอใจในการใช้แอปพลิเคชันอยู่ในระดับดี ทำให้การพัฒนาแอปพลิเคชันเป็นไปตามวัตถุประสงค์และขอบเขตที่ตั้งไว้ อีกทั้งผู้ใช้งานยังสามารถนำไปประยุกต์ใช้ในการเรียนรู้ภาษาอาเซียนต่อไป และในอนาคตมีแนวโน้มจะพัฒนาต่อยอดโดยการเพิ่มรูปแบบให้สามารถใช้งานได้ทั้งภาษาไทย และภาษาอังกฤษ ส่วนของ AR Activity Mode จะเพิ่มฟังก์ชันที่สนับสนุนครบทั้ง 8 ภาษาของประเทศอาเซียน

เอกสารอ้างอิง

[1] สำนักงานคณะกรรมการการพัฒนาการเศรษฐกิจและสังคมแห่งชาติ. 2554. การเตรียมความพร้อมเข้าสู่ประชาคมอาเซียน ปี 2558. ค้นเมื่อ 25 กันยายน 2558, จาก http://www.nesdb.go.th.

[2] Thai-AEC.com. (2558). ภาษาอาเซียน. ค้นเมื่อ 25 กันยายน 2558, จาก http://www.thai-aec.com/.

[3] Hilgard & Bower. (1981). การเรียนรู้คืออะไร. ค้น เมื่อ 21 พฤศจิกายน 2559, จาก https://candmbsri.wordpress.com/

[4] พจนานุกรมของเว็บสเตอร์. [ม.ป.ป.]. ความหมายของการเรียนรู้. ค้น เมื่อ 21 พฤศจิกายน 2559, จาก http://www.baanjomyut.com /library_2/

[5] ประดินันท์ อุปรมัย. 2540. ความหมายของการเรียนรู้. ค้น เมื่อ 21 พฤศจิกายน 2559, จากชุดวิชาพื้นฐานการศึกษา(มนุษย์กับการเรียนรู้) : นนทบุรี, พิมพ์ครั้งที่ 15, หน้า 121

[6] ภาณุพงศ์ ปัญญาดี. [ม.ป.ป.]. แอนดรอยด์ (Android). ค้นเมื่อ 28 สิงหาคม 2559, จากhttp://www. research-system.siam.edu/images/ coop/Applications_length_products_OTOP/11.%E0%B8 %9A%E0%B8%97%E0%B8%97_2.pdf

[7] Augmented Reality (AR). [ม.ป.ป.]. ค้นเมื่อ 28 สิงหาคม 2559, จาก http://www.research-system.siam.edu/.

[8] ละอองเพชร มิ่งขวัญ, ปวีณา ศิลา. [ม.ป.ป.]. การพัฒนาแอปพลิเคชันการเรียนรู้ภาษากัมพูชาบนระบบปฏิบัติการแอนดรอยด์. สาขาวิชาระบบสารสนเทศ คณะเทคโนโลยีสังคม มหาวิทยาลัยเทคโนโลยี ราชมงคลตะวันออก วิทยาเขตจันทบุรี.

[9] กิตติพงษ์ กันศร. 2555. แอปพลิเคชันการเขียนอักษรภาษาไทยสำหรับเด็กในวัยเริ่มต้น ชุด ก-ฮ. หลักสูตรปริญญา วิทยาศาสตรบัณฑิตภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยสยาม.

[10] Atchariya Darote. 2558. เก่งไทย. ค้นเมื่อ 25 กันยายน 2558, จาก https://itunes.apple.com/th/app/keng-thiy/id697559254l=th&mt=8

[11] มนตรี งามอัครกุล. 2557. คัดลายมือ Thai Handwriting. ค้นเมื่อ 25 กันยายน 2558, จาก http://roshiniblogger.blogspot.com/2012/08/

การสร้างรายงานอัตโนมัติด้วยการสกัดชื่อรายงานและโครงสร้างข้อมูล

Automatic Report Generation using Report Name and Data Structure Extraction

ไตรรัตน์ พัฒน์เอี่ยม (Trirat Pat-iam)[1] และณัฐพงศ์ ทองเทพ (Nattapong Tongtep)[2]
สาขาวิชาเทคโนโลยีสารสนเทศและการสื่อสาร คณะเทคโนโลยีและสิ่งแวดล้อม มหาวิทยาลัยสงขลานครินทร์ วิทยาเขตภูเก็ต
[1]trirat.p@phuket.psu.ac.th, [2]nattapong.t@phuket.psu.ac.th

บทคัดย่อ

การแข่งขันทางธุรกิจในประเทศไทยมีอัตราที่สูงขึ้นมากในช่วงหลายปีที่ผ่านมา มีกลุ่มธุรกิจที่เกิดขึ้นและดับไปมากมาย การจะบริหารธุรกิจให้มั่นคงได้นั้น ผู้ประกอบการจะต้องอาศัยประสบการณ์ วิสัยทัศน์ หนึ่งในสิ่งที่สำคัญต่อผู้ประกอบการคือสารสนเทศ ซึ่งมักอยู่ในรูปแบบของรายงาน สารสนเทศเหล่านี้สำคัญต่อการวางแผนงาน บริหารงานและอื่น ๆ การสร้างรายงานนั้นสำคัญที่ผู้สร้างรายงานต้องมีความรู้ความเข้าใจเกี่ยวกับข้อมูล จึงจะสามารถแปรผลได้ตอบโจทย์ธุรกิจ ในกรณีของผู้ประกอบการรายย่อยหรือรายใหม่อาจไม่สามารถเข้าถึงสารสนเทศที่ดีได้เนื่องจากงบประมาณที่จำกัด ไม่สามารถจัดหาเครื่องมือที่ดีหรือผู้วิเคราะห์ข้อมูลที่ชำนาญได้ ผู้วิจัยได้มองเห็นช่องว่างในการเข้าถึงสารสนเทศที่ดีนี้ จึงได้พัฒนาแบบจำลองที่ช่วยในการเสนอแนวทางในการสร้างรายงาน โดยมีพื้นฐานบนสถิติและความน่าจะเป็น ผู้วิจัยทำการศึกษารายงานในประเทศไทยและนำมาเป็นแนวทางในการออกแบบแบบจำลอง โดยแบบจำลองจะวิเคราะห์ข้อมูลที่ผู้ใช้งานมีและนำเสนอมุมมองในการสร้างรายงานผ่านชื่อของรายงาน ซึ่งชื่อของรายงานนั้นประกอบด้วยกลุ่มคำสี่กลุ่ม แต่ละกลุ่มคำเกิดจากความน่าจะเป็นของแต่ละประเภทของข้อมูล ผู้วิจัยได้พัฒนาระบบบนพื้นฐานของแบบจำลอง โดยเจาะจงกลุ่มธุรกิจพิเศษด้านการศึกษา ระบบสามารถแนะนำรายชื่อที่เป็นไปได้ทั้งหมดเรียงตามลำดับคะแนนให้กับผู้ใช้งานได้อย่างถูกต้องโดยชื่อเป็นชื่อที่สมเหตุสมผลตามความเป็นจริง

คำสำคัญ: การสร้างรายงาน เครื่องมือในการสร้างรายงาน
เครื่องสำหรับกลุ่มธุรกิจ ระบบสารสนเทศ

Abstract

Information has become the important thing when businesses competition has risen in Thailand. A good report is one thing from many factors that will drive the business to win the competition. Report document is the set of information that analyzed. It used to predict things that will happen in the future. To create efficiently reports, information should be analyzed by experts, so if the user is just a newcomer, they will spend more time analyzing the data. In this research, we studied about how the report is made and found that the name of the report can reflect the characteristic of the report. Name of the report can be divided into 4 parts including pre, mid-pre, mid-post and post position, each part has set of words that used frequently. We developed the model using natural language understanding concept that will give the suggestion about report naming by combining words from 4 parts based on data that user has. It allows user to spend less time on reporting process and get report quicker.

Keyword: Report generation, Reporting tool,
Organization tool, Information system.

1. บทนำ

รายงานเป็นหนึ่งในสารสนเทศที่สำคัญอย่างยิ่งต่อการดำเนินธุรกิจทั้งในด้านการวางแผน การเงิน และอื่น ๆ การสร้างรายงานที่มีประสิทธิภาพนั้นจำเป็นต้องอาศัยเครื่องมือหรือผู้วิเคราะห์ข้อมูลที่เข้าใจถึงลักษณะความสัมพันธ์ระหว่างข้อมูล รวมทั้งเข้าใจถึงจุดประสงค์ของรายงาน ซึ่งในกรณีของ

ผู้ประกอบการรายย่อยหรือรายใหม่อาจจะมีทุนน้อยไม่สามารถซื้อเครื่องมือหรือจ้างผู้วิเคราะห์ข้อมูลที่ชำนาญได้ ส่งผลให้รายงานมีสารสนเทศที่ไม่เพียงพอต่อการนำไปใช้งานหรืออาจใช้เวลาในการสร้างรายงานนาน จากปัญหาดังกล่าวทางผู้วิจัยจึงได้สร้างแบบจำลองที่จะลดช่องว่างระหว่างผู้ประกอบการและรายงานที่มีประสิทธิภาพ เพื่อให้ผู้ประกอบการเข้าถึงสารสนเทศเหล่านั้นได้ง่ายขึ้น โดยแบบจำลองออกแบบบนแนวคิดทางสถิติและความน่าเป็นเพื่อเป็นแนวทางในการพัฒนาแอปพลิเคชันสำหรับสร้างรายงาน แบบจำลองนั้นลดขั้นตอนในการสร้างรายงานโดยการวิเคราะห์ชื่อของหัวตารางและนำมาสร้างรายชื่อของชื่อที่เป็นไปได้ทั้งหมดที่สอดคล้องกับข้อมูลที่ผู้ใช้งานมี ชื่อเหล่านั้นจะให้มุมมองเกี่ยวกับข้อมูลแก่ผู้ใช้งาน ทั้งนี้ผู้วิจัยได้พัฒนาเว็บแอปพลิเคชันขึ้นมาตามแบบจำลองเพื่อพิสูจน์ผล โดยระบบสามารถแนะนำชื่อให้กับผู้ใช้งานได้ และมีขั้นตอนในการสร้างรายงานสามขั้นตอนคือ ตั้งค่ากลุ่มธุรกิจ กำหนดข้อมูลสำคัญ และสร้างรายงาน

2. วรรณกรรมที่เกี่ยวข้อง

วิทยาการคอมพิวเตอร์ได้แพร่ขยายไปในศาสตร์หลายแขนง ทั้งในงานวิเคราะห์ข้อมูล งานคำนวณทางคณิตศาสตร์ สำหรับงานวิเคราะห์ข้อมูล วิทยาการคอมพิวเตอร์ถูกนำมาประยุกต์ในฐานะตัวช่วยในการประมวลผลข้อมูล ในอดีต[1] ผู้ใช้งานทำการบันทึกข้อมูลในรูปแบบกระดาษหรือชุดเอกสาร เมื่อต้องการรายงานจะต้องทำการค้นหาเอกสารที่มีข้อมูลสำคัญและนำมาคำนวณค่าต่าง ๆ เช่นค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน ในบางครั้งอาจจำเป็นต้องเลือกใช้วิธีการคำนวณค่าบางอย่างที่ยุ่งยากและซับซ้อน จึงได้มีผู้พัฒนาระบบสำหรับสร้างรายงานขึ้นมา โดยอำนวยความสะดวกในเรื่องการคำนวณข้อมูล แต่ผู้ใช้งานจำเป็นต้องนำข้อมูลที่ได้ไปต่อยอดหรือต้องนำข้อมูลไปประยุกต์ต่อ ผู้พัฒนาจึงเพิ่มความสามารถให้กับเครื่องมือเพื่ออำนวยความสะดวกในการสร้างรูปแบบของรายงาน เครื่องมือมักอยู่ในรูปแบบของแอปพลิเคชัน แพลตฟอร์ม และอื่น ๆ เช่น SAP Crystal Reports[1] JasperReports[2] Pentaho Business Analytics[3] และอีกมากมาย เครื่องมือเหล่านี้มีความสามารถมากมายทั้งในด้านการจัดการข้อมูลขนาดใหญ่ การบริหารจัดการข้อมูลจากหลายแหล่ง การวิเคราะห์ข้อมูล ในยุคที่ระบบเครือข่ายเฟื่องฟู ผู้พัฒนาหลายท่านได้เพิ่มความสามารถให้กับแอปพลิเคชันเพื่อให้สามารถตอบสนองความต้องการของผู้ใช้งานในรูปแบบออนไลน์ เพื่อให้ผู้ใช้งานสามารถเข้าถึงระบบได้จากทุกที่ แต่นั่นไม่ได้หมายความว่าระบบมีความง่ายขึ้น ปัจจุบันได้มีผู้วิจัยหลายท่านได้ค้นคว้าหาวิธีในการสร้างรายงานแบบอัตโนมัติขึ้นมา ซึ่งเป็นศาสตร์ด้านหนึ่งชื่อว่า Knowledge-Based Report Generation

แนวคิดในการสร้างรายงานอัตโนมัติมีแนวคิดเช่นเดียวกับการสร้างชุดแนะนำคำสั่งในการดึงข้อมูลจากฐานข้อมูลด้วยภาษา SQL โดยผู้วิจัยหลายท่าน [2], [3] ได้คิดค้นกระบวนการวิธีในการแนะนำชุดคำสั่งต่าง ๆ ให้กับผู้ใช้งาน ซึ่งจากตัวอย่างเหล่านี้ได้ให้แนวคิดกับงานของผู้วิจัยในหลายด้านทั้งในด้านของการบริหารข้อมูล ความน่าเป็น การเก็บข้อมูลการใช้งานบางส่วนของผู้ใช้งานและอื่น ๆ

ระบบถามตอบอัตโนมัติเป็นอีกหนึ่งแนวความคิดที่ผู้วิจัยได้ให้ความสนใจ [4], [5] หากเปรียบแฟ้มข้อมูลให้เป็นคำถามแล้ว แบบจำลองของผู้วิจัยต้องทำความเข้าใจรูปแบบของแฟ้มข้อมูลเพื่อที่จะตอบคำถามให้กับผู้ใช้งาน

3. วิธีการดำเนินการวิจัย

งานวิจัยชิ้นนี้ผู้วิจัยให้ความสำคัญเรื่องชื่อ โดยได้รับแนวคิดมาจากเทคนิคพื้นฐานสำคัญ ๆ จากศาสตร์ด้านการประมวลผลภาษามนุษย์ในเชิงการทำความเข้าใจภาษามนุษย์ [6] เช่น เทคนิค Labelling Tagging แนวคิด Hidden Markov และอีกมากมายที่ให้ความสำคัญกับเรื่องการแทนข้อมูลด้วยคำ ชื่อ หรือสัญลักษณ์ โดยเล็งเห็นว่าชื่อนั้นเปรียบเสมือนตัวแทนที่บ่งบอกถึงความเป็นตัวตน เปรียบเสมือนตัวแทนที่แสดงให้เห็นถึงลักษณะ รูปแบบ และธรรมชาติของข้อมูล

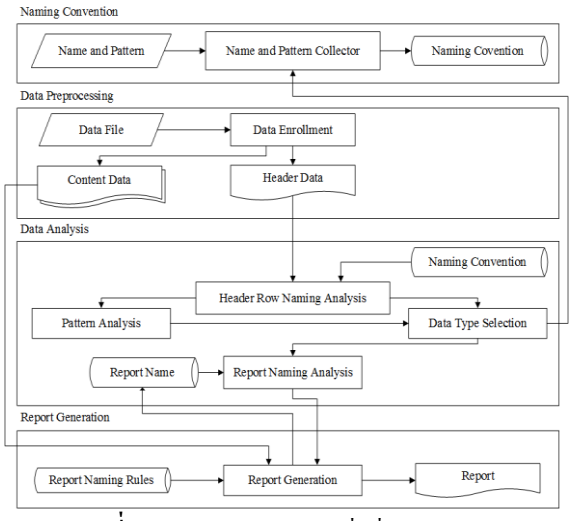
แบบจำลองของผู้วิจัยช่วยอำนวยความสะดวกในการสร้างสารสนเทศในรูปแบบรายงานให้กับผู้ประกอบการโดยการวิเคราะห์ข้อมูลที่มีอยู่แล้วนำเสนอชื่อของรายงานที่เป็นไปได้ทั้งหมดเรียงลำดับตามความเหมาะสมกับข้อมูล

ในขั้นตอนการรวบรวมข้อมูล ผู้วิจัยได้ศึกษาเกี่ยวกับลักษณะของรายงานต่าง ๆ ที่ใช้ในประเทศไทย จุดประสงค์

[1] https://www.sap.com/product/analytics/crystal-reports.html
[2] http://community.jaspersoft.com/project/jasperreports-library
[3] http://www.pentaho.com/product/product-overview

เพื่อให้ทราบถึงโครงสร้างของข้อมูลในรายงานและทำความเข้าใจความสัมพันธ์ระหว่างข้อมูลภายในรายงานกับการตั้งชื่อรายงานเหล่านั้น รายงานที่นำมาศึกษาดาวน์โหลดมาจากเว็บไซต์ www.data.go.th ซึ่งเป็นเว็บไซต์ที่รวบรวมรายงานต่าง ๆ ไว้ โดยได้ทำการเลือกรายงานที่น่าสนใจ[4]มาเป็นจำนวน 50 แฟ้มจากทั้งหมด 873 แฟ้ม โดยครอบคลุมหมวดหมู่การศึกษา คมนาคมและโลจิสติกส์ เศรษฐกิจการเงินและอุตสาหกรรม สังคมและสวัสดิการ เทคโนโลยีสารสนเทศและการสื่อสาร

ผู้วิจัยพบว่าระหว่างชื่อของรายงานกับข้อมูลภายในรายงานมีความสัมพันธ์กันในสองรูปแบบคือ 1. ชื่อและข้อมูลภายในรายงานมีความสอดคล้องกัน เช่น รายงานรายรับประจำปี ซึ่งข้อมูลภายในมักจะประกอบด้วยรายการของรายรับ ยอดเงิน วันเวลาและอื่น ๆ 2. ชื่อและข้อมูลภายในรายงานเกี่ยวข้องกันทางอ้อม เช่น รายงานสำมะโนธุรกิจ ซึ่งภายในมักจะประกอบด้วยข้อมูลธุรกิจเปิดใหม่และแนวโน้มต่าง ๆ ของธุรกิจ และพบว่าชื่อของรายงานนั้นมักจะเกิดจากการรวมกันของคำหลาย ๆ คำที่สื่อถึงประเภทข้อมูล โดยคำเหล่านั้นจะแตกต่างกันออกไปตามหมวดหมู่ของแฟ้มข้อมูล



ภาพที่ 1 แบบจำลองการแนะนำชื่อที่เหมาะสมของรายงาน

3.1 Naming Convention

ตารางที่ 1 ตัวอย่างข้อมูลสำหรับการตรวจสอบประเภทข้อมูล

b_datatype_id	t_nc_pattern
1	(^\d{1,4}(\-|\/|\\)\d{1,2}(\-|\/|\\)\d{1,4}$)
2	(^\d{1,2}(,\d{3})*(\.\d+)?$)

[4] รายงานที่น่าสนใจคือรายงานที่ข้อมูลภายในมีลักษณะเป็นแนวตั้ง ประกอบด้วยหัวตาราง และข้อมูล โดยที่ชื่อจะต้องสอดคล้องกับข้อมูลภายใน

จากตารางที่ 1 จะเป็นตัวอย่างข้อมูลที่ถูกบันทึกในระบบ โดยจะบันทึกการบันทึกชื่อ รูปแบบ และประเภทของข้อมูล สองสิ่งนี้สำคัญในการแนะนำชื่อของรายงานให้กับผู้ใช้งาน

3.1.1 Name and Pattern Collector

จุดประสงค์ของการบันทึกชื่อ รูปแบบ และประเภทของข้อมูลคือการเพิ่มความแม่นยำของการแนะนำชื่อ ยิ่งปริมาณของชื่อ รูปแบบ และประเภทข้อมูลความครอบคลุมของระบบก็จะยิ่งมากขึ้นเท่านั้น

3.2 Data Pre-processing

แบบจำลองของผู้วิจัยจำเป็นต้องใช้ข้อมูลสามส่วนด้วยกันคือ ส่วนชื่อของคอลัมน์ ส่วนข้อมูลในแถวแรก และข้อมูลทั้งหมด ในเบื้องต้นข้อมูลจะถูกแบ่งเป็นสองส่วนคือส่วนสำหรับวิเคราะห์ชื่อที่เหมาะสม และส่วนที่ใช้สำหรับการสร้างรายงาน

3.2.1 Data Enrollment

การแยกข้อมูลเป็นสองส่วนมีขั้นตอนคือข้อมูลในแถวที่หนึ่งและสองคือชุดข้อมูลที่จะใช้ในการแนะนำชื่อรายงาน และข้อมูลในแถวที่สองเป็นต้นไปคือข้อมูลที่ถูกใช้ในการคำนวณเพื่อสร้างรายงาน

3.3 Data Analysis

ขั้นตอนนี้ออกแบบสำหรับหาประเภทของข้อมูล โดยจะใช้ขั้นตอนวิธีการสองวิธีด้วยกันคือการจับคู่ชื่อกับข้อมูลชื่อในฐานข้อมูล หากพบจะนำประเภทข้อมูลที่ได้มาใช้งานและการใช้การเปรียบเทียบรูปแบบข้อมูลโดยใช้ Regular Expression

3.3.1 Header Row Naming Analysis

วิธีการแรกในการตรวจสอบประเภทข้อมูล คือการค้นหาชื่อของแต่ละคอลัมน์จากในฐานข้อมูล หากพบแสดงว่าชื่อนั้นถูกบันทึกประเภทข้อมูลไว้ในฐานข้อมูลแล้ว ยกตัวอย่างเช่น คอลัมน์ที่มีชื่อว่า เดือน โดยคำว่าเดือนเป็นคำที่สื่อถึงข้อมูลประเภทวันที่โดยตรงอยู่แล้ว ดังนั้นเมื่อทำการค้นชื่อคอลัมน์จะได้ผลลัพธ์เป็นข้อมูลประเภทวันที่

Algorithm 1: Header Row Naming Analysis

```
GET naming list AND pattern list FROM naming convention
 GET header row
 DOWHILE header row NOT EOF
   MATCH column WITH naming list
   IF MATCHED THEN RETURN THIS DATA TYPE
   ELSE PERFORM Pattern Analysis
 ENDDO
End Algorithm
```

ขั้นตอนวิธีการที่หนึ่งแสดงให้เห็นวิธีการในการหาประเภทข้อมูลโดยการนำชื่อของคอลัมน์แต่ละคอลัมน์ไปทำการค้นหาในฐานข้อมูล โดยลักษณะของข้อมูลในฐานข้อมูลจะเป็นการบันทึกกลุ่มคำแยกตามประเภทของข้อมูล ตัวอย่างเช่น ข้อมูลประเภทจำนวนจะบันทึกคำจำพวก หน่วย รวม ราคา ชิ้น และอื่น ๆ ดังนั้นหากคอลัมน์มีชื่อว่า หน่วย จะได้ผลลัพธ์จากขั้นตอนนี้คือจำนวน และหากชื่อของคอลัมน์นั้นไม่ตรงกับคำในฐานข้อมูล แบบจำลองจำเป็นจะต้องใช้วิธี Pattern Analysis ซึ่งเป็นอีกวิธีหนึ่งในการค้นหาประเภทของข้อมูล

3.3.2 Pattern Analysis

วิธีการหาประเภทของข้อมูลนี้เป็นการใช้การตรวจจับลักษณะของข้อมูลโดยใช้ Regular Expression ในการตรวจสอบ

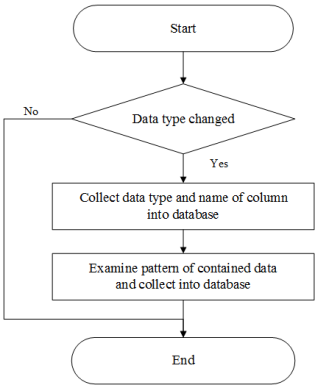
Algorithm 2: Pattern Analysis

```
GET column FROM content row of header row
DOWHILE column NOT EOF
   MATCH column WITH regular expression list
   IF MATCHED THEN RETURN THIS DATA TYPE
   ELSE RETURN NULL
ENDDO
End Algorithm
```

สำหรับขั้นตอนวิธีการที่สองนั้นจะใช้ในกรณีที่ระบบไม่สามารถหาประเภทข้อมูลจากการจับคู่ชื่อได้ โดยข้อมูลที่ถูกนำมาใช้ในขั้นตอนนี้คือข้อมูลในแถวแรกของแฟ้มข้อมูล ข้อมูลจากแต่ละคอลัมน์จะถูกนำมาทดสอบโดยใช้ Regular Expression ที่บันทึกไว้ในฐานข้อมูล ยกตัวอย่างเช่น ข้อมูลประเภทวันที่จะมีลักษณะของ Regular Expression ดังนี้ (^\d{1,4}(\-|\/|\\)\d{1,2}(\-|\/|\\)\d{1,4}$) ในกรณีที่การทดสอบสำเร็จจะได้รับประเภทข้อมูลวันที่ไปใช้งานและจะบันทึกชื่อคอลัมน์เข้าสู่ฐานข้อมูลเพื่อเพิ่มประสิทธิภาพของขั้นตอนวิธีในการตรวจสอบชื่อ ส่วนในกรณีที่ไม่สามารถตรวจสอบข้อมูลได้ ระบบจะไม่ระบุประเภทของข้อมูลโดยที่จะให้ผู้ใช้งานทำการระบุแทน หลังจากนั้นจะบันทึกชื่อและประเภทข้อมูลที่ได้มาจากผู้ใช้งานเข้าสู่ฐานข้อมูล

3.3.3 Data Type Selection

หลังจากผ่านขั้นตอนการจัดเตรียมข้อมูลเบื้องต้นแล้ว แบบจำลองของผู้วิจัยจะให้ผู้ใช้งานทำการตรวจสอบประเภทของข้อมูลอีกครั้ง เพื่อยืนยันว่าแบบจำลองได้ตรวจสอบประเภทข้อมูลได้ถูกต้อง



ภาพที่ 2 ขั้นตอนในตรวจสอบประเภทข้อมูล

ภาพที่ 2 แสดงให้เห็นถึงขั้นตอนในการเลือกประเภทของข้อมูล ในกรณีแบบจำลองตรวจสอบประเภทข้อมูลผิดพลาด ยกตัวอย่างเช่น คอลัมน์นั้นเป็นข้อมูลประเภทวันที่แต่ระบบแสดงประเภทข้อมูลเป็นข้อมูลประเภทอื่น เช่น จำนวน ผู้ใช้งานสามารถทำการเลือกข้อมูลได้และเมื่อผู้ใช้งานทำการเลือกประเภทข้อมูลใหม่ระบบจะทำการบันทึกประเภทข้อมูลและชื่อของคอลัมน์ดังกล่าวเข้าสู่ฐานข้อมูลเพื่อเพิ่มประสิทธิภาพในการตรวจสอบประเภทข้อมูลด้วยชื่อ

3.3.4 Report Naming Analysis

ขั้นตอนนี้เป็นการสร้างรายชื่อที่เหมาะสมกับข้อมูลที่ผู้ใช้งานมี โดยจากการศึกษาแฟ้มที่ได้ดาวน์โหลดมาทั้งหมดนั้นพบว่าชื่อของรายงานคือกลุ่มคำที่สามารถแยกเป็นคำได้ตั้งแต่สองคำเป็นต้นไป ซึ่งในงานวิจัยชิ้นนี้ผู้วิจัยได้ออกแบบขั้นตอนวิธีในการสร้างรายชื่อขึ้นมา โดยชื่อนั้นจะแบ่งเป็นกลุ่มคำสี่กลุ่มด้วยกัน คือ 1. กลุ่มคำนำหน้า 2. กลุ่มคำกลางลำดับแรก 3. กลุ่มคำกลางลำดับหลัง 4. กลุ่มคำลงท้าย ยกตัวอย่างเช่น รายงานชื่อ "รายงานค่าเฉลี่ยรายได้แยกตามเดือน" จะสามารถแบ่งได้ดังนี้ รายงานคือคำเริ่มต้น ค่าเฉลี่ยคือคำขยาย รายได้คือคำหลัก และแยกตามเดือนคือคำลงท้าย ซึ่งหากอ่านตามชื่อก็จะพบว่าชื่อนี้สื่อถึงการคำนวณค่าเฉลี่ยของคอลัมน์ที่เกี่ยวกับเงินแล้วทำการแสดงผลลัพธ์แยกตามคอลัมน์ที่เกี่ยวกับวันที่ ทั้งนี้ชื่ออาจไม่ได้ประกอบจากกลุ่มคำสี่กลุ่มเสมอ โดยกลุ่มคำขึ้นอยู่กับความเหมาะสมของคำในกลุ่มคำกับชื่อของรายงาน

ชื่อของรายงานแต่ละชื่อนั้นจะถูกบันทึกคะแนนไว้เพื่อใช้ในการเรียงลำดับ ชื่อที่มีคะแนนมากแสดงว่าถูกใช้งานโดย

ผู้ใช้งานมากกว่าชื่อที่มีคะแนนน้อย โดยหลังจากชื่อที่ระบบแนะนำให้ถูกนำไปใช้งาน ระบบจะเพิ่มคะแนนหนึ่งคะแนนให้กับชื่อนั้นเพราะชื่อนั้นเป็นที่สนใจสำหรับผู้ใช้งาน

ตารางที่ 2 ตัวอย่างข้อมูลของคำที่บันทึกในฐานข้อมูล

b_naming_position	t_rn_materials
1	{รายได้, รายจ่าย, ...}
1	{การยืม, การคืน, การเช่า , ...}
1	{ความพึงพอใจ, การเข้าเรียน, ...}
2	{ค่าเฉลี่ย, ผลรวม, ...}

Algorithm 3: Name of report generation
GET selected data type
GET b_datatype_id = selected data type FROM DATABASE
DOWHILE b_datatype_id NOT EOF
APPEND t_rn_materials AND weighted TO result
ENDDO
RETURN result
End Algorithm

ตารางที่ 2 แสดงให้เห็นตัวอย่างของคำที่จะใช้ในการสร้างชื่อของรายงาน สำหรับขั้นตอนวิธีที่สามนั้นจะเป็นกระบวนการในการสร้างชื่อโดยจะเริ่มจากผู้ใช้งานทำการเลือกคอลัมน์ที่เขาสนใจหลังจากนั้นระบบจะนำประเภทข้อมูลเหล่านั้นไปทำการค้นหาชื่อในฐานข้อมูลและทำการสร้างรายชื่อที่เป็นไปได้ทั้งหมดโดยชื่อจะถูกเรียงชื่อตามคะแนน

$$P(X) = P(X_{pre}) * P(X_{mid-pre}) * P(X_{mid-post}) * P(X_{post})$$

จำนวนของชื่อเกิดจากความน่าจะเป็นของแต่ละกลุ่มคำคูณกัน

$$P(X_{ep}) = \sum_{n=1}^{6} P(W)_n$$

โดยความน่าเป็นภายในกลุ่มคำจะเท่ากับจำนวนคำที่สามารถหาได้จากฐานข้อมูลโดยจะแยกตามประเภทข้อมูลซึ่งแบบจำลองของผู้วิจัยจะประกอบด้วยข้อมูล 6 ประเภทดังนี้ วันที่ จำนวน ที่อยู่ ตัวอักษร คำอธิบาย ตัวเลขคำนวณไม่ได้ โดยประเภทของข้อมูลถูกกำหนดจากการศึกษาประเภทข้อมูลที่พบในรายงานที่ทำการศึกษา

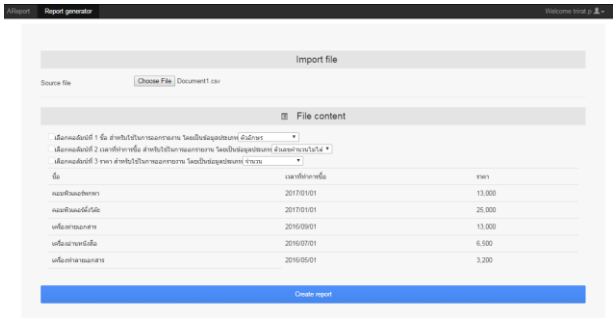
3.4 Report Generation

หลังจากระบบได้รับชื่อของรายงานแล้ว ระบบจะแยกชื่อเป็นกลุ่มคำแล้วนำไปค้นหากระบวนการในการสร้างรายงาน เช่น ผู้ใช้งานเลือกรายงานชื่อ “รายงานรายจ่ายประจำปี 2560 แยกตามเดือน” ชื่อนี้จะถูกแยกเป็น รายงาน(คำนำหน้า) รายจ่าย(คำกลางลำดับแรก) ประจำปี 2560(คำกลางลำดับหลัง) แยกตามเดือน(คำลงท้าย) โดยคำที่แยกแล้วจะนำไปค้นหาวิธีการสร้างรายงานจากฐานข้อมูล

4. ผลการดำเนินงาน

เพื่อพิสูจน์แบบจำลองว่าสามารถทำงานได้อย่างถูกต้อง ผู้วิจัยได้พัฒนาเว็บแอปพลิเคชัน[5]สำหรับสร้างรายงานขึ้นมา โดยเว็บไซต์จะมีลักษณะการทำงานแบบ Server-Client การคำนวณสำคัญ ๆ ของแอปพลิเคชันจะถูกคำนวณที่ฝั่ง Client ซึ่งพัฒนาโดยภาษา JavaScript เพราะฉะนั้นข้อมูลสำคัญ ๆ จะไม่ถูกอัพโหลดเข้าสู่ Server เว็บไซต์ประกอบด้วยส่วนจัดการข้อมูลส่วนตัว ส่วนการสร้างรายงาน โดยแอปพลิเคชันจะทำงานบนฐานข้อมูลซึ่งมีตารางสำคัญดังนี้ ตาราง b_datatype ตาราง b_datatype_pattern ตาราง b_datatype_word สำหรับเก็บข้อมูลประเภทข้อมูล รูปแบบของประเภทข้อมูลนั้น รวมทั้งคำสำคัญที่จะใช้ในการตรวจสอบประเภทข้อมูล ตาราง b_user สำหรับตั้งค่าผู้ใช้งาน ตาราง t_report_name สำหรับบันทึกและให้คะแนนชื่อของรายงานที่ถูกใช้งาน และตารางอื่น

ก่อนใช้งานระบบผู้ใช้งานต้องสมัครสมาชิกก่อน เนื้อหาของการสมัครสมาชิกจะเกี่ยวกับข้อมูลส่วนบุคคลและประเภทของธุรกิจที่กำลังทำ โดยประเภทของธุรกิจจะใช้ในการเลือกคำที่จะใช้ในการแนะนำชื่อของรายงาน สำหรับเวอร์ชั่นนี้ ผู้วิจัยได้รวบรวมข้อมูลคำเฉพาะกลุ่มธุรกิจด้านการศึกษาโดยจะครอบคลุมในเรื่อง การเรียนเช่น เข้าเรียน ส่งชิ้นงาน วัสดุเช่น การยืม การคืน และการประเมินผลต่าง ๆ

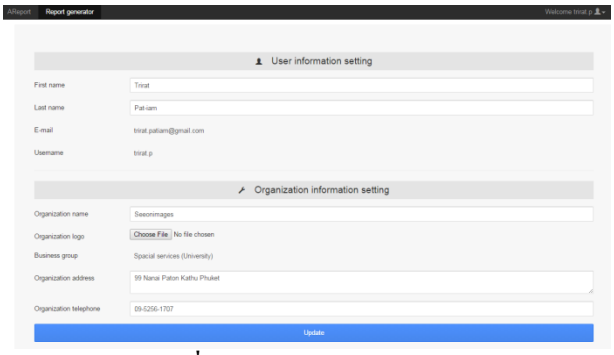


ภาพที่ 3 ลักษณะหน้าจอหลักของระบบ

ผู้ใช้งานทำการเลือกแฟ้มข้อมูลที่จะใช้สร้างรายงานโดยแฟ้มข้อมูลจะต้องอยู่ในรูปแบบของ CSV หรือ Excel หลังจากเลือกแฟ้มข้อมูล ระบบจะทำการเตรียมข้อมูลเบื้องต้นโดยการอ่านค่าข้อมูลตามแบบจำลองที่ผู้วิจัยได้ออกแบบไว้และ

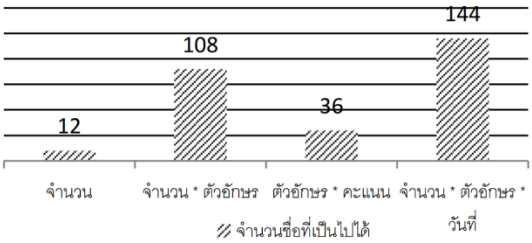
[5] www.areports.seeonimages.com (กำลังพัฒนาความสามารถเพื่อใช้งานจริง ซึ่งอาจส่งผลให้ทำงานได้ไม่ตลอด 24 ชั่วโมง)

แสดงผลจากการเตรียมข้อมูลบนหน้าจอของระบบเพื่อให้ผู้ใช้งานทำการเลือกคอลัมน์ที่จะใช้ในการสร้างรายงาน



ภาพที่ 4 หน้าจอการตั้งค่าข้อมูลส่วนตัว

ในส่วนการตั้งค่าข้อมูลส่วนตัวส่วนสำคัญที่สุดคือ Business group ซึ่งจะถูกนำไปใช้ในการเลือกคำสำหรับแนะนำชื่อรายงาน สำหรับการพัฒนาในเวอร์ชั่นแรกฐานข้อมูลคำจะรองรับเฉพาะหน่วยงานด้านการศึกษา



ภาพที่ 5 ตัวอย่างผลลัพธ์ของชื่อที่เป็นไปได้จากระบบโดยการเลือกประเภทข้อมูลแต่ละประเภท

5. สรุป

จากปัญหาในขั้นตอนการสร้างรายงาน เช่น การใช้เวลาศึกษาธรรมชาติข้อมูล การทำความเข้าใจจุดประสงค์ของรายงาน การมีมุมมองต่อข้อมูลว่าสามารถนำไปประยุกต์อะไรได้บ้าง ผู้วิจัยจึงได้คิดค้นแบบจำลองที่ช่วยในการสร้างรายงานโดยให้มุมมองเกี่ยวกับข้อมูลให้กับผู้ใช้งานและลดเวลาที่ผู้ใช้งานต้องใช้ไปกับการศึกษาข้อมูล แบบจำลองใช้หลักความน่าจะเป็นและสถิติในการสร้างรายชื่อของชื่อที่เป็นไปได้ทั้งหมดจากข้อมูลที่ผู้ใช้งานมีโดยการวิเคราะห์ประเภทของข้อมูลจากหัวตารางและสร้างรายชื่อโดยการผสมกลุ่มคำตามความน่าเป็น ชื่อเหล่านั้นจะถูกเรียงตามลำดับคะแนน คะแนนที่มีค่ามากแสดงให้เห็นว่าชื่อนั้นมีความนิยมสูง มีผู้ใช้งานจำนวนมากเลือกใช้ในการสร้างรายงาน ส่วนคะแนนที่มีค่าน้อยแสดงให้เห็นว่าชื่อรายงานนั้นถูกใช้ในกลุ่มที่จำกัด จากการพัฒนาเว็บแอปพลิเคชันตามแบบจำลองที่ได้ออกแบบไว้พบว่าระบบสามารถแนะนำชื่อให้กับผู้ใช้งานได้ และสามารถเรียงชื่อตามความนิยมได้ ส่งผลให้ผู้ใช้งานใช้เวลาน้อยลงในการสร้างรายงาน ทั้งนี้ระบบที่ทำการพัฒนายังรองรับประเภทธุรกิจแค่ประเภทเดียวคือ ธุรกิจพิเศษประเภทการศึกษา สาเหตุที่ระบบรองรับแค่หนึ่งประเภทธุรกิจเพราะกลุ่มคำที่ใช้ในธุรกิจแต่ละประเภทมีความแตกต่างกัน สำหรับงานวิจัยในอนาคต ผู้วิจัยจะทำการรวบรวมข้อมูลกลุ่มคำและพัฒนาพจนานุกรมคำที่ใช้ในการสร้างรายงานขึ้นมา และจะพัฒนากลุ่มคำเพื่อให้รองรับกับประเภทธุรกิจที่หลากหลายมากยิ่งขึ้น

เอกสารอ้างอิง

[1] Carlos Alonso Vega. “Java and reports, some solutions for the past, present and future”. *Proceedings of the 29th annual ACM SIGUCCS conference on User services (SIGUCCS '01)*. ACM, New York, NY, USA, 275-278. 2001.

[2] Veronika Thost, Konrad Voigt, and Daniel Schuster. “Query matching for report recommendation”. *Proceedings of the 22nd ACM international conference on Conference on information & knowledge management (CIKM '13)*. ACM, New York, NY, USA, 1391-1400. 2013.

[3] Matthias Nicola and Tim Kiefer. “Generating SQL/XML query and update statements”. *Proceedings of the 18th ACM conference on Information and knowledge management (CIKM '09)*. ACM, New York, NY, USA, 1187-1196. 2009.

[4] Valentin Jijkoun and Maarten de Rijke. “Retrieving answers from frequently asked questions pages on the web”. *Proceedings of the 14th ACM international conference on Information and knowledge management (CIKM '05)*. ACM, New York, NY, USA, 76-83. 2005.

[5] Anna Shtok, Gideon Dror, Yoelle Maarek, and Idan Szpektor.” Learning from the past: answering new questions with past answers”. *Proceedings of the 21st international conference on World Wide Web (WWW '12)*. ACM, New York, NY, USA, 759-768. 2012.

[6] Steven Bird, Ewan Klein, Edward Loper, *Natural Language Processing with Python*, O'Reilly Media, 2009-

ระบบรับชำระสินค้าด้วย NFC และ RFID: กรณีรับชำระสินค้าในสหกรณ์โรงเรียน
Payment System using NFC and RFID: A Case in School Cooperatives

พิชญาภา มาประจวบ (Pichayapa Maprajuab)[1] ธนวิทย์ ทัตตะกิตติยา (Tanawit Tattakittiya)[2]
สุธาสินี เอี่ยมสอาด (Suthasinee Iamsa-at)[3] และคำรณ สุนัติ (Khamron Sunat)[4]
ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น
[1]Imfame2538@gmail.com, [2]Tanawitt@kkumail.com, [3]Suthasinee@kkumail.com, [4]Khamron_sunat@yahoo.com

บทคัดย่อ

ระบบรับชำระค่าสินค้าในสหกรณ์โรงเรียน พัฒนาบนระบบปฏิบัติการแอนดรอยด์และเว็บไซต์ โดยการนำระบบฐานข้อมูลเข้ามาช่วยในการจัดเก็บข้อมูลให้มีความปลอดภัย มีการนำเทคโนโลยี RFID มาใช้ร่วมกับสมาร์ทโฟนที่รองรับเทคโนโลยี NFC ทางการเงิน โดยรับชำระค่าสินค้าจากชิป RFID ที่ฝังอยู่ในบัตรนักเรียน ซึ่งนักเรียนและผู้ปกครองสามารถติดตามรายการใช้จ่ายและยอดคงเหลือ ที่สำคัญไปกว่านั้นในส่วนของผู้ปกครองยังสามารถกำหนดวงเงินใช้จ่ายของนักเรียนผ่านเว็บไซต์ ผลที่ได้จากการพัฒนาระบบนี้จะช่วยเพิ่มความสะดวกในการชำระค่าสินค้าของนักเรียน ผู้ปกครองสามารถตรวจสอบจำนวนเงินรายการใช้จ่าย ซึ่งนับเป็นการควบคุมค่าใช้จ่ายและโภชนาการของเด็กไปในตัว อีกทั้งในส่วนของสหกรณ์โรงเรียนรวมถึงผู้บริหารนั้นสามารถนำข้อมูลรายการซื้อขายมาวิเคราะห์และวางแผนการสั่งซื้อสินค้าในครั้งถัดไปได้

คำสำคัญ: N NFC RFID สมาร์ทโฟน บัตรนักเรียน สหกรณ์โรงเรียน

Abstract

Payment system in cooperative school is developed on android and website. The data is kept by database systems to increase security. RFID technology is used in conjunction with smartphones that support NFC technology to conduct financial transactions. Receiving payments from RFID chips are embedded in a student card. Students and parents can keep track of spending and balances. More importantly, parents can set their spending limits through the website. The result of this system enhances the convenience of paying for the student's goods. Parents can check the amount paid item. This is a way to control the cost and nutrition that the children spend and receive. Also, the school cooperative administrators can use the transaction information to analyze and plan the goods orders in the future.

Keyword: NFC, RFID, Smart Phone, Student Card, School Cooperative.

1. บทนำ

ในปัจจุบันระบบการชำระค่าสินค้าของโรงเรียนเป็นแบบรับมาจ่ายไป เป็นเหตุทำให้นักเรียนต้องต่อแถวรอคิวยาว เพื่อรอรับเงินทอน บางครั้งนักเรียนพกเงินหรือมีเงินในกระเป๋าไม่เพียงพอกับค่าสินค้า การทอนเงินผิดก็ส่งผลให้นักเรียนเสียผลประโยชน์และจากการที่ไม่สามารถตรวจสอบการใช้จ่ายย้อนหลังได้ ส่งผลให้ผู้ปกครองไม่ทราบว่านักเรียนใช้จ่ายสินค้าใดบ้างในแต่ละวัน และปัญหาที่พบบ่อยคือ เด็กนักเรียนทำเงินหายระหว่างวันและเด็กบางคนไม่สามารถวางแผนการใช้เงินในแต่ละวันได้

จากปัญหาดังกล่าว ผู้จัดทำจึงสร้างระบบชำระค่าสินค้าด้วย NFC และ RFID : กรณีรับชำระค่าสินค้าในสหกรณ์โรงเรียน โดยการนำเอาเทคโนโลยี RFID และ NFC ในสมาร์ทโฟนมาช่วยอำนวยความสะดวกในการซื้อสินค้าภายในร้านค้าสหกรณ์ เพื่อเพิ่มประสิทธิภาพในการขาย ณ จุดขาย และจุดขายในแต่ละร้านค้าให้มีความถูกต้อง แม่นยำ รวดเร็วขึ้น พร้อมทั้งมีรายงานเพื่อแสดงให้ทางโรงเรียนได้ทราบถึงการเคลื่อนไหวค่าใช้จ่ายซื้อสินค้าของนักเรียน เพื่อนำมาวิเคราะห์และวางแผนการใช้จ่ายของนักเรียน นับเป็นการควบคุมโภชนาการของนักเรียนไปในตัว ซึ่งผู้ปกครองมีส่วนร่วมในการใช้เพื่อกำหนดค่าใช้จ่ายให้กับนักเรียนและตรวจสอบการใช้จ่ายของนักเรียนได้ อีกทั้งทางสหกรณ์โรงเรียนมีความเหมาะสมในเรื่องของระบบ

การเงิน ได้รับความถูกต้องในการรับชำระค่าสินค้า ได้เช็คยอดสินค้าในการขายได้ดีขึ้น

2. วรรณกรรมและงานวิจัยที่เกี่ยวข้อง

2.1 เทคโนโลยี NFC

สนามสื่อสารระยะใกล้ (Near Field Communication: NFC) ถูกพัฒนาเพื่อช่วยให้การสื่อสารระหว่างอุปกรณ์ [1] เพียงนำสมาร์ทโฟนไปใกล้กับเครื่องอ่านหรือแถบป้ายข้อมูลแสดงตัวตนทำให้สามารถส่งข้อมูลระหว่างกันได้ทันที แถบป้ายข้อมูล NFC (NFC Tag) จะมีอุปกรณ์อ่านและแปลงข้อมูลเก็บไว้ในรูปแบบข้อมูลที่แลกเปลี่ยนได้ จากความสะดวกสบายในการใช้งานดังกล่าวทำให้บริษัทผู้ผลิตมือถือสมาร์ทโฟนได้นำเทคโนโลยี NFC ไปบรรจุไว้ในสมาร์ทโฟนรุ่นมาตรฐานสูงสุดเพื่ออำนวยความสะดวกให้กับผู้ใช้ในการสื่อสารระหว่างอุปกรณ์ด้วยกัน ในปัจจุบันได้มีการประยุกต์เทคโนโลยี NFC ไปใช้ในหลากหลายด้าน เช่น ระบบการชำระเงินในร้านสะดวกซื้อ ระบบการขายตั๋วของบริษัทการบินต่าง ๆ ซึ่งทำให้สะดวกรวดเร็ว และลดการเข้าคิว [2]

ต้นแบบของระบบการจำหน่ายตั๋วเรือเฟอร์รี่ [3] โดยผู้โดยสารซื้อตั๋วออนไลน์ผ่านแอปพลิเคชัน จากนั้นจะจัดเก็บข้อมูลการสั่งซื้อตั๋วไว้ในระบบ เมื่อผู้โดยสารต้องการขึ้นเรือต้องหยุดตรวจตั๋ว โดยนำสมาร์ทโฟนที่รองรับเทคโนโลยี NFC ไปสัมผัสกับเครื่องอ่าน NFC ณ จุดตรวจ จากนั้นระบบจะทำการตรวจสอบว่าผู้โดยสารนั้นได้ทำการสั่งซื้อตั๋วหรือยัง ถ้าสั่งซื้อแล้วทางนั้นจะเปิดพร้อมส่งข้อความยืนยันมาที่โทรศัพท์มือถือของผู้โดยสาร

2.2 Contactless Payment

ปัจจุบันมีการนำเทคโนโลยีที่ทันสมัยเข้ามาช่วยในการจัดเก็บข้อมูล ทำธุรกรรมทางการเงิน จึงมีการศึกษาการทำงานของเทคโนโลยีหรือเครื่องมือที่ใช้ในการชำระค่าสินค้ามาใช้ในสหกรณ์โรงเรียน ซึ่งเป็นเทคโนโลยีการจ่ายเงินด้วยการแตะสัมผัสเครื่องโทรศัพท์เป็นเทคโนโลยีลักษณะเดียวกับบัตรเครดิตหรือบัตรสมาชิกที่ฝังชิปตระกูล RFID ไว้ภายใน ทำให้เราสามารถซื้อของต่างๆได้ โดยนำบัตรนักเรียนที่มี RFID Tag ไปสัมผัสกับสมาร์ทโฟนที่รองรับการจ่ายเงินด้วยระบบ NFC [4] เช่นระบบ AIS mPAY Rabbit [5]

2.2.1 เทคโนโลยีอาร์เอฟไอดี (RFID)

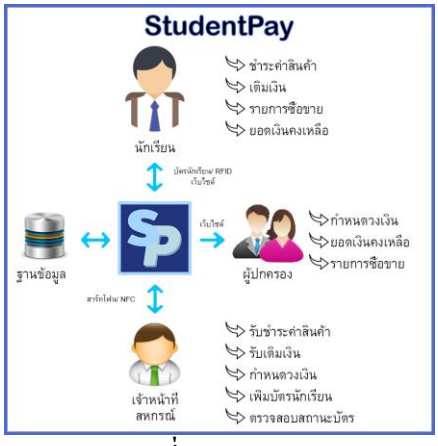
ในปัจจุบันเทคโนโลยีบ่งชี้อัตโนมัติ (Auto-ID: automatic identification) ได้เข้ามามีบทบาทสำคัญในการดำเนินชีวิตประจำวัน สังเกตได้จากการนำมาประยุกต์ใช้งานในหลายด้าน โดยมีวัตถุประสงค์เพื่อใช้แสดงตัวตนของมนุษย์, สินค้า และวัตถุดิบในกระบวนการผลิต โดยทั่วไปเทคโนโลยีบ่งชี้อัตโนมัติแบบต่าง ๆ ถูกคิดค้นขึ้นมาเพื่ออำนวยความสะดวกในการทำธุรกิจโดยเฉพาะอย่างยิ่งในการบันทึกข้อมูลแบบอัตโนมัติอย่างรวดเร็ว แทนการนับหรือจดบันทึกด้วยมือ ซึ่งอาจเกิดข้อผิดพลาดได้ง่าย ระบบต่าง ๆจึงได้อาศัย Radio frequency identification-RFID Tag บนบัตรระบุตัวตนเข้ามาร่วมใช้งานแอปพลิเคชันซอฟต์แวร์[6] ใช้ในการยืนยันตัวตนในการเข้ารับบริการและธุรกรรมต่าง ๆ

สำนักงานทรัพยากรการเรียนรู้คุณหญิงหลง อรรถกระวีสุนทร มหาวิทยาลัยสงขลานครินทร์ ได้ศึกษาระบบประตูทางเข้าอัตโนมัติ (Smart gate) [7] เพื่อเก็บข้อมูลด้านสถิติของการเข้าใช้บริการของสมาชิกห้องสมุด โดยอาศัย Sensor เป็นอุปกรณ์สำคัญในการตรวจนับและในการใช้บัตรยืนยันตัวตน โดยต้องนำบัตรเหล่านั้นไปใช้กับหัวอ่าน RFID หรือหัวอ่าน BARCODE แล้วจะมีการตรวจสอบว่าบัตรนั้นเป็นของผู้ถือบัตรหรือไม่ โดยมีรูปถ่ายใบหน้าปรากฏทุกครั้งเมื่อมีการยืนยันตัวตนและจะมีการบันทึกสถิติของผู้เข้าใช้

2.2.2 บาร์โค้ด (Barcode)

บาร์โค้ด (Barcode) เป็นรหัสแท่ง ประกอบด้วยเส้นมืดและเส้นสว่างวางเรียงกันเป็นแนวดิ่งเป็นรหัสแทนตัวเลขและตัวอักษรในลักษณะแท่งสีดำบนพื้นขาว(ช่องว่าง) ซึ่งแท่งที่เรียงกันเหล่านี้สามารถแปลเป็นตัวเลขหรืออักษรด้วยเครื่องอ่าน [8] ช่วยให้อ่านรหัสข้อมูลได้ง่ายขึ้น ซึ่งจะทำงานได้รวดเร็ว ช่วยลดความผิดพลาดในการคีย์ข้อมูลและมีความน่าเชื่อของข้อมูลสูง สร้างความสะดวกให้แก่ผู้ใช้งาน [9] เช่น การติดบาร์โค้ดลงบนสินค้า, วัสดุครุภัณฑ์ ฯลฯ

จากการศึกษาค้นคว้างานวิจัยที่เกี่ยวข้อง ผู้วิจัยได้กรอบแนวคิดในการวิจัย ดังภาพ

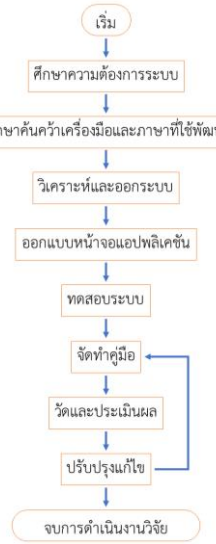


ภาพที่ 1: แผนการทำงาน

3. วิธีดำเนินงานวิจัย

3.1 การรวบรวมศึกษารวบรวมข้อมูลจากผู้ที่เกี่ยวข้องและระบบการจ่ายสินค้าของนักเรียนและสหกรณ์โรงเรียน

ขั้นตอนการดำเนินงานวิจัยของระบบ ดังภาพที่ 2

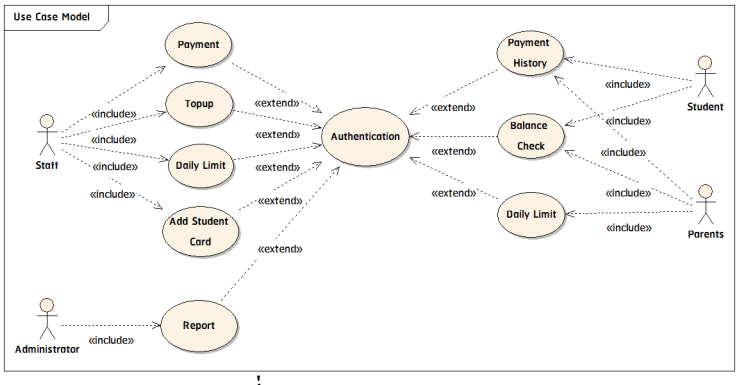


ภาพที่ 2: Flow chart ของระบบ

3.2 การวิเคราะห์และออกแบบระบบ

ระบบแบ่งผู้ใช้เป็น 4 ส่วน โดยในส่วนแรกเป็นส่วนของพนักงาน ทำงานบนแอพพลิเคชันที่ถูกติดตั้งบนสมาร์ทโฟนที่รองรับเทคโนโลยี NFC พนักงานมีหน้าที่รับชำระค่าสินค้า ณ จุดขาย จัดการข้อมูลนักเรียน ตรวจสอบสถานะบัตร รับเติมเงิน กำหนดวงเงินตามที่ผู้ปกครองต้องการให้กับเด็กนักเรียน และมีการจัดการคลังสินค้าบนเว็บไซต์ ส่วนที่ 2 ส่วนของนักเรียน สามารถดูรายการซื้อขายและตรวจสอบยอดเงินคงเหลือผ่านเว็บไซต์ ส่วนที่ 3 ส่วนของผู้ปกครอง สามารถดูรายการซื้อขาย ตรวจสอบยอดเงินคงเหลือและกำหนดวงเงินให้กับนักเรียน และส่วนที่ 4 ผู้บริหารสามารถดูสรุปรายการซื้อขาย กำไร ต้นทุน ของสินค้าแต่ละรายการบนเว็บไซต์ โดยการทำงานทั้งหมดจะต้องมีการเข้าสู่ระบบเพื่อยืนยันตัวตนก่อน

แผนภาพแสดงการใช้งานแอปพลิเคชันชำระค่าสินค้าในสหกรณ์โรงเรียน ดังภาพที่ 3

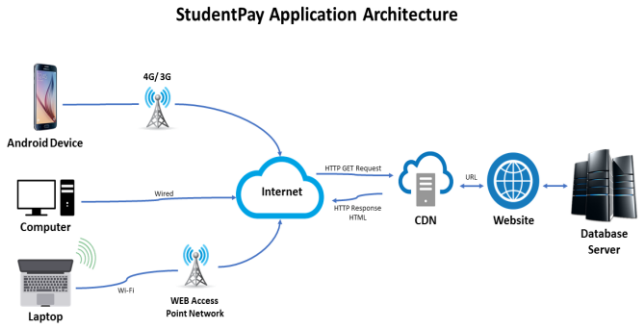


ภาพที่ 3: Use case ของระบบ

3.3 การพัฒนาระบบ

วิธีดำเนินงานการพัฒนาระบบนั้นได้ดำเนินการโดยแบ่งการพัฒนาเป็น 2 ส่วน ในส่วนแรกเป็นการพัฒนาแอปพลิเคชันใช้โปรแกรม Android Studio ในการเขียนคำสั่งและออกแบบหน้าจอแอปพลิเคชัน มีการใช้เทคโนโลยี NFC ที่อยู่บนสมาร์ทโฟนมาร่วมกับเทคโนโลยี RFID บนบัตรนักเรียน และส่วนที่ 2 เป็นการพัฒนาเว็บไซต์ ใช้โปรแกรม Sublime และ NetBean ในการออกแบบหน้าเว็บไซต์ การทำงานของเว็บไซต์เป็นไฟล์ข้อมูลชนิด HTML และใช้ PHP มีโปรแกรม phpMyAdmin ช่วยในการจัดการระบบฐานข้อมูล MySQL [10]

3.4 สถาปัตยกรรมของระบบ



ภาพที่ 4: สถาปัตยกรรมของระบบ

การทำงานของระบบจะเริ่มต้นจากไคล์เอนท์ส่งคำร้องเพื่อต้องการเข้าเว็บแอปพลิเคชัน เมื่อเว็บเซิร์ฟเวอร์ได้รับการร้องขอจะทำการเรียกไฟล์ที่ถูกร้องขอ จะนำข้อมูลบนดาต้าเบสเซิร์ฟเวอร์มาประมวลผล เมื่อประมวลผลเรียบร้อยจะส่งผลลัพธ์กลับไปยังไคล์เอนท์ โดยไคลเอนท์และเซิร์ฟเวอร์เชื่อมต่อกันด้วยอินเตอร์เน็ต

4. ผลการดำเนินงานวิจัย

ในส่วนของผลการดำเนินงานวิจัยแบ่งออกเป็น 2 ส่วน คือ ผลการพัฒนาระบบและผลการประเมินระบบ ดังนี้

4.1 ผลการพัฒนาระบบ

4.1.1 สิทธิ์การใช้งานของพนักงาน



ภาพที่ 5: หน้าแรกของแอปพลิเคชันในส่วนของพนักงาน

จากภาพที่ 5 หน้าจอ Login สำหรับพนักงาน เป็นการยืนยันตัวตนเพื่อเข้าใช้งาน มีการเก็บข้อมูลของพนักงานแต่ละบุคคลทั้งเป็นข้อมูลส่วนตัวและข้อมูลสำหรับติดต่อสื่อสารซึ่งการยืนยันตัวตนนี้จะเข้าสู่เมนูทั้งหมดของแอปพลิเคชัน ซึ่งสามารถแสดงได้ดังภาพที่ 6



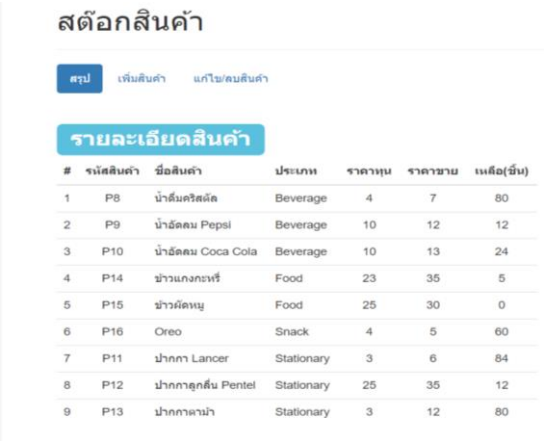
ภาพที่ 6: หน้าจอเมนูทั้งหมดของแอปพลิเคชัน

จากภาพที่ 6 เลือกเมนูได้ 4 เมนู คือเติมเงิน (Top up) ชำระสินค้า (Payment) กำหนดวงเงิน (Daily Limit) เพิ่มบัตรนักเรียน (Add Student) หากเลือกเมนูเติมเงินจะปรากฏหน้าจอดังภาพที่ 7 เจ้าหน้าที่มีหน้าที่รับเติมเงินเข้าระบบให้กับเด็ก โดยนักเรียนหรือผู้ปกครองนำบัตรและเงินมาเติมที่จุดบริการ โดยจะกดจำนวนเงินที่ขึ้นให้หรือกรอกจำนวนเงินตามที่ต้องการ



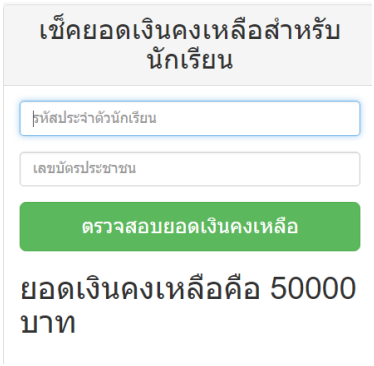
ภาพที่ 7: หน้าจอเมนูเติมเงินของแอปพลิเคชันในส่วนของพนักงาน

พนักงานสามารถตรวจสอบข้อมูลทั้งหมดของสินค้าและจำนวนคงเหลือของรายการสินค้าที่ใกล้หมดหรือสินค้าที่หมดแล้วบนเว็บไซต์ เพื่อนำไปวางแผนการสั่งซื้อสินค้าถัดไป แสดงได้ดังภาพที่ 8



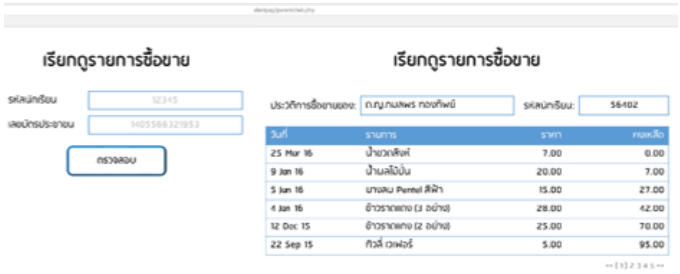
ภาพที่ 8: ข้อมูลสินค้าทั้งหมดในคลังสินค้า

4.1.2 สิทธิ์การใช้งานของนักเรียนและผู้ปกครอง



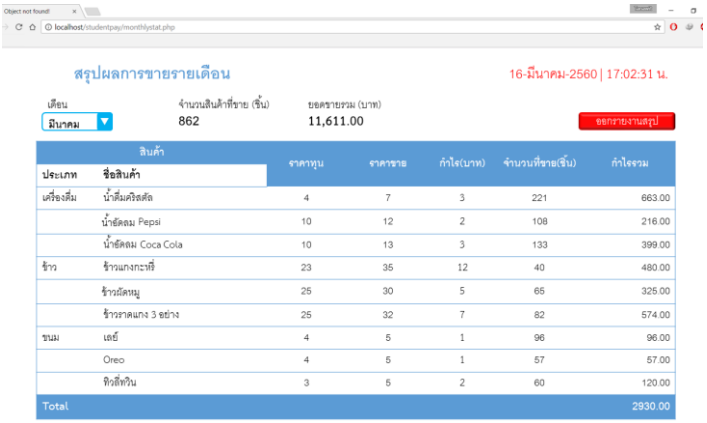
ภาพที่ 9: แสดงภาพหน้าจอเมื่อทำการเช็คยอดเงินคงเหลือบนเว็บไซต์

จากภาพที่ 9 เมื่อนักเรียนหรือผู้ปกครองกรอกรหัสนักเรียนและรหัสผ่านถูกต้อง เว็บไซต์จะแสดงจำนวนยอดเงินคงเหลือและรายการใช้จ่ายของนักเรียนบนเว็บไซต์ ดังภาพที่ 10



ภาพที่ 10: ภาพแสดงรายการซื้อขายในส่วนของนักเรียนบนเว็บไซต์

4.1.3 สิทธิ์การใช้งานของผู้บริหาร



ภาพที่ 11: ภาพแสดงหน้าสรุปยอดขายต่อเดือน

จากภาพที่ 11 ผู้บริหารสามารถดูสรุปรายการซื้อขาย กำไร ต้นทุน ของสินค้าแต่ละรายการบนเว็บไซต์

4.2 ผลการประเมินระบบ

จำแนกหัวข้อในการประเมิน 4 ด้าน ดังนี้ ด้านการตรงตามความต้องการของผู้ใช้ระบบ (Functional Requirement Test) ด้านการทำงานได้ตามฟังก์ชันงานของระบบ (Functional Test) ด้านความง่ายต่อการใช้งานระบบ (Usability Test) และด้านการรักษาความปลอดภัยของข้อมูลในระบบ (Security Test)

ตารางที่ 1 ผลการประเมินระบบด้านความสามารถในการทำงานของระบบ มีความพึงพอใจในระดับดี ซึ่งมีค่าเฉลี่ยเท่ากับ 3.87 และค่าส่วนเบี่ยงเบนมาตรฐาน 0.35 ผลการประเมินระบบด้านการทำงานตามหน้าที่ระบบ มีความพึงพอใจในระดับดี ซึ่งมีค่าเฉลี่ยเท่ากับ 3.54 และค่าส่วนเบี่ยงเบนมาตรฐาน 0.28 ผลการประเมินระบบด้านการใช้งานของโปรแกรมมีความพึงพอใจในระดับดี ซึ่งมีค่าเฉลี่ยเท่ากับ 3.87 และค่าส่วนเบี่ยงเบนมาตรฐาน 0.15 ผลการประเมินระบบด้านการรักษาความปลอดภัยของระบบ มีความพึงพอใจในระดับดี ซึ่งมีค่าเฉลี่ยเท่ากับ 3.64 และค่าส่วนเบี่ยงเบนมาตรฐาน 0.28

ระดับความพึงพอใจของระบบโดยรวมแบ่งตามสิทธิ์ผู้ใช้งาน 3 ระดับ พบว่าโปรแกรมทำงานตามสิทธิ์ใช้งานของเจ้าหน้าที่มีระดับความพึงพอใจซึ่งมีค่าเฉลี่ยเท่ากับ 4.20 และค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.45 โปรแกรมทำงานตามสิทธิ์ใช้งานของนักเรียนมีระดับความพึงพอใจซึ่งมีค่าเฉลี่ยเท่ากับ 3.80 และค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.45 และโปรแกรมทำงาน

ตามสิทธิ์ใช้งานของนักเรียนมีระดับความพึงพอใจซึ่งมีค่าเฉลี่ยเท่ากับ 3.40 และค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.89

ตารางที่1: ตารางผลการประเมินความพึงพอใจโดยผู้เชี่ยวชาญ

ด้านการประเมิน	ระดับประสิทธิภาพ		
	ค่าเฉลี่ย	S.D.	เชิงคุณภาพ
1. ความสามารถในการทำงานของระบบ	3.87	0.35	ดี
2. ด้านการทำงานตามหน้าที่ระบบ	3.54	0.28	ดี
3. ด้านการใช้งานของโปรแกรม	3.78	0.15	ดี
4. ด้านการรักษาความปลอดภัยของระบบ	3.64	0.28	ดี

สรุปผลการประเมินความพึงพอใจต่อระบบโดยรวมทุกด้านและผลการประเมินความพึงพอใจในระดับสิทธิ์ผู้ใช้งานลำดับต่าง ๆ โดยผู้เชี่ยวชาญอยู่ในระดับดี

5. สรุปผลและข้อเสนอแนะของงานวิจัย

ระบบรับชำระค่าสินค้าในสหกรณ์โรงเรียน เป็นการใช้เทคโนโลยีเอนเอฟซีในสมาร์โฟนเข้ามารับชำระค่าสินค้าแทนวิธีการจ่ายด้วยเงินสด เพิ่มความสะดวกในการใช้งาน โดยใช้บัตรนักเรียนที่มีการฝังชิฟอาร์เอฟไอดีอยู่แล้วให้เกิดประโยชน์ยิ่งขึ้น เพื่อใช้จ่ายค่าสินค้าของนักเรียน ผู้ปกครองยังสามารถตรวจสอบรายการซื้อจ่ายของเด็กได้เพื่อควบคุมค่าใช้จ่ายและโภชนาการของเด็ก ทางด้านสหกรณ์และผู้บริหารโรงเรียนสามารถตรวจสอบยอดขายเพื่อนำมาวิเคราะห์ได้ต่อไป ในการพัฒนาระบบมีการประเมินระบบโดยผู้เชี่ยวชาญ ซึ่งให้ความพึงพอใจอยู่ในระดับดี ในการพัฒนาครั้งต่อไป จะมีการเพิ่มในส่วนของผู้ใช้งานของนักเรียนและผู้ปกครองให้สามารถตรวจสอบรายการต่าง ๆ บนแอปพลิเคชันบนสมาร์ทโฟนได้ทั้งหมด

เอกสารอ้างอิง

[1] ณัฐพ แซ่ลิ้ม และ เชษฐพัฒโนทัย, “ระบบสมาชิกสำหรับแอนดรอยด์มาร์ตโฟนโดยใช้เอ็นเอฟซีการ์ดอีมูเลชั่นโหมด,” ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิทศวกรรม จุฬาลงกรณ์มหาวิทยาลัย, 2557.

[2] กมลมาลย์ เสวตวงษ์, วนิดา แก่นอากาศ และวลัยลักขณ์ แสงวรรณกูล, “เทคโนโลยี NFC และQR Code ในสมาร์ทโปสเตอร์แนะนำห้องสมุด,” PULINET ปีที่ 1 ฉบับที่ 1 มกราคม-เมษายน 2557 หน้า 28.

[3] Dino Zupanovic, “Implementation Model for Near Field Communication in Croatian Ferry Ticketing System,” University of Zadar, Maritime department, Mihovila Pavlinovica bb, Zadar, 23000 Croatia, 2015.

[4] ธนัชพร อัครบวร, “ส่วนประสมทางการตลาดที่มีอิทธิพลต่อการเลือกซื้อสินค้าผ่านทางสื่อพาณิชย์อิเล็กทรอนิกส์: กรณีศึกษาผลิตภัณฑ์ท้องถิ่นของไทย,” บริหารธุรกิจมหาบัณฑิต สาขาวิชาการจัดการสาระและการสร้างคุณค่า มหาวิทยาลัยกรุงเทพ, 2557.

[5] AIS mPay Rabbit. [ออนไลน์]. ค้นเมื่อ 28 มีนาคม 2560, จาก https://play.google.com/store/apps/details?id=com.palomar.mpay

[6] สิทธันต์ สันติชัยกูล, “โครงการศึกษาพฤติกรรมผู้บริโภค และแผนการตลาดเพื่อพัฒนา Application Software เกี่ยวกับธุรกิจท่องเที่ยวเชิงวัฒนธรรม บริเวณเกาะรัตนโกสินทร์,” นิเทศศาสตรมหาบัณฑิต สาขาวิชาการบริหารธุรกิจบันเทิงและการผลิต มหาวิทยาลัยกรุงเทพ, 2557.

[7] สมพงศ์ หุตะจูฑะ, “ประตูทางเข้าอัตโนมัติ Smart gate,” สำนักทรัพยากรการเรียนรู้คุณหญิงหลง อรรถกระวีสุนทร มหาวิทยาลัยสงขลานครินทร์, 2557.

[8] นพวรรณ เจริญกิจ, “การพัฒนากระบวนการควบคุมชิ้นงานระหว่างกระบวนการผลิตในโรงงานอุตสาหกรรมฟอก ย้อมและตกแต่งผ้าผืนด้วยการประยุกต์ใช้ระบบบาร์โค้ด” วิศวกรรมการจัดการอุตสาหกรรมเพื่อความยั่งยืน (บัณฑิตศึกษา) คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเทคโนโลยีราชมงคลพระนคร, 2557.

[9] พรศิลป์ พฤทธิวงศ์, “การพิมพ์รหัสบาร์โค้ดบนสติ๊กเกอร์สำหรับติดบนครุภัณฑ์” วิทยบริการ ปีที่ 19 ฉบับที่ 1 1 มกราคม - เมษายน 2551 หน้า 83-84.

[10] อาจารย์ สกล เลี่ยมประวัติ, “การพัฒนาระบบสารสนเทศเพื่อใช้สนับสนุนการเรียนการสอนกระบวนการวิชาระบบสารสนเทศเพื่อการวางแผนทรัพยากรองค์กร,” วิทยาลัยศิลปะ สื่อ และเทคโนโลยี มหาวิทยาลัยเชียงใหม่, 2556.

