CONFERENCE PROCEEDINGS

# nccit 2026

# THE 22nd NATIONAL Conference

## CONFERENCE ON COMPUTING AND INFORMATION TECHNOLOGY




14th - 15th MAY 2026

www.nccit.net

2026

The Association of council of IT Deans
King Mongkut's University of Technology North Bangkok
Faculty of Information Technology and Digital Innovation

# TABLE OF CONTENTS

Message from KMUTNB President....ii
Message from General Chair ....iv
Conference Organizers....viii
Conference Organization Committee ....x
Steering Committee ....xi
Technical Program Committee ....xii
Technical Program Contents....xiii
Regular Papers ....1
Author Index ....557

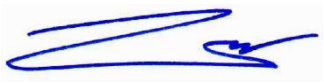
# สารจากอธิการบดี
# มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือมีความยินดีเป็นอย่างยิ่งที่ได้ให้การสนับสนุนและมีส่วนร่วมในการจัดการประชุมวิชาการระดับชาติด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศ (The National Conference on Computing and Information Technology: NCCIT) ซึ่งเป็นเวทีสำคัญสำหรับนักวิชาการ นักวิจัย คณาจารย์ นักศึกษา ตลอดจนผู้เชี่ยวชาญ ได้ร่วมแลกเปลี่ยนองค์ความรู้ ผลงานวิจัย และประสบการณ์ทางวิชาการในสาขาวิทยาการคอมพิวเตอร์และเทคโนโลยีสารสนเทศ

ในยุคดิจิทัลที่มีการเปลี่ยนแปลงอย่างรวดเร็วในปัจจุบัน ความก้าวหน้าทางเทคโนโลยีได้มีบทบาทสำคัญต่อการพัฒนาเศรษฐกิจ การเปลี่ยนแปลงทางสังคม และการยกระดับคุณภาพชีวิตของประชาชน เทคโนโลยีสมัยใหม่ อาทิ ปัญญาประดิษฐ์ การวิเคราะห์ข้อมูลขนาดใหญ่ ระบบอัตโนมัติอัจฉริยะ ความมั่นคงปลอดภัยทางไซเบอร์ และเทคโนโลยีดิจิทัลเชิงนวัตกรรมอื่น ๆ ยังคงเป็นแรงขับเคลื่อนสำคัญต่อการพัฒนาในหลากหลายภาคส่วน การประชุมวิชาการครั้งนี้จึงเป็นโอกาสอันดีสำหรับนักวิจัยและผู้นำเสนอผลงานวิจัยที่มีคุณภาพ และส่งเสริมความร่วมมือทางวิชาการ อันจะนำไปสู่การพัฒนาองค์ความรู้และนวัตกรรมเพื่อประโยชน์ต่อสังคมและประเทศชาติต่อไป

ในนามของมหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ ข้าพเจ้าขอแสดงความชื่นชมและขอบคุณคณะผู้จัดงาน คณะกรรมการวิชาการ ผู้ทรงคุณวุฒิ ตลอดจนผู้มีส่วนเกี่ยวข้องทุกท่านที่ได้ร่วมกันทุ่มเทแรงกายแรงใจในการจัดการประชุมวิชาการครั้งนี้ให้สำเร็จลุล่วงไปด้วยดี

ขออวยพรให้การประชุมวิชาการระดับชาติด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศประสบความสำเร็จตามวัตถุประสงค์ และก่อให้เกิดประโยชน์ทางวิชาการ ตลอดจนส่งเสริมความร่วมมืออย่างต่อเนื่องระหว่างนักวิจัยและผู้เชี่ยวชาญในสาขาที่เกี่ยวข้องสืบไป



(ศาสตราจารย์ ดร.ธานินทร์ ศิลป์จารุ)
อธิการบดีมหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

# MESSAGE FROM KMUTNB PRESIDENT

King Mongkut's University of Technology North Bangkok is honored to support and participate in the organization of the National Conference on Computing and Information Technology (NCCIT). This conference serves as an important platform for academician, researchers, faculty members, students, and experts to exchange knowledge, research findings, and academic experiences in the fields of computer science and information technology.

Digital era has its rapid transformation, technological advancements have played a significant role in shaping economic development, social transformation, and the improvement of people's quality of life. Emerging technologies such as Artificial Intelligence, Data analytics, Intelligent Automation, Cybersecurity, and other Innovative Digital Technologies continue to drive progress across various sectors. Therefore, this conference provides an excellent opportunity for researchers and practitioners to present high-quality research and foster academic collaboration that will contribute to the advancement of knowledge and innovation for the benefit of society and the nation.

On behalf of King Mongkut's University of Technology North Bangkok, I would like to express my sincere appreciation to the organizing committee, the academic committee, distinguished reviewers, and all contributors whose dedication and efforts have made this conference complete.

I wish the National Conference on Computing and Information Technology (NCCIT) every success and hope that it will provide valuable academic contributions and promote continued collaboration among researchers and experts in the field.



(Professor Dr.Thanin Silpcharu)

President, King Mongkut's University of Technology North Bangkok

# สารจากประธานจัดงานประชุมวิชาการ
# มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

งานประชุมวิชาการระดับชาติและระดับนานาชาติด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศ (NCCIT และ IC2IT) จัดขึ้นอย่างต่อเนื่องเป็นประจำทุกปี ตั้งแต่ พ.ศ. 2548 การจัดประชุมวิชาการนี้เป็นความร่วมมือของหน่วยงานเครือข่ายวิชาการจัดงานประชุมวิชาการ NCCIT และ IC2IT เป็นอีกหนึ่งเวทีแห่งการแลกเปลี่ยนแนวความคิด ความรู้ และการวิจัยเพื่อสร้างแรงบันดาลใจสำหรับการสรรค์สร้างนวัตกรรมดิจิทัลใหม่ ๆ เป็นเวทีที่เปิดโอกาสให้อาจารย์ นักวิจัย นักศึกษา และบุคคลทั่วไปที่สนใจได้มีโอกาสนำเสนอผลงานวิจัย แลกเปลี่ยนความคิดเห็น และสร้างเครือข่ายความร่วมมือทางวิชาการที่ก่อให้เกิดการพัฒนาทางเทคโนโลยีและนวัตกรรมที่เป็นประโยชน์ต่อสังคมและประเทศชาติ ในครั้งนี้การประชุมระดับชาติด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศ (NCCIT) ครั้งที่ 22 จัดขึ้นระหว่างวันที่ 14-15 พฤษภาคม พ.ศ. 2569

ในงานประชุมวิชาการครั้งนี้ต้องขอขอบคุณทุกท่านที่ได้ส่งบทความ เพื่อเข้าสู่กระบวนการการพิจารณาซึ่งเป็นบทความที่มีคุณภาพสูง ทำให้งานประชุมวิชาการได้รักษาคุณภาพและมาตรฐานของงาน ขอขอบพระคุณคณะกรรมการพิจารณาบทความทุกท่านที่เสียสละเวลาอันมีค่า เพื่อให้ได้ผลงานวิจัยที่มีคุณภาพได้นำเสนอในงานประชุมวิชาการในครั้งนี้ด้วย

ขอขอบคุณมหาวิทยาลัยเครือข่ายทั้ง 6 แห่ง ที่ให้การสนับสนุนเป็นอย่างดีทั้งด้านคณาจารย์ บุคลากรในการร่วมเป็นคณะกรรมการเครือข่าย คณะกรรมการดำเนินงาน และคณะกรรมการพิจารณาบทความ อีกทั้งด้านงบประมาณ

ขอขอบพระคุณ ศาสตราจารย์ ดร.ธานินทร์ ศิลป์จารุ อธิการบดี มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ สำหรับการสนับสนุนการประชุมวิชาการ NCCIT และ IC$^2$IT ขอขอบคุณ ผู้ช่วยศาสตราจารย์ ดร.ศักดิ์ชาย ตั้งวรรณวิทย์ ผู้ช่วยศาสตราจารย์ ดร.วัชรีวรรณ จิตต์สกุล นางสาวพรพิมล ฝ้ายเทศ และทีมงานจากคณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ ที่ได้ประสานงานกับนักวิจัยและผู้ที่เกี่ยวข้อง เพื่อรวบรวมผลงานการวิจัยสำหรับงานประชุมทางวิชาการในครั้งนี้ให้สำเร็จลุล่วง

ท้ายนี้หวังเป็นอย่างยิ่งว่างานประชุมวิชาการ NCCIT และ IC$^2$IT จะเป็นเวทีในการนำเสนอผลงานวิจัยและนำไปสู่การพัฒนาตนเอง เพื่อให้เป็นนักวิจัยที่ดีและมีคุณภาพสูงยิ่ง ๆ ขึ้นไป NCCIT และ IC2IT จะเป็นแพลตฟอร์มสำหรับการแลกเปลี่ยนความรู้ในด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศ สร้างแรงบันดาลใจให้กับนักวิจัยเพื่อคิดค้นนวัตกรรมและพัฒนาองค์ความรู้ใหม่ ๆ เพื่อการพัฒนาชาติอย่างยั่งยืนตลอดไป



(ผู้ช่วยศาสตราจารย์ ดร.สุนันทา สดสี)
คณบดีคณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล
มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
ประธานเครือข่ายงานประชุมวิชาการ NCCIT

# MESSAGE FROM GENERAL CHAIR

The National and International Conference on Computer and Information Technology (NCCIT and IC2IT) has been continuously held every year since 2005. This conference represents a collaborative effort among academic network organizations hosting NCCIT and IC2IT, serving as a platform for exchanging ideas, knowledge, and research to inspire the creation of new digital innovations. It provides opportunities for faculty members, researchers, students, and interested individuals to present their research work, exchange opinions, and establish academic collaboration networks that drive technological advancements and innovations beneficial to society and the nation.

This year, the 22nd National Conference on Computer and Information Technology (NCCIT) will be held from May 14-15, 2026. The proceedings of this conference compile well-curated research contributions by researchers and academics who have actively participated in this event.

A debt of gratitude is owed to our five university partners in Thailand for their financial support and for the contribution of all staff members who are serving on network committee, executive director committee, and technical program committee.

Exceptional thanks are extended to Professor Dr.Thanin Silpcharu, President of KMUTNB, for his unwavering support for NCCIT & IC2IT.

A very special thanks goes to Assistant Professor Dr. Sakchai Tangwannawit, Assistant Professor Dr. Watchareewan Jitsakul, Miss Pornpimon Faythet, and all academic staff members from the Faculty of Information Technology and Digital Innovation (ITD), KMUTNB for successfully organizing the conference.

This conference could not happen without innovative and progressive contributions from the research community at large. Special thanks are due to all our technical committee members for their diligent peer review and for continuously maintaining a high academic standard for which NCCIT & IC2IT are justifiably renowned.

Finally, looking to the future, NCCIT & IC2IT will continue to serve as a community platform for research, while raising the bar for high quality academic investigation. It will provide an enduring forum for knowledge exchange in computer and information technology, and inspire connoisseurs to develop innovations, to exchange ideas and to build up a research network geared towards sustainable development.



Assistant Professor Dr. Dr.-Ing. Sunantha Sodsee
Dean, Faculty of Information Technology and Digital Innovation
King Mongkut's University of Technology North Bangkok
NCCIT2026 General Chair

# CONFERENCE ORGANIZERS

Council of IT Deans of Thailand (CITT), Thailand

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

Fern Universität in Hagen, Germany

Oklahoma State University, USA

Edith Cowan University, Australia

Hanoi National University of Education, Vietnam

Kanchanaburi Rajabhat University, Thailand

# CONFERENCE ORGANIZERS

Ubon Ratchathani University, Thailand

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

Mahasarakham University, Thailand
Faculty of Informatics

National Institute of Development Administration, Thailand
Information Technology Management Program,
Graduate School of Applied Statistics

# CONFERENCE ORGANIZATION COMMITTEE

| | |
|---|---|
| General Chairs: | Dr. Dechanuchit Katanyutaveetip, CITT |
| | Asst. Prof. Dr. Dr.-Ing. Sunantha Sodsee, KMUTNB |
| Secretary: | Asst. Prof. Dr. Sakchai Tangwannawit, KMUTNB |
| Assistant Secretary: | Asst. Prof. Dr. Watchareewan Jitsakul, KMUTNB |

# STEERING COMMITTEE

Dr. Dechanuchit Katanyutaveetip, CITT

Prof. Krischonme Bhumkittipich, RMUTT

Prof. Dr.Siriporn Jungsuttiwong, UBU

Assoc. Prof. Dr. Jantima Polpinij, MSU

Asst. Prof. Dr. Photjanee Sukchovna, KRU

Dr. Siwiga Dusadenoad, NIDA

# TECHNICAL PROGRAM COMMITTEE

Akkarat Boonyapalanant, KMUTNB
Benjawan Intara, MSU
Chakadkit Thaenchaikun, PSU
Chatchawin Namman, UBU
Chatklaw Jareonpon, MSU
Dech Thammasiri, NPRU
Gamgarn Somprasertsri, MSU
Gritya Tongpasuk, KMUTNB
Jamornkul Laokietkul, DPU
Jaree Thongkam, MSU
Jarut Busarathid, PBRU
Jeerasak Numpradit, KMUTNB
Jiratti Tengsuthiwat, KMUTNB
Jirawat Thaenthong, PSU
Kachana Viriyapant, KMUTNB
Kairung Hengproprohm, NPRU
Kanita Saengkrajang, PCRU
Kriengsak Treeprapin, UBU
Kuljaree Tantayakul, PSU
Laddawan Champa, KRU
Lap Poomhiran, SSRU
Maleerat Maliyaem, KMUTNB
Mathuros Panmuang, RMUTT
Nadh Ditcharoen, UBU
Nalinpat Bhumpenpein, KMUTNB
Narong Phothi, KMUTNB
Nattapong Tongtep, PSU
Nattavee Utakrit, KMUTNB
Nitinan Mata, BRU
Nongluk Promthong, RMUTT
Panana Tangwannawit, PCRU
Pattaraporn Warintarawej, PSU
Phayung Meesad, KMUTNB
Pichate Kunakornvong, RMUTT
Pongsarun Boonyopakorn, KMUTNB
Pudsadee Boonrawd, KMUTNB
Rungtip Charoensak, MSU
Sakchai Tangwannawit, KMUTNB
Salintip Sudsanguan, SEC
Sarawut Plaengsorn, PBRU
Sirapat Boonkrong, SUT
Sompong Valuvanathorn, UBU
Sukuma Uamcharoen, PBRU
Sumitra Nuanmeesri, SSRU
Sunantha Sodsee, KMUTNB
Supathep Satiman, UBRU
Supawadee Hiranpongsin, UBU
Supoj Hengpraprohm, NPRU
Sutat Gammanee, KRU
Sutep Tongngam, NIDA
Tanapon Jensuttiwetchakul, KMUTNB
Thanawat Dechahiranavat, KMUTNB
Thawatchai Sarawong, RMUTK
Thitirat Siriborvornratanakul, NIDA
Thongchai Kaewkiriya, SIU
Tongpool Heeptaisong, KMUTNB
Uraiwan Inyaem, RMUTT
Wanida Saetang, PSU
Watchareewan Jitsakul, KMUTNB
Weena Janratchakool, RMUTT
Wongkot Sriurai, UBU

# TECHNICAL PROGRAM CONTENTS

# Program Schedule

| Thursday 14th May 2026 | |
|---|---|
| 08:00 - 09:00 | Registration |
| 09:00 - 09:15 | Report by **Assistant Professor Dr. Dr.-Ing. Sunantha Sodsee**,<br>*Dean, Faculty of Information Technology and Digital Innovation*<br>*King Mongkut's University of Technology North Bangkok* |
| 09:15 - 09.30 | Opening Ceremony by **Professor Dr. Thanin Silpcharu**,<br>*President of King Mongkut's University of Technology North Bangkok* |
| 09:30 - 09:45 | *Group Photo* |
| 09:45 - 10:45 | Invited Keynote Speech by **Prof. Dr. Gary G. Yen,**<br>• Chair Professor, Sichuan University, CHINA<br>Regents Professor, Oklahoma State University, USA |
| 10:45 - 11:00 | *Coffee break* |
| 11:00 - 12:00 | Invited Keynote Speech by **Dr. Soontorn Sirapisan,**<br>• Acting Head of the Research and Research Collaboration Section |
| 12:00 - 13:00 | *Lunch* |
| 13:00 - 15:00 | **Paper Presentation** |
| 15:00 - 15:20 | *Coffee break* |
| 15:20 - 17:00 | **Paper Presentation** |

| Thursday 14th May 2026 | | |
|---|---|---|
| **ROOM: I** | | **Page** |
| 13:00 - 13:20<br>NCCIT2026-088 | **Developing a Maturity Model for GRC in Hybrid Cloud Environments: A Case Study of Global Cookware in the APAC Region**<br>*Natthanun Wicheanrat and Sunantha Sodsee* | 1 |
| 13:20 - 13:40<br>NCCIT2026-023 | **An Assessment Model of Cyber Resilience Readiness and Recovery Capability**<br>*Kiattisak Luithong and Nalinpat Bhumpenpein* | 7 |
| 13:40 - 14:00<br>NCCIT2026-056 | **Content-Aware and Adaptive Lightweight Protection for UAV Video Streaming**<br>*Thitiworada Thabtimsai, Silaporn Chuntha and Chakchai So-In* | 13 |
| 14:00 - 14:20<br>NCCIT2026-091 | **Application of Microsoft Power Apps Platform for Managing and Transferring Technical Consultants' Knowledge to Sales Teams: A Case Study of a Telecommunication Company in the Lower Northern Region**<br>*Anthika Wattanakulwiwat and Sunantha Sodsee* | 30 |
| 14:20 - 14:40<br>NCCIT2026-090 | **Performance Comparison of Web Application Vulnerability Scanning Tools and Patch Validation Processes**<br>*Natthida Looksua and Sunantha Sodsee* | 36 |
| 14:40 - 15:00 | *Break* | |
| 15:00 - 15:20<br>NCCIT2026-053 | **Home-based Elderly Health Monitoring System Using IoT with Caregiver Alert Application**<br>*Panatsaya Thongdee and Tanapon Jensuttiwetchakul* | 42 |
| 15:20 - 15:40<br>NCCIT2026-092 | **Automated System for Analyzing Players' Problem-Solving Skills and Providing Feedback for CTF Competitions**<br>*Thanagrit Wutthiamornthada and Nawaporn Wisitpongphan* | 49 |
| 15:40 - 16:00<br>NCCIT2026-072 | **Factors Determining the Intention to Use AI-Microlearning Platforms in Thai Technology Enterprises**<br>*Mayurakran Chinnawong and Nattavee Utakrit* | 55 |
| 16:00 - 16:20<br>NCCIT2026-015 | **A Comparative Analysis of the Effectiveness of Single-Keyword and Multi-Keyword SEO Strategies for Walailak University Website**<br>*Wachirasak Pokakorn* | 61 |
| | *End Session* | |

| Thursday 14th May 2026 | | |
|---|---|---|
| **ROOM: II** | | **Page** |
| 13:00 - 13:20<br>NCCIT2026-029 | **Factors Influencing the Success of Requirements Elicitation for Software Development**<br>*Sutthikarn Julsuwan and Nalinpat Bhumpenpein* | 68 |
| 13:20 - 13:40<br>NCCIT2026-063 | **Automated Product Recommendation System on LINE Official Account Using Association Rule Mining: A Case of Sky Bakery and Coffee**<br>*Poramets Phetman and Tanapon Jensuttiwetchakul* | 75 |
| 13:40 - 14:00<br>NCCIT2026-068 | **Consumer Segmentation Using Data Mining: A Case Study of Information Systems for Green Products**<br>*Phattranit Vasuangsunit and Montean Rattanasiriwongwut* | 81 |
| 14:00 - 14:20<br>NCCIT2026-073 | **Classification of Social Media User Opinions on Artificial Intelligence (AI) Using Text Mining Techniques**<br>*Thanatmeth Wuttisathittham and Sakchai Tangwannawit* | 88 |
| 14:20 - 14:40<br>NCCIT2026-037 | **An Analysis of Factors Affecting the Perception of Cybersecurity Standards in Hospitals**<br>*Pongsri Prasomboon and Sakchai Tangwanawit* | 94 |
| 14:40 - 15:00 | *Break* | |
| 15:00 - 15:20<br>NCCIT2026-038 | **Forecasting Resource SAP System Utilization Using Data Mining**<br>*Chalermkiet Tongwilai and Montean Rattanasiriwongwut* | 100 |
| 15:20 - 15:40<br>NCCIT2026-013 | **Group-based Causality from IT-User Comments**<br>*Ratchada Pleanthong, Chaveevan Pechsiri and Adisorn Na Ubon* | 106 |
| 15:40 - 16:00<br>NCCIT2026-035 | **Guidelines for Promoting Artificial Intelligence Adoption in Vocational Education Teaching Practices**<br>*Ratchanat Huansin and Nalinpat Bhumpenpein* | 112 |
| 16:00 - 16:20<br>NCCIT2026-061 | **Text Mining-Based Detection of Prohibited Goods from Social Media to Support Customs Authorities**<br>*Natcharee Ariyanithitawesuk and Tanapon Jensuttiwetchakul* | 118 |
| 16:20 - 16:40<br>NCCIT2026-028 | **Guidelines for Developing Artificial Intelligence Skills among Vocational Education Students**<br>*Rewadee Aminram and Nalinpat Bhumpenpein* | 125 |
| 16:40 - 17:00<br>NCCIT2026-036 | **A Data Literacy Assessment Model**<br>*Komjak Kamsiang and Nalinpat Bhumpenpein* | 131 |
| | *End Session* | |

| Thursday 14th May 2026 | | |
|---|---|---|
| **ROOM: III** | | **Page** |
| 13:00 - 13:20<br>NCCIT2026-057 | **Three-Class GPS Spoofing Detection for UAVs Using an Attention-based CNN-LSTM**<br>*Sirikarn Netpookgana, Naravit Conjaiboon and Chakchai So-In* | 137 |
| 13:20 - 13:40<br>NCCIT2026-086 | **Applying NLP for Clustering and Visualizing Cyber Threat Intelligence from X (Twitter)**<br>*Nippich Yimporntana and Sunantha Sodsee* | 151 |
| 13:40 - 14:00<br>NCCIT2026-058 | **Risk-Triggered Hybrid Detection for DCL-based Android Malware**<br>*Chatphol Nawjampa, Korawit Sutthi-at and Chakchai So-In* | 158 |
| 14:00 - 14:20<br>NCCIT2026-089 | **Applying NIST AI RMF and NCSA AI Security Guidelines to Manage Generative AI Risks in Organizations**<br>*Phattapol Yamwankaew and Nawaporn Wisitpongphan* | 170 |
| 14:20 - 14:40<br>NCCIT2026-059 | **A Two-Phase Reactive Cyber Range Framework with Moving Target Defense and Automated Deception on Lightweight Containers**<br>*Anawin Uthaisang, Nansaphat Termsub and Chakchai So-In* | 175 |
| 14:40 - 15:00 | *Break* | |
| 15:00 - 15:20<br>NCCIT2026-060 | **Class-Incremental RAP Detection Using Deep Learning and Exemplar Replay on CSI Signals**<br>*Phutthicha Rattana, Rachata Nanta and Chakchai So-In* | 185 |
| 15:20 - 15:40<br>NCCIT2026-019 | **Benchmarking Modern CNN Architectures for Thai License Plate Character and Province Recognition Using YOLOv8n-Based Detection**<br>*Nattapon Boondang and Amnach Khawne* | 199 |
| 15:40 - 16:00<br>NCCIT2026-008 | **A Data-Driven Decision-Making Framework for Optimizing Solar Panel Cleaning in Island-Based PV Systems**<br>*Krairawee Nosri and Supakpong Jinarat* | 205 |
| 16:00 - 16:20<br>NCCIT2026-055 | **A Gray-Box Robustness Evaluation Framework for Hybrid IoT Intrusion Detection Systems Using a Global Search Genetic Algorithm**<br>*Kanoknipha Panset, Napatwan Chaiban and Chakchai So-In* | 213 |
| 16:20 - 16:40<br>NCCIT2026-031 | **A Framework for Converting Land Title Deed Images into Vector Data Using YOLOv11, Thai OCR, and Area-Preserving Adjustment**<br>*Worawat Soiphet and Supakpong Jinarat* | 223 |
| 16:40 - 17:00<br>NCCIT2026-034 | **A Weighted Cybersecurity Competency Assessment Framework**<br>*Chaiyut Mattrai and Nalinpat Bhumpenpein* | 230 |
| | *End Session* | |

| Thursday 14th May 2026 | | |
|---|---|---|
| **ROOM: IV** | | **Page** |
| 13:00 - 13:20<br>NCCIT2026-010 | **Author-based Diversification in Book Recommender Systems Using Intra-List Similarity Reduction Techniques**<br>*Laksamee Noneluecha and Supakpong Jinarat* | 236 |
| 13:20 - 13:40<br>NCCIT2026-017 | **The Decision Assistant System for Student Admission in Computer Science Curricurum, Rajamangala University of Technology Thanyaburi**<br>*Weena Janratchakool, Uraiwan Inyaem, Chayanon Premvijitr and Aimphagar Luangphusab* | 242 |
| 13:40 - 14:00<br>NCCIT2026-087 | **Development of a Real-time Anomaly Detection and Visualization System for Hospital Information Systems**<br>*Benchapol Thongchua and Nawaporn Wisitpongphan* | 248 |
| 14:00 - 14:20<br>NCCIT2026-009 | **Mining Popular Food Association Rules from Positive Customer Reviews**<br>*Ekasis Girachaipitack, Chaveevan Pechsiri and Adisorn Na Ubon* | 254 |
| 14:20 - 14:40<br>NCCIT2026-011 | **Aspect-Based Opinion Mining of User Reviews in Online Businesses**<br>*Bunyawat Pattanasakkomjon, Chaveevan Pechsiri and Adisorn Na Ubon* | 260 |
| 14:40 - 15:00 | *Break* | |
| 15:00 - 15:20<br>NCCIT2026-025 | **A Comparative Study of Quantum-inspired and Classical Machine Learning for Customer Return Prediction in Nighttime Restaurants**<br>*Tanyavat Kaewpradit and Waraporn Viyanon* | 265 |
| 15:20 - 15:40<br>NCCIT2026-014 | **Performance Comparison of Machine Learning Models for Inventory Replenishment and Supply Chain Decision-Making**<br>*Nataporn Koopipat and Supakpong Jinarat* | 271 |
| 15:40 - 16:00<br>NCCIT2026-027 | **A Machine Learning-Based Approach for the Development of Job Recommendation System**<br>*Uraiwan Inyaem, Weena Janratchakool, Benjawan Surapol and Tewarach Rakngam* | 277 |
| 16:00 - 16:20<br>NCCIT2026-022 | **E-Mail Objective Classification using Machine Learning**<br>*Theerachai Kasemphongjarat and Maleerat Maliyaem* | 283 |
| 16:20 - 16:40<br>NCCIT2026-030 | **Performance Comparison of Traditional and EMA-Based Dynamic DCA in Cryptocurrencies**<br>*Pipat Thongjapo and Supakpong Jinarat* | 289 |
| | *End Session* | |

| Thursday 14th May 2026 | | |
|---|---|---|
| **ROOM: V** | | **Page** |
| 13:00 - 13:20<br>NCCIT2026-032 | **Development of a Cultural Data Analytics Dashboard for Policy Support**<br>*Chanakit Mitsongkore, Warong Naivinit, Sutat Gammanee and Sureewan Jangjit* | 296 |
| 13:20 - 13:40<br>NCCIT2026-026 | **Development of an Intelligent Knowledge Retrieval System for Government Organizations Using Hybrid Semantic Search Techniques**<br>*Kamonlak Pinyakong and Supakpong Jinarat* | 302 |
| 13:40 - 14:00<br>NCCIT2026-083 | **Developing a Promotion Recommendation System for B2B Wholesale Retail through the Integration of Data Mining, Deep Learning, and Contextual Bandit Algorithms**<br>*Teeradaj Kanching and Phayung Meesad* | 308 |
| 14:00 - 14:20<br>NCCIT2026-093 | **Development of a Chatbot to Promote Cybersecurity Awareness**<br>*Wilaiwan Phanthura and Nawaporn Wisitpongphan* | 315 |
| 14:20 - 14:40<br>NCCIT2026-052 | **Development of a Faith-Based Tourism Recommendation System in Nakhon Pathom Province**<br>*Witthawat Chankimha, Adisak Yongpanya, Wannachai Chouegthong, Supakit Nakpomchin and Suphitcha Chanrueang* | 321 |
| 14:40 - 15:00 | *Break* | |
| 15:00 - 15:20<br>NCCIT2026-012 | **Inventory Management System Using RFID and Cloud Technology**<br>*Apichart Kanseenuan and Suphakit Awiphan* | 328 |
| 15:20 - 15:40<br>NCCIT2026-043 | **Factors Influencing Consumer Acceptance and Decision to Use Chatbot Systems**<br>*Anan Malichai and Montean Rattanasiriwongwut* | 334 |
| 15:40 - 16:00<br>NCCIT2026-039 | **Smart Retail Space Management System**<br>*Burasakorn Yoosooka, Suvil Chomchaiya, Uraiwan Inyaem, Phumsit Nanram, Bavornthat Dangtang and Nitipat Sritharachiyanont* | 341 |
| 16:00 - 16:20<br>NCCIT2026-085 | **Human Resource Analysis and Management System: A Case Study of Saint Louis College**<br>*Tadchanon Polngam and Pudsadee Boonrawd* | 349 |
| | *End Session* | |

# Program Schedule

| Friday 15th May 2026 | |
|---|---|
| 09:00 - 10:20 | **Paper Presentation** |
| 10:20 - 10:40 | *Break* |
| 10:40 - 12:00 | **Paper Presentation** |

| Friday 15th May 2026 | | |
|---|---|---|
| **ROOM: I** | | **Page** |
| 09:00 - 09:20<br>NCCIT2026-047 | **An Analysis of Factors Influencing Consumers' Impulse Buying Behavior Through Live Streaming**<br>*Phakamon Malee and Tanapon Jensuttiwetchakul* | 355 |
| 09:20 - 09:40<br>NCCIT2026-044 | **Factors Affecting Behavioral Intention to use Self-Service System via QR Code**<br>*Nopsapoj Srisuwan and Sakchai Tangwanawit* | 362 |
| 09:40 - 10:00<br>NCCIT2026-069 | **Factors Influencing Consumers' Behavioral Intention to Use Self-Drive Car Rental Services via Social Media in Thailand**<br>*Pacharin Sirirak and Sakchai Tangwannawit* | 368 |
| 10:00 - 10:20<br>NCCIT2026-070 | **Factors Influencing the Acceptance of Virtual Reality Technology in Architectural**<br>*Nattakan Tamcharoen and Montean Rattanasiriwongwut* | 375 |
| 10:20 - 10:40 | *Break* | |
| 10:40 - 11:00<br>NCCIT2026-050 | **Factors Influencing the Acceptance of Self-Checkout Systems in Modern Retail Stores: A Study of Supercenters and Supermarkets**<br>*Anothai Arunthong and Tanapon Jensuttiwetchakul* | 382 |
| 11:00 - 11:20<br>NCCIT2026-066 | **An Analysis of Factors Influencing the Use of AI in Chinese Language Learning**<br>*Prawit Ekweerapatarasakun and Montean Rattanasiriwongwut* | 388 |
| 11:20 - 11:40<br>NCCIT2026-065 | **Analyzing the Acceptance of Telemedicine Using Data Mining with Business Intelligence System**<br>*Jittakron Saetung and Montean Rattanasiriwongwut* | 394 |
| 11:40 - 12:00<br>NCCIT2026-067 | **Factors Influencing the Acceptance of Food-Serving Robots in Restaurants**<br>*Rajabhat Raksujarit and Sakchai Tangwanawit* | 401 |
| | *End Session* | |

| Friday 15th May 2026 | | |
|---|---|---|
| **ROOM: II** | | **Page** |
| 09:00 - 09:20<br>NCCIT2026-051 | **Development of a Decision Support System for Rehabilitation Planning of Drug Offenders**<br>*Supakorn Srithongjun and Nattavee Utakrit* | 408 |
| 09:20 - 09:40<br>NCCIT2026-076 | **Development of Agentic AI Models for Organizational Carbon Footprint Assessment and Governance**<br>*Sojirat Songpradit and Mahasak Ketcham* | 414 |
| 09:40 - 10:00<br>NCCIT2026-077 | **Classification of Software Development Task Complexity Levels from Requirement Documents Using Natural Language Processing and Generative AI**<br>*Wittayadet Jaikwang and Phayung Meesad* | 420 |
| 10:00 - 10:20<br>NCCIT2026-078 | **Development of an Activity Attendance Management System for the Digital Business Technology Club, Chainat Technical College**<br>*Wilasinee Chutiwat and Nattavee Utakrit* | 426 |
| 10:20 - 10:40 | *Break* | |
| 10:40 - 11:00<br>NCCIT2026-045 | **Classification of User Opinions on Healthy Meal Delivery Services on Online Platforms Using Deep Learning Techniques**<br>*Suda Phuchamnong and Watchareewan Jitsakul* | 432 |
| 11:00 - 11:20<br>NCCIT2026-064 | **Forecasting Welfare Demand Using Data Mining Techniques**<br>*Manu Boonpleum and Watchareewan Jitsakul* | 438 |
| 11:20 - 11:40<br>NCCIT2026-081 | **An Analysis of the factors that influence the use of the Cyber Security Management System of the Royal Thai Army Headquarters**<br>*Worawoot Samanpong and Montean Rattanasiriwongwut* | 444 |
| | *End Session* | |

| Friday 15th May 2026 | | |
|---|---|---|
| **ROOM: III** | | **Page** |
| 09:00 - 09:20<br>NCCIT2026-024 | **A Comparative Study of Machine Learning Models for Predicting the Cost-Effectiveness of Carbon Credits from Tree Plantation Investment in Community Enterprises, Khon Kaen Province**<br>*Papatsara Siharach and Nalinpat Bhumpenpein* | 449 |
| 09:20 - 09:40<br>NCCIT2026-041 | **Consumer Intent Analysis in the Beverage Industry on Social Media Platforms Using Deep Learning Techniques**<br>*Keerati Jantorn and Watchareewan Jitsakul* | 455 |
| 09:40 - 10:00<br>NCCIT2026-079 | **Sentiment Analysis of Banking Applications Users on Social Media Using Large Language Models**<br>*Sitawee Sirithitiwithaya and Pudsadee Boonrawd* | 461 |
| 10:00 - 10:20<br>NCCIT2026-080 | **Development of a Machine Learning-Based Log File Classification System for Problem Analysis and Escalation Support in a Game Publisher Organization**<br>*Pongpeera Chaipongpatsorn and Nattavee Utakrit* | 467 |
| 10:20 - 10:40 | *Break* | |
| 10:40 - 11:00<br>NCCIT2026-082 | **Development of a Forecasting and Decision-Making Framework for Inventory Replenishment Using Sales and Inventory Data with Data Mining Techniques and Reinforcement Learning**<br>*Teeratuch Kongyuenyongwanit and Phayung Meesad* | 473 |
| 11:00 - 11:20<br>NCCIT2026-020 | **A Comparison of Machine Learning Models for Predicting False Breakouts in Stock Markets Using Technical Indicators**<br>*Apirat Jirangkapanich and Nalinpat Bhumpenpein* | 480 |
| 11:20 - 11:40<br>NCCIT2026-062 | **Forecasting Thailand's Durian Exports Using Deep Learning Techniques**<br>*Chonnika Phuthong and Watchareewan Jitsakul* | 486 |
| 11:40 - 12:00<br>NCCIT2026-071 | **Forecasting Model for the Number of Patients Cancer in the Public Hospital Base by Time Series Technical**<br>*Run Kanchano and Nattavee Utakrit* | 493 |
| | *End Session* | |

| Friday 15th May 2026 | | |
|---|---|---|
| **ROOM: IV** | | **Page** |
| 09:00 - 09:20<br>NCCIT2026-042 | **Intelligent Fish Behavior Detection System Using Image Processing**<br>*Monthicha JImthaisong and Mahasak Ketcham* | 500 |
| 09:20 - 09:40<br>NCCIT2026-021 | **A Machine Learning Model for Predicting Royal Pharmacopoeia Drug Orders in GPO Pharmacy**<br>*Sabsin Dilokwechpanich and Nalinpat Bhumpenpein* | 507 |
| 09:40 - 10:00<br>NCCIT2026-084 | **Machine Learning-based Recommendation for Legal Opinion and Order Drafting Formats**<br>*Kritsanaphong Kimsue and Nattavee Utakrit* | 513 |
| 10:00 - 10:20<br>NCCIT2026-075 | **Development of an Assessment Tool for Personal Data Security Management in Organizations According to ISO/IEC 27701 Standard and Thailand's Personal Data Protection Act (PDPA)**<br>*Nithi Sakunphanish and Nattavee Utakrit* | 519 |
| 10:20 - 10:40 | *Break* | |
| 10:40 - 11:00<br>NCCIT2026-048 | **Predicting Academic Achievement of Vocational Certificate Using The Big Five Personality Factors and Core Subject Grades in Lower Secondary School**<br>*Uraiwan Meesang and Tanapon Jensuttiwetchakul* | 525 |
| 11:00 - 11:20<br>NCCIT2026-074 | **In-Depth Comparative Analysis of Household Savings Inequality in Thailand by Region and Income Group: A Power BI-Based Study Using NSO Data**<br>*Apinya Thongkum and Nattavee Utakrit* | 532 |
| 11:20 - 11:40<br>NCCIT2026-054 | **Forecasting Sales of Fabric Care Products Through the E-Commerce Channel**<br>*Nattaporn Kongsakunkrai and Sakchai Tangwannawit* | 545 |
| 11:40 - 12:00<br>NCCIT2026-049 | **Factors Influencing Continued Use of Generative AI**<br>*Metta Naktam and Tanapon Jensuttiwetchakul* | 551 |
| | *End Session* | |

# REGULAR PAPERS

# การพัฒนาโมเดลการวัดระดับวุฒิภาวะของการกำกับดูแล ความเสี่ยง และการปฏิบัติตามข้อกำหนดในสภาพแวดล้อมไฮบริดคลาวด์ กรณีศึกษาของบริษัทเครื่องครัวระดับโลกประจำภูมิภาคเอเชียแปซิฟิก

# Developing a Maturity Model for GRC in Hybrid Cloud Environments: A Case Study of Global Cookware in the APAC Region

*ณัฐนันท์ วิเชียรรัตน์ (Natthanun Wicheanrat)[1] และสุนันฑา สดสี (Sunantha Sodsee)[2]*
*ภาควิชาการบริหารเครือข่ายดิจิทัลและความมั่นคงปลอดภัยสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล*
*มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ*
*s6707031857131@email.kmutnb.ac.th, sunantha.s@itd.kmutnb.ac.th*

**บทคัดย่อ**

*งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาโมเดลการวัดระดับวุฒิภาวะของการกำกับดูแล ความเสี่ยง และการปฏิบัติตามข้อกำหนด (GRC Maturity Model) สำหรับสภาพแวดล้อมไฮบริดคลาวด์ โดยบูรณาการกรอบมาตรฐาน 3 ฉบับ ได้แก่ NIST Cybersecurity Framework 2.0 (2024), ISF Standard of Good Practice for Information Security 2024 และ Capability Maturity Model (CMM) ออกแบบเป็น 9 ฟังก์ชัน 200 คำถาม 5 ระดับวุฒิภาวะ ผ่านการตรวจสอบความตรงเชิงเนื้อหาโดยผู้เชี่ยวชาญ และนำไปทดสอบกับกรณีศึกษาบริษัทเครื่องครัวระดับโลกประจำภูมิภาคเอเชียแปซิฟิก (n = 10) ผลการวิจัยพบว่าค่าคะแนนวุฒิภาวะโดยรวมเท่ากับ 2.64 จาก 5.00 อยู่ในระดับ 3: Defined ฟังก์ชัน Security Controls & Protection ได้คะแนนสูงสุด (3.05) ขณะที่ Cloud Governance & Compliance ได้คะแนนต่ำสุด (1.95) มีช่องว่างวิกฤต 2.23 คะแนน งานวิจัยนำเสนอแผนพัฒนา 3 ระยะ พร้อมข้อเสนอแนะเชิงนโยบาย 6 ข้อ จากการประเมินโดยผู้บริหาร แสดงให้เห็นว่าผลการวิจัย และข้อเสนอแนะดังกล่าวสามารถนำไปใช้ในองค์กรได้*

**คำสำคัญ:** การประเมินวุฒิภาวะ ไฮบริดคลาวด์ การกำกับดูแล ความมั่นคงปลอดภัยสารสนเทศ

**Abstract**

*This research develops a GRC (Governance, Risk, and Compliance) Maturity Model for hybrid cloud environments by integrating NIST Cybersecurity Framework 2.0 (2024), ISF Standard of Good Practice for Information Security 2024, and the Capability Maturity Model (CMM). The model comprises 9 functions, 200 assessment questions, and 5 maturity levels. The model was applied to a global cookware company's APAC operation (n = 10). Results indicate an overall maturity score of 2.64/5.00 (Level 3: Defined). Security Controls & Protection scored highest (3.05), while Cloud Governance & Compliance scored lowest (1.95) with a critical gap of 2.23. A three-phase improvement roadmap and six policy recommendations are proposed. The managerial evaluation indicates that the research findings and recommendations are applicable within the organization.*

**Keywords:** Maturity Model, Governance Risk and Compliance, Hybrid Cloud, Information Security

## 1. บทนำ

ปัจจุบันองค์กรทั่วโลกนำสถาปัตยกรรมไฮบริดคลาวด์มาใช้อย่างแพร่หลาย จากรายงาน Flexera 2025 State of

the Cloud Report พบว่าร้อยละ 70 ขององค์กรนำกลยุทธ์ไฮบริดคลาวด์มาใช้ [1] และ Gartner คาดการณ์ว่าการลงทุนด้านความมั่นคงปลอดภัยคลาวด์จะเติบโตร้อยละ 25.9 ต่อปี [2] ความซับซ้อนของสภาพแวดล้อมนี้ทำให้การกำกับดูแล ความเสี่ยง และการปฏิบัติตามข้อกำหนด (GRC) เป็นความท้าทายสำคัญ IBM รายงานว่าการรั่วไหลของข้อมูลในสภาพแวดล้อมไฮบริดคลาวด์มีมูลค่าความเสียหายเฉลี่ย 4.75 ล้านดอลลาร์สหรัฐต่อครั้ง [3]

ปัญหาสำคัญคือยังไม่มีโมเดลการวัดระดับวุฒิภาวะ GRC ที่ออกแบบเฉพาะสำหรับไฮบริดคลาวด์ที่บูรณาการทั้ง 3 มิติอย่างครบถ้วน งานวิจัยที่ผ่านมาส่วนใหญ่มุ่งเน้นเฉพาะบางด้าน เช่น ความมั่นคงปลอดภัยมัลติคลาวด์ [4] หรือ GRC เชิงบูรณาการทั่วไป [5] แต่ยังขาดกรอบแนวคิดครอบคลุมไฮบริดคลาวด์โดยเฉพาะ

งานวิจัยนี้มีวัตถุประสงค์ 3 ข้อ ได้แก่ (1) พัฒนา GRC Maturity Model สำหรับไฮบริดคลาวด์โดยบูรณาการ NIST CSF 2.0, ISF SoGP 2024 และ CMM (2) นำโมเดลไปทดสอบกับกรณีศึกษาบริษัทเครื่องครัวระดับโลกประจำภูมิภาค APAC และ (3) นำเสนอแผนพัฒนาและข้อเสนอแนะเชิงนโยบาย

## 2. ทฤษฎี และงานวิจัยที่เกี่ยวข้อง

### 2.1 สภาพแวดล้อมไฮบริดคลาวด์และความท้าทาย GRC

Ali และคณะ [4] พบว่ามาตรการรักษาความมั่นคงปลอดภัยที่ใช้ได้ดีในสภาพแวดล้อมหนึ่งอาจล้มเหลวในอีกสภาพแวดล้อม ทำให้การวัดและกำกับดูแลรวมศูนย์เป็นสิ่งจำเป็น ISACA [6] ชี้ว่า Data Sovereignty ในภูมิภาค APAC ซึ่งไม่มีกรอบกฎหมายรวมศูนย์เทียบเท่า GDPR สร้างความซับซ้อนเพิ่มเติม Cloud Security Alliance ระบุว่าความซับซ้อนของ Shared Responsibility Model ข้ามผู้ให้บริการหลายรายเป็นอุปสรรคหลักในการกำกับดูแล [7]

### 2.2 โมเดลวุฒิภาวะด้านความมั่นคงปลอดภัยและ GRC

Schmitz และคณะ [8] ศึกษาความสามารถในการประเมินวุฒิภาวะ Security Controls พบว่าผู้ปฏิบัติงานมักประเมินสูงกว่าความเป็นจริง Al-Matari และคณะ [9] เสนอโมเดลวุฒิภาวะด้านความมั่นคงปลอดภัย 5 ระดับครอบคลุมเทคโนโลยี บุคลากร และโครงสร้างพื้นฐาน AlGhamdi และคณะ [10] ระบุ 7 โดเมน 27 ปัจจัยความสำเร็จของ IS Governance แต่ยังไม่ครอบคลุมมิติเฉพาะไฮบริดคลาวด์ Mahendra และคณะ [5] ระบุช่องว่างในการบูรณาการ GRC ด้วยเทคโนโลยี AI, Blockchain และ Cloud Computing สำหรับโมเดลวุฒิภาวะเฉพาะคลาวด์ Cheng และคณะ [11] เสนอ Cloud Data Governance Maturity Model แต่มุ่งเฉพาะการกำกับดูแลข้อมูล ยังไม่บูรณาการ GRC ทุกมิติ

### 2.3 กรอบแนวคิดที่ใช้ในงานวิจัย

งานวิจัยบูรณาการ 3 กรอบ ได้แก่ (1) NIST CSF 2.0 ซึ่งเพิ่มฟังก์ชัน Govern เป็นแกนกลาง ประกอบด้วย 6 ฟังก์ชัน 22 หมวดหมู่ 106 หมวดหมู่ย่อย (2) ISF SoGP 2024 ที่ให้มาตรการควบคุมเชิงปฏิบัติ 17 หมวดหมู่ รวมถึงมาตรการเฉพาะคลาวด์ผ่านการเชื่อมโยงกับ CSA CCM v4 ซึ่งองค์กรกรณีศึกษาเป็นสมาชิก ISF UK และ (3) CMM ที่ให้มาตรวัด 5 ระดับตั้งแต่ Initial ถึง Optimizing

## 3. วิธีการดำเนินการวิจัย

### 3.1 รูปแบบการวิจัยและกรอบแนวคิด

การวิจัยนี้ใช้วิธีการวิจัยเชิงผสมผสาน (Mixed Methods) ประกอบด้วย (1) การวิจัยเชิงคุณภาพ: การทบทวนวรรณกรรมและมาตรฐาน การวิเคราะห์บริบทองค์กร การออกแบบโมเดล และการตรวจสอบความถูกต้องโดยผู้เชี่ยวชาญ และ (2) การวิจัยเชิงปริมาณ: การประเมินระดับวุฒิภาวะผ่านแบบสอบถาม 200 ข้อ ด้วยเกณฑ์การให้คะแนน 1–5 ระดับ การคำนวณ Maturity Score ต่อฟังก์ชัน และการเปรียบเทียบกับ Target และ Benchmark

กรอบแนวคิดการวิจัยบูรณาการมาตรฐาน 3 ฉบับ ได้แก่ NIST CSF 2.0 (กำหนดโดเมนการประเมิน), ISF SoGP 2024 (พัฒนาคำถามและเมทริกซ์การให้คะแนน) และ CMM (กำหนด 5 ระดับวุฒิภาวะ) พร้อมบริบทองค์กร ได้แก่ Global Information Security Roadmap และ KPI ที่กำหนด ผ่านกระบวนการพัฒนาเครื่องมือ ตรวจสอบโดยผู้เชี่ยวชาญ เก็บข้อมูล และวิเคราะห์วุฒิภาวะ เพื่อผลลัพธ์ที่

ประกอบด้วย GRC Maturity Model, ผลการประเมิน, Gap Analysis, Roadmap และข้อเสนอแนะเชิงนโยบาย

**3.2 โมเดลการวัดระดับวุฒิภาวะ GRC ที่พัฒนาขึ้น**

Hybrid Cloud GRC Maturity Assessment 2.0 ที่พัฒนาขึ้นครอบคลุม 9 ฟังก์ชันหลัก ได้แก่ GS (Governance & Strategy), AR (Asset & Risk Management), SC (Security Controls & Protection), TD (Threat Detection & Monitoring), IR (Incident Response & Recovery), RC (Resilience & Continuity), ST (Supply Chain & Third Party), AC (Awareness & Culture) และ CC (Cloud Governance & Compliance) รวม 200 คำถาม โดยกำหนดระดับวุฒิภาวะ 5 ระดับตามช่วงคะแนน: 1.00–1.49 Initial, 1.50–2.49 Developing, 2.50–3.49 Defined, 3.50–4.49 Managed และ 4.50–5.00 Optimizing

**3.3 การเชื่อมโยงมาตรฐาน 3 ฉบับกับ Hybrid Cloud GRC Maturity Assessment 2.0**

Hybrid Cloud GRC Maturity Assessment 2.0 ได้รับการพัฒนาโดยบูรณาการมาตรฐาน 3 ฉบับอย่างมีโครงสร้าง โดยแต่ละมาตรฐานมีบทบาทที่แตกต่างกันและเสริมซึ่งกันและกัน ดังนี้ (1) NIST CSF 2.0 ทำหน้าที่เป็นกรอบหลักในการกำหนดโดเมนและฟังก์ชันการประเมิน โดย 6 Core Functions ของ NIST CSF 2.0 ได้แก่ Govern, Identify, Protect, Detect, Respond และ Recover ถูก Map เข้ากับ 9 ฟังก์ชันของโมเดล ยกตัวอย่างเช่น ฟังก์ชัน Govern ของ NIST CSF 2.0 สอดคล้องกับ GS (Governance & Strategy) และ CC (Cloud Governance & Compliance) ของโมเดล ฟังก์ชัน Identify สอดคล้องกับ AR (Asset & Risk Management) และ ST (Supply Chain & Third Party) ฟังก์ชัน Protect สอดคล้องกับ SC (Security Controls & Protection) และ AC (Awareness & Culture) ฟังก์ชัน Detect สอดคล้องกับ TD (Threat Detection & Monitoring) และฟังก์ชัน Respond และ Recover สอดคล้องกับ IR (Incident Response & Recovery) และ RC (Resilience & Continuity) ตามลำดับ (2) ISF SoGP 2024 ทำหน้าที่เป็นแหล่งอ้างอิงหลักในการพัฒนาคำถามการประเมิน 200 ข้อ โดยแต่ละคำถามอ้างอิงมาตรการควบคุม (Control) เฉพาะจาก ISF SoGP 2024 ซึ่งครอบคลุมทั้งมาตรการด้าน Governance, Risk Management, Compliance และมาตรการเฉพาะสำหรับ Cloud Environment รวมถึงยังใช้เป็นเกณฑ์อ้างอิงในการสร้างเมทริกซ์การให้คะแนน (Scoring Matrix) ซึ่งกำหนดหลักฐานที่คาดหวัง (Expected Evidence) ในแต่ละระดับวุฒิภาวะ (3) CMM ทำหน้าที่กำหนดนิยาม คำอธิบาย และเกณฑ์การวัดผลของระดับวุฒิภาวะทั้ง 5 ระดับ ตั้งแต่ Initial (ยังไม่มีกระบวนการ) จนถึง Optimizing (ปรับปรุงต่อเนื่อง) โดยแต่ละระดับได้รับการกำหนดช่วงคะแนนที่ชัดเจนเพื่อให้สามารถวัดผลได้อย่างเป็นรูปธรรม การบูรณาการมาตรฐานทั้งสามฉบับนี้ทำให้ Hybrid Cloud GRC Maturity Assessment 2.0 มีความครอบคลุมทั้งในมิติของโดเมนการประเมิน ความลึกของมาตรการควบคุม และกรอบการวัดระดับวุฒิภาวะ

**3.4 กลุ่มตัวอย่างและการเก็บข้อมูล**

กลุ่มตัวอย่างแบ่งออกเป็น 3 กลุ่ม ได้แก่ (1) กลุ่มพัฒนาและประเมินแบบสอบถาม: ผู้เชี่ยวชาญ APAC IT Infrastructure 3 คน และ APAC IT Director 2 คน (2) กลุ่มตรวจสอบความถูกต้อง: ผู้เชี่ยวชาญด้าน Global Information Security 3 คน และ (3) กลุ่มประเมินแบบสอบถาม: บุคลากรด้าน IT และความมั่นคงปลอดภัย 10 คน ประกอบด้วยระดับผู้บริหาร 3 คน ระดับผู้จัดการ 2 คน และพนักงานระดับปฏิบัติการ 5 คน ขั้นตอนการวิจัยแบ่งออกเป็น 5 ระยะ ตั้งแต่การวิเคราะห์บริบทองค์กร การพัฒนาโมเดลและเครื่องมือ การตรวจสอบความถูกต้อง การเก็บและวิเคราะห์ข้อมูล จนถึงการจัดทำ Roadmap และรายงาน

**3.5 ผลการตรวจสอบความถูกต้องโดยกลุ่มผู้เชี่ยวชาญ**

ผู้เชี่ยวชาญด้าน Global Information Security จำนวน 3 คน ได้ดำเนินการตรวจสอบความถูกต้องของแบบสอบถาม GRC Maturity Assessment 2.0 จำนวน 200 ข้อ และเมทริกซ์การให้คะแนน (Scoring Matrix) และให้ข้อเสนอแนะเพื่อปรับปรุง โดยสรุปผลการตรวจสอบจากกลุ่มผู้เชี่ยวชาญมีประเด็นสำคัญดังนี้ ประการแรก ผู้เชี่ยวชาญเห็นว่าโมเดลมีความครอบคลุมและเหมาะสมกับบริบทขององค์กรที่ดำเนินงานในสภาพแวดล้อม

ไฮบริดคลาวด์ โดยการบูรณาการมาตรฐาน 3 ฉบับมีความสอดคล้องกันดี ประการที่สอง ผู้เชี่ยวชาญแนะนำให้เพิ่มน้ำหนักของคำถามในฟังก์ชัน Cloud Governance & Compliance (CC) เนื่องจากเป็นด้านที่มีความสำคัญสูงและมีความท้าทายเฉพาะสำหรับสภาพแวดล้อมไฮบริดคลาวด์ในภูมิภาค APAC ประการที่สาม ผู้เชี่ยวชาญเห็นว่าเกณฑ์หลักฐานที่คาดหวัง (Expected Evidence) ในแต่ละระดับวุฒิภาวะมีความชัดเจนและสามารถตรวจสอบได้จริง ผู้วิจัยได้นำข้อเสนอแนะของผู้เชี่ยวชาญมาปรับปรุงแบบสอบถามและเมทริกซ์การให้คะแนนก่อนนำไปใช้จริง

## 4. ผลการวิจัย

### 4.1 ผลการประเมินระดับวุฒิภาวะ GRC ภาพรวม

จากการประเมินระดับวุฒิภาวะ GRC โดยผู้ตอบแบบสอบถาม 10 คน ครอบคลุม 200 คำถามใน 9 ฟังก์ชันหลัก พบว่าองค์กรมีคะแนนวุฒิภาวะโดยรวม 2.64 คะแนน (สเกล 1-5) จัดอยู่ในระดับที่ 3 Defined หมายความว่าองค์กรมีนโยบาย ขั้นตอน และมาตรฐานด้าน GRC ที่ได้รับการจัดทำและบังคับใช้ทั่วองค์กรแล้ว เมื่อเปรียบเทียบกับ RH-ISAC & Accenture Benchmark (2025) ซึ่งกลุ่มค้าปลีกระดับโลกมีค่าเฉลี่ย 3.10 คะแนน พบว่าองค์กรอยู่ต่ำกว่า 0.46 คะแนน อย่างไรก็ตาม อยู่สูงกว่า All-industry Baseline จาก Security Architects Partners (2.75) เล็กน้อย

ตารางที่ 1 แสดงผลการประเมินระดับวุฒิภาวะรายฟังก์ชัน พบว่าฟังก์ชันที่มีคะแนนสูงสุดคือ SC (Security Controls & Protection) ที่ 3.05 คะแนน ขณะที่ CC (Cloud Governance & Compliance) มีคะแนนต่ำสุดที่ 1.95 คะแนน โดยมีถึง 4 ฟังก์ชันที่ยังอยู่ในระดับ Developing ได้แก่ GS, TD, ST และ CC

**ตารางที่ 1:** ผลการประเมินระดับวุฒิภาวะรายฟังก์ชัน

| รหัส | ฟังก์ชัน | คะแนนปัจจุบัน | คะแนนเป้าหมาย | ระดับวุฒิภาวะ |
|---|---|---|---|---|
| GS | Governance & Strategy | 2.40 | 4.23 | *Developing* |
| AR | Asset & Risk Management | 2.80 | 3.80 | *Defined* |
| SC | Security Controls & Protection | 3.05 | 4.15 | *Defined* |
| TD | Threat Detection & Monitoring | 2.49 | 3.65 | *Developing* |
| IR | Incident Response & Recovery | 3.00 | 4.00 | *Defined* |
| RC | Resilience & Continuity | 2.72 | 3.95 | *Defined* |
| ST | Supply Chain & Third Party | 2.38 | 3.35 | *Developing* |
| AC | Awareness & Culture | 2.95 | 4.25 | *Defined* |
| CC | Cloud Governance & Compliance | 1.95 | 4.18 | *Developing* |

### 4.2 ความแตกต่างตามระดับตำแหน่ง

การวิเคราะห์ผลจำแนกตามระดับตำแหน่งพบความแตกต่างที่น่าสนใจ: กลุ่มผู้บริหารและผู้จัดการมีคะแนนเฉลี่ย 2.86 คะแนน (ระดับ Developing) ขณะที่กลุ่มระดับปฏิบัติการมีเพียง 2.42 คะแนน (ระดับ Initial-Developing) ส่งผลให้ช่องว่างระหว่างสองกลุ่มอยู่ที่ 0.44 คะแนน สะท้อนถึงความไม่สอดคล้องระหว่างนโยบายที่กำหนดจากระดับบนลงล่างกับการรับรู้และการปฏิบัติจริงในระดับปฏิบัติการ กล่าวคือ นโยบายด้าน GRC ที่ผู้บริหารรับรู้ว่ามีการบังคับใช้แล้วนั้น ยังไม่ได้รับการถ่ายทอดหรือนำไปปฏิบัติจริงอย่างสม่ำเสมอในระดับปฏิบัติการ ซึ่งสอดคล้องกับงานวิจัยของ Schmitz et al. [10] ที่พบว่าผู้บริหารมักประเมินระดับวุฒิภาวะสูงกว่าความเป็นจริง

เนื่องจากขาดข้อมูลเชิงลึกด้านการปฏิบัติการ ความแตกต่างดังกล่าวชี้ให้เห็นถึงความจำเป็นเร่งด่วนในการพัฒนากลไกการสื่อสารนโยบายจากระดับบนสู่ระดับล่าง (Policy Cascading) และการฝึกอบรมผู้ปฏิบัติการให้มีความเข้าใจและสามารถปฏิบัติตามนโยบายด้านความมั่นคงปลอดภัยได้อย่างมีประสิทธิภาพ

**4.3 การวิเคราะห์ช่องว่างและ Roadmap**

ผลการวิเคราะห์ช่องว่าง (Gap Analysis) พบว่า CC มีช่องว่างสูงสุด 2.23 คะแนน รองลงมาคือ GS (1.83) และ AC (1.30) ที่น่าสังเกตคือ Cloud Governance & Compliance ยังไม่ปรากฏอย่างชัดเจนใน Global Information Security Roadmap ปัจจุบัน แม้จะเป็นฟังก์ชันที่มีคะแนนต่ำสุดและช่องว่างสูงสุด ซึ่งถือเป็น Blind Spot เชิงกลยุทธ์ที่สำคัญ ช่องว่างโดยรวมอยู่ที่ 1.31 คะแนน (จาก 2.64 สู่เป้าหมาย 3.95 คะแนน)

แผนพัฒนา Roadmap 3 ระยะที่จัดทำขึ้นมีความสอดคล้องกับ Global Information Security Roadmap ของบริษัท ดังนี้: (1) Quick Wins (0-3 เดือน) มุ่งเน้นฟังก์ชัน CC และ GS ผ่านการกำหนด Cloud Security Governance Policy จัดทำ Cybersecurity Strategy ระดับ APAC ปรับปรุง Cloud Security Configuration และจัดอบรม Security Awareness (2) Medium Term (3-12 เดือน) พัฒนาระบบ SIEM/SOC สำหรับ Hybrid Cloud จัดทำ Vendor Risk Management Program ทดสอบ Incident Response Plan และ BCP/DR (3) Long Term (1-3 ปี) พัฒนา GRC Governance Framework ระดับองค์กร Implement Zero Trust Architecture สร้าง Security Culture Program และนำ AI/ML มาใช้ใน Security Analytics

**4.4 ข้อเสนอแนะเชิงนโยบาย**

ข้อเสนอแนะเชิงนโยบาย 6 ประการ ประกอบด้วย (1) กำหนด Cloud Governance & Compliance Policy อย่างเป็นทางการครอบคลุม AWS/Azure/GCP และ PDPA/GDPR (2) พัฒนา Cybersecurity Strategy ระดับ APAC บูรณาการ Hybrid Cloud สอดคล้อง NIST CSF 2.0 (3) เร่งพัฒนา Threat Detection โดยลงทุนระบบ SIEM/SOAR ที่รองรับ Hybrid Environment (4) เสริมสร้าง Supply Chain & Third-party Risk Management Program รวมถึง Vendor Risk Assessment และ Security Audit (5) จัดตั้ง Information Security Board ที่มีตัวแทนทุกระดับเพื่อลดช่องว่างการรับรู้ และ (6) กำหนด GRC Maturity Re-assessment Cycle ทุก 12 เดือน พร้อม KPI เป้าหมาย $\geq$ 3.00 ใน 12 เดือน $\geq$ 3.50 ใน 24 เดือน และ $\geq$ 3.95 ใน 36 เดือน

**4.5 ผลการประเมินจากผู้บริหาร**

ผู้วิจัยดำเนินการจัดทำ Evaluation Report แล้วเสร็จและนำเสนอผลการวิจัยต่อผู้บริหาร 5 คน ได้แก่ Chief Information Officer 1 คน Chief Technology Officer 1 คน Global Information Security Director 1 คน และ APAC IT Director 2 คน ผู้บริหารทุกท่านได้รับฟังการนำเสนอผลการประเมินระดับวุฒิภาวะ GRC รายฟังก์ชัน ผลการวิเคราะห์ช่องว่าง (Gap Analysis) แผนพัฒนา Roadmap 3 ระยะ และข้อเสนอแนะเชิงนโยบาย 6 ประการ โดยผู้บริหารทั้งหมดให้การยอมรับผลการวิจัยและเห็นชอบในหลักการของ Hybrid Cloud GRC Maturity Assessment 2.0 ว่ามีความครอบคลุมและสะท้อนสถานะจริงขององค์กรได้อย่างถูกต้อง ผู้บริหารมีความเห็นว่าผลการประเมินที่ได้คะแนนโดยรวม 2.64 (ระดับ Defined) และการที่ฟังก์ชัน Cloud Governance & Compliance (CC) มีคะแนนต่ำสุดเป็นข้อมูลที่มีคุณค่าและตรงกับความท้าทายที่องค์กรกำลังเผชิญอยู่ โดยเฉพาะในด้านนโยบาย Cloud Security Governance ที่ยังขาดความชัดเจน ผู้บริหารทั้งหมดรับทราบและมีมติให้นำผลการวิจัยและข้อเสนอแนะดังกล่าวไปปรับใช้ร่วมกับ Global Information Security Roadmap เพื่อเกิดประโยชน์สูงสุดแก่องค์กร ยกระดับความปลอดภัยของสภาพแวดล้อมไฮบริดคลาวด์ให้ทัดเทียมกับมาตรฐานอุตสาหกรรมระดับโลก

## 5. สรุป

การค้นคว้าอิสระนี้ประสบความสำเร็จในการพัฒนา Hybrid Cloud GRC Maturity Assessment 2.0 ที่มีความครบถ้วนและสามารถนำไปปฏิบัติได้จริง โดยบูรณาการ NIST CSF 2.0, ISF SoGP 2024 และ CMM เป็นโมเดลที่

ครอบคลุม 9 ฟังก์ชัน 200 คำถาม 5 ระดับวุฒิภาวะ และผ่านการตรวจสอบโดยผู้เชี่ยวชาญด้าน Global Information Security 3 คน ผลการประเมินองค์กรกรณีศึกษาได้คะแนน 2.64 (ระดับ Defined) สอดคล้องกับ CMM ที่ระบุว่าองค์กรในช่วงเปลี่ยนผ่านทางดิจิทัลมักมีวุฒิภาวะระดับ 2-3 [11]

ข้อค้นพบสำคัญที่สุดคือ Cloud Governance & Compliance (CC) ที่มีคะแนนต่ำสุด (1.95) และช่องว่างสูงสุด (2.23) ยังไม่ปรากฏอย่างชัดเจนใน Global Information Security Roadmap ปัจจุบัน สอดคล้องกับผลการวิจัยของ Mahendra et al. (2022) [8] ที่พบว่าองค์กรมักมองข้ามความท้าทายด้าน Cloud Governance เนื่องจากเป็นสาขาใหม่ที่ขาดผู้เชี่ยวชาญเฉพาะทาง นอกจากนี้ความแตกต่างในการรับรู้ระหว่างผู้บริหาร (2.86) และระดับปฏิบัติการ (2.42) เป็น 0.44 คะแนน สะท้อนถึงความจำเป็นในการลงทุนด้าน Change Management และ Security Culture ควบคู่กับ Technical Controls

โมเดลและเครื่องมือที่พัฒนาขึ้นเป็นประโยชน์ทั้งต่อองค์กรกรณีศึกษาและองค์กรอื่น ๆ ที่เผชิญความท้าทาย GRC ในสภาพแวดล้อมไฮบริดคลาวด์ ในอนาคตควรศึกษาการนำโมเดลไปประยุกต์ใช้กับองค์กรอื่น ๆ ในภูมิภาค APAC เพื่อทดสอบความสามารถในการขยายผล (Scalability) และพัฒนา AI-assisted GRC Assessment ที่สามารถทำนายความเสี่ยงแบบ Proactive Global Information Security Roadmap ปัจจุบัน สอดคล้องกับผลการวิจัยของ Mahendra et al. (2022) [8] ที่พบว่าองค์กรมักมองข้ามความท้าทายด้าน Cloud Governance เนื่องจากเป็นสาขาใหม่ที่ขาดผู้เชี่ยวชาญเฉพาะทาง นอกจากนี้ความแตกต่างในการรับรู้ระหว่างผู้บริหาร (2.86) และระดับปฏิบัติการ (2.42) เป็น 0.44 คะแนน สะท้อนถึงความจำเป็นในการลงทุนด้าน Change Management และ Security Culture ควบคู่กับ Technical Controls

โมเดลและเครื่องมือที่พัฒนาขึ้นเป็นประโยชน์ทั้งต่อองค์กรกรณีศึกษาและองค์กรอื่น ๆ ที่เผชิญความท้าทาย GRC ในสภาพแวดล้อมไฮบริดคลาวด์ ในอนาคตควรศึกษาการนำโมเดลไปประยุกต์ใช้กับองค์กรอื่น ๆ ในภูมิภาค APAC เพื่อทดสอบความสามารถในการขยายผล (Scalability) และพัฒนา AI-assisted GRC Assessment ที่สามารถทำนายความเสี่ยงแบบ Proactive

## เอกสารอ้างอิง

[1] "Cloud computing trends: Flexera 2024 State of the Cloud Report," Flexera Blog. [Online]. Available: https://www.flexera.com/blog/

[2] "Gartner Predicts 50% of Organizations Will Manage Hybrid Cloud Energy Consumption with Sustainability Monitoring By 2026," Gartner, 2024.

[3] "Hybrid Cloud Security," Cloud Security Alliance. [Online]. Available: https://cloudsecurityalliance.org/

[4] "Cost of a data breach 2024," IBM. [Online]. Available: https://www.ibm.com/think/insights/cost-of-a-data-breach-2024

[5] P. Moosavinezhad, "Hybrid Cloud Governance: Don't Risk Non-Compliance!" Devoteam. [Online]. Available: https://www.devoteam.com/

[6] "Cloud Data Sovereignty Governance and Risk Implications," ISACA, 2024.

[7] S. Ali et al., "Security and privacy in multi-cloud and hybrid cloud environments," Comput. Secur., vol. 157, p. 104599, Oct. 2025.

[8] I. Mahendra et al., "IT Challenges for Integrated GRC," in 2022 Int'l Conf. APICS, pp. 79–84, doi: 10.1109/APICS56469.2022.9918797.

[9] R. M. Boodai et al., "An Approach to Address Risk Management Challenges," in IEEE CSR 2022, pp. 184–188.

[10] C. Schmitz et al., "Maturity level assessments of information security controls," Comput. Secur., vol. 108, p. 102306, 2021.

[11] O. M. M. Al-Matari et al., "Adopting security maturity model to the organizations' capability model," Egypt. Inform. J., vol. 22, no. 2, pp. 193–199, 2021.

[12] S. AlGhamdi et al., "Information security governance challenges and critical success factors," Comput. Secur., vol. 99, p. 102030, 2020.

[13] G. Cheng et al., "Cloud data governance maturity model," in 2017 IEEE ICSESS, pp. 517–520.

# โมเดลการประเมินความพร้อมทางด้านขีดความสามารถในการฟื้นตัวและรับมือภัยคุกคามทางไซเบอร์

# An Assessment Model of Cyber Resilience Readiness and Recovery Capability

*เกียรติศักดิ์ ลุยทอง (Kiattisak Luithong)[1] และนลินภัสร์ บำเพ็ญเพียร (Nalinpat Bhumpenpein)[2]*
*ภาควิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล*
*มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ*
*[1]s6707011857152@email.kmutnb.ac.th, [2]nalinpat.b@itd.kmutnb.ac.th*

## บทคัดย่อ

*งานวิจัยนี้มีวัตถุประสงค์เพื่อนำเสนอโมเดลการประเมินความพร้อมด้านขีดความสามารถในการฟื้นตัวและการรับมือภัยคุกคามทางไซเบอร์ โดยใช้กรอบมาตรฐานของ NIST Cybersecurity Framework (CSF) Version 2.0 และมาตรฐานที่เกี่ยวข้อง จากกลุ่มตัวอย่างจำนวน 219 คน ประกอบด้วยผู้บริหาร 17 คน เจ้าหน้าที่เทคนิค 38 คน และผู้ใช้งาน 164 คน เพื่อสะท้อนมุมมองในทุกระดับ ผลการวิจัยพบว่า กลุ่มผู้บริหารมีระดับความพร้อมเชิงวุฒิภาวะเฉลี่ยเท่ากับ 3.90 คะแนน จัดอยู่ในระดับ 4 (Quantitatively Managed) ขณะที่เจ้าหน้าที่เทคนิคและผู้ใช้งานมีค่าเฉลี่ย 3.10 และ 2.93 คะแนน ตามลำดับ อยู่ในระดับ 3 (Defined) สะท้อนถึงความพร้อมเชิงปฏิบัติการระดับพื้นฐานถึงปานกลาง และทุกกลุ่มให้ความสำคัญต่อฟังก์ชันด้านการกำกับดูแล การป้องกัน และการฟื้นฟูระบบหรือการกู้คืน อยู่ในระดับสูง แสดงถึงการตระหนักรู้เชิงยุทธศาสตร์เกี่ยวกับการบริหารความเสี่ยงและการฟื้นตัว เมื่อเปรียบเทียบค่าความสำคัญกับความพร้อมเชิงวุฒิภาวะ พบว่าการระบุความเสี่ยง และการตรวจจับและเฝ้าระวังเหตุการณ์ไซเบอร์ ถึงจะได้รับการประเมินว่ามีความสำคัญสูง แต่ความพร้อมเชิงปฏิบัติยังไม่สอดคล้องกัน สำหรับผลการประเมินความยืดหยุ่นทางไซเบอร์ พบว่า กลุ่มผู้บริหารมีค่าความยืดหยุ่นสูงสุด รองลงมาคือกลุ่มเจ้าหน้าที่เทคนิค และกลุ่มผู้ใช้งานทั่วไป สะท้อนความแตกต่างของบทบาทหน้าที่ ความรับผิดชอบ และการมีส่วนร่วมในกระบวนการบริหารความมั่นคงปลอดภัยไซเบอร์*

**คำสำคัญ:** ความยืดหยุ่นทางไซเบอร์ ความพร้อมเชิงวุฒิภาวะ ค่าความสำคัญ

## Abstract

*This study proposes a readiness assessment model for cyber threat response and recovery capabilities based on the NIST Cybersecurity Framework (CSF) Version 2.0. Data were collected from 219 participants, including 17 executives, 38 technical staff, and 164 general users, to reflect multi-level organizational perspectives. Results show that executives achieved the highest average maturity score (3.90), corresponding to Level 4 (Quantitatively Managed), while technical staff (3.10) and general users (2.93) were at Level 3 (Defined), indicating foundational to moderate operational readiness. All groups rated Governance, Protect, and Recover as highly important, reflecting strong strategic awareness. However, a gap was identified in the Identify and Detect functions, where perceived importance exceeded operational maturity. Cyber resilience results further indicate that executives demonstrated the highest resilience level, followed by technical staff and general users, highlighting role-based differences in cybersecurity involvement. The findings emphasize the need to align strategic priorities with operational capability development to strengthen organizational cyber resilience systematically and sustainably.*

**Keywords:** Cyber Resilience, Maturity Level, Importance Degree.

## 1. บทนำ

ความก้าวหน้าทางเทคโนโลยีดิจิทัลและระบบเครือข่ายสารสนเทศมีบทบาทสำคัญต่อการปฏิบัติภารกิจของหน่วยงาน เช่น กองทัพบก ซึ่งเป็นโครงสร้างพื้นฐานสำคัญทางสารสนเทศตามพระราชบัญญัติการรักษาความมั่นคงปลอดภัยไซเบอร์ พ.ศ.2562 [1] การพึ่งพาเทคโนโลยีในระดับสูงส่งผลให้ภัยคุกคามทางไซเบอร์มีความซับซ้อน ส่งผลกระทบเชิงยุทธศาสตร์ต่อความลับ ความถูกต้อง และความพร้อมใช้งานของระบบ รวมถึงความเชื่อมั่นขององค์กร แม้ว่ากรอบ NIST Cybersecurity Framework (CSF) Version 2.0 [2] และมาตรฐานที่เกี่ยวข้อง [3][4][5] จะถูกนำมาใช้ในการบริหารความเสี่ยงไซเบอร์แต่ยังมีข้อจำกัดเชิงปริมาณ และไม่สามารถสะท้อนระดับความพร้อมและความยืดหยุ่นขององค์กรได้ ดังนั้น การขาดแนวทางในการประเมินที่เชื่อมโยงระดับความสำคัญ กับระดับความพร้อมเชิงวุฒิภาวะ ยังมีช่องว่างระหว่างนโยบายและการปฏิบัติ และยังไม่มีการวัดความยืดหยุ่นทางไซเบอร์อย่างเป็นระบบ งานวิจัยนี้จึงเสนอแนวคิดการประเมินวุฒิภาวะด้านความยืดหยุ่นทางไซเบอร์ (Cyber Resilience Maturity) โดยมีวัตถุประสงค์เพื่อพัฒนาโมเดลการประเมินเพื่อวิเคราะห์ระดับความพร้อมเชิงวุฒิภาวะความสำคัญ และความยืดหยุ่นขององค์กรอย่างเป็นระบบ เพื่อสนับสนุนการยกระดับความมั่นคงปลอดภัยไซเบอร์อย่างมีประสิทธิภาพ

## 2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

**2.1 พระราชบัญญัติการรักษาความมั่นคงปลอดภัยไซเบอร์ พ.ศ. 2562**

เป็นกฎหมายเพื่อรับมือภัยคุกคามไซเบอร์ โดยได้กำหนดโครงสร้างการกำกับดูแลระดับชาติ บทบาทของหน่วยงาน และมาตรการป้องกัน ตรวจสอบ เฝ้าระวัง และตอบสนองต่อเหตุการณ์ไซเบอร์ และบังคับให้หน่วยงานรัฐดำเนินมาตรการด้านความมั่นคงปลอดภัยไซเบอร์ให้สอดคล้องกับนโยบายและมาตรฐานระดับชาติ [1]

**2.2 กรอบการทำงานด้านความมั่นคงปลอดภัยไซเบอร์**

กรอบการทำงานด้านความมั่นคงปลอดภัยไซเบอร์ (Cybersecurity Framework: CSF) เวอร์ชัน 2.0 ของ NIST เป็นกรอบเชิงกลยุทธ์ในการบูรณาการความเสี่ยงไซเบอร์กับการบริหารความเสี่ยงขององค์กร ครอบคลุมการประเมินความเสี่ยง ลำดับความสำคัญ การตอบสนอง และการฟื้นคืนสภาพจากเหตุการณ์ไซเบอร์[2]

**2.3 มาตรฐานสากลด้านระบบการจัดการความมั่นคงปลอดภัยสารสนเทศ**

เป็นมาตรฐานด้านระบบการจัดการความมั่นคงปลอดภัยสารสนเทศ มีข้อกำหนด และปรับปรุงกระบวนการบริหารความมั่นคงปลอดภัยสารสนเทศ อย่างต่อเนื่อง โดยยึดแนวคิดการบริหารความเสี่ยงเป็นแกนหลัก เพื่อให้สอดคล้องกับมาตรฐานระบบการตามหลักการรักษาความลับ ความถูกต้อง และความพร้อมใช้ โดยกำหนดผู้มีส่วนได้ส่วนเสีย ข้อกำหนดทางกฎหมาย ทรัพยากร และระดับความเสี่ยงที่ยอมรับได้[3]

**2.4 ประมวลแนวทางปฏิบัติและกรอบมาตรฐานด้านการรักษาความมั่นคงปลอดภัยไซเบอร์สำหรับหน่วยงานของรัฐ และหน่วยงานโครงสร้างพื้นฐานสำคัญทางสารสนเทศ พ.ศ.2564**

เป็นมาตรฐานขั้นต่ำสำหรับหน่วยงานของรัฐและหน่วยงานโครงสร้างพื้นฐานสำคัญทางสารสนเทศ (CII) ในการป้องกัน รับมือ และฟื้นฟูจากภัยคุกคามทางไซเบอร์ กำหนดให้หน่วยงานดำเนินมาตรการตรวจสอบความมั่นคงปลอดภัยไซเบอร์ การประเมิน จัดการความเสี่ยง การจัดทำแผนรับมือเหตุการณ์ไซเบอร์ กระบวนการตอบสนอง กรอบดังกล่าวเป็นมาตรฐานบังคับใช้ที่เสริมสร้างความพร้อมความยืดหยุ่น และความต่อเนื่องของการดำเนินงาน เพื่อยกระดับความมั่นคงปลอดภัยไซเบอร์ [4], [5]

**2.5 ระดับความสามารถและวุฒิภาวะ**

เป็นแนวปฏิบัติสากลเพื่อยกระดับสมรรถนะและผลการดำเนินงานขององค์กร ออกแบบให้ยืดหยุ่น ใช้ได้ทุกอุตสาหกรรม และบูรณาการกับมาตรฐานและระเบียบ

ได้รับการยอมรับอย่างแพร่หลายจากองค์กรที่ผ่านการประเมินกว่า 14,000 แห่งทั่วโลก พร้อมกลไกสนับสนุนการนำไปใช้จริงผ่านโครงสร้างพื้นฐานการดำเนินงานและการบริหารสมรรถนะด้วยการวัดผล เพื่อให้การปรับปรุงเกิดขึ้นอย่างยั่งยืนในระดับองค์กร [6]

**2.6 ความยืดหยุ่นเชิงไซเบอร์**

เป็นความสามารถ ความทนทาน ตอบสนอง และปรับตัวต่อภัยคุกคามไซเบอร์ เช่น APT, Ransomware, Supply Chain Attacks, Cloud, IoT และ OT โดยเชื่อมโยงกับมาตรฐานสากล เช่น NIST CSF 2.0 NIST SP 800 ISO/IEC 27001 และนโยบายด้านความมั่นคงไซเบอร์ของประเทศไทย ที่เน้นความยืดหยุ่นและการปรับปรุงอย่างต่อเนื่อง [7]

**2.7 แนวคิดการประเมินความยืดหยุ่นทางไซเบอร์**

แนวคิดการประเมินความยืดหยุ่นทางไซเบอร์ (Cyber Resilience) พัฒนาจากการประเมินความมั่นคงปลอดภัยแบบดั้งเดิม มุ่งเน้นความสามารถในการป้องกัน ต้านทาน ฟื้นตัวและปรับตัวต่อภัยคุกคาม งานวิจัยของ Luo et al. (2025) ได้เสนอโมเดลการประเมินเชิงปริมาณในรูปแบบ Cybersecurity Resilience Capacity (RC) ซึ่งประกอบด้วย 4 มิติหลัก ได้แก่ P(t) คือ ความสามารถในการป้องกันและคาดการณ์ R(t) คือ ความสามารถในการต้านทานและลดผลกระทบ A(t) คือ ความสามารถในการฟื้นตัวของระบบให้กลับสู่สภาวะปกติ และ L(t) คือ ความสามารถในการเรียนรู้จากเหตุการณ์ที่ผ่านมาและปรับปรุงกลไกการป้องกันอย่างต่อเนื่อง[8]

## 3. วิธีการดำเนินงาน

**3.1 ขั้นตอนการดำเนินวิจัย**

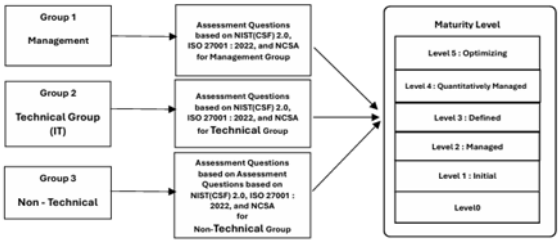
การวิจัยนี้เป็นการวิจัยเชิงสำรวจ จากกลุ่มตัวอย่างจำนวน 219 คน ประกอบด้วยผู้บริหาร เจ้าหน้าที่เทคนิค และผู้ใช้งานทั่วไป โดยใช้โมเดลการประเมิน Cyber Resilience Maturity และวิเคราะห์ข้อมูลโดยใช้ ค่าเฉลี่ย (Mean) ระดับความพร้อมเชิงวุฒิภาวะ (Maturity Level) ค่าความสำคัญ (Importance Degree) พร้อมวิเคราะห์ช่องว่าง (Gap Analysis) และการคำนวณค่าความยืดหยุ่นทางไซเบอร์ (Cyber Resilience) เพื่อสะท้อนระดับความพร้อมขององค์กรอย่างเป็นระบบ

**3.1 ประชากรและกลุ่มตัวอย่าง**

งานวิจัยในครั้งนี้ ได้เก็บข้อมูลระหว่างเดือน ธันวาคม 2568 ถึงเดือน มกราคม 2569 จากกลุ่มผู้เกี่ยวข้องกับระบบสารสนเทศ ทั้งด้าน นโยบาย ด้านเทคนิค และด้านการใช้งาน จากกลุ่มเป้าหมาย 3 กลุ่ม ได้แก่ ผู้บริหาร 17 คน เจ้าหน้าที่เทคนิค 38 คน และผู้ใช้งาน 164 คน รวมจำนวน 219 คน

**3.2 โมเดลการประเมินความพร้อมทางด้านขีดความสามารถในการฟื้นตัวและรับมือภัยคุกคามทางไซเบอร์**

โมเดลสำหรับงานวิจัยนี้ได้ออกแบบตามแนวคิดการประเมินความพร้อมทางด้านขีดความสามารถในการฟื้นตัวและรับมือภัยคุกคามทางไซเบอร์ โดยแบ่งเป็น 3 กลุ่ม ได้แก่ กลุ่มผู้บริหาร กลุ่มเจ้าหน้าที่เทคนิค และกลุ่มผู้ใช้งานทั่วไป เพื่อสะท้อนมุมมองตามมาตรฐาน NIST Cybersecurity Framework (CSF) Version 2.0 ผสานกับ ISO/IEC 27001:2022 และแนวทางของหน่วยงานกำกับดูแลด้านความมั่นคงปลอดภัยไซเบอร์ นำมาวิเคราะห์ระดับความพร้อมเชิงวุฒิภาวะ (Maturity Level) ค่าความสำคัญ (Importance Degree) และความยืดหยุ่นทางไซเบอร์ ด้วยแบบจำลอง Cyber Resilience Capacity (RC) ประกอบด้วย การป้องกัน (Prevention: P) ความต้านทาน (Destruction Resistance: R) การฟื้นฟู (Adaptive Recovery: A) และการเรียนรู้และปรับปรุงจากเหตุการณ์ที่เกิดขึ้น (Evolutionary Learning: L)




**ภาพที่ 1:** โมเดลการวิจัย

## 4. ผลการดำเนินงานวิจัย

**4.1 ผลการวิเคราะห์ความพร้อมเชิงวุฒิภาวะและค่าความสำคัญ**

จากการวิเคราะห์ข้อมูล พบว่าผู้บริหาร มีค่า Maturity เฉลี่ยสูงสุดที่ 3.90 อยู่ในระดับ 4 (Quantitatively Managed) และให้ความสำคัญสูงสุดกับ Governance และ Recover

(4.14, 0.177 และ 4.12, 0.176) สะท้อนมุมมองเชิงนโยบายและความต่อเนื่องขององค์กร ในขณะที่เจ้าหน้าที่เทคนิคมีค่าเฉลี่ย 3.10 อยู่ในระดับ 3 (Defined) โดยให้ความสำคัญกับ Protect สูงสุด (3.43, 0.185) รองลงมาคือ Detect และ Governance แสดงถึงบทบาทเชิงปฏิบัติการด้านการป้องกันและตรวจจับ ส่วนผู้ใช้งานทั่วไป มีค่าเฉลี่ย 2.93 อยู่ในระดับ 3 เช่นกัน โดยให้ความสำคัญกับ Governance (3.11, 0.177) และ Identify (3.01, 0.171) มากกว่า ด้าน Protect, Detect และ Respond (2.84, 0.161) ซึ่งสะท้อนว่าการป้องกันและตอบสนองยังถูกมองเป็นหน้าที่ของฝ่ายเทคนิค ภาพรวมแสดงให้เห็นความพร้อมระดับนโยบายแต่ยังมีช่องว่างในการบูรณาการเชิงเทคนิคและการมีส่วนร่วมของผู้ใช้งานเพื่อสร้างความสมดุลด้าน Cyber Resilience ขององค์กร

**ตารางที่ 1:** ผลการวิเคราะห์ระดับความพร้อมเชิงวุฒิภาวะ (Maturity Level) และค่าความสำคัญ (Importance Degree)

| | | Govern | Identify | Protect | Detect | Respond | Recover | Total |
|---|---|---|---|---|---|---|---|---|
| ผู้บริหาร | คะแนน | 82.75 | 78.24 | 75.00 | 72.79 | 76.47 | 82.35 | 77.9 |
| | ความพร้อมเชิงวุฒิภาวะ | 4.14 | 3.91 | 3.75 | 3.64 | 3.82 | 4.12 | 3.90 |
| | ค่าความสำคัญ | 0.177 | 0.167 | 0.16 | 0.156 | 0.164 | 0.176 | 1.00 |
| เจ้าหน้าที่เทคนิค | คะแนน | 64.33 | 57.8 | 68.6 | 64.74 | 54.31 | 61.84 | 61.90 |
| | ความพร้อมเชิงวุฒิภาวะ | 3.22 | 2.89 | 3.43 | 3.24 | 2.72 | 3.09 | 3.10 |
| | ค่าความสำคัญ | 0.173 | 0.156 | 0.185 | 0.174 | 0.146 | 0.166 | 1.00 |
| ผู้ใช้งาน | คะแนน | 62.30 | 60.25 | 56.85 | 56.83 | 56.76 | 58.92 | 58.70 |
| | ความพร้อมเชิงวุฒิภาวะ | 3.11 | 3.01 | 2.84 | 2.84 | 2.84 | 2.95 | 2.93 |
| | ค่าความสำคัญ | 0.177 | 0.171 | 0.161 | 0.161 | 0.161 | 0.167 | 1.00 |

### 4.2 ภาพรวมการวิเคราะห์ความยืดหยุ่นทางไซเบอร์

ผลการประเมินความยืดหยุ่นทางไซเบอร์โดยรวมสะท้อนให้เห็นถึงความแตกต่างแต่ละระดับ โดยกลุ่มผู้บริหารมีค่า RC(t)=0.64 อยู่ในระดับค่อนข้างดี มีจุดแข็งด้านการฟื้นฟูแบบปรับตัว A(t)=0.824 และการเรียนรู้การเรียนรู้และปรับปรุงจากเหตุการณ์ที่เกิดขึ้น L(t)=0.826 ขณะที่การป้องกัน P(t)=0.75 และความต้านทานต่อความเสียหาย R(t)=0.747 ยังสามารถพัฒนาให้สมดุลยิ่งขึ้น ส่วนกลุ่มเจ้าหน้าที่เทคนิคมีค่า RC(t)=0.57 อยู่ในระดับปานกลาง มีจุดแข็งด้านการป้องกัน P(t)=0.686 และการฟื้นฟู A(t)=0.618 แต่ยังมีข้อจำกัดด้านความต้านทานต่อความเสียหาย R(t)=0.595 และการเรียนรู้การเรียนรู้และปรับปรุงจากเหตุการณ์ที่เกิดขึ้น L(t)=0.631 ที่ยังไม่เชื่อมโยงเชิงโครงสร้างขององค์กร ทางด้านกลุ่มผู้ใช้งาน มีค่า RC(t)=0.49 อยู่ในระดับปานกลาง โดยทุกมิติอยู่ในช่วงใกล้เคียงกัน ได้แก่ P(t)=0.568, R(t)=0.568, A(t)=0.589 และ L(t)=0.606 สะท้อนบทบาทที่เน้นไปทางด้านการทำงานเชิงรับและปฏิบัติตามนโยบายมากกว่าการมีส่วนร่วม โดยภาพรวมพบว่ามีความยืดหยุ่นทางไซเบอร์ที่แข็งแกร่งในระดับนโยบายและการฟื้นฟู แต่ยังจำเป็นต้องยกระดับความต้านทานเชิงเทคนิคและการมีส่วนร่วมของบุคลากรทุกระดับ

**ตารางที่ 2:** แสดงผลวิเคราะห์ค่าความยืดหยุ่นทางไซเบอร์

| กลุ่ม | P(t) | R(t) | A(t) | L(t) | RC(t) |
|---|---|---|---|---|---|
| ผู้บริหาร | 0.752 | 0.747 | 0.824 | 0.826 | 0.64 |
| เจ้าหน้าที่เทคนิค | 0.686 | 0.595 | 0.618 | 0.631 | 0.57 |
| ผู้ใช้งาน | 0.568 | 0.568 | 0.589 | 0.606 | 0.49 |

### 4.3 ความยืดหยุ่นทางไซเบอร์ระดับผู้บริหาร

ผลการประเมินความยืดหยุ่นทางไซเบอร์ของผู้บริหาร พบว่าการป้องกัน (P(t)) มีการกำหนดนโยบาย กรอบควบคุม และมาตรการด้านความมั่นคงปลอดภัยอย่างเป็นระบบ แต่การนำไปปฏิบัติยังไม่ครอบคลุมทุกระบบอย่างสมบูรณ์ สำหรับ ความต้านทานและการตอบสนอง (R(t)) องค์กรสามารถตรวจจับและตอบสนองเหตุการณ์ได้ อย่างไรก็ตาม มาตรฐานการปฏิบัติและการบูรณาการข้อมูลจากหลายแหล่งยังไม่เป็นหนึ่งเดียวในด้าน การฟื้นฟูแบบปรับตัว (A(t)) องค์กรมีแผนความต่อเนื่องและสามารถฟื้นฟูระบบภายในระยะเวลาที่กำหนด ถือเป็นจุดแข็งสำคัญ ขณะที่ การเรียนรู้และปรับปรุง (L(t)) มีการนำบทเรียนจากเหตุการณ์มาปรับปรุงนโยบายและกระบวนการอย่างต่อเนื่อง แต่ยังเป็นลักษณะเชิงรับมากกว่าเชิงรุก

**ตารางที่ 3:** ตารางความยืดหยุ่นทางไซเบอร์ผู้บริหาร

| | วิธีการประเมิน | ผลการตรวจพบ |
|---|---|---|
| P(t) | ทบทวนนโยบาย การควบคุมการเข้าถึงและการสร้างความตระหนักรู้ | มีมาตรการกำหนดชัดเจน แต่การนำไปใช้ยังไม่ครอบคลุมทุกระบบ |
| R(t) | ทบทวนกระบวนการตรวจจับ วิเคราะห์เหตุการณ์ และแผน Incident Response | สามารถตรวจจับและตอบสนองเหตุการณ์ได้ แต่มาตรฐานการปฏิบัติยังไม่เป็นหนึ่งเดียว |
| A(t) | ตรวจสอบแผนความต่อเนื่องและการทดสอบแผน | มีแผนการฝึกซ้อม สามารถฟื้นฟูระบบภายในระยะเวลาที่กำหนด |
| L(t) | ทบทวนนโยบายระดับบริหารและกระบวนการติดตามบทเรียน | นำบทเรียนจากเหตุการณ์มาปรับปรุงนโยบายและกระบวนการ แต่ยังเป็นลักษณะเชิงรับมากกว่าเชิงรุก |

**4.4 ความยืดหยุ่นทางไซเบอร์ระดับเจ้าหน้าที่เทคนิค**

ผลการประเมินความยืดหยุ่นทางไซเบอร์ของเจ้าหน้าที่เทคนิคมีมาตรการด้านความมั่นคงปลอดภัยไซเบอร์ชัดเจน โดยเฉพาะด้านการป้องกัน (P(t)) มีการกำหนดมาตรการควบคุมสิทธิ์และความมั่นคงปลอดภัยเครือข่ายแล้ว แต่ยังไม่ครอบคลุมทุกระบบ ขณะที่ การตรวจจับและตอบสนอง (R(t)) สามารถตรวจจับเหตุการณ์ได้ แต่การรวบรวมข้อมูลยังไม่ครบทุกแหล่ง และขั้นตอนการตอบสนองยังไม่เป็นมาตรฐานเดียวกัน ด้าน การฟื้นฟูระบบ (A(t)) องค์กรสามารถกู้คืนระบบได้ภายในระยะเวลาที่กำหนด แต่ต้องพัฒนาแนวทางรองรับเหตุการณ์ที่มีความซับซ้อน ส่วนการเรียนรู้และการปรับปรุง (L(t)) มีการนำบทเรียนจากเหตุการณ์มาปรับปรุงกระบวนการแล้ว แต่ยังไม่พัฒนาไปสู่การเรียนรู้เชิงรุกหรือการปรับปรุงอย่างต่อเนื่อง

**ตารางที่ 4:** ตารางความยืดหยุ่นทางไซเบอร์เจ้าหน้าที่เทคนิค

| | วิธีการประเมิน | ผลการตรวจพบ |
|---|---|---|
| P(t) | ตรวจสอบการตั้งค่ามาตรการควบคุมสิทธิ์ และความมั่นคงปลอดภัยเครือข่าย | มีมาตรการป้องกันและการควบคุมสิทธิ์ แต่ยังไม่ครอบคลุมทุกระบบ |
| R(t) | ทบทวนระบบเฝ้าระวัง Log และขั้นตอน Incident Respons | ตรวจจับเหตุการณ์ได้ แต่ยังไม่ครบทุกแหล่ง การตอบสนองยังไม่เป็นมาตรฐานเดียวกัน |
| A(t) | ตรวจสอบแผนกู้คืนและการทดสอบตามแผน | สามารถฟื้นฟูระบบได้ในเวลาที่กำหนด แต่ต้องพัฒนากรณีเหตุการณ์ซับซ้อน |
| L(t) | ทบทวนแนวปฏิบัติ บทเรียน และการปรับปรุงกระบวนการ | มีการนำบทเรียนมาปรับปรุงกระบวนการ แต่ยังไม่พัฒนาเชิงรุก |

**4.5 ความยืดหยุ่นทางไซเบอร์ระดับผู้ใช้งานทั่วไป**

ผลการประเมินความยืดหยุ่นทางไซเบอร์ของผู้ใช้งานทั่วไปพบว่ามีความยืดหยุ่นทางไซเบอร์ในระดับพื้นฐาน โดยด้านการป้องกัน (P(t)) มีการปฏิบัติตามมาตรการพื้นฐานและควบคุมสิทธิ์ตามนโยบายองค์กร แต่ยังขาดความเข้าใจเชิงลึกด้านการป้องกันภัยคุกคาม ขณะที่ การตรวจจับและตอบสนอง (R(t)) สามารถแจ้งเหตุและปฏิบัติตามขั้นตอนที่กำหนดได้ แต่ยังไม่สามารถประเมินความรุนแรงของเหตุการณ์ได้ด้วยตนเอง ด้าน การฟื้นฟู (A(t)) ผู้ใช้งานสามารถกลับมาใช้งานระบบได้ตามเวลาและขั้นตอนที่กำหนด แต่ต้องพึ่งพาด้านเทคนิคเป็นหลัก ส่วนการเรียนรู้และการปรับปรุง (L(t)) มีการปรับพฤติกรรมตามนโยบายองค์กร แต่ยังเป็นลักษณะเชิงรับมากกว่าเชิงรุก ภาพรวมชี้ให้เห็นว่าผู้ใช้งานมีบทบาทสนับสนุนมาตรการขององค์กรในระดับปฏิบัติ แต่ต้องเสริมสร้างความตระหนักรู้และการมีส่วนร่วมเชิงรุก

**ตารางที่ 5:** ตารางความยืดหยุ่นทางไซเบอร์ระดับผู้ใช้งานทั่วไป

| | วิธีการประเมิน | ผลการตรวจพบ |
|---|---|---|
| P(t) | ตรวจสอบการใช้งานระบบตามนโยบายและพฤติกรรมการปฏิบัติตามมาตรการองค์กร | มีการปฏิบัติตามมาตรการพื้นฐานและควบคุมสิทธิ์ แต่ยังขาดความเข้าใจเชิงลึกด้านการป้องกันภัยคุกคาม |
| R(t) | ประเมินการแจ้งเหตุ การรับรู้เหตุการณ์ และการปฏิบัติตามแนวทาง Incident | สามารถแจ้งเหตุและปฏิบัติตามขั้นตอนได้ แต่ยังไม่สามารถประเมินความรุนแรงต่อเหตุการณ์ได้ด้วยตนเอง |
| A(t) | ตรวจสอบการปฏิบัติตามขั้นตอนกู้คืนหลังเหตุการณ์ | สามารถกลับมาใช้งานระบบได้ตามเวลาและขั้นตอนที่กำหนด แต่ต้องพึ่งพาหน่วยงานเทคนิคเป็นหลัก |
| L(t) | ประเมินการรับรู้บทเรียนและการปรับพฤติกรรมตามนโยบาย | มีการปรับพฤติกรรมตามนโยบายองค์กร แต่ยังเป็นลักษณะเชิงรับมากกว่าเชิงรุก |

## 5. สรุปและอภิปรายผล

งานวิจัยในครั้งนี้สะท้อนแนวคิดความยืดหยุ่นทางไซเบอร์โดยใช้กรอบ NIST Cybersecurity Framework (CSF) 2.0 เป็นโครงสร้างหลัก และพัฒนาเป็นแนวคิดการประเมินวุฒิภาวะด้านความยืดหยุ่นทางไซเบอร์ (Cyber Resilience Maturity) ช่วยให้สามารถวิเคราะห์ระดับความพร้อมเชิงวุฒิภาวะ (Maturity Level) และความสำคัญ (Importance Degree) ซึ่งผลการวิจัยพบว่าฟังก์ชัน Governance มีบทบาทสำคัญในฐานะกลไกหลักขององค์กร ซึ่งสอดคล้องกับ Bernardo et al. [9] และ Ibrahim et al. [10] ที่ชี้ให้เห็นว่า Governance เป็นรากฐานของการสร้างความยืดหยุ่นเชิงองค์กร ซึ่งองค์กรจะรับมือกับภัยคุกคามได้อย่างมีประสิทธิภาพเมื่อมีการกำกับดูแลที่ชัดเจนและสามารถตรวจจับและปรับตัวได้อย่างเป็นระบบ ในขณะที่ฟังก์ชัน Identify และ Detect ยังคงเป็นจุดอ่อนสำคัญ โดยเฉพาะ Identify ซึ่งสอดคล้องกับ López-Vasco et al. [11] ที่พบช่องว่างด้านการประเมินความเสี่ยง และชี้ให้เห็นถึงความจำเป็นในการพัฒนาความสามารถในการระบุและประเมินภัยคุกคาม นอกจากนี้ ฟังก์ชัน Detect ยังมีระดับความสามารถต่ำกว่าฟังก์ชันอื่น โดยเฉพาะในบริบทของภัยคุกคามขั้นสูง สอดคล้องกับ Liyanage et al. [12], Fadila et al. [13] และ Adeniyan [14] ที่ระบุถึงข้อจำกัดด้านการเฝ้าระวังและการตรวจจับเชิงรุก การบูรณาการตัวแปร Maturity และ

Importance จากงานวิจัยนี้สามารถสรุปได้ว่า ช่องว่างในฟังก์ชัน Identify และ Detect ไม่ได้เกิดจากการขาดความตระหนักรู้เพียงอย่างเดียว แต่เกิดจากข้อจำกัด เชิงโครงสร้าง เทคโนโลยี และการบริหารจัดการ ดังนั้น การยกระดับทั้งในเชิงนโยบาย เทคโนโลยี และบุคลากร จึงเป็นปัจจัยสำคัญในการลดช่องว่างรวมถึงการเสริมสร้างความมั่นคงปลอดภัยไซเบอร์ด้าน Respond และ Recover พบว่าความสามารถในการฟื้นตัว โดยเฉพาะแนวคิดAdaptive Recovery มีบทบาทสำคัญต่อ Cyber Resilience สอดคล้องกับ Tzavara and Vassiliadis [15], Luo et al. [16] และ Liyanage et al. [17] ที่เสนอการฟื้นตัวเชื่อมโยงกับการเรียนรู้อย่างเป็นระบบ

งานวิจัยนี้สะท้อนให้เห็นถึงช่องว่างระหว่างกลยุทธ์กับความสามารถเชิงปฏิบัติอย่างชัดเจน กล่าวคือ หากองค์กรให้ความสำคัญสูงแต่มีระดับความพร้อมต่ำ จะส่งผลให้เกิดช่องว่างที่กระทบโดยตรงต่อระดับความยืดหยุ่นทางไซเบอร์ เมื่อพิจารณาในเชิงกรอบแนวคิด NIST CSF พบว่า Tier เหมาะสำหรับการกำกับดูแลและบริบทเชิงยุทธศาสตร์ เนื่องจากเป็นการจัดระดับเชิงคุณภาพ ในขณะที่ค่า Maturity และค่า RC(t) สามารถสะท้อนความสามารถเชิงปฏิบัติการในแต่ละฟังก์ชันได้อย่างชัดเจนผลการประเมินยังแสดงให้เห็นว่ากลุ่มที่มีระดับ Maturity สูง จะมีค่า Adaptive Recovery และ Learning สูงไปด้วย

ดังนั้น งานวิจัยนี้ชี้ให้เห็นว่า การประเมิน Cyber Resilience Maturity ควรพิจารณาทั้งมิติการกำกับดูแลและความสามารถเชิงปฏิบัติร่วมกัน โดย Maturity สะท้อนความพร้อมเชิงกระบวนการ ขณะที่ RC(t) แสดงประสิทธิภาพเชิงปริมาณของการป้องกัน ต้านทาน ฟื้นตัว และการเรียนรู้ ซึ่งช่วยให้เห็นช่องว่างและแนวทางพัฒนาได้ชัดเจน

## เอกสารอ้างอิง


[1] พระราชบัญญัติการรักษาความมั่นคงปลอดภัยไซเบอร์ พุทธศักราช 2562. ราชกิจจานุเบกษา, เล่ม 136 ตอนที่ 69 ก, 27 พฤษภาคม 2562.

[2] National Institute of Standards and Technology (NIST). Cybersecurity Framework (CSF) Version 2.0. Gaithersburg, MD: NIST, 2024.

[3] International Organization for Standardization (ISO). ISO/IEC 27001:2022 Information Security,Cybersecurity and Privacy Protection Information Security Management Systems Requirements. Geneva: ISO/IEC, 2022.

[4] ประมวลแนวทางปฏิบัติและกรอบมาตรฐานด้านการรักษาความมั่นคงปลอดภัยไซเบอร์สำหรับหน่วยงานของรัฐและหน่วยงานโครงสร้างพื้นฐานสำคัญทางสารสนเทศ พ.ศ. 2564. ราชกิจจานุเบกษา, เล่ม 138 ตอน พิเศษ 208 ง, 6 กันยายน 2564.

[5] สำนักงานคณะกรรมการการรักษาความมั่นคงปลอดภัยไซเบอร์แห่งชาติ (สกมช.). นโยบายและแผนปฏิบัติการว่าด้วยการรักษาความมั่นคงปลอดภัยไซเบอร์ พ.ศ. 2565–2570.ราชกิจจานุเบกษา, เล่ม 136 ตอนพิเศษ 288 ง, 9 ธันวาคม 2565.

[6] ISACA. CMMI Model Quick Reference Guide (CMMI V3.0): An overview of the Capability Maturity Model Integration (CMMI)® Model. Schaumburg, IL: ISACA, 2024.

[7] National Institute of Standards and Technology, Developing Cyber-Resilient Systems: A Systems Security Engineering Approach, NIST SP 800-160, Vol. 2, Rev. 1.

[8] M. Luo, C. Tao, Y. Liu, S. Chen, and P. Chen, “An endogenous security-oriented framework for cyber resilience assessment in critical infrastructures,” Applied Sciences, vol. 15, no. 8342, pp. 1–17, 2025.

[9] Bernardo, L., Malta, S., & Magalhães, J. An Evaluation Framework for Cybersecurity Maturity Aligned with the NIST CSF. Electronics, 14(7), 1364. MDPI, 2025.

[10] Ibrahim, A., Valli, C., McAteer, I., & Chaudhry, J. A security review of local government using NIST CSF: a case study. The Journal of Supercomputing, 74, 5171–5186, 2018.

[11] López-Vasco, F., Cadena, B., Villarroel, Y., & co-authors. Mitigating Dual-Use Cybersecurity Risks in Latin American Higher Education: Implementing the NIST CSF 2.0 for Institutional Resilience. Preprint ResearchGate, 2025.

[12] Liyanage, L., Arachchilage, N., & Russello, G. A Novel Framework to Assess Cybersecurity Capability Maturity. Proceedings of the Thirty-Third European Conference on Information Systems (ECIS 2025), Amman, Jordan., 2025.

[13] Fadila, V., Mutiah, N., & Sari, R. P. Cyber security audit using CIS CSC, NIST CSF and COBIT 2019 framework. CESS (Journal of Computing Engineering, System and Science), 8(2), 271–283, 2023.

[14] Adeniyan, T. Cyber Risk Assessment Frameworks for Protecting U.S. Critical Infrastructure Against Emerging Threats. IRE Journals, 8(8), 821–830. ISSN 2456-8880, 2025.

[15] V. Tzavara and S. Vassiliadis, “Tracing the evolution of cyber resilience: a historical and conceptual review,” International Journal of Information Security, vol. 23, pp. 1695–1719, 2024.

[16] M. Luo, C. Tao, Y. Liu, S. Chen, and P. Chen, “An endogenous security-oriented framework for cyber resilience assessment in critical infrastructures,” Applied Sciences, vol. 15, no.8342, pp. 1–17, 2025.

[17] Sánchez-García, A., Almeida, R., & Sousa, P. Integrating ISO 27001, NIST CSF 2.0, and Cyber Defense Matrix for Enhancing Organizational Cybersecurity Governance. SSRN Preprint, 2025.

# การปกป้องวิดีโอแบบรู้บริบทและปรับตัวได้สำหรับการสตรีมจากอากาศยานไร้คนขับ
# Content-Aware and Adaptive Lightweight Protection for UAV Video Streaming

*ฐิติวรดา ทับทิมไสย์ (Thitiworada Thabtimsai), สิลาภรณ์ จันทร์ถา (Silaporn Chuntha)*
*และ จักรชัย โสอินทร์ (Chakchai So-In)*
*วิทยาลัยการคอมพิวเตอร์ มหาวิทยาลัยขอนแก่น*
*thitiworada.t@kkumail.com, silaporn.c@kkumail.com, chakso@kku.ac.th*

## บทคัดย่อ

*อากาศยานไร้คนขับถูกนำมาใช้อย่างกว้างขวางในงานเฝ้าระวัง การตรวจสอบพื้นที่ และการส่งข้อมูลภาพแบบเรียลไทม์ อย่างไรก็ตาม การสตรีมวิดีโอจากอากาศยานไร้คนขับไปยังสถานีภาคพื้นดิน ผ่านเครือข่ายไร้สายยังคงเผชิญความเสี่ยงด้านความมั่นคงปลอดภัย โดยเฉพาะเมื่อข้อมูลภาพมีเนื้อหาสำคัญหรืออ่อนไหว แม้วิธีการเข้ารหัสทั้งเฟรมจะช่วยเพิ่มการปกป้องข้อมูลได้ แต่ก็มักเพิ่มภาระของระบบและส่งผลกระทบต่อประสิทธิภาพของการสตรีมภายใต้สภาพเครือข่ายที่ไม่คงที่ งานวิจัยนี้นำเสนอแนวทางต้นแบบสำหรับการส่งวิดีโอจากอากาศยานไร้คนขับแบบมั่นคงปลอดภัยเชิงปรับตัว โดยผสานการปกป้องเฉพาะบริเวณสำคัญของภาพ (Region of Interest: ROI) เข้ากับกลไก Adaptive Security Layer (ASL) เพื่อปรับระดับการปกป้องให้เหมาะสมกับสภาพเครือข่ายและคุณภาพการสตรีม พร้อมทั้งแยกการส่งข้อมูลพื้นหลังและข้อมูล ROI ที่เข้ารหัสออกเป็นคนละช่องทางสื่อสารเพื่อลดภาระการส่งข้อมูล ผลการทดลองภายใต้สถานการณ์เครือข่ายผันผวน (S3) แสดงให้เห็นว่าโหมด ROI+ASL (M2) มีจุดเด่นในการสร้างสมดุลระหว่างระดับการปกป้อง ภาระการส่งข้อมูล ความหน่วง และความต่อเนื่องของการสตรีม โดยสามารถลดภาระการส่งข้อมูลรวมลงได้ร้อยละ 36.9 เมื่อเทียบกับ M1 และ M3 ลดอัตราการสูญหายของแพ็กเก็ตเชิงสัมพัทธ์ลงร้อยละ 3.1 และ 0.4 ตามลำดับ และให้ค่าความหน่วงต่ำที่สุดที่ 14.53 มิลลิวินาที แม้ค่าอัตราเฟรมฝั่งรับจะต่ำกว่าโหมดอ้างอิงบางโหมดอยู่บ้าง แต่ภายใต้เกณฑ์การวิเคราะห์ที่กำหนด ระบบยังสามารถรักษาการแสดงผลเฟรมที่ฝั่งรับได้อย่างต่อเนื่อง ทั้งนี้ผลลัพธ์ดังกล่าวสะท้อนประสิทธิภาพของระบบสตรีมต้นแบบเป็นหลัก โดยในระยะปัจจุบัน การคุ้มครองมุ่งเน้นข้อมูล ROI และข้อมูลกำกับที่เกี่ยวข้อง ขณะที่ข้อมูลพื้นหลังยังไม่ได้รับการป้องกันในระดับเดียวกัน*

***คำสำคัญ**: ความมั่นคงปลอดภัยของอากาศยานไร้คนขับ, การเข้ารหัสบริเวณสำคัญของภาพ, ชั้นความมั่นคงปลอดภัยเชิงปรับตัว, ChaCha20-Poly1305, การสตรีมวิดีโออย่างปลอดภัย*

## Abstract

*Unmanned Aerial Vehicles (UAVs) are widely used in surveillance, area inspection, and real-time visual data transmission. However, video streaming from a UAV to a Ground Station (GS) over wireless networks remains vulnerable to security threats, particularly when the transmitted content contains sensitive or mission-critical visual information. Although full-frame encryption can strengthen data protection, it often introduces additional computational and communication overhead, which may degrade streaming performance under unstable network conditions. This paper presents a prototype architecture for adaptive and secure UAV video streaming that integrates Region of Interest (ROI)-based content-aware protection with an Adaptive Security Layer (ASL). The proposed system dynamically adjusts the protection level, ROI size, and ROI transmission behavior according to network*

*conditions and streaming quality. In addition, background data and encrypted ROI data are transmitted through separate communication channels to reduce transmission overhead while preserving receiver-side frame-display continuity. Experimental results under the dynamic network scenario (S3) show that the proposed ROI+ASL mode (M2) provides the most favorable balance among the evaluated modes in terms of transmission load, delay, packet drop, and receiver-side continuity. Relative to the full-frame protection mode (M1) and the fixed-ROI mode (M3), M2 reduces total transmission load by 36.9% in both comparisons, reduces the packet-drop ratio by 3.1% and 0.4%, respectively, and achieves the lowest delay of 14.53 ms. Although the received frame rate remains lower than that of some reference modes, the system still maintains receiver-side frame display within the reported evaluation setting. These findings primarily reflect the performance characteristics of the proposed prototype streaming system rather than a complete security validation. At the current stage, protection is applied mainly to ROI data and related metadata, whereas background content is not protected at the same level.*

**Keywords:** UAV Security, ROI Encryption, Adaptive Security Layer, ChaCha20-Poly1305, Secure Video Streaming

## 1. บทนำ

ปัจจุบันอากาศยานไร้คนขับ (Unmanned Aerial Vehicles: UAVs) ถูกนำมาใช้อย่างแพร่หลายในงานเฝ้าระวัง การตรวจสอบพื้นที่ การค้นหาและกู้ภัย ตลอดจนภารกิจที่ต้องอาศัยการส่งข้อมูลภาพแบบเรียลไทม์ ความสามารถในการส่งวิดีโอจาก UAV ไปยังสถานีภาคพื้นดิน (Ground Station: GS) จึงเป็นองค์ประกอบสำคัญของระบบรับรู้สถานการณ์ในหลายบริบทการใช้งาน อย่างไรก็ตาม การสื่อสารผ่านเครือข่ายไร้สายยังคงเผชิญความเสี่ยงด้านความมั่นคงปลอดภัย โดยเฉพาะเมื่อข้อมูลภาพมีเนื้อหาสำคัญหรืออ่อนไหว เช่น บุคคล วัตถุเป้าหมาย หรือพื้นที่ปฏิบัติการ ขณะเดียวกัน ระบบสตรีมวิดีโอแบบเรียลไทม์สำหรับ UAV ยังต้องคำนึงถึงความหน่วง ปริมาณข้อมูลที่รับส่ง และความต่อเนื่องของการสื่อสารควบคู่กับความมั่นคงปลอดภัยของระบบ [1], [2]

แม้วิธีการเข้ารหัสข้อมูลวิดีโอทั้งเฟรมหรือทั้งสตรีมจะช่วยยกระดับการปกป้องข้อมูลได้ แต่แนวทางดังกล่าวมักเพิ่มภาระของระบบ และอาจไม่เหมาะกับแพลตฟอร์ม UAV ที่ต้องทำงานแบบเรียลไทม์ภายใต้เครือข่ายที่ไม่คงที่ งานวิจัยที่ผ่านมาได้สะท้อนให้เห็นถึงความสำคัญของกลไกความมั่นคงปลอดภัยแบบปรับเปลี่ยนได้ (adaptive security) สำหรับระบบที่มีข้อจำกัดด้านทรัพยากร โดยเฉพาะเมื่อจำเป็นต้องรักษาสมดุลระหว่างระดับการปกป้องกับคุณภาพของเครือข่าย [1], [3]

จากงานวิจัยที่มีอยู่ พบว่าแนวทางด้านการสื่อสารมัลติมีเดียแบบปลอดภัยสำหรับ UAV ส่วนหนึ่งมุ่งเน้นการเข้ารหัสแบบ lightweight โดยใช้อัลกอริทึมมาตรฐาน เช่น AES เพื่อรองรับการทำงานบนแพลตฟอร์มที่มีข้อจำกัดด้านทรัพยากร [4], [5] ขณะที่อีกส่วนหนึ่งมุ่งเน้นการปกป้องเฉพาะบริเวณสำคัญของภาพ หรือ Region of Interest (ROI) เพื่อลดปริมาณข้อมูลที่ต้องเข้ารหัสและลด overhead ของระบบ [6], [7], [8] อย่างไรก็ตาม แนวทางทั้งสองมักถูกพัฒนาแยกจากกัน กล่าวคือ งานด้าน ROI-based encryption ส่วนใหญ่มักใช้ ROI แบบคงที่และไม่ได้ปรับเปลี่ยนตามสภาวะเครือข่าย ขณะที่งานด้าน adaptive security มักทำงานในระดับระบบหรือโปรโตคอลโดยไม่ได้คำนึงถึงความสำคัญของเนื้อหาภายในภาพอย่างชัดเจน

ด้วยเหตุนี้ งานวิจัยนี้จึงนำเสนอแนวทางต้นแบบสำหรับการส่งวิดีโอ UAV แบบมั่นคงปลอดภัยเชิงปรับตัว โดยผสานการปกป้องเชิงเนื้อหาแบบอิง ROI เข้ากับกลไก Adaptive Security Layer (ASL) ภายใต้สถาปัตยกรรมแบบวงรอบปิด (closed-loop) ซึ่งสามารถปรับทั้งขนาดของ ROI ความถี่ในการส่งข้อมูล และระดับการปกป้องให้สอดคล้องกับสภาวะเครือข่ายและคุณภาพการสตรีมแบบเวลาจริง แนวทางดังกล่าวแตกต่างจากงานวิจัยเดิมตรงที่ไม่ได้พิจารณาเฉพาะการเลือกปกป้องบริเวณสำคัญของภาพหรือการปรับระดับความมั่นคงปลอดภัยแยกจากกัน แต่บูรณาการทั้งสองส่วนเข้าด้วยกันภายใต้ระบบเดียว

พร้อมใช้การสื่อสารแบบแยกช่องทางเพื่อช่วยลดภาระการส่งข้อมูลและคงความต่อเนื่องของการสตรีม ทั้งนี้ ในระยะปัจจุบัน การคุ้มครองมุ่งเน้นข้อมูลในส่วน ROI และข้อมูลกำกับเป็นหลัก ขณะที่ข้อมูลพื้นหลังยังไม่ได้รับการป้องกันในระดับเดียวกัน ผลงานหลักของบทความนี้สามารถสรุปได้ดังนี้คือ

1. นำเสนอกลไกการปรับ ROI แบบพลวัต ซึ่งสามารถปรับขนาดและความถี่ในการอัปเดต ROI ตามสภาวะเครือข่ายและตัวชี้วัดของระบบแบบเวลาจริง แตกต่างจากแนวทาง ROI แบบคงที่ในงานวิจัยก่อนหน้า
2. ออกแบบกลไกการปกป้องแบบรู้บริบท โดยผสานการเลือก ROI เข้ากับ Adaptive Security Layer (ASL) เพื่อให้สามารถปรับระดับการปกป้องตามความสำคัญของเนื้อหาภาพและสภาวะของระบบ
3. พัฒนาสถาปัตยกรรมแบบวงรอบปิดที่ใช้ข้อมูลย้อนกลับจากตัวชี้วัดของการสตรีม เช่น ความหน่วง อัตราเฟรมฝั่งรับ และการสูญหายของแพ็กเก็ต เพื่อปรับพฤติกรรมของระบบอย่างต่อเนื่อง
4. นำเสนอการสื่อสารแบบแยกช่องทางระหว่างข้อมูลพื้นหลังและข้อมูล ROI ที่เข้ารหัส เพื่อช่วยลดภาระการส่งข้อมูลและรักษาความต่อเนื่องของการสตรีมแบบเวลาจริง

บทความนี้มีโครงสร้างดังต่อไปนี้ ส่วนที่ 2 นำเสนอผลงานวิจัยที่เกี่ยวข้อง ส่วนที่ 3 อธิบายวิธีการที่เสนอ ส่วนที่ 4 นำเสนอผลการทดลองและการประเมินประสิทธิภาพของระบบ และส่วนที่ 5 สรุปผล อภิปรายผล และเสนอแนวทางสำหรับงานวิจัยในอนาคต

## 2. งานวิจัยที่เกี่ยวข้อง

การออกแบบระบบรักษาความมั่นคงปลอดภัยสำหรับการส่งข้อมูลมัลติมีเดียในอากาศยานไร้คนขับ (UAV) จำเป็นต้องคำนึงถึงทั้งความมั่นคงปลอดภัย ประสิทธิภาพการสื่อสาร และข้อจำกัดด้านทรัพยากรของอุปกรณ์ งานวิจัยที่เกี่ยวข้องในประเด็นนี้สามารถแบ่งออกเป็น 3 กลุ่มหลัก ได้แก่ ระบบความปลอดภัยแบบปรับตัว การส่งวิดีโอและการเข้ารหัสทั้งเฟรม และการเข้ารหัสเฉพาะบริเวณสำคัญของภาพ ดังนี้

### 2.1 ระบบความปลอดภัยแบบปรับตัว

แนวทางนี้มุ่งปรับระดับความมั่นคงปลอดภัยให้เหมาะสมกับบริบทการทำงานและทรัพยากรของระบบ งานของ Jamil และคณะ [3] นำเสนอแนวคิด adaptive lightweight security สำหรับ UAV ที่ขับเคลื่อนด้วย TinyML โดยออกแบบให้ระบบสามารถเลือกกลไกความมั่นคงปลอดภัยตามสถานะของอุปกรณ์ ภาระงาน และเงื่อนไขการสื่อสารได้ จุดเด่นของแนวทางนี้คือความยืดหยุ่นในการใช้ทรัพยากร และความสามารถในการรักษาสมดุลระหว่างระดับการปกป้องกับข้อจำกัดของแพลตฟอร์ม UAV

อย่างไรก็ตาม งานวิจัยในกลุ่มนี้ยังมุ่งเน้นการปรับระดับความมั่นคงปลอดภัยในระดับระบบหรือระดับการสื่อสารเป็นหลัก และยังไม่ได้เชื่อมการตัดสินใจดังกล่าวเข้ากับความสำคัญของเนื้อหาภายในภาพโดยตรง กล่าวคือ แม้ระบบจะสามารถปรับตัวตามทรัพยากรหรือสภาวะเครือข่ายได้ แต่ยังไม่สามารถกำหนดได้อย่างชัดเจนว่าส่วนใดของภาพควรได้รับการปกป้องมากกว่าส่วนอื่น ช่องว่างดังกล่าวจึงชี้ให้เห็นความจำเป็นของแนวทางที่ผสานการปรับตัวของระบบเข้ากับการปกป้องข้อมูลเชิงเนื้อหาในระดับ ROI โดยตรง

### 2.2 การส่งวิดีโอและการเข้ารหัสทั้งเฟรม

การส่งวิดีโออย่างปลอดภัยเป็นองค์ประกอบสำคัญของภารกิจ UAV โดยเฉพาะในงานเฝ้าระวังและการรับรู้สถานการณ์แบบเรียลไทม์ งานของ Psilias และคณะ [1] นำเสนอระบบส่งวิดีโอที่มีความมั่นคงปลอดภัยสำหรับการใช้งาน UAV โดยเน้นการสื่อสารระหว่างอากาศยานและสถานีภาคพื้นดินแบบเรียลไทม์

ขณะที่งานของ Cecchinato และคณะ [4] และ Cecchinato และคณะ [5] เสนอแนวทางการเข้ารหัสมัลติมีเดียแบบ Lightweight โดยใช้อัลกอริทึม AES เพื่อคุ้มครองข้อมูลที่ส่งจาก UAV ภายใต้ข้อจำกัดของอุปกรณ์ฝังตัว

**ตารางที่ 1:** แสดงการเปรียบเทียบคุณสมบัติของงานวิจัยที่เกี่ยวข้อง

| Work | Main Technique | Content-Aware | Light weight | Adaptive | Real-time | Strengths / Limitations |
|---|---|---|---|---|---|---|
| **Jamil et al. (2025) [3]** | TinyML-driven Adaptive | ✗ | ✓ | ✓ | ✓ | ปรับตัวตามภาระงานได้ดี / ไม่ครอบคลุมระดับ Pixel |
| **Psilias et al. (2021) [1]** | Secure Transmission | ✗ | ✗ | ✗ | ✓ | มุ่งเน้นความมั่นคงปลอดภัยของการส่งวิดีโอ / ยังไม่เน้น ROI และ adaptive control |
| **Cecchinato et al. (2022) [4]** | AES Encryption | ✗ | ✓ | ✗ | ✓ | ใช้การเข้ารหัสมาตรฐานสำหรับงานเวลาจริง / ยังไม่เน้นการปรับระดับตามบริบท |
| **Cecchinato et al. (2023) [5]** | Lightweight AES | ✗ | ✓ | ✗ | ✓ | เสนอการเข้ารหัสแบบ lightweight / ยังไม่ผสาน ROI กับ adaptive control |
| **Abu Taha et al. (2018) [6]** | ROI-based HEVC | ✓ | ✓ | ✗ | ✓ | ปกป้องเฉพาะบริเวณสำคัญของภาพ / ยังไม่ปรับตัวตามสภาวะเครือข่าย |
| **Proposed** | ROI + ASL + ChaCha20 | ✓ | ✓ | ✓ | ✓ | บูรณาการการปรับตัวและรู้บริบทเข้าด้วยกัน |

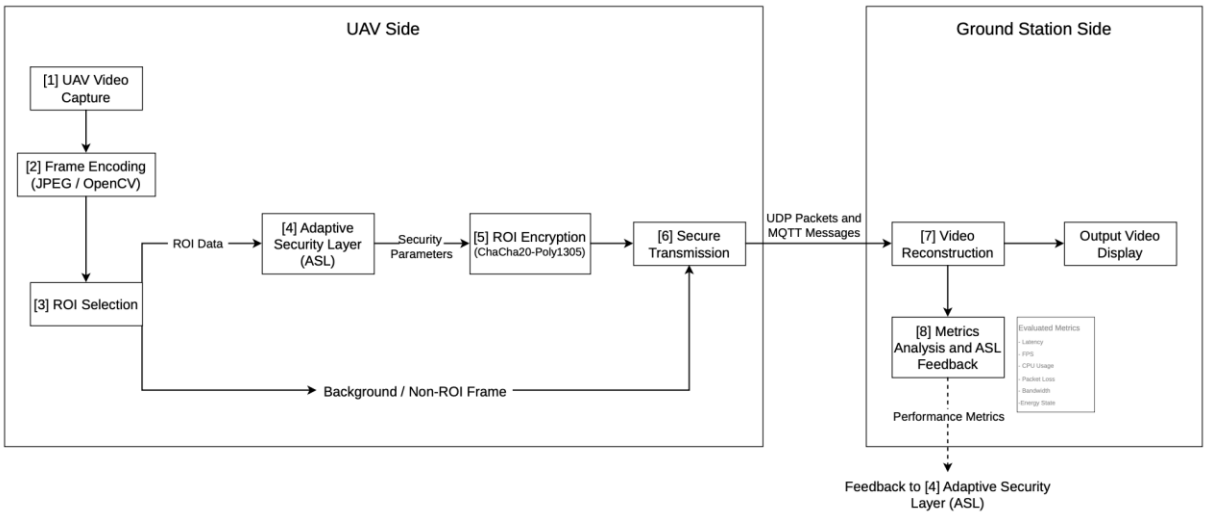
หมายเหตุ: ✓ หมายถึงมีคุณสมบัติดังกล่าว และ ✗ หมายถึงไม่มีคุณสมบัติดังกล่าว

จุดเด่นของงานในกลุ่มนี้คือการใช้กลไกเข้ารหัสมาตรฐานที่มีความน่าเชื่อถือสูง และสามารถประยุกต์ใช้กับการส่งข้อมูลมัลติมีเดียแบบเรียลไทม์ได้ อย่างไรก็ตาม แนวทางดังกล่าวมักคุ้มครองข้อมูลทั้งเฟรมหรือทั้งสตรีมในลักษณะเดียวกันทั้งหมด จึงอาจก่อให้เกิดภาระด้านการประมวลผลและการสื่อสารเกินความจำเป็น โดยเฉพาะเมื่อข้อมูลบางส่วนของภาพไม่ได้มีความสำคัญเท่ากัน

**2.3 การเข้ารหัสเฉพาะบริเวณสำคัญของภาพ**

อีกแนวทางหนึ่งที่ได้รับความสนใจคือการคุ้มครองเฉพาะบริเวณสำคัญของภาพ หรือ Region of Interest (ROI) ซึ่งมีเป้าหมายเพื่อลดปริมาณข้อมูลที่ต้องได้รับการคุ้มครองแบบเข้มข้น งานของ Abu Taha และคณะ [6] นำเสนอการเข้ารหัสแบบ ROI สำหรับวิดีโอ HEVC แบบเรียลไทม์ โดยคุ้มครองเฉพาะบริเวณสำคัญของภาพ ขณะที่ส่วนพื้นหลังยังคงสามารถถอดรหัสและแสดงผลได้ แนวทางนี้ช่วยลดต้นทุนการประมวลผลและลดผลกระทบต่ออัตราบิตของวิดีโอได้อย่างมีนัยสำคัญ

แม้งานในกลุ่มนี้จะแสดงให้เห็นข้อดีของการคุ้มครองแบบเลือกเฉพาะส่วนได้อย่างชัดเจน แต่กลไกความมั่นคงปลอดภัยส่วนใหญ่มักทำงานในลักษณะคงที่ กล่าวคือ ระดับการคุ้มครองไม่ได้ปรับเปลี่ยนตามสภาวะเครือข่าย ภาระของระบบ หรือคุณภาพการสตรีมระหว่างการทำงานจริง ดังนั้น ดังนัน งานวิจัยกลุ่ม ROI-based encryption จึงเด่นในด้านการลด overhead ผ่านการปกป้องเฉพาะส่วนสำคัญของภาพ อย่างไรก็ตาม งานในกลุ่มนี้ยังขาดกลไก adaptive control ที่ใช้ข้อมูลย้อนกลับจากระบบเพื่อปรับพฤติกรรมของ ROI และการส่งข้อมูลให้สอดคล้องกับบริบทการสื่อสารที่เปลี่ยนแปลงอยู่ตลอดเวลา




**ภาพที่ 1:** สถาปัตยกรรมของระบบต้นแบบสำหรับการสตรีมวิดีโอ UAV อย่างมั่นคงปลอดภัยเชิงปรับตัว

เพื่อให้เห็นภาพรวมของคุณลักษณะและข้อจำกัดของงานวิจัยที่เกี่ยวข้องอย่างชัดเจน ตารางที่ 1 สรุปการเปรียบเทียบแนวทางสำคัญในมิติของการปกป้องเชิงเนื้อหา การเข้ารหัสแบบน้ำหนักเบา ความสามารถในการปรับตัว และการรองรับการทำงานแบบเรียลไทม์

จากตารางที่ 1 จะเห็นได้ว่างานวิจัยก่อนหน้ามีจุดเด่นแตกต่างกันในแต่ละด้าน โดยงานด้าน adaptive security สามารถปรับระดับความมั่นคงปลอดภัยตามบริบทของระบบได้ แต่ยังไม่คำนึงถึงความสำคัญของข้อมูลในระดับ ROI อย่างชัดเจน ขณะที่งานด้าน ROI-based protection สามารถลดภาระของการคุ้มครองข้อมูลผ่านการเลือกปกป้องเฉพาะส่วนสำคัญของภาพได้ดี แต่ส่วนใหญ่มักทำงานแบบคงที่และไม่ปรับตามสภาวะเครือข่าย

ดังนั้นงานที่นำเสนอในบทความนี้จึงมุ่งเชื่อมข้อดีของทั้งสองแนวทางเข้าด้วยกัน โดยผสานการปกป้องเชิงเนื้อหาแบบอิง ROI กลไก Adaptive Security Layer (ASL) และการสื่อสารแบบแยกช่องทางไว้ในสถาปัตยกรรมเดียว เพื่อสร้างสมดุลระหว่างระดับการปกป้อง ภาระการส่งข้อมูล และความต่อเนื่องของการสตรีมแบบเวลาจริง

## 3. วิธีการดำเนินงาน

### 3.1 ภาพรวมสถาปัตยกรรมของระบบ

โครงสร้างของระบบแบ่งออกเป็น 2 ส่วนหลัก ได้แก่ ฝั่งอากาศยานไร้คนขับ (UAV Side) และฝั่งสถานีภาคพื้นดิน (Ground Station Side) ดังแสดงในภาพที่ 1 จุดมุ่งหมายของสถาปัตยกรรมนี้คือการลดภาระการประมวลผลและการสื่อสารบน UAV ขณะเดียวกันยังคงรักษาความมั่นคงปลอดภัยของข้อมูลสำคัญและความต่อเนื่องของการสตรีมวิดีโอแบบเรียลไทม์

ระบบเริ่มจากการรับภาพจากกล้องที่ฝั่ง UAV จากนั้นทำการบีบอัดเฟรม เลือกบริเวณสำคัญของภาพ กำหนดระดับการป้องกันผ่าน Adaptive Security Layer (ASL) เข้ารหัสข้อมูล ROI ส่งข้อมูลผ่านช่องทางสื่อสารที่แยกกันและประกอบภาพกลับที่ฝั่ง Ground Station พร้อมทั้งวิเคราะห์ตัวชี้วัดประสิทธิภาพเพื่อนำไปใช้เป็นข้อมูลป้อนกลับในการปรับระดับ ASL ในรอบถัดไป

- **UAV Video Capture:** ทำหน้าที่รับภาพวิดีโอจากกล้องของ UAV อย่างต่อเนื่อง เพื่อใช้เป็นข้อมูลตั้งต้นของกระบวนการทั้งหมดในระบบ
- **Frame Encoding (JPEG / OpenCV):** หลังจากรับเฟรมภาพ ระบบจะบีบอัดข้อมูลด้วย JPEG ผ่าน

OpenCV เพื่อลดขนาดข้อมูลก่อนส่งผ่านเครือข่าย และช่วยให้เหมาะกับการสตรีมแบบเรียลไทม์มากขึ้น

- **ROI Selection:** ระบบกำหนดบริเวณสำคัญของภาพหรือ ROI เพื่อใช้เป็นส่วนที่ได้รับการปกป้องเป็นพิเศษ ในโหมด M2 ระบบใช้แนวทาง user-initialized ROI with motion-based tracking กล่าวคือ ผู้ใช้กำหนดกรอบเริ่มต้นของวัตถุเป้าหมาย และระบบติดตามตำแหน่งและขนาดของ ROI ต่อเนื่องระหว่างเฟรมด้วยข้อมูลการเคลื่อนไหว ขณะที่ในโหมด M3 ใช้ ROI แบบคงที่บริเวณกึ่งกลางภาพเพื่อใช้เป็น baseline เปรียบเทียบ
- **Adaptive Security Layer (ASL):** เมื่อได้ ROI แล้ว ระบบจะประเมินสภาวะการทำงานผ่านกลไก ASL เพื่อกำหนดระดับการป้องกันที่เหมาะสมกับทรัพยากรและสภาพเครือข่ายในขณะนั้น โดยพิจารณาจากตัวชี้วัดสำคัญ เช่น Packet drop, RX FPS, Delay, CPU load and Battery level.
- **ROI Encryption:** เมื่อ ASL กำหนดระดับการป้องกันแล้ว ระบบจะเข้ารหัสข้อมูล ROI ด้วย ChaCha20-Poly1305 ในรูปแบบ AEAD (Authenticated Encryption with Associated Data) เพื่อให้ได้ทั้งความลับและความถูกต้องของข้อมูลในขั้นตอนเดียว แนวทางนี้ช่วยลดภาระของระบบเมื่อเทียบกับการเข้ารหัสทั้งเฟรม
- **Secure Transmission:** ระบบแยกการส่งข้อมูลออกเป็น 2 ช่องทาง ได้แก่ UDP สำหรับส่งข้อมูลพื้นหลังเพื่อลดความหน่วง และ MQTT สำหรับส่งข้อมูล ROI ที่เข้ารหัสแล้วพร้อมข้อมูลกำกับที่เกี่ยวข้อง วิธีนี้ช่วยให้ข้อมูลสำคัญได้รับการคุ้มครองอย่างเหมาะสม ขณะที่ภาพรวมของฉากยังคงสามารถส่งได้อย่างต่อเนื่อง
- **Video Reconstruction:** เมื่อข้อมูลมาถึงฝั่ง Ground Station ระบบจะรับข้อมูลจากทั้งสองช่องทาง ถอดรหัส ROI และประกอบกลับร่วมกับส่วนพื้นหลังเพื่อสร้างเฟรมวิดีโอสำหรับการแสดงผล
- **Metrics Analysis and ASL Feedback:** หลังจากประกอบภาพแล้ว ระบบจะวิเคราะห์ตัวชี้วัดสำคัญ เช่น ความหน่วง, อัตราเฟรมฝั่งรับ, อัตราการสูญหายของแพ็กเก็ต และภาระการส่งข้อมูล จากนั้นจึงส่งข้อมูลดังกล่าวกลับไปยัง ASL เพื่อปรับระดับการป้องกันในรอบถัดไป ทำให้ระบบมีลักษณะเป็น adaptive closed-loop

**3.2 แบบจำลองภัยคุกคามและขอบเขตความมั่นคงปลอดภัย**

งานวิจัยนี้กำหนดแบบจำลองผู้โจมตีให้อยู่ในระดับเครือข่ายระหว่าง UAV และ Ground Station (GS) โดยสมมติว่าผู้โจมตีไม่สามารถควบคุมตรรกะภายในของอุปกรณ์ปลายทางได้โดยตรง แต่สามารถสังเกตและแทรกแซงข้อมูลระหว่างการสื่อสารได้ ผู้โจมตีในแบบจำลองนี้อาจดักฟังข้อมูลที่ส่งผ่านเครือข่าย แก้ไขข้อมูลกำกับหรือ payload ระหว่างทาง ปลอมแปลงแพ็กเก็ตเพื่อแทรกเข้าสู่ระบบ ส่งข้อมูลซ้ำของเฟรมก่อนหน้า จัดลำดับการมาถึงของแพ็กเก็ตใหม่ หรือทำให้แพ็กเก็ตบางส่วนสูญหายได้

เนื่องจากระบบที่นำเสนอแยกการส่งข้อมูลพื้นหลังและข้อมูล ROI ออกเป็นคนละช่องทาง ภัยคุกคามที่พิจารณาจึงรวมถึงกรณีที่ข้อมูลจากสองช่องทางมาถึงไม่สอดคล้องกันด้วย ตัวอย่างเช่น ข้อมูลพื้นหลังของเฟรมหนึ่งมาถึงก่อน แต่ ROI ของเฟรมเดียวกันมาถึงล่าช้า สูญหาย ถูกแก้ไข หรือถูกแทนที่ด้วย ROI ของอีกเฟรมหนึ่ง สถานการณ์ดังกล่าวอาจทำให้เกิดการประกอบภาพผิดตำแหน่ง การใช้ ROI ที่ล้าสมัย หรือความต่อเนื่องของการแสดงผลลดลง แม้ว่าข้อมูลบางส่วนยังคงรับส่งได้อยู่ก็ตาม

ภายใต้ระบบต้นแบบนี้ ข้อมูลพื้นหลังถูกจัดเป็นข้อมูลที่มีความสำคัญต่ำกว่าและถูกส่งโดยไม่เข้ารหัสในระดับเดียวกับ ROI ดังนั้น ระบบจึงไม่ได้มุ่งรับประกันความลับของข้อมูลพื้นหลังภายใต้แบบจำลองผู้โจมตีนี้ โดยการคุ้มครองหลักมุ่งเน้นความลับและความถูกต้องของข้อมูล ROI และข้อมูลกำกับที่ใช้ระบุตำแหน่ง ขนาด และบริบทของเฟรมสำหรับการประกอบภาพกลับ ภัยคุกคามด้านการดักฟังและการแก้ไขข้อมูลระหว่างทางถูกลดทอนด้วยการเข้ารหัสแบบ AEAD สำหรับข้อมูล ROI และข้อมูลกำกับที่เกี่ยวข้อง ขณะที่ภัยคุกคามด้าน replay, reordering และ stale ROI ถูกจัดการในระดับระบบผ่านการตรวจสอบ

session_id, epoch_id, frame_id และกฎการจับคู่ข้อมูลข้ามช่องทางก่อนการประกอบภาพกลับ

งานวิจัยนี้ไม่ครอบคลุมกรณีที่ผู้โจมตียึดครองอุปกรณ์ปลายทาง เข้าถึงกุญแจลับภายในระบบได้โดยตรง แทรกแซงการทำงานภายในของ UAV หรือ GS ยึดอุปกรณ์ทางกายภาพ หรือก่อกวนเครือข่ายอย่างรุนแรง เช่น การ jamming หรือ denial-of-service ในระดับสูง ดังนั้นแบบจำลองนี้จึงมุ่งเน้นการคุ้มครองข้อมูลสำคัญของภาพในระหว่างการส่งผ่านเครือข่าย และการรักษาสมดุลระหว่างความมั่นคงปลอดภัยกับประสิทธิภาพของการสตรีมแบบเวลาจริงเป็นหลัก

**3.3 การกำหนดบริเวณสำคัญของภาพ (Region of Interest: ROI)**

ในโหมด M2 การกำหนดบริเวณสำคัญของภาพ (Region of Interest: ROI) ใช้แนวทาง user-initialized ROI with motion-based tracking กล่าวคือ ผู้ใช้กำหนดกรอบเริ่มต้นของวัตถุเป้าหมายในเฟรมแรก และระบบจะติดตามตำแหน่งและขนาดของ ROI ต่อเนื่องในเฟรมถัดไปโดยอาศัยข้อมูลการเคลื่อนไหวและตำแหน่งอ้างอิงจากเฟรมก่อนหน้า จากนั้นจึงสร้างกรอบล้อมรอบวัตถุเป้าหมายในลักษณะ bounding box เพื่อใช้เป็น ROI ของเฟรมนั้น วิธีดังกล่าวไม่ได้ใช้ object detection แบบ deep learning และไม่ได้ใช้การเลือก ROI แบบ manual ในทุกเฟรม แนวทางนี้ช่วยให้ระบบปกป้องเฉพาะส่วนสำคัญของภาพ แทนการเข้ารหัสทั้งเฟรม

ข้อมูลภาพถูกแบ่งออกเป็น 2 ส่วน ได้แก่ ส่วน ROI และส่วนพื้นหลัง โดยสัดส่วนพื้นที่ของ ROI มีผลโดยตรงต่อปริมาณข้อมูลที่ต้องเข้ารหัสและภาระของระบบ สัดส่วนดังกล่าวสามารถคำนวณได้จากสมการที่ (1)

$$ROI_{ratio} = \frac{Area_{ROI}}{Area_{frame}} \quad (1)$$

โดยที่ $Area_{ROI}$ คือ พื้นที่ของบริเวณ ROI และในส่วนของ $Area_{frame}$ คือพื้นที่ของเฟรมวิดีโอทั้งหมด

ในระบบที่นำเสนอ ขนาดของ ROI จะไม่คงที่ แต่ถูกปรับตามระดับของ Adaptive Security Layer (ASL) เพื่อให้สอดคล้องกับสภาวะเครือข่ายและภาระของระบบ โดยกำหนดค่าสเกลของ ROI ตามระดับ ASL ดังสมการที่ (2)

$$scale(ASL) = \begin{cases} 0.84, & ASL = HIGH \\ 0.60, & ASL = MED \\ 0.44, & ASL = LOW \end{cases} \quad (2)$$

ในงานนี้ กำหนดให้ระดับ HIGH, MED และ LOW ของ ASL สอดคล้องกับตัวคูณขนาด ROI เท่ากับ 0.84, 0.60 และ 0.44 ตามลำดับ เพื่อให้ระบบสามารถปรับขนาดบริเวณที่ต้องปกป้องให้เหมาะสมกับสภาวะเครือข่ายและภาระของระบบในแต่ละช่วงเวลา

เมื่อกำหนดค่าสเกลแล้ว ระบบจะคำนวณขนาดของกรอบ ROI ใหม่จากกรอบวัตถุที่ติดตามได้ โดยให้ $base_w$ และ $base_h$ แทนความกว้างและความสูงของกรอบอ้างอิงเดิม ขนาด ROI ที่ใช้จริงคำนวณได้จากสมการที่ (3) และ (4)

$$full_w = \max(32, base_w \cdot scale) \quad (3)$$
$$full_h = \max(32, base_h \cdot scale) \quad (4)$$

จากนั้นระบบจะวางกรอบ ROI ใหม่รอบจุดศูนย์กลางของวัตถุเป้าหมายที่ติดตามได้ โดยถ้า $C_x$ และ $C_y$ คือจุดศูนย์กลางของกรอบอ้างอิง ตำแหน่งมุมซ้ายบนของ ROI ใหม่สามารถคำนวณได้จากสมการที่ (5) และ (6)

$$x_{ROI} = c_x - \frac{full_w}{2} \quad (5)$$
$$y_{ROI} = c_y - \frac{full_h}{2} \quad (6)$$

หลังจากนั้นค่าดังกล่าวจะถูกจำกัดให้อยู่ภายในขอบของเฟรมวิดีโอ เพื่อป้องกันไม่ให้กรอบ ROI เกินขอบภาพ สำหรับโหมดอ้างอิง M3 ระบบใช้ ROI แบบคงที่บริเวณกึ่งกลางภาพเพื่อใช้เป็น baseline เปรียบเทียบกับ M2 โดยกำหนดขนาดของ ROI คงที่จากสัดส่วนของเฟรมดังสมการที่ (7)

$$s = \min(f_w, f_h) \cdot 0.28 \quad (7)$$

เมื่อ $f_w$ และ $f_h$ คือความกว้างและความสูงของเฟรม จากนั้นทำการจำกัดขนาดให้อยู่ในช่วงที่เหมาะสมด้วยสมการที่ (8)

$$s = \text{clip}(s, 96, 220) \quad (8)$$

สำหรับโหมด M3 กำหนดให้ขนาด ROI อ้างอิงจากร้อยละ 28 ของด้านที่สั้นกว่าของเฟรม และจำกัดขนาดให้อยู่ในช่วง 96 ถึง 220 พิกเซล เพื่อให้ ROI มีขนาดคงที่และเหมาะสมกับการใช้เป็น baseline เปรียบเทียบ ทั้งนี้ ค่าดังกล่าวถูกตรึงตลอดการทดลองและไม่ได้ปรับเปลี่ยนตามระดับ ASL

และกำหนดตำแหน่งของ ROI ให้อยู่ที่กึ่งกลางของเฟรมดังสมการที่ (9) และ (10)

$$x = \frac{f_w - s}{2} \quad (9)$$

$$y = \frac{f_h - s}{2} \quad (10)$$

โดยที่ $base_w$ และ $base_h$ แทนความกว้างและความสูงของกรอบอ้างอิงจากตัวติดตามวัตถุ $C_x$ และ $C_y$ แทนพิกัดจุดศูนย์กลางของกรอบดังกล่าว, $f_w$ และ $f_h$ แทนความกว้างและความสูงของเฟรมวิดีโอ และ clip($x$, $a$, $b$) หมายถึงฟังก์ชันที่ใช้จำกัดค่า $x$ ให้อยู่ในช่วง [$a$, $b$] เพื่อป้องกันไม่ให้กรอบ ROI เกินขอบภาพ

แนวทางดังกล่าวทำให้ M2 และ M3 สามารถใช้แนวคิด ROI-based protection ได้เหมือนกัน แต่แตกต่างกันในด้านความสามารถในการปรับตัว กล่าวคือ M2 ใช้ ROI ที่เปลี่ยนแปลงตามระดับ ASL ขณะที่ M3 ใช้ ROI แบบคงที่ตลอดการทำงาน

**3.4 กลไกการเข้ารหัสข้อมูล**

หลังจากกำหนดบริเวณสำคัญของภาพ (Region of Interest: ROI) แล้ว ระบบจะคุ้มครองข้อมูลในส่วนดังกล่าวด้วยอัลกอริทึม ChaCha20-Poly1305 ซึ่งเป็นการเข้ารหัสแบบยืนยันความถูกต้องร่วมกับข้อมูลกำกับ ( Authenticated Encryption with Associated Data: AEAD) ที่รองรับทั้งการรักษาความลับและการตรวจสอบความถูกต้องของข้อมูลภายในกระบวนการเดียว แนวทางนี้เหมาะสำหรับระบบสตรีมวิดีโอแบบเวลาจริงที่มีข้อจำกัดด้านทรัพยากร เนื่องจากมีภาระการประมวลผลต่ำและทำงานได้รวดเร็ว ทั้งนี้ ระบบเข้ารหัสเฉพาะข้อมูลในส่วน ROI ขณะที่ข้อมูลพื้นหลังถูกส่งผ่านช่องทางแยกต่างหาก

ในการเข้ารหัส ระบบกำหนดให้ข้อมูลกำกับที่สำคัญ เช่น หมายเลขเฟรม ตำแหน่งและขนาดของ ROI และข้อมูลที่ใช้ในการประกอบภาพกลับ ทำหน้าที่เป็น Associated Data เพื่อให้ฝั่งรับสามารถตรวจสอบความถูกต้องและความสอดคล้องของข้อมูลก่อนนำไปใช้งาน

ในระบบต้นแบบนี้ UAV และ GS มีข้อมูลลับร่วมกันล่วงหน้า (pre-shared secret) ก่อนเริ่มการสื่อสาร โดยระบบสร้างกุญแจย่อยสำหรับแต่ละช่วงเวลา (epoch-specific key) จากข้อมูลลับร่วมร่วมกับตัวระบุของ epoch ผ่านฟังก์ชันแฮช SHA-256 แล้วใช้ผลลัพธ์ขนาด 32 ไบต์เป็นกุญแจสำหรับการเข้ารหัสใน epoch นั้น แนวทางนี้ถูกใช้ในฐานะกลไกอนุพันธ์กุญแจของระบบต้นแบบ เพื่อแยกบริบทของกุญแจระหว่างช่วงการทำงานต่าง ๆ และลดความเสี่ยงจากการใช้กุญแจเดิมต่อเนื่องตลอดการสื่อสาร โดยรายละเอียดของการแลกเปลี่ยนกุญแจและการนำไปใช้งานจริงอยู่นอกขอบเขตของงานวิจัยนี้

สำหรับการสร้างค่า nonce ระบบใช้แนวทาง deterministic nonce generation โดยอาศัยบริบทของ *session_id*, *epoch_id* และ *frame_id* และกำหนดไม่ให้ชุดตัวระบุดังกล่าวถูกใช้ซ้ำภายใต้กุญแจเดียวกัน เพื่อหลีกเลี่ยงการใช้ nonce ซ้ำ ซึ่งเป็นเงื่อนไขสำคัญของ ChaCha20-Poly1305 ขณะเดียวกัน ฝั่ง GS จะตรวจสอบความสอดคล้องของตัวระบุเหล่านี้ ข้อมูลกำกับที่เกี่ยวข้องและผลการตรวจสอบความถูกต้องของ AEAD ก่อนนำข้อมูลไปใช้ในการประกอบภาพกลับ กลไกดังกล่าวช่วยลดความเสี่ยงจากข้อมูลที่ถูกแก้ไข ข้อมูลที่ไม่สอดคล้องกับบริบทของเฟรม และการใช้ ROI ที่ล้าสมัยในระดับของระบบต้นแบบ

**3.5 กลไก Adaptive Security Layer**

Adaptive Security Layer (ASL) เป็นกลไกหลักที่ใช้ปรับระดับการปกป้องของระบบให้เหมาะสมกับสภาวะการทำงานจริง โดยในระบบที่พัฒนาขึ้น งานวิจัยนี้ใช้แนวทางการตัดสินใจแบบ rule-based score aggregation กล่าวคือ ระบบจะประเมินตัวชี้วัดสำคัญหลายรายการร่วมกัน แล้วสรุปเป็นคะแนนรวมเพื่อกำหนดระดับ ASL เป้าหมาย ตัวชี้วัดที่นำมาพิจารณาประกอบไปด้วย

**ตารางที่ 2**: เกณฑ์การให้คะแนนย่อยของตัวชี้วัดใน ASL

| ตัวชี้วัด | เงื่อนไข | คะแนนเพิ่ม |
|---|---|---|
| Packet drop | (> 0.08), (> 0.14), (> 0.22) | +1, +2, +3 |
| RX FPS | (< 26), (< 20), (< 14) | +1, +2, +3 |
| Delay (ms) | (> 140), (> 220) | +1, +2 |
| Concurrent abnormality | Packet drop (> 0.10) และ RX FPS (< 20) | +1 |
| | Delay (> 160) และ Packet drop (> 0.12) | +1 |
| | หากเข้าเงื่อนไขทั้งสองข้อพร้อมกัน | +2 |
| CPU load | (> 70%) | +1 |
| Battery level | (≤ 25%) | +1 |

อัตราการสูญหายของข้อมูล อัตราเฟรมฝั่งรับ ความหน่วงของเครือข่าย ภาวะผิดปกติร่วมของหลายตัวชี้วัด ภาระของหน่วยประมวลผล และระดับแบตเตอรี่ โดยแต่ละตัวชี้วัดสะท้อนข้อจำกัดของระบบในมิติที่แตกต่างกัน เช่น คุณภาพลิงก์ ความต่อเนื่องของการสตรีม ภาระการคำนวณ และความพร้อมของทรัพยากรบนอุปกรณ์

คะแนนรวมของระบบคำนวณได้จากสมการที่ (11)

$$S = S_d + S_f + S_l + S_c + S_{cpu} + S_{battery} \qquad (11)$$

โดยที่

$S_d$ คือคะแนนจากอัตราการสูญหายของข้อมูล (Packet drop)

$S_f$ คือคะแนนจากอัตราเฟรมฝั่งรับ (RX FPS)

$S_l$ คือคะแนนจากความหน่วงของเครือข่าย (Delay)

$S_c$ คือคะแนนเสริมเมื่อเกิดอาการผิดปกติหลายตัวพร้อมกัน (Concurrent abnormalities)

$S_{cpu}$ คือคะแนนจากภาระของหน่วยประมวลผล (CPU load)

$S_{battery}$ คือคะแนนจากระดับแบตเตอรี่ (Battery level)

สมการดังกล่าวสะท้อนว่า คะแนนรวมของระบบเกิดจากการสะสมคะแนนย่อยของตัวชี้วัดแต่ละรายการตามเกณฑ์ที่กำหนดไว้ในตารางที่ 2 โดยเมื่อค่าคะแนนรวมสูงขึ้น ย่อมหมายถึงตัวชี้วัดสำคัญของระบบมีแนวโน้มอยู่ในช่วงที่สะท้อนข้อจำกัดของเครือข่ายและทรัพยากรประมวลผลมากขึ้น ดังนั้น ระบบจึงจำเป็นต้องปรับลดระดับการปกป้องลง เพื่อคงความต่อเนื่องของการสตรีมและลดภาระการทำงานโดยรวมให้เหมาะสมกับสภาวะที่เกิดขึ้น

จากเกณฑ์ในตารางที่ 2 จะเห็นได้ว่า Packet drop, RX FPS และ Delay ทำหน้าที่เป็นตัวชี้วัดหลักในการสะท้อนคุณภาพการสตรีมและสภาวะของเครือข่าย โดย Packet drop และ RX FPS มีอิทธิพลต่อคะแนนรวมมากที่สุด ขณะที่ Delay มีบทบาทรองลงมา ส่วน Concurrent abnormality, CPU load และ Battery level ทำหน้าที่เป็นตัวชี้วัดเสริม เพื่อช่วยสะท้อนความรุนแรงของสภาวะระบบเมื่อมีอาการผิดปกติหลายด้านเกิดขึ้นพร้อมกัน

เมื่อระบบคำนวณคะแนนย่อยจากตัวชี้วัดแต่ละรายการตามเกณฑ์ในตารางที่ 2 แล้ว คะแนนดังกล่าวจะถูกรวมเป็นคะแนนรวมของ ASL ระบบจะกำหนดระดับ ASL เป้าหมายตามเกณฑ์ดังสมการที่ (12)

$$ASL_{target} = \begin{cases} LOW, & S \geq 11 \\ MED, & 4 \leq S < 11 \\ HIGH, & S < 4 \end{cases} \qquad (12)$$

จากสมการที่ (12) คะแนนรวมของ ASL เป็นผลรวมของคะแนนย่อยเชิงจำนวนเต็ม ดังนั้นค่าคะแนนรวมจึงถูกตีความเป็นช่วงคะแนนไม่ต่อเนื่อง โดยกำหนดให้ช่วงคะแนนต่ำสอดคล้องกับระดับ HIGH ขณะที่คะแนนที่สูงขึ้นสะท้อนข้อจำกัดของเครือข่ายและทรัพยากรที่มากขึ้น จึงนำไปสู่การปรับระดับลงเป็น MED หรือ LOW ตามลำดับ

เพื่อป้องกันไม่ให้ระบบเปลี่ยนระดับ Adaptive Security Layer (ASL) บ่อยเกินไปเมื่อค่าตัวชี้วัดผันผวนในระยะสั้น งานวิจัยนี้จึงใช้กลไก hysteresis โดยกำหนด

จำนวนรอบต่อเนื่องที่เงื่อนไขเดิมต้องเกิดขึ้นก่อนอนุญาตให้เปลี่ยนระดับ ทั้งนี้ เกณฑ์ดังกล่าวช่วยลดอาการ flapping และทำให้การปรับระดับมีเสถียรภาพมากขึ้นระหว่างการทำงานจริง กำหนดไว้ในตารางที่ 3

**ตารางที่ 3**: เกณฑ์ hysteresis สำหรับการเปลี่ยนระดับ ASL

| การเปลี่ยนระดับ ASL | จำนวนรอบต่อเนื่องที่ต้องเกิด |
|---|---|
| HIGH → MED | 2 |
| MED → LOW | 5 |
| LOW → MED | 3 |
| MED → HIGH | 6 |

---

**Algorithm 1: Decision Mechanism and Level Adjustment for ASL**

---

**Input:** current metrics (Packet drop, RX FPS, Delay, CPU load, Battery level), current ASL level, hysteresis counters
**Output:** updated ASL level and streaming parameters

1: Collect current metrics from the UAV - GS system
2: Compute partial scores ($S_d$ , $S_f$ , $S_l$ , $S_c$ , $S_{cpu}$ , $S_{battery}$) according to Table 2
3: Compute total $Score_{ASL}$ = $S_d$ + $S_f$ + $S_l$ + $S_c$ + $S_{cpu}$ + $S_{battery}$
4: Map $Score_{ASL}$ to the target ASL level using Eq. (12)
5: Apply hysteresis to avoid frequent level switching
6: Update hysteresis counters according to the current and target ASL levels
7: If the target level persists for the required number of consecutive rounds, update the ASL level
8: Set transmission parameters according to the updated ASL level, including ROI scale, cadence, and stale threshold
9: Repeat for the next control interval

---

จากตารางที่ 3 จะเห็นได้ว่า การลดระดับการปกป้องจาก HIGH ไปเป็น MED และจาก MED ไปเป็น LOW ใช้จำนวนรอบต่อเนื่องแตกต่างกันตามระดับความรุนแรงของสภาวะระบบ ขณะที่การฟื้นกลับไปสู่ระดับที่สูงขึ้นก็ใช้เกณฑ์อีกชุดหนึ่งเพื่อให้การตัดสินใจมีเสถียรภาพมากขึ้น นอกจากนี้ ระบบยังหน่วงเวลาเริ่มต้นก่อนเข้าสู่โหมด auto ASL เป็นเวลา 3 วินาที เพื่อหลีกเลี่ยงผลกระทบจากความไม่เสถียรในช่วงเริ่มต้นของการทำงาน

เพื่อให้ลำดับการตัดสินใจและการปรับระดับของ Adaptive Security Layer (ASL) มีความชัดเจนมากขึ้นในเชิงขั้นตอนวิธี และช่วยให้สามารถทำซ้ำแนวทางที่เสนอได้สะดวกขึ้น โดยที่งานวิจัยนี้สรุปลำดับการทำงานของกลไก ASL ในรูปของขั้นตอนวิธีดังแสดงใน Algorithm 1

### 3.6 โปรโตคอลการสื่อสารของระบบ

ระบบใช้การสื่อสารแบบสองช่องทางในเชิงตรรกะ เพื่อให้เหมาะกับลักษณะของข้อมูลแต่ละประเภท โดยแยกข้อมูลพื้นหลังและข้อมูล ROI ออกจากกันเพื่อลดความหน่วงและภาระของระบบ ช่องทางแรกคือ UDP สำหรับส่งข้อมูลพื้นหลังหรือ non-ROI ที่ไม่เข้ารหัส เพื่อคงความต่อเนื่องของการสตรีมแบบเรียลไทม์ ส่วนอีกช่องทางหนึ่งคือ MQTT สำหรับส่งข้อมูล ROI ที่เข้ารหัสแล้วร่วมกับข้อมูลกำกับที่จำเป็นต่อการประกอบภาพกลับที่ฝั่ง Ground Station เช่น หมายเลขเฟรม พิกัดของ ROI ขนาดของ ROI และตัวระบุรอบกุญแจ (key epoch)

โดยที่ในการทดลองนี้กำหนดให้ช่องทาง MQTT ใช้ QoS ระดับ 0 ตลอดการทดลอง เพื่อให้พฤติกรรมของการส่งข้อมูล ROI คงที่ระหว่างการเปรียบเทียบแต่ละโหมด อย่างไรก็ตาม ในบางสถานการณ์ทดลอง เช่น S1 แม้ข้อมูลทั้งสองประเภทจะถูกแยกกันในเชิงตรรกะ แต่ยังคงใช้ลิงก์ทางกายภาพร่วมกัน จึงอาจเกิดข้อจำกัดด้านแบนด์วิดท์ร่วมได้

เมื่อข้อมูลถึงฝั่ง Ground Station ระบบจะจับคู่ข้อมูลจากทั้งสองช่องทางโดยอาศัยหมายเลขเฟรมและข้อมูลกำกับที่แนบมากับ ROI จากนั้นจึงถอดรหัส ROI และวางกลับในตำแหน่งเดิมของเฟรมเพื่อสร้างภาพสำหรับการแสดงผล หากข้อมูล ROI มาถึงล่าช้า ไม่สอดคล้องกับพื้นหลังของเฟรมนั้น หรือไม่ผ่านการตรวจสอบความถูกต้อง ระบบจะไม่ใช้ข้อมูลดังกล่าวในการประกอบภาพ และจะคงการแสดงผลจากพื้นหลังหรือ ROI ล่าสุดที่ยังใช้งานได้แทน

ในการตรวจรับข้อมูล ฝั่ง GS จะยอมรับข้อมูล ROI ก็ต่อเมื่อ (1) การตรวจสอบ AEAD ผ่านสำเร็จ (2) ค่า *session_id* และ *epoch_id* สอดคล้องกับบริบทการสื่อสารปัจจุบัน (3) ค่า *frame_id* สอดคล้องกับเฟรมพื้นหลังที่กำลังรอการประกอบ และ (4) ข้อมูล ROI ชุดดังกล่าวยังไม่เคยถูกใช้งานมาก่อน หากไม่เป็นไปตามเงื่อนไขข้อใดข้อหนึ่ง ระบบจะปฏิเสธข้อมูล ROI ชุดนั้นและไม่ใช้ในการประกอบภาพกลับ

ขณะเดียวกัน ฝั่ง Ground Station จะวิเคราะห์ตัวชี้วัดของระบบและกำหนดระดับ ASL ที่เหมาะสม ก่อนส่งค่าดังกล่าวกลับไปยังฝั่ง UAV เพื่อใช้ปรับพฤติกรรมของระบบในรอบถัดไป เช่น ขนาด ROI คุณภาพการบีบอัด และความถี่ในการส่งข้อมูล ROI ทำให้สถาปัตยกรรมโดยรวมมีลักษณะเป็นระบบการสื่อสารแบบวงรอบปิดเชิงปรับตัว (Adaptive closed-loop communication system)

นอกจากนี้ ระบบยังใช้กลไกการส่ง ROI แบบขับเคลื่อนด้วยเหตุการณ์ (event-driven) เพื่อหลีกเลี่ยงการส่งข้อมูลสำคัญทุกเฟรมโดยไม่จำเป็น โดยจะประเมินความเปลี่ยนแปลงของตำแหน่งและขนาดของ ROI จากเฟรมล่าสุดที่ส่งไปก่อนหน้า หลักการนี้ช่วยลดการส่งข้อมูลซ้ำโดยไม่จำเป็น ขณะเดียวกันยังคงความสอดคล้องของ ROI กับวัตถุเป้าหมายระหว่างการติดตาม

ให้กำหนดการเปลี่ยนตำแหน่งของ ROI จากจุดศูนย์กลางเป็นดังสมการที่ (13)

$$center_shift = \sqrt{(c_x - c_x^{last})^2 + (c_y - c_y^{last})^2} \quad (13)$$

และกำหนดการเปลี่ยนแปลงเชิงสัดส่วนของขนาด ROI เป็นดังสมการที่ (14) และ (15)

$$w_\Delta = \frac{|\, w - w_{last} \,|}{\max(1, w_{last})} \quad (14)$$

$$h_\Delta = \frac{|\, h - h_{last} \,|}{\max(1, h_{last})} \quad (15)$$

ในโหมด M2 ระบบจะพิจารณาการส่ง ROI ตามกลไก cadence ก่อน โดยระดับ HIGH พิจารณาได้ทุก 1 เฟรม ส่วนระดับ MED และ LOW พิจารณาได้ทุก 2 เฟรม เมื่อถึงรอบพิจารณา ระบบจะส่ง ROI ใหม่เมื่อเข้าเงื่อนไขอย่างน้อยหนึ่งข้อ ได้แก่ (1) จำนวนเฟรมนับตั้งแต่การส่ง ROI ครั้งล่าสุดมากกว่าหรือเท่ากับค่า max_stale (2) ค่า center_shift มากกว่าหรือเท่ากับเกณฑ์ที่กำหนด หรือ (3) ค่า $w_\Delta$ หรือ $h_\Delta$ มากกว่าหรือเท่ากับเกณฑ์ที่กำหนดตามระดับ ASL ดังแสดงในตารางที่ 4

**ตารางที่ 4**: เกณฑ์การส่ง ROI แบบ event-driven ตามระดับ ASL

| ระดับ ASL | max_stale (frames) | center_shift (pixels) | ($w_\Delta$) | ($h_\Delta$) |
|---|---|---|---|---|
| HIGH | 2 | 7.0 | 0.09 | 0.09 |
| MED | 3 | 9.0 | 0.11 | 0.11 |
| LOW | 4 | 11.0 | 0.13 | 0.13 |

จากตารางที่ 4 จะเห็นได้ว่า ระดับ HIGH ใช้เกณฑ์ที่เข้มที่สุด กล่าวคือระบบจะส่ง ROI ใหม่ได้ไวกว่าเมื่อเกิดการเปลี่ยนแปลงของตำแหน่งหรือขนาดของวัตถุ ขณะที่ระดับ MED และ LOW จะผ่อนเกณฑ์ดังกล่าวลงตามลำดับ เพื่อช่วยลดความถี่ในการส่งข้อมูลและลดภาระของเครือข่าย ทั้งนี้ ในโหมด M2 ระบบจะส่ง ROI ใหม่เมื่อเข้าเงื่อนไขอย่างน้อยหนึ่งข้อ ได้แก่ ยังไม่มี ROI ก่อนหน้า จำนวนเฟรมตั้งแต่การส่ง ROI ครั้งล่าสุดมากกว่าหรือเท่ากับค่า *max_stale* หรือค่า *center_shift*, $w_\Delta$ และ $h_\Delta$ มากกว่าหรือเท่ากับเกณฑ์ที่กำหนดตามระดับ ASL

### 3.7 ตัวชี้วัดที่ใช้ประเมินจากข้อมูลบันทึกของระบบต้นแบบ

เพื่อประเมินประสิทธิภาพของระบบ งานวิจัยนี้กำหนดตัวชี้วัดหลัก 4 รายการ ได้แก่ ภาระการส่งข้อมูลรวม อัตราเฟรมฝั่งรับ อัตราการสูญหายของข้อมูล และความหน่วงของการส่งข้อมูล โดยนิยามตัวชี้วัดเหล่านี้จากค่าที่บันทึกไว้ในไฟล์ log ของระบบโดยตรง เพื่อนำไปใช้ในการสรุปผลและคำนวณสถิติในแต่ละรอบการทดลอง

Transmission Load ใช้วัดปริมาณข้อมูลรวมที่ส่งออกจากฝั่ง UAV และมีความสำคัญในสถานการณ์ที่ลิงก์มีข้อจำกัด เช่น S1 และ S3 กำหนดให้ $T_{load}$ ดังสมการที่ )16)

$$T_{load} = uav_tx_total_mbps \quad (16)$$

RX FPS ใช้วัดอัตราเฟรมที่ฝั่ง GS รับและแสดงผลได้จริง เพื่อสะท้อนความต่อเนื่องของการสตรีม กำหนดให้ $F_{rx}$ เป็นดังสมการที่ (17)

$$F_{rx} = rx_fps_ema \quad (17)$$

Packet Drop ใช้วัดสัดส่วนของข้อมูลที่สูญหายระหว่างการส่งจาก UAV ไปยัง GS และรายงานในรูปของอัตราส่วน (ratio) ระหว่าง 0 ถึง 1 กำหนดให้ $P_{drop}$ เป็นดังสมการที่ (18)

$$P_{drop} = drop_ema \quad (18)$$

Delay ใช้วัดค่าประมาณความหน่วงของการสื่อสารระหว่างฝั่ง UAV และ GS กำหนดให้ $D_{net}$ เป็นดังสมการที่ (19)

$$D_{net} = net_delay_ema_ms \quad (19)$$

ในงานวิจัยนี้ $D_{net}$ นิยามเป็นค่าประมาณความหน่วงเชิงสัมพัทธ์จากข้อมูลบันทึกของระบบ และใช้เป็นตัวชี้วัดเชิงเปรียบเทียบระหว่างโหมดและสถานการณ์ทดลอง ไม่ใช่ one-way latency ที่ได้จากการซิงโครไนซ์นาฬิกาของสองเครื่องโดยตรงขึ้น

งานวิจัยนี้กำหนดให้ $T_{load}$, $P_{drop}$ และ $D_{net}$ เป็นตัวชี้วัดที่ยิ่งต่ำยิ่งดี ขณะที่ $F_{rx}$ เป็นตัวชี้วัดที่ยิ่งสูงยิ่งดี ดังนั้น หากโหมดใดลดภาระการส่งข้อมูลได้ แต่ทำให้ $F_{rx}$ ลดลงหรือ $D_{net}$ เพิ่มขึ้น จะถือว่าโหมดดังกล่าวมีลักษณะเป็นความสัมพันธ์เชิงแลกเปลี่ยนระหว่างประสิทธิภาพการใช้เครือข่ายกับคุณภาพการสตรีม

นอกจากนี้ งานวิจัยนี้นิยาม $TX_{ROI}$ เป็นปริมาณข้อมูลของส่วน ROI ที่ถูกส่งจากฝั่ง UAV ไปยัง GS ต่อหน่วยเวลา เพื่อใช้สะท้อนภาระเฉพาะของช่องทางข้อมูลสำคัญในระบบ ทั้งนี้ ค่า $TX_{ROI}$ ถูกคำนวณจากข้อมูลบันทึกของช่องทาง ROI ในไฟล์ log ของระบบ และใช้เป็นตัวชี้วัดรองสำหรับช่วยตีความผลการเปรียบเทียบในบางสถานการณ์

งานวิจัยนี้อาศัยข้อมูลบันทึกจากระบบต้นแบบในการคำนวณตัวชี้วัดและสรุปผลการทดลอง แทนการใช้ชุดข้อมูลมาตรฐานภายนอก นอกจากตัวชี้วัดหลักดังกล่าว งานวิจัยนี้ยังใช้การใช้งานหน่วยประมวลผลของฝั่ง UAV และสถานะของ ASL เป็นตัวชี้วัดรองสำหรับช่วยตีความ overhead ของระบบในสภาวะปกติ และพฤติกรรมการปรับตัวของโหมด M2 ภายใต้เครือข่ายที่เปลี่ยนแปลง

## 4. การประเมินประสิทธิภาพของระบบ

### 4.1 สภาพแวดล้อมและการติดตั้งระบบ

ในการประเมินประสิทธิภาพของระบบ งานวิจัยนี้กำหนดสภาพแวดล้อมการทดลองให้ประกอบด้วยฝั่งอากาศยานไร้คนขับ (UAV) และฝั่งสถานีภาคพื้นดิน (Ground Station: GS) โดยฝั่ง UAV ทำงานบนระบบปฏิบัติการ macOS และใช้ Webots R2025a สำหรับจำลองการทำงานของอากาศยานและกล้องจับภาพวิดีโอแบบต่อเนื่อง ขณะที่ฝั่ง GS ทำงานบนระบบปฏิบัติการ Windows มีหน้าที่รับข้อมูลจาก UAV ประกอบรวมข้อมูลภาพพื้นหลัง (Background) และบริเวณสำคัญของภาพ (Region of Interest: ROI) เพื่อสร้างเฟรมสำหรับแสดงผล พร้อมทั้งบันทึกตัวชี้วัดประสิทธิภาพของระบบในรูปแบบไฟล์ CSV สำหรับการวิเคราะห์ภายหลัง

ระบบต้นแบบพัฒนาด้วย Python 3.14.2 และใช้ Webots R2025a, OpenCV 4.8.1.78, NumPy 2.4.2, paho-mqtt 2.1.0, cryptography 46.0.5 และ Eclipse Mosquitto สำหรับการจำลอง การประมวลผลภาพ และการสื่อสาร โดยใช้ MQTT สำหรับรับส่งข้อมูล ROI และ UDP สำหรับส่งข้อมูลพื้นหลังตามสถาปัตยกรรมแบบแยกช่องทาง การทดลองในแต่ละโหมดและแต่ละสถานการณ์ทำซ้ำ 3 รอบภายใต้เงื่อนไขเดียวกัน เพื่อให้ผลการเปรียบเทียบมีความสม่ำเสมอ

### 4.2 รูปแบบและสถานการณ์การทดลอง

เพื่อประเมินประสิทธิภาพของระบบ งานวิจัยนี้กำหนดการทดลองออกเป็น 4 โหมด ดังนี้

- **M0 (Plain)** เป็นโหมดอ้างอิงพื้นฐานที่ไม่มีการปกป้องข้อมูลในระดับแอปพลิเคชัน และไม่มีการใช้ Adaptive Security Layer (ASL) ใช้เป็น baseline สำหรับพิจารณาพฤติกรรมของระบบเมื่อไม่มีภาระจากกลไกความมั่นคงปลอดภัยเพิ่มเติม
- **M1 (Full-frame Protection)** เป็นโหมดที่ปกป้องข้อมูลวิดีโอทั้งเฟรมอย่างสม่ำเสมอ ใช้เป็นตัวแทน

ของแนวทางที่คุ้มครองข้อมูลทุกส่วนของภาพในระดับเดียวกัน

- **M2 (Proposed ROI+ASL)** เป็นโหมดที่เสนอในงานวิจัยนี้ โดยใช้ ROI-based protection ร่วมกับกลไก ASL เพื่อปรับระดับการปกป้องตามคุณภาพเครือข่ายและคุณภาพการสตรีมระหว่างการทำงานจริง
- **M3 (Fixed ROI)** เป็นโหมดเปรียบเทียบที่ใช้แนวคิด ROI เช่นเดียวกับ M2 แต่กำหนดพฤติกรรมของ ROI แบบคงที่ และไม่มีการปรับตัวผ่าน ASL

นอกจากการเปรียบเทียบระหว่างโหมดแล้ว งานวิจัยนี้ยังประเมินระบบภายใต้ 4 สถานการณ์เครือข่าย ดังนี้

- **S0 (Normal)** เป็นสถานการณ์เครือข่ายปกติที่ไม่มีข้อจำกัดเพิ่มเติม ใช้เป็นสภาวะอ้างอิงสำหรับพิจารณา overhead และพฤติกรรมพื้นฐานของแต่ละโหมด
- **S1 (Shared Bandwidth Bottleneck)** เป็นสถานการณ์ที่จำลองคอขวด ของแบนด์วิดท์บนลิงก์ร่วม โดยข้อมูลพื้นหลังและข้อมูล ROI แม้จะแยกกันในเชิงตรรกะ แต่ยังคงใช้ลิงก์ทางกายภาพและข้อจำกัดด้านแบนด์วิดท์ร่วมกัน
- **S2 (Persistent Lossy Link)** เป็นสถานการณ์ที่จำลองคุณภาพเครือข่ายที่เสื่อมลงอย่างต่อเนื่อง โดยกำหนดให้มีการสูญหายของแพ็กเก็ตในส่วนภาพพื้นหลังและ ROI ในระดับคงที่ ร่วมกับค่าความหน่วงพื้นฐานและ jitter คงที่
- **S3 (Dynamic Burst / Congestion)** เป็นสถานการณ์ที่จำลองความแออัดของเครือข่ายแบบเปลี่ยนแปลงเป็นช่วง โดยสลับระหว่างช่วงปกติ ช่วงความแออัด และช่วงฟื้นตัวภายในรอบการทดลองเดียว

การออกแบบทั้ง 4 โหมดและ 4 สถานการณ์ดังกล่าว ช่วยให้สามารถประเมินระบบได้ครอบคลุมทั้งสภาวะพื้นฐาน ภาวะข้อจำกัดด้านแบนด์วิดท์ สภาวะลิงก์ที่มีการสูญหายต่อเนื่อง และเครือข่ายที่เปลี่ยนแปลงแบบพลวัต ทั้งนี้ ในบทความฉบับนี้ สถานการณ์ S0-S3 ถูกใช้เพื่อแทนสภาวะเครือข่ายที่แตกต่างกันในเชิงพฤติกรรม และมุ่งเน้นการเปรียบเทียบเชิงสัมพัทธ์ระหว่างโหมดภายใต้เงื่อนไขเดียวกันเป็นหลัก อย่างไรก็ตาม รายละเอียดเชิงตัวเลขของพารามิเตอร์เครือข่ายในแต่ละสถานการณ์ยังไม่ได้รายงานไว้อย่างครบถ้วน ซึ่งเป็นข้อจำกัดของงานต้นแบบและควรถูกรายงานเพิ่มเติมในงานถัดไปเพื่อสนับสนุนการทำซ้ำการทดลองอย่างสมบูรณ์

**4.3 วิธีสรุปผลจากข้อมูลบันทึกของระบบ**

ในการสรุปผลการทดลอง งานวิจัยนี้ไม่ได้รวมข้อมูลดิบจากทุกช่วงเวลาของทุกรอบการทดลองเข้าด้วยกันโดยตรง แต่ใช้แนวทางสรุปผลในระดับรอบการทดลองก่อน กล่าวคือ สำหรับแต่ละรอบจะคำนวณค่าสถิติของตัวชี้วัดจากข้อมูลบันทึกภายในรอบนั้น แล้วจึงนำค่าสถิติที่ได้จากแต่ละรอบมาสรุปรวมอีกครั้งในระดับภาพรวม แนวทางดังกล่าวช่วยให้ผลลัพธ์สะท้อนพฤติกรรมของแต่ละรอบการทดลองได้อย่างเหมาะสม และลดผลกระทบจากความผันผวนที่อาจเกิดขึ้นภายในรอบการทดลองเดียว

ในแต่ละรอบการทดลอง จะคัดเลือกเฉพาะข้อมูลที่อยู่ในช่วงวิเคราะห์ที่กำหนด โดยทั่วไปใช้ข้อมูลที่มีค่า *baseline_valid* = 1 จากนั้นจึงคำนวณค่ากลางของตัวชี้วัดแต่ละตัวด้วย ค่ามัธยฐาน (median) ของรอบนั้น หากกำหนดให้ $x_{r,1}, x_{r,2}, \dots, x_{r,nr}$ แทนค่าของตัวชี้วัดใดตัวชี้วัดหนึ่งในรอบที่ r ค่ามัธยฐานของรอบดังกล่าวคำนวณได้ดังสมการที่ (20)

$$m_r = \text{median}(x_{r,1}, x_{r,2}, \dots, x_{r,n_r}) \quad (20)$$

เมื่อได้ค่ามัธยฐานของแต่ละรอบแล้ว จะนำมาสรุปรวมอีกครั้งด้วยค่ามัธยฐานข้ามรอบการทดลอง ดังสมการที่ (21)

$$M = \text{median}(m_1, m_2, \dots, m_R) \quad (21)$$

โดยที่ $r$ คือจำนวนรอบการทดลองที่ถูกนำมาใช้จริงในแต่ละการตั้งค่า

เพื่อแสดงความแปรปรวนของผลการทดลอง งานวิจัยนี้รายงานส่วนเบี่ยงเบนมาตรฐานของค่ามัธยฐานรายรอบดังสมการที่ (22)

$$SD = \sqrt{\frac{\sum_{r=1}^{R}(m_r - \bar{m})^2}{R-1}} \quad (22)$$

**ตารางที่ 5:** ผลการทดลองภายใต้สถานการณ์ S0 (Normal)

| Mode | TX total (Mbps) | RX FPS | Drop | Delay (ms) |
|---|---|---|---|---|
| **M0** | 2.98 ± 0.01 | **57.61 ± 1.49** | **0.002 ± 0.005** | 7.70 ± 0.34 |
| **M1** | 2.99 ± 0.10 | 56.39 ± 2.08 | 0.007 ± 0.009 | 7.58 ± 1.53 |
| **M2** | **2.82 ± 0.36** | 42.01 ± 5.14 | 0.069 ± 0.014 | 8.89 ± 1.58 |
| **M3** | 3.27 ± 0.00 | 44.59 ± 1.51 | 0.115 ± 0.025 | **4.90 ± 0.23** |

**ตารางที่ 6:** ผลการทดลองภายใต้สถานการณ์ S1 (Shared Bandwidth Bottleneck)

| Mode | TX total (Mbps) | RX FPS | Drop | Delay (ms) |
|---|---|---|---|---|
| **M0** | 1.40 ± 0.00 | 26.51 ± 0.31 | 0.537 ± 0.000 | **4.10 ± 0.58** |
| **M1** | 1.40 ± 0.00 | 27.00 ± 0.45 | **0.532 ± 0.004** | 5.01 ± 0.04 |
| **M2** | 1.40 ± 0.00 | **28.83 ± 0.04** | 0.561 ± 0.015 | 8.55 ± 1.09 |
| **M3** | 1.40 ± 0.00 | 23.84 ± 0.94 | 0.580 ± 0.007 | 6.56 ± 1.37 |

**ตารางที่ 7:** ผลการทดลองภายใต้สถานการณ์ S2 (Persistent Lossy Link)

| Mode | TX total (Mbps) | RX FPS | Drop | Delay (ms) |
|---|---|---|---|---|
| **M0** | 2.84 ± 0.01 | **55.77 ± 0.99** | **0.026 ± 0.006** | 8.75 ± 0.31 |
| **M1** | 2.86 ± 0.06 | 54.10 ± 1.44 | 0.029 ± 0.011 | **5.99 ± 2.06** |
| **M2** | **2.81 ± 0.44** | 46.10 ± 5.66 | 0.064 ± 0.016 | 10.78 ± 1.53 |
| **M3** | 3.13 ± 0.03 | 47.28 ± 0.54 | 0.087 ± 0.002 | 8.63 ± 1.40 |

**ตารางที่ 8:** ผลการทดลองภายใต้สถานการณ์ S3 (Dynamic Burst / Congestion)

| Mode | TX total (Mbps) | RX FPS | Drop | Delay (ms) |
|---|---|---|---|---|
| **M0** | 1.59 ± 0.03 | **30.53 ± 0.77** | 0.507 ± 0.004 | 16.25 ± 1.88 |
| **M1** | 1.60 ± 0.01 | 29.63 ± 0.90 | 0.517 ± 0.008 | 16.05 ± 2.03 |
| **M2** | **1.01 ± 0.01** | 26.33 ± 0.59 | **0.501 ± 0.000** | **14.53 ± 1.12** |
| **M3** | 1.60 ± 0.01 | 27.92 ± 0.63 | 0.503 ± 0.002 | 17.72 ± 2.89 |

โดยที่ $\bar{m}$ คือค่าเฉลี่ยของค่ามัธยฐานรายรอบ ดังสมการที่ (23)

$$\bar{m} = \frac{1}{R}\sum_{r=1}^{R} m_r \qquad (23)$$

ดังนั้น ค่าที่รายงานในตารางผลการทดลองจึงอยู่ในรูปแบบดังสมการที่ (24)

$$Final\,Metric = M \pm SD \qquad (24)$$

นอกจากนี้ ในการจัดกลุ่มข้อมูลสำหรับการวิเคราะห์งานวิจัยนี้ยึดค่าที่บันทึกจริงของระบบ ได้แก่ *scenario_id* และ *mode_id* เป็นหลัก แทนการอ้างอิงจากชื่อแฟ้ม เพื่อป้องกันความคลาดเคลื่อนจากกรณีที่ชื่อแฟ้มอาจไม่สอดคล้องกับสถานะการทดลองที่เกิดขึ้นจริง

**4.4 ตัวชี้วัดประสิทธิภาพของระบบ**

การประเมินผลในงานวิจัยนี้ใช้ตัวชี้วัดหลัก 4 รายการ ตามที่นิยามไว้ในหัวข้อ 3.7 ได้แก่ Transmission Load, RX FPS, Packet Drop และ Delay เพื่อใช้เปรียบเทียบภาระการสื่อสาร ความต่อเนื่องของการสตรีม และผลกระทบจากสภาวะเครือข่ายที่แตกต่างกันระหว่างแต่ละโหมดและแต่ละสถานการณ์ ภายใต้เงื่อนไขการทดลองเดียวกัน

นอกจากนี้ งานวิจัยนี้ยังติดตามตัวชี้วัดรอง ได้แก่ CPU usage, TX ROI และ ASL state เพื่อใช้ประกอบการตีความพฤติกรรมของระบบต้นแบบ อย่างไรก็ตาม ในบทความฉบับนี้ผลการวิเคราะห์เชิงหลักจะเน้นที่ตัวชี้วัด 4 รายการ ได้แก่ Transmission Load, RX FPS, Packet Drop และ Delay/Latency เป็นสำคัญ

สำหรับสถานการณ์ S3 มีการกำหนดเงื่อนไขคัดเลือกข้อมูลเพิ่มเติม โดยใช้เฉพาะข้อมูลที่มี *scenario_id* = 3 และ *baseline_valid* = 1 และในกรณีของโหมด M2 จะเริ่มวิเคราะห์หลัง ROI initialization เสร็จสมบูรณ์แล้ว จากนั้นจึงใช้ข้อมูลช่วงเวลา 54 วินาทีแรกที่ต่อเนื่องกัน ด้วยความยาวเท่ากันในทุกโหมด เพื่อลดผลกระทบจาก

phase mismatch ระหว่างรอบการทดลอง ทั้งนี้ การวิเคราะห์ผลของงานวิจัยนี้มุ่งเปรียบเทียบผลภายในสถานการณ์เดียวกันเป็นหลัก เพื่อหลีกเลี่ยงการตีความข้ามสถานการณ์ที่มีเงื่อนไขเครือข่ายต่างกัน

**4.5 ผลการทดลองและการวิเคราะห์ผล**

ผลการทดลองในส่วนนี้นำเสนอในรูปแบบตารางเพื่อใช้เปรียบเทียบประสิทธิภาพของระบบในแต่ละโหมด ได้แก่ M0, M1, M2 และ M3 ภายใต้สถานการณ์การทดลองที่กำหนด โดยการวิเคราะห์ในหัวข้อนี้มุ่งเน้นการประเมินเชิงประสิทธิภาพของระบบเป็นสำคัญ ได้แก่ ภาระการส่งข้อมูลรวม ความหน่วง อัตราเฟรมฝั่งรับ และความต่อเนื่องของการสตรีม ทั้งนี้ ผลการทดลองไม่ควรถูกตีความว่าเป็นการประเมินความมั่นคงปลอดภัยอย่างครบถ้วนในทุกมิติ แต่เป็นการสะท้อนความสามารถของกลไกที่นำเสนอในการสร้างสมดุลระหว่างระดับการปกป้องข้อมูลตามการออกแบบของระบบกับประสิทธิภาพการทำงานภายใต้ข้อจำกัดของเครือข่าย

- **ตารางที่ 5** แสดงให้เห็นว่า ภายใต้สถานการณ์ S0 โหมด M2 ให้ค่าภาระการส่งข้อมูลรวมต่ำที่สุด แต่แลกกับค่าอัตราเฟรมฝั่งรับที่ลดลงอย่างชัดเจนเมื่อเทียบกับ M0 และ M1 รวมถึงค่าอัตราการสูญหายของแพ็กเก็ตและความหน่วงที่สูงขึ้นบางส่วน ผลดังกล่าวสะท้อนว่าในสภาวะเครือข่ายปกติ แนวทาง ROI+ASL ยังมีต้นทุนเชิงการทำงานของระบบต้นแบบอยู่ และจุดเด่นของ M2 ในสถานการณ์นี้จึงอยู่ที่การลดภาระการส่งข้อมูล มากกว่าการให้คุณภาพการสตรีมดีที่สุดในทุกตัวชี้วัด
- **ตารางที่ 6** แสดงให้เห็นว่า ค่าภาระการส่งข้อมูลรวมของทุกโหมดใกล้เคียงกันมากตามข้อจำกัดของแบนด์วิดท์ร่วม โดย M2 ให้ค่าอัตราเฟรมฝั่งรับสูงที่สุด แต่มีค่าอัตราการสูญหายของแพ็กเก็ตสูงกว่า M0 และ M1 และมีค่าความหน่วงสูงกว่าโหมดอื่น ผลดังกล่าวชี้ว่า กลไก ASL ของ M2 สามารถช่วยคงอัตราเฟรมฝั่งรับได้ดีขึ้นภายใต้คอขวดของลิงก์ แต่ยังมีข้อแลกเปลี่ยนด้านความหน่วงและการสูญหายของข้อมูลที่ต้องพิจารณาร่วมกัน
- **ตารางที่ 7** แสดงให้เห็นว่า M2 สามารถลดค่าภาระการส่งข้อมูลรวมได้ต่ำที่สุด และลดปริมาณข้อมูลในส่วน ROI เมื่อเทียบกับ M3 ได้ชัดเจน อย่างไรก็ตาม ค่าอัตราเฟรมฝั่งรับของ M2 ยังต่ำกว่า M0 และ M1 และค่าความหน่วงสูงกว่าโหมดอ้างอิงบางโหมด จึงสะท้อนข้อแลกเปลี่ยนระหว่างการลดภาระการส่งข้อมูลกับคุณภาพการสตรีมภายใต้ลิงก์ที่เสื่อมคุณภาพอย่างต่อเนื่อง
- **ตารางที่ 8** แสดงให้เห็นว่า M2 สะท้อนพฤติกรรมเชิงปรับตัวของกลไก ASL ระหว่างการทำงานภายใต้สภาวะเครือข่าย ที่ผันผวน โดยให้ค่าภาระการส่งข้อมูลรวม อัตราการสูญหายของแพ็กเก็ต และความหน่วงต่ำที่สุดในบรรดาโหมดที่ประเมิน แม้ค่าอัตราเฟรมฝั่งรับจะยังต่ำกว่า M0 และ M1 อยู่บ้าง แต่ผลโดยรวมชี้ว่า M2 ให้สมดุลที่เหมาะสมที่สุดในสถานการณ์นี้

## 5. สรุป อภิปรายผล และงานในอนาคต

งานวิจัยนี้นำเสนอระบบต้นแบบสำหรับการปกป้องวิดีโอจากอากาศยานไร้คนขับแบบรู้บริบทและปรับตัวได้ โดยผสาน ROI-based protection เข้ากับ Adaptive Security Layer (ASL) ภายใต้สถาปัตยกรรมแบบวงรอบปิด ผลการทดลองแสดงให้เห็นว่าแนวทางที่เสนอมีศักยภาพในการลดภาระการส่งข้อมูล และภายใต้สถานการณ์เครือข่ายผันผวน (S3) โหมด ROI+ASL (M2) สามารถสร้างสมดุลระหว่างภาระการสื่อสาร ความหน่วง และความต่อเนื่องของการสตรีมได้เด่นที่สุดในบรรดาโหมดที่ประเมิน แม้ค่าอัตราเฟรมฝั่งรับจะต่ำกว่าโหมดอ้างอิงบางโหมดอยู่บ้าง แต่ระบบยังสามารถรักษาความต่อเนื่องของการแสดงผลที่ฝั่งรับได้ภายใต้เงื่อนไขการทดลองที่กำหนด

**ตารางที่ 9:** สรุปผลการเปรียบเทียบเชิงปริมาณของระบบ

| สถานการณ์ | M2 VS M1 | | M2 VS M3 | |
|---|---|---|---|---|
| | ดีกว่า | แย่กว่า | ดีกว่า | แย่กว่า |
| **S0** | TX total ลดลง **5.7%** | RX FPS ลดลง 25.5%; Delay เพิ่มขึ้น 1.31 ms (17.3%) | TX total ลดลง 13.8%; Drop ลดลง 40.0%; TX ROI ลดลง 60.6% | RX FPS ลดลง 5.8%; Delay เพิ่มขึ้น 3.99 ms (81.4%) |
| **S1** | RX FPS เพิ่มขึ้น 6.8% | Drop เพิ่มขึ้น 5.5%; Delay เพิ่มขึ้น 3.54 ms (70.7%) | RX FPS เพิ่มขึ้น 20.9%; Drop ลดลง 3.3% | Delay เพิ่มขึ้น 1.99 ms (30.3%) |
| **S2** | TX total ลดลง 1.7% | RX FPS ลดลง 14.8%; Delay เพิ่มขึ้น 4.79 ms (80.0%) | TX total ลดลง 10.2%; Drop ลดลง 26.4%; TX ROI ลดลง 58.3% | RX FPS ลดลง 2.5%; Delay เพิ่มขึ้น 2.15 ms (24.9%) |
| **S3** | TX total ลดลง 36.9%; Drop ลดลง 3.1%; Delay ลดลง 1.52 ms (9.5%) | RX FPS ลดลง 11.1% | TX total ลดลง 36.9%; Drop ลดลง 0.4%; Delay ลดลง 3.19 ms (18.0%); TX ROI ลดลง 51.9% | RX FPS ลดลง 5.7% |

**หมายเหตุ:** ค่าร้อยละในตารางนี้คำนวณจากความแตกต่างเชิงสัมพัทธ์เมื่อเทียบกับโหมดอ้างอิง ไม่ใช่ค่าความแตกต่างแบบเปอร์เซ็นต์พอยต์

ในบริบทของงานวิจัยนี้ คำว่า “ความต่อเนื่องของการสตรีม” หมายถึงความสามารถของระบบต้นแบบในการรักษาการแสดงผลเฟรมที่ฝั่งรับภายใต้เกณฑ์การวิเคราะห์ที่กำหนด ทั้งนี้ มิได้หมายความถึงการไม่มีการสูญหายของข้อมูล หรือการคงคุณภาพของภาพในทุกมิติ

อย่างไรก็ตาม ผลดังกล่าวควรถูกตีความในฐานะการประเมินประสิทธิภาพของระบบต้นแบบภายใต้สภาพแวดล้อมจำลอง มิใช่การยืนยันความมั่นคงปลอดภัยของระบบอย่างครบถ้วนในทุกมิติ ทั้งนี้ การคุ้มครองในปัจจุบันยังมุ่งเน้น ROI และข้อมูลกำกับที่เกี่ยวข้อง ขณะที่ข้อมูลพื้นหลังยังไม่ได้รับการป้องกันในระดับเดียวกัน

สำหรับงานในอนาคต ควรขยายระบบให้รองรับหลายวัตถุ ประเมินความแม่นยำของการกำหนด ROI อย่างเป็นระบบ และทดสอบบนเครือข่ายไร้สายจริง นอกจากนี้ ควรเพิ่มการประเมินด้านความมั่นคงปลอดภัยให้ครอบคลุมยิ่งขึ้น เช่น ความทนทานต่อ replay, reordering, stale ROI assembly และผลกระทบจากการที่ข้อมูลพื้นหลังยังไม่ได้เข้ารหัส

- **ตารางที่ 9** แสดงให้เห็นว่า M2 มิได้ให้ผลดีที่สุดในทุกตัวชี้วัดและทุกสถานการณ์ หากแต่มีจุดเด่นในฐานะแนวทางเชิงปรับตัวที่มุ่งสร้างสมดุลระหว่างการปกป้องข้อมูล ภาระการส่งข้อมูล ความหน่วง และความต่อเนื่องของการสตรีม โดยเฉพาะในสถานการณ์ S3 ซึ่งสามารถลดค่าภาระการส่งข้อมูลรวมได้ร้อยละ 36.9 เมื่อเปรียบเทียบกับทั้ง M1 และ M3 พร้อมทั้งลดค่าอัตราการสูญหายของแพ็กเก็ตและความหน่วงได้ในเวลาเดียวกัน ในทางกลับกัน สถานการณ์ S0 และ S2 แสดงให้เห็นข้อแลกเปลี่ยนของระบบอย่างชัดเจน กล่าวคือ แม้ระบบจะสามารถลดภาระการส่งข้อมูลได้ แต่ยังต้องแลกกับค่าอัตราเฟรมฝั่งรับหรือความหน่วงบางส่วน

## เอกสารอ้างอิง

[1] D. Psilias, A. Milidonis, and I. Voyiatzis, “Secure video transmission system for UAV applications,” in Proc.

*25th Pan-Hellenic Conf. on Informatics (PCI 2021)*, pp. 318-322, 2021, doi: 10.1145/3503823.3503882.

[2] H. Zhou, F. Hu, M. Juras, A. B. Mehta, and Y. Deng, "Real-time video streaming and control of cellular-connected UAV system: Prototype and performance evaluation," *IEEE Wireless Communications Letters*, Vol. 10, No. 8, pp. 1657-1661, 2021, doi: 10.1109/LWC.2021.3076415.

[3] N. Jamil, K. Nega, M. H. Rozaidi, and A. M. A. Al Ghaili, "Adaptive lightweight security for TinyML driven unmanned aerial vehicles in smart cities," in Proc. *3rd Int. Conf. on Cyber Resilience (ICCR)*, pp. 1-7, 2025, doi: 10.1109/ICCR67387.2025.11291809.

[4] N. Cecchinato, A. Toma, C. Drioli, G. Oliva, G. Sechi, and G. L. Foresti, "A secure real-time multimedia streaming through robust and lightweight AES encryption in UAV networks for operational scenarios in military domain," *Procedia Computer Science*, Vol. 205, pp. 50-57, 2022, doi: 10.1016/j.procs.2022.09.006.

[5] N. Cecchinato, A. Toma, C. Drioli, G. Oliva, G. Sechi, and G. L. Foresti, "Secure real-time multimedia data transmission from low-cost UAVs with a lightweight AES encryption," *IEEE Communications Magazine*, Vol. 61, No. 5, pp. 160-165, 2023, doi: 10.1109/MCOM.001.2200611.

[6] M. Abu Taha, N. Sidaty, W. Hamidouche, O. Déforges, J. Vanne, and M. Viitanen, "End-to-end real-time ROI-based encryption in HEVC videos," in Proc. *26th European Signal Processing Conf. (EUSIPCO)*, pp. 171-175, 2018, doi: 10.23919/EUSIPCO.2018.8553038.

[7] M. Farajallah, W. Hamidouche, O. Déforges, and S. El Assad, "ROI encryption for the HEVC coded video contents," in *Proc. IEEE Int. Conf. on Image Processing (ICIP)*, pp. 3096-3100, 2015, doi: 10.1109/ICIP.2015.7351373.

[8] H. Sohn, E. T. Anzaku, W. De Neve, Y. M. Ro, and K. N. Plataniotis, "Privacy protection in video surveillance systems using scalable video coding," in Proc. *6th IEEE Int. Conf. on Advanced Video and Signal Based Surveillance (AVSS)*, pp. 424-429, 2009, doi: 10.1109/AVSS.2009.48.

# การประยุกต์ใช้แพลตฟอร์ม Microsoft Power Apps เพื่อจัดการและถ่ายทอดองค์ความรู้ของที่ปรึกษาทางด้านเทคนิคให้กับฝ่ายขาย กรณีศึกษาบริษัทโทรคมนาคม เขตภาคเหนือตอนล่าง

# Application of Microsoft Power Apps Platform for Managing and Transferring Technical Consultants' Knowledge to Sales Teams: A Case Study of a Telecommunication Company in the Lower Northern Region

*อัญธิกา วัฒนกุลวิวัฒน์ (Anthika Wattanakulwiwat)[1] และสุนันทา สดสี (Sunantha Sodsee)[2]*
*ภาควิชาการบริหารเครือข่ายดิจิทัลและความมั่นคงปลอดภัยสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล*
*มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ*
*s6707031857042@kmutnb.ac.th,sunantha.s@itd.kmutnb.ac.th*

## บทคัดย่อ

*การวิจัยนี้มีวัตถุประสงค์เพื่อศึกษาปัญหาและข้อจำกัดในการประสานงานระหว่างที่ปรึกษาทางเทคนิคและฝ่ายขายในองค์กรธุรกิจโทรคมนาคมในภาคเหนือตอนล่าง และพัฒนาระบบสนับสนุนการนำเสนอโซลูชันด้วยแอปพลิเคชันบน Microsoft Power Apps โดยใช้ระเบียบวิธีวิจัยและพัฒนา แบ่งเป็น 4 ระยะ ได้แก่ (1)การวิเคราะห์ปัญหา (2)การออกแบบและพัฒนา (3)การทดสอบและปรับปรุง และ(4)การประเมินผล กรอบแนวคิดที่ใช้ในการวิจัยคือ SECI Model โดยมุ่งเน้นกระบวนการ Externalization เพื่อแปลงความรู้ฝังลึกของที่ปรึกษาทางเทคนิคให้เป็นความรู้ชัดแจ้งในรูปแบบดิจิทัล และจัดทำเป็นฐานความรู้กลางขององค์กร ผลการวิจัยพบว่า ปัญหาหลักเกิดจากการขาดแคลนบุคลากรด้านเทคนิค ส่งผลให้ฝ่ายขายไม่สามารถนำเสนอข้อมูลเชิงเทคนิคได้อย่างถูกต้องและทันเวลา ระบบที่พัฒนาขึ้นได้รับการประเมินจากกลุ่มตัวอย่าง 13 คน ประกอบด้วยพนักงาน (n = 8) และลูกค้าองค์กร (n = 5) โดยมีระดับความพึงพอใจเฉลี่ย 4.08 และ 4.06 ตามลำดับ จากคะแนนเต็ม 5.00 ข้อค้นพบชี้ให้เห็นว่าระบบสามารถลดการพึ่งพาที่ปรึกษาทางเทคนิคในกระบวนการขาย และสนับสนุนการจัดการความรู้ภายในองค์กรได้อย่างมีประสิทธิภาพ ทั้งนี้ สามารถต่อยอดโดยบูรณาการกับ Microsoft Power Automate และ Microsoft Power BI เพื่อเพิ่มขีดความสามารถด้านการทำงานอัตโนมัติและการวิเคราะห์เชิงกลยุทธ์ในอนาคต*

**คำสำคัญ:** การจัดการองค์ความรู้ Microsoft Power Apps ฝ่ายขาย ที่ปรึกษาทางด้านเทคนิค ธุรกิจโทรคมนาคม

## Abstract

*This research aims to investigate the problems and limitations in coordination between technical consultants and sales teams in a telecommunications organization in the lower northern region, and develop a solution presentation support system using a Microsoft Power Apps application. The study adopts a Research and Development (R&D) methodology consisting of four phases: (1)problem analysis, (2)design and development, (3)testing and refinement, and (4)evaluation. The conceptual framework is based on the SECI Model, with a focus on the externalization process to transform tacit knowledge of technical consultants into explicit digital knowledge, forming a centralized organizational knowledge base. The results indicate that the primary issue is the shortage of technical personnel, which prevents the sales team from delivering accurate and*

*timely technical solutions. The developed system was evaluated by 13 participants, including employees (n = 8) and corporate clients (n = 5). The average satisfaction scores were 4.08 and 4.06 out of 5.00, respectively. The findings demonstrate that the system effectively reduces reliance on technical consultants in the sales process and enhances organizational knowledge management. Furthermore, the system can be extended by integrating with Microsoft Power Automate and Microsoft Power BI to support workflow automation and strategic data analytics in the future.*

**Keywords:** Knowledge Management, Microsoft Power Apps, Sales Team, Technical Consultant, Telecommunications Business

## 1. บทนำ

ในยุคปัจจุบัน การเปลี่ยนแปลงทางเทคโนโลยีดิจิทัลเกิดขึ้นอย่างรวดเร็วและต่อเนื่อง ส่งผลให้รูปแบบการดำเนินธุรกิจในอุตสาหกรรมโทรคมนาคมต้องปรับตัวอย่างมีนัยสำคัญ จากรายงานการสำรวจการมีการใช้เทคโนโลยีสารสนเทศและการสื่อสารในสถานประกอบการ พ.ศ. 2566 ของสำนักงานสถิติแห่งชาติพบว่าสถานประกอบการขนาดกลางและขนาดใหญ่ของไทยมีสัดส่วนการใช้งานอินเทอร์เน็ตความเร็วสูงและระบบคลาวด์เพิ่มขึ้นอย่างต่อเนื่อง [1]

จากการสำรวจหน่วยงานเอกชนแห่งหนึ่งในจังหวัดนครสวรรค์ ซึ่งรับผิดชอบพื้นที่ ภาคเหนือตอนล่าง 10 จังหวัดได้แก่ ตาก พิษณุโลก สุโขทัย เพชรบูรณ์ พิจิตร กำแพงเพชร นครสวรรค์ อุทัยธานี ชัยนาท และอุตรดิตถ์ บริบทของพื้นที่นี้มีลักษณะเฉพาะที่เป็นความท้าทายในการดำเนินงานกล่าวคือลูกค้าองค์กรกระจายตัวอยู่ในวงกว้าง โดยมีระยะทางระหว่างสำนักงานใหญ่กับลูกค้าบางรายสูงกว่า 200 กิโลเมตรทำให้การเดินทางของที่ปรึกษาทางด้านเทคนิคเพื่อสนับสนุนการขายในแต่ละครั้งต้องใช้เวลาและทรัพยากรสูง พบว่ามีที่ปรึกษาทางด้านเทคนิคเพียง 2 คน ต่อฝ่ายขาย 6 คน คิดเป็นอัตราส่วน 1:3 ส่งผลให้ฝ่ายขายไม่สามารถอธิบายในเชิงเทคนิคหรือระบุโซลูชันที่เหมาะสมได้อย่างแม่นยำ ปัญหาดังกล่าวสอดคล้องกับแนวคิดการจัดการองค์ความรู้ (Knowledge Management) ของ Nonaka และ Takeuchi (1995) ที่อธิบายถึงความสำคัญของการแปลงความรู้ฝังลึก (Tacit Knowledge) ให้เป็นความรู้ที่ชัดแจ้ง (Explicit Knowledge) เพื่อให้องค์กรสามารถถ่ายทอดความรู้ได้อย่างยั่งยืน [2]

การเลือกใช้แนวทางนี้มีเหตุผลสำคัญสองประการ ประการแรกแนวโน้มอุตสาหกรรมซอฟต์แวร์ปัจจุบันเผชิญกับความต้องการพัฒนาระบบที่เพิ่มขึ้นอย่างรวดเร็วท่ามกลางภาวะขาดแคลนนักพัฒนา (Developer Shortage) การใช้ Low-Code/No-Code จึงเป็นทางออกที่ช่วยให้หน่วยงานธุรกิจสามารถพัฒนาระบบใช้งานได้เองโดยไม่ต้องรอทรัพยากรจากฝ่ายไอที [4] ประการที่สอง ระบบ Low-Code ช่วยลดเวลาในการพัฒนาและปรับปรุงระบบทำให้สามารถตอบสนองต่อการเปลี่ยนแปลงความต้องการของธุรกิจได้รวดเร็ว เหมาะกับบริบทที่ความรู้ทางเทคนิคด้านโทรคมนาคมมีการเปลี่ยนแปลงบ่อย เพื่อใช้เป็นเครื่องมือในการจัดเก็บและถ่ายทอดองค์ความรู้ จากที่ปรึกษาทางด้านเทคนิคไปยังฝ่ายขาย ให้ฝ่ายขายสามารถเข้าถึงข้อมูลได้ทุกที่ทุกเวลา และลดการพึ่งพาที่ปรึกษาทางด้านเทคนิคในทุกขั้นตอน ของกระบวนการขาย

## 2. แนวคิด ทฤษฎี และงานวิจัยที่เกี่ยวข้อง

### 2.1 การจัดการองค์ความรู้และ SECI Model

Nonaka และ Takeuchi (1995) [2] จำแนกความรู้ออกเป็น 2 ประเภท คือ ความรู้ฝังลึก (Tacit Knowledge) ซึ่งอยู่ในตัวบุคคลและยากต่อการถ่ายทอด และความรู้ที่ชัดแจ้ง (Explicit Knowledge) ที่สามารถจัดเก็บและเผยแพร่ได้ โดยพัฒนา SECI Model อธิบายกระบวนการสร้างและถ่ายทอดความรู้ 4 ขั้นตอน ได้แก่ Socialization, Externalization, Combination และ Internalization ในการวิจัยนี้กระบวนการ Externalization มีความสำคัญสูงสุด คือการแปลงความรู้เชิงเทคนิคของที่ปรึกษาให้บันทึกเป็นฐานข้อมูลในระบบดิจิทัล Microsoft Power Apps

**2.2 การสกัดความรู้ (Knowledge Extraction) จากที่ปรึกษาทางด้านเทคนิค**

การสกัดความรู้ฝังลึกของที่ปรึกษาทางด้านเทคนิคให้กลายเป็นข้อมูลที่ฝ่ายขายเข้าใจและนำไปใช้ได้ อาศัยเทคนิค 3 ลักษณะ ได้แก่ (1) การสัมภาษณ์เชิงลึก (In-depth Interview) กับที่ปรึกษาเพื่อดึงเหตุผลเบื้องหลังของการแนะนำโซลูชันแต่ละประเภท (2) การจัดกลุ่มสถานการณ์การใช้งาน (Use-case Clustering) เพื่อแปลงความรู้เชิงบริบทให้เป็น Checklist มาตรฐาน และ (3) การทำ Solution Mapping ซึ่งเป็นการจับคู่ความต้องการของลูกค้า (pain points) กับโซลูชันของบริษัท ทำให้ความรู้เชิงซ้อนกลายเป็นรายการตัดสินใจที่ฝ่ายขายอ่านเข้าใจได้ทันที

**2.3 แพลตฟอร์ม Low-Code/No-Code และ Microsoft Power Apps**

การพัฒนาแบบ Low-Code เป็นวิธีที่ชาญฉลาดในการรับมือกับแรงกดดันจากความต้องการในการพัฒนาและปัญหาการขาดแคลนนักพัฒนา [4] Microsoft Power Apps เป็นแพลตฟอร์ม Low-Code/No-Code ที่ช่วยให้ผู้ใช้งานสร้างแอปพลิเคชันได้โดยไม่จำเป็นต้องมีทักษะการเขียนโค้ดขั้นสูง [5] รองรับ 3 รูปแบบหลัก ได้แก่ Canvas Apps สำหรับออกแบบหน้าจออิสระ Model-Driven Apps สำหรับจัดการข้อมูลซับซ้อน และ Power Pages สำหรับสร้างเว็บพอร์ทัล [6][7][8]

**2.4 หลักการออกแบบ UX/UI**

การออกแบบส่วนติดต่อผู้ใช้ที่ดีเป็นปัจจัยสำคัญในการทำให้ฐานความรู้กลางถูกใช้งานอย่างต่อเนื่อง งานวิจัยนี้อาศัยหลักการออกแบบ UX/UI ที่เผยแพร่โดย Figma [9] โดยเน้น 3 หลักการหลัก คือ (1) ความชัดเจน (Clarity) ให้ผู้ใช้เข้าใจสิ่งที่จะเกิดขึ้นได้ทันทีที่เห็นหน้าจอ (2) ความสม่ำเสมอ (Consistency) ใช้รูปแบบองค์ประกอบและสีที่เหมือนกันทั่วทั้งระบบ และ (3) ลำดับความสำคัญเชิงภาพ (Visual Hierarchy) จัดให้ข้อมูลสำคัญเด่นกว่าข้อมูลรอง เกณฑ์เหล่านี้ถูกนำไปใช้วัด "ความสะดวกในการใช้งาน" และ "ความพึงพอใจด้านอินเทอร์เฟซ" ในขั้นตอนประเมินผล

**2.5 งานวิจัยที่เกี่ยวข้อง**

งานวิจัยที่เกี่ยวข้องพบว่าการนำ Microsoft Power Apps มาใช้ในองค์กรสามารถแก้ปัญหาการดำเนินงานได้อย่างมีประสิทธิภาพ เช่น การพัฒนาระบบคืนสินค้า [10] การจัดการคลังสินค้ายางรถยนต์ที่มีผลการประเมินความพึงพอใจ 4.70/5.00 [11] และการตรวจสอบคุณภาพในกระบวนการผลิตที่สามารถลดระยะเวลาดำเนินงานได้ร้อยละ 64.89 [12] นอกจากนี้ การนำ SECI Model มาประยุกต์ใช้พบว่าฝ่ายพัฒนาทรัพยากรมนุษย์มีบทบาทสำคัญในการเชื่อมต่อช่องว่างองค์ความรู้ระหว่างฝ่ายต่าง ๆ ในองค์กร [15]

อย่างไรก็ตามงานวิจัยก่อนหน้าส่วนใหญ่เน้นการจัดการสิ่งของหรือทรัพย์สินที่จับต้องได้ เช่น คลังสินค้า การผลิต หรือกระบวนการคืนสินค้า ซึ่งเป็นความรู้ที่ชัดแจ้งอยู่แล้ว (Structured Data) ในขณะที่งานวิจัยนี้มุ่งเน้นการถ่ายทอดความรู้เชิงเทคนิคระหว่างบุคคล (Tacit-to-Explicit Knowledge Transfer) ซึ่งเป็นความรู้ที่ซับซ้อนยากต่อการเก็บและถ่ายทอด จุดแตกต่างนี้คือช่องว่างทางการวิจัย (Research Gap) ที่งานวิจัยนี้พยายามเติมเต็ม โดยประยุกต์ใช้ Microsoft Power Apps ร่วมกับกรอบ SECI Model เพื่อเปลี่ยนความรู้ฝังลึกของที่ปรึกษาให้เป็นฐานความรู้ดิจิทัลที่ฝ่ายขายใช้งานได้จริง

## 3. วิธีดำเนินการวิจัย

การดำเนินการวิจัยครั้งนี้แบ่งออกเป็น 4 ระยะ ตามกรอบแนวคิดการวิจัยและพัฒนาระบบ (Research and Development)

**3.1 ระยะที่ 1 วิจัยและวิเคราะห์**

ศึกษาปัญหาและความต้องการจากพนักงานขายและที่ปรึกษาทางด้านเทคนิคด้วยแบบสอบถาม การสัมภาษณ์เชิงลึก และการสังเกตการณ์ภาคสนาม ผลการสัมภาษณ์ (n=6) พบว่าปัญหาที่รุนแรงที่สุดคือความต้องการสื่อที่เข้าถึงและใช้งานง่าย (คะแนน 5.00/5.00) และการรอที่ปรึกษานาน (4.50/5.00) พร้อมวิเคราะห์ผลิตภัณฑ์ที่ลูกค้าให้ความสนใจสูงสุด 5 อันดับแรกดังสมการที่ 3-1

คะแนนรวม = (ความสนใจเฉลี่ย × 0.50) + (ความถี่ที่ลูกค้าสนใจแบบแปลงสเกล × 0.30) + (ระยะเวลาตอบสนอง × 0.20) (3-1)

ผลคะแนนรวมความสนใจสูงสุดได้แก่ Internet Solutions (4.37) Cyber Security Solutions (3.42) 5G Solutions (3.39) Cloud Solutions (2.98) และ ICT Solutions (2.87)

**3.2 ระยะที่ 2 การออกแบบและพัฒนา**

ออกแบบโครงสร้างฐานข้อมูล SharePoint Lists จำนวน 5 รายการ ได้แก่ Solutions, Customers, Pricelist, Checklist และ Coverage Requests พร้อมออกแบบ Wireframe ของ Canvas App ใน Microsoft Power Apps ประกอบด้วย 6 หน้าจอหลัก ได้แก่ Home Screen, Customer Screen, Solution Screen, Checklist Screen, Price List Screen และ Check Coverage Screen โดยออกแบบตามหลักการ UX/UI ด้านความชัดเจน ความสม่ำเสมอ และลำดับความสำคัญเชิงภาพ [9]

**ตารางที่ 1:** ตารางโครงสร้างหน้าจอแอปพลิเคชัน

| หน้าจอ | รูปแบบ Layout | ฟังก์ชันหลัก | แหล่งข้อมูล |
|---|---|---|---|
| Home Screen | Card Grid Layout | นำทางสู่เมนูหลักทั้ง 5 เมนู | — |
| Customer Screen | Master-Detail Layout | เพิ่ม แก้ไข ลบ และเรียกดูข้อมูลลูกค้า | Customers |
| Solution Screen | Master-Detail Layout | เรียกดูโซลูชัน ค้นหา แสดงรูปภาพและลิงก์อ้างอิง | Solutions |
| Checklist Screen | Master-Detail Layout | เรียกดูคำถามและข้อควรรู้ตามโซลูชัน | Check Lists |
| Price List Screen | Master-Detail Layout | เรียกดูข้อมูลราคาแบบ Read-Only | Price Lists |
| Check Coverage Screen | Two-Panel Layout | บันทึกและเรียกดูข้อมูลคำขอตรวจสอบพื้นที่ | Check EDS Coverage |

**3.3 ระยะที่ 3 การทดสอบและปรับปรุง**

ทดสอบฟังก์ชันการทำงานหลักกับกลุ่มตัวอย่าง 2 กลุ่ม ได้แก่ พนักงานขายและที่ปรึกษาทางด้านเทคนิค (n=8) และลูกค้าองค์กร (n=5) ทดสอบฟังก์ชันหลัก 8 รายการ ได้แก่ การเชื่อมต่อ SharePoint Lists การแสดงผลข้อมูล การค้นหา การบันทึก การลบ การนำทาง และการแสดงลิงก์ พบว่าผ่านทุกรายการ หลังจากปรับปรุงตามผลตอบรับแล้วระบบพร้อมสำหรับใช้งานจริง

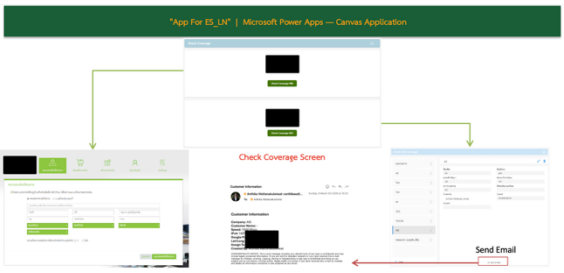
**3.4 ระยะที่ 4 การประเมินผล**

ประเมินประสิทธิภาพด้วยแบบสอบถามคะแนน 1–5 ครอบคลุมด้านความครบถ้วนของข้อมูล ความสะดวกในการใช้งาน ความเสถียร และการสนับสนุนกระบวนการขาย จากกลุ่มพนักงานขายและที่ปรึกษาทางด้านเทคนิค (n=8) และกลุ่มลูกค้าองค์กร (n=5) เนื่องจากความรู้ทางเทคนิคในธุรกิจโทรคมนาคมมีการเปลี่ยนแปลงบ่อย ผู้วิจัยจึงกำหนดโครงสร้างผู้ดูแลข้อมูล (Content Owner) โดยแบ่งความรับผิดชอบเป็น 3 ระดับ คือ (1) ที่ปรึกษาทางด้านเทคนิคเป็นเจ้าของเนื้อหาประเภท Solutions และ Checklist และรับผิดชอบทบทวนเนื้อหาไตรมาสละ 1 ครั้ง (2) หัวหน้าฝ่ายขายเป็นเจ้าของรายการ Pricelist และรายการ Customers และ (3) ผู้ดูแลระบบ (Admin) เป็นผู้ควบคุมสิทธิ์การเข้าถึงและเก็บ log การเปลี่ยนแปลงข้อมูล

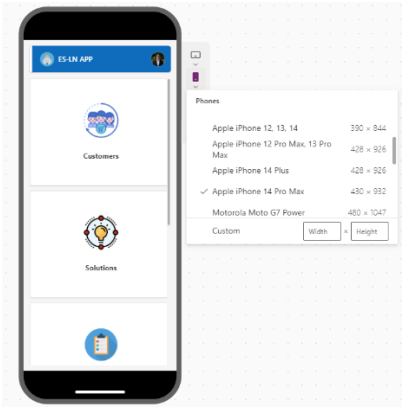
## 4. ผลการดำเนินการวิจัย

**4.1 ผลการพัฒนาแอปพลิเคชัน**

แอปพลิเคชัน "App For ES_LN" พัฒนาสำเร็จในรูปแบบ Microsoft Power Apps Canvas Application เชื่อมต่อกับ Microsoft SharePoint Lists จำนวน 5 รายการ ครอบคลุมการทำงาน 5 ด้านหลัก ได้แก่ การจัดการข้อมูลลูกค้าองค์กร การเรียกดูและค้นหาข้อมูลโซลูชัน การตรวจสอบข้อมูลราคาผลิตภัณฑ์ การเก็บข้อมูลความต้องการจากลูกค้า และการบันทึกคำขอตรวจสอบพื้นที่ให้บริการ สามารถใช้งานได้ทั้งบนคอมพิวเตอร์และอุปกรณ์เคลื่อนที่




**ภาพที่ 1:** ตัวอย่างการใช้งานการบันทึกคำขอตรวจสอบพื้นที่ให้บริการ พร้อมทั้งสามารถส่งข้อมูลไปยังที่ปรึกษาทางเทคนิค




**ภาพที่ 2:** ตัวอย่างการใช้งานบนอุปกรณ์เคลื่อนที่

**4.2 ผลการประเมินจากกลุ่มพนักงานขายและที่ปรึกษาทางด้านเทคนิค**

ผลการประเมินจากกลุ่มพนักงานขาย ที่ปรึกษาทางด้านเทคนิค และหัวหน้างาน (n=8) ด้วยแบบสอบถาม 14 ข้อ ครอบคลุมด้านความครบถ้วนของข้อมูล ความสะดวกในการใช้งาน ความเสถียร และการสนับสนุนกระบวนการขาย พบว่าค่าเฉลี่ยรวม 4.08/5.00 อยู่ในระดับมาก หัวข้อที่ได้คะแนนสูงสุดคือความสามารถใช้งานได้สะดวกบนทุกอุปกรณ์ PC/Tablet/Mobile (4.50/มากที่สุด) รองลงมาคืออินเทอร์เฟซที่ใช้งานง่ายและเข้าใจได้ทันที (4.38) และความพึงพอใจโดยรวม (4.25)และการค้นหาข้อมูลภายในแอปทำได้รวดเร็ว (4.13) ส่วนหัวข้อที่ควรพัฒนาต่อคือความถูกต้องของข้อมูลราคาและรายละเอียดผลิตภัณฑ์ (3.88) ซึ่งชี้ให้เห็นถึงความจำเป็นในการวางระบบ Content Ownership เพื่อดูแลข้อมูลอย่างต่อเนื่อง

**4.3 ผลการประเมินจากกลุ่มลูกค้า**

ผลการประเมินจากกลุ่มลูกค้าองค์กร (n=5) ด้วยแบบสอบถาม 7 ข้อ ที่ออกแบบเพื่อวัดประสบการณ์ของลูกค้าหลังพนักงานขายใช้แอปพลิเคชัน พบค่าเฉลี่ยรวม 4.06/5.00 อยู่ในระดับมาก หัวข้อที่ได้คะแนนสูงสุดคือพนักงานขายสามารถให้ข้อมูลราคาเบื้องต้นได้อย่างถูกต้อง (4.40) สะท้อนว่าลูกค้ามีความเชื่อมั่นในความแม่นยำของข้อมูลที่ได้รับ รองลงมาคือ การตอบคำถามได้อย่างรวดเร็ว (4.20) ความเหมาะสมของระยะเวลา (4.20) และการนำเสนอช่วยสนับสนุนการตัดสินใจ (4.20) ส่วนหัวข้อที่ควรพัฒนาคือการให้ข้อมูลทางเทคนิคโดยไม่ต้องรอผู้เชี่ยวชาญ (3.20) อยู่ในระดับปานกลางซึ่งชี้ให้เห็นว่ายังมีช่องว่างด้านความรู้ทางเทคนิคเชิงลึกที่ควรเสริมในแอปพลิเคชัน

**4.4 การอภิปรายผล**

ผลการวิจัยสอดคล้องกับแนวคิด SECI Model ของ Nonaka และ Takeuchi (1995) โดยเฉพาะในกระบวนการ Externalization ซึ่งเป็นการแปลงความรู้ฝังลึกของที่ปรึกษาทางเทคนิคให้กลายเป็นความรู้ที่ชัดแจ้งและจัดเก็บอย่างเป็นระบบผ่านแพลตฟอร์มดิจิทัล โดยเฉพาะอย่างยิ่งในส่วนของ Checklist Screen ที่รวบรวมคำถามและแนวทางการสอบถามลูกค้าจากที่ปรึกษาทางเทคนิคไว้อย่างเป็นระบบ ทำให้พนักงานขายสามารถนำไปใช้เก็บข้อมูลความต้องการจากลูกค้าได้อย่างครบถ้วนและถูกต้อง โดยไม่จำเป็นต้องอาศัยประสบการณ์ส่วนตัวของที่ปรึกษาทางเทคนิคในทุกสถานการณ์

**4.5 ข้อจำกัดของงานวิจัย (Limitations)**

งานวิจัยครั้งนี้มีข้อจำกัดที่สำคัญสองประการ ประการแรก กลุ่มตัวอย่างที่ใช้ในการประเมินมีจำนวนรวมเพียง 13 คน (พนักงาน 8 คน ลูกค้า 5 คน) ซึ่งเป็นขนาดตัวอย่างที่ค่อนข้างน้อยตามหลักสถิติ ผลการประเมินในงานวิจัยนี้จึงสะท้อนมุมมองของผู้ใช้ภายในหน่วยงานเอกชนแห่งเดียวในเขตภาคเหนือตอนล่าง ควรตีความเป็นผลการศึกษาเชิงกรณีศึกษา (Case Study) ในลักษณะ Pilot Implementation มากกว่าการสรุปอิงประชากรทั่วไป ประการที่สอง ยังไม่มีข้อมูลเชิงปริมาณแบบ ก่อนและหลังที่ชัดเจน เช่น ระยะเวลาที่ลดลงในการรอที่ปรึกษา หรือสัดส่วนเคสที่ฝ่ายขายสามารถปิดการขายได้โดยไม่ต้องเรียกที่ปรึกษา ซึ่งงานวิจัยในอนาคตควรออกแบบให้เก็บข้อมูลเหล่านี้อย่างเป็นระบบ

## 5. สรุป

การวิจัยครั้งนี้แสดงให้เห็นว่าการนำ Microsoft Power Apps ซึ่งเป็นแพลตฟอร์ม Low-code มาประยุกต์ใช้พัฒนาแอปพลิเคชันสนับสนุนการขายสามารถแก้ปัญหาด้านการถ่ายทอดองค์ความรู้เชิงเทคนิคระหว่างที่ปรึกษาและพนักงานขายได้อย่างมีประสิทธิภาพ โดยสอดคล้องกับกระบวนการ Externalization ใน SECI Model ของ

Nonaka และ Takeuchi ที่แปลงความรู้ฝังลึกของที่ปรึกษาให้กลายเป็นความรู้ที่ชัดแจ้งในระบบดิจิทัล

แอปพลิเคชัน "App For ES_LN" ที่พัฒนาขึ้นเป็นต้นแบบที่มีคุณค่าสำหรับนำไปพัฒนาต่อยอดและขยายผลในระดับองค์กร ข้อเสนอแนะสำหรับการวิจัยในอนาคตประกอบด้วย 4 ด้าน ดังนี้

(1) การขยายขนาดกลุ่มตัวอย่าง ควรเพิ่มจำนวนผู้ร่วมประเมินและขยายการใช้งานไปยังสาขาภูมิภาคอื่น ๆ ของบริษัท เพื่อให้ผลการศึกษาสะท้อนความเป็นตัวแทนประชากรได้กว้างขึ้น

(2) การเก็บข้อมูลเชิงปริมาณแบบก่อนและหลังเช่น ระยะเวลาเฉลี่ยในการรอข้อมูลจากที่ปรึกษา ความถูกต้องของการนำเสนอโซลูชัน และสัดส่วนของเคสที่ฝ่ายขายสามารถปิดการขายโดยไม่ต้องเรียกที่ปรึกษา เพื่อแสดงความสำเร็จเป็นรูปธรรมมากขึ้น

(3) การบูรณาการกับ Microsoft Power Automate เพื่อพัฒนาระบบแจ้งเตือนอัตโนมัติเมื่อมีการอัปเดตข้อมูลโซลูชันหรือราคา และสร้าง Workflow การอนุมัติคำขอตรวจสอบพื้นที่ให้บริการ

(4) การเชื่อมต่อกับ Microsoft Power BI เพื่อพัฒนาระบบรายงานผลเชิงวิเคราะห์ (Analytics Dashboard) ที่แสดงสถิติการใช้งานระบบ ผลิตภัณฑ์ที่ลูกค้าสนใจ และแนวโน้มคำขอตรวจสอบพื้นที่ เพื่อสนับสนุนการตัดสินใจเชิงกลยุทธ์ของผู้บริหาร

ยิ่งไปกว่านั้น การต่อยอดด้วยปัญญาประดิษฐ์หรือระบบแนะนำ (AI / Recommendation System) เพื่อแนะนำโซลูชันที่เหมาะสมกับลูกค้าโดยอัตโนมัติ จะช่วยยกระดับระบบจากเครื่องมือจัดการความรู้ ให้กลายเป็นระบบสนับสนุนการตัดสินใจเชิงกลยุทธ์ขององค์กร (Strategic Decision-Support System)

## เอกสารอ้างอิง


[1] สำนักงานสถิติแห่งชาติ (NSO). (2566). รายงานการสำรวจการมีการใช้เทคโนโลยีสารสนเทศและการสื่อสารในสถานประกอบการ. สืบค้นจาก https://surveybkk.nso.go.th/indexbkk.php

[2] Nonaka, I., & Takeuchi, H. (1995). The Knowledge-Creating Company: How Japanese Companies Create the Dynamics of Innovation. Oxford University Press.