ต้นแบบแอปพลิเคชันโค้ชชิ่งกิจกรรมนักศึกษาหลายทางเลือก ด้วยรูปแบบการใช้งานแบบปรับเหมาะ

A prototype of coaching student activities based on Adaptive Modality

พรประเสริฐ บุญนะผล (Phonepasuerth Bounnaphonh)[1] ปานจิตร์ หลงประดิษฐ์ (Panchit Longpradit)[2]

ปราโมทย์ ตงฉิน (Pramote Tongchin)[3] และวีระชัย คอนจอหอ (Veerachai Khonchoho)[4]

[1]วิทยาลัยครูหลวงพระบาง สาธารณรัฐประชาธิปไตยประชาชนลาว

[2]สาขาวิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและการสื่อสาร มหาวิทยาลัยพะเยา

[3]สาขาวิชาคอมพิวเตอร์ประยุกต์ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยราชภัฏเพชรบุรี

[4]สาขาวิชาคอมพิวเตอร์ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยราชภัฏเพชรบุรี

Phonepasuerth@gmail.com, plongpradit@yahoo.com, Pm_Pn@hotmail.com, vkhon@hotmail.com

บทคัดย่อ

ปัญหาในการบันทึกการทำกิจกรรมต่างของนักศึกษาใหม่มหาวิทยาลัยราชภัฏเพชรบุรีส่วนใหญ่มีความสามารถในการใช้งานระบบบันทึกกิจกรรมที่แตกต่างกัน งานวิจัยนี้จึงมีวัตถุประสงค์เพื่อ พัฒนาและประเมินการยอมรับต้นแบบแอปพลิเคชันโค้ชชิ่งกิจกรรมนักศึกษาหลายทางเลือก ด้วยรูปแบบการใช้งานแบบปรับเหมาะ กลุ่มเป้าหมายที่ใช้ในการประเมินเป็นนักศึกษาที่กำลังศึกษาอยู่ในภาคเรียนที่ 2558/1 โดยการสุ่มตัวอย่างอย่างง่ายจำนวน 30 คน เครื่องมือที่ใช้ในการวิจัยประกอบด้วย ต้นแบบและแบบประเมินการยอมรับต้นแบบแอปพลิเคชันโค้ชชิ่งกิจกรรมนักศึกษาหลายทางเลือกด้วยรูปแบบการใช้งานแบบปรับเหมาะ สถิติที่ใช้ในการวิเคราะห์ข้อมูลได้แก่ ค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐาน

ผลการวิจัยพบว่า ต้นแบบแอปพลิเคชันโค้ชชิ่งกิจกรรมนักศึกษาหลายทางเลือกด้วยรูปแบบการใช้งานแบบปรับเหมาะประกอบด้วย 1) ส่วนการเข้าระบบ ที่สามารถตรวจสอบสิทธิผู้ใช้ที่แตกต่างกัน 2) การเข้ารายการกิจกรรม 3) การนำทางมีหน้าที่ในการระบุตำแหน่ง และนำทางไปสู่อาคารต่าง ๆ ของมหาวิทยาลัย 4) ส่วนติดตามกิจกรรม และนักศึกษายอมรับต้นแบบแอปพลิเคชันโค้ชชิ่งกิจกรรมนักศึกษาหลายทางเลือกด้วยรูปแบบการใช้งานแบบปรับเหมาะโดยรวมอยู่ในระดับมาก ($\overline{X}$ =4.09,S.D.=0.46)

คำสำคัญ: แอปพลิเคชัน โค้ชชิ่ง ปรับเหมาะ

Abstract

The usability issues in tracking activities for new students at the Phetchaburi Rajabhat University result in the need for an application allowing for activity tracking system with modality. The objectives of this research were therefore to develop and evaluate a prototype of coaching student activities application based on adaptive modality. The sample group for this research were 30 students selected by means of simple sampling method from the Phetchaburi Rajabhat University in the first semester of the academic year 2016. The research instruments were a prototype and an evaluation questionnaire of coaching student activities based on adaptive modality. The statistical methods for data analysis were mean and standard deviation.

The research results were as follows:

1) The coaching student activities application based on adaptive modality developed consists of (1) authentication function, (2) activities selection function, (3) navigation function to pin the location and guide the users to navigate from one building to another building in Phetchaburi Rajabhat University, and (4) tracking activities function with adaptive modality.

2)The user evaluation was performed and the users were overall accepted the application at a high level ($\overline{X}$ =4.09,S.D.=0.46).

Keyword: student activity application, coaching, adaptive modality

1. บทนำ

การทำกิจกรรมเสริมหลักสูตรต่าง ๆ ได้ถูกนำมาเป็นส่วนหนึ่งในการสำเร็จการศึกษา เพื่อเพิ่มทักษะชีวิตที่เกี่ยวข้องกับการทำกิจกรรมต่างๆในรั้วมหาวิทยาลัย อันเป็นปัจจัยหนึ่งที่สำคัญในการพัฒนาคุณภาพของนักศึกษาคู่กับการเรียนในชั้นเรียน เป็นการฝึกที่ต้องการส่งเสริมนักศึกษาให้เป็นคนเก่ง และคนดีในสังคม ซึ่งนักศึกษาต้องปฏิบัติตามข้อกำหนดที่ได้ตั้งเอาไว้ โดยแยกออกจากรายวิชาที่เรียนตามหลักสูตร เมื่อนักศึกษาสำเร็จการศึกษาจะได้รับรายงานผลการศึกษาและรายงานผลกิจกรรม เพื่อประโยชน์ในการสมัครงาน โดยได้มีการตั้งเกณฑ์ของกิจกรรมเสริมหลักสูตรทั้งหมด 24 กิจกรรม ประกอบด้วยเครดิตขั้นต่ำคือ 100 เครดิตสำหรับหลักสูตร4 ปี และ 5 ปี ถ้านักศึกษาทำกิจกรรมไม่ครบตามจำนวนที่กำหนดจะไม่สามารถสำเร็จการศึกษาได้[1]

ในปัจจุบันความสะดวกทางด้านเทคโนโลยีของได้รับความนิยมอย่างแพร่หลาย มีอิทธิพลกับการใช้งานเพราะสามารถกระจายข้อมูลข่าวสารที่ต้องการความสะดวกและรวดเร็วไม่ว่าจะเป็นทางด้านความบันเทิงและการศึกษา ผ่านทางสังคมออนไลน์รวมไปถึงการใช้งานด้านต่างๆ เนื่องจากความสามารถในการใช้งานบางอย่างที่ใกล้เคียงกับคอมพิวเตอร์ของอุปกรณ์เคลื่อนที่เพราะมีความสะดวกในการพกพาจึงทำให้นักพัฒนาซอฟต์แวร์หันมาสนใจในการพัฒนาซอฟต์แวร์กันอย่างแพร่หลายมากขึ้น รวมทั้งซอฟต์แวร์ที่มีสามารถใช้งานบนระบบปฏิบัติการต่างๆ เช่น ระบบปฏิบัติการแอนดรอยด์ (Android OS) และ ระบบปฏิบัติการไอโอเอส (Apple OS)

จากหลักการและเหตุผลที่กล่าวมาข้างต้น ผู้วิจัยได้มองเห็นความสำคัญของการจัดการกับสารสนเทศโดยเฉพาะการติดตามกิจกรรมของนักศึกษา โดยพัฒนาต้นแบบแอปพลิเคชันโค้ชชิ่งกิจกรรมนักศึกษาหลายทางเลือก ด้วยรูปแบบการใช้งานแบบปรับเหมาะ อันเป็นอีกทางเลือกหนึ่งในการใช้งานของนักศึกษาเพื่อให้ได้รูปแบบของโปรแกรมต้นแบบมีความสอดคล้องกับกลุ่มเป้าหมาย จึงมีวัตถุประสงค์เพื่อพัฒนาและประเมินการยอมรับต้นแบบแอปพลิเคชันโค้ชชิ่งกิจกรรมนักศึกษาหลายทางเลือกด้วยรูปแบบการใช้งานแบบปรับเหมาะ

2. วรรณกรรมที่เกี่ยวข้อง

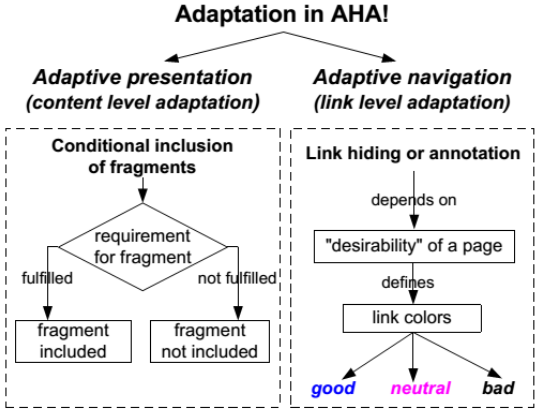
2.1 การโค้ชชิ่ง (coaching)

การเรียนรู้ผ่านการโค้ชชิ่งเป็นการเรียนรู้ที่สามารถทำได้โดยทั่วไปจะเป็นการใช้งานเพื่ออธิบายงานที่เกี่ยวข้องกับความสัมพันธ์ของทักษะ รวมไปถึงกระบวนการคิดที่สมเหตุสมผลของการพัฒนาอย่างมืออาชีพโดยการโค้ชชิ่งมีหลายหลายวิธีที่แตกต่างกันออกไปแต่สิ่งหนึ่งที่คนจะเข้าใจเสมอคือคำว่าโค้ชชิ่งจะเป็นการทำสิ่งที่เป็นอยู่ให้ง่ายมากกว่าเดิม[2]การเรียนรู้ผ่านการโค้ชชิ่งเป็นการเพิ่มศักยภาพของคนเพื่อสร้างประสิทธิภาพของเขาโดยได้พัฒนาทักษะความสัมพันธ์อาจเป็นในรูปแบบการสอนหรือสาธิต รวมไปถึงการนำเสนอเพื่อช่วยเหลือให้เป็นแรงจุงใจในการแก้ไขเพื่อรรลุเป้าหมายในการฝึกร่วมกัน เพื่อให้ผลลัพธ์ที่ได้จากการทำงานมีประสิทธิภาพเสริมสร้างความหมั้นใจให้กับผู้ที่ถูกฝึกให้มีความเชื่อมั่นในการทำงานให้มากขึ้นโดยได้บ่งทักษะการโค้ชชิ่งออกเป็น3 ประเภทคือ 1) เอ็กคูชิฟโค้ชชิ่ง(Executive coaching) เป็นการโค้ชชิ่งที่เป็นศูนย์กลางในการพัฒนาอย่างรวดร็ว ที่มีประสิทธิภาพสูงและมุ่งเน้นไปทางดด้านเทคนิคการแก้ไขปัญหาโดยตระหนักถึงความจำเป็นในการปรับปรุงอย่างต่อเนื่องโดยทุกคนสามารถมีโอกาสในการพัฒนาได้ 2) ทรานเซ็กชันนล แอน ทรานฟรอมเมทิลเนลโค้ชชิ่ง (Transactional and transformational coaching) เป็นการเปลี่ยนแปลงรูปแบบที่เกี่ยวข้องกับสมรรถนะการเรียนรู้ทักษะและเทคนิคมุ่งเน้นไปที่การขยับมุมมองบุคคลค่านิยมและความรู้สึกของวัตถุประสงค์มีสามารถอำนวยความสะดวกที่มีประสิทธิภาพเพื่อสร้างการเปลี่ยนแปลง 3) เซลฟโค้ชชิ่ง (Self-coaching) เป็นลักษณะพื้นฐานของการฝึก มันเป็นสิ่งจำเป็นสำหรับคนที่ต้องการจะฝึกตัวเองโดยอาศัยประสบการณ์การทำอย่างต่อเนื่องจนเกิดให้มีการนำพาตัวเองไปสู่ความสำเร็จที่มากที่สุด[3]โดยมีสภาพแวดล้อมมีหลากลายวิธีการเรียนรู้ที่แตกต่างออกไปเช่น สถานที่ที่คนไปเรียนรู้วิธีการจัดพื้นที่การเรียนรู้ตำแหน่งของการมุ่งเน้นในห้อง บางครั้งนักเรียนจะถูกจัดเรียงเป็นผู้ชมสำหรับบางชนิดของการเรียนการสอนและการเรียนรู้ที่ได้แบ่งการออกแบบมาให้เหมาะสมกับผู้เรียน เทคโนโลยีสารสนเทศก็เป็นอีกอย่างหนึ่งที่ได้ถูกมาเป็นส่วนหนึ่งในการออกแบบการจัดการเรียนการสอนหรือรวมเป็นสภาพแวดล้อมของการเรียนเช่นเดียวกัน[4]

2.2 การปรับเหมาะ (Adaptive)

เทคนิคการปรับเหมาะเป็นเทคนิคที่นำมาใช้ในการนำเสนอด้านข้อความหรือการลิงก์นำทางซึ่งสามารถนำไปประยุกต์ใช้ในการค้นหาข้อมูลที่สามารถลดเวลาในการค้นหาข้อมูล และได้มีการพัฒนาให้การปรับเปลี่ยนทางด้านการนำเสนอให้มีประสิทธิภาพยิ่งขึ้นด้วยการเพิ่มในส่วนของการนำเสนอไม่ใช่แค่มีสองหน้าที่หลัก แต่ได้เพิ่มในส่วนที่เป็นการเสนอด้านไกด์ไลน์(Direct Guidance) เสนอข้อมูลที่สมควรนำเสนอ (Adaptive Sorting of Links) ปรับเหมาะการซ่อนข้อมูลบางส่วนที่ไม่จำเป็น(Adaptive hiding of links) แนะนำข้อมูลที่ควรจะทำ(Adaptive annotation of links) สามารถกำหนดแผนผังการปรับเหมาะให้เหมาะสมได้ (Map Adaptation) นอกนั้นยังมีเทคนิคการปรับเปลี่ยนเกิดขึ้นมากมายในปัจจุบันโดยได้มีการอาศัยทางด้านการพัฒนา การปรับเหมาะทางด้านการใช้งานของผู้ใช้บนระบบสารสนเทศ[5]

เดอบราและคณะ[6] ได้แบ่งเทคนิคการปรับเหมาะออกเป็น 2 แบบหลักคือ 1) การนำเนอแบบปรับเหมาะ (Adaptive Presentation) เป็นการให้ลำดับความสำคัญของเนื้อหาโดยจะมีการแบ่งเนื้อหาการนำเสนอออกเป็นส่วน ๆ ที่แตกต่างกันขึ้นอยู่กับระดับของผู้ใช้ที่แตกต่างกัน 2) การนำทางแบบปรับเหมาะ(Adaptive Navigation) เป็นการนำไปสู่เนื้อหาที่เชื่อมโยงกันโดยที่ผู้ใช้ไม่รู้โดยจะไม่แสดงหน้า หรือ เนื้อหาที่ไม่พึงประสงค์โดยจะมีการปรับเหมาะให้เหมาะสมกับความต้องการของผู้ใช้ว่ามีความต้องการย่างไรสามารถสร้างความเชื่อมโยงให้กับผู้ใช้ได้อย่างเหมาะสมและนำทางไปสู้เนื้อหาข้อมูลที่ต้องการเทคนิคการปรับเหมาะสามารถทำได้ด้วยขันตอนดังภาพ



ภาพที่ 1: เทคนิคการปรับเหมาะ(Adaptive Techniques used AHA,De Bra, et.al, 2002)

3. วิธีการดำเนินการวิจัย

ในการดำเนินการวิจัยเพื่อให้ได้ต้นแบบแอปพลิเคชันโค้ชชิ่งกิจกรรมนักศึกษาหลายทางเลือกด้วยรูปแบบการใช้งานแบบปรับเหมาะตามความต้องการของนักศึกษาผู้วิจัยได้วางแนวทางในการวิจัยตามขั้นตอน ดังต่อไปนี้

3.1 ศึกษาความต้องการของนักศึกษาเพื่อทำการพัฒนาต้นแบบแอปพลิเคชันโค้ชชิ่งกิจกรรมนักศึกษาหลายทางเลือกด้วยรูปแบบการใช้งานแบบปรับเหมาะ

3.2 ทำการออกแบบระบบต้นแบบแอปพลิเคชันโค้ชชิ่งกิจกรรมนักศึกษาหลายทางเลือกด้วยรูปแบบการใช้งานแบบปรับเหมาะ

3.3 ทำการพัฒนาระบบต้นแบบแอปพลิเคชันโค้ชชิ่งกิจกรรมนักศึกษาหลายทางเลือกด้วยรูปแบบการใช้งานแบบปรับเหมาะตามที่ออกแบบไว้ตามขั้นตอนที่ 3.2

3.4 นำระบบที่ได้ไปทำการทดสอบการทำงานของระบบกับนักศึกษาที่กำลังศึกษาอยู่ในภาคเรียนที่ 1/2558 ของมหาวิทยาลัยราชภัฏเพชรบุรีที่ไม่ใช้กลุ่มตัวอย่าง

3.5 นำระบบที่ได้ไปทำการประเมินการยอมรับกับนักศึกษาที่กำลังศึกษาอยู่ในภาคเรียนที่ 1/2558 โดยการสุ่มตัวอย่างอย่างง่ายจำนวน 30 คน โดยใช้สถิติคือ ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐาน ตั้งค่าระดับคะแนนการยอมรับ ดังนี้

ระดับคะแนน	ระดับการยอมรับ
5.00 – 4.50	มากที่สุด
4.49 – 3.50	มาก
3.49 – 2.50	ปานกลาง
2.49 – 1.50	น้อย
1.49 – 1.00	น้อยที่สุด

3.6 สรุปผลการวิจัย

4. ผลการวิจัย

ผลการพัฒนาและประเมินการยอมรับต้นแบบแอปพลิเคชันโค้ชชิ่งกิจกรรมนักศึกษาหลายทางเลือกด้วยรูปแบบการใช้งานแบบปรับเหมาะสามารถแสดงได้ดังต่อไปนี้

4.1 ผลพัฒนาต้นแบบ

ผลพัฒนาต้นแบบแอปพลิเคชันโค้ชชิ่งกิจกรรมนักศึกษาหลายทางเลือกด้วยรูปแบบการใช้งานแบบปรับเหมาะ ทำให้ได้ระบบที่ประกอบด้วย 6 ส่วนคือ 1) ส่วนโปรไฟลผู้ใช้ 2) ช่องทางในการเข้าถึงโปรแกรม 3) การปรับเหมาะ 4) การนำเสนอสารสนเทศ และ 5) การแจ้งเตือน ดังภาพตัวอย่างต่อไปนี้



ภาพที่ 2 หน้าหลักเข้าสู่ระบบต้นแบบ



ภาพที่ 3 หน้าแรกของนักศึกษาเมื่อผ่านการเข้าสู้ระบบ



ภาพที่ 4 แสดงการใช้งานนำทางสถานที่



ภาพที่ 5 การนำเส้นทางของต้นแบบฯ



ภาพที่ 6 หน้าแสดงการเข้าเลือกลงตารางกิจกรรม

4.2 ผลการประเมินการยอมรับ

ผลการประเมินการยอมรับของนักศึกษาที่กำลังศึกษาอยู่ในภาคเรียนที่ 1/2558 โดยการสุ่มตัวอย่างอย่างง่ายจำนวน 30 คน ที่มีต่อต้นแบบแอปพลิเคชันโค้ชชิ่งกิจกรรมนักศึกษาหลายทางเลือกด้วยรูปแบบการใช้งานแบบปรับเหมาะ 4 ด้าน คือ ด้านความสามารถการทำงาน (Function Requirement) ด้านหน้าที่ของต้นแบบ(Functional Test) ด้านการใช้งาน (Usability test) และ ด้านความปลอดภัย (Security Test) มีรายละเอียดดังนี้

ตารางที่ 1: แสดงผลประเมินการยอมรับของนักศึกษาที่มีต่อแอปพลิเคชันโค้ชชิ่งกิจกรรมนักศึกษาหลายทางเลือกด้วยรูปแบบการใช้งานแบบปรับเหมาะ

รายการประเมิน	$\bar{X}$	S.D.	ระดับการยอมรับ
1.ด้านความสามารถการทำงาน	4.03	0.40	มาก
2.ด้านหน้าที่ของต้นแบบ	4.13	0.51	มาก
3.ด้านการใช้งาน	4.10	0.47	มาก
4.ด้านความปลอดภัย	4.11	0.46	มาก
รวม	4.09	0.46	มาก

จากตารางที่ 1 พบว่านักศึกษายอมรับในต้นแบบแอปพลิเคชันโค้ชชิ่งกิจกรรมนักศึกษาหลายทางเลือกด้วยรูปแบบการใช้งานแบบปรับเหมาะโดยรวมอยู่ในระดับมาก ซึ่งด้านหน้าที่ของต้นแบบมีค่ามากที่สุดอยู่ในระดับ $\bar{X}$= 4.13 S.D.= 0.51

5. สรุปผลการวิจัย

จากการวิจัยเพื่อพัฒนาและประเมินการยอมรับต้นแบบแอปพลิเคชันโค้ชชิ่งกิจกรรมนักศึกษาหลายทางเลือก ด้วยรูปแบบการใช้งานแบบปรับเหมาะ ทำให้ได้ระบบต้นแบบที่ประกอบด้วย 6 ส่วนคือ 1) ส่วนโปรไฟลผู้ใช้ 2) ช่องทางในการเข้าถึงโปรแกรม 3) การปรับเหมาะ 4) การนำเสนอสารสนเทศ และ 5) การแจ้งเตือน และได้รับการยอมรับโดยรวมอยู่ในระดับมากทั้ง 4 ด้านคือ 1) ด้านความสามารถการทำงาน 2) ด้านหน้าที่ของต้นแบบ 3) ด้านการใช้งาน และ4) ด้านความปลอดภัย

เอกสารอ้างอิง

[1] กองพัฒนานักศึกษา มหาวิทยาลัยราชภัฏเพชรบุรี. (2558). การดำเนินกิจกรรมเสริมสร้างหลักสูตร. สืบค้นเมื่อ 19 มีนาคม 2558. แหล่งข้อมูลจากhttp://pbruact.pbru.ac.th/pbruact

[2] Morel. (2014). Learning from coaching. ASCD Alexandaria, VA USA. www.ascd.org. Copyright in 2014 by ASCD

[3] Sandra fielden. (2014). Coaching effectiveness a summary. A summary of a report for the NHS leadership center, Manchester Business school, University of Manchester

[4] Warger Tom, Eduserve and Gregory Dobbin. (2009). Learning environments. ELI paper 1: October, 2009

[5] P. Brusilovsky. “Methods and techniques of adaptive hypermedia”. User modeling and user adapted interaction 1996

[6] De Bra Paul, Ad Aserts, Bart Berden, Barend de Lange, Brendan Rousseau, Tomi Santic, David Smits, Natalia Stash. (2002) .The Adaptive Hypermedia Architecture. Department of Computer Science Eindhoven University of Technology, Eindhoven, the Netherlands

ระบบรักษาความปลอดภัยในบ้านด้วยราสเบอรีพายผ่านช่องทางเฟซบุ๊กเมสเซนเจอร์
Home Security System using Raspberry PI through Facebook Messenger

เดชาวุฒิ วานิชสรรพ์ (Dechawut Wanichsan)[1] คมศักดิ์ ผลพฤกษา (Komsak Ponpruksa)[2]
และญานี ลำจอง (Yanee Lamjong)[3]
[1]คณะวิทยาการคอมพิวเตอร์และเทคโนโลยีสารสนเทศ [2,3]คณะครุศาสตร์ มหาวิทยาลัยราชภัฏรำไพพรรณี
[1]dechawut.w@rbru.ac.th, [2]5615261004@rbru.ac.th, [3]5615261010@rbru.ac.th

บทคัดย่อ

วัตถุประสงค์หลักของงานวิจัยชิ้นนี้คือ เพื่อพัฒนาระบบรักษาความปลอดภัยในบ้านด้วยการใช้ราสเบอรีพาย 3 โมเดลบี ร่วมกับตัวรับรู้พีไออาร์ โดยสามารถทำการแจ้งเตือนเป็นข้อความผ่านทางเฟซบุ๊ก เมสเซนเจอร์เมื่อตัวรับรู้สามารถตรวจพบความเคลื่อนไหวของมนุษย์ หรือวัตถุที่สามารถแผ่รังสีความร้อน ผลการทดลองพบว่าระบบรักษาความปลอดภัยที่พัฒนาขึ้นได้รับการควบคุมการเปิดปิดการใช้งานผ่านการพิมพ์ข้อความที่ระบุไว้เพื่อเป็นคำสั่ง ผ่านช่องทางเฟซบุ๊ก เมสเซนเจอร์ นอกจากนี้การใช้ตัวรับรู้พีไออาร์หนึ่งตัวสามารถตรวจจับการเคลื่อนไหวในทิศทางและระยะทางที่แตกต่างกันได้ค่อนข้างมีประสิทธิผล

คำสำคัญ: ระบบรักษาความปลอดภัย ราสเบอรีพาย ตัวรับรู้พีไออาร์ เฟซบุ๊กเมสเซนเจอร์ เอพีไอ

Abstract

The main purpose of this research was to develop a home security system using Raspberry pi 3 model b together with PIR motion sensor. The system could alarm with messages through Facebook messenger when there were some movement of human or object that could emit heat. The experimental result revealed that the developed system could be controlled to be on-off status by typing a predefine text as command in the messenger. In addition, sensor could effectively detect motion status in different directions and different distances.

Keyword: Security System, Raspberry Pi, PIR sensor, Facebook Messenger, API

1. บทนำ

ด้วยสภาพปัญหาด้านเศรษฐกิจและปัญหาสังคมที่มีอยู่ในปัจจุบันมีส่วนก่อให้เกิดอาชญากรรมเพิ่มมากขึ้น ในปัจจุบันข่าวตามหน้าหนังสือพิมพ์นำเสนอข่าวที่เป็นคดีโจรกรรมมากมาย โดยเฉพาะคดีเกี่ยวกับการขโมยทรัพย์สินในบ้าน [1], [2], [3] ดังนั้นทางออกในการแก้ไขปัญหาเบื้องต้นสำหรับการบรรเทาปัญหาโจรกรรมคือ การติดตั้งระบบรักษาความปลอดภัยด้วยการติดตั้งกล้องวงจรปิดเพื่อแก้ปัญหา อย่างไรก็ตามกล้องวงจรปิดรุ่นก่อน ๆ ที่ติดตั้งอยู่ตามบ้านเรือนนั้นสามารถบันทึกได้เพียงภาพเหตุการณ์เท่านั้น เมื่อเกิดการโจรกรรมขึ้นเจ้าของบ้านมาทราบภายหลังเมื่อถูกโจรกรรมไปแล้ว ทำได้เพียงเปิดกล้องวงจรปิดเพื่อดูเหตุการณ์ย้อนหลังเท่านั้น แม้ว่ากล้องจะสามารถจับภาพคนร้าย บางเหตุการณ์ที่เป็นคดียังตามจับคนร้ายไม่ได้มีให้เห็นเป็นจำนวนมาก หากจับคนร้ายได้ทรัพย์สินที่ถูกโจรกรรมไปอาจไม่ได้คืน ในปัจจุบันมีการพัฒนากล้องวงจรปิดให้สามารถจับความเคลื่อนไหวได้ เมื่อมีผู้บุกรุกจะส่งข้อความเข้าโทรศัพท์มือถือให้เจ้าของบ้านทราบ แต่ยังไม่นิยมมากนักเนื่องจากมีราคาสูง

จากปัญหาดังกล่าว ผู้วิจัยจึงมีแนวคิดที่จะสร้างระบบรักษาความปลอดภัยในบ้านที่มีงบประมาณที่ไม่สูงนัก ซึ่งสามารถแจ้งเตือนเจ้าของบ้านได้อย่างรวดเร็วและทันเวลาเมื่อมีผู้บุกรุกเข้ามาภายในบ้าน เจ้าของบ้านจะได้หาทางรับมือและยับยั้งเหตุได้ทันเวลา ระบบดังกล่าวมีการแจ้งเตือนด้วยการส่งข้อความผ่านทาง เฟซบุ๊กเมสเซนเจอร์ (Facebook Messenger) ซึ่งเป็นโปรแกรมที่ได้รับความนิยมและแพร่หลายมากในทุกเพศ ทุกวัย นับเป็นการตอบโจทย์ของผู้ใช้ในปัจจุบัน หากติดตั้งระบบนี้ไว้ภายในบ้านทำให้เจ้าของบ้านมีความรู้สึกปลอดภัยและมีความมั่นใจมากขึ้นเมื่อไม่อยู่บ้าน

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 เฟซบุ๊ก

เฟซบุ๊ก คือ เว็บไซต์เครือข่ายสังคม (Social Networking Website) สำหรับการติดต่อแลกเปลี่ยนข้อมูล ข่าวสาร ระหว่างสมาชิก เฟซบุ๊ก วันที่ 28 มกราคม 2555 พบข้อมูลจาก socialbakers.com ว่าในประเทศไทยมีผู้ใช้เฟซบุ๊กจำนวนประมาณ 13,686,960 คน เป็นประเทศอันดับที่ 16 ของทั่วโลก ช่วงอายุ 18 – 24 ปี มีมากที่สุดคือ 34% และ ช่วงอายุ 25 – 34 ปี มีรองลงมาคือ 29% ซึ่งสองกลุ่มนี้รวมกันได้ถึง 63% ของทั้งหมด และข้อมูลจาก 213 ประเทศพบว่า อเมริกาเป็นอันดับหนึ่งมี 155,701,780 คน จากทั้งหมดประมาณ 800 ล้านคน [4]

2.2 ราสเบอรีพาย

ราสเบอรีพาย (Raspberry Pi) คือ บอร์ดคอมพิวเตอร์ขนาดเล็กที่สามารถเชื่อมต่อกับจอมอนิเตอร์ คีย์บอร์ด และเมาส์ได้ และมีช่องเสียบ พอร์ตยูเอสบี (Universal Serial Bus: USB) ถึงสี่พอร์ต เชื่อมต่อกับสัญญาณไวไฟและสายแลน สามารถนำมาประยุกต์ใช้ในการทำโครงงานทางด้านอิเล็กทรอนิกส์ การเขียนโปรแกรม หรือเป็นเครื่องคอมพิวเตอร์ตั้งโต๊ะขนาดเล็ก

2.3 ตัวรับรู้พีไออาร์

พีไออาร์ (Passive Infrared Receiver: PIR) คือ ตัวรับรู้รังสีอินฟราเรด จากมนุษย์ สัตว์ หรือสิ่งของต่าง ๆ มีการแผ่รังสีอินฟราเรดอยู่ตลอดเวลา ดังนั้นจึงทำมาเพื่อใช้ในการตรวจจับ ซึ่งหลักการทำงานคือ ตรวจจับการเปลี่ยนแปลงของรังสีอินฟราเรด ที่แผ่ออกมา ราคาของพีไออาร์อยู่ที่ ประมาณ 65 บาท ซึ่งแต่ละรุ่นจะมีราคาแตกต่างกันออกไป

2.4 ระบบการเชื่อมต่อสัญญาณไวไฟ

เป็นระบบการสื่อสารข้อมูลที่ใช้คลื่นความถี่วิทยุ (Radio Frequency: RF) และคลื่นอินฟราเรด ในการรับและส่งข้อมูลระหว่างอุปกรณ์ ผ่านอากาศ ทะลุกำแพง เพดานหรือสิ่งก่อสร้างอื่น ๆ โดยปราศจากการใช้สายเพื่อนำสัญญาณ นอกจากนั้นระบบเครือข่ายไร้สายยังมีคุณสมบัติครอบคลุมทุกอย่างเหมือนกับระบบ แลน แบบใช้สาย ที่สำคัญคือ การที่ไม่ต้องใช้สายทำให้การเคลื่อนย้ายการใช้งานทำได้โดยสะดวก

2.5 โหนดเจเอส

โหนดเจเอส (Node js) เป็นจาวาสคริปต์ที่ทำงานอยู่บนฝั่งเซิร์ฟเวอร์เพื่อช่วยให้นักพัฒนาสามารถเขียนโปรแกรมควบคุมการทำงานบนฝั่งเซิร์ฟเวอร์ได้สะดวกและมีประสิทธิภาพ จึงสามารถประมวลผลได้รวดเร็ว คล้าย ๆ กับการทำงานหลายอย่างพร้อมกันในเวลาเดียวกัน ในที่จริงเป็นการทำงานแบบทีละอย่าง แต่มีการปรับกระบวนการโดยเอาแต่ละงานไปเข้าคิวเวลาเอาไว้ก่อน แล้วค่อยนำไปประมวลผลงานตามคิวเวลา ซึ่งกระบวนการเข้าคิวเวลา ใช้เวลาน้อยมาก ๆ จึงทำให้เหมือนว่าระบบทำงานได้ทีละหลาย ๆ อย่าง ทำให้เป็นที่นิยมกันอย่างแพร่หลายสำหรับการพัฒนาเว็บเซิร์ฟเวอร์ในยุคปัจจุบัน

2.6 เฟซบุ๊ก แชท เอพีไอ

เฟซบุ๊ก แชท เอพีไอ (Application Programming Interface: API) [5] เป็นช่องทางการเชื่อมต่อระหว่างการแชทด้วยเฟซบุ๊ก เมสเซนเจอร์ กับระบบหรือเว็บไซต์ภายนอก โดยปกติ นักพัฒนาโปรแกรมจะใช้งานเอพีไอดังกล่าวสำหรับการโต้ตอบข้อความโดยอัตโนมัติ ในทางเทคนิค เฟซบุ๊ก แชท เอพีไอ จะใช้การร้องขอแบบ GET และ POST ด้วยรูปแบบพื้นฐานของภาษาเอชทีเอ็มแอล เพื่อเข้าถึงเว็บไซต์เฟซบุ๊ก ดังนั้น นักพัฒนาจำเป็นต้องใช้ชื่อบัญชีเฟซบุ๊ก เป็นข้อมูลที่ระบุในโปรแกรมที่พัฒนาขึ้นด้วย อย่างไรก็ตาม การสั่งให้เอพีไอพิมพ์ข้อความด้วยความถี่เกินไป จนมีพฤติกรรมคลายแสปม อาจทำให้ชื่อบัญชีนั้นถูกระงับการใช้งานได้

2.7 งานวิจัยที่เกี่ยวข้อง

Narkhede และ Kjadke [6], [7] ได้ออกแบบระบบเฝ้าระวังโดยใช้ราสเบอร์รีพายร่วมกับตัวรับรู้พีไออาร์ เพื่อตรวจสอบความเคลื่อนไหว เมื่อมีผู้บุกรุก ระบบจะส่งข้อมูลผ่านโมเด็มสามจี และแสดงผลผ่านทางเว็บแอปพลิเคชัน

Gajham และคณะ [8] ได้ออกแบบระบบเฝ้าระวังโดยใช้ราสเบอรีพาย ร่วมกับกล้อง โดยมีการบันทึกภาพต่อเนื่องเป็นวิดีโอลงในฮาร์ดดิสก์ เมื่อมีผู้บุกรุก ระบบจะส่งข้อมูลการแจ้งเตือนผู้ใช้งานผ่านทางอีเมล

Bhavani และคณะ [9] ได้ออกแบบระบบเฝ้าระวังต้นทุนต่ำโดยใช้ราสเบอรีพาย ตัวรับรู้พีไออาร์ ร่วมกับกล้องเว็บแคมที่ทำงานด้วยขั้นตอนวิธีสำหรับการกรองภาพให้มีความสามารถในการมองเห็นได้ดีขึ้นเวลากลางคืน เมื่อมีผู้บุกรุก ระบบจะส่งข้อมูลภาพถ่ายจากกล้องให้ผู้ใช้งานผ่านทางอีเมล

จากงานวิจัยข้างต้น พบว่าระบบรักษาความปลอดภัยที่ได้มีการนำเอาเซ็นเซอร์พีไออาร์มาใช้งานสามารถทำงานได้ดี

เมื่อพบผู้บุกรุก ระบบเหล่านี้จะทำการแจ้งเตือนผ่านทางเว็บแอปพลิเคชัน และทางอีเมล จากสไตล์การใช้ชีวิตของคนไทยในปัจจุบัน โซเชียลเน็ตเวิร์กเป็นสิ่งสำคัญและเป็นส่วนหนึ่งในชีวิตประจำวัน งานวิจัยชิ้นนี้จึงมีแนวคิดที่ต้องการพัฒนาระบบรักษาความปลอดภัยออนไลน์ โดยใช้งานราสเบอรีพายร่วมกับพีไออาร์ ให้มีความสามารถแจ้งเตือนเมื่อพบผู้บุกรุก ผ่านทางเฟซบุ๊กเมสเซนเจอร์ของผู้ใช้ เพื่อเป็นการตอบโจทย์ของผู้ใช้ในปัจจุบันด้วยความสะดวกสบายรวดเร็วและทันสมัย

3. วิธีดำเนินงานวิจัย

หัวข้อนี้กล่าวถึงส่วนของการดำเนินงานและการออกแบบการสร้างระบบรักษาความปลอดภัยในบ้านโดยมีขั้นตอนต่าง ๆ ดังนี้

3.1 ฮาร์ดแวร์ที่ใช้ในการวิจัย

1) ราสเบอรี พาย 3 โมเดล บี (Raspberry Pi 3 Model B) เป็นบอร์ดคอมพิวเตอร์แต่มีขนาดเล็ก ซึ่งเป็นบอร์ดไมโครคอนโทรลเลอร์ชนิดหนึ่ง โดยผู้วิจัยได้นำมาเป็นตัวกลางในการควบคุมการทำงานของตัวรับรู้ความเคลื่อนไหว ความเคลื่อนไหวพีไออาร์ (PIR Motion Sensor) และทำหน้าที่ในการเชื่อมต่อกับเฟซบุ๊ก

2) ตัวรับรู้ความเคลื่อนไหวพีไออาร์ ใช้เป็นตัวตรวจจับการเคลื่อนไหวเมื่อมีการเคลื่อนไหวจะทำการส่งค่าที่ได้ไปยังบอร์ด ราสเบอรี พาย 3 โมเดล บี เพื่อทำการตรวจเช็คตามขั้นตอนของโปรแกรมที่กำหนดไว้

3.2 ซอฟต์แวร์ที่ใช้ในการวิจัย

1) ราสเบียน (Raspbian) เป็นระบบปฏิบัติการ (Operating System: OS) ที่ผู้ใช้ส่วนใหญ่นิยมนำมาใช้กับราสเบอรี พาย เนื่องจากมีความเหมาะสม และทางผู้ผลิตราสเบอรีให้คำแนะนำ

2) เฟซบุ๊ก เมสเซนเจอร์ (Facebook Messenger) ใช้เป็นตัวควบคุมการเปิด-ปิด การทำงานของระบบรักษาความปลอดภัย หากผู้ใช้ทำการเปิดระบบแล้ว เมื่อเซ็นเซอร์ตรวจพบความเคลื่อนไหว ระบบจะทำการแจ้งเตือนข้อความให้ผู้ใช้งานทราบ

3.2 การเชื่อมต่อราสเบอรี พาย กับพีไออาร์

การต่อพีไออาร์กับบอร์ดราสเบอรี พาย 3 โมเดล บี สามารถทำการต่อได้หลายวิธีแล้ว แต่ผู้วิจัยใช้ขา 1 ซึ่งสามารถทำการจ่ายไฟได้ไม่เกิน 3.3V ซึ่งในบอร์ดราสเบอรี พาย 3 โมเดล บี นี้มีขาที่จ่ายกระแสไฟ 3.3V อยู่ 2 ขา ส่วนในขา กราวด์ (Ground) ในโครงการวิจัยนี้จะเลือกใช้ขาที่ 6 ซึ่งในบอร์ดราสเบอรี พาย3 โมเดล บี นี้มีขากราวด์ อยู่ 8 ขาด้วยกัน และสุดท้ายขา จีพีไอโอ (GPIO) ในโครงการวิจัยนี้จะเลือกใช้ขาที่ 12 ซึ่งในบอร์ดราสเบอรี พาย 3 โมเดล บี มีขา จีพีไอโอ ทั้งหมด 28 ขา โดยในแต่ละขาจะมีฟังก์ชันแตกต่างกันไป ได้แก่ เอสพีไอ (SPI) ไอทูซี (I2C) ยูอาร์ท (UART) จีพีไอโอ คือขาอเนกประสงค์สามารถรับและส่งข้อมูลต่าง ๆ ได้ผ่านทางขานี้ ส่วนมากนำไปใช้กับการควบคุมการเปิดปิดอุปกรณ์ภายในบ้าน

3.3 การติดตั้งคำสั่งเบื้องต้น

3.3.1 การติดตั้ง โหนดเจเอส (NodeJS)

1) เข้าสู่หน้าต่างเทอร์มินอลของราสเบอรี พาย 3

2) พิมพ์คำสั่งต่อไปนี้ sudo su - รอจนกว่าโปรแกรมจะรันเสร็จ หลังจากนั้นพิมพ์คำสั่ง #apt-get remove nodered -y รอจนกว่าโปรแกรมจะรันเสร็จ หลังจากนั้นพิมพ์คำสั่ง apt-get remove nodejs nodejs-legacy -y รอจนกว่าโปรแกรมจะรันเสร็จ หลังจากนั้นพิมพ์คำสั่ง apt-get remove npm -y รอจนกว่าโปรแกรมจะรันเสร็จ หลังจากนั้นพิมพ์คำสั่ง curl -sL https://deb.nodesource.com/setup_6.x | sudo bash - รอจนกว่าโปรแกรมจะรันเสร็จ หลังจากนั้นพิมพ์คำสั่ง apt-get install nodejs -y รอจนกว่าโปรแกรมจะรันเสร็จ หลังจากนั้นพิมพ์คำสั่ง node -v จากนั้นพิมคำสั่ง npm -v เป็นการเสร็จสิ้นการลง โหนดเจเอส หลังจากนั้นสามารถลงแพ็กเกจของแชทเฟซบุ๊กในขั้นตอนต่อไป

3.3.2 การติดตั้งแพ็กเกจ เฟซบุ๊ก-แชท-เอพีไอ

1) ย้ายโฟลเดอร์ปัจจุบันมาที่ เดสท็อป (Desktop) ก่อน โดยการใช้คำสั่ง cd Desktop

2) สร้างโฟลเดอร์เพื่อเก็บงาน โดยตั้งชื่อว่า Project ด้วยคำสั่ง mkdir Project

3) หลังจากนั้นย้ายโฟลเดอร์ปัจจุบันมาที่ Project โดยใช้คำสั่ง cd Project

4) ทำการติดตั้ง เฟซบุ๊ก-แชท-เอพีไอ โดยใช้คำสั่ง npm install facebook-chat-api

3.3.3 การติดตั้งแพ็กเกจ ออนออฟ (onoff) ซึ่งช่วยให้โหนดเจเอสควบคุมขาจีพีไอโอ บนราสเบอรี พาย 3 ได้

1) โฟลเดอร์อยู่ที่ Project

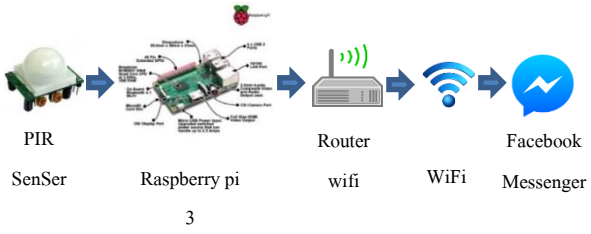
2) ทำการติดตั้งแพ็กเกจ ออนออฟ ด้วยคำสั่ง npm install onoff

3.4 ตำแหน่งติดตั้งตัวรับรู้

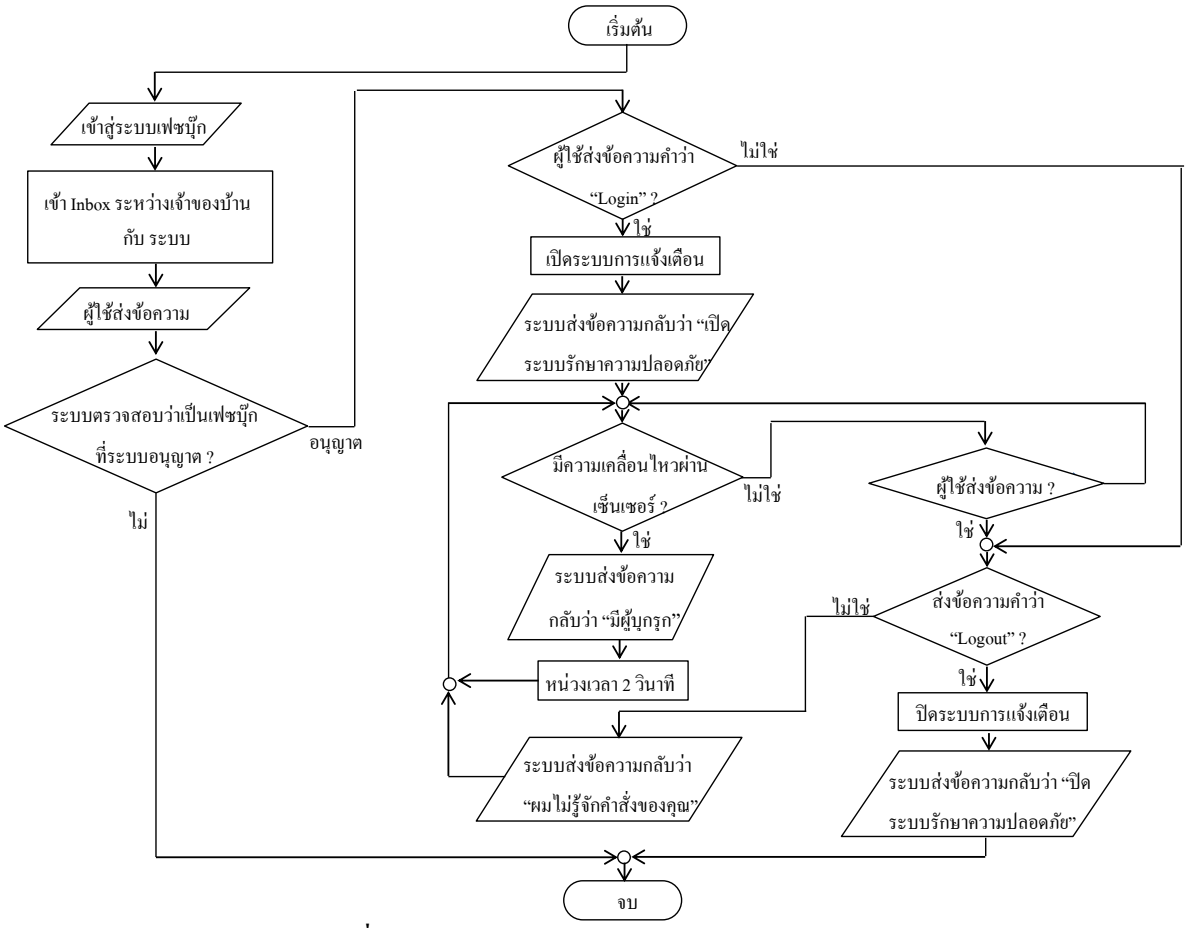
ผู้วิจัยจะติดตั้งตัวรับรู้ที่ใช้ในการตรวจจับความเคลื่อนไหวบริเวณมุมบนด้านในของประตู ด้วยเหตุผลที่ว่า การติดตั้งในบริเวณดังกล่าว ในระยะที่เหนือศีรษะพอดี ทำให้ผู้บุกรุกไม่สามารถทราบได้ว่ามีอุปกรณ์ติดตั้งอยู่

3.5 การกำหนดเฟซบุ๊กที่สามารถเปิด – ปิด ระบบได้

การเปิด-ปิดการใช้งานระบบสามารถทำได้ผ่านเฟซบุ๊ก ข้อมูลที่จำเป็นต้องใช้ในการอนุญาตคือ เฟซบุ๊กไอดี ซึ่งหาได้โดยการใช้ยูอาร์แอลของเฟซบุ๊กในหน้าจอผู้ใช้งาน เป็นข้อมูลนำเข้า มาวางใน http://findmyfbid.com และเว็บดังกล่าวจะให้ผลลัพธ์เป็นเฟซบุ๊กไอดีคืนกลับมา ตัวอย่างเช่น นำ https://www.facebook.com/kookiess มาวางในเว็บไซต์ดังกล่าวจะได้ผลลัพธ์เป็น 100010588024420 เป็นต้น



ภาพที่ 1: ภาพรวมการทำงานของระบบที่พัฒนาขึ้น



ภาพที่ 2: ขั้นตอนการทำงานของระบบรักษาความปลอดภัยในบ้าน

3.6 การทำงานของระบบรักษาความปลอดภัยในบ้าน

ภาพรวมของระบบทั้งหมดแสดงในภาพที่ 1 เมื่อตัวรับรู้พีไออาร์ตรวจจับการเคลื่อนไหวได้ จะส่งข้อมูลไปยังราสเบอรีพาย ซึ่งจะส่งข้อมูลเป็นสัญญาณไร้สายไปยังเราเตอร์ จากนั้นเราเตอร์จะทำการส่งข้อมูลต่อผ่านสัญญาณไร้สายเป็นข้อความเตือนไปยังเฟซบุ๊ก เมสเซนเจอร์ ผ่านทางเฟซบุ๊ก ไอดีที่ตั้งค่าไว้เท่านั้น สำหรับบอร์ดราสเบอรี พาย 3 โมเดล บี ตัวเชื่อมต่อสัญญาณไร้สายของบอร์ดดังกล่าว ไม่จำเป็นต้องป้อนค่าใด ๆ เช่น

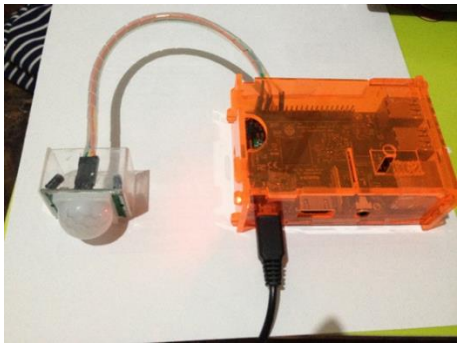
เอสเอสไอดี (Service Set Identifier: SSID) หรือไอพีแอดเดรส (IP Address) เป็นต้น เพื่อทำความรู้จักกับระบบเครือข่าย ทำให้ระบบที่พัฒนาขึ้นสามารถทำงานกับระบบเครือข่ายใด ๆ ที่มีการเชื่อมต่อกับอินเทอร์เน็ต

การทำงานของระบบรักษาความปลอดภัยในบ้านมีขั้นตอนการทำงาน ดังภาพที่ 2 ที่ผู้ใช้งานเปิดใช้งานระบบได้ต้องเคยระบุเฟซบุ๊กไอดีให้ราสเบอรีพายรู้จัก หลังจากนั้น การเปิดระบบ ทำได้ด้วยการพิมพ์ข้อความตั้งค่าไว้ ในที่นี้คือ “Login” หากพิมพ์ข้อความที่ไม่ตรงกับที่ระบุไว้ ระบบจะส่งข้อความเพื่อบอกว่าไม่รู้จักคำสั่ง และไม่อนุญาตในการเปิดระบบ หากพิมพ์ข้อความได้ถูกต้อง ระบบจะเริ่มตรวจจับความเคลื่อนไหว เมื่อมีการเคลื่อนไหวเกิดขึ้น ระบบจะแจ้งเตือนด้วยการส่งข้อความผ่านเฟซบุ๊ก เมสเซนเจอร์ให้ผู้ใช้ทราบ เมื่อผู้ใช้ได้รับข้อความจะทราบโดยทันทีว่ามีผู้บุกรุกและหาทางยับยั้งเหตุได้ทันเวลา หากผู้ใช้ต้องการปิดระบบทำได้ด้วยการพิมพ์ข้อความตั้งค่าไว้ ในที่นี้คือข้อความ “Logout”

4. ผลการดำเนินงาน

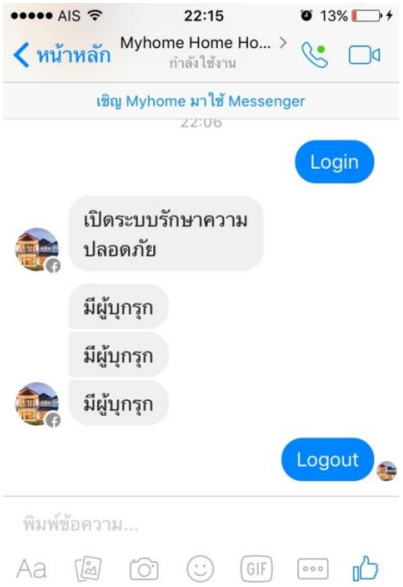
4.1 ผลการพัฒนาระบบรักษาความปลอดภัย

ผู้วิจัยได้ประกอบอุปกรณ์ฮาร์ดแวร์ต่าง ๆ เข้าด้วยกัน ก่อนนำไปติดตั้งใช้งานในสถานที่จริง ดังแสดงในภาพที่ 3



ภาพที่ 3: การเชื่อมต่อตัวรับรู้พีไออาร์กับราสเบอรีพาย

เมื่อต้องการสั่งให้ระบบรักษาความปลอดภัยเริ่มทำงาน ผู้ใช้งานต้องเข้าสู่ระบบบัญชีเฟซบุ๊ก จากนั้นเปิดกล่องข้อความของเฟซบุ๊ก เมสเซนเจอร์ เพื่อคุยกับบัญชีที่สร้างขึ้น เพื่อพิมพ์ข้อความ “Login” เพื่อเปิดระบบ ระบบจะตอบกลับข้อความว่า “เปิดระบบรักษาความปลอดภัย” เป็นการแจ้งให้ทราบว่าระบบเริ่มทำงาน เมื่อระบบตรวจพบความเคลื่อนไหว ระบบทำการส่งข้อความแจ้งเตือนว่า “มีผู้บุกรุก” เมื่อต้องการปิดระบบให้พิมพ์ข้อความว่า “Logout” ระบบตอบกลับมาว่า “ปิดระบบรักษาความปลอดภัย” เพื่อแจ้งให้ทราบว่าระบบได้ปิดการใช้งาน ตัวอย่างหน้าจอบนเฟซบุ๊ก เมสเซนเจอร์แสดงดังภาพที่ 4



ภาพที่ 4: การพิมพ์ข้อความใช้งานระบบผ่านเฟซบุ๊ก เมสเซนเจอร์

4.2 ประสิทธิภาพการตรวจจับของพีไออาร์

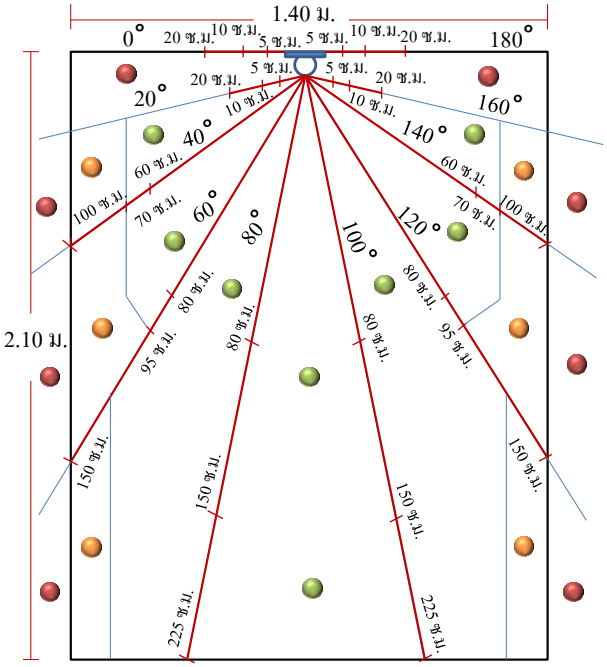
เมื่อผู้วิจัยทำการทดสอบประสิทธิภาพของระบบรักษาความปลอดภัยในบ้าน โดยได้เดินผ่านจุดที่ติดตั้ง โดยแบ่งระยะการตรวจจับในช่วง 1-3 เมตร โดยแบ่งช่วงย่อยละ 10 เซนติเมตร ที่ระดับมุม 90° (0° 90° และ 180° อยู่ที่ตำแหน่ง 9 6 และ 3 นาฬิกาตามลำดับ) โดยทำการทดลองช่วงย่อยละ 30 ครั้ง เปอร์เซนต์การตรวจจับการเคลื่อนไหวได้แสดงดังตารางที่ 1

ตารางที่ 1: การตรวจจับการเคลื่อนไหวตามระยะทางที่มุม 90°

ระยะการตรวจวัด (เซนติเมตร)	การตรวจจับได้ (เปอร์เซนต์)
10-250	100
250	100
260	87
270	53
280	27
290	13
300	0

จากตารางที่ 1 แสดงให้เห็นว่าระยะทำการที่ดีที่สุดที่ระดับมุม 90° อยู่ในระยะไม่เกิน 250 เซนติเมตร อย่างไรก็ตาม

เปอร์เซนต์การตรวจจับได้ที่ลดลง เมื่อระยะวัตถุอยู่ไกลกว่า 250 เซนติเมตร ไม่ส่งผลต่อประสิทธิภาพของระบบเพราะความสูงของประตูบ้านส่วนมากไม่เกิน 250 เซนติเมตร ผู้วิจัยยังได้ทำการทดลองเพิ่มเติมเพื่อวัดประสิทธิภาพของพีไออาร์ในการตรวจจับในระดับมุมที่แตกต่างกัน โดยผลการทดลองแสดงดังภาพที่ 4



ภาพที่ 4: แบบจำลองรัศมีการตรวจจับของตัวรับรู้พีไออาร์

จากภาพที่ 4 ประตูกว้าง 1.40 เมตร และสูง 2.10 เมตร บริเวณที่เป็นจุดสีเขียว หมายถึงบริเวณที่ตัวรับรู้สามารถตรวจจับการเคลื่อนไหวได้ 100 เปอร์เซนต์ บริเวณที่เป็นจุดสีเหลือง หมายถึงบริเวณที่ตัวรับรู้สามารถตรวจจับการเคลื่อนไหวได้เป็นบางครั้ง และบริเวณที่เป็นจุดสีแดง หมายถึงบริเวณที่ตัวรับรู้ไม่สามารถทำการตรวจจับการเคลื่อนไหวได้ โดยแต่ละระดับมุม และระยะที่แตกต่างกัน จะทำการทดลองซ้ำองศาละ 30 ครั้ง และขยับระยะออกไปเรื่อย ๆ จนถึงขอบประตู ผลที่ได้แสดงให้เห็นว่าตัวรับรู้จะมีประสิทธิภาพการตรวจจับที่ดีที่สุดในระดับมุม 80°-100° ส่วนในระดับมุม 0° 20° 160° และ 180° ตัวรับรู้ไม่สามารถตรวจจับได้ อย่างไรก็ตาม เนื่องจากบานประตูมีขนาดไม่ใหญ่นัก การที่มนุษย์ต้องเคลื่อนตัวผ่านบริเวณที่ตรวจจับไม่ได้หรือตรวจจับได้บางครั้ง ซึ่งมีขนาดพื้นที่แคบ จึงเป็นเรื่องที่เป็นไปได้ยาก

5. สรุปผลการดำเนินงาน

การศึกษาวิจัยครั้งนี้ เป็นการวิจัยที่ประยุกต์เฟซบุ๊กเมสเซนเจอร์ร่วมกับตัวรับรู้พีไออาร์และบอร์ดราสเบอรีพาย เพื่อสร้างระบบรักษาความปลอดภัยในบ้าน ให้สามารถแจ้งเตือนเมื่อมีผู้บุกรุก โดยการทำงานของระบบคือ เมื่อพีไออาร์ตรวจจับความเคลื่อนไหวได้จะส่งค่า 1 ไปยังบอร์ด จากนั้นบอร์ดจะประมวลผล และส่งข้อความแจ้งเตือนผ่านเฟสบุ๊กเมสเซนเจอร์ นอกจากนี้ผลการวัดประสิทธิภาพของการใช้งานตัวรับรู้พีไออาร์หนึ่งตัวพบว่าสามารถตรวจจับการเคลื่อนไหวได้ค่อนข้างแม่นยำ อย่างไรก็ตาม การเพิ่มความสามารถให้ระบบที่พัฒนาขึ้นในอนาคต อาจเป็นไปในแนวทางที่ใช้ตัวรับรู้พีไออาร์มากกว่าหนึ่งตัว ติดตั้งในบริเวณใกล้กัน เพื่อทำงานร่วมกันในการตรวจจับ เพราะรัศมีที่ได้ในการตรวจจับของตัวรับรู้สองตัวจะกว้างกว่าการใช้ตัวรับรู้เพียงตัวเดียว โดยตัวรับรู้สองตัวสามารถทำงานร่วมกันโดยใช้บูลีนลอจิกออร์ (OR) ทำให้ระบบมีความแม่นยำในการตรวจจับเพิ่มมากขึ้น

เอกสารอ้างอิง

[1] dailynews.co.th. “โจรขึ้นบ้านผู้พิพากษา 6 ครั้ง ฉกทรัพย์ เกลี้ยงยังจับไม่ได้”. Internet: http://www.dailynews.co.th/regional/371044. [Mar. 24, 2017].

[2] dailynews.co.th. “รวบตืนแมวย่องลักทรัพย์ นับสิบครั้งหาเงินเสพยา”. Internet: http://www.dailynews.co.th/regional/504211. [Mar. 24, 2017].

[3] dailynews.co.th. “ตร.รวบโจรโชว์ห่วย ลักทรัพย์นับสิบล้าน”. Internet: http://www.dailynews.co.th/regional/526870. [Mar. 24, 2017].

[4] บุรินทร์ รุจจนพันธุ์. “เฟซบุ๊ก (อังกฤษ: Facebook)”. Internet: http://www.thaiall.com/facebook/.[Mar. 3, 2017].

[5] Avery Morin. “facebook-chat-api”. Internet: https://github.com /Schmavery/facebook-chat-api. [Aug. 25, 2016].

[6] Yogita V. Narkhede, Prof. S. G. Khadke, “Application of Raspberry PI and PIR Sensor for Monitoring of Smart Surveillance Systems,” *IJEECS*. vol. 5, no. 5, pp.145-148, 2016.

[7] Sanjana Prasad1, P. Mahalakshmi, A. John Clement Sunder, R. Swathi, “Smart Surveillance Monitoring System Using Raspberry PI and PIR Sensor,” *JCSIT*. vol. 5, no. 6, pp.7107-7109, 2014.

[8] Nikhil Gajjam, P. S. R. Patnaik, D. P. Gandhmal and V. A. Bingi, “Low Cost Surveillance Using Raspberry Pi through Email,” *IJSART*. vol. 2, no. 12, pp.11-14, 2016.

[9] Anabham Bhavani, Tulasi Jami, Gajjala Ashok, “Low Cost Smart Security Camera with Night Vision Capability Using Raspberry Pi and PIR Sensor,” *IJATIR*. vol. 8, no. 21, pp.4053-4056, 2016.

ระบบแลกเปลี่ยนข้อมูลสุขภาพส่วนบุคคลด้านผลตรวจทางห้องปฏิบัติการ
Personal Medical Record for Laboratory Data Exchange System

ปฏิพัทธ์ ตุ้มสังข์ทอง(Patipat Tumsangthong)[1] วิศิษฎ์ วงศ์วิไล(Wisit Wongwilai)[2] วีระชัย จันทร์สุด(Weerachai Chansud)[3]
นิธิภัทร ว่องชิงชัย(Nithipat Wongchingchai)[4] และอสมาภรณ์ ฉัตรัตติกรณ์(Asamaporn Chatrattikorn)[5]
ห้องปฏิบัติการวิจัยนวัตกรรมบริการอัจฉริยะ ศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ
[1]patipat.tumsangthong@nectec.or.th, [2]wisit.wongwilai@nectec.or.th, [3]weerachai.chansud@nectec.or.th,
[4]nithipat.wongchingchai@nectec.or.th, [5]asamaporn.chatrattikorn@nectec.or.th

บทคัดย่อ

ปัญหาการแลกเปลี่ยนข้อมูลผลตรวจทางห้องปฏิบัติการของผู้ป่วย ระหว่างระบบสารสนเทศโรงพยาบาลที่มีการใช้ฐานข้อมูล และสภาพแวดล้อมระบบที่แตกต่างกัน ทำให้ข้อมูลเมื่อถูกจัดเก็บไว้ในระบบสารสนเทศของโรงพยาบาลหนึ่งแล้วไม่สามารถนำไปใช้กับโรงพยาบาลอื่นได้ งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาระบบแลกเปลี่ยนข้อมูลสุขภาพส่วนบุคคลด้านผลตรวจทางห้องปฏิบัติการ (Personal Medical Record for Laboratory Data Exchange System: PMR-LabDES) ตามหลักสถาปัตยกรรมซอฟต์แวร์เชิงบริการ โดยใช้ Web Services ผ่าน Enterprise Service Bus (ESB) ในระบบ Cloud Computing มีการออกแบบโครงสร้างเอกสารอิเล็กทรอนิกส์ที่ใช้ในการรับส่งข้อมูลในรูปแบบ HL7 Clinical Document Architecture (HL7 CDA) ร่วมกับแนวทางการเชื่อมโยงข้อมูล Thailand e-Government Interoperability Framework (TH e-GIF) โดย PMR-LabDES นี้จะช่วยให้ระบบสารสนเทศของแต่ละโรงพยาบาลสามารถแลกเปลี่ยนข้อมูลผลตรวจทางห้องปฏิบัติการระหว่างกันได้ ผู้ป่วยสามารถสืบค้นข้อมูลผลตรวจย้อนหลังของตนเองจากหลายโรงพยาบาลได้ในที่เดียว โดยผลการทดสอบระบบฯ พบว่าผู้ป่วยและบุคลากรโรงพยาบาลมีความพึงพอใจต่อการใช้งานในระดับดี

คำสำคัญ: การแลกเปลี่ยนข้อมูล สถาปัตยกรรมเชิงบริการ ข้อมูลสุขภาพส่วนบุคคล

Abstract

The exchange problems of Patient's Laboratory test data between Hospital Information System with the different databases and environment of system, cause the data is stored in the hospital's information system can not be used with other hospitals. "Personal Medical Record for Laboratory Data Exchange System: PMR-LabDES)" work on Service Oriented Architecture (SOA) by using Web Services in Enterprise Service Bus (ESB) in the cloud computing system. An electronic document structure is used to transmit data in the form of HL7 Clinical Document Architecture (HL7 CDA) combined with Thailand e-Government Interoperability Framework (TH e-GIF)'s data linking. This PMR-LabDES will allow the Information System of each hospital can be exchanged the Patient's Laboratory test data. In summary, the results indicated that the all users' satisfaction was at good level.

Keyword: Data Exchange, Service Oriented Architecture, Personal Medical Record.