[3] Microsoft Corporation. (2024). What is Power Apps? Retrieved from https://learn.microsoft.com/en-us/power-apps/powerapps-overview

[4] Microsoft. (2567). การพัฒนาแบบ Low-Code คืออะไร. สืบค้นจาก https://www.microsoft.com/th-th/power-platform/products/power-apps/topics/low-code-no-code/what-is-low-code

[5] Microsoft. (2023). What is Power Apps? Microsoft Learn. Retrieved from https://learn.microsoft.com/en-us/power-apps/powerapps-overview

[6] Microsoft. (2024). What are canvas apps? Microsoft Learn. Retrieved from https://learn.microsoft.com/en-us/power-apps/maker/canvas-apps/getting-started

[7] Microsoft. (2024). Model-driven apps. Microsoft Learn. Retrieved from https://learn.microsoft.com/en-us/power-apps/maker/model-driven-apps/model-driven-app-overview

[8] Microsoft. (2024). Power Pages. Retrieved from https://www.microsoft.com/th-th/power-platform/products/power-pages

[9] Figma. (2024). UI design principles. Retrieved from https://www.figma.com/resource-library/ui-design-principles/

[10] สถาบันเทคโนโลยีไทย-ญี่ปุ่น. (2563). การพัฒนาระบบคืนสินค้าของลูกค้าโดยใช้โปรแกรม Power Apps. สืบค้นจาก https://e-research.siam.edu

[11] มหาวิทยาลัยราชภัฏอุดรธานี. (2565). งานวิจัยการพัฒนาระบบด้วย Power Apps. สืบค้นจาก https://ms.udru.ac.th

[12] มหาวิทยาลัยธรรมศาสตร์. (2567). วิทยานิพนธ์ปริญญาโท มหาวิทยาลัยธรรมศาสตร์ ปีการศึกษา 2567. สืบค้นจาก https://ethesisarchive.library.tu.ac.th

[13] พิชิต งามจรัสศรีวิชัย และ จักร ดิงศภัทิย์. (2560). วารสารวิจัยและพัฒนา มจธ. ปีที่ 40 ฉบับที่ 1. TCI กลุ่ม 1.

# การเปรียบเทียบประสิทธิภาพเครื่องมือสแกนช่องโหว่เว็บแอปพลิเคชัน และกระบวนการยืนยันผลการแก้ไข

# Performance Comparison of Web Application Vulnerability Scanning Tools and Patch Validation Processes

*ณัฐธิดา ลูกเสือ (Natthida Looksua)[1] และสุนันทา สดสี (Sunantha Sodsee)[2]*
*ภาควิชาการบริหารเครือข่ายดิจิทัลและความมั่นคงปลอดภัยสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล*
*มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ*
*s6707031857182@email.kmutnb.ac.th, sunantha.s@itd.kmutnb.ac.th*

## บทคัดย่อ

*การวิจัยนี้มีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพของเครื่องมือสแกนช่องโหว่เว็บแอปพลิเคชันและพัฒนากระบวนการยืนยันผลการแก้ไขช่องโหว่ให้เป็นระบบ โดยการนำเครื่องมือสแกนช่องโหว่จำนวน 3 ชนิด ได้แก่ OWASP ZAP, Arachni และ Burp Suite มาทำการทดสอบประสิทธิภาพผ่านเกณฑ์การประเมิน 4 มิติ รวม 9 หัวข้อสำคัญ จากผลการวิจัยพบว่า เครื่องมือ OWASP ZAP มีประสิทธิภาพและความเหมาะสมสูงสุด และแม่นยำที่สุด ผู้วิจัยจึงได้จัดทำคู่มือแนวทางการแก้ไขช่องโหว่ และทำการตรวจสอบยืนยันผลพบว่าสามารถลดจำนวนช่องโหว่ระดับสูงและปานกลางได้ร้อยละ 100 จากนั้นนำเสนอผลการประเมินคุณภาพคู่มือโดยผู้เชี่ยวชาญซึ่งมีความพึงพอใจในระดับมากที่สุดที่ค่าเฉลี่ย 4.54 เพื่อใช้เป็นแนวปฏิบัติมาตรฐาน และฐานข้อมูลความรู้สำหรับนักพัฒนาในการยกระดับความมั่นคงปลอดภัยของเว็บแอปพลิเคชันได้อย่างมีประสิทธิภาพต่อไป*

**คำสำคัญ:** ความมั่นคงปลอดภัยเว็บแอปพลิเคชัน การแสกนช่องโหว่ การจัดการช่องโหว่

## Abstract

*This research aims to compare the performance of web application vulnerability scanning tools and develop a systematic remediation validation process. Three vulnerability scanning tools—OWASP ZAP, Arachni, and Burp Suite—were utilized to test performance through evaluation criteria across four dimensions, totaling nine key items. The research findings revealed that OWASP ZAP demonstrated the highest efficiency, suitability, and accuracy. Consequently,the researcher developed a vulnerability remediation manual and conducted validation, which indicated that high-level and medium-level vulnerabilities were reduced by 100 percent. Furthermore, the manual's quality was evaluated by experts, who expressed the highest level of satisfaction with an average score of 4.54. This manual serves as a standard operating procedure (SOP) and a knowledge base for developers to effectively enhance web application security in the future.*

**Keywords:** Web Application Security, Vulnerability Scanning, Vulnerability Management.

## 1. บทนำ

ในยุคดิจิทัล เว็บแอปพลิเคชันเป็นเป้าหมายหลักของการโจมตีไซเบอร์ โดยปี 2024 ไทยเผชิญการโจมตีเฉลี่ยสูงถึง 5,811 ครั้งต่อวัน(Fortinet, 2024) [1] ซึ่งสร้างความเสียหายมหาศาลต่อองค์กร ในปี 2025 ค่าใช้จ่ายเฉลี่ยทั่วโลกจากเหตุการณ์ข้อมูลรั่วไหลสูงถึง 4.88 ล้านดอลลาร์สหรัฐ ซึ่งเพิ่มขึ้น 10% จากปีก่อนหน้า(IBM, 25025) [2] เครื่องมือสแกนช่องโหว่จึงเป็นสิ่งจำเป็น แต่การเลือกใช้ที่ไม่เหมาะสมอาจนำไปสู่ปัญหาผลบวกลวง ทำให้เสีย

ทรัพยากร หรือผลลบลวงที่ปล่อยให้ช่องโหว่ร้ายแรงคงอยู่ นอกจากนี้ กระบวนการยืนยันผลการแก้ไข มักถูกมองข้าม ทำให้การจัดการความปลอดภัยไม่สมบูรณ์

บทความนี้จึงมุ่งวิเคราะห์เปรียบเทียบประสิทธิภาพเครื่องมือสแกน พร้อมเสนอกระบวนการตรวจสอบที่เป็นระบบ เพื่อสร้างวงจรการจัดการช่องโหว่ที่ช่วยให้องค์กรและนักพัฒนาตัดสินใจเลือกเครื่องมือได้เหมาะสม และยกระดับความปลอดภัยของเว็บแอปพลิเคชันได้อย่างยั่งยืน

## 2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ผู้วิจัยได้ศึกษาค้นคว้าเอกสารและงานวิจัยที่เกี่ยวข้องต่าง ๆ เพื่อนำทฤษฎี แนวคิด มาประยุกต์ใช้กับงานวิจัย เรียงตามลำดับหัวข้อดังต่อไปนี้

**2.1 แนวคิดและทฤษฎีเกี่ยวกับความมั่นคงปลอดภัยของแอพพลิเคชั่น** แนวคิดด้านความมั่นคงปลอดภัยที่สำคัญประกอบด้วย

2.1.1 หลักการพื้นฐาน CIA Triad การรักษาความมั่นคงปลอดภัยสารสนเทศมุ่งเน้นที่การคุ้มครององค์ประกอบ 3 ประการ ความลับ (Confidentiality), ความถูกต้องครบถ้วน (Integrity), ความพร้อมใช้งาน (Availability)

2.1.2 มาตรฐาน OWASP Top 10[3] โครงการของ OWASP ที่จัดอันดับความเสี่ยงทางด้านความมั่นคงปลอดภัยของเว็บแอปพลิเคชันไว้ 10 อันดับ

**2.2 แนวคิดเกี่ยวกับเทคโนโลยีการตรวจหาช่องโหว่เว็บแอปพลิเคชัน** เทคโนโลยีที่ใช้ในการค้นหาและวิเคราะห์ช่องโหว่ สามารถจำแนกตามลักษณะการทดสอบได้เป็น 3 ประเภทหลัก ดังนี้

2.2.1 Static Application Security Testing (SAST) โดยการวิเคราะห์ซอร์สโค้ด หรือไบนารีของแอปพลิเคชันโดยไม่มีการรันระบบ ช่วยให้ตรวจพบข้อผิดพลาดในการเขียนโปรแกรมได้ตั้งแต่ระยะเริ่มต้น

2.2.2 Dynamic Application Security Testing (DAST) โดยจะจำลองบทบาทเป็นผู้โจมตีเพื่อส่งชุดคำสั่งทดสอบ เข้าไปยังแอปพลิเคชันที่กำลังทำงานอยู่ ผ่านอินเทอร์เฟซของเว็บ เพื่อสังเกตการตอบสนองและระบุช่องโหว่ที่เกิดขึ้นจริงในสภาวะการใช้งานจริง

2.2.3 Interactive Application Security Testing (IAST) เป็นเทคโนโลยีแบบผสมผสาน โดยการติดตั้งโปรแกรมตรวจวัด ไว้ภายในแอปพลิเคชัน เพื่อวิเคราะห์การทำงานภายในร่วมกับการทดสอบแบบ Dynamic ซึ่งช่วยเพิ่มความแม่นยำและลดอัตราการแจ้งเตือนที่ผิดพลาด

**2.3 วงจรการบริหารจัดการช่องโหว่ (Vulnerability Management Lifecycle)** วงจรการบริหารจัดการช่องโหว่ประกอบด้วย 6 ขั้นตอนหลัก ขั้นตอนที่ 1 การสำรวจและจัดทำบัญชีสินทรัพย์, ขั้นตอนที่ 2 การประเมินช่องโหว่อัตโนมัติ, ขั้นตอนที่ 3 การจัดลำดับความสำคัญตามระดับความเสี่ยง, ขั้นตอนที่ 4 การการแก้ไขและตอบสนอง, ขั้นตอนที่ 5 การยืนยันผลและการเฝ้าระวังอย่างต่อเนื่อง, ขั้นตอนที่ 6 การรายงานผลและการปรับปรุงอย่างต่อเนื่อง

**2.4 งานวิจัยที่เกี่ยวข้อง**

Belay และคณะ[4] ได้ทำการศึกษาการนำเสนอแนวทางการประเมินโดยการนำเครื่องมือสแกนหลายตัวมาทดสอบกับเว็บแอปที่จงใจสร้างให้มีช่องโหว่เดียวกัน Escarrone และคณะ[5] ได้ทำการศึกษาการสแกนเปรียบเทียบเครื่องมือ 10 ตัวบนเว็บเป้าหมายมาตรฐาน OWASP และเว็บที่ใช้ JS เป็นหลัก โดยศึกษากรณีตัวอย่างจากสภาพแวดล้อมในองค์กร และใช้คนตรวจสอบผลซ้ำเพื่อให้ผลลัพธ์ Potti และคณะ [6] ได้ทำการศึกษาแนวทางการประเมินโดยนำโปรแกรม OWASP ZAP เวอร์ชัน 2.12.0 และ 2.13.0 มาสแกนเปรียบเทียบกันโดยตรงบนชุดทดสอบมาตรฐานของ OWASP ซึ่งช่วยแสดงให้เห็นถึงการเปลี่ยนแปลงประสิทธิภาพของเครื่องมือในแต่ละเวอร์ชัน Jha, T. K. [7] ได้ทำการศึกษาโดยมีข้อจำกัดอยู่บนเว็บแอปเพียง 2 ตัวและขอบเขตประเภทช่องโหว่จำกัด Kondraciuk และคณะ[8] ได้ทำการศึกษาเทียบผลลัพธ์ "ก่อนและหลัง" การนำระบบ JARVIS มาใช้งานร่วมกับเครื่องมือสแกน 5 ตัว บนเว็บแอปพลิเคชันที่ไม่ปลอดภัย 7 ตัว Mburano, B., & Si, W. [9] ได้ทำการศึกษาการทดสอบแบบจำลองโดยนำเสนอกรอบการทดสอบเปรียบเทียบที่พัฒนาขึ้นมาใหม่และครอบคลุมกว่าเดิม Alshreef, และคณะ [10] ได้ทำการศึกษางานวิจัย 90 ฉบับ เพื่อรวบรวมข้อมูลประเภทเครื่องมือสแกน ตัวชี้วัด

ประสิทธิภาพ และสร้างแผoภาพรวมที่ชี้ให้เห็นถึงช่องว่างและความเอนเอียงในงานวิจัยที่ผ่านมา

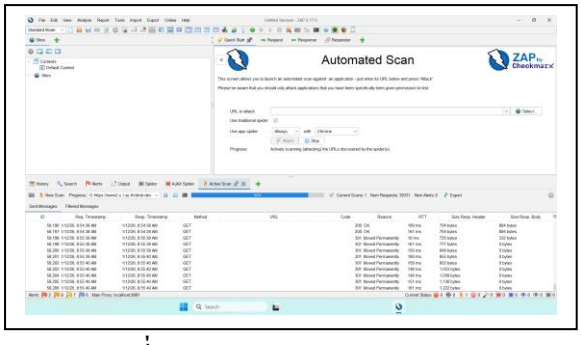
## 3. วิธีการดำเนินการวิจัย

งานวิจัยนี้เป็นการวิจัยเชิงทดลองและแก้ปัญหา มุ่งเน้นการประเมินสมรรถนะเครื่องมือที่เหมาะสมกับบริบทข้อจำกัดขององค์กร พร้อมพัฒนากระบวนการยืนยันความถูกต้องของการจัดการช่องโหว่การเสนอแนะการแก้ไข

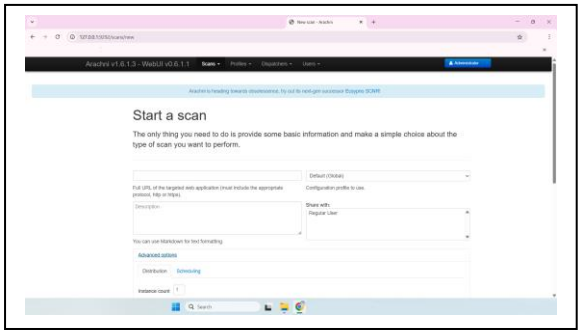
### 3.1 การศึกษาและพัฒนากรอบแนวคิด

กรอบแนวคิดการวิจัยนี้ได้รับการพัฒนาขึ้นโดยบูรณาการองค์ความรู้จากมาตรฐานความมั่นคงปลอดภัยสากล วงจรการบริหารจัดการช่องโหว่ และขั้นตอนการทดลองเชิงประยุกต์ โดยแบ่งออกเป็น 3 ส่วนหลัก ดังนี้

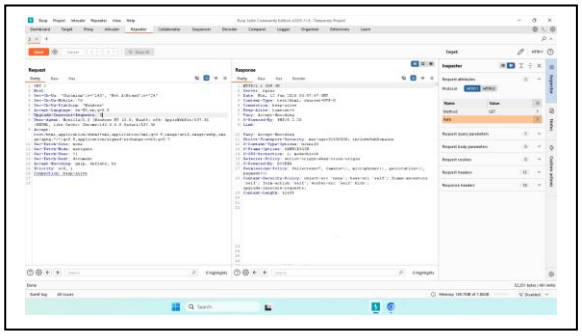
3.1.1 ปัจจัยนำเข้า (Input) ประกอบด้วย มาตรฐานความปลอดภัยสากล OWASP Top 10 (2021) และ NIST SP 800-115 [11] เครื่องมือสแกนช่องโหว่ OWASP ZAP, Arachni และ Burp Suite เกณฑ์การวัดผลประสิทธิภาพ ตัวชี้วัดเปอร์เซ็นต์การลดลงของช่องโหว่เว็บแอปพลิเคชันเป้าหมายภายใต้บริบทขององค์กร และรายการช่องโหว่ตั้งต้นที่ตรวจพบ




**ภาพที่ 1:** OWASP ZAP version 2.17.0




**ภาพที่ 2:** Archni version 1.6.1.3 -Web UI version 0.6.1.1



**ภาพที่ 3:** Burp Suite Version Community Edition 2025.11.6

3.1.2 กระบวนการ (Process) ประกอบด้วย

การทดลองสแกนเพื่อประเมินประสิทธิภาพเชิงเปรียบเทียบของเครื่องมือทั้ง 3 ชนิด การคัดเลือกเครื่องมือที่มีความเหมาะสมกับบริบทขององค์กรมากที่สุด การแก้ไขให้ตรงกับประเภทของช่องโหว่ การจัดทำคู่มือและดำเนินการแก้ไขช่องโหว่ในระดับซอร์สโค้ดและการตั้งค่าระบบและการสแกนซ้ำเพื่อตรวจสอบและยืนยันผล

3.1.3 ผลลัพธ์ (Output) ประกอบด้วย

รายงานผลการเปรียบเทียบประสิทธิภาพของเครื่องมือสแกนช่องโหว่เว็บแอปพลิเคชัน คู่มือปฏิบัติงานเพื่อการแก้ไขช่องโหว่สำหรับนักพัฒนาและผู้ดูแลระบบ สรุปผลดัชนีชี้วัดความสำเร็จในรูปแบบเปอร์เซ็นต์การลดลงของช่องโหว่ภายหลังการดำเนินการแก้ไข

### 3.2 การพัฒนาแนวทางการประยุกต์ใช้

เป็นการนำผลการทดสอบประสิทธิภาพของเครื่องมือมาจัดทำ "คู่มือแนวทางการแก้ไขช่องโหว่" โดยผู้วิจัยได้ออกแบบโครงสร้างเนื้อหาให้เป็นมาตรฐาน 7 หัวข้อหลัก (เช่น ระดับความเสี่ยง, สาเหตุของช่องโหว่, วิธีการแก้ไข และตัวอย่างซอร์สโค้ด) เพื่อเปลี่ยนรายงานทางเทคนิคที่ซับซ้อนให้เป็นแนวปฏิบัติที่ชัดเจนสำหรับนักพัฒนา โดยมุ่งเน้นการแก้ปัญหาที่ต้นเหตุด้วยหลักการ Secure Coding และการตั้งค่าระบบ ตามมาตรฐาน OWASP

### 3.3 การทดลองใช้งานจริงและการประเมินผล

เป็นการนำคู่มือที่พัฒนาขึ้นไปประยุกต์ใช้ในการแก้ไขช่องโหว่บนระบบเว็บแอปพลิเคชันจริง โดยระบบที่ใช้ในการทดสอบครั้งนี้คือ เว็บแอปพลิเคชันประเภทการจัดการรายงานข้อมูล ซึ่งพัฒนาด้วยภาษา PHP และรันอยู่บน

สภาพแวดล้อม UAT ขององค์กร เพื่อให้การทดลองสะท้อนต่อระบบงานจริงในสภาพแวดล้อมที่ควบคุมได้ และได้ดำเนินการสแกนซ้ำ เพื่อตรวจสอบยืนยันผล โดยใช้ดัชนีชี้วัดเปอร์เซ็นต์การลดลงของช่องโหว่ จากนั้นนำคู่มือและผลการดำเนินงานเสนอต่อผู้เชี่ยวชาญจำนวน 4 ท่าน เพื่อประเมินคุณภาพและความพึงพอใจ

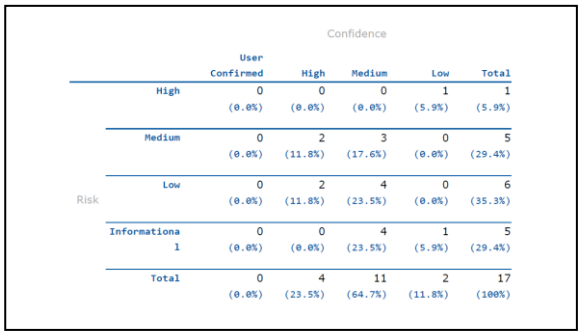
## 4. ผลการวิจัย

**4.1 ผลการเปรียบเทียบและคัดเลือกเครื่องมือสแกนช่องโหว่** จากการวิเคราะห์เชิงเปรียบเทียบในแบบประเมินทั้ง 9 หัวข้อ ผู้วิจัยพบว่า OWASP ZAP มีคะแนนประเมินอยู่ในเกณฑ์ 'ดีเยี่ยม' และสามารถตอบโจทย์การวิจัยภายใต้ข้อจำกัดขององค์กรได้ครบถ้วนที่สุด โดยเฉพาะด้านความแม่นยำและการนำข้อมูลรายงานไปใช้งานต่อ ดังนั้น ผู้วิจัยจึงตัดสินใจคัดเลือก OWASP ZAP เป็นเครื่องมือหลักในการดำเนินงานวิจัยในขั้นตอนถัดไป

**ตารางที่ 1:** ผลการเปรียบเทียบและคัดเลือกเครื่องมือสแกนช่องโหว่

| คุณสมบัติการทำงาน | OWASP ZAP | Arachni | Burp Suite |
|---|---|---|---|
| ใบอนุญาต (License) | 1 | 1 | 0.5 |
| สถานะการอัปเดต | 1 | 0 | 1 |
| การสแกนอัตโนมัติ | 1 | 1 | 0 |
| การทดสอบแบบแมนนวล | 1 | 0 | 1 |
| การรองรับ Modern Web(SPA/AJAX) | 1 | 0.5 | 0.5 |
| ความครอบคลุมของประเภทช่องโหว่ | 1 | 0.5 | 0.5 |
| การขยายขีดความสามารถ | 1 | 0 | 1 |
| การเชื่อมต่อ CI/CD | 1 | 0.5 | 0 |
| ความสามารถในการนำข้อมูลรายงานออกมาใช้งาน | 1 | 0.5 | 0 |
| รวม | 9 | 4 | 4.5 |

ทำการแสกนเพื่อดึงรีพอร์ตผลการแสกนช่องโหว่ จาก OWASP ZAP ที่แสดงการจัดหมวดหมู่ตามระดับความรุนแรงของช่องโหว่ เพื่อทำการแก้ไขในขั้นตอนต่อไป



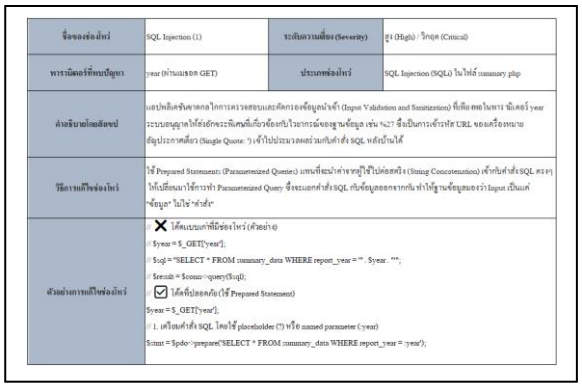
Confidence

| Risk | User Confirmed | High | Medium | Low | Total |
|---|---|---|---|---|---|
| High | 0 (0.0%) | 0 (0.0%) | 0 (0.0%) | 1 (5.9%) | 1 (5.9%) |
| Medium | 0 (0.0%) | 2 (11.8%) | 3 (17.6%) | 0 (0.0%) | 5 (29.4%) |
| Low | 0 (0.0%) | 2 (11.8%) | 4 (23.5%) | 0 (0.0%) | 6 (35.3%) |
| Informational | 0 (0.0%) | 0 (0.0%) | 4 (23.5%) | 1 (5.9%) | 5 (29.4%) |
| Total | 0 (0.0%) | 4 (23.5%) | 11 (64.7%) | 2 (11.8%) | 17 (100%) |

**ภาพที่ 4:** Report ผลการแสกนก่อนได้รับการแก้ไขช่องโหว่ จาก OWASP ZAP

**4.2 ผลการวิเคราะห์สาเหตุและแนวทางการแก้ไข**

มุ่งเน้นการเปลี่ยนข้อมูลเชิงเทคนิคจากรายงานการสแกน ให้กลายเป็นแนวทางปฏิบัติที่เป็นรูปธรรม วิเคราะห์สาเหตุเชิงลึก จัดทำเป็นคู่มือแนวทางการแก้ไข ที่ครอบคลุมทั้งหลักการทางทฤษฎีและตัวอย่างการปรับปรุง โดยมีเป้าหมายเพื่อสื่อสารองค์ความรู้ด้านความปลอดภัยไปยังทีมพัฒนาซอฟต์แวร์ได้อย่างแม่นยำและลดข้อผิดพลาด อ้างอิงรายละเอียดดังนี้ 1) ชื่อของช่องโหว่ 2) ระดับความรุนแรง จัดตามผลจากเครื่องมือ เพื่อลำดับความสำคัญ 3) ตำแหน่งที่พบปัญหา ที่เครื่องมือสแกนตรวจพบ 4) ประเภทของช่องโหว่ 5) คำอธิบายโดยสังเขป อธิบายลักษณะและกลไกของช่องโหว่นั้นให้เข้าใจง่าย 6) แนวทางการแก้ไข อธิบายวิธีแก้ปัญหาในระดับตรรกะและโครงสร้าง 7) ตัวอย่างการแก้ไข แสดงตัวอย่างซอร์สโค้ดแบบ ก่อนและหลังแก้ไขเพื่อให้เห็นภาพชัดเจน



| | | | |
|---|---|---|---|
| ชื่อของช่องโหว่ | SQL Injection (1) | ระดับความเสี่ยง (Severity) | สูง (High) / วิกฤต (Critical) |
| พารามิเตอร์ที่พบปัญหา | year (ค่าจากเมธอด GET) | ประเภทช่องโหว่ | SQL Injection (SQLi) ในไฟล์ summary.php |
| คำอธิบายโดยสังเขป | แอปพลิเคชันขาดกลไกการตรวจสอบและคัดกรองข้อมูลนำเข้า (Input Validation and Sanitization) ที่เพียงพอในพารา มิเตอร์ year ระบบยอมรับค่าใดๆ ที่ส่งเข้ามาเพื่อใช้ค้นหาข้อมูล เช่น %27 ซึ่งเป็นการเข้ารหัส URL ของเครื่องหมายอัญประกาศเดี่ยว (Single Quote: ') เข้าไปปะปนรวมกับคำสั่ง SQL หลังบ้านได้ | | |
| วิธีการแก้ไขช่องโหว่ | ใช้ Prepared Statements (Parameterized Queries) แทนที่จะนำค่าจากผู้ใช้ไปต่อสตริง (String Concatenation) เข้ากับคำสั่ง SQL ตรงๆ ให้เปลี่ยนมาใช้การทำ Parameterized Query ซึ่งจะแยกคำสั่ง SQL กับข้อมูลออกจากกัน ทำให้ฐานข้อมูลมองว่า Input เป็นแค่ "ข้อมูล" ไม่ใช่ "คำสั่ง" | | |
| ตัวอย่างการแก้ไขช่องโหว่ | ❌ โค้ดแบบเก่าที่มีช่องโหว่ (ตัวอย่าง)<br>// $year = $_GET['year'];<br>// $sql = "SELECT * FROM summary_data WHERE report_year = '" . $year . "'";<br>// $result = $conn->query($sql);<br>✅ โค้ดที่ปลอดภัย (ใช้ Prepared Statement)<br>$year = $_GET['year'];<br>// 1. เตรียมคำสั่ง SQL โดยใช้ placeholder (?) หรือ named parameter (:year)<br>$stmt = $pdo->prepare('SELECT * FROM summary_data WHERE report_year = :year'); | | |

**ภาพที่ 5:** ตัวอย่างคู่มือแนะนำการแก้ไขช่องโหว่

**ตารางที่ 2:** ผลการเปรียบเทียบจำนวนช่องโหว่ก่อนและหลังการปรับปรุง

| ระดับความเสี่ยง (Severity) | จำนวนก่อนแก้ไข | จำนวนหลังแก้ไข | ผลลัพธ์ |
|---|---|---|---|
| สูง(High/Critical) | 1 | 0 | ลดลง 100% |
| ปานกลาง )Medium) | 5 | 0 | ลดลง 100% |
| ต่ำ (Low) | 6 | 5 | คงเดิม |
| แจ้งเพื่อทราบ )Informational) | 5 | 4 | ลดลง 20% |
| รวม | 17 | 9 | ลดลง 8 รายการ |

จากตารางข้างต้นอัตราการลดลงในภาพรวมจะอยู่ที่ 47.06% แต่หากพิจารณาเฉพาะช่องโหว่ที่มีผลกระทบต่อระบบอย่างรุนแรง พบว่ามีอัตราการลดลงถึง 100% ซึ่งถือว่าบรรลุเป้าหมายหลักในการป้องกันการโจมตีที่สำคัญ

**4.3 ผลการประเมินคุณภาพและความเหมาะสมของคู่มือแนวทางการแก้ไขช่องโหว่**

**ตารางที่ 3:** ผลคะแนนเฉลี่ย

| รายการประเมิน | ค่าเฉลี่ย | ระดับความคิดเห็น |
|---|---|---|
| ความชัดเจนของเนื้อหา | 4.5 | มากที่สุด |
| ลำดับขั้นตอนการปฏิบัติ | 4.5 | มากที่สุด |
| ความถูกต้องของตัวอย่างโค้ดและวิธีแก้ | 4 | มาก |
| ความสามารถในการลดเวลาการทำงาน | 5 | มากที่สุด |
| รูปแบบการนำเสนอและการค้นหาข้อมูล | 4.75 | มากที่สุด |
| ประโยชน์ในภาพรวมต่อองค์กร | 4.5 | มากที่สุด |
| รวม | 4.54 | มากที่สุด |

ในภาพรวม ผลการประเมินคุณภาพคู่มือแนวทางการแก้ไขช่องโหว่มีค่าเฉลี่ยรวมเท่ากับ 4.54 ซึ่งจัดอยู่ในระดับ "มากที่สุด" สถิตินี้เป็นเครื่องยืนยันว่างานค้นคว้าอิสระฉบับนี้ ประสบความสำเร็จในการสร้างองค์ความรู้ที่ใช้งานได้จริง โดยตอบโจทย์เกณฑ์การประเมินประสิทธิภาพของเครื่องมือใน 4 มิติหลัก ดังนี้ มิติที่ 1 ด้านการบริหารจัดการและลิขสิทธิ์ หัวข้อระดับความเสี่ยงและคำอธิบายเครื่องมือที่มีการอัปเดตอย่างต่อเนื่องจะให้ฐานข้อมูลช่องโหว่ที่ทันสมัย คู่มือจึงมีการนำค่าจากรีพอร์ตเหล่านั้นมาจัดลำดับความสำคัญเพื่อให้องค์กรบริหารจัดการทรัพยากรในการแก้ไขได้อย่างคุ้มค่าตามสิทธิ์การใช้งาน มิติที่ 2 ด้านความสามารถทางเทคนิค หัวข้อพารามิเตอร์ที่พบปัญหาและหัวข้อประเภทช่องโหว่ ช่องโหว่ที่ถูกตรวจพบด้วยวิธีอัตโนมัติ ไม่ว่าจะเป็นเทคโนโลยีใหม่อย่าง Modern Web หรือไม่ คู่มือจะทำหน้าที่ระบุโดยใช้ข้อมูลจากเครื่องมือมาเขียนในหัวข้อ "พารามิเตอร์ที่พบปัญหา" เพื่อให้นักพัฒนารู้ทันทีว่าต้องไปแก้ไขที่จุดใดของแอปพลิเคชัน มิติที่ 3 ด้านประสิทธิภาพเชิงลึก หัวข้อชื่อช่องโหว่และวิธีการแก้ไขเมื่อเครื่องมือมีความครอบคลุมช่องโหว่สูง คู่มือจึงถูกออกแบบมาให้รองรับข้อมูลเหล่านั้นโดยระบุ "ชื่อช่องโหว่" ตามมาตรฐานสากล เช่น OWASP และเสนอ วิธีการแก้ไขเชิงโครงสร้างที่สามารถปิดช่องโหว่ที่ซับซ้อนที่ตรวจเจอผ่านการขยายขีดความสามารถของเครื่องมือ มิติที่ 4 ด้านการนำไปใช้งานและการบูรณาการ การเชื่อมโยงกับคู่มือ หัวข้อตัวอย่างการแก้ไข นี่คือจุดเชื่อมโยงที่สำคัญที่สุด โดยคู่มือจะเปลี่ยนข้อมูลจาก Reporting ให้กลายเป็น คู่มือปฏิบัติงานที่มี ตัวอย่างการแก้ไข ซึ่งจำเป็นอย่างยิ่งในการบูรณาการเข้ากับระบบ CI/CD เพื่อให้นักพัฒนาสามารถแก้ไข Code ได้ทันทีในกระบวนการพัฒนา

## 5. สรุป

จากการดำเนินงานวิจัยนี้ ผู้วิจัยสามารถสรุปผลสัมฤทธิ์ตามวัตถุประสงค์ที่ตั้งไว้ใน 3 มิติหลัก ดังนี้

**5.1 ความสำเร็จด้านประสิทธิภาพในการคัดเลือกเครื่องมือ** ผู้วิจัยพิสูจน์ให้เห็นว่า OWASP ZAP เป็น

เครื่องมือที่มีความเหมาะสมที่สุดสำหรับการใช้งานในองค์กรกรณีศึกษา โดยมีดัชนีชี้วัดความสำเร็จที่ชัดเจน คือ ความคุ้มค่าด้านเวลา ความแม่นยำและความลึก

**5.2 ความสำเร็จด้านการบริหารจัดการความเสี่ยง** ผลการตรวจสอบยืนยันภายหลังการแก้ไข แสดงให้เห็นถึงการยกระดับความมั่นคงปลอดภัยของระบบอย่างเป็นรูปธรรม สามารถปิดช่องโหว่ระดับ สูง และ ปานกลาง ได้อย่างสมบูรณ์ แม้จะมีช่องโหว่คงเหลือ 9 แต่ผลการวิจัยชี้ให้เห็นชัดเจนว่าเป็นความเสี่ยงในระดับที่ยอมรับได้ และเกี่ยวข้องกับการตั้งค่าระดับโครงสร้างพื้นฐาน ซึ่งได้ถูกนำเสนอเป็นแนวทางในการพัฒนาลำดับต่อไป

**5.3 ความสำเร็จด้านการสร้างองค์ความรู้และมาตรฐานการปฏิบัติงาน** ผลผลิตที่มีค่าที่สุดของการวิจัยนี้คือ "คู่มือการแก้ไขช่องโหว่ (Remediation Manual)" ที่ผู้วิจัยสร้างขึ้น สามารถเปลี่ยนรายงานสแกนที่ซับซ้อนให้กลายเป็นขั้นตอนการแก้ไขที่นักพัฒนานำไปใช้ได้จริง เป็นฐานข้อมูลความรู้ องค์กรสามารถนำไปใช้เป็นมาตรฐานในการเขียนโค้ดที่ปลอดภัย เพื่อป้องกันการเกิดช่องโหว่ซ้ำในอนาคตได้ และการยืนยันคุณภาพและความน่าเชื่อถือจากมุมมองผู้ใช้งาน คุณภาพของคู่มือได้รับการพิสูจน์ผ่านกระบวนการประเมินจากผู้เชี่ยวชาญจำนวน 4 ท่าน ซึ่งให้คะแนนความพึงพอใจเฉลี่ยสูงถึง 4.54 (มากที่สุด) โดยเฉพาะอย่างยิ่งในด้านความสามารถในการลดระยะเวลาการทำงานที่ได้รับคะแนนเต็ม 5.00 ผลลัพธ์นี้เป็นเครื่องยืนยันเชิงประจักษ์ว่าคู่มือที่ผู้วิจัยพัฒนาขึ้นมีความสมบูรณ์ทั้งในด้านความถูกต้องเชิงเทคนิคและความสะดวกในการนำไปปฏิบัติจริง

## เอกสารอ้างอิง


[1] Fortinet. (2024). 2024 Global Threat Landscape Report from FortiGuard Labs [Online]. Available: https://www.fortinet.com/resources/reports/threat-landscape-report.

[2] IBM. (2025). Cost of a Data Breach Report 2025.IBM.[Online].Available: https://www.ibm.com/reports/data-breach

[3] “Top 10 Web Application Security Risks” Open Worldwide Application Security Project Blog. Accessed: Sept, 2025. [Online]. Available: https://owasp.org/www-project-top-ten/

[4] Belay, T. E., Gupta, S., & Kumar, A. (2025). Systematic Evaluation And Model Based Selection Of Web Vulnerability Scanners: Toward A Prior Guided Assessment Framework. International Journal of Environmental Sciences.

[5] Escarrone, T., Rosa, R. L., Kreutz, D., et al. (2025). Sob a Lupa Corporativa Brasileira: Avaliação da Cobertura de Scanners de Vulnerabilidades em Aplicações Atuais.

[6] Potti, U.-S., Huang, H.-S., Chen, H.-T., et al. (2025). Security Testing Framework for Web Applications: Benchmarking ZAP V2.12.0 and V2.13.0 by OWASP as an example.

[7] Jha,T.K.(2024).Modern Web Application Penetration Testing Tools: A Comparative Analysis. Deleted Journal.

[8] Kondraciuk, A., Bartos, A., & Pańczyk, B. (2022)Comparative analysis of the effectiveness of OWASP ZAP, Burp Suite, Nikto and Skipfish in testing the security of web applications. Journal of Computer Sciences Institute.

[9] Mburano, B., & Si, W. (2018). Evaluation of Web VulnerabilityScannersBased on OWASP Benchmark. 2018 IEEE International Conference on Systems Engineering.

[10] Alsaleh, M., Alomar, N., Alshreef, M., et al. (2017). Performance-Based Comparative Assessment of Open-Source Web Vulnerability Scanners. Security and Communication Networks.

[11] NIST SP 800-115 (2008) Technical Guide to Information Security Testing and Assessment. NIST.[Online].Available: https://www.ibm.com/reports/data-breach

# ระบบติดตามสุขภาพผู้สูงอายุในบ้านโดยใช้ IoT ร่วมกับระบบแจ้งเตือนผู้ดูแลผ่านแอปพลิเคชัน

# Home-based Elderly Health Monitoring System Using IoT with Caregiver Alert Application

*ปนัสยา ทองดี (Panatsaya Thongdee)[1] และธนพล เจนสุทธิเวชกุล (Tanapon Jensuttiwetchakul)[2]*

*ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล*

*มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ*

*[1]s6707021856051@email.kmutnb.ac.th, [2]tanapon.j@itd.kmutnb.ac.th*

## บทคัดย่อ

*งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาระบบติดตามสุขภาพเชิงรุกสำหรับผู้สูงอายุโดยประยุกต์ใช้เทคโนโลยีอินเทอร์เน็ตของสรรพสิ่ง (IoT) ร่วมกับแอปพลิเคชันบนสมาร์ตโฟน โดยใช้สมาร์ตวอทช์เป็นอุปกรณ์หลักในการตรวจวัดสัญญาณชีพ อัตราการเต้นของหัวใจ, ระดับออกซิเจนในเลือด, ความเครียด และคุณภาพการนอนหลับ ข้อมูลถูกส่งผ่านระบบคลาวด์เพื่อประมวลผลและแจ้งเตือนเหตุผิดปกติในลักษณะใกล้เคียงเวลาจริง (Near Real-time) ไปยังผู้ดูแลผ่านระบบ Android ผลการทดสอบระบบด้วยวิธีการทดสอบแบบกล่องดำ (Black-box Testing) และการทดสอบการยอมรับจากผู้ใช้งาน (UAT) พบว่าระบบมีความถูกต้องแม่นยำในการทำงานร้อยละ 98.4 โดยผลการประเมินความพึงพอใจจากกลุ่มตัวอย่างผู้ดูแล 10 ราย ในระยะเวลา 1 สัปดาห์ พบว่าผู้ใช้งานมีความพึงพอใจด้านทัศนคติต่อการใช้งานและประโยชน์ที่ได้รับในระดับมากที่สุด ($\bar{x}$ = 4.68 และ 4.53 ตามลำดับ) ขณะที่ด้านการใช้งานแอปพลิเคชันอยู่ในระดับมาก ($\bar{x}$ = 4.28) ระบบต้นแบบนี้จึงมีประสิทธิภาพในการลดภาระการเฝ้าระวังและเพิ่มความปลอดภัยในการดูแลผู้สูงอายุที่บ้านได้อย่างมีนัยสำคัญ*

**คำสำคัญ:** อินเทอร์เน็ตสรรพสิ่ง (IoT), สมาร์ตวอทช์, การดูแลผู้สูงอายุ, การติดตามสุขภาพระยะไกล, แสดงผล, การแจ้งเตือนแบบพุช

## Abstract

*This research aims to develop a proactive health monitoring system for the elderly by integrating Internet of Things (IoT) technology with a mobile application. Smartwatches were employed as the primary sensing devices to monitor vital signs, including heart rate, blood oxygen levels, stress levels, and sleep quality. The collected data is transmitted via a cloud system for Near Real-time processing and anomaly detection, with push notifications sent to caregivers through a custom-developed Android application. System performance was evaluated using Black-box testing and User Acceptance Testing (UAT). The results demonstrated a high functional accuracy of 98.4 %. Furthermore, a satisfaction assessment conducted with 10 caregivers over a one-week pilot period revealed that users were highly satisfied with the system's overall attitude and perceived benefits ($\bar{x}$ = 4.68 and 4.53, respectively), while the application's usability was rated at a high level ($\bar{x}$ = 4.28). This prototype system demonstrates significant potential in reducing caregiving burdens and enhancing the safety of elderly individuals living at home.*

**Keywords:** Internet of Things (IoT), Smartwatch, Elderly care, Remote health monitoring, Visualization, Push notifications

## 1. บทนำ

ประเทศไทยกำลังเผชิญกับการเปลี่ยนแปลงทางโครงสร้างประชากรครั้งสำคัญ โดยคาดว่าจะก้าวเข้าสู่สังคมสูงวัยระดับสุดยอด (Super-Aged Society) อย่างเต็มรูปแบบภายในปี ค.ศ. 2030 สถานการณ์ดังกล่าวส่งผลให้รูปแบบการอยู่อาศัยเปลี่ยนไป ผู้สูงอายุจำนวนมากต้องใช้ชีวิตอยู่ตามลำพังและขาดการเฝ้าระวังด้านสุขภาพอย่างต่อเนื่อง ในขณะเดียวกัน ระบบการดูแลที่ใช้กันอยู่ทั่วไปยังคงพึ่งพาการสังเกตด้วยสายตาของผู้ดูแลเป็นหลัก ซึ่งวิธีการนี้มีข้อจำกัดอย่างมากทั้งในแง่ของเวลา ความต่อเนื่องในการติดตามอาการ และความแม่นยำของการประเมินความเสี่ยงเมื่อเกิดเหตุฉุกเฉิน แม้ในปัจจุบันจะมีการนำระบบติดตามอาการผู้ป่วยทางไกล (Remote Patient Monitoring: RPM) มาใช้งานบ้างแล้ว แต่งานวิจัยและเทคโนโลยีส่วนใหญ่ยังคงมุ่งเน้นไปที่การใช้งานกับผู้ป่วยในโรงพยาบาลหรือในบริบทของคลินิก จึงทำให้เกิดช่องว่างสำคัญในการดูแลสุขภาพระยะยาว นั่นคือการขาดแคลนระบบที่ถูกออกแบบมาอย่างเจาะจงสำหรับผู้สูงอายุ

ที่พักอาศัยอยู่ที่บ้าน ซึ่งต้องการเทคโนโลยีที่ใช้งานง่าย มีต้นทุนต่ำ และสามารถเชื่อมโยงข้อมูลสุขภาพส่งตรงถึงผู้ดูแลได้แบบออนไลน์ เพื่อปิดช่องว่างดังกล่าว งานวิจัยนี้จึงมีวัตถุประสงค์เพื่อพัฒนาระบบติดตามสุขภาพเชิงรุกสำหรับผู้สูงอายุในที่พักอาศัย โดยบูรณาการการทำงานร่วมกันระหว่างนาฬิกาอัจฉริยะ (Smartwatch) เทคโนโลยีอินเทอร์เน็ตของสรรพสิ่ง (IoT) และแอปพลิเคชันแจ้งเตือน พร้อมทั้งประเมินประสิทธิภาพการทำงานและระดับความพึงพอใจของผู้ดูแล ผลลัพธ์จากการศึกษาครั้งนี้มุ่งหวังที่จะช่วยลดภาระความเครียดของผู้ดูแล ยกระดับความปลอดภัย และสนับสนุนแนวคิดการดูแลผู้สูงอายุให้สามารถใช้ชีวิตในบ้านของตนเองได้อย่างมีคุณภาพและยั่งยืน (Aging in Place) อย่างเป็นรูปธรรม

## 2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

### 2.1 งานวิจัยที่พัฒนาระบบ IoT สำหรับผู้สูงอายุ (System Development Works)

จากงานวิจัยเรื่อง IoT-Based Elderly Health Monitoring System Using Firebase Cloud Computing นำเสนอแนวทางการพัฒนาระบบติดตามสุขภาพที่อาศัยเซนเซอร์แบบสวมใส่ (Wearable sensors) เพื่อเก็บข้อมูลสัญญาณชีพ เช่น อัตราการเต้นของหัวใจและระดับออกซิเจนในเลือด ข้อมูลจะถูกส่งผ่านเกตเวย์ไปยังระบบคลาวด์เพื่อประมวลผลแบบเวลาจริงและแจ้งเตือนเมื่อพบความผิดปกติ บางระบบมีการบูรณาการการเรียนรู้ของเครื่อง (Machine Learning) เพื่อประเมินสถานะสุขภาพ โครงสร้างระบบส่วนใหญ่ทำงานบนสถาปัตยกรรม IoT แบบ 3 ระดับชั้น คือ ชั้นกายภาพ ชั้นเครือข่าย และชั้นแอปพลิเคชัน ซึ่งช่วยให้ระบบมีความยืดหยุ่น รองรับการขยายตัว และมีความปลอดภัยในการใช้งาน

### 2.2 การพัฒนาแอปพลิเคชันมือถือ

1) แบบจำลองการสร้างต้นแบบต้องอาศัยกระบวนการพัฒนาและทดสอบแบบวนซ้ำ โดยเปิดโอกาสให้กลุ่มผู้สูงอายุและผู้ดูแลเข้ามามีส่วนร่วมในการประเมินผล เพื่อลดความคลาดเคลื่อนของการออกแบบและให้ระบบตอบสนองต่อข้อจำกัดทางกายภาพของผู้ใช้งานได้อย่างแม่นยำซึ่งได้กล่าวในงานวิจัย Human-Centered Design of Mobile Health Apps for Older Adults: Systematic Review and Narrative Synthesis 2) การพัฒนาบนแพลตฟอร์มแอนดรอยด์ เนื่องจากเป็นแพลตฟอร์มที่มีสัดส่วนผู้ใช้งานสูงสุดในประเทศไทยและอุปกรณ์มีราคาที่เข้าถึงง่าย โดยประยุกต์ใช้เครื่องมือ Android Studio (ที่มาจาก Statcounter) และเฟรมเวิร์ก .NET MAUI เพื่อสนับสนุนการทำงานข้ามแพลตฟอร์มและการจัดการระบบแจ้งเตือน 3) การออกแบบ UI/UX สำหรับผู้สูงอายุจากงานวิจัย Design Guidelines of Mobile Apps for Older Adults: Systematic Review and Thematic Analysis จะเป็นการมุ่งเน้นการลดภาระการรับรู้ข้อมูล โดยใช้ตัวอักษรแบบไม่มีเชิงขนาดใหญ่ที่ระบบสามารถปรับสัดส่วนได้อัตโนมัติ ใช้คู่สีที่มีความเปรียบต่างสูง ขยายพื้นที่ตอบสนองของปุ่มกดเพื่อลดความผิดพลาดในการสัมผัส ตลอดจนออกแบบโครงสร้างการนำทางให้กระชับและแสดงผลเพียงภารกิจเดียวต่อหนึ่งหน้าจอ

### 2.3 ทฤษฎีพฤติกรรมการตอบสนองต่อการแจ้งเตือน (Notification Response Theory)

การพัฒนาระบบแจ้งเตือนทางสุขภาพที่มีประสิทธิภาพ ต้องอาศัยความเข้าใจกลไกการตอบสนองต่อสิ่งเร้าเพื่อป้องกันความเหนื่อยล้าและการเพิกเฉยต่อข้อมูลสำคัญ โดยมีหลักการบูรณาการ 4 ประการ ได้แก่ 1) การบริหารความสนใจ ลดสภาวะข้อมูลล้นและภาระทางปัญญา โดยคัดกรองส่งเฉพาะข้อมูลวิกฤต ใช้ข้อความกระชับ และหลีกเลี่ยงการแจ้งเตือนในเวลาที่ไม่เหมาะสม 2) การสื่อสารเชิงรุก ประยุกต์ใช้กลไกการส่งข้อมูลแบบผลัก (Push Model) ผสานกับจิตวิทยาการโน้มน้าวใจ เพื่อส่งมอบข้อมูลที่ตรงกับเป้าหมายและกระตุ้นให้ผู้ดูแลตอบสนองต่อเหตุฉุกเฉินทันที 3) เสถียรภาพของระบบอาศัยสถาปัตยกรรมเชิงวิศวกรรมแบบขับเคลื่อนด้วยเหตุการณ์ (Event-Driven) ที่มีระบบจัดการคิวและข้อความตกหล่น เพื่อให้การแจ้งเตือนทำงานได้แม่นยำแม้เครือข่ายไม่เสถียร 4) การออกแบบปฏิสัมพันธ์ เน้นความเรียบง่ายตามข้อจำกัดทางกายภาพของผู้สูงอายุ สื่อสารตรงประเด็น และเปิดสิทธิ์ให้ผู้ใช้งานสามารถปรับแต่งเงื่อนไขหรือช่วงเวลาห้ามรบกวนได้ด้วยตนเอง [1]

### 2.4 ทฤษฎีการยอมรับเทคโนโลยี (Technology Acceptance Theories)

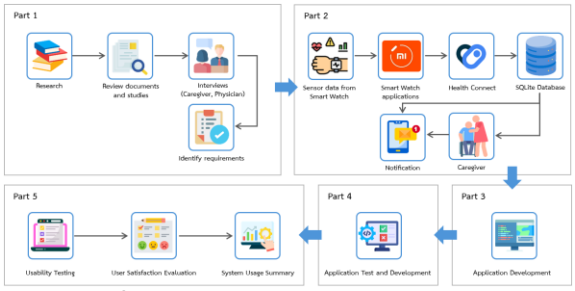
ความสำเร็จของระบบสารสนเทศขึ้นอยู่กับระดับการยอมรับเทคโนโลยีของผู้ใช้งาน ซึ่งเป็นปัจจัยชี้วัดการนำระบบไปใช้งานจริงอย่างยั่งยืน โดยมีทฤษฎีและแบบจำลองที่ใช้เป็นกรอบแนวคิดหลัก 3 แนวทาง ได้แก่ 1) แบบจำลองการยอมรับเทคโนโลยี Technology Acceptance Model (TAM) จากวิทยานิพนธ์ A technology acceptance model for empirically testing new end-user information systems: Theory and results อธิบายว่าพฤติกรรมการยอมรับเทคโนโลยีเกิดขึ้นจากการประเมินปัจจัยหลัก 2 ด้าน คือ การรับรู้ถึงประโยชน์ที่จะได้รับ และการรับรู้ถึงความง่ายในการใช้งานระบบ 2) ทฤษฎีรวมการยอมรับและการใช้เทคโนโลยี Unified Theory of Acceptance and Use of Technology (UTAUT) จากงานวิจัย User Acceptance of Information Technology: Toward a Unified View ของ Venkatesh และคณะ ได้ขยายขอบเขตปัจจัยที่มีอิทธิพลต่อผู้ใช้งาน โดยครอบคลุม

เพิ่มเติมในส่วนของแรงผลักดันทางสังคมและความพร้อมของโครงสร้างพื้นฐานที่เกื้อหนุนต่อการใช้งาน 3) ทฤษฎีการแพร่กระจายนวัตกรรม Diffusion of Innovations (DOI) ในหนังสือ Diffusion of Innovation ของ Everett M. Rogers ชี้ให้เห็นว่าอัตราความเร็วในการยอมรับนวัตกรรมแปรผันตามคุณลักษณะ 5 ประการ ได้แก่ ความได้เปรียบกว่าแนวทางเดิม ความสอดคล้องกับค่านิยมผู้ใช้ ระดับความซับซ้อน โอกาสในการทดลองปฏิบัติจริง และความสามารถในการสังเกตเห็นผลลัพธ์ได้อย่างเป็นรูปธรรม

## 3. วิธีการดำเนินการวิจัย

### 3.1 กรอบแนวความคิดของงานวิจัย

ผู้วิจัยได้ทำการศึกษาค้นคว้า ทฤษฎี แนวคิด และงานวิจัยที่เกี่ยวข้องต่าง ๆ จึงได้นำมาประยุกต์ใช้ในงานวิจัย โดยแบ่งออกเป็น 5 ส่วนการทำงาน โดยมีกรอบแนวคิดแสดงดังภาพที่ 1




**ภาพที่ 1:** กรอบแนวความคิดการพัฒนาระบบ

การวิจัยนี้ขับเคลื่อนด้วยกระบวนการออกแบบที่เน้นมนุษย์เป็นศูนย์กลาง Human-Centered Design (HCD) เป็นแนวทางหลักในการพัฒนาระบบ โดยเริ่มต้นจากการทำความเข้าใจผู้ใช้งานอย่างลึกซึ้งผ่านการศึกษาทฤษฎีที่เกี่ยวข้อง อาทิ เทคโนโลยี IoT, Smart Health และหลักการออกแบบ UX/UI ควบคู่ไปกับการสัมภาษณ์เชิงลึก (In-depth Interview) กับกลุ่มเป้าหมายหลัก ซึ่งประกอบด้วยผู้ดูแล (Caregivers) และแพทย์เฉพาะทาง เพื่อระบุปัญหาและข้อกำหนดของระบบ (System Requirements) ให้สอดคล้องกับสรีรวิทยาและพฤติกรรมของผู้สูงอายุในไทยอย่างแท้จริง จากนั้นการพัฒนาระบบจะดำเนินการตามรูปแบบการสร้างต้นแบบ (Prototyping Model) โดยนำข้อกำหนดที่ได้มาพัฒนาเป็นระบบจำลองเบื้องต้น และนำไปทดสอบเพื่อรับฟังข้อเสนอแนะ (Feedback) จากผู้ใช้จริง ทั้งกลุ่มผู้ดูแลและแพทย์ การนำข้อเสนอแนะดังกล่าวมาวิเคราะห์และปรับปรุงต้นแบบซ้ำอย่างต่อเนื่องเป็นวงรอบ (Iterative Process) จะช่วยให้มั่นใจได้ว่าระบบที่พัฒนาขึ้นขั้นสุดท้ายนั้น สามารถตอบสนองความต้องการลดข้อจำกัดในการใช้งาน และยกระดับประสิทธิภาพในการดูแลผู้สูงอายุได้อย่างเป็นรูปธรรม

### 3.2 ขั้นตอนการรวบรวม และเตรียมข้อมูล

ผู้วิจัยเริ่มต้นจากการสัมภาษณ์เชิงลึก (In-depth Interview) กับแพทย์เฉพาะทางจำนวน 7 ท่าน เพื่อค้นหาปัญหาจริงที่พบในกระบวนการดูแลผู้สูงอายุ พร้อมทั้งระบุข้อกำหนดของระบบ (System Requirements) ให้สอดคล้องกับบริบทการใช้งาน จากข้อค้นพบดังกล่าว ผู้วิจัยจึงได้กำหนดพารามิเตอร์ทางสุขภาพหลักที่สามารถติดตามได้ผ่านสมาร์ตวอทช์ ประกอบด้วย อัตราการเต้นของหัวใจ (Heart Rate), ระดับออกซิเจนในเลือด (SpO2), ระดับความเครียด (Stress), คุณภาพการนอนหลับ (Sleep), จำนวนก้าว (Steps) และค่า VO2 Max เพื่อตรวจสอบความเที่ยงตรงเชิงเนื้อหาและร่วมกันพิจารณากำหนดจุดตัดเกณฑ์ตัดสิน (Cut-off points) สำหรับจำแนกความเสี่ยงทางสุขภาพของผู้สูงอายุออกเป็น 3 ระดับ ได้แก่ ระดับต่ำ ระดับปานกลาง และระดับสูง อย่างมีมาตรฐานทางการแพทย์ ดังแสดงในตารางที่ 1

**ตารางที่ 1:** เกณฑ์การประเมินการเกิดโรค/ภาวะ โดยแบ่งออกเป็น 3 ระดับ

| โรคหัวใจและหลอดเลือด | | | |
|---|---|---|---|
| **ค่าจากนาฬิกา** | **Normal** | **Medium** | **High** |
| Steps | ≥7,000 ก้าว ≥5 วัน/สัปดาห์ | 3,000–6,999 ก้าว ≥3 วัน/สัปดาห์ | <3,000 ก้าว ≥3 วันติดต่อกัน |
| $VO_2$ Max | - เพศชาย ≥31 ml/kg/min<br>- เพศหญิง ≥27 ml/kg/min | ต่ำกว่ามาตรฐาน ≥7 วัน<br>- เพศชาย 22 – 30 ml/kg/min<br>- เพศหญิง 19 – 26 ml/kg/min | ต่ำกว่ามาตรฐาน >30% ≥14 วัน<br>- เพศชาย <22 ml/kg/min<br>- เพศหญิง <19 ml/kg/min |
| Heart Rate | 60–80 bpm ต่อเนื่อง ≥5 วัน | >85 bpm ต่อเนื่อง ≥3 วัน | >100 bpm หรือ <45 bpm ≥2 วัน |
| **โรคเบาหวาน** | | | |
| **ค่าจากนาฬิกา** | **Low** | **Medium** | **High** |
| Steps | ≥7,000 ก้าว ≥5 วัน/สัปดาห์ | 3,000–6,999 ก้าว ≥3 วัน | <3,000 ก้าว ≥3 วันติดต่อกัน |
| Heart Rate | 60–80 bpm | >85 bpm ≥3 วัน | >95 bpm ≥3 วัน |
| Sleep | นอน ≥7 ชม./คืน | นอน <6 ชม. ≥2–3 วัน | นอน <5 ชม. ≥2 วัน |
| **ภาวะความเครียดเรื้อรัง** | | | |
| **ค่าจากนาฬิกา** | **Low** | **Medium** | **High** |
| Heart Rate | 60–80 bpm | >85 bpm ขณะพัก ≥3 วัน | >95 bpm ≥3 วัน + ใจสั่น/นอนไม่หลับ |
| Stress | ต่ำ–กลาง สลับปกติ คะแนนอยู่ระหว่าง 0 – 39 ใน ≥70% ของวัน | สูงต่อเนื่อง ≥3 วัน คะแนนอยู่ระหว่าง 40 – 69 | สูงมาก ≥ 3 วัน โดยไม่มีแนวโน้มลดลง คะแนนอยู่ระหว่าง 70-100 |
| Sleep | นอน ≥7 ชม., หลับลึกเพียงพอ | นอน <6 ชม. ≥3 วัน | นอน <5 ชม. ≥2 วัน + ตื่นบ่อย |

| ภาวะนอนไม่เพียงพอ / คุณภาพการนอนต่ำ | | | |
|---|---|---|---|
| ค่าจากนาฬิกา | Low | Medium | High |
| Stress | ต่ำ–กลาง สลับปกติ คะแนนอยู่ระหว่าง 0 – 39 ใน ≥70% ของวัน | สูงต่อเนื่อง ≥3 วัน คะแนนอยู่ระหว่าง 40 – 69 | สูงมาก ≥ 3 วัน โดยไม่มีแนวโน้มลดลง คะแนนอยู่ระหว่าง 70-100 |
| Sleep | นอน ≥7 ชม./คืน ≥5 วัน | นอน 5–6 ชม. ≥3 วัน | นอน <5 ชม. ≥2 คืนติดต่อกัน |
| $SpO_2$ | นอน ≥95% | นอน 90–94% ≥2 คืน | นอน <90% ≥2 คืน + ง่วงกลางวัน |
| ภาวะขาดการออกกำลังกาย | | | |
| ค่าจากนาฬิกา | Low | Medium | High |
| Steps | ≥7,000 ก้าว/วัน ≥5 วัน/สัปดาห์ | 3,000–6,999 ก้าว ≥3 วัน | <3,000 ก้าว ≥3 วันติดต่อกัน |
| Calories | เผาผลาญเฉลี่ย >200 kcal/วัน | 100–200 kcal ≥3 วัน | <100 kcal ≥3 วันต่อเนื่อง |
| $VO_2$ Max | - เพศชาย ≥31 ml/kg/min<br>- เพศหญิง ≥27 ml/kg/min | ต่ำกว่ามาตรฐาน ≥7 วัน<br>- เพศชาย 22 – 30 ml/kg/min<br>- เพศหญิง 19 – 26 ml/kg/min | ต่ำกว่ามาตรฐาน >30% ≥14 วัน<br>- เพศชาย <22 ml/kg/min<br>- เพศหญิง <19 ml/kg/min |

### 3.3 การออกแบบและพัฒนาระบบ

การออกแบบระบบอ้างอิงสถาปัตยกรรมอินเทอร์เน็ตของสรรพสิ่ง 3 ชั้น [3] โดยใช้แผนภาพบริบทและแผนภาพกระแสข้อมูลกำหนดเส้นทางการแลกเปลี่ยนข้อมูล กระบวนการเริ่มต้นที่ชั้นรับรู้ซึ่งใช้สมาร์ตวอทช์วัดชีวสัญญาณและส่งผ่านบลูทูธพลังงานต่ำไปยังสมาร์ตโฟนเกตเวย์ จากนั้นชั้นเครือข่ายจะส่งข้อมูลไปจัดเก็บในฐานข้อมูล SQLite เพื่อประมวลผลแบบใกล้เวลาจริง และเชื่อมโยงการแสดงผลบนแอปพลิเคชัน Android ในชั้นแอปพลิเคชัน ระบบประกอบด้วย 4 ฟังก์ชันหลัก ได้แก่ 1) การจัดการสิทธิ์เข้าถึงบัญชีอย่างปลอดภัย 2) การประมวลผลความเสี่ยงเทียบกับจุดตัดเกณฑ์มาตรฐานของแพทย์ โดยประเมินความต่อเนื่องร่วมด้วยเพื่อลดการแจ้งเตือนผิดพลาด 3) ระบบแจ้งเตือนอัตโนมัติแบบขับเคลื่อนด้วยเหตุการณ์ ซึ่งจะส่งคำแนะนำไปยังผู้ดูแลเฉพาะเมื่อพบค่าวิกฤตตามเงื่อนไขในตาราง 1 เพื่อป้องกันภาวะความเหนื่อยล้าจากการรับแจ้งเตือน และ 4) การแสดงกราฟรายงานประวัติสุขภาพเพื่อสนับสนุนการประเมินแนวโน้มในระยะยาว

### 3.4 การทดสอบระบบ

การประเมินประสิทธิภาพของระบบแบ่งออกเป็น 2 ส่วนหลัก ได้แก่ 1) การทดสอบแบบ Black-box Testing เพื่อตรวจสอบความถูกต้องของผลลัพธ์เทียบกับข้อมูลนำเข้า โดยมีผู้เชี่ยวชาญด้านระบบสารสนเทศจำนวน 3 ท่าน ทดสอบการทำงาน 4 หมวดหลัก ได้แก่ การเชื่อมต่อฮาร์ดแวร์ การแสดงผล การจัดการข้อมูลผู้ใช้ และระบบแจ้งเตือนอัตโนมัติ โดยทำการทดสอบฟังก์ชันย่อยรายการละ 10 ครั้ง และคำนวณประสิทธิภาพเป็นร้อยละความถูกต้อง ตามสมการดังนี้

$$\text{เปอร์เซ็นต์ความถูกต้อง} = \frac{A}{C} \times 100$$

(เมื่อ A คือจำนวนครั้งที่ระบบให้ผลลัพธ์ถูกต้องตามที่คาดหวัง และ C คือจำนวนครั้งของการทดสอบทั้งหมด)

2) การทดสอบการยอมรับของผู้ใช้งาน User Acceptance Testing (UAT) หลังจากระบบผ่านการทดสอบเชิงฟังก์ชัน ผู้วิจัยได้คัดเลือกกลุ่มตัวอย่างแบบเจาะจง ซึ่งเป็นผู้ดูแลผู้สูงอายุจริงจำนวน 10 คน ให้ทดลองใช้งานแอปพลิเคชันในที่พักอาศัยจริงเป็นเวลา 1 สัปดาห์ จากนั้นประเมินความพึงพอใจด้านการใช้งาน ความชัดเจนของข้อมูล และการแจ้งเตือน ผ่านแบบสอบถามมาตราส่วนประมาณค่า 5 ระดับ (Likert Scale) โดยนำผลมาวิเคราะห์ด้วยสถิติค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐาน

### 3.5 การวิเคราะห์ข้อมูลและสถิติที่ใช้ในการวิจัย

การวิเคราะห์ข้อมูลแบ่งออกเป็น 2 ส่วน ได้แก่ 1) การวิเคราะห์ผลการทดสอบความถูกต้องของระบบแบบ Black-box Testing ด้วยสถิติเชิงพรรณนา ได้แก่ ความถี่และค่าร้อยละ เพื่อประเมินสัดส่วนกรณีทดสอบที่ผ่านเกณฑ์ตามข้อกำหนดของระบบ และ 2) การวิเคราะห์ระดับความพึงพอใจของผู้ดูแลจำนวน 10 คน ผ่านแบบสอบถามมาตราส่วนประมาณค่า 5 ระดับ (Likert Scale) ซึ่งครอบคลุมด้านการทำงานของระบบ ด้านการใช้งาน และด้านประโยชน์ที่ได้รับ โดยใช้ค่าเฉลี่ย (Mean) และส่วนเบี่ยงเบนมาตรฐาน (S.D.) ทั้งนี้ ผลการวิเคราะห์จะถูกนำเสนอในรูปแบบตารางประกอบความเรียงเพื่อใช้เป็นข้อมูลพื้นฐานในการอภิปรายผลต่อไป

## 4. ผลการดำเนินงาน

### 4.1 ผลการพัฒนาระบบติดตามสุขภาพผู้สูงอายุในบ้านโดยใช้ IoT ร่วมกับระบบแจ้งเตือนผู้ดูแลผ่านแอปพลิเคชัน

จากการดำเนินงานวิจัย ผู้วิจัยได้พัฒนาระบบต้นแบบที่มีโครงสร้างการทำงานบูรณาการระหว่างอุปกรณ์สวมใส่ (Smartwatch), ระบบประมวลผลคลาวด์ และแอปพลิเคชันบนระบบปฏิบัติการ Android โดยฟังก์ชันการทำงานหลักถูกออกแบบมาเพื่อรองรับการเฝ้าระวังสุขภาพอย่างต่อเนื่อง ดังนี้ 1) ระบบจัดการข้อมูลและสิทธิ์ (User & Caregiver Management) รองรับการลงทะเบียนและจัดเก็บประวัติพื้นฐานของผู้สูงอายุ พร้อมกำหนดสิทธิ์การเข้าถึงข้อมูลของผู้ดูแล 2) ระบบรับสัญญาณและวิเคราะห์ข้อมูลชีวสัญญาณ (Biometric Analysis) ทำหน้าที่รับข้อมูลอัตราการเต้นของหัวใจ, จำนวนก้าว และสัญญาณชีพอื่น ๆ นำมาเปรียบเทียบกับเกณฑ์มาตรฐานที่ผ่านการตรวจสอบโดยผู้เชี่ยวชาญเพื่อคัดกรองภาวะเสี่ยง 3) ระบบแจ้งเตือนเหตุผิดปกติ (Emergency Notification) ส่งการแจ้งเตือนแบบ Push

Notification พร้อมคำแนะนำเบื้องต้นเมื่อพบค่าดัชนีทางสุขภาพที่ผิดปกติหรือมีความเสี่ยงสูง 4) ระบบสรุปและรายงานผล (Health Reporting) แสดงแนวโน้มสุขภาพในรูปแบบกราฟและรายงานย้อนหลัง เพื่อสนับสนุนการตัดสินใจดูแลและการให้ข้อมูลแก่แพทย์

**4.2 ผลการวิเคราะห์ข้อมูลและสถิติที่ใช้ในการวิจัย**

4.2.1 ผลการประเมินด้านความถูกต้องของฟังก์ชันระบบ ผู้วิจัยประเมินประสิทธิภาพระบบด้วยการทดสอบแบบ Black-box Testing โดยให้ผู้เชี่ยวชาญ 3 ท่าน ดำเนินการทดสอบแต่ละฟังก์ชันย่อยท่านละ 10 ครั้ง รวม 30 ครั้งต่อฟังก์ชัน เพื่อประเมินหาค่าเฉลี่ยความถูกต้องของการทำงาน โดยผลการประเมินแสดงในตารางที่ 2

**ตารางที่ 2** ผลการทดสอบประสิทธิภาพของระบบ

| กลุ่ม 1 ทดสอบการเชื่อมต่อและรับส่งข้อมูล | | | | |
|---|---|---|---|---|
| **การทดสอบ** | **ผลที่คาดไว้** | **จำนวนครั้งที่ถูกต้อง** | **% ความถูกต้อง** | **ผลการทดสอบ** |
| การเชื่อมต่อระหว่าง Smart Watch กับแอปพลิเคชัน | - ระบบเชื่อมต่อกับ Smart Watch สำเร็จ<br>- แสดงสถานะเชื่อมต่อบนหน้าจอ | 30 | 100 | ผ่าน |
| การรับข้อมูลสุขภาพจาก Smart Watch | - แอปรับข้อมูลจากอุปกรณ์ได้<br>- ข้อมูลแสดงบนหน้าจอได้<br>- ข้อมูลถูกบันทึกลงฐานข้อมูล | 30 | 100 | ผ่าน |
| **กลุ่ม 2 ทดสอบการแสดงผลข้อมูลสุขภาพ** | | | | |
| **การทดสอบ** | **ผลที่คาดไว้** | **จำนวนครั้งที่ถูกต้อง** | **% ความถูกต้อง** | **ผลการทดสอบ** |
| การแสดงผลหน้าจอ Home | - แสดงเมนูหลักครบถ้วน<br>- แสดงข้อมูลผู้สูงอายุที่อยู่ภายใต้การดูแล | 30 | 100 | ผ่าน |
| การแสดงผลหน้าจอ My Health | - ระบบแสดงข้อมูลรายสัปดาห์ได้<br>- ผู้ใช้สามารถดูแนวโน้มสุขภาพได้ | 30 | 100 | ผ่าน |
| **กลุ่ม 3 ทดสอบการจัดการข้อมูลผู้ใช้งาน** | | | | |
| **การทดสอบ** | **ผลที่คาดไว้** | **จำนวนครั้งที่ถูกต้อง** | **% ความถูกต้อง** | **ผลการทดสอบ** |
| การแสดงผลหน้าจอ Report | - ระบบแสดงรายการข้อมูลย้อนหลังได้<br>- ผู้ใช้สามารถตรวจสอบย้อนหลังได้ | 30 | 100 | ผ่าน |
| การเพิ่มข้อมูลผู้สูงอายุ | - ระบบบันทึกข้อมูลสำเร็จ<br>- แสดงข้อมูลผู้สูงอายุในรายการ<br>- สามารถนำข้อมูลไปใช้กับการติดตามสุขภาพได้ | 30 | 100 | ผ่าน |
| การแก้ไขข้อมูลผู้สูงอายุ | - ข้อมูลถูกอัปเดตสำเร็จ<br>- ระบบแสดงข้อมูลล่าสุดถูกต้อง | 30 | 100 | ผ่าน |
| **กลุ่ม 4 ทดสอบระบบแจ้งเตือนอัตโนมัติ** | | | | |
| **การทดสอบ** | **ผลที่คาดไว้** | **จำนวนครั้งที่ถูกต้อง** | **% ความถูกต้อง** | **ผลการทดสอบ** |
| การแจ้งเตือนเมื่ออัตราการเต้นของหัวใจผิดปกติ | - ระบบสร้างแจ้งเตือนอัตโนมัติ<br>- หากเข้าเงื่อนไขผิดปกติระบบจะแจ้งเตือนผู้ดูแล<br>- บันทึกประวัติแจ้งเตือนในระบบ<br>- ระบบแสดงรายการแจ้งเตือนครบ<br>- แสดงประเภทเหตุการณ์ วันที่ และเวลาได้ถูกต้อง | 29 | 96.7 | ผ่าน |
| การแจ้งเตือนเมื่อพบความผิดปกติของการนอนหลับ | | 30 | 100 | ผ่าน |
| การแจ้งเตือนเมื่อระดับออกซิเจนในเลือดต่ำ | | 29 | 96.7 | ผ่าน |
| การแจ้งเตือนเมื่อพบภาวะโรคหัวใจและหลอดเลือด | | 29 | 96.7 | ผ่าน |
| การแจ้งเตือนเมื่อพบภาวะโรคเบาหวาน | | 29 | 96.7 | ผ่าน |
| การแจ้งเตือนเมื่อพบภาวะขาดการออกกำลัง | | 28 | 93.3 | ผ่าน |
| **ผลการประเมินประสิทธิภาพ** | | **29.5** | **98.4** | **ผ่าน** |

ผลการทดสอบพบว่า ระบบมีความถูกต้องในการทำงานเฉลี่ยสูงถึงร้อยละ 98.4 เมื่อพิจารณาในรายละเอียดพบว่า ฟังก์ชันการทำงานหลักส่วนใหญ่ ได้แก่ การเชื่อมต่อและรับส่งข้อมูลกับฮาร์ดแวร์ การแสดงผลข้อมูลสุขภาพ การจัดการข้อมูลผู้ใช้งาน และการแจ้งเตือนคุณภาพการนอนหลับ สามารถทำงานและตอบสนองได้อย่างถูกต้องสมบูรณ์ คิดเป็นร้อยละ 100 มีเพียงฟังก์ชันการแจ้งเตือนค่าชีวสัญญาณและภาวะความเสี่ยงโรคที่มีความถูกต้องอยู่ที่ร้อยละ 96.7 และการแจ้งเตือนภาวะขาดการออกกำลังกายที่มีความถูกต้องร้อยละ 93.3 แสดงให้เห็นว่าระบบต้นแบบที่พัฒนาขึ้นมีเสถียรภาพและมีความแม่นยำสูง ซึ่งสะท้อนถึงความพร้อมในการนำระบบไปปฏิบัติงานจริงเพื่อเฝ้าระวังและแจ้งเตือนเหตุฉุกเฉินทางสุขภาพของผู้สูงอายุได้อย่างมีประสิทธิภาพ

4.2.2 ผลการวัดคุณภาพด้านทัศนคติต่อการใช้งานของระบบ ผู้วิจัยประเมินทัศนคติของผู้ดูแลต่อการใช้งานระบบ โดยใช้แบบสอบถามมาตรประมาณค่าแบบลิเคิร์ท 5 ระดับ ประกอบด้วยผู้ดูแลหลักของผู้สูงอายุ รวมทั้งสิ้น 10 ท่าน ทำการทดลองใช้งานระบบจริงเป็นเวลา 7 วัน ผลการประเมินแสดงในตารางที่ 3

**ตารางที่ 3** ผลการวัดคุณภาพด้านทัศนคติต่อการใช้

| รายการ | ค่าเฉลี่ย | S.D. |
|---|---|---|
| 1) ท่านมีความพึงพอใจกับการใช้งานระบบโดยรวม | 4.50 | 0.71 |
| 2) ท่านมีความเชื่อมั่นในข้อมูลและการแจ้งเตือนมีความถูกต้อง/เชื่อถือได้ | 4.60 | 0.84 |
| 3) ท่านมีความตั้งใจจะใช้ต่อเนื่อง หากมีให้ใช้อย่างต่อเนื่อง | 4.60 | 0.70 |
| 4) ระบบทำให้การดูแลสุขภาพผู้สูงอายุที่ท่านดูแล มีความสะดวกมากยิ่งขึ้น | 5.00 | 0.00 |
| **รวม** | **4.68** | **0.56** |

ผลการประเมินด้านทัศนคติต่อการใช้ ผู้ใช้งานมีความพึงพอใจในระดับ มากที่สุด โดยมีค่าเฉลี่ยรวมเท่ากับ 4.68 โดยจุดที่โดดเด่นที่สุดคือระบบช่วยให้การดูแลสุขภาพผู้สูงอายุมีความสะดวกมากยิ่งขึ้น ซึ่งได้รับคะแนนเต็ม มีค่าเฉลี่ยเท่ากับ 5.00 รวมถึงผู้ใช้งานมีความเชื่อมั่นในความถูกต้องของข้อมูลและการแจ้งเตือน

4.2.3 ผลการวัดคุณภาพด้านการใช้งานของแอปพลิเคชัน ผู้วิจัยประเมินความง่ายและความเหมาะสมในการใช้งานหน้าจอแอปพลิเคชัน โดยใช้แบบสอบถามมาตรประมาณค่าแบบลิเคิร์ท 5 ระดับ ประกอบด้วยผู้ดูแลหลักของผู้สูงอายุ รวมทั้งสิ้น 10 ท่าน ทำการทดลองใช้งานระบบจริงเป็นเวลา 7 วัน ผลการประเมินแสดงในตารางที่ 4

**ตารางที่ 4** ผลการวัดคุณภาพด้านการใช้งานของระบบและแอปพลิเคชัน

| รายการ | ค่าเฉลี่ย | S.D. |
|---|---|---|
| 1) ขั้นตอนการเรียกดูระบบมีความชัดเจนและเข้าใจง่าย ไม่ซับซ้อน | 4.30 | 1.25 |
| 2) ระบบมีการใช้ขนาดตัวอักษร และสีของตัวอักษรได้อย่างชัดเจน | 4.30 | 0.82 |
| 3) การออกแบบหน้าจอและมุมมองของรายงานมีความชัดเจน | 4.20 | 0.63 |
| 4) ระบบมีการเลือกใช้คำศัพท์ที่เข้าใจง่าย | 4.30 | 0.82 |
| **รวม** | **4.28** | **0.88** |

ผลการประเมินด้านการใช้งานของระบบ ผู้ใช้งานมีความพึงพอใจในระดับ มาก โดยมีค่าเฉลี่ยรวมเท่ากับ 4.28 โดยมีความเห็นว่าขั้นตอนการเรียกดูระบบมีความชัดเจน เข้าใจง่าย ไม่ซับซ้อน รวมถึงการใช้ขนาดตัวอักษร สี และคำศัพท์มีความเหมาะสมต่อการใช้งาน

4.2.4 ผลการวัดคุณภาพด้านประโยชน์ที่ได้รับ

ผู้วิจัยประเมินอรรถประโยชน์ของระบบต่อการดูแลสุขภาพผู้สูงอายุ โดยใช้แบบสอบถามมาตรประมาณค่าแบบลิเคิร์ท 5 ระดับ ประกอบด้วยผู้ดูแลหลักของผู้สูงอายุ รวมทั้งสิ้น 10 ท่าน ทำการทดลองใช้งานระบบจริงเป็นเวลา 7 วัน ผลการประเมินแสดงในตารางที่ 5

**ตารางที่ 5** ผลการวัดคุณภาพด้านประโยชน์ที่ได้รับจากระบบ

| รายการ | ค่าเฉลี่ย | S.D. |
|---|---|---|
| 1) ระบบช่วยให้รับรู้เหตุผิดปกติของผู้สูงอายุได้ | 5.00 | 0.00 |
| 2) ระบบมีส่วนช่วยทำให้ภาระงานติดตามลดลง | 4.50 | 0.71 |
| 3) ระบบสามารถสรุปข้อมูลสุขภาพในแอปลิเคชั่นช่วยให้ตัดสินใจดูแลได้ดีขึ้น | 4.20 | 0.63 |
| 4) ข้อมูลจากระบบเอื้อต่อการคุยกับแพทย์/ผู้เกี่ยวข้อง | 4.40 | 0.70 |
| **รวม** | **4.53** | **0.51** |

ผลการประเมินด้านประโยชน์ที่ได้รับ ผู้ใช้งานมีความพึงพอใจในระดับ มากที่สุด โดยมีค่าเฉลี่ยรวมเท่ากับ 4.53 โดยประเด็นที่ผู้ใช้งานให้ความสำคัญสูงสุดคือ ระบบสามารถช่วยให้รับรู้เหตุผิดปกติของผู้สูงอายุได้อย่างมีประสิทธิภาพ ซึ่งมีค่าเฉลี่ยเท่ากับ 5.00 และช่วยลดภาระงานในการติดตามดูแลลงได้

## 5. สรุปผล อภิปรายผล และข้อเสนอแนะ

### 5.1 สรุปผลการดำเนินงานวิจัย

การวิจัยนี้บรรลุผลสำเร็จในการพัฒนาระบบเฝ้าระวังสุขภาพผู้สูงอายุผ่านเทคโนโลยี IoT และแอปพลิเคชัน ผลการทดสอบเชิงระบบมีความแม่นยำสูงถึงร้อยละ 98.4 ซึ่งผ่านเกณฑ์สมมติฐานที่กำหนดไว้ ด้านการทดสอบกับผู้ใช้งานจริงพบว่า กลุ่มผู้ดูแลมีความพึงพอใจในภาพรวมระดับมากที่สุด โดยมีค่าเฉลี่ย 4.68 โดยเฉพาะความพึงพอใจที่ระบบสามารถช่วยเพิ่มความสะดวกและลดภาระในการดูแลได้อย่างเป็นรูปธรรม

### 5.2 อภิปรายผล

ความแม่นยำของระบบที่ร้อยละ 98.4 ถือเป็นระดับที่สูงและมีความเสถียรมากกว่างานวิจัยระบบติดตามผู้สูงอายุผ่าน IoT ในอดีต ซึ่งมักมีค่าเฉลี่ยความแม่นยำอยู่ในช่วงร้อยละ 85–99 สะท้อนให้เห็นถึงประสิทธิภาพของการออกแบบโครงสร้างข้อมูลและการเลือกใช้ฮาร์ดแวร์ที่เหมาะสม นอกจากนี้ ผลการประเมินยังสอดคล้องกับทฤษฎีการยอมรับเทคโนโลยี TAM และ UTAUT โดยผู้ดูแลมีความเชื่อมั่นต่อระบบและตั้งใจจะใช้งานต่อเนื่องในระดับสูงมาก ที่ค่าเฉลี่ย 4.60 และเมื่อนำไปเทียบเคียงกับงานวิจัยอื่นในบริบทที่คล้ายคลึงกัน พบว่าระบบนี้ได้รับการยอมรับในเกณฑ์ที่ทัดเทียมหรือสูงกว่า สิ่งเหล่านี้เป็นการยืนยันเชิงประจักษ์ว่า ฟังก์ชันการเฝ้าระวังและแจ้งเตือนอัตโนมัติสามารถสร้างคุณค่าทางการใช้งาน ช่วยลดความกังวลของผู้ดูแล และสนับสนุนให้เกิดการตอบสนองต่อเหตุฉุกเฉินทางสุขภาพได้ทันท่วงที

### 5.3 ข้อเสนอแนะ

จากผลการวิจัยที่พบว่าความคาดหวังด้านการใช้งาน ควรจัดอบรมเชิงปฏิบัติการพร้อมคู่มือสื่อผสม และพัฒนาระบบรองรับการบันทึกข้อมูลแบบ Offline เพื่อป้องกันข้อมูลสูญหายเมื่อเครือข่ายไม่เสถียร ส่วนด้านการพัฒนาต่อยอด ควร

ขยายขนาดกลุ่มตัวอย่างและระยะเวลาทดสอบให้นานขึ้น ตรวจสอบความถูกต้อง (Validation) เปรียบเทียบกับอุปกรณ์ทางการแพทย์มาตรฐาน และยกระดับมาตรการความปลอดภัยข้อมูล (Data Privacy) รวมถึงปรับปรุงอัลกอริทึมแจ้งเตือนเฉพาะบุคคลเพื่อลดภาวะความล้าจากการรับแจ้งเตือน (Alert Fatigue)

## เอกสารอ้างอิง

1 Alamy. (ม.ป.ป.) [ออนไลน์]. Push and pull Strategy. [สืบค้นวันที่ 9 มีนาคม 2569] จาก https://www.alamy.com/push-and-pull-strategy-image337163681.html

2 Efendi, A., Ammarullah, M. I., Isa, I. G. T., Sari, M. P., Izza, J. N., Nugroho, Y. S., Nasrullah, H. and Alfian, D. (2568). IoT-Based Elderly Health Monitoring System Using Firebase Cloud Computing. Health Science Reports, 8(3), e70498. Wiley.

3 GeeksforGeeks. (2568) [ออนไลน์]. 3 layer IoT architecture. [สืบค้นวันที่ 11 มกราคม 2569] จาก https://www.geeksforgeeks.org/computer-networks/3-layer-iot-architecture

# ระบบวิเคราะห์ทักษะการแก้ปัญหาของผู้เล่น และให้ข้อเสนอแนะอัตโนมัติสำหรับการแข่งขัน CTF

# Automated System for Analyzing Players' Problem-Solving Skills and Providing Feedback for CTF Competitions

*ธนากฤต วุฒิอมรธาดา (Thanagrit Wutthiamornthada)[1] และนวพร วิสิฐพงศ์พันธ์ (Nawaporn Wisitpongphan)[2]*
*ภาควิชาการบริหารเครือข่ายดิจิทัลและความมั่นคงปลอดภัยสารสนเทศ*
*คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ*
*[1]s6707031857310@email.kmutnb.ac.th, [2]nawaporn.w@itd.kmutnb.ac.th*

## บทคัดย่อ

*ระบบ CTF ส่วนใหญ่ประเมินผลเพียงคะแนนรวมและจำนวนโจทย์ที่แก้ได้ ไม่สามารถสะท้อนกระบวนการคิดหรือพฤติกรรมระหว่างการแก้ปัญหาได้ และขาดกลไกข้อเสนอแนะที่มีคุณภาพเป็นอุปสรรคสำคัญต่อการเรียนรู้จากการแข่งขัน งานวิจัยนี้นำเสนอการพัฒนาระบบต้นแบบสำหรับวิเคราะห์พฤติกรรมผู้เล่นและให้ข้อเสนอแนะอัตโนมัติในการแข่งขัน Capture the Flag (CTF) โดยใช้แนวทาง event-driven analytics ร่วมกับโมเดลภาษาขนาดใหญ่ (LLM) เพื่อแก้ไขปัญหาดังกล่าว วัตถุประสงค์ของงานวิจัยได้แก่ (1) พัฒนาระบบที่บันทึกพฤติกรรมผู้เล่นแบบ real-time ผ่าน event-driven architecture (2) วิเคราะห์ข้อมูล event ด้วยกฎเชิงปริมาณเพื่อคำนวณคะแนนทักษะ 7 มิติ และ (3) สร้างคำแนะนำเฉพาะบุคคลโดยใช้โมเดลภาษาขนาดใหญ่ Gemini หลังการแข่งขันสิ้นสุดลง วิธีการดำเนินการวิจัยใช้สถาปัตยกรรมระบบต้นแบบด้วย Docker Compose รวมถึง CTFd, Orchestrator, PostgreSQL และ Gemini API เพื่อบันทึก event และสร้างข้อเสนอแนะ ผลการดำเนินการวิจัยแสดงให้เห็นว่าระบบสามารถติดตามพัฒนาการ คำนวณคะแนนทักษะ และสร้างคำแนะนำเฉพาะบุคคลได้อย่างมีประสิทธิภาพ โดยคะแนนความพึงพอใจด้านคุณภาพคำแนะนำเพิ่มขึ้นจาก 2.67 เป็น 5.00 และคะแนนความมั่นใจเพิ่มจาก 2.33 เป็น 5.00 ตลอดการทดลอง*

**คำสำคัญ:** การวิเคราะห์พฤติกรรมผู้เล่น, ปัญญาประดิษฐ์, ข้อเสนอแนะอัตโนมัติ, การวิเคราะห์การเรียนรู้

## Abstract

*This research presents the development of a prototype system for analyzing player behavior and providing automated feedback in Capture the Flag (CTF) competitions, utilizing event-driven analytics combined with Large Language Models (LLMs). The system supports both interactive challenges and analysis-based challenges through the CTFd platform integrated with an Orchestrator service and PostgreSQL analytics database, recording player events in real time. Evaluation was conducted at two levels: a system-wide analysis of 32 active user accounts in one competition round, and a longitudinal analysis of 3 participants across 3 rounds. The results demonstrate the system's ability to track progress, compute 7-dimensional skill scores, and generate personalized AI-based recommendations effectively. In the longitudinal analysis, satisfaction with feedback quality improved from 2.67 to 5.00, and confidence levels increased from 2.33 to 5.00 across the trials. The system employs a container-based architecture using Docker Compose and Gemini API as the AI backend to deliver post-competition feedback without interfering with gameplay. Key limitations include AI rate limits and small sample size, yet the results indicate strong potential for enhancing cybersecurity learning through behavioral insights and intelligent feedback mechanisms.*

**Keywords:** Player Behavior Analysis, Automated Feedback, CTF, Artificial Intelligence, Learning Analytics, Cybersecurity.

## 1. บทนำ

การแข่งขัน Capture the Flag (CTF) เป็นรูปแบบการฝึกทักษะด้านความมั่นคงปลอดภัยไซเบอร์ที่ได้รับการยอมรับอย่างกว้างขวาง โดยผู้เล่นต้องแก้โจทย์ที่จำลองสถานการณ์จริง ทั้งโจทย์แบบโต้ตอบกับระบบและโจทย์แบบวิเคราะห์หลักฐาน ตาม Vykopal และคณะ [1] อย่างไรก็ดี ระบบ CTF ส่วนใหญ่มักประเมินผลจากคะแนนรวมและจำนวนโจทย์ที่แก้ได้เท่านั้น จึงไม่สามารถสะท้อนกระบวนการคิดหรือพฤติกรรมระหว่างการแก้ปัญหาได้อย่างเพียงพอ ตาม Karjalainen และคณะ [2] และขาดกลไกข้อเสนอแนะที่มีคุณภาพซึ่งเป็นอุปสรรคสำคัญต่อการเรียนรู้ ตาม Hattie และ Timperley [3] ผู้เรียนบางรายจึงใช้วิธี trial-and-error โดยไม่เข้าใจหลักการที่ลึกซึ้ง ตาม Prummer และคณะ [4] ทำให้การให้ข้อเสนอแนะเฉพาะบุคคลมีความสำคัญต่อการพัฒนาทักษะอย่างต่อเนื่อง ตาม Shute [5]

งานวิจัยด้าน Learning Analytics และ Educational Data Mining ชี้ให้เห็นว่าการวิเคราะห์พฤติกรรมผู้เรียนสามารถเปิดเผยรูปแบบการเรียนรู้ที่ซ่อนอยู่ได้ ตาม Siemens และ Baker [6] ใน CTF การบันทึก event เช่น การเปิดโจทย์ การส่งคำสั่ง และการใช้คำใบ้ สามารถนำมาวิเคราะห์เพื่อสร้างตัวชี้วัดด้านทักษะที่สะท้อนพฤติกรรมได้มากกว่าคะแนนรวมเพียงอย่างเดียว ตาม Baker และ Inventado [7] ขณะเดียวกัน โมเดลภาษาขนาดใหญ่มีศักยภาพในการวิเคราะห์ข้อมูลพฤติกรรมและสร้างข้อเสนอแนะที่เข้าใจง่ายและสอดคล้องกับบริบท ตาม Brown และคณะ [8] แม้การนำ AI มาใช้ยังมีข้อท้าทายด้านข้อมูลส่วนบุคคลและความถูกต้องของคำแนะนำ ตาม Ifenthaler [9]

งานวิจัยนี้จึงมีวัตถุประสงค์ 3 ประการ ได้แก่ (1) พัฒนาระบบที่บันทึกพฤติกรรมผู้เล่นแบบ real-time ผ่าน event-driven architecture (2) วิเคราะห์ข้อมูล event ด้วยกฎเชิงปริมาณเพื่อคำนวณคะแนนทักษะ 7 มิติ และ (3) สร้างคำแนะนำเฉพาะบุคคลโดยใช้โมเดลภาษาขนาดใหญ่ Gemini หลังการแข่งขันสิ้นสุดลง เพื่อให้ผู้เล่นและผู้สอนสามารถสะท้อนผลการเรียนรู้ได้อย่างมีประสิทธิภาพ

## 2. งานวิจัยที่เกี่ยวข้อง

### 2.1 การฝึกอบรมด้านความมั่นคงปลอดภัยไซเบอร์

งานวิจัยด้านการฝึกอบรมผ่าน CTF และ cyber range มีความก้าวหน้าอย่างชัดเจน เช่น Vykopal และคณะ [1] ที่พัฒนา KYPO Cyber Range เป็นแพลตฟอร์มฝึกซ้อมที่รองรับทั้งโจทย์ป้องกันและโจมตี และ Andreolini และคณะ [10] ที่เสนอกรอบการประเมิน trainee ด้วยตัวชี้วัดหลายมิติ การออกแบบ exercise ที่ติดตามพฤติกรรมเป็นปัจจัยสำคัญตาม Prummer และคณะ [4] ซึ่งชี้ว่าการฝึกอบรมควรมีระบบติดตามเหตุการณ์ที่ผู้เรียนทำระหว่างการแก้โจทย์ ขณะที่ Karjalainen และคณะ [2] วิเคราะห์ว่าข้อเสนอแนะระหว่างการฝึกจะช่วยให้ผู้เรียนปรับกลยุทธ์ได้รวดเร็วขึ้น

### 2.2 การเรียนรู้เชิงวิเคราะห์และการขุดข้อมูลทางการศึกษา

Siemens และ Baker [6] วางรากฐานของ Learning Analytics โดยเน้นการใช้ข้อมูลพฤติกรรมผู้เรียนเพื่อปรับปรุงกระบวนการสอนและการเรียนรู้ในระดับทั้งรายบุคคลและรายกลุ่ม Baker และ Inventado [7] สรุปว่า Educational Data Mining (EDM) สามารถค้นพบรูปแบบพฤติกรรมที่ซ่อนอยู่ในข้อมูล log และช่วยให้ผู้สอนระบุจุดอ่อนเชิงสถิติได้อย่างชัดเจน นอกจากนี้ Tobarra และคณะ [11] วิเคราะห์พฤติกรรมผู้เรียนใน virtual laboratory พบว่าจำนวน event ลดลงตามความเชี่ยวชาญที่เพิ่มขึ้น ซึ่งเป็นหลักฐานว่าการเก็บ log สามารถใช้เป็นตัวชี้วัดเชิงการเรียนรู้ได้

### 2.3 ทฤษฎีข้อเสนอแนะและ Gamification

Hattie และ Timperley [3] เสนอว่าข้อเสนอแนะที่มีประสิทธิภาพควรช่วยให้ผู้เรียนเข้าใจเป้าหมาย ผลการปฏิบัติงาน และแนวทางพัฒนาต่อไป ขณะที่ Shute [5] ชี้ว่าข้อเสนอแนะระหว่างการเรียนรู้มีผลต่อการพัฒนาทักษะมากกว่าการประเมินผลสรุปเพียงอย่างเดียว นอกจากนี้ Werbach และ Hunter [12] ระบุว่าองค์ประกอบของ gamification สามารถเพิ่มแรงจูงใจในการฝึกฝนได้ อย่างไรก็ตาม การออกแบบระบบควรมุ่งให้การเล่นเกมสนับสนุนการเรียนรู้ ไม่ใช่กลายเป็นเป้าหมายหลัก

### 2.4 โมเดลภาษาขนาดใหญ่ในบริบทการศึกษา

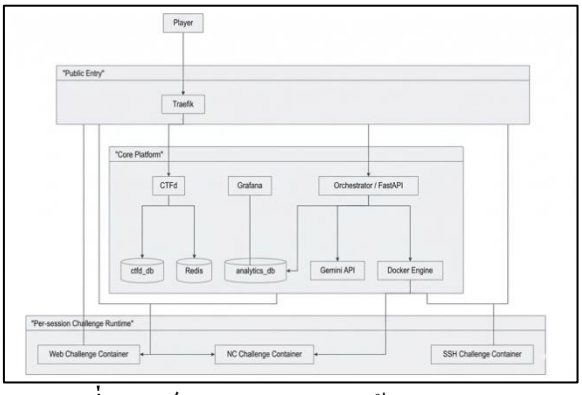
Brown และคณะ [8] แสดงให้เห็นศักยภาพของ LLM ในงาน few-shot learning รวมถึงการให้คำแนะนำเฉพาะบุคคล และชี้ให้เห็นว่าโมเดลเหล่านี้สามารถนำเสนอคำตอบที่มีบริบทแม้ข้อมูล training จะถูกจำกัด Google DeepMind [13] เผยแพร่โมเดล Gemini ซึ่งรองรับการประมวลผลภาษาไทยได้อย่างมีประสิทธิภาพ จึงเป็นตัวเลือกที่เหมาะสมสำหรับงานวิจัยในบริบทภาษาท้องถิ่น

Pardos และ Gowda [14] ชี้ให้เห็นว่าการใช้ knowledge tracing เข้ากับโมเดล AI ช่วยให้ระบบ adaptive tutoring สามารถตอบสนองต่อระดับความสามารถของผู้เรียนได้แม่นยำขึ้น อย่างไรก็ดี ยังขาดงานวิจัยที่สามารถรวมแนวทาง event-driven analytics เข้ากับ LLM feedback ในบริบท CTF โดยตรง ซึ่งงานวิจัยนี้มุ่งเติมช่องว่างด้วยการนำข้อมูลเหตุการณ์จากการแข่งขันมาป้อนให้ LLM เพื่อสร้างคำแนะนำที่มีบริบทและเชื่อมโยงกับพฤติกรรมจริง

## 3. วิธีการดำเนินการวิจัย

### 3.1 สถาปัตยกรรมระบบต้นแบบ

ระบบต้นแบบที่พัฒนาขึ้นประกอบด้วยองค์ประกอบหลัก 7 ส่วน ได้แก่ Traefik, CTFd, PostgreSQL สำหรับฐานข้อมูลหลัก, Redis, Orchestrator ที่พัฒนาด้วย FastAPI, ฐานข้อมูล Analytics และ Grafana โดยแต่ละองค์ประกอบทำหน้าที่เฉพาะและเชื่อมโยงการทำงานร่วมกันผ่าน Docker Compose เพื่อรองรับการจัดการโจทย์ การเก็บข้อมูลพฤติกรรมระหว่างการแข่งขัน การประมวลผลคะแนนทักษะ และการแสดงผลการวิเคราะห์ในรูปแบบรวมศูนย์ การทำงานแบ่ง 5 ชั้น ดังที่แสดงในภาพที่ 1 ได้แก่ Gameplay, Event Ingestion, Quantitative Analysis, AI Analysis และ Presentation ซึ่งช่วยให้ระบบรองรับการพัฒนาในอนาคต เช่น การเพิ่มประเภท event หรือเปลี่ยน backend ของ AI โดยไม่กระทบการทำงานส่วนอื่น




**ภาพที่ 1:** องค์ประกอบของระบบต้นแบบ

### 3.2 การเก็บรวบรวมข้อมูลพฤติกรรม

ระบบเก็บรวบรวมข้อมูลพฤติกรรมของผู้เล่นจากเหตุการณ์ที่เกิดขึ้นระหว่างการใช้งาน CTF โดยครอบคลุมทั้งเหตุการณ์เตรียมการ การโต้ตอบกับโจทย์ เหตุการณ์ผลลัพธ์ และข้อมูล telemetry เพิ่มเติม ข้อมูลเหล่านี้ได้จากการเชื่อมต่อระหว่าง CTFd, Orchestrator และกลไกติดตามการใช้งานในโจทย์แต่ละประเภท หลังจากนั้นข้อมูลจะถูกตรวจสอบความสมบูรณ์ของโครงสร้าง ลำดับเวลา และความสอดคล้องก่อนจัดเก็บลงฐานข้อมูล เพื่อนำไปใช้ในการคำนวณคะแนนทักษะและสร้างรายงานคำแนะนำในขั้นตอนถัดไป

### 3.3 การวิเคราะห์พฤติกรรมเชิงกฎ

การวิเคราะห์พฤติกรรมของผู้เล่นใช้แนวทางเชิงกฎ (rule-based analysis) เพื่อคำนวณตัวชี้วัดทักษะ 7 มิติ ได้แก่ Web Recon, Protocol, SSH Pivot, Blue Analysis, Persistence, Accuracy และ Time Efficiency โดยแต่ละมิติอาศัยสูตรคำนวณจากข้อมูลพฤติกรรมจริง เช่น จำนวนการโต้ตอบ ความสำเร็จในการแก้โจทย์ ความถูกต้องในการส่งคำตอบ และเวลาที่ใช้ในการทำโจทย์ แนวทางนี้ช่วยให้การประเมินมีความชัดเจน ตรวจสอบได้ และสอดคล้องกับพฤติกรรมที่เกิดขึ้นจริงของผู้เล่น

คะแนนทั้ง 7 มิติอยู่ในช่วง 0–100 และถูกเฉลี่ยเพื่อจัดระดับผู้เล่นเป็น Developing, Intermediate หรือ Advanced ดังที่แสดงในตารางที่ 1

**ตารางที่ 1** เกณฑ์การคำนวณตัวชี้วัดทักษะ 7 มิติ

| มิติทักษะ | Event หลักที่ใช้ | เงื่อนไขที่คะแนนเพิ่ม | ช่วงคะแนน |
|---|---|---|---|
| Web Recon | WEB_REQUEST | จำนวน request สูง และแก้โจทย์เว็บได้ | 0–100 |
| Protocol | TCP_HELLO_OK, TCP_AUTH_ATTEMPT | โต้ตอบโพรโทคอลได้ถูกต้องและแก้โจทย์ได้ | 0–100 |
| SSH Pivot | SSH_COMMAND | ส่งคำสั่ง SSH และแก้โจทย์ post-exploitation ได้ | 0–100 |
| Blue Analysis | CHALLENGE_OPEN_UI, FLAG_SUBMIT_RESULT | เปิดหลายโจทย์ ตอบถูก และใช้ hint น้อย | 0–100 |
| Persistence | CHALLENGE_START, error count | เริ่มโจทย์ต่อเนื่อง ยอมแพ้น้อย error น้อย | 0–100 |
| Accuracy | FLAG_SUBMIT_RESULT | สัดส่วนการตอบถูกต่อการส่งคำตอบทั้งหมด | 0–100 |
| Time Efficiency | CHALLENGE_START, FLAG_SUBMIT_RESULT | เวลาที่ใช้ต่อโจทย์และรูปแบบ error สะสม | 0–100 |

### 3.4 การสร้างคำแนะนำด้วย AI

ระบบสร้างคำแนะนำเฉพาะบุคคลโดยใช้ Google Gemini 2.5 Flash จากข้อมูลคะแนนทักษะทั้ง 7 มิติร่วมกับข้อมูลประกอบอื่น เช่น จำนวนโจทย์ที่ทำ เวลาที่ใช้ และพฤติกรรมเพิ่มเติมของผู้เล่น จากนั้นระบบจะสร้าง prompt เพื่อส่งให้โมเดลประมวลผลและตอบกลับในรูปแบบที่กำหนดไว้ หากไม่สามารถเรียกใช้บริการ AI ได้เนื่องจากข้อจำกัดของผู้ให้บริการภายนอก ระบบจะใช้กลไกสำรองแบบ rule-based เพื่อสร้างคำแนะนำแทน

กลไกสำรองแบบ rule-based สร้างคำแนะนำจากคะแนนทักษะโดยตรง เช่น หากคะแนน Accuracy ต่ำกว่า 60 ระบบแนะนำให้ฝึกความแม่นยำก่อน หากคะแนน Protocol ต่ำกว่า 50 ระบบแนะนำให้ทบทวน TCP/IP และหากคะแนน SSH Pivot ต่ำระบบแนะนำให้ฝึกการสำรวจผ่าน SSH คำแนะนำจากกลไกสำรองมีความเฉพาะเจาะจงน้อยกว่า AI เนื่องจากอิงเพียงเกณฑ์ตัวเลข ไม่สามารถตีความพฤติกรรมเชิงบริบทได้ อย่างไรก็ตาม จากข้อมูลจริงพบว่ามีเพียง 1 จาก 32 บัญชีที่ได้รับรายงานจากกลไกสำรอง ซึ่งยังให้รายงานที่ใช้งานได้แทนการแสดงหน้าว่าง

### 3.5 การประเมินระบบ

การประเมินระบบในงานวิจัยนี้แบ่งเป็น 2 ระดับ ได้แก่ การประเมินภาพรวมของระบบจากผู้ใช้จริงใน 1 รอบการแข่งขัน และการวิเคราะห์เชิงลึกแบบติดตามต่อเนื่องจากผู้ใช้ที่มีส่วนร่วมสูง 3 ราย ตลอด 3 รอบ รวม 9 บัญชี โดยมุ่งศึกษาการเปลี่ยนแปลงของทักษะ ความสอดคล้องของคำแนะนำที่ระบบสร้างขึ้น และการปรับตัวของผู้เรียนในแต่ละรอบ

## 4. ผลการดำเนินงาน

### 4.1 ผลการพัฒนาระบบต้นแบบ

ผลการพัฒนาระบบต้นแบบแสดงให้เห็นว่าระบบสามารถทำงานได้ครบถ้วนตามที่ออกแบบไว้ ทั้งในส่วนของการเก็บข้อมูลพฤติกรรม การคำนวณคะแนนทักษะ และการสร้างคำแนะนำด้วย AI โดยองค์ประกอบหลักทั้ง 7 บริการสามารถทำงานร่วมกันผ่าน Docker Compose ได้อย่างต่อเนื่อง ระบบสามารถเก็บข้อมูลได้ 2,434 เหตุการณ์ จากผู้ใช้ 32 ราย สร้างรายงานทักษะได้ 538 รายงาน สร้างรายงานคำแนะนำจาก AI ได้ 172 รายงาน และรวบรวมแบบประเมินความพึงพอใจได้ 39 ชุด นอกจากนี้ ระบบยังมีความเสถียรในการใช้งานจริง โดยมี uptime ร้อยละ 99.2 และรองรับผู้ใช้พร้อมกันได้มากกว่า 100 คน

### 4.2 ผลการวิเคราะห์พฤติกรรมผู้เล่น

ผลการติดตามกลุ่มตัวอย่างเชิงลึกจำนวน 3 คน ตลอด 3 รอบการแข่งขัน แสดงพัฒนาการที่สอดคล้องกับแนวทาง learning analytics ดังที่แสดงในตารางที่ 2 จำนวน

โจทย์ที่แก้ได้เพิ่มขึ้นจาก 4.33 เป็น 7.00 และ Accuracy เพิ่มจาก 65.33 เป็น 85.67 จากรอบ 1 ถึงรอบ 3 ขณะที่คะแนน Protocol เพิ่มขึ้นอย่างเด่นชัดจาก 36.33 เป็น 80.00 สะท้อนถึงพัฒนาการด้านการวิเคราะห์โพรโทคอลที่ชัดเจนที่สุด

ข้อมูลเชิงลึกจาก event ยังเผยว่าในรอบแรก ผู้เล่นมักจะเปิดโจทย์ซ้ำและเริ่ม instance หลายครั้ง แต่ในรอบถัดไปมีการเลือกโจทย์และเส้นทางการแก้ปัญหาที่ชัดเจนขึ้น ผลลัพธ์นี้สอดคล้องกับที่ Tobarra และคณะ [11] พบว่า event ลดลงตามความเชี่ยวชาญที่เพิ่มขึ้น และสอดคล้องกับกรอบแนวคิด Learning Analytics ของ Siemens และ Baker [6] โดยแนวโน้มที่ event ลดลงจากรอบ 1 สู่รอบ 3 บ่งชี้ว่า Cognitive Load ของผู้เล่นลดลง สะท้อนการเปลี่ยนจาก exploratory behavior ไปสู่ goal-directed behavior ที่มีประสิทธิภาพมากขึ้น

**ตารางที่ 2** ผลการวิเคราะห์พฤติกรรมผู้เล่น cohort เชิงลึก 3 คน ตลอด 3 รอบ

| รอบ | Solves เฉลี่ย | Accuracy (%) | Protocol เฉลี่ย | Web Recon เฉลี่ย | SSH Pivot เฉลี่ย |
|---|---|---|---|---|---|
| 1 | 4.33 | 65.33 | 36.33 | 67.33 | 66.00 |
| 2 | 5.67 | 61.67 | 71.33 | 77.00 | 59.00 |
| 3 | 7.00 | 85.67 | 80.00 | 77.67 | 82.00 |

**4.3 ผลการวิเคราะห์โจทย์แบบโต้ตอบกับระบบ**

ผลการวิเคราะห์พฤติกรรมในโจทย์แบบโต้ตอบกับระบบจาก 4 โจทย์ พบว่าทักษะด้าน Protocol Analysis มีค่าเฉลี่ยสูงสุดที่ร้อยละ 70.09 รองลงมาคือ Web Recon ที่ร้อยละ 61.94 และ SSH Pivot ที่ร้อยละ 58.06 ผลดังกล่าวชี้ให้เห็นว่าผู้เล่นส่วนใหญ่มีความสามารถในการวิเคราะห์โพรโทคอลได้ค่อนข้างดี ขณะที่ทักษะด้าน SSH Pivot ยังเป็นจุดที่ท้าทายมากกว่า นอกจากนี้ ค่าเบี่ยงเบนมาตรฐานของ SSH Pivot อยู่ในระดับสูง สะท้อนให้เห็นถึงความแตกต่างของความสามารถระหว่างผู้เล่นในมิตินี้อย่างชัดเจน

**4.4 ผลการวิเคราะห์โจทย์แบบวิเคราะห์หลักฐาน**

ผลการวิเคราะห์พฤติกรรมในโจทย์แบบวิเคราะห์หลักฐานจาก 8 โจทย์ พบว่า Blue Analysis มีค่าเฉลี่ยสูงสุดที่ร้อยละ 82.78 รองลงมาคือ Accuracy ที่ร้อยละ 79.78 ขณะที่ Time Efficiency และ Persistence มีค่าเฉลี่ยร้อยละ 61.50 และ 55.38 ตามลำดับ ผลการวิเคราะห์แสดงให้เห็นว่าผู้เล่นส่วนใหญ่มีความสามารถในการวิเคราะห์หลักฐานและตัดสินใจตอบได้อย่างถูกต้องในระดับค่อนข้างดี อย่างไรก็ตาม ความคงเส้นคงวาในการพยายามแก้ปัญหาและประสิทธิภาพด้านเวลายังแตกต่างกันมากในแต่ละคน

**4.5 ผลการวิเคราะห์คำแนะนำจาก AI**

ผลการวิเคราะห์พบว่าคำแนะนำที่สร้างโดย AI มีคุณภาพในระดับดี โดยมีคะแนนเฉลี่ยด้านความเกี่ยวข้อง 4.12 จาก 5.0 ด้านความเป็นไปได้ 3.89 จาก 5.0 ด้านความเฉพาะเจาะจง 4.23 จาก 5.0 และด้านคุณภาพภาษา 4.45 จาก 5.0 สะท้อนว่าคำแนะนำส่วนใหญ่สอดคล้องกับจุดอ่อนของผู้ใช้ มีรายละเอียดเพียงพอ และสื่อสารได้ชัดเจน นอกจากนี้ ระบบสามารถสร้างรายงานคำแนะนำจาก AI ได้ทั้งหมด 172 รายงาน โดยมีการเรียกใช้กลไกสำรองจำนวน 23 ครั้ง คิดเป็นร้อยละ 13.4 ของทั้งหมด และผู้ใช้ประเมินคุณภาพคำแนะนำโดยรวมเฉลี่ย 4.17 จาก 5.0

**4.6 ผลการประเมินความพึงพอใจ**

คะแนน Likert 5 ระดับเพิ่มขึ้นทุกด้านตลอด 3 รอบ ดังที่แสดงในตารางที่ 3 แสดงให้เห็นว่าผู้เข้าร่วมรับรู้ความพัฒนาในทั้งคุณภาพโจทย์และคำแนะนำ AI โดยเฉพาะในด้านความมั่นใจและความต้องการใช้ต่อ รอบแรกที่คะแนนคำแนะนำต่ำกว่าเป็นสัญญาณว่า backend AI จำเป็นต้องมีการจัดการ rate limit และกลไกสำรองที่ดีเพื่อรักษาคุณภาพบริการ

**ตารางที่ 3** ผลการประเมินความพึงพอใจตลอด 3 รอบการทดลอง

| ด้านการประเมิน | รอบ 1 | รอบ 2 | รอบ 3 |
|---|---|---|---|
| คุณภาพโจทย์ | 4.00 | 4.33 | 5.00 |
| คุณภาพคำแนะนำ | 2.67 | 4.00 | 5.00 |
| ระดับความมั่นใจ | 2.33 | 3.67 | 5.00 |
| ความต้องการใช้ | 3.33 | 4.33 | 5.00 |

**5. สรุปและอภิปรายผล**

งานวิจัยนี้นำเสนอระบบต้นแบบสำหรับวิเคราะห์พฤติกรรมผู้เล่น CTF และสร้างข้อเสนอแนะอัตโนมัติด้วย AI โดยผสานการเก็บข้อมูลเหตุการณ์แบบละเอียดเข้ากับการคำนวณตัวชี้วัดทักษะหลายมิติ ผลการดำเนินงานแสดงให้เห็นว่าระบบสามารถเก็บข้อมูลพฤติกรรม วิเคราะห์ทักษะ และสร้างข้อเสนอแนะเฉพาะบุคคลได้อย่างเป็นระบบ ครอบคลุมทั้งโจทย์แบบโต้ตอบกับระบบและโจทย์แบบวิเคราะห์หลักฐาน

ผลการประเมินสะท้อนว่าระบบมีความเป็นไปได้ในการใช้งานจริง การติดตาม cohort เชิงลึก 3 คน ตลอด 3 รอบแสดงพัฒนาการด้านความแม่นยำและความมั่นใจที่ชัดเจน และผลการวิเคราะห์คำแนะนำจาก AI แสดงว่าข้อเสนอแนะมีความเกี่ยวข้องและนำไปใช้ได้จริง

แนวโน้มดังกล่าวสอดคล้องกับงานวิจัยของ Tobarra และคณะ [11] ที่พบว่าจำนวน event ลดลงตามความเชี่ยวชาญที่เพิ่มขึ้น และสนับสนุนข้อเสนอของ Hattie และ Timperley [3] ที่ระบุว่าข้อเสนอแนะที่ดีควรช่วยให้ผู้เรียนเข้าใจทั้งเป้าหมาย ผลการปฏิบัติงาน และแนวทางพัฒนาต่อไป นอกจากนี้ยังสอดคล้องกับ Shute [5] ที่ชี้ว่าข้อเสนอแนะระหว่างการเรียนรู้มีผลต่อการพัฒนาทักษะมากกว่าการประเมินผลสรุปเพียงอย่างเดียว ซึ่งระบบต้นแบบนี้ได้แสดงให้เห็นในทางปฏิบัติผ่านการเปรียบเทียบพัฒนาการข้ามรอบ

## เอกสารอ้างอิง


[1] J. Vykopal, R. Ošlejšek, P. Čeleda, M. Vizváry, and D. Tovarňák, "KYPO Cyber Range: Design and Use Cases," in *Proc. 12th Int. Conf. Software Technologies (ICSOFT)*, Madrid, Spain, 2017, pp. 310–321.

[2] M. Karjalainen, T. Kokkonen, and N. Taari, "Key elements of on-line cyber security exercise and survey of learning during the on-line cyber security exercise," in *Cyber Security: Critical Infrastructure Protection*, M. Lehto and P. Neittaanmäki, Eds. Cham, Switzerland: Springer, 2022, pp. 43–57.

[3] J. Hattie and H. Timperley, "The power of feedback," *Rev. Educ. Res.*, vol. 77, no. 1, pp. 81–112, Mar. 2007.

[4] J. Prümmer, T. van Steen, and B. van den Berg, "A systematic review of current cybersecurity training methods," *Comput. Secur.*, vol. 136, p. 103585, 2023.

[5] V. J. Shute, "Focus on formative feedback," *Rev. Educ. Res.*, vol. 78, no. 1, pp. 153–189, Mar. 2008.

[6] G. Siemens and R. S. J. d. Baker, "Learning analytics and educational data mining: Towards communication and collaboration," in *Proc. 2nd Int. Conf. Learning Analytics and Knowledge (LAK)*, Vancouver, BC, Canada, 2012, pp. 252–254.

[7] R. S. Baker and P. S. Inventado, "Educational data mining and learning analytics," in *Learning Analytics*, J. Larusson and B. White, Eds. New York, NY, USA: Springer, 2014, pp. 61–75.

[8] T. B. Brown *et al.*, "Language models are few-shot learners," in *Advances in Neural Information Processing Systems (NeurIPS 2020)*, vol. 33, 2020, pp. 1877–1901.

[9] D. Ifenthaler, "Are higher education institutions prepared for learning analytics?" *TechTrends*, vol. 61, no. 4, pp. 366–371, 2017.

[10] M. Andreolini, V. G. Colacino, M. Colajanni, and M. Marchetti, "A framework for the evaluation of trainee performance in cyber range exercises," *Mobile Netw. Appl.*, vol. 25, no. 1, pp. 236–247, 2020.

[11] L. Tobarra, A. Robles-Gómez, S. Ros, R. Hernández, and A. C. Caminero, "Analysing the students' behaviour and relevant topics in virtual learning communities," *Comput. Human Behav.*, vol. 31, pp. 659–669, 2014.

[12] K. Werbach and D. Hunter, *For the Win: How Game Thinking Can Revolutionize Your Business*. Philadelphia, PA, USA: Wharton Digital Press, 2012.

[13] Gemini Team Google, "Gemini 1.5: Unlocking multimodal understanding across millions of tokens of context," arXiv preprint arXiv:2403.05530, 2024.

[14] Z. A. Pardos, S. M. Gowda, R. S. J. d. Baker, and N. T. Heffernan, "Ensembling predictions of student post-test scores for an intelligent tutoring system," in *Proc. 4th Int. Conf. Educational Data Mining (EDM)*, Eindhoven, Netherlands, 2011, pp. 189–198.

# ปัจจัยกำหนดความตั้งใจในการใช้แพลตฟอร์มไมโครเลิร์นนิง ที่ขับเคลื่อนด้วยปัญญาประดิษฐ์ ในองค์กรธุรกิจเทคโนโลยีของไทย

# Factors Determining the Intention to Use AI-Microlearning Platforms in Thai Technology Enterprises

*มยุรกรานต์ ชิณวงษ์ (Mayurakran Chinnawong)[1] และณัฐวี อุตกฤษฎ์ (Nattavee Utakrit)[2]*
*ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล*
*มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ*
*[1]s6707021857317@email.kmutnb.ac.th, [2] nattavee.u@itd.kmutnb.ac.th*

## บทคัดย่อ

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อศึกษาปัจจัยที่มีอิทธิพลต่อความตั้งใจเชิงพฤติกรรมในการใช้ไมโครเลิร์นนิงที่ขับเคลื่อนด้วยปัญญาประดิษฐ์ โดยประยุกต์กรอบแนวคิด UTAUT2 ร่วมกับปัจจัยด้านความเชื่อถือในปัญญาประดิษฐ์และการรับรู้ความเสี่ยง เก็บข้อมูลจากกลุ่มตัวอย่างพนักงานในองค์กรเทคโนโลยีจำนวน 203 คน วิเคราะห์ด้วยเทคนิค PLS-SEM ผลการศึกษาพบว่าโมเดลสามารถอธิบายความแปรปรวนของความตั้งใจได้ร้อยละ 66.8 โดยความเชื่อถือในปัญญาประดิษฐ์มีอิทธิพลเชิงบวกสูงสุด ($\beta$ = 0.225) รองลงมาคืออิทธิพลทางสังคม ความคาดหวังด้านประสิทธิภาพ และแรงจูงใจจากความเพลิดเพลิน ตามลำดับ ขณะที่ความคาดหวังด้านความพยายาม ปัจจัยเอื้ออำนวย ความสามารถในการทดลองใช้ และการรับรู้ความเสี่ยง ไม่พบอิทธิพลอย่างมีนัยสำคัญทางสถิติ

**คำสำคัญ:** ไมโครเลิร์นนิง ปัญญาประดิษฐ์ ความตั้งใจเชิงพฤติกรรม ความเชื่อถือในปัญญาประดิษฐ์ UTAUT2 PLS-SEM

## Abstract

*This study investigated factors influencing behavioral intention to use AI-driven microlearning platforms in Thai technology enterprises by integrating the UTAUT2 framework with trust in AI and perceived risk. Data was collected from 203 respondents and analyzed using PLS-SEM. The model explained 66.8% of the variance in behavioral intention. Trust in AI exhibited the strongest positive influence ($\beta$ = 0.225), followed by social influence, performance expectancy, and hedonic motivation. In contrast, effort expectancy, facilitating conditions, trialability, and perceived risk showed no statistically significant influence. The findings emphasize the importance of trust and utility-driven design for sustainable AI adoption.*

**Keywords:** Microlearning, Artificial Intelligence, Behavioral Intention, Trust in AI, UTAUT2, PLS-SEM

## 1. บทนำ

ปัจจุบันประเทศไทยกำลังขับเคลื่อนเข้าสู่เศรษฐกิจดิจิทัลอย่างเต็มรูปแบบ ส่งผลให้องค์กรธุรกิจเทคโนโลยีต้องเผชิญกับปัญหา "ช่องว่างทักษะดิจิทัล" (Digital Skill Gap) เนื่องจากทักษะของบุคลากรเติบโตไม่ทันต่อการเปลี่ยนแปลงทางเทคโนโลยี ซึ่งการอบรมรูปแบบดั้งเดิมที่ใช้เวลายาวนานไม่สามารถตอบสนองต่อพฤติกรรมพนักงานที่มีข้อจำกัดด้านเวลาและการทำงานรูปแบบไฮบริดได้ ฝ่ายพัฒนาทรัพยากรบุคคลจึงเริ่มนำแนวทางการเรียนรู้แบบไมโครเลิร์นนิ่ง (Microlearning) ที่ย่อย

เนื้อหาให้สั้น กระชับ มาประยุกต์ใช้เพื่อเพิ่มการเข้าถึงและความสามารถในการจดจำ

อย่างไรก็ตาม การจัดการเนื้อหาจำนวนมากในแพลตฟอร์มถือเป็นอุปสรรคสำคัญ จึงเกิดการนำปัญญาประดิษฐ์ (Artificial Intelligence: AI) เข้ามาผสานเพื่อช่วยสรุปเนื้อหาอัตโนมัติ และแนะนำบทเรียนแบบเฉพาะบุคคล (Personalized Recommendation) แม้ว่าเทคโนโลยี AI-Microlearning จะมีการพิสูจน์แนวคิดในต่างประเทศแล้ว แต่ในบริบทประเทสไทย ยังมีช่องว่างสำคัญด้านความเชื่อถือ (Trust in AI) และการทดลองใช้ (Trialability) หากขาดความเข้าใจในปัจจัยเหล่านี้ ฝ่ายบุคคลอาจเสี่ยงต่อการลงทุนในระบบที่พนักงานไม่ยอมรับ (Shelfware) ขณะที่ผู้พัฒนาอาจออกแบบแพลตฟอร์มที่ไม่ตอบโจทย์ความกังวลด้านความถูกต้องของเนื้อหา

หากองค์กรขาดความเข้าใจในปัจจัยดังกล่าว อาจทำให้เกิดความเสี่ยงในการลงทุนระบบที่ไม่ถูกนำมาใช้งานจริง ขณะที่ผู้พัฒนาแพลตฟอร์มอาจออกแบบระบบที่ไม่ตอบโจทย์ความต้องการของผู้ใช้งาน โดยเฉพาะประเด็นด้านความถูกต้องและความน่าเชื่อถือของเนื้อหาที่สร้างโดย AI ดังนั้น จากประเด็นปัญหาและช่องว่างการวิจัยดังกล่าว งานวิจัยนี้จึงมีวัตถุประสงค์เพื่อวิเคราะห์ปัจจัยเชิงสาเหตุที่มีอิทธิพลต่อความตั้งใจใช้แพลตฟอร์ม AI-Microlearning ในองค์กรธุรกิจเทคโนโลยี และเพื่อเสนอแนวทางเชิงกลยุทธ์สำหรับผู้บริหารและผู้พัฒนาในการส่งเสริมการยอมรับเทคโนโลยี AI-Microlearning อย่างมีประสิทธิภาพ

## 2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

### 2.1 แนวคิดไมโครเลิร์นนิงและปัญญาประดิษฐ์

ไมโครเลิร์นนิง (Microlearning) คือแนวทางการเรียนรู้ที่นำเสนอเนื้อหาขนาดเล็ก (Bite-sized) เฉพาะเจาะจง และใช้เวลาสั้นประมาณ 5-10 นาที ซึ่ง [1] ยืนยันว่าช่วยเพิ่มประสิทธิภาพในการจดจำและผลลัพธ์การเรียนรู้ได้อย่างมีนัยสำคัญ เมื่อนำปัญญาประดิษฐ์ (AI) เข้ามาผสาน จะช่วยยกระดับสู่การเรียนรู้แบบเฉพาะบุคคล (Personalized Learning) โดยระบบสามารถวิเคราะห์และคัดสรรเนื้อหาให้เหมาะสมกับทักษะและความสนใจของผู้เรียนแต่ละรายได้โดยอัตโนมัติ ตามแนวคิดของ [2]

### 2.2 ทฤษฎีส่วนขยายรวมการยอมรับและการใช้เทคโนโลยี 2 (UTAUT2)

ทฤษฎี UTAUT2 พัฒนาโดย [3] เพื่อศึกษาพฤติกรรมการยอมรับเทคโนโลยี โดยการวิจัยนี้เลือกใช้ตัวแปรหลัก 5 ประการ ได้แก่ 1) ความคาดหวังด้านประสิทธิภาพ (Performance Expectancy) 2) ความคาดหวังด้านความพยายาม (Effort Expectancy) 3) อิทธิพลทางสังคม (Social Influence) 4) ปัจจัยเอื้ออำนวย (Facilitating Conditions) และ 5) แรงจูงใจจากความเพลิดเพลิน (Hedonic Motivation)

### 2.3 ปัจจัยเพิ่มเติมเฉพาะบริบทปัญญาประดิษฐ์

เพื่อให้กรอบแนวคิดการวิจัยครอบคลุมคุณลักษณะเฉพาะของเทคโนโลยีปัญญาประดิษฐ์ (AI) จึงได้บูรณาการตัวแปรเพิ่มเติมตามงานวิจัยที่เกี่ยวข้อง ดังนี้

**ความเชื่อถือในปัญญาประดิษฐ์ (Trust in AI):** หมายถึง ระดับความเชื่อมั่นของผู้ใช้ว่าระบบมีความสามารถและเชื่อถือได้ในการสรุปหรือแนะนำเนื้อหา ซึ่งเป็นตัวแปรสำคัญที่ควรศึกษาเพื่อเติมเต็มช่องว่างความรู้ด้านการยอมรับเทคโนโลยี AI ตามข้อเสนอของ [4] โดยชุดข้อคำถามอ้างอิงจาก [5] และ [6]

**ความสามารถในการทดลองใช้ (Trialability):** เป็นแนวคิดจากทฤษฎีการแพร่กระจายนวัตกรรม (Diffusion of Innovations) ของ Rogers (2003) ซึ่งอธิบายว่าการที่ผู้ใช้มีโอกาสได้ทดลองใช้เทคโนโลยีก่อนตัดสินใจใช้งานจริง จะช่วยลดความกังวลและเพิ่มอัตราการยอมรับนวัตกรรม โดยข้อคำถามพัฒนามาจาก [7]

**การรับรู้ความเสี่ยง (Perceived Risk):** หมายถึง ความกังวลของผู้ใช้ต่อผลเสียที่อาจเกิดขึ้น เช่น ความเสี่ยงด้านความเป็นส่วนตัวหรือคุณภาพของเนื้อหา ซึ่ง [4] ชี้ว่าเป็นปัจจัยสำคัญที่ส่งผลต่อการตัดสินใจยอมรับเทคโนโลยีสมัยใหม่ โดยใช้ข้อคำถามที่พัฒนามาจาก [8] และ [9]

### 2.4 โมเดลสมการเชิงโครงสร้างแบบ PLS-SEM

โมเดลสมการเชิงโครงสร้าง (Structural Equation Modeling: SEM) เป็นเทคนิคทางสถิติที่ใช้ในการวิเคราะห์ความสัมพันธ์เชิงสาเหตุระหว่างตัวแปรแฝง (Latent Variables) ที่วัดผ่านตัวแปรสังเกตได้ (Observed

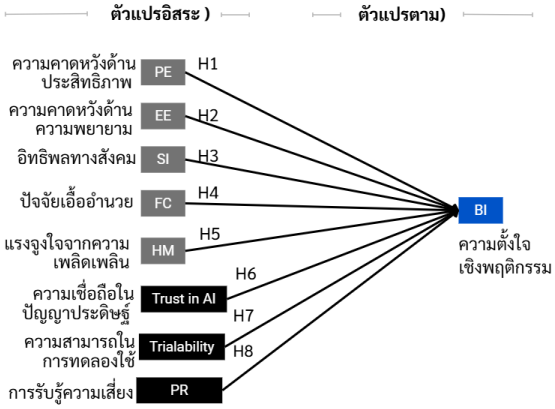
Variables) ปัจจุบัน SEM มี 2 แนวทางหลัก คือ 1) Covariance-Based SEM (CB-SEM) และ 2) Partial Least Squares SEM (PLS-SEM) ซึ่ง [10] ได้อธิบายว่าซึ่งมุ่งเน้นการทำนายและเหมาะสำหรับโมเดลที่มีความซับซ้อน ข้อมูลมีการแจกแจงไม่ปกติ หรือมีขนาดกลุ่มตัวอย่างไม่มากนัก งานวิจัยนี้จึงเลือกใช้เทคนิค PLS-SEM เพื่อทำนายความตั้งใจเชิงพฤติกรรม (Behavioral Intention) ภายใต้กรอบแนวคิด UTAUT2 เนื่องจากมีความสามารถในการประมาณค่าพารามิเตอร์ทั้งหมดในโมเดลได้พร้อมกัน และสอดคล้องกับแนวทางการรายงานผลการวิจัยที่มุ่งเน้นความแม่นยำในการพยากรณ์ตามข้อเสนอแนะของ [11]

## 3. วิธีการดำเนินการวิจัย

### 3.1 กรอบแนวคิดและสมมติฐาน

จากการทบทวนวรรณกรรมที่เกี่ยวข้อง งานวิจัยนี้ได้บูรณาการแนวคิด UTAUT2 ร่วมกับปัจจัยด้านความเชื่อถือในปัญญาประดิษฐ์ ความสามารถในการทดลองใช้ และการรับรู้ความเสี่ยง เพื่อกำหนดเป็น**กรอบแนวคิด**ในการศึกษาอิทธิพลเชิงสาเหตุที่มีต่อความตั้งใจเชิงพฤติกรรมในการใช้แพลตฟอร์ม AI-Microlearning โดยสามารถสรุปสมมติฐานการวิจัยได้ ดังนี้

- H1: ความคาดหวังด้านประสิทธิภาพ (Performance Expectancy) มีอิทธิพลเชิงบวกต่อความตั้งใจเชิงพฤติกรรม
- H2: ความคาดหวังด้านความพยายาม (Effort Expectancy) มีอิทธิพลเชิงบวกต่อความตั้งใจเชิงพฤติกรรม
- H3: อิทธิพลทางสังคม (Social Influence) มีอิทธิพลเชิงบวกต่อความตั้งใจเชิงพฤติกรรมในการใช้งาน
- H4: ปัจจัยเอื้ออำนวย (Facilitating Conditions) มีอิทธิพลเชิงบวกต่อความตั้งใจเชิงพฤติกรรม
- H5: แรงจูงใจจากความเพลิดเพลิน (Hedonic Motivation) มีอิทธิพลเชิงบวกต่อความตั้งใจเชิงพฤติกรรม
- H6: ความเชื่อถือในปัญญาประดิษฐ์ (Trust in AI) มีอิทธิพลเชิงบวกต่อความตั้งใจเชิงพฤติกรรม
- H7: ความสามารถในการทดลองใช้ (Trialability) มีอิทธิพลเชิงบวกต่อความตั้งใจเชิงพฤติกรรม
- H8: การรับรู้ความเสี่ยง (Perceived Risk) มีอิทธิพลเชิงลบต่อความตั้งใจเชิงพฤติกรรม




**ภาพที่ 1:** กรอบแนวคิดการวิจัย (Conceptual Framework)

### 3.2 การออกแบบการวิจัย

งานวิจัยนี้เป็นการวิจัยเชิงปริมาณ ที่ใช้ระเบียบวิธีวิจัยเชิงสำรวจ (Survey Research) ในการทดสอบและอธิบายความสัมพันธ์ของระหว่างปัจจัยต่าง ๆ โดยเก็บรวบรวมข้อมูลโดยใช้แบบสอบถามเป็นเครื่องมือหลัก

### 3.3 ประชากรและกลุ่มตัวอย่าง

ประชากรที่ใช้ในการวิจัย คือ พนักงานที่ปฏิบัติงานในองค์กรธุรกิจเทคโนโลยีในประเทศไทย กำหนดขนาดกลุ่มตัวอย่างด้วยการวิเคราะห์อำนาจทดสอบ (Power Analysis) ตามแนวทางของ [12] โดยมีระดับนัยสำคัญ ($\alpha$) ที่ 0.05, อำนาจทดสอบ (Power) ที่ 0.80 และขนาดอิทธิพล (Effect size) ระดับปานกลางที่ 0.15 ผลการคำนวณพบว่า ต้องการกลุ่มตัวอย่างขั้นต่ำจำนวน 62 คน อย่างไรก็ตาม เพื่อให้โมเดล PLS-SEM มีความเสถียรตามเกณฑ์ของ Hair [11] ผู้จึงดำเนินการเก็บรวบรวมข้อมูลด้วยวิธีการเลือกตัวอย่างแบบเฉพาะเจาะจง (Purposive Sampling) และแบบสะดวก (Convenience Sampling) จนได้กลุ่มตัวอย่างที่สมบูรณ์ทั้งสิ้นจำนวน 203 คน ซึ่งถือว่าครอบคลุมและเพียงพอต่อการวิเคราะห์ระดับลึก

### 3.4 เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัยประกอบด้วย 2 ส่วน ดังนี้

ระบบต้นแบบ (Functional Prototype): ได้พัฒนาต้นแบบแพลตฟอร์ม AI-Microlearning เพื่อใช้เป็นเครื่องมือกระตุ้น (Stimulus) ให้กลุ่มตัวอย่างเห็นภาพการ

ทำงานจริงก่อนตอบแบบสอบถาม โดยเน้นคุณสมบัติหลัก ได้แก่ 1) ระบบจัดการเนื้อหาขนาดเล็ก 2) ปัญญาประดิษฐ์ในการสรุปเนื้อหาและแนะนำบทเรียนแบบเฉพาะบุคคล และ 3) ระบบกลไกเกม เช่น การสะสมแต้มเพื่อแลกรางวัล

แบบสอบถาม (Questionnaire): แบ่งเป็น 3 ส่วน คือ 1) ข้อมูลทั่วไปของตอบแบบสอบถาม 2) ปัจจัยที่มีอิทธิพลต่อความตั้งใจใช้งาน จำนวน 9 ตัวแปร รวม 32 ข้อคำถาม โดยปรับปรุงมาจากชุดข้อคำถามที่เป็นมาตรฐานสากลตาม [3], [5], [7] และ [9] โดยใช้มาตรวัดแบบลิเคิร์ท (Likert Scale) 7 ระดับ และ 3) ข้อเสนอแนะเพิ่มเติม

**3.4 การตรวจสอบคุณภาพเครื่องมือ**

ดำเนินการได้ตรวจสอบคุณภาพเครื่องมือตามขั้นตอนมาตรฐาน ดังนี้

**ความตรงเชิงเนื้อหา (Content Validity):** ตรวจสอบโดยผู้เชี่ยวชาญจำนวน 3 ท่าน เพื่อหาค่าดัชนีความสอดคล้อง (IOC) ซึ่งผลการวิเคราะห์พบว่าข้อคำถามทุกข้อมีค่า IOC อยู่ระหว่าง 0.67 - 1.00 (ค่าเฉลี่ย 0.96) ซึ่งถือว่ามีความเที่ยงตรงผ่านเกณฑ์ที่กำหนด

**ความเชื่อมั่น (Reliability):** ดำเนินการทดสอบนำร่อง (Pilot Test) กับกลุ่มตัวอย่างที่มีลักษณะใกล้เคียงกับประชากรเป้าหมายจำนวน 32 คน ผลการวิเคราะห์ค่าสัมประสิทธิ์อัลฟาของครอนบาค (Cronbach's Alpha Coefficient) ของทุกตัวแปรมีค่าอยู่ระหว่าง 0.812 - 0.941 ซึ่งสูงกว่าเกณฑ์ 0.70 แสดงว่าเครื่องมือมีความเชื่อมั่นในระดับสูงมาก

## 4. ผลการดำเนินงาน

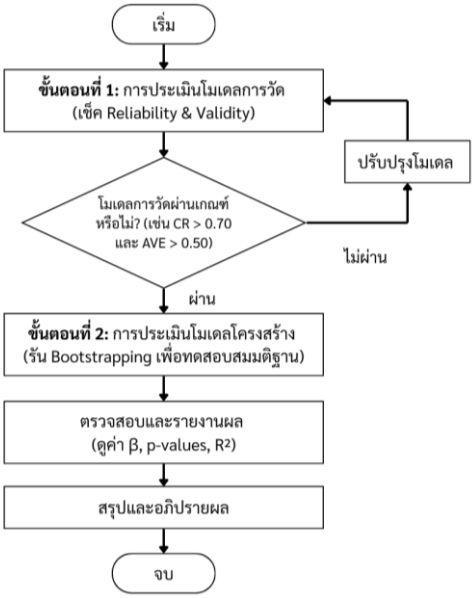
**4.1 ผลการวิเคราะห์ข้อมูลทั่วไปของกลุ่มตัวอย่าง**

จากการเก็บรวบรวมข้อมูลจากกลุ่มตัวอย่างจำนวน 203 คน พบว่าเป็นเพศชาย 51.72% ส่วนใหญ่อยู่ในช่วงอายุ 31-40 ปี (44.33%) สำเร็จการศึกษาระดับปริญญาตรี (5.86%) และปฏิบัติงานในกลุ่มผู้ให้บริการโครงสร้างพื้นฐานดิจิทัลมากที่สุด (40.39%)

ในการรายงานผลการวิเคราะห์ข้อมูล ได้จัดลำดับการนำเสนอตามแนวทาง Two-Step Approach เพื่อยืนยันความถูกต้องและน่าเชื่อถือของผลการวิจัย โดยเริ่มจากการประเมินคุณภาพของโมเดลการวัดก่อนดำเนินการทดสอบสมมติฐานในลำดับถัดไป โดยมีขั้นตอนดังภาพภาพที่ 2

**4.2 การประเมินโมเดลการวัด (Measurement Model)**

ผลการประเมินคุณภาพของเครื่องมือวัดพบว่า ตัวชี้วัดทุกรายการมีค่าน้ำหนักองค์ประกอบ (Outer Loadings) อยู่ระหว่าง 0.859 ถึง 0.956 ซึ่งสูงกว่าเกณฑ์ 0.708 ค่าความเชื่อมั่นรวม (CR) อยู่ระหว่าง 0.919 ถึง 0.964 และค่าเฉลี่ยความแปรปรวนที่สกัดได้ (AVE) อยู่ระหว่าง 0.791 ถึง 0.900 ซึ่งสูงกว่าเกณฑ์ 0.50 ทุกตัวแปร ดังแสดงในตารางที่ 1 ยืนยันว่าโมเดลการวัดมีความเชื่อมั่นและความเที่ยงตรงเชิงเข้า (Convergent Validity) ในระดับดีมาก นอกจากนี้ผลการทดสอบความเที่ยงตรงเชิงจำแนก (Discriminant Validity) ตามเกณฑ์ HTMT ดังตารางที่ 2 พบว่าทุกคู่มีค่าต่ำกว่า 0.85 จึงมั่นใจได้ว่าตัวแปรไม่มีความซ้ำซ้อนกัน




**ภาพที่ 2:** ขั้นตอนการวิเคราะห์แบบ Two-Step Approach

**ตารางที่ 1:** ผลการประเมินความเชื่อมั่นและเที่ยงตรงของโมเดลการวัด

| ตัวแปรแฝง | Outer Loadings | α | CR |
|---|---|---|---|
| Performance Expectancy | 0.865 - 0.927 | 0.885 | 0.929 |
| Effort Expectancy | 0.859 - 0.940 | 0.894 | 0.933 |
| Social Influence | 0.930 - 0.936 | 0.927 | 0.953 |
| Facilitating Conditions | 0.866 - 0.917 | 0.868 | 0.919 |
| Hedonic Motivation | 0.934 - 0.955 | 0.939 | 0.961 |
| Trust in AI | 0.935 - 0.956 | 0.945 | 0.964 |
| Trialability | 0.884 - 0.954 | 0.902 | 0.939 |
| Perceived Risk | 0.869 - 0.941 | 0.901 | 0.937 |
| Behavioral Intention | 0.923 - 0.938 | 0.922 | 0.950 |

**ตารางที่ 2:** ผลการตรวจสอบความเที่ยงตรงเชิงจำแนกด้วย HTMT

| | Performance Expectancy | Effort Expectancy | Social Influence | Facilitating Conditions | Hedonic Motivation | Trust in AI | Trialability | Perceived Risk | Behavioral Intention |
|---|---|---|---|---|---|---|---|---|---|
| PE | | | | | | | | | |
| EE | 0.68 | | | | | | | | |
| SI | 0.79 | 0.71 | | | | | | | |
| FC | 0.66 | 0.74 | 0.75 | | | | | | |
| HM | 0.68 | 0.72 | 0.67 | 0.71 | | | | | |
| TR | 0.70 | 0.67 | 0.72 | 0.71 | 0.75 | | | | |
| TB | 0.64 | 0.63 | 0.68 | 0.71 | 0.69 | 0.75 | | | |
| PR | 0.07 | 0.09 | 0.12 | 0.18 | 0.11 | 0.07 | 0.23 | | |
| BI | 0.77 | 0.68 | 0.77 | 0.67 | 0.74 | 0.77 | 0.70 | 0.14 | |

### 4.3 การประเมินโมเดลโครงสร้าง/ทดสอบสมมติฐาน

ในการวิเคราะห์ปัจจัยเชิงสาเหตุเพื่อตอบ **วัตถุประสงค์การวิจัยข้อที่ 1** งานวิจัยนี้ได้ประเมินโมเดลโครงสร้างเพื่อทดสอบสมมติฐาน โดยโมเดลการวิจัยสามารถอธิบายความแปรปรวนของความตั้งใจเชิงพฤติกรรม ($R^2$) ได้ร้อยละ 66.8 ดังตารางที่ 3 ซึ่งถือเป็นระดับอำนาจการทำนายที่อยู่ในเกณฑ์ปานกลางค่อนข้างสูงตามเกณฑ์ของ Hair [11]

ทั้งนี้ ผลการทดสอบสมมติฐานด้วยเทคนิค Bootstrapping จำนวน 5,000 รอบ พบปัจจัยที่มีอิทธิพลอย่างมีนัยสำคัญทางสถิติจำนวน 4 ปัจจัย สรุปได้ดังตารางที่ 4 และภาพที่ 3

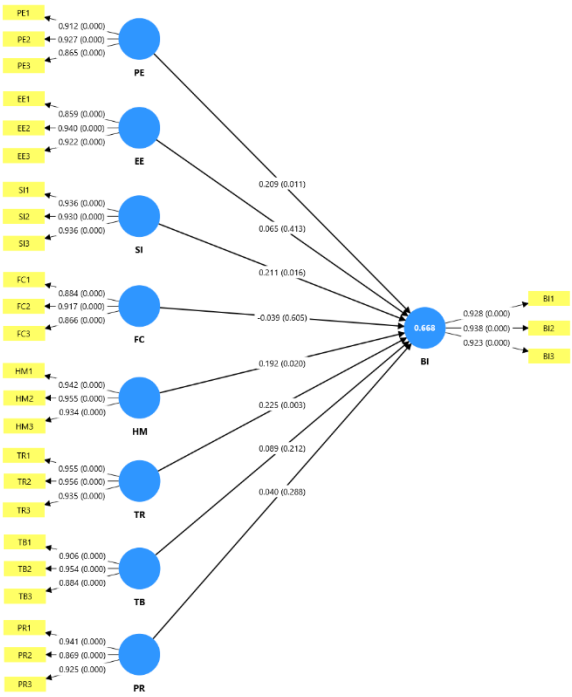
**ตารางที่ 3:** สัมประสิทธิ์การทำนาย

| ตัวแปรแฝงภายใน | $R^2$ | $R^2$Adjusted |
|---|---|---|
| Behavioral Intention | 0.668 | 0.654 |

**ตารางที่ 4:** ผลการทดสอบสมมติฐาน

| สมมติฐาน | เส้นทาง | $\beta$ | $t$-value |
|---|---|---|---|
| H1 | PE --> BI | 0.209 | 2.546 |
| H2 | EE --> BI | 0.065 | 0.818 |
| H3 | SI --> BI | 0.211 | 2.407 |
| H4 | FC --> BI | -0.039 | 0.517 |
| H5 | HM --> BI | 0.192 | 2.324 |
| H6 | TR --> BI | 0.225 | 3.004 |
| H7 | TB --> BI | 0.089 | 1.249 |
| H8 | PR --> BI | 0.040 | 1.062 |

จากภาพที่ 2 พบว่าโมเดลการวิจัยมีความสอดคล้องกับข้อมูลเชิงประจักษ์ในระดับดีมาก สำหรับผลการทดสอบสมมติฐานพบว่ามีสมมติฐานได้รับการสนับสนุน 4 ข้อ จากทั้งหมด 8 ข้อ โดย **ความเชื่อถือในปัญญาประดิษฐ์ (Trust in AI)** เป็นปัจจัยที่มีอิทธิพลเชิงบวกสูงสุด ($\beta = 0.225$, $p < 0.01$) รองลงมาคืออิทธิพลทางสังคม ($\beta = 0.211$) ความคาดหวังด้านประสิทธิภาพ ($\beta = 0.209$) และแรงจูงใจจากความเพลิดเพลิน ($\beta = 0.192$) ตามลำดับ




**ภาพที่ 3:** ผลการวิเคราะห์โมเดลสมการโครงสร้าง

## 5. สรุปผลและอภิปราย

### 5.1 สรุปผล

การศึกษาครั้งนี้ยืนยันว่าโมเดล UTAUT2 ที่บูรณาการร่วมกับปัจจัยด้านความเชื่อถือและการรับรู้ความเสี่ยง มีพลังในการทำนายความตั้งใจในการใช้แพลตฟอร์ม AI-Microlearning ได้อย่างมีนัยสำคัญ โดยพบว่า ปัจจัยด้านจิตวิทยาและความเชื่อมั่น (Trust in AI) และ ปัจจัยด้านประสบการณ์ (Hedonic Motivation) ร่วมกับ ปัจจัยเชิงหน้าที่ (Performance Expectancy) และ อิทธิพลเชิงสังคม (Social Influence) คือ แรงขับเคลื่อนสำคัญที่ทำให้พนักงานในองค์กรเทคโนโลยีตัดสินใจยอมรับเทคโนโลยี

นี้ ในขณะที่ปัจจัยด้านความยากง่ายและทรัพยากรสนับสนุนไม่ใช่เงื่อนไขหลักสำหรับกลุ่มตัวอย่างนี้

**5.2 อภิปรายผล**

1) **บทบาทของความเชื่อถือในปัญญาประดิษฐ์:** ผลการวิจัยที่พบว่าความเชื่อถือ (Trust) มีอิทธิพลสูงสุดสะท้อนว่าในยุคของ Generative AI ผู้ใช้งานไม่ได้มองแค่ความเก่ง ของระบบเท่านั้น แต่ยังให้ความสำคัญกับความถูกต้องและความโปร่งใสของเนื้อหา (AI Transparency) เนื่องจากพนักงานสายเทคโนโลยีมีความตระหนักถึงความเสี่ยงจากข้อมูลที่ผิดพลาด (AI Hallucination) ได้ดีกว่ากลุ่มผู้ใช้ทั่วไป

2) **คุณลักษณะเฉพาะของกลุ่มตัวอย่าง:** การที่ปัจจัยด้านความพยายาม (EE) และปัจจัยเอื้ออำนวย (FC) ไม่ส่งผลต่อความตั้งใจ อธิบายได้จากระดับตระหนักทางดิจิทัล (Digital Literacy) ที่สูงของพนักงานในองค์กรธุรกิจเทคโนโลยี ซึ่งมีความพร้อมเชิงอุปกรณ์และทักษะพื้นฐานอยู่แล้ว ทำให้ความง่ายในการใช้งานไม่ได้เป็นแรงจูงใจพิเศษในการตัดสินใจเหมือนกลุ่มประชากรอื่น

**5.3 ข้อเสนอแนะ**

เพื่อให้บรรลุตาม **วัตถุประสงค์การวิจัยข้อที่ 2** ในการจัดทำข้อเสนอเชิงกลยุทธ์สำหรับการนำเทคโนโลยีไปปรับใช้ งานวิจัยนี้จึงสรุปแนวทางปฏิบัติที่สำคัญจากผลการศึกษาเชิงประจักษ์ ดังนี้

จากผลการศึกษา ได้เสนอแนะให้องค์กรธุรกิจเทคโนโลยีให้ความสำคัญกับกลยุทธ์ Trust-First เป็นลำดับแรก โดยการสร้างความโปร่งใสของระบบปัญญาประดิษฐ์ (AI Transparency) และมีกระบวนการตรวจสอบความถูกต้องของเนื้อหาโดยผู้เชี่ยวชาญก่อนนำเสนอสู่ผู้ใช้ (Man in the loop) ควบคู่ไปกับการสื่อสารอรรถประโยชน์เชิงหน้าที่ที่ชัดเจนและการใช้กลไกเกม (Gamification) เพื่อสร้างความสนุกในการเรียนรู้

สำหรับการวิจัยในอนาคต ควรขยายขอบเขตการศึกษาไปยังกลุ่มประชากรที่มีระดับทักษะดิจิทัลที่หลากหลายในอุตสาหกรรมอื่น และควรเปลี่ยนมาวัดผลที่พฤติกรรมการใช้งานจริง (Actual Use Behavior) เพื่อให้เห็นผลกระทบที่แท้จริงของเทคโนโลยีต่อการพัฒนาบุคลากรในระยะยาว

## เอกสารอ้างอิง


[1] W. K. Monib et al., “The impact of microlearning on knowledge acquisition and retention: A systematic literature review,” *Educational Technology Research and Development*, 2025.

[2] S. M. Khan, *Fundamentals of AI for Medical Education, Research and Practice (1st ed.)*, Academic Press, 2025.

[3] V. Venkatesh, J. Y. L. Thong, and X. Xu, “Consumer acceptance and use of information technology: Extending the unified theory of acceptance and use of technology,” *MIS Quarterly*, vol. 36, no. 1, pp. 157–178, 2012.

[4] S. C. Kong, Y. Yang, and C. Hou, “Examining teachers' behavioral intention of using generative artificial intelligence tools for teaching and learning based on the extended technology acceptance model,” *Computer and Education: Artificial Intelligence*, Vol. 7, 2024.

[5] D. H. McKnight, V. Choudhury, and C. Kacmar, “Developing and validating trust measures for e-commerce: An integrative typology,” *Information Systems Research*, Vol. 13, No. 3, pp. 334-359, 2002.

[6] V. Venkatesh and H. Bala, “Technology acceptance model 3 and a research agenda on interventions,” *Decision Sciences*, Vol. 39, No. 2, pp. 273-315, 2008.

[7] E. M. Rogers, *Diffusion of innovations (5th ed.)*, Free Press: New York, 2003.

[8] G. C. Moore and I. Benbasat, “Development of an instrument to measure the perceptions of adopting an information technology innovation,” *Information Systems Research*, Vol. 2, No. 3, pp. 192-222, 1991.

[9] M. S. Featherman and P. A. Pavlou, “Predicting e-services adoption: a perceived risk facets perspective,” *International Journal of Human-Computer Studies*, Vol. 59, No. 4, pp. 451-474, 2003.

[10] J. F. Hair, C. M. Ringle, and M. Sarstedt, “PLS-SEM: Indeed a silver bullet,” *Journal of Marketing Theory and Practice*, Vol. 19, No. 2, pp. 139-152, 2011.

[11] J. F. Hair, G. T. M. Hult, C. M. Ringle, and M. Sarstedt, *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM) (2nd ed.)*, Sage Publications: Thousand Oaks, CA, 2017.

[12] J. Cohen, *Statistical power analysis for the behavioral sciences (2nd ed.)*, Lawrence Erlbaum Associates: Hillsdale, NJ, 1988.

# การวิเคราะห์เชิงเปรียบเทียบประสิทธิภาพของกลยุทธ์ SEO แบบคีย์เวิร์ดเดียว และหลายคีย์เวิร์ดสำหรับเว็บไซต์มหาวิทยาลัยวลัยลักษณ์

# A Comparative Analysis of the Effectiveness of Single-Keyword and Multi-Keyword SEO Strategies for Walailak University Website

*วชิรศักดิ์ โภคากรณ์ (Wachirasak Pokakorn)*
*ศูนย์นวัตกรรมดิจิทัล มหาวิทยาลัยวลัยลักษณ์*
*pwachira@wu.ac.th*

**บทคัดย่อ**

*งานวิจัยนี้มีวัตถุประสงค์เพื่อศึกษาและเปรียบเทียบประสิทธิภาพของกลยุทธ์การทำ Search Engine Optimization (SEO) ระหว่างแนวทางการใช้คีย์เวิร์ดเดียวต่อหนึ่งหน้าเว็บ (Single Keyword per Page) และการใช้หลายคีย์เวิร์ดต่อหนึ่งหน้าเว็บ (Multiple Keywords per Page) ในบริบทของเว็บไซต์สถาบันอุดมศึกษา โดยใช้กรณีศึกษาเว็บไซต์ Walailak University Research Portal และเว็บไซต์ Walailak University Graduate School การวิจัยใช้ข้อมูลจาก Google Search Console (GSC), Google Analytics 4 (GA4) และเครื่องมือ Ahrefs ในช่วงเดือนตุลาคม พ.ศ. 2567 ถึงเดือนกรกฎาคม พ.ศ. 2568 รวมระยะเวลา 10 เดือน ตัวชี้วัดที่ใช้ในการวิเคราะห์ประกอบด้วย จำนวนคีย์เวิร์ดที่ติดอันดับ ค่าเฉลี่ยตำแหน่งการแสดงผล (Average Position) อัตราการคลิก (CTR) ปริมาณผู้เข้าชมจากการค้นหาแบบธรรมชาติ (Organic Traffic) ระยะเวลาเฉลี่ยที่ผู้ใช้งานอยู่บนหน้าเว็บ และอัตราการออกจากหน้าเว็บ ผลการศึกษาพบว่า เว็บไซต์ที่ใช้กลยุทธ์ Single Keyword per Page มีค่า CTR และคุณภาพการมีส่วนร่วมของผู้ใช้งานสูงกว่า ขณะที่เว็บไซต์ที่ใช้ Multiple Keywords per Page สามารถสร้างปริมาณผู้เข้าชมและขยายขอบเขตการมองเห็นในผลการค้นหาได้มากกว่า งานวิจัยนี้เสนอแนวทางการใช้กลยุทธ์ SEO แบบผสมผสาน (Hybrid SEO Strategy) เพื่อเพิ่มทั้งประสิทธิภาพด้านคุณภาพและปริมาณการเข้าถึงเว็บไซต์สถาบันอุดมศึกษาอย่างยั่งยืน*

**คำสำคัญ:** การปรับแต่งเว็บไซต์เพื่อเครื่องมือค้นหา, SEO, คีย์เวิร์ด, เว็บไซต์มหาวิทยาลัย, Webometrics

**Abstract**

*This study aims to compare the effectiveness of two Search Engine Optimization (SEO) strategies, namely Single Keyword per Page and Multiple Keywords per Page, within the context of university websites. The case studies include the Walailak University Research Portal and the Walailak University Graduate School website. Data were collected from Google Search Console (GSC), Google Analytics 4 (GA4), and Ahrefs over a ten-month period from October 2024 to July 2025. The analysis employed key SEO performance indicators, including the number of ranking keywords, average search position, click-through rate (CTR), organic traffic, average time on page, and bounce rate. The results indicate that the Single Keyword per Page strategy achieves higher CTR and better user engagement, reflecting stronger content and query relevance. In contrast, the Multiple Keywords per Page strategy generates greater organic traffic and broader search visibility by capturing a wider range of search queries. Based on these findings, this study proposes a Hybrid SEO Strategy that integrates both approaches to balance ranking quality and traffic*

*volume. The proposed strategy provides practical implications for improving the visibility and performance of academic and institutional websites, particularly in relation to Webometrics rankings.*

**Keywords:** Search Engine Optimization, Keyword Strategy, University Websites, Organic Search, Webometrics.

## 1. บทนำ

มหาวิทยาลัยวลัยลักษณ์มีนโยบายในการขับเคลื่อนการจัดอันดับ Webometrics Ranking อย่างเป็นระบบ เพื่อเพิ่มการมองเห็นในระดับสากล โดยมุ่งพัฒนาเว็บไซต์หลักของมหาวิทยาลัยและหน่วยงานภายใน เช่น Walailak University Research Portal และ Graduate School ให้มีประสิทธิภาพและเข้าถึงได้ง่าย เว็บไซต์เหล่านี้เป็นแหล่งรวบรวมข้อมูลบุคลากร ผลงานวิจัย และองค์ความรู้ที่สำคัญ ซึ่งมีบทบาทโดยตรงต่อการสืบค้นผ่านเครื่องมือค้นหา เช่น Google ดังนั้น การพัฒนาโครงสร้างและเนื้อหาให้สามารถแสดงผลในลำดับต้นของผลการค้นหา จึงเป็นปัจจัยสำคัญในการเพิ่มโอกาสการเข้าถึงข้อมูลและสนับสนุนคะแนน Webometrics ของมหาวิทยาลัย

การจัดอันดับผลการค้นหาอาศัยอัลกอริทึมของเครื่องมือค้นหา ผู้พัฒนาเว็บไซต์จึงจำเป็นต้องเข้าใจหลักการดังกล่าวเพื่อปรับปรุงโครงสร้างและเนื้อหาให้สอดคล้องกับเกณฑ์การจัดอันดับ การทำ Search Engine Optimization (SEO) จึงเป็นกลยุทธ์สำคัญที่ช่วยเพิ่มจำนวนผู้เข้าชมและยกระดับการมองเห็นของเว็บไซต์ อย่างไรก็ตาม การศึกษาเชิงเปรียบเทียบเกี่ยวกับการใช้คีย์เวิร์ดเดียวต่อหนึ่งหน้าเว็บและการใช้หลายคีย์เวิร์ดต่อหนึ่งหน้าเว็บในบริบทของเว็บไซต์สถาบันอุดมศึกษา ยังมีอยู่จำกัด เนื่องจากงานวิจัยส่วนใหญ่มุ่งเน้นเว็บไซต์เชิงพาณิชย์

ดังนั้น งานวิจัยนี้จึงมุ่งศึกษาผลของกลยุทธ์การใช้คีย์เวิร์ดในระดับหน้าเว็บของเว็บไซต์สถาบันอุดมศึกษา โดยเปรียบเทียบประสิทธิผลระหว่างการใช้คีย์เวิร์ดเดียวและหลายคีย์เวิร์ดต่อหน้าเว็บ ผลการศึกษาที่ได้จะเป็นข้อมูลเชิงประจักษ์สำหรับใช้เป็นแนวทางในการพัฒนาเว็บไซต์ของหน่วยงานภายในมหาวิทยาลัยให้มีประสิทธิภาพในการปรากฏในลำดับต้นของผลการค้นหา Google และสนับสนุนการยกระดับอันดับ Webometrics ของมหาวิทยาลัยอย่างยั่งยืน

## 2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

### 2.1 แนวคิดเกี่ยวกับการทำงานของโปรแกรมค้นหา (Search Engine)

Search Engine คือเครื่องมือสำหรับสืบค้นข้อมูลบนเครือข่ายอินเทอร์เน็ต โดยแสดงผลลัพธ์ตามคำค้นหา (Keyword) ที่ผู้ใช้งานป้อนเข้าไป ซึ่งผลลัพธ์อาจอยู่ในภาพของเว็บไซต์ ภาพ วิดีโอ หรือข้อมูลเชิงโครงสร้างอื่น ๆ การดำเนินการด้าน Search Engine Optimization (SEO) อย่างมีประสิทธิภาพจึงจำเป็นต้องอาศัยความเข้าใจเกี่ยวกับกลไกและกระบวนการทำงานของโปรแกรมค้นหา เพื่อปรับปรุงโครงสร้างและเนื้อหาของเว็บไซต์ให้สอดคล้องกับหลักเกณฑ์การจัดอันดับ

งานวิจัยนี้มุ่งศึกษากระบวนการทำงานของ Google Search ซึ่งเป็น Search Engine ที่ได้รับความนิยมสูงสุดในปัจจุบัน โดยกระบวนการทำงานสามารถแบ่งออกเป็น 3 ขั้นตอนหลัก ได้แก่ (1) การรวบรวมข้อมูล (Crawling) ซึ่ง Google ใช้โปรแกรม Googlebot ในการเข้าถึงและเก็บข้อมูลจากหน้าเว็บ (2) การจัดทำดัชนี (Indexing) โดยวิเคราะห์เนื้อหาและองค์ประกอบเชิงโครงสร้างของหน้าเว็บ และ (3) การจัดอันดับและแสดงผลการค้นหา (Ranking) ซึ่งเป็นขั้นตอนที่อัลกอริทึมของ Google ประมวลผลคำค้นหาและจัดลำดับผลลัพธ์บนหน้า Search Engine Results Pages (SERPs)

จากการศึกษาของศุภณัฐ สุขโข (2564) [1] ระบุว่าขั้นตอนการจัดอันดับและแสดงผลการค้นหามีความสำคัญอย่างยิ่งต่อประสิทธิภาพของหน้าเว็บ เนื่องจาก Google ใช้ปัจจัยในการจัดอันดับมากกว่า 200 ปัจจัย และมีการปรับปรุงอัลกอริทึมอย่างต่อเนื่องเพื่อให้สอดคล้องกับเจตนาการค้นหาของผู้ใช้งาน (User Intent) (Schwartz, 2024) [4]

**2.2 หลักการทำงานของ Search Engine**

Search Engine อาศัยโปรแกรมรวบรวมข้อมูลที่เรียกว่า Crawler หรือ Spider ในการเข้าถึงและเก็บข้อมูลจากเว็บไซต์ต่าง ๆ ทั้งเว็บไซต์ใหม่และเว็บไซต์ที่มีการปรับปรุงเนื้อหาอย่างสม่ำเสมอ (Grzywaczewshi et al., 2010) [2] แม้งานวิจัยบางส่วนจะเรียกโปรแกรมดังกล่าวแตกต่างกัน เช่น Web Crawler, Spider หรือ Robot (Jing & Shi et al., 2010) [3] แต่มีวัตถุประสงค์ร่วมกันคือการรวบรวมและประมวลผลข้อมูลเพื่อนำไปใช้ในการจัดอันดับผลการค้นหา

ในขั้นตอนการจัดทำดัชนี เว็บไซต์จะถูกวิเคราะห์ทั้งด้านเนื้อหาและโครงสร้าง ก่อนจัดเก็บข้อมูลไว้ในฐานข้อมูลของ Search Engine โดยปัจจัยสำคัญที่ใช้ในการจัดทำดัชนี ได้แก่ ชื่อเรื่องของหน้าเว็บ (Title), โครงสร้าง URL, เนื้อหาในหน้าเว็บ และภาพแบบการเชื่อมโยงแบบไฮเปอร์ลิงก์ (Zhu & Wu, 2021) [6]

**2.3 แนวคิดเกี่ยวกับ SEO**

SEO (Search Engine Optimization) หมายถึง กระบวนการเพิ่มประสิทธิภาพของเว็บไซต์เพื่อให้สามารถแสดงผลในตำแหน่งที่เหมาะสมบนหน้าผลลัพธ์การค้นหา (SERPs) โดยมีเป้าหมายเพื่อเพิ่มทั้งปริมาณและคุณภาพของการเข้าชมจากผลการค้นหาแบบธรรมชาติ [5] การทำ SEO ที่มีประสิทธิภาพจำเป็นต้องอาศัยความเข้าใจเกี่ยวกับกลไกการทำงานของ Search Engine เพื่อปรับโครงสร้างและเนื้อหาของเว็บไซต์ให้สอดคล้องกับหลักเกณฑ์การจัดอันดับ โดยงานวิจัยนี้เลือกศึกษา Google Search เป็นหลัก เนื่องจากเป็นเครื่องมือค้นหาที่มีผู้ใช้งานมากที่สุดในปัจจุบัน

## 3. วิธีการดำเนินการวิจัย

**3.1 การกำหนดกลุ่มเว็บไซต์เป้าหมาย**

การวิจัยนี้เลือกเว็บไซต์เป้าหมายจำนวน 2 เว็บไซต์ภายใต้สังกัดมหาวิทยาลัยวลัยลักษณ์ ซึ่งมีแนวทางการดำเนินการ SEO ที่แตกต่างกันอย่างชัดเจน เพื่อใช้ในการเปรียบเทียบกลยุทธ์การใช้คีย์เวิร์ดในระดับหน้าเว็บ ได้แก่ 1. เว็บไซต์ Walailak University Research Portal ซึ่งใช้กลยุทธ์ Single Keyword per Page โดยแต่ละหน้าเว็บมุ่งเน้นคีย์เวิร์ดหลักเพียงหนึ่งคำหรือกลุ่มคำที่มีความหมายใกล้เคียงกัน และ 2. เว็บไซต์ Walailak University Graduate School ซึ่งใช้กลยุทธ์ Multiple Keywords per Page โดยแต่ละหน้าเว็บมีการใช้คีย์เวิร์ดหลายคำที่มีความเกี่ยวข้องกันเพื่อรองรับการค้นหาที่หลากหลายและเพิ่มโอกาสในการปรากฏในผลการค้นหา

**3.2 ตัวชี้วัดที่ใช้ในการเปรียบเทียบ**

ตัวชี้วัดด้าน SEO ที่ใช้ในการศึกษา ประกอบด้วย

(1) จำนวนคีย์เวิร์ดที่ติดอันดับ Google จำแนกตาม Top 3, Top 10 และ Top 20 รวมถึงค่าเฉลี่ยตำแหน่งแสดงผล

(2) ค่าเฉลี่ยตำแหน่งแสดงผล (Average Position) จาก Google Search Console

(3) อัตราการคลิก (Click-Through Rate: CTR) คำนวณจาก (Clicks ÷ Impressions) × 100

(4) ปริมาณผู้เข้าชมจาก Organic Traffic โดยใช้ข้อมูลจาก Google Analytics 4 และเปรียบเทียบค่าเฉลี่ยรายเดือน

(5) ตัวชี้วัดด้านการมีส่วนร่วมของผู้ใช้งาน ได้แก่ Average Time on Page และ Bounce Rate เพื่อประเมินคุณภาพและความน่าสนใจของเนื้อหาเว็บไซต์

**3.3 การวิเคราะห์ข้อมูล** การวิเคราะห์ข้อมูลมีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพของเว็บไซต์ที่ใช้กลยุทธ์ Single Keyword per Page และ Multiple Keywords per Page โดยดำเนินการดังนี้

**3.3.1 การเตรียมข้อมูล (Data Preparation)** ดึงข้อมูลจาก Google Search Console, Google Analytics 4 และ Ahrefs/SEMrush โดยใช้ค่าเฉลี่ยรายเดือนตลอดระยะเวลา 10 เดือน เพื่อลดผลกระทบจากความผันผวนระยะสั้น กำหนดช่วงเวลาของข้อมูลให้ตรงกันทั้งสองเว็บไซต์ และคัดกรองเฉพาะข้อมูลจาก Organic Search รวมถึงคีย์เวิร์ดที่มี Impressions อย่างน้อย 1 ครั้ง

**3.3.2 การวิเคราะห์เชิงพรรณนา (Descriptive Statistics)** ใช้สถิติเชิงพรรณนาเพื่ออธิบายลักษณะทั่วไปของข้อมูลในแต่ละตัวชี้วัด โดยรายงานค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน ค่าสูงสุด ค่าต่ำสุด และแนวโน้มการเปลี่ยนแปลงของข้อมูล

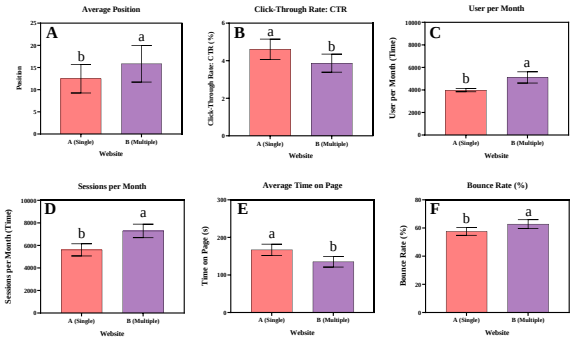
**3.3.3 การวิเคราะห์เชิงอนุมาน (Inferential Statistics)** ทดสอบความแตกต่างของค่าเฉลี่ยตัวชี้วัดระหว่างสองกลยุทธ์ SEO โดยใช้สถิติ Independent Samples t-test ภายใต้สมมติฐานทางสถิติที่เหมาะสม

**3.3.4 การวิเคราะห์เชิงพหุตัวแปร** ใช้การวิเคราะห์องค์ประกอบหลัก (Principal Component Analysis: PCA) การวิเคราะห์สหสัมพันธ์ (Correlation) และการวิเคราะห์ถดถอยเชิงเส้นพหุคูณ (Multiple linear regression) เพื่ออธิบายความสัมพันธ์ของตัวชี้วัดด้าน SEO ได้แก่ ค่าเฉลี่ยจำนวนคีย์เวิร์ดต่อเดือน, Top 3 (%), Top 10 (%), Top 20 (%), Average Position, CTR (%), Users/เดือน, Sessions/เดือน, Average Time on Page และ Bounce Rate (%) ภายใต้แนวทางการดำเนินการ SEO ที่แตกต่างกัน โดยการวิเคราะห์ทั้งหมดดำเนินการด้วยโปรแกรม R (เวอร์ชัน 4.3.1; R Foundation for Statistical Computing, Vienna, Austria)

## 4. ผลการดำเนินงาน

การเก็บรวบรวมและวิเคราะห์ข้อมูลดำเนินการในช่วงเดือนตุลาคม พ.ศ. 2567 ถึงเดือนกรกฎาคม พ.ศ. 2568 รวมระยะเวลา 10 เดือน โดยใช้เครื่องมือ Google Search Console (GSC) สำหรับวิเคราะห์จำนวนคีย์เวิร์ดที่ติดอันดับ ค่าเฉลี่ยตำแหน่งแสดงผล (Average Position) และอัตราการคลิก (Click-Through Rate: CTR) ใช้ Google Analytics 4 (GA4) สำหรับวิเคราะห์ปริมาณการเข้าชมเว็บไซต์จากการค้นหาแบบธรรมชาติ (Organic Traffic) รวมถึงค่า Average Time on Page และ Bounce Rate และใช้เครื่องมือ Ahrefs เพื่อช่วยตรวจสอบและยืนยันจำนวนคีย์เวิร์ดและตำแหน่งการจัดอันดับในผลการค้นหา

กลุ่มตัวอย่างที่ใช้ในการศึกษาประกอบด้วย 2 เว็บไซต์ ได้แก่ เว็บไซต์ A คือ Walailak University Research Portal ซึ่งใช้กลยุทธ์ Single Keyword per Page และเว็บไซต์ B คือ Walailak University Graduate School ซึ่งใช้กลยุทธ์ Multiple Keywords per Page การเปรียบเทียบตำแหน่งแสดงผล, อัตราการคลิก (CTR) จากการค้นหาแบบธรรมชาติระหว่างเว็บไซต์ที่ใช้กลยุทธ์ Single Keyword per Page และ Multiple Keywords per Page ตลอดระยะเวลา 10 เดือน (ภาพที่ 1 A-F) ผลการวิเคราะห์พบว่าเว็บไซต์ที่ใช้กลยุทธ์ Single Keyword per Page มีค่า CTR สูงกว่าอย่างสม่ำเสมอตลอดช่วงเวลาการศึกษา โดยมีค่าเฉลี่ยเท่ากับ 4.6% ขณะที่เว็บไซต์ที่ใช้กลยุทธ์ Multiple Keywords per Page มีค่าเฉลี่ย 3.9% ทั้งนี้ ค่า CTR ของกลยุทธ์ Single Keyword per Page มีค่าสูงสุดในเดือนธันวาคม พ.ศ. 2567 เดือนเมษายน พ.ศ. 2568 และเดือนกรกฎาคม พ.ศ. 2568 (4.7%) สะท้อนให้เห็นว่าการมุ่งเน้นคีย์เวิร์ดที่เฉพาะเจาะจงต่อหนึ่งหน้าเว็บช่วยเพิ่มโอกาสในการคลิกจากผลการค้นหาได้อย่างมีประสิทธิภาพ




**ภาพที่ 1** : การเปรียบเทียบจำนวนผู้ใช้จาก Organic Search ระหว่าง Single Keyword และ Multiple Keywords

ภาพที่ 1 ยังแสดงการเปรียบเทียบจำนวนผู้ใช้จากการค้นหาแบบธรรมชาติระหว่างสองกลยุทธ์ SEO ผลการวิเคราะห์พบว่าเว็บไซต์ที่ใช้กลยุทธ์ Multiple Keywords per Page มีจำนวนผู้ใช้เฉลี่ยต่อเดือนสูงกว่าเว็บไซต์ที่ใช้กลยุทธ์ Single Keyword per Page ประมาณ 1,113 คน และมีจำนวนผู้ใช้สูงสุดในเดือนกรกฎาคม พ.ศ. 2568 เท่ากับ 5,220 คน แสดงให้เห็นว่าการใช้หลายคีย์เวิร์ดต่อหน้าเว็บช่วยเพิ่มการเข้าถึงและขยายฐานผู้ใช้งานได้ดีกว่า (ภาพที่ 2)




**ภาพที่ 2**: แสดงข้อมูลการเปรียบเทียบจำนวนผู้ใช้ระหว่าง Single Keyword และ Multiple Keywords ตลอด 10 เดือน

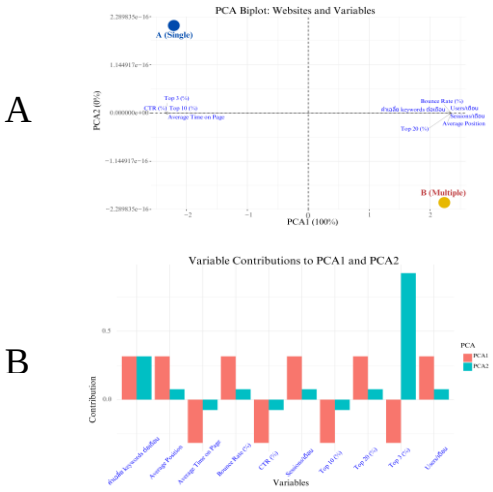
นอกจากนี้ ยังพบว่าเว็บไซต์ B มีจำนวนคีย์เวิร์ดรวมที่ติดอันดับในผลการค้นหามากกว่าเว็บไซต์ A อย่างมีนัยสำคัญทางสถิติ ($p < 0.05$) อย่างไรก็ตาม เว็บไซต์ A มีสัดส่วนคีย์เวิร์ดที่ติดอันดับในกลุ่ม Top 3 และ Top 10 สูงกว่า สะท้อนถึงความแม่นยำและคุณภาพของการจัดอันดับคีย์เวิร์ดภายใต้กลยุทธ์ Single Keyword per Page

**ตารางที่ 1** จำนวนคีย์เวิร์ดที่ติดอันดับ Google

| เว็บไซต์ | ค่าเฉลี่ยต่อเดือน (คำ) | Top 3 (%) | Top 10 (%) | Top 20 (%) |
|---|---|---|---|---|
| Single | 192 | 16.2% | 36.8% | 58.4% |
| Multiple | 254 | 11.4% | 28.7% | 61.5% |

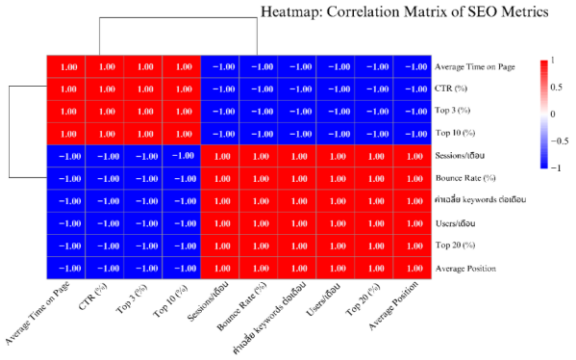
**การวิเคราะห์เชิงพหุตัวแปร**

ผลการวิเคราะห์องค์ประกอบหลัก (Principal Component Analysis: PCA) พบว่า องค์ประกอบหลักที่หนึ่ง (PCA1) สามารถอธิบายความแตกต่างระหว่างเว็บไซต์ทั้งสองได้อย่างชัดเจน (ภาพที่ 3) โดยเว็บไซต์ A (Single) มีค่า PCA1 เป็นลบ ขณะที่เว็บไซต์ B (Multiple) มีค่า PCA1 เป็นบวก แสดงให้เห็นว่า PCA1 เป็นแกนที่สะท้อนความแตกต่างเชิงกลยุทธ์ของการดำเนินการ SEO ในภาพรวม ทั้งนี้ PCA1 มีความสัมพันธ์เชิงบวกกับตัวแปรด้านปริมาณการเข้าชมและการมองเห็นของเว็บไซต์ ได้แก่ ค่าเฉลี่ยจำนวนคีย์เวิร์ดต่อเดือน, Top 20 (%), Average Position, Users/เดือน, Sessions/เดือน และ Bounce Rate (%) และมีความสัมพันธ์เชิงลบกับตัวแปรด้านคุณภาพของการจัดอันดับและการมีส่วนร่วมของผู้ใช้งาน ได้แก่ Top 3 (%), Top 10 (%), CTR (%) และ Average Time on Page ผลดังกล่าวสะท้อนให้เห็นว่าเว็บไซต์ที่ใช้หลายคีย์เวิร์ดต่อเพจมีแนวโน้มเพิ่มปริมาณการเข้าชมและขอบเขตการมองเห็นได้มากกว่า แต่มีข้อจำกัดด้านคุณภาพของการมีส่วนร่วมของผู้ใช้งาน สำหรับองค์ประกอบหลักที่สอง (PCA2) พบว่ามีความสัมพันธ์เด่นชัดกับตัวแปร Top 3 (%) เพียงตัวแปรเดียว ขณะที่ตัวแปรอื่นมีอิทธิพลในระดับต่ำ และตำแหน่งของเว็บไซต์ทั้งสองบนแกน PCA2 อยู่ใกล้เคียงศูนย์ แสดงให้เห็นว่าความแตกต่างหลักของแนวทาง SEO ระหว่างเว็บไซต์เกิดขึ้นตามแกน PCA1 เป็นสำคัญ




**ภาพที่ 3:** การวิเคราะห์องค์ประกอบหลัก (Principal Component Analysis: PCA) เพื่อศึกษาความสัมพันธ์เชิงพหุตัวแปรของตัวชี้วัดด้าน SEO ภายใต้แนวทางการดำเนินการ SEO ที่แตกต่างกัน

ผลการวิเคราะห์เมทริกซ์สหสัมพันธ์ในรูปแบบ heatmap (ภาพที่ 4) แสดงให้เห็นว่าตัวแปรสามารถจัดกลุ่มความสัมพันธ์ได้เป็นสองกลุ่มหลัก ได้แก่ กลุ่มตัวแปรด้านปริมาณการเข้าชมและการมองเห็นของเว็บไซต์ และกลุ่มตัวแปรด้านคุณภาพของการจัดอันดับและการมีส่วนร่วมของผู้ใช้งาน โดยกลุ่มตัวแปรด้านปริมาณ ได้แก่ ค่าเฉลี่ยจำนวนคีย์เวิร์ดต่อเดือน, Users/เดือน และ Sessions/เดือน มีความสัมพันธ์เชิงบวกซึ่งกันและกัน รวมถึงกับ Top 20 (%) และ Bounce Rate (%) สะท้อนว่าการเพิ่มปริมาณการเข้าชมอาจมาพร้อมกับคุณภาพการมีส่วนร่วมที่ลดลงในทางตรงกันข้าม กลุ่มตัวแปรด้านคุณภาพ ได้แก่ Top 3 (%), Top 10 (%), CTR (%) และ Average Time on Page มีความสัมพันธ์เชิงบวกซึ่งกันและกัน และมีความสัมพันธ์เชิงลบกับ Bounce Rate (%) แสดงถึงความสอดคล้องระหว่างคำค้นหาและเนื้อหาที่ช่วยยกระดับการมีส่วนร่วมของผู้ใช้งาน




**ภาพที่ 4:** แสดงเมทริกซ์สหสัมพันธ์ในภาพแบบ heatmap

เมื่อพิจารณาผลจากภาพที่ 3 และภาพที่ 4 ร่วมกัน พบว่า Correlation matrix (heatmap) สนับสนุนการตีความผล PCA อย่างสอดคล้องกัน โดย PCA1 สามารถอธิบายลักษณะการแยกกลุ่มของตัวแปรด้านปริมาณและด้านคุณภาพได้อย่างเป็นระบบ ผลการวิเคราะห์นี้ชี้ให้เห็นถึงลักษณะการแลกเปลี่ยนระหว่างการเพิ่มปริมาณการเข้าชมเว็บไซต์กับการรักษาคุณภาพของการจัดอันดับและการมีส่วนร่วมของผู้ใช้งาน ซึ่งมีนัยสำคัญต่อการกำหนดกลยุทธ์ SEO สำหรับเว็บไซต์เชิงวิชาการและสถาบันอุดมศึกษา

เพื่อเสริมการตีความเชิงพหุตัวแปร ผู้วิจัยได้ใช้การวิเคราะห์ถดถอยเชิงเส้นพหุคูณ โดยกำหนดตัวแปรตามเป็น Users/เดือน, CTR (%) และ Average Time on Page และตัวแปรอิสระเป็นตัวชี้วัดด้าน SEO ผลการวิเคราะห์พบว่า ตัวแปรด้านปริมาณการมองเห็น (เช่น จำนวนคีย์เวิร์ด และ Top 20 (%)) มีความสัมพันธ์เชิงบวกกับ Users/เดือน ($p < 0.05$) ขณะที่ตัวแปรด้านคุณภาพ (เช่น Top 3 (%), CTR (%)) มีความสัมพันธ์เชิงบวกกับ Average Time on Page และเชิงลบกับ Bounce Rate (%)

ผลดังกล่าวสอดคล้องกับ PCA โดยตัวแปรด้านปริมาณสัมพันธ์กับ ปริมาณการเข้าชมเว็บไซต์ ขณะที่ตัวแปรด้านคุณภาพสัมพันธ์กับ การมีส่วนร่วมของผู้ใช้งาน อย่างไรก็ตาม ผลลัพธ์สะท้อนเพียงความสัมพันธ์เชิงสถิติ และไม่สามารถตีความเชิงเหตุและผลได้โดยตรง แม้ผลจาก PCA, correlation และ regression จะสอดคล้องกัน แต่เนื่องจากเป็นข้อมูลเชิงสังเกต การตีความเชิงเหตุและผลยังมีข้อจำกัด ดังนั้น งานวิจัยในอนาคตควรใช้วิธีที่รองรับ causal inference เช่น regression ที่ควบคุมตัวแปรแทรกซ้อน หรือ quasi-experimental design (เช่น Difference-in-Differences) ได้

## 5. สรุป

การศึกษาเชิงเปรียบเทียบกลยุทธ์ SEO จากกรณีศึกษาเว็บไซต์ Walailak University Research Portal และ Graduate School ในช่วงเดือนตุลาคม พ.ศ. 2567 ถึงเดือนกรกฎาคม พ.ศ. 2568 โดยใช้ข้อมูลจาก Google Search Console (GSC), Google Analytics 4 (GA4) และเครื่องมือ Ahrefs สามารถสรุปข้อค้นพบที่สำคัญได้ดังนี้ประการแรก ผลการวิเคราะห์ข้อมูลจาก GSC แสดงให้เห็นว่า เว็บไซต์ที่ใช้กลยุทธ์ Single Keyword per Page มีค่าเฉลี่ยตำแหน่งการแสดงผลใน SERP ที่ดีกว่าอย่างต่อเนื่อง และมีอัตราการคลิก (CTR) สูงกว่าเว็บไซต์ที่ใช้หลายคีย์เวิร์ดต่อหน้าเพจ สะท้อนถึงความตรงประเด็นระหว่างคำค้นหาและเนื้อหาที่นำเสนอ ผลการวิเคราะห์จาก GSC พบว่า เว็บไซต์ที่ใช้กลยุทธ์ Single Keyword per Page มีค่าเฉลี่ยตำแหน่งการแสดงผลใน SERP และอัตราการคลิก (CTR) สูงกว่าอย่างต่อเนื่อง สะท้อนถึงความตรงประเด็นระหว่างคำค้นหาและเนื้อหา ในขณะที่ข้อมูลจาก GA4 แสดงให้เห็นว่า เว็บไซต์ที่ใช้กลยุทธ์ Multiple Keywords per Page มีจำนวนผู้ใช้งานและจำนวนครั้งเข้าชมเว็บไซต์เฉลี่ยต่อเดือนสูงกว่า โดยเกิดจากการครอบคลุมคำค้นแบบ Long-tail keywords ที่หลากหลาย ในด้านการมีส่วนร่วมของผู้ใช้งาน เว็บไซต์ที่ใช้ Single Keyword per Page มีค่า Average Time on Page สูงกว่าและมีค่า Bounce Rate ต่ำกว่า แสดงถึงความสอดคล้องระหว่างคำค้นหาและเนื้อหาที่ช่วยยกระดับคุณภาพประสบการณ์ผู้ใช้งาน นอกจากนี้ ผลการวิเคราะห์จาก Ahrefs พบว่า กลยุทธ์ Single Keyword per Page มีค่า Keyword Difficulty ต่ำกว่าและมีเสถียรภาพของอันดับการจัดอันดับที่ดีกว่า ขณะที่กลยุทธ์ Multiple Keywords per Page มีจำนวน Backlinks และการปรากฏในคำค้นที่หลากหลายมากกว่า ซึ่งช่วยเพิ่มการมองเห็นในวงกว้าง

โดยสรุป ผลการศึกษาชี้ให้เห็นถึงลักษณะการแลกเปลี่ยนระหว่างคุณภาพและปริมาณของผลลัพธ์ด้าน

SEO กล่าวคือ ความตรงประเด็นของคีย์เวิร์ดช่วยเพิ่มอัตราการคลิกและการมีส่วนร่วมของผู้ใช้งาน ขณะที่ความหลากหลายของคีย์เวิร์ดช่วยเพิ่มปริมาณ Organic Traffic ดังนั้น งานวิจัยนี้จึงเสนอการใช้กลยุทธ์ SEO แบบผสมผสาน โดยเน้น Single Keyword per Page สำหรับหน้าเพจหลัก และ Multiple Keywords per Page สำหรับหน้าเพจสนับสนุน เพื่อเพิ่มประสิทธิภาพการเข้าถึงเว็บไซต์อย่างสมดุลและยั่งยืน

## เอกสารอ้างอิง

[1] ศุภณัฐ สุขโข. (2564). ดันเว็บไซต์ให้เป็นที่ 1 ในใจลูกค้าด้วย SEO (พิมพ์ครั้งที่ 3). นนทบุรี: ไอดีซีฯ.

[2] Grzywaczewski, A., et al. (2010). E-marketing strategy for businesses. Proceedings of the 7th International Conference on e-Business Engineering (ICEBE), 428–434. IEEE. https://doi.org/10.1109/ICEBE.2010.96

[3] Jing, S., Shi, J., et al. (2010). Research on SEO strategies of university journal websites. Proceedings of the 2nd International Conference on Information Science and Engineering, 3060–3063. IEEE. https://doi.org/10.1109/ICISE.2010.5691519

[4] Schwartz, B. (2024, January 15). Google: Number one search ranking factor is awesomeness. Search Engine Roundtable. https://www.seroundtable.com/google-awesomeness-24418.html

[5] Ellis, M. (2024, January 15). What is SEO? Moz. https://moz.com/learn/seo/what-is-seo

[6] Zhu, C., & Wu, G. (2021). Research and analysis of search engine optimization factors based on reverse engineering. Proceedings of the 3rd International Conference on Multimedia Information Networking and Security, 225–228. IEEE. https://doi.org/10.1109/MINES52562.2021.00056

# ปัจจัยที่ส่งผลต่อความสำเร็จในการสกัดความต้องการสำหรับการพัฒนาซอฟต์แวร์
# Factors Influencing the Success of Requirements Elicitation for Software Development

*สุทธิกานต์ จุลสุวรรณ์ (Sutthikarn Julsuwan)[1] และนลินภัสร์ บำเพ็ญเพียร (Nalinpat Bhumpenpein)[2]*
*ภาควิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล*
*มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ*
*[1]s6707011857055@email.kmutnb.ac.th, [2]nalinpat.b@itd.kmutnb.ac.th*

**บทคัดย่อ**

*งานวิจัยนี้มีวัตถุประสงค์เพื่อศึกษาปัจจัยที่ส่งผลต่อความสำเร็จในการสกัดความต้องการด้านซอฟต์แวร์ โดยใช้ระเบียบวิธีวิจัยเชิงคุณภาพ เก็บข้อมูลผ่านการสัมภาษณ์เชิงลึกแบบกึ่งโครงสร้างจากผู้ปฏิบัติงานที่มีประสบการณ์จำนวน 12 คน และวิเคราะห์ข้อมูลด้วยการวิเคราะห์เนื้อหา ผลการวิจัยพบ 7 ปัจจัยสำคัญ จัดกลุ่มเป็น 3 มิติ ดังนี้ (1) มิติด้านมนุษย์ ได้แก่ การสื่อสาร ความร่วมมือ ทักษะนักวิเคราะห์ และความชัดเจนของความต้องการ (2) มิติด้านองค์กร ได้แก่ นโยบาย วัฒนธรรม ระยะเวลา และงบประมาณ และ (3) มิติด้านข้อกำหนดและบริบทภายนอก ได้แก่ ข้อจำกัดทางกฎหมาย ผลการศึกษาชี้ว่าการบูรณาการทั้งสามมิติช่วยยกระดับคุณภาพข้อกำหนดและลดความคลาดเคลื่อนในระยะพัฒนาซอฟต์แวร์*

**คำสำคัญ :** การสกัดความต้องการ การพัฒนาซอฟต์แวร์ วิศวกรรมความต้องการ ระบบสารสนเทศ ผู้มีส่วนได้ส่วนเสีย

**Abstract**

*This study aims to investigate the factors influencing the success of software requirements elicitation. A qualitative research methodology was employed. Data were collected through semi-structured in-depth interviews with 12 practitioners who had experience in software requirements elicitation, and the data were analyzed using content analysis. The findings identified seven key factors, which were categorized into three main dimensions: (1) the human dimension, including communication, collaboration, analysts' skills, and clarity of requirements (2) the organizational dimension, including policies, organizational culture, time, and budget; and (3) the regulatory and external context dimension, including legal constraints. The results indicate that integrating these three dimensions enhances the quality of system requirements and reduces misalignment risks during the software development phase.*

**Keywords:** Requirements Elicitation, Software Development, Requirement Engineering, Information Systems, Stakeholder

## 1. บทนำ

ในยุคดิจิทัล การพัฒนาซอฟต์แวร์ให้มีคุณภาพและตอบสนองผู้ใช้งานขึ้นอยู่กับความถูกต้องของกระบวนการสกัดความต้องการ ซึ่งเป็นขั้นตอนสำคัญในวงจรการพัฒนาระบบ [1] ความล้มเหลวในการกำหนดข้อกำหนดที่ชัดเจนมักนำไปสู่ความล่าช้า งบประมาณบานปลาย และระบบที่ไม่ตรงกับความคาดหวังของผู้ใช้และองค์กร [2] งานวิจัยสมัยใหม่ได้ขยายขอบเขตการสกัดความต้องการไปสู่บริบทปัญญาประดิษฐ์ การวิเคราะห์ข้อมูลจากแอปพลิเคชัน และการใช้ แบบจำลองปัญญาประดิษฐ์ด้านภาษาเพื่อสนับสนุนทักษะการสกัดความต้องการ [3] การสกัดความต้องการเป็นกระบวนการสำคัญในการพัฒนาซอฟต์แวร์ที่มีความไม่แน่นอนและไม่สามารถกำหนดข้อกำหนดได้ชัดเจนตั้งแต่ต้น รวมถึงการสื่อสาร

อย่างมีประสิทธิภาพ จำต้องอาศัยการแลกเปลี่ยนข้อมูลอย่างต่อเนื่องเพื่อสร้างความเข้าใจร่วมกันระหว่างผู้เกี่ยวข้องและตอบสนองความต้องการของผู้ใช้งานได้อย่างแท้จริง

## 2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

### 2.1 ทักษะด้านการสื่อสาร

ความสามารถในการสื่อสารอย่างมีประสิทธิภาพเป็นกลไกพื้นฐานที่ช่วยลดความคลาดเคลื่อนของข้อกำหนดและส่งเสริมความสอดคล้องระหว่างเป้าหมายทางเทคนิคกับความคาดหวังของผู้ใช้ [4] การเสนอแนวทางพัฒนาวิธีการและกลไกสนับสนุน เช่น การใช้ ระบบแนะนำ (Recommendation System) เพื่อจัดการปัญหาข้อมูลความต้องการจำนวนมากหรือกระจัดกระจาย และช่วยให้การตัดสินใจมีความแม่นยำมากขึ้น [5] นอกจากนี้ ความร่วมมือระหว่างผู้ใช้งานและทีมพัฒนาเป็นอีกปัจจัยสำคัญ การใช้การวิเคราะห์ความคิดเห็นจากรีวิวผู้ใช้เป็นอีกแนวทางเพื่อสะท้อนความต้องการที่แท้จริง อันเป็นการเสริมประสิทธิภาพการสื่อสารและความร่วมมือระหว่างผู้ใช้กับทีมพัฒนา

### 2.2 วิศวกรรมความต้องการ

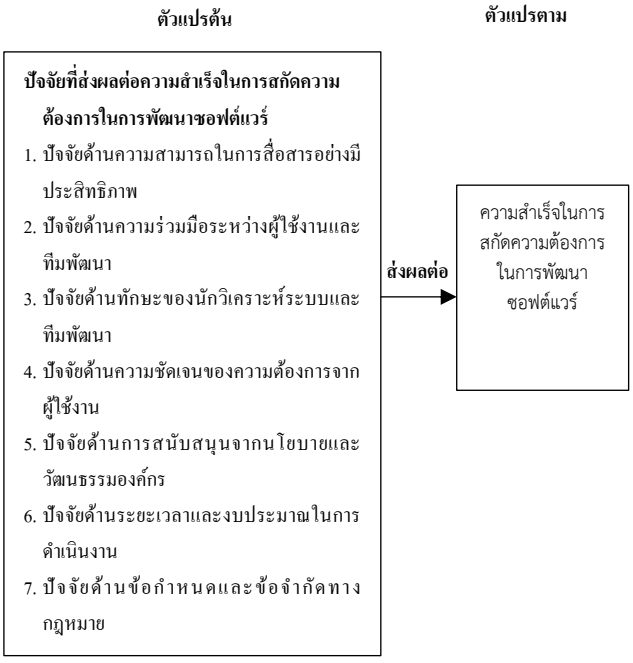
วิศวกรรมความต้องการเป็นกระบวนการสำคัญในวงจรการพัฒนาซอฟต์แวร์ โดยมุ่งเน้นการระบุ วิเคราะห์ กำหนด และตรวจสอบความต้องการของผู้มีส่วนได้ส่วนเสีย เพื่อให้ซอฟต์แวร์ที่พัฒนาขึ้นสามารถตอบสนองวัตถุประสงค์ขององค์กรได้อย่างเหมาะสม [6] กระบวนการนี้ครอบคลุมตั้งแต่การสกัดความต้องการ การวิเคราะห์และจัดการความต้องการ การจัดทำเอกสารข้อกำหนดอย่างเป็นทางการ ไปจนถึงการตรวจสอบความถูกต้องของข้อกำหนด ซึ่งคุณภาพของกระบวนการในระยะเริ่มต้นมีผลโดยตรงต่อความสำเร็จของโครงการซอฟต์แวร์ในภาพรวม

### 2.3 กระบวนการสกัดความต้องการ

กระบวนการสกัดความต้องการ มุ่งเน้นไปที่การค้นหาและทำความเข้าใจความต้องการของผู้ใช้งานและผู้มีส่วนได้ส่วนเสีย ตั้งแต่การทำความเข้าใจบริบท ไปจนถึงการเลือกและดำเนินการใช้เทคนิคการสกัดความต้องการ และจัดทำเอกสารและการตรวจสอบความถูกต้อง การเลือกเทคนิคและการดำเนินการสกัดความต้องการที่เหมาะสมกับบริบทของโครงการมีผลอย่างยิ่งต่อความสำเร็จของโครงการ [7] ขณะเดียวกัน การตรวจสอบความถูกต้อง การจัดลำดับความสำคัญ และการติดตามข้อกำหนดอย่างเป็นระบบช่วยลดความเสี่ยงจากการเปลี่ยนแปลงข้อกำหนดในระยะท้าย [8] การผสมผสานเทคนิคการสกัดความต้องการหลายรูปแบบยังช่วยเพิ่มความครอบคลุมและสะท้อนมุมมองของผู้มีส่วนได้ส่วนเสียได้ดียิ่งขึ้น [9]

### 2.4 ความสำเร็จในการสกัดความต้องการ

ในบริบทขององค์กรยุคดิจิทัล ความสำเร็จในการสกัดความต้องการของระบบสารสนเทศมีบทบาทสำคัญต่อการพัฒนาและการใช้งานระบบให้สอดคล้องกับทิศทางขององค์กร ท่ามกลางการเปลี่ยนแปลงทางเทคโนโลยีที่เกิดขึ้นอย่างรวดเร็ว กระบวนการสกัดความต้องการจำเป็นต้องสามารถรองรับการเปลี่ยนแปลงดังกล่าวได้อย่างยืดหยุ่นและมีประสิทธิภาพ




**ภาพที่ 1:** โมเดลแนวคิดการศึกษาปัจจัยที่ส่งผลต่อความสำเร็จในการสกัดความต้องการในการพัฒนาซอฟต์แวร์

## 3. วิธีการดำเนินงานวิจัย

งานวิจัยนี้เป็นการวิจัยเชิงคุณภาพ (Qualitative Research) มีวัตถุประสงค์เพื่อศึกษาปัจจัยที่ส่งผลต่อความสำเร็จของกระบวนการสกัดความต้องการสำหรับการพัฒนาซอฟต์แวร์ โดยเก็บรวบรวมข้อมูลผ่านการสัมภาษณ์เชิงลึกแบบกึ่งโครงสร้าง (Semi-structured Interview)

### 3.1 กลุ่มตัวอย่างและการคัดเลือก

กลุ่มตัวอย่างถูกคัดเลือกโดยวิธีการเลือกแบบเจาะจง (Purposive Sampling) จากผู้ที่มีประสบการณ์เกี่ยวข้องกับกระบวนการสกัดความต้องการซอฟต์แวร์ ได้แก่ นักวิเคราะห์ระบบ ผู้จัดการโครงการ และผู้ใช้งานระบบ จำนวน 12 คน เป็นระยะเวลา 2 สัปดาห์ ในการสัมภาษณ์แต่ละครั้งอยู่ในช่วงเวลา 18.00 น. – 19.00 น. ระยะเวลาที่ใช้สัมภาษณ์ประมาณ 10-15 นาทีต่อคนเพื่อให้ได้ข้อมูลเชิงลึกจากผู้มีประสบการณ์ตรงในบริบทการพัฒนาระบบสารสนเทศ

### 3.2 เครื่องมือวิจัย

เครื่องมือที่ใช้ในการเก็บข้อมูลคือแบบสัมภาษณ์กึ่งโครงสร้าง ซึ่งประกอบด้วยคำถามปลายเปิดที่ครอบคลุมประเด็นด้านการสื่อสาร ความร่วมมือ ทักษะของทีมพัฒนา ปัจจัยด้านองค์กร และข้อจำกัดภายนอก โดยคำถามถูกออกแบบให้สอดคล้องกับวัตถุประสงค์ของการวิจัย และผ่านการตรวจสอบความเหมาะสมของเนื้อหา

### 3.3 การเก็บรวบรวมข้อมูล

ผู้วิจัยดำเนินการเก็บข้อมูลในช่วงเดือนกุมภาพันธ์ โดยใช้ระยะเวลา 2 สัปดาห์ การสัมภาษณ์แต่ละครั้งใช้เวลาประมาณ 10–15 นาที ดำเนินการในรูปแบบออนไลน์และออฟไลน์ตามความสะดวกของผู้ให้ข้อมูล พร้อมทั้งมีการจดบันทึกและถอดความข้อมูลอย่างเป็นระบบ

### 3.4 การวิเคราะห์ข้อมูล

ข้อมูลจากการสัมภาษณ์ถูกนำมาวิเคราะห์ด้วยวิธีการวิเคราะห์เชิงเนื้อหา (Content Analysis) โดยผู้วิจัยดำเนินการเข้ารหัสข้อมูล (Coding) และจัดกลุ่มประเด็น (Thematic Analysis) เพื่อสกัดประเด็นสำคัญและจัดหมวดหมู่เป็นปัจจัยที่ส่งผลต่อความสำเร็จของกระบวนการสกัดความต้องการ

### 3.5 การตรวจสอบความน่าเชื่อถือของข้อมูล

เพื่อเพิ่มความน่าเชื่อถือของผลการวิจัย ผู้วิจัยใช้วิธีการตรวจสอบความถูกต้องของข้อมูลโดยการทวนสอบกับผู้ให้ข้อมูล (Member Checking) และตรวจสอบความสอดคล้องของการตีความข้อมูลกับบริบทจริง

## 4. ผลการดำเนินงานวิจัย

### 4.1 ข้อมูลทั่วไปของผู้สัมภาษณ์

ผู้ให้ข้อมูลสำคัญจำนวน 12 คน ได้แก่ นักวิเคราะห์ระบบ ผู้จัดการโครงการ ผู้ใช้งานระบบ และผู้ที่เกี่ยวข้องในโครงการพัฒนาซอฟต์แวร์ ผู้ให้ข้อมูลมีประสบการณ์ในการทำงานด้านการพัฒนาซอฟต์แวร์แตกต่างกัน ตั้งแต่น้อยกว่า 1 ปี ไปจนถึงมากกว่า 7 ปี ซึ่งช่วยให้ข้อมูลที่ได้มีความหลากหลายทั้งในด้านมุมมองและบริบทการทำงาน

### 4.2 ผลวิเคราะห์จากข้อมูลจากการสัมภาษณ์

จากการเก็บรวบรวมข้อมูลด้วยวิธีการสัมภาษณ์เชิงลึก (In-depth Interview) ผู้วิจัยได้ดำเนินการถอดความข้อมูลจากบทสัมภาษณ์อย่างเป็นระบบ จากนั้นนำข้อมูลที่ได้มาทำการวิเคราะห์เนื้อหา (Content Analysis) เพื่อค้นหาประเด็นสำคัญที่สะท้อนถึงปัจจัยที่ส่งผลต่อความสำเร็จของกระบวนการสกัดความต้องการสำหรับการพัฒนาซอฟต์แวร์ รายละเอียด ดังนี้

**ตารางที่ 1:** การวิเคราะห์บทสัมภาษณ์

| ตัวอย่างข้อความจากผู้ให้สัมภาษณ์ (Raw Data) | รหัสข้อมูล (Coding) | หมวดหมู่หลัก (Theme) |
|---|---|---|
| การสื่อสารไม่ตรงกันระหว่างผู้ใช้กับผู้พัฒนา ทำให้เข้าใจความต้องการผิด | ความคลาดเคลื่อนในการสื่อสาร | ความสามารถในการสื่อสาร |
| ควรใช้ Visual Tools เช่น Prototype / Wireframe เพื่อให้เห็นภาพเดียวกัน | การใช้เครื่องมือเชิงภาพเพื่อสร้างความเข้าใจร่วม | ความชัดเจนของความต้องการ |
| ควรให้ผู้เกี่ยวข้องมีส่วนร่วมตั้งแต่ต้นและทบทวนงานหลังจบโครงการ | การมีส่วนร่วมตั้งแต่ระยะเริ่มต้น | ความร่วมมือระหว่างผู้ใช้กับทีมพัฒนา |

| ตัวอย่างข้อความจากผู้ให้สัมภาษณ์ (Raw Data) | รหัสข้อมูล (Coding) | หมวดหมู่หลัก (Theme) |
|---|---|---|
| ผู้ใช้งานไม่รู้ความต้องการของตนเองอย่างแท้จริง และให้ข้อมูลไม่ครบถ้วน | ความต้องการไม่ชัดเจนของผู้ใช้งาน | ความชัดเจนของความต้องการ |
| ควรพัฒนาทักษะการวิเคราะห์ การสื่อสาร และความเข้าใจธุรกิจ | การพัฒนาทักษะเชิงวิเคราะห์และการสื่อสาร | ทักษะของนักวิเคราะห์ระบบ |
| เน้นทักษะการตั้งคำถามเชิงลึก และการคิดเชิงวิเคราะห์ | การตั้งคำถามเชิงลึกเพื่อค้นหาปัญหาที่แท้จริง | ทักษะของนักวิเคราะห์ระบบ |
| องค์กรที่เน้น Agile หรือ User-Centered Design ช่วยให้ตัดสินใจรวดเร็วและยืดหยุ่น | วัฒนธรรมองค์กรที่สนับสนุนความยืดหยุ่น | การสนับสนุนจากวัฒนธรรมองค์กร |
| กฎหมาย เช่น PDPA และ Accessibility ต้องกำหนดเป็น ตั้งแต่ต้น | การกำหนดข้อจำกัดทางกฎหมายเป็นข้อกำหนดระบบ | ข้อกำหนดและข้อจำกัดทางกฎหมาย |
| งบประมาณและเวลาที่เหมาะสมช่วยให้ใช้วิธีวิจัยคุณภาพสูง และครอบคลุมฟังก์ชันที่เหมาะสมต่อความต้องการ | งบประมาณสนับสนุนการใช้วิธีการที่มีคุณภาพ | ระยะเวลาและงบประมาณ |

### 4.3 การวิเคราะห์ปัจจัยที่ส่งผลต่อความสำเร็จของกระบวนการสกัดความต้องการ

สามารถจัดกลุ่มเป็น 3 มิติหลัก และ แบ่งออกเป็น 7 ปัจจัยที่สำคัญ ดังนี้

**ตารางที่ 2:** ปัจจัยที่ส่งผลต่อความสำเร็จของกระบวนการสกัดความต้องการตามมิติ

| ลำดับ | มิติ | ปัจจัย |
|---|---|---|
| 1 | ด้านมนุษย์ | 1) ด้านความสามารถในการสื่อสารอย่างมีประสิทธิภาพ<br>2) ด้านความร่วมมือระหว่างผู้ใช้งานและทีมพัฒนา<br>3) ด้านทักษะของนักวิเคราะห์ระบบและทีมพัฒนา<br>4) ด้านความชัดเจนของความต้องการจากผู้ใช้งาน |
| 2 | ด้านองค์กร | 5) ด้านการสนับสนุนจากนโยบายและวัฒนธรรมองค์กร<br>6) ด้านระยะเวลาและงบประมาณในการดำเนินงาน |
| 3 | ด้านข้อกำหนดและบริบทภายนอก | 7) ด้านข้อกำหนดและข้อจำกัดทางกฎหมาย |

ผลการศึกษาพบว่า ความสำเร็จของการสกัดความต้องการขึ้นอยู่กับ 3 มิติหลัก ได้แก่ ด้านมนุษย์ (การสื่อสาร ความร่วมมือ และทักษะ) ด้านองค์กร (นโยบาย วัฒนธรรม เวลา และงบประมาณ) และด้านข้อกำหนด/บริบทภายนอก (กฎหมายและมาตรฐาน) ซึ่งต้องบูรณาการร่วมกันอย่างเป็นระบบ รายละเอียดปัจจัยแต่ละด้าน ดังนี้

**1) ความสามารถในการสื่อสารอย่างมีประสิทธิภาพ** ผู้ให้ข้อมูลส่วนใหญ่เห็นสอดคล้องกันว่า ความสามารถในการสื่อสารอย่างมีประสิทธิภาพระหว่างผู้พัฒนาระบบและผู้ใช้งานเป็นปัจจัยพื้นฐานที่กำหนดคุณภาพของความต้องการที่ได้จากกระบวนการสกัดความต้องการ ความมีประสิทธิภาพในที่นี้ครอบคลุมถึงความสามารถในการใช้ภาษาที่เข้าใจง่าย การตั้งคำถามเชิงลึกเพื่อค้นหาความต้องการแฝง การฟังอย่างตั้งใจ และการยืนยันความเข้าใจร่วมกันก่อนสรุปเป็นข้อกำหนด หากขาดองค์ประกอบเหล่านี้ อาจเกิดการตีความคลาดเคลื่อน ทำให้ข้อกำหนดที่จัดทำขึ้นสะท้อนเพียง “สิ่งที่ผู้ใช้กล่าว” แต่ไม่สะท้อน “ความต้องการที่แท้จริง” ส่งผลให้เกิดการแก้ไขซ้ำในระยะพัฒนา เพิ่มต้นทุนและระยะเวลาโครงการ

**2) ความร่วมมือระหว่างผู้ใช้งานและทีมพัฒนา** ผลการสัมภาษณ์พบว่า คุณภาพของความต้องการขึ้นอยู่กับความร่วมมือของผู้ใช้งาน โดยหากผู้ใช้งานมีส่วนร่วมและให้ข้อมูลอย่างชัดเจน จะช่วยลดความคลาดเคลื่อนของข้อกำหนด ในทางกลับกัน หากขาดความร่วมมือหรือให้ข้อมูลไม่ครบถ้วน อาจนำไปสู่การพัฒนาฟังก์ชันที่ไม่ตรงกับการใช้งานจริง ดังนั้น การหารือเชิงลึกและการใช้ต้นแบบร่วมกันจึงมีบทบาทสำคัญในการเพิ่มความถูกต้องของข้อกำหนด

**3) ทักษะของนักวิเคราะห์ระบบและทีมพัฒนา** ผลการสัมภาษณ์พบว่า ทักษะของนักวิเคราะห์ระบบและทีมพัฒนา เช่น การฟัง การวิเคราะห์ และการตั้งคำถามเชิงลึก มีบทบาทสำคัญในการแยกแยะระหว่าง “สิ่งที่ผู้ใช้พูด” กับ “ปัญหาที่แท้จริง” ซึ่งช่วยแปลงความต้องการที่คลุมเครือให้เป็นข้อกำหนดที่ชัดเจนและนำไปใช้ได้จริง [10] หากขาดทักษะดังกล่าว อาจทำให้พัฒนาฟีเจอร์ที่ไม่ตรงกับ

ปัญหาจริง ส่งผลให้ข้อกำหนดไม่ครบถ้วนและเพิ่มต้นทุนการแก้ไขในระยะหลัง

**4) ความชัดเจนของความต้องการจากผู้ใช้งาน** ผลการสัมภาษณ์พบว่า ผู้ใช้งานมักไม่สามารถอธิบายความต้องการได้ชัดเจนในระยะแรก เนื่องจากข้อจำกัดในการถ่ายทอดความคิดเชิงนามธรรม ส่งผลให้ข้อกำหนดมีความคลุมเครือ ทีมพัฒนาจึงต้องใช้เทคนิค เช่น Prototype และ Use Case เพื่อช่วยทำให้ความต้องการมีความชัดเจน ลดความคลาดเคลื่อน และเพิ่มความเข้าใจร่วมกันระหว่างผู้ใช้งานกับทีมพัฒนาสามารถสะท้อนความต้องการของตนเองได้อย่างเป็นระบบและตรวจสอบได้

**5) การสนับสนุนจากนโยบายและวัฒนธรรมองค์กร**

ผลการสัมภาษณ์พบว่า โครงสร้างการบริหารโครงการและวัฒนธรรมองค์กรมีผลต่อคุณภาพของกระบวนการสกัดความต้องการ โดยองค์กรที่มีความยืดหยุ่นและเปิดโอกาสให้มีการตัดสินใจแบบกระจายอำนาจ จะช่วยให้ทีมพัฒนาสามารถรับข้อเสนอแนะจากผู้ใช้งานได้อย่างต่อเนื่อง ลดความล่าช้า และทำให้การยืนยันข้อกำหนดมีความรวดเร็วและชัดเจนมากขึ้น ในทางตรงกันข้าม องค์กรที่มีลำดับชั้นสูงหรือไม่สนับสนุนการมีส่วนร่วมของผู้ใช้งาน อาจทำให้ข้อมูลความต้องการถูกบิดเบือนหรือไม่ครบถ้วน ส่งผลให้ข้อกำหนดไม่สอดคล้องกับการใช้งานจริงและเพิ่มความเสี่ยงในกระบวนการพัฒนาระบบ

**6) ระยะเวลาและงบประมาณในการดำเนินงาน** ผลการสัมภาษณ์พบว่า ระยะเวลาและงบประมาณมีผลโดยตรงต่อคุณภาพของการสกัดความต้องการ โดยหากมีทรัพยากรเพียงพอจะสามารถใช้วิธีการที่มีคุณภาพได้ ขณะที่การเร่งรัดมากเกินไปอาจทำให้ขั้นตอนสำคัญถูกลดทอน ส่งผลให้ข้อกำหนดไม่สมบูรณ์ เพิ่มความเสี่ยงต่อการเปลี่ยนแปลงความต้องการ และทำให้ต้นทุนการแก้ไขในระยะหลังสูงขึ้น

**7) ข้อกำหนดและข้อจำกัดทางกฎหมาย** ผู้ให้สัมภาษณ์ระบุว่ากฎหมายและมาตรฐานที่เกี่ยวข้อง เช่น กฎหมายคุ้มครองข้อมูลส่วนบุคคล (PDPA) หรือมาตรฐานด้านการเข้าถึง (Accessibility) เป็นข้อจำกัดที่ต้องถูกนำมาพิจารณาตั้งแต่เริ่มต้นการพัฒนาซอฟต์แวร์ เพื่อหลีกเลี่ยงปัญหาด้านกฎหมายในภายหลัง ตัวอย่างสถานการณ์ที่สะท้อนผลกระทบของการละเลยข้อกำหนดทางกฎหมาย ได้แก่ โครงการพัฒนาระบบสมัครสมาชิกออนไลน์ที่ไม่ได้กำหนดกลไกการถอนความยินยอมในการใช้ข้อมูลส่วนบุคคลไว้ตั้งแต่ต้น เมื่อมีการร้องเรียนจากผู้ใช้งาน องค์กรต้องปรับแก้โครงสร้างฐานข้อมูลและกระบวนการทำงานใหม่ทั้งหมด ส่งผลให้เกิดต้นทุนเพิ่มเติมและความเสียหายต่อภาพลักษณ์องค์กร

## 5. สรุปและอภิปรายผล

จากการวิเคราะห์ข้อมูลการสัมภาษณ์เชิงลึก สามารถสรุปปัจจัยที่ส่งผลต่อความสำเร็จของกระบวนการสกัดความต้องการได้ 7 ปัจจัย ซึ่งจัดกลุ่มเป็น 3 มิติหลัก รายละเอียด ดังนี้

**มิติด้านมนุษย์** (1) ความสามารถในการสื่อสารอย่างมีประสิทธิภาพ ผลการสัมภาษณ์สะท้อนว่าการสื่อสารที่ชัดเจน การตั้งคำถามเชิงลึก และการยืนยันความเข้าใจร่วมกันเป็นปัจจัยพื้นฐานที่กำหนดคุณภาพของข้อกำหนด สอดคล้องกับ Singh et al. [11] ที่ระบุว่าปัญหาการสื่อสารเป็นสาเหตุหลักของความล้มเหลวในกระบวนการสกัดความต้องการ และ Bormane et al. [12 ]พบว่ากระบวนการสื่อสารที่มีโครงสร้างชัดเจนช่วยเพิ่มอัตราความสำเร็จของโครงการระบบสารสนเทศ (2) ความร่วมมือระหว่างผู้ใช้งานและทีมพัฒนา ผู้ให้สัมภาษณ์เห็นตรงกันว่าการมีส่วนร่วมอย่างต่อเนื่องของผู้ใช้งานช่วยลดความคลาดเคลื่อนของข้อกำหนด สอดคล้องกับ Wiegers and Beatty [13] ที่เน้นบทบาทของผู้มีส่วนได้ส่วนเสียตลอดวงจรความต้องการ และการให้ความสำคัญของการตรวจสอบความถูกต้องของข้อกำหนดร่วมกันเพื่อลดต้นทุนในระยะหลังซึ่งสอดคล้องกับ Maalem & Derivry [14] (3) ทักษะของนักวิเคราะห์ระบบและทีมพัฒนา ผลการวิจัยพบว่าทักษะในการแยกแยะปัญหาที่แท้จริง การวิเคราะห์เชิงลึก และการสังเคราะห์ข้อกำหนด เป็นกลไกสำคัญที่ลดความคลาดเคลื่อนของความต้องการสอดคล้องกับ Lojo et al. [10] ที่ชี้ว่าการพัฒนาทักษะการสกัดความต้องการ โดยเฉพาะการตั้งคำถามเชิงวิเคราะห์ เป็นปัจจัยที่

ช่วยยกระดับคุณภาพข้อกำหนด และ (4) ความชัดเจนของความต้องการจากผู้ใช้งาน การใช้ต้นแบบ (Prototype) ช่วยลดความคลุมเครือของข้อกำหนด ผลดังกล่าวสอดคล้องกับ Pacheco, García & Reyes [7] ที่ระบุว่าการเลือกใช้เทคนิคที่เหมาะสมกับบริบทช่วยเพิ่มประสิทธิภาพกระบวนการสกัดความต้องการ และสอดคล้องกับ Lim, Henriksson & Zdravkovic [8] ที่เสนอการใช้ข้อมูลสนับสนุนการตัดสินใจในกระบวนการดังกล่าว

**มิติด้านองค์กร** (5) การสนับสนุนจากนโยบายและวัฒนธรรมองค์กร ผลการสัมภาษณ์สะท้อนว่าองค์กรที่มีโครงสร้างยืดหยุ่น และสนับสนุนการมีส่วนร่วมของผู้ใช้งาน ช่วยลดความล่าช้าและความคลาดเคลื่อนของข้อกำหนด ซึ่งสอดคล้องกับ Kotonya & Sommerville [6] กระบวนการและโครงสร้างองค์กรมีผลต่อคุณภาพข้อกำหนด และสอดคล้องกับแนวทางการบริหารโครงการ PMI Guide [2] และ (6) ระยะเวลาและงบประมาณในการดำเนินงาน ข้อมูลจากการสัมภาษณ์ชี้ว่าข้อจำกัดด้านเวลาและงบประมาณส่งผลต่อความสมบูรณ์ของข้อกำหนด หากเร่งรัดมากเกินไปอาจเพิ่มความเสี่ยงในระยะพัฒนา [2] ข้อจำกัดด้านทรัพยากรเป็นปัจจัยกำหนดความสำเร็จของโครงการ ที่เน้นความสำคัญของการตรวจสอบความถูกต้องของข้อกำหนดเพื่อลดต้นทุนที่อาจเกิดขึ้นภายหลัง [14]

**มิติด้านข้อกำหนดและบริบทภายนอก** (7) ข้อกำหนดและข้อจำกัดทางกฎหมาย ผลการวิจัยพบว่าการพิจารณาข้อกำหนดทางกฎหมายตั้งแต่ระยะต้นช่วยลดความเสี่ยงด้านการปรับแก้ในภายหลัง ข้อกำหนดเชิงกฎระเบียบและบริบทภายนอกควรถูกให้ความสำคัญในการเข้าสู่กระบวนการตั้งแต่ระยะเริ่มต้น หากละเลยในช่วงต้น อาจทำให้โครงสร้างระบบไม่สอดคล้องกับกฎหมาย ส่งผลให้ต้องปรับแก้สถาปัตยกรรม ระบบฐานข้อมูล หรือกระบวนการทำงานใหม่ในภายหลัง [4], [6]

## เอกสารอ้างอิง


[1] A. van Lamsweerde. (2009). Requirements Engineering: From System Goals to UML Models to Software Specifications. Chichester, UK: John Wiley & Sons.

[2] PMI.(2021). A Guide to the Project Management Body of Knowledge (PMBOK Guide), 7th ed. Newtown Square. PA: Project Management Institute.

[3] Wan, Hongyan; An, Zhiquan; Wang, Bangchao; Xiong, Teng. (2023). Sentiment Analysis for Requirements Elicitation from App Reviews: A Systematic Mapping Study. 2023 30th Asia-Pacific Software Engineering Conference.

[4] Elvira, Timothy; Procko, Tyler Thomas; Ochoa, Omar. (2024). Requirements Elicitation for Machine Learning Applications: A Research Preview. Conference on AI, Science, Engineering, and Technology.

[5] Shambour, Qusai Y.; Qandeel, Nazem N.; Shlool, Hatem A. (2024). An Enhanced Recommendation Approach for Efficient Requirements Elicitation. 11th International Conference on Electrical and Electronics Engineering.

[6] G. Kotonya and I. Sommerville. (1998). Requirements Engineering: Processes and Techniques. Chichester, UK: John Wiley & Sons

[7] C. Pacheco, I. García and M. Reyes. (2018). Requirements elicitation techniques: a systematic literature review based on the maturity of the techniques. 365–378 IET Software.

[8] Sachiko Lim1,Aron Henriksson1,Jelena Zdravkovic. (2021). Data Driven Requirements Elicitation: A Systematic Literature Review. Department of Computer and Systems Sciences, Stockholm University, DSV, PO Box 7003, 164 07 Kista, Stockholm, Sweden

[9] Ian Alexander. (2006). 10 small steps to better requirements.IEEE Software.

[10] Lojo, Nelson; González, Rafael; Philip, Rohan; Parejo, José Antonio; Toro, Amador Durán; Fox, Armando; Fernández, Pablo. (2025). Using Large Language Models to Develop Requirements Elicitation Skills. arXiv.

[11] Singh, Virendra; et al. (2017). Problems Associated with Requirement Elicitation Process: An Overview. International Journal of Advanced Research in Computer Science.

[12] Bormane, Līga; Gržibovska, Jūlija; Bērziša, Solvita; Grabis, Jānis. (2016). Impact of Requirements Elicitation Processes on Success of Information System Development Projects. Information Technology and Management Science, Riga Technical University.

[13] K. E. Wiegers and J. Beatty. (2013). Software Requirements, 3rd ed. Redmond.WA: Microsoft Press.

[14] W. Maalem and D. Derivry. (2016). 31-46. A survey on requirements validation techniques International Journal of Software Engineering and Its Applications.