1. บทนำ

ปัจจุบันข้อมูลผลการตรวจรักษาของผู้ป่วย เช่น ข้อมูลการวินิจฉัยโรค ภาพถ่ายทางรังสี ผลตรวจทางห้องปฏิบัติการ ฯลฯ มักถูกจัดเก็บไว้ ณ โรงพยาบาลที่ผู้ป่วยเข้ารับบริการเท่านั้น โดยที่ผู้ป่วยไม่สามารถเก็บรวบรวมประวัติผลการตรวจรักษาของตนเอง หรือนำข้อมูลไปเป็นพื้นฐานในการเข้ารับบริการที่โรงพยาบาลอื่นได้โดยง่าย โดยเฉพาะอย่างยิ่งผู้ป่วยที่จำเป็นต้องเข้ารับการรักษาในต่างจังหวัดหรือต่างประเทศ ปัจจุบันเมื่อมีความต้องการผลการตรวจย้อนหลัง ผู้ป่วยจำเป็นต้องเดินทางไปแต่ละโรงพยาบาลเพื่อขอข้อมูลดังกล่าวด้วยตนเอง ทำให้เสียเวลา เสียค่าใช้จ่าย และเกิดความไม่สะดวกแก่ผู้ป่วย

ในส่วนระบบแลกเปลี่ยนข้อมูลระหว่างโรงพยาบาลที่มีการใช้งานอยู่ในปัจจุบันนั้น จะจำกัดเพียงให้สามารถแลกเปลี่ยนข้อมูลกันเฉพาะระหว่างระบบสารสนเทศเดียวกันหรือโดยบริษัทผู้พัฒนาเดียวกันเท่านั้น ด้วยเหตุผลด้านความแตกต่างของระบบฯ ทั้งในด้านฐานข้อมูล สภาพแวดล้อมการพัฒนาและการทำงานของระบบ ตลอดจนมาตรฐานข้อมูลที่ใช้ในแต่ละโรงพยาบาล จึงเป็นการยากที่จะนำข้อมูลมาบูรณาการรับส่งระหว่างกัน เพื่อให้สื่อสารเข้าใจตรงกันและนำไปใช้ประโยชน์ได้ อีกทั้งประเทศไทยยังขาดการกำหนดมาตรฐานในการแลกเปลี่ยนข้อมูลด้านการแพทย์และสุขภาพระหว่างระบบสารสนเทศโรงพยาบาล

เพื่อแก้ไขปัญหาดังกล่าวจึงได้มีการออกแบบและพัฒนาต้นแบบระบบแลกเปลี่ยนข้อมูลสุขภาพส่วนบุคคลด้านผลตรวจทางห้องปฏิบัติการ หรือ Personal Medical Record for Laboratory Data Exchange System (PMR-LabDES) เพื่อให้ข้อมูลผลตรวจทางห้องปฏิบัติการของผู้ป่วยนำไปสู่การเป็นข้อมูลสุขภาพส่วนบุคคลมากยิ่งขึ้น สามารถสืบค้นข้อมูลประวัติผลตรวจย้อนหลังของตนเองจากหลายโรงพยาบาลได้ในที่เดียว แสดงรายงาน และแลกเปลี่ยนข้อมูลผลการตรวจของผู้ป่วยระหว่างโรงพยาบาลได้ งานวิจัยนี้จะกล่าวถึงกระบวนการและวิธีการ ตลอดจนมาตรฐานและโครงสร้างเอกสารอิเล็กทรอนิกส์ที่ใช้ในการแลกเปลี่ยนข้อมูล รวมถึงการออกแบบและพัฒนาระบบ PMR-LabDES ที่เป็นเว็บแอพลิเคชั่นเพื่ออำนวยความสะดวกให้แก่ผู้ป่วยและบุคลากรที่เกี่ยวข้องของโรงพยาบาลด้วย

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 HL7 CDA

HL7 Clinical Document Architecture (HL7 CDA) เป็นมาตรฐานของเอกสารที่กำหนดโครงสร้าง (Structure) และความหมาย (Semantics) ของ Clinical Documents หรือเอกสารทางการแพทย์ และกำหนดมาตรฐานในการแลกเปลี่ยนเอกสารดังกล่าวให้เป็นสากล ซึ่งเกี่ยวข้องกับการวินิจฉัย การตรวจ และการให้บริการทางการแพทย์ [1]

P Marcheschi, A Mazzarisi, S Dalmiani และ A Benassi [2] พัฒนาระบบรายงานข้อมูลโรคหัวใจ (Cardiology) เพื่อทดสอบมาตรฐานใหม่ทางด้านสุขภาพ (health standards) ในการแลกเปลี่ยนข้อมูลทางคลินิกโดยใช้ HL7 Clinical Document Architecture Release2 และมาตรฐาน ebXML 3.0 (Electronic Business Extensible Markup Language) ซึ่งพัฒนาเป็น Web Services เพื่อให้สามารถแลกเปลี่ยนข้อมูลที่มีความหลากหลายของแพลตฟอร์ม โดยระบบฯ สามารถเชื่อมโยงข้อมูลจากระบบ Magnetic Resonance, Digital Angiography device หรือระบบ DICOM อื่นๆ เพื่อแสดงภาพผลตรวจทางรังสีได้

2.2 TH e-GIF

กรอบแนวทางการเชื่อมโยงรัฐบาลอิเล็กทรอนิกส์แห่งชาติ หรือ Thailand e-Government Interoperability Framework (TH e-GIF) เป็นชุดของข้อเสนอแนะ แนวทางการพัฒนาระบบมาตรฐานกลาง ที่กำหนดรายการข้อมูล ข้อกำหนดเอกสารอิเล็กทรอนิกส์ และมาตรฐานทางเทคนิคที่ใช้ในการเชื่อมโยงระหว่างระบบสารสนเทศที่มีความแตกต่างกัน [3]

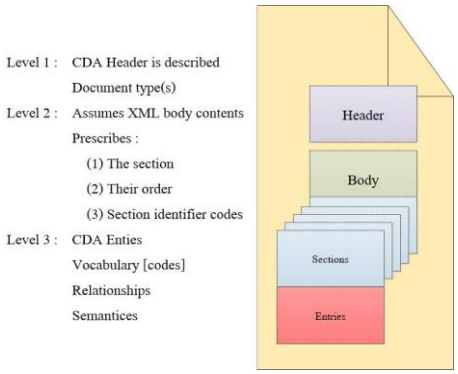
2.3 Enterprise Service Bus (ESB)

ESB เป็นโครงสร้างพื้นฐานสำคัญในการขับเคลื่อนสถาปัตยกรรมซอฟต์แวร์เชิงบริการ (SOA) ซึ่งเป็น platform มาตรฐานที่ใช้ในการรวบรวมหรือสนับสนุนให้องค์ประกอบบริการต่างๆในระบบ เช่น messaging, web services, data transformation หรือ intelligent routing สามารถทำงานร่วมกันได้และมีการเชื่อมโยงอย่างมีประสิทธิภาพ โดย ESB สามารถจัดการ connectivity, data transformation, routing, service, security, monitoring and logging [4]

3. วิธีดำเนินงานวิจัย

3.1 ออกแบบโครงสร้างเอกสารอิเล็กทรอนิกส์

เอกสารอิเล็กทรอนิกส์ที่ใช้ในการแลกเปลี่ยนข้อมูลผลตรวจทางห้องปฏิบัติการ มีการออกแบบโดยใช้โครงสร้างตามมาตรฐาน HL7 CDA® Release 2 ในรูปแบบไฟล์ XML ซึ่ง HL7 CDA นี้แบ่งโครงสร้างเป็น 2 ส่วนได้แก่ ส่วน Header และ Body ดังภาพที่ 1 และตัวอย่างเอกสารแสดงดังภาพที่ 2



ภาพที่ 1: โครงสร้างเอกสารอิเล็กทรอนิกส์ HL7 CDA® Release 2

```
<?xml version="1.0" encoding="UTF-8"?>
- <ClinicalDocument xsi:schemaLocation="urn:hl7-org:v3 CDA.xsd"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xmlns:voc="urn:hl7-
org:v3/voc" xmlns="urn:hl7-org:v3">
    <typeId extension="POCD_HD000040" root="2.16.840.1.113883.1.3"/>
    <id displayable="false"/>
    <code displayName="Lab Report"/>
    <effectiveTime value="20131030160000"/>
    <confidentialityCode codeSystem="2.16.840.1.113883.5.25" code="N"/>
    <setId extension="11470.5189086.HE001" root="DocumentID"/>
    <versionNumber value="1"/>
  + <recordTarget>
  + <author>
  + <custodian>
  - <component>
      - <structuredBody>
          + <component>
        </structuredBody>
    </component>
</ClinicalDocument>
```

ภาพที่ 2: ตัวอย่างเอกสาร HL7 CDA ส่วน Header

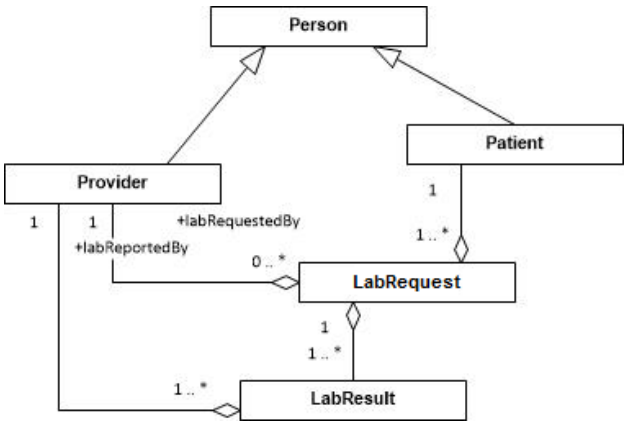
ส่วนที่ 1 Header เป็นข้อมูลทั่วไปที่อธิบายประเภทของเอกสาร การจัดสร้าง ผู้รับผิดชอบ ปลายทางที่ต้องการส่ง และข้อมูลอื่นๆ ตามจุดประสงค์ของผู้สร้าง ตามโครงสร้าง XML HL7 CDA Normative Schema

ส่วนที่ 2 Body แสดงข้อมูลทาง Clinical เพื่อให้สามารถนำไปใช้งานเบื้องต้นได้ แต่อาจจะยังไม่สามารถใช้งานประมวลผลข้ามระบบ โดยส่วนนี้ เรียกว่า “human readable” ซึ่งผู้ใช้เอกสารสามารถเข้าใจได้ โดยไม่ต้องแปลผลด้วยเครื่องคอมพิวเตอร์

ส่วนที่ 3 เป็นส่วนเพิ่มเติมซึ่งออกแบบมาเพื่อให้สามารถมีการประมวลผลข้ามระบบได้ โดยการกำหนดโครงสร้างและความหมายของข้อมูล ส่วนนี้เรียกว่า “machine processable” สามารถเข้าใจได้ก็ต่อเมื่อผ่านการประมวลผลด้วยคอมพิวเตอร์ และนำไปจัดเก็บในฐานข้อมูลของระบบสารสนเทศเพื่อนำไปใช้งานต่อไปได้ ในส่วนนี้ทางผู้วิจัยเลือกใช้โครงสร้างเอกสารตามกรอบแนวทาง TH-eGIF เพื่อระบุมาตรฐานที่ใช้ของแต่ละ element โดยสามารถกำหนด XML Namespace ที่แตกต่างกันได้ คือ 1) elements ที่อยู่ในมาตรฐาน HL7 จะถูกกำหนดให้อยู่ภายใต้ XML Namespace xmlns=”urn:hl7-org:v3” เป็น default namespace และ 2) elements ที่อยู่ในมาตรฐาน TH-eGIF จะถูกกำหนดให้อยู่ภายใต้ XML Namespace xmlns:th-eGIF=”urn:th.gov.th-eGIF:v1” (แสดงในส่วน machine processable)

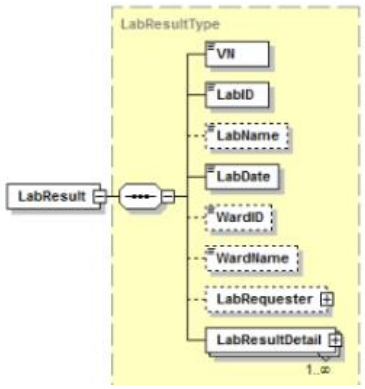
3.2 ออกแบบฐานข้อมูล

โครงสร้างข้อมูลระบบรายงานผลตรวจทางห้องปฏิบัติการประกอบด้วย ข้อมูลของผู้ป่วย (Patient) ข้อมูลการสั่งตรวจและข้อมูลผลการตรวจทางห้องปฏิบัติการ (LabRequest & LabResult) ซึ่งในแต่ละผลการตรวจจะมีส่วน Provider เป็นข้อมูลของผู้ให้บริการ ซึ่งได้แก่ แพทย์ผู้สั่งตรวจ ผู้ตรวจสอบผลการตรวจ และผู้อนุมัติผลการตรวจ ดังแสดงความสัมพันธ์ของกลุ่มข้อมูล ดังภาพที่ 4



ภาพที่ 4: ER-Diagram ระบบ PMR-LabDES

ในข้อมูล LabResult จะมีรายการข้อมูลผลตรวจทางห้องปฏิบัติการหลายๆ รายการย่อย (LabResultDetail) ของการสั่งตรวจแต่ละครั้ง (LabRequest) ดังภาพที่ 5 และเมื่อถูกใช้ในเอกสาร HL7 CDA จะถูกเรียกใช้ผ่าน elements ที่อยู่ในมาตรฐาน TH-eGIF

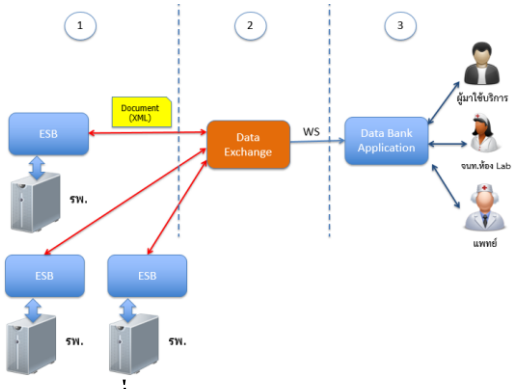



```
- <th-eGIF:labReport xmlns:th-eGIF="urn:th.gov.th-eGIF:v1">
  - <patient>
      <citizenID>1234567890123</citizenID>
    - <personName>
        <prefix>???</prefix>
        <givenName>abc</givenName>
        <middleName>hij</middleName>
        <familyName>defg</familyName>
      </personName>
      <birthdate>1984-06-29</birthdate>
      <birthJurisdictionCountrySubDivision/>
      <nationality/>
      <gender>M</gender>
      <religion/>
      <lifeStatus/>
      <telephone>0808080808</telephone>
      <email/>
      <hospCode>11470</hospCode>
      <age/>
      <hospitalNumber>22/51</hospitalNumber>
      <admissionNumber/>
      <bloodTypeABO/>
      <bloodTypeRh/>
      <preDiagnosis/>
      <insType/>
    </patient>
  - <labResult>
      <VN>5189086</VN>
      <labID>HE001</labID>
      <labName/>
      <labDate/>
      <wardID>7508</wardID>
      <wardName/>
    - <labRequester>
        <citizenID/>
      - <personName>
```

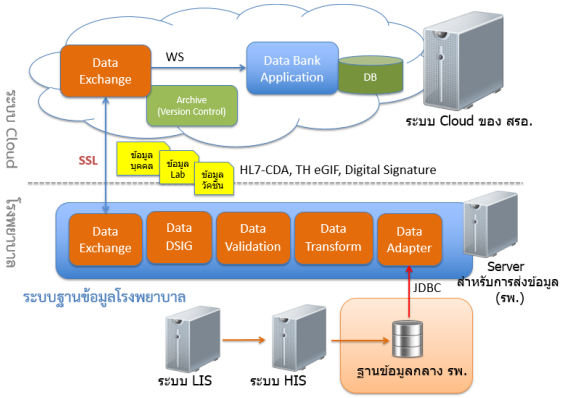
ภาพที่ 5: ส่วนประกอบข้อมูล LabResult และ LabResultDetail

3.3 ออกแบบระบบในภาพรวมและสถาปัตยกรรมระบบ

ระบบมีการออกแบบให้รองรับการเชื่อมโยงกันในแต่ละระบบสารสนเทศโรงพยาบาล ให้สามารถแลกเปลี่ยนข้อมูลระหว่างกันได้ โดยผู้ป่วยสามารถเข้าถึงข้อมูลผลตรวจทางห้องปฏิบัติการของตนเองจากหลายโรงพยาบาลได้ในที่เดียวผ่านทางเว็บแอปพลิเคชั่น มีส่วนประกอบดังภาพที่ 6



ภาพที่ 6: แสดงส่วนประกอบหลักของระบบ



ภาพที่ 7: แสดงรายละเอียดสถาปัตยกรรมภายในของระบบ

ในแต่ละระบบสารสนเทศของโรงพยาบาลจะมีการส่งออกข้อมูลและสร้างเอกสารอิเล็กทรอนิกส์ตามมาตรฐานที่กำหนดเพื่อให้สามารถแลกเปลี่ยนข้อมูลระหว่างกันได้ โดยใช้เทคโนโลยี Enterprise Service Bus (ESB) ในการควบคุมการเชื่อมโยงระหว่างเว็บเซอร์วิสและคอมโพเนนต์ต่างๆ เข้าด้วยกัน ดังภาพที่ 7 ประกอบด้วย

1) Data Adapter ใช้เชื่อมโยงและควบคุมการส่งผ่านข้อมูลกับฐานข้อมูลของระบบสารสนเทศโรงพยาบาล จะเป็นส่วนที่ต้องสามารถปรับรูปแบบเทคโนโลยีให้เหมาะสมกับบริบทของแต่ละโรงพยาบาลได้ เช่น นโยบายของผู้บริหาร/ผู้ดูแลระบบสารสนเทศโรงพยาบาล เทคโนโลยีที่ใช้อยู่ในปัจจุบัน ความสามารถของบุคลากรในการจัดการข้อมูลและความปลอดภัยของข้อมูล เป็นต้น ซึ่งรูปแบบที่ใช้ในงานวิจัยนี้จะใช้การติดต่อกับฐานข้อมูลที่ทางโรงพยาบาลจัดเตรียมไว้ให้ส่งออกข้อมูลตามเวลาที่กำหนด ซึ่งควรจะรองรับการเชื่อมต่อทั้งในรูปแบบแฟ้มข้อมูล (File System) และรองรับโปรโตคอล HL7 version 2 ได้ด้วย

2) Data Transformer ใช้แปลงข้อมูลและสร้างเอกสารมาตรฐาน HL7 CDA อ้างอิง TH e-GIF เป็นส่วนของการจัดรูปแบบข้อมูลให้อยู่ในโครงสร้างที่ต้องการ โดยงานวิจัยนี้ได้เลือกใช้มาตรฐาน HL7 CDA ที่เป็นมาตรฐานการรับส่งข้อมูลทางการแพทย์ ที่ง่ายต่อการนำไปแสดงผล (Human Readable) และเพิ่มส่วนการแลกเปลี่ยนข้อมูลตามกรอบแนวทาง TH e-GIF ซึ่งออกแบบมาเพื่อใช้สำหรับการแลกเปลี่ยนข้อมูลระหว่างหน่วยงานภาครัฐของประเทศไทย เพื่อให้ข้อมูลมีความสะดวกในการใช้งานโดยไม่ต้องทำการ

แปลงข้อมูลหลายๆ แบบเมื่อต้องการแลกเปลี่ยนข้อมูลไปยังหน่วยงานอื่น เช่น ข้อมูลบุคคลของกรมการปกครอง เป็นต้น

3) Data Validator ใช้ตรวจสอบความถูกต้องและรับรองเอกสารอิเล็กทรอนิกส์ โดยจะตรวจสอบโครงสร้างของเอกสารที่ได้จากส่วน Data Transformer ที่อยู่ในรูปแบบเอกสาร XML ด้วย XML Schema เพื่อให้แน่ใจว่าเอกสารที่สร้างขึ้นสามารถนำไปใช้งานได้ และระบบจะรับรองเอกสารอิเล็กทรอนิกส์ด้วยลายมือชื่อดิจิตอล (Digital Signature) ตามมาตรฐาน XML Signature แบบ Enveloped Signatures เพื่อเป็นการยืนยันว่าเอกสารนั้นมาจากโรงพยาบาลต้นทางที่ถูกต้องจริง สามารถพิสูจน์และตรวจสอบย้อนกลับได้

4) Data Exchange ใช้รับส่งข้อมูลเอกสารมาตรฐานจากโรงพยาบาลเข้าสู่ระบบฐานข้อมูลกลาง ซึ่งมีการเข้ารหัสเอกสารระหว่างการส่ง โดยในส่วนของการแลกเปลี่ยนข้อมูลได้ใช้เทคโนโลยีของ Web Services ด้วย Message Transmission Optimization Mechanism (MTOM) และได้ทำการบีบอัดข้อมูลเอกสาร XML และแนบเอกสารที่บีบอัดแล้วในรูปแบบ Binary แทนรูปแบบ Base64 ซึ่งจะทำให้ลดแบนด์วิดท์การใช้อินเตอร์เน็ตของโรงพยาบาลและเพิ่มความสำเร็จในการรับส่งข้อมูล ถ้าการรับส่งข้อมูลไม่สำเร็จระบบจะทำการส่งเอกสารเดิมซ้ำ และในการรับส่งได้มีการใช้โปรโตคอล SSL ในการเข้ารหัสข้อมูลระหว่างการรับส่งเพื่อเพิ่มความปลอดภัยของข้อมูลด้วย

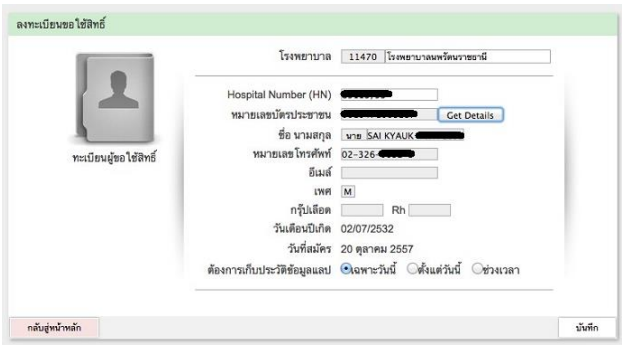
5) Data Monitor ใช้จัดการตรวจสอบและเฝ้าระวังสถานการณ์ทำงานของระบบ โดยจะทำการสร้าง Log การทำงานของระบบ ESB และสามารถตรวจสอบสถานะของระบบ ผ่านทาง Web Services

การแลกเปลี่ยนข้อมูลที่อยู่บนคลาวด์นั้นใช้เทคโนโลยีของ ESB เช่นเดียวกับฝั่งโรงพยาบาลแต่ใช้ในจุดประสงค์เพื่อเป็นตัวรับข้อมูลจากโรงพยาบาลต่างๆ มาเก็บไว้ที่คลาวด์ โดยมีฟังก์ชั่นหลัก ได้แก่ 1) คลายการบีบอัดข้อมูล 2) ตรวจสอบลายมือชื่อของเอกสารว่ามาจากโรงพยาบาลถูกต้องหรือไม่ ถ้าเอกสารถูกต้องจะถูกส่งไปยังระบบแสดงผลและจัดเก็บเอกสารไว้เป็นหลักฐาน และ 3) ส่วนการแสดงผลและติดต่อกับผู้ใช้งาน (User Interface)

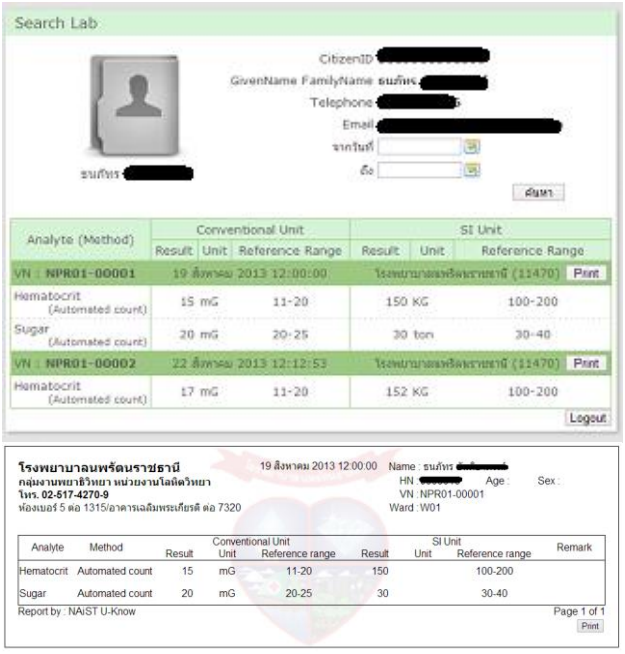
4. ผลการดำเนินงาน

4.1 ผลการพัฒนาระบบ

การพัฒนาเว็บแอพลิเคชั่นใช้ภาษา PHP และฐานข้อมูล MySQL สำหรับการจัดการข้อมูลผู้ใช้งานระบบ การแสดงผลข้อมูลผลตรวจทางห้องปฏิบัติการ และการแชร์ข้อมูลผลการตรวจ ตัวอย่างหน้าจอระบบ ดังภาพที่ 8 แสดงหน้าจอการขอแสดงความจำนงอนุญาตให้แลกเปลี่ยนข้อมูลผลการตรวจทางห้องปฏิบัติการของผู้ป่วยได้ และภาพที่ 9 แสดงผลการตรวจทางห้องปฏิบัติการและสามารถจัดพิมพ์ผลการตรวจด้วยตนเองได้



ภาพที่ 8: หน้าจอแสดงความจำนงอนุญาตให้แลกเปลี่ยนข้อมูล



ภาพที่ 9: หน้าจอแสดงผลการตรวจและจัดพิมพ์ผลการตรวจ

4.2 การทดสอบระบบ

ผู้วิจัยได้สร้างกรณีการทดสอบ และดำเนินการทดสอบตามฟังก์ชั่น (Functional Testing) ได้ผลการทดสอบดังภาพที่ 10

ที่	ผลการทดสอบ scenario1	ผลทดสอบ
1	[ST-DB-001] ทดสอบการลงทะเบียนขอใช้สิทธิ์ (Registering) แบบขอใช้สิทธิ์เฉพาะวันนี้ และใช้วิธีสร้างบัญชีผู้ใช้	ผ่าน
2	[ST-DB-006] ทดสอบขั้นตอนการเข้าใช้ระบบฐานข้อมูลสุขภาพ	ผ่าน
3	[ST-DB-007] ทดสอบขั้นตอนเรียกดูผลตรวจห้องปฏิบัติการจากระบบฐานข้อมูลสุขภาพ	ผ่าน
ที่	ผลการทดสอบ scenario2	ผลทดสอบ
1	[ST-DB-003] ทดสอบการลงทะเบียนขอใช้สิทธิ์ (Registering) แบบขอใช้สิทธิ์เฉพาะวันนี้ และยืนยันสิทธิ์ผ่าน E-Mail	ผ่าน
2	[ST-DB-005] ทดสอบขั้นตอนการยืนยันสิทธิ์ผ่าน E-Mail	ผ่าน
3	[ST-DB-006] ทดสอบขั้นตอนการเข้าใช้ระบบฐานข้อมูลสุขภาพ	ผ่าน
4	[ST-DB-007] ทดสอบขั้นตอนเรียกดูผลตรวจห้องปฏิบัติการจากระบบฐานข้อมูลสุขภาพ	ผ่าน
ที่	ผลการทดสอบ scenario3	ผลทดสอบ
1	[ST-DB-006] ทดสอบขั้นตอนการเข้าใช้ระบบฐานข้อมูลสุขภาพ	ผ่าน
2	[ST-DB-008] ทดสอบขั้นตอนแก้ไขข้อมูลผลตรวจห้องปฏิบัติการจากระบบฐานข้อมูลสุขภาพ	ผ่าน

ภาพที่ 10: แสดงกรณีการทดสอบและผลการทดสอบระบบ

4.3 การประเมินความพึงพอใจในการใช้ระบบ

จากผลการประเมินความพึงพอใจในการใช้ระบบ PMR-LabDES จากผู้ใช้ 2 กลุ่ม ได้แก่ 1) ผู้ป่วย ประเมินในส่วนการใช้งานระบบฯ เพื่อลงทะเบียนและดูผลการตรวจของตนเอง และ 2) บุคลากรโรงพยาบาล ประเมินในส่วนการใช้งานระบบฯ เพื่อบริหารจัดการข้อมูลผู้ป่วย ข้อมูลการสั่งตรวจและผลการตรวจ ทั้งสองกลุ่มผู้ใช้รวมจำนวน 30 คน ของโรงพยาบาลนพรัตนราชธานีเป็นโรงพยาบาลนำร่อง สรุปผลได้ดังตารางที่ 1

ตารางที่ 1: สรุปผลการประเมินความพึงพอใจในการใช้ระบบ

หัวข้อการประเมิน	$\bar{x}$	S.D.	ระดับคุณภาพ
1. ด้านความสะดวกในการใช้งาน	4.27	0.61	ดี
2. ด้านลดเวลาในกระบวนการ	4.26	0.64	ดี
3. ด้านคุณภาพที่ดีขึ้นในการให้บริการ	4.21	0.63	ดี
รวม	4.25	0.62	ดี

จากผลสรุปการประเมินจากผู้ป่วยและบุคลากรโรงพยาบาล พบว่า ผลการประเมินความพึงพอใจอยู่ในระดับดี ซึ่งค่าเฉลี่ยโดยรวม เท่ากับ 4.25 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.62

5. สรุปผล

งานวิจัยนี้นำเสนอแนวทางในการพัฒนาระบบเพื่อแก้ปัญหาและลดข้อจำกัดของการแลกเปลี่ยนข้อมูลผลตรวจทางห้องปฏิบัติการของผู้ป่วย ระหว่างระบบสารสนเทศโรงพยาบาลที่มีความแตกต่างกันให้สามารถเชื่อมโยงกันได้อย่างมีประสิทธิภาพ ผ่านระบบแลกเปลี่ยนข้อมูลสุขภาพส่วนบุคคลด้านผลตรวจทางห้องปฏิบัติการ (PMR-LabDES) ที่มีการออกแบบโครงสร้างเอกสารอิเล็กทรอนิกส์ที่เอื้อต่อการแลกเปลี่ยนข้อมูลโดยใช้แนวคิดการรับส่งข้อมูลในรูปแบบเอกสาร HL7 Clinical Document Architecture (HL7-CDA) ร่วมกับกรอบแนวทางการเชื่อมโยงข้อมูล Thailand e-Government Interoperability Framework (TH e-GIF) ตามหลักสถาปัตยกรรมซอฟต์แวร์เชิงบริการ ผลการทดสอบและการประเมินความพึงพอใจระบบฯ พบว่า ผู้ป่วยและบุคลากรโรงพยาบาล มีความพึงพอใจต่อการใช้งานในการดูประวัติข้อมูลผลตรวจและออกรายงานต่างๆ ในระดับดี ทั้งนี้ PMR-LabDES ยังคงต้องมีการพัฒนาเพิ่มเติมในส่วนการติดตามเพื่อตรวจสอบสถานะการทำงานของเครื่องแม่ข่ายที่ให้บริการ Web services รวมถึงการพัฒนาระบบแจ้งเตือน (Alert System) ความผิดพลาดต่างๆ เพื่อให้ง่ายต่อการดูแลรักษาระบบ นอกจากนี้ผู้วิจัยยังมีแนวคิดที่จะขยายผลการนำเอกสารอิเล็กทรอนิกส์ตามโครงสร้าง HL7-CDA ไปใช้ร่วมกับกรอบแนวทางการเชื่อมโยงข้อมูลในเรื่องอื่นๆ ของไทยและสากลนอกเหนือจากกรอบแนวทาง TH-eGIF อีกด้วย

เอกสารอ้างอิง

[1] http://www.hl7.org/implement/standards/product_brief.cfm?product_id=7

[2] P Marcheschi, A Mazzarisi, S Dalmiani and A Benassi, “New Standards for Cardiology Report and Data Communication: An Experience with HL7 CDA Release 2 and EbXML” *Comput. Cardiol.* 383–386, 2005

[3] สมนึก คีรีโต, “ภาพรวมการปรับปรุงกรอบแนวทางเชื่อมโยงรัฐบาลอิเล็กทรอนิกสแห่งชาติ เวอรชั่น 2.0” สืบค้นเมื่อ 29 กรกฎาคม 2557.

[4] https://en.wikipedia.org/wiki/Enterprise_service_bus

[5] Supachai Parchariyanon, Kavin Asavanant, Sireerat Srisiriratanakul, Chaiwiwat Tongtaweechaikit, Nawanan Theera-Ampornpunt, Chusak Okaschareon and Artit Ungkanont, “Implementation of Thailand’s First Prototype for Exchanging of Laboratory Results Using HL7 Version 3 and LOINC” *AMIA Annu Symp Proc.* 2012.

การพัฒนาระบบจัดการข้อมูลภูมิปัญญาท้องถิ่นที่บูรณาการร่วมกับเส้นทางท่องเที่ยวพื้นที่จังหวัดนครปฐม ในรูปแบบข้อมูลเชิงพื้นที่บนระบบสารสนเทศภูมิศาสตร์

Development of local knowledge information management system integrated with Nakhon Pathom tourist routes in spatial data format on geographic information systems.

ไพศาล สิมาเลาเต่า (PAISAN SIMALAOTAO) [1] และ อุบลรัตน์ ศิริสุขโภคา (UBONRAT SIRISUKPOCA) [2]

สาขาวิชาวิทยาการคอมพิวเตอร์[1] สาขาวิชาคอมพิวเตอร์(ค.บ.)[2] คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏนครปฐม

Paisan.smlt@gmail.com, Ubonrat76@gmail.com

บทคัดย่อ

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อ 1) พัฒนาระบบจัดการข้อมูลภูมิปัญญาท้องถิ่นที่บูรณาการร่วมกับเส้นทางท่องเที่ยวพื้นที่จังหวัดนครปฐม ในรูปแบบข้อมูลเชิงพื้นที่บนระบบสารสนเทศภูมิศาสตร์ 2) ประเมินประสิทธิภาพและความเหมาะสมของระบบโดยผู้เชี่ยวชาญ 3) ประเมินความพึงพอใจของผู้ใช้ กลุ่มตัวอย่างในการวิจัย ได้แก่ ผู้ใช้ทั่วไปในสถานที่ท่องเที่ยวโดยวิธีการสุ่ม จำนวน 80 คน เครื่องมือในการวิจัยประกอบด้วย 1) ระบบจัดการข้อมูลภูมิปัญญาท้องถิ่นที่บูรณาการร่วมกับเส้นทางท่องเที่ยวพื้นที่จังหวัดนครปฐม ในรูปแบบข้อมูลเชิงพื้นที่บนระบบสารสนเทศภูมิศาสตร์ ที่สามารถเข้าใช้งานด้วยอุปกรณ์พกพาบนระบบปฏิบัติการแอนดรอยด์ ซึ่งประกอบด้วย 3 โมดูล คือ โมดูลข้อมูล โมดูลแผนที่ และโมดูลแสดงผล 2) แบบประเมินประสิทธิภาพและความเหมาะสมของระบบ 3) แบบประเมินความพึงพอใจของผู้ใช้ระบบ ผลการวิจัยพบว่า การประเมินประสิทธิภาพของระบบโดยผู้เชี่ยวชาญ จำนวน 3 คน มีผลอยู่ในระดับ “ดีมาก” ($\bar{X}$ = 4.56, S.D. = 0.48) มีความเหมาะสมในการใช้งานทุกด้าน เมื่อนำไปทดลองใช้กับกลุ่มตัวอย่างโดยการสุ่มจากผู้ใช้งานอุปกรณ์พกพาระบบปฏิบัติการแอนดรอยด์และประเมินความพึงพอใจจากผู้ใช้ พบว่า ผู้ใช้มีความพึงพอใจทุกด้านอยู่ในระดับดีมาก ($\bar{X}$ = 4.57, S.D. = 0.50) โดยมีความพึงพอใจความสามารถในด้านการทำงานของระบบ ($\bar{X}$ = 4.60, S.D. = 0.51) และด้านการออกแบบหน้าจอ ($\bar{X}$ = 4.66, S.D. = 0.50) ผู้ใช้มีความพึงพอใจอยู่ในระดับดีมาก และด้านการทดสอบการใช้งาน ($\bar{X}$ = 4.46, S.D. = 0.50) ผู้ใช้มีความพึงพอใจอยู่ในระดับดี

คำสำคัญ: *ภูมิปัญญาท้องถิ่น เส้นทางท่องเที่ยว ระบบสารสนเทศภูมิศาสตร์*

Abstract

This research aims to 1) Development of local knowledge information management system integrated with Nakhon Pathom tourist routes in spatial data format on geographic information systems. 2) evaluate performance and appropriability of the system by the experts. 3) evaluate user satisfaction. The research sample consisted of 80 random users in the tourist destinations. Research tool include 1) Local Wisdom Information Management System integrated with Nakhon Pathom Tourist Route In spatial data format on GIS Accessible with android portable devices consists of 3 modules: Data Module, Map API Module, and Presentation Module, 2) evaluate the performance and suitability of the system, 3) user satisfaction assessment form. The result of research is performance evaluation by 3 experts were in "very good" ($\bar{X}$ =4.56, S.D. = 0.48). Are suitable for use on all sides. When applied to the sample by randomly from an android mobile device user and evaluate the user satisfaction found. The user is satisfied all aspects is very good ($\bar{X}$ = 4.57, S.D. = 0.50), The satisfaction of the ability of the system ($\bar{X}$ = 4.60, S.D. = 0.51) and the aspect of the screen design. ($\bar{X}$ =4.66, S.D. = 0.50). The user satisfaction is very good. And the use of testing ($\bar{X}$ =4.46, S.D. = 0.50). The users are satisfied at a good level.

Keyword: *local knowledge, travel route, geographic information system*

1. บทนำ

สังคมไทยเป็นสังคมที่มีการประยุกต์ใช้ความรู้จากสิ่งแวดล้อมและธรรมชาติในการดำรงชีวิต จนเกิดเป็นภูมิปัญญาชาวบ้านหรือภูมิปัญญาท้องถิ่น แต่เนื่องจากการโยกย้ายของประชากรจากท้องถิ่นเข้าสู่เมืองเพื่อทำงาน ทำให้ภูมิปัญญาท้องถิ่นบางอย่างขาดผู้สืบทอดและกำลังจะสูญหายไป จึงมีความจำเป็นอย่างยิ่งในการอนุรักษ์ภูมิปัญญาท้องถิ่นไว้ อีกทั้งภูมิปัญญาท้องถิ่นในหลาย ๆ ด้านยังสามารถสร้างรายได้ให้กับคนในท้องถิ่นทำให้เกิดการสร้างงานและอาชีพได้เป็นอย่างดี แต่ขาดการบริหารจัดการและประชาสัมพันธ์ที่เหมาะสม ทำให้ภูมิปัญญาท้องถิ่นหรือผลิตภัณฑ์จากภูมิปัญญาท้องถิ่นหรือชุมชนต่าง ๆ ไม่เป็นที่รู้จักและถูกมองข้ามจากคนนอกพื้นที่ หากมีการนำข้อมูลภูมิปัญญาท้องถิ่นมาบูรณาการร่วมกับเส้นทางท่องเที่ยวจะช่วยให้บุคคลทั่วไปสามารถเข้าถึงข้อมูลภูมิปัญญาท้องถิ่นได้สะดวกขึ้น

การพัฒนาระบบจัดการข้อมูลภูมิปัญญาท้องถิ่นที่บูรณาการร่วมกับเส้นทางท่องเที่ยวพื้นที่จังหวัดนครปฐม ในรูปแบบข้อมูลเชิงพื้นที่บนระบบสารสนเทศภูมิศาสตร์ เป็นการจัดเก็บข้อมูลเส้นทางการท่องเที่ยวของจังหวัดนครปฐมไว้ จำนวน 4 เส้นทาง ได้แก่ 1) เส้นทางท่องเที่ยวในอำเภอนครชัยศรี 2) เส้นทางท่องเที่ยวในอำเภอสามพราน 3) เส้นทางท่องเที่ยวในอำเภอกำแพงแสน 4) เส้นทางท่องเที่ยวในอำเภอเมือง โดยจัดเก็บข้อมูลภูมิปัญญาท้องถิ่นรวม 7 ประเภท ได้แก่ 1) ด้านเกษตรกรรม 2) ด้านอุตสาหกรรมและหัตถกรรม 3) ด้านศิลปกรรม 4) ด้านปรัชญา ศาสนา และประเพณี 5) ด้านโภชนาการ 6) ด้านการแพทย์แผนไทย 7) ด้านการจัดการทรัพยากรธรรมชาติและสิ่งแวดล้อม เพื่อประโยชน์ในการส่งเสริมการท่องเที่ยวควบคู่กับการอนุรักษ์วัฒนธรรมและภูมิปัญญาท้องถิ่นตลอดจนการเผยแพร่ภูมิปัญญาท้องถิ่นผ่านสื่อออนไลน์ ทำให้ผู้ใช้สามารถเข้าถึงข้อมูลภูมิปัญญาท้องถิ่นบนเส้นทางท่องเที่ยวได้อย่างสะดวกและรวดเร็ว

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

จากแนวคิดการแก้ปัญหางานวิจัย มีทฤษฎีที่เกี่ยวข้อง ดังนี้

ระบบสารสนเทศภูมิศาสตร์ หรือ GIS : Geographic Information System เป็นการกำหนดข้อมูลเชิงพื้นที่ด้วยการระบุพิกัดทางภูมิศาสตร์โดยเก็บข้อมูลละติจูดและลองติจูดเพื่อกำหนดข้อมูลที่ไม่ใช่เชิงพื้นที่ หรือข้อมูลเชิงคุณลักษณะที่อธิบายรายละเอียดต่างๆของข้อมูลเชิงพื้นที่ เช่น ชุมชน แหล่งน้ำ เป็นต้น [1] ซึ่งสามารถใช้บริการ Google Maps API ของบริษัท Google เป็นเครื่องมือในการพัฒนาแอพพลิเคชันที่ใช้งานผ่านระบบเครือข่ายอินเทอร์เน็ต ซึ่งนำเสนอข้อมูลบนแผนที่ในลักษณะของการทำเครื่องหมายหรือสัญลักษณ์ลงในตำแหน่งละติจูดและลองติจูดที่ต้องการ (place marker) [2]

2) งานวิจัยที่เกี่ยวข้อง

การวิจัยต่อยอดภูมิปัญญาท้องถิ่นเพื่อพัฒนาแหล่งท่องเที่ยวและท้องถิ่นที่ยั่งยืน พบว่า ศักยภาพที่มีส่วนสำคัญได้แก่ ศักยภาพด้านทรัพยากรมนุษย์ ด้านผู้นำชุมชน ด้านการมีส่วนร่วมในกิจกรรมชุมชน ด้านเครือข่ายชุมชน โดยทรัพยากรด้านภูมิปัญญามีศักยภาพมากที่สุด รูปแบบการต่อยอดภูมิปัญญาเพื่อนำมาพัฒนาการท่องเที่ยว ได้แก่ การพัฒนาแหล่งท่องเที่ยว การสร้างเครือข่าย การพัฒนาผลิตภัณฑ์ การมีส่วนร่วม และการเรียนรู้ร่วมกัน [3]

การวิจัยเพื่อพัฒนาการจัดการท่องเที่ยวโดยชุมชนในจังหวัดยะลา พบว่า ศักยภาพด้านการท่องเที่ยว ได้แก่ สิ่งโดดเด่นของแหล่งท่องเที่ยวของชุมชน การจัดการด้านความปลอดภัย การจัดกิจกรรมและการมีส่วนร่วมของชุมชนในการจัดการท่องเที่ยว ซึ่งการมีส่วนร่วมของชุมชนที่สำคัญ ได้แก่ การพัฒนาด้านผลิตภัณฑ์ ด้านการจัดการ ด้านการตลอด และด้านทรัพยากรมนุษย์ จากการประเมินสภาพแวดล้อมและการจัดระบบท่องเที่ยวในชุมชนทำให้มีการท่องเที่ยวเพิ่มขึ้น [4]

การรวบรวมข้อมูลสำหรับสร้างฐานข้อมูลในงานระบบสารสนเทศภูมิศาสตร์ พบว่า ข้อมูลที่ใช้ในงานระบบสารสนเทศภูมิศาสตร์ ประกอบด้วย 1) ข้อมูลเชิงพื้นที่ เป็นข้อมูลที่สามารถใช้อ้างอิงตำแหน่งทางภูมิศาสตร์ได้อย่างถูกต้อง คือ จุด เส้น และขอบเขตพื้นที่ 2) ข้อมูลไม่ใช่เชิงพื้นที่ เป็นข้อมูลคุณลักษณะรายละเอียดของพื้นที่ ได้แก่ การถือครองที่ดิน สภาวะเศรษฐกิจและสังคม เป็นต้น ซึ่งมีความสำคัญต่อการจัดทำฐานข้อมูลเพื่อใช้ในการจัดการพื้นที่ การทราบถึงแหล่งที่มาช่วยให้สามารถรวบรวมข้อมูลที่แตกต่างกันได้อย่างเหมาะสมทำให้สามารถจัดการข้อมูลได้อย่างรวดเร็วและมีประสิทธิภาพ [5]

การพัฒนาระบบจัดเก็บข้อมูลด้วยโทรศัพท์ระบบปฏิบัติการแอนดรอยด์และสร้างข้อมูลเส้นทางการเดินรถด้วย Google

และสร้างเส้นทางการเดินรถ ซึ่งใช้เวลาในการคำนวณเส้นทางเฉลี่ย 5.75 ms ต่อหนึ่งเส้นทาง โดยนำผลไปใช้พัฒนาเอพีไอและโปรแกรมบนโทรศัพท์เคลื่อนที่ระบบปฏิบัติการแอนดรอยด์ พบว่า โปรแกรมประยุกต์ที่พัฒนาสามารถตอบกลับการทำงานภายในเวลา 4 วินาทีซึ่งผู้ใช้ยอมรับได้ โดยมีความพึงพอใจในระดับ 4.16 [6] การพัฒนาโมเดลการแพร่ของโรคระบาดบนระบบสารสนเทศทางภูมิศาสตร์ผ่านไดนามิคเว็บเซอร์วิส มีการใช้ข้อมูลตำแหน่งละติจูดและตำแหน่งลองติจูดในการแสดงผลร่วมกับข้อมูลเชิงพื้นที่บนแผนที่ภูมิศาสตร์ ซึ่งผลการประเมินคุณภาพของระบบจากผู้เชี่ยวชาญทุกด้านอยู่ในระดับดี โดยมีค่าเฉลี่ยอยู่ที่ 4.30 [7]

จากการค้นคว้าทฤษฎีและงานวิจัยที่เกี่ยวข้อง การพัฒนาระบบจัดการข้อมูลภูมิปัญญาท้องถิ่นที่บูรณาการร่วมกับเส้นทางท่องเที่ยวพื้นที่จังหวัดนครปฐม ในรูปแบบข้อมูลเชิงพื้นที่บนระบบสารสนเทศภูมิศาสตร์ จึงเป็นการพัฒนาเว็บแอพพลิเคชันด้วยภาษา PHP เชื่อมต่อกับฐานข้อมูล MySQL ซึ่งติดตั้งบนเว็บเซิร์ฟเวอร์ Apache โดยใช้บริการ Google Maps API เป็นเครื่องมือในการกำหนดตำแหน่งของภูมิปัญญาท้องถิ่นและแหล่งท่องเที่ยวบนแผนที่ภูมิศาสตร์และสามารถใช้บริการผ่านสมาร์ทโฟนทั้งในรูปแบบโมบายเว็บและแอนดรอยด์แอพพลิเคชัน

3. วิธีดำเนินการวิจัย

1) ประชากรและกลุ่มตัวอย่าง

ประชากร คือ ผู้ใช้ทั่วไปตามเส้นทางท่องเที่ยว ได้แก่ 1) เส้นทางท่องเที่ยวในอำเภอนครชัยศรี 2) เส้นทางท่องเที่ยวในอำเภอสามพราน 3) เส้นทางท่องเที่ยวในอำเภอกำแพงแสน 4) เส้นทางท่องเที่ยวในอำเภอเมือง โดยมีกลุ่มตัวอย่าง ได้แก่ ผู้ใช้ทั่วไปที่ใช้งานอุปกรณ์พกพาสมาร์ทโฟนบนระบบปฏิบัติการแอนดรอยด์ จำนวน 80 คน

2) เครื่องมือที่ใช้ในการวิจัย

1. ระบบจัดการข้อมูลภูมิปัญญาท้องถิ่นที่บูรณาการร่วมกับเส้นทางท่องเที่ยวพื้นที่จังหวัดนครปฐม ในรูปแบบข้อมูลเชิงพื้นที่บนระบบสารสนเทศภูมิศาสตร์ ที่สามารถเข้าใช้งานด้วยอุปกรณ์พกพาบนระบบปฏิบัติการแอนดรอยด์

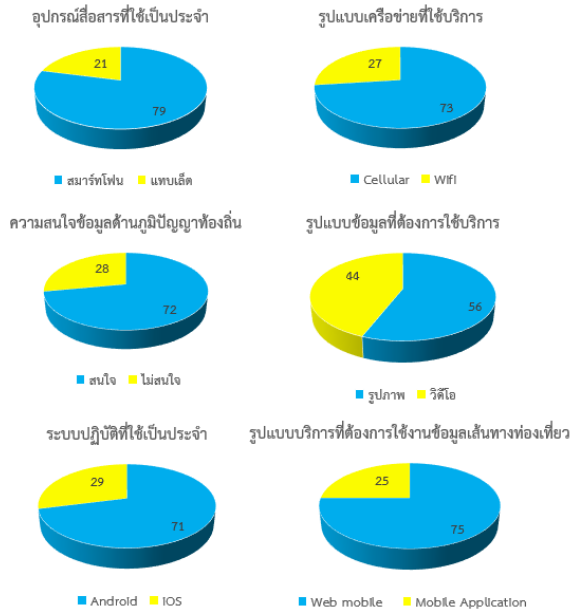
2. แบบประเมินประสิทธิภาพและความเหมาะสมของระบบ ที่ประเมินโดยผู้เชี่ยวชาญจำนวน 3 คน

3. แบบประเมินความพึงพอใจ 5 ด้าน หลังใช้งานระบบจัดการข้อมูลภูมิปัญญาท้องถิ่นที่บูรณาการร่วมกับเส้นทางท่องเที่ยวพื้นที่จังหวัดนครปฐม ในรูปแบบข้อมูลเชิงพื้นที่บนระบบสารสนเทศภูมิศาสตร์ ที่ประเมินโดยผู้ใช้ รวม 80 คน

3) ขั้นตอนการดำเนินการวิจัย

ใช้วิธีดำเนินการแบบ SDLC ซึ่งมีขั้นตอน ดังนี้

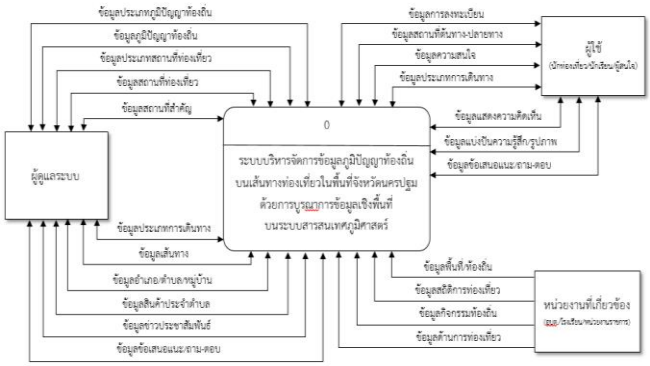
1. วิเคราะห์ความต้องการ (Requirement) ด้วยการแบ่งกลุ่มของผู้ใช้ออกเป็น 3 กลุ่ม ประกอบด้วย 1) ผู้ดูแลระบบทำหน้าที่บริหารจัดการข้อมูลภูมิปัญญาท้องถิ่น ข้อมูลสถานที่ท่องเที่ยว และข้อมูลเชิงพื้นที่บนแผนที่ภูมิศาสตร์ 2) หน่วยงานที่เกี่ยวข้อง ได้แก่ หน่วยงานราชการ ชุมชน เป็นต้น ทำหน้าที่จัดการข้อมูลสถานที่ท่องเที่ยว ข้อมูลภูมิปัญญาท้องถิ่น และข้อมูลในพื้นที่ 3) ผู้ใช้งานทั่วไป คือ นักท่องเที่ยวหรือผู้ที่เดินทางไปในพื้นที่ให้บริการ ผลจากการวิเคราะห์ความต้องการด้วยการสำรวจข้อมูลจากผู้ใช้ทั่วไปและประชาชนในพื้นที่สามารถแสดงได้ดังภาพที่ 1



ภาพที่ 1 ผลการสำรวจและวิเคราะห์ข้อมูลเบื้องต้น

จากภาพที่ 1 พบว่า ผู้ใช้ส่วนใหญ่ใช้อุปกรณ์พกพาเป็นสมาร์ทโฟนที่มีการทำงานบนระบบปฏิบัติการแอนดรอยด์ ซึ่งในระหว่างเดินทางจะใช้บริการผ่านเครือข่ายผู้ให้บริการโทรศัพท์มือถือ มีความสนใจข้อมูลด้านภูมิปัญญาท้องถิ่นในรูปแบบของรูปภาพพร้อมคำอธิบาย โดยสะดวกในการใช้บริการผ่านเว็บโมบายเนื่องจากสามารถให้ผลลัพธ์ได้สะดวกและรวดเร็ว

2. การวิเคราะห์และออกแบบระบบ (Analysis & Design) โดยนำผลจากการวิเคราะห์ข้อมูลเบื้องต้นมาออกแบบเป็นแผนภาพบริบท (Context diagrams) ดังตัวอย่างภาพที่ 2



ภาพที่ 2 แผนภาพบริบทการทำงานของระบบเบื้องต้น

จากภาพที่ 2 แสดงแผนภาพความสัมพันธ์ในการใช้บริการระบบจัดการข้อมูลภูมิปัญญาท้องถิ่นที่บูรณาการร่วมกับเส้นทางท่องเที่ยวพื้นที่จังหวัดนครปฐม ในรูปแบบข้อมูลเชิงพื้นที่บนระบบสารสนเทศภูมิศาสตร์ โดยผู้ใช้ 3 กลุ่ม

3. การพัฒนาระบบ (Development/coding)

ส่วนที่ 1 สำหรับผู้ดูแลระบบและหน่วยงานที่เกี่ยวข้อง แบ่งออกเป็น 1) พัฒนาระบบจัดการข้อมูลที่สามารถใช้บริการด้วยเครื่องคอมพิวเตอร์ส่วนบุคคลและโน้ตบุ๊คผ่านโปรแกรมเว็บบราวเซอร์ด้วยภาษา PHP เชื่อมต่อฐานข้อมูล MySQL และ 2) พัฒนาระบบจัดการข้อมูลที่สามารถใช้บริการด้วยอุปกรณ์พกพาด้วย JQuery ร่วมกับภาษา PHP เชื่อมต่อฐานข้อมูล MySQL ติดต่อกับแผนที่ภูมิศาสตร์ด้วย Google API และให้บริการผ่านระบบเครือข่ายอินเทอร์เน็ต

ส่วนที่ 2 สำหรับผู้ใช้ทั่วไป พัฒนาส่วนเข้าใช้บริการที่สามารถใช้งานด้วยอุปกรณ์พกพาสมาร์ทโฟน แบ่งออกเป็น 1) พัฒนาเว็บโมบาย (web mobile) ด้วย JQuery ร่วมกับภาษา PHP เชื่อมต่อฐานข้อมูล MySQL และ 2) พัฒนาแอนดรอยด์แอพพลิเคชัน (android application) ด้วย thunkable ร่วมกับภาษา PHP เชื่อมต่อฐานข้อมูล MySQL และแผนที่ภูมิศาสตร์

4. การทดสอบและบูรณาการระบบ (Testing/System integration) ประกอบด้วย 1) ทดสอบการทำงานของระบบย่อย ได้แก่ ทดสอบส่วนการจัดการข้อมูลภูมิปัญญาท้องถิ่น ทดสอบส่วนการจัดการข้อมูลสถานที่ท่องเที่ยว ทดสอบส่วนการจัดการข้อมูลเชิงพื้นที่บนแผนที่ภูมิศาสตร์ ทดสอบส่วนติดต่อกับผู้ใช้ผ่านอุปกรณ์พกพาสมาร์ทโฟน 2) บูรณาการระบบให้สามารถนำเสนอข้อมูลภูมิปัญญาท้องถิ่นและข้อมูลสถานที่ท่องเที่ยวบนระบบสารสนเทศภูมิศาสตร์

ผู้วิจัยนำระบบที่ผ่านการทดสอบและการบูรณาการระบบไปประเมินประสิทธิภาพและความเหมาะสมของระบบโดยผู้เชี่ยวชาญ จำนวน 3 คน เพื่อนำผลการประเมินและข้อเสนอแนะมาปรับปรุงก่อนนำไปทดลองจริง (ผลการประเมินดังตารางที่ 1)

5. การดำเนินงานและการบำรุงรักษา (Operation & maintenance) แบ่งการดำเนินงานเป็น 3 ขั้นตอน ประกอบด้วย 1) ขั้นตอนการนำเข้าข้อมูลเบื้องต้น ได้แก่ ข้อมูลภูมิปัญญาท้องถิ่น ข้อมูลสถานที่ท่องเที่ยว ข้อมูลเส้นทางท่องเที่ยว ข้อมูลเชิงพื้นที่ 2) ขั้นตอนการนำไปทดลองกับกลุ่มตัวอย่างด้วยวิธีการสุ่มผู้ใช้ที่เดินทางมายังสถานที่ท่องเที่ยวในพื้นที่จังหวัดนครปฐม จำนวน 80 คน โดยแบ่งเป็น 4 เส้นทาง คือ เส้นทางท่องเที่ยวในอำเภอนครชัยศรี เส้นทางท่องเที่ยวในอำเภอสามพราน เส้นทางท่องเที่ยวในอำเภอกำแพงแสน เส้นทางท่องเที่ยวในอำเภอเมือง และ 3) ขั้นตอนการประเมินความพึงพอใจโดยผู้ใช้งาน

5) การวิเคราะห์และการประเมินผล

การประเมินประสิทธิภาพของระบบและการประเมินความพึงพอใจใช้สถิติที่เกี่ยวข้อง ดังนี้ [9]

1. ค่าเฉลี่ยเลขคณิต (Arithmetic Mean) ดังสูตร

$$\bar{x} = \frac{\sum x}{n}$$

โดย $\bar{x}$ คือ ค่าเฉลี่ย
$\sum x$ คือ ผลรวมทั้งหมด
n คือ จำนวนประชากร

2. ค่าเบี่ยงเบนมาตรฐาน (Standard Deviation) ดังสูตร

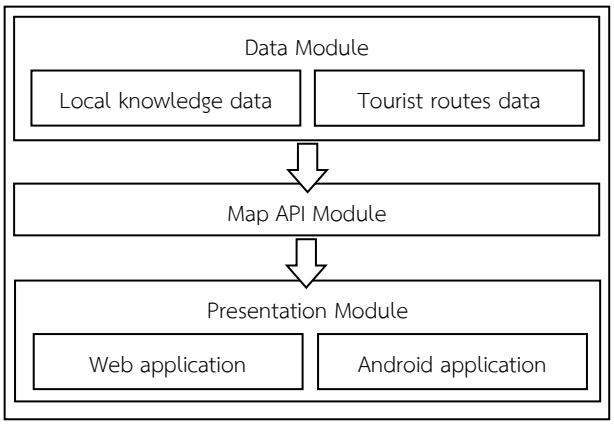
$$S.D. = \sqrt{\frac{n\sum x^2 - (\sum x)^2}{n(n-1)}}$$

โดย S.D. คือ ค่าเบี่ยงเบนมาตรฐาน
$\sum x$ คือ ผลบอกของค่าแต่ละตัว
n คือ จำนวนประชากร

4. ผลการดำเนินงาน

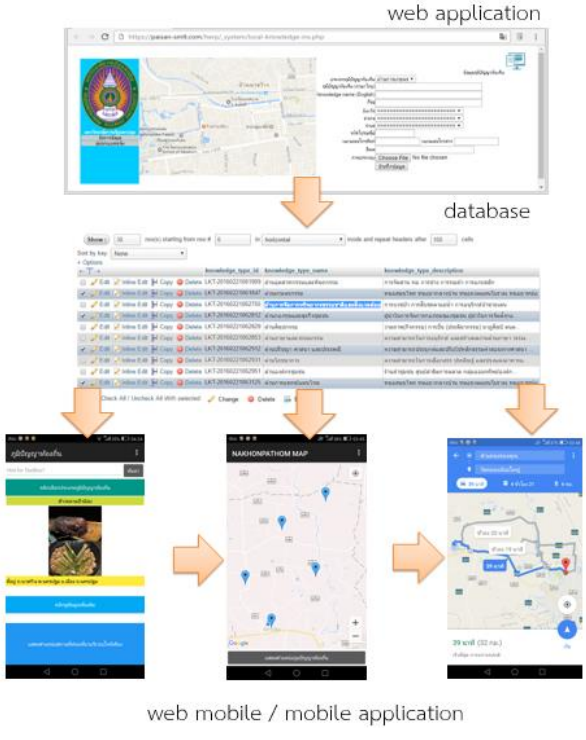
1) ผลการพัฒนาระบบ

ระบบจัดการข้อมูลภูมิปัญญาท้องถิ่นที่บูรณาการร่วมกับเส้นทางท่องเที่ยวพื้นที่จังหวัดนครปฐม ในรูปแบบข้อมูลเชิงพื้นที่บนระบบสารสนเทศภูมิศาสตร์ มีโมดูลระบบจัดการข้อมูลบนระบบสารสนเทศภูมิศาสตร์ ดังภาพที่ 3



ภาพที่ 3 โมดูลระบบจัดการข้อมูลบนระบบสารสนเทศภูมิศาสตร์

จากภาพที่ 3 โมดูลข้อมูล (Data Module) เป็นส่วนของการนำเข้าข้อมูลภูมิปัญญาท้องถิ่นและข้อมูลสถานที่ท่องเที่ยว เพื่อนำไปเชื่อมโยงกับข้อมูลเชิงพื้นที่บนระบบสารสนเทศภูมิศาสตร์ของโมดูลแผนที่ (Map API Module) ซึ่งสามารถแสดงผลในโมดูลแสดงผล (Presentation Module) ซึ่งสามารถเข้าใช้ได้ทั้งในรูปแบบของเว็บแอพพลิเคชันและแอนดรอยด์แอพพลิเคชัน ผลการพัฒนาระบบและขั้นตอนการทำงานของระบบ ดังภาพที่ 4



ภาพที่ 4 ผลการพัฒนาระบบและขั้นตอนการทำงานของระบบ

จากภาพที่ 4 ผลการพัฒนาระบบ พบว่า 1) เว็บแอพพลิเคชันสามารถนำเข้าข้อมูลและจัดการข้อมูลภายในฐานข้อมูล MySQL 2) โมบายแอพพลิเคชันหรือแอนดรอยด์แอพพลิเคชันสามารถเข้าถึงข้อมูลภายในฐานข้อมูล MySQL ด้วยอุปกรณ์พกพา เช่น สมาร์ทโฟน หรือ แท็บเล็ต ได้ตลอดเวลาผ่านเครือข่ายอินเทอร์เน็ต

2) ผลการประเมินประสิทธิภาพโดยผู้เชี่ยวชาญ

การประเมินประสิทธิภาพโดยผู้เชี่ยวชาญจำนวน 3 คน แสดงผลดังตารางที่ 1

ตารางที่ 1 แสดงการวิเคราะห์หาประสิทธิภาพของระบบ

ข้อคำถามของแบบประเมินประสิทธิภาพ	$\bar{X}$	S.D.
1. ประสิทธิภาพในการทำงานของระบบ	4.44	0.38
2.ประสิทธิภาพของการออกแบบหน้าจอ	4.50	0.58
3. ประสิทธิภาพระบบด้านการทดสอบใช้งาน	4.78	0.38
4. ประสิทธิภาพระบบด้านความปลอดภัย	4.50	0.58
ค่าเฉลี่ย	4.56	0.48

จากตารางที่ 1 พบว่า ประสิทธิภาพโดยรวมของระบบมีค่าเฉลี่ยเท่ากับ 4.56 ส่วนเบี่ยงเบนมาตรฐาน 0.48 อยู่ในระดับดีมาก โดยมีประสิทธิภาพของการออกแบบหน้าจอ ประสิทธิภาพระบบด้านการทดสอบใช้งาน ประสิทธิภาพระบบด้านความปลอดภัย อยู่ในระดับดีมาก และประสิทธิภาพในการทำงานของระบบ อยู่ในระดับดี

3) ผลการประเมินความเหมาะสมโดยผู้เชี่ยวชาญ

การประเมินความเหมาะสมของรายละเอียดภูมิปัญญาท้องถิ่นโดยผู้เชี่ยวชาญจำนวน 3 คน แสดงผลดังตารางที่ 2

ตารางที่ 2 แสดงผลการประเมินความเหมาะสมของระบบ

ข้อคำถามของการประเมินความเหมาะสม	ผลการประเมิน
1. รูปแบบระบบบริหารจัดการข้อมูลภูมิปัญญาท้องถิ่นบนเส้นทางท่องเที่ยว	มีความเหมาะสม
2. รูปแบบของเนื้อหาภูมิปัญญาท้องถิ่นและชนิดของสื่อที่นำเสนอ	มีความเหมาะสม
3. รูปแบบของเนื้อหาสถานที่ท่องเที่ยวและชนิดของสื่อที่นำเสนอ	มีความเหมาะสม
4. รูปแบบของซอฟต์แวร์/แอพพลิเคชันมีความเหมาะสมในการใช้งาน	มีความเหมาะสม
5. เทคโนโลยี/ซอฟต์แวร์/เครื่องมือในการพัฒนามีความเหมาะสม	มีความเหมาะสม

จากตารางที่ 2 พบว่า ระบบจัดการข้อมูลภูมิปัญญาท้องถิ่นที่บูรณาการร่วมกับเส้นทางท่องเที่ยวพื้นที่จังหวัดนครปฐม ในรูปแบบข้อมูลเชิงพื้นที่บนระบบสารสนเทศภูมิศาสตร์มีความเหมาะสมทุกด้าน

4) ผลการประเมินความพึงพอใจจากกลุ่มผู้ใช้

การประเมินความพึงพอใจจากกลุ่มผู้ใช้ ดังตารางที่ 3

ตารางที่ 3 แสดงผลการประเมินความพึงพอใจของผู้ใช้

ข้อคำถามของแบบประเมินความพึงพอใจ	$\bar{X}$	S.D.
1. ความสามารถในการทำงานของระบบ	4.60	0.51
1.1 การนำเสนอข้อมูลและเนื้อหา	4.68	0.47
1.2 การจัดการข้อมูลของระบบ	4.40	0.61
1.3 การทำงานกับอุปกรณ์ที่หลากหลาย	4.71	0.46
2. การออกแบบหน้าจอ	4.66	0.50
2.1 ความเหมาะสมของขนาดตัวอักษร สีที่นำเสนอ	4.58	0.50
2.2 ความเหมาะสมของการใช้ภาษาและการสื่อสาร	4.75	0.44
2.3 ความสวยงาม และการออกแบบหน้าจอ	4.59	0.63
2.4 ความง่ายในการใช้งานระบบ	4.71	0.46
3. การทดสอบการใช้งาน	4.46	0.50
3.1 ความเร็วในการประมวลผลของระบบ	4.57	0.50
3.2 ความถูกต้องในการประมวลผลของระบบ	4.41	0.50
3.3 ความปลอดภัยและการให้สิทธิ์เข้าใช้งานระบบ	4.40	0.49
ค่าเฉลี่ย	4.57	0.50

จากตารางที่ 3 พบว่า ผู้ใช้มีความพึงพอใจในการใช้งานระบบจัดการข้อมูลภูมิปัญญาท้องถิ่นที่บูรณาการร่วมกับเส้นทางท่องเที่ยวพื้นที่จังหวัดนครปฐม ในรูปแบบข้อมูลเชิงพื้นที่บนระบบสารสนเทศภูมิศาสตร์ทุกด้านอยู่ในระดับดีมาก มีค่าเฉลี่ยเท่ากับ 4.57 มีส่วนเบี่ยงเบนมาตรฐาน 0.50 โดยพบว่าความสามารถในด้านการทำงานของระบบ ($\bar{X}$ = 4.60, S.D. = 0.51) และด้านการออกแบบหน้าจอ ($\bar{X}$ = 4.66, S.D. = 0.50) ผู้ใช้มีความพึงพอใจอยู่ในระดับดีมาก และด้านการทดสอบการใช้งาน ($\bar{X}$ = 4.46, S.D. = 0.50) ผู้ใช้มีความพึงพอใจในระดับดี

5. สรุป

ผลการประเมินประสิทธิภาพของพัฒนาระบบจัดการข้อมูลภูมิปัญญาท้องถิ่นที่บูรณาการร่วมกับเส้นทางท่องเที่ยวพื้นที่จังหวัดนครปฐม ในรูปแบบข้อมูลเชิงพื้นที่บนระบบสารสนเทศภูมิศาสตร์ โดยผู้เชี่ยวชาญจำนวน 3 คน อยู่ในระดับ “ดีมาก” ($\bar{X}$ =4.56, S.D. = 0.48) และมีความเหมาะสมในทุกด้าน เมื่อนำไปทดลองใช้กับกลุ่มตัวอย่างโดยการสุ่มจากผู้ใช้งานอุปกรณ์พกพาระบบปฏิบัติการแอนดรอยด์และประเมินความพึงพอใจจากผู้ใช้ พบว่า ผู้ใช้มีความพึงพอใจทุกด้านอยู่ในระดับดีมาก ($\bar{X}$ = 4.57, S.D. = 0.50) โดยมีความพึงพอใจความสามารถในด้านการทำงานของระบบ ($\bar{X}$ = 4.60, S.D. = 0.51) และด้านการออกแบบหน้าจอ ($\bar{X}$ = 4.66, S.D. = 0.50) ผู้ใช้มีความพึงพอใจอยู่ในระดับดีมาก และด้านการทดสอบการใช้งาน ($\bar{X}$ = 4.46, S.D. = 0.50) ผู้ใช้มีความพึงพอใจอยู่ในระดับดี

6. ข้อเสนอแนะ

ข้อเสนอแนะทางการวิจัย ควรมีการต่อยอดทางการวิจัยด้วยการให้ชุมชนมีส่วนร่วมในการพัฒนารูปแบบการท่องเที่ยวเชิงอนุรักษ์ด้านภูมิปัญญาท้องถิ่นอย่างเต็มที่

ข้อเสนอแนะทางการนำไปใช้ ผลักดันให้หน่วยงานที่เกี่ยวข้องนำระบบไปใช้อย่างจริงจังเพื่อประโยชน์ส่วนรวมของคนในชุมชน

เอกสารอ้างอิง

[1] GIS. Available http://www.siamgis.com/features (วันที่ค้นคว้าข้อมูล 25 ธันวาคม 2559)

[2] Google Maps API. http://www.mcs.most.go.th/map_api/index.php (วันที่ค้นคว้าข้อมูล 25 ธันวาคม 2559)

[3] วราภรณ์ ดวงแสง, อนุวัติ เชื้อเย็น, วลัยลดา ถาวรมงคลกิจ, วัชรีวรรณ ศศิผลิน และณฤพล อัฐวงศ์. การต่อยอดภูมิปัญญาท้องถิ่นเพื่อพัฒนาแหล่งท่องเที่ยวอย่างยั่งยืนภายใต้ปรัชญาเศรษฐกิจพอเพียง: กรณีบ้านหม้อ ตำบลป่าไผ่ อำเภอสันทราย จังหวัดเชียงใหม่. เชียงใหม่, มหาวิทยาลัยแม่โจ้ สำนักวิจัยและส่งเสริมวิชาการการเกษตร. 2558.

[4] วีรภรณ์ โตคีรี. การพัฒนาการจัดการท่องเที่ยวโดยชุมชน: กรณีศึกษา หมู่บ้านจุฬาภรณ์พัฒนา 10 อำเภอเบตง จังหวัดยะลา. เชียงใหม่. มหาวิทยาลัยแม่โจ้. สำนักหอสมุด. 2558.

[5] ธีระ ลาภิศชยางกูล. การรวบรวมข้อมูลสำหรับสร้างฐานข้อมูลในงานระบบสารสนเทศภูมิศาสตร์. กรุงเทพฯ. มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี. คณะวิศวกรรมศาสตร์. 2559.

[6] เนื่องวงศ์ ทวยเจริญ. การพัฒนาโปรแกรมประยุกต์บนโทรศัพท์เคลื่อนที่ระบบปฏิบัติการแอนดรอยด์เพื่อค้นหาเส้นทางรถประจำทางในกรุงเทพมหานครที่มีระยะทางสั้นที่สุด. กรุงเทพฯ. มหาวิทยาลัยธุรกิจบัณฑิตย์. วิทยาลัยนวัตกรรมด้านเทคโนโลยีและวิศวกรรมศาสตร์. 2559.

[7] Simalaotao P, Sirisukpoca U. Development of Disease Dispersion Model on Geographic Information System via Dynamic Web Service, Information Technology Journal. Vol. 10, No. 2, July - December 2014, PP. 1-6, 2016.

แอปพลิเคชันจัดการหนังสือเสียงสำหรับผู้พิการทางสายตาบนอุปกรณ์เคลื่อนที่

Audio book management application for the blind on mobile device

กิ่งกาญจนา โคตรแก้ว (Kingkanjana Khotkaew)[1] ดลยา สุริยวงค์ (Donlaya Suriyawong)[2]
เพชร อิ่มทองคำ (Phet Imtongkhum) [3] และจักรชัย โสอินทร์ (Chakchai So-In)[4]
ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น
[1]yuan_yuan12@hotmail.com, [2]Su.donlaya@hotmail.com, [3]phet@kkumail.com, [4]chakso@kku.ac.th

บทคัดย่อ

โครงงานนี้มีวัตถุประสงค์ในการจัดทำแอปพลิเคชันจัดการหนังสือเสียงสำหรับผู้พิการทางสายตาบนอุปกรณ์เคลื่อนที่ โดยอ้างอิงกับความต้องการของโรงเรียนการศึกษาคนตาบอดจังหวัดขอนแก่น สืบเนื่องจากสื่อสิ่งพิมพ์ต่าง ๆ และหนังสือส่วนใหญ่เป็นแบบพิมพ์ซึ่งทำให้ผู้พิการทางสายตาไม่สามารถอ่านหนังสือเองได้ นอกจากนี้สื่อที่เป็นอักษรเบรลล์ยังมีไม่เพียงพอจึงทำให้โอกาสในการรับข้อมูลข่าวสารและการอ่านหนังสือนั้นน้อยลง ดังนั้นผู้จัดทำจึงพัฒนาแอปพลิเคชันเพื่อช่วยในการสร้างหนังสือเสียงและรวบรวมหนังสือเสียงไว้ให้ผู้พิการทางสายตาเพื่อสามารถเข้าถึงหนังสือได้ง่าย โดยใช้เครื่องมือ Android Studio, Speech Recognition และภาษา Java ในส่วนรู้จำเสียงพูดเพื่อให้ผู้ใช้งานใช้คำสั่งเสียงและการจัดเก็บข้อมูล เช่น ข้อมูลหนังสือเสียงหมวดต่าง ๆ โดยผลการทดสอบปรากฏผลผ่านในทุกฟังก์ชัน และจากผู้ใช้งานที่ โรงเรียนการศึกษาคนตาบอดปรากฏความพึงพอใจอยู่ในระดับดี

คำสำคัญ: หนังสือเสียง, หนังสือเสียงระบบเดซี, การรู้จำเสียงพูด.

Abstract

This project aims to develop an audio book management application for visually impaired people on mobile devices based on the needs of the Khon Kaen School for the blind. Since most print media and books are in print format, the blinds can not read books themselves. In addition, the braille media is not enough, and so the chances of getting information and reading are limited. Thus, the authors aim to develop the application help to create audiobooks and to collect them for the disabled to easily access. Android Studio, Speech Recognition, and Java language tools are used to allow users to use voice commands and data storage, such as audiobooks classified in each category. The functionality evaluation yields pass for all functions and from users at the Khon Kaen School for the blind, the satisfaction score is good.