# ระบบแนะนำสินค้าอัตโนมัติบนบัญชีไลน์ทางการด้วยการทำเหมืองกฎความสัมพันธ์ กรณีศึกษา ร้านสกายเบเกอรี่แอนด์คอฟฟี่

# Automated Product Recommendation System on LINE Official Account using Association Rule Mining: A Case of Sky Bakery and Coffee

*ปรเมศวร์ เพชรมั่น (Poramets Phetman)[1] และธนพล เจนสุทธิเวชกุล (Tanapon Jensuttiwetchakul)[2]*

*[1]สาขาวิชาระบบสารสนเทศเพื่อการจัดการ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล*

*มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ*

*s6707021857198@email.kmutnb.ac.th, tanapon.j@itd.kmutnb.ac.th*

## บทคัดย่อ

*LINE Official Account (LINE OA) เป็นช่องทางหลักที่ SMEs ในประเทศไทยนิยมใช้สื่อสารและรับคำสั่งซื้อจากลูกค้า ก่อให้เกิดข้อมูลบทสนทนาสะสมในปริมาณมากซึ่งยังไม่ได้ถูกนำมาใช้ประโยชน์อย่างเป็นระบบ ขณะที่กระบวนการแนะนำสินค้าเพิ่มเติมยังอาศัยประสบการณ์ของผู้ดูแลระบบเป็นหลัก งานวิจัยนี้นำเสนอกรอบการพัฒนาระบบแนะนำสินค้าอัตโนมัติแบบ End-to-End โดยใช้ n8n เป็น Central Orchestration Layer เชื่อมต่อตั้งแต่การรับข้อมูล การสร้างกฎความสัมพันธ์ จนถึงการส่งคำแนะนำกลับแบบเรียลไทม์ วิเคราะห์ข้อมูล 28,742 ข้อความจากผู้ใช้งาน 15,695 บัญชี ประยุกต์ใช้ FP-Growth (Support=0.02, Confidence=0.95) ได้กฎ 13,046 กฎ คัดกรองด้วย Lift $\geq$ 1.50 เหลือ 12,095 กฎ จัดเก็บลง Pinecone Vector Store และอัปเดตอัตโนมัติทุกสัปดาห์ผ่าน n8n Scheduler ผลการทดสอบกับ 600 Hold-out Session ผ่าน n8n Workflow จริง พบ H1: Item Extraction Accuracy ร้อยละ 96.88, H2: Rule Applicability Rate ร้อยละ 92.7 และ H3: Recommendation Precision ร้อยละ 84.6 ซึ่งสูงกว่าเกณฑ์ที่กำหนดทุกข้อ*

**คำสำคัญ:** *การทำเหมืองกฎความสัมพันธ์ ระบบแนะนำสินค้า LINE Official Account FP-Growth ระบบอัตโนมัติเวิร์กโฟลว์*

## Abstract

*LINE Official Account (LINE OA) is the primary communication and ordering channel for SMEs in Thailand, generating large volumes of conversational data that remain largely unutilized while cross-selling relies on administrator experience. This paper presents an end-to-end automated product recommendation framework on LINE OA, using n8n as a Central Orchestration Layer connecting data ingestion, association rule generation, and real-time recommendation delivery. The study analyzed 28,742 messages from 15,695 customer accounts of Sky Bakery & Coffee. The FP-Growth algorithm (Minimum Support 0.02, Minimum Confidence 0.95) yielded 13,046 rules, of which 12,095 with Lift $\geq$ 1.50 were deployed on Pinecone Vector Store and updated automatically weekly via n8n Scheduler. Evaluation on 600 hold-out sessions through the live n8n Workflow demonstrated H1: Item Extraction Accuracy 96.88%, H2: Rule Applicability Rate 92.7%, and H3: Recommendation Precision 84.6%, all exceeding their thresholds.*

**Keywords:** Association Rule Mining, Recommendation System, LINE Official Account, FP-Growth, Workflow Automation.

## 1. บทนำ

ในปัจจุบัน LINE Official Account เป็นช่องทางหลักที่ SMEs ในประเทศไทยนิยมใช้สื่อสารและรับคำสั่งซื้อจากลูกค้า ซึ่งมีผู้ใช้งานกว่า 54 ล้านบัญชี [12] [3] พบว่า LINE สามารถเพิ่มช่องทางการขายและกระตุ้นการซื้อซ้ำได้ ในขณะที่ [2] ยืนยันว่า LINE เป็นองค์ประกอบเชิงกลยุทธ์ของตราสินค้า อย่างไรก็ตาม แม้การดำเนินงานดังกล่าวจะก่อให้เกิดข้อมูลบทสนทนาสะสมในปริมาณมาก แต่ข้อมูลเหล่านี้กลับถูกทิ้งไว้โดยไม่ได้นำมาใช้ประโยชน์อย่างเป็นระบบ กระบวนการแนะนำสินค้า

เพิ่มเติม (Cross-selling) ยังอาศัยประสบการณ์ผู้ดูแลระบบเป็นหลัก ส่งผลให้พลาดโอกาสการขายในช่วงที่มีข้อความสูง งานวิจัยของ [13] ยืนยันว่าข้อมูลบทสนทนาเป็นแหล่งข้อมูลเชิงพฤติกรรมที่มีคุณค่าสูง แต่มีลักษณะเป็นข้อความที่ไม่มีโครงสร้าง ใช้ภาษาพูดและคำย่อ [10] งานวิจัยนี้ไม่ใช่แค่การวิเคราะห์ข้อมูล แต่เป็นการสร้าง Workflow อัตโนมัติแบบ End-to-End บน n8n ที่สามารถนำไปใช้งานจริงได้ทันทีในบริบท SMEs ไทย โดยใช้ n8n เป็นศูนย์กลางการประสานงาน (Central Orchestration Layer) ตั้งแต่การดึงข้อมูลจาก LINE OA การประมวลผล การสร้างกฎ จนถึงการส่งคำแนะนำกลับ งานวิจัยนี้มีวัตถุประสงค์สามประการ ได้แก่ (1) วิเคราะห์รูปแบบความสัมพันธ์ของสินค้าจากข้อมูลบทสนทนาด้วยเทคนิค Association Rule Mining (2) พัฒนาระบบแนะนำสินค้าอัตโนมัติที่เชื่อมต่อ LINE OA ผ่าน n8n และ (3) ประเมินประสิทธิภาพของระบบในสภาพแวดล้อมการใช้งานจริง

## 2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

### 2.1 ระบบแนะนำสินค้าและแชทบอทบนแพลตฟอร์มสนทนา

จำแนกระบบแนะนำสินค้าออกเป็นสามประเภท ได้แก่ Collaborative Filtering, Content-Based Filtering และ Rule-based Recommendation [4] สำหรับ SMEs ที่ขาดโปรไฟล์ผู้ใช้เชิงลึกแนวทาง Rule-based Recommendation มีความเหมาะสมที่สุดเนื่องจากไม่ต้องพึ่งพาประวัติผู้ใช้รายบุคคล และนำไปใช้งานบน Workflow Automation ได้โดยตรง [14] [15]แสดงให้เห็นว่าระบบแนะนำกำลังเปลี่ยนไปสู่การเข้าใจบริบทจากข้อความ [16] โดยใช้กรอบแนวคิด TAM ของ [9] พบว่าความสามารถในการใช้งานและการตอบสนองอย่างรวดเร็วมีผลเชิงบวกต่อประสบการณ์ลูกค้า [17] เสริมว่าการบูรณาการ AI กับ CRM ช่วยเพิ่มประสิทธิภาพการตลาดอย่างมีนัยสำคัญ

### 2.2 การทำเหมืองกฎความสัมพันธ์และการวิเคราะห์พฤติกรรมผู้บริโภค

พัฒนาการทำเหมืองกฎความสัมพันธ์ในบริบทการวิเคราะห์ตะกร้าสินค้า [5] ประเมินด้วยค่า Support, Confidence และ Lift [7] อัลกอริทึม FP-Growth ของ [11] เหมาะกับข้อมูล Long-tail ซึ่งพบในข้อมูลบทสนทนา [1] ยืนยันว่า FP-Growth สะท้อนพฤติกรรมผู้บริโภคไทยในบริบทการค้าออนไลน์ได้จริง

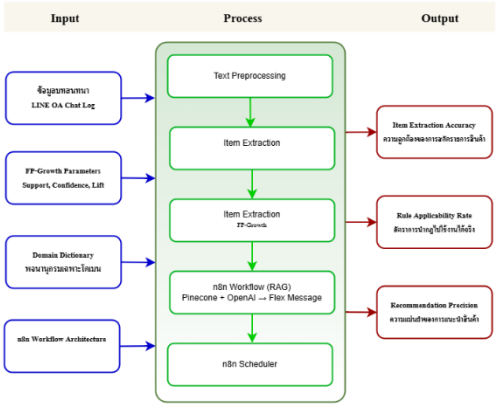
### 2.3 เวิร์กโฟลว์อัตโนมัติในระบบแนะนำสินค้า

n8n คือแพลตฟอร์มระบบอัตโนมัติเวิร์กโฟลว์แบบโอเพ่นซอร์สที่ใช้แนวคิด Node-based Workflow โดยแต่ละโหนดทำหน้าที่เฉพาะ เช่น การดึงข้อมูล การประมวลผล หรือการส่งต่อผลลัพธ์ไปยังปลายทาง n8n มีความเหมาะสมกับงานวิจัยนี้ด้วยเหตุผลสี่ประการ ได้แก่ โอเพ่นซอร์สที่ไม่มีค่าใช้จ่ายรายเดือนตามปริมาณการใช้งาน รองรับการเชื่อมต่อ API ภายนอกผ่าน HTTP Request และ Webhook รองรับ AI/LLM Nodes เช่น OpenAI และ Embeddings สำหรับประมวลผล RAG และรองรับสถาปัตยกรรมเชิงเหตุการณ์ (Event-driven Architecture) ที่เหมาะกับการประมวลผลข้อความแบบเรียลไทม์ [8] ในบริบท SMEs ที่ใช้ n8n ยังสามารถติดตั้งบนเซิร์ฟเวอร์ของธุรกิจเอง (Self-hosting) ซึ่งควบคุมข้อมูลได้เต็มที่ตามหลักการ PDPA และไม่ต้องพัฒนา Backend แยกต่างหาก จากการทบทวนวรรณกรรมในปัจจุบัน พบข้อจำกัดสำคัญที่ปรากฏชัดเจน 3 ประการ ประกอบด้วย 1) ความล่าช้าของข้อมูล (Data Latency) ซึ่งเกิดจากกระบวนการประมวลผลข้อมูลแบบ Batch ทำให้ไม่สามารถตอบสนองต่อพฤติกรรมของลูกค้าบนแพลตฟอร์มสนทนาได้อย่างทันท่วงที 2) อุปสรรคด้านการปฏิบัติงาน (Operational Barrier) เนื่องจากผู้ประกอบการ SMEs ส่วนใหญ่ยังขาดโครงสร้างพื้นฐานทางเทคโนโลยีที่จะรองรับโมเดลขนาดใหญ่ และ 3) ช่องว่างในการบูรณาการเทคโนโลยี ซึ่งแม้จะมีการศึกษาพฤติกรรมผู้บริโภคและการยอมรับเทคโนโลยีมาบ้าง แต่ยังคงขาดการวิจัยที่มุ่งเน้นการสกัดกฎความสัมพันธ์จากข้อมูลบทสนทนาบน LINE OA เพื่อเชื่อมโยงเข้ากับระบบแนะนำสินค้าแบบเรียลไทม์ผ่าน Workflow Automation อย่างครบวงจร ด้วยเหตุนี้ งานวิจัยนี้จึงมุ่งเติมเต็มช่องว่างดังกล่าวด้วยการนำเสนอกรอบการพัฒนาระบบแบบ End-to-End ผ่านสถาปัตยกรรม Stateful Conversation Architecture บนเครื่องมือ n8n ทำงาน

ร่วมกับ PostgreSQL เพื่อเพิ่มประสิทธิภาพในการจัดเก็บและรวมข้อมูล (Data Aggregation) จากข้อความที่กระจัดกระจายให้กลายเป็นธุรกรรมที่สมบูรณ์ ซึ่งจะช่วยยกระดับความแม่นยำในการแนะนำสินค้าและตอบโจทย์การใช้งานจริงในภาคธุรกิจได้อย่างมีประสิทธิภาพ

## 3. วิธีดำเนินการวิจัย

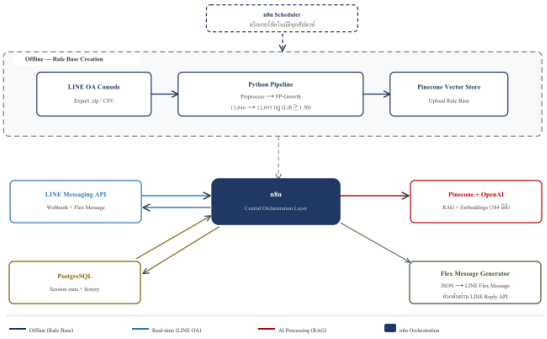
งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาและประเมินประสิทธิภาพของระบบแนะนำสินค้าอัตโนมัติบน LINE OAโดยใช้เทคนิคการทำเหมืองกฎความสัมพันธ์ด้วยอัลกอริทึม FP-Growth ร่วมกับแพลตฟอร์ม n8n ในฐานะ Central Orchestration Layer เพื่อนำมาใช้งานจริงในบริบทธุรกิจขนาดกลางและขนาดย่อม (SMEs) ของไทย โดยมีกรอบแนวคิดการวิจัยแสดงดังภาพที่ 1




**ภาพที่ 1:** กรอบแนวคิดการวิจัย

### 3.1 สถาปัตยกรรมระบบโดยรวม

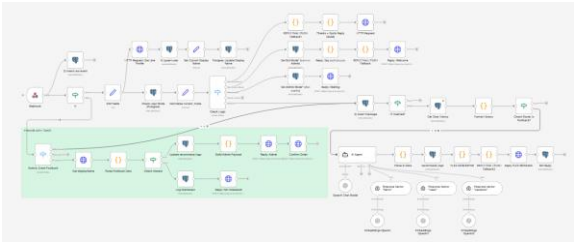
ระบบที่พัฒนาขึ้นใช้ n8n เป็น Central Orchestration Layer เชื่อมต่อ LINE Messaging API (รับข้อความ) PostgreSQL (เก็บประวัติบทสนทนา) Python Script / ML Model (สร้าง Association Rules จากข้อมูลเก่า) Pinecone / OpenAI (RAG สำหรับบริบท) และ Flex Message Generator (ส่งคำแนะนำกลับ) แบ่งออกเป็นสองเส้นทาง ได้แก่ เส้นทางออฟไลน์ซึ่ง n8n Scheduler กระตุ้น Python Script ทุกสัปดาห์เพื่อสร้างและอัปเดต Rule Base ลง Pinecone อัตโนมัติ และเส้นทางเรียลไทม์ ซึ่ง n8n รับข้อความจาก LINE Messaging API ค้นหากฎผ่าน Pinecone ร่วมกับ OpenAI แล้วส่งคำแนะนำกลับผ่าน LINE Flex Message ดังแสดงใน ภาพที่ 2



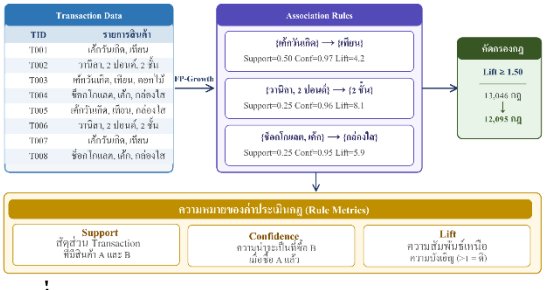

**ภาพที่ 2:** ระบบแนะนำสินค้าอัตโนมัติบน LINE OA

### 3.2 ขั้นตอนการประมวลผลบน n8n

แหล่งข้อมูลคือบันทึกบทสนทนาจาก LINE OA ของร้านสกายเบเกอรี่แอนด์คอฟฟี่ ประกอบด้วย 28,742 ข้อความจากผู้ใช้งาน 15,695 บัญชี 38,799 Session ช่วงเวลา 1 ม.ค. 2567 – 24 ม.ค. 2569 ข้อมูลส่วนบุคคลถูกปกปิดตาม PDPA ประกอบด้วยห้าขั้นตอนหลักดังแสดงในตารางที่ 2 จุดเด่นสำคัญคือ Stateful Conversation Architecture ที่บันทึกสถานะบทสนทนา (Current mode) ของผู้ใช้แต่ละรายใน PostgreSQL อย่างต่อเนื่อง ทำให้รวบรวมข้อมูลสินค้าข้ามหลายรอบสนทนาก่อนแนะนำ และรองรับการสลับระหว่าง Bot Mode กับ Admin Mode ตามหลักการ Human-in-the-Loop [6]



**ภาพที่ 3:** กระบวนการการทำงานWorkflow ของ n8n




**ภาพที่ 4**: ตัวอย่างกระบวนการสร้างและตีความ Association Rules จากข้อมูลบทสนทนา

ตัวอย่างการอ่านกฎ {เค้กวันเกิด} → {เทียน} Conf=0.97, Lift=4.2 "ใน 100 ครั้งที่ลูกค้าสั่งเค้กวันเกิด มี 97 ครั้งที่สั่งเทียนด้วย" Lift = 4.2 หมายถึง ความน่าจะ

เป็นสูงกว่าการซื้อโดยบังเอิญถึง 4.2 เท่า ⟶ กฎนี้มีความหมายเชิงธุรกิจ⇒ ระบบจะแนะนำ "เทียน" อัตโนมัติเมื่อลูกค้าพิมพ์ "เค้กวันเกิด" ใน LINE OA

**ตารางที่ 1:** ลำดับการทำงานของ n8n Workflow

| ขั้น | โหนดหลัก | กระบวนการ | วินาที |
|---|---|---|---|
| 1 | Webhook Node | รับ HTTP POST จาก LINE API + บันทึกลง PostgreSQL แบบขนาน | 0.1–0.2 |
| 2 | PostgreSQL Node | ตรวจสอบ current mode + ดึงประวัติ Session 30 นาทีล่าสุดของผู้ใช้ | 0.2–0.3 |
| 3 | Function Node | รวมข้อความใน 30 นาทีล่าสุดเป็น 1 Transaction + สกัดรายการสินค้า | 0.1–0.2 |
| 4 | AI Agent (RAG) | Embedding 384 มิติ + Pinecone ค้นหา Association Rules ที่ตรงกับ Transaction | 0.5–0.8 |
| 5 | Flex Generator | JSON ⟶ LINE Flex Message + ส่งกลับผ่าน LINE Reply API | 0.1–0.2 |

### 3.3 การสร้าง Rule Base แบบจากข้อมูลบทสนทนา

กระบวนการนี้ไม่ใช่เพียงการวิเคราะห์ข้อมูลครั้งเดียว แต่เป็นการสร้าง Rule Base แบบอัตโนมัติที่อัปเดตตัวเองทุกสัปดาห์ผ่าน n8n Scheduler ดำเนินการสี่ขั้นตอน ขั้นแรกผู้ดูแลระบบส่งออกข้อมูลบทสนทนาจาก LINE OA Console ในรูปแบบไฟล์ .zip ซึ่งภายในประกอบด้วยไฟล์ CSV แยกตามผู้ใช้งาน ขั้นที่สอง n8n Scheduler กระตุ้น Python Script ผ่าน Webhook ให้การประมวลผลคำสั่งผ่าน build_preprocessed_outputs.py, step1_create_sessions.py และ step2_run_apriori.py สำหรับรัน FP-Growth บนข้อมูล 28,742 ข้อความ ได้กฎ 13,046 กฎ คัดกรองด้วยเกณฑ์ Lift $\geq$ 1.50 ตาม [7] เหลือ 12,095 กฎ ขั้นที่สาม Script อัปโหลดกฎลง Pinecone Vector Store อัตโนมัติพร้อม Metadata เพื่อให้ n8n ค้นหาแบบ Semantic Search ในขณะให้บริการเรียลไทม์ ขั้นที่สี่ n8n ส่งรายงานสรุปไปยังผู้ดูแลระบบโดยอัตโนมัติ Rule Base นี้ถูกอัปเดตอัตโนมัติทุกสัปดาห์ผ่าน n8n Scheduler โดยไม่ต้องดำเนินการด้วยมนุษย์

### 3.4 การทดสอบระบบ

การทดสอบดำเนินการผ่าน n8n Workflow จริงทั้งหมด ไม่ใช่เพียงการทดสอบโมเดลในสภาพแวดล้อมจำลอง แต่ทดสอบระบบอัตโนมัติทั้งหมดที่ทำงานบน n8n ตั้งแต่การรับข้อความผ่าน Webhook Node จนถึงการส่ง Flex Message ออก โดยใช้ชุดข้อมูลทดสอบ (Hold-out Session) จำนวน 600 Session ที่แยกจากชุดข้อมูลที่ใช้สร้างกฎ จากนั้นผู้เชี่ยวชาญโดเมน (เจ้าของร้าน) ประเมินความสมเหตุสมผลเชิงธุรกิจของคำแนะนำ การประเมินประสิทธิภาพของระบบในงานวิจัยนี้ใช้ตัวชี้วัดสี่ประการ ได้แก่ ค่าความถูกต้อง (Accuracy) ค่าความแม่นยำ (Precision) ค่าความครอบคลุม (Recall) และค่า F1-Score โดยมีนิยามและสูตรการคำนวณ ดังนี้

$$\text{Accuracy}=\frac{TP+TN}{TP+TN+FP+FN} \quad (3\text{-}1)$$

$$\text{Precision}=\frac{TP}{TP+FP} \quad (3\text{-}2)$$

$$\text{Recall}=\frac{TP+TN}{TP+TN+FP+FN} \quad (3\text{-}3)$$

$$F1-Score = 2\times\frac{Precision\times Recall}{Precision+Recall} \quad (3\text{-}4)$$

โดยที่ TP (True Positive) คือจำนวน Session ที่มีรายการสินค้าในข้อความและระบบระบุได้ถูกต้อง TN (True Negative) คือจำนวน Session ที่ไม่มีรายการสินค้าและระบบระบุได้ถูกต้อง FP (False Positive) คือจำนวน Session ที่ระบบระบุว่ามีรายการสินค้าแต่ความจริงไม่มี และ FN (False Negative) คือจำนวน Session ที่มีรายการสินค้าแต่ระบบตรวจไม่พบ [18]




**ภาพที่ 5:** ตัวอย่างการทำงานของระบบบน LINE OA

## 4. ผลการดำเนินการวิจัย

ก่อนนำเสนอผลการทดสอบสมมติฐาน ผู้วิจัยขอชี้แจงนิยามของประสิทธิภาพสองระดับระดับแรกคือประสิทธิภาพของกระบวนการทำเหมืองข้อความและกฎความสัมพันธ์ (Text Mining & Association Rule) ได้แก่ ความถูกต้องของการสกัดรายการสินค้าและอัตราการใช้งานได้จริงของกฎ ระดับที่สองคือประสิทธิภาพของการนำระบบไปใช้งานจริงบน n8n ได้แก่ ความแม่นยำของการแนะนำสินค้าผ่าน LINE OA

### 4.1 ผลการประเมินความถูกต้องของการสกัด Item

ประเมิน Item Extraction Accuracy จากชุดข้อมูล Session ทั้งหมด 38,799 รายการด้วยวิธีการตรวจสอบเชิงตรรกะอัตโนมัติ (Rule-based Audit) ผลการประเมินแสดงในตารางที่ 3

**ตารางที่ 2:** ผลการประเมิน Item Extraction Accuracy (n = 38,799)

| ตัวชี้วัด | ค่าที่ได้ |
|---|---|
| Accuracy | 96.88% |
| Precision | 95.22% |
| Recall | 99.22% |
| F1-Score | 97.18% |

ค่า Recall สูงถึงร้อยละ 99.22 แสดงว่าระบบแทบไม่พลาด Session ที่มีสินค้าในข้อความ โดยมีผลบวกที่แท้จริง (True Positive) 20,859 Session ผลบวกเท็จ (False Positive) 1,047 Session และผลลบเท็จ (False Negative) เพียง 165 Session ความแม่นยำดังกล่าวเป็นผลจากพจนานุกรมเฉพาะโดเมน 34 รายการร่วมกับ Exact Match ที่ป้องกัน noise token [18] สรุป การประเมินความถูกต้องของการสกัดรายการสินค้าได้รับการสนับสนุน

### 4.2 ผลการทำเหมืองกฎความสัมพันธ์และ Rule Applicability

**ตารางที่ 3:** ผลการทดลอง Parameter Sensitivity Analysis

| Use case | Min Support | Frequent Itemsets | กฎทั้งหมด | กฎที่ใช้จริง | อัตราใช้งาน |
|---|---|---|---|---|---|
| 1 | 0.10 | 37 | 7 | 7 | 100% |
| 2 | 0.05 | 199 | 99 | 99 | 100% |
| 3 | 0.02 | 4,399 | 13,046 | 12,095 | 92.7% |
| 4 | 0.015 | 5,469 | 15,430 | 14,216 | 92.1% |
| 5 | 0.01 | 6,488 | 16,256 | 14,985 | 92.2% |

*ที่มา: ผลการวิจัย กฎที่ใช้จริง = กฎที่มี Lift $\geq$ 1.50*

จากการวิเคราะห์พบจุด Knowledge Saturation ระหว่าง Use case 2 และ 3 การลด Support จาก 0.05 เป็น 0.02 ทำให้กฎเพิ่มขึ้น 13,078% ในขณะที่การลดจาก 0.02 เป็น 0.015 ให้กฎเพิ่มขึ้นเพียง 18.3% ผู้วิจัยจึงเลือก Use case 3 ตาม [11] อัตราการใช้งานได้จริงร้อยละ 92.7 สูงกว่าเกณฑ์ H2 ที่ร้อยละ 30 อย่างมีนัยสำคัญ ตัวอย่างกฎแสดงในตารางที่ 5 สรุปผลการทำเหมืองกฎความสัมพันธ์และ Rule Applicability ได้รับการสนับสนุน

**ตารางที่ 4:** ตัวอย่างกฎความสัมพันธ์ที่ใช้ใน Deployment

| Antecedent | Consequent | Conf | Lift | ความหมายเชิงธุรกิจ |
|---|---|---|---|---|
| เค้กวันเกิด | แฮปปี้เบิร์ธเดย์ | 0.97 | ~4.2 | สั่งเค้กวันเกิดมักต้องการข้อความอวยพร |
| วานิลา, 2 ปอนด์ | 2 ชั้น | 0.96 | ~8.1 | เค้กวานิลา 2 ปอนด์มักสั่งแบบ 2 ชั้น |
| ช็อกโกแลต, เค้ก | กล่องใส | 0.95 | ~5.9 | เค้กช็อกโกแลตมักต้องการกล่องใส |
| เรดเวลเวท | กล่องกระดาษสีขาว | 0.96 | ~6.3 | เรดเวลเวทนิยมบรรจุกล่องกระดาษสีขาว |
| 2 ปอนด์, เทียน | แฮปปี้เบิร์ธเดย์ | 0.97 | ~4.1 | เค้กพร้อมเทียนมักต้องการข้อความวันเกิด |

นอกเหนือจากผลเชิงตัวเลขงานวิจัยนี้พบประเด็นสำคัญเชิงทฤษฎีสองประการประการแรกคือ Knowledge Saturation Point กล่าวคือการลดค่า Support ต่ำกว่า 0.02 ไม่ได้ช่วยเพิ่มกฎที่มีนัยสำคัญเชิงธุรกิจ แต่กลับเพิ่ม Noise ในระบบ ดังเห็นจากตารางที่ 4 ที่อัตราการใช้งานจริงลดลงเล็กน้อยใน Use Case 4 และ 5 ประการที่สองคือ High Rule Applicability อัตราร้อยละ 92.7 สะท้อนว่าในบริบทร้านเบเกอรี่ ลูกค้ามีรูปแบบการสั่งซื้อที่เกาะกลุ่มชัดเจน เช่น เค้กคู่กับเทียนหรือข้อความอวยพร ซึ่งกฎเหล่านี้สามารถเปลี่ยนประสบการณ์ผู้ใช้จากการสอบถามเป็นการตัดสินใจซื้อได้ทันที

### 4.3 ผลการประเมินระบบแนะนำสินค้าอัตโนมัติ

ประเมินระบบกับ Hold-out Session 600 รายการผ่าน n8n Workflow จริงครบวงจร Stateful Architecture ให้ Recommendation Precision ร้อยละ 84.6 สูงกว่าเกณฑ์ที่ร้อยละ 80 และสูงกว่า Stateless Architecture 13.4

เปอร์เซ็นต์พอยท์ (84.6% เทียบกับ 71.2%) Rule Match Rate อยู่ที่ร้อยละ 72.8 เจ้าของร้านยืนยันว่าระบบช่วยลดภาระการตอบแชทซ้ำ คำแนะนำมีความสม่ำเสมอ และลูกค้าตัดสินใจซื้อเร็วขึ้น สรุปผลการประเมินระบบแนะนำสินค้าอัตโนมัติได้รับการสนับสนุน ผลการทดสอบสมมติฐานทั้งสามข้อสรุปได้ดังนี้ H1: ความถูกต้องของการสกัดรายการสินค้าร้อยละ 96.88 สูงกว่าเกณฑ์ร้อยละ 80 H2: อัตราการใช้งานได้จริงของกฎร้อยละ 92.7 สูงกว่าเกณฑ์ร้อยละ 30 และ H3: ความแม่นยำของการแนะนำสินค้าร้อยละ 84.6 สูงกว่าเกณฑ์ร้อยละ 80 ยืนยันว่ากรอบการพัฒนาระบบ End-to-End บน n8n มีประสิทธิภาพเพียงพอสำหรับการนำไปใช้งานจริงในบริบท SMEs ไทย แม้ว่างานวิจัยนี้มิได้กำหนดตัวชี้วัดเชิงธุรกิจเป็นวัตถุประสงค์หลัก แต่จากการทดสอบในสภาพแวดล้อมจริงร่วมกับการสัมภาษณ์เจ้าของธุรกิจ พบผลกระทบเชิงบวกสามด้าน ได้แก่ (1) ด้านประสิทธิภาพการดำเนินงาน ระบบลดภาระผู้ดูแลในการตอบคำถามซ้ำซ้อนและทำหน้าที่ Cross-selling แทนมนุษย์ (2) ด้านความรวดเร็ว ระบบใช้เวลาเพียง 1.1–1.7 วินาทีต่อ Session ช่วยรักษาโมเมนตัมการสั่งซื้อไม่ให้ขาดตอน และ (3) ด้านประสบการณ์ผู้บริโภค คำแนะนำที่สม่ำเสมอช่วยให้ลูกค้าตัดสินใจซื้อเร็วขึ้น ทั้งนี้ การวัดผลกระทบเชิงปริมาณต่อยอดขายและอัตราการแปลงคำสั่งซื้อ (Conversion Rate) ถือเป็นแนวทางวิจัยในอนาคตที่สำคัญ

## 5. สรุปและข้อเสนอแนะ

งานวิจัยนี้นำเสนอกรอบการพัฒนาระบบแนะนำสินค้าอัตโนมัติบน LINE OA โดยใช้กฎความสัมพันธ์ที่สกัดจากข้อมูลบทสนทนา และใช้ n8n เป็น Central Orchestration Layer ตั้งแต่การรับข้อความ การสร้างและอัปเดต Rule Base ด้วย FP-Growth ทุกสัปดาห์ จนถึงการส่งคำแนะนำผ่าน Flex Message จุดสำคัญสองประการคือ (1) การใช้พจนานุกรมเฉพาะโดเมนในการสกัด Item เพื่อรับประกันคุณภาพ Transaction ก่อน Mining และ (2) การออกแบบ Stateful Architecture ที่เพิ่ม Recommendation Precision อย่างมีนัยสำคัญเมื่อเทียบกับ Stateless ข้อจำกัดคือใช้ข้อมูลจากร้านเบเกอรี่เพียงแห่งเดียว แนวทางในอนาคต ได้แก่ (1) ขยายการทดสอบไปยัง SMEs ประเภทอื่นเพื่อยืนยัน Generalizability (2) ผสาน Sequential Pattern Mining เพื่อวิเคราะห์ลำดับการตัดสินใจซื้อ และ (3) พัฒนาไปสู่ Hybrid AI ที่ผสานกฎความสัมพันธ์กับ LLM ตามแนวทางของ [14]

## เอกสารอ้างอิง


[1] ธีรวิทย์ ตันติอธิมงคล และคณะ. (2568). การวิเคราะห์พฤติกรรมการซื้อสินค้าออนไลน์ด้วย FP-Growth. Science & Technology Asia, 30(4), 290–302.

[2] ภูริวิชญ์ วีระจิตต์. (2565). ทัศนคติของผู้บริโภคต่อสื่อโฆษณาออนไลน์ฯ มหาวิทยาลัยมหิดล.

[3] อิสริยาภรณ์ วรรณา. (2566). การนำเทคโนโลยีใน LINE App มาประยุกต์ใช้. มหาวิทยาลัยธรรมศาสตร์.

[4] Adomavicius, G., & Tuzhilin, A. (2005). Toward the next generation of recommender systems. IEEE TKDE, 17(6), 734–749.

[5] Agrawal, R., & Srikant, R. (1994). Fast algorithms for mining association rules. Proc. VLDB, 487–499.

[6] Amershi, S., et al. (2019). Software engineering for machine learning. Proc. ICSE-SEIP, 291–300.

[7] Brin, S., et al. (1997). Dynamic itemset counting. Proc. ACM SIGMOD, 255–264.

[8] Davenport, T. H. (2018). The AI advantage. MIT Press.

[9] Davis, F. D. (1989). Perceived usefulness, perceived ease of use. MIS Quarterly, 13(3), 319–340.

[10] Feldman, R., & Sanger, J. (2007). The text mining handbook. Cambridge University Press.

[11] Han, J., Pei, J., & Yin, Y. (2000). Mining frequent patterns without candidate generation. ACM SIGMOD Record, 29(2), 1–12.

[12] LINE Corporation. (2024). LINE official account: Business solutions. https://www.linebiz.com

[13] Ay, F. C., et al. (2025). Conversation logs as a source of insight:Predicting user satisfaction for customer service chatbots. Telenor Research.

[14] Wang, Q., Li, J., Wang, S., Xing, Q., Niu, R., Kong, H., Li, R., Long, G., Chang, Y., & Zhang, C. (2024). Towards next-generation LLM-based recommender systems: A survey and beyond. arXiv preprint arXiv:2410.19744.

[15] Resnick, P., & Varian, H. R. (1997). Recommender systems. Communications of the ACM, 40(3), 56–58.

[16] Hasanova, K. (2023). Consumers' adoption of chatbots in the e-commerce industry (Master's thesis). Tallinn University of Technology.

[17] Muralidharan, S., et al. (2024). Role of AI in marketing through CRM integration with specific reference to chatbots. IJIRSS, 8(1), 1902–1909.

[18] Tan, P.-N., Steinbach, M., & Kumar, V. (2019). Introduction to data mining (2nd ed.). Pearson.

# การวิเคราะห์กลุ่มผู้บริโภคโดยใช้เทคนิคเหมืองข้อมูล: กรณีศึกษาการใช้ระบบสารสนเทศเพื่อผลิตภัณฑ์สีเขียว

# Consumer Segmentation Using Data Mining: A Case Study of Information Systems for Green Products

*ภัทรานิษฐ์ วสุอังศุนิตย์ (Phattranit Vasuangsunit)[1] และมณเฑียร รัตนศิริวงศ์วุฒิ (Montean Rattanasiriwongwut)[2]*

*สาขาวิชาระบบสารสนเทศเพื่อการจัดการ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ*

*[1]s607021857295@email.kmutnb.ac.th, [2] montean.r@itd.kmutnb.ac.th*

## บทคัดย่อ

*การค้นคว้าอิสระนี้มีวัตถุประสงค์เพื่อวิเคราะห์และจัดกลุ่มผู้บริโภคตามระดับการรับรู้ด้านสิ่งแวดล้อมและพฤติกรรมการใช้ระบบสารสนเทศเพื่อสนับสนุนการตัดสินใจเลือกซื้อผลิตภัณฑ์สีเขียวงานวิจัยนี้ใช้ข้อมูลที่ผ่านการเก็บรวบรวมจากแบบสอบถามของกลุ่มตัวอย่างจำนวน 600 คน ตัวแปรที่ใช้ในการวิเคราะห์ประกอบด้วยตัวแปรด้านการรับรู้สิ่งแวดล้อมและตัวแปรด้านพฤติกรรมการใช้ระบบสารสนเทศ มาทำการวิเคราะห์ด้วยวิธีการจัดกลุ่มอัลกอริทึม K-Means ซึ่งได้ค่า K ที่เหมาะสม เท่ากับ 3 โดยได้ทำการเปรียบเทียบอัลกอริทึม K-Means X-Mean และ K-Medoids ผลการทดสอบพบว่าอัลกอริทึม K-Means นั้นให้ผลลัพธ์ในด้านของประสิทธิภาพดีที่สุด โดยได้ค่าจุดศูนย์กลางกลุ่มข้อมูล (Avg. within centroid distance) เท่ากับ 0.267 ค่าเฉลี่ยความคล้ายกันระหว่างกลุ่ม (Davies Bouldin) เท่ากับ 0.116 และระยะเวลาการประมวลผล (Execution Time) เท่ากับ 0.124 วินาที สรุปได้ว่าอัลกอริทึม K-Means เหมาะสมที่สุดในการจัดกลุ่มการรับรู้ด้านสิ่งแวดล้อมและพฤติกรรมการใช้ระบบสารสนเทศเพื่อผลิตภัณฑ์สีเขียว*

**คำสำคัญ:** การทำเหมืองข้อมูล, การจัดกลุ่มข้อมูล, ผลิตภัณฑ์สีเขียว, K-Means, ระบบสารสนเทศ

## Abstract

*This independent study aimed to analyze and cluster consumers based on their environmental perception and information system usage behavior in supporting green product purchase decisions. Data were collected from 600 respondents using questionnaires and analyzed using the K-Means algorithm. The optimal number of clusters was three. The performance of K-Means was also compared with X-Means and K-Medoids. The results showed that K-Means achieved the best performance, with an average within-centroid distance of 0.267, a Davies–Bouldin Index of 0.116, and an execution time of0.124 seconds. Therefore, K-Means was considered the most suitable algorithm for clustering consumers according to environmental perception and information system usage behavior toward green products.*

**Keywords:** Data Mining, Clustering, Green Products, K-Means, Green Products, Information Systems

## 1. บทนำ

ประเด็นด้านสิ่งแวดล้อม เช่น ภาวะโลกร้อนการเปลี่ยนแปลงสภาพภูมิอากาศและการลดลงของทรัพยากรธรรมชาติได้กลายเป็นความท้ายทายสำคัญระดับโลกโดยพฤติกรรมการบริโภคของมนุษย์เป็นปัจจัยหนึ่งที่ส่งผลต่อปัญหาดังกล่าวอย่างมีนัยสำคัญ

งานวิจัยด้านการบริโภคบริโภคอย่างยั่งยืนจึงให้ความสำคัญกับการทำความเข้าใจว่าผู้บริโภครับรู้ประเด็นสิ่งแวดล้อมอย่างไรและนำข้อมูลจากระบบสารสนเทศมาใช้ประกอบการตัดสินใจเลือกซื้อผลิตภัณฑ์สีเขียวมากน้อยเพียงใด แม้ว่าระบบสารสนเทศด้านสิ่งแวดล้อมและแพลตฟอร์มผลิตภัณฑ์สีเขียวจะมีบทบาทมากขึ้น แต่ผู้บริโภคแต่ละกลุ่มยังมีระดับการรับรู้และรูปแบบการใช้งานที่แตกต่างกันงานวิจัยนี้จึงมุ่งใช้เทคนิคการจัดกลุ่มข้อมูลเพื่อค้นหารูปแบบผู้บริโภคจาก 2 มิติร่วมกัน ได้แก่ การรับรู้ด้านสิ่งแวดล้อมและพฤติกรรมการใช้ระบบสารสนเทศ ซึ่งเป็นมุมมองที่ช่วยเติมเต็มช่องว่างของงานวิจัยเดิมที่มักศึกษาทั้งสองด้านแบบแยกส่วน

## 2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

จากการศึกษาทฤษฎี แนวคิด และงานวิจัยที่เกี่ยวข้อง พบว่าการแบ่งกลุ่มข้อมูลเป็นแนวทางที่ช่วยอธิบายความแตกต่างของผู้บริโภคตามระดับการรับรู้ด้านสิ่งแวดล้อมและพฤติกรรมการใช้ระบบสารสนเทศเพื่อเลือกซื้อผลิตภัณฑ์สีเขียวได้อย่างชัดเจน การจัดกลุ่มทำให้สามารถจำแนกลักษณะเฉพาะของผู้บริโภคแต่ละกลุ่มได้อย่างเป็นระบบ ทั้งด้านการรับรู้ผลกระทบต่อสิ่งแวดล้อม คุณค่าของผลิตภัณฑ์สีเขียว และพฤติกรรมการใช้ระบบสารสนเทศ

ผลการวิเคราะห์ดังกล่าวเป็นประโยชน์ต่อการออกแบบระบบสารสนเทศ การสื่อสารข้อมูล และการกำหนดกลยุทธ์ทางการตลาดให้เหมาะสมกับผู้บริโภคแต่ละกลุ่มได้อย่างมีประสิทธิภาพ

### 2.1 การทำเหมืองข้อมูล (Data Mining)

การทำเหมืองข้อมูลเป็นกระบวนการสกัดรูปแบบหรือความรู้จากข้อมูลจำนวนมาก โดยเทคนิคการจัดกลุ่ม (Clustering) เป็นหนึ่งในวิธีสำคัญสำหรับการค้นหาโครงสร้างของข้อมูลแบบไม่มีตัวแปรกำกับ [1], [2] อัลกอริทึมที่นิยมใช้ ได้แก่ K-Means ซึ่งมีจุดเด่นด้านความรวดเร็วและเหมาะกับข้อมูลเชิงปริมาณ และเป็นอัลกอริทึมที่พยายามจัดกลุ่มข้อมูลให้สมาชิกภายในคลัสเตอร์เดียวกันมีความใกล้เคียงกันมากที่สุด K-Medoids ใช้วิธีการเลือกจุด Medoid เป็นศูนย์กลางเพื่อลดผลกระทบจากข้อมูลผิดปกติ (Outliers) และ X-Means พัฒนาจาก K- Meansซึ่งสามารถช่วยประมาณจำนวนกลุ่มที่เหมาะสมได้ [2], [3]

### 2.2 ผลิตภัณฑ์สีเขียวและระบบสารสนเทศ

งานวิจัยด้านผู้บริโภคและผลิตภัณฑ์สีเขียวชี้ว่า การรับรู้ผลกระทบต่อสิ่งแวดล้อม คุณค่าของผลิตภัณฑ์สีเขียว และการรับรู้บทบาทของตนเอง ล้วนมีผลต่อเจตคติและพฤติกรรมการเลือกซื้อ ขณะเดียวกัน ระบบสารสนเทศสามารถช่วยเพิ่มความโปร่งใสของข้อมูล ช่วยเปรียบเทียบทางเลือก และสนับสนุนการตัดสินใจได้อย่างมีประสิทธิภาพ [4]–[6] อย่างไรก็ตาม ยังมีงานจำนวนไม่มากที่บูรณาการตัวแปรด้านการรับรู้และพฤติกรรมการใช้ระบบสารสนเทศเข้าด้วยกันเพื่อจำแนกกลุ่มผู้บริโภคอย่างเป็นระบบ

### 2.3 งานวิจัยที่เกี่ยวข้อง

ณรงค์ศักดิ์ ชูสินชินภัทร (2558) ได้ทำการวิจัยการจัดกลุ่มปัจจัยเสี่ยงของนักเรียนที่มีโอกาสออกกลางคันโดยประยุกต์ใช้อัลกอริทึมเอ็กซ์มีนส์ (X-Means) เพื่อแก้ปัญหาการกำหนดค่า K ของวิธีเคมีนส์ (K-Means) ผลการศึกษาพบว่าแบบจำลองที่สร้างจากเอ็กซ์มีนส์ให้ความแม่นยำในการประเมินสูงถึงร้อยละ 95.12 ซึ่งสูงกว่าการใช้วิธีเคมีนส์เดิม

ทิชากร บุญเอี่ยม (2562) ได้พัฒนาระบบธุรกิจอัจฉริยะเพื่อสนับสนุนการตัดสินใจสำหรับผู้แนะนำการลงทุน โดยทำการศึกษาเปรียบเทียบ 3 อัลกอริทึม ได้แก่ K-Means, X-Means และ K-Medoids ผลการทดสอบพบว่าการจัดกลุ่มด้วยข้อมูลแบบ K-Means ให้ประสิทธิภาพดีที่สุด โดยมีค่าเฉลี่ยระยะห่างของทุกจุดจากจุดศูนย์กลางของคลัสเตอร์เท่ากับ 2.618 และใช้เวลาในการประมวลผลเพียง 3 วินาที

กิตติพงษ์ ศรีสุข และ ปาริชาติ บุญยืน (2564) ได้ประยุกต์ใช้เทคนิคการทำเหมืองข้อมูลโดยใช้อัลกอริทึม K-Means Clustering เพื่อวิเคราะห์และจัดกลุ่มพฤติกรรมผู้บริโภค โดยมีวัตถุประสงค์เพื่อค้นหารูปแบบและความ

แตกต่างของผู้บริโภคในแต่ละกลุ่ม ผลการวิจัยพบว่าสามารถแบ่งกลุ่มผู้บริโภคได้อย่างชัดเจนตามลักษณะพฤติกรรม ซึ่งช่วยสนับสนุนการวางแผนกลยุทธ์ทางการตลาดและการตัดสินใจเชิงธุรกิจได้อย่างมีประสิทธิภาพ

## 3. วิธีดำเนินการวิจัย

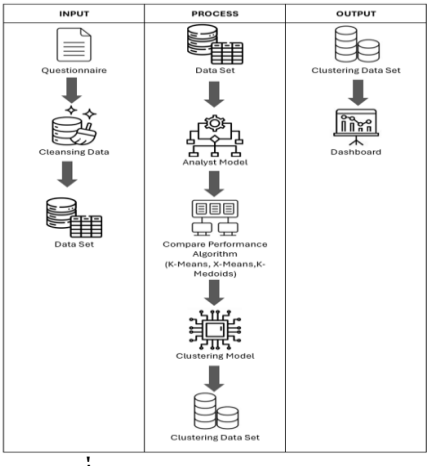
การวิจัยครั้ง มีวัตถุประสงค์การจัดกลุ่มการรับรู้ด้านสิ่งแวดล้อม และพฤติกรรมในการใช้ระบบสารสนเทศเพื่อผลิตภัณฑ์สีเขียว โดยดำเนินการตามกรอบแนวคิดของงานวิจัยดังนี้

### 3.1 การเก็บรวบรวมข้อมูล

งานวิจัยนี้เก็บข้อมูลจากกลุ่มตัวอย่างจำนวน 600 คน โดยใช้แบบสอบถามมาตราส่วนลิเคิร์ต 5 ระดับ กลุ่มตัวอย่างเป็นผู้บริโภควัยทำงานที่สามารถเข้าถึงหรือใช้งานระบบสารสนเทศได้ เครื่องมือวิจัยคือแบบสอบถามที่ใช้ประเมินข้อมูลสำคัญ ประกอบด้วยแบบสอบถามการรับรู้ด้านสิ่งแวดล้อม 3 ด้าน ซึ่งแต่ละด้านประด้วยข้อคำถามด้านละ 5 คำถาม รวมทั้งสิ้น 15 คำถาม ได้แก่ ข้อมูลด้านการรับรู้สิ่งแวดล้อม ประกอบด้วย การรับรู้ผลกระทบต่อสิ่งแวดล้อม (EnvImpact), การรับรู้คุณค่าของผลิตภัณฑ์สีเขียว (GreenValue) และ การรับรู้บทบาทของตนเองต่อสิ่งแวดล้อม (SelfRole) สำหรับพฤติกรรมการใช้ระบบสารสนเทศเพื่อผลิตภัณฑ์สีเขียว ประกอบด้วยแบบสอบถามพฤติกรรมการใช้ระบบสารสนเทศ 5 พฤติกรรม ซึ่งประกอบด้วยข้อคำถาม 5 คำถาม ในแต่ละพฤติกรรม รวมทั้งสิ้น 25 คำถาม ประกอบด้วย ความถี่ในการใช้ระบบสารสนเทศ (ISFreq), การใช้ระบบเพื่อเปรียบเทียบข้อมูล (ISCompare), การมีส่วนร่วมกับระบบ (ISEngage) การใช้ข้อมูลประกอบการตัดสินใจ, (ISDecision) และความตั้งใจใช้ระบบในอนาคต (ISIntention)

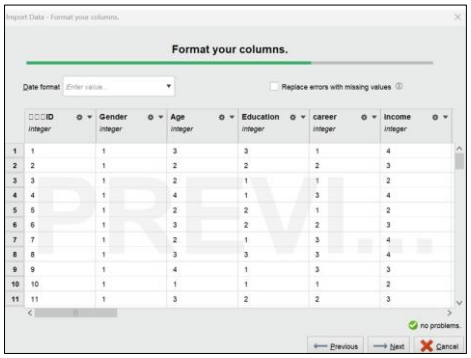
### 3.2 กรอบแนวความคิดของงานวิจัย

การวิจัยครั้งนี้ได้กำหนดกรอบแนวคิดของการดำเนินงานวิจัยเพื่อใช้เป็นแนวทางในการวิเคราะห์และประมวลผลข้อมูล โดยกรอบแนวคิดดังกล่าวแสดงถึงกระบวนการดำเนินงานตั้งแต่การนำเข้าข้อมูล การวิเคราะห์ข้อมูลด้วยแบบจำลองทางการทำเหมืองข้อมูลไปจนถึงการนำเสนอผลลัพธ์ในรูปแบบของรายงานเชิงวิเคราะห์ ทั้งนี้กรอบแนวคิดของการวิจัยสามารถอธิบายได้ตามลำดับขั้นตอนดังนี้



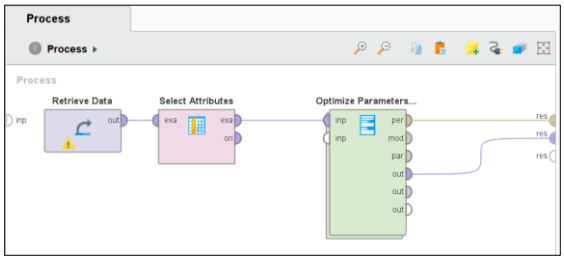

**ภาพที่ 1 :** กรอบแนวความคิดในการดำเนินการวิจัย

ส่วนที่ 1 การนำเข้า (Input) ข้อมูลสำหรับการพัฒนาแบบจำลอง หลังจากที่ผู้วิจัยได้ดำเนินการตรวจสอบและเตรียมข้อมูลให้มีความถูกต้องและเหมาะสม

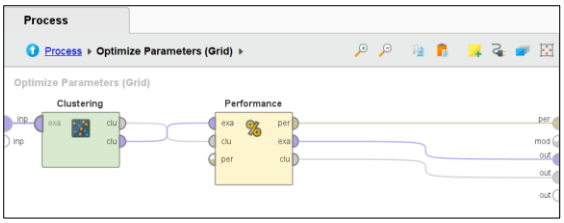



**ภาพที่ 2** : นำเข้าข้อมูลในรูปแบบตารางข้อมูล

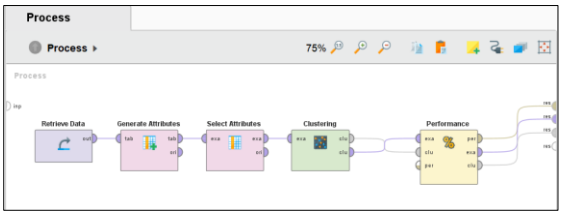
ส่วนที่ 2 กระบวนการวิเคราะห์ (Process) หาจำนวนกลุ่มที่เหมาะสมด้วย Elbow Method เปรียบเทียบอัลกอริทึม K-Means, X-Means และ K-Medoids โดยประเมินด้วย Avg. Within Centroid Distance, Davies–Bouldin Index (DBI) และExecution Time



**ภาพที่ 3 :** ปรับค่าตัวแปรของอัลกอริทึมใน Program RapidMiner




**ภาพที่ 4 :** การวัดประสิทธิภาพของการจัดกลุ่มข้อมูล



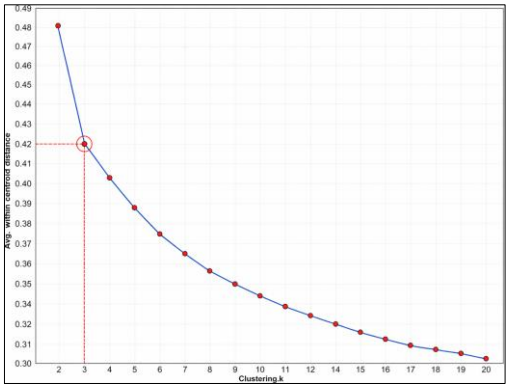

**ภาพที่ 5 :** การแบ่งกลุ่มข้อมูลด้วย K-Means Clustering

ส่วนที่ 3 ผลลัพธ์ (Output) วิเคราะห์ข้อมูลเพื่อนำมาจัดทำเป็นรายงานและนำเสนอในรูปแบบของการแสดงผลเชิงภาพ เพื่อใช้ในการนำเสนอข้อมูลเชิงลึกจากการวิเคราะห์ข้อมูล การนำเสนอในรูปแบบแดชบอร์ดช่วยให้สามารถสื่อสารข้อมูลที่มีความซับซ้อนให้อยู่ในรูปแบบที่เข้าใจง่ายผ่านโปรแกรม Power BI Desktop

## 4. ผลการวิจัย

### 4.1 การทดสอบประสิทธิภาพของการจัดกลุ่มข้อมูล

ผลการหาจำนวนกลุ่มที่เหมาะสมพบว่า K = 3 เป็นค่าที่เหมาะสมสำหรับการอธิบายโครงสร้างข้อมูล โดยให้ค่า Avg. Within Centroid Distance เท่ากับ 0.421 และ DBI เท่ากับ 0.073




**ภาพที่ 6 :** ผลการหาจำนวนกลุ่มที่เหมาะสมด้วย Elbow Method

เมื่อนำ K = 3 ไปใช้เปรียบเทียบอัลกอริทึม พบว่า K-Means ให้ผลดีที่สุดทั้งในด้านความกระชับของกลุ่ม ความแตกต่างระหว่างกลุ่ม และเวลาประมวลผล

**ตารางที่ 1:** เปรียบเทียบประสิทธิภาพของอัลกอริทึมการจัดกลุ่ม

| Algorithm | Performance summary |
|---|---|
| K-Means | Avg.=0.267, DBI=0.116, Time=0.124 s |
| X-Means | Avg.=0.287, DBI=0.142, Time=0.226 s |
| K-Medoids | Avg.=0.516, DBI=0.138, Time=11.434 s |

เมื่อใช้ K-Means เพื่อวิเคราะห์โครงสร้างข้อมูล พบว่าสามารถแบ่งผู้ตอบแบบสอบถามออกเป็น 3 กลุ่ม ได้แก่ กลุ่มที่ 1 จำนวน 101 คน (16.8%) กลุ่มที่ 2 จำนวน 320 คน (53.3%) และกลุ่มที่ 3 จำนวน 179 คน (29.8%) โดยกลุ่มที่ 2 เป็นกลุ่มขนาดใหญ่ที่สุด

**Cluster Model**

```
Cluster 0: 101 items
Cluster 1: 320 items
Cluster 2: 179 items
Total number of items: 600
```

**ภาพที่ 7 :** โมเดลการจัดกลุ่ม (Cluster Model)

| Attribute | cluster_0 | cluster_1 | cluster_2 |
|---|---|---|---|
| Gender | 1.584 | 1.597 | 1.631 |
| Age | 2.366 | 2.334 | 1.989 |
| Education | 1.465 | 1.816 | 1.832 |
| Carrer | 3.752 | 3.112 | 4.989 |
| Income | 2.198 | 2.569 | 2.145 |
| EnvImpact | 3.617 | 4.312 | 4.343 |
| GreenValue | 3.165 | 4.148 | 4.194 |
| SelfRole | 3.241 | 4.117 | 4.181 |
| ISFreq | 2.670 | 3.958 | 3.994 |
| ISCompare | 2.815 | 4.097 | 4.143 |
| ISIntention | 2.294 | 3.862 | 3.907 |
| ISEngage | 2.683 | 4.034 | 4.074 |
| ISDecision | 2.703 | 4.094 | 4.060 |

**ภาพที่ 8 :** ผลของการจัดกลุ่ม K = 3

### 4.2 การวิเคราะห์ระดับการรับรู้ด้านสิ่งแวดล้อมและพฤติกรรมการใช้ระบบสารสนเทศ

เมื่อพิจารณาค่าเฉลี่ยของการรับรู้ด้านสิ่งแวดล้อมและพฤติกรรมการใช้ระบบสารสนเทศเพื่อผลิตภัณฑ์สีเขียว สามารถอธิบายรายละเอียดของแต่ละกลุ่มได้ดังนี้

กลุ่มที่ 1 เมื่อพิจารณาค่าเฉลี่ยของตัวแปรที่ใช้ในการวิเคราะห์ พบว่ากลุ่มนี้มีระดับการรับรู้ด้านสิ่งแวดล้อมอยู่ในระดับปานกลาง ได้แก่ EnvImpact = 3.62, GreenValue = 3.17 และ SelfRole = 3.24 อย่างไรก็ตาม พฤติกรรมการใช้ระบบสารสนเทศเพื่อผลิตภัณฑ์สี

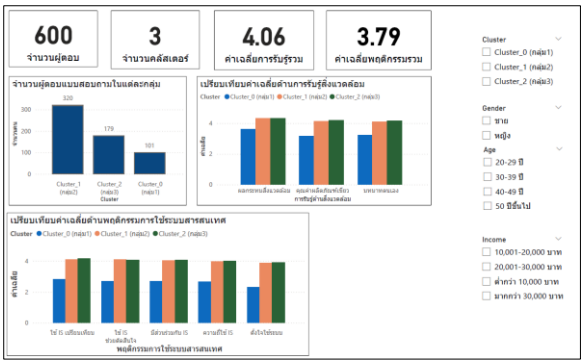
เขียวยังอยู่ในระดับค่อนข้างต่ำ ได้แก่ ISFreq = 2.67, ISCompare = 2.82, ISEngage = 2.68, ISDecision = 2.70 และ ISIntention = 2.29 โดยเฉพาะตัวแปรความตั้งใจใช้ระบบอย่างต่อเนื่องที่มีค่าค่อนข้างต่ำ ผลการวิเคราะห์สะท้อนลักษณะของผู้ใช้กลุ่มนี้ว่าเป็นกลุ่มที่ “มีการรับรู้แต่ยังไม่เกิดพฤติกรรมการใช้งานจริงอย่างต่อเนื่อง” ซึ่งอาจเกิดจากการยังไม่เห็นประโยชน์ของระบบอย่างชัดเจน หรือประสบการณ์การใช้งานที่ยังไม่เอื้อต่อการเริ่มต้นใช้ระบบสารสนเทศ ดังนั้น แนวทางในการยกระดับการใช้งานของกลุ่มนี้ควรมุ่งเน้นการลดความซับซ้อนของข้อมูล การนำเสนอข้อมูลในรูปแบบที่เข้าใจง่าย เช่น การใช้เนื้อหาขนาดสั้น (Micro-content) การใช้ไอคอนหรือคะแนนด้านความยั่งยืนที่ตีความได้ง่าย รวมถึงการสร้างแรงจูงใจเริ่มต้นในการใช้งาน เช่น การให้แต้มสะสมหรือคูปองส่วนลด เพื่อกระตุ้นให้เกิดการทดลองใช้ระบบและพัฒนาไปสู่การใช้งานอย่างต่อเนื่อง

กลุ่มที่ 2 เมื่อพิจารณาด้านการรับรู้และพฤติกรรม พบว่ากลุ่มนี้มีค่าเฉลี่ยของตัวแปรทั้งด้านการรับรู้สิ่งแวดล้อมและพฤติกรรมการใช้ระบบสารสนเทศอยู่ในระดับสูง ได้แก่ EnvImpact = 4.31, GreenValue = 4.15 และ SelfRole = 4.12 สำหรับตัวแปรด้านพฤติกรรมการใช้ระบบสารสนเทศพบว่าอยู่ในระดับค่อนข้างสูงเช่นกัน ได้แก่ ISFreq = 3.96, ISCompare = 4.10, ISEngage = 4.03, ISDecision = 4.09 และ ISIntention = 3.86 ผลการวิเคราะห์สะท้อนให้เห็นว่าผู้ใช้ในกลุ่มนี้เป็นกลุ่มที่มีพฤติกรรมการใช้งานระบบสารสนเทศเชิงรุก โดยมีการค้นหาข้อมูล เปรียบเทียบข้อมูล และใช้ข้อมูลจากระบบสารสนเทศประกอบการตัดสินใจเกี่ยวกับผลิตภัณฑ์สีเขียวอย่างสม่ำเสมอ ดังนั้น หากระบบหรือแพลตฟอร์มสามารถพัฒนาเครื่องมือที่ช่วยสนับสนุนการตัดสินใจได้อย่างมีประสิทธิภาพ เช่น ฟังก์ชันการเปรียบเทียบผลิตภัณฑ์ที่ชัดเจน ระบบรีวิวที่มีความน่าเชื่อถือ หรือข้อมูลฉลากและมาตรฐานด้านสิ่งแวดล้อมที่สามารถตรวจสอบย้อนกลับได้ ก็จะช่วยเพิ่มประสิทธิภาพในการใช้งานระบบและส่งเสริมให้เกิดการใช้งานอย่างต่อเนื่องมากยิ่งขึ้น

กลุ่มที่ 3 เมื่อพิจารณาค่าเฉลี่ยของตัวแปรที่ใช้ในการวิเคราะห์ พบว่ากลุ่มนี้มีระดับการรับรู้ด้านสิ่งแวดล้อมอยู่ในระดับสูง ได้แก่ EnvImpact = 4.34, GreenValue = 4.19 และ SelfRole = 4.18 และมีพฤติกรรมการใช้ระบบสารสนเทศเพื่อผลิตภัณฑ์สีเขียวในระดับสูงอย่างสม่ำเสมอ ได้แก่ ISFreq = 3.99, ISCompare = 4.14, ISEngage = 4.07, ISDecision = 4.06 และ ISIntention = 3.91 ผลการวิเคราะห์สะท้อนให้เห็นว่าผู้ใช้ในกลุ่มนี้เป็นกลุ่มที่มีความตระหนักและความมุ่งมั่นด้านความยั่งยืนในระดับสูง รวมทั้งมีพฤติกรรมการใช้ระบบสารสนเทศอย่างต่อเนื่องเพื่อสนับสนุนการตัดสินใจเลือกผลิตภัณฑ์ที่เป็นมิตรต่อสิ่งแวดล้อม ดังนั้น การพัฒนาระบบสารสนเทศสำหรับกลุ่มผู้ใช้นี้ควรมุ่งเน้นการออกแบบประสบการณ์การใช้งานที่สามารถรักษาการใช้งานในระยะยาว เช่น ฟีเจอร์การติดตามผลลัพธ์ด้านความยั่งยืน (เช่น การลดการปล่อยคาร์บอนหรือการลดปริมาณขยะ) ระบบสมาชิก หรือกิจกรรมและภารกิจด้านความยั่งยืน รวมถึงระบบรางวัลตามพฤติกรรมการใช้งาน ซึ่งจะช่วยเสริมสร้างความผูกพันของผู้ใช้กับระบบและส่งเสริมการใช้งานอย่างต่อเนื่อง

### 4.3 การนำเสนอผลการวิเคราะห์ผ่านโปรแกรม Power BI

ผลการจัดกลุ่มผู้บริโภคโดยแสดงข้อมูลสำคัญ ได้แก่ จำนวนผู้ตอบทั้งหมด จำนวนคลัสเตอร์ ค่าเฉลี่ยการรับรู้ด้านสิ่งแวดล้อม และค่าเฉลี่ยพฤติกรรมการใช้ระบบสารสนเทศเพื่อเลือกซื้อผลิตภัณฑ์สีเขียว พร้อมทั้งแสดงจำนวนผู้ตอบในแต่ละกลุ่ม เพื่อให้เห็นสัดส่วนการกระจายตัวของผู้บริโภคอย่างชัดเจน นอกจากนี้ยังมีกราฟเปรียบเทียบค่าเฉลี่ยด้านการรับรู้สิ่งแวดล้อมและพฤติกรรมการใช้ระบบสารสนเทศของแต่ละกลุ่ม รวมถึงตัวกรองด้านประชากรศาสตร์ เช่น เพศ อายุ และรายได้ เพื่อช่วยวิเคราะห์ความแตกต่างของผู้บริโภคในแต่ละคลัสเตอร์




**ภาพที่** 9 **:** รายงานภาพรวมของการจัดกลุ่ม

แสดงผลการวิเคราะห์เชิงลึกของผู้บริโภคในแต่ละกลุ่ม โดยครอบคลุมทั้งขนาดกลุ่ม ระดับการรับรู้ด้านสิ่งแวดล้อม พฤติกรรมการใช้ระบบสารสนเทศ และข้อมูลประชากรศาสตร์ที่สำคัญมีการแสดงการกระจายตัวของผู้ตอบตามอาชีพ รายได้ เพศ การศึกษา และช่วงอายุ เพื่อให้เห็นลักษณะเด่นของแต่ละกลุ่มอย่างชัดเจน พร้อมทั้งเปรียบเทียบคะแนนการรับรู้ด้านสิ่งแวดล้อมและคะแนนพฤติกรรมการใช้ระบบสารสนเทศในแต่ละด้านทำให้สามารถอธิบายลักษณะเฉพาะและความแตกต่างของผู้บริโภคในแต่ละคลัสเตอร์




**ภาพที่ 10** : การเปรียบเทียบลักษณะของแต่ละคลัสเตอร์

## 5. สรุป

### 5.1 สรุปผลการวิจัย

ผู้วิจัยได้พัฒนาแบบจำลองการจัดกลุ่มข้อมูลด้วยโปรแกรม RapidMiner Studio โดยใช้ข้อมูลที่ผ่านการเตรียมแล้วจำนวน 600 รายการ และวิเคราะห์ด้วยเทคนิค K-Means Clustering ผลการวิเคราะห์สามารถแบ่งข้อมูลออกได้เป็น 3 กลุ่มอย่างเหมาะสม เมื่อเปรียบเทียบประสิทธิภาพระหว่างอัลกอริทึม K-Means, X-Means และ K-Medoids พบว่า K-Means ให้ผลลัพธ์ดีที่สุด โดยมีค่า Avg. within centroid distance เท่ากับ 0.267 ค่า Davies-Bouldin Index เท่ากับ 0.116 และใช้เวลาในการประมวลผล 0.124 วินาที จึงสรุปได้ว่า K-Means เป็นอัลกอริทึมที่เหมาะสมที่สุดสำหรับการจัดกลุ่มการรับรู้ด้านสิ่งแวดล้อมและพฤติกรรมการใช้ระบบสารสนเทศเพื่อผลิตภัณฑ์สีเขียวในงานวิจัยนี้ เพื่อค้นหาโครงสร้างเชิงพฤติกรรมของผู้บริโภคโดยไม่ต้องมีการกำหนดกลุ่มล่วงหน้า โดยบูรณาการตัวแปรด้านการรับรู้สิ่งแวดล้อมและพฤติกรรมการใช้ระบบสารสนเทศเข้าด้วยกัน ทำให้สามารถจำแนกผู้บริโภคออกเป็นกลุ่มที่มีลักษณะเฉพาะได้อย่างชัดเจน ผลลัพธ์ไม่เพียงช่วยอธิบายความแตกต่างของผู้บริโภค แต่ยังสามารถนำไปใช้ในการออกแบบระบบและกำหนดกลยุทธ์ที่เหมาะสมกับแต่ละกลุ่มได้อย่างมีประสิทธิภาพ

### 5.2 อภิปรายผล

การรับรู้ด้านสิ่งแวดล้อมมีความสัมพันธ์กับพฤติกรรมการใช้ระบบสารสนเทศเพื่อสนับสนุนการเลือกซื้อผลิตภัณฑ์สีเขียวอย่างชัดเจน โดยผู้บริโภคที่มีระดับการรับรู้สูงมีแนวโน้มใช้ระบบสารสนเทศมากกว่าในด้านการค้นหา เปรียบเทียบ และตัดสินใจเลือกซื้อสินค้า ผลการเปรียบเทียบอัลกอริทึมยังชี้ให้เห็นว่า K-Means มีความเหมาะสมในการจำแนกกลุ่มข้อมูลลักษณะนี้ เนื่องจากให้ผลลัพธ์ที่มีความกระชับภายในกลุ่มและมีความแตกต่างระหว่างกลุ่มอย่างชัดเจน ดังนั้นผลการศึกษานี้จึงเป็นประโยชน์ทั้งในเชิงวิชาการสำหรับการอธิบายพฤติกรรมผู้บริโภค และในเชิงปฏิบัติสำหรับการออกแบบระบบสารสนเทศและกลยุทธ์การส่งเสริมผลิตภัณฑ์สีเขียวให้สอดคล้องกับลักษณะของผู้บริโภคแต่ละกลุ่ม ผลลัพธ์สามารถนำไปใช้สนับสนุนการออกแบบระบบสารสนเทศ การพัฒนาฟังก์ชันการสื่อสาร และการกำหนดกลยุทธ์เพื่อส่งเสริมการบริโภคอย่างยั่งยืนได้อย่างมีประสิทธิภาพ

### 5.3 ข้อเสนอแนะ

ด้านการทำเหมืองข้อมูลสามารถพัฒนาการวิเคราะห์พฤติกรรมการจัดกลุ่มโดยวิเคราะห์จาก แอททริบิวต์ที่น่าสนใจ หรือการทำเหมืองข้อมูลเพื่อหาความสัมพันธ์ที่ชัดเจนมากขึ้น และสำหรับเทคนิคในการแบ่งกลุ่มแต่ละวิธีอาจให้ผลลัพธ์ที่แตกต่างกัน ดังนั้นการเลือกใช้วิธีการใดขึ้นอยู่กับข้อมูลและวัตถุประสงค์ของการนำไปใช้งานและการประยุกต์ใช้ต่อไป

## เอกสารอ้างอิง

[1] J. Han, M. Kamber, and J. Pei, Data Mining: Concepts and Techniques, 3rd ed. Waltham, MA: Morgan Kaufmann, 2012.

[2] P.-N. Tan, M. Steinbach, and V. Kumar, Introduction to Data Mining, 2nd ed. Boston, MA: Pearson, 2019.

[3] D. Pelleg and A. Moore, “X-means: Extending K-means with Efficient Estimation of the Number of Clusters,” in Proc. ICML, 2000, pp. 727–734.

[4] United Nations Environment Programme (UNEP), Sustainable Consumption and Production: A Handbook for Policymakers, 2021.

[5] OECD, Empowering Consumers for the Green Transition, Paris: OECD Publishing, 2020.

[6] IPCC, Climate Change 2022: Mitigation of Climate Change. Cambridge: Cambridge University Press, 2022.

[7] ณรงค์ศักดิ์ ชูสินชินภัทร. (2558). การจัดกลุ่มนักเรียนที่มีโอกาสออกกลางคันโดยใช้ขั้นตอนวิธีเอ็กซ์มีนส์.

[8] กิตติพงษ์ ศรีสุข และ ปาริชาติ บุญยืน. (2564). “การประยุกต์ใช้เทคนิคการทำเหมืองข้อมูลในการวิเคราะห์พฤติกรรมผู้บริโภค.

[9] ทิชากร บุญเอี่ยม. (2562). การพัฒนาระบบธุรกิจอัจฉริยะเพื่อสนับสนุนการติดสินใจสำหรับผู้แนะนำการลงทุน.

# การจำแนกความคิดเห็นของผู้ใช้งานบนโซเชียลมีเดียเกี่ยวกับเทคโนโลยีปัญญาประดิษฐ์ (AI) ด้วยเทคนิคเหมืองข้อความ

# Classification of Social Media User Opinions on Artificial Intelligence (AI) Using Text Mining Techniques

*ธนัทเมศร์ วุฒิสถิตย์ธรรม (Thanatmeth Wuttisathittham)[1] และศักดิ์ชาย ตั้งวรรณวิทย์ (Sakchai Tangwannawit)[2]*
*ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล*
*มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ*
*[1]s6707021857252@email.kmutnb.ac.th, [2] sakchai.t@itd.kmutnb.ac.th*

## บทคัดย่อ

*งานวิจัยนี้มีวัตถุประสงค์เพื่อ 1) วิเคราะห์และจำแนกทัศนคติของผู้ใช้งานโซเชียลมีเดียในประเทศไทยต่อเทคโนโลยีปัญญาประดิษฐ์ (AI) ออกเป็นเชิงบวกและเชิงลบ 2) เปรียบเทียบประสิทธิภาพของอัลกอริทึมการเรียนรู้ของเครื่องจำนวน 8 อัลกอริทึม เพื่อหาโมเดลที่เหมาะสมที่สุด 3) นำเสนอผลการวิเคราะห์ข้อมูลในรูปแบบสารสนเทศเชิงภาพโดยใช้เทคนิคเหมืองข้อความ (Text Mining) และเปรียบเทียบประสิทธิภาพของ 8 อัลกอริทึม ได้แก่ Random Forest, Naïve Bayes, SVM, Deep Learning, Decision Tree, K-Nearest Neighbors (KNN), Neural Network และ Logistic Regression ข้อมูลถูกรวบรวมจาก Facebook, YouTube, Pantip และ TikTok ระหว่างเดือนมกราคม 2566 ถึงสิงหาคม 2568 เฉพาะข้อความภาษาไทย ดำเนินการวิเคราะห์ด้วย Python, RapidMiner และนำเสนอผลผ่าน Power BI ผลการวิจัยพบว่า อัลกอริทึม KNN มีประสิทธิภาพสูงสุดในการจำแนก โดยให้ค่าความถูกต้อง (Accuracy) 82.09% ความแม่นยำ (Precision) 88.68% การระลึก (Recall) 84.77% และค่าความถ่วงดุล (F-Measure) 86.68% ผลการศึกษาชี้ให้เห็นว่า KNN เป็นโมเดลที่มีความเสถียรและเหมาะสมที่สุดสำหรับการวิเคราะห์ความรู้สึกจากข้อมูลโซเชียลมีเดียในบริบทของงานวิจัยนี้*

**คำสำคัญ:** ปัญญาประดิษฐ์ เหมืองข้อความ การจำแนกความคิดเห็น

## Abstract

*The objectives of this research are: 1) to analyze and classify the sentiments of social media users in Thailand toward Artificial Intelligence (AI) technology into positive and negative categories; 2) to compare the performance of eight machine learning algorithms to identify the most suitable model; and 3) to present the data analysis results through information visualization. Using text mining techniques, the study evaluates eight algorithms: Random Forest, Naïve Bayes, SVM, Deep Learning, Decision Tree, K-Nearest Neighbors (KNN), Neural Network, and Logistic Regression. The dataset was collected from Thai-language posts on Facebook, YouTube, Pantip, and TikTok between January 2023 and August 2025. The analytical process was conducted using Python and RapidMiner, with the final findings presented via Power BI. The results indicate that the KNN algorithm achieved the highest performance in classification, with an Accuracy of 82.09%, Precision of 88.68%, Recall of 84.77%, and an F-Measure of 86.68%. The study concludes that KNN is a stable and the most appropriate model for sentiment analysis of social media data within the context of this research.*

**Keywords:** Artificial Intelligence, Text Mining, Opinion Classification.

## 1. บทนำ

ในปัจจุบันเทคโนโลยีปัญญาประดิษฐ์ (Artificial Intelligence: AI) ได้เข้ามามีบทบาทสำคัญในการขับเคลื่อนเศรษฐกิจและสังคมไทยอย่างแพร่หลาย ทั้งในภาคการแพทย์ การเงิน และการศึกษา การขยายตัวนี้ส่งผลให้เกิดการแสดงความคิดเห็นที่หลากหลายบนสื่อสังคมออนไลน์ อาทิ Facebook, YouTube, Pantip และ TikTok ซึ่งสะท้อนถึงการยอมรับเทคโนโลยี (Technology

Acceptance) ทั้งในแง่บวกด้านการเพิ่มประสิทธิภาพการดำเนินชีวิต และในแง่ลบด้านจริยธรรมรวมถึงความมั่นคงในอาชีพ

อย่างไรก็ตาม ข้อมูลบนสื่อสังคมออนไลน์มีลักษณะเป็นข้อมูลขนาดใหญ่ (Big Data) ที่มีความหลากหลายและไม่เป็นระเบียบ ทำให้การวิเคราะห์ด้วยวิธีการดั้งเดิมมีข้อจำกัดด้านระยะเวลาและความแม่นยำ งานวิจัยนี้จึงมุ่งเน้นการประยุกต์ใช้เทคนิคเหมืองข้อความ (Text Mining) เพื่อวิเคราะห์ความรู้สึก (Sentiment Analysis) โดยทำการเปรียบเทียบประสิทธิภาพของอัลกอริทึมการเรียนรู้ของเครื่อง (Machine Learning) 8 โมเดล ได้แก่ Random Forest, Naïve Bayes, SVM, Deep Learning, Decision Tree, K-Nearest Neighbors (KNN), Neural Network และ Logistic Regression เพื่อจำแนกความคิดเห็นออกเป็นกลุ่มเชิงบวก (Positive) และเชิงลบ (Negative)

ผลการวิจัยนี้จะเป็นข้อมูลเชิงยุทธศาสตร์สำหรับหน่วยงานที่เกี่ยวข้อง ในการทำเข้าใจทัศนคติของผู้ใช้งาน เพื่อนำไปประกอบการตัดสินใจเชิงนโยบาย การพัฒนาผลิตภัณฑ์ และการวางรากฐานทางจริยธรรม เพื่อเตรียมความพร้อมให้สังคมไทยก้าวสู่ยุคปัญญาประดิษฐ์ได้อย่างยั่งยืน

## 2. เอกสารและงานวิจัยที่เกี่ยวข้อง

ในการดำเนินการวิจัยเรื่องการจำแนกความคิดเห็นของผู้ใช้งานบนโซเชียลมีเดียเกี่ยวกับเทคโนโลยีปัญญาประดิษฐ์ ผู้วิจัยได้ทำการรวบรวมทฤษฎีและงานวิจัยที่เกี่ยวข้องเพื่อใช้เป็นแนวทางในการดำเนินงานวิจัย ดังนี้

**2.1 เทคโนโลยีปัญญาประดิษฐ์ (Artificial Intelligence: AI)**

เป็นเทคโนโลยีที่มีบทบาทสำคัญในการเปลี่ยนแปลงรูปแบบการทำงานและการใช้ชีวิตในปัจจุบัน โดยในบริบทของสังคมไทย AI ถูกนำมาประยุกต์ใช้ในหลากหลายอุตสาหกรรม เช่น การแพทย์ การศึกษา และการเงิน ส่งผลให้เกิดการแสดงความคิดเห็นที่หลากหลายบนแพลตฟอร์มโซเชียลมีเดีย ทั้งในแง่ของการสนับสนุนเพื่อเพิ่มประสิทธิภาพการทำงาน และความกังวลด้านจริยธรรมหรือผลกระทบต่อตลาดแรงงาน [6]

**2.2 การทำเหมืองข้อความ (Text Mining)**

เหมืองข้อความ เป็นกระบวนการสกัดความรู้หรือข้อมูลที่เป็นประโยชน์จากข้อความที่ไม่มีโครงสร้าง (Unstructured Data) งานวิจัยนี้มุ่งเน้นการทำ Sentiment Analysis เพื่อจำแนกความคิดเห็นออกเป็นกลุ่มเชิงบวก (Positive) และเชิงลบ (Negative) โดยใช้ข้อมูลภาษาไทยจากแพลตฟอร์ม Facebook, YouTube, Pantip และ TikTok ซึ่งต้องผ่านกระบวนการจัดการข้อมูลด้วยเครื่องมือ เช่น Excel และ Google Sheets ก่อนนำไปสร้างโมเดล [7]

**2.3 การวิเคราะห์ความรู้สึก (Sentiment Analysis)**

การวิเคราะห์ความรู้สึกเป็นสาขาหนึ่งของเหมืองข้อความที่มุ่งเน้นการจำแนกอารมณ์หรือทัศนคติที่ปรากฏในข้อความ งานวิจัยนี้กำหนดการจำแนกออกเป็น 2 กลุ่มหลัก คือ ความคิดเห็นเชิงบวก (Positive) และความคิดเห็นเชิงลบ (Negative) เพื่อสะท้อนให้เห็นถึงแนวโน้มการยอมรับเทคโนโลยี AI ของคนไทยในปัจจุบัน [8]

**2.4 ระบบธุรกิจอัจฉริยะ (Business Intelligence)**

ระบบธุรกิจอัจฉริยะเป็นเครื่องมือที่ช่วยในการรวบรวม วิเคราะห์ และนำเสนอข้อมูลเพื่อประกอบการตัดสินใจเชิงกลยุทธ์ งานวิจัยนี้ได้นำซอฟต์แวร์ Power BI มาใช้ในการสร้าง Dashboard เพื่อนำเสนอผลการวิเคราะห์ข้อมูลความรู้สึกเชิงลึก ช่วยให้หน่วยงานที่เกี่ยวข้องสามารถมองเห็นภาพรวมของทัศนคติสังคมไทยต่อ AI ได้อย่างชัดเจนและมีประสิทธิภาพ [9]

**2.5 งานวิจัยที่เกี่ยวข้อง**

จากการทบทวนงานวิจัยที่เกี่ยวข้องกับการวิเคราะห์ความรู้สึก (Sentiment Analysis) ด้วยเทคนิคเหมืองข้อความ พบว่ามีการประยุกต์ใช้ในหลากหลายบริบทและใช้อัลกอริทึมที่แตกต่างกันตามลักษณะของชุดข้อมูล ดังนี้

วสวัตติ์ อินทร์แปลง [1] ศึกษาความคิดเห็นต่อเกมมือถือภาษาไทย โดยใช้เทคนิค TF-IDF ร่วมกับ SMOTE เพื่อแก้ปัญหาความไม่สมดุลของข้อมูล พบว่าอัลกอริทึม SVM ให้ค่าความถูกต้องสูงสุดถึงร้อยละ 99.87

นิธิกร เลิศชาญวุฒิ [2] ที่วิเคราะห์บทวิจารณ์แหล่งท่องเที่ยวจากเว็บไซต์ TripAdvisor และพบว่าโมเดล SVM แบบ RBF ให้ผลการทำนายดีที่สุดด้วยค่าความถูกต้องร้อยละ 83.90

ศราวุฒิ เกิดถาวร [3] ศึกษาความคิดเห็นต่อโรคโควิด-19 โดยเปรียบเทียบ 6 เทคนิคจำแนกข้อมูล พบว่าอัลกอริทึม MLP และ KNN ให้ประสิทธิภาพสูงสุดร้อยละ 99.20 เมื่อใช้ร่วมกับการคัดเลือกคุณลักษณะด้วยวิธี TF-IDF และ BM25

ภานุมาศ ทุกข์จาก [4] ศึกษาการลงทุนในสกุลเงินดิจิทัลผ่านความคิดเห็นจากวิดีโอออนไลน์ พบว่าแบบจำลอง BERT ให้ค่าความถูกต้องสูงสุดที่ร้อยละ 79 และการนำเทคนิค PCA มาช่วยลดมิติข้อมูลสามารถเพิ่มประสิทธิภาพของแบบจำลองได้ดีกว่าการใช้ข้อมูลดิบทั้งหมด

กมลทิพย์ ตะสะเวทย์ [5] ได้ศึกษารัศมีทัศนคติต่อกัญชาเพื่อการแพทย์บน Facebook, Pantip และ TikTok โดยเปรียบเทียบ 4 อัลกอริทึมหลัก ผลการวิจัยพบว่าอัลกอริทึมป่าสุ่ม (Random Forest) ให้ประสิทธิภาพดีที่สุดด้วยค่าความถูกต้องร้อยละ 84 รองลงมาคือ SVM และนาอีฟเบย์ ตามลำดับ ซึ่งแสดงให้เห็นว่าการเลือกใช้อัลกอริทึมที่เหมาะสมร่วมกับการแสดงผลผ่านระบบธุรกิจอัจฉริยะ (BI) มีส่วนช่วยสำคัญในการสนับสนุนการวิเคราะห์เชิงนโยบายและการตัดสินใจ

## 3. วิธีดำเนินงานวิจัย

ผู้วิจัยได้ทำการศึกษาและค้นคว้างานวิจัยที่เกี่ยวข้องและนำมาประยุกต์ใช้กับงานวิจัย โดยมีกระบวนการดังภาพที่ 1




**ภาพที่ 1** กรอบแนวความคิด

จากภาพที่ 1 กรอบแนวคิดของงานวิจัย การศึกษาวิจัยครั้งนี้ดำเนินการตามกรอบแนวคิดที่กำหนดไว้ โดยแบ่งขั้นตอนการทำงานออกเป็น 4 ส่วนหลัก ดังนี้

### 3.1 การเก็บรวบรวมข้อมูลและการจัดเตรียมข้อมูล

ผู้วิจัยดำเนินการรวบรวมข้อมูลความคิดเห็นภาษาไทยเกี่ยวกับเทคโนโลยีปัญญาประดิษฐ์จากแพลตฟอร์ม Pantip, Facebook, TikTok และ YouTube ระหว่างเดือนมกราคม 2566 ถึงสิงหาคม 2568 โดยใช้เครื่องมือ Instant Data Scraper สืบค้นผ่านคำสำคัญ (Keywords) เช่น “เอไอ”, “AI”, “Copilot”, “Gemini”, “ChatGPT” และวิดีโอที่เกี่ยวข้อง ข้อมูลที่รวบรวมได้รวมทั้งสิ้น 2,177 ข้อความ ประกอบด้วยความคิดเห็นเชิงบวก 1,497 ข้อความ และเชิงลบ 680 ข้อความ โดยกระบวนการจัดเก็บอยู่ภายใต้หลักจริยธรรมการวิจัยที่ไม่เปิดเผยข้อมูลส่วนบุคคลและใช้เพื่อวัตถุประสงค์ทางวิชาการเท่านั้น

จากนั้นข้อมูลจะเข้าสู่กระบวนการเตรียมข้อมูล (Data Pre-processing) เพื่อเพิ่มประสิทธิภาพในการวิเคราะห์ ประกอบด้วย การตัดคำ (Word Segmentation), การลบคำหยุด (Stop Words Removal), การกำจัดข้อความซ้ำ และการปรับรูปแบบคำให้เป็นมาตรฐาน (Text Normalization) เพื่อให้ข้อมูลพร้อมสำหรับนำไปสร้างโมเดลการเรียนรู้ของเครื่องในขั้นตอนถัดไป ดังแสดงกระบวนการในภาพที่ 2

| comment | clean_comment | Class |
|---|---|---|
| ChatGPT ไม่เก็บข้อมูลของผู้ใช้เพื่อนำไปพัฒนาโมเดลโดยอัตโนมัติ ข้อมูลการสนทนาของคุณจะไม่ถูกนำไปใช้เพื่อฝึกหรือปรับปรุงโมเดล เว้นแต่ว่าคุณจะอนุญาตโดยสมัครใจ | ข้อมูล,ผู้ใช้,เพื่อนำ,พัฒนา,โมเดล,โดยอัตโนมัติ,ข้อมูล,การสนทนา,ถูกนำไปใช้,ฝึก,ปรับปรุง,โมเดล,เว้นแต่ว่า,อนุญาต,โดยสมัครใจ | Positive |
| AI ที่สามารถเรียนรู้เรื่องจริยธรรม และสามารถยกระดับจิตใจของมนุษย์ให้สูงขึ้น ก็น่าจะสมควรค่าในการกระทำและเป็นไปได้ครับ | ที่สามารถ,เรียนรู้เรื่อง,จริยธรรม,ยกระดับจิตใจ,ของมนุษย์,สมควรค่า,การกระทำ,เป็นไปได้ | Positive |
| AI เป็นปัญญาประดิษฐ์ที่ฉลาดล้ำมากๆ ไม่ว่าจะเป็น AI Cover, AI Disney Pixar Posters และอื่นๆอีกมากมาย จนเป็นกระแสและโด่งดังมาก | ปัญญาประดิษฐ์,ฉลาด,ล้ำ,อีกมากมาย,กระแส,โด่งดัง | Positive |
| Chat GPT ก็มีคุณธรรมอยู่นะ | คุณธรรม | Positive |
| Claude เก่งภาษาไทยสุด | เก่ง,ภาษาไทย | Positive |
| Claude จาก Anthropic ตอนนี้ครองโลกไปแล้วครับ ตอนนี้ไม่มีใครสู้ไหวจริงๆด้านความฉลาด | ตอนนี้,ครองโลก,ตอนนี้,สู้,ไหว,ฉลาด | Positive |
| Gemini ฉลาดขึ้นมากจากจุดเริ่มต้น เพราะทรัพยากรในมือ google มหาศาล | ฉลาด,จุดเริ่มต้น,ทรัพยากร,มือ,มหาศาล | Positive |

**ภาพที่ 2** ตัวอย่างข้อมูลความคิดเห็นต่อเทคโนโลยี AI และผลการจัดเตรียมข้อมูล

### 3.2 การประเมินผลการจำแนกความคิดเห็นโดยผู้เชี่ยวชาญ

เพื่อยืนยันความถูกต้องของระบบ ผู้วิจัยได้แต่งตั้งผู้เชี่ยวชาญด้านภาษาไทยเพื่อตรวจสอบความถูกต้องในการจำแนกความคิดเห็นเชิงบวก (Positive) และเชิงลบ (Negative) โดยใช้วิธีการสุ่มข้อมูลร้อยละ 5 จากชุดข้อมูลทั้งหมดมาให้ผู้เชี่ยวชาญทำการกำหนดป้ายกำกับ (Label) เพื่อใช้เป็นค่าอ้างอิง (Ground Truth) จากนั้นนำผลลัพธ์ที่ได้จากระบบมาเปรียบเทียบเพื่อคำนวณหาค่าร้อยละความถูกต้องและค่าเฉลี่ยรวม ตามเกณฑ์การแปลผลระดับประสิทธิภาพที่กำหนดไว้

### 3.3 ทดสอบประสิทธิภาพและเลือกอัลกอริทึมที่เหมาะสม

กระบวนการคัดเลือกตัวแบบจำลองที่เหมาะสมที่สุดดำเนินการโดยการเปรียบเทียบประสิทธิภาพของอัลกอริทึมการเรียนรู้ของเครื่องจำนวน 8 รูปแบบ ได้แก่ Random Forest, Naïve Bayes, SVM, Deep Learning, Decision Tree, KNN, Neural Network และ Logistic Regression โดยใช้เทคนิค k-fold Cross Validation เพื่อลดความเอนเอียงของข้อมูล การวัดประสิทธิภาพพิจารณาจากตัวชี้วัดหลัก ได้แก่ Accuracy, Precision, Recall และ F-Measure เพื่อคัดเลือกอัลกอริทึมที่มีความแม่นยำสูงสุดในการวิเคราะห์ทัศนคติที่มีต่อเทคโนโลยี AI ตารางที่ 3

**ตารางที่ 1**: Confusion Matrix

| | Predicted | |
|---|---|---|
| Actual | Positive | Negative |
| Positive | True Positive | False Negative |
| Negative | False Positive | True Negative |

3.4 การแสดงผลข้อมูลในรูปแบบ Dashboard ผ่านโปรแกรม Power BI ในขั้นตอนสุดท้าย ข้อมูลที่ผ่านการประมวลผลและจำแนกเรียบร้อยแล้วจะถูกนำมาวิเคราะห์แนวโน้มความคิดเห็นและประเด็นสำคัญที่สังคมให้ความสนใจ โดยนำเสนอผ่านระบบธุรกิจอัจฉริยะ Power BI ในรูปแบบของ Dashboard ที่ประกอบด้วยแผนภูมิวงกลม แผนภูมิแท่ง และตารางสรุปผล เพื่อให้สามารถตีความข้อมูลทัศนคติของผู้ใช้งานโซเชียลมีเดียได้อย่างชัดเจนและเป็นรูปธรรม

## 4. ผลการดำเนินงานวิจัย

4.1 ผลการรวบรวมและเตรียมข้อมูล จากการรวบรวมข้อมูลผ่านแพลตฟอร์ม Facebook, YouTube, Pantip และ TikTok พบข้อมูลความคิดเห็นที่เกี่ยวข้องกับเทคโนโลยี AI จำนวนทั้งสิ้น 2,177 รายการ เมื่อผ่านกระบวนการ Data Pre-processing ได้แก่ การตัดคำ (Tokenization) การลบคำหยุด (Stop Words) และการกำจัดข้อความซ้ำ ผลลัพธ์ที่ได้คือชุดข้อมูลที่มีความพร้อมสำหรับการวิเคราะห์ โดยมีการกระจายตัวของข้อมูลแบ่งเป็นความคิดเห็นเชิงบวก (Positive) และเชิงลบ (Negative) เพื่อนำไปใช้เป็นชุดข้อมูลฝึกฝน (Training Set) และชุดข้อมูลทดสอบ (Testing Set) ต่อไป

4.2 ผลการประเมินความถูกต้องของการจำแนกความคิดเห็นโดยผู้เชี่ยวชาญ จากการสุ่มข้อมูลจำนวนร้อยละ 5 เพื่อให้ผู้เชี่ยวชาญด้านภาษาไทยทำการตรวจสอบความสอดคล้องระหว่างผลลัพธ์จากระบบกับการพิจารณาของมนุษย์ พบว่าระบบมีความสามารถในการจำแนกได้ถูกต้องสอดคล้องกับความเห็นของผู้เชี่ยวชาญในระดับที่น่าพอใจ โดยผลลัพธ์จากการคำนวณร้อยละความถูกต้องแสดงให้เห็นว่ากระบวนการเตรียมข้อมูลและการกำหนดเกณฑ์การ

จำแนกมีความน่าเชื่อถือ และสามารถนำไปใช้เป็นบรรทัดฐาน (Baseline) ในการประเมินประสิทธิภาพของอัลกอริทึมได้

**ตารางที่ 2**: ผลการเปรียบเทียบร้อยละความถูกต้องระหว่างระบบและผู้เชี่ยวชาญ

| รายการ | จำนวน | สอดคล้อง | ร้อยละ |
|---|---|---|---|
| เชิงบวก | 50 | 48 | 96 |
| เชิงลบ | 50 | 49 | 98 |
| รวมเฉลี่ย | 100 | 97 | 97 |

**4.3 ผลการเปรียบเทียบประสิทธิภาพของแบบจำลองและอัลกอริทึม**

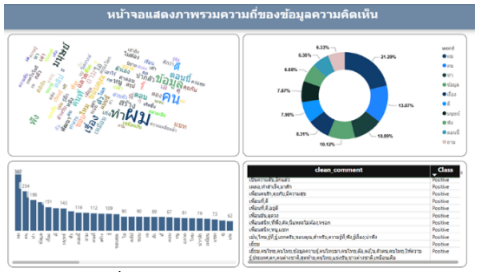
ผลการทดสอบเปรียบเทียบประสิทธิภาพของอัลกอริทึมทั้ง 8 รูปแบบ ผ่านเทคนิค k-fold Cross Validation แสดงให้เห็นว่าอัลกอริทึม K-Nearest Neighbors (KNN) มีประสิทธิภาพสูงสุดในการวิเคราะห์ความคิดเห็นต่อเทคโนโลยี AI โดยมีค่าความถูกต้อง (Accuracy) อยู่ที่ร้อยละ 82.09 ค่าความแม่นยำ (Precision) ร้อยละ 88.68 และค่าการระลึก (Recall) ร้อยละ 84.77 ส่งผลให้ค่า F-Measure อยู่ที่ร้อยละ 86.68 ซึ่งเมื่อพิจารณาจาก Confusion Matrix พบว่าระบบสามารถระบุความคิดเห็นเชิงบวกได้อย่างถูกต้องจำนวน 1,269 รายการ และเชิงลบจำนวน 518 รายการ ในขณะที่อัลกอริทึมอื่น ๆ ให้ประสิทธิภาพที่ลดหลั่นกันไปตามลักษณะของชุดข้อมูล

**ตารางที่ 3:** สรุปผลการเปรียบเทียบประสิทธิภาพของแบบจำลอง

| อัลกอริทึม | Acc. | Prec. | Rec. | F1 |
|---|---|---|---|---|
| KNN | 82.09 | 88.68 | 84.77 | 86.68 |
| Deep Learning | 81.58 | 86.01 | 87.44 | 86.72 |
| SVM | 81.12 | 94.36 | 77.15 | 84.89 |
| Neural Network | 81.12 | 94.73 | 76.82 | 84.84 |
| Random Forest | 80.89 | 95.08 | 76.15 | 84.57 |
| Decision Tree | 79.10 | 96.11 | 72.55 | 82.68 |
| Logis. Regression | 71.70 | 70.89 | 99.87 | 82.92 |
| Naïve Bayes | 71.43 | 97.30 | 60.12 | 74.32 |

**4.4 การนำเสนอข้อมูลสารสนเทศเชิงภาพผ่านโปรแกรม Power BI**

ผู้วิจัยได้นำผลการจำแนกที่ได้จากแบบจำลองที่มีประสิทธิภาพสูงสุด (KNN) มาสร้างการแสดงผลในรูปแบบสารสนเทศเชิงภาพ (Data Visualization) ผ่านโปรแกรม Power BI โดยนำเสนอในลักษณะ Dashboard ที่ประกอบด้วยสถิติจำนวนความคิดเห็นแยกตามประเภท แผนภูมิแสดงแนวโน้มความรู้สึกต่อประเด็นต่าง ๆ ของ AI เช่น ChatGPT หรือการใช้ AI ในตลาดแรงงาน ซึ่งช่วยให้การตีความข้อมูลทัศนคติของสังคมไทยมีความชัดเจนเข้าถึงง่าย และสามารถนำไปประยุกต์ใช้ในการวางแผนเชิงยุทธศาสตร์ได้อย่างมีประสิทธิภาพ




**ภาพที่ 3** หน้าจอแสดงภาพรวมความถี่ของคำ

## 5. สรุปผลการวิจัย

ผู้วิจัยได้ดำเนินการรวบรวมและวิเคราะห์ข้อมูลความคิดเห็นภาษาไทยจากแพลตฟอร์ม Facebook, YouTube, Pantip และ TikTok จำนวนทั้งสิ้น 2,177 รายการ โดยแบ่งเป็นความคิดเห็นเชิงบวก 1,497 รายการ และเชิงลบ 680 รายการ เพื่อนำมาสร้างตัวแบบการเรียนรู้ของเครื่องผ่านโปรแกรม RapidMiner และภาษา Python

ผลการประเมินความถูกต้องโดยผู้เชี่ยวชาญด้านภาษาไทยจากชุดข้อมูลสุ่มร้อยละ 5 พบว่าระบบมีความถูกต้องเฉลี่ยร้อยละ 97 ซึ่งอยู่ในระดับประสิทธิภาพดีมาก เมื่อทำการเปรียบเทียบประสิทธิภาพของอัลกอริทึมจำนวน 8 รูปแบบด้วยวิธี k-fold Cross Validation พบว่าอัลกอริทึมเพื่อนบ้านใกล้เคียงที่สุด (KNN) ให้ค่าความถูกต้อง (Accuracy) สูงที่สุดที่ร้อยละ 82.09 โดยมีค่า F-measure อยู่ที่ร้อยละ 86.68 รองลงมาคือ Deep Learning (ร้อยละ 81.58) และ SVM (ร้อยละ 81.12) ตามลำดับ

นอกจากนี้ ผู้วิจัยได้นำผลการวิเคราะห์มาสร้างการแสดงผลเชิงภาพ (Data Visualization) ผ่านโปรแกรม Power BI ซึ่งช่วยให้สามารถวิเคราะห์แนวโน้มทัศนคติของสังคมต่อเทคโนโลยี AI ได้อย่างมีประสิทธิภาพและชัดเจนผ่าน Dashboard รูปแบบต่าง ๆ

## 6. อภิปรายผลวิจัย

ผลการวิจัยพบว่าอัลกอริทึม KNN มีประสิทธิภาพสูงสุดในทุกมิติ โดยมีค่าความถูกต้อง (Accuracy) ที่ 82.09% สอดคล้องกับแนวทางของศราวุฒิ เกิดถาวร (2565) ที่ชี้ให้เห็นว่าเทคนิค KNN มีความโดดเด่นเมื่อใช้ร่วมกับการสกัดคุณลักษณะที่เหมาะสม ผู้วิจัยอภิปรายว่าความสำเร็จนี้เกิดจากการเตรียมข้อมูลภาษาไทยและการแปลงเป็นเวกเตอร์ที่แม่นยำ ทำให้ข้อความที่มีคำศัพท์เฉพาะทางด้าน AI สามารถเกาะกลุ่มและวัดระยะห่างได้ชัดเจน แม้ว่าในบริบทอื่น KNN อาจมีประสิทธิภาพรองจาก Random Forest แต่ในงานวิจัยนี้พิสูจน์ให้เห็นว่าลักษณะเฉพาะของชุดข้อมูลและการสกัดคุณลักษณะมีผลสำคัญต่อการเพิ่มขีดความสามารถของโมเดล KNN อย่างมีนัยสำคัญ

## เอกสารอ้างอิง

[1] วสวัตติ์ อินทร์แปลง (2563). “การวิเคราะห์ความคิดเห็นต่อเกมมือถือพับจี ด้วยเหมืองข้อความ” วิทยานิพนธ์ปริญญามหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ มหาวิทยาลัยมหาสารคาม.

[2] นิธิกร เลิศชาญวุฒิ (2564). “การวิเคราะห์ความสนใจของนักท่องเที่ยวด้วยเทคนิคการเรียนรู้ของเครื่อง” วิทยานิพนธ์ปริญญามหาบัณฑิต สาขาวิชาสถิติและการวิเคราะห์ธุรกิจ คณะวิทยาศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง.

[3] ศราวุฒิ เกิดถาวร (2565). “การคัดเลือกคุณลักษณะความรู้สึกของคนไทยต่อโรคโควิด 19 บนสื่อสังคมออนไลน์ด้วยเหมืองข้อความ” วิทยานิพนธ์ปริญญามหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ มหาวิทยาลัยมหาสารคาม.

[4] ภานุมาศ ทุกข์จาก (2565). “การขุดเหมืองข้อความและการวิเคราะห์ความรู้สึกสำหรับการลงทุนในสกุลเงินดิจิทัลโดยใช้ความคิดเห็นจากเว็บไซต์คลังคลิปวีดีโอ” การค้นคว้าอิสระปริญญามหาบัณฑิต สาขาวิชาวิทยาการข้อมูลและการวิเคราะห์ ศูนย์วิเคราะห์ข้อมูลดิจิทัลอัจฉริยะ คณะวิทยาศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง.