Keywords: audiobooks, DAISY audiobooks, speech recognition

1. บทนำ

จากการสำรวจพบว่าผู้พิการในประเทศไทยมีจำนวน 2,137,105 คน มีผู้พิการทางสายตาประมาณ 178,049 คน ของผู้พิการทั้งหมด และปัญหาของผู้พิการทางสายตาส่วนใหญ่พบว่า 90% ของผู้พิการทางสายตายังเข้าไม่ถึงสื่อสิ่งพิมพ์แม้ว่าจะมีเทคโนโลยีแปลงเป็นอักษรเบรลล์ ก็ยังพบว่ามีปัญหาด้านลิขสิทธิ์ สำหรับประเทศไทยโอกาสของผู้พิการทางสายตาที่เข้าถึงการศึกษามีเพียง 15% เท่านั้น [1] อีกทั้งสื่อการสอนยังไม่เพียงพอต่อความต้องการ และยังขาดเงินทุนสนับสนุนจึงทำให้ผู้พิการทางสายตาขาดโอกาสในการเรียนรู้และเข้าทำงาน ซึ่งบริษัทส่วนใหญ่ก็ยังปฏิเสธผู้พิการทางสายตา เนื่องจากอาจจะมีสมมุติฐานในเรื่องของสมรรถนะการทำงานที่ไม่เท่าเทียมกัน

ปัจจุบันเทคโนโลยีได้เข้ามามีบทบาทในชีวิตประจำวัน และผู้พิการทางสายตาส่วนใหญ่เริ่มมีการใช้สมาร์ตโฟนหรือแท็บเล็ตมากยิ่งขึ้น [2] ดังนั้นผู้จัดทำจึงเล็งเห็นช่องทางในการผนวกเทคโนโลยีเข้ากับความเป็นไปได้ในการเข้าถึงสื่อของผู้พิการทางสายตาโดยพัฒนาแอปพลิเคชันจัดการหนังสือเสียงสำหรับผู้พิการทางสายตาโดยใช้เทคโนโลยีสั่งงานด้วยเสียงและเป็น

แอปพลิเคชันที่รวบรวมหนังสือเสียงเพื่อให้ผู้พิการทางสายตาได้เข้าถึงหนังสือและแหล่งความรู้ได้สะดวกยิ่งขึ้น อีกทั้งยังสามารถแบ่งปันความรู้ให้กับผู้พิการทางสายตาคนอื่น ๆ ได้ นอกจากนั้นยังมีในส่วนของอาสาสมัครที่สามารถสร้างหนังสือเสียงให้ผู้พิการทางสายตาได้อีกด้วย

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 ทฤษฎีที่เกี่ยวข้อง

2.1.1 ความบกพร่องทางการมองเห็น [3]

บุคคลที่สูญเสียการเห็นตั้งแต่ระดับเล็กน้อยจนถึงตาบอดสนิท แบ่งออกเป็น 2 ประเภท คือ 1) ผู้พิการทางสายตาสนิท หรือคนที่สูญเสียการเห็นจนไม่สามารถอ่านหนังสือจึงต้องใช้อักษรเบรลล์หรือใช้วิธีการฟังเทปหรือแผ่นเสียง และ 2) ผู้พิการทางสายตาประเภทเลือนราง หรือคนที่สูญเสียการเห็นแต่ยังสามารถอ่านอักษรตัวพิมพ์ที่ขยายใหญ่ได้ สำหรับตัวอย่างของสาเหตุของความบกพร่องทางการเห็นมีดังนี้คือ

- ความผิดปกติของดวงตา คือ เกิดจากความเสื่อมสภาพของกล้ามเนื้อตาเป็นเหตุให้สายตาสั้น สายตายาวหรือเกิดมีปัญหาจากการปรับภาพที่เลนส์ในดวงตา เป็นต้น
- ความผิดปกติของสายตา คือ เกิดจากโรคภัยไข้เจ็บจากอุบัติเหตุต่าง ๆ ที่เป็นอันตรายต่อดวงตา
- จากฤทธิ์ยาบางประเภทตลอดจนใช้ยาผิด โรคบางอย่างที่ไม่สามารถป้องกันได้ เช่น เนื้องอกที่ตา โรคเหล่านี้อาจทำให้ตาบอดหรือมีความบกพร่องทางสายตาอย่างรุนแรง

2.1.2 หนังสือเสียงระบบเดซี [4]

DAISY (Digital Accessible Information System) หรือหนังสือเสียงระบบเดซีเป็นอีกหนึ่งช่องทางของการผลิตหนังสือเสียง เนื่องจากเทคโนโลยีการผลิตหนังสือเสียงระบบเดิมยังเป็นระบบ Analog และจัดเก็บไว้บนม้วนเทปคาสเซต

หนังสือ Daisy ทำให้การสืบค้นและเข้าถึงข้อมูลได้โดยง่าย ดังนั้นจึงมีประสิทธิภาพในการเข้าถึงข้อมูลอย่างรวดเร็วทจึงมีประโยชน์ต่อผู้ใช้เป็นอย่างมากและอุปกรณ์ที่รองรับการอ่านหนังสือ Daisy มีมากมายทั้งโปรแกรมคอมพิวเตอร์ อุปกรณ์แบบตั้งโต๊ะ และแบบพกพา หนังสือDaisy เป็นมาตรฐานหนังสือเสียง เนื่องจากใช้ได้อย่างสะดวก และทำให้ผู้พิการสายตาเข้าถึงหนังสือเสียงได้ง่ายขึ้น

2.1.3 Speech Recognition [5]

การรู้จำเสียง โดยรับรู้เสียงของมนุษย์อัตโนมัติ โดยการแปลงข้อมูลนำเข้าในรูปแบบเสียงพูดเป็นข้อความตัวอักษร

2.1.4 Text To Speech และ Speech To Text [6]

Text To Speech เป็นการส่งข้อความให้ระบบแปลงข้อความเป็นเสียง ซึ่งถือว่าเป็นพื้นฐานบนอุปกรณ์แอนดรอยด์ และในส่วนของ Speech To Text มีคล้ายคลึงกับ Text To Speech คือการพูดผ่าน Microphone แล้วจึงแปลงเป็นข้อความอักษร

2.1.5 แอนดรอยด์ (Android) [7]

เป็นระบบปฏิบัติการสำหรับอุปกรณ์พกพา เช่น โทรศัพท์มือถือ แท็บเล็ต คอมพิวเตอร์ ซึ่งเป็นระบบที่มีพื้นฐานการทำงานบนลินุกซ์ เคอร์เนล ปัจจุบันเป็นของบริษัทกูเกิลและได้พัฒนาในนามของ Open Handset Alliance ทางกูเกิลได้เปิดให้นักพัฒนาสามารถแก้ไขโค้ดต่าง ๆ ด้วยภาษา Java และควบคุมอุปกรณ์ผ่านทางชุด Java libraries ที่กูเกิลพัฒนาขึ้น แอนดรอยด์จึงเป็นซอฟต์แวร์ Open Source ที่ได้รับความนิยมเป็นอย่างมาก และยังมีฟังก์ชันที่หลากหลาย

2.2 งานวิจัยที่เกี่ยวข้อง

2.2.1 Be My Eyes เป็นแอปพลิเคชันสำหรับช่วยเหลือผู้พิการทางสายตา โดยให้ผู้พิการทางสายตาใช้กล้องบนโทรศัพท์มือถือถ่ายเหตุการณ์ที่กำลังเกิดขึ้น แล้วส่งไปยังอาสาสมัครซึ่งเป็นคนสายตาปกติช่วยตรวจสอบในสิ่งที่ผู้พิการทางสายตาต้องการความช่วยเหลือ เช่น ตรวจสอบวันหมดอายุบนฉลากอาหาร โดยสามารถเห็นภาพแบบวีดีโอและเป็นแบบถ่ายทอดสดแบบเวลาจริงแต่ใช้ได้เฉพาะในพื้นที่ที่มีความเร็วของอินเตอร์เน็ตสูง [8]

2.2.2 Audible เป็นแอปพลิเคชันที่รวบรวมหนังสือเสียงโดยสร้างขึ้นเพื่อผู้ใช้งานทั่วไปที่ต้องการความสะดวกสบายในการอ่านหนังสือซึ่งเป็นการฟังแทนการอ่าน โดยสามารถส่งหนังสือไปยังอุปกรณ์พกพาผ่าน WiFi ได้ แต่มีค่าใช้จ่าย [9]

2.2.3 Read for the Blind เป็นแอปพลิเคชันช่วยเหลือผู้พิการทางสายตาด้วยการร่วมกันสร้างหนังสือเสียงให้ผู้พิการทางสายตาฟัง โดยหนังสือเสียงที่สมบูรณ์จะถูกแปลงเป็นไฟล์เสียงและส่งไปยังผู้พิการทางสายตา ซึ่งแอปพลิเคชันนี้ไม่เสียค่าใช้จ่ายในการฟังหนังสือเสียง แต่ฟังหนังสือเสียงได้เฉพาะผู้พิการทางสายตาและไม่สามารถแบ่งปันหนังสือเสียงได้ [10]

2.2.4 เว็บจุฬาเดซี เป็นโปรแกรมสร้างหนังสือเสียงดิจิทัลในมาตรฐานเดซี เพื่อให้อาสาสมัครสามารถสร้างหนังสือเสียงด้วยตนเอง โดยจัดทำขึ้นเพื่อเปิดโอกาสให้กับอาสาสมัครที่มีความประสงค์จะสร้างหนังสือเสียงมาตรฐานเดซี ซึ่งมีข้อดีคือสามารถสร้างหนังสือเสียงเอง อีกทั้งยังสามารถสังเคราะห์เสียงพูดจากข้อความที่นำเข้ามาเป็นภาษาไทยได้ แต่ยังไม่สามารถใช้งานได้ในอุปกรณ์เคลื่อนที่อย่างสมบูรณ์ [11]

2.2.5 Bus Alert เป็นแอปพลิเคชันสำหรับผู้พิการทางสายตา โดย Bus Alert จัดการเรื่องการเดินทาง เช่น แจ้งพิกัดบริเวณจุดรับและส่งผู้โดยสาร ซึ่งมีข้อดีคือ สามารถแจ้งเตือนด้วยเสียงเมื่อถึงปลายทางแต่ใช้งานได้ภายในประเทศบราซิลเท่านั้น [12]

จากที่ได้สอบถามจากโรงเรียนการศึกษาคนตาบอดและเว็บไซต์พบว่างานวิจัยต่าง ๆ มีข้อดีและข้อเสียแตกต่างกันออกไป แต่โดยสรุปยังไม่มีแอปพลิเคชันที่สามารถสั่งงานผ่านเสียงได้อย่างสมบูรณ์ ดังนั้นทางผู้จัดทำจึงนำส่วนนี้มาพัฒนาเพื่อจัดการหนังสือเสียงผ่านคำสั่งเสียงบนอุปกรณ์เคลื่อนที่



ภาพที่ 1 ภาพรวมของแอปพลิเคชัน

3. วิธีดำเนินการวิจัย

3.1 วิเคราะห์ปัญหาและความต้องการ

วิเคราะห์และระบุถึงความต้องการจากระบบการทำงานเดิม

3.2 ค้นคว้า ศึกษาทฤษฎี และงานวิจัยที่เกี่ยวข้อง

ศึกษาจากเว็บไซต์และค้นหาแอปพลิเคชันที่เกี่ยวข้องกับหนังสือเสียงและผู้พิการทางสายตาที่มีลักษณะคล้ายคลึงกัน เช่น Be My Eyes, Audible, Read for the Blind, เว็บไซต์จุฬาเดซี และ Bus Alert เพื่อระบุถึงความสามารถของแอปพลิเคชัน

3.3 การออกแบบระบบ และวิเคราะห์ระบบ

การออกแบบระบบ จะแบ่งเป็น 3 หมวดผู้ใช้งานหลัก คือ ผู้ใช้ผู้พิการทางสายตา, อาสาสมัคร และผู้ดูแลระบบ

3.3.1 Use Case Diagram (แผนภาพแบบจำลอง)

Use Case Diagram แบ่งเป็น 3 ส่วนผู้ใช้งานคือ ผู้ใช้ผู้พิการทางสายตา, ผู้ใช้งานทั่วไป, ผู้ดูแลระบบ

ผู้ใช้งานผู้พิการทางสายตาสามารถใช้คำสั่งเสียงในการสั่งงานต่าง ๆ เช่น ดูหนังสือแนะนำ, ดูหนังสือเสียงใหม่, ดูหมวดหมู่, ฟังหนังสือเสียง, ค้นหาหนังสือเสียง, เล่นหนังสือเสียง, ร้องขอหนังสือเสียง, หนังสือโปรด, แชร์หนังสือ, ดาวน์โหลด และดูประวัติการใช้งาน

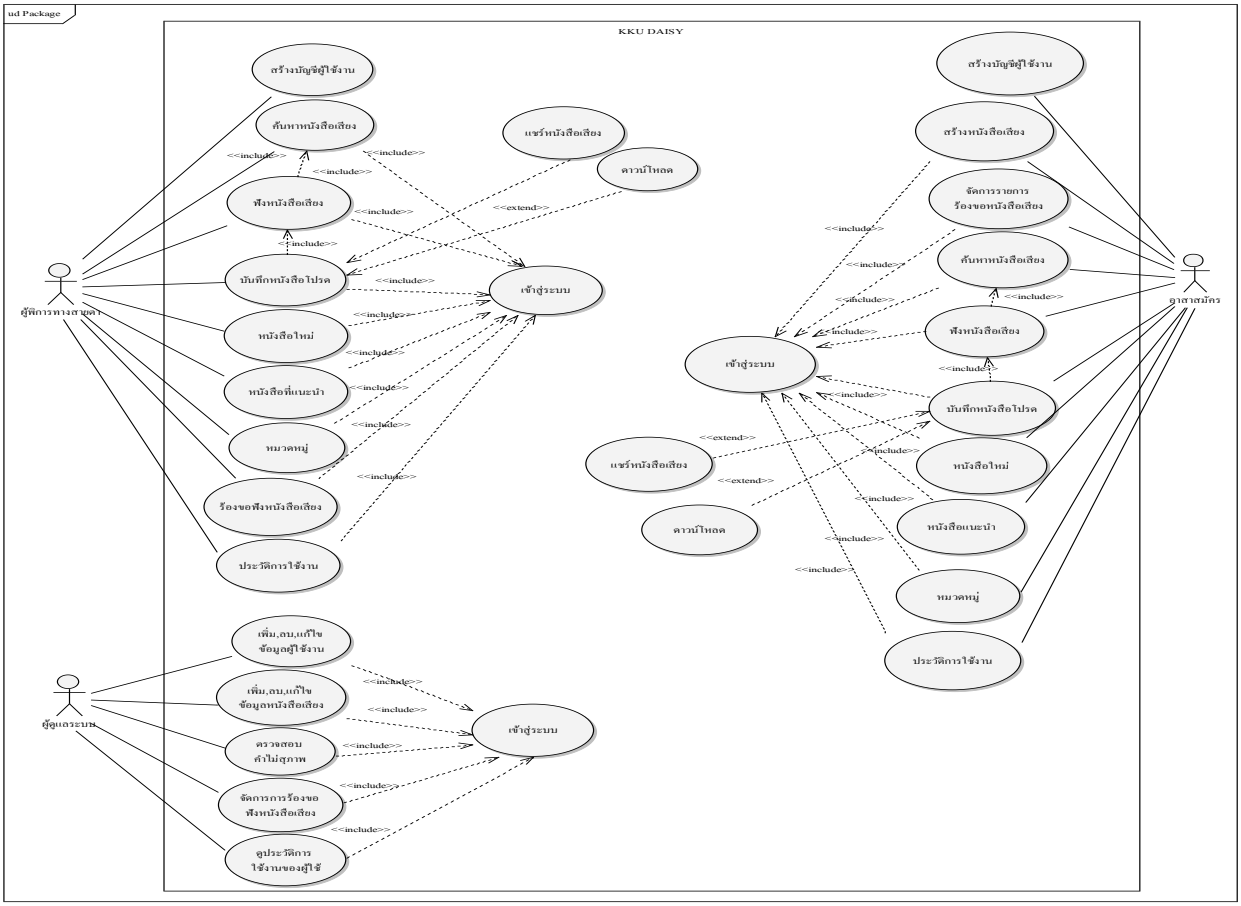
ส่วนของอาสาสมัครก็สามารถใช้งานแอปพลิเคชันเหมือนในส่วนของผู้พิการทางสายตาแต่จะมีข้อแตกต่างกันคือผู้ใช้ทั่วไปสามารถสร้างหนังสือเสียงได้ และดูคำร้องขอจากผู้พิการทางสายตา และส่วนของ Server ก็จะใช้จัดการบัญชีผู้ใช้งานและรวบรวมหนังสือเพิ่มลบข้อมูล ของหนังสือเสียงและตรวจคำไม่สุภาพของหนังสือเสียง

3.3.2 เทคนิคที่ใช้ในการพัฒนาแอปพลิเคชัน

ใช้เทคโนโลยี Text to Speech มาใช้ในส่วนของการใช้คำสั่งเสียงทำให้แอปพลิเคชันเอื้อต่อผู้พิการทางสายตา ทำให้ผู้พิการทางสายตาเข้าถึงหนังสือเสียงได้ง่ายมากขึ้นโดยการใช้งานผ่านคำสั่งเสียง ตัวอย่างคำสั่งเสียงในแอปพลิเคชัน เช่น การค้นหาหนังสือเสียง สั่งเล่น และหยุด เป็นต้น ในส่วนของการสร้างหนังสือเสียงอาสาสมัครสามารถบันทึกเสียงบนอุปกรณ์เคลื่อนที่ และจัดเก็บเป็นไฟล์เสียง

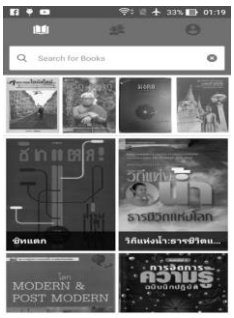
4. ผลการดำเนินงาน

การทดสอบระบบโดยให้ผู้พิการทางสายตาทดลองใช้แอปพลิเคชันและทำแบบสอบถามว่าเป็นไปตามวัตถุประสงค์หรือขอบเขตที่ได้วางไว้ รวมไปถึงการตรวจสอบข้อผิดพลาดของระบบและลงมือแก้ไขข้อผิดพลาด

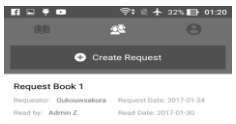


ภาพที่ 2 Use Case Diagram ของระบบ

4.1 ตัวอย่างการพัฒนาแอปพลิเคชัน



(ก)



(ข)

ภาพที่ 5 (ก) เมื่อทำการ Log in ก็จะแสดงหน้าจอหลักโดยแสดงรายชื่อหนังสือ ที่สามารถค้นหาหนังสือเสียงและใช้คำสั่งเสียงในการค้นหา **(ข)** เมื่อเลื่อนหน้าจอไปทางซ้ายจะปรากฏหน้าจอร้องขอหนังสือเสียงในการใช้คำสั่งเสียงเพื่อร้องขอฟังหนังสือเสียง



(ก)

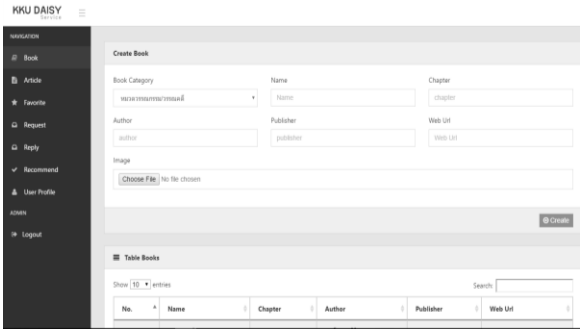


(ข)

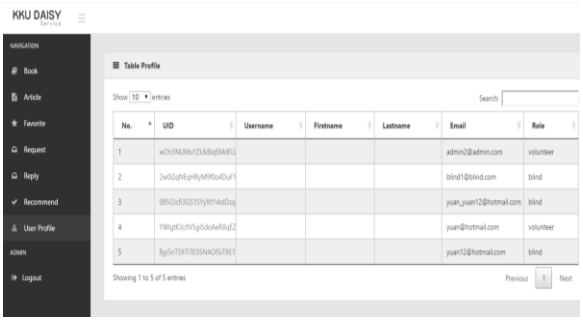
ภาพที่ 6 (ก) เมื่อเลื่อนหน้าจอไปทางซ้ายครั้งที่ 3 จะแสดงหน้าประวัติและหน้าเมนูต่าง ๆ **(ข)** หน้าเล่นหนังสือเสียง โดยสามารถเลือกแบ่งปัน, หนังสือโปรด และดาวน์โหลด



ภาพที่ 9 เมื่อกดเลือกหนังสือเสียงจะแสดงรายชื่อแต่ละบทของหนังสือเสียง



ภาพที่ 10 ส่วนของผู้ดูแลระบบแสดงหน้าจัดการหน้าหนังสือเสียง



ภาพที่ 11 ส่วนของผู้ดูแลระบบแสดงหน้าจัดการบัญชีผู้ใช้งาน

4.2 การทดสอบการใช้งานแอปพลิเคชัน

4.2.1 ผลการประเมินประสิทธิภาพของระบบ โดยมีทั้งหมด 14 ฟังก์ชันตาม Use Case Diagram โดยทดสอบจำนวน 10 ครั้ง (ผ่าน/ไม่ผ่าน) ได้ผลดังต่อไปนี้

- เข้าสู่ระบบแสดงหน้าเข้าสู่ระบบโดยให้กรอก User และ Password สรุปว่าผ่าน ทั้ง 10 ครั้ง
- ใช้คำสั่งเสียง เลื่อนหน้าจอเพื่อให้ระบบรับคำสั่งเสียง สรุปว่าผ่าน ทั้ง 10 ครั้ง
- เลือกเมนูหนังสือใหม่ แสดงหน้ารายชื่อหนังสือเสียง สรุปว่าผ่าน ทั้ง 10 ครั้ง
- เลือกเมนูแนะนำหนังสือ แสดงหน้ารายชื่อหนังสือเสียง สรุปว่าผ่าน ทั้ง 10 ครั้ง
- เลือกเมนูหมวดหนังสือ แสดงหน้าหมวดหนังสือและรายชื่อหนังสือเสียง สรุปว่าผ่าน ทั้ง 10 ครั้ง
- เลือกเมนูเล่นหนังสือเสียง เล่นหนังสือเสียง สรุปว่าผ่าน ทั้ง 10 ครั้ง
- เลือกเมนูค้นหา แสดงหน้ารายชื่อหนังสือเสียง สรุปว่าผ่าน ทั้ง 10 ครั้ง
- เลือกเมนูหนังสือโปรด บันทึกหนังสือที่โปรด สรุปว่าผ่าน ทั้ง 10 ครั้ง
- ประวัติการใช้งาน แสดงหน้าประวัติการใช้งาน สรุปว่าผ่าน ทั้ง 10 ครั้ง
- เลือกเมนูสร้างหนังสือเสียง แสดงหน้าสร้างหนังสือเสียงโดยอัดเสียงและบันทึกได้ สรุปว่าผ่าน ทั้ง 10 ครั้ง
- เลือกเมนูร้องขอฟังหนังสือเสียง สามารถส่งคำร้องขอฟังหนังสือเสียงได้ สรุปว่าผ่าน ทั้ง 10 ครั้ง
- เลือกแบ่งปัน แบ่งปันหนังสือไปยังเฟสบุ๊ค สรุปว่าผ่าน ทั้ง 10 ครั้ง
- จัดการการร้องขอหนังสือเสียง สามารถจัดการการร้องขอฟังหนังสือเสียงของผู้พิการทางสายตาได้ สรุปว่าผ่าน ทั้ง 10 ครั้ง
- ผู้ดูแลระบบแสดงหน้าจัดการลบ, เพิ่ม, แก้ไขข้อมูลผู้ใช้และหนังสือ สรุปว่าผ่าน ทั้ง 10 ครั้ง

4.2.2 ทดสอบประสิทธิภาพของการใช้คำสั่งเสียง โดยผลจากการทดสอบของคำเสียงที่จำนวน 10 ครั้งได้ผลดังนี้

ตารางที่ 1 การทดสอบการใช้คำสั่งเสียงหลัก

Scenario	Expected Results	Pass/Fail
คำสั่งค้นหา	แสดงหน้ารายชื่อหนังสือเสียง	Pass
คำสั่งเลือกเมนูหนังสือ	แสดงรายชื่อเมนูหนังสือเสียง	Pass
คำสั่งส่วนของการเล่นหนังสือเสียง	เล่นหนังสือเสียง	Pass
คำสั่งบันทึกเป็นหนังสือโปรด	บันทึกหนังสือที่โปรด	Pass
คำสั่งดาวน์โหลดหนังสือ	ดาวน์โหลดหนังสือ	Pass

4.3 การทดสอบการประเมินความพึงพอใจ

ผลการประเมินของผู้พิการทางสายตาที่โรงเรียนศึกษาคนตาบอด จังหวัดขอนแก่น จำนวน 10 คนและผู้ใช้งานทั่วไปจำนวน 10 คนดังตารางที่ 2 เกณฑ์คะแนนที่ใช้ประเมินระดับความพึงพอใจของแบบสอบถามคือค่าเฉลี่ย 4.50 - 5.00 อยู่ในระดับดีมาก 3.50 – 4.49 อยู่นระดับดี 2.50 – 3.49 อยู่ในระดับปานกลาง 1.50 – 2.49 อยู่ในระดับพอใจ และ 0.00 – 1.49 อยู่ในระดับควรปรับปรุง โดยสรุปผลการประเมินประสิทธิภาพของแอปพลิเคชันโดยผู้พิการทางสายตาซึ่งได้ค่าเฉลี่ยเท่ากับ 4.18 และ ผู้ใช้ทั่วไปได้ค่าเฉลี่ยเท่ากับ 4.26 ซึ่งแสดงถึงความพึงพอใจของผู้ใช้งานอยู่ในระดับดี

ตารางที่ 2 ผลสรุปความพึงพอใจของผู้ใช้งาน

รายการประเมิน	ค่าเฉลี่ยผู้พิการทางสายตา	ค่าเฉลี่ยผู้ใช้ทั่วไป
1. รูปแบบแอปพลิเคชัน		
1.1 รูปแบบหน้าจอมีความเหมาะสม	4.1	4.4
1.2 ตำแหน่งการจัดวางปุ่ม	4.3	4.1
1.3 ขนาดและรูปแบบของตัวอักษร	3.8	4.6
2. การใช้งานแอปพลิเคชัน		
2.1 สามารถทำงานและแสดงผลได้ถูกต้อง	3.6	4.6
2.2 ความเร็วของในการตอบสนอง	3.9	4.2
2.3 แอปพลิเคชันมีประโยชน์ต่อผู้ใช้งาน	4.5	4.1
2.4 การใช้งานผ่านคำสั่งเสียงเป็นที่พึงพอใจมากน้อยเพียงใด	3.7	4.4
3. ภาพรวมของแอปพลิเคชัน		
3.1 ความน่าสนใจของแอปพลิเคชัน	4.2	4.2
3.2 ความทันสมัยของแอปพลิเคชัน	4.8	4.1
3.3 ผู้ใช้งานสามารถใช้งานและสามารถเข้าใจการใช้งานได้อย่างง่าย	4.9	3.9
สรุป	**4.18**	**4.26**

5. สรุปผลและข้อเสนอแนะ

แอปพลิเคชันจัดการหนังสือเสียงสำหรับผู้พิการทางสายตาบนอุปกรณ์เคลื่อนที่โดยเป็น Mobile Application บนระบบปฏิบัติการแอนดรอยด์สามารถสั่งงานด้วยเสียงได้แบ่งการใช้งานออกเป็น 3 ส่วน คือ อาสาสมัคร ผู้ใช้ผู้พิการทางสายตา และผู้ดูแลระบบ ประกอบด้วย 14 ฟังก์ชันหลัก ซึ่งแอปพลิเคชันสามารถทำงานได้เฉพาะระบบปฏิบัติการแอนดรอยด์ 5.1 ขึ้นไป ใช้งานคำสั่งเสียงได้เฉพาะภาษาไทย โดยผู้ใช้งานต้องมีบัญชีผู้ใช้งานก่อน นอกจากนี้ต้องทำการเชื่อมต่ออินเตอร์เน็ตเพื่อทำให้แอปพลิเคชันทำงานได้อย่างสมบูรณ์ตามที่ผู้ใช้งานต้องการ โดยส่วนการตอบแบบสอบถามจากโรงเรียนการศึกษาคนตาบอดนั้นมีผลความพึงพอใจในการใช้แอปพลิเคชันอยู่ในระดับดี ทำให้การพัฒนาแอปพลิเคชันเป็นไปตามวัตถุประสงค์และขอบเขตที่ตั้งไว้ และในอนาคตมีแนวโน้มจะพัฒนาต่อยอดโดยการเพิ่มรูปแบบให้สามารถใช้งานได้ทั้งภาษาไทยและภาษาอังกฤษ

6. เอกสารอ้างอิง

[1] กรมส่งเสริมคนพิการ. (2559). สถิติข้อมูลคนพิการ.[ออนไลน์]. ค้นเมื่อ 17 ตุลาคม 2559, จาก https://www.m-society.go.th/ewt_news.php?nid=17397

[2] กรมส่งเสริมและพัฒนาคุณภาพชีวิตคนพิการ. (2557). เทคโนโลยีกับคนพิการ. [ออนไลน์]. ค้นเมื่อ 21 เมษายน 60, จาก http://dep.go.th/th/disabilities-technology

[3] ศูนย์การศึกษาพิเศษ เขตการศึกษา 5 จังหวัดสุพรรณบุรี. (2556).ความบกพร่องทางการมองเห็น. [ออนไลน์]. ค้นเมื่อ 23 สิงหาคม 2559, จาก http://www.specialcenter5.org/

[4] webmaster. (2557).หนังสือเสียงระบบเดซี. [ออนไลน์]. ค้นเมื่อ 23 สิงหาคม 2559, จาก http://www.bluerollingdot.org/articles/scoop/279

[5] ณัฐวุธ อวยชัยพรเลิศ. (2558). Speech Recognition. [ออนไลน์]. ค้นเมื่อ 17 ตุลาคม 2559, จาก https://biometricskmit.wordpress.com/

[6] Tlecoco. (2557). Text-To-Speech/Speech-To-Text. [ออนไลน์]. ค้นเมื่อ 17 พฤษภาคม 2559, จาก http://php662.blogspot.com/

[7] จุมพล สุวรรณจันทร์. (2558). Android Studio. [ออนไลน์]. ค้นเมื่อ 17 พฤษภาคม 2559, จาก http://biggaf.blogspot.com

[8] Be My Eyes. (2558).ค้นเมื่อ 28 สิงหาคม 2558, จาก http://bit.ly/19ezbAo

[9] Audible. (2558, 28 กันยายน). ค้นเมื่อ 28 สิงหาคม 2558, จาก https://itunes.apple.com/th

[10] Read For The Blind. (2557, 2 มิถุนายน). ค้นเมื่อ 28 สิงหาคม 2558, จาก https://itunes.apple.com/

[11] จุฬาเดซี. (2554). ค้นเมื่อ 28 สิงหาคม 2558, จาก http://161.200.80.21/cudaisy/about.aspx

[12] Bus Alert. (2558). ค้นเมื่อ 28 สิงหาคม 2558, จาก http://men.kapook.com/view115731.html

แอปพลิเคชันช่วยเหลือผู้ป่วยอัลไซเมอร์ระยะแรกบนอุปกรณ์เคลื่อนที่
Alzheimer aiding application in early stage on mobile device

ชนนิกา ปัญจันทร์สิงห์ (Chonnika Punjunsing)[1] ปิยะพร อภิสุวรรณ (Piyaporn Apisuwan)[2]
เพชร อิ่มทองคำ (Phet Imtongkhum) [3] และจักรชัย โสอินทร์ (Chakchai So-In)[4]
ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น
[1]Chonnika.p@gmail.com, [2]piyaporn9916@gmail.com, [3]phet@kkumail.com, [4]chakso@kku.ac.th

บทคัดย่อ

โรคอัลไซเมอร์เป็นอาการสมองเสื่อมชนิดหนึ่ง ซึ่งจะส่งผลกระทบโดยตรงต่อความทรงจำของผู้ป่วยในระยะแรกมักจะมีอาการหลงลืมบ่อย จำในสิ่งที่เพิ่งเกิดขึ้นไม่ได้ มีปัญหาในการใช้คำสนทนา ชอบเดินออกนอกบ้านบ่อย ๆ ทำให้ผู้ป่วยหลงทางจำทางกลับบ้านไม่ได้ อาจก่อให้เกิดอุบัติเหตุ และคลาดจากผู้ดูแล ปัญหาทั้งหมดทำให้มีผลกระทบต่อการใช้ชีวิตประจำวันของผู้ป่วย ในปัจจุบันพบว่าประชากรผู้สูงอายุที่ป่วยเป็นโรคอัลไซเมอร์มีจำนวนมากเพิ่มขึ้น ดังนั้นโครงงานนี้จึงพัฒนาแอปพลิเคชันส่งเสริมทักษะด้านความจำสำหรับผู้ป่วยอัลไซเมอร์บนโทรศัพท์เคลื่อนที่ร่วมกับนาฬิกาอัจฉริยะ เพื่อช่วยให้ผู้ป่วยโรคอัลไซเมอร์ระยะแรกใช้ชีวิตประจำวันได้อย่างปกติ โดยร่วมมือกับแพทย์ผู้เชี่ยวชาญจากโรงพยาบาลศรีนครินทร์ ภาควิชาจิตเวชศาสตร์ คณะแพทยศาสตร์ มหาวิทยาลัยขอนแก่น ซึ่งประยุกต์ใช้ JAVA, CSS6, Android Wear และ Android Studio ในการพัฒนา โดยผลการทดสอบปรากฏผลผ่านในทุกฟังก์ชัน และจากผู้ใช้งานที่โรงพยาบาลศรีนครินทร์ ปรากฏถึงความพึงพอใจเฉลี่ยในระดับดี

คำสำคัญ: อัลไซเมอร์, แอปพลิเคชันช่วยเหลือ, นาฬิกาอัจฉริยะ

Abstract

Alzheimer's disease is a type of dementia. This will directly affect the memories of patients in the early stages, often with frequent forgetfulness; they can not remember what just happened. There was a problem using the conversation; like to walk out often causing the patient to miss the way home. This could cause an accident; perhaps out of the caretaker. All these issues have an impact on the patient's daily lives. In addition, there is an increasing number of elderly patients with Alzheimer's disease. Thus, this project develops an application for promoting memory skills for Alzheimer's patients on mobile phones with intelligent watches so as to help patients with Alzheimer's disease in early life as normal as possible in collaboration with a specialist from the Department of Psychiatry, Faculty of Medicine, Khon Kaen University. This application uses JAVA, CSS6, Android Wear, and Android Studio. The functionality evaluation yields pass for all functions and from users at Srinagarind Hospital, the result appears to average satisfaction at a good level.

Keywords: Alzheimer, Helper Application, Smart Watch.

1. บทนำ

เนื่องจากในปัจจุบันจำนวนประชากรผู้สูงอายุ หรือประชากรที่มีอายุตั้งแต่ 60 ปีขึ้นไป มีจำนวนเพิ่มมากขึ้น เช่นจากผลการสำรวจปี พ.ศ. 2557 พบว่ามีจำนวนผู้สูงอายุคิดเป็นร้อยละ 14.9 ของประชากรทั้งหมด [1] และพบว่าส่วนมากป่วยเป็นโรคอัลไซเมอร์ ซึ่งโรคนี้เป็นอาการสมองเสื่อมชนิดหนึ่ง เกิดจากการตายของเซลล์สมองทำให้การทำงานของสมองเสื่อมลงและอาการที่เด่นชัดที่สุดคือสูญเสียความทรงจำ

โรคนี้ไม่สามารถรักษาให้หายขาดได้ แต่สามารถรักษาเพื่อให้อาการดีขึ้น เพื่อให้ผู้ป่วยสามารถช่วยเหลือตัวเองได้ ซึ่งจะขึ้นอยู่กับความเอาใจใส่ของผู้ดูแลผู้ป่วย ผู้ป่วยในระยะแรกมักจะจำสิ่งที่เพิ่งเกิดขึ้นไม่ได้ ชอบหลงทางกลับบ้าน มีปัญหาการใช้คำสนทนา ไม่ได้ หรืออาจก่อให้เกิดอุบัติเหตุ ในบางครั้ง

อาจคลาดจากสายตาผู้ดูแล ซึ่งปัญหาทั้งหมดเหล่านี้ทำให้มีผลกระทบต่อการใช้ชีวิตประจำวันของผู้ป่วยเป็นอย่างมาก

ปัจจุบันเทคโนโลยีโทรศัพท์เคลื่อนที่ เข้ามามีบทบาทสำคัญในการใช้ชีวิตประจำวันเป็นอย่างมาก [2] การนำเทคโนโลยีนี้มาใช้ให้เกิดประโยชน์จึงเป็นสิ่งสำคัญ นอกจากนี้ยังมีนาฬิกาอัจฉริยะ ที่สามารถใช้ในการติดต่อสื่อสารเชื่อมต่อกับโทรศัพท์ได้ โดยมีขนาดเล็ก พกพาสะดวก จึงเป็นที่นิยมเพิ่มมากขึ้น

ดังนั้นผู้จัดทำจึงได้พัฒนาแอปพลิเคชันบนโทรศัพท์ เคลื่อนที่และนาฬิกาอัจฉริยะในการช่วยเหลือผู้ป่วยอัลไซเมอร์ระยะแรกบนอุปกรณ์เคลื่อนที่ เพื่อเป็นอีกทางเลือกหนึ่งในการเพิ่มความสะดวกในการดูแล และการใช้ชีวิตประจำวันของผู้ป่วย โดยประยุกต์ใช้ JAVA, CSS6, Android Wear และ Android Studio

2. ทฤษฎีและผลงานที่เกี่ยวข้อง

2.1 ทฤษฎีที่เกี่ยวข้อง

2.1.1 โรคอัลไซเมอร์ [3]

เป็นอาการสมองเสื่อมชนิดหนึ่ง เกิดจากการตายของเซลล์สมอง ทำให้การทำงานของสมองเสื่อมสภาพลง จนทำให้มีผลกระทบต่อการทำกิจกรรมในชีวิตประจำวันของผู้ป่วย หลังจากที่ผู้ป่วยเป็นโรคนี้ในระยะเวลา 8 – 10 ปี จะทำให้ผู้ป่วยมีอาการสมองเสื่อมรุนแรงขึ้นเรื่อย จนไม่สามารถทำกิจกรรมในชีวิตประจำวันได้อย่างปกติ

2.1.2 แอนดรอยด์ [4]

เป็นซอฟต์แวร์ที่มีโครงสร้างแบบเรียงทับซ้อนหรือ แบบสแต็ก (Stack) ซึ่งรวมเอาระบบปฏิบัติการ มิดเดิลแวร์ และแอปพลิเคชันที่สำคัญเข้าไว้ด้วยกัน สำหรับทำงานบนอุปกรณ์พกพาเคลื่อนที่โดยเฉพาะ เช่น โทรศัพท์เคลื่อนที่ หรือ แท็บเล็ต เป็นต้น การทำงานของแอนดรอยด์มีพื้นฐานอยู่บนระบบลินุกซ์เคอร์เนล ซึ่งใช้ Android SDK (Software Development Kit) เป็นเครื่องมือสำหรับการพัฒนาแอปพลิเคชันบนระบบปฏิบัติการ Android และใช้ภาษา Java ในการพัฒนา

2.1.3 ระบบกำหนดตำแหน่งบนโลก [5]

คือระบบบอกตำแหน่ง โดยอาศัยการคำนวณพิกัดจากความถี่สัญญาณนาฬิกาที่ส่งมาจากดาวเทียมที่โคจรอยู่รอบโลกซึ่งทราบตำแหน่ง ทำให้ระบบนี้สามารถบอกตำแหน่ง ณ จุดที่สามารถรับสัญญาณได้ทั่วโลก

2.2 ผลงานที่เกี่ยวข้อง

2.2.1 Medicgeo controller [6] เป็นแอปพลิเคชันที่ใช้ในการติดตามผู้ป่วยอัลไซเมอร์ จากการระบุตำแหน่งของ GPS บนอุปกรณ์ต่างๆ เช่น โทรศัพท์เคลื่อนที่ และแท็บแล็ต ซึ่งแอปพลิเคชันนี้จะมีการติดตามผู้ป่วยตลอดเวลา มีความแม่นยำสูง สามารถระบุขอบเขตพื้นที่ที่ผู้ป่วยอยู่ โดยมีข้อดีคือ มีการแสดงตำแหน่งผู้ป่วย แต่ไม่มีการนำทางให้กับผู้ป่วย

2.2.2 Alzheimer Caretaker [7] เป็นแอปพลิเคชันที่ช่วยดูแลผู้ป่วยอัลไซเมอร์ที่เกิดอาการหลงลืมของผู้ป่วยในการใช้ชีวิตประจำวัน ซึ่งแอปพลิเคชันนี้มีการเพิ่มข้อมูลผู้ป่วย เพิ่มข้อมูลผู้ดูแล เพิ่มการนำทางสำหรับการทำกิจกรรมต่าง ๆ ของผู้ป่วยได้ มีการติดตามผู้ป่วยเมื่อผู้ป่วยเกิดการสูญหาย โดยใช้ระบุตำแหน่งผู้ป่วยจากตำแหน่ง GPS บนโทรศัพท์เคลื่อนที่ โดยมีข้อดีคือนำทาง แสดงตำแหน่งของผู้ป่วย มีข้อมูลของผู้ป่วย แต่ไม่มีการโทรฉุกเฉินให้กับผู้ป่วย

2.2.3 Alzheimer Patient [8] เป็นแอปพลิเคชันสำหรับที่ช่วยในการนำทางของผู้ป่วยไปยังสถานที่ต่าง ๆ เพื่อไม่ให้ผู้ป่วยหลงทาง มีตำแหน่งที่อยู่ปัจจุบันของตนเพื่อช่วยให้ผู้ป่วยรู้ตำแหน่งปัจจุบันได้ พร้อมทั้งกับสามารถตั้งค่าตำแหน่งของบ้านได้เพื่อให้นำทางกลับมาตำแหน่งบ้านได้ โดยมีข้อดีคือ แสดงตำแหน่งของผู้ป่วย แสดงตำแหน่งบ้าน นำทาง และมีปุ่มช่วยเหลือที่แจ้งไปยังผู้ดูแล แต่ไม่มีการแจ้งเตือนเมื่อผู้ป่วยออกนอกบริเวณที่กำหนด

2.2.4 Tweri Localizador Alzheimer [9] เป็นแอปพลิเคชันที่ใช้ระบุตำแหน่งของผู้ป่วยอัลไซเมอร์ โดยผู้ดูแลสามารถกำหนดขอบเขตของผู้ป่วยได้ มีระบบโทรฉุกเฉิน ช่วยให้ผู้ป่วยสามารถคุยกับผู้ดูแลได้ในกรณีฉุกเฉิน มีระบบแจ้งเตือนขอความช่วยเหลือเพื่อให้ช่วยเหลือผู้ป่วยได้ทันเวลา โดยมีข้อดีคือ โทรฉุกเฉิน แสดงตำแหน่งผู้ป่วย แจ้งขอความช่วยเหลือ แต่ไม่มีการแจ้งเตือนเมื่อผู้ป่วยออกนอกบริเวณที่กำหนด

2.2.5 Alzheimer Desease [10] เป็นแอปพลิเคชันที่ให้ความรู้เกี่ยวกับโรคอัลไซเมอร์ ประกอบไปด้วย 10 อาการเตือนของโรค แนวทางการป้องกันโรครวมถึงการดูแลผู้ป่วย ซึ่งในแอปพลิเคชันนี้จะมีเกม ช่วยฝึกทักษะทางด้านสมองสำหรับผู้ป่วย โดยมีข้อดีคือมีความรู้เกี่ยวกับโรค เกม แต่ไม่มีข้อมูลประวัติของผู้ป่วยและผู้ดูแล

จากแอปพลิเคชันที่ได้กล่าวมา เมื่อเปรียบเทียบกับแอปพลิเคชันช่วยเหลือผู้ป่วยอัลไซเมอร์ระยะแรกบนอุปกรณ์เคลื่อนที่ มีความแตกต่างกันที่สำคัญคือ การแจ้งเตือนการกินยา กิจกรรม แบบทดสอบความจำ แจ้งเตือนเมื่อผู้ป่วยออกนอกบริเวณ และค้นหาโทรศัพท์เคลื่อนที่

3. วิธีดำเนินงาน

3.1 วิเคราะห์ปัญหาและความต้องการ

การวิเคราะห์และหาความต้องการจากผู้ป่วยอัลไซเมอร์ระยะแรก และแพทย์ผู้เชี่ยวชาญจากโรงพยาบาลศรีนครินทร์

3.2 ค้นคว้า ศึกษาทฤษฎี และงานวิจัยที่เกี่ยวข้อง

- ศึกษาค้นคว้าทฤษฎีที่เกี่ยวข้องเกี่ยวกับ วิธีการดูแลผู้ป่วยอัลไซเมอร์ รวมถึงลักษณะพฤติกรรม และวิธีการป้องกันต่าง ๆ
- ศึกษาค้นคว้าเกี่ยวกับเครื่องมือที่ใช้สร้างแอปพลิเคชัน
- ศึกษาข้อมูลและวิธีใช้งานเครื่องมือต่าง ๆ ที่จะใช้ในงานวิจัย รายละเอียดทางด้านซอฟต์แวร์

3.3 การออกแบบและวิเคราะห์ระบบ

การออกแบบระบบ จะมีผู้ใช้งานหลัก 2 ส่วน คือส่วนของผู้ป่วยและผู้ดูแล โดยฟังก์ชันหลัก คือข้อมูลผู้ป่วยผู้ดูแล ตั้งค่า การกินยา กิจกรรม ความรู้เบื้องต้น แบบทดสอบความจำ นำทาง และโทรฉุกเฉิน (ดังภาพที่ 1)



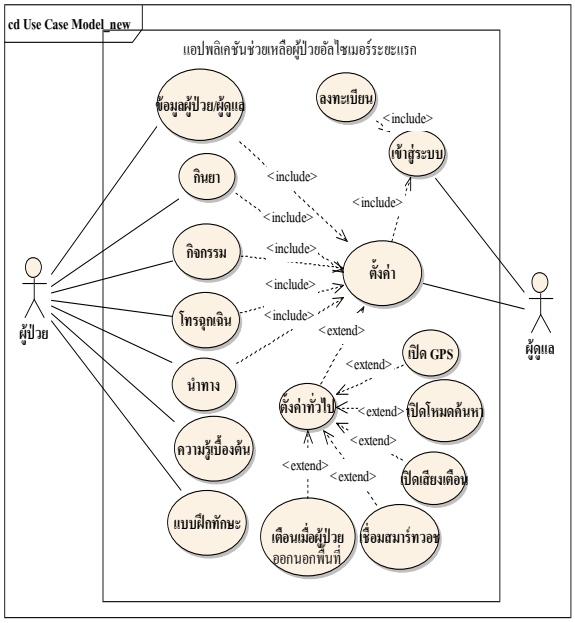
ภาพที่ 1 ภาพรวมของแอปพลิเคชัน [11]

3.4 Use Case Diagram (แผนภาพแบบจำลอง)

แผนภาพแสดงการใช้งานของผู้ใช้ แบ่งออกเป็น 2 ส่วนดังนี้ ส่วนสำหรับโทรศัพท์เคลื่อนที่และนาฬิกาอัจฉริยะ

3.4.1 Use Case Diagram โทรศัพท์เคลื่อนที่

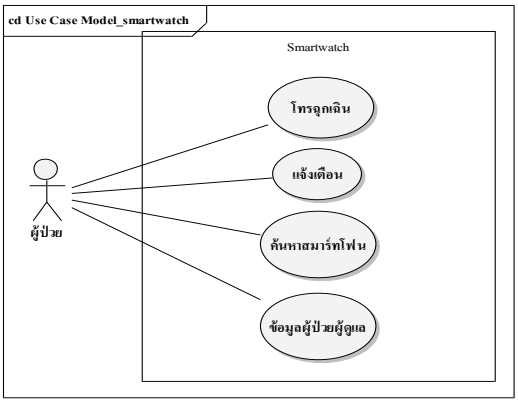
ผู้ใช้งานมี 2 ส่วน คือ ส่วนของผู้ดูแล สามารถเลือกใช้งานฟังก์ชันตั้งค่า ข้อมูลผู้ป่วยผู้ดูแล การกินยา กิจกรรม ผู้ป่วยอยู่ไหน และโทรหาผู้ป่วย และ ส่วนของผู้ป่วย สามารถเลือกใช้งานฟังก์ชันข้อมูลผู้ป่วยผู้ดูแล การกินยา กิจกรรม ความรู้เบื้องต้น แบบทดสอบความจำ นำทาง และโทรฉุกเฉิน



ภาพที่ 2 Use Case Diagram โทรศัพท์เคลื่อนที่

3.4.2 Use Case Diagram นาฬิกาอัจฉริยะ

ผู้ป่วยสามารถใช้งานการแจ้งเตือน โทรฉุกเฉิน ค้นหาสมาร์ทโฟน และดูข้อมูลผู้ป่วยผู้ดูแล



ภาพที่ 3 Use Case Diagram นาฬิกาอัจฉริยะ

3.4.3 เทคนิคที่ใช้ในการพัฒนา

- ฟังก์ชันกินยาและกิจกรรม ใช้เทคนิค notification ของ firebase ในการแจ้งเตือน
- ฟังก์ชันความรู้เบื้องต้น ใช้เทคนิค Text to Speech ในการอ่านเสียงข้อความ
- ฟังก์ชันนำทาง ผู้ป่วยอยู่ไหน และแจ้งเตือนเมื่อผู้ป่วยออกนอกบริเวณ ใช้เทคนิค GPS ในการระบุตำแหน่ง

4. ผลการดำเนินงาน

4.1 ตัวอย่างการพัฒนาแอปพลิเคชัน

4.1.1 หน้าเข้าสู่ระบบของแอปพลิเคชัน โดยผู้ใช้จะต้องลงทะเบียนชื่อผู้ป่วยและผู้ดูแลก่อนจึงจะเข้าสู่ระบบได้ (ดังภาพที่ 4 (ก))

4.1.2 หน้าหลักของผู้ป่วย ประกอบด้วย 7 ฟังก์ชัน คือข้อมูลผู้ป่วยผู้ดูแล กินยา กิจกรรม นำทาง แบบทดสอบความจำความรู้เบื้องต้น และโทรฉุกเฉิน (ดังภาพที่ 4 (ข))



(ก) (ข)

ภาพที่ 4 (ก) หน้าเข้าสู่ระบบของแอปพลิเคชัน **(ข)** หน้าหลักของแอปพลิเคชันบนเครื่องผู้ป่วย

4.1.3 หน้าหลักของผู้ป่วย ถ้าผู้ดูแลเข้าสู่ระบบเรียบร้อยแล้วจะแสดงหน้าหลักของผู้ป่วย โดยมี 6 ฟังก์ชัน คือข้อมูลผู้ป่วยผู้ดูแล กินยา กิจกรรม ผู้ป่วยอยู่ไหน โทรหาผู้ป่วย และตั้งค่า (ดังภาพที่ 5 (ก))

4.1.4 หน้าตั้งค่า กรณีผู้ดูแลเลือกฟังก์ชันตั้งค่า จะแสดงหน้าตั้งค่า มี 6 ฟังก์ชัน คือ ทั่วไป กิจกรรม กินยา นำทาง เบอร์โทรฉุกเฉิน และข้อมูลผู้ป่วยผู้ดูแล (ดังภาพที่ 5 (ข))

4.1.5 หน้าแผนที่นำทาง กรณีที่ผู้ป่วยเลือกฟังก์ชันนำทางจะแสดงภาพแผนที่ และสถานที่ให้เลือกไปได้อีก 3 แห่ง (ดังภาพที่ 6 (ก))

4.1.6 หน้าแสดงข้อมูลผู้ป่วย กรณีที่ผู้ป่วยหรือผู้ดูแลเลือกฟังก์ชันข้อมูลผู้ป่วยผู้ดูแลจะแสดงข้อมูลของผู้ป่วยและผู้ดูแล (ดังภาพที่ 6 (ข))



(ก) (ข)

ภาพที่ 5 (ก) หน้าหลักของแอปพลิเคชันบนเครื่องผู้ดูแล **(ข)** หน้าตั้งค่าของแอปพลิเคชันบนเครื่องผู้ดูแล



(ก) (ข)

ภาพที่ 6 (ก) หน้านำทางของแอปพลิเคชันบนเครื่องผู้ป่วย **(ข)** หน้าข้อมูลผู้ป่วย/ผู้ดูแลของแอปพลิเคชัน



(ก) (ข)

ภาพที่ 7 (ก) หน้าความรู้ทั่วไปของแอปพลิเคชัน **(ข)** หน้าแบบทดสอบความจำของแอปพลิเคชัน

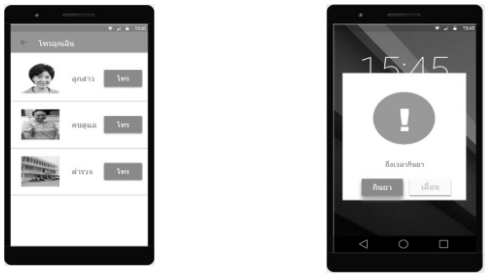
4.1.7 หน้าข้อมูลความรู้เบื้องต้น กรณีที่ผู้ป่วยเลือกฟังก์ชันความรู้เบื้องต้น จะแสดงหน้าข้อมูลความรู้เบื้องต้น (ดังภาพที่ 7 (ก))

4.1.8 หน้าแบบทดสอบ กรณีที่ผู้ป่วยเลือกเมนูแบบทดสอบความจำจะแสดงวิธีเล่นและปุ่มเริ่มเล่น (ดังภาพที่ 7 (ข))

4.1.9 หน้าโทรฉุกเฉิน กรณีที่ผู้ใช้เลือกฟังก์ชันโทรฉุกเฉิน จะแสดงรายชื่อและรูปภาพของผู้ดูแลและสถานีตำรวจ (ดังภาพที่ 8 (ก))

4.1.10 แจ้งเตือนผู้ดูแล เมื่อผู้ดูแลตั้งค่ากิจกรรมให้กับผู้ป่วยแล้ว แอปพลิเคชันจะแจ้งเตือนไปยังโทรศัพท์ของผู้ดูแล (ดังภาพที่ 8 (ข))

4.1.11 นาฬิกาอัจฉริยะ ประกอบด้วย 4 ฟังก์ชัน คือ โทรฉุกเฉิน แจ้งเตือน ค้นหาโทรศัพท์เคลื่อนที่ และข้อมูลผู้ป่วยผู้ดูแล (ดังภาพที่ 14)



(ก) (ข)

ภาพที่ 8 (ก) หน้าหลักของแอปพลิเคชันบนนาฬิกาอัจฉริยะ **(ข)** หน้าแจ้งเตือนของแอปพลิเคชันบนเครื่องผู้ป่วย



ภาพที่ 9 หน้าหลักของแอปพลิเคชันบนนาฬิกาอัจฉริยะ

4.2 การทดสอบฟังก์ชันหลักในแอปพลิเคชัน

ทดสอบโดยการเข้าใช้งานฟังก์ชันหลักของแอปพลิเคชันหลักตาม Use Case Diagram โดยทดสอบจำนวน 10 ครั้ง (ผ่าน/ไม่ผ่าน) ได้แก่ ฟังก์ชันหน้าหลักของผู้ป่วยทั้งหมด 6 ฟังก์ชัน ฟังก์ชันหน้าหลักของผู้ดูแลทั้งหมด 6 ฟังก์ชัน และฟังก์ชันหน้าบนนาฬิกาอัจฉริยะทั้งหมด 4 ฟังก์ชัน เข้าใช้งานได้ทุกฟังก์ชันตามที่กำหนด สรุปผลว่าผ่านจำนวน 10 ครั้งทุกฟังก์ชัน

4.3 แบบสอบถาม

เกณฑ์คะแนนที่ใช้ประเมินระดับความพึงพอใจของแบบสอบถาม มีทั้งหมด 5 ระดับ โดยแบ่งเกณฑ์คะแนนดังต่อไปนี้

- ค่าเฉลี่ย 4.50 – 5.00 อยู่ในระดับดีมาก
- ค่าเฉลี่ย 3.50 – 4.49 อยู่ในระดับดี
- ค่าเฉลี่ย 2.50 – 3.49 อยู่ในระดับปานกลาง
- ค่าเฉลี่ย 1.50 – 2.49 อยู่ในระดับพอใจ
- ค่าเฉลี่ย 0.00 – 1.49 อยู่ในระดับ ควรปรับปรุง

4.3.1 การทำแบบสอบถามของผู้ป่วยอัลไซเมอร์ระยะแรกและผู้ดูแลจากโรงพยาบาลศรีนครินทร์ จำนวน 10 คน ผลอยู่ในระดับดี ค่าเฉลี่ยเท่ากับ 3.74 (ดังตารางที่ 1)

4.3.2 การทำแบบสอบถามของแพทย์ผู้เชี่ยวชาญ จากโรงพยาบาลศรีนครินทร์ จำนวน 3 คน ค่าเฉลี่ยเท่ากับ 3.58 (ดังตารางที่ 2)

6. สรุปผลและข้อเสนอแนะ

แอปพลิเคชันช่วยเหลือผู้ป่วยอัลไซเมอร์ระยะแรกบนอุปกรณ์เคลื่อนที่ แบ่งออกเป็น 2 ส่วน คือโทรศัพท์เคลื่อนที่และนาฬิกาอัจฉริยะโดยที่ (1) ส่วนของโทรศัพท์เคลื่อนที่แบ่งผู้ใช้งานออกเป็น 2 ส่วน คือส่วนของผู้ป่วย ประกอบด้วย 7 ฟังก์ชัน ได้แก่ ข้อมูลผู้ป่วยผู้ดูแล กินยา กิจกรรม โทรฉุกเฉิน นำทาง ความรู้เบื้องต้น และแบบทดสอบความจำ ส่วนของผู้ดูแล ประกอบด้วย 6 ฟังก์ชันหลัก ได้แก่ ตั้งค่า ข้อมูลผู้ป่วย ผู้ดูแล กิจกรรม กินยา ผู้ป่วยอยู่ไหน และโทรหาผู้ป่วย (2) นาฬิกาอัจฉริยะ ประกอบด้วย 4 ฟังก์ชัน ได้แก่ ข้อมูลผู้ป่วย ดูแล แจ้งเตือนกิจกรรม การกินยา โทรฉุกเฉิน และค้นหาโทรศัพท์เคลื่อนที่

ทั้งนี้จากการทดสอบฟังก์ชันหลักทั้ง 2 ส่วน แอปพลิเคชันสามารถทำงานได้ตามที่ผู้ใช้งานต้องการ เช่น ผู้ดูแลสามารถตั้งค่าการกินยา กิจกรรมให้กับผู้ป่วย และผู้ป่วยสามารถดูการแจ้งเตือนการกินยา และกิจกรรมได้ ส่วนของการประเมินความพึง

พอใจต่อใช้งานแอปพลิเคชัน ได้สำรวจผู้ตอบแบบสอบถามจากโรงพยาบาลศรีนครินทร์ ซึ่งผลการตอบแบบสอบถามของผู้ป่วยและผู้ดูแล อยู่ในระดับดี ค่าเฉลี่ยเท่ากับ 3.74 และผลการตอบแบบสอบถามของแพทย์ผู้เชี่ยวชาญ อยู่ในระดับดี ค่าเฉลี่ยเท่ากับ 3.58

การพัฒนาแอปพลิเคชันเป็นไปตามวัตถุประสงค์และขอบเขตที่ตั้งไว้ คือ สามารถใช้งานโทรศัพท์เคลื่อนที่ได้โดยไม่จำเป็นต้องมีนาฬิกาอัจฉริยะ มีการแจ้งเตือนกิจกรรม และการกินยาผ่านหน้าจอโทรศัพท์เคลื่อนที่และนาฬิกาอัจฉริยะของผู้ป่วย หากผู้ป่วยออกนอกบริเวณที่กำหนดไว้ จะมีการแจ้งเตือนไปยังผู้ดูแลทันทีและผู้ดูแลสามารถดูตำแหน่งปัจจุบันของผู้ป่วย สามารถค้นหาโทรศัพท์เคลื่อนที่ผ่านฟังก์ชันบนนาฬิกาอัจฉริยะ ข้อจำกัดของแอปพลิเคชัน แอปพลิเคชันใช้งานกับโทรศัพท์เคลื่อนที่และนาฬิกาอัจฉริยะ บนระบบปฏิบัติการแอนดรอยด์ รุ่นที่ 6.0 ขึ้นไป โทรศัพท์เคลื่อนที่และนาฬิกาอัจฉริยะต้องเชื่อมต่ออินเทอร์เน็ต ในการพัฒนาระบบต่อไป ผู้พัฒนาควรศึกษารายละเอียดและวิธีการใช้เครื่องมือ และวิธีการพัฒนาแอปพลิเคชันให้มีประสิทธิภาพมากยิ่งขึ้น

7. เอกสารอ้างอิง

[1] ฐานข้อมูลประชากร (ฝ่ายวิจัยและศูนย์สารสนเทศฯ). [ออนไลน์]. ค้นเมื่อ 6 เมษายน 2559. ที่มา http://www.cps.chula.ac.th

[2] รู้จัก Smart Watch นาฬิกาอัจฉริยะคอมพิวเตอร์สวมใส่ที่เริ่มมาไทย. [ออนไลน์]. ค้นเมื่อ 4 เมษายน 2559 . ที่มา http://www.it24hrs.com

[3] โรคอัลไซเมอร์ (Alzheimer). [ออนไลน์]. ค้นเมื่อ 4 เมษายน 2559. ที่มา http://guru.sanook.com

[4] แอนดรอยด์. [ออนไลน์]. ค้นเมื่อ 29 สิงหาคม 2559. ที่มา https:// beer kung .wordpress.com/ระบบปฏิบัติการรุ่นล่าสุดระบบปฏิบัติการ

[5] ระบบกำหนดตำแหน่งบนโลก. [ออนไลน์]. ค้นเมื่อ 29 สิงหาคม 2559. ที่มา https://th.wikipedia.org/wiki/ระบบกำหนดตำแหน่งบนโลก

[6] Android กับ SQLite Database การเขียนแอนดรอยด์เพื่อติดต่อกับฐานข้อมูลของ SQLite. [ออนไลน์]. ค้นเมื่อ 7 เมษายน 2559. ที่มา http:// www.thaicreate.com/mobile/android-sqlite-database.html

[7] บลูทูธ "ฟันสีฟ้า" เทคโนโลยีไร้สายสำหรับอนาคต. [ออนไลน์]. ค้น เมื่อ 4 เมษายน 2559. ที่มา http://www.siamphone.com

[8] ภาพรวมของ Photoshop CS6. [ออนไลน์]. ค้นเมื่อ 5เมษายน 2559. ที่ มา http: / / www. ninetechno. com/ a/ photoshop-cs6/ 715-all-photoshopcs6.html

[9] เทคโนโลยีเสียง. [ออนไลน์]. ค้นเมื่อ 20 พฤษภาคม 2559. ที่มา https:// educationradio.wikispaces.com/เทคโนโลยีเสียง.html

[10] อารี ตัณฑ์เจริญรัตน์ วารสารศึกษาศาสตร์ มหาวิทยาลัยนเรศวร. [ออนไลน์]. ค้นเมื่อ 18 พฤษภาคม 2559. ที่มา https://www.tci-thaijo.org/index.php/edujournal_nu/article/view/9285/8396

[11] รูปผู้ป่วยอัลไซเมอร์. [ออนไลน์]. ค้นเมื่อ 20 พฤษภาคม 2559. ที่มา https://www.google.co.th/search?qรูปผู้ป่วยอัลไซเมอร์

ตารางที่ 1 ค่าคะแนนจากการทำแบบสอบถามของผู้ป่วยและผู้ดูแล

รายการประเมิน	ระดับความพึงพอใจ					
	มากที่สุด (5)	มาก (4)	พอใช้ (3)	น้อย (2)	น้อยที่สุด (1)	ค่าเฉลี่ย (คะแนน)
1. ด้านการใช้งานโทรศัพท์มือถือ						
โทรศัพท์มือถือมีความจำเป็นต่อคุณ	2	5	1	2	0	3.70
การใช้งานใช้งานในแต่ละวัน	1	6	2	1	0	3.70
2. ด้านการออกแบบแอปพลิเคชัน						
การออกแบบมีความเข้าใจง่าย	0	6	2	2	0	3.40
ขนาดของตัวอักษรมีความเหมาะสม และชัดเจน	0	3	6	1	0	3.20
เมนูง่ายต่อการใช้งาน	0	4	5	1	0	3.30
3. ด้านเจตคติต่อการใช้แอปพลิเคชัน						
คุณมีความชอบต่อแอปพลิเคชันมากเพียงใด	3	5	1	1	0	4.00
แอปพลิเคชันมีประโยชน์ต่อการใช้ชีวิตประจำวันของคุณมากเพียงใด	3	4	2	1	0	4.30
แอปพลิเคชันมีความจำเป็นต่อคุณมากเพียงใด	2	6	2	0	0	4.00
ภาพรวมของแอปพลิเคชันนี้	3	5	2	0	0	4.10
รวม	**8**	**48**	**24**	**9**	**0**	**3.74**

ตารางที่ 2 ค่าคะแนนจากการทำแบบสอบถามของแพทย์ผู้เชี่ยวชาญ

รายการประเมิน	ระดับความพึงพอใจ					
	มากที่สุด (5)	มาก (4)	พอใช้ (3)	น้อย (2)	น้อยที่สุด (1)	ค่าเฉลี่ย (คะแนน)
1. ด้านข้อมูลผู้ป่วย/ผู้ดูแล						
ความเหมาะสมการจัดวางตำแหน่งรูป ข้อความ	0	2	1	0	0	3.67
มีความชัดเจน เข้าใจง่าย	0	2	1	0	0	3.67
2. ด้านแจ้งเตือนกิจกรรม						
การออกแบบมีความเข้าใจง่าย	0	1	2	0	0	3.33
เสียงแจ้งเตือนมีความเหมาะสม	0	1	2	0	0	3.33
เมนูง่ายต่อการใช้งาน	0	2	1	0	0	3.67
3. ด้านการเรียนรู้						
ขนาดของตัวอักษรมีความเหมาะสม และชัดเจน	0	1	2	0	0	3.33
มีการใช้ภาษาที่เหมาะสม	0	2	1	0	0	3.67
เนื้อหาเข้าใจได้ง่าย	0	3	0	0	0	4.00
4. ด้านนำทาง						
การออกแบบมีความเหมาะสม	0	3	0	0	0	4.00
เส้นทางมีความเข้าใจง่ายไม่ซับซ้อน	0	3	0	0	0	4.00
เมนูมีความสะดวกต่อการใช้งาน	0	1	2	0	0	3.33
5. ด้านการออกแบบของแบบทดสอบ						
การออกแบบสื่อให้ผู้ใช้เข้าใจง่าย	0	1	2	0	0	3.33
ความเหมาะสมขนาดของตัวอักษร และชัดเจน	0	1	2	0	0	3.33
เมนูง่ายต่อการใช้งาน	0	1	2	0	0	3.33
สีสันในการออกแบบเกมมีความเหมาะสม	0	2	1	0	0	3.67
6. ด้านโทรฉุกเฉิน						
การออกแบบมีความเข้าใจง่าย	0	2	1	0	0	3.67
เมนูง่ายต่อการใช้งาน	0	1	2	0	0	3.33
โทรเชื่อมต่อกับเบอร์โทรฉุกเฉินได้ง่าย ไม่ซับซ้อน	0	2	1	0	0	3.67
ความพึงพอใจโดยรวม	0	2	1	0	0	3.67
รวม	**0**	**33**	**24**	**0**	**0**	**3.58**

เกมจำลองชนิดสามมิติเพื่อพัฒนาความรู้เกี่ยวกับคำศัพท์ภายในบ้าน
A 3D Simulation Game for Improving Knowledge of Vocabulary in the House

ชุติพนธ์ บำรุงจิตต์ (Chutipon Blumrungjit)[1] ปรีชา คางงูเหลือม (Preecha Kangngulam)[2]
เดชาวุฒิ วานิชสรรพ์ (Dechawut Wanichsan)[3] และเอื้อมพร รุ่งศิริ (Auemporn Rungsiri)[4]
[1,2]คณะครุศาสตร์ [3]คณะวิทยาการคอมพิวเตอร์และเทคโนโลยีสารสนเทศ [4]คณะมนุษยศาสตร์และสังคมศาสตร์
มหาวิทยาลัยราชภัฏรำไพรรณี
[1]chutiponblumrungjit@gmail.com, [2]preechadekcom@gmail.com, [3]dechawut.w@rbru.ac.th, [4]auemporn.r@rbru.ac.th

บทคัดย่อ

งานวิจัยชิ้นนี้มีวัตถุประสงค์เพื่อพัฒนาเกมสามมิติเกี่ยวกับคำศัพท์ภายในบ้านสำหรับนักเรียนชั้นประถมศึกษาปีที่ 1-3 โดยเกมจะให้สถานการณ์จำลองให้ผู้เล่นได้เดินสำรวจภายในห้องว่ามีสิ่งของใดอยู่ในห้องบ้าง เมื่อเจอสิ่งของแล้วให้สะกดคำศัพท์ให้ถูกต้อง หากผู้เล่นตอบผิด เกมจะมีระบบช่วยเหลือที่หลากหลาย เช่น การออกเสียงคำศัพท์ และการเผยตัวอักษรที่เป็นสระ เป็นต้น จากนั้นจึงทำการวัดผลสัมฤทธิ์ทางการเรียนและความพึงพอใจของผู้ใช้งานจำนวน 15 คน โดยผลการทดลองแสดงให้เห็นว่าผู้ใช้งานที่เล่นเกมมีผลสัมฤทธิ์ทางการเรียนรายวิชาภาษาอังกฤษเรื่องคำศัพท์ภายในบ้านสูงขึ้น และผู้ใช้งานมีความพึงพอใจในเกมที่พัฒนาขึ้นในระดับดีมาก

คำสำคัญ: การจำลอง สามมิติ เกม วิชาภาษาอังกฤษ คำศัพท์ ศึกษาบันเทิง

Abstract

This research aimed to develop a three-dimensional game about vocabularies inside the house for grades 1-3. The game provided a virtual situation that a user could investigate several rooms to see what were in those. When an object was clicked, a user was challenged by putting alphabets to construct a word. If a wrong answer was put, the game could offer a support system such as pronouncing the word, and hinting vowel alphabets, etc. Additionally, the learning achievement and satisfaction of fifteen participants were evaluated. The experimental results showed that it could help learners to enhance their learning achievement, and satisfaction level was very high.

Keyword: Simulation, Three Dimension, Game, English, Vocabulary, Edutainment

1. บทนำ

ภาษาอังกฤษมีบทบาทสำคัญในชีวิตประจำวันของการดำรงชีวิตในฐานะพลโลกเป็นอย่างมาก โลกปัจจุบันนี้มีการเปลี่ยนแปลงอย่างรวดเร็วในด้านต่าง ๆ โดยเฉพาะการติดต่อสื่อสารที่ได้มีการเปลี่ยนแปลงอย่างรวดเร็ว การขยายตัวของอินเทอร์เน็ต และเทคโนโลยีบนอุปกรณ์เคลื่อนที่ ในปัจจุบันข้อมูลบนอินเทอร์เน็ตมากกว่าห้าสิบเปอร์เซ็นต์นำเสนอข้อมูลเป็นภาษาอังกฤษ [1] นอกจากนี้ภาษาอังกฤษยังเป็นภาษาสำคัญในการก้าวสู่ประชาคมอาเซียน เนื่องจากได้รับการยอมรับให้เป็นภาษากลางที่ใช้ในการสื่อสาร

ในด้านเทคโนโลยีสารสนเทศ สถาบันวิชาชีพวิศวกรไฟฟ้าและอิเล็กทรอนิกส์ (Institute of Electrical and Electronics Engineers: IEEE) ได้ทำการสำรวจโปรแกรมภาษาคอมพิวเตอร์ที่เป็นนิยมที่สุดสิบอันดับแรกในปี 2559 พบว่าทุกภาษาใช้ภาษาอังกฤษเป็นหลักในการสร้างชุดคำสั่ง [2] โดยรวมถึงภาษาไพธอน (Python) และรูบี (Ruby) ที่พัฒนาขึ้นโดยโปรแกรมเมอร์ที่เป็นพลเมืองของประเทศที่ไม่ได้ใช้ภาษาอังกฤษเป็นภาษาหลักในการสื่อสาร โดยที่ภาษาคอมพิวเตอร์ทั้งสองพัฒนาขึ้นโดยชาวเนเธอร์แลนด์และชาวญี่ปุ่นตามลำดับ

อย่างไรก็ตาม ในปี 2559 ความสามารถในการใช้ภาษาอังกฤษของคนไทยถูกจัดอันดับเป็นอันดับที่ 56 จาก 72 ประเทศที่ไม่ได้ใช้ภาษาอังกฤษเป็นภาษาหลักในการสื่อสาร และถูกจัดอยู่กลุ่มที่ความสามารถในการใช้ภาษาอังกฤษอยู่ในระดับต่ำมาก [3] และจำเป็นต้องได้รับการพัฒนาอย่างเร่งด่วน หากต้องการดำรงชีวิตในฐานะพลโลกที่มีคุณภาพ

นักวิจัยของมหาวิทยาลัยวอชิงตัน [4] กล่าวว่าการเรียนรู้ด้านภาษามีช่วงเวลาที่ดีที่สุดในการเรียนรู้ คือตั้งแต่ผู้เรียนยังเป็นทารกจนกระทั่งอายุเจ็ดขวบ และความสามารถในการเรียนรู้ด้านภาษาจะค่อย ๆ ลดลงตามลำดับ

จากข้อมูลและปัญหาดังกล่าว งานวิจัยชิ้นนี้จึงมีแนวคิดที่จะพัฒนาเกมจำลองชนิดสามมิติเพื่อการศึกษาในรายวิชาภาษาอังกฤษเกี่ยวกับคำศัพท์ภายในบ้านในระดับชั้นประถมศึกษาปีที่ 1-3 เพื่อช่วยส่งเสริมการเรียนรู้ในรายวิชาภาษาอังกฤษให้ผู้เรียนได้มีความรู้ในคำศัพท์ภาษาอังกฤษมากยิ่งขึ้น อีกทั้งผู้เรียนในช่วงอายุดังกล่าวอยู่ในช่วงอายุที่เหมาะสมต่อการจัดการเรียนการสอนภาษาอังกฤษ

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 โปรแกรมประยุกต์ที่ใช้ในการพัฒนา

1) โปรแกรมยูนิตี (Unity) คือเครื่องมือที่ใช้ในการพัฒนาเกมเพื่อใช้งานบนเครื่องพีซี (Personal Computer: PC) เกมคอนโซล อุปกรณ์เคลื่อนที่ต่าง ๆ รวมถึงเว็บไซต์ เกมที่พัฒนาขึ้นสามารถทำงานได้บนระบบที่หลากหลาย เช่น ระบบปฏิบัติการวินโดวส์ (Windows) ไอโอเอส (iOS) และแอนดรอยด์ (Android) เป็นต้น ภาษาคอมพิวเตอร์ที่ใช้ในการพัฒนาโปรแกรมบนยูนิตีคือภาษาซีชาร์ป (C#) และยูนิตีสคริปต์

2) โปรแกรมเบลนเดอร์ (Blender) เป็นฟรีแวร์สำหรับงานคอมพิวเตอร์กราฟิกสามมิติ สามารถใช้สร้างโมเดลสามมิติ ทำพื้นผิว จัดการการเคลื่อนไหวแบบใช้กระดูก จำลองการไหลของน้ำ จำลองผิวหนัง สร้างคอมพิวเตอร์แอนิเมชัน เบลนเดอร์มีคุณลักษณะทัดเทียมกับโปรแกรมสามมิติระดับสูงอื่น ๆ เช่น ซอฟต์อิมเมจ (Softimage) ไลท์เวฟ (Lightwave) ออโต้เดสก์สามดีเอสแมกซ์ (Autodesk 3ds Max) เป็นต้น สามารถส่งออกแบบจำลองสามมิติเพื่อให้โปรแกรมอื่นนำไปใช้งานต่อได้โดยการบันทึกไฟล์เป็นนามสกุล .fbx

2.2 ทฤษฎีที่เกี่ยวข้องด้านวิธีการเรียนรู้

1) การจัดการเรียนรู้ด้วยเกม (Game-based Learning) [5] คือการจัดกิจกรรมการเรียนการสอนด้วยการใช้เกมโดยมีวัตถุประสงค์หลักคือเพื่อให้ผู้เรียนเกิดการเรียนรู้ในขณะเล่นเกม พร้อมกับได้รับความเพลิดเพลินไปด้วย โดยทั่วไปเกมประเภทนี้ได้รับการพัฒนาขึ้นโดยกำหนดเป้าหมายในการเรียนรู้ไว้ เกมได้รับการออกแบบมาเพื่อให้เนื้อหาที่ต้องการสอดแทรก กับระยะเวลาที่ใช้ในการเล่นเกมเป็นไปอย่างสมดุล ทำให้ผู้เรียนเกิดการเรียนรู้อย่างมีความหมาย ผู้เรียนต้องเล่นเกมตามกติกาเพื่อให้บรรลุภารกิจ จนท้ายที่สุดผู้เล่นสามารถเชื่อมโยงความรู้ที่ได้รับกับเนื้อหาในบทเรียนได้

2) การจัดการเรียนรู้ด้วยสถานการณ์จำลอง (Simulation-based Learning) [6] คือ กระบวนการที่ผู้สอนใช้ในการช่วยให้ผู้เรียนเกิดการเรียนรู้ตามวัตถุประสงค์ที่กำหนด โดยให้ผู้เรียนลงไปเล่นในสถานการณ์ที่มีบทบาท ข้อมูล และกติกาการเล่น ที่สะท้อนความเป็นจริง และมีปฏิสัมพันธ์กับสิ่งต่าง ๆ ที่อยู่ในสถานการณ์นั้น ๆ โดยใช้ข้อมูลที่มีสภาพคล้ายกับข้อมูลในความเป็นจริงในการตัดสินใจและแก้ปัญหา ซึ่งการตัดสินใจนั้นจะส่งผลถึงผู้เล่นในลักษณะเดียวกันที่เกิดขึ้นในสถานการณ์จริง

2.3 งานวิจัยที่เกี่ยวข้อง

สุวิชัย พรรษา และอภินันท์ รังสูงเนิน [7] ได้ศึกษาและพัฒนาเกมคอมพิวเตอร์สามมิติ ที่ให้ความรู้เกี่ยวยาเสพติดเจ็ดชนิด เพื่อส่งเสริมให้เยาวชนได้รับรู้ถึงโทษของมันอันส่งผลให้ไม่ต้องการยุ่งเกี่ยวกับยาเสพติด โดยที่เกมดังกล่าวได้นำไปทดลองใช้กับเยาวชนอายุตั้งแต่เจ็ดปีขึ้นไปจำนวน 35 คน ผลการวิจัยพบว่า ผู้ร่วมโครงการมีความพึงพอใจเฉลี่ยอยู่ในระดับดี

Gulec, U. และ Yilmaz, M. [8] ได้ศึกษาเกมที่ใช้สมาธิเล่นอย่างจริงจัง ซึ่งใช้เพื่อการพัฒนาทักษะการตัดสินใจ และระดับความรู้เกี่ยวกับการเล่นฟุตบอลในตุรกี โดยที่เกมดังกล่าวได้นำไปทดลองใช้กับผู้ตัดสิน 54 คน ผลการวิจัยที่เปรียบเทียบความรู้ด้วยการทดสอบ *t*-test พบว่าผู้เรียน

ที่เรียนรู้ด้วยเกมมีผลสัมฤทธิ์ในการเรียนที่สูงกว่าผู้เรียนที่เรียนรู้ด้วยหนังสือคู่มือ

Amoia, M. และ Gardent, A. [9] ได้สร้างเกมเพื่อการเรียนรู้ภาษาฝรั่งเศสในฐานะภาษาที่สอง ซึ่งเป็นเกมจำลองสังคมสามมิติ และใช้ประโยชน์จากสภาพแวดล้อมเสมือนจริงเพื่อนำมาใช้ในการเรียนรู้ด้วยการลงมือทำ (Learning-by-Doing) ผลการทดลองแสดงให้เห็นว่าวิธีการเรียนรู้แบบนี้เป็นวิธีที่มีประสิทธิภาพในการเรียนรู้ภาษาที่สอง

Rankin, Y. และคณะ [10] พัฒนาเกม Ever Quest2 ที่เป็นเกมสวมบทบาทแบบที่มีผู้เล่นหลายคน เพื่อเป็นเครื่องมือในการสอนภาษาอังกฤษในฐานะภาษาที่สอง ผลการศึกษาเบื้องต้นแสดงให้เห็นว่านักเรียนระดับกลางและระดับสูงมีความรู้ภาษาอังกฤษที่พัฒนาขึ้น ผลการทดลองแสดงให้เห็นว่าโลกเสมือนจริงสามารถให้แรงบันดาลใจที่เพียงพอในการสนับสนุนการเรียนภาษาของผู้เรียนได้

งานวิจัยทุกเรื่องข้างต้นแสดงให้เห็นว่าเกมเป็นสื่อการสอนที่มีประสิทธิภาพในการใช้สนับสนุนการเรียนรู้โดยเฉพาะอย่างยิ่งกับการเรียนรู้ด้านภาษาด้วยการใช้เกมเป็นส่วนประกอบ จากการสรุปใจความสำคัญทั้งจากทฤษฎีที่เกี่ยวข้องกับวิธีการจัดการเรียนรู้ ทำให้ได้แนวคิดในการพัฒนาเกม ที่มีคุณสมบัติดังนี้ 1) เกมต้องทำให้ผู้ใช้จำลองบทบาทเสมือนเดินอยู่ในบ้านและพบของใช้ที่สามารถพบเจอได้ในชีวิตประจำวัน เนื่องจากวิธีการเรียนรู้คำศัพท์ที่ดีควรมีความสัมพันธ์กับประสบการณ์ที่พบเจอในชีวิตประจำวัน อันจะทำให้เพิ่มความสามารถในการจดจำได้ดียิ่งขึ้น 2) ของใช้ที่เจอในบ้านต้องมีรูปร่างและขนาดที่ใกล้เคียงกับวัตถุจริงมากที่สุด เกมนี้จึงออกแบบในรูปแบบของเกมสามมิติ เพื่อหลีกเลี่ยงความเข้าใจผิดในเนื้อหาที่เกมต้องการจะสื่อ และ3) เกมเพื่อการศึกษาที่ดีต้องมีระบบช่วยเหลือผู้เล่นเมื่อไม่สามารถทำภารกิจได้สำเร็จ หากเกมมีความยากเกินไปผู้เล่นจะถอดใจและไม่เล่นเกมต่อทำให้เสียโอกาสในการเรียน

3. วิธีดำเนินงานวิจัย

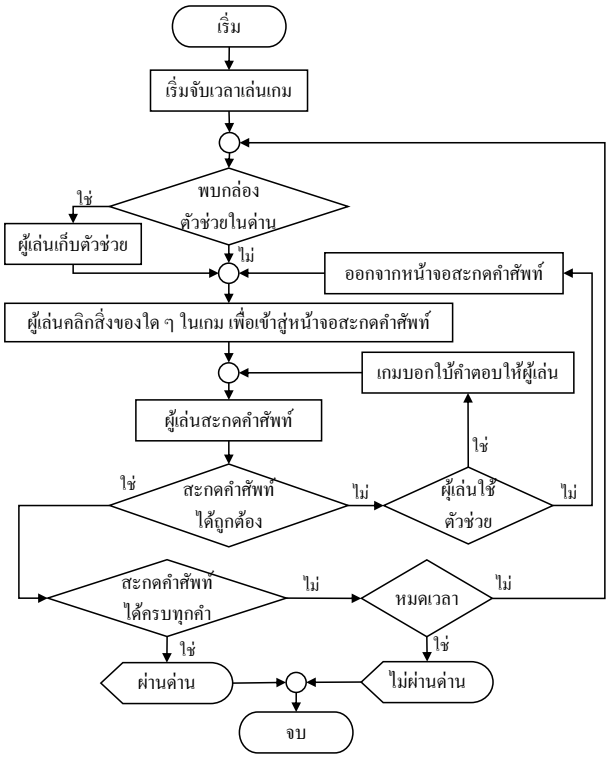
3.1 เกมและส่วนประกอบของเกม

เกมจำลองชนิดสามมิติได้รับการพัฒนาขึ้นโดยใช้โปรแกรมยูนิตี ในส่วนโมเดลสามมิติของใช้ภายในบ้านได้รับการขึ้นรูปและใส่พื้นผิว (Texture) ด้วยโปรแกรมเบลนเดอร์ เนื้อหาคำศัพท์ภายในเกม เป็นเนื้อหาเกี่ยวคำศัพท์ภายในบ้านสำหรับนักเรียนชั้นประถมศึกษาปีที่ 1-3 ซึ่งนำมาจากหนังสือแบบเรียนภาษาอังกฤษ เรื่องคำศัพท์ภาษาอังกฤษภายในบ้าน [11] อย่างไรก็ตาม คำศัพท์ภายในบ้านมีจำนวนมาก ดังนั้นผู้วิจัยได้ปรึกษากับผู้เชี่ยวชาญซึ่งเป็นอาจารย์ผู้สอนในสาขาวิชาภาษาอังกฤษธุรกิจ ที่มีประสบการณ์สอนในระดับอุดมศึกษาไม่น้อยกว่าห้าปี และให้ผู้เชี่ยวชาญเป็นผู้เลือกคำศัพท์โดยมีเกณฑ์การเลือกดังนี้ 1) เลือกคำศัพท์จากของใช้ภายในบ้านที่สามารถพบทั่วไปในบ้าน และพบเจอในชีวิตประจำวัน 2) เลือกคำศัพท์ที่มีความยากในระดับเหมาะสมกับกลุ่มตัวอย่าง โดยปริมาณคำศัพท์จะไม่มากเกินไป เพราะอาจทำให้เกิดความรู้สึกเบื่อหน่าย คำศัพท์ทั้งหมดที่เลือกมาแบ่งเป็นหมวดของใช้ได้สี่หมวดหมู่ได้แก่ ในห้องนอน ในห้องน้ำ ในห้องครัว และในห้องนั่งเล่น

3.2 รูปแบบการเล่น

ผู้เล่นจะต้องสะกดคำศัพท์ภาษาอังกฤษให้สอดคล้องกับวัตถุสามมิติที่เลือกไว้ เช่นต้องสะกด PILLOW เมื่อเลือกวัตถุสามมิติที่เป็นรูปหมอน ในขั้นเริ่มต้นเมื่อเข้าสู่เกมผู้เล่นสามารถเลือกด่านในเกม โดยด่านแต่ละด่านคือห้องสี่ห้องได้แก่ ห้องนอน ห้องน้ำ ห้องครัว และห้องนั่งเล่น จากนั้นเกมจะแสดงหนังสืออิเล็กทรอนิกส์ให้กับผู้เล่นสำหรับศึกษาข้อมูลคำศัพท์ภาษาอังกฤษ การอ่านออกเสียงคำ และคำแปลภาษาไทย จากนั้นผู้เล่นจึงเข้าสู่เกมโดยขั้นตอนการเล่นเกมทั้งหมดแสดงด้วยผังงานในภาพที่ 1

จากภาพที่ 1 ตัวช่วยเหลือในเกมมีอยู่สามชนิด ที่ผู้เล่นควรใช้เมื่อไม่สามารถสะกดคำศัพท์ได้ถูกต้อง ตัวช่วยแรกใช้เพื่อเปล่งเสียงการสะกดคำศัพท์นั้น ตัวช่วยนี้เปรียบเหมือนครูเอาการคำศัพท์นักเรียน ตัวช่วยที่สองใช้เพื่อแสดงอักขระที่เป็นสระทั้งหมดที่ปรากฏอยู่ในคำศัพท์ เพื่อให้ผู้เล่นสะกดคำศัพท์ได้ง่ายขึ้น เช่น กรณีคำศัพท์ “หมอน” ผลลัพธ์ที่ได้บนจอภาพจะเป็น **_I__O_** ตัวช่วยที่สามใช้เพื่อแสดงอักขระที่เป็นพยัญชนะทั้งหมดที่ปรากฏอยู่ในคำศัพท์ เช่นกรณีคำศัพท์ “หมอน” ผลลัพธ์ที่ได้จะเป็น ผลลัพธ์ที่ได้บนจอภาพจะเป็น **P_LL_W** ตัวช่วยทั้งสามทำให้ระดับความยากของโจทย์ข้อนั้นลดลง เพื่อให้ผู้เล่นสามารถเล่นเกมต่อได้



ภาพที่ 1: ขั้นตอนการเล่นเกมที่ออกแบบไว้

3.3 เครื่องมือที่ใช้ในการวัดประสิทธิภาพของเกม

เพื่อศึกษาประสิทธิภาพของเกมที่พัฒนาขึ้น เครื่องมือที่พัฒนาขึ้นมีดังนี้

1) แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียน เป็นข้อสอบจำนวน 40 ข้อ แต่ละข้อมีสี่ตัวเลือก ก่อนนำไปใช้งานจริง ข้อสอบชุดนี้ถูกนำไปทดสอบกับผู้ทดสอบจำนวนสิบคน เพื่อวิเคราะห์หาค่าความยากง่ายของข้อสอบ [12] โดยผลที่ได้พบว่าข้อสอบแต่ละข้อมีค่า P อยู่ระหว่าง 0.20-0.80 ซึ่งหมายความว่าข้อสอบแต่ละข้อมีความยากง่ายในระดับเหมาะสมที่จะนำไปใช้จริง ข้อสอบชุดนี้ได้รับการนำไปสลับลำดับของคำถาม และคำตอบเพื่อนำไปสร้างเป็นข้อสอบก่อนเรียน และข้อสอบหลังเรียน เพื่อทำให้ข้อสอบทั้งสองชุดมีระดับความยากที่ไม่แตกต่างกัน

2) แบบสอบถามวัดความพึงพอใจจำนวน 10 ข้อ ที่ใช้เพื่อการวัดความพึงพอใจสองด้าน ได้แก่ ด้านการออกแบบเกม และด้านความเหมาะสมของเนื้อหาในเกม

เมื่อเครื่องมือที่ใช้ในการวัดผลเสร็จสมบูรณ์ จะนำไปทดลองใช้กับผู้ร่วมวิจัยระดับประถมศึกษาปีที่สาม โรงเรียนวัดหนองบัว (ปทุมาภิพัฒน์) จำนวน 15 คน

4. ผลการดำเนินงาน

4.1 เกมจำลองชนิดสามมิติเพื่อพัฒนาความรู้เกี่ยวกับคำศัพท์ภายในบ้าน

เมื่อผู้ใช้งานเปิดเกมคำศัพท์ภาษาอังกฤษจะพบหน้าจอหลักของเกมดังแสดงในภาพที่ 2 ในหน้าจอนี้จะมีปุ่มให้ผู้ใช้งานกดอยู่สี่ปุ่ม สามปุ่มอยู่บริเวณด้านขวาของจอภาพที่ใช้สำหรับการตั้งค่าเสียง แสดงวิธีการเล่น และออกจากเกมตามลำดับบนลงล่าง

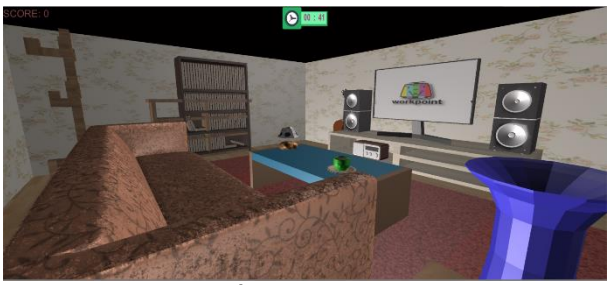


ภาพที่ 2 หน้าจอแรกของเกม

จากหน้าแรก เมื่อผู้เล่นคลิกที่ปุ่ม “PLAY” เกมจะแสดงหน้าจอประตูสี่ประตู เพื่อผู้เล่นเลือกเข้าสู่ห้องแต่ละห้อง ได้แก่ห้องนอน ห้องครัว ห้องน้ำ และห้องนั่งเล่น ซึ่งหมายถึงด่านในเกมทั้งสี่ด่าน โดยหน้าจอห้องนอน ห้องนั่งเล่น ห้องครัวและห้องน้ำ แสดงดังภาพที่ 3-6 ตามลำดับ



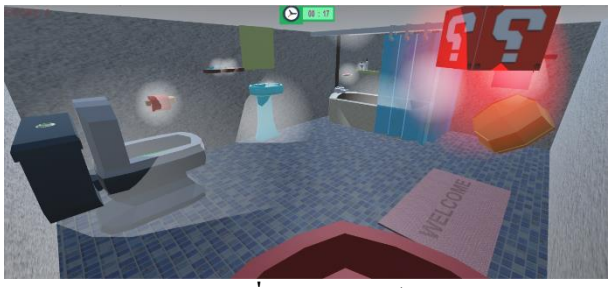
ภาพที่ 3 หน้าจอห้องนอน



ภาพที่ 4 หน้าจอห้องนั่งเล่น



ภาพที่ 5 หน้าจอห้องครัว



ภาพที่ 6 หน้าจอห้องน้ำ

เมื่อเข้าสู่ด่านที่เลือก ผู้เล่นมีเวลา 5 นาทีในการสะกดคำศัพท์ให้ถูกต้องจำนวน 10 คำ การควบคุมตัวละครในเกมทำได้โดยการกดปุ่ม w a s d บนคีย์บอร์ดเพื่อควบคุมให้ตัวละครในเกมเคลื่อนที่ไปด้านหน้า ด้านซ้าย ด้านขวา และด้านล่าง ตามลำดับ ทุก 30 วินาที กล่องตัวช่วยจะถูกสุ่มตำแหน่งให้ปรากฏในฉาก กล่องสีเหลืองเป็นตัวช่วยออกเสียงคำศัพท์ ซึ่งสร้างขึ้นจากโปรแกรมสร้างเสียงจากข้อความของกูเกิ้ล (Google Text-to-Speech) กล่องสีน้ำเงินเป็นตัวช่วยแสดงอักขระที่เป็นสระ และกล่องสีแดงเป็นตัวช่วยแสดงอักขระที่เป็นพยัญชนะ กล่องแต่ละสีจะถูกสุ่มด้วยโอกาสที่ไม่เท่ากัน ในสัดส่วน 0.5 0.4 และ 0.1 ตามลำดับ เหตุผลที่กล่องสีแดงมีโอกาสเกิดน้อยที่สุดเพราะเป็นตัวช่วยที่ทำให้ผู้เล่นทายคำศัพท์ได้ง่ายกว่าการใช้ตัวช่วยอื่น ๆ เพราะข้อมูลนำเข้าที่เป็นไปได้เหลือเพียงอักขระที่เป็นสระจำนวนห้าตัว

การคลิกเมาส์ปุ่มซ้ายไปยังวัตถุใด ๆ ในฉากเป็นการเลือกเพื่อเข้าสู่หน้าจอสะกดคำศัพท์ภาษาอังกฤษ ตัวอย่างเช่นเมื่อผู้เล่นคลิกที่วัตถุ “หนังสือ” หน้าจอสะกดคำจะแสดงดังภาพที่ 7 ซึ่งในภาพดังกล่าวบริเวณมุมจอด้านขวาใช้เพื่อออกจากหน้าจอสะกดคำ บริเวณด้านบนตำแหน่งกึ่งกลางของหน้าจอแสดงรูปหนังสือ เพื่อบอกให้ผู้ใช้ทราบว่าผู้ใช้ได้คลิกวัตถุ “หนังสือ” ก่อนผ่านมายังหน้าจอนี้ บริเวณด้านซ้ายของจอภาพเป็นปุ่มที่ให้ผู้ใช้เลือกกดเพื่อใช้งานตัวช่วยทั้งสามกล่องข้อความบริเวณกลางค่อนมาทางล่างจอภาพสี่กล่องใช้เพื่อให้ผู้ใช้พิมพ์อักขระ โดยแต่ละกล่องสามารถพิมพ์อักขระได้เพียงหนึ่งอักขระเท่านั้น เหตุผลที่จำนวนกล่องข้อความมีสี่กล่องคือ เท่ากับจำนวนอักขระคำตอบที่ถูกต้องนั่นคือ “BOOK” ซึ่งแปลว่า “หนังสือ” ปุ่มด้านล่างกล่องข้อความใช้เพื่อตรวจคำตอบว่าถูกต้องหรือไม่

การเล่นเกมจะจบลงในสองกรณี กรณีแรกคือเมื่อเวลาห้านาทีหมดลง ซึ่งเกมจะแสดงหน้าจอที่บอกผู้เล่นว่าไม่ผ่านด่าน และกรณีที่สองคือผู้เล่นสะกดคำศัพท์ได้ครบสิบคำ ซึ่งเกมจะแสดงหน้าจอที่บอกผู้เล่นว่าผ่านด่าน



ภาพที่ 7 หน้าต่างสะกดคำศัพท์ “หนังสือ”

4.2 ผลสัมฤทธิ์ทางการเรียนและระดับความพึงพอใจของผู้ใช้งาน

งานวิจัยชิ้นนี้ได้วัดประสิทธิภาพของเกมที่พัฒนาขึ้นด้วยการวัดผลสัมฤทธิ์ทางการเรียนของผู้เล่นเกมด้วยการทดสอบก่อนเรียนและหลังเรียน ผลที่ได้แสดงในตารางที่ 1

ตารางที่ 1: ผลสัมฤทธิ์ทางการเรียนของผู้ใช้งาน

แบบทดสอบ	คะแนนเต็ม	$\bar{X}$	SD	*t*	*p*
ก่อนเรียน	40	18.00	4.855	15.629	.000*
หลังเรียน	40	35.80	1.935		

$p^* \leq 0.05$

ตารางที่ 1 แสดงให้เห็นว่าผู้ใช้งานมีผลสัมฤทธิ์ทางการเรียนที่สูงขึ้นอย่างมีนัยสำคัญทางสถิติที่ 0.05 โดยได้รับคะแนนสอบเฉลี่ยเมื่อทดสอบหลังเรียน 35.80 คะแนน ซึ่งมากกว่าคะแนนสอบเฉลี่ยเมื่อทดสอบก่อนเรียน 18.00 คะแนน นอกจากนี้งานวิจัยชิ้นนี้ยังได้ศึกษาความพึงพอใจ

ของผู้ใช้งานในกลุ่มดังกล่าว โดยผลระดับความพึงพอใจที่ได้แสดงในตารางที่ 2 ที่แสดงให้เห็นว่าผลความพึงพอใจเฉลี่ยของผู้ใช้งานอยู่ในระดับดีมาก โดยมีค่าเฉลี่ย 4.57 และส่วนเบี่ยงเบนมาตรฐาน 0.66 ตามลำดับ

ตารางที่ 2: ความพึงพอใจของผู้ใช้งาน

รายการประเมิน	ความพึงพอใจ		
	$\bar{x}$	SD	ระดับ
ด้านการออกแบบเกม			
1. ความสนุกสนาน	4.47	0.64	ดี
2. ความสวยงาม	4.60	0.74	ดีมาก
3. ความยากง่ายของเกม	4.47	0.64	ดี
4. การสอดแทรกความรู้	4.80	0.56	ดีมาก
5. ความเหมาะสมกับวัย	4.53	0.64	ดีมาก
ค่าเฉลี่ยด้านเกม	4.57	0.64	ดีมาก
ด้านความเหมาะสมของเนื้อหาภายในเกม			
1. ความรู้ที่ได้รับ	4.53	0.64	ดีมาก
2. ความถูกต้องของเนื้อหา	4.53	0.64	ดีมาก
3. ความเหมาะสมของคำศัพท์	4.67	0.62	ดีมาก
4. ความยากง่ายของคำศัพท์	4.40	0.83	ดี
5. ความสอดคล้องของคำศัพท์กับห้องภายในบ้าน	4.67	0.62	ดีมาก
ค่าเฉลี่ยด้านความรู้	4.56	0.67	ดีมาก
ค่าเฉลี่ยรวม	4.57	0.66	ดีมาก

5. สรุป

การรู้ภาษาที่สองที่เป็นภาษาที่เป็นที่ยอมรับในสากลนับเป็นทักษะสำคัญในการดำรงชีวิตในศตวรรษที่ 21 และเนื่องจากความสามารถในการใช้ภาษาอังกฤษของคนไทยยังต้องพัฒนาให้ดียิ่งขึ้น งานวิจัยชิ้นนี้จึงพัฒนาเกมจำลองชนิดสามมิติเพื่อพัฒนาความรู้คำศัพท์ภาษาอังกฤษเกี่ยวกับของใช้ภายในบ้าน สำหรับนักเรียนระดับประถมศึกษาปีที่ 3 ผู้เล่นจะได้รับบทบาทจำลองเพื่อเดินสำรวจสิ่งของภายในห้องระหว่างเล่นเกม ผู้เล่นจะได้เรียนรู้การสะกดคำศัพท์ไปพร้อมกัน นอกจากนี้ภายในเกมยังมีตัวช่วยสนับสนุนการเล่นเกมในกรณีที่ผู้เล่นไม่สามารถสะกดคำศัพท์ได้ถูกต้อง ได้แก่ ตัวช่วยเสียง ตัวช่วยแสดงอักขระที่เป็นสระ และตัวช่วยที่แสดงอักขระที่เป็นพยัญชนะ ผลการทดลองพบว่า ผู้เล่นที่ได้เล่นเกมที่พัฒนาขึ้นมีผลสัมฤทธิ์ในการเรียนดีกว่าก่อนเล่นเกมอย่างมีนัยสำคัญ และมีระดับความพึงพอใจในระดับดีมาก อย่างไรก็ตาม เกมดังกล่าวยังมีข้อจำกัดในการใช้งานบางประการ นั่นคือคอมพิวเตอร์ที่ใช้งานควรมีประสิทธิภาพในระดับกลาง เพื่อลดปัญหาการกระตุกขณะเล่นเกม

เอกสารอ้างอิง

[1] ef.co.uk. “English and Technology”. Internet: http://www.ef.co.uk/epi/insights/english-and-technology/. [Jan. 14, 2017].

[2] spectrum.ieee.org. “The 2016 Top Programming Languages”. Internet: http://spectrum.ieee.org/computing/software/the-2016-top-programming-languages. [Feb. 15, 2017].

[3] ef.co.uk. “The World's Largest Ranking of Countries by English Skills”. Internet: http://www.ef.co.uk/epi/. [Jan. 15, 2017].

[4] Kuhl, P. “Early Language Acquisition: Cracking the Speech Code”, *Nature Reviews Neuroscience*. vol. 18, no. 5, pp. 831-843, 2017.

[5] Tobias, S., Fletcher, J.D. and Wind, A.P. *Handbook of Research on Educational Communications and Technology.* Springer New York: New York, 2013, pp. 485-503.

[6] Frasson, C. and Blanchard, E.G. *Encyclopedia of the Sciences of Learning.* Springer US: Boston, 2012, pp. 3076-3080.

[7] สุวิชัย พรรษา และอภินันท์ รังสูงเนิน, “การพัฒนาเกมคอมพิวเตอร์ 3 มิติ เกมโซล่าผจญภัย” *การประชุมทางวิชาการระดับชาติ ด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศ ครั้งที่ 4 (NCCIT 2008)*, มหาสารคาม, 23-24 พฤษภาคม 2551.

[8] Gulec, U. and Yilmaz, M. “A Serious Game for Improving the Decision Making Skills and Knowledge of Turkish Football Referees According to the Laws of the Game”, *Springer Plus*. vol. 5, no. 1, pp. 622, 2016.

[9] Amoia, M. “A Serious Game for Second Language Acquisition”, *Proceedings of the 3rd International Conference on Computer Supported Education,* Netherlands, 2011.

[10] Rankin, Y., Gold, R. and Gooch, B. “3D Role - Playing Games as Language Learning Tools”, *Proceedings of Euro Graphics,* USA, 2006.

[11] อาจิน มารีประสิทธิ์. *Active English Primer New Edition book 3*. สำนักพิมพ์ศูนย์ส่งเสริมวิชาการ :กรุงเทพฯ, 2559.

[12] Matlock-Hetzel, S. “Basic Concepts in Item and Test Analysis” *The annual meeting of the Southwest Educational Research Association*, Austin, January 1997.

ระบบคัดเลือกข้อสอบสำหรับผู้เรียนที่มีความแตกต่างกันแบบอัตโนมัติ
Automatic Exam Analysis for the Different Learner System

สุทธิกานต์ บ่อจักรพันธ์ (Sutthikarn Bojukrapan)[1] และปิยสุดา ตันเลิศ (Piyasuda Tanloet)[2]
สาขาวิชาวิทยาการคอมพิวเตอร์และเทคโนโลยีสารสนเทศ คณะวิทยาศาสตร์ มหาวิทยาลัยราชภัฏอุดรธานี
[1]sutthikarn@udru.ac.th, [2]piyasuda@udru.ac.th

บทคัดย่อ

งานวิจัยนี้เป็นการพัฒนาระบบคัดเลือกข้อสอบสำหรับผู้เรียนที่มีความรู้แตกต่างกันแบบอัตโนมัติ ระบบงานนี้ทำงานบนเครือข่ายอินเทอร์เน็ต จัดทำเป็นคลังข้อสอบที่มีการวิเคราะห์หาคุณภาพของข้อสอบเพื่อสุ่มเลือกข้อสอบให้กับผู้เรียนที่มีความรู้พื้นฐานแตกต่างกัน เครื่องมือในการพัฒนาประกอบด้วย ภาษา HTML5 และ PHP ฐานข้อมูลใน MySQL แบบทดสอบมี 3 วิชาได้แก่ ภาษาไทย ภาษาอังกฤษ และคอมพิวเตอร์ วิชาละ 50 ข้อรวม 150 ข้อ สุ่มข้อสอบโดยใช้ Cuckoo Search Algorithm วิชาละ 10 ข้อ รวม 30 ข้อ ข้อสอบมีค่าความยากง่ายระหว่าง 0.20-0.80

ผลการศึกษาพบว่า ผู้เรียน ในสาขาวิชาภาษาไทยจะได้ข้อสอบวิชาภาษาไทยที่มีความยาก ง่ายระหว่าง 0.20-0.45 ผู้เรียนในสาขาวิชาภาษาอังกฤษ ได้ ข้อสอบวิชาภาษาอังกฤษที่มีความยากง่ายระหว่าง 0.22-0.43 และผู้ที่เรียนในสาขาวิชาคอมพิวเตอร์จะได้ข้อสอบวิชา คอมพิวเตอร์ที่มีความยากง่ายระหว่าง 0.22-0.40 และทำการวัดความพึงพอใจการทำงานของระบบอยู่ในระดับดี ค่าเฉลี่ยที่ 4.20 ส่วนเบี่ยงเบนมาตรฐาน 0.61

คำสำคัญ: การวิเคราะห์ข้อสอบ,คุณภาพข้อสอบ,ความรู้แตกต่างกัน, อัลกอริทึมการค้นหา

Abstract

This research aims to develop an automatic system of choosing a test for a learner who has an exactly different knowledge. The proposed system can work via the Internet network and has the corpus of test for analysis of test's quality in order to randomly select the test sets used for the learners who have different knowledge. The tools used to develop the proposed system are composed of script language including HTML5 as well as PHP and MySQL served as the database system. The set of test has three subjects: Thai, English and Computer subjects. Each subject has 50 items and total of all subjects are 150 items. Each subject of the test set is randomized by using the Cuckoo search algorithm for 10 items and the total of three subjects has 30 items. The test set has the level of difficulty between 0.20 and 0.80.

The results of study found that the learners of Thai, English and Computer subjects have the level of difficulty in the test set with 0.20-0.45, 0.22-0.43, and 0.22-0.40, respectively and satisfaction result of system usage is good level with the average and standard deviation of 4.20 and 0.61, respectively.

Keyword: Item Analysis, Qualifying Items, Different Knowledge, Searching Algorithm.

1. บทนำ

ปัจจุบันเทคโนโลยีสารสนเทศ ได้เข้ามามีบทบาทในการดำเนินงานด้านต่างๆ มากยิ่งขึ้น การให้บริการด้านการศึกษาในปัจจุบันก็จำเป็นที่จะต้องให้เทคโนโลยี สารสนเทศเพื่ออำนวยความสะดวก การทำงานของระบบบนเครือข่ายอินเทอร์เน็ตจะอำนวยความสะดวกในการเพิ่ม แก้ไข และลบข้อมูลผ่านเครือข่าย ซึ่งสามารถลดต้นทุนการผลิตได้ ทั้งในแง่เวลาและทรัพยากร อาทิเช่น กระดาษ ค่าจัดพิมพ์ แรงงาน หรือแม้แต่พื้นที่จัดเก็บ นอกจากนี้ยังสามารถสรุปผลคะแนนการสอบของแต่ละวิชา ออกมาได้โดยอัตโนมัติ ซึ่งทำให้ลดขั้นตอนการทำงานลงได้ ทำให้เกิดความประหยัด รวดเร็ว และสะดวก

ในการออกข้อสอบ จัดเก็บข้อสอบ และวัดผลการสอบได้มีประสิทธิภาพมากยิ่งขึ้น [1] ได้มีผู้พัฒนาระบบคลังข้อสอบแบบอัตโนมัติไว้ อาทิเช่น

ดำรงค์ สุพล [3] ได้ศึกษาวิจัยเรื่องระบบคลังข้อสอบออนไลน์ โดยใช้วิธีการวิเคราะห์ข้อสอบรายข้อ ผลการวิจัยพบว่า การประเมินคุณภาพระบบโดยผู้เชี่ยวชาญ จำนวน 5 คน ซึ่งผลการประเมินคุณภาพระบบโดยรวมอยู่ในระดับดีมาก ($\bar{X}$ = 4.80, SD = 0.40) และการประเมินความพึงพอใจของผู้ใช้งาน จำนวน 30 คน ที่มีต่อระบบที่พัฒนาขึ้น ซึ่งผลการประเมินความพึงพอใจต่อระบบโดยรวมอยู่ในระดับดี

อรวรรณ สุวรรณรัตน์ [4] ได้ศึกษางานวิจัยเรื่องการหาคุณภาพข้อสอบเพื่อบริหารจัดการด้านการเรียนการสอน เป็นงานรูปแบบการวิจัยงานวิจัยเรื่องการหาคุณภาพข้อสอบ ข้อสอบปรนัยปีการศึกษา 2549 สำหรับนักศึกษาแพทย์ชั้นปีที่ 4 จำนวน 3 ชุด ชุดละ 40 ข้อจำนวน 5 ตัวเลือก ทำการวิเคราะห์สถิติที่ใช้ในคือ ค่าเฉลี่ย ค่าความแปรปรวน ค่าสหสัมพันธ์ ค่าความคลาดเคลื่อนมาตรฐานในการวัด ค่าความเชื่อมั่น ค่าความยากง่าย ค่าอำนาจจำแนกมาคำนวณ

ประดิษฐ์ ทรเทพ [2] ได้ศึกษาวิจัยเรื่องการพัฒนาระบบคลังข้อสอบออนไลน์ด้วยเทคนิคในการวิเคราะห์ตามแนวคิดอิงเกณฑ์ สำหรับมหาวิทยาลัยฟาร์อีสเทอร์น จากการทดสอบโดยกำหนดให้มีผู้ดูแลระบบ 1 คน อาจารย์ 1 คน นักศึกษา 15 คน จากนั้นจึงรวบรวมข้อมูลจากนักศึกษาและผู้ใช้งาน โดยใช้แบบสอบถามประเมินระดับความพึงพอใจ ซึ่งแบ่งเป็นส่วนการใช้งานทั่วไปและส่วนการใช้งานโปรแกรมย่อยต่าง ๆ ซึ่งผลจากแบบสอบถามพบว่า นักศึกษามีระดับความพอใจมาก ผู้ใช้งานที่เป็นอาจารย์ประจำวิชามีระดับความพึงพอใจค่อนข้าง มาก และผู้ใช้งานที่เป็นผู้ดูแลระบบมีระดับความพึงพอใจมาก

ระบบการสอบ exit-exam สำหรับนักศึกษามหาวิทยาลัยราชภัฏอุดรธานี นักศึกษาทุกสาขาวิชาจะต้องทำการทดสอบวิชา ภาษาไทย ภาษาอังกฤษ และคอมพิวเตอร์ แบบทดสอบไม่มีการแยกข้อสอบให้นักศึกษาใน 3 สาขาวิชานี้ทำให้ทั้ง 3 สาขาส่วนใหญ่จะผ่านในวิชาที่ตนเองเรียนกล่าวคือนักศึกษาสาขาวิชาคอมพิวเตอร์มักจะทำข้อสอบวิชาคอมพิวเตอร์ได้ระดับคะแนนที่ดี แต่จะสอบไม่ผ่านวิชาภาษาไทยและภาษาอังกฤษ หรือนักศึกษาที่เรียนสาขาวิชาอื่น มักจะทำข้อสอบทั้ง 3 วิชานี้ไม่ได้ทำให้เกิดความแตกต่างในการสอบ

ระบบจัดการข้อสอบสำหรับผู้เรียนที่มีความแตกต่างกันเป็นอีกระบบหนึ่งที่ประยุกต์เอาเทคโนโลยีสารสนเทศมาใช้ อีกทั้งยังเป็นวิธีการในการพัฒนาตัวข้อสอบให้ได้มาตรฐาน จัดเก็บในรูปแบบของฐานข้อมูล แบ่งหมวดหมู่ของแต่ละวิชา แต่ละจุดประสงค์ เป็นการมุ่งพัฒนาเพื่อเพิ่มประสิทธิภาพในการจัดทำเว็บและเข้าถึงฐานข้อมูล ด้วยเหตุนี้ผู้วิจัยได้เห็นถึงความสำคัญในการพัฒนาระบบจัดการข้อสอบสำหรับผู้เรียนที่มีความแตกต่างกันแบบอัตโนมัติ โปรแกรมช่วยอำนวยความสะดวก ลดขั้นตอน เพิ่มประสิทธิภาพในการจัดทำข้อสอบ การวิเคราะห์หาคุณภาพข้อสอบ การจัดสอบ การตรวจทานข้อสอบ การคัดข้อคำถามที่มีคุณภาพเก็บไว้ในคลังข้อสอบเพื่อนำไปใช้ในการออกข้อสอบในครั้งต่อไป ครูผู้สอนสามารถนำผลการสอบไปใช้ในการปรับปรุง และพัฒนาการจัดการเรียนการสอน และระบบนี้เป็นการทำงานบนเครือข่ายอินเทอร์เน็ต จะช่วยอำนวยความสะดวกในการเข้าถึงอย่างรวดเร็ว ทำให้สามารถใช้งานระบบนี้ได้จากทุกที่ ทุกเวลา

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 การวิเคราะห์คุณภาพของแบบทดสอบวัดผลสัมฤทธิ์ ในการวิเคราะห์ข้อสอบ วิเคราะห์ได้ 2 ลักษณะ คือ วิเคราะห์ข้อสอบเป็นรายข้อและวิเคราะห์ข้อสอบทั้งฉบับ วิธีการวิเคราะห์สามารถทำได้หลายแนวทางขึ้นกับทฤษฎีการวัดผลที่นำมาใช้ ทฤษฎีการวัดผลที่นำมาใช้โดยทั่วไปในการวิเคราะห์ข้อสอบ ได้แก่ ทฤษฎีการทดสอบแบบคลาสสิก หรือเรียกว่าทฤษฎีการทดสอบแบบดั้งเดิม และอีกทฤษฎีหนึ่งก็คือทฤษฎีการตอบสนองข้อสอบ [3]

1) ค่าความยากง่าย คือ สัดส่วนหรือร้อยละของผู้สอบที่ตอบข้อสอบนั้นถูก คำนวณได้จากสมการ (1) [1]

$$P = \frac{R}{N} \qquad (1)$$

เมื่อกำหนดให้ P คือค่าความยากง่าย R คือจำนวนผู้เรียนที่ตอบคำถามข้อนั้นถูกต้อง และ N คือจำนวนผู้เรียนทั้งหมด เกณฑ์การแปลความหมายความยากง่ายของข้อสอบดังแสดงได้ในตารางที่ 1[8]

ตารางที่ 1: เกณฑ์การแปลความหมายของความยากของข้อสอบ

ค่าความยาก(p)	ความหมาย
0.91 - 1.00	ง่ายมาก
0.81 - 0.90	ง่าย
0.61 - 0.80	ค่อนข้างง่าย
0.40 – 0.60	ยากง่ายพอเหมาะ
0.20 - 0.39	ค่อนข้างยาก
0.10 - 0.19	ยาก
0.01 - 0.09	ยากมาก

2) อำนาจจำแนกข้อสอบ คือ ความสามารถของข้อสอบที่จะจำแนกความแตกต่างของสิ่งที่ต้องการวัด โดยสามารถจำแนกกลุ่มผู้สอบที่ได้คะแนนรวมสูง ออกจากกลุ่มผู้ที่ได้คะแนนรวมต่ำ การคำนวณค่าอำนาจจำแนก คำนวณได้โดยใช้สูตรอย่างง่าย (r)

$$r = \frac{H - L}{N_H} \qquad (2)$$

เมื่อกำหนดให้ H คือ จำนวนผู้สอบในกลุ่มที่ได้คะแนนรวมสูงที่ตอบตัวเลือกนั้น N คือ จำนวนผู้สอบในกลุ่มที่ได้คะแนนรวมต่ำที่ตอบตัวเลือกนั้น N_H คือ จำนวนคนในกลุ่มที่ได้คะแนนรวมสูง N_L คือ จำนวนคนในกลุ่มที่ได้คะแนนรวมต่ำ

สูตรสัมประสิทธิสหสัมพันธ์แบบพอยท์ไบซีเรียล (Point biserial correlation coefficient หรือ r_{pb})

$$r_{pb} = \frac{\overline{X_p} - \bar{X}_q}{S_x} \qquad (3)$$

เมื่อกำหนดให้ X_p คือ ค่าเฉลี่ยของคะแนนของผู้สอบที่ตอบข้อสอบข้อนั้นถูก X_q คือ ค่าเฉลี่ยของคะแนนของผู้สอบที่ตอบข้อสอบข้อนั้นผิด S_x คือ ค่าเบี่ยงเบนมาตรฐานของคะแนนรวมของผู้สอบทั้งหมด P คือ สัดส่วนของผู้สอบที่ตอบข้อสอบนั้นถูก และ q คือ สัดส่วนของผู้สอบที่ตอบข้อสอบนั้นผิด

4) ความเที่ยงหรือความเชื่อมั่น หมายถึง ความคงที่ของคะแนนที่วัดได้แต่ละครั้งคำนวณได้จากสูตร

$$r = \frac{n}{n-1}\left[1 - \frac{\sum pq}{s^2}\right] \qquad (4)$$

เมื่อกำหนดให้ n คือ จำนวนข้อของแบบทดสอบ S^2 คือ ความแปรปรวนของคะแนนรวม p คือ สัดส่วนของผู้ตอบข้อสอบถูกในแต่ละข้อ q คือ สัดส่วนของผู้ตอบข้อสอบผิดในแต่ละข้อ (1-p)

2.2 อัลกอริทึมการค้นหาแบบนกดุเหว่า

Cuckoo Search Algorithm อัลกอริทึมการค้นหาแบบนกดุเหว่า [7] เป็นวิธีการค้นหาค่าที่เหมาะสมที่สุด โดยมีแรงบันดาลใจจากพฤติกรรมการวางไข่ของนกดุเหว่าซึ่งการวางไข่ลงไปในรังจะถูกสุ่มเลือกรังว่า มีคุณภาพสูงหรือไม่ ถ้าเป็นรังที่มีคุณภาพสูงก็ทำการวางไข่แล้วเลือกให้เป็นรุ่นถัดไป และในขั้นตอนการบินหารังใหม่นั้นจะใช้หลักการของ Lévy flight ที่มีการเคลื่อนที่ไปได้ไกลกว่าปกติ รหัสเทียมของอัลกอริทึมการค้นหาแบบนกดุเหว่าซึ่งมีขั้นตอนหลักดังนี้

```
Begin  Objective function  f(x).x= ( x……x_d )^T ;
  Generate initial population of n host nests
  X_i ( I = 1,2,…,n );
While (t < MaxGeneration) or (stop criterion);
  Get a Cuckoo (say, i ) randomly by Levy flight;
  Evaluate its quality/fitness  F_j ;
  Choose a nest among n (say j ) randomly ;
if  (F_i > F_j ) ,
  Replace  j  by the new solution;
end
  Abandon a fraction  (pa) of worse nests
  [and build new ones at new locations via Levy
  flight];
   Keep the best solutions (or nests with quality
   solution );
  Rank the solutions and find the current best;
end  while
  Post process results and visualization;
End
```

ภาพที่ 1: รหัสเทียมอัลกอริทึมการค้นหาแบบนกกาเหว่า [9]

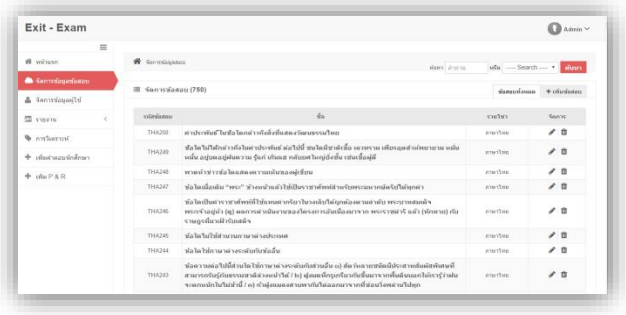
การเคลื่อนที่แบบ Lévy flights สามารถเคลื่อนที่ได้ไกลกว่าปกติ ดังนั้น Lévy flights จึงนิยมนำมาใช้แก้ปัญหา Optimization เพื่อหา Global Optima ให้ได้ผลลัพธ์ของคำตอบดีที่สุด

ขั้นตอนวิธีการค้นหาแบบนกดุเหว่าสามารถนำมาใช้ประมวลผลในการคัดเลือกข้อสอบสำหรับชุดข้อสอบที่มีค่าความยากง่ายอยู่ในเกณฑ์ที่ตั้งไว้คือ High level (0.20-0.40) และเกณฑ์ Medium level (0.41-0.80) [7]

3. วิธีดำเนินการวิจัย

3.1 ขั้นตอนการดำเนินงาน

1) คัดเลือกข้อสอบสามวิชา คือ ภาษาไทย ภาษาอังกฤษ คอมพิวเตอร์ วิชาละ 50 ข้อ รวม 150 ข้อนำเข้าสู่ระบบ มีระดับความยากง่ายตั้งแต่ 0.20-0.90



ภาพที่ 2: การจัดการคลังข้อสอบ

2) ทำการวิเคราะห์ข้อสอบรายข้อจากกลุ่มตัวอย่างจำนวน 30 คน เพื่อหาค่าความยากง่าย อำนาจจำแนกและค่าความเชื่อมั่นของแบบทดสอบ

3) ผู้ดูแลระบบกำหนดสิทธิ์เข้าใช้งาน โดยผู้ดูแลระบบผู้เรียนทำแบบทดสอบออนไลน์โดยแยกเป็นสามสาขาวิชา ได้แก่ ภาษาไทย ภาษาอังกฤษ และคอมพิวเตอร์

4) ทำการสุ่มข้อสอบจากคลังข้อสอบสำหรับการนำเข้า โดยใช้ Cuckoo Search คัดเลือกข้อสอบนำเข้าเพื่อนำมาจัดชุดข้อสอบนำเข้า ดังสมการที่ 5

$$\text{ค่าสูงสุด} \quad P_{sys} = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{N_i} X_{ij}P_{ij} \qquad (5)$$

n แทน จำนวนข้อสอบที่ต้องการทั้งหมด

N_i แทน จำนวนแบบทดสอบที่มีหลายชุดสำหรับให้ข้อสอบข้อที่ i คัดเลือกได้

X_{ij} แทน ค่าตำแหน่ง ข้อสอบข้อที่ i ตำแหน่งที่ j

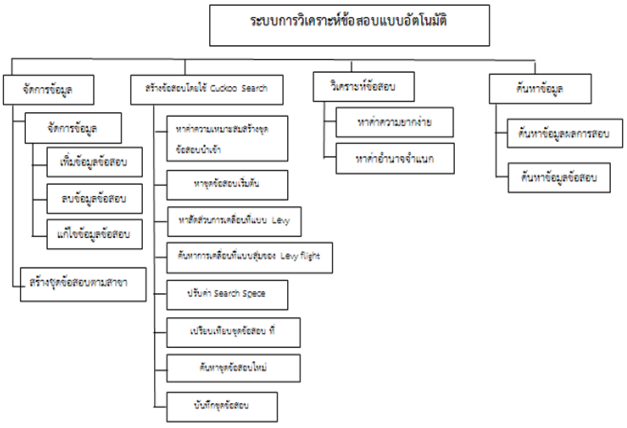
P_{ij} แทน ค่าความยากง่ายของชุดแบบทดสอบ j ที่ข้อสอบข้อที่ i คัดเลือกไปใช้

P_{sys} แทน ค่าความยากง่ายสูงสุดของแบบทดสอบทั้งฉบับในแต่ละรอบ

มีการกำหนดค่าพารามิเตอร์ในการทดลองที่ช่วยทำให้หาผลลัพธ์ที่ดีที่สุด ดังนี้ จำนวนชุดข้อสอบ (n) เท่ากับ 5 ความน่าจะเป็นในการค้นพบ ค่า (pa) เท่ากับ 0.70 รอบการค้นหา เท่ากับ 50 รอบ α เท่ากับ 0.01 และ β เท่ากับ 1.5 ในการกำหนดฟังก์ชั่นวัตถุประสงค์สำหรับการแก้ปัญหาการหาค่าที่เหมาะสมของค่าอำนาจจำแนกของแบบทดสอบทั้งฉบับ เงื่อนไขของสมการคือ ค่าความยากง่าย P_{ij} ต้องมีค่าอยู่ระหว่าง 0.20- 0.80

5) ผู้เรียนทดลองทำแบบทดสอบโดยเลือกมา 3 วิชาแต่ละวิชาจำนวน 10 ข้อ รวม 30 ข้อ วิเคราะห์และสรุปผลลัพธ์ที่ได้

ระบบจัดการข้อสอบสำหรับผู้เรียนที่มีความรู้แตกต่างกันแบบอัตโนมัติในระบบประกอบด้วยการจัดการข้อมูลข้อสอบ การวิเคราะห์ข้อสอบ การสุ่มข้อสอบตามความแตกต่างของผู้เรียน และการค้นหาข้อมูลเพื่อรายงานผลการสอบได้ ดังแสดงภาพที่ 3



ภาพที่ 3: ภาพรวมการทำงานระบบ

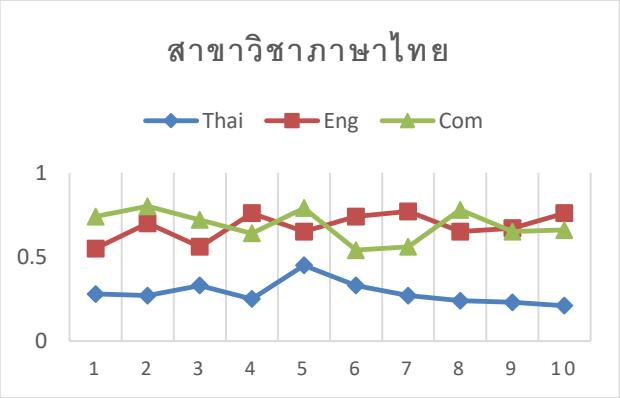
4. ผลการวิจัย

ผลการพัฒนาระบบจัดการข้อสอบสำหรับผู้เรียนที่มีความรู้แตกต่างกันแบบอัตโนมัติ ผู้เรียนจะต้องลงทะเบียนเข้าสู่ระบบดังภาพที่ 4 เพื่อทราบว่าผู้เรียนอยู่ในสาขาวิชาใด



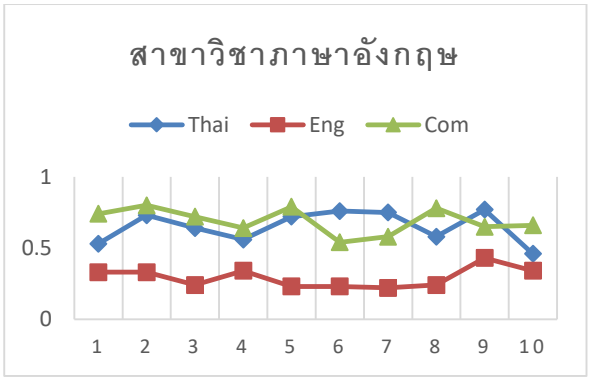
ภาพที่ 4: หน้าหลักเข้าสู่ระบบ

ในการคัดเลือกข้อสอบมี 3 วิชาคือ ภาษาอังกฤษ ภาษาไทย คอมพิวเตอร์ ระบบจะทำการสุ่มข้อสอบขึ้นมาวิชาละ 10 ข้อ รวม 30 ข้อ ลักษณะข้อสอบทุกข้อจะมีค่าความยากง่ายแตกต่างกันขึ้นอยู่กับความรู้พื้นฐานในสาขาวิชา จากภาพที่ 5-7 แกน x คือจำนวนข้อสอบที่สุ่ม แกน y คือค่าความยากง่าย ผู้เรียนที่อยู่ในแต่ละสาขาวิชาจะได้แบบทดสอบที่มีค่าความยากง่ายแตกต่างกัน ผลจากการคัดเลือกข้อสอบที่นำไปใช้กับผู้เรียน 3 สาขาวิชาได้แก่ สาขาวิชาภาษาไทย สาขาวิชาภาษาอังกฤษ และสาขาวิชาคอมพิวเตอร์ ผู้เรียนที่อยู่สาขาวิชาภาษาไทยสอบวิชาภาษาไทยจะได้ข้อสอบที่มีความยากมากว่าสาขาวิชาภาษาอังกฤษและสาขาวิชาคอมพิวเตอร์ สาขาวิชาภาษาอังกฤษได้ข้อสอบวิชาภาษาอังกฤษยากกว่าสาขาวิชาภาษาไทยและสาขาวิชาคอมพิวเตอร์ และสาขาวิชาคอมพิวเตอร์จะได้วิชาคอมพิวเตอร์ยากกว่าสาขาวิชาภาษาไทยและสาขาวิชาภาษาอังกฤษ โดยได้ข้อมูลข้อสอบแยกตามระดับความยากง่ายตามสาขาวิชาแบ่งออกเป็นสองระดับคือ ระดับยากจะอยู่ในช่วง (0.10 - 0.40) และระดับปานกลางถึงง่ายจะอยู่ในช่วง (0.60 - 0.90) จากผลการทดลองจะพบว่าข้อสอบแต่ละชุดของแต่ละสาขาวิชาจะมีค่าความเหมาะสมที่พอเหมาะและมีค่าความยากง่ายที่ใกล้เคียงกัน แสดงดังภาพที่ 5-7 ดังนี้



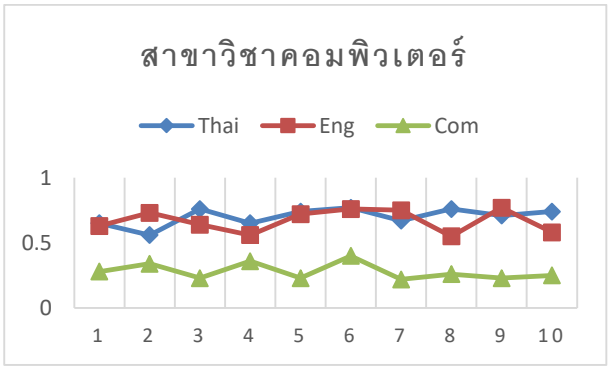
ภาพที่ 5: ผลการคัดเลือกวิชาภาษาไทย

จากภาพที่ 5 สาขาวิชาภาษาไทยจะได้แบบทดสอบวิชาภาษาไทยที่มีค่าความยากง่ายระหว่าง 0.21-0.45 วิชาภาษาอังกฤษมีค่าความยากง่ายระหว่าง 0.55-0.77 และวิชาคอมพิวเตอร์มีค่าความยากง่ายระหว่าง 0.54-0.80



ภาพที่ 6: ผลการคัดเลือกวิชาภาษาอังกฤษ

จากภาพที่ 6 ผู้เรียนในสาขาวิชาภาษาอังกฤษ ได้ข้อสอบวิชาภาษาอังกฤษที่มีความยากง่ายระหว่าง 0.22-0.43 ได้แบบทดสอบวิชาภาษาไทยมีความยากง่ายระหว่าง 0.46-0.77 และวิชาคอมพิวเตอร์มีค่าความยากง่ายระหว่าง 0.54-0.88



ภาพที่ 7: ผลการคัดเลือกวิชาคอมพิวเตอร์

จากภาพที่ 7 ผู้ที่เรียนในสาขาวิชาคอมพิวเตอร์จะได้ข้อสอบวิชาคอมพิวเตอร์ที่มีความยากง่ายระหว่าง 0.22-0.40 ได้แบบทดสอบวิชาภาษาไทยที่มีค่าความยากง่ายระหว่าง 0.56-0.77 วิชาภาษาอังกฤษมีค่าความยากง่ายระหว่าง 0.55-0.77

ตารางที่ 4.2 สรุปผลการประเมินประสิทธิภาพของระบบจากผู้ใช้

รายการประเมิน	$\bar{X}$	S.D	แปลผล
1. ความสวยงามของระบบ	4.60	0.50	ดีมาก
2. ความยาก-ง่ายการใช้งานของระบบ	4.40	0.56	ดี
3. ความถูกต้องในการทำงานของระบบ	4.50	0.67	ดี
4. ความถูกต้องสมบูรณ์ของรายงาน	3.90	0.68	ปานกลาง
5. ความถูกต้องของระบบ	4.00	0.71	ดี
6. ระดับความยาก-ง่ายของชุดข้อสอบ	3.80	0.50	ปานกลาง
ผลรวม	4.20	0.61	ดี

จากตารางที่ 4.2 ทำการประเมินความพึงพอใจของผู้ใช้งานด้านความถูกต้องของระบบและประเมินประสิทธิภาพของระบบจากผู้ใช้งานจำนวน 30 คน โดยสามารถสรุปผลการประเมินความพึงพอใจอยู่ในระดับดี มีค่าเฉลี่ยเท่ากับ 4.20 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.61

5. สรุปผลการดำเนินงาน

ระบบการจัดการข้อสอบสำหรับผู้เรียนที่มีความแตกต่างกันแบบอัตโนมัตินี้ได้จัดทำเป็นคลังข้อสอบสำหรับสอบจบตามเกณฑ์การศึกษาในมหาวิทยาลัยราชภัฏอุดรธานี สามารถวิเคราะห์ข้อสอบอัตโนมัติและคัดเลือกข้อสอบโดยใช้อัลกอริทึมการค้นหาแบบนกคุเหว่า เพื่อใช้ในการตัดสินใจคัดเลือกข้อสอบตามระดับ ค่าความยากง่ายของข้อสอบแต่ล่ะข้อ ผลการวิจัย แสดงให้เห็นว่าผู้เข้าสอบในแต่ล่ะสาขาวิชาได้รับข้อสอบที่มีความแตกต่างกันตามความยากง่ายของสาขาวิชา ทำให้มีความเท่าเทียมกันในการสอบ สามารถนำมาใช้ทดสอบประมวลผลในการคัดเลือกชุดข้อสอบที่มีค่าความยากง่ายที่เหมาะสมในแต่ละระดับและเป็นอีกหนึ่งแนวทางให้ผู้สอนสามารถนำไปประยุกต์ใช้กับหาคุณภาพของแบบทดสอบและนำไปใช้กับผู้เรียนที่มีความแตกต่างกันได้

เอกสารอ้างอิง

[1] มนต์ชัย เทียนทอง, “การหาคุณภาพแบบทดสอบของบทเรียนคอมพิวเตอร์ช่วยสอน,” ใน การออกแบบและพัฒนาคอร์สแวร์สำหรับบทเรียนคอมพิวเตอร์ช่วยสอน, กรุงเทพฯ: ศูนย์ผลิตตำราเรียนสถาบันเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2545 .

[2] ประดิษฐ์ ทรเทพ. การพัฒนาระบบคลังข้อสอบออนไลน์ด้วยเทคนิคในการวิเคราะห์ตามแนวคิดอิงเกณฑ์สำหรับ มหาวิทยาลัยฟาร์อีสเทอร์น. วิทยาการคอมพิวเตอร์ มหาวิทยาลัยเชียงใหม่, 2554.

[3] ดำรงค์ สุพล , ระบบคลังข้อสอบออนไลน์โดยใช้วิธีการวิเคราะห์ข้อสอบรายข้อ, เทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ , มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2554.

[4] อรวรรณ สุวรรณรัตน์, “ข้อสอบเพื่อบริหารจัดการด้านการเรียนการสอน,” สงขลานครินทร์เวชสาร ปีที่ 27 ฉบับที่ 5. หน้า 382 - 388.

[5] กิตติพงษ์ จรัญศิริไพศาล สิรภัทร เชี่ยวชาญวัฒนา และคำรณ สุนัติ “ เพิ่มความถูกต้องของตัวแบบ ซัพพอร์ตเวกเตอร์แมทชีนแบบค่ากำลังสองน้อยที่สุดด้วยขั้นตอนวิธีการค้นหาแบบนกคุเหว่า” ในระดับบัณฑิต ครั้งที่ 13 ขอนแก่น: บัณฑิตวิทยาลัย มหาวิทยาลัยขอนแก่น ,2555.

[6] วิรวรรณ์ สืบสายทองคำ. การพัฒนาระบบการจัดการและวิเคราะห์ข้อสอบผ่านเครือข่ายอินเทอร์เน็ต. เทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2549.

[7] สุทธิกานต์ บ่อจักรพันธ์ ,มานะ โสภา ,ขนิษฐา ทะคะลิง และมนัสนันท์ กองสุวรรณ์.” ระบบคัดเลือกข้อสอบสำหรับการสอบเพื่อจบการศึกษาแบบออนไลน์ด้วยวิธีค้นหาแบบนกคุเหว่า ”ในการประชุมทางวิชาการระดับชาติด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศ ครั้งที่12 (NCCIT2016) ขอนแก่น: คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยมหาสารคาม2559 , หน้า 380-385.

[8] บุญชม ศรีสะอาด. “เครื่องมือและเทคนิคในการเก็บรวบรวมข้อมูลในการวิจัยเบื้องต้น”. พิมพ์ครั้งที่ 7. กรุงเทพฯ:สุริวิริยาสาส์น, 2545.

[9] X.-S. Yang, and S. Deb, “Engineering Optimisation by Cuckoo Search”. *Int. J. Mathematical Modelling and Numerical Optimization*, Vol. l, No. 4, pp. 330–343, 2010.

การพัฒนาระบบจัดการเรียนรู้ทัศนศึกษานอกสถานที่
The Development of Field Trip Learning Management System

สุชิตา บุญร่วม (Sukita Boonraum)[1] และสมคิด แซ่หลี (Somkid Saelee)[2]
คณะครุศาสตร์อุตสาหกรรม มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
sukita2534@gmail.com, somkid.l@fte.kmutnb.ac.th

บทคัดย่อ

การวิจัยครั้งนี้เป็นการพัฒนาระบบจัดการเรียนรู้ทัศนศึกษานอกสถานที่โดยนำเทคโนโลยี GPS มาประยุกต์ใช้เพื่อติดตามตำแหน่งนักเรียน โดยครูสามารถจัดการเนื้อหาผ่านเฟรมเวิร์คของ Wordpress สามารถปักหมุดสำหรับจุดการทำกิจกรรม การสร้างโซนสำหรับจุดนัดหมายก่อนทำกิจกรรมและตรวจสอบการเข้า-ออกโซน ของนักเรียน นักเรียนลงพื้นที่ทัศนศึกษาเรียนรู้ข้อมูลด้วยตนเองในรูปแบบ (Self Directed Learning) โดยนักเรียนใช้อุปกรณ์สมาร์ทโฟนเป็นสื่อกลางในการเรียนรู้รูปแบบโมบายเลิร์นนิง (Mobile Learning) ในการเข้าถึงเนื้อหาและการทำกิจกรรม ซึ่งรูปแบบของระบบเป็นเว็บแอพพลิเคชั่นแบบเรซสปอนต์ซีฟ (Responsive Design) เพื่อรองรับการทำงานของสมาร์ทโฟนระบบปฏิบัติการ Android และ iOS

การทดลองใช้พื้นที่ในมหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ ใช้กลุ่มตัวอย่างทดลองจำนวน 5 คน ใช้ระยะเวลา 5 วัน และทดลองวันละ 5 เวลา เพื่อทดสอบประสิทธิภาพด้านตำแหน่งโดยแบ่งย่อยออกเป็น 3 ด้านคือ การเช็คอินตำแหน่ง การเช็คอินโซนและการแจ้งเตือนเข้า-ออกโซน ซึ่งผลการวิจัยสรุปออกมาทั้ง 3 ด้าน มีความถูกต้องร้อยละ 87.39% ซึ่งสูงกว่าสมมติฐานที่ตั้งไว้ 80%

คำสำคัญ: ทัศนศึกษานอกสถานที่ โมบายเลิร์นนิง การเรียนรู้ด้วยตนเอง เรซสปอนต์ซีฟ

Abstract

This research was developed the learning management system for field trips that apply GPS technology to track many students, teacher can manage content by using wordpress framework, and create marker on map, zone for appointing and checking student. Students walk as self-directed learning to outdoor exploratory learning in field trip activities with mobile devices for learning and do any activities. The system was designed with responsive web application for support most Android and iOS devices.

This system was proved at King Mongkut's University of Technology North Bangkok by 5 students on 5 days and 5 times per day. To find the efficiency that divided into 3 phases 1) check-in to location, 2) check-in to zone and 3) alert when student walk out from assigned zone. The accuracy from 3 phase as 87.39 percent that higher than hypothesis 80 percent.

Keyword: field trips, mobile-learning, self-directed learning, responsive web

1. บทนำ

การทัศนศึกษาเป็นรูปแบบการเรียนรู้ที่สร้างความกระตือรือร้นและเป็นการค้นพบความรู้ด้วยตนเองจากประสบการณ์ตรง ด้วยการสังเกต การสำรวจ การทดลอง การปฏิบัติจริง และเชื่อมโยงกับความรู้ในห้องเรียน เพราะนักเรียนจะเรียนรู้ได้ดีจากการลงมือปฏิบัติ ทดลอง สำรวจ ค้นคว้า และพบความรู้ด้วยตนเองจากประสบการณ์ที่เขาได้ปฏิบัติโดยใช้ประสาทสัมผัสทั้งห้า [1] แต่ปัญหาการทัศนศึกษาในรูปแบบเดิมส่วนใหญ่จะเป็นการเรียนรู้โดยครูเป็นศูนย์กลางซึ่งทำให้จำนวนครูไม่เพียงพอต่อการดูแลและให้ข้อมูลกับนักเรียน รวมถึงไม่มีแหล่งจัดการข้อมูลเพื่อให้ความรู้แก่นักเรียน จึงส่งผลให้นักเรียนไม่สามารถที่จะกำหนดการเรียนรู้ได้ด้วยตนเองและขาดแรงจูงใจในการเรียนรู้ [2]

จากการศึกษางานวิจัยที่เกี่ยวข้อง Zurita และ Baloian ได้พัฒนาต้นแบบระบบ “แท็บเล็ต” การเรียนรู้ทางภูมิศาสตร์ ครู

สามารถกำหนดคำแนะนำ คำชี้แนะ บนแผนที่ดิจิทัลให้กับนักเรียนของเขา โดยก่อนที่นักเรียนจะเรียนรู้ต้องเปิดฟังก์ชันการทำงานของ GPS ในมือถือ [3] Jongได้พัฒนาและออกแบบ EagleEye เพื่อบูรณาการระบบการศึกษาแบบกลางแจ้ง โดยการใช้ GPS สนับสนุนการเรียนรู้กลางแจ้งด้วยการทำกิจกรรมในช่วงทัศนศึกษาเพื่อแก้ปัญหาการทัศนศึกษาแบบดั้งเดิมด้วยระบบปฏิบัติการ iOS โดยเป้าหมายของการทัศนศึกษาคือส่งเสริมให้นักเรียนสำรวจสถานที่ในเขตเมืองเพื่ออนุรักษ์วัฒนธรรมจีนดั้งเดิมในฮ่องกง ซึ่งผลจากการทดลองปรากฏว่านักเรียนได้รับประสบการณ์ในการทัศนศึกษามากยิ่งขึ้นจากการใช้ EagleEye [2] Jong และ Luk ได้นำ EagleEye ไปศึกษาเชิงปริมาณด้านคุณภาพ กับครูทางด้านภูมิศาสตร์จำนวน 98 คน เพื่อประเมินการทำงานและความสามารถในการสร้างประสบการณ์ให้กับนักเรียนในการทำกิจกรรมระหว่างทัศนศึกษาและการเพิ่มประสิทธิภาพในการทัศนศึกษา ผลปรากฏว่าครูมีการรับรู้ในเชิงบวกต่อ EagleEye ซึ่งสามารถแก้ปัญหาการทัศนศึกษาแบบดั้งเดิมได้ [4]

ดังนั้นผู้วิจัยจึงมีแนวความคิดที่จะพัฒนาระบบการจัดการเรียนรู้ทัศนศึกษานอกสถานที่ เพื่อเข้ามาช่วยส่งเสริมสนับสนุนการทัศนศึกษากลางแจ้ง โดยมีเครื่องมือที่ใช้พัฒนาระบบดังนี้

1) เฟรมเวิร์ค Wordpress นำมาพัฒนาเว็บแอพพลิเคชั่น เพราะเป็นโอเพนซอร์ส (Open Source) และเป็นระบบจัดการเนื้อของเว็บไซต์ ซึ่งผู้วิจัยพัฒนาธีม (Theme) เพื่อใช้สำหรับแสดงเนื้อหาของระบบตามที่ต้องการ [5]

2) ติดตั้งปลั๊กอิน (Plugin) HP5 เพื่อใช้ในการสร้างเนื้อหาที่มีลักษณะการโต้ตอบกับผู้ใช้งาน เป็นปลั๊กอินแบบโอเพนซอร์สที่พัฒนาด้วยภาษา HTML5 การแสดงผลของเนื้อหาสามารถแสดงผลในรูปแบบแบบเรซสปอนต์ซีฟได้ [6]

3) ติดตั้งปลั๊กอิน Buddypress ซึ่งมีรูปแบบเป็นโซเชียลมีเดียที่มีการทำงานคล้ายกับ Facebook [7] โดยฟังก์ชันหลักที่นำมาใช้งานคือ ฟังก์ชันการสร้างกลุ่มสำหรับเพิ่มนักเรียนเข้ามาทำกิจกรรม โดยผู้วิจัยได้พัฒนาฟังก์ชันเพิ่มเข้าไปในกลุ่มคือ ฟังก์ชันการปักหมุด ฟังก์ชันการสร้างโซน ฟังก์ชันการเช็คอินตำแหน่งหมุด ฟังก์ชันการเช็คอินโซน ฟังก์ชันการแจ้งเตือนเข้า-ออกโซนและฟังก์ชันการแสดงผลการเช็คอินตำแหน่งหมุด ผลการเช็คอินโซนและผลการเข้า-ออกโซน โดยฟังก์ชันดังกล่าวจะแสดงต่อเมื่อมีการสร้างกลุ่มโดยผู้ใช้ที่สิทธิ์เป็นครู ฟังก์ชันที่พัฒนาเพิ่มเติม ได้ประยุกต์ใช้เทคโนโลยี GPS (Global Positioning System) สำหรับใช้ระบุตำแหน่งของนักเรียนในระหว่างทำกิจกรรม ซึ่งเป็นระบบนำทางด้วยดาวเทียมทั้งหมด 24 ดวง ที่ให้บริการฟรีและมีความแม่นยำที่เชื่อถือได้เกี่ยวกับข้อมูลทางภูมิศาสตร์ [8] และใช้งานร่วมกับไลบรารี่ (Library) Google Map API ซึ่งเป็น API ที่ให้บริการเซอร์วิสต่างๆเกี่ยวกับแผนที่ [9]

ซึ่งการพัฒนาระบบเพื่อเป็นการทดสอบประสิทธิภาพการทำงานของระบบที่มีความถูกต้อง มากกว่า 80% ตามสมมุติฐานที่ตั้งไว้หรือไม่

2. การดำเนินการวิจัย

การพัฒนาระบบจัดการเรียนรู้ทัศนศึกษานอกสถานที่ มีขั้นตอนดำเนินการวิจัยดังนี้

2.1 ขั้นตอนการพัฒนาระบบ

การพัฒนาระบบใช้เฟรมเวิร์คของ Wordpress ในการจัดการเนื้อหา ติดตั้งปลั๊กอิน H5P , Buddypress และพัฒนาฟังก์ชันเพิ่มเติมในปลั๊กอิน Buddypress โดยใช้ภาษา PHP, HTML, CSS, Javascript และไลบรารี่ (Library) ของ Google Map API เชื่อมต่อฐานข้อมูล MySQL ซึ่งขั้นตอนการพัฒนาระบบแบ่งย่อยเป็นฟังก์ชันดังนี้

2.1.1 การพัฒนาธีม ผู้วิจัยใช้วิธีการพัฒนาธีมของ Wordpress เพื่อให้รูปแบบของธีมเหมาะสมกับเนื้อหาที่ต้องการแสดง พัฒนาด้วยวิธี Responsive Web Design ด้วยเฟรมเวิร์ค Bootstrap [10] เพื่อรองรับการทำงานของสมาร์ทโฟนทุกรุ่น

2.1.2 พัฒนาฟังก์ชันเพิ่มเติมในปลั๊กอิน Buddypress ซึ่งฟังก์ชันที่พัฒนามีดังนี้

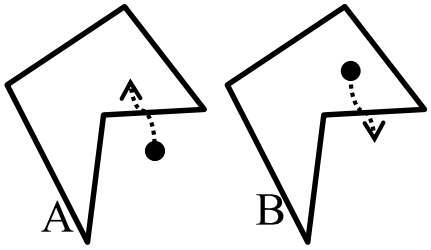
1) ฟังก์ชันการปักหมุด โดยใช้ไลบรารี่ Drawing tools ประเภท Marker [9] ซึ่งกำหนดให้ 1 กลุ่มครูสามารถปักหมุดได้ไม่เกิน 10 จุด

2) ฟังก์ชันเช็คอินตำแหน่ง ซึ่งกำหนดค่ารัศมีที่ห่างจากหมุดไว้ 20 เมตร เมื่อนักเรียนเดินมาในระยะรัศมีของหมุด หมุดจะแจ้งเตือนด้วยวิธีการ Marker Animations ซึ่งเป็นลักษณะการเคลื่อนไหวของหมุดในรูปแบบที่กำหนด [9] ด้วยลักษณะเด้งขึ้น-ลง เพื่อแจ้งเตือนในนักเรียนทราบ สำหรับการเช็คอินตำแหน่ง กรณีกรองตำแหน่งข้อมูล GPS ที่คลาดเคลื่อน ระบบจะใช้วิธีการ Sliding Window ซึ่งเป็นการเก็บชุดข้อมูลตามจำนวนจำกัดและจะมีชุดข้อมูลไหลเข้ามาแทนที่ชุดข้อมูลเก่าเรื่อยๆ [11] โดยนำมาใช้ในรูปแบบการเก็บชุดข้อมูลตำแหน่งละติจูด ลองติจูด ของนักเรียนทั้งหมด 5 ชุดล่าสุด แต่ละชุดห่างกัน 2 วินาที เพื่อนำมาคำนวณหาค่าเฉลี่ยตำแหน่งละติจูด ลองติจูด ให้ออกมาเป็นระยะทาง ด้วยสูตรการหาค่า Vector Distance [12] ซึ่งถ้าระยะทางที่คำนวณออกมามากกว่า 10 เมตร ระบบจะไม่บันทึกข้อมูลการเช็คอินลงฐานข้อมูล จนกว่าสมาร์ทโฟนจะจับสัญญาณ GPS ได้ โดยนักเรียนสามารถกดปุ่มเช็คอินตำแหน่งได้จุดละ 1 ครั้งเท่านั้น

3) ฟังก์ชันสร้างโซนโดยใช้ไลบรารี่ Drawing tools ประเภท Polygon ในแผนที่ [9] ซึ่งกำหนดให้ 1 กลุ่มสร้างโซนได้ 1 โซนสามารถวาดโซนเป็นรูปทรงหลายเหลี่ยมได้

4) ฟังก์ชันเช็คอินโซนโดยใช้ฟังก์ชัน containsLocation ของ Google Map API [9] ซึ่งใช้วิธีการ Point in polygon [13] เพื่อคำนวณเปรียบเทียบตำแหน่งละติจูด ลองติจูดของนักเรียนกับตำแหน่งตะลิจูด ลองติจูดของโซน ถ้าตำแหน่งละติจูด ลองติจูดของนักเรียนอยู่ภายในโซนแปลว่าเช็คอินผ่าน

5) ฟังก์ชันแจ้งเตือนโซน การแจ้งเตือนเข้า-ออกโซนเพื่อให้เกิดความถูกต้องของการแจ้งเตือน จึงใช้วิธีการเก็บชุดข้อมูลตำแหน่งละติจูด ลองติจูด ของนักเรียนทั้งหมด 10 ชุดล่าสุด แต่ละชุดห่างกัน 6 วินาที โดยนำข้อมูลทั้ง 10 ชุด มาเปรียบเทียบกับค่าตำแหน่งละติจูด ลองติจูดของโซน โดยใช้ฟังก์ชัน containsLocation ของ Google Map API [9] ซึ่งใช้วิธีการ Point in Polygon [13] ด้วยการหาค่าเฉลี่ยชุดข้อมูล ทั้ง 10 จุด ถ้าค่าเฉลี่ยอยู่ในโซนมากกว่าอยู่นอกโซน แสดงว่าการเคลื่อนที่ของผู้ใช้จะมีลักษณะดังภาพ A ซึ่งจะแสดงป๊อปอัพแจ้งเตือนอยู่ในโซน แต่ถ้าค่าเฉลี่ยอยู่นอกโซนมากกว่าอยู่ในโซน แสดงว่าการเคลื่อนที่ของผู้ใช้มีลักษณะดังภาพ B ซึ่งจะแสดงป๊อปอัพแจ้งเตือนอยู่นอกโซนโดยใช้ระยะเวลา 1 นาทีก่อนจะแจ้งเตือน



ภาพที่ 1: ภาพ A แสดงการอยู่นอกโซน
ภาพ B แสดงการอยู่ในโซน

6) ฟังก์ชันแสดงตำแหน่งผู้ใช้ โดยใช้ฟังก์ชัน watchPosition()และ geocode() ของ Google Map API [9] ซึ่งระบบจะอัพเดท (Update) ตำแหน่งของผู้ใช้ทุก 10 วินาที ในฟังก์ชันการแสดงตำแหน่งผู้ใช้ Google Map API จะบังคับให้ใช้ HTTPS (Hypertext Transfer Protocol Secure) เก็บข้อมูลไว้เป็นความลับระหว่างคอมพิวเตอร์ของผู้ใช้กับระบบ [9]

7) ฟังก์ชันหน้าจอการแสดงผล โดยหน้าจอเมนูการทำกิจกรรมจะแสดงผลของการเช็คอินตำแหน่ง หน้าจอเมนูการเข้าร่วมกิจกรรมจะแสดงผลของการเช็คอินโซนและหน้าจอเมนูการแจ้งเตือนโซนจะแสดงผลของการเข้า-ออกโซน