[5] กมลทิพย์ ตะสะเวทย์ (2566). “การจำแนกความคิดเห็นที่มีต่อกัญชาเพื่อการแพทย์บนโซเชียลมีเดียด้วยเทคนิคเหมืองข้อความ” การค้นคว้าอิสระปริญญามหาบัณฑิต สาขาวิชาระบบสารสนเทศเพื่อการจัดการ ภาควิชาการ

[6] “Artificial Intelligence (AI) คืออะไร” [ออนไลน์]. สืบค้นเมื่อ 01 มีนาคม 2569, จาก: https://www.dia.co.th/articles/what-is-artificial-intelligence/

[7] “Text Mining คืออะไร,” [ออนไลน์]. สืบค้นเมื่อ 01 มีนาคม 2569, จาก: https://text2 mining.blogspot.com/2 0 1 0 / 1 0 / text-mining.html

[8] “การวิเคราะห์ความรู้สึกคืออะไร,” [ออนไลน์]. สืบค้นเมื่อ 01 มีนาคม 2569, จาก: https://aws.amazon.com/th/what-is/sentiment-analysis/

[9] “Business Intelligence (BI) คืออะไร,” [ออนไลน์]. สืบค้นเมื่อ 01 มีนาคม 2569, จาก: https://aws.amazon.com/th/what-is/business-intelligence/

# การวิเคราะห์ปัจจัยที่มีผลต่อการรับรู้มาตรฐานความมั่นคงปลอดภัยทางไซเบอร์สำหรับโรงพยาบาล

# An Analysis of Factors Affecting the Perception of Cybersecurity Standards in Hospitals

*ผ่องศรี ประสมบูรณ์ (Pongsri Prasomboon)[1] และศักดิ์ชาย ตั้งวรรณวิทย์ (Sakchai Tangwannawit)[2]*

*[1]ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล*

*มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ*

*[1]s6707021857325@email.kmutnb.ac.th, [2]sakchai.t@itd.kmutnb.ac.th*

## บทคัดย่อ

*การวิจัยเรื่อง การวิเคราะห์ปัจจัยการรับรู้มาตรฐานความมั่นคงปลอดภัยทางไซเบอร์สำหรับโรงพยาบาล มีวัตถุประสงค์เพื่อศึกษาระดับการรับรู้ความพร้อมด้านความมั่นคงปลอดภัยทางไซเบอร์ และวิเคราะห์ปัจจัยที่มีอิทธิพลต่อการรับรู้ดังกล่าว การวิจัยใช้ระเบียบวิธีแบบผสมผสาน โดยเก็บข้อมูลเชิงปริมาณด้วยแบบสอบถามมาตราส่วน 5 ระดับ และวิเคราะห์ด้วยสถิติเชิงพรรณนา การวิเคราะห์องค์ประกอบ และการถดถอยเชิงเส้นพหุคูณ ผลการวิจัยพบว่า ระดับความพร้อมด้านความมั่นคงปลอดภัยทางไซเบอร์อยู่ในระดับมาก (ค่าเฉลี่ย = 3.85) โดยความตระหนักรู้ด้านภัยคุกคามมีค่าเฉลี่ยสูงที่สุด (ค่าเฉลี่ย = 4.12) ผลการทดสอบสมมติฐานพบว่า ความตระหนักรู้ด้านภัยคุกคาม (SA) และการสนับสนุนจากองค์กร (SP) มีอิทธิพลเชิงบวกต่อการรับรู้ประโยชน์ (PU) และความต้องการพัฒนาระบบ (Dev) อย่างมีนัยสำคัญทางสถิติ ขณะที่ความตระหนักรู้ด้านภัยคุกคามมีอิทธิพลเชิงลบต่อความสามารถในการควบคุม (CT) แต่การสนับสนุนจากองค์กรมีอิทธิพลเชิงบวกต่อความสามารถในการควบคุม*

## Abstract

*This study, entitled "An Analysis of Perceptual Factors of Cybersecurity Standards for Hospitals," aimed to (1) examine the level of perceived cybersecurity readiness and (2) analyze the factors influencing such perceptions in the hospital context. A quantitative research approach was employed. Data were collected using a five-point Likert scale questionnaire and analyzed using descriptive statistics, exploratory factor analysis (EFA), and multiple linear regression. The findings revealed that the overall level of cybersecurity readiness was high (mean = 3.85). Among the examined dimensions, security awareness (SA) demonstrated the highest mean score (mean = 4.12). Hypothesis testing indicated that security awareness (SA) and organizational support (SP) had significant positive effects on perceived usefulness (PU) and system development intention (Dev). However, SA showed a significant negative effect on controllability (CT), whereas SP had a significant positive effect on CT. The results suggest that enhancing cybersecurity standards in alignment with ISO/IEC 27001:2022 and HAIT Plus should emphasize not only improving security awareness but also strengthening organizational support and governance mechanisms to ensure sustainable cybersecurity development in hospitals.*

**Keywords**: Cybersecurity, Cybersecurity Readiness, Hospitals; Security Awareness; Organizational Support; Perceived Usefulness

## 1. บทนำ

ปัจจุบันภัยคุกคามทางไซเบอร์เป็นความเสี่ยงสำคัญที่ส่งผลกระทบต่อหน่วยงานทั้งภาครัฐและเอกชน โดยเฉพาะภาคสาธารณสุขซึ่งมีการจัดเก็บข้อมูลส่วนบุคคลที่มีความอ่อนไหว รายงานความเสี่ยงระดับโลกปี 2025 ชี้ให้เห็นว่าความเสี่ยงด้านข้อมูลบิดเบือนและการโจมตีทางไซเบอร์มีแนวโน้มเพิ่มสูงขึ้นอย่างต่อเนื่อง ขณะเดียวกัน สถิติการโจมตีทางไซเบอร์ต่อสถานพยาบาลทั่วโลกเพิ่มขึ้นอย่างมีนัยสำคัญ และในประเทศไทยพบว่าหน่วยงานด้านสาธารณสุขเป็นหนึ่งในกลุ่มที่ถูกโจมตีสูง โดยมีสาเหตุสำคัญจากข้อจำกัดด้านความรู้และความตระหนักรู้ของบุคลากร โรงพยาบาลเป็นหน่วยงานที่มีบทบาทสำคัญต่อชีวิตประชาชนและมีการใช้เทคโนโลยีดิจิทัลอย่างเข้มข้น ทั้งระบบสารสนเทศทางการแพทย์และอุปกรณ์อัจฉริยะ ซึ่งหากขาดมาตรการด้านความมั่นคงปลอดภัยที่เหมาะสม อาจก่อให้เกิดความเสียหายต่อข้อมูลและการให้บริการได้ แม้ว่าจะมีการกำหนดมาตรฐานและนโยบายด้านความมั่นคงปลอดภัย เช่น ISO/IEC 27001 และ HAIT Plus แต่ยังคงพบเหตุการณ์การโจมตีอย่างต่อเนื่อง

ดังนั้น งานวิจัยนี้จึงมุ่งศึกษาระดับความพร้อมด้านความมั่นคงปลอดภัยทางไซเบอร์ และวิเคราะห์ปัจจัยที่มีอิทธิพลต่อการรับรู้ดังกล่าวในบริบทของโรงพยาบาล เพื่อสนับสนุนการพัฒนามาตรการด้านความมั่นคงปลอดภัยให้มีประสิทธิภาพและยั่งยืน

## 2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

### 2.1 มาตรฐานความมั่นคงปลอดภัยทางไซเบอร์

มาตรฐาน ISO/IEC 27001:2022 เป็นมาตรฐานสากลด้านระบบบริหารจัดการความมั่นคงปลอดภัยสารสนเทศ (Information Security Management System: ISMS) ภายใต้ชื่อ Information security, cybersecurity and privacy protection - Information security management systems ประกาศใช้เมื่อวันที่ 10 ตุลาคม พ.ศ. 2565 โดยเป็นการปรับปรุงจากฉบับเดิมคือ ISO/IEC 27001:2013 เพื่อให้สอดคล้องกับบริบทภัยคุกคามไซเบอร์ที่เปลี่ยนแปลงอย่างรวดเร็ว เทคโนโลยีดิจิทัลที่มีความซับซ้อนมากขึ้น และข้อกำหนดด้านการคุ้มครองข้อมูลส่วนบุคคลในระดับสากล [1-4]

### 2.2 มาตรฐานความมั่นคงปลอดภัยทางไซเบอร์สำหรับโรงพยาบาล

แนวทางการดำเนินงานด้านการรักษาความมั่นคงปลอดภัยไซเบอร์สำหรับโรงพยาบาลของรัฐ (HAIT Plus) เป็นกรอบมาตรฐานเฉพาะบริบทของหน่วยบริการสาธารณสุขภาครัฐในประเทศไทย ที่พัฒนาขึ้นเพื่อบูรณาการข้อกำหนดตามกฎหมายไทยและมาตรฐานสากลด้านความมั่นคงปลอดภัยสารสนเทศเข้าด้วยกัน โดยมีเป้าหมายเพื่อยกระดับความปลอดภัยไซเบอร์ของโรงพยาบาลให้มีความมั่นคง ยั่งยืน และสอดคล้องกับความเสี่ยงเฉพาะด้านสุขภาพ [5-7]

### 2.3 ข้อกฎหมาย ระเบียบปฏิบัติด้านความมั่นคงปลอดภัยสารสนเทศที่เกี่ยวข้องกับโรงพยาบาล

การบริหารจัดการความมั่นคงปลอดภัยไซเบอร์ของโรงพยาบาลในประเทศไทยอยู่ภายใต้กรอบกฎหมายและระเบียบข้อบังคับหลายฉบับ โดยเฉพาะกฎหมายที่เกี่ยวข้องกับการคุ้มครองโครงสร้างพื้นฐานสำคัญทางสารสนเทศ (Critical Information Infrastructure: CII) การคุ้มครองข้อมูลส่วนบุคคล และการกำกับดูแลความมั่นคงปลอดภัยไซเบอร์ในระดับชาติ [5 - 7]

### 2.4 ทฤษฎีปัจจัยด้านการรับรู้และพฤติกรรมของบุคลากร

ทฤษฎีปัจจัยด้านการรับรู้และพฤติกรรมของบุคลากรมีรากฐานจาก Theory of Planned Behavior (TPB) อธิบายว่าพฤติกรรมของบุคคลเกิดจาก “ความตั้งใจ” (Behavioral Intention) ที่ได้รับอิทธิพลจากการรับรู้ ทัศนคติ และแรงกดดันทางสังคม แนวคิดนี้สามารถประยุกต์ใช้กับบริบทความมั่นคงปลอดภัยไซเบอร์ในโรงพยาบาลได้อย่างชัดเจน [8-12]

### 2.5 ทฤษฎีที่เกี่ยวข้องกับการยอมรับเทคโนโลยีสารสนเทศ

ทฤษฎีการยอมรับเทคโนโลยี (TAM) พัฒนาโดย Fred Davis (1989) เพื่ออธิบายเหตุผลที่บุคคลตัดสินใจยอมรับหรือปฏิเสธการใช้เทคโนโลยี โดยเน้นปัจจัยด้านการรับรู้ของผู้ใช้เป็นสำคัญ แบบจำลองนี้เสนอว่า “การรับรู้” จะ

ส่งผลต่อ “ทัศนคติ” และนำไปสู่ “ความตั้งใจ” และ “การใช้งานจริง” ตามลำดับ [8-12]

## 3. วิธีการดำเนินการวิจัย

### 3.1 กำหนดปัญหาและวัตถุประสงค์การวิจัย

ภัยคุกคามทางไซเบอร์เป็นความเสี่ยงระดับโลกที่ทวีความรุนแรงขึ้นอย่างต่อเนื่อง โดยรายงาน *The Global Risk Report 2025* ระบุว่า ความไม่มั่นคงทางไซเบอร์เป็นหนึ่งในความเสี่ยงสำคัญทั้งในระยะสั้นและระยะยาว โดยเฉพาะในภาคสาธารณสุขซึ่งมีการโจมตีเพิ่มขึ้นอย่างมีนัยสำคัญ แม้จะมีการกำหนดนโยบายและมาตรฐานด้านความมั่นคงปลอดภัย ได้แก่ ISO/IEC 27001:2022 และ HAIT Plus แต่ยังคงพบเหตุการณ์การโจมตีอย่างต่อเนื่อง สะท้อนถึงความจำเป็นในการยกระดับมาตรการด้านความมั่นคงปลอดภัยทั้งในมิติของบุคลากร กระบวนการ และเทคโนโลยี

ดังนั้น งานวิจัยนี้มีวัตถุประสงค์เพื่อศึกษาระดับความพร้อมด้านความมั่นคงปลอดภัยทางไซเบอร์ และวิเคราะห์ปัจจัยที่มีอิทธิพลต่อการรับรู้ดังกล่าวในบริบทของโรงพยาบาล เพื่อสนับสนุนการพัฒนามาตรการด้านความมั่นคงปลอดภัยให้มีประสิทธิภาพและยั่งยืน

### 3.2 ประชากรและกลุ่มตัวอย่าง

การวิจัยนี้ใช้ระเบียบวิธีแบบผสมผสาน (Mixed Methods) โดยการวิจัยเชิงปริมาณเก็บข้อมูลจากบุคลากรในโรงพยาบาล ครอบคลุมสายงานด้านการแพทย์ พยาบาล เทคโนโลยีสารสนเทศ และฝ่ายสนับสนุน สำหรับการวิเคราะห์แบบจำลองสมการโครงสร้าง และใช้วิธีการสุ่มแบบแบ่งชั้น (Stratified Sampling) เพื่อให้ได้กลุ่มตัวอย่างที่เป็นตัวแทนของแต่ละสายงาน โดยงานวิจัยนี้ใช้กลุ่มตัวอย่างจำนวน 250 คน

เครื่องมือที่ใช้ในการเก็บข้อมูลคือแบบสอบถามมาตราส่วนประมาณค่า 5 ระดับ ซึ่งผ่านการตรวจสอบความตรงเชิงเนื้อหาและความเชื่อมั่นด้วยค่าสัมประสิทธิ์ Cronbach’s Alpha สำหรับการวิจัยเชิงคุณภาพ ใช้วิธีการวิเคราะห์เอกสาร (Document Research) จากนโยบายและเอกสารที่เกี่ยวข้องด้านความมั่นคงปลอดภัยสารสนเทศ เพื่อสนับสนุนและอธิบายผลการวิจัยเชิงปริมาณ

### 3.3 เครื่องมือที่ใช้ในการวิจัย

การวิจัยครั้งนี้ใช้เครื่องมือในการเก็บรวบรวมข้อมูลทั้งเชิงปริมาณและเชิงคุณภาพ โดยพัฒนาให้สอดคล้องกับกรอบแนวคิดการวิจัย และผ่านกระบวนการตรวจสอบคุณภาพเครื่องมือตามหลักวิชาการ ดังรายละเอียดต่อไปนี้

3.3.1 การวิจัยเชิงปริมาณ (Quantitative Research) โดยใช้ระเบียบวิธีวิจัยเชิงสำรวจ (Survey) เครื่องมือในการวิจัย ได้แก่ แบบสอบถาม (Questionnaire) ทั้งแบบออไลน์และรูปแบบเอกสาร

3.3.2 การวิจัยเชิงคุณภาพ (Qualitative Research) ผู้วิจัยได้ทำการวิจัยด้วยวิธีการวิจัยเชิงเอกสาร (Documentary Research) โดยศึกษาจากเอกสารที่เกี่ยวข้องกับการบริหารจัดการความมั่นคงปลอดภัยสารสนเทศของโรงพยาบาล

### 3.4 วิธีดำเนินการวิจัยและการเก็บรวบรวมข้อมูล

การดำเนินการวิจัยครั้งนี้เป็นการวิจัยแบบผสมผสาน (Mixed Methods Research) โดยมุ่งศึกษาเชิงปริมาณควบคู่กับการวิเคราะห์เชิงคุณภาพ เพื่อให้ได้ผลการวิจัยที่มีความครอบคลุมและรอบด้าน ผู้วิจัยได้กำหนดขั้นตอนการดำเนินการวิจัยและการเก็บรวบรวมข้อมูลอย่างเป็นระบบ

### 3.5 วิธีการวิเคราะห์ข้อมูล

ภายหลังจากรวบรวมข้อมูลจากแบบสอบถามเรื่อง “ปัจจัยการรับรู้มาตรฐานความมั่นคงปลอดภัยทางไซเบอร์สำหรับโรงพยาบาล” เรียบร้อยแล้ว ผู้วิจัยจะดำเนินการประมวลผลและวิเคราะห์ข้อมูลด้วยโปรแกรมคอมพิวเตอร์สำเร็จรูป โดยแบ่งการวิเคราะห์ออกเป็น 2 ส่วนหลัก ได้แก่ สถิติเชิงพรรณนา และสถิติเชิงอนุมาน

## 4. ผลการดำเนินงาน

### 4.1 ผลการวิเคราะห์ข้อมูลทั่วไป

ในการวิเคราะห์ข้อมูลทั่วไปของผู้ตอบแบบสอบถาม ผู้วิจัยได้นำเสนอข้อมูลลักษณะทางประชากรศาสตร์ของกลุ่มตัวอย่าง ได้แก่ เพศ อายุ ระดับการศึกษา ฝ่าย และประสบการณ์การทำงาน โดยใช้สถิติเชิงพรรณนาในการ

วิเคราะห์ ได้แก่ การแจกแจงความถี่ (Frequency) และค่าร้อยละ (Percentage) นอกจากนี้ ผู้วิจัยได้จัดกลุ่มองค์กรตามลักษณะอุตสาหกรรมหลัก โดยรวมประเภทที่มีจำนวนผู้ตอบน้อยเข้ากลุ่ม Other Technology-related เพื่อความเหมาะสมในการนำเสนอข้อมูล ดังแสดงในตารางที่ 1

**ตารางที่ 1:** การวิเคราะห์ข้อมูลทั่วไปของกลุ่มตัวอย่าง

| ข้อมูลทั่วไป | จำนวน (คน) | ร้อยละ (%) |
|---|---|---|
| **เพศ (Gender)** | | |
| ชาย | 90 | 36.00 |
| หญิง | 133 | 53.20 |
| ไม่ระบุ / ทางเลือกอื่น | 27 | 10.80 |
| **รวม** | 250 | 100.00 |
| **อายุ (Age)** | | |
| 18 - 30 ปี | 45 | 18.00 |
| 31 - 40 ปี | 105 | 42.00 |
| 41 - 50 ปี | 62 | 24.80 |
| 51 - 60 ปี | 37 | 14.80 |
| 60 - 70 ปี | 1 | 0.40 |
| มากกว่า 70 ปี | 0 | 0.00 |
| **รวม** | 250 | 100.00 |
| **ระดับการศึกษา (Education)** | | |
| ปริญญาเอก | 6 | 2.40 |
| ปริญญาโท | 72 | 28.80 |
| ปริญญาตรี | 157 | 62.80 |
| มัธยมศึกษาตอนปลาย/อนุปริญญา | 13 | 5.20 |
| มัธยมศึกษาตอนต้น | 2 | 0.80 |
| ต่ำกว่ามัธยมศึกษา | 0 | 0.00 |
| **รวม** | 250 | 100.00 |
| สังกัดฝ่าย | | |
| ฝ่ายบริหาร | 27 | 10.80 |
| ฝ่ายการแพทย์/ห้องปฏิบัติการ | 54 | 21.60 |
| ฝ่ายการพยาบาล | 74 | 29.60 |
| ฝ่ายเทคโนโลยีสารสนเทศ | 34 | 13.60 |
| ฝ่ายการเงิน/ธุรการ | 2 | 0.80 |
| ฝ่ายเภสัชกรรม | 19 | 7.60 |
| ฝ่ายคุณภาพ | 20 | 8.00 |
| **รวม** | 250 | 100.00 |
| **ประสบการณ์ทำงาน/ความรู้ด้านความมั่นคงปลอดภัยทางไซเบอร์** | | |
| 1-3 ปี | 88 | 35.20 |
| 4-6 ปี | 96 | 38.40 |
| 7 ปีขึ้นไป | 42 | 16.80 |
| ไม่มีประสบการณ์ | 24 | 9.60 |
| **รวม** | 250 | 100.00 |

**4.2 ผลการวิเคราะห์ตัวแปรแต่ละด้าน**

ผู้วิจัยได้ทำการวิเคราะห์ระดับความคิดเห็นของกลุ่มตัวอย่างที่มีต่อปัจจัยตามกรอบแนวคิดการวิจัย โดยใช้สถิติเชิงพรรณนา ได้แก่ ค่าเฉลี่ย (Mean) ใช้ในการแสดงระดับความคิดเห็นโดยรวมของกลุ่มตัวอย่างต่อแต่ละปัจจัย ขณะที่ส่วนเบี่ยงเบนมาตรฐาน (S.D.) ใช้เพื่อแสดงระดับความแปรปรวนของความคิดเห็นหากค่าส่วนเบี่ยงเบนมาตรฐาน มีค่าน้อย แสดงว่าผู้ตอบมีความคิดเห็นไปในทิศทางเดียวกันค่อนข้างมากแต่หากค่าส่วนเบี่ยงเบนมาตรฐานมีค่าสูง แสดงว่าความคิดเห็นมีความแตกต่างกันมากขึ้น โดยผลการวิเคราะห์แสดงรายละเอียดดังตารางที่ 2

**ตารางที่ 2:** ผลสรุปปัจจัยด้านความพร้อมของเทคโนโลยีสารสนเทศ

| ข้อคำถาม | ค่าเฉลี่ย | S.D. | ระดับความคิดเห็น |
|---|---|---|---|
| นโยบายและระเบียบปฏิบัติ | 3.94 | 0.72 | สูง |
| โครงสร้างพื้นฐานเทคโนโลยี | 3.88 | 0.76 | สูง |
| ทรัพยากรและการสนับสนุน | 3.72 | 0.81 | สูง |
| รวม | 3.85 | 0.69 | สูง |

**ตารางที่ 3:** ผลสรุปปัจจัยด้านบุคลากร

| ข้อคำถาม | ค่าเฉลี่ย | S.D. | ระดับความคิดเห็น |
|---|---|---|---|
| สมรรถนะและความรู้ | 3.67 | 0.83 | สูง |
| ความตระหนักรู้ด้านภัยคุกคามทางไซเบอร์ | 4.12 | 0.64 | สูง |
| **รวม** | **3.90** | **0.68** | **สูง** |

**ตารางที่ 4:** ผลสรุปปัจจัยด้านการรับรู้มาตรฐานความมั่นคงปลอดภัยทางไซเบอร์

| ข้อคำถาม | ค่าเฉลี่ย | S.D. | ระดับความคิดเห็น |
|---|---|---|---|
| การรับรู้มาตรฐาน HAIT Plus | 3.95 | 0.70 | สูง |
| การรับรู้มาตฐาน ISO/IEC 27001 | 3.89 | 0.75 | สูง |
| **รวม** | **3.92** | **0.68** | **สูง** |

**ตารางที่ 5:** ผลสรุปปัจจัยด้านความพร้อมในการบริหารจัดการตามมาตรฐานความมั่นคงปลอดภัยทางไซเบอร์

| ข้อคำถาม | ค่าเฉลี่ย | S.D. | ระดับความคิดเห็น |
|---|---|---|---|
| ความสามารถในการนำมาตรฐานไปปฏิบัติ | 3.79 | 0.77 | สูง |
| ความพร้อมด้านระบบและกระบวนการ | 3.82 | 0.74 | สูง |
| ความมั่นใจในการดำเนินงานตามมาตรฐาน | 3.69 | 0.81 | สูง |
| **รวม** | **3.77** | **0.77** | **สูง** |

### 4.3 การวิเคราะห์องค์ประกอบเชิงสำรวจ

ผู้วิจัยได้ดำเนินการวิเคราะห์องค์ประกอบเชิงสำรวจ (Exploratory Factor Analysis: EFA) ซึ่งเป็นขั้นตอนการยืนยันเครื่องมือวัดมีความสอดคล้องกับโครงสร้างทฤษฎีและสามารถนำไปใช้ทดสอบสมมติฐานได้อย่างเหมาะสม

**ตารางที่ 6:** สรุปน้ำหนักองค์ประกอบและค่า Uniquenes จาก EFA ครั้งที่ 4

| มิติ/ปัจจัย | ข้อคำถามที่คงอยู่ | ช่วง Loading เด่น (โดยสรุป) |
|---|---|---|
| SA | SA(6), SA(8) | 0.644-0.751 |
| Sp | SP(10), SP(11), SP(12), SP(13) | 0.354-0.699 |
| PU | PU(7) | 0.767 |
| Ct | Ct(15) | 0.692 |
| Dev | Dev(16), Dev(17), Dev(18), Dev(19), Dev(20), Dev(22) | 0.634-0.979 |

### 4.4 ผลการทดสอบสมมติฐานด้วยการวิเคราะห์การถดถอยเชิงเส้น

การทดสอบสมมติฐานในบทนี้มุ่งเน้นตรวจสอบความสัมพันธ์เชิงเหตุและผลระหว่าง ตัวแปรต้น (Independent Variables) และ ตัวแปรตาม (Dependent Variables) ที่กำหนดในกรอบแนวคิดการวิจัย โดยอาศัยวิธีการ ถดถอยเชิงเส้นพหุคูณ (Multiple Linear Regression Analysis) ซึ่งเหมาะสมกับบริบทการวิจัยที่มีตัวแปรต้นมากกว่าหนึ่งตัวแปร และต้องการอธิบายอิทธิพลเชิงเส้นต่อผลลัพธ์ที่เป็นเชิงปริมาณ ในงานวิจัยนี้ ตัวแปรต้นที่ใช้ในการวิเคราะห์ ประกอบด้วย ความตระหนักรู้ด้านภัยคุกคาม (Security Awareness: SA) และ การสนับสนุนจากองค์กร (Support from Organization: Sp)

**ตารางที่ 7:** สรุปผลการทดสอบสมมติฐานโดยรวม

| สมมติฐาน | ความสัมพันธ์ที่ทดสอบ | ผลเชิงสถิติ (สรุป) | ข้อสรุป |
|---|---|---|---|
| H1 | SA, Sp → PU | - $R^2$ = 0.46<br>- SA และ Sp มีอิทธิพลเชิงบวก อย่างมีนัยสำคัญที่ .05 | ยอมรับสมมติฐาน |
| H2 | SA, Sp → Ct | - $R^2$ = 0.132<br>- Sp มีอิทธิพลเชิงบวกอย่างมีนัยสำคัญที่ .05<br>- SA มีอิทธิพลเชิงลบอย่างมีนัยสำคัญที่ .05 | ยอมรับสมมติฐานบางส่วน |
| H3 | SA, Sp → Dev | - $R^2$ = 0.296<br>- SA และ Sp มีอิทธิพลเชิงบวกอย่างมีนัยสำคัญที่ .05 | ยอมรับสมมติฐาน |

## 5. สรุปผล

การวิจัยเรื่อง “แนวทางพัฒนาการออกแบบเทคโนโลยีความมั่นคงปลอดภัยทางไซเบอร์สำหรับโรงพยาบาล เพื่อสอดคล้องกับมาตรฐาน HAIT Plus และ ISO/IEC 27001:2022” มีเป้าหมายสำคัญ 2 ประการ ได้แก่ (1) ศึกษาระดับการรับรู้มาตรฐานความมั่นคงปลอดภัยทางไซเบอร์ของผู้บริหารโรงพยาบาลหลายสายงาน และ (2) วิเคราะห์ปัจจัยที่มีอิทธิพลต่อการรับรู้มาตรฐานดังกล่าวตามแนวทาง HAIT Plus และ ISO/IEC 27001:2022

ผลสรุปการรับรู้มาตรฐานความมั่นคงปลอดภัยทางไซเบอร์ของผู้บริหารโรงพยาบาล ผลการวิเคราะห์ระดับการรับรู้โดยรวม พบว่า “การรับรู้ด้านความพร้อมความมั่นคงปลอดภัยเทคโนโลยีสารสนเทศ )IT Security Readiness) อยู่ในระดับมาก โดยมีค่าเฉลี่ยรวม 3.85 และส่วนเบี่ยงเบนมาตรฐาน 0.78 สะท้อนว่ากลุ่มตัวอย่างโดยภาพรวมมีทัศนะ/การรับรู้เชิงบวกต่อสภาพความพร้อมด้านความมั่นคงปลอดภัยในองค์กรโรงพยาบาล

ผลสรุปการวิเคราะห์ปัจจัย พบว่าความตระหนักรู้ด้านภัยคุกคาม (SA) และการสนับสนุนจากองค์กร (SP) มีอิทธิพลเชิงบวกต่อการรับรู้ประโยชน์ (PU) อย่างมีนัยสำคัญ โดยแบบจำลองสามารถอธิบายความแปรปรวนได้ร้อยละ 46.4 ($R^2 = 0.464$) นอกจากนี้ SA และ SP ยังมีอิทธิพลต่อความสามารถในการควบคุม (CT) โดย SP มีอิทธิพลเชิงบวก ขณะที่ SA มีอิทธิพลเชิงลบอย่างมีนัยสำคัญ ($R^2 = 0.132$) ในส่วนของความต้องการพัฒนาระบบ (Dev) พบว่าทั้ง SA และ SP มีอิทธิพลเชิงบวกอย่างมีนัยสำคัญทางสถิติ โดยแบบจำลองสามารถอธิบายความแปรปรวนได้ร้อยละ 29.6 ($R^2 = 0.296$) โดยสรุป การสนับสนุนจากองค์กรเป็นปัจจัยสำคัญที่ส่งผลเชิงบวกในทุกมิติ ขณะที่ความตระหนักรู้ด้านภัยคุกคามมีบทบาททั้งเชิงบวกและเชิงลบต่อการรับรู้ของบุคลากร

## เอกสารอ้างอิง


[1] International Organization for Standardization. (2022). ISO/IEC 27001:2022 Information security, cybersecurity and privacy protection — Information security management systems — Requirements. ISO.

[2] National Institute of Standards and Technology. (2018). Framework for improving critical infrastructure cybersecurity (Version 1.1). U.S. Department of Commerce.

[3] National Institute of Standards and Technology. (2020). Security and privacy controls for information systems and organizations (SP 800-53 Rev. 5). U.S. Department of Commerce.

[4] สำนักงานคณะกรรมการการรักษาความมั่นคงปลอดภัยไซเบอร์แห่งชาติ. (2564). กรอบการรักษาความมั่นคงปลอดภัยไซเบอร์แห่งชาติ. กรุงเทพฯ: ผู้แต่ง.

[5] สำนักงานพัฒนาธุรกรรมทางอิเล็กทรอนิกส์. (2563). แนวปฏิบัติด้านความมั่นคงปลอดภัยสารสนเทศสำหรับหน่วยงานภาครัฐ. กรุงเทพฯ: ผู้แต่ง.

[6] สถาบันรับรองคุณภาพสถานพยาบาล. (2565). มาตรฐานโรงพยาบาลและบริการสุขภาพ ฉบับ HAIT Plus. กรุงเทพฯ: ผู้แต่ง.

[7] Tornatzky, L. G., & Fleischer, M.. (1990). The processes of technological innovation. Lexington Books.

[8] DiMaggio, P. J., & Powell, W. W.. (1983). The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. American Sociological Review, 48(2), 147–160.

[9] Herath, T., & Rao, H. R.. (2009). Encouraging information security behaviors in organizations: Role of penalties, pressures and perceived effectiveness. Decision Support Systems, 47(2), 154–165.

[10] Ifinedo, P.. (2012). Understanding information systems security policy compliance: An integration of the theory of planned behavior and the protection motivation theory. Computers & Security, 31(1), 83–95.

[11] Bulgurcu, B., Cavusoglu, H., & Benbasat, I.. (2010). Information security policy compliance: An empirical study of rationality-based beliefs and information security awareness. MIS Quarterly, 34(3), 523–548.

[12] Siponen, M., Pahnila, S., & Mahmood, M. A.. (2010). Compliance with information security policies: An empirical investigation. Computer, 43(2), 64–71.

# การพยากรณ์การใช้งานทรัพยากรระบบ SAP โดยใช้เหมืองข้อมูล
# Forecasting Resource SAP System Utilization Using Data Mining

*เฉลิมเกียรติ ทองวิลัย (Chalermkiet Tongwilai)* [1] *และ มณเฑียร รัตนศิริวงศ์วุฒิ (Montean Rattanasiriwongwut)* [2]
ภาควิชาการจัดการเทคโนโลยีสารสนเทศ
คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
[1]*s6707021857040@email.kmutnb.ac.th*, [2]*montean.r@itd.kmutnb.ac.th*

## บทคัดย่อ

*การค้นคว้าอิสระนี้มีวัตถุประสงค์เพื่อคัดเลือกแบบจำลองสำหรับพยากรณ์การใช้งานทรัพยากรระบบ SAP ทั้งหน่วยประมวลผลกลาง (CPU) และหน่วยความจำหลัก (Memory) เพื่อนำไปพัฒนาแดชบอร์ดเชิงธุรกิจสำหรับสนับสนุนการตัดสินใจของผู้บริหาร โดยรวบรวมข้อมูลการใช้งานของบริษัท พีทีจี เอ็นเนอยี จำกัด (มหาชน) ช่วงเดือน พ.ย. 2567 ถึง ต.ค. 2568 ผ่านเครื่องมือ PRTG ผู้วิจัยประยุกต์ใช้อัลกอริทึม LSTM, ARIMA และ SARIMA ในการวิเคราะห์ข้อมูลอนุกรมเวลา ผลการทดสอบพบว่าแบบจำลอง SARIMA มีความแม่นยำสูงสุดในทรัพยากรทั้ง 2 ประเภท โดยพยากรณ์ CPU ได้ค่า RMSE ต่ำที่สุดที่ 1.549 ($R^2$ = 0.830) และพยากรณ์ Memory ได้ค่า RMSE ที่ 1.531 ($R^2$= 0.834) ผลลัพธ์ดังกล่าวได้ถูกนำไปสร้างแดชบอร์ดเพื่อช่วยวางแผนการจัดสรรทรัพยากรและลดปัญหาคอขวดของระบบได้อย่างมีประสิทธิภาพ*

**คำสำคัญ:** การพยากรณ์ทรัพยากรระบบ SAP, การเรียนรู้ของเครื่อง, LSTM, ARIMA, SARIMA, แดชบอร์ด

## Abstract

*This independent study aims to select a forecasting model for SAP system resource utilization, specifically the Central Processing Unit (CPU) and Memory, to develop a Business Intelligence Dashboard supporting executive decision-making. Usage data from PTG Energy Public Company Limited, spanning November 2024 to October 2025, were collected via PRTG. The researcher applied LSTM, ARIMA, and SARIMA algorithms for time-series analysis. The results indicate that the SARIMA model achieved the highest accuracy for both resource types. For CPU forecasting, SARIMA yielded the lowest RMSE of 1.549 ($R^2$ = 0.830), and for Memory forecasting, it achieved an RMSE of 1.530 ($R^2$ = 0.834). These findings were utilized to create a dashboard to effectively aid resource allocation planning and mitigate system bottlenecks.*

**Keywords**: *SAP Resource Forecasting, Machine Learning, LSTM, ARIMA, SARIMA, Business Intelligence Dashboard.*

## 1. บทนำ

บริษัท พีทีจี เอ็นเนอยี จำกัด (มหาชน) มีการใช้งานระบบ SAP (Systems, Applications, and Products) ซึ่งเป็นระบบ ERP หลักที่รวบรวมข้อมูลจากทุกหน่วยงานไว้ในฐานข้อมูลกลาง อย่างไรก็ตาม เมื่อปริมาณข้อมูลและธุรกรรมเพิ่มสูงขึ้น องค์กรต้องเผชิญกับปัญหาด้านประสิทธิภาพ (Performance Issues) ซึ่งมีสาเหตุหลักมาจากการใช้งานทรัพยากรระบบ ทั้งหน่วยประมวลผลกลาง (CPU) ที่ทำงานหนักจนเกิดภาวะคอขวด และหน่วยความจำหลัก (Memory) ที่อาจมีการจัดสรรไม่เพียงพอจนระบบตอบสนองช้าหรือหยุดทำงาน (Downtime)

ผู้วิจัยจึงมุ่งเน้นพัฒนาแบบจำลองการพยากรณ์การใช้ทรัพยากรระบบ SAP โดยประยุกต์ใช้เทคนิคเหมืองข้อมูล (Data Mining) เพื่อคาดการณ์ความต้องการใช้งาน CPU และ Memory ล่วงหน้า พร้อมทั้งนำผลลัพธ์มาพัฒนาเป็นแดชบอร์ดรายงานผลเชิงธุรกิจ เพื่อสนับสนุนการตัดสินใจเชิงรุกในการวางแผนเพิ่มขยายขีดความสามารถ (Capacity Planning) ขององค์กร

## 2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ผู้วิจัยได้ศึกษาค้นคว้าเอกสาร ทฤษฎี และงานวิจัยที่เกี่ยวข้องต่าง ๆ โดยศึกษาข้อมูลต่าง ๆ ดังต่อไปนี้

### 2.1 แนวคิดและทฤษฎีเกี่ยวกับระบบ SAP และการจัดการทรัพยากร

ระบบ SAP ERP เป็นซอฟต์แวร์ที่ใช้ในการบูรณาการกระบวนการทางธุรกิจขององค์กรเข้าสู่ฐานข้อมูลกลาง ทำให้ข้อมูลมีความถูกต้องแม่นยำและทำงานแบบเรียลไทม์ องค์กรขนาดใหญ่มักพบปัญหาจากภาระงานที่เพิ่มขึ้น ส่งผลให้การใช้งาน CPU และ Memory ทำให้เกิดภาวะคอขวด (System Bottleneck) ระบบตอบสนองช้า หรืออาจหยุดทำงาน (Downtime)

### 2.2 แนวคิดและทฤษฎีเกี่ยวกับเหมืองข้อมูลและการพยากรณ์

1. กระบวนการเหมืองข้อมูลประกอบด้วยการรวบรวมข้อมูล การเตรียมข้อมูล (จัดการค่าว่าง ปรับเสกล) การสร้างแบบจำลอง และการประเมินผลประสิทธิภาพ

2. การวิเคราะห์ข้อมูลอนุกรมเวลา (Time Series Analysis) เน้นวิเคราะห์ข้อมูล CPU และ Memory ที่บันทึกตามลำดับเวลา โดยพิจารณาแนวโน้มและรูปแบบฤดูกาล

### 2.3 อัลกอริทึมที่ใช้ในการพยากรณ์

1. ARIMA เป็นแบบจำลองเชิงสถิติที่อาศัยความสัมพันธ์ของข้อมูลในอดีต (p), การทำผลต่างเพื่อปรับให้ข้อมูลคงที่ (d), และค่าเฉลี่ยเคลื่อนที่ (q) เหมาะกับข้อมูลที่ไม่มีฤดูกาลชัดเจน [1]

2. SARIMA เป็นการขยายความสามารถของ ARIMA ให้รองรับข้อมูลที่มีรอบฤดูกาล โดยเพิ่มพารามิเตอร์ส่วนฤดูกาล (P, D, Q, s) ซึ่งเหมาะกับพฤติกรรมระบบ SAP ที่มีการใช้งานซ้ำเป็นรอบสัปดาห์หรือรอบปิดงบการเงิน [2]

3. Long Short-Term Memory (LSTM) เป็นแบบจำลองโครงข่ายประสาทเทียมเชิงลึก (Deep Learning) ที่สามารถจดจำความสัมพันธ์ระยะยาวและจัดการกับข้อมูลที่ไม่เป็นเส้นตรงได้ดีเยี่ยม โดยมีกลไกควบคุม 3 ส่วน คือ Forget Gate Input Gate และ Output Gate เพื่อคัดกรองเฉพาะข้อมูลที่สำคัญ [3]

### 2.4 การวัดประสิทธิภาพ

1. MAE (Mean Absolute Error) ใช้วัดค่าเฉลี่ยความคลาดเคลื่อนสัมบูรณ์ระหว่างค่าที่พยากรณ์ได้กับค่าจริง ค่า MAE ยิ่งต่ำแสดงว่าแบบจำลองยิ่งมีความแม่นยำสูง แสดงดังสมการที่ [2]

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|$$

2. RMSE (Root Mean Squared Error) ใช้วัดระดับความผิดพลาดของการพยากรณ์ โดยให้ความสำคัญกับความคลาดเคลื่อนขนาดใหญ่ ค่า RMSE ยิ่งต่ำแสดงว่าแบบจำลองยิ่งมีประสิทธิภาพ แสดงดังสมการที่ [3]

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$$

3. MAPE (Mean Absolute Percentage Error) ใช้วัดความคลาดเคลื่อนในรูปแบบร้อยละ ช่วยให้เปรียบเทียบข้อมูลต่างมาตราส่วนได้ง่าย โดยผู้วิจัยตั้งสมมติฐานให้มีความคลาดเคลื่อนน้อยกว่าร้อยละ 20 แสดงดังสมการที่ [3]

$$MAPE = \frac{100}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right|$$

4. $R^2$ (Coefficient of Determination) ใช้วัดความสามารถในการอธิบายความแปรปรวนของข้อมูล ค่า $R^2$ ยิ่งเข้าใกล้ 1 แสดงว่าแบบจำลองยิ่งพยากรณ์ได้แม่นยำและเหมาะสม แสดงดังสมการที่ [3]

$$R^2 = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y})^2}$$

### 2.5 งานวิจัยที่เกี่ยวข้อง

Qinyun, etc. [1] วิเคราะห์ปรากฏการณ์แส้ม้าในนโยบาย POUT ภายใต้ความต้องการไม่คงที่ (ARIMA(1,1,2)) โดยใช้ผลต่างความแปรปรวนเป็นตัวชี้วัด พบว่าอาจสร้างความผันผวนรุนแรงกว่านโยบาย OUT ซึ่งขัดแย้งกับทฤษฎีเดิม โดยมีการยืนยันผลด้วยชุดข้อมูล M4 Competition

S. Huang and etc. [2] นำเสนอแบบจำลองผสม SARIMA-WOA-LSTM เพื่อพยากรณ์ปริมาณน้ำท่าในพื้นที่ราบ โดยผสานจุดเด่นในการจัดการข้อมูลทั้งแบบเชิงเส้น ฤดูกาล และไม่เชิงเส้นเข้าด้วยกัน ผลการศึกษาพบว่าแบบจำลองมีความแม่นยำสูงมาก (MAE = 1.29, RMSE = 2.14 และ $R^2$ = 0.986)

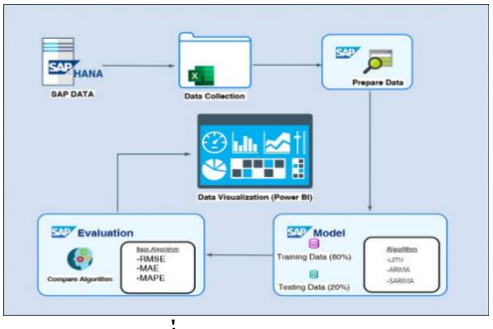
Falah and Mesut [3] งานวิจัยนี้เสนอแบบจำลองผสม LSTM-XGBoost เพื่อแก้ปัญหาความซับซ้อนของข้อมูล

และเพิ่มความแม่นยำในการทำนายความต้องการไฟฟ้า โดยผลลัพธ์มีความคลาดเคลื่อน (MAPE) เพียง 1.18% ซึ่งสามารถนำไปประยุกต์ใช้จัดการ Smart Grid ได้อย่างมีประสิทธิภาพ

Yong Wang [4] เปรียบเทียบแบบจำลองดั้งเดิมกับ LSTM, RF และ SVR เพื่อพยากรณ์ผลตอบแทนตลาดหุ้นไทย โดยพบว่าโมเดล RF มีประสิทธิภาพดีที่สุด รองลงมาคือ LSTM

J. Blanco and etc. [5] ทบทวนแบบจำลองพยากรณ์ผู้ป่วยฉุกเฉิน พบว่าเทคนิค ML/DL (เช่น XGBoost RF LSTM,CNN) มีความแม่นยำสูงกว่าโมเดลดั้งเดิม โดยมีค่าความคลาดเคลื่อน MAPE อยู่ที่ 3% - 18% และการใช้ตัวแปรภายนอกช่วยเพิ่มประสิทธิภาพได้ แม้ยังมีข้อจำกัดด้านการตีความและการนำไปใช้จริง

## 3. วิธีการดำเนินการวิจัย




**ภาพที่ 1** : กรอบแนวทางการดำเนินงานวิจัย

จากภาพที่ 1 การศึกษานี้เป็นงานวิจัยเชิงประยุกต์ที่บูรณาการกระบวนการวิทยาการ เพื่อแปลงข้อมูลประวัติการใช้งานระบบ SAP ERP ให้เป็นแบบจำลองพยากรณ์ทรัพยากร (CPU และ Memory) ที่มีความแม่นยำสูง โดยมุ่งนำผลลัพธ์ไปประยุกต์ใช้สนับสนุนการตัดสินใจของผู้บริหารอย่างมีประสิทธิภาพ ซึ่งแบ่งการดำเนินงานหลักออกเป็น 4 ขั้นตอน ดังนี้

```python
def clean_resource_data(df, target_col):
    """
    ทำความสะอาดข้อมูล SAP Resource (CPU/Memory)
    """
    # 1. จัดการ Index และความถี่ (Resampling)
    # ป้องกันปัญหาวันที่ขาดหายไปซึ่งทำให้ R2 ต่ำ
    df = df.sort_index()
    df = df.resample('D').mean()

    # 2. จัดการค่าที่ผิดปกติ (System Down/Reset)
    median_val = df[target_col].median()
    df.loc[df[target_col] < (median_val * 0.3), target_col] = np.nan

    # 3. เติมค่าว่างด้วย Interpolation (ดีกว่า ffill ในเชิงสถิติ)
    df[target_col] = df[target_col].interpolate(method='time')
    df[target_col] = df[target_col].ffill().bfill()

    return df
```

**ภาพที่ 2** : ภาพแสดงการจัดเตรียมข้อมูล

### 3.1 การรวบรวมและจัดเตรียมข้อมูล

เก็บข้อมูลประวัติการใช้งานจริงของทรัพยากร CPU และ Memory จากเซิร์ฟเวอร์ SAP ผ่านโปรแกรม PRTG Network Monitoring ระหว่างวันที่ 1 พ.ย. 2567 ถึง 31 ต.ค. 2568 (รวม 365 วัน) จำนวน 8,760 ข้อมูล จาก 4 โมดูล ได้แก่ FICO MM SD IS-OIL มีจำนวนผู้ใช้งาน 300 ราย จากนั้นดำเนินการจัดการค่าว่าง ปรับสเกลข้อมูล ดังภาพที่ 2 และแบ่งข้อมูลเป็นชุดฝึกสอน (Training Set) ร้อยละ 80 และชุดทดสอบ (Testing Set) ร้อยละ 20

### 3.2 การสร้างแบบจำลองและการปรับพารามิเตอร์

```python
# ARIMA Grid Search
'arima_p_range': range(0, 4),
'arima_d_range': range(0, 3),
'arima_q_range': range(0, 4),
```

**ภาพที่ 3** : ภาพแสดงการกำหนดพารามิเตอร์ของ ARIMA

1. ARIMA กำหนดช่วงค่า (p, d, q) เพื่อทดสอบด้วยวิธี Grid Search ในการหาสมดุลความแม่นยำ ดังภาพที่ 3

```python
# SARIMA Grid Search
'sarima_seasonal_period': 7,  # Weekly seasonality
'sarima_P_range': range(0, 3),
'sarima_D_range': range(0, 2),
'sarima_Q_range': range(0, 3),
```

**ภาพที่ 4** : ภาพแสดงการกำหนดพารามิเตอร์ของ SARIMA

2. SARIMA กำหนดช่วงค่าพารามิเตอร์ทั้งส่วนที่ไม่เป็นฤดูกาลและส่วนฤดูกาล โดยกำหนดคาบฤดูกาลเท่ากับ 7 เพื่อรองรับพฤติกรรมรายสัปดาห์และหาความแม่นยำสูงสุด ดังภาพที่ 4

```python
# LSTM Optuna
'optuna_trials': 20,
'lstm_lookback_range': [7, 14, 21, 30, 60],
'lstm_units_range': [32, 64, 128, 256],
'lstm_layers_range': [1, 2, 3],
'lstm_dropout_range': [0.1, 0.2, 0.3, 0.4],
'lstm_epochs': 100,
'lstm_batch_sizes': [16, 32, 64],
```

**ภาพที่ 5** : ภาพแสดงการกำหนดพารามิเตอร์ของ LSTM

3. LSTM กำหนดช่วงค่าและปรับจูนพารามิเตอร์แบบอัตโนมัติด้วย Optuna (Layers , Lookback, Dropout) ให้มีความแม่นยำ ดังภาพที่ 5

### 3.3 การพัฒนาแดชบอร์ดรายงานผลเชิงธุรกิจ

ผู้วิจัยนำผลลัพธ์ไปพัฒนาแดชบอร์ดรายงานผลเชิงธุรกิจด้วยโปรแกรม Power BI เพื่อสนับสนุนการตัดสินใจของผู้บริหาร โดยแบ่งการแสดงผลเป็น 4 ส่วนหลัก ได้แก่ (1) ส่วนแสดงสถานะปัจจุบัน (2) ส่วนเปรียบเทียบขณะสร้างแบบจำลอง (3) ส่วนของการทำนายใช้แบบจำลอง SARIMA พยากรณ์ล่วงหน้า 1 ปี (ถึง เม.ย. 2570) และ (4) ส่วนสรุปภาพรวมแบบข้อมูลตั้งแต่อดีตถึงอนาคต เพื่อสนับสนุนการวางแผนขยายทรัพยากรได้อย่างทันท่วงที

## 4. ผลการดำเนินงานวิจัย

### 4.1 ข้อมูลที่ใช้ในการศึกษาและการเตรียมข้อมูล

| Date | PercentUsedMemory |
|---|---|
| 2024-11-01 | 65.993428 |
| 2024-11-02 | 68.660101 |
| 2024-11-03 | 71.224962 |
| 2024-11-04 | 70.297896 |
| 2024-11-05 | 62.472165 |
| ... | ... |
| 2025-10-27 | 78.098222 |
| 2025-10-28 | 75.813642 |
| 2025-10-29 | 69.852895 |
| 2025-10-30 | 71.866794 |
| 2025-10-31 | 76.380288 |

365 rows × 2 columns

**ภาพที่ 6 :** แสดงผลหลังการเตรียมข้อมูล

จากภาพที่ 6 แสดงผลหลังจากการเตรียมข้อมูล ผลปรากฏว่ามีข้อมูลครบทั้ง 365 วัน ไม่มีข้อมูลที่มีค่าว่าง

### 4.2 ผลการกำหนดค่าพารามิเตอร์ที่เหมาะสมที่สุด

**ตารางที่** 1 : ตารางแสดงค่าพารามิเตอร์ที่ดีที่สุดของ Arima

| รายการ | p | d | q | RMSE |
|---|---|---|---|---|
| Memory | 2 | 1 | 3 | 2.4009 |
| CPU | 2 | 1 | 3 | 2.4133 |

ค่าพารามิเตอร์ที่ดีที่สุดของแบบจำลอง ARIMA จากการทดสอบด้วยวิธี Grid Search แสดงดังตารางที่ 1

**ตารางที่ 2** : ตารางแสดงค่าพารามิเตอร์ที่ดีที่สุดของ Sarima

| รายการ | Parameter | | | | | | | Mean RMSE |
|---|---|---|---|---|---|---|---|---|
| | Order | | | Seasonal | | | | |
| | p | d | q | P | D | Q | S | |
| Memory | 2 | 0 | 0 | 1 | 0 | 1 | 7 | 2.1332 |
| CPU | 1 | 0 | 0 | 1 | 0 | 1 | 7 | 2.1697 |

ค่าพารามิเตอร์ที่ดีที่สุดของแบบจำลอง SARIMA จากการทดสอบด้วยวิธี Grid Search แสดงดังตารางที่ 2

**ตารางที่** 3 :ตารางแสดงค่าพารามิเตอร์ที่ดีที่สุดของ LSTM

| รายการ | Parameter | | | | | | RMSE |
|---|---|---|---|---|---|---|---|
| | Batch size | dropout | Learning rate | Look back | layers | units | |
| Memory | 16 | 0.3 | 0.007 | 7 | 1 | 256 | 1.7448 |
| CPU | 16 | 0.4 | 0.005 | 7 | 3 | 256 | 1.7226 |

ค่าพารามิเตอร์ที่ดีที่สุดของแบบจำลอง LSTM จากการทดสอบด้วยวิธี Grid Search แสดงดังตารางที่ 3

### 4.3 ผลการเปรียบเทียบประสิทธิภาพแบบจำลอง

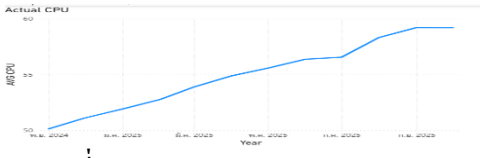
**ตารางที่** 4 **:** ตารางแสดงประสิทธิภาพแบบจำลอง

| ตัวแปร | แบบจำลอง | MSE | RMSE | MAE | MAPE | $R^2$ |
|---|---|---|---|---|---|---|
| CPU | ARIMA | 5.045 | 2.246 | 1.868 | 0.031 | 0.643 |
| | SARIMA | 2.400 | 1.549 | 1.283 | 0.022 | 0.830 |
| | LSTM | 2.921 | 1.709 | 1.388 | 0.023 | 0.792 |
| Memory | ARIMA | 5.063 | 2.250 | 1.871 | 0.025 | 0.641 |
| | SARIMA | 2.343 | 1.531 | 1.271 | 0.017 | 0.834 |
| | LSTM | 3.143 | 1.730 | 1.416 | 0.019 | 0.769 |

จากตารางที่ 4 ผลการเปรียบเทียบพบว่าแบบจำลอง SARIMA มีประสิทธิภาพสูงสุด ทั้งในการพยากรณ์ CPU และ Memory โดยให้ค่าความคลาดเคลื่อน (MSE, RMSE, MAE, MAPE) ต่ำที่สุดทุกตัวชี้วัด และมีค่าสัมประสิทธิ์การตัดสินใจ ($R^2$) สูงที่สุด (0.830 และ 0.834 ตามลำดับ) รองลงมาคือแบบจำลอง LSTM และ ARIMA

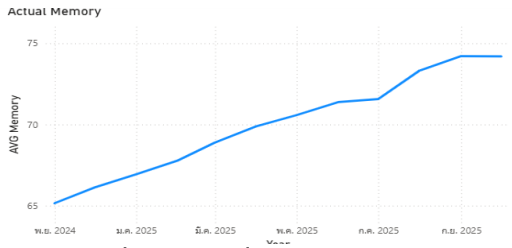
### 4.4 ผลการแสดงผล

1.การแสดงผลการใช้ข้อมูลจริงของ CPU และ Memory



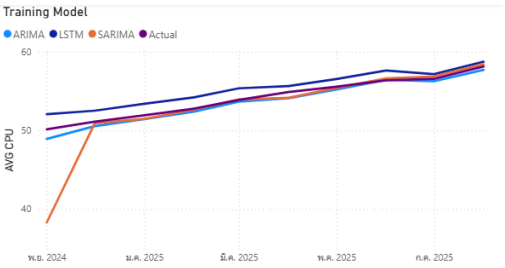

**ภาพที่ 7** : ภาพแสดงการใช้ข้อมูลของ CPU

**ภาพที่ 8** : ภาพแสดงการใช้ข้อมูลของ Memory

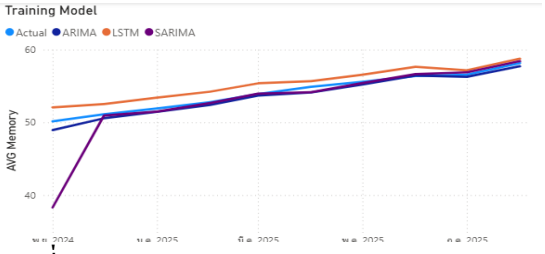



จากภาพที่ 7 และภาพที่ 8 แสดงข้อมูลการใช้งานจริงของหน่วยประมวลผลกลาง (CPU) และหน่วยความจำหลัก (Memory) ตั้งแต่เดือนพฤศจิกายน พ.ศ. 2567 ถึงตุลาคม พ.ศ. 2568 โดยพบว่าทรัพยากรทั้งสองส่วนมีแนวโน้มเพิ่มขึ้นอย่างต่อเนื่อง จากระดับร้อยละ 50 ไปแตะที่ระดับเกือบร้อยละ 60 ซึ่งสะท้อนให้เห็นถึงภาระงานและปริมาณธุรกรรมในระบบ SAP ที่เพิ่มสูงขึ้นตามการขยายตัวขององค์กรอย่างชัดเจน

2.ผลการ Training อัลกอริทึมของ CPU และ Memory



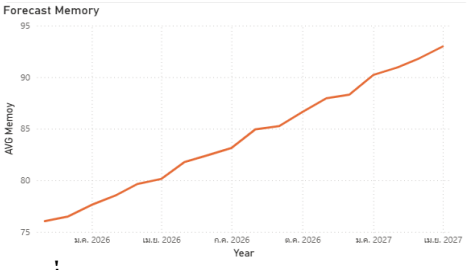

**ภาพที่ 9** : ภาพแสดงข้อมูลของ CPU ขณะ Training อัลกอริทึม



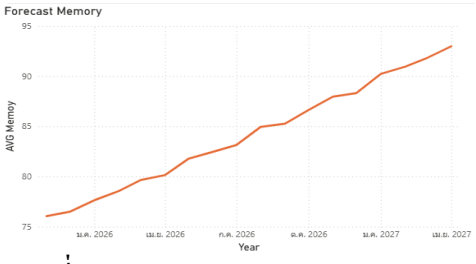

**ภาพที่ 10 :** ภาพแสดงข้อมูลของ Memory ขณะ Training อัลกอริทึม

จากภาพที่ 9 และภาพที่ 10 แสดงการเปรียบเทียบผลลัพธ์ขณะสร้างแบบจำลอง (Training Model Comparison) ของการใช้งาน CPU และ Memory ตามลำดับ โดยนำเสนอกราฟเปรียบเทียบระหว่างค่าการใช้งานจริง (Actual) กับค่าพยากรณ์ (Predicted) จากอัลกอริทึม ARIMA SARIMA และ LSTM ในช่วงเวลาเดียวกัน เพื่อประเมินประสิทธิภาพและความแม่นยำของแต่ละแบบจำลอง

3.ผลการพยากรณ์อัลกอริทึมของ CPU และ Memory



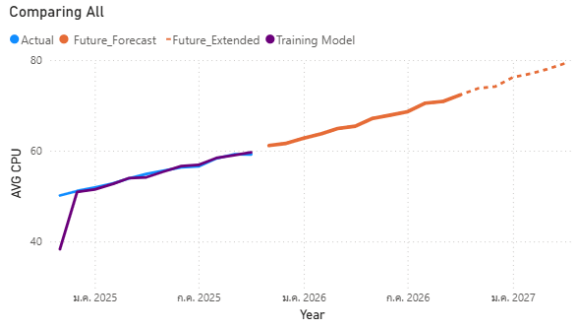

**ภาพที่ 11 :** ภาพแสดงการพยากรณ์ในอนาคตของ CPU



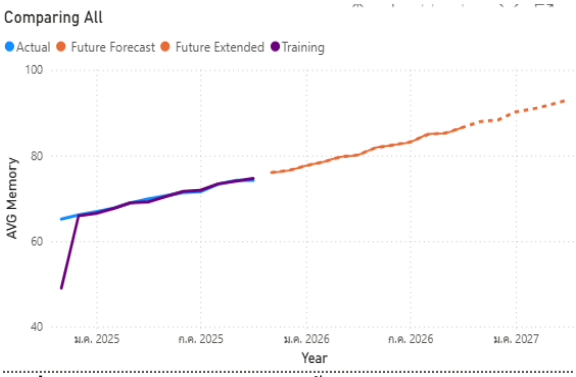

**ภาพที่ 12 :** แนวโน้มการพยากรณ์ในอนาคตของ Memory

ภาพที่ 11 CPU และ ภาพที่ 12 Memory แสดงผลพยากรณ์ล่วงหน้าด้วย SARIMA ถึงเดือน เม.ย. 2570 ซึ่งชี้ให้เห็นแนวโน้มการใช้ทรัพยากรที่เพิ่มสูงขึ้นอย่างรวดเร็ว โดยเฉพาะ Memory ที่มีความเสี่ยงเข้าใกล้ขีดจำกัดของระบบ

4.ผลการพยากรณ์อัลกอริทึมของ CPU และ Memory




**ภาพที่ 13 :** แสดงการเปรียบเทียบทั้งหมดของ CPU




**ภาพที่ 14 :** แสดงการเปรียบเทียบทั้งหมดของ Memory

ภาพที่ 13 CPU และ ภาพที่ 14 Memory แสดงความต่อเนื่องของข้อมูลตั้งแต่ปัจจุบัน และข้อมูลขณะฝึกสอนของ SARIMA คือโมเดลที่ดีที่สุดไปจนถึงการพยากรณ์อนาคต

ภาพรวมนี้ช่วยให้องค์กรเฝ้าระวังความเสี่ยงที่ทรัพยากรจะเต็มขีดจำกัด และนำไปใช้วางแผนขยายขีดความสามารถของทรัพยากรได้อย่างทันท่วงที

## 5. สรุป

### 5.1 สรุปผลการวิจัย

การศึกษานี้ประสบความสำเร็จในการพัฒนาแบบจำลองพยากรณ์การใช้งานทรัพยากรระบบ SAP จากการวิเคราะห์ข้อมูลอนุกรมเวลา 1 ปี (พ.ย. 2567 – ต.ค. 2568) ผลการเปรียบเทียบพบว่า SARIMA คือแบบจำลองที่มีประสิทธิภาพสูงสุดในการพยากรณ์ทั้ง CPU และ Memory โดยให้ค่าความคลาดเคลื่อน RMSE ต่ำที่สุด (1.549 และ 1.530 ตามลำดับ) และมีค่าสัมประสิทธิ์การตัดสินใจ ($R^2$) สูงที่สุด (0.829 และ 0.834 ตามลำดับ) นอกจากนี้ ผลการทดสอบยังบรรลุสมมติฐานที่ตั้งไว้ โดยมีค่า MAPE เพียงร้อยละ 1.7–2.1 ซึ่งต่ำกว่าเกณฑ์ที่ร้อยละ 20 อย่างมีนัยสำคัญ

เมื่อนำแบบจำลองไปพยากรณ์ล่วงหน้าถึงเดือนเมษายน พ.ศ. 2570 และแสดงผลผ่านแดชบอร์ดพบการแจ้งเตือนความเสี่ยงว่าการใช้งาน Memory จะพุ่งสูงแตะระดับวิกฤตที่ร้อยละ 93 ในขณะที่ CPU จะเพิ่มขึ้นถึงร้อยละ 79.2

### 5.2 อภิปรายผลการวิจัย

แบบจำลอง SARIMA มีความเหมาะสมกับบริบทข้อมูลของระบบ SAP ที่มีรูปแบบการทำงานซ้ำเดิมตามรอบเวลา (Seasonality) จึงสามารถเรียนรู้ได้แม่นยำกว่าแบบจำลอง LSTM ที่ต้องอาศัยข้อมูลขนาดใหญ่และระยะยาวกว่า 1 ปี โดยผลพยากรณ์ชี้ให้เห็นว่าการใช้งาน Memory มีแนวโน้มพุ่งสูงแตะระดับวิกฤตที่ร้อยละ 93 ในขณะที่ CPU มีแนวโน้มเพิ่มขึ้นสู่ระดับ ร้อยละ 79.2 ตัวเลขดังกล่าวสะท้อนถึงการเติบโตของปริมาณธุรกรรมในระบบที่อาจนำไปสู่ภาวะคอขวดซึ่งสอดคล้องกับข้อจำกัดเชิงโครงสร้างไอทีขององค์กร เมื่อพิจารณาจากทรัพยากรปัจจุบันที่มี Memory 160 GB และ CPU 6 Core อาจไม่เพียงพอต่อการรองรับในอนาคต ดังนั้น เพื่อดำเนินมาตรการเชิงรุกและป้องกันระบบหยุดทำงาน ผู้วิจัยจึงเสนอแนะให้องค์กรพิจารณาลงทุนขยายโครงสร้างพื้นฐาน โดยปรับเพิ่ม Memory เป็น 256 GB และ CPU เป็น 12 Core ซึ่งเป็นการลงทุนที่เหมาะสมและสอดคล้องกับผลพยากรณ์มีทรัพยากรสำรองที่เพียงพอต่อการปฏิบัติงานและสามารถรองรับการขยายตัวของข้อมูลได้อย่างต่อเนื่อง

### 5.3 ข้อเสนอแนะ

องค์กรควรพิจารณาเพิ่มหน่วยความจำเชิงรุกก่อนถึงเดือนเมษายน พ.ศ. 2570 เพื่อลดความเสี่ยงระบบหยุดทำงาน ควบคู่กับการใช้แดชบอร์ดติดตามช่วงเวลาที่มีภาระงานหนาแน่น เพื่อโยกย้ายงานประมวลผลหนัก หรือการสำรองข้อมูลไปทำในช่วงเวลา Off-Peak รวมถึงควรปรับแต่งคำสั่ง เพื่อลดภาระการทำงานของทรัพยากรระบบ

1. ควรพัฒนาแบบจำลองผสม เช่น ARIMA-LSTM หรือ LSTM-XGBoost เพื่อรองรับข้อมูลที่มีความซับซ้อนทั้งเชิงเส้นและไม่เชิงเส้น

2. ควรเพิ่มตัวแปรนำเข้า เช่น ปริมาณธุรกรรม หรือ จำนวนผู้ใช้พร้อมกัน เพื่อให้สะท้อนบริบทแวดล้อมได้ครอบคลุมยิ่งขึ้น

3.ควรขยายระยะเวลาการเก็บข้อมูลให้มากกว่า 1 ปี เพื่อให้แบบจำลองสามารถเรียนรู้รูปแบบฤดูกาลรายปี ได้อย่างสมบูรณ์

## เอกสารอ้างอิง


[1] Q. Li, G. Gaalman, and S. M. Disney, "Dynamic analysis of the proportional order-up-to policy with damped trend forecasts," Appl. Energy, vol. 307, p. 118092, 2022.

[2] S. Huang, Y. Lv, H. Nie, T. Jin, Z. Ye, and J. Guo, "Machine learning–enhanced runoff prediction in plain river networks: Insights into recurrent river segments," J. Hydrol. Eng., 2024, doi: 10.1061/JHYEFF.HEENG-6586.

[3] F. Dakheel and M. Cevik, "Optimizing smart grid load forecasting via a hybrid long short-term memory-XGBoost framework: Enhancing accuracy, robustness, and energy management," Energies, vol. 18, no. 11, p. 2842, 2025.

[4] Y. Wang, "Forecasting SET Index Returns Using ARIMA, SVR, RNN and Random Forest Models," Independent Study, Faculty of Commerce and Accountancy, Thammasat University, Bangkok, Thailand, 2022.

[5] J. Blanco, M. Ferreras, and O. Cosido, "Predictive modeling of hospital emergency department demand using artificial intelligence: A systematic review," Int. J. Med. Inform., vol. 207, p. 106215, 2026, doi: 10.1016/j.ijmedinf.2025.106215.

# การสกัดความสัมพันธ์ระหว่างเหตุและผลเชิงกลุ่มจากความคิดเห็นผู้ใช้ไอที
# Group-based Causality from IT-User Comments

*รัชดา เปลี่ยนทอง (Ratchada Pleanthong) ฉวีวรรณ เพ็ชรศิริ (Chaveevan Pechsiri)*
*และอดิศร ณ อุบล (Adisorn Na Ubon)*
*สาขาวิชาเทคโนโลยีสารสนเทศ วิทยาลัยวิศวกรรมศาสตร์และเทคโนโลยี มหาวิทยาลัยธุรกิจบัณฑิตย์*
*67130073@dpu.ac.th, chaveevan.pet@dpu.ac.th and adisorn.nan@dpu.ac.th*

## บทคัดย่อ

*งานวิจัยนี้มีจุดประสงค์เพื่อสกัดความสัมพันธ์ระหว่างเหตุและผลเชิงกลุ่มจากความคิดเห็นของผู้ใช้งานระบบเทคโนโลยีสารสนเทศ ซึ่งเป็นข้อความที่ไม่มีโครงสร้าง เพื่อใช้วิเคราะห์ปัญหาการใช้งานระบบในเชิงกลุ่มของเหตุและกลุ่มของผล การวิเคราะห์ดำเนินการในระดับหน่วยวาทกรรมย่อย (Elementary Discourse Unit: EDU) งานวิจัยมุ่งแก้ไขปัญหาหลัก 3 ประการ ได้แก่ การระบุ EDU ที่มีแนวความคิดเหตุ และ EDU ที่มีแนวความคิดผล การกำหนดขอบเขตของกลุ่มเหตุและกลุ่มผล และการสกัดความสัมพันธ์ระหว่างเหตุและผลเชิงกลุ่ม โดยเสนอวิธีการที่ใช้เซตแนวความคิดกริยาแสดงเหตุและกริยาแสดงผลร่วมกับการพิจารณาบริบทของ EDU ที่อยู่ก่อนหน้าและถัดไป นอกจากนี้ ได้ประยุกต์ใช้การเรียนรู้ของเครื่องแบบมีผู้สอนด้วยวิธี Naïve Bayes เพื่อเรียนรู้ความสัมพันธ์เชิงเหตุและผลจากคู่กริยา ผลการประเมินด้วยข้อมูลทดสอบจำนวน 100 EDU โดยผู้เชี่ยวชาญ 3 ท่าน แสดงให้เห็นว่าวิธีการ ที่นำเสนอสามารถระบุความสัมพันธ์ระหว่างเหตุและผลเชิงกลุ่มได้ด้วยค่าความถูกต้องร้อยละ 90 วิธีการนี้ มีประโยชน์สำหรับการวิเคราะห์ความคิดเห็นของผู้ใช้งานเพื่อระบุสาเหตุรากของปัญหาและสนับสนุนการแก้ไขปัญหาการใช้งานระบบอย่างเป็นระบบ รวมทั้งสามารถประยุกต์ใช้กับการวิเคราะห์บทวิจารณ์ในโดเมนอื่นที่มีลักษณะข้อมูลไม่เป็นโครงสร้างได้*

**คำสำคัญ:** ความสัมพันธ์ระหว่างเหตุและผลเชิงกลุ่ม หน่วยวาทกรรมย่อย บทวิจารณ์ ขอบเขตเหตุ ขอบเขตผล

## Abstract

*This research aims to extract group-based causality from users' comments on information technology systems, which are expressed as unstructured text, in order to analyze system usage problems at the level of cause groups and effect groups. The analysis is conducted at the level of Elementary Discourse Units (EDUs). The research addresses three main challenges: identifying EDUs that represent causative concepts and effect concepts, determining the boundaries of cause groups and effect groups, and extracting group-based causal relationships. The proposed approach employs sets of causative verb concepts and effect verb concepts, together with contextual information from preceding and subsequent EDUs. In addition, supervised machine learning using the Naïve Bayes method is applied to learn causal relationships from verb pairs. Experimental evaluation using a test set of 100 EDUs annotated by three experts demonstrates that the proposed method can identify group-based causality with an accuracy of 90%. The proposed approach is useful for analyzing user comments to identify root causes of system usage problems and to support systematic problem resolution. Furthermore, it can be applied to the analysis of reviews in other domains that involve unstructured textual data.*

**Keywords:** Core Asset Development, Common Asset, Control Flow Graph, Path Testing.

## 1. บทนำ

ระบบเทคโนโลยีสารสนเทศขององค์กรต่าง ๆ (เช่น สำนักงานคณะกรรมการการรักษาความมั่นคงปลอดภัยไซเบอร์แห่งชาติ เป็นต้น) ส่วนใหญ่มักพบปัญหาที่เกิดขึ้นจากระบบรวมทั้งการใช้งานระบบ ปัญหาเหล่านี้

สามารถแสดงออกในรูปของเหตุและผลที่สามารถสังเกตุได้ เช่น ความล่าช้าในการทำงาน ความไม่พึงพอใจของผู้ใช้งาน หรือประสิทธิภาพการดำเนินงานที่ลดลง อย่างไรก็ตามปัญหาที่ปรากฏนี้สามารถแสดงออกมาในรูปของการบรรยายเชิงเหตุและผลในเอกสารต่าง ๆ เช่น คำวิจารณ์ของผู้ใช้ ดังนั้นงานวิจัยนี้จึงมีจุดประสงค์การสกัดความสัมพันธ์ระหว่างเหตุและผลเชิงกลุ่ม (Group-based Causality, GC) เพื่อใช้สำหรับการวิเคราะห์ปัญหาการใช้ไอที (IT, Information Technology) จากความคิดเห็นผู้ใช้ โดย GC แสดงให้เห็นถึงกลุ่มสาเหตุ (Causative Group) ของปัญหาการใช้ไอที และกลุ่มผลลัพธ์ (Effect Group) ที่เกิดขึ้นทำให้สามารถเข้าใจปัญหาได้ง่ายสำหรับการวิเคราะห์ปัญหาที่ซับซ้อน นอกจากนี้ การวิเคราะห์ปัญหาเป็นกระบวนการอย่างเป็นระบบในการตรวจสอบประเด็นปัญหาเพื่อทำความเข้าใจอย่างถูกต้อง นำไปสู่การหาสาเหตุที่แท้จริงและผลที่เกิดขึ้นทำให้สามารถวิเคราะห์ปัญหาได้ชัดเจน โดยเฉพาะการวิเคราะห์ต้นไม้ปัญหา (Problem Tree Analysis) ซึ่งเป็นสิ่งสำคัญสำหรับการพัฒนาวิธีการแก้ปัญหาอย่างมีประสิทธิภาพและยั่งยืนในสาขาต่าง ๆ (https://www.who.int/publications/i/item/9789240068032) โดยส่วนใหญ่ GC ที่ปรากฏในข้อความหรือเอกสารวิจารณ์จะแสดงออกมาในรูปแบบความสัมพันธ์ระหว่างกลุ่มหน่วยข้อความระดับวาทกรรมย่อย (Elementary Discourse Unit, EDU, คือประโยคความเดียว [1]) 2 กลุ่ม คือ กลุ่ม EDU ที่มีแนวความคิดเหตุ (a Causative-Concept EDU group ที่แทนด้วย C ; C = $EDUc_i$ ; *i=1,2,..numc; numc* คือจำนวน EDU ที่มีแนวความคิดเหตุ; *numc>=1*) และกลุ่ม EDU ที่มีแนวความคิดผล (an Effect Concept EDU group) ที่แทนด้วย *E; E = $EDUe_j$; j=1,2,..nume; nume* คือจำนวน EDU ที่มีแนวความคิดผล; *nume>=1*) ดังต่อไปนี้

$$EDUc_1 EDUc_2 \ldots, EDUc_{numc} \xrightarrow{GC} EDUe_1\ EDUe_2 \ldots EDUe_{nume}$$

การแสดงออกของ $EDUc_i$ และ $EDUe_j$ อิงตามหลักภาษาไทยทั่วไป (General Thai Linguistic Expression) ดังต่อไปนี้ หลังจากได้ทำการหาแก่นคำ (stemming words) และกำจัดคำหยุด (stop words)

EDU → NP1 VP | VP

VP → Verb NP2 | Verb adv | Verb

Verb → $Verb_{weak}$ Word | $Verb_{strong}$

NP1 → Noun | Noun Adjphrase

NP2 → Noun | Noun Adjphrase | Noun Prepphrase

Word→ Noun| Adj| Adv| $Verb_{strong}$

$Verb_{weak}$ → {'*เป็น/be*','*ไม่เป็น/notBe*','*มี/have*','*ไม่มี/notHave*','*ใช้/use*',,'*รู้สึก/feel*'}

$Verb_{strong}$→ { '*เกิด/*appear', '*ไม่ตรง/beIrrelevant*', '*ไม่เสถียร/unstable*','*ซับซ้อน/beComplex*','*หมดอายุ/expire*''*ค้าง/freeze*,'*ช้า/beslow*','*ลด/reduce*', '*รวดเร็ว/fast*', '*ยาก/be difficult*', '*รอ/wait*'..}

Noun → {'*ปัญหา/Problem*' '*เวลา/ time*',...}

ตัวอย่างเช่น

ตัวอย่างที่1

$EDUc_1$:"*วิธีการอธิบายของผู้สอนไม่ชัดเจนเพียงพอ*"

"(*วิธีการอธิบายของผู้สอน*)/NP1 (*ไม่ชัดเจน*/$Verb_{strong}$ *เพียงพอ*/Adv)/VP"

$EDUc_2$: "*และโครงสร้างภาษาโปรแกรมซับซ้อน*"

"*และ*/Conj (*โครงสร้างภาษาโปรแกรม*)/NP1 (*ซับซ้อน*/ $Verb_{strong}$))/VP"

$EDUe_1$: "*ทำให้ผู้เรียนรู้สึกสับสน*"

"*ทำให้*/causalConj (*ผู้เรียน*)/NP1 (*รู้สึก*/$Verb_{weak}$ *สับสน*/$Verb_{strong}$)/VP"

$EDUe_2$: "*และใช้เวลาในการทำความเข้าใจ*"

"*และ*/Conj((*ใช้*/$Verb_{weak}$ *เวลา*/Noun) /Verb(*การทำความเข้าใจ*)/NP2)/VP"

อย่างไรก็ตามในงานวิจัยนี้การสกัด GC จากเอกสารมีปัญหาหลัก 3 ประการดังนี้

1. การระบุ EDU ที่ที่มีแนวความคิดเหตุ ($EDUc_i$) และ EDU ที่ที่มีแนวความคิดผล ($EDUe_j$)
2. การหาขอบเขตของเหตุ (*numc*) และขอบเขตของผล(*nume*)
3. การหา GC ระหว่าง C และ E

เนื่องจากงานวิจัยก่อนหน้า [2] - [7] แนวความคิดเหตุและแนวความคิดผล ส่วนใหญ่แสดงออกมาในรูปของคำนาม /นามวลี เช่น Earthquake และ Tsunami เป็นต้น โดยที่ Earthquake เป็นเหตุการณ์เหตุและ Tsunami เป็นเหตุการณ์ผลภายใต้ประโยคความเดียวต่อไปนี้ "Earthquake generates Tsunami" แต่แนวความคิดเหตุและแนวความคิดผลของงานวิจัยนี้แสดงออกมาในรูปของคำกริยา หรือกริยาวลี เพราะภาษาไทยส่วนใหญ่จะละนามวลี และคำกริยาเช่น *'ทำให้'* สามารถทำหน้าที่เป็นคำเชื่อมระหว่างประโยคเหตุและประโยคผล (Causal

conjunction) ดังแสดงในตัวอย่างที่ 1 และบางครั้งก็เป็นกริยาของประโยค เช่น *"กลิ่นกุหลาบทำให้รู้สึกสดชื่น"*

ในงานวิจัยของ [2] ใช้กริยาหลักของ EDU ที่มีแนวความคิดเหตุ และกริยาหลักของ EDU ที่มีแนวคิดผลแทนเหตุการณ์เหตุ และเหตุการณ์ผล ตามลำดับ ดังนั้นงานวิจัยนี้จึงเสนอใช้เซตแนวความคิดกริยาแสดงเหตุ (a causative verb concept set, CV) และเซตแนวความคิดกริยาแสดงผล (a effect verb concept set, EV) มาทำการระบุ EDU$c_i$ และ EDU$e_j$ ตามลำดับ พร้อมหาขอบเขตของเหตุ และขอบเขตของผล ตามลำดับ

CV= {('ขาด / lack of')/Verb$_{strong}$, ('มี + ปัญหา / have+Problem')/Verb$_{weak}$+Word,('ไม่ตรง beIrrelevant')/Verb$_{strong}$, ('ไม่เสถียร / unstable') /Verb$_{strong}$, ('ซับซ้อน beComplex')/Verb$_{strong}$, ('หมดอายุ/expire')/ Verb$_{strong}$,...}

EV= {('ค้าง / freeze')/ Verb$_{strong}$, ('เกิด / appear')/ Verb$_{strong}$,('ช้า / beSlow')/ Verb$_{strong}$,('ลด / reduce')/Verb$_{strong}$, ('รวดเร็ว/fast') Verb$_{strong}$, ('ยาก / beDifficult')/Verb$_{strong}$, ('รอ / wait') /Verb$_{strong}$,('ใช้ + เวลา / takeTime')/Verb$_{weak}$ + Word,...}

แต่เกิดมีความทับซ้อนกัน (Overlap) ระหว่างเซตแนวความคิดกริยาเหตุและเซตแนวความคิดกริยาผล เกิดเป็นเซตกริยาซับซ้อน ( Overlap Verb set ) คือ $CV \cap EV = OV$ และ $OV \neq \emptyset$ ; $OV = \{$*'ค้าง/freeze', 'เกิด/appear', 'ช้า/beslow',...}* ดังนั้นจึงใช้บริบท EDU (Context EDU) ที่อยู่ก่อนหน้าและหลัง EDU$_{OV}$ (คือ EDU ที่มี Verb อยู่ใน OV; Verb∈OV) มาเป็นตัวกำหนด Verb ของ EDU$_{OVerb}$ ว่าแสดงแนวความคิดเหตุหรือแนวความคิดผล ตามลำดับกฎต่อไปนี้ หลังจากกำหนดให้

$$CCV = CV - OV \quad (1)$$

$$EEV = EV - OV \quad (2)$$

CCV คือเซตแนวความคิดกริยาแสดงเหตุที่ไม่มีเซตกริยาซับซ้อน EEV คือเซตแนวความคิดกริยาแสดงผลที่ไม่มีเซตกริยาซับซ้อน

เมื่อ Verb$_{EDUbefore}$ Verb$_{EDU_OV}$ Verb$_{EDUafter}$

1. ถ้า Verb$_{EDUbefore}$ (คือ Verb ของ EDU ก่อนหน้า EDU$_{OV}$) ∈ CCV or ∈ EEV , แล้วทำให้ Verb$_{EDU_OV}$ (คือ Verb ของ EDU$_{OV}$) มีแนวความคิดเป็นผล (*verb$_{effect}$*) หรือเหตุ (*verb$_{cause}$*) ตามลำดับ
2. ถ้า (Verb$_{EDUbefore}$ ∉ CCV or ∉ EEV) และ Verb$_{EDUafter}$ (คือ Verb ของ EDU หลัง EDU$_{OV}$) ∈ CV or ∈ EV , แล้วทำให้ *Verb$_{EDU_OV}$* เป็น *verb*$_{effect}$ หรือ *verb$_{cause}$* ตามลำดับ

นอกจากนี้เราได้ประยุกต์ใช้การเรียนรู้ของเครื่องที่มีผู้เชี่ยวชาญกำกับ (Supervised Machine Learning) ที่เป็น เนอีฟเบย์ (Naïve Bayes) [8] มาหา GC โดยใช้ NB เรียนรู้ความสัมพันธ์ระหว่างคู่กริยา (Verb pair) $v_{ci}$ $v_{ej}$ เป็น Causality ในทุกคู่กริยา $v_{ci}$ $v_{ej}$ ระหว่าง C และ E เมื่อ $v_{ci}$ คือ Verb$_{EDUci}$ เป็นตัวแทน EDU$c_i$ และ $v_{ej}$ คือ Verb$_{EDUej}$ เป็นตัวแทน EDU$_{ej}$

ในส่วนที่ 2 สรุปงานวิจัยที่เกี่ยวข้อง ส่วนที่ 3 แสดงภาพรวมของระบบการสกัด GC จากเอกสารคำวิจารณ์ ส่วนที่ 4 ประเมินและอภิปรายวิธีการที่เสนอและข้อสรุป

## 2. งานวิจัยที่เกี่ยวข้อง

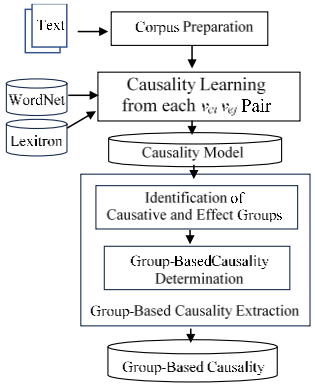
มีงานวิจัยด้านการสกัดความสัมพันธ์เชิงเหตุและผลจากข้อความในเอกสารจำนวนมาก [2]-[7] ทั้งในระดับความเดี่ยว (Simple Sentence) และระดับความซ้อน (Complex Sentence) [3] ,[4] ได้ทำการสกัดความสัมพันธ์เชิงเหตุ–ผล )Causal Relation) จากคลังข้อมูลข่าว LA Times และ Washington Post ตามลำดับ โดยใช้กฎทางภาษาศาสตร์ ร่วมกับคูเวิร์ด (Cue Word) เช่น '*cause*' '*effect to*' '*due to*' '*because*' เป็นต้น และการเรียนรู้ของเครื่องแบบ C4.5 Decision Tree หาความสัมพันธ์เชิงเหตุ-ผลภายใต้ประโยคความเดียวได้ค่าความแม่นยำ (Precision) 73.91% และค่าระลึก (Recall) 88.69% [3] หรือใช้คีย์เฟส (Key Phase) เช่น '*..because of…*''*..as the result of..*' เป็นต้น ร่วมกับค่าความน่าจะเป็นของคู่คำศัพท์ [4] มาระบุความสัมพันธ์เชิงเหตุ-ผล ภายใต้ประโยคเชิงซ้อนที่แสดงความสัมพันธ์ระหว่างสองคำนาม/นามวลีที่คำนาม นามวลีหนึ่งมีแนวความคิดเป็นเหตุ และอีกคำนาม/นามวลีที่มีแนวความคิดเป็นผล ได้ค่า F1-Score เฉลี่ย เป็น 0.75 [4] ,[5] ได้ทำการสกัดความสัมพันธ์เชิงเหตุ–ผลจากคลังข้อมูลไทยโดเมนโรคพืช การดูแลสุขภาพและภาวะโลกร้อน โดยใช้การเรียนรู้ของเครื่องแบบต่าง ๆ เช่น เนอีฟเบย์ (Naïve Bayes), ซัพพอร์ตเวกเตอร์ แมชชีน

(Support Vector Machine) และโครงข่ายประสาทเทียม (Neural Network) เป็นต้น กับคู่กริยา (ที่ได้จาก เซตกริยาที่มีแนวความคิดเหตุ และเซตกริยาที่มีแนวความคิดผล; เซตกริยาที่มีแนวความคิดเหตุ ∩ เซตกริยาที่มีแนวความคิดผลได้ค่าว่าง) แทนคู่เหตุการณ์ระหว่างเหตุการณ์เหตุและเหตุการณ์ผลจาก คู่ประโยคความเดียว หรือคู่ EDU ที่มีอยู่ต่อเนื่องกันหลาย ๆ EDUs โดยมีหนึ่ง EDU เป็น EDU ที่มีแนวความคิดเป็นเหตุ และหลาย EDUs ที่เหลือมีแนวความคิดเป็นผล ได้ค่า F1-Score เฉลี่ยเป็น 0.77 สำหรับการสกัดความสัมพันธ์เชิงเหตุ–ผล ของคู่กริยาดังกล่าวจากคู่ EDU ที่อยู่ต่อเนื่องกัน [6] ได้ระบุเหตุการณ์ความสัมพันธ์เชิงเหตุ-ผล (Event Causality Identification) ด้วยเหตุการณ์ที่แสดงด้วยนามวลีจากโดเมนข่าวสาธารณะของจีน โดยใช้ทรานฟอร์มเมอร์โมเดลหลักคือ BERT ที่เป็น Encoder ระบุคู่เหตุการณ์ที่มีความสัมพันธ์เชิงเหตุ-ผลทั้งประโยคความเดียวและประโยคความซ้อน ได้ค่าความถูกต้อง 68% [7] ได้เสนอการตีความเชิงสาเหตุโดยใช้ Self-attention ในโครงข่ายประสาทเทียมแบบทรานฟอร์มเมอร์กับข้อมูล 25,000 ตัวอย่าง/รีวิว เพื่อหา causal relation ระหว่าง token สำหรับสร้าง causal graph สำหรับใช้ในระบบแนะนำ

งานวิจัยข้างต้น [2]-[7] เน้นในเรื่องการหาความสัมพันธ์เชิงเหตุ-ผลจากคู่นามวลี หรือ คู่กริยาวลีโดยปราศจากการพิจารณาระหว่างคู่กลุ่มเหตุและกลุ่มผล ซึ่งส่วนใหญ่จะปรากฏในเอกสารความคิดเห็น

## 3. ภาพรวมของระบบ

การสกัดความสัมพันธ์เชิงเหตุและผลแบบเป็นกลุ่มประกอบด้วย 3 ขั้นตอนหลักคือ การเตรียมคลังข้อมูล (Corpus Preparation) การเรียนรู้ความสัมพันธ์เชิงเหตุและผลสำหรับคู่เหตุ-ผล (Causality Learning from each Cause-Effect Pair) และสกัดความสัมพันธ์ระหว่างเหตุและผลเชิงกลุ่ม (Group Based Causality Extraction) (ดูภาพที่ 1)




**ภาพที่ 1:** กรอบงานระบบ

### 3.1 การเตรียมคลังข้อมูล

ขั้นตอนนี้มีวัตถุประสงค์เพื่อเตรียมคลังข้อมูล EDU (EDU corpus) จากเอกสารความคิดเห็นของผู้ใช้ที่ได้จากหน่วยงานภาครัฐคือ สำนักงานคณะกรรมการการรักษาความมั่นคงปลอดภัยไซเบอร์แห่งชาติ ขั้นตอนดังกล่าวประกอบด้วยการใช้เครื่องมือตัดคำภาษาไทย [9] และการแบ่งหน่วย EDU (EDU Segmentation) [10] เพื่อสร้างคลังข้อมูลจำนวน 600 EDUs จากเอกสารความคิดเห็นจำนวน 100 ฉบับ โดยคลังข้อมูลดังกล่าวมีการหาแก่นคำ (stemming) และการตัดคำหยุด (stop word removal) และถูกแบ่งออกเป็น 3 ส่วน ได้แก่

(1) ส่วนแรกจำนวน 100 EDUs ใช้เป็นคลังข้อมูลที่มีการกำกับ GC (annotated GC corpus) (ดูภาพที่ 2) โดยแนวคิดของคำกริยาอ้างอิงจาก WordNet [11] หลังจากแปลจากภาษาไทยเป็นภาษาอังกฤษด้วย Lexitron (http://lexitron. nectec.or.th/; [12]) เพื่อใช้ในการรวบรวมเซต CV และเซต EV จากป้ายกำกับ cv และ ev

*"วิธีการอธิบายของผู้สอนไม่ชัดเจนเพียงพอ"*[EDU1] *"และโครงสร้างภาษาโปรแกรมซับซ้อน"*[EDU2]
*"ทำให้ผู้เรียนรู้สึกสับสน"*[EDU3] *"และใช้เวลาในการทำความเข้าใจ"*[EDU4]
<GC: ID 1 ; type: 1 (Group-basedCausality=1 Non-Group-basedCausality=0) >
<CauseGroup><EDUc: C_no=C1><NP1><HeadNoun concept= 'method' >*วิธี*</HeadNoun>*การอธิบายของผู้สอน*</NP1>
<cv type: strong verb; concept='unclear'>*ไม่ชัดเจน*</ cv >*เพียงพอ*
</EDUc>
<EDUc: C_no=C2><NP1><HeadNoun concept='structure'>*โครงสร้าง*
</HeadNoun>*ภาษาโปรแกรม*</NP1>
< cv type: strong verb; concept='solution'>*ซับซ้อน*</ cv >
</EDUc></CauseGroup>
<EffectGroup><EDUe: E_no=E1>*ทำให้*<NP1><HeadNoun concept='person'>*ผู้เรียน*</HeadNoun></NP1>
< ev type: weak verb +word; concept='confuse'>*รู้สึก+สับสน*</ev>
</EDUe>
<EDUe: E_no=E2>และ< NP1 Ø='person'> Ø </NP1>
< ev type: weak verb +word; concept='take time'>*ใช้+เวลา*</ ev >*ในการทำความเข้าใจ*</EDUe></EffectGroup>
<List: EDUc-EDUePair with Causality> C1-E1, C1-E2, C2-E1, C2-E2</List>
<List: EDUc-EDUePair with nonCausality> - </List>
</GC>
Where GC = Group-based Causality Tag; EDUc/EDUe = a causative/effect EDU tag (EDU is Elementary Discourse Unit); cv = Causative Verb Concept Tag; ev = Effect Verb Concept Tag

**ภาพที่ 2:** การกำกับคลังข้อมูล

(2) ส่วนที่สองจำนวน 400 EDUs ใช้สำหรับการเรียนรู้ความสัมพันธ์เชิงเหตุและผลจากคู่เหตุ–ผลที่ได้มีการกำกับ

(3) ส่วนที่สามจำนวน 100 EDUsใช้สำหรับการทดสอบและ/หรือการสกัดความสัมพันธ์เชิงเหตุและผลแบบกลุ่ม

**3.2 การเรียนรู้ความสัมพันธ์เชิงเหตุและผลสำหรับคู่กริยา $v_{ci}\, v_{ej}$**

ขั้นตอนนี้เป็นการเรียนรู้ความสัมพันธ์ ระหว่างคู่กริยา $v_{ci}\, v_{ej}$ เป็น Causality ในทุกคู่กริยา $v_{ci}v_{ej}$ ระหว่าง C และ E ด้วยวิธี Naive Bayes (NB) [8] เพื่อเรียนรู้คู่คำกริยา (verb pairs) โดยจำแนกความสัมพันธ์ระหว่างคู่กริยาเป็นคลาสความสัมพันธ์เชิงเหตุและผล จากการกำกับคู่ EDU (C_no - E_no) ที่มีการกำกับ Causality/ nonCausality จากขั้นตอนการเตรียมคลังข้อมูล ผลลัพธ์การเรียนรู้ในขั้นตอนนี้ โดยใช้เครื่องมือ Weka [13], [14] คือค่าความน่าจะเป็นของคู่คำกริยาที่มีการแสดงแนวคิดจากคลังข้อมูลสำหรับการเรียนรู้ที่มีการกำกับคำประโยคคู่เหตุ-ผลดังแสดงในตารางที่ 1

**ตารางที่ 1** ค่าความน่าจะเป็น $v_{ci}\, v_{ej}$ จาก NB

| $v_{ci}$ | Causality | NonCausality |
|---|---|---|
| Crash | 0.0303 | 0.0457 |
| haveProblem | 0.0964 | 0.0718 |
| open | 0.0468 | 0.0326 |
| explain | 0.0303 | 0.0196 |
| haveQuantity | 0.0358 | 0.0065 |
| … | … | … |
| $v_{ej}$ | Causality | NonCausality |
| Wait | 0.0605 | 0.0176 |
| wasteTime | 0.0394 | 0.0294 |
| appear | 0.0394 | 0.0294 |
| affect | 0.0342 | 0.0294 |
| notWork | 0.0236 | 0.0411 |
| … | … | … |

**3.3 สกัดความสัมพันธ์ระหว่างเหตุและผลเชิงกลุ่ม**

ขั้นตอนนี้ประกอบด้วย 2 ขั้นตอนต่อไปนี้

3.3.1 การระบุกลุ่มเหตุ (C) และกลุ่มผล (E)

ขั้นตอนนี้ใช้ CCV, EEV และ OV ระบุ $EDUc_i$ และ $EDUe_j$ จนได้แต่ละคู่ C E ใน Test Corpus

3.3.2 การหาความสัมพันธ์เชิงเหตุและผลจากทุกคู่กริยาระหว่าง C และ E ของแต่ละคู่ C E

นำแต่ละคู่ C E ทุกคู่กริยา $Verb_{EDUci}$ $Verb_{EDUej}$ $(v_{ci}\, v_{ej})$ หาความสัมพันธ์เชิงเหตุและผล ด้วย NB ในสมการ (3) กับค่าความน่าจะเป็นของคู่กริยา $v_{ci}\, v_{ej}$ ที่ได้จากการเรียนรู้ความสัมพันธ์เป็น CausalityRelation หรือ NonCausality-Relation (ตารางที่ 1) ด้วย การตรวจสอบแบบไขว้ 10 เท่า (ten-fold cross validation)

$$VerbPairRelation = \underset{class \in Class}{\operatorname{argmax}}\ P(class|v_{ci}v_{ej})$$

$$\underset{class \in Class}{\operatorname{argmax}}\ argmaxP(v_{ci}|class)P(v_{ej}|class)P(class) \quad (3)$$

เมื่อ $v_{ci} \in$CCV ; $v_{ej} \in$EEV; Class= {‘CausalityRelation’ , ‘NonCausalityRelation’}

ความสัมพันธ์ ระหว่าง C และ E ในแต่ละ คู่ C E เป็น GC เมื่อทุกคู่กริยา $Verb_{EDUci}$ $Verb_{EDUej}$ ใน C และ E มีความสัมพันธ์เป็น Causality ดังอัลกอริทึมในภาพที่ 3

```
Algorithm: GROUP_BASED_CAUSALITY_EXTRACTION
Input: EDU sequence L
Output: results (list of CE pairs)
Initialize:
j ← 1, k ← 1
results ← empty list
Initialize empty lists: C, E, OL
      //C is Cause; E is Effect, OL is Overlap
Reset flags: flagC, flagE, flagO ← 0
// Step 1: Get First verb Group
While j ≤ Length(L):
BeginWhile
   v ← getVerb(EDUj)
   Case: v ∈ CCV:
     While v ∈ CCV and j ≤ Length(L): //GetListC_byDoWhileLoop
      BeginWhile Add v to C ; j ← j + 1; flagC ← 1 EndWhile
   Case: v ∈ EEV:
     While v ∈ EEV and j ≤ Length(L): //GetListE_byDoWhileLoop
      BeginWhile Add v to E; j ← j + 1; flagE ← 1 EndWhile
   Case: v ∈ OV:
     While v ∈ OV and j ≤ Length(L): //GetListOL_byDoWhileLoop
      BeginWhile Add v to OL ; j ← j + 1; flagO ← 1 EndWhile
   Break
 // Adjust OL based on context
  If flagO = 1 and next verb ∈ CCV :
    BeginIf Move OL → C ; flagO ← 0 ; flagC ← 1 EndIf
  Else if next verb ∈ EEV:
    BeginIf Move OL → E ; flagO ← 0 ; flagE ← 1 EndIf
 // Step 2: Get Second verb Group
  If flagC=1 then
     BeginIf GetListE_byDoWhileLoop ; flagC← 0 EndIf
  If flagE=1 then
     BeginIf GetListC_byDoWhileLoop ; flagE← 0 EndIf
 // Step 3: Check C-E relation
  flagnonCE ← 0
  For each vc in C:
    For each ve in E:
      BeginFor class ← argmax P(class | vc, ve)
                If class = NonCausality:
                   flagnonCE ← 1
      EndFor
// Step 4: Store result
  If flagnonCE = 0 and C ≠ ∅ and E ≠ ∅:
     results[k] ← (C, E);
  k ← k + 1
EndWhile
Return results
```

ภาพที่ 3 อัลกอริทึมการสกัดความสัมพันธ์ระหว่างเหตุและผลเชิงกลุ่ม

## 4. การประเมินผลการอภิปรายและข้อสรุป

งานวิจัยที่เสนอมีการประเมินผล 3 ส่วนกับข้อมูลทดสอบ (Test Corpus) มี 100 EDUs โดยผู้เชี่ยวชาญ 3 ท่าน ผ่านการลงคะแนนแบบชนะสูงสุด (Max win Voting)

การประเมินส่วนแรกคือการระบุ EDU ที่ที่มีแนวความคิดเหตุ ($EDUc_i$) และ EDU ที่ที่มีแนวความคิดผล ($EDUe_j$) และการหาขอบเขตของเหตุ (*numc*) และขอบเขตของผล (*nume*) การประเมินส่วนที่สองคือ การหาความสัมพันธ์เหตุและผลจากคู่ $EDUc_i$-$EDUe_j$ และการประเมินส่วนที่สามคือการหา GC ระหว่างกลุ่ม EDU ที่ที่มีแนวความคิดเหตุ และกลุ่ม EDU ที่ที่มีแนวความคิดผล การประเมินครั้งแรกและครั้งที่สองนั้นอิงตามค่าความแม่นยำ (Precision,PC) และค่าระลึก (Recall,RC) ส่วนการประเมินส่วนที่สามอิงตามความถูกต้อง (Accuracy,AC) ดังแสดงในตารางที่ 2

**ตารางที่ 2** ผลการประเมิน

| การประเมินผล | ค่าความแม่นยำ (PC) | ค่าระลึก (RC) | ความถูกต้อง (AC) |
|---|---|---|---|
| การระบุ $EDUc_i$ | 0.79 | 0.74 | - |
| การระบุ $EDUe_j$ | 0.77 | 0.72 | - |
| การหาขอบเขตของ *numc* | 0.75 | 0.70 | - |
| การหาขอบเขตของ *nume* | 0.74 | 0.69 | - |
| Causality ($EDUc_i$- $EDUe_j$) | 0.8 | 0.75 | - |
| Group-based Causality | - | - | 0.90 |

ผลการประเมินแสดงให้เห็นว่าวิธีการที่นำเสนอสามารถสกัดความสัมพันธ์เชิงเหตุและผลได้อย่างมีประสิทธิภาพ ทั้งในระดับคู่เหตุ–ผล และระดับความสัมพันธ์เชิงเหตุและผลแบบกลุ่ม โดยให้ค่าความแม่นยำ (Precision) เท่ากับ 0.80 และค่าระลึก (Recall) เท่ากับ 0.75 ในระดับคู่ และให้ค่าความถูกต้องร้อยละ 90 ในการระบุความสัมพันธ์เชิงเหตุและผลแบบกลุ่ม ในขณะที่ [6] ได้ทำนายคู่ Cause-Efface ความถูกต้อง 68% โดยใช้โมเดล BERT กับข้อมูล 7,000 ประโยค

ผลลัพธ์ดังกล่าวสะท้อนว่าวิธีการสามารถระบุองค์ประกอบเหตุและผล รวมถึงโครงสร้างความเชื่อมโยงระหว่างกลุ่มเหตุและกลุ่มผลได้อย่างเหมาะสม แม้ข้อมูลความคิดเห็นจะเป็นข้อความที่ไม่มีโครงสร้างและมีความซับซ้อนของภาษา ที่สำคัญ วิธีการนี้ช่วยให้การวิเคราะห์ปัญหามีความชัดเจนและเป็นระบบมากขึ้น โดยยกระดับจากการพิจารณาเหตุ–ผลแบบแยกส่วนไปสู่การวิเคราะห์ปัญหาเชิงกลุ่ม ซึ่งสอดคล้องกับวัตถุประสงค์ของงานวิจัย และยืนยันศักยภาพในการประยุกต์ใช้กับการวิเคราะห์ความคิดเห็นในบริบทอื่นที่มีลักษณะข้อมูลไม่เป็นโครงสร้างได้อย่างเป็นระบบ

## เอกสารอ้างอิง


[1] Carlson L, Marcu D, Okurowski M E. “Building a Discourse-Tagged Corpus in the Framework of Rhetorical Structure Theory.” *Current Directions in Discourse and Dialogue*, van Kup-pevelt J, Smith R (eds.), Kluewer Academic Publishers, 2003, pp. 85-112.

[2] Pechsiri C, Kawtrakul A. “Mining causality for explanation knowledge from text.” *Journal of Computer Science and Technology*, 2007, 22 (6): 877-889.

[3] Girju R. “Automatic detection of causal relations for question answering.” In Proc. *The 41st Annual Meeting of the Association for Computational Linguistics, Workshop on Multilingual Summarization and Question Answering - Machine Learning and Beyond*, Sapporo, Japan, Jul. 7-12, 2003, pp.76-83.

[4] Chang D S, Choi K S. “Causal relation extraction using cue phrase and lexical pair probabilities.”*Proc.of IJCNLP*, Hainan Island, China, Mar. 22-24, 2004, pp.61-70.

[5] Pechsiri C, Piriyakul R. “Explanation knowledge graph construction through causality extraction from texts.” *Journal of Computer Science and Technology*., 25 (5) (2010), pp. 1055-1070

[6] Cui S, Sheng J, Cong X, Li QG, Liu T, Shi J. “Event Causality Extraction with Event Argument Correlations.” *Proc. of the 29th International Conference. on Computational Linguistics*, 2022, Republic of Korea, pp 2300 - 2312.

[7] Rohekar RY, Gurwicz Y, Nisimov S. “Causal interpretation of self-attention in pre-trained transformers” *Proc. Of the 37th Inter. Conf. on Neural Info. Proc. Sys.*, Article No.: 1366, pp 31450 - 31465

[8] Mitchell T M. *Machine Learning*. The McGraw-Hill Companies Inc. and MIT Press, Singapore, 1997.

[9] Sudprasert S, Kawtrakul A. “Thai word segmentation based on global and local unsupervised learning.” *In Proc. NCSEC 2003*, Chonburi, Thailand, 2003, pp. 1-8.

[10] Chareonsuk J, Sukvakree T, Kawtrakul A. “Elementary discourse unit segmentation for Thai using discourse cue and syntactic information.” *In Proc. NCSEC 2005*, Thailand, Oct. 27-28, 2005, pp. 85-90.

[11] Miller G.A. “WordNet: a lexical database for English” *Commun. ACM*, 38 (11) (1995), pp. 39-41

[12] Trakultaweekoon K, Porkaew P, Supnithi T. (2007). “LEXiTRON vocabulary suggestion system with recommend -dation and vote mechanism,” *in Proc. SNLP 2007*, Pattaya, Thailand, pp. 43–48.

[13] R.R. Bouckaert, et al. “WEKA Manual for Version 3.9.1 Univ.” Waikato Hamilt. N. Z. (2016), pp. 1-341

[14] Hall M., Frank E., Holmes G., Pfahringer B., Reutemann P., and Witten I. H., “The WEKA data mining software: an update” *ACM SIGKDD Explor. Newsl.*, 11 (1) (2009), pp. 10-18

# แนวทางการส่งเสริมการนำเทคโนโลยีปัญญาประดิษฐ์ไปใช้ในการจัดการเรียนการสอนในอาชีวศึกษา

# Guidelines for Promoting Artificial Intelligence Adoption in Vocational Education Teaching Practices

*รัชนาท หวนสิน (Ratchanat Huansin)[1] และนลินภัสร์ บำเพ็ญเพียร (Nalinpat Bhumpenpein)[2]*
*ภาควิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล*
*มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ*
*[1] 6707011857047@kmutnb.ac.th, [2]nalinpat.b@itd.kmutnb.ac.th*

## บทคัดย่อ

*บทความนี้เป็นการวิจัยเชิงสำรวจ มีวัตถุประสงค์เพื่อศึกษาปัจจัยที่ส่งผลต่อการยอมรับและการประยุกต์ใช้ปัญญาประดิษฐ์ และเพื่อเสนอแนวทางการส่งเสริมการนำเทคโนโลยีดังกล่าวไปใช้ในการจัดการเรียนการสอนของครูในสถาบันอาชีวศึกษา กลุ่มตัวอย่างคือครูสังกัดสำนักงานคณะกรรมการการอาชีวศึกษา จำนวน 447 คน เก็บข้อมูลด้วยแบบสอบถามและวิเคราะห์ผลด้วยสถิติเชิงพรรณนาและการวิเคราะห์การถดถอยพหุคูณแบบขั้นตอน ผลการวิจัยพบว่ามี 4 ปัจจัยที่ส่งผลต่อความสำเร็จอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 ได้แก่ ศักยภาพการเรียนรู้เฉพาะบุคคลของครู ความคาดหวังผลสัมฤทธิ์การสอน ความพร้อมของเทคโนโลยีและโครงสร้างพื้นฐาน และการสนับสนุนเชิงยุทธศาสตร์จากผู้บริหาร โดยตัวแปรทั้งหมดสามารถร่วมกันพยากรณ์ความสำเร็จได้ร้อยละ 74.5 นำไปสู่ข้อเสนอแนะในการพัฒนาความฉลาดรู้ด้าน AI ของครู การปรับโครงสร้างพื้นฐานสู่ระบบคลาวด์ และการสนับสนุนเชิงนโยบายจากผู้บริหารอย่างเป็นรูปธรรม*

**คำสำคัญ:** ปัญญาประดิษฐ์ อาชีวศึกษา การจัดการเรียนการสอน ปัจจัยความสำเร็จ การยอมรับเทคโนโลยี

## Abstract

*This survey research aims to study the factors affecting the acceptance and application of Artificial Intelligence (AI) and to propose guidelines for promoting its use in the teaching and learning management of vocational education teachers. The sample consisted of 447 teachers under the Office of the Vocational Education Commission. Data were collected via questionnaires and analyzed using descriptive statistics and Stepwise Multiple Linear Regression. The findings revealed that four factors significantly impacted success at the 0.05 level: teachers' personalized learning capabilities, performance expectancy, technology and infrastructure readiness, and strategic support from administrators. These variables combined could predict success by 74.5%. The results lead to recommendations for developing teachers' AI literacy, upgrading cloud infrastructure, and providing concrete policy support from executives.*

**Keywords:** Artificial Intelligence, Vocational Education, Teaching and Learning, Success Factors Technology adoption

## 1. บทนำ

การเปลี่ยนผ่านสู่สังคมดิจิทัลทำให้อุตสาหกรรม 4.0 ต้องการกำลังคนสมรรถนะสูง เทคโนโลยีปัญญาประดิษฐ์ (Artificial Intelligence: AI) จึงเข้ามามีบทบาทสำคัญต่อการจัดการศึกษาและการฝึกอบรมด้านเทคนิคและอาชีวศึกษา (TVET) โดย AI เข้ามาช่วยสนับสนุนการเรียนรู้เฉพาะบุคคล ระบบผู้ช่วยสอนอัจฉริยะ และการวิเคราะห์ข้อมูลการเรียน ซึ่งมีส่วนช่วยยกระดับคุณภาพ

การสอน ลดความเหลื่อมล้ำ และแก้ปัญหาการขาดแคลนครูได้อย่างเป็นรูปธรรม [1][2] แม้สำนักงานคณะกรรมการการอาชีวศึกษา (สอศ.) จะมีนโยบายขับเคลื่อนการทรานส์ฟอร์มดิจิทัล แต่ในระดับปฏิบัติการและชั้นเรียนยังคงเผชิญความท้าทายสำคัญ ได้แก่ ความแตกต่างด้านความฉลาดรู้เรื่อง AI ของครูผู้สอน ข้อจำกัดด้านเวลา โครงสร้างพื้นฐานที่ไม่พร้อม รวมถึงความกังวลด้านจริยธรรมข้อมูล และ ความไว้วางใจของผู้ใช้ต่อความผิดพลาดของระบบ AI [3] ด้วยเหตุนี้ การวิจัยครั้งนี้จึงมุ่งสังเคราะห์และทดสอบแบบจำลองปัจจัยความสำเร็จของการนำปัญญาประดิษฐ์ไปใช้ในการจัดการเรียนการสอนของครูอาชีวศึกษา ภายใต้กรอบ ทฤษฎีเอกภาพว่าด้วยการยอมรับและการใช้เทคโนโลยี ฉบับที่ 2 (UTAUT2) เพื่อสร้างหลักฐานเชิงประจักษ์สำหรับใช้กำหนดนโยบายและต่อยอดเป็นแนวทางส่งเสริมที่ปฏิบัติได้จริงสำหรับสถานศึกษาต่อไป โดยผู้วิจัยได้กำหนดวัตถุประสงค์ของการวิจัย ดังนี้ เพื่อศึกษาปัจจัยที่ส่งผลต่อการยอมรับและการประยุกต์ใช้ปัญญาประดิษฐ์ในการจัดการเรียนการสอนของครูในสถาบันอาชีวศึกษา และเพื่อเสนอแนวทางการส่งเสริมการนำเทคโนโลยีปัญญาประดิษฐ์ไปใช้ในการจัดการเรียนการสอนในอาชีวศึกษา

## 2. วรรณกรรมที่เกี่ยวข้อง

### 2.1 ปัญญาประดิษฐ์

ปัญญาประดิษฐ์ (Artificial Intelligence: AI) หมายถึงสาขาวิชาวิทยาการคอมพิวเตอร์ที่มุ่งเน้นการสร้างระบบหรือเครื่องจักรที่มีความสามารถในการเลียนแบบสติปัญญาและพฤติกรรมของมนุษย์ เช่น การเรียนรู้ (Learning) การให้เหตุผล (Reasoning) การแก้ปัญหา (Problem Solving) การจดจำรูปแบบ (Pattern Recognition) และการเข้าใจภาษาธรรมชาติ (Natural Language Processing) [3][4] นอกจากนี้ AI ยังถูกนิยามว่าเป็นระบบที่สามารถประมวลผลข้อมูล วิเคราะห์เชิงลึก และตัดสินใจได้คล้ายหรือดีกว่ามนุษย์ในบางภารกิจ เพื่อเพิ่มประสิทธิภาพและลดข้อจำกัดของมนุษย์

### 2.2 แนวทางการส่งเสริมการใช้ปัญญาประดิษฐ์ในอาชีวศึกษา

เพื่อยกระดับคุณภาพการเรียนการสอนและการบริหารจัดการ จำเป็นต้องมีแนวทางที่ครอบคลุมทั้งด้านนโยบาย โครงสร้างพื้นฐาน และบุคลากร ดังนี้

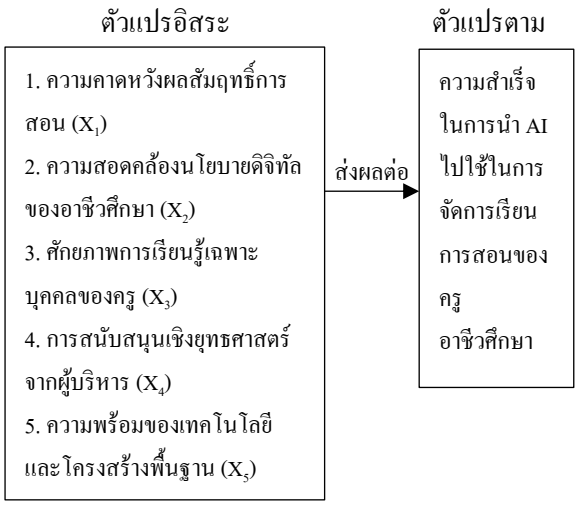
2.2.1 ความสำเร็จของการใช้ AI ในอาชีวศึกษา ขึ้นอยู่กับการกำหนดนโยบายองค์กรที่ชัดเจน (Organizational Policy) และการสนับสนุนจากผู้บริหาร (Administrative Support) ซึ่งต้องสอดคล้องกับแผนยุทธศาสตร์ชาติและบริบทของพื้นที่ เช่น เขตพัฒนาพิเศษภาคตะวันออก (EEC) ผู้บริหารต้องมีวิสัยทัศน์ในการนำ AI มาใช้เพื่อการตัดสินใจ การพัฒนาหลักสูตรที่ทันสมัย และการบริหารจัดการที่โปร่งใส [5] รวมถึงต้องมีการกำหนดจริยธรรมในการใช้ AI และนโยบายความเป็นส่วนตัวของข้อมูลเพื่อป้องกันความเสี่ยง

2.2.2 การพัฒนาหลักสูตรและการเรียนการสอนสถานศึกษา โดยหลักสูตรควรปรับปรุงให้บูรณาการทักษะ AI Data Science และทักษะดิจิทัล เพื่อให้ผู้เรียนมีสมรรถนะตรงตามความต้องการของตลาดแรงงานในยุคอุตสาหกรรม 4.0 [6][7] ส่งเสริมการใช้ AI Tools เช่น ระบบติวเตอร์อัจฉริยะ (Intelligent Tutoring Systems) และแพลตฟอร์มการเรียนรู้แบบปรับตัว (Adaptive Learning) ที่สามารถวิเคราะห์พฤติกรรมผู้เรียนและปรับเนื้อหาให้เหมาะสมกับรายบุคคล ซึ่งจะช่วยเพิ่มผลสัมฤทธิ์ทางการเรียนและลดอัตราการออกกลางคัน [8][9]

2.2.3 การเตรียมความพร้อมบุคลากร องค์กรควรพัฒนาทักษะความฉลาดทางดิจิทัลและ AI (AI Literacy) ให้แก่ครูผู้สอน เพื่อให้สามารถเลือกใช้เครื่องมือ AI ได้อย่างเหมาะสมและมีวิจารณญาณ ครูต้องเปลี่ยนบทบาทจากผู้สอนเป็นผู้ชี้แนะ (Facilitator) ที่ใช้ AI เป็นผู้ช่วยสอน [10]

2.2.4 การสร้างระบบนิเวศและความร่วมมือ ความสำเร็จของการนำ AI มาใช้ต้องอาศัยความร่วมมือระหว่างรัฐบาล ภาคเอกชน และสถานศึกษา (Public-Private Partnership) [10] ในการสนับสนุนงบประมาณและโครงสร้างพื้นฐาน เช่น อินเทอร์เน็ตความเร็วสูงและ

อุปกรณ์ที่ทันสมัยเพื่อลดความเหลื่อมล้ำทางดิจิทัล (Digital Divide)




**ภาพที่ 1:** โมเดลแนวคิดการศึกษาปัจจัยที่ส่งผลต่อความสำเร็จในการนำ AI ไปใช้ในการจัดการเรียนการสอนของครูอาชีวศึกษา

## 3. วิธีการดำเนินงาน

โมเดลการวิจัยส่วนของตัวแปรอิสระมี 5 ตัวแปรได้แก่ (1) ความคาดหวังผลสัมฤทธิ์การสอน (Performance Expectancy: PE) (2) ความสอดคล้องนโยบายดิจิทัลของอาชีวศึกษา (Digital Policy Alignment: DA) (3) ศักยภาพการเรียนรู้เฉพาะบุคคลของครู (Personal Learning Capability: PLC) (4) การสนับสนุนเชิงยุทธศาสตร์จากผู้บริหาร (Strategic Support from Administrators: SS) (5) ความพร้อมของเทคโนโลยีและโครงสร้างพื้นฐาน (Technological Readiness: TR) และ 1 ตัวแปรตาม คือ ความสำเร็จในการนำ AI ไปใช้ในการจัดการเรียนการสอนของครูอาชีวศึกษา

การวิจัยนี้เป็นการวิจัยเชิงสำรวจ (Survey Research) โดยประชากร คือ ข้าราชการครู ครูอัตราจ้าง และพนักงานราชการทั่วไป ที่ปฏิบัติหน้าที่สอนในสถานศึกษาภาครัฐสังกัดสำนักงานคณะกรรมการการอาชีวศึกษา (สอศ.) จำนวนทั้งสิ้นประมาณ 30,000 คน ขนาดกลุ่มตัวอย่าง ตามสูตรของ Cochran ที่ระดับความเชื่อมั่นร้อยละ 95 และค่าความคลาดเคลื่อนที่ยอมรับได้ร้อยละ 5 ได้ขนาดกลุ่มตัวอย่างขั้นต่ำจำนวน 384 คน และเพื่อป้องกันความคลาดเคลื่อนจึงเก็บข้อมูลจริงได้จำนวน 447 คน ช่วงระหว่างเดือน มกราคม ถึงเดือน กุมภาพันธ์ 2569

เครื่องมือที่ใช้ในการสำรวจ คือ แบบสอบถาม ซึ่งผ่านการตรวจสอบความตรงเชิงเนื้อหา จากผู้เชี่ยวชาญจำนวน 3 ท่าน ผลการตรวจสอบพบว่าทุกข้อคำถามมีค่า IOC มากกว่า 0.50 และตรวจสอบความเชื่อมั่น ด้วยค่าสัมประสิทธิ์แอลฟาของครอนบาค พบว่ามีค่าแอลฟ่าทั้งฉบับเท่ากับ 0.970 ผลจากการเก็บรวบรวมข้อมูลจะถูกประมวลผลด้วยโปรแกรม SPSS

## 4. วิธีการดำเนินงาน

### 4.1 ข้อมูลทั่วไปและพฤติกรรมการใช้ AI

กลุ่มตัวอย่างกลุมตัวอย่างจำนวน 447 คน ส่วนใหญ่เป็นเพศหญิง (ร้อยละ 55.0) อายุ 30–39 ปี (ร้อยละ 41.4) จบการศึกษาระดับปริญญาตรี (ร้อยละ 78.3) มีประสบการณ์สอนในอาชีวศึกษา 5-9 ปี (ร้อยละ 33.1) และปฏิบัติหน้าที่ในตำแหน่งข้าราชการครู (ร้อยละ 63.8) ด้านพฤติกรรมการใช้ AI พบว่า กลุ่มตัวอย่างเลือกใช้เครื่องมือ AI ประเภท Chatbot เช่น ChatGPT, Gemini สูงที่สุด คิดเป็นร้อยละ 93.3 รองลงมาคือ AI ช่วยสร้างแบบฝึกหัดกิจกรรม (ร้อยละ 63.3) โดยส่วนใหญ่นำ AI ไปใช้ในการสอนจำนวน 3 รายวิชา (ร้อยละ 31.1) และมีความถี่ในการใช้งานอยู่ในระดับ ใช้เกือบทุกครั้งที่มีการเรียนการสอนสูงที่สุด ร้อยละ 36.9

### 4.2 ผลการวิเคราะห์การถดถอยพหุคูณ

ผลการตรวจสอบข้อตกลงเบื้องต้น ไม่พบปัญหาความสัมพันธ์กันเองของตัวแปรอิสระ (Multicollinearity) สามารถแสดงผลได้ดังตารางที่ 1

**ตารางที่ 1:** ผลการตรวจสอบปัญหาความสัมพันธ์กันเองของตัวแปรอิสระ (Multicollinearity)

| โมเดล | | Tolerance | VIF |
|---|---|---|---|
| 4 | PLC | 0.431 | 2.318 |
| | PE | 0.426 | 2.349 |
| | TR | 0.407 | 2.456 |
| | SS | 0.345 | 2.901 |

จากตารางที่ 1 ค่า Tolerance ทุกตัวสูงกว่า 0.1 และค่า VIF น้อยกว่า 5 แสดงให้เห็นว่าตัวแปรอิสระทั้ง 4 สามารถสร้างสมการณ์พยากรณ์ ไม่เกิดปัญหาความสัมพันธ์กันเองในระดับสูง เป็นไปตามข้อตกลงเบื้องต้นทางสถิติ และใช้ผลการวิเคราะห์การถดถอยพหุคูณ ด้วยวิธี Stepwise สรุปผลได้ดังตารางที่ 2

**ตารางที่ 2:** การสรุปผลการวิเคราะห์การถดถอยพหุคูณแบบขั้นตอน (Stepwise) เพื่อพยากรณ์ความสำเร็จในการนำปัญญาประดิษฐ์ (AI) ไปใช้ในการจัดการเรียนการสอนในอาชีวศึกษาต่อไปได้

| โมเดล | | b | $SE_b$ | β | t | p-value |
|---|---|---|---|---|---|---|
| 4 | PLC | 0.395 | 0.036 | 0.399 | 10.956 | 0.000 |
| | PE | 0.348 | 0.042 | 0.306 | 8.338 | 0.000 |
| | TR | 0.116 | 0.028 | 0.154 | 4.097 | 0.000 |
| | SS | 0.113 | 0.033 | 0.142 | 3.477 | 0.000 |

หมายเหตุ: * มีนัยสำคัญทางสถิติที่ระดับ 0.05 ค่าคงที่ (a) = 0.115 $R = 0.864$ $R^2 = 0.747$\$ $R^2_{adj} = 0.745$ $F = 326.361$ $SE_{est} = 0.275$

จากตารางที่ 2 พบว่า มี 4 ปัจจัยที่ส่งผลต่อความสำเร็จในการนำ AI ไปใช้อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 โดยสามารถร่วมกันพยากรณ์ความสำเร็จได้ร้อยละ 74.5 ($R^2_{adj} = 0.745$) สามารถเขียนเป็นสมการพยากรณ์ได้ดังนี้ (1) สมการพยากรณ์ในรูปแบบคะแนนดิบ $\hat{Y} = 0.115 + 0.395\ (X_3) + 0.348\ (X_1) + 0.116\ (X_5) + 0.113\ (X_4)$ (2) สมการพยากรณ์ในรูปแบบคะแนนมาตรฐาน $\hat{Z} = 0.399\ (X_3) + 0.306\ (X_1) + 0.154\ (X_5) + 0.142\ (X_4)$

**4.3 แนวทางการส่งเสริมการนำเทคโนโลยีปัญญาประดิษฐ์ไปใช้ในการจัดการเรียนการสอนในอาชีวศึกษา**

ผู้วิจัยสามารถสรุปผลแนวทางการส่งเสริมการนำเทคโนโลยีปัญญาประดิษฐ์ (AI) ไปใช้ในการจัดการเรียนการสอนในอาชีวศึกษาในรูปแบบของ กลไกการขับเคลื่อนแบบบูรณาการ 4 ด้าน ได้แก่ ศักยภาพการเรียนรู้เฉพาะบุคคลของครู ความคาดหวังผลสัมฤทธิ์การสอน ความพร้อมของเทคโนโลยีและโครงสร้างพื้นฐาน และการสนับสนุนเชิงยุทธศาสตร์จากผู้บริหาร ดังนี้

**4.3.1 การขับเคลื่อนด้านศักยภาพการเรียนรู้เฉพาะบุคคลของครู**

การส่งเสริมในมิตินี้มุ่งเน้นการเปลี่ยนผ่านทัศนคติของครูจากความกังวลต่อเทคโนโลยีไปสู่การทำงานร่วมกับ AI อย่างชาญฉลาด [1][2] โดยสถานศึกษาต้องจัดกระบวนการพัฒนาบุคลากรอย่างเป็นลำดับขั้นขั้น เริ่มจากการสร้างชัยชนะระยะสั้น ผ่านการใช้ AI ลดภาระงานธุรการ เพื่อสร้างความมั่นใจ [3] จากนั้นจึงยกระดับสู่การสร้างคลังความรู้เฉพาะทางวิชาชีพ และการบูรณาการ AI เข้ากับการออกแบบการเรียนรู้ฐานสมรรถนะซึ่งสอดคล้องกับแนวคิดที่ว่าความฉลาดรู้ด้าน AI และความเชื่อมั่นในตนเอง

**4.3.2 การขับเคลื่อนด้านความคาดหวังผลสัมฤทธิ์การสอน**

การที่ครูจะยอมรับและใช้งาน AI อย่างยั่งยืน จำเป็นต้องเห็นผลลัพธ์เชิงประจักษ์ว่า AI สามารถยกระดับคุณภาพการสอนได้จริง แนวทางการดำเนินการจึงต้องเริ่มจากการกำหนดเป็นฐานและเป้าหมายที่วัดผลได้ จากนั้นนำแพลตฟอร์มการเรียนรู้มาใช้เพื่อลดช่องว่างทางการเรียนรู้ของผู้เรียน [4][5] ควบคู่ไปกับการปฏิรูประบบการวัดและประเมินผล ที่เปลี่ยนจากการสอบท่องจำไปสู่การประเมินกระบวนการคิดและชิ้นงานนวัตกรรมทั้งนี้ ต้องมีการกำหนดมาตรการป้องกันการพึ่งพา AI มากเกินไปอย่างรัดกุม เพื่อให้ผู้เรียนอาชีวศึกษายังคงเกิดสมรรถนะทางวิชาชีพที่ตรงตามความต้องการของตลาดแรงงาน [6]

**4.3.3 การขับเคลื่อนด้านความพร้อมของเทคโนโลยีและโครงสร้างพื้นฐาน**

โครงสร้างพื้นฐานถือเป็นปัจจัย ที่ขาดไม่ได้ [6][7] การดำเนินการต้องครอบคลุมการยกระดับเครือข่ายอินเทอร์เน็ตความเร็วสูงให้ครอบคลุมทุกพื้นที่เรียนและโรงฝึกงาน การปรับนโยบายสู่ระบบคลาวด์ เพื่อลดข้อจำกัดด้านฮาร์ดแวร์ การจัดหาลิขสิทธิ์ซอฟต์แวร์ AI ระดับองค์กรตลอดจนการจัดตั้งหน่วยสนับสนุนทางเทคนิคที่ตอบสนองอย่างรวดเร็ว [8][9] และการปฏิบัติตามกฎหมาย PDPA เพื่อสร้างสภาพแวดล้อมดิจิทัลที่ปลอดภัย

**4.3.4 การขับเคลื่อนด้านการสนับสนุนเชิงยุทธศาสตร์จากผู้บริหาร**

ผู้บริหารระดับสูงทำหน้าที่เป็นผู้ส่งเสริมการดำเนินงานสำคัญที่สุดในการผลักดันให้เกิดวัฒนธรรมนวัตกรรม แนวทางที่ต้องดำเนินการคือการจัดทำแผนแม่บทดิจิทัลที่สอดคล้องกับยุทธศาสตร์ระดับชาติ[8] การสื่อสารวิสัยทัศน์อย่างโปร่งใส การจัดสรรงบประมาณเฉพาะสำหรับการวิจัยและนวัตกรรม รวมถึงการสร้างเครือข่ายความร่วมมือกับภาคอุตสาหกรรม [9] ที่สำคัญที่สุดคือการสร้าง ความปลอดภัยทางจิตวิทยาโดยส่งเสริมวัฒนธรรมที่ยอมรับความผิดพลาดจากการทดลองเรียนรู้ และมีการสร้างระบบแรงจูงใจและรางวัลเกียรติยศ เพื่อให้ครูกล้าที่จะก้าวออกจากพื้นที่ปลอดภัยและนำ AI มาใช้พัฒนาการสอนอย่างต่อเนื่อง [10] ในการนี้ จากแนวทางการส่งเสริมดังกล่าวเพื่อต้องการให้ครูผู้สอน รวมไปถึงผู้ที่เกี่ยวข้องนำไปสู่การปฏิบัติใน สถานศึกษาและชั้นเรียน

## 5. สรุปและอภิปรายผล

จากผลการวิเคราะห์ทางสถิติร่วมกับวรรณกรรมที่ผู้วิจัยได้ศึกษาขอเสนอการอภิปรายผลโดยการศึกษาเชิงเปรียบเทียบกับงานวิจัยที่เกี่ยวข้องได้ ดังนี้

**5.1 ศักยภาพการเรียนรู้เฉพาะบุคคลของครู**

พบว่าเป็นปัจจัยที่มีอิทธิพลสูงสุด ต่อความสำเร็จสะท้อนว่าครูคือแกนกลางสำคัญในการบูรณาการ AI เข้ากับเนื้อหาวิชาชีพ โดยมีความสอดคล้อง กับผลงานของ [11] ที่พบว่า ความรู้ด้าน AI (AI Literacy) เป็นตัวแปรทำนายที่แข็งแกร่งที่สุดในบริบทครูและสอดคล้องกับ [12] ที่ระบุว่าความรู้ความสามารถในการใช้ส่งผลโดยตรงต่อความตั้งใจใช้งาน ในส่วนที่เหนือกว่า งานวิจัยนี้ขยายขอบเขตความรู้เดิมโดยระบุเจาะจงไปที่ ศักยภาพการเรียนรู้เฉพาะบุคคล ในบริบทอาชีวศึกษาซึ่งต่างจากงานทั่วไปที่เน้นเพียงทักษะทางเทคนิค งานวิจัยนี้ชี้ให้เห็นว่าครูอาชีวะไม่ได้ต้องการแค่รู้ว่า AI ทำงานอย่างไร แต่ต้องมีความสามารถในการ บูรณาการ (Integrate) เพื่อออกแบบการสอนรายบุคคล ซึ่งเป็นมิติที่ลึกซึ้งกว่างานของ [13]

**5.2 ความคาดหวังผลสัมฤทธิ์**

เป็นปัจจัยที่มีอิทธิพลรองลงมา ผลยืนยันว่าครูจะใช้ AI เมื่อมั่นใจว่ามันช่วยยกระดับผลลัพธ์ผู้เรียนได้จริง ซึ่งสนับสนุน ทฤษฎี UTAUT2 และงานวิจัยของ [11] ที่ยืนยันว่า Performance Expectancy เป็นปัจจัยหลักในการยอมรับเทคโนโลยีในภาคการศึกษา ทั้งพบส่วนที่แตกต่าง จากของ [12] ที่พบว่าในบางบริบท ความง่ายในการใช้งาน (Effort Expectancy) อาจสำคัญกว่า แต่ในบริบทอาชีวศึกษาของไทย งานวิจัยนี้พิสูจน์แล้วว่า ประโยชน์ที่จับต้องได้ มีน้ำหนักมากกว่าความยากง่าย สอดคล้องกับ [13] ที่พบว่าการใช้ AI เพื่ออำนวยความสะดวกและเพิ่มประสิทธิภาพได้เป็นปัจจัยสำคัญ ซึ่งชี้ให้เห็นว่าครูอาชีวะยอมรับความซับซ้อนได้หากผลลัพธ์คุ้มค่า

**5.3 ความพร้อมของโครงสร้างพื้นฐาน**

มีผลการวิจัยสอดคล้อง ที่ระบุว่าโครงสร้างพื้นฐานเป็นสิ่งจำเป็นสำหรับการสร้างนวัตกรรมในโรงเรียนอาชีวศึกษาจาก [14] ที่พบว่า Facilitating Conditions ส่งผลต่อพฤติกรรม แม้จะมีอิทธิพลน้อยกว่าปัจจัยด้านบุคคล แต่ยังคงมีนัยสำคัญ ซึ่งต่างจากบางงานวิจัยในประเทศพัฒนาแล้วที่อาจมองข้ามปัจจัยนี้เนื่องจากมีความพร้อมอยู่แล้ว งานวิจัยนี้จึงสะท้อนบริบทของ ประเทศกำลังพัฒนา ที่ความเหลื่อมล้ำทางดิจิทัล ยังเป็นประเด็นสำคัญ จากสิ่งที่รายงานได้เสนอไว้

**5.4 ความสอดคล้องนโยบายดิจิทัลของอาชีวศึกษา**

ไม่มีนัยสำคัญทางสถิติและถูกตัดออกจากสมการพยากรณ์ การมีนโยบายเพียงอย่างเดียวไม่เพียงพอ แต่ผู้บริหารสถานศึกษาจำเป็นต้องแปลงนโยบายส่วนกลางให้เป็นวิสัยทัศน์และแนวทางปฏิบัติที่ชัดเจน เพื่อให้บุคลากรเกิดความเข้าใจและมองเห็นแนวทางในการนำไปประยุกต์ใช้ [13] นอกจากนี้ พบที่ว่านโยบายที่สนับสนุนไม่มีนัยสำคัญทางสถิตินั้น ยังสอดคล้องกับงานวิจัยในระดับสากลหลายฉบับที่พบว่า ปัจจัยด้านโครงสร้างนโยบาย มักไม่มีอิทธิพลทางตรงต่อความตั้งใจใช้งานของครูหากนโยบายเหล่านั้นไม่ได้แก้ปัญหาในวิถีการทำงานประจำวัน หรือขาดการบูรณาการให้เข้ากับบริบทการสอนจริง [14]

สรุป ความสำเร็จในการประยุกต์ใช้เทคโนโลยีปัญญาประดิษฐ์ในการจัดการเรียนการสอนอาชีวศึกษายุค

ใหม่ ไม่ใช่ปรากฏการณ์เชิงเส้นตรงที่เกิดขึ้นได้เพียงเพราะสถานศึกษามีเทคโนโลยีที่ทันสมัยที่สุด ทว่ามันคือปฏิกิริยาของความสัมพันธ์เชิงระบบ ที่เกิดจากการประสานกันอย่างลงตัวระหว่าง องค์ประกอบด้านบุคคลทัศนคติและความฉลาดรู้ของครู องค์ประกอบด้านเทคโนโลยี ความพร้อมของโครงสร้างพื้นฐาน และองค์ประกอบด้านองค์กรการบริหารจัดการเชิงยุทธศาสตร์ จึงก่อให้เกิดการผลักดันการใช้ในวงการอาชีวศึกษาได้อย่างยั่งยืน

ผลการวิจัยนี้จึงเสนอในอีกมุมมองจาก งานวิจัยอื่นในแง่ของการเสนอปัญหาหน้างานว่าการมีนโยบายเพียงอย่างเดียวไม่เพียงพอ หากขาดการ แปลงนโยบาย สู่แนวปฏิบัติที่ชัดเจนนั้น จะไม่สามารถทำให้ครูผู้สอนสามารถนำเทคโนโลยีดังกล่าวไปใช้ และพัฒนาตัวครูผู้สอนเองให้สามารถนำไปประกอบการจัดการเรียนการสอนได้อย่างเต็มประสิทธิภาพ

## เอกสารอ้างอิง


[1] Jaya, H., Haryoko, S., Anwar, B., Lu'mu, & Kasau, S. (2024). Integration of artificial intelligence in learning at vocational schools: Building sustainable collaboration for educational innovation in the Industrial Era 4.0. https://doi.org/10.2991/978-2-38476-335-1_12

[2] อธิป เกตุสิริ, ณัฐยา เกตุสิริ, ศุภชัย เส้นเกษ, & วารุณี เรียนรู้. (2568). การใช้เทคโนโลยีปัญญาประดิษฐ์ (AI) เพื่อการจัดการเรียนการสอนในศตวรรษที่ 21. *วารสารนวัตกรรมการวิจัยเพื่อสังคม*, *1*(1), 51-64.