2.2 ขั้นเตรียมการก่อนการทดลอง

การทดลองเพื่อวัดประสิทธิภาพด้านตำแหน่งของระบบผู้วิจัยใช้พื้นที่มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ ซึ่งมีพื้นที่ประมาณ 100 ไร่ รวมระยะทางในการเดินประมาณ 1 กิโลเมตร ระยะเวลาในการทดลอง 5 วัน วันละ 5 เวลา โดยภายใน 5 วัน เวลาจะแตกต่างกันและมีจำนวนนักเรียนที่ทดลองทั้งหมด 5 คน แต่ละคนจะใช้สมาร์ทโฟนรุ่นที่แตกต่างกันโดยมีทั้งระบบปฏิบัติการ Android และ iOS ดังนี้ Sumsung Gaxy Tab A6, iPad mini 2, Asus Note6, iPhone6 และ iPhone5 โดยแต่ละช่วงเวลานักเรียนจะต้องเดินไปเช็คอินตำแหน่งหมุด เช็คอินโซนและเข้า-ออกโซน เพื่อวัดประสิทธิภาพด้านตำแหน่ง เพื่อให้เกิดความสอดคล้องกับสภาพแวดล้อมการใช้งานจริงในการทดลองครั้งต่อไป กับกลุ่มตัวอย่างที่มีการใช้สมาร์ทโฟนรุ่นที่แตกต่างกันและตามสภาพแวดล้อมจริงผู้วิจัยไม่สามารถควบคุมการถือสมาร์ทโฟนของกลุ่มตัวอย่างได้จึงได้ออกแบบการทดลองตามตารางที่ 1

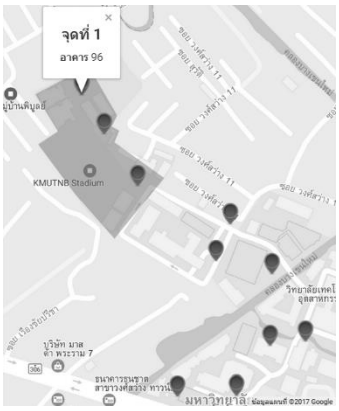
ตารางที่ 1 :สมาร์ทโฟนแต่ละรุ่นและรูปแบบการถือสมาร์ทโฟนของนักเรียนทั้ง 5 คน

นักเรียนคนที่	ยี่ห้อสมาร์ทโฟน	ลักษณะการถือสมาร์ทโฟน
นักเรียนคนที่ 1	Sumsung Galaxy Tab A6	ใช้วิธีการถือสมาร์ทโฟนไว้ที่มือ
นักเรียนคนที่ 2	iPad mini 2	ใช้วิธีการเก็บสมาร์ทโฟนไว้ในกระเป๋า
นักเรียนคนที่ 3	Asus Note6	ใช้วิธีการถือสมาร์ทโฟนไว้ที่มือ
นักเรียนคนที่ 4	iPhone6	ใช้วิธีการเอาสมาร์ทโฟนไว้ในกระเป๋ากางเกง
นักเรียนคนที่ 5	iPhone5	ใช้วิธีการเอาสมาร์ทโฟนไว้ในกระเป๋ากางเกง

2.3 ขั้นเตรียมข้อมูลก่อนทัศนศึกษาสำหรับครู

ก่อนที่จะเริ่มทัศนศึกษาครูจะต้องเตรียมเนื้อหาเพื่อให้นักเรียนทำกิจกรรมระหว่างทัศนศึกษาโดยใช้ปลั๊กอิน H5P เพื่อสร้างเนื้อหาและสร้างกลุ่มโดยใช้ปลั๊กอิน Buddypress เพื่อปักหมุดและสร้างโซนจุดที่ต้องการให้นักเรียนเดินไปทำกิจกรรมโดยปักหมุด 10 จุด ดังนี้ จุดที่ 1 อาคาร คณะพัฒนาธุรกิจและอุตสาหกรรม จุดที่ 2 อาคารพักอาศัยเพื่อการเรียนรู้และสันทนาการ จุดที่ 3 หน้า อาคารปฏิบัติการวิศวกรรมโยธา เคมีและไฟฟ้าอุตสาหกรรม จุดที่ 4 อาคารสถาบันสหกิจศึกษาและพัฒนาสื่ออิเล็กทรอนิกส์ ไทย-เยอรมัน จุดที่ 5 หน้าคณะวิศวกรรมศาสตร์ จุดที่ 6 อาคาร 40 ปี มจพ จุดที่ 7 ลานสวนปาล์ม มจพ จุดที่ 8 พระบรมราชานุสาวรีย์ รัชกาลที่ 4 จุดที่ 9 หอพระ (หลวงพ่อสิงห์) และจุดที่ 10 คณะสถาปัตยกรรมและ

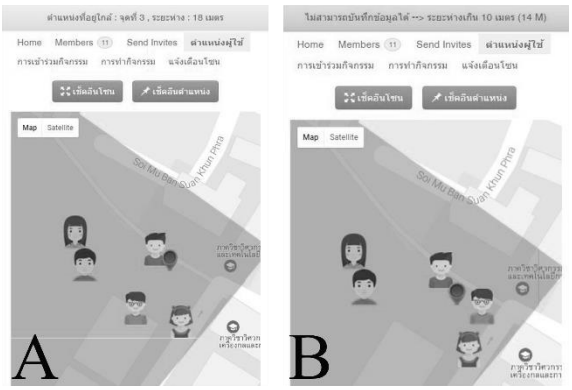
การออกแบบ และสร้างโซน 1 โซน บริเวณ อาคารปฏิบัติการวิทยาศาสตร์การกีฬาและรอบๆ ทุกครั้งก่อนการทดลอง ครูจะปักหมุดตำแหน่งเดิมและสร้างโซนรูปแบบเดิมแสดงดังภาพที่ 2



ภาพที่ 2: การแสดงหมุดและโซนที่ครูสร้างขึ้นเพื่อเป็นจุดสำหรับทำกิจกรรมของนักเรียน

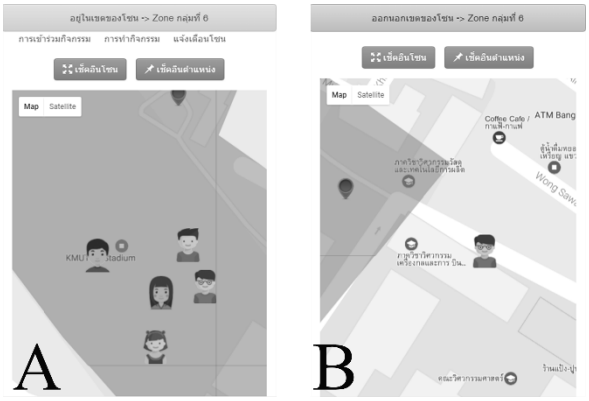
2.4 ขั้นทดลองเก็บข้อมูล

การลงพื้นที่ทัศนศึกษาสมาร์ทโฟนของนักเรียนจะต้องเปิดฟังก์ชัน GPS เมื่อเข้าสู่ระบบและเข้าไปยังกลุ่มที่กำหนด ให้เปิดเมนูตำแหน่งผู้ใช้ ซึ่งจะแสดงตำแหน่งของเพื่อนในกลุ่ม ตำแหน่งหมุดและโซน หลังจากนั้นเดินไปทำกิจกรรมตามหมุดในแผนที่เมื่ออยู่บริเวณหมุดและหมุดแจ้งเตือนด้วยการเด้งขึ้น-ลง ให้นักเรียนกดปุ่มเช็คอินตำแหน่งและเมื่อเดินไปยังบริเวณโซนที่ครูกำหนดให้นักเรียนกดปุ่มเช็คอินโซน เพื่อเป็นการยืนยันการได้ทำกิจกรรมจุดนั้นแล้ว แสดงดังภาพที่ 3



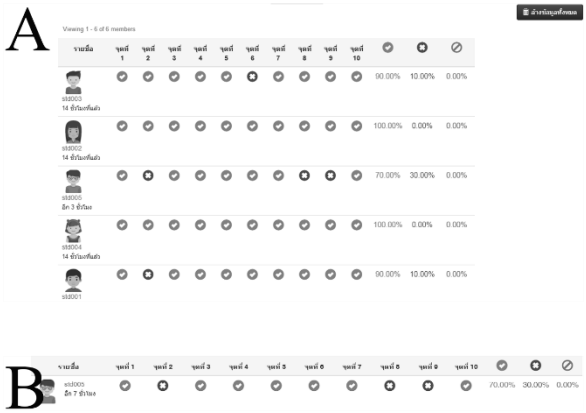
ภาพที่ 3: ภาพ A แสดงป๊อปอัพแจ้งเตือนระยะทางระหว่างนักเรียนกับหมุด ภาพ B แสดงป๊อปอัพแจ้งเตือน กรณีสมาร์ทโฟนยังจับตำแหน่ง GPS ไม่ได้

เมื่อนักเรียนเดินไปทำกิจกรรมภายในโซนจะมีป๊อปอัพแจ้งเตือนในนักเรียนทราบว่าอยู่ในโซน แต่กรณีที่นักเรียนรีเฟรส (Refresh) หน้าจอระบบ ขณะอยู่ในโซนระบบก็จะแสดงป๊อปอัพแจ้งเตือนว่าอยู่ในโซนและเมื่อนักเรียนเดินออกจากโซน ระบบจะแสดงป๊อปอัพแจ้งเตือนออกนอกโซน ซึ่งแสดงดังภาพที่ 4

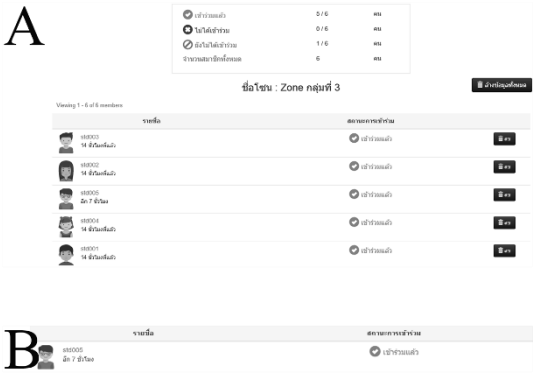


ภาพที่ 4: ภาพ A การแสดงป๊อปอัพการแจ้งเตือนอยู่ในโซน ภาพ B การแสดงป๊อปอัพการแจ้งเตือนอยู่นอกโซน

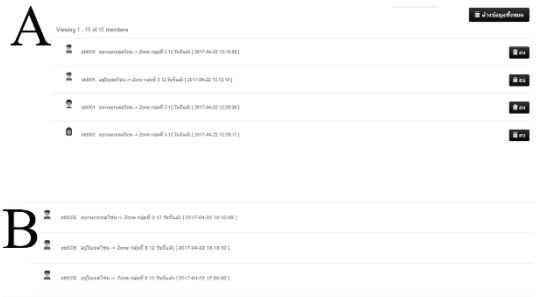
หลังจากที่นักเรียนทำกิจกรรมครบทุกจุดแล้ว ผลของการเช็คอินตำแหน่ง เช็คอินโซนและการเข้า–ออกโซนข้อมูลทั้งหมดจะถูกเก็บไว้และส่งไปให้ครูที่สร้างกลุ่มทราบ ซึ่งนักเรียนสามารถเข้ามาดูผลการทำกิจกรรมของตนเองได้ โดยหน้าจอผลการทำกิจกรรม หน้าจอผลการเข้าร่วมกิจกรรมและหน้าจอการแจ้งเตือนโซนสำหรับครูและนักเรียน แสดงภาพที่ 5, 6 และ7



ภาพที่ 5: ภาพ A แสดงหน้าจอผลการทำกิจกรรมสำหรับครู ภาพ B แสดงหน้าจอผลการทำกิจกรรมสำหรับนักเรียน



ภาพที่ 6 : ภาพ A แสดงหน้าจอผลการเข้าร่วมกิจกรรมสำหรับครู ภาพ B 7 แสดงหน้าจอผลการเข้าร่วมกิจกรรมสำหรับนักเรียน



ภาพที่ 7: ภาพ A แสดงหน้าจอผลการแจ้งเตือนโซนสำหรับครู ภาพ B 7 แสดงหน้าจอผลการแจ้งเตือนโซนสำหรับนักเรียน

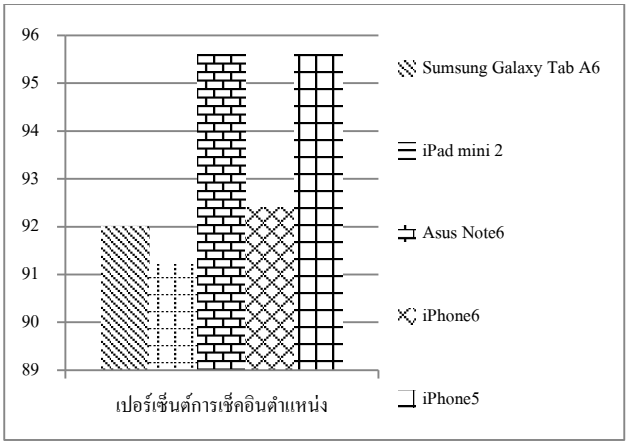
3. ผลการดำเนินการวิจัย

หลังจากพัฒนาระบบและทดสอบประสิทธิภาพด้านตำแหน่งของระบบสามารถสรุปออกมาเป็นประเด็นดังนี้

3.1 ผลการทดสอบประสิทธิภาพการเช็คอินตำแหน่ง

สรุปผลการเช็คอินตำแหน่งทั้ง 5 วันของนักเรียนที่ใช้สมาร์ทโฟนทั้ง 5 รุ่น โดยการเดินไปทำกิจกรรมแต่ละรอบไม่เหมือนกัน ซึ่งวันที่ 1 ถึงวันที่ 4 นักเรียนจะใช้วิธีการเดินไปถึงบริเวณหมุดแล้วจึงเช็คอินตำแหน่ง แต่วันที่ 5 นักเรียนใช้วิธีการเช็คอินตำแหน่งก่อนถึงบริเวณหมุดและใช้วิธีการเดินผ่านหมุดไปสักพักจึงเช็คอินตำแหน่ง

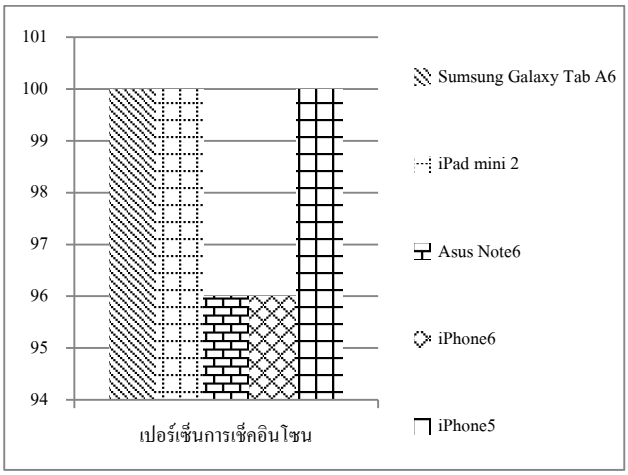
ผลปรากฏว่าสมาร์ทโฟนรุ่น iPhone 5 และ Asus Note6 มีความถูกต้องสูงสุดคือ 95.6% รองลงมาคือ iPhone6 92.4% Sumsung Gaxy Tab A6 92% และ iPad mini2 91.2% และความถูกต้องทั้งหมดคิดเป็น 93.36% จากการสังเกตระหว่างทดลองลักษณะการถือสมาร์ทโฟนมีผลต่อความแม่นยำของการรับสัญญาณ GPS แสดงผลการเช็คอินตำแหน่งดังภาพที่ 8



ภาพที่ 8: เปอร์เซ็นต์ความถูกต้องของการเช็คอินตำแหน่งสมาร์ทโฟน 5 รุ่น

3.3 ผลการทดสอบประสิทธิภาพการเช็คอินโซน

ผลการเช็คอินโซนตลอดระยะเวลาการทดลอง 5 วัน ผลปรากฏว่าสมาร์ทโฟนรุ่น Sumsung Gaxy Tab A6, iPad mini2 และ iPhone5 มีความถูกต้องสูงสุดคือ 100% รองลงมาคือ Asus Note6 และ iPhone6 มีความถูกต้อง 96% และความถูกต้องรวมทั้งหมดคิดเป็น 96% จากผลสรุปค่าความถูกต้องสูงเนื่องจากโซนเป็นพื้นที่กว้างจึงทำให้เกิดความผิดพลาดได้น้อย



ภาพที่ 9: เปอร์เซ็นต์ความถูกต้องของการเช็คอินโซนสมาร์ทโฟน 5 รุ่น

3.4 ผลการทดสอบประสิทธิภาพการแจ้งเตือนเข้า-ออกโซน

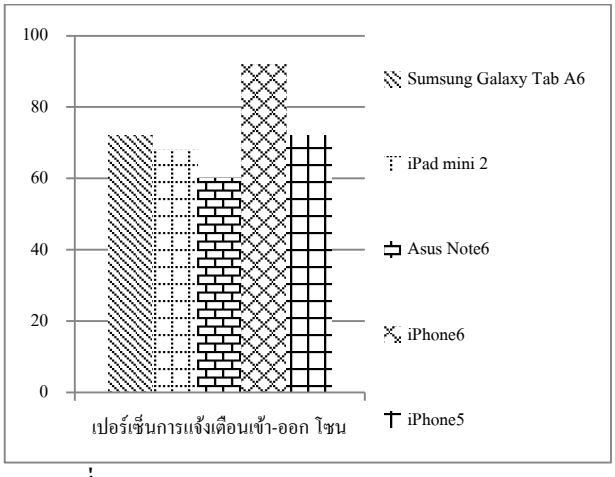
ผลการแจ้งเตือนเข้า-ออกโซนตลอดระยะเวลาการทดลอง 5 วัน ผลปรากฏว่าสมาร์ทโฟนรุ่น iPhone6 มีความถูกต้องสูงสุดคือ 92% รองลงมาคือ iPhone5และSumsung Gaxy Tab A6 72%, iPad mini2 68% และAsus Note6 60% และความถูกต้องรวมทั้งหมดคิดเป็น 72.8% จากการสังเกตระหว่างการทดลองซึ่งมีปัจจัยภายนอกที่ส่งผลให้เกิดปัญหาต่างๆดังนี้

1) เครื่องสมาร์ทโฟนจับสัญญาณ GPS ไม่ได้ แม้อยู่ในที่โล่ง ดังนั้นระบบจึงใช้วิธีการระบุตำแหน่งด้วย AGPS ทำให้ตำแหน่งผู้ใช้เกิดความคลาดเคลื่อนในระดับที่ยอมรับไม่ได้

2) นักเรียนอยู่บริเวณอาคาร ซึ่งรับสัญญาณ GPS ไม่ได้ ทำให้ตำแหน่งของนักเรียนคลาดเคลื่อน จากการระบุตำแหน่งด้วย AGPS

3) นักเรียนเดินบริเวณขอบโซนที่ชิดกับอาคารปฏิบัติการวิทยาศาสตร์การกีฬา ซึ่งเป็นตะแกรงเหล็กสูง จึงทำให้มาร์ทโฟนรับสัญญาณ GPS ไม่ได้

4) นักเรียนไม่ได้เปิดเมนูหน้าตำแหน่งผู้ใช้ จึงทำให้ระบบไม่สามารถจับตำแหน่งการเข้า-ออกโซน เนื่องจากระบบไม่สามารถประมวลผลเบื้องหลังได้ ถ้าไม่เปิดใช้งานเมนู จึงส่งผลให้ไม่มีข้อความแจ้งเตือนการออกนอกโซน



ภาพที่ 10: เปอร์เซ็นต์ความถูกต้องของการแจ้งเตือนเข้า-ออกโซน สมาร์ทโฟน 5 รุ่น

4. สรุป

จากการพัฒนาและทดสอบระบบจัดการเรียนรู้ทัศนศึกษานอกสถานที่ ซึ่งสรุปผลรวมความถูกต้องทั้ง 3 ด้านคิดเป็นร้อยละ 87.39% ซึ่งเป็นไปตามสมมุติฐานที่ตั้งไว้คือ ประสิทธิภาพด้านตำแหน่งมากกว่า 80% ระหว่างเก็บผลการทดลองปัญหาที่พบ ส่งผลให้เกิดอุปสรรคในการทดลองดังนี้ 1) ปัญหาเว็บเซิร์ฟเวอร์ล่มจึงทำให้ระบบใช้งานไม่ได้ ซึ่งใช้ระยะเวลาในการแก้ปัญหาประมาณ 30 นาที จึงจะสามารถทำการทดลองต่อได้ 2) ปัญหาจำนวนการ Request ของ Google Maps Geocoding API เพื่อแสดงตำแหน่งของนักเรียนเกินจำนวนที่จำกัดไว้ จึงใช้วิธีการเปลี่ยนเพิ่มโควต้า (Quota) ให้กับจำนวนการ Request ซึ่งปัญหาที่ได้กล่าวมีผลต่อความแม่นยำของ GPS และการเก็บผลการทดลอง

5. การศึกษาขั้นต่อไป

ในอนาคตจะพัฒนาเพิ่มเติมเกี่ยวกับส่วนของการนำเสนอข้อมูลหลังจากที่นักเรียนเช็คอินตำแหน่งผ่านแล้ว ระบบจะนำนักเรียนไปยังลิงค์ (Link) ข้อมูลกิจกรรมที่ครูเตรียมไว้เพื่อให้นักเรียนเรียนรู้และทำกิจกรรมตอบคำตอบและส่งไปยังครู โดยครูสามารถเข้าไปดูการตอบคำถามและให้คะแนนนักเรียนได้ และระบบสามารถนำไปใช้กับกลุ่มตัวอย่างจริงและใช้งานในสถานที่จริงได้

เอกสารอ้างอิง

[1] K. T. Henson and Dean," FOUNDATIONS FOR LEARNER-CENTERED EDUCATION: A NOWLEDGE BASE," *Education*, Vol.124 No.1,pp 5-16,2003.

[2] M. S. Y. Jong," DESIGN AND IMPLEMENTATION OF EAGLEEYE —AN INTEGRATED OUTDOOR EXPLORATORY EDUCATIONAL SYSTEM" *Research and Practice in Technology Enhanced Learning*, Vol. 8, No. 1,pp 43－64 ,2013

[3] G. Zurita and N. Baloian, "Context, Patterns and Geo-Collaboration to Support Situated Learning," *Ubiquitous Computing and Ambient Intelligence, J. Bravo, D. López-de-Ipiña, and F. Moya, eds., Springer-Verlag*, pp. 503-511, 2012.

[4] M. S. Y. Jong and E.T.H.Luk," Adopting EagleEye in Outdoor Exploratory Learning from the Teacher Perspective" *IEEE 14th International Conference on Advanced Learning Technologies*, pp. 617－621,2014.

[5] H. Hou-Sandí, D. Hulse, M.Jaquith, et al. (2003).wordpress. สืบค้นเมื่อวันที่ 2 ธันวาคม 2559, จาก: https://wordpress.org/

[6] G. With, P.Jørgensen, F.Petterson, et al.(2013).plugin H5P. สืบค้นเมื่อวันที่ 25 มกราคม 2560, จาก: https://h5p.org/

[7] J. J. Jacoby, B. B. Gorges, P.Gibbs, et al.plugin buddypress. สืบค้นเมื่อวันที่ 2 ธันวาคม 2559, จาก: https://buddypress.org/

[8] D.Ashbrook and T.Starner,"Learning significant locations and predicting user movement with GPS," *Proceedings of the 6th International Symposium on Wearable Computers*, pp.101–108, Retrieved December 30, 2012.

[9] L. Page and S. Brin.(1988).Google Map API. สืบค้นเมื่อวันที่ 2 กุมภาพันธ์ 2560, จาก: https://developers.google.com/maps/web/

[10] M.Otto, J.Thornton, C.Rebert , et al .(2010).bootstrap. สืบค้นเมื่อวันที่ 25 มกราคม 2560, จาก: http://getbootstrap.com/

[11] C. Song and Q. Zhang. "Sliding-Window Algorithm for Asynchronous Cooperrative Sensing in Wireless Cognitive Network," *IEEE Communications Society subject matter experts for publication in the ICC 2008 proceedings*,pp.3432-3426,2008

[12] C.Veness .(1993).Vecter Distance.สืบค้นเมื่อวันที่ 20 กุมภาพันธ์ 2560, จาก: http://www.movable-type.co.uk/scripts/latlong.html

[13] B.Zalik, and I. Kolingerova, " A cell-based point-in-polygon algorithm suitable for large sets of points," *Computers & Geosciences 27*,pp. 1135–1145,2001.

ต้นแบบระบบควบคุมอุปกรณ์ไฟฟ้าอัตโนมัติจากพฤติกรรมของผู้ใช้
Prototype of Automated Appliance Control System from User Behavior

ธีระ ศิริธีรากุล[1], เพชรไพลิน แซ่อุ้ย[2], พายุ พลูศรี[3], และ ศิริกุล ศิริธีรากุล[4]
[1234]ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
ภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
[1]teera.si@kmitl.ac.t[1], [2]ploy5315@gmail.com, [3]whatzupbeguys@gmail.com, [4]yok.math@gmail.com

บทคัดย่อ

ระบบบ้านอัจฉริยะได้รับความสนใจ และได้รับการพัฒนาในด้านต่างๆ เป็นอันมาก งานวิจัยนี้ มีแนวคิดที่จะเรียนรู้พฤติกรรมของผู้ใช้งาน และนำมาควบคุมอุปกรณ์ไฟฟ้าภายในบ้าน โดยงานวิจัยชิ้นนี้เป็นการพัฒนาระบบต้นแบบที่สามารถควบคุมเครื่องปรับอากาศ พฤติกรรมการเดินทางของผู้ใช้ ระบบต้นแบบได้รับการพัฒนาด้วยระบบแอนดรอยด์ ซึ่งเป็นระบบที่สามารถควบคุมเครื่องปรับอากาศ ผ่านเครือข่ายอินเตอร์เน็ตได้ โดยระบบจะเน้นไปที่ความสามารถในการวัดระยะการเดินทางจากผู้ใช้งานไปยังบ้านด้วยระบบจีพีเอส และใช้ข้อมูลดังกล่าวเพื่อทำการเปิดปิดเครื่องปรับอากาศ โดยอัตโนมัติ ได้อย่างเหมาะสม

***คำสำคัญ**: บ้านอัจฉริยะ, พฤติกรรมผู้ใช้, อินเตอร์เน็ตของสรรพสิ่ง, แอนดรอยด์*

Abstract

Smart home system has become popular and has been developed from different point of views. This research focuses on learning user behaviors and use these information for automatically controlling home appliance. In this work, we developed an android application which can remotely control an air conditioner unit via the Internet. The prototype focuses on tracking user's distance from home via GPS and uses this information to remotely and automatically control an air conditioner unit.

Keyword: Smart home, user behavior, Internet of Things, android

1. บทนำ

ในปัจจุบัน ระบบบ้านอัจฉริยะ (smart home) หรือระบบบ้านอัตโนมัติ (home automation) ได้รับการพัฒนาในแง่มุมต่างๆ เป็นอันมาก ไม่ว่าจะเป็นในด้านความปลอดภัย (smart security) [1, 2] ซึ่งมีการนำกล้องวงจรปิดและเซนเซอร์ต่างๆ มาเชื่อมต่อกับเครือข่าย ให้สามารถทำการแจ้งเตือนไปยังโทรศัพท์หรือหน่วยงานที่เหมาะสม หรือในด้านการแพทย์ การดูแลผู้สูงอายุ [3] ที่นำเทคโนโลยีที่หลากหลาย ไม่ว่าจะเป็นการวัดค่า จัดเก็บ สื่อสารข้อมูล การวิเคราะห์ข้อมูลในแง่ต่างๆ มาช่วยดูแลผู้ป่วยหรือผู้สูงอายุ ซึ่งในปัจจุบันมีมุมมองที่ดีขึ้นเกี่ยวกับการใช้เทคโนโลยีในด้านนี้ [4]

การใช้เทคโนโลยีระบบบ้านอัจฉริยะ มาอำนวยความสะดวกและช่วยประหยัดพลังงาน ยังเป็นอีกหนึ่งด้านที่ได้รับความนิยมเป็นอย่างมาก อีกทั้งการศึกษาทางด้านอินเตอร์เน็ตของสรรพสิ่ง (Internet of Things – IoT) (ศึกษาข้อมูลเพิ่มเติมได้ใน [5]) ยังช่วยสร้างเทคโนโลยีใหม่ๆ เพื่อใช้กับระบบบ้านอัจฉริยะได้อย่างหลากหลายอีกด้วย

ในบทความวิจัยนี้ จะนำเสนอต้นแบบระบบอัจฉริยะ ที่สามารถควบคุมเครื่องปรับอากาศ โดยอาศัยข้อมูลจากพฤติกรรมของผู้ใช้งาน ระบบจะแบ่งออกเป็นสองส่วน โดยส่วนแรกจะติดตั้งอยู่ภายในบ้าน และทำหน้าที่รับข้อมูลผ่านอินเตอร์เน็ต และทำการควบคุมเครื่องปรับอากาศผ่านสัญญาณอินฟราเรด และส่วนที่สองจะเป็นโปรแกรมที่ทำงานอยู่บนอุปกรณ์สื่อสารเคลื่อนที่ระบบแอนดรอยด์ ทำหน้าที่ตรวจจับพฤติกรรมผู้ใช้งาน (ระยะห่างจากบ้าน) และทำการส่งสัญญาณควบคุมไปยังส่วนที่หนึ่งได้อย่างอัตโนมัติ

ทั้งนี้ อุปกรณ์ส่วนที่หนึ่งนั้น ได้รับการออกแบบโดยใช้หลักการของ universal remote ที่สามารถเรียนรู้สัญญาณจาก

รีโมทควบคุมอื่นๆ ได้ ด้วยเหตุนี้ ต้นแบบที่ได้จึงสามารถนำไปประยุกต์ได้กับอุปกรณ์ไฟฟ้าทุกชนิด ที่มีรีโมทควบคุมโดยสัญญาณอินฟราเรด แม้ว่าต้นแบบที่รายงานในงานวิจัยชิ้นนี้จะใช้การควบคุมเครื่องปรับอากาศเป็นตัวอย่างก็ตาม

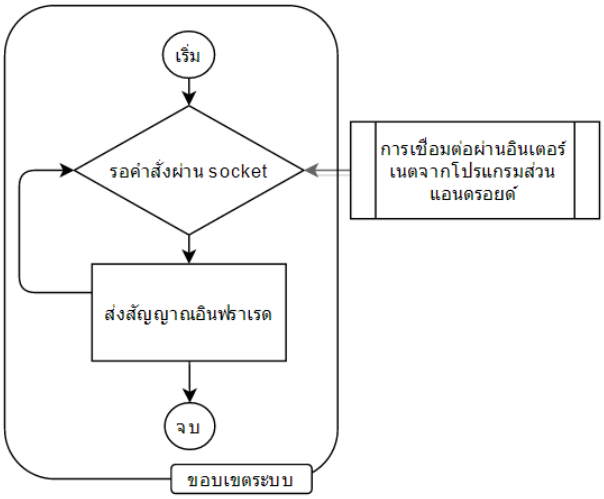
บทความวิจัยนี้ จะเป็นประโยชน์ต่อวงการศึกษาและวงการอุตสาหกรรมไอที กล่าวคือ จะเป็นต้นแบบในการประยุกต์ใช้ IoT ในระบบบ้านอัจฉริยะ เพื่อเป็นประโยชน์ต่อการศึกษาวิจัยเพิ่มเติม อีกทั้งยังเป็นต้นแบบในการพัฒนาอุปกรณ์ชนิดนี้ในเชิงพาณิชย์ต่อไปด้วย

2. ระบบควบคุมอุปกรณ์ไฟฟ้าอัตโนมัติจากพฤติกรรมของผู้ใช้

ในส่วนนี้ จะกล่าวถึงรายละเอียดของระบบ ทั้งในส่วนควบคุมอุปกรณ์ไฟฟ้าด้วยสัญญาณอินฟราเรด และในส่วนโปรแกรมบนอุปกรณ์เคลื่อนที่ รวมทั้งจะกล่าวถึงภาพรวมของระบบอีกด้วย

2.1 ส่วนควบคุมอุปกรณ์ไฟฟ้าด้วยสัญญาณอินฟราเรด

ภาพที่ 1 แสดงถึงผังการทำงานของส่วนควบคุมอุปกรณ์ไฟฟ้าด้วยสัญญาณอินฟราเรด โดยระบบในส่วนนี้ จะเริ่มทำงาน โดยการเปิด TCP Socket เพื่อรอรับการเชื่อมต่อจากโปรแกรมในส่วนแอนดรอยด์



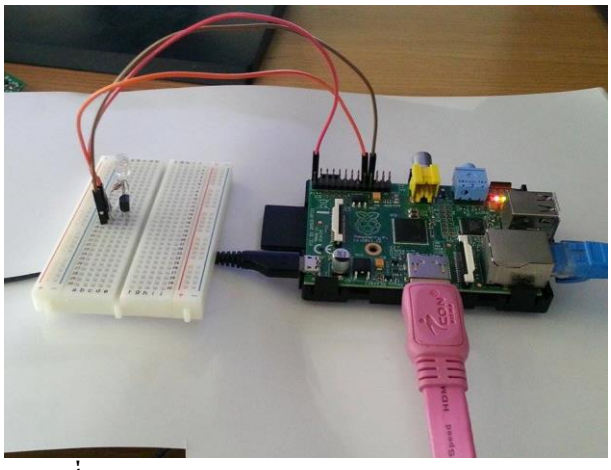
ภาพที่ 1: ระบบควบคุมอุปกรณ์ไฟฟ้าด้วยสัญญาณอินฟราเรด

โดยหลังจากมีการเชื่อมต่อจากโปรแกรม ในส่วนแอนดรอยด์แล้ว ระบบจะอ่านคำสั่ง และส่งสัญญาณอินฟราเรดออกไปยังเครื่องปรับอากาศ หรืออุปกรณ์ไฟฟ้าอื่นๆ ตามที่ได้มีการสั่งงาน

อุปกรณ์ในส่วนนี้ ออกแบบโดยใช้ ราสเบอรีพาย (Raspberry Pi) ซึ่งเป็นบอร์ดคอมพิวเตอร์ขนาดเล็ก ราคาประหยัด ที่สามารถติดตั้งระบบปฏิบัติการลินุกซ์ (Linux) ได้ ทั้งนี้ ระบบต้นแบบได้รับการพัฒนาโดยใช้ LIRC (Linux Infrared remote control) ซึ่งเป็นคลังโปรแกรมที่สามารถส่งสัญญาณอินฟราเรดเพื่อควบคุมอุปกรณ์ไฟฟ้าที่เป็นที่นิยมส่วนใหญ่ผ่าน GPIO (General-purpose input/output) ของราสเบอรีพายได้

ภาพที่ 2 แสดงตัวอย่างการต่อวงจรส่วนควบคุมอุปกรณ์ไฟฟ้าด้วยสัญญาณอินฟราเรด ซึ่งจะเห็นได้ว่าเป็นวงจรอย่างง่าย โดยในภาพจะ เห็นการต่อสายอินเตอร์เน็ต(ขวา) ซึ่งสามารถเปลี่ยนเป็นอุปกรณ์ WIFI แบบ USB ได้ และสายจอภาพ(ล่าง) ซึ่งไม่จำเป็นในขณะใช้งาน

นอกจากนี้ แม้ว่าต้นแบบจะใช้ราสเบอรีพายโมเดล B แต่เนื่องด้วยระบบการควบคุมอุปกรณ์ด้วยสัญญาณอินฟราเรดนี้ มีความซับซ้อนและความต้องการความสามารถในการคำนวณต่ำ เราจึงสามารถใช้ราสเบอรีพายโมเดล A ซึ่งมีขนาดเล็กกว่า ราคาถูกกว่า และประหยัดพลังงานขึ้นอีกด้วย

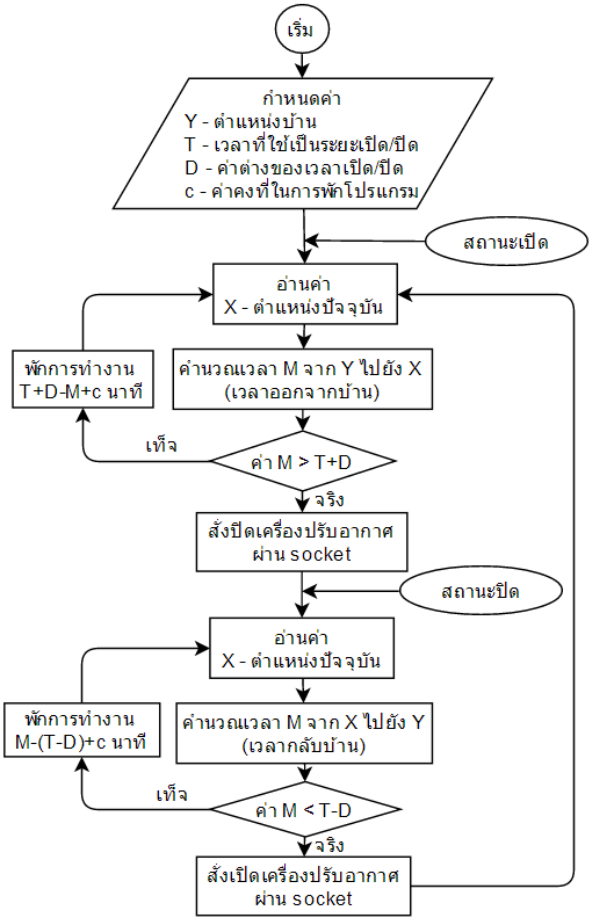


ภาพที่ 2: ตัวอย่างส่วนควบคุมอุปกรณ์ไฟฟ้าด้วยสัญญาณอินฟราเรด

2.2 ส่วนตรวจจับและส่งสัญญาณควบคุมผ่านอินเตอร์เน็ตจากอุปกรณ์ที่ใช้ระบบปฏิบัติการแอนดรอยด์

โปรแกรมในส่วนนี้ ได้รับการพัฒนาขึ้นเป็นโปรแกรมประยุกต์บนระบบปฏิบัติการแอนดรอยด์ โดยจะมีการ

กำหนดค่า Y เป็นตำแหน่ง (พิกัด GPS) ของบ้าน ค่า T เป็นเวลาที่ใช้เป็นระยะเปิดปิด ซึ่งจะใช้ร่วมกับค่าต่างของเวลาเปิดปิด D เช่น T=10 D=2 หมายถึง การเดินทางออกจากบ้าน 10+2=12 นาที จะทำให้ระบบสั่งปิดเครื่องปรับอากาศ ในขณะที่หากกำลังเดินทางกลับบ้าน หากระบบคำนวณแล้วว่าเหลือเวลาอีก 10-2=8 นาทีจะถึงบ้าน ระบบจะสั่งเปิดเครื่องปรับอากาศไว้รอ ซึ่งในต้นแบบนี้ จะใช้ค่า T เดียวกันในการเปิดและปิด เพื่อความสะดวกในการคำนวณ ทั้งนี้ ค่า D มีไว้เพื่อป้องกันมิให้เกิดการเปิดปิดบ่อยครั้งขณะที่อยู่ในระยะระหว่างเปิด/ปิด การกำหนดค่า T สามารถทำได้โดยคำนึงถึง 2 ปัจจัยหลัก ได้แก่ ความต้องการในการประหยัดพลังงาน (ปิดเร็ว/ช้า) และความต้องการความสะดวกสบาย (เปิดเร็ว/ช้า) ส่วนการกำหนดค่า D นั้น ควรคำนึงถึงความผิดพลาดที่ได้จากสัญญาณ GPS เป็นหลัก



ภาพที่ 3: ระบบตรวจจับและส่งสัญญาณควบคุมผ่านอินเตอร์เน็ต

จากภาพที่ 3 จะเห็นได้ว่า การทำงานของระบบจะมีสองสถานะหลักคือ เปิด กับ ปิด ซึ่งทั้งสองสถานะมีขั้นตอนการทำงานในลักษณะเดียวกัน คือ ระบบจะทำการอ่านตำแหน่งปัจจุบันโดยอาศัยระบบ GPS ที่มีอยู่บนอุปกรณ์แอนดรอยด์ หลังจากนั้น ระบบจะนำค่าตำแหน่งปัจจุบัน (X) และตำแหน่งของบ้าน (Y) ไปคำนวณเวลา (M) ที่ใช้ในการขับรถออกจากบ้าน (สถานะเปิด) หรือเวลาที่ใช้ในการขับรถเข้าบ้าน (สถานะปิด) ด้วยความสามารถของ Google Map API หลังจากนั้น ระบบจะนำเวลา M ที่คำนวณได้ ไปเทียบกับค่า T และ D ดังที่กล่าวไว้ตอนต้น เพื่อคำนวณว่า สมควรที่จะสั่งปิดหรือเปิดผ่าน TCP Socket ไปยังอุปกรณ์ในส่วนควบคุมอุปกรณ์ไฟฟ้าด้วยสัญญาณอินฟราเรดหรือยัง ซึ่งหากสั่งปิดหรือเปิดไปแล้ว ระบบจะย้ายสถานะตัวเอง ทั้งนี้ ระบบต้นแบบยังได้รับการออกแบบให้ระบบเปลี่ยนสถานะระหว่างเปิดและปิดด้วยมือ หรือด้วยการตรวจจับสถานะปัจจุบันของเครื่องปรับอากาศได้อีกด้วย

ทั้งนี้ หากระบบตรวจสอบแล้วพบว่า ยังไม่สมควรที่จะสั่งปิดหรือเปิดเครื่องปรับอากาศ ระบบจะทำการหยุดเป็นเวลา (T+D)-M+c ในกรณีที่อยู่ในสถานะเปิด หรือ M-(T-D)+c ในกรณีที่อยู่ในสถานะปิด โดยเวลาที่พักการทำงานนั้น คำนวณจากเวลาที่เหลือจากการเดินทาง จากตำแหน่งปัจจุบัน จนถึงระยะที่จะทำการเปลี่ยนสถานะ

3. ผลการดำเนินงานและการใช้งานระบบ

ในส่วนนี้ จะกล่าวถึงการใช้งานระบบ รวมถึงตัวอย่างหน้าจอของโปรแกรมในส่วนตรวจจับและส่งสัญญาณควบคุมผ่านอินเตอร์เน็ตจากอุปกรณ์ที่ใช้ระบบปฏิบัติการแอนดรอยด์

3.1 การใช้งานระบบ

จากแผนภาพการทำงานของผู้ใช้งานระบบ (use case diagram) ที่แสดงในภาพที่ 4 จะเห็นได้ว่า ระบบที่สร้างขึ้นมีฟังก์ชันการใช้งานอยู่ 2 ส่วนหลัก คือ การเปิด/ปิดเครื่องปรับอากาศด้วยตัวเอง และการตั้งค่าและใช้งานระบบเปิด/ปิดเครื่องปรับอากาศอัตโนมัติ ซึ่งหากเปิดใช้งานอัตโนมัติแล้ว ระบบจะทำการตรวจจับตำแหน่ง รวมทั้งดำเนินการคำนวณและส่งสัญญาณเปิดหรือปิดเครื่องปรับอากาศให้ โดยที่ผู้ใช้ไม่จำเป็นต้องสั่งการใดอีก

3.2 ตัวอย่างหน้าจอของระบบ

ภาพที่ 5 (a) แสดงหน้าจอการทำงานเมื่อผู้ใช้เข้าสู่ระบบในอุปกรณ์แอนดรอยด์ ซึ่งหน้าจอหลักในการทำงานจะมี 2 หน้าจอ อันได้แก่หน้าจอสำหรับการสั่งการเปิด/ปิดเครื่องปรับ

อากาศด้วยมือ ภาพที่ 5 (b) และหน้าจอสำหรับเริ่มต้นหรือหยุดการทำงานอัตโนมัติของระบบ ภาพที่ 5 (c) ส่วนหน้าจอที่แสดงในภาพที่ 5 (d) นั้น จะเป็นหน้าจอที่สามารถกำหนดตำแหน่งของบ้านได้ โดยระบบจะเริ่มต้นจากจุดปัจจุบันที่ตรวจจับได้ด้วยระบบ GPS และผู้ใช้งานสามารถเปลี่ยนจุดที่จะกำหนดให้เป็นบ้านได้ตามความเหมาะสม

4. สรุปผลการดำเนินงานและข้อเสนอแนะ

งานวิจัยชิ้นนี้ได้นำเสนอต้นแบบของระบบควบคุมอุปกรณ์ไฟฟ้าอัตโนมัติจากพฤติกรรมของผู้ใช้ ซึ่งในระบบที่นำเสนอนั้น สามารถตรวจจับพฤติกรรมการเข้าออกบ้านของผู้ใช้งานและทำการสั่งงานเปิด/ปิดเครื่องปรับอากาศได้อย่างเหมาะสม อีกทั้งยังมีราคาประหยัด และติดตั้งได้ง่ายอีกด้วย

นอกจากนี้ ระบบที่ได้ยังเป็นต้นแบบที่ดีที่จะนำไปพัฒนาต่อยอด เพื่อให้ใช้งานได้หลากหลายมากขึ้น เช่น การพัฒนาส่วนตรวจจับและส่งสัญญาณในระบบปฏิบัติการอื่นๆ นอกจากแอนดรอยด์ (ระบบไอโอเอส วินโดว์โฟน ฯลฯ) หรือจะนำไปเชื่อมต่อกับอุปกรณ์เครื่องใช้ไฟฟ้าในบ้านอื่นๆ เพื่อช่วยอำนวยความสะดวกและช่วยประหยัดพลังงานแก่ผู้ใช้ได้อย่างหลากหลายมากขึ้นได้อีกด้วย

เอกสารอ้างอิง

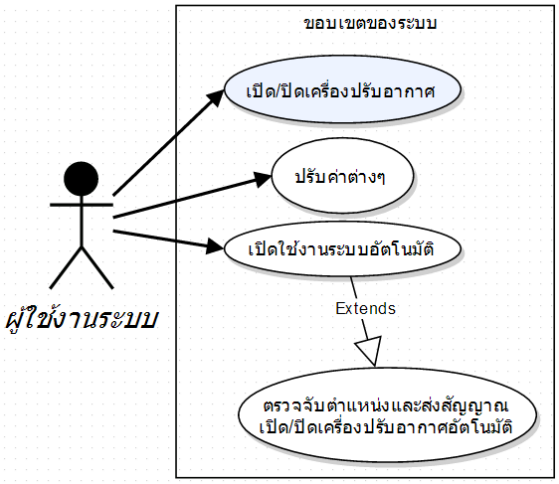
[1] R. J. Robbles and T. Kim, “A Review on Security in Smart Home Development”, International Journal of Advance Science and Technology, Vol. 15, February 2010.

[2] M. Griffiths, “Smart Home Security”, Homebuilding & Renovating, Retrieved July 2015.

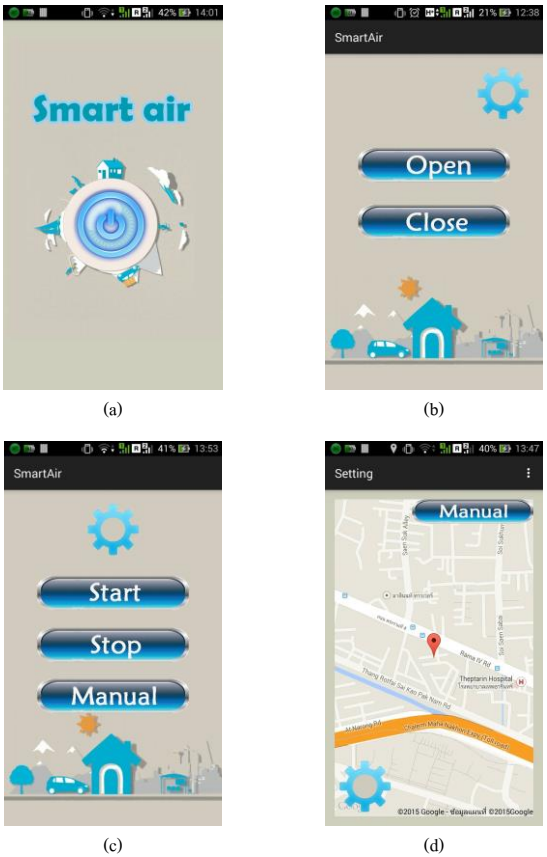
[3] V. Rialle, F. Duchene, N. Noury, L. Bajolle, and J. Demongeot, “Health “Smart” Home: Information Technology for Patients at Home”, Telemedicine Journal and e-Health, Vol. 8 Issue 4 July 9, 2004.

[4] G. Demiris, M. J. Rantz, M. A. Aud, K. D. Marek, H. W. Tyrer, M. Skubic, and A. A. Hussam, “Older adults’ attitudes towards and perceptions of ‘smart home’ technologies: a pilot study”, Informatics for Health and Social Care, Vol. 29, No. 2, Pages 87-94, 2004.

[5] J. Gubbi, R. Buyya, S. Marusic, and M. Palaniswami, “Internet of Things (IoT): A vision, architectural elements, and future directions”, Future Generation Computer Systems, Vo. 29, Issue 7, Pages 1645-1660, September 2013.



ภาพที่ 4: แผนภาพการทำงานของผู้ใช้งานระบบ



(a) (b) (c) (d)

ภาพที่ 5: (a) หน้าจอเริ่มต้น (b) หน้าจอสั่งการเปิดปิดด้วยมือ (c) หน้าจอเปิดปิดระบบอัตโนมัติ และ (d) หน้าจอการตั้งค่าตำแหน่งบ้าน

จิ๊กซอว์: ตารางธาตุแบบความเป็นจริงเสริมสำหรับการสนับสนุนการเรียนรู้แบบร่วมมือ
The Jigsaw: An Augmented Reality Periodic Table for Supporting Collaborative Learning

วัชรพงษ์ โถรัตน์ (Watcharapong Thorat)[1] เดชาวุฒิ วานิชสรรพ์ (Dechawut Wanichsan)[2]
พลวัฒน์ วัฒนศิริยานนท์ (Phollawat Wattanasiriyanon)[3] และวัชรี วรัจฉรียกุล (Watcharee Waratchareeyakul)[4]
[1,3]คณะครุศาสตร์ [2]คณะวิทยาการคอมพิวเตอร์และเทคโนโลยีสารสนเทศ [4]คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏรำไพรรณี
[1]watcharapong.thorat@gmail.com, [2]dechawut.w@rbru.ac.th, [3]phollawat21@gmail.com, [4]waratcha@yahoo.com

บทคัดย่อ

งานวิจัยชิ้นนี้มีวัตถุประสงค์เพื่อพัฒนาสื่อการเรียนวิชาเคมีเรื่องตารางธาตุ ซึ่งมีส่วนประกอบสำคัญสองประการ คือ 1) การนำเทคโนโลยีเออาร์มาใช้งานเพื่อนำเสนอข้อมูลในรูปแบบโมเดลสามมิติพร้อมเสียงบรรยายประกอบแต่ละธาตุ และ 2) สร้างมาร์คเกอร์ในรูปแบบจิ๊กซอว์ เพื่อสามารถนำไปจัดกิจกรรมการเรียนรู้แบบร่วมมือได้ ผลการทดลองแสดงให้เห็นว่า สื่อการสอนดังกล่าวสามารถนำไปใช้งานได้จริง ระยะที่เหมาะสมที่สุดในการตรวจจับมาร์คเกอร์อยู่ที่ 17-20 เซนติเมตร และผู้ใช้งานมีความพึงพอใจในในการใช้งานตารางธาตุที่สร้างขึ้นในระดับดีมาก

คำสำคัญ: ความเป็นจริงเสริม ตารางธาตุ การเรียนรู้แบบร่วมมือ เคมีศึกษา

Abstract

This research aimed to develop Chemistry learning media in the topic of periodic table which was composed of two important parts. First, it adopted an augmented reality to visualize three-dimensional model together with sound for describing characteristics of each element. Second, Markers were developed as jigsaw in order that they could be used for setting an activity which supports a collaborative learning. The experimental result showed that the media could be practically used, the optimal distance for marker detection was ranging from 17 to 20 centimeters, and the user felt satisfied after using the developed periodic table at very high level.

Keyword: Augmented Reality, Periodic Table, Collaborative Learning, Chemistry Education

1. บทนำ

บทเรียนเรื่องตารางธาตุ เป็นเนื้อหาพื้นฐานในการเรียนศาสตร์ทางด้านเคมี ทั้งในระดับช่วงชั้นมัธยมปลายสายวิทย์-คณิต และระดับอุดมศึกษาในสาขาวิชาที่เกี่ยวข้องกับเคมี รายละเอียดของธาตุใด ๆ มีเนื้อหาที่ควรรู้อยู่หลายประการ เช่น จำนวนอิเล็กตรอน การจัดเรียงอิเล็กตรอน กลุ่มและหมู่ของธาตุ เป็นต้น อย่างไรก็ตาม การที่เนื้อหาเหล่านี้มีปริมาณมาก ทำให้ผู้เรียนต้องใช้ความสามารถในการจดจำสูง และสื่อการเรียนส่วนมากยังใช้หนังสือเป็นอุปกรณ์ประกอบการสอนหลัก ซึ่งขาดการกระตุ้นความสนใจให้แก่ผู้เรียน และผู้เรียนไม่ได้เรียนรู้เนื้อหาอื่นที่เกี่ยวข้องกับธาตุแต่ละตัว เช่น รูปร่างของธาตุ และการนำไปใช้ประโยชน์ เป็นต้น

ในปัจจุบัน มีงานวิจัยหลายชิ้นที่นำเทคโนโลยีมาช่วยสนับสนุนการสอนในวิชาเคมี ดังที่ปรากฏในงานวิจัยระดับชาติ เช่น การสร้างเกมในการเรียนเรื่องตารางธาตุ [1] การสร้างหนังสือตารางธาตุที่ทำงานร่วมกับเทคโนโลยีความจริงเสริม [2] เป็นต้น และดังที่ปรากฏในงานวิจัยระดับนานาชาติ เช่น การศึกษาผลกระทบของบทเรียนคอมพิวเตอร์ช่วยสอนวิชาเคมี เรื่อง ส่วนประกอบของอะตอม อากาศ และคาร์บอน [3] เป็นต้น จากการศึกษางานวิจัยที่เกี่ยวข้องทั้งสามชิ้น พบว่า ยังไม่มีงานวิจัยชิ้นไหนทำการสร้างสื่อตารางธาตุที่มีการแยกส่วนได้เพื่อนำไปประยุกต์ในการจัดการเรียนรู้ ผู้วิจัยจึงได้แนวคิดโดยการนำเทคโนโลยีความจริงเสมือน มาใช้งานร่วมกับ โปรแกรมประยุกต์ที่พัฒนาขึ้นและทำงานบนระบบปฏิบัติการแอนดรอยด์ เพื่อนำเสนอในรูปแบบของโมเดลสามมิติพร้อมเสียงบรรยาย ทำให้เกิดความน่าสนใจ

และทำให้ผู้เรียนได้รู้ข้อมูลเพิ่มเติมแต่ละธาตุ เช่น การนำไปใช้ประโยชน์ และรูปร่างของธาตุ เป็นต้น นอกจากนี้ยังได้นำความรู้จากผู้เชี่ยวชาญด้านเคมี ว่าธาตุใดมีอันตรายและไม่มีอันตรายต่อมนุษย์ บรรจุรวมในเนื้อหาบรรยายด้วย

ตารางธาตุที่พัฒนาขึ้นมีอยู่สองรูปแบบ ได้แก่ รูปแบบของจิ๊กซอว์ ที่มาร์คเกอร์แต่ละตัวสามารถแยกส่วนได้ ทำให้ผู้สอนสามารถนำไปประยุกต์กับการจัดกิจกรรมการสอนแบบมีส่วนร่วม อีกรูปแบบหนึ่งคือรูปแบบไฟล์ ที่ให้ความสะดวกในการพกพาและใช้งาน

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 โปรแกรมประยุกต์ที่ใช้ในการพัฒนา

1) โปรแกรมออโตเดสก์มายา (Autodesk Maya) เป็นโปรแกรมที่ใช้สร้างแบบจำลอง หรือ ชิ้นงานสามมิติสำหรับงานออกแบบผลิตภัณฑ์ โดยมีความโดดเด่นในเรื่องการสร้างรูปทรงที่ต้องการความละเอียด ความแม่นยำ และสามารถสร้างแบบจำลองสามมิติให้โปรแกรมอื่นนำไปใช้งานต่อได้ด้วยการบันทึกไฟล์เป็นนามสกุล .fbx

2) โปรแกรมยูนิตี (Unity) คือเกมเอนจิน (Game Engine) ที่ทำงานบนระบบปฏิบัติการวินโดวส์ (Windows) ระบบปฏิบัติการโอเอสเอ็กซ์ (OS X) และเว็บ (Web) เท่านั้น แต่ปัจจุบันได้เพิ่มความสามารถในการทำงานโปรแกรมบนระบบปฏิบัติการไอโอเอส (IOS) ระบบปฏิบัติการแอนดรอยด์ (Android) และแพลตฟอร์ม (Platfrom) อื่น ๆ

3) โปรแกรมวูฟอเรีย (Vuforia) เป็นเครื่องมือที่ใช้งานร่วมกับโปรแกรมยูนิตี ให้สามารถทำงานร่วมกับเทคโนโลยีความเป็นจริงเสริมได้

4) โปรแกรมอโดบีโฟโตชอป (Adobe Photoshop) เป็นโปรแกรมที่ใช้ในการสร้างและแก้ไขรูปภาพ และมีเครื่องมือในการสนับสนุน การสร้างงานประเภทสิ่งพิมพ์งานมัลติมีเดีย จนถึงงานออกแบบและพัฒนาเว็บไซต์

2.2 เทคโนโลยีความเป็นจริงเสริม

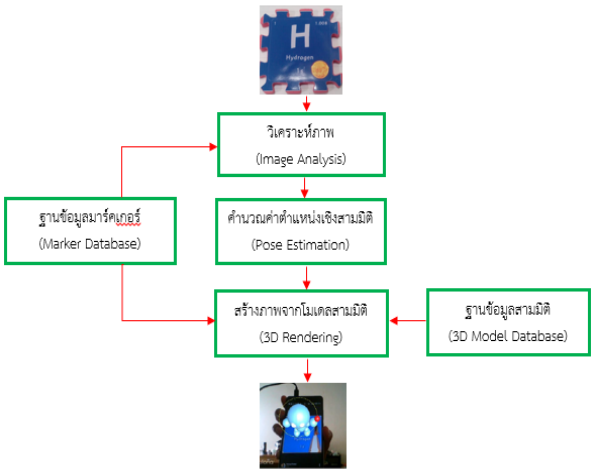
เทคโนโลยีความจริงเสริม (Augmented Reality: AR) คือการรวมเอาความจริงและความเสมือนเข้าด้วยกัน โดยผ่านโปรแกรมและอุปกรณ์เชื่อมต่อต่าง ๆ และแสดงผลในรูปแบบรูปทรงสามมิติ วิดีโอ และมีการปฏิสัมพันธ์ในเวลาจริง หลักการทำงานของเทคโนโลยีความเป็นจริงเสริม[4] มีขั้นตอนดังนี้

1) การวิเคราะห์ภาพ (Image Analysis) เป็นขั้นตอนการค้นหามาร์คเกอร์ (Marker) จากภาพที่ได้จากกล้องแล้วสืบค้นจากฐานข้อมูลมาร์คเกอร์ (Marker Database) ที่มีการเก็บข้อมูลขนาดและรูปแบบของมาร์คเกอร์ เพื่อนำมาวิเคราะห์รูปแบบ

2) การคำนวณค่าตำแหน่งเชิงสามมิติ (Pose Estimation) คือการคำนวณตำแหน่งของมาร์คเกอร์เทียบกับกล้อง

3) กระบวนการสร้างภาพสองมิติ จากโมเดลสามมิติ (3D Rendering) เป็นการเพิ่มข้อมูลเข้าไปในภาพ โดยใช้ค่าตำแหน่งเชิงสามมิติ ที่คำนวณได้จนได้ภาพเสมือนจริง

ภาพรวมทั้งหมดของการทำงานตั้งแต่เริ่มต้นจนถึงขั้นแสดงผล สรุปได้ดังข้อมูลในภาพที่ 1



ภาพที่ 1: การทำงานของเทคโนโลยีความเป็นจริงเสริม

2.3 การเรียนรู้แบบร่วมมือ

การเรียนรู้แบบร่วมมือ (Collaborative Learning) [5] เป็นการจัดการเรียนการสอนโดยเน้นให้ผู้เรียนลงมือปฏิบัติงานเป็นกลุ่มย่อย และสมาชิกของกลุ่มมีความสามารถแตกต่างกัน ให้ช่วยเหลือพึ่งพากัน เพื่อเสริมสร้างสมรรถภาพการเรียนรู้ของแต่ละคนและการทำงานเป็นกลุ่ม โดยผู้วิจัยได้ทำการสร้างจิ๊กซอว์เพื่อนำไปใช้กับการจัดกิจกรรมการเรียนรู้

แบบร่วมมือ เช่น การทำกิจกรรมจับกลุ่มต่อจิ๊กซอว์ธาตุให้สมบูรณ์ตามหมู่หรือ การแบ่งกลุ่มตามกลุ่มธาตุ เป็นต้น และยังสามารถนำไปจัดกิจกรรมสอดคล้องกับการเรียนรู้แบบร่วมมือแบบอื่นได้อีกด้วย

2.4 งานวิจัยที่เกี่ยวข้อง

Su Cai [6] ได้พัฒนาโปรแกรมประยุกต์ที่ใช้เทคโนโลยีความจริงเสริมในรายวิชาเคมี เรื่อง องค์ประกอบของสาร เพื่อให้นักเรียนสามารถโต้ตอบกับแบบจำลองสามมิติ ในการทดลองวิชาเคมี โดยผลการทดลองที่ได้พบว่า สื่อการสอนดังกล่าวมีผลต่อการเพิ่มผลสัมฤทธิ์ทางการเรียนอย่างมีนัยสำคัญ และผู้เรียนมีความพึงพอใจที่ดีในการใช้โปรแกรมประยุกต์

Pribeanu C. [7] ได้ทำการประเมินผลแรงจูงใจในการใช้งานเทคโนโลยีเสมือนจริงโดยเน้นผู้เรียนเป็นศูนย์กลางในวิชาเคมี จากการทดสอบเก็บข้อมูลเชิงปริมาณและคุณภาพ และทำการประเมินแรงจูงใจที่มีต่อการใช้งานเทคโนโลยีเสมือนจริง พบว่านักเรียนมีความสนใจต่อกระบวนการเรียนในวิชาเคมีเพิ่มขึ้น

Boletsis K. [8] ได้ทำการศึกษาเกมรูปแบบร่วมมือชื่อ “Table Mystery” โดยใช้เทคโนโลยีเสมือนจริงในการเรียนวิชาเคมี โดยมีวัตถุประสงค์เพื่อศึกษาผลกระทบของเทคโนโลยีเสมือนจริงมีผลต่อการสร้างประสบการณ์เรียนและความตื่นเต้น

ทฤษฎีและงานวิจัยที่เกี่ยวข้องข้างต้น ได้ให้แนวคิดที่ทำให้ผู้วิจัยทำการศึกษาเทคโนโลยีเสมือนจริงในทางวิชาเคมี หลายแนวคิดดังนี้ 1) การใช้งานเทคโนโลยีเสมือนจริงในการสร้างสื่อการเรียนรู้ 2) การวัดผลการวิจัยในรูปแบบต่าง ๆ ในงานวิจัยนี้ผู้วิจัยเลือกเป็นการวัดความพึงพอใจ 3) การเน้นผู้เรียนเป็นศูนย์กลางสามารถทำการทดลองใช้งาน เพื่อให้ผู้เรียนเกิดความสนใจในการเรียนยิ่งขึ้น 4) การนำเกมมาประยุกต์ในการจัดการเรียนรู้ ในงานวิจัยนี้ผู้วิจัยออกแบบตารางธาตุในรูปแบบจิ๊กซอว์

3. วิธีดำเนินงานวิจัย

การดำเนินงานวิจัยเรื่อง ตารางธาตุอัจฉริยะ ได้ศึกษาค้นคว้าดำเนินการตามขั้นตอนดังต่อไปนี้

3.1 การออกแบบสื่อการเรียนรู้

มีขั้นตอนหลักที่จำเป็นต้องใช้โปรแกรมประยุกต์ที่เกี่ยวข้องทั้งหมดสี่ขั้นตอนดังนี้

1) เริ่มด้วยการหาข้อมูลและสร้างบทบรรยายเนื้อหาเพื่อการสร้างโมเดลมีความสอดคล้องกัน และทำการบันทึกเสียงบทบรรยาย โดยเนื้อหาในบทบรรยายจะได้รับคำแนะนำจากผู้ทรงคุณวุฒิ ซึ่งเป็นอาจารย์ที่มีคุณวุฒิระดับปริญญาเอก สาขาวิชาเคมีอินทรีย์ ที่มีประสบการณ์สอนในระดับอุดมศึกษาไม่น้อยกว่าห้าปี

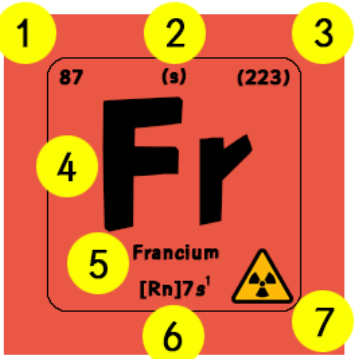
2) ทำการออกแบบมาร์คเกอร์ โดยการออกแบบในโปรแกรมอโดบี โฟโตชอป ซีเอสหก

3) การสร้างโมเดลสามมิติ โดยมีรูปโมเดลเป็นรูปร่างที่เกี่ยวข้องเฉพาะแต่ละธาตุ โดยใช้โปรแกรมออโตเดสก์มายาในการสร้าง

4) นำงานทั้งสามส่วนได้แก่ เสียงบรรยาย มาร์คเกอร์ และโมเดลสามมิติ มารวมการทำงานกัน โดยใช้โปรแกรมยูนิตีและวูฟอเรีย

3.2 การออกแบบและพัฒนามาร์คเกอร์

ผู้วิจัยได้ออกแบบและสร้างมาร์คเกอร์ด้วยโปรแกรมอโดบี โฟโตชอป ซีเอสหก ดังภาพที่ 2



ภาพที่ 2: ส่วนประกอบของมาร์คเกอร์

ภาพที่ 2 แสดงมาร์คเกอร์ที่ข้อมูลบนนั้นเป็นของธาตุแฟรนเซียม โดยส่วนประกอบหลักของมาร์คเกอร์มีอยู่เจ็ดส่วนดังนี้

1) เลขอะตอม ในธาตุแฟรนเซียมมี 87

2) สถานะ ในอุณหภูมิปกติ โดยตัวเอส คือ ของแข็ง

3) เลขมวล น้ำหนักของธาตุ

4) สัญลักษณ์แทนชื่อธาตุ โดยธาตุแฟรนเซียม คือ Fr

5) ชื่อธาตุ แสดงชื่อธาตุของแฟรนเซียม

6) การจัดเรียงอิเล็กตรอน

7) สัญลักษณ์รังสี หรือภาพประกอบแสดงถึงการนำธาตุนั้น ๆ ไปใช้

3.3 การออกแบบและพัฒนารูปแบบสามมิติ

ผู้วิจัยได้ใช้โปรแกรมออโตเดสก์มายาสร้างแบบจำลองสามมิติของธาตุต่างๆ พร้อมทั้งตกแต่งสีให้กับแบบจำลอง เมื่อตกแต่งเสร็จแล้ว จึงบันทึกแบบจำลองเป็นไฟล์ที่มีนามสกุล .fbx เพื่อนำไปใช้งานต่อในโปรแกรมยูนิตี ตัวอย่างโมเดลสามมิติที่พัฒนาขึ้นแสดงดังภาพที่ 3



ภาพที่ 3: การสร้างโมเดลสามมิติที่เพิ่มแมททีเรียล

จากภาพที่ 3 เครื่องมือที่ใช้ในโปรแกรมส่วนใหญ่ คือเครื่องมือโพลีกอน (Polygon) เพราะง่ายต่อการขึ้นรูปแบบจำลองต่างๆ เมื่อขึ้นรูปแบบเสร็จ จะใช้ฟังก์ชัน (Function) เพิ่มแมททีเรียล (Material) เพื่อทำการใส่สีให้กับแบบจำลอง

3.4 การออกแบบและพัฒนาจิ๊กซอว์ตารางธาตุ

ผู้วิจัยได้เลือกโฟมที่มีรูปร่างเป็นจิ๊กซอว์ ขนาดต่อชิ้นคือ 6.7 เซนติเมตรซึ่งเป็นขนาดที่มีความเหมาะสมต่อการตรวจจับมาร์คเกอร์ของกล้องมือถือ เหตุผลในการใช้วัดุที่ทำจากโฟมเนื่องจากมีความยืดหยุ่น ไม่พังบุบสลายโดยง่าย สามารถนำไปประกอบกิจกรรมได้อย่างคงทน จากนั้นได้ทำการพิมพ์มาร์คเกอร์ที่ออกแบบไว้เพื่อนำมาติดบนตัวต่อจิ๊กซอว์โฟม A B C หนึ่งจิ๊กซอว์โฟม ต่อหนึ่งธาตุ ดังภาพที่ 4



ภาพที่ 4: โฟมที่ใช้ในการต่อเป็นตารางธาตุ

3.5 แบบสอบถามวัดความพึงพอใจ

ผู้วิจัยได้จัดทำแบบสอบถามวัดความพึงพอใจ สื่อการเรียนรู้ความเป็นจริงเสริม เรื่อง ตารางธาตุ เพื่อนำไปใช้กับผู้เข้าร่วมวิจัย คือ นักศึกษาครุศาสตรบัณฑิต สาขาวิชาเคมี และนักศึกษาวิทยาศาสตรบัณฑิต สาขาวิชาเคมี มหาวิทยาลัยราชภัฏรำไพพรรณี จังหวัดจันทบุรี จำนวน 30 คน โดยแบบสอบถามแบ่งออกเป็นสองด้าน คือ ด้านเนื้อหา และด้านรูปแบบการใช้งาน

4. ผลการดำเนินงาน

4.1 ส่วนประกอบแผ่นจิ๊กซอว์ตารางธาตุ

เมื่อนำมาร์คเกอร์ชิ้นเล็ก ๆ ของธาตุทุกธาตุมาต่อกันจะเกิดเป็นตารางธาตุที่มีขนาดประมาณ 80×134 เซนติเมตร และส่วนที่เป็นไฟล์ตารางธาตุที่มีขนาดเท่ากัน โดยการสร้างกระดานตารางธาตุที่ประกอบจากจิ๊กซอว์ จำเป็นต้องใช้วัสดุอื่นๆ ประกอบเพิ่มเติมอีก ดังนี้ กรอบของจิ๊กซอว์ตารางธาตุมีสามชั้น โดยชั้นล่างสุดหรือฐาน ใช้วัสดุฟิวเจอร์บอร์ดในการสร้าง เนื่องจากต้องการให้มีน้ำหนักน้อย ชั้นกลางหรือพื้นสำหรับจิ๊กซอว์ใช้โฟมยางในการสร้างเพื่อเพิ่มความยืดหยุ่นและเพิ่มความหนืดเวลาวางจิ๊กซอว์ลงพื้นชั้นนี้จะได้ไม่ลื่นไหลง่าย ในส่วนชั้นบนสุดใช้วัสดุเป็นจิ๊กซอว์โฟมยางเอ บี ซี และใช้กระดาษพิมพ์เป็นรูปภาพในการทำเป็นหน้ามาร์คเกอร์ การพิมพ์โดยใช้กระดาษสำหรับรูปภาพ (Photo Paper) ทำให้ภาพมีสีสันสวยงาม และชัดเจน การใช้สีสันที่แตกต่างกันมีส่วนช่วยในการแยกหมวดหมู่ของธาตุให้ดูง่ายยิ่งขึ้น ตารางธาตุที่สร้างจากการประกอบวัสดุทั้งหมดแสดงดังภาพที่ 5 ซึ่งแสดงถึงส่วนประกอบของจิ๊กซอว์ตารางธาตุ ประกอบไปด้วยจิ๊กซอว์แต่ละธาตุ บริเวณท้ายข้อความ “ตารางธาตุอัจฉริยะ” มีสัญลักษณ์คิวอาร์โค้ด (Quick Response: QR Code) จัดวางอยู่

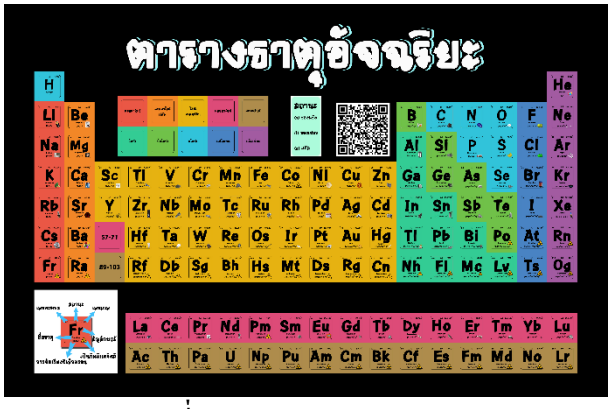
เพื่อให้ผู้ใช้งานได้ส่องเพื่อโหลดไฟล์นามสกุล .apk มาติดตั้งโปรแกรมสำหรับตรวจจับมาร์คเกอร์เพื่อแสดงผลความจริงเสริมบนจิ๊กซอว์แต่ละตัว



ภาพที่ 5: กระดานตารางธาตุที่สร้างจากจิ๊กซอว์ที่นำมาต่อกัน

4.2 ส่วนประกอบไฟล์ตารางธาตุ

ไฟล์ตารางธาตุ สร้างด้วยโปรแกรมอโดบี โฟโตชอป ซีเอสหก เมื่อตกแต่งภาพเสร็จแล้ว จะบันทึกภาพเป็นไฟล์นามสกุล .png โดยไม่มีการบีบอัดความละเอียดของภาพแต่อย่างใด เพื่อให้ภาพมีความคมชัด ทำให้เพิ่มประสิทธิภาพในการอ่านมาร์คเกอร์ได้ดีขึ้น เหตุผลในการสร้างตารางธาตุในรูปแบบของไฟล์ นอกเหนือจากแบบจิ๊กซอว์ เนื่องจากต้องการเพิ่มทางเลือกในกรณีที่ต้องการความสะดวกสบายในกรณีไม่ต้องการพกพาแผ่นจิ๊กซอว์ตารางธาตุ ไฟล์ตารางธาตุที่ออกแบบเสร็จแล้วแสดงดังภาพที่ 6 ซึ่งมีข้อมูลเหมือนข้อมูลในแผ่นจิ๊กซอว์ตารางธาตุทุกประการ



ภาพที่ 6: ตารางธาตุในรูปแบบไฟล์ .png

4.3 ผลการทดสอบประสิทธิภาพการส่องติดมาร์คเกอร์

ผู้วิจัยได้ทำการทดลองเพื่อทดสอบการส่องติดมาร์คเกอร์โดยใช้สมาร์ทโฟนสองยี่ห้อ ได้แก่ เอไอเอส ลาวา แกรนด์ และ วีโว วีไฟว์ ในการทดลองครั้งนี้ โดยผลการทดลองที่ได้แสดงดังตารางที่ 1

ตารางที่ 1: ประสิทธิภาพการส่องติดมาร์คเกอร์

ระยะห่างระหว่างกล้องและมาร์คเกอร์	เปอร์เซนต์การตรวจจับมาร์คเกอร์สำเร็จ	
	เอไอเอส ลาวา แกรนด์	วีโว วีไฟว์
0-4	0	0
5-8	85.9605	86.6384
9-12	21.2994	97.8814
13-16	0	97.6554
17-20	0	100
21-24	0	93.7006
25-28	0	62.2034
29-32	0	39.6893
33-36	0	0

จากตารางที่ 1 เปอร์เซ็นต์การส่องติด คำนวณค่าได้จากการส่องมาร์คเกอร์ 118 ตัว ซึ่งเท่ากับจำนวนธาตุ โดยทำการส่องตัวละ 30 ครั้ง เพื่อลดความคลาดเคลื่อนในการวัดผลการทดลอง ตัวอย่างการคำนวณค่าดังเช่น การใช้มือถือเอไอเอส ลาวา แกรนด์ ส่องมาร์คเกอร์ด้วยระยะ 5-8 เซนติเมตร เป็นดังนี้

$$85.9605 = \frac{3043}{118 \times 30} \times 100$$

โดยที่ 3043 คือจำนวนครั้งที่ส่องติดทั้งหมด 118 คือ จำนวนธาตุทั้งหมดที่ส่องซ้ำธาตุละ 30 ครั้ง การทดลองที่ได้แสดงให้เห็นว่าระยะทำการที่ดีที่สุดในการตรวจจับมาร์คเกอร์อยู่ในช่วง 17-20 เซนติเมตร ที่สามารถส่องติดได้ทุกครั้ง

4.4 ระดับความพึงพอใจของผู้ใช้งาน

งานวิจัยชิ้นนี้ยังได้ศึกษาความพึงพอใจจากผู้ใช้งานสื่อการเรียนเรื่องตารางธาตุ โดยผลที่ได้แสดงดังตารางที่ 2 ซึ่งแสดงให้เห็นผลจากการประเมินความพึงพอใจเฉลี่ยของ

ผู้ใช้งานอยู่ในระดับดีมาก โดยมีค่าเฉลี่ย 4.56 และค่าส่วนเบี่ยงเบนมาตรฐาน 0.21 ตามลำดับ

ตารางที่ 2: ความพึงพอใจของผู้ใช้งาน

รายการประเมิน	ความพึงพอใจ		
	$\bar{X}$	S.D.	ระดับ
ด้านเนื้อหา			
1.สื่อตารางธาตุแสดงเนื้อหาที่เข้าใจง่าย กระชับ ชัดเจน	4.46	0.34	ดี
2.สื่อตารางธาตุมีเนื้อหาที่สอดคล้องกับการเรียนวิชาเคมี	4.63	0.33	ดีมาก
3.ผู้เรียนสามารถนำสื่อตารางธาตุไปใช้ประโยชน์กับการเรียน	4.76	0.25	ดีมาก
4.สื่อตารางธาตุช่วยให้ง่ายต่อการจดจำ	4.40	0.24	ดี
5.เสียงบรรยายประกอบของสื่อตารางธาตุ	4.36	0.43	ดี
ด้านรูปแบบการใช้งาน			
1.สื่อตารางธาตุ (ส่วนจิ๊กซอว์) มีสีสันสวยงาม	4.80	0.16	ดีมาก
2.สื่อตารางธาตุ (ส่วนโมเดลสามมิติ) มีสีสันสวยงาม	4.66	0.29	ดีมาก
3.ตัวอักษรในสื่อตารางธาตุมีความคมชัด	4.63	0.27	ดีมาก
4.สื่อตารางธาตุมีความเหมาะสมต่อการใช้งาน	4.56	0.31	ดีมาก
5.เสียงบรรยายมีความน่าสนใจ	4.33	0.66	ดี
รวม	4.56	0.21	ดีมาก

5. สรุป

ตารางธาตุเป็นเนื้อหาพื้นฐานในการเรียนวิชาเคมีขั้นสูง การมีสื่อการสอนที่ดี เป็นเรื่องที่เป็นประโยชน์ต่อผู้เรียนโดยตรงในการเรียน งานวิจัยชิ้นนี้จึงพัฒนาสื่อการสอนตารางธาตุที่ทำงานร่วมกับเทคโนโลยีความจริงเสริม ที่สามารถทำงานบนสมาร์ทโฟนระบบปฏิบัติการแอนดรอยด์ โดยตารางธาตุที่พัฒนาขึ้นนี้มีอยู่สองรูปแบบ ได้แก่แบบตารางธาตุที่ประกอบขึ้นจากจิ๊กซอว์ที่ผู้สอนสามารถนำไปใช้แยกส่วนเพื่อสนับสนุนการจัดการเรียนรู้แบบร่วมมือ และอีกแบบเป็นไฟล์ตารางธาตุที่เพิ่มความสะดวกให้ผู้ใช้ในกรณีที่ไม่ต้องการพกพาตารางธาตุไปด้วย ผลการทดลองการส่องติดมาร์คเกอร์แสดงให้เห็นว่าระยะที่เหมาะสมที่สุดในการตรวจจับมาร์คเกอร์อยู่ที่ระยะ 17-20 เซนติเมตร และระดับความพึงพอใจของผู้ใช้งานอยู่ในระดับดีมาก หากผู้ใดสนใจข้อมูลตารางธาตุที่ผู้วิจัยพัฒนาขึ้น สามารถติดต่อกับผู้วิจัยได้โดยตรง ผู้วิจัยยินดีเผยแพร่ผลงาน พร้อมให้คำแนะนำในการใช้งานอย่างเต็มที่

เอกสารอ้างอิง

[1] กัญญาณี ใบเนียม และมะยูโซ๊ะ กูโน, การพัฒนาบทเรียนคอมพิวเตอร์ช่วยสอนประเภทเกมการศึกษา เรื่องตารางธาตุสำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 4. *Veridian E-Journal,Silpakorn University*. ปีที่ 8 ฉบับที่ 1 มกราคม – เมษายน พ.ศ.2558 หน้า 109-121.

[2] พลยุทธ พุดตาน และจักกริช พฤษการ, ตารางธาตุเสมือนจริง. ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์. มหาวิทยาลัยเกษตรศาสตร์ วิทยาเขตกำแพงแสน, 2558.

[3] Vinita, B. Shikha, “The retention effect of computer assisted instruction on student’s achievement for teaching the chemistry topic of class VIII student” *International Journal of Scientific and Research Publications*. Vol 5, pp.485-488, 2015.

[4] สุกัญญา. “หลักการทำงานของ Augmented Reality”. Internet: https://sukunya055.wordpress.com. [Jul. 20, 2016].

[5] kroobannok.com. “การเรียนรู้แบบร่วมมือ collaborative learning”. Internet: www.kroobannok.com/25557. [Oct. 22, 2016].

[6] Su Cai, Xu Wang, Feng-Kuang Chiang, “A case study of Augmented Reality simulation system application in a chemistry course” *Comput. Hum. Behav.* Vol. 37, pp.31–40, 2014.

[7] C. Pribeanu, D.D Iordache, “Evaluating the Motivational Value of an Augmented Reality System for Learning Chemistry” *Symposium of the Austrian HCI and Usability Engineering Group (USAB 2008)*, Graz, Austria, November 2008.

[8] C. Boletsis, S. McCallum, “The Table Mystery: An Augmented Reality Collaborative Game for Chemistry Education” *International Conference on Serious Games Development and Applications (SGDA 2013)*, Trondheim, Norway, September 2013.

เกมหนุมานตะลุยกรุงลงกา เพื่อส่งเสริมการศึกษาวรรณคดีไทย
Game Hanuman burnt Longka to promote the study of Thai literature.

ณัชพล สุรรัตนากร (Natchapol Surarattanagon)[1] พัชรี เจริญลาภ (Patcharee Chareonlap)[2]
และนที ยงยุทธ (Nathee Yongyut)[3]
[1,2]สาขาวิชาคอมพิวเตอร์ศึกษา คณะครุศาสตร์ มหาวิทยาลัยราชภัฏรำไพพรรณี จังหวัดจันทบุรี
[3]คณะวิทยาการคอมพิวเตอร์และเทคโนโลยีสารสนเทศ มหาวิทยาลัยราชภัฏรำไพพรรณี จังหวัดจันทบุรี
5615261011@rbru.ac.th, 5615261022@rbru.ac.th, [3]nathee.y@rbru.ac.th

บทคัดย่อ

เกมหนุมานตะลุยกรุงลงกา มีจุดประสงค์ในการพัฒนาเพื่อส่งเสริมการศึกษาวรรณคดีไทยที่ได้ทั้งความรู้ ความสนุกและความท้าทาย เกมประกอบด้วยห้าด่านโดยแต่ละด่านจะบอกเรื่องราวก่อนที่จะให้ผู้เล่นเข้าไปทำภารกิจ เครื่องมือที่ใช้ในการวิจัยประกอบด้วย 1) เกมหนุมานตะลุยลงกา 2) แบบประเมินสำหรับผู้เชี่ยวชาญและ 3) แบบประเมินความพึงพอใจ กลุ่มตัวอย่าง คือ นักศึกษาสาขาวิชาคอมพิวเตอร์ศึกษา จำนวน 33 คน ขั้นตอนการดำเนินงานวิจัย ได้แก่ การศึกษาข้อมูลเกี่ยวกับการสร้างเกมเพื่อการศึกษา การศึกษาวรรณคดีเรื่องรามเกียรติ์ตอน หนุมานเผากรุงลงกา การเขียนสตอรีบอร์ด การอกแบบพัฒนาโมเดลสามมิติเพื่อสร้างเป็นเกมให้สามารถทำงานได้บนระบบปฏิบัติการแอนดรอยด์ แล้วนำไปทดลองใช้กับกลุ่มตัวอย่าง ซึ่งผลการประเมินความพึงพอใจอยู่ในระดับ 4.56 มีค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.55 อยู่ในระดับดีมาก

คำสำคัญ: เกมเพื่อการศึกษา วรรณคดีไทย รามเกียรติ์

Abstract

Game Hanuman burnt Longka the purpose of the development was to promote the study of Thai literature,to gain knowledge fun and challenge. The game consists of five scene, each of which explain the story before entering the mission. Research tools include: 1)Game Hanuman burnt Longka 2)Evaluation form for experts and 3)Satisfaction evaluation form. The sample group consisted of 33 computer education students. The research process was to study information about creating games for educational. Ramayana literature study when Hanuman burnt Longka, storyboard, develop of three-dimensional model to create a game for Android. Then try the sample. The satisfaction rating was 4.56. The standard deviation of 0.55 was very good.

Keyword: Educational Games, Thai literature, Ramayana

1. บทนำ

ในปัจจุบันความสนใจของคนไทยต่อวรรณคดีของชาติมีลดน้อยลงเพราะมีสิ่งที่น่าดึงดูดความสนใจอยู่หลายอย่าง โดยเฉพาะกิจกรรมที่ต้องใช้สมาร์ทโฟนซึ่งกำลังเป็นที่นิยมใช้กันอย่างแพร่หลายและเข้ามามีส่วนร่วมในชีวิตประจำวันของเรามากขึ้นในการทำกิจกรรมต่าง ๆ ไม่ว่าจะเป็นการค้นหาข้อมูล การรับชมข่าวสาร การสื่อสารกับผู้อื่น เนื่องจากมีคุณสมบัติการใช้งานที่หลากหลาย หรือที่เรียกกันว่าแอปพลิเคชันที่มีให้เลือกใช้เป็นจำนวนมาก โดยถ้าหากมองในมุมมองด้านแอปพลิเคชันเพื่อความบันเทิงแล้ว เกมก็เป็นแอปพลิเคชันที่มีจำนวนมากในตลาดแอปพลิเคชันของสมาร์ทโฟน ดังนั้นการใช้เกมเป็นเครื่องมือที่ช่วยแทรกเรื่องราวของวรรณคดีลงไปก็จะเป็นการง่ายที่จะทำให้มีผู้ให้ความสนใจต่อวรรณคดีเพิ่มมากขึ้นอีกทางหนึ่ง

ระบบปฏิบัติการแอนดรอยด์เป็นระบบปฏิบัติการบนสมาร์ทโฟนอีกตัวหนึ่งที่ได้รับความสนใจจากบุคคลทุกระดับชั้นในสังคมเนื่องจากระบบปฏิบัติการแอนดรอยด์เป็นระบบที่มีเสถียรภาพสูงและตอบรับความต้องการของผู้ใช้งานในระดับต่าง ๆ และด้วยคุณสมบัติที่ตอบรับทุกความต้องการใน

สังคมจึงทำให้แอปพลิเคชันบนระบบปฏิบัติการแอนดรอยด์เป็นที่ต้องการของบุคคลในยุคใหม่นี้ ทั้งแอปพลิเคชันที่เกี่ยวกับการทำงาน การเชื่อมต่ออินเทอร์เน็ต รวมไปถึงแอปพลิเคชันที่เกี่ยวกับความบันเทิงในรูปแบบของสื่อมัลติมีเดีย เกม ฯลฯ โดยเฉพาะมัลติมีเดียในรูปแบบของเกมจะได้รับความนิยมจากผู้ใช้ระบบปฏิบัติการแอนดรอยด์ในทุกระดับ เนื่องจากสามารถให้ความเพลิดเพลินระหว่างการเดินทางหรือคลายเครียดในระหว่างการทำงาน ด้วยความหลากหลายของเกมที่ถูกสร้างขึ้นจึงทำให้ผู้ใช้สนุกได้

เกมที่ใช้เล่นบนสมาร์ทโฟนในรูปแบบสามมิติเป็นที่นิยมของผู้เล่นในปัจจุบันเหมาะสำหรับผู้เล่นทุก ๆ วัย ด้วยความสวยของภาพกราฟิคสามารถดึงดูดความสนใจของผู้เล่นได้อย่างดี รูปแบบของเกมที่ผู้วิจัยพัฒนาจะเป็นแนวผจญภัยโดยจะสอดแทรกเนื้อหาสาระไว้ในแต่ละด่าน เพื่อทำให้ผู้เล่นเกิดความรู้สึกอยากติดตามและเข้าใจเรื่องราวที่เกิดขึ้นได้

ดังนั้นผู้วิจัยจึงมีความสนใจสร้างแอปพลิชันเกมในรูปแบบสามมิติที่ทำงานอยู่บนสมาร์ทโฟนระบบปฏิบัติการแอนดรอยด์เพื่อส่งเสริมการศึกษาวรรณคดีไทย ที่ให้ทั้งความบันเทิง สร้างความสนุกสนานและส่งเสริมการเรียนรู้วรรณกรรมของไทยได้เป็นอย่างดี

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 เนื้อเรื่องจากวรรณคดีไทย รามเกียรติ์ ตอน หนุมานเผากรุงลงกา

อัมพร เบื่อขุนทด [1] ได้กล่าวถึงเนื้อหา ตอนหนุมานเผากรุงลงกาว่า เมื่อหนุมานลานางสีดาออกมาข้างนอกแล้วมาเจอกับพวกทหารยักษ์และได้ต่อสู้กัน ยักษ์ที่รอดตายรีบไปทูลสหัสกุมาร เมื่อทราบเรื่องจึงรีบมาดู แล้วบอกหนุมานว่าจะจับตัวไปจึงเกิดการสู้รบระหว่างหนุมานกับสหัสกุมาร หนุมานฆ่าพวกยักษ์และสหัสกุมารตาย ที่เหลือจึงหนีมาทูลทศกัณฐ์ ทศกัณฐ์จึงสั่งให้อินทรชิตไปจัดการ อินทรชิตออกไปสู้รบกับหนุมานเห็นสหัสกุมารตายอยู่กับพื้นจึงสั่งกำลังพลที่มาด้วยเข้าต่อสู้ อินทรชิตเห็นท่าไม่ดีจึงยิงศรนาคบาศกลายเป็นพญานาคมาพันตัวของหนุมาน หนุมานจึงแกล้งยอมให้จับตัวแล้วได้ไปพบกับทศกัณฐ์เห็นอินทรชิตจับหนุมานมาได้ ก็ชื่นชมความเก่งกล้าของอินทรชิต แล้วจึงสั่งให้เอาช้างบรรลัยจักรวาลไปเหยียบหนุมานให้ตาย หนุมานจัดการหักคอช้างบรรลัยจักรวาลจนตาย เมื่อทศกัณฐ์ได้เห็นหน้าหนุมานก็คิดว่า ลิงตัวนี้เก่งกล้า ต้องเลี้ยงไว้เป็นทหารรับใช้ หนุมานได้ฟังก็คิดจะเผาเมืองลงกาจึงออกอุบายว่าให้ใช้ไฟเผาตนเองถึงจะตายทศกัณฐ์หลงเชื่อทำตามเมื่อไฟติดหนุมานก็วิ่งทั่วกรุงลงกา ไฟจึงไหม้เผาเมืองจนวอดวายแล้วหนุมานก็หนีออกจากกรุงลงกา

2.2 ทฤษฎีเกี่ยวกับเกม

เกมสวมบทบาท (Role Playing Game: RPG) คือ เกมที่ผู้เล่นต้องสวมบทบาทเป็นตัวละครต่าง ๆ เข้ามาผจญภัยรูปแบบของการ พัฒนาตัวละครที่เกมสวมบทบาท ทุกเกมจะต้องมีไม่ว่าจะมีในรูปแบบใดก็ตามเมื่อเล่นไปเรื่อย ๆ ตัวละครก็จะพัฒนาความสามารถไปเรื่อย ๆ เรื่องราวภายในเกมจะกล่าวว่าเกมที่จะทำให้เราเข้าใจโลกของเกมนั้น ๆ ได้อย่างชัดเจนเข้าถึงโดยตรงที่สุด หนึ่งในนั้นจะต้องมีเกมสวมบทบาทอยู่เพราะจะลงลึกไปถึงขนาดที่ว่าผู้เล่นจะเป็นผู้ดำเนินเรื่องราวต่างๆ ด้วยตัวเอง คำพูดและบทสนทนาต่าง ๆ จะทำให้ผู้เล่นเข้าใจโลกของเกมมากขึ้น อย่างเช่นการคุยกับ NPC (ตัวละครที่ควบคุมโดยคอมพิวเตอร์) การดำเนินเนื้อเรื่องของบางเกมจะมีภารกิจต่าง ๆ ให้ทำด้วย ซึ่งเป็นการตอกย้ำถึงการสวมบทบาทไปที่ตัวละครนั้น ๆ อย่างชัดเจน

2.3 ภาษาซีชาร์ป (C#)

เนื่องจากผู้วิจัยได้ใช้โปรแกรมยูนิตีในการพัฒนาเกม ดังนั้นจึงมีความจำเป็นที่จะต้องทำความเข้าใจกับการใช้ภาษาซีชาร์ปซึ่งเป็นภาษาโปรแกรมแบบหลายโมเดล ที่ใช้ระบบชนิดข้อมูลแบบรัดกุม (strong typing) และสนับสนุนการเขียนโปรแกรมเชิงคำสั่ง การเขียนโปรแกรมเชิงประกาศ การเขียนโปรแกรมเชิงฟังก์ชัน การเขียนโปรแกรมเชิงกระบวนการ การเขียนโปรแกรมเชิงวัตถุ (แบบคลาส) และการเขียนโปรแกรมเชิงส่วนประกอบ มีรากฐานมาจากภาษาซีพลัสพลัสและภาษาอื่นๆ โดยเฉพาะภาษาจาวา โดยมีจุดมุ่งหมายให้เป็นภาษาสมัยใหม่ที่ไม่ซับซ้อน ใช้งานได้ทั่วไป (general-purpose) และเป็นเชิงวัตถุเป็นหลัก

ผู้วิจัยได้นำเอาภาษาซีชาร์ปมาใช้ในการเขียนโปรแกรมการทำงานของระบบเกมฟังก์ชั่นที่ต้องใช้คือการในการลดเลือดเมื่อมีการต่อสู้กันของตัวละคร เป็นต้น

2.4 หลักการ 3 มิติ

หลักการของการมองภาพให้เกิดภาพเป็น 3 มิติ คือการที่มนุษย์เรามี 2 ตา การที่สามารถมองเห็นภาพต่าง ๆ เป็น 3 มิติได้นั้นเกิดจากมุมมองของสายตาที่เห็นภาพของวัตถุ เมื่อมองวัตถุด้วยตาข้างใดข้างหนึ่งเพียงข้างเดียว จะไม่สามารถมองให้เป็น 3 มิติได้ เพราะจะขาดส่วนลึกของภาพอีกด้านหนึ่งไป

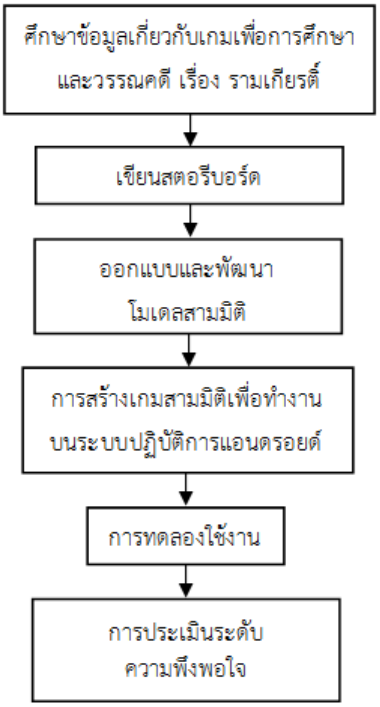
ในการทำเกมหนุมานตะลุยกรุงลงกาครั้งนี้ ผู้วิจัยได้ทำการศึกษาถึงทฤษฎีและงานวิจัยที่เกี่ยวข้องในการนำแนวคิดมาปรับใช้ จากงานวิจัยที่ได้ทำการวิจัยที่เกี่ยวกับการพัฒนาเกมเพื่อศึกษา ดังต่อไปนี้ จรินทร อุ่มไกร [2] ได้ทำการวิจัย เรื่อง การพัฒนาบทเรียนคอมพิวเตอร์ช่วยสอนแบบเกม 3 มิติ ชั้นประถมศึกษาปีที่ 2 โรงเรียนลาดปลาเค้า จังหวัดนครปฐม มีวัตถุประสงค์เพื่อศึกษาผลสัมฤทธิ์ทางการเรียนด้วยบทเรียนคอมพิวเตอร์ช่วยสอนแบบเกม 3 มิติ เพื่อหาประสิทธิภาพของบทเรียนคอมพิวเตอร์ช่วยสอนแบบเกม 3 มิติที่พัฒนาขึ้นและแบบทดสอบก่อนเรียนและหลังเรียน ผลการวิจัยพบว่าบทเรียนคอมพิวเตอร์ช่วยสอนแบบเกม 3 มิติสามารถนำไปใช้ในวิชาคอมพิวเตอร์ สำหรับนักเรียนชั้นประถมศึกษาปีที่ 2 ได้จริงและทำให้นักเรียนเกิดความสนใจในวิชาที่เรียน อีกงานวิจัยเป็นของ สุวิชัย พรรษา [3] ได้ทำการวิจัยเรื่อง การพัฒนาเกมคอมพิวเตอร์ 3 มิติ “เกมโซล่าผจญภัย” เพื่อพัฒนาเกมคอมพิวเตอร์ 3 มิติ และศึกษาความพึงพอใจในการเล่นเกมโดยภายในเกมจะให้ความรู้เกี่ยวกับยาเสพติด 7 ชนิด ได้แก่ เฮโรอีน สารระเหย กัญชา ยาบ้า ฝิ่น บุหรี่ และแอลกอฮอล์ ผลการวิจัยพบว่า การนำเกมไปทดลองเล่นเกมได้ผลสรุป การดำเนินเนื้อเรื่อง ตัวละคร รูปแบบการเล่น ความคิดสร้างสรรค์อยู่ในระดับดี เด็กและเยาวชนในกลุ่มตัวอย่างได้มีความพึงพอใจอยู่ในระดับดี เพราะได้รับความรู้ที่เกี่ยวกับยาเสพติดจากการเล่นเกม และกัญญาณี ใบเนียม [4] ได้ทำการวิจัยเรื่อง การพัฒนาบทเรียนคอมพิวเตอร์ช่วยสอนประเภทเกมการศึกษา เรื่อง ตารางธาตุ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 4 อีกทั้งได้ศึกษาการออกแบบโมเดลสามมิติจากตัวอย่างงานวิจัยของ อนุสรณ์ บรรเทิง และคณะ [5] เพื่อใช้เป็นแนวทางในการพัฒนาเกม และงานวิจัยของกานต์ธิดา นานอนและคณะ [6] เพื่อเป็นแนวทางในการสร้างเกมส่งเสริมการเรียนรู้ จากการศึกษางานวิจัยเกี่ยวกับเกมคอมพิวเตอร์ที่อยู่ในรูปแบบ 3 มิติ เป็นสิ่งที่มีความน่าสนใจสามารถนำมาใช้ได้จริงเหมาะแก่การส่งเสริมการเรียนรู้ ผู้วิจัยจึงนำมาใช้ในการพัฒนเกมหนุมานตะลุยกรุงลงกา เพื่อส่งเสริมการศึกษาวรรณคดีไทย

3. วิธีการดำเนินงาน

ในการพัฒนาเกมเพื่อสนับสนุนการศึกษาดังกล่าว มีขั้นตอนดังต่อไปนี้

3.1 การศึกษาข้อมูล

โดยทำการศึกษาข้อมูลเกี่ยวกับการการพัฒนาเกมเพื่อการเรียนรู้โดยสิ่งสำคัญที่สุดของการสร้างเกมเพื่อการเรียนรู้ [7] คือ การพยายามสร้างการเรียนรู้ด้วยตนเองให้กับผู้เรียน และศึกษาเรื่องราวของวรรณคดีไทยและรวบรวมข้อมูลที่ได้ไปให้ผู้เชี่ยวชาญตรวจสอบเนื้อหาและภาพที่นำมาใช้



ภาพที่ 1: ขั้นตอนการดำเนินงานวิจัย

3.2 การออกแบบและพัฒนา

แบ่งออกเป็น 3 ส่วน คือ 1) การเขียนสตอรีบอร์ด 2) การออกแบบและพัฒนารูปแบบสามมิติ 3) การออกแบบเกมบนระบบปฏิบัติการแอนดรอยด์

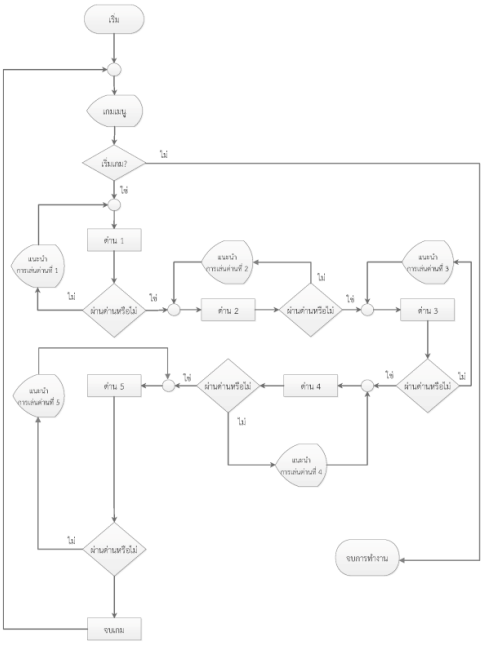
3.2.1 การเขียนสตอรีบอร์ด

ผู้วิจัยได้เขียนสตอรีบอร์ดของเกมหนุมานตะลุยกรุงลงกานี้ เป็นเนื้อเรื่องของหนุมานทำภารกิจต่าง ๆ ในเนื้อเรื่องของเกมเริ่มตั้งแต่หนุมานเมื่อทูลลานางสีดาแล้ว คิดว่าเมื่อเข้า

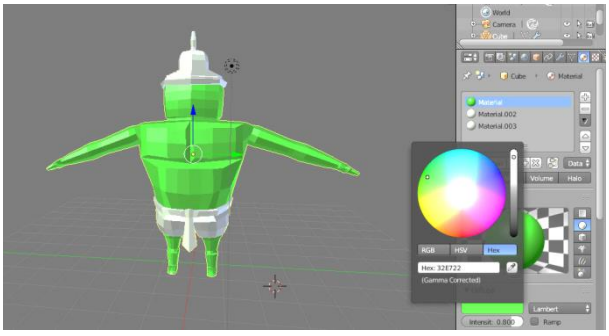
มาเหยียบกรุงลงกาแล้ว ก็ควรจะได้ลองฤทธิ์กับเหล่ายักษ์ จึงแผลงฤทธิ์จนต้องโดนจับไปเข้าพบกับทศกัณฐ์ จนสิ้นสุดที่หนุมานเผากรุงลงกาไหม้จนวอดวาย [8]

3.2.2 การออกแบบและพัฒนารูปแบบสามมิติ

ผู้วิจัยใช้โปรแกรมคอมพิวเตอร์เพื่อขึ้นรูปแบบโมเดลสามมิติแล้วส่งออกไฟล์ไปพัฒนาต่อเป็นรูปแบบสามมิติที่มีลักษณะที่สมบูรณ์มากขึ้นด้วยโปรแกรมเบลนเดอร์ (Blender) เพราะเป็นโปรแกรมที่มีความโดดเด่นในการสร้างภาพสามมิติและการสร้างโมเดลได้อย่างรวดเร็ว สร้างความสมจริงให้กับวัตถุ และใช้ทรัพยากรของระบบน้อย สร้างงานได้หลากหลายรูปแบบ



ภาพที่ 2: ขั้นตอนการทำงานของเกม

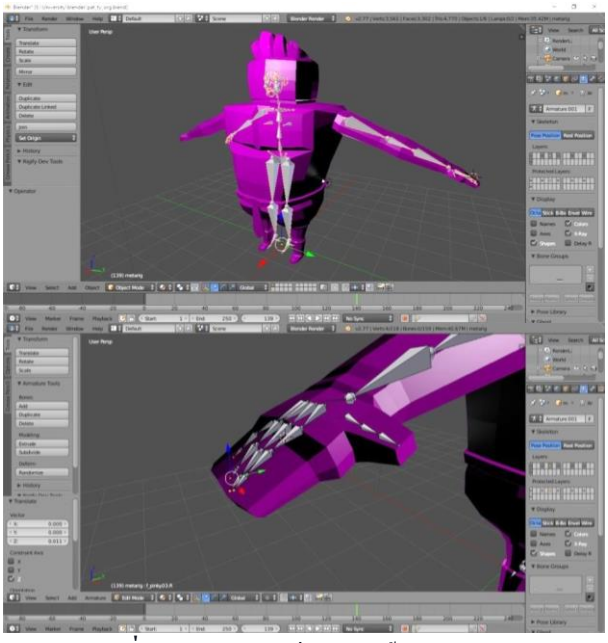


ภาพที่ 3: การออกแบบโมเดลด้วยโปรแกรมเบลนเดอร์

3.2.3 การออกแบบเกมบนระบบปฏิบัติการแอนดรอยด์

1) ทำการนำออกไฟล์โมเดลจากโปรแกรมเบลนเดอร์ (Blender) เป็นนามสกุล เอฟบีเอ็กซ์ (fbx) ในการกำหนดค่า ให้ทำการเลือกการฝังสื่อ (Embed media) ด้วย

2) เปิดหรือสร้างไฟล์งานภายในโปรแกรมยูนิตี ทำการนำเข้าโมเดลที่ทำการนำออกจากเบลนเดอร์ หากไฟล์โมเดลนั้นเป็นตัวละครที่จะต้องมีการควบคุมท่าทาง เช่น การเดิน การวิ่ง การกระโดด การต่อสู้ เป็นต้น จะต้องใส่กระดูกให้กับตัวโมเดลให้ทำการเลือกประเภทโมเดลเป็นรูปแบบจำลองมนุษย์ (Humanoid) แล้วนำไปกำหนดค่าของตัวโมเดล ตัวอย่างของการใส่กระดูกที่บริเวณแขนและนิ้วมือดังภาพที่ 3 ของโมเดลเป็นจุดที่สามารถกำหนดลักษณะท่าทางการเคลื่อนไหวได้ จำนวนของกระดูกมีผลต่อความละเอียดของการควบคุมการเคลื่อนไหวของโมเดล แต่อย่างไรก็ตามต้องอ้างอิงกับต้นแบบของโมเดลที่ได้ออกแบบไว้เพื่อให้โมเดลที่สร้างขึ้นมาแต่ละตัวมีความสอดคล้องกัน มีรูปแบบเป็นไปในทางเดียวกันด้วย



ภาพที่ 4: การใส่กระดูกที่แขนและนิ้วมือของโมเดล

3) การควบคุมตัวละคร เช่น การเดิน การกระโดด เป็นต้น

การควบคุมตัวละครจะมีการใช้ตัวกำหนดการเล่นท่าทางการเคลื่อนไหว และควบคุมทำทางการเคลื่อนไหว ซึ่งทำงานร่วมกันกับสคริปต์ที่เขียนขึ้น ซึ่งในงานวิจัยนี้ได้ใช้ตัวกำหนดการการเล่นท่าทางการเคลื่อนไหวและท่าทางการเคลื่อนไหวของยูนิตี ซึ่งให้ใช้ได้ฟรีพร้อมทั้งสคริปต์ของยูนิตีในบางส่วน การควบคุมตัวละครงานวิจัยนี้ใช้ตัวสคริปต์ของทางยูนิตีและเขียนขึ้นเองในส่วนของโหมดต่อสู้ โดยใช้ภาษาซีชาร์ปตัวละครหรือตัวโมเดลนั้นจะต้องทำการใส่วัตถุแข็งเกร็ง (Rigid Body) เพื่อให้ตัวโมเดลนั้น

ไม่สามารถทะลุวัตถุต่าง ๆ ได้ จากนั้นจึงนำไฟล์ออกมาเป็นนามสกุลจุดเอพีเค (.apk) เพื่อนำไปติดตั้งบนระบบปฏิบัติการแอนดรอยด์ได้ ผู้วิจัยได้ทำการติดตั้งตัวเกมบนมือถือสมาร์ทโฟนที่เป็นระบบ ปฏิบัติการแอนดรอยด์ เวอร์ชั่น 4.2 ขึ้นไป

3.2.4 นำเกมหนุมานตะลุยกรุงลงกาไปให้ผู้เชี่ยวชาญประเมินด้านเนื้อหาและเทคนิคการสร้าง จำนวน 3 ท่าน เพื่อให้เกิดความถูกต้องเหมาะสม

3.2.5 นำเกมหนุมานตะลุยกรุงลงกาไปทดลองใช้

3.3 กลุ่มเป้าหมายในการวิจัย

นักศึกษาสาขาคอมพิวเตอร์ศึกษา คณะครุศาสตร์ มหาวิทยาลัยราชภัฏรำไพพรรณี จ.จันทบุรี จำนวน 33 คน

3.4 เครื่องมือที่ใช้ในการวิจัยประกอบด้วย

3.4.1 เกมหนุมานตะลุยกรุงลงกา

3.4.2 แบบประเมินด้านสำหรับผู้เชี่ยวชาญ

3.4.3 ประเมินความพึงพอใจ

4. ผลการดำเนินงาน

4.1 การแสดงผลออกมาเป็นเกมสามมิติ

เกมหนุมานตะลุยกรุงลงกา ประกอบไปด้วยโมเดล 3 มิติ ที่เป็นตัวละครหลักตามเนื้อเรื่องของเกม ซึ่งประกอบไปด้วย หนุมาน ทหารยักษ์ สหัสกุมาร อินทรชิต และทศกัณฑ์ เมื่อเข้าสู่เกมหน้าแรกดังภาพที่ 5 จะพบกับปุ่มเมนูให้เลือกเมื่อกดเริ่มเกมจะพบกับหน้าต่างการบรรยายเนื้อเรื่องของเกม ในรูปแบบของภาพอินโฟกราฟิคทำให้ผู้เล่นสามารถที่จะเข้าใจเรื่องราวได้อย่างรวดเร็ว ดังภาพที่ 6 พร้อมกับมีบทสนทนาก่อนที่จะเริ่มเล่นเกม ดังภาพที่ 7 ผู้เล่นต้องบังคับหนุมานให้ทำภารกิจต่าง ๆ โดยสัมผัสปุ่มจอยสติ๊กที่หน้าจอแสดงโดยใช้นิ้วหัวแม่มือซ้ายกดค้างไว้เพื่อบังคับทิศทางและใช้นิ้วหัวแม่มือขวาเพื่อกดปุ่มโจมตีตัวละครฝ่ายตรงข้าม ให้ผ่านภารกิจที่จะต้องต่อต่อสู้กับเหล่าสมุนของทหารยักษ์เข้าไปยังกรุงลงกา ดังภาพที่ 8



ภาพที่ 5: หน้าแรกของเกม



ภาพที่ 6: หน้าการบรรยายเนื้อเรื่องก่อนเริ่มเล่นเกม



ภาพที่ 7: ฉากก่อนเข้าสู่ด่านแรกของเกม

ด่านของเกมหนุมานตะลุยกรุงลงกา มีดังต่อไปนี้

ด่านที่ 1 หนุมานแผลงฤทธิ์ ผู้เล่นจะได้ควบคุมหนุมานเพื่อต่อสู้กับไพร่พลทหารยักษ์ในกรุงลงกา ทำให้เข้าใจเรื่องราวที่เป็นสาเหตุให้เกิดเหตุการณ์เผาลงกาในที่สุด

ด่านที่ 2 ต่อสู้กับสหัสกุมาร ด่านนี้เป็นการนำเสนอตัวละคร คือ สหัสกุมารซึ่งเป็นลูกชายของทศกัณฑ์ที่ต้องการจะจับตัวหนุมานหลังจากเหล่าทหารยักษ์ที่รอดตายมาเข้าทูล ผู้เล่นมีภารกิจต้องต่อสู้กับสหัสกุมารเพื่อผ่านไปยังด่านต่อไป

ด่านที่ 3 อินทรชิตยิงลูกศรนาคบาศที่กลายเป็นพญานาคมาพันรอบตัวหนุมาน และหนุมานได้ยอมให้จับตัวไปพบกับทศกัณฑ์

ด่านที่ 4 ต่อสู้กับช้างบรรลัยจักรวาล ที่จะเข้ามาเหยียบหนุมานให้ตาย แต่สุดท้ายช้างก็ถูกหนุมานฆ่าตาย ในด่านนี้ผู้เล่นจะได้ควบคุมหนุมานให้ต่อสู้กับช้างบรรลัยจักรวาลที่มีพละกำลังมาก เพื่อผ่านไปยังด่านสุดท้ายให้ได้

ด่านที่ 5 หนุมานเผากรุงลงกา ภายหลังจากหนุมานออกอุบายให้ใช้ไฟเผาตนเองถึงจะตายและทศกัณฑ์หลงเชื่อทำให้หนุมานวิ่งกระโดดเผารอบเมืองไหม้ทั้งเมืองในที่สุด ในด่านสุดท้ายนี้ผู้เล่นต้องควบคุมหนุมานให้เข้าไปเผาเมืองตามจุดต่างๆ โดยรอบให้สำเร็จ เมื่อทำได้ครบตามที่กำหนดแล้วจะสิ้นสุดเกม



ภาพที่ 8: ด่านที่ 5 หนุมานเผากรุงลงกา

4.2 ผลการประเมินด้านเนื้อหาและการทำงานของระบบจากผู้เชี่ยวชาญ

ตารางที่ 1: ผลจากการประเมินทั้ง 3 ด้าน โดยผู้เชี่ยวชาญ

รายการประเมิน	ค่าเฉลี่ย	ค่าเบี่ยงเบนมาตรฐาน	ระดับคุณภาพ
1. ด้านภาพรวมของเกม	4.00	0.70	ดี
2. ด้านเนื้อหามีสอดคล้อง	4.50	0.50	ดี
3. ด้านการทำงานของระบบ	3.50	0.50	ปานกลาง
รวม	**4.07**	0.50	ดี

จากผลการประเมินในด้านการทำงานของระบบอยู่ในระดับปานกลางเนื่องจากว่าในขั้นตอนของการออกแบบโมเดลและการสร้างด่านแต่ละด่านมีการสร้างวัตถุขึ้นมาจำนวนมาก ทำให้ต้องมีการประมวลผลมากขึ้นส่งผลให้การทำงานของเกมจะช้าลง การแก้ไขโดยลดจำนวนวัตถุในจุดที่เป็นส่วนประกอบของฉากการแสดงผลให้น้อยลง ผลปรากฏว่าทำให้ทำงานของเกมโดยรวมเร็วขึ้น

4.3 ผลการประเมินระดับความพึงพอใจ

ระดับความพึงพอใจต่อการใช้งานของกลุ่มผู้ร่วมทดลองโดยแบ่งออกเป็น 3 ด้าน คือ 1) ด้านภาพรวม 2) ด้านข้อมูล 3) ด้านการทำงานของระบบ

ตารางที่ 2: ค่าเฉลี่ย ค่าเบี่ยงเบนมาตรฐาน และระดับความพึงพอใจของผู้เล่นที่มีต่อเกม

รายการประเมิน	ค่าเฉลี่ย	ค่าเบี่ยงเบนมาตรฐาน	ระดับ
1. ด้านภาพรวมของเกม	4.59	0.50	ดีมาก
2. ด้านเนื้อหาที่ได้รับ	4.86	0.34	ดีมาก
3. ด้านการทำงานของระบบ	3.87	0.47	ดี
รวม	4.56	0.55	ดีมาก

จากตารางที่ 2 ค่าเฉลี่ยที่ได้จากการประเมินทั้ง 3 ด้าน โดยนักศึกษากลุ่มตัวอย่างมีค่า 4.56 และค่าส่วนเบี่ยงเบนมาตรฐานมีค่า 0.55 ดังนั้นเกมหนุมานตะลุยลงกา 3 มิติ บนระบบปฏิบัติการแอนดรอยด์ ที่ได้จัดทำขึ้นมีความพึงพอใจในภาพรวมอยู่ในระดับ ดีมาก

5. สรุปและอภิปรายผล

การพัฒนาเกมหนุมานตะลุยกรุงลงกาเป็นการนำเทคโนโลยีมาประยุกต์เพื่อพัฒนาเป็นสื่อส่งเสริมการเรียนรู้ที่สามารถใช้งานบนสมาร์ทโฟนทำให้เกิดความสะดวกต่อการใช้งานจากผลการทดลองสภาพการใช้งานจริงทำให้เห็นได้ว่าเมื่อผู้ใช้เล่นเกมแล้วสามารถทำให้มีความเข้าใจในเรื่องราวของวรรณคดีไทยมากขึ้นพร้อมกับได้รับความสนุกสนานกับการเล่นเกม ส่งผลให้ผลการประเมินความพึงพอใจอยู่ในระดับที่สูงมาก ดังนั้นจึงควรส่งเสริมให้มีการพัฒนาเกมในรูปแบบสามมิติเพื่อเป็นสื่อในการเรียนรู้ให้มีเพิ่มมากขึ้น

เอกสารอ้างอิง

[1] อัมพร เบื่อขุนทด,*หนุมานเผากรุงลงกา*,มูลนิธิเด็ก: กรุงเทพฯ,2552.

[2] จรินทร อุ่มไกร, “การพัฒนาบทเรียนคอมพิวเตอร์ช่วยสอนแบบเกม 3 มิติ ในรายวิชาคอมพิวเตอร์ชั้นประถมศึกษาปีที่ 2 กรณีศึกษาโรงเรียนวัดลาดปลาเค้า” *คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏนครปฐม* : นครปฐม, 2557.

[3] สุวิชัย พรรษา, “การพัฒนาเกมคอมพิวเตอร์ 3 มิติ เกมโซล่าผจญภัย” *สาขาวิชาสื่อนฤมิต คณะวิทยาการสารสนเทศ มหาวิทยาลัยมหาสารคาม*:มหาสารคาม,2558.

[4] กัญญาณี ใบเนียม, “การพัฒนาบทเรียนคอมพิวเตอร์ช่วยสอนประเภทเกมการศึกษาเรื่องตารางธาตุ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 4” *มหาวิทยาลัยศิลปากร* : กรุงเทพฯ, 2558.

[5] อนุสรณ์ บรรเทิง และคณะ, “การพัฒนาและปรับปรุงประสิทธิภาพเกม 3 มิติเพื่อจำลองสภาพแวดล้อมตึกวิจัยคณะวิทยาศาสตร์ มหาวิทยาลัยอุบลราชธานี” *การประชุมวิชาการระดับชาติด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศครั้งที่ 12 (NCCIT 2016)*, ขอนแก่น, 6-7 กรกฎาคม 2559.

[6] กานต์ธิดา นานอน และคณะ, “กรอบแนวคิดสำหรับซีเรียสเกมเพื่อส่งเสริมการเรียนรู้ในพิพิธภัณฑ์”*การประชุมวิชาการระดับชาติด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศครั้งที่ 12 (NCCIT 2016)*, ขอนแก่น, 6-7 กรกฎาคม 2559.

[7] E. Klopfer, S. Osterweil, and K. Salen, “ Moving learning games forward,” The Education Arcade Massachusetts Institute of Technology, USA, 2009.

[8] กรกช เพ็งสมบูรณ์ และคณะ, “โปรแกรมประยุกต์บนโทรศัพท์เคลื่อนที่เพื่อนำชมเรื่องรามเกียรติ์ บนกำแพงวัดพระแก้ว”*การประชุมวิชาการระดับชาติด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศ ครั้งที่ 10 (NCCIT 2014)*, ภูเก็ต, 8-9 พฤษภาคม 2557.

การออกแบบระบบการเรียนการสอนตามโมเดลดิกและคาเร่ย์โดยใช้การ์ตูนมัลติมีเดีย 2 มิติ ร่วมกับคลาวด์เทคโนโลยี ในรายวิชาการวิเคราะห์และออกแบบระบบสารสนเทศ สำหรับนักศึกษาวิชาชีพครู สาขาวิชาคอมพิวเตอร์ศึกษา มหาวิทยาลัยราชภัฏนครปฐม

The Design of Instructional Systems Based on Dick and Carey Model. By using 2D via Cloud Computing in the Course of the Information Systems Analysis and Design. For Teaching Profession. Computer Education Major. Nakorn Pathom Rajaphat University.

จรินทร อุ่มไกร (Charinthorn Aumgri)[1] ณพงศ์ วรรณพิรุณ (Naphong Wannapiroon)[2]
วาสนา แย้มพราย (Wassana Yaemprai)[3] ปาวีณา ฉ่ำกิ่ง (Paveena Chamking)[4]
และสิวาลัย จินเจือ (Siwalai Chinchua)[5]
[1]สาขาวิชาคอมพิวเตอร์ศึกษา คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏนครปฐม
[2]สาขาวิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยเทคโนโลยีราชมงคลสุวรรณภูมิ
[3]สาขาวิชาคอมพิวเตอร์ศึกษา คณะครุศาสตร์ มหาวิทยาลัยราชภัฏภูเก็ต
[4]สาขาวิชาคอมพิวเตอร์ศึกษา คณะครุศาสตร์ มหาวิทยาลัยราชภัฏศรีสะเกษ
[5]ภาควิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีและการจัดการอุตสาหกรรม
มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ วิทยาเขตปราจีนบุรี
[1]aor030@hotmail.com.com

บทคัดย่อ

งานวิจัยนี้เป็นการวิจัยเชิงทดลอง ซึ่งมีวัตถุประสงค์ดังนี้ 1) เพื่อนำเสนอแนวการออกแบบระบบการเรียนการสอนตามโมเดลดิกและคาเร่ย์โดยใช้การ์ตูนมัลติมีเดีย 2 มิติ ร่วมกับคลาวด์เทคโนโลยี ในรายวิชาการวิเคราะห์และออกแบบระบบสารสนเทศ สำหรับนักศึกษาชั้นปีที่ 3 สาขาวิชาคอมพิวเตอร์ศึกษา มหาวิทยาลัยราชภัฏนครปฐม 2) เพื่อศึกษาความพึงพอใจของผู้เรียนที่มีต่อการออกแบบระบบการเรียนการสอนตามโมเดลดิกและคาเร่ย์โดยใช้การ์ตูนมัลติมีเดีย 2 มิติ ร่วมกับคลาวด์เทคโนโลยี ในภาคเรียนที่ 2/2559 ได้มาด้วยวิธีการเลือกแบบเจาะจงจำนวน 87 คนจากนักศึกษาหมู่เรียน 57/12 จำนวน 42 คน และหมู่เรียน 57/13 จำนวน 45 คน เครื่องมือที่ใช้ บทเรียนการ์ตูนมัลติมีเดีย 2 มิติ ร่วมกับคลาวด์เทคโนโลยี, เอกสารประกอบการสอน และแบบสอบถามความพึงพอใจ

ผลการวิจัยพบว่า ผลจากการประเมินความพึงพอใจของผู้เรียนที่มีต่อการออกแบบระบบการเรียนการสอนตามโมเดลดิกและคาเร่ย์โดยใช้การ์ตูนมัลติมีเดีย 2 มิติ ร่วมกับคลาวด์เทคโนโลยี นั้นอยู่ในระดับดีมาก โดยมีค่าเฉลี่ยรวมทั้งหมด 4.71 และมีค่าส่วนเบี่ยงเบนมาตรฐาน 0.35 ดังนั้นสามารถสรุปได้ว่า การออกแบบระบบการเรียนการสอนตามโมเดลดิกและคาเร่ย์โดยใช้การ์ตูนมัลติมีเดีย 2 มิติ ร่วมกับคลาวด์เทคโนโลยี ในรายวิชาการวิเคราะห์และออกแบบระบบสารสนเทศ สำหรับนักศึกษาชั้นปีที่ 3 สาขาวิชาคอมพิวเตอร์ศึกษา มหาวิทยาลัยราชภัฏนครปฐมที่พัฒนาขึ้นสามารถนำไปใช้ร่วมกับการเรียนการสอนในรายวิชาเพื่อช่วยในการแก้ปัญหาได้ในระดับดีมาก

คำสำคัญ : โมเดลดิกและคาเร่ย์ คลาวด์เทคโนโลยี การวิเคราะห์และออกแบบระบบสารสนเทศ

Abstract

This research was an experimental research. The purposes of this study were 1) to design of instructional systems based on Dick and Carey model. By using 2D via cloud computing in the course of the information systems analysis and design. For teaching profession. computer education major. Nakorn Pathom Rajaphat University and 2) to study satisfaction of the developed instructional systems based on Dick and Carey model. By using 2D via cloud computing in the course of the information systems analysis and design . The sample, selected by purposive sampling, were 87 students who studied in academic year 2016 . The research tools consisted of the developed instructional systems based on Dick and Carey model. By using 2D via cloud computing in the course of the information systems analysis and design, teaching Documents and satisfaction .

The results showed that the evaluation of the satisfaction is excellent level and the standard deviation deviation is 4.71, and the average 0.35, we can use this CAI on Tablet to the target group.

Keyword: Dick and Carey Model, Cloud Computing, The Information Systems Analysis and Design.

1. บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

จากแผนพัฒนาดิจิทัลไทยแลนด์ (Digital Thailand) เรื่อง การแก้ไขปัญหาความเหลื่อมล้ำในสังคม ซึ่งมีหลากหลายมิติ ทั้งด้านการพัฒนาคุณภาพคน ด้านการศึกษา ด้านรายได้ ด้านโอกาสทางสังคมและการได้รับสิทธิประโยชน์ต่างๆ รวมถึงบริการของภาครัฐ และยังรวมถึงความเหลื่อมล้ำทางดิจิทัล (digital divide) หรือความแตกต่างและช่องว่างระหว่างผู้ที่สามารถเข้าถึงและใช้ประโยชน์จากเทคโนโลยีดิจิทัลกับผู้ที่เข้าไม่ถึง ไม่เข้าใจ และไม่สามารถใช้ประโยชน์จากไอซีทีได้[1]

การพัฒนานวัตกรรมต้นแบบที่เกี่ยวกับการเรียนการสอนนั้นถือเป็นแบบอย่างให้กับผู้เรียนในสาขาวิชาชีพครูได้ซึ่งจะเป็นแบบอย่างเบื้องต้นและสร้างสรรค์ความคิดให้กับผู้เรียนได้ความรู้สืบต่อไปได้จากการได้ทดลองใช้และปฏิบัติได้เองทำให้ผู้เรียนสามารถนำถ่ายทอดได้ซึ่งสามารถนำไปใช้ในชีวิตจริงได้ซึ่งตรงกับการศึกษาในศตวรรษที่ 21

จากการจัดการเรียนการสอนในชั้นเรียนของรายวิชาการวิเคราะห์และออกแบบระบบสารสนเทศนั้น ผู้วิจัยได้มีการสอนในรายวิชานี้มาแล้วจำนวน 3 ปี พบว่า การจัดการเรียนการสอนในชั้นเรียนนั้นผู้เรียนไม่สามารถวิเคราะห์และสังเคราะห์การถ่ายทอดเนื้อหาให้สามารถผสานกับการจัดการเรียนการสอนทางด้านคอมพิวเตอร์ศึกษาได้ซึ่งจำเป็นต้องใช้สื่อทางดิจิทัลรวมกับการจัดการศึกษาที่ตรงกับศตวรรษที่ 21

ดังนั้นผู้วิจัยจึงได้ออกแบบระบบการเรียนการสอนตามโมเดลดิกและคาเร่ย์โดยใช้การ์ตูนมัลติมีเดีย 2 มิติ ร่วมกับคลาวด์เทคโนโลยี ในรายวิชาการวิเคราะห์และออกแบบระบบสารสนเทศ สำหรับนักศึกษาชั้นปีที่ 3 สาขาวิชาคอมพิวเตอร์ศึกษา มหาวิทยาลัยราชภัฏนครปฐม สืบเนื่องจากในรายวิชาการวิเคราะห์และออกแบบระบบสารสนเทศนั้นผู้เรียนจะไม่สามารถดึงความคิดไปผนวกกับการพัฒนาการเรียนการสอนตามโมเดลทางการศึกษาได้ ดังนั้นผู้วิจัยจึงนำโมเดลต้นแบบนี้เมื่อมีการมาทดลองใช้นักศึกษาวิชาชีพครู ชั้นปีที่ 3 สาขาวิชาคอมพิวเตอร์ศึกษา เพื่อเป็นโมเดลต้นแบบในการจัดการเรียนการสอนในรายวิชาต่อไป

1.2 วัตถุประสงค์ของการวิจัย

1.2.1 เพื่อพัฒนาการออกแบบระบบการเรียนการสอนตามโมเดลดิกและคาเร่ย์โดยใช้การ์ตูนมัลติมีเดีย 2 มิติ ร่วมกับคลาวด์เทคโนโลยี ในรายวิชาการวิเคราะห์และออกแบบระบบสารสนเทศ สำหรับนักศึกษาชั้นปีที่ 3 สาขาวิชาคอมพิวเตอร์ศึกษา มหาวิทยาลัยราชภัฏนครปฐม

1.2.2 เพื่อศึกษาความพึงพอใจของผู้เรียนที่มีการออกแบบระบบการเรียนการสอนตามโมเดลดิกและคาเร่ย์โดยใช้การ์ตูนมัลติมีเดีย 2 มิติ ร่วมกับคลาวด์เทคโนโลยี ในรายวิชาการวิเคราะห์และออกแบบระบบสารสนเทศ สำหรับนักศึกษาชั้นปีที่ 3 สาขาวิชาคอมพิวเตอร์ศึกษา มหาวิทยาลัยราชภัฏนครปฐม

1.3 สมมติฐานของการวิจัย

1.3.1 ความพึงพอใจของผู้เรียนที่มีต่อการออกแบบระบบการเรียนการสอนตามโมเดลดิกและคาเร่ย์โดยใช้การ์ตูนมัลติมีเดีย 2 มิติ ร่วมกับคลาวด์เทคโนโลยี ในรายวิชาการ

วิเคราะห์และออกแบบระบบสารสนเทศ สำหรับนักศึกษาชั้นปีที่ 3 สาขาวิชาคอมพิวเตอร์ศึกษา มหาวิทยาลัยราชภัฏนครปฐม อยู่ในระดับดี

2. วรรณกรรมที่เกี่ยวข้อง

2.1 รูปแบบการสอนของดิคค์และคาเรย์ (Dick and Carey model) [2]

รูปแบบการสอน (Model) ของดิคและคาเรย์ ประกอบด้วยองค์ประกอบด้วย 10 ขั้นด้วยกัน คือ

1. การกำหนดเป้าหมายของการเรียนการสอน (Identify Instructional Goals)
2. ดำเนินการวิเคราะห์การเรียนการสอน (Conduct Instructional Analysis)
3. กำหนดพฤติกรรมก่อนเรียนและลักษณะผู้เรียน (Identify Entry Behaviors, Characteristics)
4. เขียนจุดมุ่งหมายเชิงพฤติกรรม (Write Performance Objective)
5. พัฒนาข้อสอบอิงเกณฑ์ (Develop Criterion - Referenced Test Items)
6. พัฒนายุทธวิธีการสอน (Develop Instructional Strategies)
7. พัฒนาและเลือกวัสดุการเรียนการสอน (Develop and Select Instructional Materials)
8. ออกแบบและดำเนินการประเมินเพื่อการปรับปรุง (Design and Conduct Formative Evaluation)
9. การปรับปรุงการสอน (Revise Instruction)
10. การออกแบบและดำเนินการประเมินระบบการสอน (Design and Conduct Summative E valuation)

2.2 คลาวด์เทคโนโลยี

Cloud Computing คือ บริการที่ครอบคลุมถึงการให้ใช้กำลังประมวลผล หน่วยจัดเก็บข้อมูล และระบบออนไลน์ต่างๆ จากผู้ให้บริการ เพื่อลดความยุ่งยากในการติดตั้ง ดูแลระบบ ช่วยประหยัดเวลา และลดต้นทุนในการสร้างระบบคอมพิวเตอร์และเครือข่ายเอง ซึ่งก็มีทั้งแบบบริการฟรีและแบบเก็บเงิน [3]

2.3 รายวิชาการวิเคราะห์และออกแบบระบบสารสนเทศ

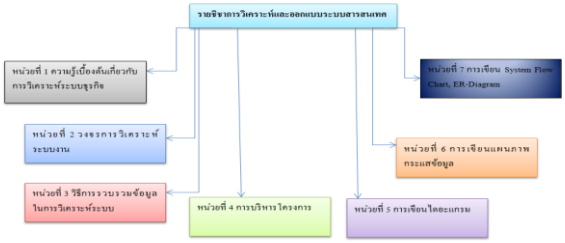
ยุทธศาสตร์และวิถีทางในการพัฒนาระบบสารสนเทศ วงจรการพัฒนาระบบซึ่งประกอบไปด้วยการศึกษาถึงความเป็นไปได้ในการพัฒนาระบบ ความต้องการของระบบสารสนเทศ การวิเคราะห์โครงสร้างวิธีการออกแบบระบบ การออกแบบระบบ การพัฒนาโปรแกรม การติดตั้งและการทดสอบระบบ ฝึกปฏิบัติการวิเคราะห์และออกแบบระบบสารสนเทศทางการศึกษา

3. วิธีการดำเนินการวิจัย

3.1 กรอบแนวคิดในการวิจัย

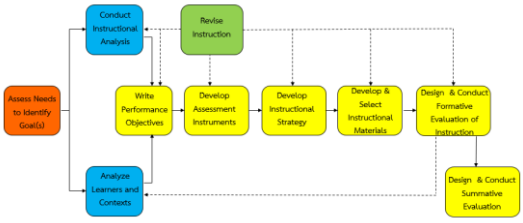
3.1.1 กรอบแนวคิดด้านเนื้อหา

ศึกษาคำอธิบายรายวิชา และจุดประสงค์รายวิชาการวิเคราะห์และออกแบบระบบสารสนเทศ ของนักศึกษาวิชาชีพครู ชั้นปีที่3 สาขาวิชาคอมพิวเตอร์ศึกษาโดยมีขอบเขตรายละเอียดเนื้อหาในรายวิชาดังนี้ ดังนี้



ภาพที่ 1: เนื้อหาในรายวิชาการวิเคราะห์และออกแบบระบบสารสนเทศ

3.1.2 ศึกษาหลักการการออกแบบระบบการเรียนการสอนตามโมเดลดิกและคาเร่ย์โดยใช้การ์ตูนมัลติมีเดีย 2 มิติร่วมกับคลาวด์เทคโนโลยี ในรายวิชาการวิเคราะห์และออกแบบระบบสารสนเทศ สำหรับนักศึกษาชั้นปีที่ 3 สาขาวิชาคอมพิวเตอร์ศึกษา มหาวิทยาลัยราชภัฏนครปฐม มีขั้นตอนดังนี้



ภาพที่ 2: การออกแบบระบบการเรียนการสอนตามโมเดลดิกและคาเร่ย์

3.2 ขอบเขตของการวิจัย

ประชากรที่ใช้การวิจัยครั้งนี้ คือ นักศึกษาวิชาชีพครู ชั้นปีที่ 3 สาขาวิชาคอมพิวเตอร์ศึกษา มหาวิทยาลัยราชภัฏนครปฐม ภาคเรียนที่ 2/2559

กลุ่มเป้าหมายที่ใช้การวิจัยครั้งนี้คือนักศึกษาวิชาชีพครู ชั้นปีที่ 3 สาขาวิชาคอมพิวเตอร์ศึกษา มหาวิทยาลัยราชภัฏนครปฐม ภาคเรียนที่ 2/2559 ได้มาด้วยวิธีการเลือกแบบเจาะจง จำนวน 87 คน ได้แก่หมู่เรียน 57/12 จำนวน 42 คน และหมู่เรียน 57/13 จำนวน 45 คน โดยผู้วิจัยได้ใช้การออกแบบเนื้อหาให้ตรงกับบทเรียนของโรงเรียนอัมพรไพศาลเพื่อเป็นต้นแบบให้กับผู้เรียนนักศึกษาวิชาชีพครู ชั้นปีที่ 3

3.3 เครื่องมือที่ใช้ในการวิจัย

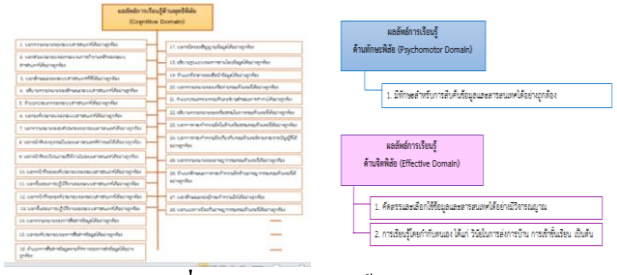
3.3.1 บทเรียนการ์ตูนมัลติมีเดีย 2 มิติ ร่วมกับคลาวด์เทคโนโลยี โดยแบ่งออกเป็น 10 ขั้นตอนดังนี้

ขั้นตอนที่ 1 การประเมินความต้องการจำเป็นเพื่อบ่งชี้เป้าหมายการเรียนการสอนเพื่อประเมินความต้องการจำเป็น



ภาพที่ 3: การประเมินความต้องการจำเป็น

ขั้นตอนที่ 2 การวิเคราะห์การเรียนการสอน เพื่อให้ผู้เรียนบรรลุเป้าหมายตามทักษะ 3 ด้านที่กำหนดไว้ คือ ด้านพุทธิพิสัย ด้านด้านทักษะพิสัย และด้าน ผลลัพธ์การเรียนรู้ด้านจิตพิสัย เพื่อให้สอดคล้องกับผลลัพธ์การเรียนรู้



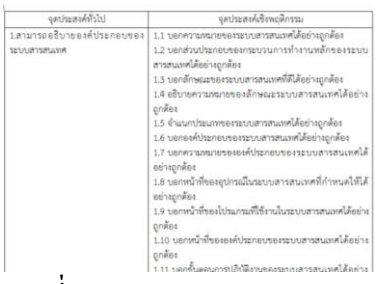
ภาพที่ 4: การวิเคราะห์ตัวชี้วัดของรายวิชา

ขั้นตอนที่ 3 วิเคราะห์การเรียนรู้และบริบท ผู้ออกแบบได้จัดทำแบบวัดพฤติกรรมนำเข้า โดยการวิเคราะห์ผู้เรียน เป็นรายบุคคล โดยจำแนกเป็น 5 ด้าน ดังต่อไปนี้

1) ความรู้ ความสามารถ และประสบการณ์
   (1) ความรู้พื้นฐานของวิชาที่จะทำการสอนในระดับชั้นนั้น ๆ
   (2) ความสามารถในการแก้ปัญหา
   (3) ความสนใจและสมาธิในการเรียนรู้
2) ความพร้อมด้านสติปัญญา
   (1) ความคิดริเริ่ม สร้างสรรค์
   (2) ความมีเหตุผล
   (3) ความสามารถในการเรียนรู้ / การลำดับความ
3) ความพร้อมด้านพฤติกรรม
   (1) การแสดงออก
   (2) การควบคุมอารมณ์
   (3) ความมุ่งมั่น ขยันหมั่นเพียร อดทน
   (4) ความรับผิดชอบ
4) ความพร้อมด้านร่างกายและจิตใจ
   (1) ด้านสุขภาพร่างกายสมบูรณ์
   (2) การเจริญเติบโตสมวัย
   (3) ความสมบูรณ์ด้านสุขภาพจิต
5) ความพร้อมด้านสังคม
   (1) การปรับตัวเข้ากับคนอื่น
   (2) การช่วยเหลือ เสียสละ แบ่งปัน
   (3) การเคารพ ครู กติกา และมีระเบียบวินัย

ภาพที่ 5: การวิเคราะห์การเรียนรู้และบริบท

ขั้นตอนที่ 4 เขียนวัตถุประสงค์เชิงปฏิบัติการ ผู้ออกแบบกำหนดจุดมุ่งหมายเฉพาะ หรือจุดมุ่งหมายเชิงพฤติกรรมและสอดคล้องกับความมุ่งหมายการสอน จุดมุ่งหมายการเรียน ดังนี้



ภาพที่ 6: การเขียนวัตถุประสงค์เชิงปฏิบัติการ

ขั้นตอนที่ 5 สร้างเครื่องมือการประเมินผล เป็นการสร้างเครื่องมือการประเมินผลที่ต้องอธิบายความสัมพันธ์ระหว่างแบบประเมินและจุดประสงค์ของการเรียนรู้ และการจัดกิจกรรมการเรียนการสอน



ภาพที่ 7: ตัวอย่างการสร้างเครื่องมือการประเมินผล

ขั้นตอนที่ 6 พัฒนากลยุทธ์การเรียนการสอน โดยนำเอาแนวแนวความคิด 9 ประการ มาใช้ประกอบการออกแบบบทเรียนการ์ตูนมัลติมีเดีย2มิติ เพื่อให้ได้บทเรียนที่เกิดจากการออกแบบในลักษณะการเรียนการสอนจริง โดยยึดหลักการนำเสนอเนื้อหาและจัดกิจกรรมการเรียนรู้จากการมีปฏิสัมพันธ์



ภาพที่ 8: การออกแบบทเรียนแบบปฏิสัมพันธ์

ขั้นตอนที่ 7 การพัฒนาสื่อการเรียนการสอน การพัฒนาสื่อการเรียนการสอนนี้ เป็นกรณีที่นักออกแบบไม่ได้เป็นผู้สอน ทีมต้องสื่อสารและการทำงานร่วมแบ่งปันส่วนที่เป็นทักษะถ้าไม่มีสิ่งดังกล่าวเป็นขั้นตอนการดำเนินงานมาตรฐานสำหรับการสื่อสารที่เกิดขึ้นคือระหว่างผู้ออกแบบและพัฒนาวัสดุมันก็จะทำงานร่วมกันไม่ซ้ำกันที่จะถูกกำหนดโดยผสมผสานระหว่างการออกแบบและการพัฒนาทักษะการครอบงำโดยผู้เข้าร่วมแต่ละคน โดยใช้วิธีการสอนแบบไฮบริดคือเลือกผู้สอนและปรับวัสดุที่มีอยู่เพื่อให้เหมาะกับกลยุทธ์การเรียนการสอน

ขั้นตอนที่ 8 ออกแบบและประเมินระหว่างการดำเนินการ การประเมินระหว่างการดำเนินการ จะมีการวิเคราะห์เนื้อหาการสอนโดยผู้เชี่ยวชาญด้านเนื้อหา และด้านเทคโนโลยีมัลติมีเดีย และมีการนำสื่อไปทดลองใช้

ขั้นตอนที่ 9 การปรับปรุงสื่อการเรียนการสอน เป็นกระบวนการปรับปรุงสื่อการเรียนการสอน

ขั้นตอนที่ 10 การออกแบบและการประเมินผลสรุป เป็นการออกแบบและการประเมินผลสรุป

3.3.2 เอกสารประกอบการเรียนการสอนในรายวิชา ผู้วิจัยได้แทรกเนื้อหาบทเรียนการพัฒนาการออกแบบระบบการเรียนการสอนตามโมเดลดิกและคาเร่ย์โดยใช้การ์ตูนมัลติมีเดีย 2 มิติ ร่วมกับคลาวด์เทคโนโลยีเข้าไปในเอกสารประอบการสอนในรายวิชา เพื่อเป็นต้นแบบสำหรับผู้เรียนได้ศึกษา



ภาพที่ 9: แสดงเอกสารประกอบการสอนที่ใช้เรียนในรายวิชา

3.3.3 ความพึงพอใจของผู้เรียนที่มีการออกแบบระบบการเรียนการสอนตามโมเดลดิกและคาเร่ย์โดยใช้การ์ตูนมัลติมีเดีย 2 มิติ ร่วมกับคลาวด์เทคโนโลยี ในรายวิชาการวิเคราะห์และออกแบบระบบสารสนเทศ

แบบที่ 6 ประเมินบทเรียน

1. ให้พิจารณาความสอดคล้องระหว่างลักษณะที่สังเกตได้กับรายการประเมินบทเรียนของผู้เรียน โดยใช้เกณฑ์ดังนี้

1 = ปรับปรุง 2 = พอใช้ 3 = ปานกลาง 4 = ดี 5 = ดีมาก

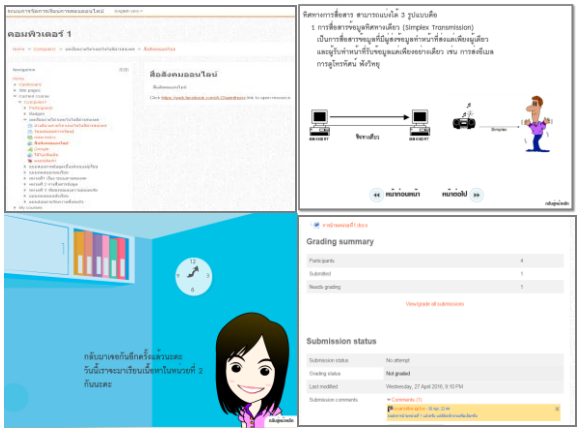
2. ให้เขียนวงกลมล้อมรอบผลการประเมิน

รายการ	ผลการประเมิน
1. ประกอบด้วยเนื้อหาและสาระที่เหมาะสม	1 2 3 4 5
2. มีความประณีต สะอาด และจัดสาระอย่างเป็นระบบ	1 2 3 4 5
3. เป็นบทเรียนมีสีสัน และประกอบด้วยความคิดสร้างสรรค์	1 2 3 4 5
4. การสะกดคำ เครื่องหมาย และการใช้ภาษาของข้อความในบทเรียนมีความถูกต้อง	1 2 3 4 5
5. การออกแบบงานศิลปะในบทเรียน มีความเหมาะสมและละเอียด	1 2 3 4 5
6. สาระที่เสนอสอดคล้องกับหัวข้อ	1 2 3 4 5
7. หัวข้อร้อมรายละเอียดสื่อสารได้เข้าใจ	1 2 3 4 5
8. บทเรียนประกอบด้วยสาระครบตามวัตถุประสงค์ที่กำหนด	1 2 3 4 5
9. การนำเสนอบทเรียนผู้เรียนทำได้เหมาะสม	1 2 3 4 5
10. บทเรียนดึงดูดความสนใจผู้อื่น	1 2 3 4 5
ความคิดเห็นเพิ่มเติม ----------	

ภาพที่ 10: ตัวอย่างแบบสอบถามความพึงพอใจ

4. ผลการดำเนินงาน

4.1 ผลของการออกแบบระบบการเรียนการสอน ตามโมเดลดิกและคาเร่ย์โดยใช้การ์ตูนมัลติมีเดีย 2 มิติ ร่วมกับคลาวด์เทคโนโลยี ในรายวิชาการวิเคราะห์และออกแบบระบบสารสนเทศ



ภาพที่ 11: ตัวอย่างบทเรียน

4.2 ผลสำรวจความพึงพอใจของผู้เรียน โดยมีรายละเอียดดังนี้

ตารางที่ 1: ผลความพึงพอใจการออกแบบระบบการเรียนการสอนตามโมเดลดิกและคาเร่ย์โดยใช้การ์ตูนมัลติมีเดีย 2 มิติ ร่วมกับคลาวด์เทคโนโลยี ในรายวิชาการวิเคราะห์และออกแบบระบบสารสนเทศ

รายการประเมิน	ค่าเฉลี่ย	S.D.	ระดับคุณภาพ
1. ประกอบด้วยเนื้อหาและสาระที่เหมาะสม	4.83	0.27	ดีมาก
2. มีความประณีต สะอาด และจัดสาระอย่างเป็นระบบ	4.56	0.44	ดีมาก
3. เป็นบทเรียนมีสีสัน และประกอบด้วยความคิดสร้างสรรค์	4.76	0.28	ดีมาก
4. การสะกดคำ เครื่องหมาย และการใช้ภาษาของข้อความในบทเรียนมีความถูกต้อง	4.68	0.45	ดีมาก
5. การออกแบบงานศิลปะในบทเรียน มีความเหมาะสมและละเอียด	4.47	0.62	ดี
6. สาระที่เสนอสอดคล้องกับหัวข้อ	4.87	0.21	ดีมาก
7. หัวข้อพร้อมรายละเอียดสื่อสารได้เข้าใจ	4.62	0.37	ดีมาก
8. บทเรียนประกอบด้วยสาระครบตามวัตถุประสงค์ที่กำหนด	4.76	0.22	ดีมาก
9. การนำเสนอบทเรียนทำได้เหมาะสม	4.81	0.34	ดีมาก
10. บทเรียนดึงดูดความสนใจผู้อื่น	4.78	0.34	ดีมาก
ค่าเฉลี่ยรวม	4.71	0.35	ดีมาก

จากตารางที่ 2 พบว่ามีค่าเฉลี่ยรวมมีค่าเท่ากับ 4.71 และมีส่วนเบี่ยงเบนมาตรฐาน 0.35 อยู่ในระดับดีมาก

5. สรุปและอภิปรายผล

5.1 สรุปผลการวิจัย

ผลการวิจัยพบว่า 1) ผลจากการประเมินความพึงพอใจของผู้เรียนที่มีต่อการออกแบบระบบการเรียนการสอนตามโมเดลดิกและคาเร่ย์โดยใช้การ์ตูนมัลติมีเดีย 2 มิติ ร่วมกับคลาวด์เทคโนโลยี ในรายวิชาการวิเคราะห์และออกแบบระบบสารสนเทศ สำหรับนักศึกษาชั้นปีที่ 3 นั้นอยู่ในระดับดีมาก โดยมีค่าเฉลี่ยรวมทั้งหมด 4.71 และมีค่าส่วนเบี่ยงเบนมาตรฐาน 0.35 ซึ่งสอดคล้องกับ [4] ซึ่งใช้แนวทางในการพัฒนาสื่อเพื่อการเรียนรู้ตามโมเดลดิกและค่าเร่ย์ทำให้การพัฒนามีคุณภาพ และ[5] ได้กล่าวว่าโมเดลดิกและค่าเร่ย์เป็นรูปแบบการสอนที่นิยมนำไปประยุกต์ใช้ และยังสอดคล้องกับ [6] ได้กล่าวว่า คลาวด์เทคโนโลยีก็เป็นส่วนหนึ่งของ การจัดการเรียนการสอนในศตวรรษที่ 21 ได้เพียงแต่เรา เลือกใช้ระบบคลาวด์ที่เหมาะสมและคุ้มค่าต่อการจัดการเรียน การสอนให้มากที่สุด เพื่อเป็นพัฒนาความรู้อย่างต่อเนื่องและ ตลอดชีวิต และยังพัฒนาความสามารถในการดารงชีวิตในการ เปลี่ยนแปลงทางสังคมในโลกเทคโนโลยีได้อย่างยั่งยืนในสภาวะ การเรียนรู้แบบทุกที่ทุกเวลา ดังนั้นสามารถสรุปได้ว่า การออกแบบระบบการเรียนการสอนตามโมเดลดิกและคาเร่ย์โดยใช้การ์ตูนมัลติมีเดีย 2 มิติ ร่วมกับคลาวด์เทคโนโลยี ในรายวิชาการวิเคราะห์และออกแบบระบบสารสนเทศที่พัฒนาขึ้นสามารถนำไปใช้ในการจัดการเรียนการสอนในชั้นเรียนได้จริงอีกทั้งยังเป็นต้นแบบของการสร้างนวัตกรรมทางการศึกษาที่ทำให้ผู้เรียนเข้าใจในบริบทได้มากยิ่งขึ้นทำให้ผู้เรียนสามารถพัฒนาผลงานของตนเองตามรูปแบบโมเดลที่ตนเองสนใจได้ด้วย

5.2 ข้อเสนอแนะในการวิจัยครั้งต่อไป

5.2.1 ควรเพิ่มหัวข้อสนับสนุนการวิจัยให้มากกว่านี้ เช่น ผลสัมฤทธิ์ทางการเรียน,การหาประสิทธิภาพทางการเรียน, หรือทักษะด้านICT

5.2.2 ควรเพิ่มลักษณะของรูปแบบการจัดการเรียนการสอนให้มากยิ่งขึ้น

5.2.3 ควรเพิ่มลักษณะของลักษณะการจัดการเรียนการสอนให้น่าสนใจมากขึ้น

เอกสารอ้างอิง

[1] กระทรวงเทคโนโลยีสารสนเทศและการสื่อสาร, “ **แผนพัฒนาดิจิทัลเพื่อเศรษฐกิจและสังคม**” , พิมพ์ครั้งที่ 1. 2559

[2] Walter Dick, Lou Carey, James O.Carey. **The Systematic Design of Instruction**.2013

[3] _____, “คลาวด์เทคโนโลยี” สืบค้นเมื่อวันที่ 30 มกราคม 2560 จากออนไลน์http://sc2.kku.ac.th/office/sci-.php/29-cloud-computing

[4] พงษ์พิพัฒน์ สายทอง, “ **การออกแบบอินโฟกราฟิกแอนิเมชันเพื่อการเรียนการสอน**” , วารสารวิชาการ. ศิลปะสถาปัตยกรรมศาสตร์. มหาวิทยาลัยนเรศวร. ปีที่ 5. ฉบับที่ 2 กรกฎาคม – ธันวาคม 2557

[5] จตุพร ตันติรังสีและคณะ, “ **การออกแบบอินโฟกราฟิกแอนิเมชันเพื่อการเรียนการสอน**” , **การพัฒนาบทเรียนคอมพิวเตอร์ช่วยสอนบนเครือข่ายอินเทอร์เน็ต โดยใช้การสอนแบบ ADDIE Model วิชาการใช้โปรแกรมกราฟิก**. งานประชุมวิชาการระดับชาติ. NCCIT2012.