[3] Kertsawat, W., Homchan, S., & Naruechokekittikirati, K. (2024). Supporting factors affecting work performance of staff in public and private organizations in the AI era. Journal of Interdisciplinary Research and Educational Innovation, 3(2), 31–50.

[4] ภคอร จตุพรธนภัทร. (2567). นโยบายการใช้นวัตกรรมปัญญาประดิษฐ์เพื่อเพิ่มขีดความสามารถในการแข่งขันของประเทศ. *วารสารการบริหาร การจัดการ และการพัฒนาที่ยั่งยืน*, *2*(2), 331-344.

[5] Soeykrathoke, P., Sukwiphat, P., Saengkaew, P., Korthavat, V., & Phakamach, P. (2025). Strategic Alignment for AI Success: Vocational Education in Thailand's Eastern Economic Corridor. *Journal of Education and Learning Reviews*, 2(4), 43-62.

[6] Berigel, D. S., Şilbir, L., & Şilbir, G. M. (2025). Integrating artificial intelligence (AI) into technical and vocational education and training (TVET): A PRISMA-based systematic review. *Revista Calitatea Vieţii*.

[7] Treve, M. (2024). Integrating artificial intelligence in education: Impacts on student learning and innovation. *International Journal of Vocational Education and Training Research*, *10*(2), 61-69. https://doi.org/10.11648/j.ijvetr.20241002.14

[8] อธิป เกตุสิริ, ณัฐยา เกตุสิริ, ศุภชัย เส้นเกษ, & วารุณี เรียนรู้. (2568). การใช้เทคโนโลยีปัญญาประดิษฐ์ (AI) เพื่อการจัดการเรียนการสอนในศตวรรษที่ 21. *วารสารนวัตกรรมการวิจัยเพื่อสังคม*, *1*(1), 51-64.

[9] Wang, J., & Zhang, C. (2025). Aiming to enhance higher education practice through generative AI integration: A theoretical exploration of critical success factors. *SAGE Open*, April-June 2025, 1-19.

[10] Pakdee, S., & Phetmalhkul, T. (2024). Guidelines for application Artificial Intelligence (AI) in education under the Secondary Educational Service Area Office Bangkok 2. Interdisciplinary Academic and Research Journal, https://doi.org/10.60027/iarj.2024.278760

[11] Al-Amri, N. A., & Al-Abdullatif, A. M. (2024). Drivers of chatbot adoption among K-12 teachers in Saudi Arabia. *Education Sciences*, *14*(1), 1-19. https://doi.org/10.3390/educsci14010001

[12] Moradi, H. (2025). Integrating AI in higher education: Factors influencing ChatGPT acceptance among Chinese university EFL students. International Journal of Educational Technology in Higher Education, 22, Article 30. https://doi.org/10.1186/s41239-025-00530-4

[13] คุณาพร โฉมจิตร, & รัชกฤช ปัทมโสภาสกุล. (2568). ปัจจัยที่ส่งผลต่อการใช้ปัญญาประดิษฐ์เพื่อ การเรียนรู้ของผู้เรียนในวิทยาลัยอาชีวศึกษาเอกชน. *Journal of Social Science and Cultural*, *9*(2), 380-395.

[14] Wangdi, T., Dhendup, S., & Gyelmo, T. (2023). Factors influencing teachers' intention to use technology: Role of TPACK and facilitating conditions. International Journal of Instruction

# การตรวจจับของต้องห้ามจากสื่อสังคมออนไลน์ด้วยเทคนิคเหมืองข้อความเพื่อสนับสนุนงานศุลกากร

# Text Mining-Based Detection of Prohibited Goods from Social Media to Support Customs Authorities

*ณัชชารีย์ อริยนิธิทวีสุข (Natcharee Ariyanithitawesuk)* [1] *และธนพล เจนสุทธิเวชกุล (Tanapon Jensuttiwetchakul)* [2]
*ภาควิชาการจัดการเทคโนโลยีสารสนเทศ*
*คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ*
[1]*s6707021857058@email.kmutnb.ac.th,* [2] *Tanapon.j@itd.kmutnb.ac.th*

## บทคัดย่อ

*การศึกษาวิจัยเชิงประยุกต์นี้มุ่งเปรียบเทียบประสิทธิภาพอัลกอริทึมจำแนกข้อความเพื่อตรวจจับการค้าสินค้าต้องห้ามบนแพลตฟอร์ม Twitter จากชุดข้อมูลสมดุล 2,000 รายการ โดยประยุกต์ใช้เทคนิค NLP (Hybrid Tokenization) ร่วมกับวิธี TF-IDF ผ่านสถาปัตยกรรมจำแนกข้อมูล 2 ขั้นตอน เพื่อทดสอบ 4 แบบจำลอง ได้แก่ Naïve Bayes, Random Forest, Logistic Regression และ Linear SVM ผลการวิจัยขั้นตอนที่ 1 การคัดกรองสินค้าต้องห้าม พบว่าแบบจำลองนาอีฟเบย์แสดงประสิทธิภาพดีที่สุดตามตัวชี้วัด (Accuracy, Precision, Recall, F1-score) ที่ระดับ 0.9921 สำหรับขั้นตอนที่ 2 การจำแนกประเภทยาเสพติด พบว่าแบบจำลองการป่าสุ่มประสิทธิภาพดีที่สุด โดยให้ค่าความถูกต้อง (Accuracy) และความระลึก (Recall) ที่ 0.9190 ความแม่นยำ (Precision) 0.9170 และความถ่วงดุล (F1-score) 0.9110 นอกจากนี้ ผู้วิจัยได้พัฒนาเทคนิค Weighted Keyword Scoring ร่วมกับกฎเชิงบริบท (Contextual Rules) เพื่อขจัดข้อจำกัดการตรวจจับรหัสแฝงและ Emoji ซึ่งสามารถนำไปบูรณาการเป็นเครื่องมือเฝ้าระวังเชิงรุก (Proactive Surveillance) เพื่อสนับสนุนการปฏิบัติงานของกรมศุลกากรได้อย่างมีประสิทธิภาพ*

**คำสำคัญ:** เหมืองข้อความ, การประมวลผลภาษาธรรมชาติ, สินค้าต้องห้าม, กรมศุลกากร, นาอีฟเบย์, การป่าสุ่ม

## Abstract

*This study investigates text classification methods for detecting prohibited goods on X (formerly Twitter) using a balanced dataset of 2,000 records. A two-stage architecture incorporating hybrid tokenization and TF-IDF was employed to evaluate Naïve Bayes, Random Forest, Logistic Regression, and Linear SVM. In Stage 1 (screening), Naïve Bayes achieved the highest performance (0.9921 across all metrics). In Stage 2 (drug categorization), Random Forest outperformed other models, with 0.9190 accuracy and recall, 0.9170 precision, and an F1-score of 0.9110. Additionally, a weighted keyword scoring approach with contextual rules was developed to address limitations in detecting coded language and emojis, thereby enhancing its utility as a proactive surveillance tool for Customs operations.*

**Keywords**: Text Mining, Natural Language Processing, Prohibited Goods, Customs Department, Naive Bayes, Random Forest.

## 1. บทนำ

สื่อสังคมออนไลน์แพลตฟอร์ม Twitter ถูกนำมาใช้เป็นช่องทางลักลอบจำหน่ายยาเสพติดและบุหรี่ไฟฟ้าซึ่งเป็นของต้องห้ามตามกฎหมายศุลกากร ทั้งนี้ผู้กระทำผิดใช้รหัสแฝง อิโมจิ เพื่อหลีกเลี่ยงการกระทำผิด ทำให้การตรวจสอบโดยบุคลากรนั้นไม่เพียงพอต่อการรับมือในสื่อสังคมออนไลน์และพฤติกรรมเพื่อปิดช่องโหว่จึงประยุกต์ใช้เทคนิคการทำเหมืองข้อความและอัลกอริทึมการเรียนรู้ของเครื่องด้วยสถาปัตยกรรมจำแนกแบบ 2 ขั้นตอน (Two-Stage Classification) ประกอบด้วย 1) การคัดกรองเพื่อแยกกลุ่มสินค้าต้องห้ามออกจากกลุ่มสินค้าทั่วไป และ 2) การจำแนกประเภทสินค้าต้องห้ามเชิงลึกอย่างเจาะจง (เช่น ยาบ้า, กัญชา, บุหรี่ไฟฟ้า)

ทั้งนี้ยังได้ประยุกต์ใช้ทฤษฎีการให้คะแนนคำสำคัญแบบถ่วงน้ำหนัก (Weighted Keyword Scoring) ร่วมกับการใช้กฎเชิง

บริบท เพื่อก้าวข้ามข้อจำกัดของรหัสแฝงและอิโมจิ ซึ่งเป็นประเด็นที่ท้าทายของงานวิจัยก่อนหน้ามักไม่ให้ความสำคัญ ผลการศึกษาในครั้งนี้สามารถนำไปประยุกต์ใช้เป็นกลไกหลักในการขับเคลื่อนระบบเครื่องมือเฝ้าระวังเชิงรุกเพื่อสนับสนุนการปฏิบัติงานและลดภาระของเจ้าหน้าที่กรมศุลกากรได้อย่างมีประสิทธิภาพ

## 2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ผู้วิจัยได้ศึกษาค้นคว้าเอกสาร ทฤษฎี และงานวิจัยที่เกี่ยวข้องต่าง ๆ เพื่อนำไปประยุกต์ใช้ในงานวิจัย ดังต่อไปนี้

### 2.1 การทำเหมืองข้อความ (Text Mining)

Text Mining คือกระบวนการแปลงข้อมูลข้อความที่ไม่มีโครงสร้างให้เป็นข้อมูลเชิงโครงสร้างเพื่อการประมวลผลของคอมพิวเตอร์ โดยผ่านการรวบรวมและเตรียมข้อมูลด้วยเทคนิคการตัดคำแบบลูกผสมระหว่าง newmm และ deepcut เพื่อระบุรหัสแฝงให้แม่นยำ และใช้เทคนิค TF-IDF แปลงข้อความเป็นตัวเลขเพื่อกำหนดน้ำหนักคำเชิงสถิติ แม้งานวิจัยที่เกี่ยวข้อง (อาทิ ชวรัส[1]ใช้ Random Forest พยากรณ์การเบิกสิทธิประกันสังคม ปรางค์ปราณี[2] ใช้ Naïve Bayes จำแนกการช่วยเหลือเด็ก อุลัยลักษ์ [3] ใช้ Logistic Regression จำแนกปัญหาไอที และสิริวุฒิ[4]ใช้ Linear SVM จำแนกพิกัดศุลกากร) มีการใช้การเรียนรู้ของเครื่อง แต่งานวิจัยนี้ได้สร้างความแตกต่างเชิงนวัตกรรมด้วยสถาปัตยกรรม 2 ขั้นตอนร่วมกับการวิเคราะห์ข้อความแฝงและสัญลักษณ์อีโมจิ นอกจากนี้ ยังประยุกต์ใช้กฎเชิงบริบทและการให้คะแนนแบบถ่วงน้ำหนักอิงพจนานุกรม เพื่อตีความรหัสแฝงและเจตนาที่ซ่อนอยู่ ซึ่งเป็นกลไกสำคัญในการปิดช่องโหว่ทางภาษาและเพิ่มสมรรถนะการเฝ้าระวังได้อย่างมีประสิทธิภาพ

### 2.2 เทคนิคการจำแนกข้อมูลและการเรียนรู้ของเครื่อง

2.2.1 Naïve Bayes เป็นการจำแนกข้อมูลเชิงความน่าจะเป็นที่รวดเร็วและประหยัดทรัพยากร ซึ่งประสิทธิภาพจะขึ้นอยู่กับความสมดุลของชุดข้อมูล โดยมีสมการดังนี้

$$P(C|X) = \frac{(P(X|C) \cdot P(C))}{P(X)}$$

2.2.2 Logistic Regression (LR) เป็นการประยุกต์ใช้ฟังก์ชันซิกมอยด์เพื่อคำนวณความน่าจะเป็นช่วง 0 -1 มีจุดเด่นด้านความเร็วและการตีความค่าน้ำหนักคุณลักษณะ แต่มีข้อจำกัดต่อความไวค่าผิดปกติและความสัมพันธ์ระหว่างตัวแปร โดยมีสมการพื้นฐานดังนี้

$$P(y = 1 \mid x) = \frac{1}{1 + e^{-z}} \text{ โดยที่ } z = w \cdot x + b$$

2.2.3 Random Forest (RF) เป็นกลุ่มผสานการทำงานของต้นไม้ตัดสินใจจำนวนมากผ่านการสุ่มตัวอย่างค่าเฉลี่ยที่ดีที่สุดและคุณลักษณะเพื่อลดความแปรปรวน สามารถทนทานต่อสัญญาณรบกวนและค่าผิดปกติได้ดีแต่ใช้ทรัพยากรสูง โดยมีสมการพื้นฐานดังนี้

$$\hat{Y} = \text{mode}\{h_{1(x)}, h_{2(x)}, \dots, h_{n(x)}\}$$

2. 2.4 Support Vector Machine (Linear SVM) เป็นการสร้างระนาบแบ่งสามารถจัดการข้อมูลมิติสูงและประหยัดหน่วยความจำแต่มีความไวต่อสัญญาณรบกวน โดยมีสมการพื้นฐานดังนี้

$$w \cdot x + b = 0$$

### 2.3 การประเมินประสิทธิภาพ

2.3.1 เมทริกซ์ความสับสน (Confusion Matrix) เป็นการเปรียบเทียบผลการทำนายกับข้อมูลจริงมี 4 สถานะ ได้แก่ ทำนายสินค้าต้องห้ามถูก (TP) ทำนายสินค้าทั่วไปถูก (TN) ทำนายสินค้าทั่วไปผิดเป็นของต้องห้าม (FP) และ ทำนายของต้องห้ามผิดเป็นสินค้าทั่วไป (FN)

2.3.2 ตัวชี้วัดประสิทธิภาพเชิงสถิติ

1) ค่าความถูกต้อง (Accuracy)

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

2) ค่าความเที่ยงตรง(Precision)

$$Precision = \frac{TP}{TP + FP}$$

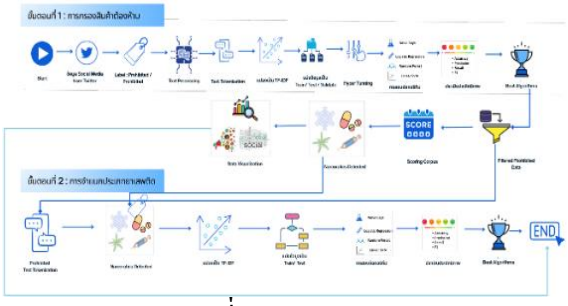
3) ค่าระลึก (Recall)

$$Recall = \frac{TP}{TP + FN}$$

4) ค่าความสมดุลระหว่างค่าความเที่ยงตรงและค่าระลึก (F1-Score)

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

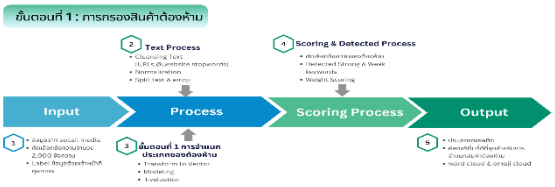
## 3. วิธีการดำเนินการวิจัย



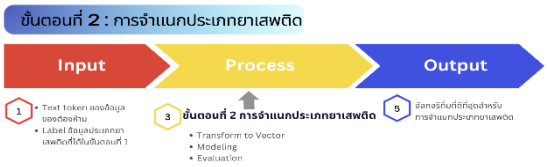
**ภาพที่ 1** : กรอบแนวคิดงานวิจัย

### 3.1 การสร้างกรอบแนวคิดของงานวิจัย

ผู้วิจัยกำหนดสถาปัตยกรรมระบบแบบสองขั้นตอน โดยเริ่มจากการเตรียมข้อมูลและใช้การเรียนรู้ของเครื่องเพื่อคัดกรองสินค้าต้องห้ามออกจากสินค้าทั่วไปในขั้นตอนที่ 1 จากนั้นจึงนำข้อมูลที่ผ่านการคัดกรองเข้าสู่ระบบให้คะแนนคำสำคัญและแบบจำลองเชิงลึกเพื่อระบุประเภทยาเสพติดในขั้นตอนที่ 2 ทั้งนี้การแบ่งขั้นตอนดังกล่าวมีวัตถุประสงค์เพื่อลดความซับซ้อนของโครงสร้างข้อมูลและแก้ไขปัญหาความไม่สมดุลของคลาสซึ่งส่งผลให้การจำแนกมีความแม่นยำและมีประสิทธิภาพสูงขึ้น




**ภาพที่ 2** : แสดง input output ในขั้นตอนที่ 1




**ภาพที่ 3** : แสดง input output ในขั้นตอนที่ 2

### 3.2 การเก็บรวบรวมข้อมูล



**ภาพที่ 4** : ภาพแสดงตัวอย่างลักษณะของ Dataset

ผู้วิจัยดำเนินการรวบรวมชุดข้อมูลข้อความสาธารณะเชิงพาณิชย์จากแพลตฟอร์ม Twitter ระหว่างเดือนช่วงเดือน ม.ค. 2564 ถึง ก.พ. 2565 จำนวน 2,000 ข้อความผ่านเทคนิค Web Scraping จากนั้นได้ดำเนินการ Label ข้อมูลโดยอ้างอิงตามคำนิยามของต้องห้ามและเจ้าหน้าที่กรมศุลกากร ข้อมูลที่รวบรวมมีลักษณะเป็นข้อมูลที่ไม่มีโครงสร้าง มีความซับซ้อนประกอบด้วยคำสแลง Emoji ภาษาไม่เป็นทางการสัญญาณรบกวน รวมถึงกลวิธีทางภาษาที่จงใจสร้างขึ้นเพื่อหลบเลี่ยงการตรวจจับ ทั้งนี้ เพื่อให้ชุดข้อมูลมีความสมดุล ผู้วิจัยจึงกำหนดสัดส่วนข้อมูลแบ่งเป็นกลุ่มสินค้าต้องห้าม 1,000 รายการ และกลุ่มสินค้าทั่วไป 1,000 รายการ

### 3.3 การจัดเตรียมข้อมูล

```python
# --- Emoji Part ---
def split_emoji_from_text(text: str) -> str:
    emoji_pattern = regex.compile(r'(\p{Emoji_Presentation}|\p{Emoji}\uFE0F)
    text = emoji_pattern.sub(r' \1 ', text)
    return regex.sub(r'\s+', ' ', text).strip()
# --- Clean Part ---
def clean_text(text: str):
    """Clean unwanted characters while keeping #word"""
    text = normalize(text)
    text = regex.sub(r"#(\w+)", r"\1", text)
    text = regex.sub(r"https?://\S+", "", text)
    text = regex.sub(r"www.\S+", "", text)
    text = regex.sub(r"@\w+", "", text)
    text = split_emoji_from_text(text)
    text = text.replace("/", " ")
    text = regex.sub(r"[.,\-—/|()]", " ", text)  # จุด/ขีด/สแลช/วงเล็
    text = regex.sub(r"\s+", " ", text).strip()
    return text
```

**ภาพที่ 5** : แสดงการจัดเตรียมข้อมูลและการแยก Emoji

1. การทำความสะอาดด้วยการลบ URLs และ @mention แต่เก็บ Emoji เอาไว้เพราะผู้ขายสินค้าใช้เป็นรหัสแทนชื่อสินค้า ดังแสดงในภาพที่ 5

2. กระบวนการตัดคำใช้ฟังก์ชันจาก PyThaiNLP โดยใช้เครื่องมือ Attacut และ Newmm เพื่อตัดคำภาษาไทยและคำแสลง

3. Emoji Extraction ผู้วิจัยได้ประยุกต์ใช้ Regular Expression ในการแยก Emoji ออกจากข้อความไว้เป็น Feature ดังแสดงในภาพที่ 5

4. การแปลงข้อมูลเป็นเวกเตอร์ใช้เทคนิค TF-IDF และแบ่งข้อมูลเป็น ชุดฝึกสอน (70%) ชุดตรวจสอบ (15%) และชุดทดสอบ (15%) ในขั้นตอนที่ 1 ส่วนขั้นตอนที่ 2 แบ่งชุดฝึกสอน (70%) ชุดตรวจสอบ (30%)

### 3.4 โมเดลการจำแนก (Classification Models)

ผู้วิจัยเปรียบเทียบประสิทธิภาพของอัลกอริทึม 4 รูปแบบ ได้แก่ Naïve Bayes Random Forest Logistic Regression และ Linear SVM โดยประยุกต์ใช้ Grid Search ร่วมกับการทดสอบแบบไขว้ 5 ส่วน เพื่อหาพารามิเตอร์ที่เหมาะสมและวัดประสิทธิภาพแบบจำลองผ่าน Accuracy Precision Recall และ F1-score เพื่อคัดเลือกโมเดลที่มีความแม่นยำที่สุด

### 3.5 การให้คะแนนคำสำคัญและกฎเชิงบริบท (Keyword Scoring & Contextual Rules)



**ภาพที่ 6** : ภาพตัวอย่างแสดงคำบ่งชี้ชัดเจน (Strong Keywords)



**ภาพที่ 7** : ภาพตัวอย่างแสดงคำบ่งชี้บริบท(Weak Keywords)

1. ผู้วิจัยสร้างพจนานุกรมแบ่งเป็นคำบ่งชี้ชัดเจน Strong Keywords ดังภาพที่ 6 เช่น ยาบ้า บุหรี่ไฟฟ้า และ คำบ่งชี้บริบท Weak Keywords ดังภาพที่ 7 เช่น น้ำแข็ง, ตัวเล็ก, ตัวใหญ่ เนื่องจากผู้ขายอาจใช้คำแสลงจาก กนิษฐา [5] ในการจำหน่ายสินค้าต้องห้ามบนแพลตฟอร์มออนไลน์

2. ส่วนที่สำคัญของงานวิจัยนี้ คือการประยุกต์ใช้แนวคิด การตัดสินใจแบบหลายหลักเกณฑ์ (MCDA) ภายใต้กระบวนการวิเคราะห์เชิงพจนานุกรม (Lexicon-based Approach) โดยระบบจะดำเนินการคำนวณคะแนนรวม จากการสะสมน้ำหนักของคำสำคัญและ Emoji เทคนิคดังกล่าวถูกนำมาบูรณาการในขั้นตอนที่ 2 เพื่อช่วยในการจำแนกประเภทสินค้าแม่นยำขึ้น โดยรูปแบบสมการพื้นฐาน ดังนี้

$$Score_{total} = (N_{strong} \times W_{strong} \times P_{cat}) + (N_{weak} \times W_{weak}) + Bonus$$

โดยที่

$N_{strong}$ , $N_{weak}$ คือ จำนวนคำ Strong และ Weak ที่ตรวจพบในข้อความ

$W_{strong}$ , $W_{weak}$ คือ น้ำหนักของคำ Strong และ Weak (กำหนดให้ Strong เป็น 3 และ Weak เป็น 1)

$P_{cat}$ คือ ค่าลำดับความสำคัญของหมวดหมู่

$Bonus$ คือ คะแนนพิเศษกรณีพบ คำชี้ชัด และ คำบริบท ปรากฏร่วมกัน

## 4. ผลการดำเนินงาน

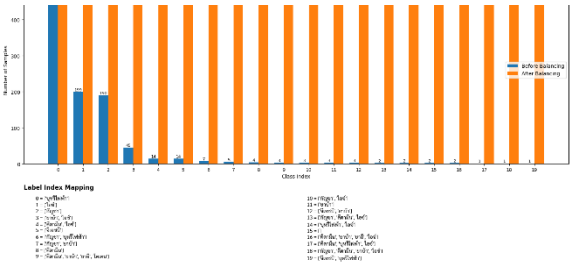
### 4.1 ผลการจำแนกของต้องห้ามกับสินค้าทั่วไป

**ตารางที่ 1** : ประสิทธิภาพอัลกอริทึมในขั้นตอนที่ 1

| Algorithm | Accuracy | Precision | Recall | F1 |
|---|---|---|---|---|
| Naive Bayes | 0.9921 | 0.9921 | 0.9921 | 0.9921 |
| Random Forest | 0.9914 | 0.9915 | 0.9914 | 0.9914 |
| Logistic Regression | 0.9914 | 0.9914 | 0.9914 | 0.9914 |
| Linear SVM | 0.9906 | 0.9907 | 0.9906 | 0.9906 |

ตารางที่ 1 แบบจำลอง Naïve Bayes มีประสิทธิภาพสูงสุด 0.9921 ในทุกตัวชี้วัด เนื่องจากข้อสมมติฐานความเป็นอิสระของคุณลักษณะสอดคล้องกับโครงสร้างข้อมูล ทำให้มีความเสถียรและลดความเสี่ยงการเกิด Overfitting ได้ดีกว่า Random Forest ซึ่งผลลัพธ์นี้ยังสะท้อนถึงคุณภาพของกระบวนการจัดเตรียมข้อมูลเบื้องต้นอีกด้วย

### 4.2 ผลการจำแนกประเภทสินค้าต้องห้าม



**ภาพที่ 8** : ภาพแสดงปรับข้อมูลประเภทยาเสพติดที่ไม่สมดุล

ผู้วิจัยได้ทำการปรับสมดุลของประเภทยาเสพติดด้วยเทคนิค SMOTE ก่อนทำการประเมินประสิทธิภาพในขั้นตอนที่ 2

**ตารางที่ 2** : ผลการเปรียบเทียบประสิทธิภาพในขั้นตอนที่ 2

| Algorithm | Accuracy | Precision | Recall | F1 |
|---|---|---|---|---|
| Random Forest | 0.9190 | 0.9170 | 0.9190 | 0.9110 |
| Logistic Regression | 0.9040 | 0.9230 | 0.904 | 0.9090 |
| Linear SVM | 0.8810 | 0.9140 | 0.8810 | 0.8910 |
| Naive Bayes | 0.8930 | 0.8740 | 0.8930 | 0.8760 |

จากตารางที่ 2 แสดงผลการจำแนกประเภทยาเสพติด พบว่า Random Forest มีประสิทธิภาพสูงสุด โดยให้ค่า Accuracy และ Recall ที่ 0.9190 Precision ที่ 0.9170 และ F1-score ที่ 0.9110 ปัจจัยสำคัญเกิดจากการทำงานร่วมกับเทคนิค SMOTE ซึ่งช่วยให้แบบจำลองสามารถจัดการกับข้อมูลที่ซับซ้อน ไร้โครงสร้าง และทนทานต่อปัญหาความไม่สมดุลของปริมาณข้อมูลในแต่ละประเภทได้อย่างมีประสิทธิภาพ

### 4.3 แสดงผลการทำ Weighted Keyword Scoring

```
================================================================================
WITH vs WITHOUT WEIGHTED KEYWORD SCORING
================================================================================
                                Method  Accuracy  Precision_weighted  Recall_weighted  F1_weighted  Num_Samples
1  With Weighted Keyword Scoring  0.875250  0.885499  0.875250  0.879734  1002
A  Without Weighted Keyword Scoring  0.525948  0.895139  0.525948  0.659191  1002
================================================================================
ABLATION STUDY RESULTS
================================================================================
                                Method  Accuracy  Precision_weighted  Recall_weighted  F1_weighted
0  Full Method  0.875250  0.885499  0.875250  0.879734
2  No Bonus  0.875250  0.885499  0.875250  0.879734
3  No Override  0.875250  0.885499  0.875250  0.879734
5  No Global Threshold  0.864271  0.885216  0.864271  0.873789
4  No Priority  0.860279  0.882196  0.860279  0.870173
1  Baseline (No Weight/Bonus/Priority/Override)  0.525948  0.895139  0.525948  0.659191
```

**ภาพที่ 10** วัดผลการทำ ABLATION STUDY

| | text_tokens | true_narcotics | pred_baseline | pred_full |
|---|---|---|---|---|
| | [โค้ก, พร้อมส่ง, ไทย, 30g, สนใจ, ทัก, คลิ๊กลิง... | คีตามีน, โคเคน, ยาบ้า, ยาอี | คีตามีน\|บุหรี่ไฟฟ้า\|ยาบ้า\|ยาอี\|โคเคน\|ไอซ์ | คีตามีน\|ยาบ้า\|ยาอี\|โคเคน |
| | [, พอด, ดูด, ทั้ง, fof, , ราคา, 200, ฿, , ค... | บุหรี่ไฟฟ้า | บุหรี่ไฟฟ้า\|ยาอี | บุหรี่ไฟฟ้า |
| | [ด่วน, 1, ขั้วโบง, คุณภาพ, ดู, ลูกค้า, ร้, วิว... | ไอซ์ | ยาบ้า\|ไอซ์ | ไอซ์ |
| | [งาน, ขาเขียว, นัด, ขนมือ, ADD, Line, or, DM, ... | ไอซ์ | ยาบ้า\|ไอซ์ | ไอซ์ |
| | [ลดล, , , 1G, 1000, 2G, 1499, 3G, 1999,... | ไอซ์ | คีตามีน\|ยาบ้า\|ไอซ์ | คีตามีน\|ไอซ์ |

**ภาพที่ 11** ผลการแสดงผลการจัดกลุ่มประเภทยาเสพติด

จากภาพที่ 9-11 แม้ผลทดสอบ Weighted Keyword Scoring จะไม่ต่างกันมากนัก แต่หากขาดกระบวนการนี้ อาจทำให้การทำนายกลุ่มสินค้าทั่วไปคลาดเคลื่อนได้

### 4.4 ตัวอย่างการวิเคราะห์จริง (Case Study)

```
================================================================================
TEXT: มาครับ ❄️ G 850 5 G 3500 🦄 เส้น 750 100 🦄 3500 200 🦄 600 ส่งจริง ส่ง
TOKENS: ['❄', '', 'G', '850', '5', 'G', '3500', '🦄', 'เส้น', '750', '100',
การจำแนกในขั้นตอนที่ 1 : Is Prohibited True
การจำแนกในขั้นตอนที่ 2 : Detected ['ยาบ้า', 'ไอซ์']
Strong Keywords: ['🦄', 'งาน', 'ไฮ', 'ไฮเงี่ยน', 'งานเงี่ยน']
Weak Keywords  : ['ตัวใหญ่', '❄', 'น้ำแข็ง']
```

**ภาพที่ 12** : ตัวอย่างแสดงการวิเคราะห์ 2 ขั้นตอน

ภาพที่ 12 แสดงตัวอย่างผลการวิเคราะห์ข้อมูล โดยขั้นตอนที่ 1 ระบบสามารถคัดกรองของต้องห้ามได้อย่างถูกต้อง และในขั้นตอนที่ 2 สามารถระบุประเภทยาเสพติดได้อย่างแม่นยำผ่านกลไกการตรวจจับ Strong Keywords (เช่น ไฮ) ร่วมกับ Weak Keywords (เช่น ตัวใหญ่, น้ำแข็ง, ❄ )

### 4.5 การแสดงผล

ภาพที่ 13 แผนภาพแสดงกลุ่มคำที่ใช้ซื้อขายบนสังคมออนไลน์ ส่วนภาพที่ 14 แสดง emoji แทนการซื้อขายเพื่อหลีกเลี่ยงการตรวจสอบของสินค้าต้องห้าม

1. ภาพแสดงแผนภาพคำ Word Cloud




**ภาพที่ 13** แสดง Word Cloud สินค้าทั่วไปและสินค้าต้องห้าม

2. ภาพแสดงแผนกลุ่มคำ Emoji Cloud



**ภาพที่ 15** แสดง Confusion Matrix ในการจำแนกข้อมูลของต้องห้ามในตอนที่ 1

จาก Confusion Matrix ในขั้นตอนที่ 1 บนชุดทดสอบ 300 ตัวอย่าง พบว่า Naïve Bayes และ LR ให้ผลดีที่สุด โดยมีค่า F1-score เท่ากับ 0.9967 และจำแนกผิดเพียง 1 ตัวอย่าง ขณะที่ Linear SVM และ Random Forest มีความคลาดเคลื่อนสูงกว่าเล็กน้อย

```
                precision    recall  f1-score   support

None_Prohibited       1.00      0.99      0.99       150
Prohibited_Goods      0.99      1.00      0.99       150

      accuracy                           0.99       300
     macro avg      0.99      0.99      0.99       300
  weighted avg      0.99      0.99      0.99       300
```

**ภาพที่ 16** แสดงผล Per-class performance

ผลการประเมินรายคลาสจากชุดทดสอบที่มีกลุ่มละ 150 ตัวอย่าง พบว่าแบบจำลองสามารถจำแนกสินค้าทั่วไปและสินค้าต้องห้ามได้อย่างถูกต้อง โดยมีค่า Accuracy และค่าเฉลี่ย ของ Precision Recall และ F1-score อยู่ที่ประมาณ 0.99 ในทุกตัวชี้วัด สะท้อนถึงประสิทธิภาพที่ดีในการจำแนกทั้งสองคลาส

### 4.6 ผลการเปรียบเทียบประสิทธิภาพกับงานวิจัยอื่น ๆ

**ตารางที่ 3** : ผลการเปรียบเทียบประสิทธิภาพงานวิจัยอื่นๆ

| Reference | Algorithm | Accuracy | Precision | Recall | F1 |
|---|---|---|---|---|---|
| ผู้วิจัย[Stage 1] | Naive Bayes | 0.9921 | 0.9921 | 0.9921 | 0.9921 |
| ผู้วิจัย[Stage 2] | Random Forest | 0.9190 | 0.9170 | 0.9190 | 0.9110 |
| ชวรัส [1] | Random Forest | 0.9709 | 0.8800 | 0.9000 | 0.8800 |
| ปรางค์ปราณี [2] | Naive Bayes | - | 0.9231 | 0.9834 | 0.9839 |
| อุลัยลักษ์ [3] | Logistic regression | 0.8600 | 0.8600 | 0.8600 | 0.8600 |
| สิริวุฒิ [4] | Linear SVM | 0.9845 | 0.9845 | 0.9845 | 0.9845 |

## 5. สรุป

### 5.1 สรุปผลการวิจัย

ผลการเปรียบเทียบประสิทธิภาพของแบบจำลองการเรียนรู้เพื่อตรวจจับข้อความสินค้าต้องห้ามบนสื่อสังคมออนไลน์แพลตฟอร์ม Twitter พบว่าทุกแบบจำลองให้ประสิทธิภาพผ่านเกณฑ์ร้อยละ 80 ให้สัมบูรณ์ สอดคล้องกับสมมติฐานที่ตั้งไว้ โดยในขั้นตอนที่ 1 แบบจำลอง Naïve Bayes มีประสิทธิภาพสูงสุด (99.21%) ในการคัดกรองสินค้าต้องห้าม และในขั้นตอนที่ 2 แบบจำลอง Random Forestมีประสิทธิภาพสูงสุด (91.90%) สำหรับการจำแนกประเภทยาเสพติดกลุ่มยาเสพติด

### 5.2 อภิปรายผลการวิจัย

1. ผลการทดลองชี้ให้เห็นว่าความเหมาะสมของแบบจำลองขึ้นอยู่กับลักษณะข้อมูล โดยในขั้นตอนที่ 1 Naïve Bayes ให้ผลลัพธ์ได้ดีกับชุดข้อมูลที่สอดคล้องกับ ปรางค์ปราณี[2] จากการจัดการชุดข้อมูลที่สมดุล

ขณะที่ขั้นตอนที่ 2 Random Forest มีประสิทธิภาพที่ดีที่สุดในการจัดการกับข้อมูลที่มีโครงสร้างซับซ้อนและไม่สมดุลมีเสถียรภาพกว่าแบบจำลองอื่นโดยเฉพาะข้อความจากสื่อสังคมออนไลน์สอดคล้องกับ ชวรัส [1] พบว่า Random Forest เป็นแบบจำลองที่มีประสิทธิภาพสูงสุดในการจัดการกับข้อมูลที่มีความซับซ้อนและจัดการกับข้อมูลไม่ได้มีความสมดุลได้ดี

2. แบบจำลอง SVM ทำงานได้ดีกับข้อมูลที่มีโครงสร้างเป็นแบบแผนและสมดุลต่างจากกรณีของสิริวุฒิ [4] ที่จำแนกชื่อสารเคมี IUPAC ได้แม่นยำร้อยละ 98.45 สาเหตุหลักมาจากความท้าทายของข้อความบนสื่อสังคมออนไลน์ที่ขาดความสมดุล ไร้โครงสร้างชัดเจน และเต็มไปด้วยสัญญาณรบกวน เช่น คำสแลง Emoji และกลวิธีทางภาษาที่จงใจหลบเลี่ยงการตรวจจับ

3. งานวิจัยนี้ก้าวข้ามข้อจำกัดเดิมด้วยการรักษา Emoji ไว้เป็นคุณลักษณะสำคัญ พร้อมพัฒนาเทคนิค Weighted Keyword Scoring และ Contextual Rules ซึ่งเป็นนวัตกรรมทางเทคนิคที่ช่วยให้แบบจำลองสามารถตรวจจับรหัสแฝงที่จงใจใช้เพื่ออำพรางข้อความได้อย่างแม่นยำ

4. พบข้อจำกัดสำคัญ 2 ประการ คือ ต้นทุน ในการดึงข้อมูลผ่าน API ของแพลตฟอร์ม Twitter ที่มีราคาสูง และปัญหาความกำกวมข้ามหมวดหมู่ เช่น คำเฉพาะที่อาจเป็นได้ทั้งชื่อยี่ห้อและชื่อร้านค้าในบริบทที่ต่างกัน ซึ่งอาจส่งผลให้การจำแนกหมวดหมู่คลาดเคลื่อนได้ เช่น ร้านขายกัญชาชื่อร้านเป็นยี่ห้อบุหรี่ไฟฟ้า เป็นต้น

### 5.3 ข้อเสนอแนะ

กรมศุลกากรสามารถนำแบบจำลองนี้ไปบูรณาการเป็นเครื่องมือเฝ้าระวังเชิงรุกเพื่อลดภาระงาน และควรปรับปรุงพจนานุกรมรหัสแฝงอยู่เสมอ รวมถึงขยายการตรวจจับไปยังหมวดต้องห้าม และแพลตฟอร์มอื่นๆ (Facebook, TikTok, Telegram) พร้อมเพิ่มขนาดชุดข้อมูลให้แบบจำลองอีกทั้งยังพัฒนา Web Application หรือ Dashboard อัตโนมัติพร้อมเชื่อมโยงกับฐานข้อมูลการนำเข้า-ส่งออกของศุลกากร เพื่อวิเคราะห์เครือข่ายผู้กระทำผิดแบบครบวงจร

## เอกสารอ้างอิง


[1] ชวรัส เกรอต. การตรวจสอบพฤติกรรมการเบิกสิทธิประโยชน์กรณีทันตกรรมของสำนักงานประกันสังคมด้วยเทคนิคเหมืองข้อมูล. สารนิพนธ์ปริญญาวิทยาศาสตรมหาบัณฑิต มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2565.

[2] ปรางค์ปราณี ล้ำเลิศพระคุณ. การจำแนกประเภทการให้ความช่วยเหลือเด็กที่ถูกทารุณโดยเทคนิคการทำเหมืองข้อความ. วิทยานิพนธ์ปริญญาวิทยาศาสตรมหาบัณฑิต มหาวิทยาลัยธุรกิจบัณฑิตย์, 2561

[3] อุลัยลักษ์ นาดี. การจำแนกข้อมูลการรับแจ้งปัญหาด้านสารสนเทศด้วยเทคนิคเหมืองข้อมูล. สารนิพนธ์ปริญญาวิทยาศาสตรมหาบัณฑิต มหาวิทยาลัยเทคโนโลยี พระจอมเกล้าพระนครเหนือ, 2565

[4] สิริวุฒิ คงนคร. การจำแนกพิกัดศุลกากรของสินค้าเคมีภัณฑ์อินทรีย์ด้วยเทคนิคการทำเหมืองข้อมูล. สารนิพนธ์ปริญญาวิทยาศาสตรมหาบัณฑิตมหาวิทยาลัยธรรมศาสตร์, 2565

[5] กนิษฐา ไทยกล้า.(2566) ภาษาชื่อเรียกอิโมจิ ที่ใช้ในการขายยาเสพติดบนอินเทอร์เน็ต. เชียงใหม่: สถาบันวิจัยวิทยาศาสตร์สุขภาพ มหาวิทยาลัยเชียงใหม่.

[6] กรมศุลกากร (2567) [ออนไลน์] ของต้องห้ามต้องกำกัด [สืบค้นวันที่ 1 กรกฎาคม 2567] จาก https://www.customs.go.th/

[7] Analysis Function. (ม.ป.ป.). [ออนไลน์]. An introductory guide to MCDA. [สืบค้นวันที่ 24 มีนาคม 2569] จาก https://analysisfunction.civilservice.gov.uk/policy-store/an-introductory-guide-to-mcda

[8] ScienceDirect. (ม.ป.ป.). [ออนไลน์]. Lexicon-based approach. [สืบค้นวันที่ 24 มีนาคม 2569] จาก https://www.sciencedirect.com/topics/computer-science/lexicon-based-approach

[9] Bundidth Dangsawang and Siranee Nuchitprasitchai. Amachine learning approach for detecting customs fraud through unstructured data analysis in social media. Decision Analytics Journal. [serial online] 2024. DOI: 10.1016/j.dajour.2024.100408

# แนวทางการพัฒนาทักษะการใช้ปัญญาประดิษฐ์ของนักเรียนและนักศึกษาอาชีวศึกษา

# Guidelines for Developing Artificial Intelligence Skills among Vocational Education Students

*เรวดี อะมินรัมย์ (Rewadee Aminram)[1] และนลินภัสร์ บำเพ็ญเพียร (Nalinpat Bhumpenpein)[2]*

*[1]ภาควิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล*

*มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ*

*[1]s6707011857080@email.kmutnb.ac.th, [2]nalinpat.b@itd.kmutnb.ac.th*

**บทคัดย่อ**

*งานวิจัยนี้มีวัตถุประสงค์เพื่อศึกษาปัจจัยที่ส่งผลต่อการพัฒนาทักษะการใช้ปัญญาประดิษฐ์ (AI) ของนักเรียนและนักศึกษาอาชีวศึกษา พร้อมเสนอแนวทางการพัฒนาทักษะดังกล่าว กลุ่มตัวอย่าง ได้แก่ นักเรียนและนักศึกษาอาชีวศึกษาจำนวน 402 คน ได้มาโดยการสุ่มตัวอย่างอย่างง่าย เครื่องมือที่ใช้คือแบบสอบถามมาตราส่วนประมาณค่า 5 ระดับ วิเคราะห์ข้อมูลด้วยสถิติเชิงพรรณนา สหสัมพันธ์เพียร์สัน และการถดถอยพหุคูณแบบขั้นตอน ผลการวิจัยพบว่า มี 5 ปัจจัยที่ส่งผลต่อการพัฒนาทักษะการใช้ปัญญาประดิษฐ์ของนักเรียนและนักศึกษาอาชีวศึกษาอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ได้แก่ ทักษะการสื่อสารและโต้ตอบกับ AI ทัศนคติที่ดีและความมั่นใจในการใช้ AI ทักษะการประเมินและตรวจสอบผลลัพธ์ของ AI ทักษะการเรียนรู้ด้วย AI และทักษะการใช้เครื่องมือหรือแอปพลิเคชัน AI ซึ่งร่วมกันพยากรณ์การพัฒนาทักษะการใช้ AI ได้ร้อยละ 81.0 ผลการวิจัยนี้สามารถนำไปใช้เป็นแนวทางในการพัฒนาทักษะการใช้ปัญญาประดิษฐ์ โดยควรมุ่งเน้นการจัดการเรียนรู้เชิงปฏิบัติ การฝึกเขียนคำสั่งพรอมต์อย่างมีประสิทธิภาพ การประเมินความน่าเชื่อถือของผลลัพธ์ และการประยุกต์ใช้ AI ในงานวิชาชีพจริงเพื่อยกระดับสมรรถนะผู้เรียนให้สอดคล้องกับตลาดแรงงานยุคดิจิทัล*

**คำสำคัญ:** ปัญญาประดิษฐ์ นักเรียนและนักศึกษาอาชีวศึกษา ทักษะการใช้ปัญญาประดิษฐ์ สมรรถนะดิจิทัล

**Abstract**

*This research aimed to investigate the factors affecting the development of artificial intelligence (AI) usage skills among vocational students and to propose guidelines for enhancing such skills. The sample consisted of 402 vocational students selected through simple random sampling. The research instrument was a five-point rating scale questionnaire. Data were analyzed using descriptive statistics, Pearson's correlation coefficient, and stepwise multiple regression analysis. The findings revealed that five factors significantly affected the development of AI usage skills among vocational students at the .05 level of statistical significance: AI communication and interaction skills, positive attitudes and confidence in using AI, skills in evaluating and verifying AI outputs, AI-assisted learning skills, and skills in using AI tools or applications. These factors jointly predicted 81.0% of the variance in AI usage skill development. The findings can be applied as guidelines for developing AI usage skills by emphasizing practical learning, effective prompt writing, evaluation of the reliability of AI outputs, and the application of AI in authentic vocational tasks to enhance learners' competencies in alignment with the demands of the digital labor market.*

**Keywords:** Artificial Intelligence, Vocational Student, AI Use Skills, Digital Competency.

## 1. บทนำ

ปัญญาประดิษฐ์ (Artificial Intelligence: AI) เป็นเทคโนโลยีสำคัญที่มีบทบาทต่อการเปลี่ยนรูปแบบการเรียนรู้ การทำงาน รวมถึงเข้ามาเป็นส่วนหนึ่งในชีวิตประจำวัน

ส่งผลให้สมรรถนะด้าน AI เป็นทักษะจำเป็นสำหรับผู้เรียนยุคดิจิทัล [1] หลายประเทศจึงให้ความสำคัญกับการพัฒนาสมรรถนะด้าน AI เพื่อเตรียมกำลังคนให้สอดคล้องกับรูปแบบงานและทักษะแห่งอนาคต [2] ประเทศไทยได้กำหนดยุทธศาสตร์ปัญญาประดิษฐ์แห่งชาติ พ.ศ. 2565–2570 เพื่อส่งเสริมการพัฒนาและการประยุกต์ใช้ AI อย่างมีประสิทธิภาพ ซึ่งจะนำไปสู่การยกระดับเศรษฐกิจและคุณภาพชีวิต [3] ในบริบทการศึกษา AI ถูกนำมาใช้เพิ่มขึ้นอย่างต่อเนื่อง ผลสำรวจของ OECD พบว่า นักเรียนอายุ 8–18 ปี ในสวิตเซอร์แลนด์ ผู้เรียนสายอาชีวะร้อยละ 42 ใช้ Generative AI ในชั้นเรียน และร้อยละ 41 ใช้เพื่อการเรียนที่บ้าน [4] สะท้อนให้เห็นว่า AI กลายเป็นเครื่องมือสำคัญของผู้เรียนในยุคดิจิทัล อย่างไรก็ตาม การใช้ AI ที่เพิ่มขึ้นไม่ได้หมายความว่าผู้เรียนมีทักษะการใช้ AI อย่างเหมาะสมเสมอไป เนื่องจากยังพบข้อจำกัด เช่น การขาดทักษะการเขียนคำสั่งพรอมต์ การไม่ตรวจสอบความถูกต้องของข้อมูล การไม่เข้าใจข้อจำกัดของ AI และการนำผลลัพธ์ไปใช้โดยขาดการพิจารณาด้านจริยธรรม สถานศึกษาอาชีวศึกษาเป็นกลไกสำคัญในการพัฒนากำลังคนสู่ภาคอุตสาหกรรม [5] จึงจำเป็นต้องส่งเสริมให้ผู้เรียนสามารถประยุกต์ใช้ AI ในงานวิชาชีพได้จริงแต่ในทางปฏิบัติยังพบข้อจำกัด ทั้งด้านความเข้าใจ AI ทักษะการใช้ การประเมินผลลัพธ์ และด้านจริยธรรม [6] แม้ว่าจะมีงานวิจัยเกี่ยวกับสมรรถนะด้าน AI ของผู้เรียนในภาพรวม แต่การศึกษาปัจจัยที่ส่งผลต่อการพัฒนาทักษะการใช้ AI ของนักเรียนและนักศึกษาอาชีวศึกษาโดยตรงยังมีอย่างจำกัด ด้วยเหตุนี้ งานวิจัยนี้มุ่งศึกษาปัจจัยที่ส่งผลต่อการพัฒนาทักษะการใช้ปัญญาประดิษฐ์ของนักเรียนและนักศึกษาอาชีวศึกษา พร้อมเสนอแนวทางการพัฒนาที่เหมาะสม เพื่อสนับสนุนการพัฒนากำลังคนสายอาชีพให้พร้อมสู่ตลาดแรงงานยุคดิจิทัลที่ขับเคลื่อนด้วยปัญญาประดิษฐ์

## 2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

### 2.1 แนวทางการพัฒนาทักษะการใช้ปัญญาประดิษฐ์ของนักเรียนและนักศึกษาอาชีวศึกษา

การพัฒนาทักษะการใช้ปัญญาประดิษฐ์ ควรดำเนินการอย่างเป็นระบบ โดยเริ่มจากการสร้างความรู้พื้นฐานไปสู่การประยุกต์ใช้งานจริง ผ่านกรอบแนวคิด Five Big Ideas [7] ซึ่งเน้นองค์ประกอบที่สัมพันธ์กันด้วยการจัดการเรียนรู้ 3 ด้าน คือ ด้านเนื้อหา ความก้าวหน้าตามช่วงชั้น และการจัดการเรียนรู้ผ่านการลงมือทำ เพื่อให้ผู้เรียนเกิดทักษะจากการสร้างชิ้นงานจริง และเปลี่ยนบทบาทจากผู้ใช้งานสู่ผู้ร่วมสร้างเทคโนโลยี ด้วยการบูรณาการวงจรการเรียนรู้ 4 ประการ ได้แก่ การรู้เท่าทัน AI ทักษะการออกแบบพรอมต์ การเรียนรู้เชิงโต้ตอบ และการประเมินผล โดยการสะท้อนผลลัพธ์อย่างต่อเนื่อง เป็นการสร้างวงจรการพัฒนาทักษะด้วยตนเองอย่างต่อเนื่องให้แก่ผู้เรียน [8], [9]

### 2.2 ปัจจัยที่ส่งผลการพัฒนาทักษะการใช้ปัญญาประดิษฐ์ของนักเรียนและนักศึกษาอาชีวศึกษา

2.2.1 ความรู้พื้นฐานเกี่ยวกับ AI เป็นฐานที่ช่วยให้ผู้เรียนเข้าใจแนวคิดสำคัญของ AI มากกว่าการใช้เครื่องมือแบบกดตามขั้นตอน โดยความฉลาดรู้ด้าน AI เน้นการเข้าใจหลักการทั่วไปหรือบทบาทของข้อมูล และตระหนักเหตุผลที่ปัญญาประดิษฐ์ อาจให้คำตอบผิดได้ เพื่อให้ผู้เรียนไม่ยึดผลลัพธ์จาก AI เป็นข้อสรุปสุดท้าย และสามารถนำความรู้ไปประยุกต์ใช้กับงานจริงได้ [10]

2.2.2 ความเข้าใจขอบเขตความสามารถและข้อจำกัดของ AI ช่วยให้ผู้เรียนใช้ AI อย่างระมัดระวัง และมีวิจารณญาณ เช่น การตระหนักถึงความคลาดเคลื่อน อคติ หรือคำตอบที่ดูน่าเชื่อถือแต่ผิดจากข้อเท็จจริง ซึ่งเป็นเงื่อนไขสำคัญของการใช้ AI อย่างน่าเชื่อถือและสอดคล้องกับหลักมนุษย์เป็นศูนย์กลางในบริบทการศึกษา [11], [12]

2.2.3 ทักษะการใช้เครื่องมือหรือแอปพลิเคชัน AI เป็นสมรรถนะในการเลือกใช้เครื่องมือและฟังก์ชัน ให้เหมาะกับบริบทการเรียนและงานวิชาชีพ โดยความชำนาญในการใช้เครื่องมือ AI ขั้นพื้นฐานจะเป็นฐานสำคัญในการต่อยอดสู่ทักษะขั้นสูง เช่น การออกแบบคำสั่งและการควบคุมคุณภาพผลลัพธ์ [8], [9]

2.2.4 ทักษะการสื่อสารและโต้ตอบกับ AI เป็นสมรรถนะในการเปลี่ยนความต้องการของผู้เรียนไปสู่ผลลัพธ์ที่มีคุณภาพผ่านการออกแบบคำสั่ง โดยผู้เรียนจำเป็นต้องกำหนดบริบท เป้าหมาย เงื่อนไข และรูปแบบคำตอบอย่างชัดเจน และสามารถปรับปรุงคำสั่งซ้ำจากผล

ย้อนกลับ เพื่อให้ผู้เรียนได้ฝึกลงมือทำ โดยผลลัพธ์จะดีขึ้นอย่างเห็นได้ชัดช่วยให้ผู้เรียนควบคุมทิศทางคำตอบได้แม่นยำมากขึ้น [8]

2.2.5 ทักษะการประเมินและตรวจสอบผลลัพธ์ของ AI เป็นทักษะสำคัญของการใช้ AI อย่างมีวิจารณญาณ ซึ่งจะช่วยลดความเสี่ยงจากข้อมูลที่คลาดเคลื่อนและเพิ่มคุณภาพการตัดสินใจ ผู้เรียนควรตรวจสอบความถูกต้อง ความสอดคล้องกับบริบท และความน่าเชื่อถือของแหล่งข้อมูลก่อนนำไปใช้จริง ซึ่งเป็นแนวทางที่กรอบสมรรถนะสากลด้าน AI เน้นย้ำ [10]

2.2.6 ความตระหนักด้านจริยธรรมและการใช้ AI เป็นเงื่อนไขที่ใช้กำกับการใช้ AI ให้มีความรับผิดชอบ ครอบคลุมทั้งด้านความเป็นธรรม ความโปร่งใส ความเป็นส่วนตัว และคำนึงถึงผลกระทบต่อผู้อื่น กรอบสมรรถนะด้าน AI ในการศึกษาเน้นย้ำว่าการใช้งานที่ดีต้องไม่ละเมิดสิทธิ ไม่สร้างความเหลื่อมล้ำ และต้องมีความรับผิดชอบต่อผลลัพธ์ที่นำไปใช้ [12]

2.2.7 ทักษะการเรียนรู้ด้วย AI สะท้อนการใช้ AI เป็นเสมือนผู้ช่วยในการเรียนรู้ส่วนบุคคล เช่น อธิบาย หรือสรุปบทเรียน ประยุกต์โจทย์ปัญหา และสนับสนุนการกำกับตนเองในการเรียนรู้ การจัดกิจกรรมกระบวนการเรียนรู้แบบ “ลงมือ–สะท้อน–ปรับปรุง” ช่วยให้ผู้เรียนเห็นคุณค่าการใช้ AI และพัฒนาทักษะการเรียนรู้ด้วย AI ได้ดีขึ้นสามารถนำไปใช้ได้จริง [8], [9]

2.2.8 ทัศนคติที่ดีและความมั่นใจในการใช้ AI เป็นแรงขับสำคัญที่ส่งผลต่อพฤติกรรมการใช้งานจริง เมื่อผู้เรียนมีทัศนคติเชิงบวกและมีความเชื่อมั่นในความสามารถของตนเองต่อการใช้ AI ทำให้ผู้เรียนเกิดความกล้าในการทดลองใช้ เรียนรู้ปรับคำสั่ง ตรวจสอบผลลัพธ์และฝึกใช้อย่างต่อเนื่องจนเกิดความชำนาญ โดยหลังจากผู้เรียนได้รับแนวทางการเขียนคำสั่งและเห็นผลลัพธ์ที่ดีขึ้น จะช่วยเพิ่มความมั่นใจและทัศนคติเชิงบวกต่อการใช้งานจริง [9]

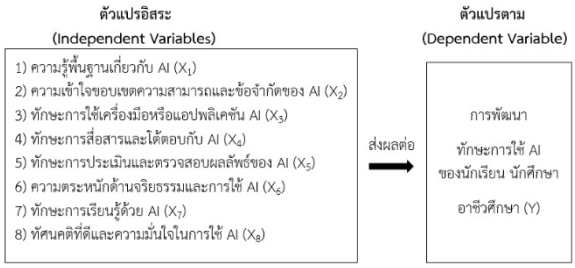
2.2.9 การพัฒนาทักษะการใช้ AI ของนักเรียนและนักศึกษาอาชีวศึกษา เป็นสมรรถนะที่ต้องพัฒนาด้วยการบูรณาการร่วมกันหลายมิติ ทั้งด้านความรู้ ความเข้าใจข้อจำกัด การสื่อสาร การประเมินผลลัพธ์ และจริยธรรม ควบคู่กับการสร้างทัศนคติและความมั่นใจ ผ่านการเรียนรู้เชิงปฏิบัติ และการสะท้อนผลอย่างต่อเนื่อง มุ่งเน้นการสร้างความฉลาดรู้ด้าน AI เพื่อเตรียมความพร้อม และตอบสนองความต้องการของตลาดแรงงาน ให้กำลังคนอาชีวศึกษาสามารถปรับตัวเข้าสู่การทำงานร่วมกับ AI ได้อย่างปลอดภัยและมีประสิทธิภาพ

## 3. วิธีการดำเนินงาน

### 3.1 ประชากรและกลุ่มตัวอย่าง

การวิจัยนี้เป็นการวิจัยเชิงปริมาณ เพื่อศึกษาปัจจัยที่ส่งผลต่อการพัฒนาทักษะการใช้ปัญญาประดิษฐ์ของนักเรียนและนักศึกษาอาชีวศึกษา และเสนอแนวทางการพัฒนาทักษะดังกล่าว ประชากรที่ใช้ในการวิจัยคือ นักเรียนและนักศึกษาสังกัดสำนักงานคณะกรรมการการอาชีวศึกษา จากสถานศึกษาทั้งหมด 433 แห่ง รวมจำนวนทั้งสิ้น 637,485 คน [13] กำหนดขนาดกลุ่มตัวอย่างด้วยสูตรคำนวณของ Cochran ที่ระดับความเชื่อมั่นร้อยละ 95 และค่าความคลาดเคลื่อนที่ยอมรับได้ร้อยละ 5 ได้ขนาดกลุ่มตัวอย่างที่เหมาะสมจำนวน 384 คน ทั้งนี้เพื่อให้ข้อมูลมีความแม่นยำ ผู้วิจัยจึงกำหนดขนาดของกลุ่มตัวอย่างเป็นอย่างน้อย 400 คน และเก็บข้อมูลจริงจากกลุ่มตัวอย่างจำนวน 402 คน โดยวิธีการสุ่มตัวอย่างอย่างง่าย ระหว่างเดือนธันวาคม 2568 ถึงเดือนมกราคม 2569 ด้วยแบบสอบถามออนไลน์ ซึ่งมีข้อคำถามทั้งสิ้น 43 ข้อ แบ่งเป็น 3 ส่วน ได้แก่ ส่วนที่ 1 ข้อมูลทั่วไป จำนวน 8 ข้อ ส่วนที่ 2 ปัจจัยที่ส่งผลต่อการพัฒนาทักษะการใช้ AI จำนวน 34 ข้อ ส่วนที่ 3 ข้อเสนอแนะเพิ่มเติม จำนวน 1 ข้อ การตรวจสอบคุณภาพเครื่องมือ แบบสอบถามผ่านการตรวจสอบความตรงเชิงเนื้อหาจากผู้เชี่ยวชาญ จำนวน 5 ท่าน มีค่าดัชนีความสอดคล้อง อยู่ระหว่าง 0.80 ถึง 1.00 และมีค่าความเชื่อมั่นสัมประสิทธิ์แอลฟาของครอนบาคทั้งฉบับเท่ากับ 0.980 จากนั้นนำข้อมูลประมวลผลด้วยโปรแกรมสำเร็จรูปทางสถิติ SPSS โดยผู้วิจัยเลือกใช้สถิติการวิเคราะห์การถดถอยพหุคูณแบบขั้นตอน Stepwise เพื่อคัดเลือกตัวแปรอิสระที่สามารถพยากรณ์ตัวแปรตามได้อย่างมีนัยสำคัญทางสถิติ และเพื่อให้ได้สมการพยากรณ์ที่เหมาะสม

**3.2 โมเดลการวิจัย**




**ภาพที่ 1:** โมเดลการวิจัย

## 4. ผลการดำเนินการวิจัย

**4.1 ข้อมูลทั่วไปของผู้ตอบแบบสอบถาม**

จากการวิเคราะห์ข้อมูลผู้ตอบแบบสอบถามจำนวน 402 คน พบว่า ส่วนใหญ่เป็นเพศชาย จำนวน 229 คน คิดเป็นร้อยละ 57.0 กำลังศึกษาอยู่ในระดับ ปวช. จำนวน 286 คน คิดเป็นร้อยละ 71.1 ด้านประสบการณ์การใช้ AI พบว่า ส่วนใหญ่มีประสบการณ์การใช้งานระหว่าง 6 เดือน ถึง 1 ปี จำนวน 237 คน คิดเป็นร้อยละ 59.0 และเครื่องมือ AI ที่นิยมใช้มากที่สุดคือ ChatGPT คิดเป็นร้อยละ 45.6 รองลงมาคือ Gemini คิดเป็นร้อยละ 38.1 และ Gamma คิดเป็นร้อยละ 5.8 ตามลำดับ โดยสรุปกลุ่มตัวอย่างส่วนใหญ่เป็นนักศึกษาระดับ ปวช. เพศชาย มีประสบการณ์ใช้งาน AI ในระดับปานกลาง และนิยมใช้ ChatGPT มากที่สุด

**4.2 ปัจจัยที่ส่งผลต่อการพัฒนาทักษะการใช้ AI**

ก่อนทำการวิเคราะห์การถดถอยพหุคูณแบบ Stepwise ผู้วิจัยได้ทำการตรวจสอบความสัมพันธ์ระหว่างตัวแปรเพื่อป้องกันปัญหาภาวะร่วมเส้นตรงพหุ (Multicollinearity) ผลการตรวจสอบพบว่า ตัวแปรอิสระทุกตัวในสมการที่ 5 อยู่ในเกณฑ์ยอมรับได้ โดยมีค่า Tolerance อยู่ระหว่าง 0.206 ถึง 0.314 ซึ่งมีค่ามากกว่า 0.1 และมีค่าสถิติวัดการขยายตัวของความแปรปรวน (VIF) อยู่ระหว่าง 3.186 ถึง 4.853 ซึ่งมีค่าน้อยกว่า 5 สะท้อนให้เห็นว่าตัวแปรอิสระไม่มีความสัมพันธ์กันเองสูงเกินไปจึงสามารถนำมาวิเคราะห์การถดถอยพหุคูณเพื่อทดสอบสมมติฐานในลำดับต่อไปได้

จากการวิเคราะห์การถดถอยพหุคูณแบบ Stepwise พบว่า มี 5 ปัจจัยที่ร่วมกันพยากรณ์การพัฒนาทักษะการใช้ AI (DAI) ได้อย่างมีนัยสำคัญทางสถิติที่ระดับ .001 โดยปัจจัยพยากรณ์ทั้ง 5 ด้านดังกล่าว สามารถร่วมกันอธิบายความแปรปรวนของตัวแปรตามคือ การพัฒนาทักษะการใช้ AI ได้ร้อยละ 81.00 ($R^2$ = .810) ได้แก่ ทักษะการสื่อสารและโต้ตอบกับ AI ทัศนคติและความมั่นใจในการใช้ AI ทักษะการประเมินและตรวจสอบผลลัพธ์ของ AI ทักษะการเรียนรู้ด้วย AI และทักษะการใช้เครื่องมือหรือแอป-พลิเคชัน AI ดังตารางที่ 1 ผลการศึกษาชี้ให้เห็นว่า ทักษะการสื่อสารและโต้ตอบกับ AI เป็นปัจจัยสำคัญที่ช่วยในการส่งเสริมการพัฒนาทักษะการใช้ AI ของผู้เรียน ขณะที่ปัจจัยด้านทัศนคติและความมั่นใจ ความสามารถในการประเมินผล และการใช้งานเครื่องมือ มีบทบาทสนับสนุนรองลงมา แต่ยังคงมีความสำคัญทางสถิติ

**ตารางที่ 1:** ผลการวิเคราะห์ปัจจัยที่ส่งผลต่อการพัฒนาทักษะการใช้ปัญญาประดิษฐ์ของนักเรียนและนักศึกษาอาชีวศึกษา

n=402

| ตัวแปรพยากรณ์ | B | S.E. | Beta | t | Sig |
|---|---|---|---|---|---|
| ค่าคงที่(Constant) | .116 | .099 | | 1.180 | .239 |
| ทัศนคติที่ดีและความมั่นใจในการใช้ AI (ATAI) | .222 | .045 | .229 | 4.978 | .000** |
| ทักษะการสื่อสารและโต้ตอบกับ AI (CAI) | .241 | .043 | .243 | 5.567 | .000** |
| ทักษะการประเมินและตรวจสอบผลลัพธ์ของ AI (EAI) | .183 | .045 | .187 | 4.078 | .000** |
| ทักษะการใช้เครื่องมือหรือแอปพลิเคชัน AI (SKAI) | .148 | .038 | .152 | 3.899 | .000** |
| ทักษะการเรียนรู้ด้วย AI (LAI) | .176 | .049 | .172 | 3.573 | .000** |
| R = .900 | $R^2$ = .810 | Adj.$R^2$=.808 | F = 338.250 | Sig.=.000 | |
| a. Dependent Variable: การพัฒนาทักษะการใช้ AI (DAI) | | | | | |

** มีนัยสำคัญทางสถิติที่ระดับ .01

**4.3 แนวทางการพัฒนาทักษะการใช้ปัญญาประดิษฐ์ของนักเรียนและนักศึกษาอาชีวศึกษา**

จากการศึกษาปัจจัยที่ส่งผลต่อการพัฒนาทักษะการใช้ AI ของนักเรียนและนักศึกษาอาชีวศึกษา ผู้วิจัยเสนอแนวทางสำหรับสถานศึกษาอาชีวศึกษาเพื่อยกระดับสมรรถนะผู้เรียน โดยให้ความสำคัญกับปัจจัยที่มีน้ำหนักการพยากรณ์ที่มีอิทธิพลสูงที่สุด ดังนี้

4.3.1 บูรณาการทักษะการสื่อสารและโต้ตอบกับ AI โดยสถานศึกษาควรเน้นบูรณาทักษะการเขียนคำสั่งร่วมกับรายวิชาชีพ มุ่งเน้นการจัดกิจกรรมการเรียนรู้ให้ผู้เรียนสามารถให้บริบท กำหนดเป้าหมาย และการเขียนเงื่อนไขที่ชัดเจน เพื่อให้ผู้เรียนสามารถควบคุมทิศทาง

ผลลัพธ์ที่ได้จาก AI เนื่องจากเป็นสมรรถนะหลัก ที่ส่งผลต่อการพัฒนาทักษะสูงที่สุด [10], [14]

4.3.2 สร้างทัศนคติเชิงบวกและความมั่นใจในการใช้ AI โดยครูจัดสภาพแวดล้อมเปิดโอกาสให้ผู้เรียนได้เรียนรู้ด้วยตนเอง ได้ทดลองใช้เครื่องมือ AI ในการแก้ปัญหาจริง เพื่อเสริมสร้างและเพิ่มความมั่นใจในศักยภาพของตนเองต่อการใช้ AI จัดกิจกรรมเน้นการสร้างแรงจูงใจที่แสดงให้เห็นว่า AI เป็นเสมือนผู้ช่วยในการทำงานไม่ใช่คนทำงานแทน [15]

4.3.3 สร้างวิจารณญาณการประเมินและตรวจสอบความถูกต้องของผลลัพธ์ โดยบูรณาการร่วมกับรายวิชาชีพ ให้มีการกำหนดขั้นตอนการตรวจสอบความถูกต้องของผลลัพธ์และการวิเคราะห์อคติของข้อมูลที่ได้จาก AI เพื่อป้องกันความเสี่ยงจากข้อมูลที่คลาดเคลื่อนหรือเกิดจากอาการหลอนของ AI [7], [16]

4.3.4 สนับสนุนการเรียนรู้ด้วยตนเอง จัดกิจกรรมส่งเสริมให้ผู้เรียนใช้ AI เป็นเครื่องมือสนับสนุนการเรียนรู้ เน้นการ "ลงมือทำ–สะท้อนผล–ปรับปรุง" อย่างต่อเนื่อง ให้ผู้เรียนฝึกความคล่องแคล่วในการเข้าถึงแอปพลิเคชัน AI ต่าง ๆ ซึ่งเป็นพื้นฐานสำคัญของการยอมรับเทคโนโลยีก่อนการพัฒนาไปสู่ทักษะที่สูงขึ้น [17]

4.3.5 ส่งเสริมการใช้งานปัญญาประดิษฐ์อย่างต่อเนื่องและเชื่อมโยงสู่การปฏิบัติงานจริง สถานศึกษาควรจัดพื้นที่หรือทรัพยากรด้านซอฟต์แวร์ AI โดยจัดกิจกรรมที่สร้างความถี่ และโครงสร้างพื้นฐานในการเข้าถึง AI [3] เนื่องจากกลุ่มที่ใช้งาน AI เป็นประจำทุกวันมีระดับทักษะสูงกว่ากลุ่มอื่น [3], [11] แสดงให้เห็นว่า ความถี่ในการใช้งานส่งผลโดยตรงต่อระดับทักษะการใช้ AI จึงควรสร้างความถี่ให้ผู้เรียนเข้าถึง AI ได้ง่ายในชีวิตประจำวัน ควบคู่ไปกับการออกแบบโจทย์ปัญหาหรือแบบฝึกหัดที่จำลองสถานการณ์จริงในวิชาชีพ และส่งเสริมการใช้งานอย่างสม่ำเสมอจนเกิดความคล่องแคล่วในการเลือกใช้เครื่องมือให้เหมาะสมกับบริบทของแต่ละงานในวิชาชีพ [6]

## 5. สรุปและอภิปรายผล

งานวิจัยนี้มีวัตถุประสงค์เพื่อศึกษาปัจจัยที่ส่งผลต่อการพัฒนาทักษะการใช้ปัญญาประดิษฐ์ของนักเรียนและนักศึกษาอาชีวศึกษา และเสนอแนวทางการพัฒนาที่เหมาะสม ผลการศึกษาวิจัยพบว่า นักเรียนและนักศึกษาอาชีวศึกษาจำนวน 402 คน มีระดับทักษะการใช้ปัญญาประดิษฐ์โดยรวมอยู่ในระดับมาก ปัจจัยที่ส่งผลอย่างมีนัยสำคัญได้แก่ ทักษะการสื่อสารและโต้ตอบกับ AI ทัศนคติและความมั่นใจในการใช้ AI ทักษะการประเมินและตรวจสอบผลลัพธ์ของ AI ทักษะการเรียนรู้ด้วย AI และทักษะการใช้เครื่องมือหรือแอปพลิเคชัน AI นอกจากนี้ยังพบว่า กลุ่มที่ใช้งาน AI ทุกวันมีระดับทักษะสูงกว่ากลุ่มที่ใช้งานนาน ๆ ครั้ง (เดือนละ 1–2 ครั้ง) และกลุ่มที่ใช้งานบางครั้ง (สัปดาห์ละ 1–2 ครั้ง) อย่างมีนัยสำคัญ ขณะที่คู่อื่นไม่แตกต่างอย่างมีนัยสำคัญ ดังนั้น การพัฒนาทักษะการใช้ปัญญาประดิษฐ์ของนักเรียนและนักศึกษาอาชีวศึกษา ควรมุ่งเน้นฝึกทักษะการสื่อสารกับ AI เพื่อเสริมความมั่นใจควบคู่กับทักษะตรวจสอบผลลัพธ์และการใช้ AI ในการเรียนรู้อย่างมีวิจารณญาณ เพื่อเตรียมความพร้อมสำหรับการทำงานในตลาดแรงงานยุคดิจิทัล

ผลการวิจัยแสดงให้เห็นว่า ทักษะการสื่อสารและโต้ตอบกับ AI เป็นปัจจัยที่ส่งผลต่อการพัฒนาทักษะมากที่สุด สอดคล้องกับ Ng et al. [14] ที่ระบุว่าความสามารถในการสร้างคำสั่ง ทักษะการสื่อสารและโต้ตอบเป็นสมรรถนะสำคัญที่สุดในการดึงศักยภาพของ AI มาใช้ในการเรียนรู้ รองลงมาคือ ทัศนคติและความมั่นใจในการใช้ AI สะท้อนว่าทัศนคติเชิงบวกและความกล้าทดลองใช้งานเป็นแรงผลักดันสำคัญที่ทำให้ผู้เรียนกล้าเรียนรู้ และใช้งานอย่างต่อเนื่องจนเกิดความชำนาญ สอดคล้องกับ Chai et al. [15] ที่พบว่าความเชื่อมั่นในความสามารถของตนเองเป็นแรงจูงใจภายในที่ส่งผลโดยตรงต่อความตั้งใจในการพัฒนาทักษะและการใช้งานจริง นอกจากนี้ ทักษะการประเมินและตรวจสอบผลลัพธ์ สอดคล้อง Baidoo-Anu & Ansah [16] และทักษะการเรียนรู้ AI สะท้อนว่านักเรียนและนักศึกษาอาชีวศึกษาไม่ได้เป็นเพียงผู้ใช้งาน แต่ยังมีความสามารถในการตรวจสอบความถูกต้องของข้อมูลจาก AI ซึ่งเป็นทักษะการใช้วิจารณญาณที่จำเป็นอย่างยิ่งในการป้องกันความคลาดเคลื่อนของข้อมูล สอดคล้องกับ UNESCO [1] และ Molenaar [17] ที่ให้ความสำคัญทักษะนี้อย่างยิ่งในยุค Generative AI และทักษะการใช้เครื่องมือหรือแอปพลิเคชัน

AI แม้จะมีอิทธิพลต่ำที่สุด แต่ยังสะท้อนให้เห็นว่า ความคล่องแคล่วในการเข้าถึงเครื่องมือหรือฟีเจอร์ต่างๆ เป็นเพียงทักษะพื้นฐาน การพัฒนาทักษะไม่ได้ขึ้นอยู่กับการรู้จักเครื่องมือเพียงอย่างเดียว แต่ต้องอาศัยความสามารถในการสื่อสารกับ AI การประเมินผลลัพธ์ และความมั่นใจ การเข้าถึงเครื่องมือเป็นปัจจัยเริ่มต้นที่ทำให้เกิดการยอมรับเทคโนโลยีก่อนที่จะพัฒนาไปสู่ทักษะขั้นสูง ต่อไป

อย่างไรก็ตาม งานวิจัยนี้มีข้อจำกัดด้านการเก็บรวบรวมข้อมูลซึ่งใช้การประเมินตนเอง (Self-report) ของผู้เรียนเป็นหลัก อาจมีข้อจำกัดเกี่ยวกับความลำเอียงจากการประเมินที่เข้าข้างตนเอง ส่งผลให้ระดับทักษะหรือค่าการพยากรณ์ที่ได้อาจสูงกว่าระดับความสามารถในการปฏิบัติงานจริงเล็กน้อย ดังนั้น ผู้นำผลการวิจัยไปใช้จึงควรพิจารณาข้อมูลควบคู่ไปกับการประเมินเชิงประจักษ์จากผลการปฏิบัติงานจริงของผู้เรียนประกอบด้วย

## เอกสารอ้างอิง


[1] UNESCO, "Guidance for generative AI in education and research," 2023. Available online at https://doi.org/10.54675/EWZM9535.

[2] OECD, "Bridging the AI skills gap: Is training keeping up?," OECD Publishing, Paris, 2025. Available online at https://doi.org/10.1787/66d0702e-en.

[3] กระทรวงการอุดมศึกษา วิทยาศาสตร์ วิจัย และนวัตกรรม และกระทรวงดิจิทัลเพื่อเศรษฐกิจและสังคม, "แผนปฏิบัติการด้านปัญญาประดิษฐ์แห่งชาติเพื่อการพัฒนาประเทศไทย (พ.ศ. 2565–2570)," 2565. Available online at https://nationalai.in.th.

[4] OECD, "AI adoption in the education system: International insights and policy considerations for Italy," OECD Publishing, Paris, 2025. Available online at https://doi.org/10.1787/69bd0a4a-en.

[5] OECD, "Training Supply for the Green and AI Transitions: Equipping Workers with the Right Skills," Getting Skills Right, OECD Publishing,Paris, 2024. Available online at https://doi.org/10.1787/7600d16d-en.

[6] X. Lin, G. Xu, and B. Xiong, "Artificial intelligence literacy, sustainability of digital learning and practice achievement: A study of vocational college students," *PLOS ONE*, vol. 20, no. 10, e0332175, 2025. Available online at https://doi.org/10.1371/journal.pone.0332175.

[7] D. Touretzky, et al., "Envisioning AI for K-12: What should every child know about AI?," in *Proceedings of the Thirty-Third AAAI Conference on Artificial Intelligence (AAAI-19)*, 2019. p. 9795-9799.

[8] D. Ognibene, et al., "Use me wisely: AI-driven assessment for LLM prompting skills development," *Educational Technology & Society*, vol. 28, no. 3, 2025. p. 184–201.

[9] E. Theophilou, et al., "Learning to prompt in the classroom to understand AI limits: A pilot study," in *AIxIA 2023 – Advances in Artificial Intelligence*, Springer Nature Switzerland, 2023. p. 481–496.

[10] D. Long and B. Magerko, "What is AI literacy? Competencies and design considerations," in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20)*, 2020. p. 1-16.

[11] UNESCO, "AI competency framework for students," 2024. Available online at https://doi.org/10.54675/JKJB9835.

[12] UNESCO, "AI competency framework for teachers," 2024. Available online at https://doi.org/10.54675/ZJTE2084.

[13] ศูนย์เทคโนโลยีสารสนเทศและกำลังคนอาชีวศึกษา, "สถิติข้อมูลนักเรียนนักศึกษา ปีการศึกษา 2566," 2566. Available online at https://techno.vec.go.th/Default.aspx?tabid=659

[14] D. T. K. Ng, et al., "Design and validation of the AI Literacy Questionnaire: The AILQ," *British Journal of Educational Technology*, vol. 55, no. 3, 2024. p. 1082–1104.

[15] C. S. Chai, et al., "Perceptions of and behavioral intentions towards learning artificial intelligence in primary school students," *Educational Technology & Society*, vol. 24, no. 3, 2021. p. 89–101.

[16] D. Baidoo-Anu and L. O. Ansah, "Education in the Era of Generative Artificial Intelligence (AI): Understanding the Potential Benefits of ChatGPT in Promoting Teaching and Learning," *SSRN Electronic Journal*, 2023. Available online at https://doi.org/10.2139/ssrn.4337484.

[17] I. Molenaar, "Towards hybrid human-AI learning technologies," *European Journal of Education*, vol. 57, no. 4, 2022. p. 632–645.

# โมเดลการประเมินทักษะการรู้เท่าทันข้อมูล
# A Data Literacy Assessment Model

*คมจักร คำเสียง (Komjak Kamsiang)[1] และนลินภัสร์ บำเพ็ญเพียร (Nalinpat Bhumpenpein)[2]*
*ภาควิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล*
*มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ*
*[1]s6707011857110@email.kmutnb.ac.th, [2]nalinpat.b@itd.kmutnb.ac.th*

## บทคัดย่อ

ข้อมูลเป็นทรัพยากรสำคัญที่ขับเคลื่อนองค์กร แต่บุคลากรจำนวนมากยังขาดสมรรถนะที่จำเป็นในการจัดการและวิเคราะห์ข้อมูลอย่างมีประสิทธิภาพ งานวิจัยนี้จึงนำเสนอโมเดลการรู้เท่าทันข้อมูล ซึ่งมีการประเมิน 5 ด้านหลัก ได้แก่ 1) การเข้าถึงและการจัดการข้อมูล 2) การวิเคราะห์และการตีความข้อมูล 3) การสื่อสารและการนำเสนอข้อมูล 4) การใช้ข้อมูลอย่างมีจริยธรรมและถูกต้องตามกฎหมาย และ 5) ทัศนคติและความตระหนักรู้ต่อข้อมูล โมเดลนี้นำไปทดสอบกับกลุ่มตัวอย่างบุคลากรในมหาวิทยาลัยแห่งหนึ่งเป็นจำนวน 168 คน เครื่องมือวิจัยเป็นแบบทดสอบปรนัยชนิดเลือกตอบ 4 ตัวเลือก จำนวน 60 ข้อ แบ่งเป็น 5 ด้าน ด้านละ 12 ข้อ โดยผลการประเมินพบว่าบุคลากรที่ได้รับการทดสอบมีระดับสมรรถนะการรู้เท่าทันข้อมูลในภาพรวมอยู่ในระดับทักษะดีมาก โมเดลนี้สามารถใช้เป็นเครื่องมือที่ช่วยระบุช่องว่างทางทักษะและวางแผนพัฒนาบุคลากรให้สอดคล้องกับยุทธศาสตร์การเป็นองค์กรที่ขับเคลื่อนด้วยข้อมูลได้อย่างเป็นระบบ

**คำสำคัญ**: การรู้เท่าทันข้อมูล โมเดลการประเมินสมรรถนะองค์กรที่ขับเคลื่อนด้วยข้อมูล

## Abstract

*Data is a critical resource that drives organizations forward; however, many personnel still lack the necessary competencies to manage and analyze data effectively. This research therefore presents a data literacy model encompassing five core assessment dimensions: 1) data access and management, 2) data analysis and interpretation, 3) data communication and presentation, 4) ethical and legally compliant data use, and 5) attitudes and awareness toward data. The model was tested with a sample of 168 university personnel. The research instrument consisted of a 60-item multiple-choice test with four answer choices, divided equally into five dimensions of 12 items each. Assessment results indicated that the personnel demonstrated an overall data literacy competency at the excellent level. The findings suggest that this model has the potential to serve as a standardized tool enabling universities to systematically identify skill gaps and develop personnel in alignment with their data-driven organization strategy.*

**Keywords:** data literacy, competency assessment model, data-driven organization

## 1. บทนำ

ในยุคปัจจุบัน ข้อมูลได้กลายเป็นทรัพยากรที่มีมูลค่าสูงสุดและเป็นกลไกสำคัญในการขับเคลื่อนองค์กรสู่ความได้เปรียบทางการแข่งขัน [1] โดยเฉพาะอย่างยิ่งในบริบทของสถาบันอุดมศึกษาที่กำลังปรับตัวสู่การเป็นองค์กรที่ขับเคลื่อนด้วยข้อมูล (Data-Driven Organization) การตัดสินใจเชิงนโยบายและการบริหารจัดการจำเป็นต้องอาศัยข้อมูลที่ถูกต้องและแม่นยำ [2] อย่างไรก็ตาม การขับเคลื่อนองค์กรด้วยข้อมูลไม่ได้ขึ้นอยู่กับโครงสร้างพื้นฐานด้านเทคโนโลยีเพียงอย่างเดียว แต่หัวใจสำคัญคือความพร้อมของทรัพยากรบุคคลในการเข้าถึง จัดการ และใช้ประโยชน์จากข้อมูลได้อย่างมีประสิทธิภาพ [3]

ทักษะการรู้เท่าทันข้อมูล (Data Literacy: DL) จึงกลายเป็นสมรรถนะที่สำคัญยิ่งสำหรับบุคลากรในศตวรรษที่ 21 ซึ่งครอบคลุมทั้งความสามารถในการอ่าน การทำงาน การวิเคราะห์ และการสื่อสารข้อมูลเพื่อการตัดสินใจ หากบุคลากรขาดทักษะดังกล่าวจะส่งผลให้เกิดความเหลื่อมล้ำในการนำข้อมูลมาใช้ประโยชน์ และอาจนำไปสู่ความผิดพลาดในการตีความข้อมูลที่ส่งผลกระทบต่อภาพรวม

ขององค์กร [4] แม้ที่ผ่านมาจะมีการตื่นตัวในเรื่องดิจิทัลทรานส์ฟอร์เมชัน แต่การประเมินระดับทักษะการรู้เท่าทันข้อมูลที่เป็นรูปธรรมและสอดคล้องกับบริบทของมหาวิทยาลัยในประเทศไทยยังมีความจำกัด [5]

ด้วยเหตุนี้ ผู้วิจัยจึงมีวัตถุประสงค์ในการพัฒนาโมเดลการประเมินทักษะการรู้เท่าทันข้อมูล โดยมุ่งเน้นการประเมินสมรรถนะใน 5 มิติหลัก ได้แก่ 1) การเข้าถึงและการจัดการข้อมูล 2) การวิเคราะห์และการตีความข้อมูล 3) การสื่อสารและการนำเสนอข้อมูล 4) การใช้ข้อมูลอย่างมีจริยธรรมและถูกต้องตามกฎหมาย และ 5) ทัศนคติและความตระหนักรู้ต่อข้อมูล กรณีศึกษาบุคลากรมหาวิทยาลัยเทคโนโลยีราชมงคลพระนคร ผลการวิจัยนี้จะเป็นเครื่องมือสำคัญในการวินิจฉัยช่องว่างทางทักษะ (Skill Gap) เพื่อเป็นแนวทางในการกำหนดนโยบายการพัฒนาบุคลากรให้มีสมรรถนะที่สอดคล้องกับความต้องการขององค์กรในยุคเศรษฐกิจดิจิทัลต่อไป

## 2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

### 2.1 แนวคิดเกี่ยวกับการรู้เท่าทันข้อมูล (Data Literacy)

การรู้เท่าทันข้อมูลถูกนิยามว่าเป็นสมรรถนะในการอ่าน การทำงาน การวิเคราะห์ และการสื่อสารข้อมูลเพื่อให้เกิดความเข้าใจและสามารถนำไปตัดสินใจได้อย่างถูกต้อง [6] ซึ่งในบริบทขององค์กรดิจิทัล ทักษะนี้ไม่ได้จำกัดอยู่เพียงบุคคลที่ทำงานด้านสถิติเท่านั้น แต่เป็นทักษะพื้นฐานที่บุคลากรทุกระดับจำเป็นต้องมีเพื่อขับเคลื่อนองค์กรให้ก้าวไปสู่การเป็นองค์กรที่ขับเคลื่อนด้วยข้อมูล (Data-Driven Organization) [7]

### 2.2 KSA Model

โมเดล KSA เป็นโมเดลสำหรับประเมินสมรรถนะบุคลากร โดยงานวิจัยนี้นำมาใช้ประเมินบุคลากรให้ครอบคลุมทั้งพฤติกรรมและทัศนคติต่อการใช้ข้อมูลในการทำงาน [8]

### 2.3 รูปแบบการประเมินการรู้เท่าทันข้อมูล

ผู้วิจัยได้ทำการสังเคราะห์กรอบแนวคิดจากมาตรฐานสากล โดยเฉพาะแนวคิดของ Kang and Kim [9] และกรอบมาตรฐานทักษะดิจิทัลของประเทศไทย (Thailand Digital Competency Framework) [10] ซึ่งแบ่งมิติการประเมินออกเป็น 5 ด้านหลัก ได้แก่ การเข้าถึงและการจัดการข้อมูล การวิเคราะห์และการตีความข้อมูล การสื่อสารและการนำเสนอข้อมูล การใช้ข้อมูลอย่างมีจริยธรรมและถูกต้องตามกฎหมาย และทัศนคติและความตระหนักรู้ต่อข้อมูล

### 2.4 งานวิจัยที่เกี่ยวข้อง

จากการศึกษาพบว่าบุคลากรในมหาวิทยาลัยไทยยังมีระดับทัศนคติและการจัดการข้อมูลวิจัยที่แตกต่างกันตามสายงาน สอดคล้องกับแนวคิดของ Selwyn [11] ที่ระบุว่าการส่งเสริมการรู้เท่าทันข้อมูลต้องคำนึงถึงบริบททางสังคมและการใช้งานจริงในองค์กร ผลจากการศึกษาวรรณกรรมทั้งหมดจึงนำไปสู่การสร้างเครื่องมือวัดระดับวุฒิภาวะ (Maturity Level) ที่แบ่งออกเป็น 5 ระดับ เพื่อใช้ในการวินิจฉัยและพัฒนาบุคลากรได้อย่างแม่นยำ

## 3. วิธีการดำเนินงาน

### 3.1 ประชากรและกลุ่มตัวอย่าง

ประชากร คือ ประชาชนทั่วไป ที่สนใจหรือสามารถประเมินการรู้เท่าทันข้อมูล ซึ่งมีเป็นจำนวนมาก กลุ่มตัวอย่างมีจำนวน 168 คน ประเมินการรู้เท่าทันข้อมูล

### 3.2 โมเดลการวิจัย

โมเดลในการวิจัยที่ผู้วิจัยนำมาใช้งานในการประเมินทักษะการรู้เท่าทันข้อมูลโดยยึดตามกรอบ Data Literacy ทั้ง 5 ด้าน ได้แก่ ด้านที่ 1 การเข้าถึงและการจัดการข้อมูล ด้านที่ 2 การวิเคราะห์ข้อมูลและการตีความ ด้านที่ 3 การใช้ข้อมูลเชิงจริยธรรมและกฎหมาย ด้านที่ 4 การสื่อสารและการนำเสนอข้อมูล ด้านที่ 5 ทัศนคติและความตระหนักต่อข้อมูล โดยเครื่องมือวิจัยเป็นแบบทดสอบปรนัยชนิดเลือกตอบ 4 ตัวเลือก จำนวน 60 ข้อ แบ่งเป็น 5 ด้าน ด้านละ 12 ข้อ

เครื่องมือวิจัยถูกนำไปทดลองใช้กับบุคลากรที่มีลักษณะใกล้เคียงกับกลุ่มตัวอย่างจำนวน 30 คนเพื่อทดสอบคุณภาพของเครื่องมือและวิเคราะห์ค่าสัมประสิทธิ์ความเชื่อมั่นโดยใช้สูตร Kuder-Richardson (KR-20) เนื่องจากเป็นแบบทดสอบที่ให้คะแนนแบบ 0 และ 1 และ

ควรมีการตรวจสอบคุณภาพเครื่องมือด้านความเที่ยงตรง (Validity) ด้วยผู้เชี่ยวชาญ จำนวน 3 ท่าน โดยใช้ค่าดัชนีความสอดคล้อง (IOC) โดย ผลการวิเคราะห์ ได้ค่า IOC เฉลี่ยเท่ากับ 0.96 (สูงกว่าเกณฑ์ 0.50) โดยมีการปรับปรุงข้อคำถามตามคำแนะนำของผู้เชี่ยวชาญ เพื่อให้มีความชัดเจนและตรงตามสมรรถนะที่ต้องการวัด

เนื่องจากงานวิจัยนี้มีตัวแปรสังเกตได้ (Observed Variables) ถึง 60 ข้อ ตามหลักการของ CFA ที่ควรมีกลุ่มตัวอย่างอย่างน้อย 10-20 คนต่อพารามิเตอร์ (ซึ่งต้องใช้กลุ่มตัวอย่าง 600 คนขึ้นไป) ผู้วิจัยจึงเลือกใช้ระบบถ่วงน้ำหนักในระยะเริ่มต้นนี้ก่อน และได้ระบุการทำ CFA ไว้เป็นข้อเสนอแนะสำหรับการวิจัยในอนาคตเมื่อมีการขยายขอบเขตกลุ่มตัวอย่างให้ครอบคลุมตามเกณฑ์ดังกล่าว

จากนั้นนำคะแนนที่ได้มาคำนวณเป็นคะแนนรวมของความพร้อม โดยใช้แนวคิด การให้คะแนนแบบถ่วงน้ำหนักตามระดับความยากของข้อคำถาม ดังนี้

3.2.1 การกำหนดค่าน้ำหนักจากค่าความยาก (p) แบบทดสอบจำนวน 60 ข้อถูกนำไปทดสอบกับกลุ่มตัวอย่างจำนวน 30 คนเพื่อวิเคราะห์คุณภาพข้อสอบ โดยคำนวณค่าความยากรายข้อ (p-value) และจัดระดับความยากของข้อคำถาม จากนั้นกำหนดค่าน้ำหนัก (w) ให้กับข้อคำถามตามระดับความยาก โดยข้อที่มีความยากมากจะได้รับน้ำหนักสูงกว่า และข้อที่มีความยากน้อยจะได้รับน้ำหนักต่ำกว่า [12] ทั้งนี้กำหนดค่าน้ำหนักเป็น 4 ระดับ ได้แก่ 0.8, 0.6, 0.4 และ 0.2 และกำหนดจำนวนข้อในแต่ละระดับความยากจำแนกตามทั้ง 5 ด้านของ Data Literacy ตามตารางค่าน้ำหนักในส่วนวิธีวิจัย

3.2.2 คะแนนรวมถ่วงน้ำหนักของบุคคลที่ $i$ คำนวณจากการนำจำนวนข้อที่ตอบถูกในแต่ละระดับความยาก ($n_ij$) คูณกับค่าน้ำหนักประจำระดับนั้น ($W_j$) แล้วจึงนำผลรวมมาแปลผลตามสมการที่ (1) [13]

$$TotalScore_i = \sum_{j=1}^{n} (n_{ij} \times W_j) \quad (1)$$

3.2.3 ค่าคะแนนเต็มถ่วงน้ำหนักและแปลงเป็นร้อยละในแปลผลจะใช้คะแนนเต็มถ่วงน้ำหนักที่ 30.4 แทนคะแนนดิบ 60 คะแนน ซึ่งคำนวณจากเพดานคะแนนสูงสุดตามโครงสร้างค่าน้ำหนักที่กำหนดไว้ จากนั้นจึงนำมาแปลงเป็นค่าร้อยละเพื่อความสะดวกในการตีความตามสมการที่ (2) [13]

$$PercentageScore_i = \left( \frac{TotalScore_i}{30.4} \right) \times 100 \quad (2)$$

3.2.4 เพื่อให้ผลการประเมินสะท้อนจุดแข็งและข้อควรพัฒนาในแต่ละสมรรถนะได้อย่างละเอียด งานวิจัยจึงกำหนดคะแนนเต็มถ่วงน้ำหนักแยกตามองค์ประกอบทั้ง 5 ด้าน ซึ่งรวมเป็นคะแนนเต็มถ่วงน้ำหนักรวม 30.4 คะแนน ดังรายละเอียดต่อไปนี้ ด้านที่ 1 เท่ากับ 6.2 ด้านที่ 2 เท่ากับ 6.2 ด้านที่ 3 เท่ากับ 6.0 ด้านที่ 4 เท่ากับ 6.0 และด้านที่ 5 เท่ากับ 6.0 สำหรับการคำนวณคะแนนร้อยละรายด้านใช้สมการที่ (3) [14]

$$SubScorePercentage_k = \left( \frac{WeightedScore_k}{FullWeightedScore_k} \right) \times 100 \quad (3)$$

3.2.5 เกณฑ์การแปลผลระดับสมรรถนะเมื่อได้ผลการคำนวณคะแนนร้อยละทั้งในภาพรวมและรายด้านแล้ว จะนำค่าที่ได้มาเปรียบเทียบกับเกณฑ์มาตรฐานเพื่อจำแนกสภาวะสมรรถนะออกเป็น 5 ระดับ โดยมีรายละเอียดการแปลผลดังนี้ ร้อยละ 80.00 – 100.00 หมายถึง มีสมรรถนะในระดับ ทักษะดีมาก (Excellent) ร้อยละ 70.00 – 79.99 หมายถึง ในระดับ ทักษะดี (Good) ร้อยละ 60.00 – 69.99 หมายถึง ในระดับ ทักษะปานกลาง (Moderate) ร้อยละ 50.00 – 59.99 หมายถึง ในระดับ ทักษะน้อย (Low) [ต่ำกว่าร้อยละ 50.00 หมายถึง มี ทักษะน้อยมาก (Very Low) [15]

ทั้งนี้ ผู้วิจัยได้ตรวจสอบคุณภาพของเครื่องมือ (โมเดล) พบว่า มีค่าสัมประสิทธิ์ความเชื่อมั่น (Reliability) เท่ากับ 0.88 ซึ่งสูงกว่าเกณฑ์มาตรฐานที่ยอมรับได้ (0.70) สะท้อนถึงความสอดคล้องภายในและความคงเส้นคงวาของเครื่องมือในระดับสูงมาก เหมาะสมสำหรับนำไปใช้ประเมินสมรรถนะด้านทักษะความเข้าใจและใช้ข้อมูลอย่างแม่นยำ

## 4. ผลการดำเนินการวิจัย

### 4.1 ผลการประเมินทักษะการรู้เท่าทันข้อมูล

ผลการวิเคราะห์ภาพรวม พบว่าคะแนนเฉลี่ยถ่วงน้ำหนักเท่ากับ 25.9 คะแนน จากคะแนนเต็ม 30.40

คะแนน (คิดเป็นร้อยละ 85.09 ) ซึ่งแสดงให้เห็นว่าระดับสมรรถนะโดยรวมของผู้เข้ารับการทดสอบอยู่ในระดับทักษะดีมาก (Excellent) ตามเกณฑ์การแปลผลที่กำหนดไว้

เมื่อจำแนกผลการประเมินสมรรถนะด้านทักษะความเข้าใจและใช้ข้อมูล (Data Literacy) ออกเป็น 5 ด้าน ตามกรอบงานวิจัยที่กำหนด พบว่าค่าเฉลี่ยคะแนนถ่วงน้ำหนักจากกลุ่มตัวอย่างจำนวน 168 คน มีผลการประเมินอยู่ในระดับ ทักษะดีมาก (Excellent) ในทุกมิติ โดยมีรายละเอียดรายด้านดังนี้ ด้านที่ 1 และด้านที่ 2 มีผลการประเมินสูงที่สุดโดยมีคะแนนเต็มถ่วงน้ำหนักด้านละ 6.2 คะแนน ด้านที่ 3 ด้านที่ 4 และด้านที่ 5 มีคะแนนเต็มถ่วงน้ำหนักด้านละ 6.0 คะแนน ทั้งนี้ ผลการวิเคราะห์พบว่ากลุ่มตัวอย่างส่วนใหญ่ มีระดับสมรรถนะรายด้าน ที่สอดคล้องกันอย่างมีนัยสำคัญ โดยรายละเอียดคะแนนเฉลี่ยและค่าร้อยละของแต่ละด้านแสดงดังตารางที่ 1 เมื่อพิจารณาและสรุปข้อมูลทักษะรายบุคคลพบว่าส่วนมากอยู่ในระดับที่สูง ผู้วิจัยจึงมุ่งเน้นการวิเคราะห์กลุ่มที่มีระดับความพร้อมต่ำกว่าเกณฑ์มาตรฐาน (ร้อยละ 70) เพื่อระบุจุดอ่อนที่ควรได้รับการพัฒนา ดังแสดงในตารางที่ 2

**ตารางที่ 1** คะแนนเฉลี่ยและร้อยละของสมรรถนะ

| ด้าน | คะแนนเต็มถ่วงน้ำหนัก | คะแนนเฉลี่ย | ร้อยละ |
|---|---|---|---|
| ด้านที่ 1 | 6.2 | 5.5 | 89.25 |
| ด้านที่ 2 | 6.2 | 4.8 | 78.65 |
| ด้านที่ 3 | 6.0 | 5.3 | 88.43 |
| ด้านที่ 4 | 6.0 | 5.0 | 83.08 |
| ด้านที่ 5 | 6.0 | 5.2 | 86.07 |
| ภาพรวม | 30.4 | 25.9 | 85.09 |

**ตารางที่ 2** จำนวนและร้อยละจำแนกรายด้าน

| ด้าน | จำนวนคนที่ไม่ถึงเกณฑ์ 70% | ร้อยละ |
|---|---|---|
| ด้านที่ 1 | 11 | 6.55 |
| ด้านที่ 2 | 35 | 20.83 |
| ด้านที่ 3 | 20 | 11.90 |
| ด้านที่ 4 | 30 | 17.86 |
| ด้านที่ 5 | 25 | 14.88 |
| ภาพรวม | 19 | 11.31 |

**ตารางที่ 3** การวิเคราะห์ช่องว่างความพร้อม (Gap Analysis) และระดับความรุนแรงในกลุ่มวิกฤต

| ด้าน | กลุ่มวิกฤต (< 60%) (คน) | กลุ่มเฝ้าระวัง (60-69%) (คน) | คะแนนเฉลี่ยกลุ่มไม่ผ่าน (%) | ช่องว่างความพร้อม (Gap) |
|---|---|---|---|---|
| ด้านที่ 1 | 4 | 7 | 60.12 | 9.88 |
| ด้านที่ 2 | 26 | 9 | 50.32 | 19.68 |
| ด้านที่ 3 | 12 | 8 | 49.50 | 20.50 |
| ด้านที่ 4 | 18 | 12 | 52.11 | 17.89 |
| ด้านที่ 5 | 15 | 10 | 48.67 | 21.33 |

จากตารางที่ 2 พบว่า ในภาพรวมมีผู้เข้าทดสอบที่มีความพร้อมไม่ถึงเกณฑ์ร้อยละ 70 เมื่อพิจารณารายด้าน พบว่า ด้านที่ 2 การวิเคราะห์ข้อมูลและการตีความมีจำนวนผู้ไม่ผ่านเกณฑ์สูงที่สุด คือ 35 คน รองลงมาคือ ด้านที่ 4 การสื่อสารและการนำเสนอข้อมูล จำนวน 30 คน และด้านที่ 5 ทัศนคติและความตระหนักต่อข้อมูล จำนวน 25 คน ตามลำดับ แม้ภาพรวมสมรรถนะอยู่ในเกณฑ์ดีมาก แต่พบช่องว่าง (Gap) ต่ำกว่าเกณฑ์ร้อยละ 70 โดยเฉพาะ ด้านการวิเคราะห์และสื่อสารข้อมูล ที่ต้องพัฒนาเร่งด่วน ผู้วิจัยจึงเสนอแนวทางขับเคลื่อนดังนี้ (1) วินิจฉัยจุดบกพร่องตรวจสอบผลทดสอบรายข้อเพื่อระบุทักษะที่ขาดแคลนและออกแบบหลักสูตรให้ตรงจุด (2) สร้างระบบนิเวศการเรียนรู้ จัดทำ Roadmap พัฒนาตามความจำเป็นรายกลุ่ม เน้นฝึกปฏิบัติ (Hands-on) เพื่อสร้างวัฒนธรรมการใช้ข้อมูลอย่างยั่งยืน และจากตารางที่ 3 ผลการวิเคราะห์เชิงลึกพบประเด็นสำคัญคือ ด้านที่ 5 ทัศนคติและความตระหนักต่อข้อมูลแม้จะมีจำนวนคนไม่ผ่านเกณฑ์น้อยกว่าด้านที่ 2 แต่กลับมีค่าช่องว่างความพร้อม (Gap) กว้างที่สุดถึงร้อยละ 21.33 สะท้อนว่าผู้ที่ไม่ผ่านเกณฑ์ในด้านนี้มีระดับความพร้อมที่ต่ำกว่าเป้าหมายมากที่สุด นอกจากนี้ยังพบว่า ด้านที่ 2 การวิเคราะห์ข้อมูล มีจำนวนผู้ที่อยู่ในระดับวิกฤต (คะแนนต่ำกว่าร้อยละ 60) สูงถึง 26 คน ซึ่งถือเป็นสัดส่วนที่สูงที่สุดเมื่อเทียบกับด้านอื่น จึงสะท้อนให้เกิดแนวทางในการแก้ปัญหาไว้ ดังนี้ (1) เร่งด่วนใน (ด้านที่ 2) จัดอบรมทักษะการวิเคราะห์ข้อมูลขั้นพื้นฐานแบบเข้มข้น

สำหรับกลุ่มวิกฤต 26 คน เพื่อลดจำนวนผู้ไม่ผ่านเกณฑ์ในภาพรวม (2.) เจาะลึก (ด้านที่ 5) พัฒนาทัศนคติและความตระหนักผ่านกิจกรรมสร้างแรงจูงใจ (Awareness Campaign) หรือ Case Study เชิงจริยธรรมข้อมูล เพื่อปิดช่องว่างความพร้อม (Gap) ที่กว้างให้แคบลง (3) ติดตามผลใช้ระบบการให้คำปรึกษาแบบตัวต่อตัวหรือวัดซ้ำเมื่อมีการวัดทักษะรายครั้งต่อ ๆ ไป สำหรับผู้ที่มีค่า Gap สูง เพื่อให้การพัฒนาเป็นไปตามศักยภาพรายบุคคล

**5 สรุปผลและอภิปรายผล**

จากการวิเคราะห์ความพร้อมด้านทักษะความเข้าใจและใช้ข้อมูล (Data Literacy) ของบุคลากรจำนวน 168 คน โดยการใช้โมเดลในการประเมินทักษะการรู้เท่าทันข้อมูลที่ผู้วิจัยได้พัฒนาขึ้นสามารถสะท้อนระดับทักษะของบุคคล–ด้อย่างเป็นระบบ โดยใช้การวัดผลค่าแนนถ่วงน้ำหนักตามระดับความยาก (0.2 - 0.8) สรุปผลได้ดังนี้

ภาพรวมความพร้อมกลุ่มเป้าหมายมีความพร้อมเฉลี่ยอยู่ในระดับสูง (ร้อยละ 85.10) โดยมิติด้านการเข้าถึงข้อมูลมีคะแนนสูงที่สุด (ร้อยละ 89.25) กลุ่มที่ควรได้รับการพัฒนามีผู้เข้าทดสอบร้อยละ 11.31 (19 คน) ที่มีผลการประเมินไม่ถึงเกณฑ์มาตรฐาน (ร้อยละ 70) มิติวิกฤต ด้านการวิเคราะห์ข้อมูลและการตีความ (ด้านที่ 2) เป็นมิติที่มีจำนวนคนตกเกณฑ์วิกฤต (ต่ำกว่าร้อยละ 60) มากที่สุดถึง 26 คน และด้านทัศนคติข้อมูล (ด้านที่ 5) มีค่าช่องว่างความพร้อม (Gap) กว้างที่สุดถึง ร้อยละ 21.33 [16]

จากผลการวิจัยมีประเด็นสำคัญที่ผู้วิจัยนำมาอภิปรายผล ดังนี้ (1) ปัญหาทักษะการวิเคราะห์ข้อมูลในระดับวิกฤตการที่มิติด้านการวิเคราะห์ข้อมูลและการตีความมีจำนวนผู้ที่อยู่ในระดับวิกฤตสูงที่สุด (26 คน) สอดคล้องกับรายงานของ สำนักงานคณะกรรมการดิจิทัลเพื่อเศรษฐกิจและสังคมแห่งชาติ ที่ระบุว่าทักษะการคิดเชิงวิเคราะห์และการจัดการข้อมูลขนาดใหญ่เป็นสมรรถนะหลักที่บุคลากรในภาครัฐและเอกชนของไทยยังขาดความชำนาญ เนื่องจากเป็นทักษะที่ต้องอาศัยทั้งทรรศนะเชิงสถิติและการตีความบริบทของข้อมูล ซึ่งซับซ้อนกว่าทักษะการใช้งานดิจิทัลทั่วไป [17] และ (2) ช่องว่างความพร้อมในด้านทัศนคติ (ด้านที่ 5) ประเด็นที่น่าสนใจคือมิติด้านทัศนคติมีความถี่ของคนตกเกณฑ์น้อยกว่าด้านอื่นแต่มีค่า Gap กว้างที่สุด สะท้อนว่าคนที่ไม่ผ่านในด้านนี้มีความเข้าใจคลาดเคลื่อนต่อการใช้ข้อมูลอย่างรุนแรง "ทัศนคติและความตระหนัก (Mindset)" คือหัวใจสำคัญในการขับเคลื่อนองค์กรด้วยข้อมูล (Data-Driven Organization) หากบุคลากรขาดความเข้าใจในคุณค่าของข้อมูล แม้จะมีความสามารถเชิงเทคนิคก็ไม่สามารถนำข้อมูลไปสร้างประโยชน์สูงสุดให้กับองค์กรได้ [18]

สำหรับข้อเสนอแนะเพื่อการนำไปใช้การพัฒนาบุคลากรรายบุคคล (Individual Development Plan) องค์กรควรจัดโปรแกรมฝึกอบรมแบบเร่งด่วน (Fast-track) สำหรับกลุ่มวิกฤตจำนวน 10 รายที่มีคะแนนภาพรวมต่ำกว่าร้อยละ 60 โดยเฉพาะผู้เข้าทดสอบที่มีคะแนนวิกฤตในทุกด้าน (ID 19, 21, 125 จากตารางเก็บข้อมูล) เพื่อยกระดับทักษะพื้นฐานให้ทันต่อการปฏิบัติงาน การเสริมสร้างทักษะการตีความข้อมูล ควรจัดทำคู่มือหรือ Workshop ที่เน้นการนำข้อมูลจริงในหน่วยงานมาฝึกวิเคราะห์และตีความผลลัพธ์ แทนการอบรมเชิงทฤษฎีเพียงอย่างเดียว เพื่อลดจำนวนคนตกเกณฑ์ในด้านที่ 2

สำหรับข้อเสนอแนะเพื่อการวิจัยครั้งต่อไปควรมีการศึกษารูปแบบการเรียนรู้ (Learning Styles) ของกลุ่มที่มีระดับความพร้อมในขั้นวิกฤต เพื่อออกแบบเครื่องมือการสอนที่เหมาะสมกับข้อจำกัดของบุคลากรเฉพาะกลุ่ม ควรศึกษาปัจจัยเชิงสภาพแวดล้อมที่ส่งผลต่อทัศนคติและความตระหนักในการใช้ข้อมูลของบุคลากรเพื่อหาแนวทางในการสร้างวัฒนธรรมข้อมูล (Data Culture) ในหน่วยงานอย่างยั่งยืน ตามแนวทางของ Digital Economy Promotion Agency [19]

## เอกสารอ้างอิง

[1] พนาสิน จึงสวนันทน์. (2568). กลยุทธ์ความได้เปรียบทางการแข่งขันด้วยเทคโนโลยีสารสนเทศสำหรับธุรกิจยุคเศรษฐกิจดิจิทัล. วารสาร มจร พุทธปัญญาปริทรรศน์, 10(1), 751–762.