[6] นาวิน คงรักษา, “ **คลาวด์คอมพิวติ้งกับการจัดการเรียนการสอนในศตวรรษที่ 21**” , วารสารการอาชีวะและเทคนิคศึกษา. ปีที่ 4. ฉบับที่ 7. มกราคม – มิถุนายน 2557.

แบบจำลองทางคณิตศาสตร์สำหรับการพัฒนาระบบตรวจสอบการทำงานระบบผลิตน้ำร้อนพลังงานแสงอาทิตย์

Mathematical model for the development of monitoring systems for solar water heating systems.

พงศ์พิชญ์ ต่วนภูษา(Pongpith Tuenpusa)[1] นิติ วิทยาวิโรจน์(Niti Witthayawiroj)[1] สถาพร ทองวิค(Sathaporn Thongwik)[2]

สุริยะ สุกรินทร์(Suriya Sukarin)[3]

[1]สำนักวิทยบริการและเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี

[2]ภาควิชาวิศวกรรมเครื่องกล คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี

[3]E-mail suriyasukarin@gmail.com

บทคัดย่อ

บทความฉบับนี้กล่าวถึงการนำแบบจำลองทางคณิตศาสตร์ มาประยุกต์ใช้ในการพัฒนาแอพพลิเคชั่น เพื่อช่วยวิเคราะห์ความเสียหายและความผิดปกติของแผงผลิตน้ำร้อนพลังงานแสงอาทิตย์แสงอาทิตย์ เพื่อลดความผิดพลาดในการตรวจสอบการทำงานของระบบ ซึ่งแบบจำลองทางคณิตศาสตร์นี้สามารถประยุกต์ใช้ในการพัฒนาแอพพลิเคชั่นได้หลายแพลตฟอร์ม นอกจากนั้นบทความฉบับนี้ยังมีการยกตัวอย่างการเชื่อมต่อระบบฮาร์ดแวร์เพื่อใช้ในการรับส่งข้อมูลสำหรับการประมวลผล และส่วนของการเข้าถึงข้อมูลด้วย ซึ่งทางผู้เขียนหวังว่าบทความฉบับนี้จะเป็นแนวทางในการพัฒนาแอพพลิเคชั่นด้านการจัดการงานซ่อมบำรุงเครื่องจักรต่อไป

คำสำคัญ: น้ำร้อนแสงอาทิตย์ แอพพลิเคชั่น การซ่อมบำรุง

Abstract

This article discusses the implementation of mathematical models for application development. This application analyzes the damage and malfunctions of the solar water heater and reduces the failure of the system to check the operation of the system. This mathematical model can be used to develop multiple applications. In addition, this article also provides an example of a hardware-based interface for processing data and the data access. The author hopes that this article will guide the development of maintenance management applications

Keyword: Solar hot water system, Maintenance, Data monitoring, Modbus protocol

1. บทนำ

ปัจจุบันการติดตั้งและใช้งานระบบผลิตน้ำร้อนพลังงานแสงอาทิตย์ภาคธุรกิจ นั้นมีการติดตั้งเป็นจำนวนมาก นอกจากนี้ยังมีการพัฒนาผลิตภัณฑ์ที่นำระบบผลิตน้ำร้อนพลังงานแสงอาทิตย์มาใช้มากขึ้น เช่น การนำมาใช้ปรับอุณหภูมิของน้ำในสระว่ายน้ำ



ภาพที่ 1 ระบบปรับอุณหภูมิของน้ำในสระว่ายน้ำด้วยแสงอาทิตย์ (ที่มา: sunflower-solar.com)

ซึ่งระบบดังกล่าวสามารถทำให้ผู้ประกอบการลดต้นทุนด้านพลังงานได้เป็นอย่างมาก แต่การติดตั้งระบบดังเมื่อมีการใช้งานไปในระยะหนึ่งจะต้องมีการบำรุงรักษาเพื่อให้ระบบสามารถทำงานได้อย่างมีประสิทธิภาพดังเดิม จากศึกษา

ผู้วิจัยพบว่าการตรวจสอบการทำงานของระบบนั้นโดยส่วนมากเป็นหน้าที่ของช่างเทคนิคเป็นผู้ตรวจสอบความผิดปกติของแผงรับแสงอาทิตย์และอุปกรณ์ส่วนควบทีละจุดด้วยสายตา หรือ การใช้เครื่องมือวัดหน้างาน เป็นการทำงานในสภาพแวดล้อมกลางแจ้งที่มีอุณหภูมิสูง เป็นเวลานานจึงมีความเสี่ยงสูงที่ทำให้เกิดความผิดพลาดในการตรวจสอบที่มาจาก ตัวผู้ตรวจสอบเอง และจากเครื่องมือ จากบัญหาดังกล่าวผู้วิจัยได้เล็งเห็นว่าหากมีการนำระบบแอพพลิเคชั่นมาประยุกต์ใช้ จะสามารถช่วยให้การแก้ปัญหาดังกล่าวได้เป็นอย่างดี

แต่ปัจจุบันการพัฒนาแอพพลิเคชั่นเพื่อใช้ในแพลตฟอร์มต่างนั้นจะต้องใช้เครื่องมือในการพัฒนาที่ต่างกัน แต่ทุกเครื่องมือนั้นจะมีสิ่งที่เหมือนกันคืออัลกอริทึมในการวิเคราะห์และติดสินใจของระบบซอร์ฟแวร์

บทความฉบับนี้เป็นการนำเสนอแบบจำลองทางคณิตศาสตร์เพื่อใช้เป็นต้นแบบในการพัฒนาอัลกอริทึม สำหรับการพัฒนาชุดคำสั่งของซอร์ฟแวร์ตรวจสอบความเสียหายของระบบผลิตน้ำร้อนพลังงานแสงอาทิตย์ ซึ่งใช้การสื่อสารผ่าน industrial protocols ในการรับส่งข้อมูลระหว่างเครื่องมือวัดและระบบประมวลผล แล้วนำข้อมูลที่ได้รับนั้นมาทำการวิเคราะห์เชิงเปรียบเทียบแล้วแสดงผลแจ้งเตือนการทำงาน และลดเหตุการณ์ที่จะต้องบำรุงรักษาแบบซ่อมเมื่อเสีย (Breakdown Maintenance) ซึ่งส่งผลให้ต้องหยุดการทำงานของระบบทั้งหมด

รูปแบบการวิเคราะห์ของแบบจำลองทางคณิตศาสตร์ชุดนี้ได้ใช้กลยุทธ์การบำรุงรักษาอุปกรณ์ในรูปแบบของการบำรุงรักษาเชิงป้องกัน (Planned Maintenance or Preventive Maintenance) และ การบำรุงรักษาเชิงพยากรณ์ (Predictive Maintenance) ซึ่งกลยุทธ์การซ่อมบำรุงทั้งสองนี้เป็นรูปแบบของการซ่อมบำรุงเชิงรุก

2. การวิเคราะห์ประสิทธิภาพระบบผลิตน้ำร้อนพลังงานแสงอาทิตย์ด้วยแบบจำลองคณิตศาสตร์

ในการพิจารณาประสิทธิภาพของแผงผลิตน้ำร้อนจากแสงอาทิตย์ ตัวแปลในการพิจารณาหลายตัวเปล เช่น แผงผลิตน้ำร้อนนั้นจากแสงอาทิตย์นั้นดูดกลืนพลังงานได้อย่างไร มีการสูญเสียความร้อนในรูปแบบใด เป็นต้น ซึ่งการพิจารณานั้นจะเป็นไปตามมาตรฐาน ASHRAE 93-77 โดยการหาประสิทธิภาพชั่วขณะนั้นจะสามารถหาได้จากสมการที่ 1

$$\eta_c = \frac{Q_u}{A_C G_T} = F_R(\tau\alpha)_e - F_R U_L \frac{(T_i - T_a)}{G_T} \quad (1)$$

ในการพิจารณาประสิทธิภาพตามมาตรฐาน ASHRAE 93-77 ในรูปแบบของการสร้างกราฟ ความสัมพันธ์ระหว่างประสิทธิภาพของแผงรับแสงอาทิตย์ (η_C) กับผลลัพธ์ของของเทอม $\frac{(T_i - T_a)}{G_T}$ โดยกราฟความสัมพันธ์นั้นจะเป็นกราฟเส้นตรง โดยค่าความชันซึ่งมีค่าติดลบนั้นคือค่า $F_R U_L$ และจุดตัดของกราฟ ในแนวแกน y นั้นคือ $F_R(\tau\alpha)$ ซึ่งสมการเส้นตรงที่ได้จากการพิจารณาจะเป็นไปตามสมการที่ 2

$$y = (-F_R U_L)(x) + F_R(\tau\alpha)_e \quad (2)$$

โดยที่

y คือ ประสิทธิภาพของแผงรับแสงอาทิตย์ (η_C)

x คือ $\frac{(T_i - T_a)}{G_T}$

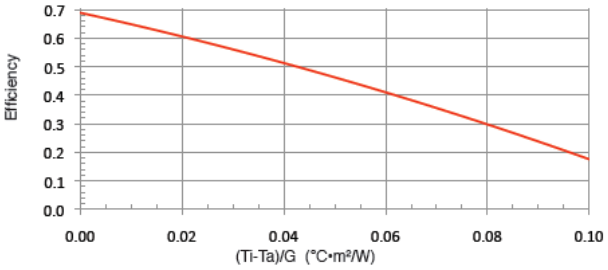
ภายหลังการทดสอบแผงผลิตน้ำร้อนพลังงานแสงอาทิตย์จากสถาบันทดสอบแล้วนั้น แผงดังกล่าวจะถูกนำไปใช้งานจริง ซึ่งการใช้งานนั้นผู้ใช้งานอาจมีการปรับอัตราการไหลไม่เป็นไปตามผลการทดสอบของผู้ผลิตกำหนด ซึ่งจะส่งผลให้เกิดการเปลี่ยนแปลงของค่า $F_R U_L$ และ $F_R(\tau\alpha)$ ซึ่งการเปลี่ยนแปลงดังกล่าวจะเป็นไปตามสมการที่ 3 และ 4

$$r = \frac{F_R U_L|_{test}}{F_R U_L|_{use}} = \frac{F_R(\tau\alpha)_n|_{use}}{F_R(\tau\alpha)_n|_{test}} \quad (3)$$

$$r = \frac{\left.\frac{\dot{m} C_p}{A_C F' U_L}\left[1 - \exp\left(\frac{-A_C F' U_L}{\dot{m} C_p}\right)\right]\right|_{Use}}{\left.\frac{\dot{m} C_p}{A_C F' U_L}\left[1 - \exp\left(\frac{-A_C F' U_L}{\dot{m} C_p}\right)\right]\right|_{Test}} \quad (4)$$

โดยค่า $F'U_L$ หาได้จากสมการที่ 5

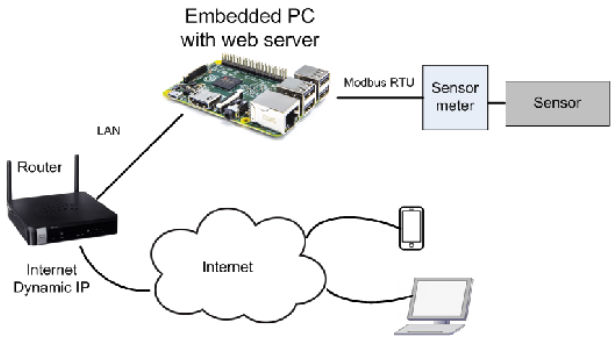
$$F'U_L = -\frac{\dot{m}C_p}{A_C}\ln\left(1-\frac{F_R U_L A_c}{\dot{m}C_p}\right) \quad (5)$$



ภาพที่ 2 กราฟที่ได้จากวิเคราะห์ด้วยแบบจำลองทางคณิตศาสตร์

3. การออกแบบระบบการรับส่งข้อมูล

เนื่องจากการระบบการตรวจสอบจะต้องอาศัยการทำงานร่วมกันระหว่างฮาร์ดแวร์ ซึ่งประกอบด้วยเครื่องมือวัด (Sensor) ที่ติดตั้งตามจุดต่างๆ กับซอร์ฟแวร์ประมวลผล ซึ่งจะต้องมีการรับส่งข้อมูลไปมาตลอดเวลา ซึ่งกระบวนการดังกล่าวผู้วิจัยได้ประยุกต์การใช้ Protocol Modbus ถือเป็น Industrial Protocol รูปแบบหนึ่งที่นิยมใช้แพร่หลายในโรงงานอุตสาหกรรม และยังมีผู้ผลิตอุปกรณ์ที่รองรับการทำงานของระบบการสื่อสารในรูปแบบนี้ด้วย



ภาพที่ 3 การเชื่อมต่อฮาร์ดแวร์

จากภาพเป็นการเชื่อมต่อสัญญาณระว่างเครื่องมือวัด (Sensor) และอุปกรณ์ต่าง โดยที่ข้อมูลที่ได้จากเครื่องมือวัดนั้นจะถูกแปลงสัญญาณให้อยู่ในรูปแบบของ Protocol Modbus แล้วส่งข้อมูลไปยังเครื่องคอมพิวเตอร์ผ่านทางช่องทางสื่อสารแบบ RS 485 ซึ่งในเครื่องคอมพิวเตอร์ดังกล่าวได้มีการติดตั้งแอพพลิเคชั่นสำหรับประมวลผลไว้ และยังทำหน้าที่เป็น Web server อีกหน้าที่หนึ่งด้วย ซึ่งการเข้าถึงข้อมูลดังกล่าวนั้นสามารถเข้าได้จากอุปกรณ์หลายแบบ เช่น เครื่องคอมพิวเตอร์บุคคล หรือการการใช้ Smart Phone ด้วย

4. การใช้อัลกอลิทึม และแบบจำลองทางคณิตศาสตร์ของระบบสูบน้ำในระบบผลิตน้ำร้อนพลังงานแสงอาทิตย์

ในการออกแบบซอฟต์แวร์นั้นการไหลของข้อมูลขาเข้านั้นจะประกอบด้วยประกอบด้วยข้อมูล 2 ส่วนคือ

1. ตัวแปรจากเครื่องมือวัด
2. ตัวแปรเชิงเปรียบเทียบ

เมื่อซอฟต์แวร์ได้รับข้อมูลทั้ง 2 ชุดแล้วนั้น ระบบจะมีชุดคำสั่งเพื่อทำการประมวลผลตามเงื่อนไขการทำงาน 2 ส่วนคือ อันประกอบด้วย

1. การคำนวณประสิทธิภาพ
2. การวิเคราะห์ข้อมูลเชิงเปรียบเทียบ

ในการวิเคราะห์การทำงานของระบบสูบจ่ายน้ำร้อนนั้นผู้ออกแบบซอร์ฟแวร์สามารถใช้แบบจำลองทางคณิตศาสตร์โดยกำหนดหนดให้

P_H คือกำลังที่ได้จากน้ำที่ออกจากเครื่องสูบ

γ คือน้ำหนักจำเพาะของน้ำที่ไหลในระบบผลิตน้ำร้อนพลังงานแสงอาทิตย์

Q คือ อัตราการสูบของเครื่องสูบน้ำร้อน

δ คือเฮดรวมของเครื่องสูบน้ำ

P_{ip} คือกำลังทางกลที่ส่งเข้าสู่เครื่องสูบ

P_{me} คือกำลังทางกลจากเครื่องต้นกำลังที่ส่งให้เครื่องสูบ

P_E คือกำลังที่ให้กับเครื่องต้นกำลัง

τ ทอร์คจากการหมุนของมอเตอร์

ω ความเร็วเชิงมุมในการหมุนมอเตอร์ต้นกำลัง

V คือแรงเคลื่อนไฟฟ้าที่ส่งให้แก่มอเตอร์

I คือกระแสไฟฟ้าที่ไหลเข้าสู่มอเตอร์

$cos\theta$ คือ ตัวประกอบกำลัง

η_T คือประสิทธิภาพของระบบส่งกำลัง

η_P คือประสิทธิภาพของเครื่องสูบ



ภาพที่ 4 การส่งกำลังในระบบเครื่องสูบ

การวิเคราะห์กำลังที่ส่งให้กับเครื่องต้นกำลัง ซึ่งในระบบนี้ใช้มอเตอร์ไฟฟ้าเป็นเครื่องต้นกำลัง ดังนั้นจึงสามารถหากำลังที่ให้กับเครื่องต้นกำลังได้จากสมการ

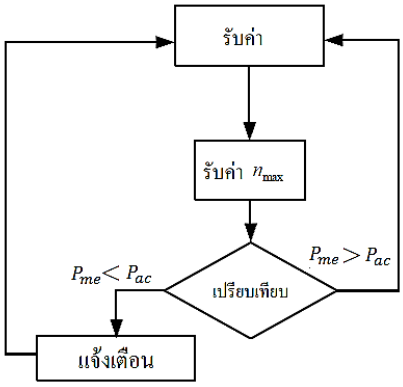
$$P_E = VIcos\theta \qquad (6)$$

ค่าทั้งสองนั้นได้มาจากการวัดค่าผ่านอุปกรณ์วัดข้อมูลทางไฟฟ้า

เมื่อมอเตอร์ต้นกำลังนั้นได้รับพลังงานเข้าไปแล้วจะเปลี่ยนพลังงานดังกล่าวเป็นพลังงานกล ซึ่งสามารถหาได้จากสมการ

$$P_{me} = \tau\omega \qquad (7)$$

ซึ่งความสัมพันธ์ของกำลังทางกลจากเครื่องต้นกำลังที่ส่งให้เครื่องสูบ ระบบจะต้องมีการวิเคราะห์กำลังที่แท้จริงและกำลังที่เครื่องสูบต้องการ โดยที่กำลังที่แท้จริงนั้นจะต้องไม่ต่ำกว่ากำลังที่เครื่องสูบต้องการ หากเกิดเหตุการณ์ที่กำลังแท้จริงนั้นต่ำกว่ากำลังที่เครื่องสูบต้องการแล้วนั้นหมายความมีแนวโน้มที่จะเกิดความผิดปกติในระบบเครื่องต้นกำลัง ดังนั้นเพื่อป้องกันความเสียหายระบบซอร์ฟแวร์ต้องมีระบบแจ้งเตือนความผิดปกตินี้ ให้ผู้ควบคุมทราบทันที



รูปที่ 5 เงื่อนไขการเปรียบเทียบระบบส่งกำลัง

เมื่อเครื่องสูบน้ำได้รับพลังงานกลจากมอเตอร์แล้วนั้น พลังงานกลดังกล่าวจะใช้เป็นเพลังงานในการสูบน้ำในระบบผลิตน้ำร้อนพลังงานแสงอาทิตย์ โดยที่กำลังที่เครื่องสูบได้รับนั้นสามารถแบ่งรูปแบบการส่งกำลังได้ เป็น 2 แบบคือ

1. การส่งสำลังโดยตรง หากเครื่องสูบน้ำในระบบผลิตน้ำร้อนพลังงานแสงอาทิตย์รูปแบบการส่งกำลังชนิดนี้จะไม่มีการคิดประสิทธิภาพของระบบส่งกำลัง จึงทำให้ระบบระบบซอร์ฟแวร์สามารถข้ามการตรวจสอบในส่วนนี้ได้

2. การส่งกำลังผ่านระบบส่งกำลัง หากเครื่องสูบน้ำในระบบผลิตน้ำร้อนพลังงานแสงอาทิตย์รูปแบบการส่งกำลังชนิดนี้ระบบซอร์ฟแวร์จะต้องมีการวิเคราะห์ประสิทธิภาพของระบบนี้ด้วย โดยใช้เงื่อไขการวิเคราะห์ดังนี้

$$\eta_T = \frac{P_{ip}}{P_{me}} \qquad (8)$$

โดยมีเงื่อไขในการแจ้งเตือนดังนี้

- ระบบส่งกำลังแบบเฟืองทดนั้นมีประสิทธิภาพลดลงเหลือ 98%
- ระบบส่งกำลังแบบสายพานแบนมีค่าประสิทธิภาพลดลงเหลือ 97 %
- ระบบส่งกำลังด้วยสายพานรูปตัววี มีประสิทธิภาพลดลงเหลือ 95 %

ดังนั้นหากประสิทธิภาพของระบบส่งกำลังลดลงต่ำกว่าเงื่อนไขดังกล่าวระบบควรที่จะมีการแจ้งเตือนทันที อันเนื่องมาการอาจมีความผิดปกติของระบบส่งกำลัง



รูปที่ 6 ระบบเครื่องสูบในระบบผลิตน้ำร้อนพลังงานแสงอาทิตย์

การวิเคราะห์กำลังการไหลของน้ำในระบบผลิตน้ำร้อนพลังงานแสงอาทิตย์ ซึ่งการวิเคราะห์ดังกล่าวนี้ระบบจะต้องแบ่งการวิเคราะห์ออกเป็น 2 ส่วน คือ

1. การวิเคราะห์กำลังของน้ำในระบบท่อ สามารถใช้เงื่อไขการวิเคราะห์ดังนี้

$$P_H = \frac{\gamma Q \delta}{1000} \quad (9)$$

2. การวิเคราะห์ประสิทธิภาพของเครื่องสูบ ใช้เงื่อไขการวิเคราะห์ดังนี้

$$\eta_p = \frac{P_H}{P_{ip}} \times 100 \quad (10)$$

หากผลลัพธ์ที่ได้จากการวิเคราะห์นี้มีค่าน้อยกว่าค่าประสิทธิภาพที่ระบุในคู่มือของเครื่องสูบ ระบบซอร์ฟแวร์จะต้องมีการแจ้งเตือนให้ผู้ควบคุมทราบทันที

5.บทสรุป

จากที่ได้กล่าวมาต้นแบบของอัลกอลิทึมนี่มีรูปแบบกระบวกการคำนวณประสิทธิภาพตามมาตรฐาน ASHRAE 93:77 และการวิเคราะห์ข้อมูลเชิงเปรียบเทียบเพื่อการแจ้งเตือนความผิดปกติของระบบเครื่องสูบน้ำร้อน ซึ่งเป็นประโยชน์ต่อผู้ปฏิบัติงานเป็นอย่างมาก ผู้เขียนหวังว่าแบบจำลองทางคณิตศาสตร์นี้จะเป็นแนวทางสำหรับนักพัฒนาแอพพลิเคชั่นต่อไป

เอกสารอ้างอิง

[1] สุริยะ สุกรินทร์. 2554 .การพัฒนาระบบจัดการงานซ่อมบำรุงระบบผลิตน้ำร้อนพลังงาน แสงอาทิตย์ด้วยคอมพิวเตอร์ .วิทยานิพนธ์ปริญญามหาบัณฑิต. มหาวิทยาลัยเทคโนโลยี ราชมงคลธัญบุรี

[2] จงจิตร์ หิรัญลาภ. 2547. กระบวนการพลังงานรังสีอาทิตย์ในรูปความร้อน. กรุงเทพมหานคร: สำนักพิมพ์ดวงกมล.

[3] กรมพัฒนาพลังงานทดแทนและอนุรักษ์พลังงาน. 2553 . คู่มือผู้รับผิดชอบด้านพลังงาน พ.ศ. 2553

[4] จรัส แซ่เล้า.การพัฒนาระบบมิเตอร์ไฟฟ้าอัตโนมัติด้วยโปรโตคอล Modbus RTU, เว็ปไซต์ meath-co.com

[5] E. H. AMER, P. JADEJA, J. K. NAYAK and G. K. SHARMA ,Comparison of Two Dynamic Test Methods for Solar Flat-Plate Collector, Energy Convers. Mgmt Vol. 39, No. 3/4, pp. 285-293, 1998

[6] S.Farahat, F.Sarhaddi, H. Ajam ,Exergetic optimization of flat plate solar collectors, Renewable Energy 34 (2009) 1169–1174

[7] J. K. NAYAK and E. H. AMER, Experimental and Theoretical Evaluation of Dynamic Test Procedures for Solar Flat-Plate Collector, Solar Energy Vol. 69, No. 5, pp. 377–401, 2000

[8] Solar Engineering Of Thermal Processes,John A. Duffie ,Willam A.Beckman,John Wiley& Sons,Inc.

โมบายแอปพลิเคชันเพื่ออ่านฉลากยาสำหรับผู้บกพร่องทางการมองเห็น โดยใช้เทคโนโลยีเอ็นเอฟซี

Mobile Application to Reading Drug Labels for Visually Impairment Using NFC Technology

ณัฐรุจา สุวรรณพรรค (Natruja Suwannapak)[1] และผุสดี บุญรอด (Pudsadee Boonrawd)[2]

[1,2]ภาควิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ

มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

little.rainn@gmail.com[1], pudsadee.b@it.kmutnb.ac.th[2]

บทคัดย่อ

งานวิจัยนี้นำเสนอโมบายแอปพลิเคชัน บนระบบปฏิบัติการแอนดรอยด์ เพื่ออ่านฉลากยาสำหรับผู้บกพร่องทางการมองเห็น โดยใช้เทคโนโลยีเอ็นเอฟซี ที่เป็นการสื่อสารไร้สายระยะสั้น ซึ่งช่วยในการบันทึกและอ่านข้อมูลยาในฉลากยาเอ็นเอฟซี และนำมาแสดงผลในรูปแบบเสียงผ่านสมาร์ทโฟน ได้แก่ ชื่อยา สรรพคุณ วิธีการรับประทาน คำเตือนและวันหมดอายุ ทำให้ผู้บกพร่องทางการมองเห็นสามารถรับประทานยาได้อย่างสะดวกและถูกต้อง นอกจากนั้นงานวิจัยได้นำแอนดรอยด์ทอล์คแบค ซึ่งเป็นโปรแกรมช่วยอ่านหน้าจอสมาร์ทโฟนมาประยุกต์ใช้ เพื่อให้คนที่มีอาการบกพร่องทางการมองเห็นสามารถใช้งานได้ และผลการทดสอบประสิทธิภาพเพื่อวัดค่าความถูกต้องของการอ่านฉลากยาเอ็นเอฟซี พบว่าค่าความถูกต้องมีค่าเท่ากับ 91.33% และผลการประเมินความพึงพอใจของผู้บกพร่องทางการมองเห็นอยู่ในระดับดี ดังนั้นจึงสามารถนำงานวิจัยนี้ไปพัฒนาต่อยอดเพื่อใช้งานจริงในอนาคต

คำสำคัญ: ฉลากยา เอ็นเอฟซี ผู้บกพร่องทางการมองเห็น โปรแกรมอ่านหน้าจอ

Abstract

This research presents the Android mobile application to read drug labels for visually impaired persons by adopting Near Field Communication (NFC), a standards-based short-range wireless connectivity technology. The application is able to read and write drug information into NFC labels and then represents them in a voice format via smartphone such as drug name, indication, directions, precautions, and expiration dates so visually impaired persons can take the medicine conveniently and correctly. It also provides accessibility to persons with other disabilities by employing Android Talkback. From the experiment, the accuracy rate is 91.33%, and the satisfaction survey among visually impaired persons was in a good level. Accordingly, further development of the application could lead to the adoption of this research.

Keyword: Drug Label, NFC, Visually Impaired People, Android TalkBack

1. บทนำ

จากข้อมูลการจดทะเบียนคนพิการทั่วประเทศไทยของกระทรวงการพัฒนาสังคมและความมั่นคงของมนุษย์ ตั้งแต่วันที่ 1 พฤศจิกายน พ.ศ. 2537 ถึงวันที่ 31 มีนาคม พ.ศ. 2560 พบว่ามีจำนวนผู้พิการทุกประเภทรวมทั้งสิ้น 1,756,849 ราย โดยจำนวนผู้บกพร่องทางการมองเห็นเท่ากับ 184,802 ราย ซึ่งมีจำนวนมากเป็นอันดับที่ 3 รองจากจำนวนผู้บกพร่องทางการได้ยิน คิดเป็นร้อยละ 10.52 ของจำนวนผู้พิการทั้งหมดทั่วประเทศไทย [1] และข้อมูลขององค์การอนามัยโลกรายงานว่าในปี พ.ศ. 2557 มีผู้บกพร่องทางการมองเห็นในโลกกว่า 285 ล้านคน โดย 39 ล้านคนตาบอดสนิท และ 246 ล้านคนมองเห็นเลือนราง [2] ซึ่งถ้าบุคคลเหล่านี้มีอาการเจ็บป่วยและต้องรับประทานยาด้วยตนเอง อาจทำให้เกิดโอกาสในการรับประทานยาผิด จึงมีความจำเป็นอย่างยิ่งที่จะต้องอาศัยญาติ

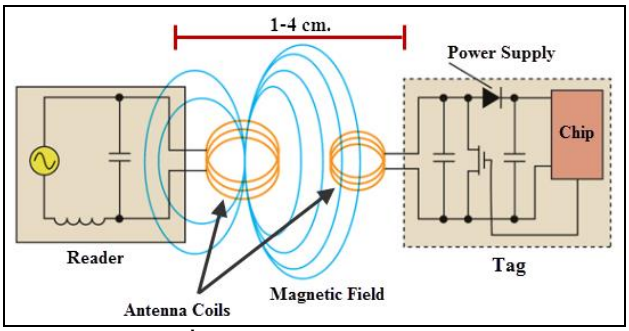
หรือผู้ดูแลให้ความช่วยเหลือ เนื่องจากไม่สามารถอ่านฉลากยาหรือจำแนกยาที่มีลักษณะคล้ายกันได้ ดังนั้นจึงควรหาวิธีเพื่อช่วยเหลือผู้บกพร่องทางการมองเห็นให้สามารถรับประทานยาได้ด้วยตนเองอย่างสะดวกและถูกต้อง

ปัจจุบันอินเทอร์เน็ตและสมาร์ทโฟนได้เข้ามามีบทบาทในการดำรงชีวิตประจำวัน และจากปัญหาดังกล่าวงานวิจัยนี้จึงมีวัตถุประสงค์เพื่อช่วยเหลือผู้บกพร่องทางการมองเห็นให้สามารถรับประทานยาได้ด้วยตนเองอย่างสะดวกและถูกต้อง โดยงานวิจัยได้พัฒนาในรูปแบบโมบายแอปพลิเคชันและแสดงผลในรูปแบบเสียงผ่านสมาร์ทโฟน เช่น ชื่อยา สรรพคุณ คำเตือน วันหมดอายุ เป็นต้น ซึ่งจะช่วยลดความเสี่ยงจากการใช้ยาผิดพลาด โดยในบทความวิจัยนี้ประกอบด้วยเนื้อหาหลัก 5 ส่วน ได้แก่ บทนำ ทฤษฎีและงานวิจัยที่เกี่ยวข้อง วิธีการดำเนินการวิจัย ผลการดำเนินงานวิจัย และสรุปผลการวิจัย

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 เทคโนโลยีเอ็นเอฟซี

เทคโนโลยีเอ็นเอฟซี (Near Field Communication: NFC) [3] เป็นเทคโนโลยีสื่อสารไร้สายระยะสั้น (Short-Range Wireless Connectivity) ใช้คลื่นวิทยุความถี่ 13.56 เมกะเฮิร์ต มีความเร็วในการรับส่งข้อมูล 424 กิโลบิตต่อวินาที และมีความเร็วในการติดต่อต่ำกว่า 0.1 วินาที ระยะทางที่ติดต่อสื่อสารถึงกันมีระยะทางประมาณ 1-4 เซนติเมตร ซึ่งมีหลักการทำงานโดยการสร้างสนามแม่เหล็กจากเสาอากาศแบบขดของอุปกรณ์สองตัวที่อยู่ใกล้กัน สะท้อนและเหนี่ยวนำซึ่งกันและกัน ทำให้เกิดกระแสไฟฟ้าและคลื่นวิทยุ จากนั้นเครื่องอ่านก็จะแปลงคลื่นออกมาเป็นข้อมูล ดังแสดงในภาพที่ 1



ภาพที่ 1: การทำงานของเทคโนโลยีเอ็นเอฟซี

ลักษณะการทำงานของเทคโนโลยีเอ็นเอฟซี แบ่งเป็น 3 ลักษณะ ได้แก่

2.1.1 โหมดเพียร์ทูเพียร์ (Peer-to-Peer Mode) จะทำการแลกเปลี่ยนข้อมูลระหว่างอุปกรณ์เอ็นเอฟซีด้วยกัน เช่น นามบัตร รูปถ่าย แฟ้มข้อมูลอื่น ๆ เพียงแค่เลือกข้อมูลที่ต้องการแลกเปลี่ยน แล้วนำอุปกรณ์เอ็นเอฟซีที่รองรับโหมดนี้มาแตะหรือสัมผัสกัน ข้อมูลดังกล่าวก็จะถูกทำการถ่ายโอน

2.1.2 โหมดอ่านและเขียน (Reader and Writer Mode) อุปกรณ์เอ็นเอฟซีจะทำตัวเสมือนเป็นเครื่องอ่านเขียนแบบไร้สัมผัส โดยสามารถอ่านข้อมูลจากแท็กที่ติดอยู่ในจุดให้บริการข้อมูล จากนั้นข้อมูลก็จะปรากฏขึ้นมาบนโทรศัพท์มือถือ

2.1.3 โหมดเอ็นเอฟซีการ์ดอีมูเลชั่น (NFC Card Emulation Mode) อุปกรณ์เอ็นเอฟซีจะทำงานเสมือนบัตรไร้สัมผัส (Contactless) เพื่อใช้ในการทำธุรกรรมด้านการเงิน เช่น การจ่ายเงินชำระค่าผ่านทาง เป็นต้น

จากลักษณะการทำงานของเทคโนโลยีเอ็นเอฟซีดังกล่าว งานวิจัยนี้ได้ประยุกต์ใช้โหมดอ่านและเขียน โดยโมบายแอปพลิเคชันที่พัฒนาขึ้นจะทำหน้าที่ในการอ่านและเขียนข้อมูลยาลงในฉลากยาเอ็นเอฟซี

2.2 วรรณกรรมที่เกี่ยวข้อง

จากการศึกษางานวิจัยที่เกี่ยวข้องกับการพัฒนาเครื่องอ่านฉลากยาอัตโนมัติ สำหรับผู้บกพร่องทางการมองเห็น พบว่ามีการนำเทคโนโลยีต่าง ๆ มาใช้ ตัวอย่างเช่น การใช้เทคโนโลยีอาร์เอฟไอดี ซึ่งจะบันทึกข้อมูลยาลงในแท็กอาร์เอฟไอดีและนำแท็กดังกล่าวไปติดไว้ที่ซองยาหรือขวดยา จากนั้นจึงใช้เครื่องอ่านฉลากยาที่พัฒนาขึ้นเพื่ออ่านข้อมูลยา แต่ด้วยขนาดของเครื่องอ่านที่มีขนาดใหญ่และต้องมีการชาร์ตแบตเตอรี่ ทำให้ไม่สะดวกต่อการพกพาไปยังสถานที่ต่าง ๆ [4], [5] ต่อมามีงานวิจัยที่นำเลเซอร์สแกนและกล้องเว็บแคมมาใช้ ซึ่งได้ทำการเขียนโปรแกรมอาดูโน (Arduino) เพื่อควบคุมการทำงาน และนำเครื่องมือที่พัฒนาขึ้นไปส่องที่ฉลากยา จากนั้นเครื่องมือก็จะอ่านฉลากยาดังกล่าวให้ผู้บกพร่องทางการมองเห็นได้รับฟัง [6] นอกจากนั้นมีผู้วิจัยได้นำเทคโนโลยีเอ็นเอฟซี [7] มาใช้ในการพัฒนาเครื่องอ่านฉลากยา โดยจะเชื่อมต่อเครื่องอ่านฉลากยาเข้ากับคอมพิวเตอร์พกพาและ

แสดงผลในรูปแบบเสียงผ่านลำโพงของคอมพิวเตอร์ ซึ่งทำให้ไม่สะดวกในการพกพาและมีข้อจำกัดในการอ่านฉลากยาบนบรรจุภัณฑ์ที่มีส่วนผสมของโลหะ โดยจะส่งผลให้กระแสไฟฟ้าในฉลากยาเอ็นเอฟซีลดต่ำลงและไม่สามารถทำงานได้ [8]

สำหรับงานวิจัยนี้ได้พัฒนาโมบายแอปพลิเคชันเพื่ออ่านฉลากยาสำหรับผู้บกพร่องทางการมองเห็น ซึ่งมุ่งเน้นให้เห็นถึงการนำเทคโนโลยีเอ็นเอฟซีที่มีอยู่บนสมาร์ทโฟนมาประยุกต์ใช้ให้เกิดประโยชน์ และเป็นการเพิ่มความสะดวกในการพกพาให้แก่ผู้บกพร่องทางการมองเห็นอีกด้วย

3. วิธีการดำเนินการวิจัย

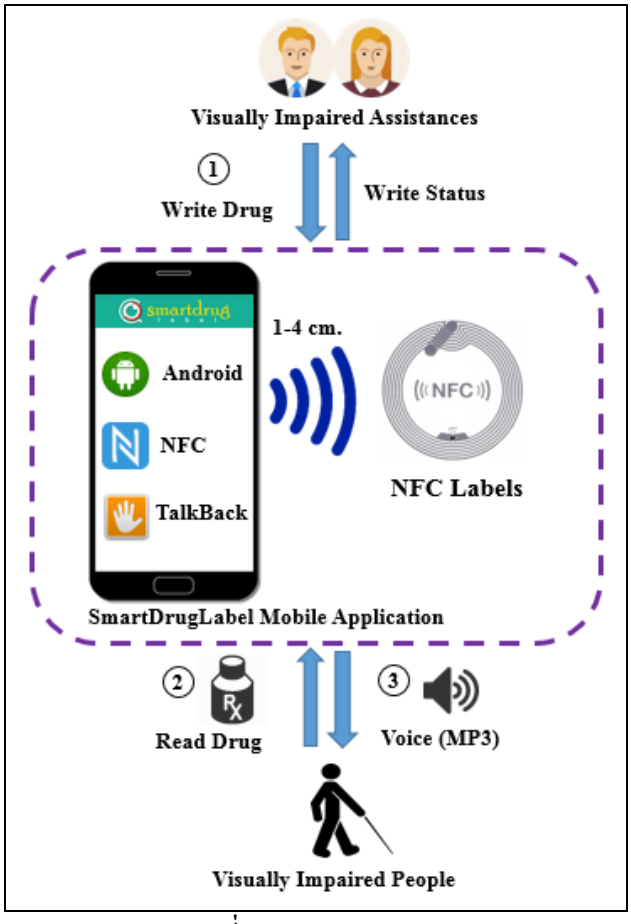
จากปัญหาความไม่สะดวกในการพกพาเครื่องอ่านฉลากยาที่มีขนาดใหญ่และปัญหาที่ผู้บกพร่องทางการมองเห็นไม่สามารถอ่านฉลากยาหรือจำแนกยาที่มีลักษณะคล้ายกันได้ คณะผู้วิจัยจึงได้นำเสนอการพัฒนาโมบายแอปพลิเคชันเพื่ออ่านฉลากยาสำหรับผู้บกพร่องทางการมองเห็น โดยใช้เทคโนโลยีเอ็นเอฟซี เพื่อช่วยให้การรับประทานยามีความสะดวกและลดความเสี่ยงจากการใช้ยาผิดพลาด ซึ่งมีวิธีการดำเนินการวิจัย 3 ขั้นตอน ดังต่อไปนี้

3.1 การศึกษาและรวบรวมข้อมูล

งานวิจัยใช้ข้อมูลยาสามัญประจำบ้านแผนปัจจุบันตามประกาศกระทรวงสาธารณสุข ฉบับที่ 5 พ.ศ. 2555 [9] โดยเลือกยาสามัญประจำบ้านแผนปัจจุบัน จำนวน 10 รายการ ได้แก่ ยาธาตุน้ำแดง ผงน้ำตาลเกลือแร่ ยาหม่องชนิดขี้ผึ้ง ยาแก้ไอน้ำดำ ไดเมนไฮดริเนท พาราเซตามอล คลอร์เฟนิรามีน ทิงเจอร์ไอโอดีน น้ำเกลือล้างแผล และคาลาไมน์

3.2 การวิเคราะห์และออกแบบแอปพลิเคชัน

จากปัญหาดังกล่าว คณะผู้วิจัยจึงได้วิเคราะห์และออกแบบกรอบแนวคิดของการพัฒนาโมบายแอปพลิเคชันเพื่ออ่านฉลากยาสำหรับผู้บกพร่องทางการมองเห็นโดยใช้เทคโนโลยีเอ็นเอฟซี ซึ่งสามารถแสดงได้ดังภาพที่ 2



ภาพที่ 2: กรอบแนวคิดของการวิจัย

3.2.1 การบันทึกรหัสยาลงในฉลากยาเอ็นเอฟซี (Write Drug) กระทำโดยผู้ดูแลผู้บกพร่องทางการมองเห็น ด้วยโมบายแอปพลิเคชันที่พัฒนาขึ้นบนระบบปฏิบัติการแอนดรอยด์ ซึ่งมีชื่อว่า SmartDrugLabel โดยจะทำการเลือกยาที่ต้องการ และนำฉลากยาเอ็นเอฟซีมาแตะที่สมาร์ทโฟนในระยะประมาณ 1-4 เซนติเมตร เมื่อบันทึกข้อมูลเรียบร้อยแล้ว แอปพลิเคชันจะแสดงสถานะการบันทึกให้ผู้ใช้ทราบ จากนั้นจึงนำฉลากยาดังกล่าวไปติดไว้ที่บรรจุภัณฑ์ยา

3.2.2 การอ่านข้อมูลในฉลากยาเอ็นเอฟซี (Read Drug) ด้วยโมบายแอปพลิเคชันดังกล่าว โดยผู้บกพร่องทางการมองเห็นจะนำผลิตภัณฑ์ยาที่ติดฉลากยาเอ็นเอฟซีแล้ว มาแตะกับสมาร์ทโฟน พร้อมทั้งเปิดการใช้งานทอล์คแบคเพื่อช่วยในการอ่านหน้าจอ ทำให้การใช้งานแอปพลิเคชันง่ายและสะดวกยิ่งขึ้น

3.2.3 การแสดงผลของโมบายแอปพลิเคชัน จะแสดงผลในรูปแบบเสียง (Voice) โดยแอปพลิเคชันจะทำการอ่านรหัสยาในฉลากยาเอ็นเอฟซี และใช้อินเทอร์เน็ตเพื่อค้นหาไฟล์เสียง

นามสกุลเอ็มพีสาม (MP3) ของรหัสยานั้น ๆ ในฐานข้อมูล และนำมาแสดงให้แก่ผู้บกพร่องทางการมองเห็นได้รับฟัง เช่น ชื่อยา สรรพคุณ วิธีการรับประทาน คำเตือน และวันหมดอายุ ซึ่งจะช่วยลดความเสี่ยงจากการใช้ยาผิดพลาดได้

3.3 การพัฒนาแอปพลิเคชัน

งานวิจัยนี้พัฒนาในรูปแบบโมบายแอปพลิเคชันบนระบบปฏิบัติการแอนดรอยด์ โดยใช้ภาษาจาวา (Java Language) และเครื่องมือ Android Studio ซึ่งเป็นเครื่องมือมาตรฐานสำหรับพัฒนาซอฟต์แวร์ของแอนดรอยด์ และใช้มายเอสคิวแอล (MySQL) ในการจัดการฐานข้อมูล โดยรายละเอียดของการพัฒนาแอปพลิเคชัน แบ่งออกเป็น 3 ส่วน ดังต่อไปนี้

3.3.1 การประยุกต์ใช้แอนดรอย์ทอล์คแบค ซึ่งเป็นโปรแกรมอ่านหน้าจอสมาร์ทโฟน เพื่อช่วยให้ผู้บกพร่องทางการมองเห็นสามารถโต้ตอบกับสมาร์ทโฟนได้ โดยทอล์คแบคจะใช้การอธิบายและอ่านออกเสียงการทำงานต่าง ๆ ซึ่งคณะผู้วิจัยได้พัฒนาโมบายแอปพลิเคชันตามมาตรฐานการเข้าถึงที่ระบบปฏิบัติการแอนดรอยด์ได้กำหนดไว้ ด้วยการเรียกใช้เมธอด ContentDescription() เพื่อให้แอนดรอยด์ทอล์คแบคสามารถอ่านรูปภาพในแอปพลิเคชันได้ ซึ่งสามารถแสดงได้ดังภาพที่ 3

```
<ImageButton
  android:layout_width="wrap_content"
  android:layout_height="wrap_content"
  android:contentDescription="กรุณาแตะฉลากยาเอ็นเอฟซี
  เพื่ออ่านข้อมูล"
  android:src="@drawable/nfc_icon"
/>
```

ภาพที่ 3: เมธอดที่ใช้สำหรับอ่านรูปภาพของแอนดรอยด์ทอล์คแบค

3.3.2 การประยุกต์ใช้เทคโนโลยีเอ็นเอฟซีในการบันทึกและอ่านรหัสยาที่อยู่ในฉลากยาเอ็นเอฟซี โดยงานวิจัยนี้มีลักษณะการทำงานแบบอ่านและเขียน (Reader and Writer Mode) ในการพัฒนาแอปพลิเคชันมีการเรียกใช้งานไลบารี่ android.nfc ซึ่งประกอบด้วยคลาส NfcAdapter NdefMessage และ NdefRecord จากภาพที่ 4 แสดงตัวอย่างเมธอด (Method) ที่ใช้สำหรับอ่านข้อมูลในฉลากยาเอ็นเอฟซี

```
private String readText(NdefRecord record) throws
UnsupportedEncodingException
{
        byte[] payload = record.getPayload();
        String textEncoding = ((payload[0] & 128)
           == 0) ? "UTF-8" : "UTF-16";
        int languageCodeLength = payload[0] &
           0063;
        return new String(payload,
           languageCodeLength + 1, payload.length -
           languageCodeLength - 1, textEncoding);
}
```

ภาพที่ 4: เมธอดที่ใช้สำหรับอ่านข้อมูลในฉลากยาเอ็นเอฟซี

3.3.3 การแสดงผลในรูปแบบเสียง แอปพลิเคชันจะอ่านไฟล์เสียงนามสกุลเอ็มพีสาม (MP3) ให้แก่ผู้บกพร่องทางการมองเห็นได้รับฟังผ่านลำโพงหรือหูฟังของสมาร์ทโฟน ซึ่งข้อมูลยาจะประกอบด้วย ชื่อยา สรรพคุณ วิธีการรับประทาน คำเตือน และวันหมดอายุ โดยการพัฒนาในส่วนนี้เรียกใช้งานไลบรารี่ android.media คลาส MediaPlayer จากภาพที่ 5 แสดงตัวอย่างเมธอด (Method) ที่ใช้สำหรับแสดงผลในรูปแบบเสียง

```
String urlAudio =
"http://smartdruglabel.com/uploaded/" +AudioName;
    MediaPlayer mediaPlayer = new MediaPlayer();
    mediaPlayer.setAudioStreamType
    (AudioManager.STREAM_MUSIC);
try
{    mediaPlayer.setDataSource(urlAudio);
     mediaPlayer.prepare();
     mediaPlayer.start();
} catch (IOException ex)
{
     Log.e(TAG, "Could not open file", ex);
}
```

ภาพที่ 5: เมธอดที่ใช้สำหรับแสดงผลในรูปแบบเสียง

4. ผลการดำเนินงานวิจัย

4.1 ผลการพัฒนาแอปพลิเคชัน

โมบายแอปพลิเคชันบนระบบปฏิบัติการแอนดรอยด์ที่พัฒนาขึ้น สามารถบันทึกรหัสยาลงในฉลากยาเอ็นเอฟซีได้ ซึ่งกระทำโดยผู้ดูแลผู้บกพร่องทางการมองเห็น ด้วยการเลือกยาที่ต้องการ จากนั้นนำฉลากยาเอ็นเอฟซีมาแตะที่สมาร์ทโฟนเพื่อบันทึกรหัสผลิตภัณฑ์ยานั้น ๆ สามารถแสดงได้ดังภาพที่ 6



ภาพที่ 6: หน้าจอการบันทึกรหัสยาลงในฉลากยาเอ็นเอฟซี

จากนั้นผู้บกพร่องทางการมองเห็นจะนำผลิตภัณฑ์ยาที่ติดฉลากยาเอ็นเอฟซีมาแตะกับสมาร์ทโฟนเพื่ออ่านรหัสยาและส่งผลลัพธ์ออกมาในรูปแบบเสียง สามารถแสดงได้ดังภาพที่ 7



ภาพที่ 7: หน้าจอการอ่านฉลากยาเอ็นเอฟซี

4.2 ผลการทดสอบความถูกต้องในการอ่านฉลากยา

ทำการทดสอบความถูกต้องของการอ่านข้อมูลยาจากฉลากยาเอ็นเอฟซี ด้วยสูตรการคำนวณค่าความถูกต้อง (Accuracy) ดังสมการที่ 1

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FN + FP} \quad (1)$$

เมื่อ TP คือ จำนวนผลการอ่านฉลากยาที่มีข้อมูลในฐานข้อมูล ว่ามีค่าถูกต้อง

TN คือ จำนวนผลการอ่านฉลากยาที่ไม่มีข้อมูลในฐานข้อมูล ว่ามีค่าไม่ถูกต้อง

FN คือ จำนวนผลการอ่านฉลากยาที่มีข้อมูลในฐานข้อมูล ว่ามีค่าไม่ถูกต้อง

FP คือ จำนวนผลการอ่านฉลากยาที่ไม่มีข้อมูลในฐานข้อมูล ว่ามีค่าถูกต้อง

ในการทดสอบความถูกต้องของการอ่านข้อมูลยาจากฉลากยาเอ็นเอฟซี ทดสอบด้วยการสุ่มหยิบซองยาหรือขวดยาจำนวน 10 ชนิด ชนิดละ 30 ครั้ง ดังแสดงในตารางที่ 1

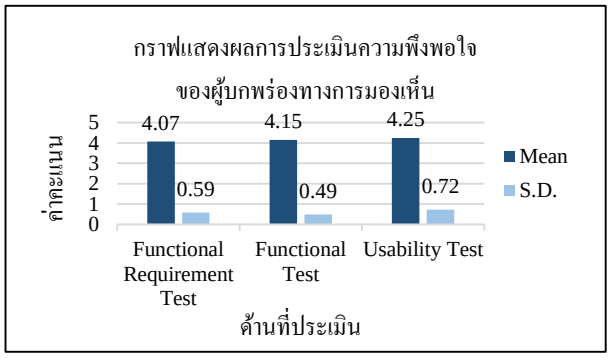
ตารางที่ 1: ผลการทดสอบความถูกต้องในการอ่านฉลากยา

ลักษณะบรรจุภัณฑ์	ชื่อยา	จำนวนครั้งในการทดสอบ	จำนวนครั้งที่อ่านถูกต้อง	ความถูกต้อง (%)
ขวดหรือซองพลาสติก	น้ำเกลือล้างแผล	30	30	100
	พาราเซตามอล	30	30	100
	คลอร์เฟนิรามีน	30	30	100
	ทิงเจอร์ไอโอดีน	30	30	100
	คาลาไมน์	30	30	100
ขวดแก้ว	ยาธาตุน้ำแดง	30	30	100
	ยาแก้ไอน้ำดำ	30	30	100
	ยาหม่องขี้ผึ้ง	30	30	100
แผงหรือซองฟอยล์	ผงน้ำตาลเกลือแร่	30	15	50
	ไดเมนไฮดริเนท	30	19	63.33
รวม		300	274	91.33

ผลการทดสอบความถูกต้องในการอ่านข้อมูลยาจากฉลากยาเอ็นเอฟซี พบว่าแอปพลิเคชันอ่านข้อมูลได้ถูกต้องจำนวน 274 ครั้ง จากจำนวนทั้งหมด 300 ครั้ง มีค่าความถูกต้องเท่ากับ 91.33% ซึ่งความผิดพลาดที่เกิดขึ้นเนื่องจากเทคโนโลยีเอ็นเอฟซีไม่สามารถทำงานได้บนบรรจุภัณฑ์ที่ผลิตจากวัสดุที่มีส่วนผสมของโลหะ ได้แก่ ซองยาหรือแผงยาที่ผลิตจากอลูมิเนียมฟอยล์ ซึ่งเป็นโลหะชนิดหนึ่ง

4.3 ผลการประเมินความพึงพอใจของแอปพลิเคชัน

ผลการประเมินความพึงพอใจของผู้บกพร่องทางการมองเห็น จำนวน 5 ท่าน จากสมาคมคนตาบอดแห่งประเทศไทย ที่มีต่อแอปพลิเคชัน ด้านความสามารถของแอปพลิเคชันตามความต้องการของผู้ใช้งาน (Functional Requirement Test) ด้านการทำงานได้ตามฟังก์ชัน (Functional Test) และด้านความสะดวกและง่ายต่อการใช้งาน (Usability Test) อยู่ในระดับดี มีค่าเฉลี่ยเท่ากับ 4.16 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.60 แสดงรายละเอียดดังภาพที่ 8



ภาพที่ 8: ผลการประเมินความพึงพอใจของผู้บกพร่องทางการมองเห็น

5. สรุปผล

จากการพัฒนาโมบายแอปพลิเคชันเพื่ออ่านฉลากยาสำหรับผู้บกพร่องทางการมองเห็น โดยใช้เทคโนโลยีเอ็นเอฟซีร่วมกับแอนดรอยด์ทอล์คแบค พบว่าแอปพลิเคชันสามารถแสดงผลในรูปแบบเสียงให้แก่ผู้บกพร่องทางการมองเห็นได้รับฟัง ได้แก่ ชื่อยา สรรพคุณ วิธีการรับประทาน คำเตือน และวันหมดอายุ ทำให้สามารถรับประทานยาได้ด้วยตนเองอย่างสะดวกและช่วยลดความเสี่ยงจากการใช้ยาผิดพลาด จากผลการทดสอบหาค่าความถูกต้องในการอ่านฉลากยาเอ็นเอฟซี พบว่ามีค่าความถูกต้องเท่ากับ 91.33% เนื่องจากซองผงน้ำตาลเกลือแร่และแผงยาไดเมนไฮดริเนทมีส่วนผสมของโลหะ ซึ่งจะส่งผลให้กระแสไฟฟ้าในฉลากยาเอ็นเอฟซีลดต่ำลงและไม่สามารถทำงานได้ โดยปัญหานี้สามารถแก้ไขด้วยการติดฉลากยาเอ็นเอฟซีที่ซองยาตรงบริเวณที่ไม่ใช่โลหะหรือนำซองยาแบบกระดาษมาใส่ผลิตภัณฑ์ยานั้น ๆ ก่อน แล้วจึงติดฉลากยาเอ็นเอฟซี เพื่อเพิ่มความถูกต้องในการอ่านฉลากยาให้มากขึ้น และผลการประเมินความพึงพอใจต่อการใช้งานแอปพลิเคชันของผู้บกพร่องทางการมองเห็นอยู่ในเกณฑ์ดีในทุกด้าน

การวิจัยในอนาคตสามารถพัฒนาต่อยอดให้แอปพลิเคชันแปลงข้อความให้เป็นไฟล์เสียงอัตโนมัติ โดยการนำ Google-Text-To-Speed มาประยุกต์ใช้เพื่อลดขั้นตอนในการนำเข้าข้อมูลยา หรือพัฒนาแอปพลิเคชันให้มีการเชื่อมต่อกับฐานข้อมูลของเว็บไซต์ Ya&You ที่พัฒนาโดยศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ ซึ่งให้บริการข้อมูลด้านยาและสุขภาพ หรือพัฒนาแอปพลิเคชันให้มีการแจ้งเตือน (Notification) เมื่อถึงเวลารับประทานยา เพื่อช่วยเหลือผู้บกพร่องทางการมองเห็นให้ได้รับความสะดวกมากขึ้น

เอกสารอ้างอิง

[1] กรมการส่งเสริมและพัฒนาคุณภาพชีวิตคนพิการ. (2560). [ออนไลน์]. *สถิติข้อมูลคนพิการที่มีบัตรประจำตัวคนพิการจำแนกตามจังหวัด ประเภทความพิการ และเพศ.* [สืบค้นวันที่ 16 เมษายน 2560]. จาก https://www.m-society.go.th

[2] World Health Organization. (2014). [online]. *Visual impairment and blindness.* [cite 25 March 2017]. Available from: URL: http://www.who.int/mediacentre/factsheets/fs282/en

[3] S. Lermthong, P. Suwanna and S. Airphaiboon, "Bedside patient monitoring by NFC," *2016 9th Biomedical Engineering International Conference (BMEiCON)*, Laung Prabang, 2016, pp. 1-4.

[4] พีระพงศ์ ฟักเขียว. (2556). *การพัฒนาเครื่องอ่านฉลากยาสำหรับผู้ที่มีปัญหาทางด้านการมองเห็น โดยใช้ RFID.* วิทยานิพนธ์ครุศาสตร์อุตสาหกรรมมหาบัณฑิต สาขาวิชาวิศวกรรมไฟฟ้า ภาควิชาครุศาสตร์ไฟฟ้า คณะครุศาสตร์อุตสาหกรรม มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ.

[5] T. Kornsingha and P. Punyathep, "A voice system, reading medicament label for visually impaired people," *RFID SysTech 2011; 7th European Workshop on Smart Objects: Systems, Technologies and Applications*, Dresden, Germany, 2011, pp. 1-6.

[6] สุชาครีย์ เลวรรณ์ รวิรุจ บุตโคษา และนิลาวัลย์ หมีกุละ. (2559). *เครื่องอ่านสลากยาสำหรับผู้พิการทางสายตา.* ภาควิชาวิศวกรรมเมคาทรอนิกส์ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยมหาสารคาม.

[7] M. Harjumaa, M. Isomursu, S. Muuraiskangas and A. Konttila, "HearMe: A touch-to-speech UI for medicine identification," *2011 5th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth) and Workshops*, Dublin, 2011, pp. 85-92.

[8] T. Sekiguchi, Y. Okano, N. Ohmura and S. Ogino, "Study on effective pattern of magnetic sheet attached on NFC antenna," *2015 IEEE International Symposium on Radio-Frequency Integration Technology (RFIT)*, Sendai, 2015, pp. 208-210.

[9] ราชกิจจานุเบกษา. (วันที่ 29 พฤศจิกายน 2555). ประกาศกระทรวงสาธารณสุข. *ยาสามัญประจำบ้านแผนปัจจุบัน (ฉบับที่ 5) พ.ศ.2555.* เล่ม 130 ตอนพิเศษ 9 ง : 2.

AUTHOR INDEX

	Page
Aimaschana Niruntasukrat	680
Amnat Gongsuk	731
Anan Chanlun	515
Anawat Leatongkam	274
Angsuma Thongthiew	521
Apichat Buatoom	76
Apisit Saengsai	287
Araya Pudtal	409
Arnon Chonrawut	618
Asamaporn Chatrattikorn	767
Auemporn Rungsiri	792
Aziz Nanthaamornphong	267, 274
Boonchom Sudjit	331
Boonprasert Surakratanasakul	624
Buncha Joompol	88
Bundit Thanasopon	534, 569
Bunthit Watanapa	293, 521
Burasakorn Yoosooka	612
Burhan Wanglem	361
Chakchai So-In	427, 515, 737, 779, 785
Chaloemchai Lowongtrakool	70
Chanida Tawornrach	515
Chaovalit Somboonpattanakit	217
Charinthorn Aumgri	826
Charnsak Srisawatsakul	299
Chartchai Leenawong	521, 674
Charun Sanrch	293
Chatsayam Philam	637
Chatsuda Ninsakul	587
Chayakorn kongtad	465
Chayaporn Kaensar	229

AUTHOR INDEX

	Page
Chisanupong Sooksawang	624
Chonnika Punjunsing	785
Chotika Iamprik	32
Chutipon Blumrungjit	792
Dechawut Wanichsan	344, 761, 792, 814
Donlaya Suriyawong	779
Donruedee Rapisuwan	385
Duangkamol Sodalee	421
Dunyarat kawviset	581
Ekkalak Chimjan	217
Esther Sangiamkul	606
Guntapich Anantasak	551
Hasuenee Masalong	247
Ittikon Thammachantuek	217
Janjao Mongkolnavin	540
Jantima Polpinij	88, 261, 397
Jaree Thongkam	8, 311, 421
Jarunan Boonwittayalikit	692
Jedsada Boonsiri	205
Jetsada Meela	261
Jirasak Jullawat	100
Jiruttigarn Gunteemoon	737
Kaewjai Apornpisarn	1
Kasidech Tapang	540
Khamron Sunat	253, 483, 749
Khongthep Boonme	612
Khwanchanok Srimool	680
Kiattikun Kawila	100
Kingkanjana Khotkaew	779
Kitsana Waiyamai	439, 546
Kitti Mala	663

AUTHOR INDEX

	Page
Komsak Ponpruksa	761
Krerk Piromsopa	667
Krich Sintanakul	305, 355
Krisada Khrueawan	427
Kulit Na Nakorn	100
Kultida Rojviboonchai	100
Kunlaya Charoenmongkonvilai	350
Kuntpong Woraratpanya	163, 169
Kwan Sitathani	338
Kwanruetai Sudadet	32
Luxsanan Ploywattanawong	144
Mahasak Ketcham	137
Maleerat Sodanil	26, 205, 215, 657, 713
Mananchaya Amornpalang	379
Manoch Pracha	199
Manutchai Preechahan	121
Massaya Samnienggam	57
Monlica Wattana	427
Monsawan Ploymookda	76
Montean Rattanasiriwongwut	527, 692
Nakornthip Prompoon	126
Nalinpat Porrawatpreyakorn	637
Naphong Wannapiroon	826
Narisara Nakmaetee	63
Narongdech Keeratipranon	14, 433
Narongdech Rattananonsathien	76
Narroup Rak-ngam	14
Narumon Orachon	344
Natchapol Surarattanagon	820
Nathee Yongyut	820
Natruja Suwannapak	837

AUTHOR INDEX

	Page
Nattapong Tongtep	247, 361, 744
Nattavadee Hongboonmee	241, 587
Nattavee Utakrit	551, 593, 719, 725, 731
Nattaya Jitpongtae	54
Nawaporn Wisitpongphan	637, 650
Ngammij Arch-Int	515, 643, 706
Nichakorn Pankong	63
Nichamon Uraisriphong	534
Nitat Ninchawee	281
Nithaya Sakolpisit	163, 169
Nithipat Wongchingchai	767
Niti Witthayawiroj	832
Nitinan Mata	70
Nitit Wangno	483
Nittaya Phoda	253
Nivet Chirawivhitchai	324
Noppamas Pukkhem	453
Nuansri Denwattana	287
Nunnapus Benjamas	51, 427
Nuttakit Vatcharatiansakul	18
Nuttanont Hongwarittorrn	175, 497
Orawan Liabsiri	367
Paisan Simalaotao	773
Panchit Longpradit	756
Panida Lorwongtrakool	70
Panita Pongpaibool	680
Panjarat Hanpanich	509
Pannapa Nuanplord	713
Panwit Tuwanut	106
Paphonput Sopon	397
Paravee Pattarakavin	667

AUTHOR INDEX

	Page
Parinya Natho	150
Parisa Pattarathanawat	719
Patcharaon Petchsamutr	137
Patcharee Chareonlap	820
Patcharee Oankaew	38
Patipat Tumsangthong	767
Pattamaporn Kengkarnrob	211
Pattamavadee Kroeiram	311
Pattanaphong Donratand	8
Pattanaphong Pothipassa	477
Pattanee Arboonngam	706
Pattarapat Waidee	156
Pattaraporn Chotnasiaw	44
Paveena Chamking	826
Paweena Wanchai	51, 643
Payu Poonsri	810
Petchpailin Saeauy	810
Phatchareeya Meetha	600
Phayung Meesad	1, 20, 373
Phet Imtongkhum	737, 779
Phinyawat Sookgaew	453
Phollawat Wattanasiriyanon	814
Phonepasuerth Bounnaphonh	756
Phongdanai Jittavisuttikul	305
Phongphet Thaporn	235
Phuttasak Somchai	223
Pichamon Sinlalab	187
Pichayapa Maprajuab	749
Pisut Norkaew	569
Pita Jarupunphol	433
Piyanoot Voraboot	20

AUTHOR INDEX

	Page
Piyaporn Apisuwan	785
Piyasuda Tanloet	798
Pongkun Pankerd	181
Pongpith Tuenpusa	832
Pongpon Nilaphruek	612
Pongsakorn Naklamai	725
Ponrudee Netisopakul	534
Porawat Visutsak	187
Porawat Visutsak	193
Pornrudee Sittipong	344
Pornsak preelakha	205
Porntheera Chanta	403
Porntip Boontem	247
Pornvimon Klongsangson	38
Pramote Tongchin	756
Prasong Praneetpolgrang	311, 324
Pratchaya Dokmaingam	490
Pratchaya Jaisudthi	564, 663
Pratya Sutee	261
Prayoot Surasena	217
Preecha Kangngulam	792
Pudsadee Boonrawd	837
Pusadee Seresangtakul	557
Puttida Sakulviriyakitkul	674
Puttipong Leakkaw	82
Rujika Booncherd	344
Rungsan Suwannahong	699
Ruttasiri Tessiri	126
Saiyan Saiyod	44, 379, 515, 575
Sakchai Tangwannawit	385, 447, 459, 600
Sakkarin Phupanna	51

AUTHOR INDEX

	Page
Salyapong Wichaidit	350
Samlee Mankhetkorn	439
San Ratanasanya	187, 193
Saowalak Arampongsanuwat	63
Saowaluk Watanapa	211, 367, 471
Sathaporn Thongwik	832
Sinapas Juanngarm	415
Sippawit Saewang	163, 169
Sirapat Boonkrong	144, 150
Sirapat Chiewchanwattana	253
Sirichai Triamlumlerd	199
Sirikul Siriteerakul	810
Sirikunya Nilpanich	132
Sirimongkon Meesiri	156
Siripa Jantabon	318
Sirivimon Voraditpatthana	490
Siriwan Moolmao	593
Sithikorn Charoensub	106
Sittipon Prajan	241
Siwalai Chinchua	826
Soawanee Prachayagringkai	205
Somjit Arch-int	44, 643, 706
Somkid Saelee	686, 804
Sompong Valuvanathon	223
Sonchai Boonluea	663
Songsak Rongviriyapanich	630
Sooksan Panichpapiboon	82
Sorawit Yotalai	415
Sucha Smanchat	113, 121
Suchart Rommaneeyarak	509
Sudarat Worawetwichai	650

AUTHOR INDEX

	Page
Sukanchalika Boonmatham	205, 373
Sukanya Khayantham	706
Sukita Boonraum	804
Sukree Sinthupinyo	409
Sukumal Kitisin	680
Sumet Prabhavat	94
Sunee Pongpinigpinyo	181
Supaluck Mongkhonrat	737
Supansa Najaitrong	483
Supathep Satiman	20, 217
Supawadee Klommit	447
Supitcha Saengchat	527
Supranee Horma	205
Surachet Yuwarat	564
Suraiya Charornreh	24
Sureechay Meemoo	38
Sureerut Inmor	699
Suriya Sukarin	832
Sutat Gammanee	217
Suthasinee Iamsa-at	749
Suthat Banchong	88
Sutthikarn Bojukrapan	798
Suwimon Huangkasem	344
Tanagrit Chansaeng	433
Tanakul Pakkawesa	557
Tananop Thongthavorn	94
Tanaporn Panrungsri	606
Tanasanee Phienthrakul	57, 391
Tanawat Gajaseni	503
Tanawit Tattakittiya	749
Taranop Sairoon	471

AUTHOR INDEX

	Page
Tareenichar Leepeeravitit	338
Tawa Khampachua	650, 686
Tawatchai Kovicsakul	94
Teera Siriteerakul	810
Teerapat Khunpech	113
Thanakit Tantithanakorn	624
Thanchanok Naowaratkul	459
Thanisa Numnonda	503
Thanrat Sintanakul	674
Thanyarat Kitpanich	267
Thapanee Chayakul	680
Theanchai Ounoi	439
Thearasak Phaladisailoed	503
Theraphat Satheanphong	630
Thiphawan Sawaengmee	483
Thongchai Chuachan	156
Thotsapon Tamduang	181
Tosaporn Saiyim	465, 490
Trirat Pat-iam	744
Trirat Soemsap	38
Trithep Ratanapipop	618
Tuangsit Tirabutr	181
Ubonrat Sirisukpoca	773
Unchalee Rungthian	657
Vatinee Nuipian	1
Veerachai Khonchoho	756
Wandee Chokchueypattanakit	281
Wannapa Phopli	205, 373
Wanvisa Krutthai	293
Waransanang Boontarig	299
Wararat Songpan	44, 51

AUTHOR INDEX

	Page
Warisara Thumwirat	534
Warunee Tamkoo	546
Wassana Yaemprai	826
Watchapon Suksakul	569
Watcharapong Thorat	814
Watcharathorn Laikhunakhorn	94
Watcharee Waratchareeyakul	814
Watcharin Nampanya	235
Wattanachai Wattanasingha	465
Weerachai Chansud	767
Weerayuth Sukkhamduang	575
Wichian Chutimaskul	338
Wilawan Rukpakavong	132
Winai Phengpinyo	355
Wiriya Lawanpreda	175
Wisan Tangwongcharoen	403
Wisit Wongwilai	767
Wongkot Sriurai	235
Worapoj Suwanpipob	643
Worasit Choochaiwattana	318, 581
Yanee Lamjong	761
Yuttana Lungkatoong	26
Yuttasak Jinaporn	193



THE 14TH NATIONAL CONFERENCE ON COMPUTING AND INFORMATION TECHNOLOGY

5TH – 6TH JULY 2018

FACULTY OF INFORMATION TECHNOLOGY
KING MONGKUT'S UNIVERSITY OF TECHNOLOGY NORTH BANGKOK

WWW.NCCIT.NET

Conference Organizers



ECU



OSU


FernUni



HNUE



KMUTNB


RMUTT


TU Chemnitz


Gesellschaft für Informatik




UBU


SSKRU


KRU


NPRU


PBRU


RMUTK


MSU


PSU


NIDA



King Mongkut's University of Technology North Bangkok
Rajamangala University of Technology Thanyaburi
FernUniversitaet in Hagen, Germany
Chemnitz University of Technology, Germany
Oklahoma State University, USA
Edith Cowan University, Australia
Hanoi National University of Education, Vietnam
Gesellschaft für Informatik
Ubon Ratchathani University
Kanchanaburi Rajabhat University
Nakhon Pathom Rajabhat University
Sisaket Rajabhat University
Phetchaburi Rajabhat University
Rajamangala University of Technology Krungthep
Mahasarakham University
Prince Of Songkla University, Phuket Campus
National Institute of Development Administration
Council of IT Deans of Thailand
IEEE Computer Intelligence Society Thailand Chapter

ISBN 978-616-368-044-0