[2] ภัทธิรา แก้วเกิด, & ปราโมช ธรรมกรณ์. (2566). บิ๊กดาต้ากับการดำเนินธุรกิจยุคดิจิทัล. วารสารพัฒนาธุรกิจและอุตสาหกรรม, 3(1), 92–99.

[3] นิศาชล จำนนงศรี และคณะ. (2567). พฤติกรรมและทัศนคติด้านการจัดการและการแบ่งปันข้อมูลวิจัยของนักวิจัยในมหาวิทยาลัยไทย. วารสารวิจัยสารสนเทศ, 42(1), 114–132.

[4] Pangrazio, L., & Selwyn, N. (2023). Critical data literacies: Rethinking data and everyday life. MIT Press.*

[5] Open Data Institute (ODI). (2024). Empowering Thailand's digital government with open data. ODI, London.*

[6] ธนกฤต พรหมเมศ (2565): ความฉลาดรู้ด้านสถิติ ความฉลาดรู้ด้านข้อมูล และความฉลาดรู้ด้านสารสนเทศ: สิ่งที่แตกต่างและเหมือนกัน.

[7] วีระยุทธ วรเวชกุล (2565): "การพัฒนาทักษะดิจิทัลของบุคลากรภาครัฐไทยสู่การเป็นรัฐบาลดิจิทัล" เน้นการเปลี่ยนผ่านองค์กรภาครัฐให้ใช้ข้อมูลในการทำงาน (Data-Driven Government)

[8] A KSA system for competency-based assessment of clinicians' professional development in China and quality gap analysis

[9] Kang, S., & Kim, J. (2024). Synthesizing Digital Competency Frameworks: A Systematic Review of Data Literacy in Higher Education. Journal of Educational Technology & Society, 27(1), 45-62.

[10] Wongchai, S., & Limthong, W. (2023). Development of digital literacy indicators for Thai teachers. Journal of Education Studies, 51(2), 85–103.*

[11] ธีรเดช กาญจนสำราญวงศ์, และ คณะ. (2568). การวิเคราะห์องค์ประกอบและระดับสมรรถนะดิจิทัลของบุคลากรในสถาบันอุดมศึกษา: การศึกษาเปรียบเทียบระหว่างสายวิชาการและสายสนับสนุน. วารสารนวัตกรรมการเรียนรู้และเทคโนโลยี, 5(1), 12-28.

[12] Nguyen, T., & Liao, J. (2023). Data literacy assessments A systematic literature review. Assessment in Education, 30(4), 432–456.

[13] บุญชม ศรีสะอาด. (2560). การวิจัยเบื้องต้น (พิมพ์ครั้งที่ 10). สุวีริยาสาส์น.

[14] พลอยไพลิน กรีธาพล, และ วิศรุต บัวทอง. (2568). การพัฒนาตัวชี้วัดและค่าน้ำหนัก องค์ประกอบสมรรถนะด้านข้อมูลสำหรับบุคลากรสายสนับสนุนใน สถาบันอุดมศึกษา. วารสารศิลปากรศึกษาศาสตร์วิจัย, 17(1), 45–60.

[15] สำนักงานคณะกรรมการดิจิทัลเพื่อเศรษฐกิจและสังคมแห่งชาติ. (2565). รายงานดัชนีชี้วัดระดับความพร้อมทักษะดิจิทัลของประเทศไทย ประจำปี 2565. กระทรวงดิจิทัลเพื่อเศรษฐกิจและสังคม.

[16] สำนักงานคณะกรรมการดิจิทัลเพื่อเศรษฐกิจและสังคมแห่งชาติ. (2568). รายงานดัชนีชี้วัดระดับความพร้อมทักษะดิจิทัลของประเทศไทย ประจำปี 2567-2568. กระทรวงดิจิทัลเพื่อเศรษฐกิจและสังคม.

[17] Tanaka, Y., & Li, H. (2024). Proposed data literacy competency framework through literature analysis. Journal of Information Literacy Research, 12(1), 77–98.*

[18] ไพโรจน์ ภู่ต้อง. (2564). การพัฒนาทรัพยากรมนุษย์ในยุคดิจิทัล. วารสารการบริหารและนวัตกรรมท้องถิ่น, 3(2), 15-30.

[19] Open Data Institute (ODI). (2024). Empowering Thailand's digital government with open data. ODI, London

# การตรวจจับ GPS Spoofing แบบสามคลาสสำหรับ UAV ด้วยโมเดล Attention-based CNN-LSTM

# Three-Class GPS Spoofing Detection for UAVs Using an Attention-based CNN-LSTM

*สิริกาญจน์ เนตรพุกกณะ (Sirikarn Netpookgana) ณราวิชญ์ คนใจบุญ (Naravit Conjaiboon)*
*และจักรชัย โสอินทร์ (Chakchai So-In)*
*วิทยาลัยการคอมพิวเตอร์ มหาวิทยาลัยขอนแก่น*
*paansirikarn@kkumail.com, naravit.c@kkumail.com, chakso@kku.ac.th*

## บทคัดย่อ

*อากาศยานไร้คนขับพึ่งพาสัญญาณ GPS ในการระบุตำแหน่งและควบคุมการบิน อย่างไรก็ตาม สัญญาณ GPS มีความเสี่ยงต่อการถูกปลอมแปลง ซึ่งอาจทำให้อากาศยานไร้คนขับรับรู้ตำแหน่งผิดพลาดและส่งผลต่อความถูกต้องของการนำทาง งานวิจัยนี้จึงนำเสนอโมเดล Attention-based CNN-LSTM สำหรับตรวจจับการปลอมแปลงสัญญาณ GPS แบบสามคลาส ได้แก่ BENIGN, SPOOF และ SUSPECTED โดยผสานเลเยอร์ CNN สำหรับสกัดคุณลักษณะจากข้อมูลบันทึกสถานะการบิน เลเยอร์ LSTM สำหรับเรียนรู้ความสัมพันธ์เชิงเวลา และกลไก Attention เพื่อถ่วงน้ำหนักให้กับช่วงที่มีนัยสำคัญต่อการจำแนก ในการประเมินผลใช้ 5-fold cross-validation สำหรับฝึกและประเมินประสิทธิภาพของโมเดล พร้อมทั้งเปรียบเทียบกับ PCA-CNN-LSTM ที่พัฒนาขึ้นใหม่ภายใต้ระเบียบวิธีการทดลองเดียวกัน ผลการทดลองแสดงให้เห็นว่าโมเดลที่นำเสนอมีค่า Accuracy 0.9812 และ Macro F1 0.9570 จากการประเมินแบบ 5-fold และมีค่า Accuracy 0.9885 และ Macro F1 0.9829 บนชุดทดสอบสุดท้าย โดยผลลัพธ์ดังกล่าวบ่งชี้ว่าโมเดลที่นำเสนอสามารถจำแนกสถานะการบินทั้งสามคลาสได้ดีภายใต้ข้อมูลจำลองที่ใช้ในการศึกษา*

**คำสำคัญ:** การปลอมแปลงสัญญาณ GPS, โดรน, ArduPilot SITL, CNN-LSTM, กลไกความสนใจตนเอง, การจำแนกประเภทอนุกรมเวลา

## Abstract

*Unmanned Aerial Vehicles (UAVs) rely on GPS signals for positioning and flight control. However, GPS signals are vulnerable to spoofing attacks, which may lead to position estimation errors and reduced navigation accuracy. This study proposes an attention-based CNN-LSTM model for three-class GPS spoofing detection in UAVs. The model integrates CNN layers to extract salient features from flight-log data, LSTM layers to learn temporal dependencies, and an attention mechanism to emphasize informative time steps, enabling three-class classification of flight states into BENIGN, SPOOF, and SUSPECTED. The model is evaluated using 5-fold cross-validation and compared with a re-implemented PCA-CNN-LSTM baseline under the same dataset and experimental protocol. Experimental results show that the proposed model achieves an accuracy of 0.9812 and a macro F1-score of 0.9570 under 5-fold cross-validation, and an accuracy of 0.9885 and a macro F1-score of 0.9829 on the final test set. These results indicate that the proposed model is effective for three-class GPS spoofing detection under the simulated conditions used in this study.*

**Keywords:** GPS spoofing, UAV, ArduPilot SITL, CNN-LSTM, Self-Attention, Time-Series Classification.

## 1. บทนำ

อากาศยานไร้คนขับ (UAV) ถูกนำมาใช้งานอย่างกว้างขวางในหลายด้าน ไม่ว่าจะเป็นการตรวจการณ์ การเกษตร การสำรวจพื้นที่ และการขนส่ง การทำงานของ UAV นั้นพึ่งพาสัญญาณ GPS เป็นหลักในการระบุตำแหน่งและควบคุมเส้นทางบิน อย่างไรก็ตาม สัญญาณ GPS สามารถถูกปลอมแปลงได้ ซึ่งอาจทำให้ UAV รับรู้ตำแหน่งผิดพลาดหรือบินออกนอกเส้นทางโดยไม่รู้ตัว

ดังนั้น การตรวจจับการปลอมแปลงสัญญาณ GPS จึงเป็นเรื่องสำคัญที่ส่งผลโดยตรงต่อความปลอดภัยและความน่าเชื่อถือของระบบ UAV [1]–[10] งานวิจัยที่ผ่านมามีแนวทางที่หลากหลาย ดังสรุปในตารางที่ 1

งานวิจัยที่ผ่านมาได้เสนอวิธีตรวจจับการปลอมแปลงสัญญาณ GPS ไว้หลายแนวทาง ทั้งการวิเคราะห์สัญญาณ GPS โดยตรง การใช้ข้อมูลบันทึกสถานะการบิน และการประยุกต์ใช้ Machine Learning หรือ Deep Learning เพื่อเรียนรู้รูปแบบความผิดปกติจากข้อมูลบันทึกสถานะการบิน ตัวอย่างเช่น Sun et al. [1] เสนอโมเดล PCA-CNN-LSTM โดยคัดเลือกฟีเจอร์จากสัญญาณ GPS ดิบด้วย Spearman’s rank correlation เหลือ 36 รายการ และใช้ PCA (Principal Component Analysis) ลดมิติข้อมูลลงเหลือ 15 มิติ

ดังนั้นงานวิจัยนี้จึงใช้รายการฟีเจอร์ดังกล่าวเป็นแนวทางตั้งต้น โดยเทียบกับฟีเจอร์ที่มีนิยามใกล้เคียงกันและตัดรายการที่ไม่เกี่ยวกับพฤติกรรมการบินออก จนเหลือ 33 รายการที่ครอบคลุม GPS, IMU (Inertial Measurement Unit), EKF (Extended Kalman Filter), barometer และ magnetometer ก่อนนำมาใช้พัฒนาโมเดลอ้างอิงใหม่ภายใต้ชุดข้อมูลและเงื่อนไขการทดลองเดียวกัน เพื่อเปรียบเทียบกับแบบจำลองที่นำเสนอในงานวิจัยนี้

Agyapong et al. [2] ใช้ข้อมูลบันทึกสถานะการบินร่วมกับโมเดล LSTM และ Wei et al. [6] ใช้ข้อมูลที่สะท้อนลักษณะการควบคุมอากาศยานเพื่อตรวจจับความผิดปกติในเชิงพฤติกรรม แม้งานเหล่านี้จะแสดงให้เห็นถึงศักยภาพในการตรวจจับการโจมตีได้ดี แต่ส่วนใหญ่ยังคงกำหนดปัญหาในรูปแบบการจำแนกสองกลุ่ม และยังไม่ได้แยกช่วงเริ่มต้นของการโจมตีออกเป็นคลาสเฉพาะอย่างชัดเจน ด้วยเหตุนี้ งานวิจัยนี้จึงกำหนดปัญหาเป็นการจำแนกแบบสามคลาส เพื่อให้โมเดลสามารถแยกช่วงบินปกติ ช่วงเริ่มต้นของการโจมตี และช่วงที่การโจมตีปรากฏชัดออกจากกันได้อย่างชัดเจน

ในการพัฒนาโมเดล งานวิจัยนี้นำเสนอโมเดล Attention-based CNN-LSTM โดยใช้เลเยอร์ CNN สำหรับสกัดฟีเจอร์สำคัญจากข้อมูลบันทึกสถานะการบิน ใช้เลเยอร์ LSTM สำหรับเรียนรู้การเปลี่ยนแปลงของข้อมูลตามเวลา และใช้กลไก attention เพื่อให้โมเดลให้น้ำหนักกับช่วงที่มีนัยสำคัญต่อการจำแนก พร้อมเปรียบเทียบผลกับโมเดล PCA-CNN-LSTM ภายใต้เงื่อนไขการทดลองเดียวกัน ทั้งในด้านชุดข้อมูล การสร้างหน้าต่างเวลา การกำหนดป้ายกำกับ และการตั้งค่าหลักในการฝึกสอน พร้อมรายงานผลทั้งแบบ 5-fold cross-validation และบนชุดทดสอบขั้นสุดท้าย

## 2. งานวิจัยที่เกี่ยวข้อง

### 2.1 งานวิจัยด้าน Machine Learning แบบดั้งเดิม

งานวิจัยกลุ่มแรกใช้วิธีการบน Machine Learning แบบดั้งเดิมสำหรับตรวจจับ GPS spoofing ของ UAV โดยอาศัยการออกแบบคุณลักษณะจากข้อมูล GPS หรือข้อมูลการบิน ตัวอย่างเช่น Meng et al. [7] เสนอแนวทางบนพื้นฐานของ linear regression รองรับทั้งการตรวจจับและการรับมือ ขณะที่ Shafique et al. [8] ศึกษาแบบจำลอง Machine Learning หลายชนิดสำหรับตรวจจับ signal spoofing attack นอกจากนี้ Talaei Khoei et al. [9] เสนอ dynamic selection techniques สำหรับเลือกตัวจำแนกที่เหมาะสม และ Nayfeh et al. [10] ศึกษาการสร้างโมเดลจากคุณลักษณะของ GPS เพื่อรองรับทั้งการตรวจจับและการจำแนก

งานในกลุ่มนี้ชี้ให้เห็นว่าการออกแบบคุณลักษณะอย่างเหมาะสมยังคงมีบทบาทสำคัญต่อการตรวจจับการโจมตี อย่างไรก็ตาม วิธีการเหล่านี้อาศัยการออกแบบคุณลักษณะโดยผู้วิจัยเป็นหลัก

**ตารางที่ 1** เปรียบเทียบงานวิจัยที่เกี่ยวข้อง

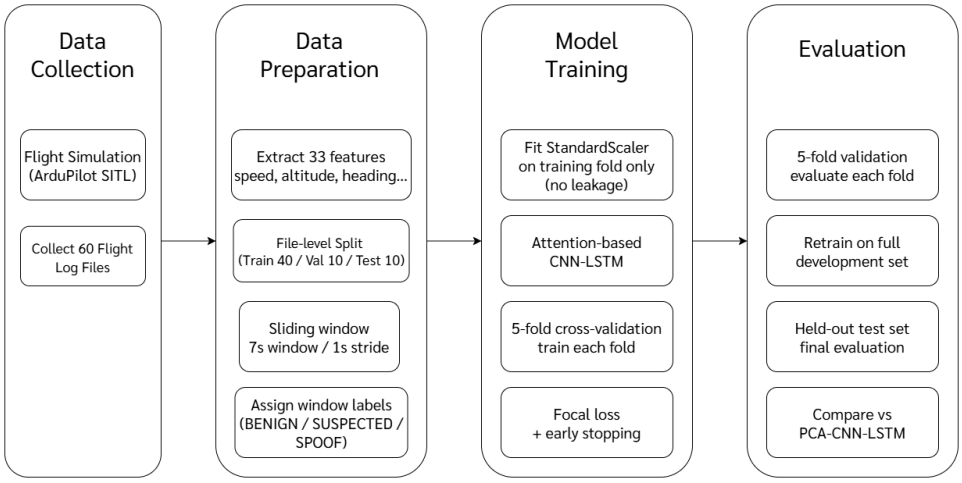
| อ้างอิง | วิธีหลัก | ข้อมูลนำเข้าหลัก | ขอบเขตของงานวิจัย |
|---|---|---|---|
| [1] Sun et al. | PCA-CNN-LSTM | Real GPS signal data from small UAVs | Binary spoofing detection |
| [2] Agyapong et al. | LSTM-based binary classifier and LSTM autoencoder for one-class classification | Telemetry and flight-log benchmark data | Binary detection and one-class detection |
| [3] Eshmawi et al. | Stacked ensemble | GPS spoofing data for small UAVs | Spoofing detection |
| [4] Al-Sabbagh et al. | GA-driven Deep Learning framework | GPS spoofing dataset | Spoofing detection |
| [5] Sung et al. | 1D-CNN | UAV GPS spoofing data | Early spoofing detection |
| [6] Wei et al. | ConstDet (control-semantics-based ML) | Real flight data and control semantics | Spoofing detection |
| [7] Meng et al. | Linear regression-based approach | UAV route / trajectory information | Spoofing detection and countermeasure |
| [8] Shafique et al. | Machine Learning models | GPS signal characteristics | Spoofing detection |
| [9] Talaei Khoei et al. | Dynamic selection techniques | Engineered GPS spoofing features | Spoofing detection |
| [10] Nayfeh et al. | Machine-learning modeling of GPS features | GPS features in controlled test scenarios | Spoofing detection and classification |
| งานวิจัยนี้ | Attention-based CNN-LSTM | GPS features and sensor features | 3-class classification (BENIGN / SPOOF / SUSPECTED) |

นอกจากนี้ยังสามารถจำแนกได้เพียงภาวะปกติกับภาวะถูกโจมตี จึงอธิบายช่วงที่เริ่มถูกโจมตีได้ไม่ดีนัก เมื่อเทียบกับโมเดลที่เรียนรู้จากข้อมูล time-series โดยตรง

**2.2 งานวิจัยด้าน Deep Learning**

งานวิจัยในช่วงหลังหันมาประยุกต์ใช้ Deep Learning มากยิ่งขึ้น เพื่อให้โมเดลสามารถเรียนรู้รูปแบบเชิงเวลาและความสัมพันธ์ระหว่างตัวแปรจากข้อมูลการบินได้ดีขึ้น โดยโมเดลสามารถสกัดคุณลักษณะสำคัญได้เองโดยไม่จำเป็นต้องกำหนดล่วงหน้า (Manual Feature Engineering) เช่น Sun et al. [1] ได้นำเสนอโมเดล PCA-CNN-LSTM สำหรับตรวจจับ GPS Spoofing ใน UAV ขนาดเล็ก ขณะที่ Agyapong et al. [2] ได้ใช้ LSTM ทั้งในรูปแบบการจำแนกสองกลุ่ม (binary classification) และการฝึกฝนจากข้อมูลสภาวะปกติเพียงอย่างเดียว (one-class classification) ในขณะที่ Eshmawi et al. [3] ใช้วิธีการรวมโมเดล (ensemble method) และ Al-Sabbagh et al. [4] ได้ประยุกต์ใช้ Genetic Algorithm ในการปรับแต่งพารามิเตอร์ของโมเดลโดยอัตโนมัติ นอกจากนี้ Sung et al. [5] ได้นำ 1D-CNN มาใช้ในการตรวจจับระยะเริ่มต้น

แม้ว่างานวิจัยเหล่านี้จะแสดงให้เห็นว่า Deep Learning ช่วยเพิ่มประสิทธิภาพการตรวจจับได้ แต่ยังไม่ได้กำหนดให้ช่วงเริ่มต้นของการโจมตีเป็นคลาสเฉพาะอย่างชัดเจนในโจทย์การจำแนกหลายคลาส ดังนั้น งานวิจัยนี้จึงแบ่งปัญหาออกเป็น 3 คลาส ได้แก่ BENIGN, SPOOF และ SUSPECTED โดยนำเสนอโมเดล Attention-based CNN-LSTM และใช้เกณฑ์การแบ่งข้อมูลในระดับไฟล์ (file-level split) เพื่อป้องกันการเกิดข้อมูลรั่วไหล (Data Leakage)




**ภาพที่ 1**: ภาพรวมสถาปัตยกรรมของระบบ

## 3. วิธีการดำเนินงาน

งานวิจัยนี้พัฒนาโมเดลตรวจจับการปลอมแปลงสัญญาณ GPS ของ UAV โดยใช้โปรแกรมจำลองการบิน ArduPilot SITL

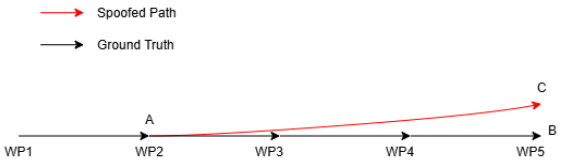
### 3.1 ภาพรวมสถาปัตยกรรมของระบบ

ระบบประกอบด้วย 6 ขั้นตอน ได้แก่ การจำลองการบินและเก็บข้อมูล การสกัดคุณลักษณะจากข้อมูลการบิน การแบ่งชุดข้อมูลแบบแยกไฟล์และการปรับมาตรฐาน การสร้างหน้าต่างเวลาและกำหนดป้ายกำกับ การฝึกและจำแนกด้วยโมเดล และการประเมินผล ดังแสดงในภาพที่ 1

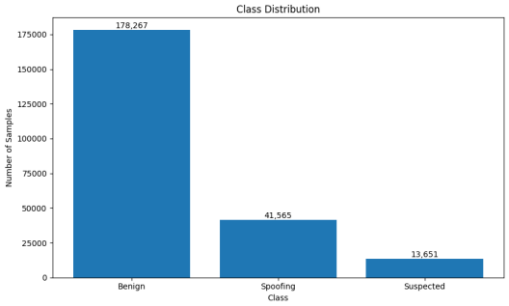
### 3.2 การเก็บข้อมูลและชุดข้อมูล

ข้อมูลที่ใช้ในงานวิจัยนี้ถูกสร้างขึ้นจากสภาพแวดล้อมจำลอง ArduPilot Software-In-The-Loop (SITL) ซึ่งเป็นเครื่องมือจำลองระบบควบคุมอากาศยานไร้คนขับ ช่วยให้สามารถควบคุมเงื่อนไขการบินและกำหนดช่วงเริ่มต้นของการโจมตีได้อย่างเป็นระบบ จึงเหมาะสำหรับการสร้างชุดข้อมูล ดังตัวอย่างเส้นทางการบินและการจำลองสถานการณ์โจมตี GPS Spoofing แบบ Drift Attack ในภาพที่ 2

ชุดข้อมูลประกอบด้วยบันทึกการบินทั้งสิ้น 60 ไฟล์ โดยแต่ละไฟล์บันทึกข้อมูลเซนเซอร์ของอากาศยานในรูปแบบอนุกรมเวลา ในระดับจุดข้อมูล ป้ายกำกับถูกกำหนดจากเหตุการณ์การโจมตีที่ควบคุมได้ในระหว่างการจำลอง โดยช่วงก่อนเริ่มต้นโจมตีถูกกำหนดเป็น BENIGN ช่วงเริ่มต้นของการโจมตี ซึ่งมีระยะเวลาประมาณ 6–10 วินาที ขึ้นอยู่กับลักษณะของสถานการณ์จำลอง ถูกกำหนดเป็น SUSPECTED และช่วงที่ผลของการโจมตีปรากฏอย่างชัดเจนถูกกำหนดเป็น SPOOF ทั้งนี้ ข้อมูลทั้งหมดมี 233,483 จุด จากบันทึกการบิน 60 ไฟล์ ดังสรุปในตารางที่ 2 และในภาพที่ 3




**ภาพที่ 2:** ตัวอย่างเส้นทางการบินและการจำลองสถานการณ์โจมตี GPS Spoofing แบบ Drift Attack




**ภาพที่ 3**: จำนวนตัวอย่างของข้อมูล แยกตามคลาส BENIGN, SPOOF และ SUSPECTED

**ตารางที่ 2**: การกระจายข้อมูลต่อคลาส

| คลาส | Dev Set (จุด) | Held-out Test (จุด) | รวม |
|---|---|---|---|
| BENIGN | 160,310 | 17,957 | 178,267 |
| SPOOF | 35,025 | 6,540 | 41,565 |
| SUSPECTED | 10,330 | 3,321 | 13,651 |
| รวม | 205,665 | 27,818 | 233,483 |

**ตารางที่ 3**: ข้อมูลบันทึกสถานะการบิน 33 ฟีเจอร์

| กลุ่ม | คุณลักษณะ |
|---|---|
| Time/Kinematics | dt_s, vdx_mps, vdy_mps, vxy_mps, acc_xy_mps2 |
| GPS/Altitude | Alt, dAlt_mps, Spd, Acc_mps2, VZ |
| Heading | crs_sin, crs_cos, turnrate_dps |
| Attitude | ATT_Roll, ATT_Pitch, ATT_Yaw |
| IMU | IMU_GyrX, IMU_GyrY, IMU_GyrZ, IMU_AccX, IMU_AccY, IMU_AccZ |
| Barometer | BARO_Alt |
| XKF1 Velocity | XKF1_VN, XKF1_VE, XKF1_VD |
| Magnetometer | MAG_MagX, MAG_MagY, MAG_MagZ |
| RC Channels | C1, C2, C3, C4 |

### 3.3 การเตรียมข้อมูล (Data Preparation)

#### 3.3.1 การสกัดคุณลักษณะ

งานวิจัยนี้คัดเลือกฟีเจอร์ 33 รายการจากข้อมูลบันทึกการบิน ดังแสดงในตารางที่ 3 ครอบคลุมทั้งข้อมูลการเคลื่อนที่ เช่น ความเร็วและความสูง และค่าจากเซ็นเซอร์ต่าง ๆ เพื่อให้โมเดลตรวจจับทั้งความผิดปกติของเส้นทางบินและความขัดแย้งระหว่างเซ็นเซอร์ที่มักเกิดขึ้นเมื่อถูกโจมตี

คุณลักษณะในกลุ่ม kinematics และ altitude ใช้สะท้อนความเบี่ยงเบนของเส้นทางและความเร็ว ขณะที่ heading และ turn rate ช่วยอธิบายการเปลี่ยนทิศทางที่ผิดปกติ ส่วนข้อมูลจาก IMU (Inertial Measurement Unit), attitude, EKF (Extended Kalman Filter), barometer, magnetometer และ XKF1 มีบทบาทสำคัญในการสะท้อนความไม่สอดคล้องข้ามเซ็นเซอร์ ซึ่งเป็นลักษณะที่มักปรากฏเมื่อระบบนำทางได้รับผลกระทบจากการปลอมแปลงสัญญาณ GPS โดย XKF1 คือกลุ่มข้อมูล velocity estimate ที่ได้จาก EKF ภายใน ArduPilot

#### 3.3.2 การคำนวณคุณลักษณะอนุพันธ์จากพิกัด GPS

คุณลักษณะเชิงจลน์ในกลุ่ม Time/Kinematics ถูกคำนวณจากพิกัดละติจูด/ลองจิจูดดิบผ่านการแปลงพิกัดทรงกลมโดยประมาณ โดยกำหนดจุดอ้างอิง $\varphi_0$ เป็นค่าละติจูดของตัวอย่างแรก แล้วแปลงพิกัดเชิงพื้นที่ (x, y) หน่วยเมตร ดังนี้:

$$x = \Delta\lambda cos(\varphi^0) \cdot R, y = \Delta\varphi \cdot R \quad (1)$$

เมื่อ $R$ = 6,371,000 เมตร คือรัศมีโลก, $\Delta\lambda$ และ $\Delta\varphi$ คือผลต่างของลองจิจูดและละติจูดในหน่วยเรเดียนเมื่อเทียบกับจุดอ้างอิง, และ x กับ y คือพิกัดเชิงพื้นที่ในหน่วยเมตร จากนั้นคำนวณความเร็วเชิงระนาบ $v_{dx}$, $v_{dy}$, $v_{xy}$ และอัตราเร่ง $A_{xy}$ โดยใช้ผลต่างอันดับหนึ่งหารด้วยช่วงเวลา $\Delta t$:

$$v_{dx} = \frac{\Delta x}{\Delta t} \quad (2)$$

$$v_{dy} = \frac{\Delta y}{\Delta t} \quad (3)$$

$$v_{xy} = \sqrt{(v_{dx}^2 + v_{dy}^2)} \quad (4)$$

$$A_{xy} = \frac{\Delta v_{xy}}{\Delta t} \quad (5)$$

ทิศทางการบิน (course) ถูกแปลงเป็นคู่ $crs_sin$ และ $crs_cos$ เพื่อหลีกเลี่ยงความไม่ต่อเนื่องที่ 0°/360° ส่วน turnrate_dps คือ อัตราการเปลี่ยนทิศทางในหน่วยองศาต่อวินาที คำนวณจากผลต่างเชิงมุมที่ห่อกลับ (wrap-around) ในช่วง [−180°, +180°] หารด้วย $\Delta t$

การแปลงพิกัดดังกล่าวเป็นการประมาณเชิงระนาบสำหรับพื้นที่รอบจุดอ้างอิงที่มีระยะเคลื่อนที่จำกัด จึงเหมาะกับข้อมูลการบินในสภาพแวดล้อมจำลอง

#### 3.3.3 การแบ่งชุดข้อมูลระดับไฟล์

การแบ่งชุดข้อมูลดำเนินการในระดับไฟล์ (file-level) ไม่ใช่ระดับหน้าต่าง (window-level) เพื่อป้องกันการรั่วไหลของข้อมูล (data leakage) อันเกิดจากการที่หน้าต่างที่ทับซ้อนกัน (overlapping windows) จากไฟล์เดียวกันปรากฏทั้งในชุดฝึกและชุดทดสอบ จากบันทึกการบินทั้งหมด 60 ไฟล์ แยกชุดทดสอบขั้นสุดท้าย (held-out test set) จำนวน 10 ไฟล์ถูกแยกออกไว้ตั้งแต่ต้น ส่วนที่เหลืออีก 50 ไฟล์ถูกใช้เป็นชุดพัฒนา (development set)

สำหรับการฝึกโมเดลด้วยวิธี 5-fold cross-validation การแบ่งข้อมูลในระดับไฟล์ทำให้จำนวนไฟล์ต่อ fold มีจำกัด หากใช้ 10-fold แต่ละ fold จะมีชุดตรวจสอบเพียง 5 ไฟล์ ทำให้จำนวนไฟล์ในชุดตรวจสอบต่อ Fold อาจไม่เพียงพอสำหรับการประเมินทั้งสามคลาสอย่างมีเสถียรภาพ จึงเลือกใช้ 5-fold แทน โดยแต่ละ fold แบ่งออกเป็นชุดฝึก 40 ไฟล์ และชุดตรวจสอบ 10 ไฟล์ เมื่อสิ้นสุดกระบวนการ cross-validation จึงนำชุดพัฒนาทั้ง 50 ไฟล์มาฝึกโมเดลใหม่อีกครั้ง แล้วประเมินผลขั้นสุดท้ายบนชุดทดสอบ 10 ไฟล์ที่แยกสำรองไว้ ทั้งนี้จำนวนตัวอย่างข้อมูลในแต่ละไฟล์บันทึกการบินมีความแตกต่างกัน เนื่องจากระยะเวลาและรูปแบบของสถานการณ์จำลองในแต่ละครั้งไม่เท่ากัน

**3.3.4 การสร้างหน้าต่างเวลา**

เนื่องจากข้อมูลการบินมีลักษณะเป็นอนุกรมเวลา การสกัดบริบทเชิงเวลาจึงอาศัยการสร้างหน้าต่างเวลาแบบเลื่อนซ้อนทับกัน (overlapping sliding window) โดยกำหนดขนาดหน้าต่างเวลาเท่ากับ 7.0 วินาที และระยะเลื่อน (stride) เท่ากับ 1.0 วินาที การใช้หน้าต่างแบบซ้อนทับกันแทนหน้าต่างแบบไม่ซ้อนทับ (non-overlapping) ช่วยให้โมเดลได้เห็นการเปลี่ยนผ่านระหว่างช่วงเวลาในหลายมุมมอง และเพิ่มจำนวนตัวอย่างฝึกสอนจากข้อมูลที่มีอยู่จำกัด ซึ่งมีความสำคัญอย่างยิ่งสำหรับคลาส SUSPECTED ที่มีช่วงเวลาสั้นและจำนวนตัวอย่างน้อยกว่าคลาสอื่น ความถี่การสุ่มตัวอย่างถูกกำหนดโดย ArduPilot SITL อยู่ที่ 5 Hz ทำให้แต่ละหน้าต่างมีข้อมูลประมาณ 35 ตัวอย่าง การซ้อนทับดังกล่าวช่วยให้โมเดลติดตามการเปลี่ยนแปลงเชิงเวลาได้อย่างละเอียดและเหมาะกับการแยกช่วงเริ่มต้นของการโจมตี

**3.3.5 การกำหนดป้ายกำกับของหน้าต่าง**

การกำหนดป้ายกำกับระดับหน้าต่างทำโดยคำนวณสัดส่วนของจุดข้อมูลที่มีป้ายกำกับเป็น SPOOF และ SUSPECTED ภายในหน้าต่าง W ดังสมการ (6)–(8)

$$r_{spoof}(W) = \left(\frac{1}{|W|}\right)\sum_{i\in W} I(y_i = \{SPOOF\}) \quad (6)$$

$$r_{sus}(W) = \left(\frac{1}{|W|}\right)\sum_{i\in W} I(y_i = \{SUSPECTED\}) \quad (7)$$

โดยที่ $|W|$: จำนวนตัวอย่างข้อมูลทั้งหมดภายในหน้าต่างสังเกตการณ์

$I(\cdot)$:ฟังก์ชันบ่งชี้ ซึ่งมีค่าเป็น 1 เมื่อเงื่อนไขเป็นจริง

$y_i$: ป้ายกำกับจริง (ground-truth label) ของจุดข้อมูลที่ $i$

$$\begin{cases} \{SPOOF\}; if\ r_{\{spoof\}} \geq \tau_{\{spoof\}} \\ \{SUSPECTED\}; if\ r_{\{spoof\}} < \tau_{\{spoof\}}, r_{\{sus\}} \geq \tau_{\{sus\}} \\ \{BENIGN\};\ otherwise \end{cases} \quad (8)$$

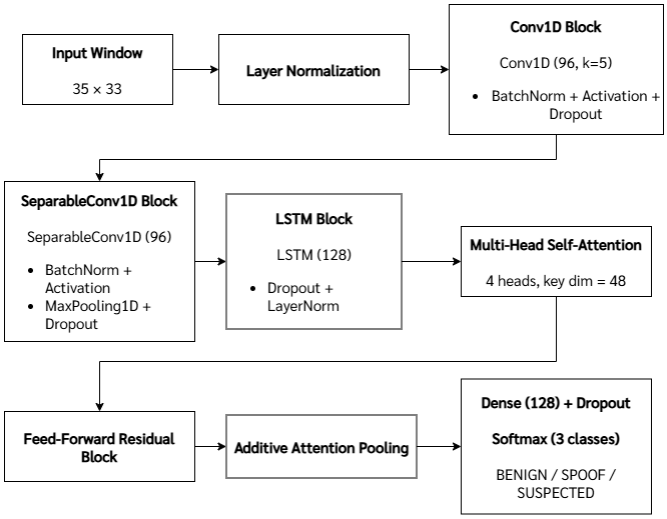
$\tau_{\{spoof\}}$, $\tau_{\{sus\}}$: ค่าเกณฑ์การตัดสินใจ (thresholds) ที่กำหนดไว้สำหรับแต่ละประเภทการตรวจจับ

ในการทดลองนี้กำหนด $\tau_{\{spoof\}}$= 0.20 หมายถึงต้องมีอย่างน้อยประมาณ 7 จุด หรือราว 1.4 วินาที ที่แสดงพฤติกรรม spoofing อย่างต่อเนื่องภายในหน้าต่างเวลาจึงจะจัดหน้าต่างนั้นว่าเป็น SPOOF เกณฑ์นี้ถูกเลือกใช้เพื่อให้มั่นใจว่าความผิดปกติที่ตรวจพบนั้นเกิดขึ้นจริง ไม่ใช่แค่สัญญาณรบกวนชั่วคราวที่อาจทำให้ระบบแจ้งเตือนผิดพลาด ในขณะเดียวกัน ค่า $\tau_{\{sus\}}$ = 0.10 หมายความว่าต้องมีจุดข้อมูลที่แสดงสัญญาณผิดปกติอย่างน้อยประมาณ 3–4 จุด (ราว 0.7 วินาที เมื่อหน้าต่างมีประมาณ 35 จุดข้อมูล) แต่ยังไม่ต่อเนื่องมากพอจะจัดเป็น SPOOF เกณฑ์นี้ทำให้สามารถแยกช่วง SUSPECTED ออกจาก BENIGN ได้ โดยยังคงรักษาความแตกต่างจากคลาส SPOOF อย่างชัดเจน

นอกจากนี้ ค่า threshold ทั้งสองถูกกำหนดเชิง heuristic และตรึงค่าไว้ตลอดกระบวนการทดลอง โดยไม่มีการปรับค่าเพิ่มเติมจากผลของ validation หรือ held-out test set ระหว่างการทดลอง ภายใต้เป้าหมายเพื่อแยกหน้าต่าง BENIGN, SPOOF และ SUSPECTED ให้มีความสมเหตุสมผลเชิงเวลาและเชิงพฤติกรรม ทั้งนี้ ค่าที่ต่ำกว่านี้จะเพิ่มความเสี่ยงต่อการแจ้งเตือนผิดพลาด ขณะที่ค่าสูงกว่านี้อาจทำให้พลาดสัญญาณในช่วงเริ่มต้นของการโจมตี

**3.4 การปรับมาตรฐานข้อมูลด้วย StandardScaler**

ฟีเจอร์จะถูกแปลงด้วย StandardScaler โดยลบค่าเฉลี่ยและหารด้วยส่วนเบี่ยงเบนมาตรฐานที่คำนวณจากชุดฝึกเท่านั้น ทำให้แต่ละฟีเจอร์มีค่าเฉลี่ยใกล้ 0 และส่วนเบี่ยงเบนมาตรฐานใกล้ 1 ดังสมการ (9)




**ภาพที่ 4:** แผนภาพลำดับการทำงานของโมเดล Attention-based CNN-LSTM

ทั้งนี้มีการยกเว้นฟีเจอร์ crs_sin และ crs_cos เนื่องจากมีค่าอยู่ในช่วง [−1, 1] ตามธรรมชาติของฟังก์ชันตรีโกณมิติ:

$$\hat{x_j} = \frac{x_j - \mu_j}{\sigma_j} \quad (9)$$

เมื่อ $\hat{x_j}$ คือค่าที่ผ่านการปรับมาตรฐานแล้ว, $\mu_j$และ $\sigma_j$ คือค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของคุณลักษณะที่ $j$

**3.5 สถาปัตยกรรมของโมเดลที่นำเสนอ**

โมเดลที่นำเสนอรับอินพุตเป็นหน้าต่างเวลาขนาด $T$ × 33 ภายใต้ความถี่การสุ่มตัวอย่างอยู่ที่ 5 Hz โดย $T$ มีค่า 35 ตัวอย่างต่อหน้าต่าง และประมวลผลผ่าน 4 ส่วนหลัก ได้แก่ CNN, LSTM, Multi-Head Self-Attention, Additive Attention Pooling พร้อม Classifier Head ดังแสดงในภาพที่ 4

- **ส่วนที่ 1 — CNN Block:** ขั้นตอนแรกใช้ Layer Normalization เพื่อปรับมาตรฐานค่าภายในแต่ละลำดับเวลา ทำให้การฝึกมีเสถียรภาพมากขึ้น จากนั้นใช้ Conv1D ใช้ kernel size = 5 ตามด้วย SeparableConv1D เพื่อสกัดรูปแบบเฉพาะที่ของอนุกรมเวลาด้วยจำนวนพารามิเตอร์ที่น้อยกว่าการคอนโวลูชันแบบมาตรฐาน (standard convolution) แต่ละชั้นมี Batch Normalization และ Swish activation พร้อม MaxPool1D ใช้pool size = 2 และ Dropout เอาต์พุตของส่วนนี้มีขนาด ($T$/2) × 96
- **ส่วนที่ 2 — LSTM Block:** ส่ง feature map จาก CNN เข้า LSTM เพื่อเรียนรู้บริบทเชิงเวลาและการพึ่งพาระยะยาวในลำดับการบิน ตามด้วย Dropout และ Layer Normalization เอาต์พุตมีขนาด ($T$/2) × 128 ซึ่งเป็นลำดับของ hidden state สำหรับทุกตำแหน่งเวลา และส่งต่อไปยัง attention
- **ส่วนที่ 3 — Multi-Head Self-Attention Block:** ลำดับ hidden state จาก LSTM ถูกป้อนเข้า Multi-Head Self-Attention เพื่อให้แต่ละตำแหน่งในลำดับเวลาสามารถประเมินและถ่วงน้ำหนักความสำคัญของตำแหน่งอื่นๆ ได้โดยตรง โดยที่ H ∈ ℝ^($T$/2 × 128) คือลำดับ hidden state จาก LSTM Block การคำนวณนิยามดังสมการ (10)–(11)

$$Q_i = HW_i^Q, K_i = HW_i^K, V_i = HW_i^V$$
$$head_i = softmax\left(\frac{(Q_i K_i^T)}{\sqrt{d_k}}\right)V_i \quad (10)$$
$$MultiHead(H) = Concat(head_1, \dots, head_h)W^O \quad (11)$$

โดยที่ $Q$, $K$ และ $V$ แทนเมทริกซ์ *Query*, *Key* และ *Value* ที่ได้จากการแปลง Hidden States ผ่านเมทริกซ์น้ำหนัก $HW_i^Q$, $HW_i^K$, $HW_i^V$ ของหัวที่ $i$, $d_k$ คือมิติของเวกเตอร์ *Key* ในแต่ละหัว, $h$ คือจำนวนหัวของ Attention, $W^O$ คือเมทริกซ์น้ำหนักสำหรับฉายผลลัพธ์กลับสู่มิติเดิม

ทั้งนี้ผลลัพธ์จาก attention ถูกรวมกลับด้วย residual connection และ Layer Normalization จากนั้นผ่าน position-wise feed-forward sublayer ซึ่งประกอบด้วย Dense(128×2 units, Swish) เพื่อขยายมิติเป็นสองเท่า ตามด้วย Dropout และ Dense(128 units) เพื่อฉายกลับสู่มิติเดิม พร้อม residual connection อีกชั้น ช่วยให้โมเดลสามารถคำนวณน้ำหนักความสำคัญของบริบทเชิงเวลา (Temporal Context) ที่เกี่ยวข้องก่อนเข้าสู่ขั้นตอน pooling

- **ส่วนที่ 4 — Additive Attention Pooling และ Classifier:** แทนที่จะใช้ Global Average Pooling ซึ่งให้น้ำหนักเท่ากันทุกตำแหน่ง โมเดลที่นำเสนอใช้ Additive Attention Pooling เพื่อเรียนรู้น้ำหนักเฉพาะต่อตำแหน่งแบบ soft selection ดังสมการ (12)–(14)

$$e_t = v^T tanh\left(W\hat{h}_t + b\right) \quad (12)$$

โดยที่ $e_t$ คือคะแนนความสำคัญของตำแหน่งเวลา t, $\hat{h}_t$ คือเวกเตอร์ข้อมูลอินพุต ณ ตำแหน่งหรือเวลาที่ $t$, $W$ และ $b$ คือพารามิเตอร์ค่าน้ำหนัก (Weights) และค่าอคติ (Bias) ที่แปลง $\hat{h}_t$ ให้อยู่ในรูปที่พร้อมจะนำไปวัดค่าความสำคัญ, $tanh$ คือฟังก์ชัน Hyperbolic Tangent ที่แปลงข้อมูลให้อยู่ในค่าระหว่าง [-1, 1], $v^T$ คือเวกเตอร์ค่าน้ำหนักที่ใช้แปลงผลลัพธ์ให้เป็นค่าสเกลาร์ (Scalar)

$$a_t = \frac{exp(e_t)}{\sum_j exp(e_j)} \quad (13)$$

โดยที่ $a_t$: ค่าน้ำหนักความสนใจ (Attention Weight) ซึ่งเป็นค่าความน่าจะเป็นที่ได้จากการทำ *Softmax* เพื่อระบุว่าต้องให้น้ำหนักที่ตำแหน่ง t มากน้อยเพียงใด

$$z = \sum_t a_t \hat{h}_t, p = softmax(W_c z + b_c) \quad (14)$$

โดยที่ $z$ (context vector) คือเวกเตอร์บริบทที่ได้จากการรวม hidden states แบบถ่วงน้ำหนักด้วยค่า attention และถูกส่งต่อเข้าสู่ชั้นการจำแนกประเภท (classification layer) เพื่อคำนวณผลลัพธ์สุดท้าย โดยเวกเตอร์ผลลัพธ์ $p$ คือเวกเตอร์ความน่าจะเป็นของทั้งสามคลาส ได้แก่ BENIGN, SPOOF และ SUSPECTED ซึ่งได้จากการแปลง $\boldsymbol{z}$ ผ่านชั้น Dense ตามด้วยฟังก์ชัน softmax, $W_c, b_c$: คือพารามิเตอร์ของชั้นการจำแนกประเภท (Classification Layer)

### 3.6 ฟังก์ชันความสูญเสียและการจัดการความไม่สมดุลของคลาส

เนื่องจากคลาส SUSPECTED มีจำนวนน้อยกว่าคลาส BENIGN อย่างมีนัยสำคัญ งานวิจัยนี้จึงใช้ Focal Loss ร่วมกับการถ่วงน้ำหนักคลาสภายนอก (class weighting) เพื่อจัดการความไม่สมดุลของข้อมูล โดย Focal Loss ทำหน้าที่ลดอิทธิพลของตัวอย่างที่โมเดลทำนายได้ง่ายและเพิ่มความสำคัญของตัวอย่างที่โมเดลยังทำนายได้ไม่ดี ขณะที่ class weight ทำหน้าที่เพิ่มน้ำหนักให้กับตัวอย่างที่อยู่ในคลาสส่วนน้อย

ในการศึกษานี้ Focal Loss ถูกคำนวณก่อนในระดับตัวอย่าง ดังนี้

$$FL_i = \left(1 - \hat{p}_{y_i}\right)^{\gamma}\left(-log\hat{p}_{y_i}\right) \quad (15)$$

โดยที่ $\hat{p}_{y_i}$ คือความน่าจะเป็นของคลาสจริงที่โมเดลทำนายได้สำหรับตัวอย่างที่ $i$ และ $\gamma$ เป็นพารามิเตอร์ที่ใช้ลดน้ำหนักของตัวอย่างที่ทำนายได้ง่าย เพื่อให้โมเดลให้ความสำคัญกับตัวอย่างที่ยากและคลาสที่มีจำนวนน้อยมากขึ้น

จากนั้นค่า Focal Loss ของแต่ละตัวอย่างจะถูกคูณด้วยค่าน้ำหนักของคลาสจริง

$$L_i = w_{y_i} FL_i \quad (16)$$

โดยที่ $w_{y_i}$ คือค่าน้ำหนักของคลาสจริงของตัวอย่างที่ $i$ ซึ่งคำนวณแบบผกผันกับความถี่ของคลาสในชุดฝึก ดังนี้

$$w_c = \frac{N}{KN_c} \quad (17)$$

เมื่อ $N$ คือจำนวนตัวอย่างทั้งหมดในชุดฝึก, $K$ คือจำนวนคลาสทั้งหมด และ $N_c$ คือจำนวนตัวอย่างในคลาส $c$

**ตารางที่ 4**: การตั้งค่าพารามิเตอร์ทั้งสองโมเดล

| พารามิเตอร์ | โมเดลที่นำเสนอ | PCA-CNN-LSTM |
|---|---|---|
| **Window size** | 7.0 s | 7.0 s |
| **Stride** | 1.0 s | 1.0 s |
| **Spoof threshold** | $\tau_spoof \geq 0.20$ | $\tau_spoof \geq 0.20$ |
| **Suspect threshold** | $\tau_sus \geq 0.10$ | $\tau_sus \geq 0.10$ |
| **CNN channels** | 96 | 96 |
| **LSTM units** | 128 | 128 |
| **PCA components** | — | 15 |
| **Attention heads** | 4 (key_dim=48) | — |
| **Attention pooling** | Additive (128 units) | — |
| **Dense units** | 128 | 128 |
| **Dropout** | 0.3 | 0.3 |
| **L2 regularization** | 0.0005 | 0.0005 |
| **Focal Loss (γ)** | 2.0 | 2.0 |
| **Optimizer** | Adam (lr=0.0005) | Adam (lr=0.0005) |
| **Batch size** | 64 | 64 |
| **Early stopping (patience)** | 5 (val_F1_macro) | 5 (val_F1_macro) |
| **Max epochs** | 90 | 90 |

---

**Algorithm 1: Window Construction and Window Label Assignment**

---

**Input:** flight log files F, window size = 7.0 s, stride = 1.0 s
**Output:** labeled windows $\{(X_w, y_w)\}$
**for** each file f **in** F **do**
read timestamps, telemetry, GPS, and sensor variables
build 33-dimensional point-level feature vector
clean invalid values: forward-fill → backward-fill → zero-fill
**if** time gap > 2.0 s **or** time reset detected: split segment
set $f_s = 5.0$Hz; win = round(7.0 × $f_s$); stride_n = round(1.0 × $f_s$)
slide overlapping window over each valid segment
**for** each window W **do**
compute $r_{spoof}(W)$ and $r_{sus}(W)$ using Eq. (6) and Eq. (7)
assign $y_w$ using Eq. (8)
store the labeled window ($X_w$, $y_w$)
**end for**
**end for**
**return** all labeled windows

---

ดังนั้น ตัวอย่างที่ได้รับน้ำหนักมากที่สุดในการปรับปรุงโมเดลคือตัวอย่างที่อยู่ในคลาสส่วนน้อยและยังถูกทำนายได้ไม่ดี ซึ่งช่วยให้โมเดลให้ความสำคัญกับคลาส SUSPECTED มากขึ้นโดยไม่ละเลยคลาสอื่น โดยที่น้ำหนักของแต่ละคลาสถูกคำนวณใหม่จากข้อมูลชุดฝึกของแต่ละ fold

---

**Algorithm 2: Model Training**

---

**Input:** development files D, held-out test files T
**Output:** cross-validation metrics and final test predictions
partition D into 5 near-equal folds at the file level
**for** each fold j = 1..5 **do**
define Dtrain(j) and Dval(j) from the 5-fold partition
fit StandardScaler on training-side point-level data only
transform training and validation data using the fitted scaler
build training windows and validation windows via Algorithm 1
train model with Adam (lr=5e-4), Focal Loss (γ=2), and external class weighting
use early stopping: patience=5, restore_best_weights=True, monitor=val_F1_macro
evaluate on validation fold; record metrics
**end for**
retrain the final model on the full development set D
evaluate on held-out test set T
report Accuracy, Precision, Recall, F1-score, Macro F1, Confusion Matrix

---

## 4. การประเมินผลและการอภิปรายผล

### 4.1 การตั้งค่าการทดลอง

งานวิจัยนี้เปรียบเทียบโมเดล Attention-based CNN-LSTM กับ PCA-CNN-LSTM ที่นำมาพัฒนาใหม่ภายใต้ชุดข้อมูลและเกณฑ์การทดลองเดียวกัน โดยไม่

เปรียบเทียบกับผลจากงานต้นฉบับโดยตรง เนื่องจากใช้ชุดข้อมูลต่างกัน โดยงานต้นฉบับใช้ข้อมูล GPS จริง ขณะที่งานวิจัยนี้ใช้ข้อมูลจำลอง SITL

ทั้งสองโมเดลใช้ชุดข้อมูล การสร้างหน้าต่างเวลา เกณฑ์การกำหนดป้ายกำกับ และการตั้งค่าการฝึกสอนชุดเดียวกันทุกประการ ดังแสดงในตารางที่ 4 ความแตกต่างหลักคือโมเดลที่นำเสนอใช้ attention mechanism ขณะที่โมเดล PCA-CNN-LSTM ใช้ PCA ลดมิติข้อมูลก่อนเข้า CNN-LSTM การทดลองทั้งหมดดำเนินการบน Acer Nitro AN515-55 (Intel Core i5-10300H 2.50 GHz, NVIDIA GeForce GTX 1650 Ti, RAM 32 GB, Windows 11 Home)

**4.2 การประเมินผลโมเดล**

การประเมินประสิทธิภาพแบ่งเป็นสองขั้นตอน ขั้นตอนแรกใช้ชุดพัฒนา 50 ไฟล์ผ่านกระบวนการ 5-fold cross-validation โดยในแต่ละ fold แบ่งเป็นชุดฝึก 40 ไฟล์ และชุดตรวจสอบ 10 ไฟล์ ขั้นตอนที่สองนำชุดพัฒนาทั้ง 50 ไฟล์มาฝึกโมเดลใหม่ แล้วทดสอบบนชุดทดสอบขั้นสุดท้าย (held-out test set) อีก 10 ไฟล์ที่ถูกแยกไว้ตั้งแต่เริ่มต้นและไม่เคยถูกใช้ในกระบวนการ cross-validation

การวัดประสิทธิผลใช้ Accuracy รวมของโมเดล และ Precision, Recall และ F1-score ในระดับรายคลาส โดยคำนวณในลักษณะ one-vs-rest ก่อนสรุปเป็นค่า Macro-average เพื่อให้ทุกคลาสได้รับน้ำหนักเท่าเทียมกัน โดยตัวชี้วัดแต่ละรายการนิยามจาก TP (True Positive), TN (True Negative), FP (False Positive) และ FN (False Negative) แล้วนำค่าที่ได้ไปใช้คำนวณ Precision, Recall และ F1-score ของแต่ละคลาส ก่อนสรุปเป็นค่า Macro-average ดังสมการ (18)–(22) รายงานทั้งค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานจาก 5-fold ร่วมกับผลบนชุดทดสอบขั้นสุดท้าย (held-out test) เนื่องจากเป้าหมายหลักคือการแยกช่วงเริ่มต้นของการโจมตี การวิเคราะห์จึงให้ความสำคัญกับประสิทธิผลของคลาส SUSPECTED เป็นพิเศษ

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (18)$$

$$Precision = \frac{TP}{TP + FP} \quad (19)$$

$$Recall = \frac{TP}{TP + FN} \quad (20)$$

$$F1 - score = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \quad (21)$$

$$Macro - Average = \frac{1}{N}\sum_{i=1}^{N} M_i \quad (22)$$

โดย $N$ คือจำนวนคลาสทั้งหมด และ $M_i$ คือค่าตัวชี้วัดของคลาสที่ $i$ การใช้ Macro-average ทำให้ทุกคลาสได้รับน้ำหนักเท่าเทียมกัน โดยไม่คำนึงถึงความไม่สมดุลของจำนวนข้อมูล

**4.3 ผลการทดลองของโมเดลที่นำเสนอ**

ตารางที่ 5 แสดงผลการประเมินของโมเดล Attention-based CNN-LSTM ภายใต้ 5-fold ซึ่งให้ค่าเฉลี่ย Accuracy เท่ากับ 0.9812, Macro Precision เท่ากับ 0.9571, Macro Recall เท่ากับ 0.9576 และ Macro F1 เท่ากับ 0.9570 ผลดังกล่าวสะท้อนว่าโมเดลที่นำเสนอมีความแปรปรวนระหว่างแต่ละ fold อยู่ในระดับไม่สูงนักตามค่าส่วนเบี่ยงเบนมาตรฐานที่รายงานในตารางที่ 5

**4.4 การเปรียบเทียบโมเดลที่นำเสนอกับ PCA-CNN-LSTM**

จากผลการทดลองแบบ 5-fold ที่แสดงในภาพที่ 5–8 พบว่าโมเดลที่นำเสนอให้ผลสูงกว่า PCA-CNN-LSTM โดยมีค่าเฉลี่ย Accuracy เท่ากับ 0.9812 เทียบกับ 0.9771 และค่าเฉลี่ย Macro F1 เท่ากับ 0.9570 เทียบกับ 0.9437 ตัวชี้วัดที่สะท้อนความแตกต่างได้เด่นชัดที่สุดคือค่าเฉลี่ย Macro Recall ซึ่งโมเดลที่นำเสนอได้ 0.9576 ขณะที่ PCA-CNN-LSTM ได้ 0.9364 แสดงว่าโมเดลที่นำเสนอสามารถตรวจจับตัวอย่างจริงของแต่ละคลาสได้ดีกว่า

จากภาพที่ 5–8 ซึ่งแสดงผลแบบ fold-by-fold พบว่าโมเดลที่นำเสนอมีแนวโน้มให้ประสิทธิภาพดีกว่าโมเดล PCA-CNN-LSTM และมีความแปรปรวนต่ำกว่าโดยเฉพาะอย่างยิ่งในด้าน Macro Recall และ Macro F1

**4.5 การเปรียบเทียบการประเมินผลบนชุดทดสอบขั้นสุดท้าย**

**ตารางที่ 5:** ผลการทดลอง Macro-average บน 5-fold ของโมเดล Attention-based CNN-LSTM

| Fold | Accuracy | Macro Precision | Macro Recall | Macro F1 |
|---|---|---|---|---|
| **1** | 0.9857 | 0.9683 | 0.9736 | 0.9708 |
| **2** | 0.9782 | 0.9573 | 0.9483 | 0.9527 |
| **3** | 0.9840 | 0.9642 | 0.9607 | 0.9623 |
| **4** | 0.9766 | 0.9533 | 0.9411 | 0.9471 |
| **5** | 0.9816 | 0.9421 | 0.9642 | 0.9523 |
| **เฉลี่ย** | **0.9812** | **0.9571** | **0.9576** | **0.9570** |
| **±SD** | **±0.0040** | **±0.0100** | **±0.0130** | **±0.0090** |

**ตารางที่ 6:** การเปรียบเทียบ Macro-average บนชุดทดสอบขั้นสุดท้าย

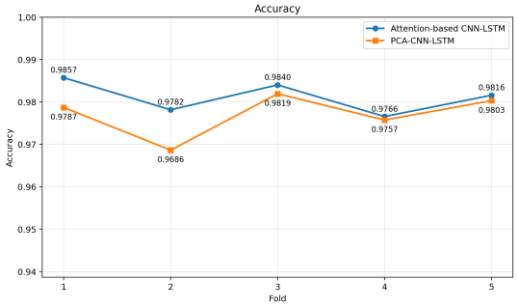
| โมเดล | Accuracy | Macro Precision | Macro Recall | Macro F1 |
|---|---|---|---|---|
| โมเดลที่นำเสนอ | **0.9885** | **0.9824** | **0.9834** | **0.9829** |
| PCA-CNN-LSTM | 0.9765 | 0.9719 | 0.9596 | 0.9650 |

**ตารางที่ 7:** ผลรายคลาสของโมเดล Attention-based CNN-LSTM บนชุดทดสอบขั้นสุดท้าย

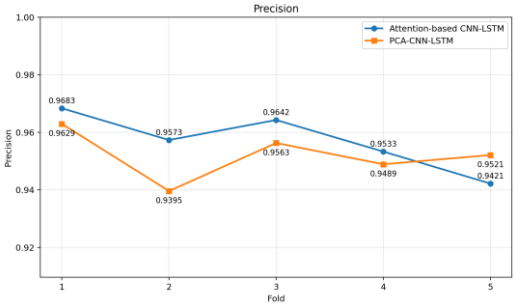
| Class | Precision | Recall | F1-score |
|---|---|---|---|
| BENIGN | **0.9971** | **0.9940** | **0.9955** |
| SPOOF | **0.9808** | 0.9869 | **0.9838** |
| SUSPECTED | 0.9692 | **0.9692** | **0.9692** |

**ตารางที่ 8**: ผลรายคลาสของโมเดล PCA-CNN-LSTM บนชุดทดสอบขั้นสุดท้าย

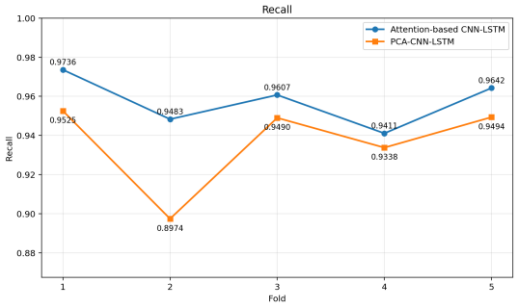
| Class | Precision | Recall | F1-score |
|---|---|---|---|
| BENIGN | 0.9961 | 0.9873 | 0.9917 |
| SPOOF | 0.9385 | **0.9906** | 0.9639 |
| SUSPECTED | **0.9810** | 0.9009 | 0.9393 |



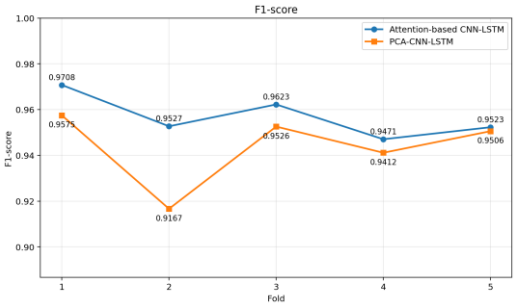

**ภาพที่ 5**: การเปรียบเทียบ Accuracy



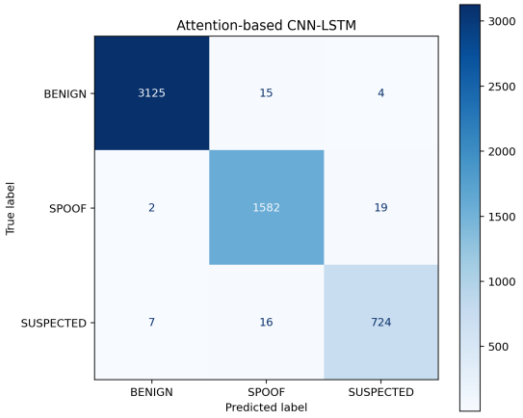

**ภาพที่ 6**: การเปรียบเทียบ Macro Precision



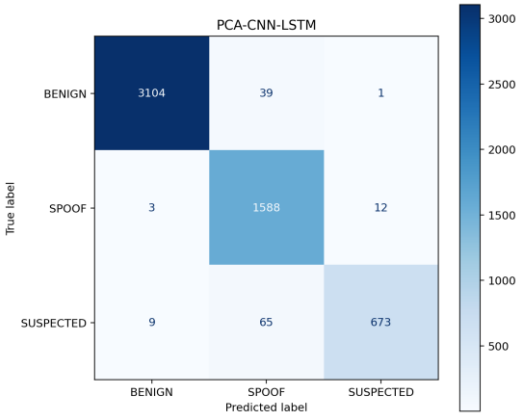

**ภาพที่ 7:** การเปรียบเทียบ Macro Recall




**ภาพที่ 8:** การเปรียบเทียบ Macro F1




**ภาพที่ 9:** Confusion Matrix ของโมเดล Attention-based CNN-LSTM




**ภาพที่ 10:** Confusion Matrix ของโมเดล PCA-CNN-LSTM

ผลการทดสอบบนชุดทดสอบขั้นสุดท้าย (held-out test) แสดงให้เห็นว่าโมเดลที่นำเสนอให้ผลสูงกว่า PCA-CNN-LSTM ดังแสดงในตารางที่ 6 โดยให้ค่า Accuracy เท่ากับ 0.9885, Macro Precision เท่ากับ 0.9824, Macro Recall เท่ากับ 0.9834 และ Macro F1 เท่ากับ 0.9829 ขณะที่ PCA-CNN-LSTM ให้ค่า Accuracy เท่ากับ 0.9765, Macro Precision เท่ากับ 0.9719, Macro Recall เท่ากับ 0.9596 และ Macro F1 เท่ากับ 0.9650

เมื่อพิจารณาผลการจำแนกรายคลาสจากตารางที่ 7 และตารางที่ 8 จะเห็นได้ว่า โมเดลที่นำเสนอมีความสามารถในการจำแนกคลาส SUSPECTED ได้ดีกว่า PCA-CNN-LSTM โดยให้ค่า recall ของคลาสนี้เท่ากับ 0.9692 ขณะที่ PCA-CNN-LSTM ทำได้เพียง 0.9009

**4.6 การวิเคราะห์ Confusion Matrix**

เมื่อพิจารณา Confusion Matrix ของชุดทดสอบขั้นสุดท้าย จากภาพที่ 9 และภาพที่ 10 พบว่า โมเดล PCA-CNN-LSTM ยังมีแนวโน้มสับสนระหว่างคลาส SUSPECTED กับ SPOOF มากกว่าอย่างชัดเจน โดยรูปแบบที่จำแนกผิดพลาดที่สำคัญที่สุดคือการที่โมเดลทำนายหน้าต่างข้อมูลที่เป็น SUSPECTED ให้เป็น SPOOF ซึ่ง PCA-CNN-LSTM ทำผิดพลาดนี้ 65 หน้าต่างจาก 747 หน้าต่างของคลาส SUSPECTED ในชุดทดสอบขั้นสุดท้ายคิดเป็น 8.7% เทียบกับโมเดลที่นำเสนอที่ผิดพลาดเพียง 16 หน้าต่าง คิดเป็น 2.1%

อย่างไรก็ตาม โมเดลที่นำเสนอไม่ได้มีประสิทธิภาพเหนือกว่า PCA-CNN-LSTM ในทุกตัวชี้วัดรายคลาส ดังแสดงในตารางที่ 7 และตารางที่ 8 ในชุดทดสอบขั้นสุดท้าย PCA-CNN-LSTM ให้ค่า recall ของคลาส SPOOF สูงกว่าเล็กน้อย (0.9906 เทียบกับ 0.9869) และให้ค่า precision ของคลาส SUSPECTED สูงกว่า (0.9810 เทียบกับ 0.9692) อย่างไรก็ตาม เมื่อพิจารณาภาพรวม โมเดลที่นำเสนอให้ผลสูงกว่าในด้าน Macro F1 และโดยเฉพาะอย่างยิ่งด้าน recall ของคลาส SUSPECTEDซึ่งเป็นประเด็นหลักของงานวิจัยนี้ ทั้งนี้

attention mechanism ช่วยให้โมเดลให้ความสำคัญกับช่วงเวลาที่สัญญาณเซนเซอร์เริ่มผิดปกติ ซึ่งเป็นจุดที่ PCA อาจสูญเสียข้อมูลไปในขั้นตอนการลดมิติ

## 5. สรุปและงานในอนาคต

งานวิจัยนี้นำเสนอแนวทางการตรวจจับ GPS spoofing สำหรับ UAV ในสภาพแวดล้อมจำลอง ArduPilot SITL โดยใช้โมเดล Attention-based CNN-LSTM บนข้อมูลบันทึกสถานะการบิน 33 ฟีเจอร์ ภายใต้การแบ่งข้อมูลระดับไฟล์ (file-level split) และการประเมินผลทั้งแบบ 5-fold และชุดทดสอบขั้นสุดท้าย (held-out test) ผลการทดลองแสดงให้เห็นว่า โมเดลที่นำเสนอให้ค่า Accuracy และ Macro F1 สูงกว่าโมเดลอ้างอิงแบบ PCA-CNN-LSTM ทั้งในการประเมินแบบ 5-fold cross-validation และบนชุดทดสอบขั้นสุดท้าย โดยบนชุดทดสอบขั้นสุดท้ายเพิ่ม Macro F1 จาก 0.9650 เป็น 0.9829 และเพิ่ม recall ของคลาส SUSPECTED จาก 0.9009 เป็น 0.9692 ผลดังกล่าวสะท้อนว่าโมเดลที่นำเสนอสามารถจำแนกช่วงเริ่มต้นของการโจมตีได้ดีกว่าโมเดลอ้างอิง ภายใต้ชุดข้อมูลและเงื่อนไขการทดลองของงานวิจัยนี้

อย่างไรก็ตาม งานวิจัยนี้ยังมีข้อจำกัดสำคัญ เนื่องจากการประเมินผลทั้งหมดดำเนินการบนข้อมูลจำลองจากสภาพแวดล้อม ArduPilot SITL จึงยังไม่สามารถยืนยันได้อย่างชัดเจนว่าโมเดลจะทำงานได้ดีเพียงใดกับข้อมูลการบินจริง ดังนั้นงานในอนาคตควรขยายการประเมินไปยังข้อมูลจากสภาพแวดล้อมจริงและเส้นทางการบินที่ซับซ้อนมากขึ้น รวมทั้งทดสอบโมเดลกับรูปแบบการโจมตีที่หลากหลายยิ่งขึ้น นอกจากนี้ ควรเพิ่มการเปรียบเทียบกับโมเดลในกลุ่ม Machine Learning แบบดั้งเดิม และปรับปรุงแบบจำลองให้มีขนาดเล็กลงและใช้ทรัพยากรคำนวณน้อยลง เพื่อให้เหมาะสมต่อการนำไปใช้งานบน UAV ขนาดเล็ก อีกทั้งควรมีการศึกษาเชิง ablation เพื่อประเมินผลขององค์ประกอบแต่ละส่วนของโมเดล เช่น CNN, LSTM และ attention mechanism ต่อประสิทธิภาพโดยรวม รวมทั้งวิเคราะห์ความสามารถในการอธิบายผลของโมเดล

## เอกสารอ้างอิง

[1] Y. Sun, M. Yu, L. Wang, T. Li, and M. Dong, “A Deep-Learning-Based GPS Signal Spoofing Detection Method for Small UAVs,” *Drones*, vol. 7, no. 6, p. 370, 2023, doi:10.3390/drones7060370.

[2] R. A. Agyapong, M. Nabil, A.-R. Nuhu, M. I. Rasul, and A. Homaifar, “Efficient Detection of GPS Spoofing Attacks on Unmanned Aerial Vehicles Using Deep Learning,” *in Proc. 2021 IEEE Symposium Series on Computational Intelligence (SSCI)*, pp. 1–8, 2021, doi:10.1109/SSCI50451.2021.9659972.

[3] A. A. Eshmawi, M. Umer, I. Ashraf, and Y. Park, “Enhanced Machine Learning Ensemble Approach for Securing Small Unmanned Aerial Vehicles From GPS Spoofing Attacks,” *IEEE Access*, vol. 12, pp. 27344–27355, 2024, doi:10.1109/ACCESS.2024.3359700.

[4] A. Al-Sabbagh, A. El-Bokhary, S. El-Koussa, A. Jaber, and M. Elkhodr, “Enhancing UAV Security Against GPS Spoofing Attacks Through a Genetic Algorithm-Driven Deep Learning Framework,” *Information*, vol. 16, no. 2, p. 115, 2025, doi:10.3390/info16020115.

[5] Y.-H. Sung, S.-J. Park, D.-Y. Kim, and S. Kim, “GPS Spoofing Detection Method for Small UAVs Using 1D Convolution Neural Network,” *Sensors*, vol. 22, no. 23, p. 9412, 2022, doi:10.3390/s22239412.

[6] X. Wei, C. Sun, M. Lyu, Q. Song, and Y. Li, “ConstDet: Control Semantics-Based Detection for GPS Spoofing Attacks on UAVs,” *Remote Sensing*, vol. 14, no. 21, p. 5587, 2022, doi:10.3390/rs14215587.

[7] L. Meng, L. Yang, S. Ren, G. Tang, L. Zhang, F. Yang, and W. Yang, “An Approach of Linear Regression-Based UAV GPS Spoofing Detection,” *Wireless Communications and Mobile Computing*, vol. 2021, p. 5517500, 2021,

doi:10.1155/2021/5517500.

[8] A. Shafique, A. Mehmood, and M. Elhadef, "Detecting Signal Spoofing Attack in UAVs Using Machine Learning Models," *IEEE Access*, vol. 9, pp. 93803–93815, 2021, doi:10.1109/ACCESS.2021.3089847.

[9] T. Talaei Khoei, S. Ismail, and N. Kaabouch, "Dynamic Selection Techniques for Detecting GPS Spoofing Attacks on UAVs," *Sensors*, vol. 22, no. 2, p. 662, 2022, doi:10.3390/s22020662.

[10] M. Nayfeh, Y. Li, K. Al Shamaileh, V. Devabhaktuni, and N. Kaabouch, "Machine Learning Modeling of GPS Features with Applications to UAV Location Spoofing Detection and Classification," *Computers & Security*, vol. 126, p. 103085, 2023, doi:10.1016/j.cose.2022.103085.

# การประยุกต์ใช้การประมวลผลภาษาธรรมชาติ (NLP) สำหรับการจัดกลุ่มและการแสดงภาพข่าวกรองภัยคุกคามทางไซเบอร์จาก X (Twitter)

# Applying NLP for Clustering and Visualizing Cyber Threat Intelligence from X (Twitter)

*นิพพิชฌน์ ยิ้มพรธนา (Nippich Yimporntana)* [1] *และสุนันทา สดสี (Sunantha Sodsee)* [2]

[1,2] *สาขาวิชาการบริหารเครือข่ายดิจิทัลและความมั่นคงปลอดภัยสารสนเทศ*

*คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล*

*มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ*

*s6707031857034@email.kmutnb.ac.th* [1], *sunantha.s@itd.kmutnb.ac.th* [2]

## บทคัดย่อ

*การจัดทำโครงงานวิจัยในครั้งนี้ มีวัตถุประสงค์เพื่อประยุกต์ใช้การประมวลผลภาษาธรรมชาติ (NLP) สำหรับการจัดกลุ่มและการแสดงภาพข่าวกรองภัยคุกคามทางไซเบอร์จาก X (Twitter) โดยเลือกใช้ชุดข้อมูล CybAttT ซึ่งเป็นชุดข้อมูลที่ประกอบด้วย Tweets ที่เกี่ยวข้องกับข่าวหรือเหตุการณ์โจมตีทางไซเบอร์จำนวน 36,080 แถว และรวมไปถึงชุดข้อมูลจริงที่ดำเนินการเก็บเองด้วยวิธี craping แหล่งที่มาที่น่าเชื่อถือจำนวน 7,560 แถว รวมทั้งหมด 43,640 แถว และได้มีการทำการเตรียมข้อมูล และวิศวกรรมคุณลักษณะ (Data Preprocessing & Feature Engineering) เหลือออกมาเป็นข้อมูลจำนวน 43,623 แถว เพื่อนำไปทำการฝึกแบบจำลองทั้งหมด 3 ประเภท คือ Bisecting K-Means, HDBSCAN และ Fuzzy C-Means พร้อมกับการทำ Text Representation 2 แบบ (BERT และ TF-IDF) และ การลดมิติของข้อมูลอีก 2 แบบ (PCA และ UMAP) และเมื่อได้ผลลัพธ์ออกมาให้มีการดำเนินการแสดงผลผ่าน Web-Based Dashboard สรุปผลการดำเนินงานได้ว่า แบบจำลองที่ทางผู้จัดทำเลือกให้เป็นแบบจำลองที่ดีและเหมาะสมที่สุดคือแบบจำลอง TF-IDF + UMAP + HDBSCAN มีค่า Silhouette Score เท่ากับ0.70 มีค่า Intra-Clustering Distance เท่ากับ 0.03 และ ค่า Inter-Clustering Distance เท่ากับ 9.54*

**คำสำคัญ:** ข่าวกรองภัยคุกคามไซเบอร์, การประมวลผลภาษาธรรมชาติ, การจัดกลุ่มข้อมูล, การแสดงผลข้อมูล

## Abstract

*The purpose of this independent study is to apply Natural Language Processing (NLP) techniques for clustering and visualizing cyber threat intelligence from the X (formerly Twitter) platform. The study utilizes the CybAttT dataset, which consists of 36,080 tweets related to cyberattack news and incidents, along with 7,560 real-world tweets collected through web scraping from reliable cybersecurity sources. In total, 43,640 tweets were collected. After performing data preprocessing and feature engineering, the dataset was refined to 43,623 records for model training. Three clustering algorithms were employed, namely Bisecting K-Means, HDBSCAN, and Fuzzy C-Means. Additionally, two text representation techniques, BERT and TF-IDF, were applied, along with two dimensionality reduction methods, PCA and UMAP. And once the results are obtained, they should be displayed via a web-based dashboard. The experimental results indicate that the TF-IDF + UMAP + HDBSCAN model achieved the best overall performance, with a*

*Silhouette Score of 0.70, an Intra-Clustering Distance of 0.03, and an Inter-Clustering Distance of 9.54. These results demonstrate the effectiveness of the proposed approach in identifying meaningful clusters and enhancing the analysis of cyber threat intelligence.*

**Keywords:** Cyber Threat Intelligence, Natural Language Processing, Clustering, Visualization

## 1. บทนำ

ในยุคที่การโจมตีทางไซเบอร์เกิดขึ้นทุกวินาที การรับรู้และวิเคราะห์ภัยคุกคามทางไซเบอร์ (Cyber Threat) อย่างทันท่วงทีนับเป็นสิ่งสำคัญอย่างยิ่งต่อการรักษาความมั่นคงปลอดภัยทางไซเบอร์ อย่างไรก็ตามข้อมูลที่เกี่ยวข้องมีปริมาณมหาศาล และไม่ใช่ทุกองค์กรที่สามารถจัดซื้ออุปกรณ์ที่จะสามารถหาข่าวที่เกี่ยวข้องได้ โดยอ้างอิงจากผลสำรวจจาก Bitsight ซึ่งได้ทำการเก็บผลสำรวจจากบริษัททางด้าน Cybersecurity กว่า 1000 แห่งที่มีจำนวนพนักงานมากกว่า 500 คน รายงานว่าการเฝ้าระวังอย่างต่อเนื่อง (Continuous Monitoring) คือสิ่งที่ผู้นำด้านความปลอดภัยจัดเป็นลำดับความสำคัญสูงที่สุด แต่มีเพียง 17% เท่านั้นที่ทำได้จริง [1] และจากรายงาน Cost of a Data Breach Report 2025 ที่จัดทำโดย IBM ในหน้าที่ 42 – 43 ระบุไว้เช่นกันว่า Threat Intelligence เป็นหนึ่งในเครื่องมือสำคัญที่ช่วยลดค่าใช้จ่ายจากเหตุข้อมูลรั่วไหล ซึ่งช่วยลดได้เฉลี่ยกว่า USD 211,906 ต่อเหตุการณ์ [2]

ด้วยเหตุนี้ การนำเทคโนโลยีการประมวลผลภาษาธรรมชาติ (Natural Language Processing: NLP) มาประยุกต์ใช้ในการจัดกลุ่ม (Clustering) และการแสดงภาพ (Visualization) ข้อมูลภัยคุกคามจากข้อความจาก X (Twitter) จึงเป็นอีกแนวทางหนึ่งที่จะสามารถช่วยให้นักวิเคราะห์เหตุการณ์สามารถมองเห็นแนวโน้มภัยคุกคามและตัดสินใจได้อย่างรวดเร็ว และมีประสิทธิภาพมากยิ่งขึ้น

วัตถุประสงค์ของงานวิจัยคือ

1. พัฒนาเครื่องมือสำหรับสืบค้นข้อมูลข่าวภัยคุกคามทางไซเบอร์ ผ่านการจัดกลุ่มข้อมูลภัยคุกคามที่รวบรวมจากทวีต พร้อมการแสดงผลข้อมูล (Visualization) ที่ช่วยให้ผู้ใช้สามารถเข้าใจแนวโน้ม และ บริบทของภัยคุกคามได้อย่างรวดเร็วและมีประสิทธิภาพ

2. เพื่อทราบผลลัพธ์ของประสิทธิภาพของกระบวนการจัดกลุ่มและการแสดงผลข้อมูลภัยคุกคามผ่านค่าความแม่นยำของกลุ่มคำ

## 2. เอกสาร งานวิจัย และทฤษฎีที่เกี่ยวข้อง

### 2.1 เอกสาร และงานวิจัยที่เกี่ยวข้อง (Literature Review)

งานวิจัยด้าน Cyber Threat Intelligence (CTI) จาก Social Media ส่วนใหญ่ใช้เทคนิค NLP และ Machine Learning ในการวิเคราะห์ข้อมูลภัยคุกคาม โดยมีพื้นฐานจากแนวคิด NLP [3] และ Unsupervised Learning [4] ซึ่งเป็นหลักการสำคัญที่งานวิจัยนี้นำมาประยุกต์ใช้

ในเชิงการใช้ข้อมูล งานวิจัยของ Arnolnt Spyros et. Al. [5], Favour Olaoye & Kaledio Potter [6], Kanti Singh Sangher et Al. [7], Sagar Bharat Shah [8] เสนอกระบวนการทำ Cyber Threat Intelligence แต่ไม่ได้มีการระบุว่าประยุกต์ใช้ข้อมูลจาก X (Twitter)

ในเชิงการตรวจจับและจำแนกประเภท งานวิจัยของ Hina Rashid et. Al. [9] ไม่ได้มีการประยุกต์ใช้ NLP, Mohamed El Amine Bekhouche & Kamel Adi [10], Omar Alsodi et. Al. [11], Prasasthy Balasubramanian et. Al. [12] และ Victor Obojo et. Al. [13] ส่วนใหญ่เลือกใช้โมเดล Classification ในการจัดกลุ่ม

ด้านชุดข้อมูลและแพลตฟอร์ม Lughbi et al. [14] สร้างชุดข้อมูล CybAttT จำนวน 36,071 Tweets พร้อม Dashboard สำหรับ Classification (F1-Score 87.6%) ซึ่งงานวิจัยนี้ได้นำชุดข้อมูลดังกล่าวมาใช้ เพื่อเพิ่มปริมาณข้อมูลสำหรับการฝึกชุดแบบจำลอง

จากการทบทวนงานวิจัยข้างต้นพบว่า งานส่วนใหญ่เน้น Classification และ Detection เป็นหลัก แต่ยังมีน้อยที่ลองใช้แนวทาง Clustering เพื่อวิเคราะห์รูปแบบภัยคุกคามในเชิงบริบท รวมถึง Visualization Dashboard สำหรับผู้

วิเคราะห์ ซึ่งเป็นช่องว่างที่งานวิจัยนี้มุ่งเติมเต็มช่องว่างดังกล่าว

**2.2 ทฤษฎีที่เกี่ยวข้อง (Related Theories)**

1) Natural Language Processing (NLP)

เป็นเทคนิคที่ใช้ในการประมวลผลและวิเคราะห์ข้อความ โดยในงานวิจัยนี้ใช้สำหรับการเตรียมข้อมูล การแปลงข้อความเป็นตัวเลข และการสกัดคุณลักษณะจากข้อความ

2) Text Representation

BERT เป็นแบบจำลองแบบ Transformer ที่สามารถเข้าใจบริบทของคำจากทั้งด้านหน้าและด้านหลัง

TF-IDF เป็นเทคนิคการให้ค่าน้ำหนักคำตามความสำคัญในเอกสาร

3) Dimensionality Reduction

PCA เป็นอัลกอริทึมที่ใช้ในการลดมิติข้อมูลเชิงเส้นเพื่อลดความซับซ้อน

UMAP: เป็นอัลกอริทึมที่ใช้ในการลดมิติแบบไม่เชิงเส้น เหมาะสำหรับข้อมูลข้อความที่มีโครงสร้างซับซ้อน

4) Clustering Algorithms

Bisecting K-Means เป็นแบบจำลองการจัดกลุ่มแบบลำดับชั้นที่พัฒนาจาก K-Means

HDBSCAN เป็นแบบจำลองการจัดกลุ่มแบบอิงความหนาแน่น สามารถจัดการ noise ได้ดี

Fuzzy C-Means เป็นแบบจำลองการจัดกลุ่มแบบ soft clustering ที่ข้อมูลสามารถอยู่ได้หลายกลุ่ม

5) Evaluation Metrics

Silhouette Score เป็นค่าวัดคุณภาพของการจัดกลุ่ม (ค่ายิ่งสูงยิ่งดี)

Intra-Clustering Distance เป็นค่าวัดความกระชับภายในกลุ่ม ค่ายิ่งต่ำแสดงว่าข้อมูลในกลุ่มมีความใกล้เคียงกันสูง และมีการรวมตัวกันอย่างมีประสิทธิภาพ

Inter-Clustering Distance เป็นค่าวัดความแตกต่างระหว่างกลุ่ม ค่ายิ่งสูงแสดงว่าแต่ละคลัสเตอร์แยกออกจากกันได้ชัดเจน และมีโครงสร้างกลุ่มที่เหมาะสม

## 3. รายละเอียดการดำเนินงาน

งานวิจัยนี้ดำเนินการตามลำดับขั้นตอนดังต่อไปนี้

**3.1 การเก็บรวบรวมข้อมูล (Data Collection)**

เก็บรวบรวมข้อมูลข่าวกรองภัยคุกคามทางไซเบอร์จาก 2 แหล่งหลัก ได้แก่ ข้อมูลทวีตจากแพลตฟอร์ม X (Twitter) ที่รวบรวมด้วยวิธี scraping จำนวน 7,560 Tweets จากทั้งหมด 6 บัญชี

1. Dark Web Intelligence
2. Dark Web Informer
3. FalconFeeds.io
4. H4ckmanac
5. Venarix
6. DarkWebSonar.io

และข้อมูลจากชุดข้อมูล CybAttT จำนวน 36,080 Tweets เพื่อนำมารวมเป็นชุดข้อมูลสำหรับการวิเคราะห์

**3.2 การเตรียมข้อมูลและการทำความสะอาดข้อมูล (Data Cleaning & Preprocessing)**

นำข้อมูลที่รวบรวมได้มาทำความสะอาดและเตรียมข้อมูล เช่น การรวมชุดข้อมูล (concat) การลบข้อมูลซ้ำ (drop_duplicates) การแปลงข้อความเป็นตัวพิมพ์เล็ก (text.lower) การลบสัญลักษณ์หรือข้อมูลรบกวน และการตัดคำรวมถึงลบ stopwords เพื่อให้ข้อมูลมีความเหมาะสมสำหรับการประมวลผล

**3.3 การสร้างคุณลักษณะ (Feature Engineering)**

แปลงข้อความให้อยู่ในรูปแบบที่สามารถนำไปวิเคราะห์ได้ โดยเปรียบเทียบเทคนิคการแทนค่าข้อความ 2 รูปแบบ ได้แก่ BERT และ TF-IDF

```python
# ----- Feature Extraction -----
def extract_tfidf(texts):
    tfidf = TfidfVectorizer(max_features=1000)
    return tfidf.fit_transform(texts).toarray()

def extract_bert(texts):
    model = SentenceTransformer("all-MiniLM-L6-v2")
    return model.encode(texts, show_progress_bar=True)
```

จากนั้นลดมิติข้อมูลด้วย PCA และ UMAP เพื่อช่วยลดความซับซ้อนของข้อมูลก่อนเข้าสู่ขั้นตอนการจัดกลุ่ม

```python
# ----- Dimensionality Reduction -----
def reduce_pca(X, n_components=50):
    return PCA(n_components=n_components).fit_transform(X)

def reduce_umap(X, n_neighbors=15, min_dist=0.1):
    reducer = umap.UMAP(n_neighbors=n_neighbors, min_dist=min_dist, metric="cosine")
    return reducer.fit_transform(X)
```

### 3.4 การสร้างแบบจำลองการจัดกลุ่มข้อมูล (Clustering Modeling)

นำข้อมูลที่ผ่านการแปลงและลดมิติแล้วไปจัดกลุ่มด้วยอัลกอริทึม 3 รูปแบบ ได้แก่ Bisecting K-Means, HDBSCAN และ Fuzzy C-Means และกำหนดค่า cluster_size=5, random_state=42 เพื่อเปรียบเทียบประสิทธิภาพของแต่ละแนวทาง ทำให้จะเกิดการดำเนินการ Cross Modeling ทั้งหมด 12 ชุดแบบจำลองเพื่อหาชุดแบบจำลองที่ดีที่สุด

```
# ----- Cross Evaluation -----
feature_methods = {
    "TF-IDF": extract_tfidf,
    "BERT": extract_bert,
}

reduce_methods = {
    "PCA": reduce_pca,
    "UMAP": reduce_umap,
}

cluster_methods = {
    "KMeans": cluster_bisecting_kmeans,
    "HDBSCAN": cluster_hdbscan,
    "Fuzzy C-Means": cluster_fuzzy_cmeans,
}

results = []

for feat_name, feat_func in feature_methods.items():
    print(f"🔍 Extracting features with: {feat_name}")
    X_feat = feat_func(texts)

    for red_name, red_func in reduce_methods.items():
        print(f"↘ Reducing with: {red_name}")
        X_red = red_func(X_feat)

        for clus_name, clus_func in cluster_methods.items():
            model_name = f"{feat_name} + {red_name} + {clus_name}"
            print(f"⏳ Clustering with: {model_name}")
            start = time.time()
            labels = clus_func(X_red)
            duration = time.time() - start
            sil, intra, inter = evaluate_clustering(X_red, labels)
            results.append({
                "Model": model_name,
                "Silhouette Score": sil,
                "Intra-cluster": intra,
                "Inter-cluster": inter,
                "Execution Time (sec)": duration
            })
```

### 3.5 การจัดทำหน้าเว็บไซต์แสดงผลแผนภาพ (Visualization)

นำผลลัพธ์จากแบบจำลองที่ดีที่สุดไปพัฒนาเป็นการแสดงผลเชิงภาพ (Visualization) เพื่อช่วยให้ผู้ใช้สามารถมองเห็นแนวโน้มและบริบทของภัยคุกคามได้ง่ายขึ้น โดยโครงสร้างการทำงานของระบบประกอบด้วย 4 ส่วนหลัก ได้แก่

1. Data Collector ทำหน้าที่รวบรวม Tweet จากแหล่งข้อมูล

2. Preprocessing Module ทำความสะอาดและเตรียมข้อมูล

3. Clustering Module วิเคราะห์และจัดประเภทภัยคุกคาม

4. Visualization Dashboard แสดงผลลัพธ์ผ่าน Web Application

## 4. ผลการการดำเนินงาน

### 4.1 การวัดผลแบบจำลอง (Model Evaluation)

| ประเภทโมเดล | Evaluation Value | | | ระยะเวลาในการฝึกแบบจำลอง (วินาที) |
|---|---|---|---|---|
| | Silhouette Score | Intra-Clustering | Inter-Clustering | |
| TF-IDF + UMAP + HDBSCAN | 0.7047439 | 0.037046 | 9.5394239 | 0.522322416 |
| TF-IDF + PCA + HDBSCAN | 0.6451835 | 0.0598969 | 0.7350232 | 26.81583691 |
| BERT + UMAP + HDBSCAN | 0.6005765 | 0.0391573 | 8.394578 | 0.519563913 |
| BERT + PCA + HDBSCAN | 0.4942867 | 0.1624132 | 0.8416144 | 37.00231266 |
| BERT + UMAP + Fuzzy C-Means | 0.347804 | 4.4067945 | 8.7608147 | 1.172604799 |
| BERT + UMAP + KMeans | 0.3380458 | 5.083333 | 10.306252 | 0.026004314 |
| TF-IDF + UMAP + Fuzzy C-Means | 0.3341235 | 5.8355036 | 12.683556 | 0.911996841 |
| TF-IDF + UMAP + KMeans | 0.2876431 | 6.7518873 | 13.871046 | 0.029025316 |
| TF-IDF + PCA + Fuzzy C-Means | 0.1639293 | 0.7203543 | 0.3192778 | 1.358226538 |
| BERT + PCA + Fuzzy C-Means | 0.1275072 | 0.8208115 | 0.4577311 | 1.894693851 |
| BERT + PCA + KMeans | 0.1176269 | 0.7272354 | 0.5715468 | 0.032516956 |
| TF-IDF + PCA + KMeans | 0.0711639 | 0.5922002 | 0.5478734 | 0.058053493 |

ทางผู้จัดทำโครงงานเลือกค่าสำหรับวัดประสิทธิภาพของแบบจำลองประกอบไปด้วยค่า Silhouette Score, Intra-clustering Distance และ Inter-Clustering Distance ซึ่งหลังจากทำนายผลเสร็จสิ้น สามารถสรุปได้ว่าแบบจำลองที่ดีที่สุด ทางผู้จัดทำจึงเลือกจาก 1. ค่า Silhouette Score ใกล้เคียง 1 มากที่สุด 2. ค่า Intra-Clustering ต่ำ 3. ค่า Inter-Clustering สูง จึงสรุปได้ว่าแบบจำลองที่ทางผู้จัดทำเลือกให้เป็นแบบจำลองที่ดีและเหมาะสมที่สุดคือแบบจำลอง TF-IDF + UMAP + HDBSCAN มีค่า Silhouette Score เท่ากับ 0.70 ซึ่งถือว่าอยู่ในเกณฑ์ที่ยอมรับได้ [15]

โดยเหตุผลสนับสนุนเริ่มต้นจาก Silhouette Score (0.70) ค่าดังกล่าวอยู่ในช่วง 0.51–0.70 ซึ่งตามเกณฑ์ของ Rousseeuw [15] จัดเป็นระดับ "Reasonable" หรือยอมรับได้ และเข้าใกล้ช่วง "Strong" (0.71–1.00) นอกจากนี้ยัง

เป็นค่า Silhouette Score ที่สูงที่สุดในบรรดาแบบจำลองทั้ง 12 ชุดที่ทดลอง สะท้อนว่าการจัดกลุ่มมีโครงสร้างชัดเจน ข้อมูลภายในกลุ่มมีความใกล้เคียงกันและแยกจากกลุ่มอื่นได้ดีที่สุด

Intra-Clustering Distance (0.03) ค่าดังกล่าวถือว่าต่ำที่สุดในบรรดาแบบจำลองทั้งหมดเมื่อเทียบกับแบบจำลองที่ใช้ Bisecting K-Means และ Fuzzy C-Means ที่มีค่าสูงถึง 4–6 สะท้อนว่าข้อมูล Tweet ที่ถูกจัดให้อยู่ในกลุ่มเดียวกันเกาะกลุ่มกันอย่างกระชับมากที่สุด จึงเป็นค่าที่ยอมรับได้

Inter-Clustering Distance (9.54) แม้ว่าค่าดังกล่าวจะไม่ใช่ค่าสูงสุดในตาราง โดยแบบจำลอง TF-IDF + UMAP + Bisecting K-Means มีค่าสูงถึง 13.87 และ TF-IDF + UMAP + Fuzzy C-Means มีค่า 12.68 แต่แบบจำลองทั้งสองดังกล่าวมีค่า Silhouette Score เพียง 0.29 และ 0.33 ตามลำดับ ซึ่งอยู่ในระดับ "Weak" ตามเกณฑ์ของ Rousseeuw และมีค่า Intra-Clustering Distance ที่สูงถึง 5.83 และ 6.75 แสดงว่าแม้กลุ่มจะอยู่ห่างกันก็จริง แต่ข้อมูลภายในกลุ่มกลับกระจายตัวมากจนไม่สามารถจัดกลุ่มได้อย่างมีความหมาย ในขณะที่แบบจำลองที่เลือกมีค่า Inter-Clustering Distance ที่สูงในระดับที่ยอมรับได้ (9.54) ควบคู่ไปกับค่า Intra-Clustering Distance ที่ต่ำมาก จึงเป็นค่าที่สะท้อนความสมดุลระหว่างความกระชับภายในกลุ่มและความห่างระหว่างกลุ่มได้ดีที่สุด

ระยะเวลาในการฝึกแบบจำลอง (0.52 วินาที) แม้ว่าแบบจำลอง Bisecting K-Means บางชุดจะมีระยะเวลาในการฝึกที่เร็วกว่า (เช่น BERT + PCA + Bisecting KMeans ใช้เวลาเพียง 0.032 วินาที) แต่แบบจำลองเหล่านั้นมีค่า Silhouette Score ต่ำมาก (0.07–0.34) ซึ่งไม่สามารถยอมรับได้ในเชิงคุณภาพการจัดกลุ่ม ในทางกลับกัน แบบจำลองที่มีคุณภาพใกล้เคียงกันอย่าง TF-IDF + PCA + HDBSCAN และ BERT + PCA + HDBSCAN กลับใช้ระยะเวลาในการฝึกสูงถึง 26.82 และ 37.00 วินาทีตามลำดับ ซึ่งช้ากว่าแบบจำลองที่เลือกประมาณ 50–70 เท่า ดังนั้นระยะเวลา 0.52 วินาทีจึงถือเป็นค่าที่สมดุลที่สุดระหว่างคุณภาพและประสิทธิภาพเชิงเวลากับงานวิจัยในครั้งนี้

จากการพิจารณาค่าประเมินผลทั้งสี่ตัวประกอบกัน จะเห็นได้ว่าแม้บางแบบจำลองจะมีค่าบางตัวที่ดีกว่า แต่เมื่อพิจารณาภาพรวมแล้ว แบบจำลอง TF-IDF + UMAP + HDBSCAN เป็นเพียงแบบจำลองเดียวที่ให้ผลลัพธ์ที่ดีในทุกมิติพร้อมกัน จึงถือเป็นแบบจำลองที่เหมาะสมที่สุดสำหรับงานวิจัยนี้

**4.2 ระบบแสดงผล (Visualization Platform)**

ผลลัพธ์จากแบบจำลองถูกนำไปพัฒนาเป็นระบบแสดงผล (Visualization) ในรูปแบบ Dashboard ซึ่งช่วยให้ผู้ใช้งานสามารถวิเคราะห์แนวโน้มของภัยคุกคามได้อย่างชัดเจน เช่น การแสดงกลุ่มของข้อมูลภัยคุกคาม การกระจายตัวของข้อมูล และแนวโน้มของเหตุการณ์ในช่วงเวลา

การแสดงผลในรูปแบบภาพช่วยลดความซับซ้อนของข้อมูล และเพิ่มความสามารถในการตีความข้อมูลสำหรับผู้วิเคราะห์ (Analyst) ทำให้สามารถระบุรูปแบบของภัยคุกคามและตัดสินใจได้อย่างรวดเร็วมากยิ่งขึ้น




## 5. สรุปผลการดำเนินงาน

**5.1 สรุปผลการดำเนินงาน**

การดำเนินงานโครงงานหัวข้อ “การประยุกต์ใช้การประมวลผลภาษาธรรมชาติ (NLP) สำหรับการจัดกลุ่มและการแสดงภาพข่าวกรองภัยคุกคามทางไซเบอร์จาก X (Twitter)” ได้ดำเนินงานโดยเลือกใช้ชุดข้อมูล CybAttT ซึ่งเป็นชุดข้อมูลที่ประกอบด้วย Tweets ที่เกี่ยวข้องกับข่าวหรือเหตุการณ์โจมตีทางไซเบอร์จำนวน 36,080 แถว และรวมไปถึงชุดข้อมูลจริงที่ดำเนินการเก็บเองด้วยวิธี Scraping แหล่งที่มาที่น่าเชื่อถือจำนวน 7,560 แถว รวมทั้งหมด 43,640 แถว และได้มีการทำการเตรียมข้อมูล และวิศวกรรมคุณลักษณะ (Data Preprocessing & Feature

Engineering) เหลือออกมาเป็นข้อมูลจำนวน 43,623 แถว เพื่อนำไปทำการฝึกแบบจำลองทั้งหมด 3 ประเภท คือ Bisecting K-Means, HDBSCAN และ Fuzzy C-Means พร้อมกับการทำ Text Representation 2 แบบ (BERT และ TF-IDF) และ การลดมิติของข้อมูลอีก 2 แบบ (PCA และ UMAP) และเมื่อได้ผลลัพธ์ออกมาให้มีการดำเนินการแสดงผลผ่าน Web-Based Dashboard

สรุปผลการดำเนินงานได้ว่า แบบจำลองที่ทางผู้จัดทำเลือกให้เป็นแบบจำลองที่ดีและเหมาะสมที่สุดคือ แบบจำลอง TF-IDF + UMAP + HDBSCAN มีค่า Silhouette Score เท่ากับ 0.70 มีค่า Intra-Clustering Distance เท่ากับ 0.03 และ ค่า Inter-Clustering Distance เท่ากับ 9.54

## 6. เอกสารอ้างอิง


[1] BITSIGHT. Bitsight Report Finds Lack of Visibility, AI, and Third-Party Risk Driving Cybersecurity Workforce Strain. Website, 2025 //// from https://www.bitsight.com/press-releases/bitsight-report-finds-lack-visibility-ai-and-third-party-risk-driving-cybersecurity?utm_medium=organic_social

[2] IBM. Cost of a Data Breach Report 2025 The AI Oversight Gap. PDF, 2025 //// from https://www.ibm.com/downloads/documents/us-en/131cf87b20b31c91

[3] Cole Stryker & Jim Holdsworth. What is NLP (natural language processing)?. IBM //// from https://www.ibm.com/think/topics/natural-language-processing

[4] Cole Stryker, Fangfang Lee, Dave Bergmann & Mark Scapicchio. What is unsupervised learning?. IBM //// from https://www.ibm.com/think/topics/natural-language-processing

[5] Arnolnt Spyros, Ilias Koritsas, Angelos Papoutsis, Panos Panagiotou, Despoina Chatzakou & Dimitrios Kavallieros. AI-Based Holistic Framework for Cyber Threat Intelligence Management. IEEE, 2025 //// from https://ieeexplore.ieee.org/abstract/document/ 10851288

[6] Favour Olaoye & Kaledio Potter. Natural Language Processing (NLP) for Social Media Threat Intelligence. Preprints, 2024 //// from https://www.preprints.org/frontend/manuscript/d1de84ac92fea8eb9074119acc554d49/download_pub

[7] Kanti Singh Sangher, Archana Singh & Hari Mohan Pandey. LSTM and BERT based transformers models for cyber threat intelligence for intent identification of social media platforms exploitation from darknet forums. Springer Nature, 2024 //// from https://link.springer.com/article/10.1007/s41870-024-02077-5

[8] Sagar Bharat Shah. Machine Learning for Cyber Threat Detection and Prevention in Critical Infrastructure. ResearchGate, 2025 //// from https://www.researchgate.net/profile/Sagar-Shah-40/publication/389579485_Machine_Learning_for_Cyber_Threat_Detection_and_Prevention_in_Critical_Infrastructure/links/67c862a5cc05504 3ce6de1ff/Machine-Learning-for-Cyber-Threat-Detection-and-Prevention-in-Critical-Infrastructure.pdf

[9] Hina Rashid, Hannan Bin Liaqat, Muhammad Usman Sana, Tayybah Kiren, Hanen Karamti & Imran Ashraf. Framework for detecting phishing crimes on Twitter using selective features and machine learning. ScienceDirect, 2025 //// from https://www.sciencedirect.com/science/article/abs/pii/S0045790625003064

[10] Mohamed El Amine Bekhouche & Kamel Adi. Advanced Real-Time Detection of Cyber Threat Information from Tweets. Spring Nature, 2025 //// from https://link.springer.com/chapter/10.1007/978-3-031-87499-4_2

[11] Omar Alsodi, Xujuan Zhou, Raj Gururajan, Anup Shrestha & Eyad Btoush. Cyber Threat Detection on Twitter Using Deep Learning Techniques: IDCNN and BiLSTM Integration. IEEE, 2025 //// from https://ieeexplore.ieee.org/abstract/document/10858790

[12] Prasasthy Balasubramanian, Sadaf Nazari, Danial Khosh Kholgh, Alireza Mahmoodi, Justin Seby & Panos Kostakos. A cognitive platform for collecting cyber threat intelligence and real-time detection using cloud computing. ScienceDirect, 2025 //// from https://www.sciencedirect.com/science/article/pii/S2772662225000013

[13] Victor Obojo, Haula Galadima & Richard Ikuesan. Social Media Threat Intelligence: A Framework for Collecting and Categorizing Threat-Related Data. Springer Nature, 2025 ////

from https://link.springer.com/chapter/10.1007/978-3-031-89360-5_6

[14] Huda Lughbi, Mourad Mars & Khaled Almotairi. CybAttT: A Dataset of Cyberattack News Tweets for Enhanced Threat Intelligence. MDPI, 2024 /// from https://www.mdpi.com/2306-5729/9/3/39

[15] Rousseeuw, P.J. A Graphical Aid to the Interpretation and Validation of Cluster Analysis. Comput. Appl. Math. 20, 53-65, 1987 /// from https://www.sciencedirect.com/science/article/pii/S2772662225000013

# การตรวจจับมัลแวร์ Android ที่ใช้ Dynamic Code Loading ด้วยการวิเคราะห์แบบไฮบริดที่กระตุ้นตามความเสี่ยง
# Risk-Triggered Hybrid Detection for DCL-based Android Malware

*ฉัตรพล แนวจำปา (Chatphol Nawjampa),กรวิทย์ สุทธิอาจ (Korawit Sutthi-at)*
*และจักรชัย โสอินทร์ (Chakchai So-In)*
*วิทยาลัยการคอมพิวเตอร์ มหาวิทยาลัยขอนแก่น*
*chatphol.n@kkumail.com, korawit.su@kkumail.com, chakso@kku.ac.th*

## บทคัดย่อ

*งานวิจัยนี้นำเสนอกรอบการตรวจจับมัลแวร์บนระบบปฏิบัติการ Android ที่ใช้เทคนิค Dynamic Code Loading (DCL) โดยใช้แนวทางการวิเคราะห์แบบไฮบริดที่มุ่งลดการใช้ทรัพยากร ผ่านการผสานการวิเคราะห์เชิงสถิต (Static Analysis) และเชิงไดนามิก (Dynamic Analysis) แบบจำลอง XGBoost ถูกนำมาใช้เป็นตัวจำแนกหลักเพื่อระบุพฤติกรรมที่เป็นอันตรายซึ่งซ่อนอยู่ในระหว่างการทำงาน (runtime) โดยแบบจำลองถูกฝึกและประเมินผลจากชุดข้อมูลแอปพลิเคชัน Android ทั้งแบบปกติและมัลแวร์ ซึ่งมีการสกัดคุณลักษณะจาก permissions และ API calls ผลการทดลองแสดงให้เห็นว่า Static Analysis ให้ค่า Accuracy = 0.919, ROC-AUC = 0.976 และ Recall = 0.946 ขณะที่ Dynamic Analysis ให้ค่า Accuracy = 0.906, ROC-AUC = 0.956 และ Recall = 0.838 เมื่อรวมผลลัพธ์จากทั้งสองขั้นตอนด้วยวิธี Combined Risk Score แนวทางที่เสนอสามารถใช้ผลจาก static stage เพื่อคัดกรองแอปพลิเคชันก่อนเข้าสู่ dynamic stage จึงช่วยลดจำนวนที่ต้องทำการวิเคราะห์เชิงไดนามิกเต็มรูปแบบได้ ภายใต้เงื่อนไขการทดลอง ดังนั้น แนวทางนี้จึงเหมาะสมสำหรับการรับมือกับมัลแวร์ Android ที่มีความซับซ้อนในปัจจุบัน*

**คำสำคัญ:** Android Malware, Dynamic Code Loading (DCL), การวิเคราะห์มัลแวร์แบบไฮบริด, XGBoost.

## Abstract

*This study proposes a framework for detecting Android malware that utilizes Dynamic Code Loading (DCL) through a resource-efficient hybrid analysis approach. The proposed method integrates Static Analysis and Dynamic Analysis, employing XGBoost as the primary classifier to identify malicious behaviors hidden during application runtime. The model is trained and evaluated on datasets consisting of both benign and malicious Android applications, with features extracted from permissions and API calls. Experimental results show that Static Analysis achieves an Accuracy of 0.919, ROC-AUC of 0.976, and Recall of 0.946, while Dynamic Analysis achieves an Accuracy of 0.906, ROC-AUC of 0.956, and Recall of 0.838. By combining both stages using a Combined Risk Score, the proposed approach uses the static stage to screen applications before invoking dynamic analysis, thereby reducing the need for full dynamic analysis under the reported experimental setting. Therefore, this framework provides an effective solution for addressing complex Android malware threats in modern environments.*

**Keywords:** Android Malware, Dynamic Code Loading (DCL), Hybrid Malware Analysis, XGBoost

## 1. บทนำ

ปัจจุบันระบบปฏิบัติการ Android เป็นแพลตฟอร์มอุปกรณ์เคลื่อนที่ที่ได้รับความนิยมสูง ทำให้กลายเป็น

เป้าหมายสำคัญของผู้พัฒนามัลแวร์บนอุปกรณ์เคลื่อนที่ มัลแวร์จำนวนมากใช้เทคนิคต่างๆ เพื่อหลบเลี่ยงการตรวจจับของระบบรักษาความปลอดภัย หนึ่งในเทคนิคที่ถูกนำมาใช้อย่างแพร่หลายคือ Dynamic Code Loading (DCL) ซึ่งเป็นกลไกที่อนุญาตให้แอปพลิเคชันโหลดโค้ดเพิ่มเติมในระหว่างการทำงาน ทำให้โค้ดอันตรายสามารถถูกซ่อนไว้และถูกเรียกใช้งานได้ในภายหลัง ส่งผลทำให้การตรวจจับด้วยวิธีการวิเคราะห์เชิงสถิตเพียงอย่างเดียวมีข้อจำกัด [1]

งานวิจัยก่อนหน้า [1] ได้ชี้ให้เห็นว่าการใช้เทคนิค DCL ทำให้มัลแวร์สามารถซ่อนพฤติกรรมอันตรายและหลบเลี่ยงกระบวนการตรวจจับแบบดั้งเดิมได้อย่างมีประสิทธิภาพ ดังนั้น การผสานการวิเคราะห์เชิงสถิตและเชิงไดนามิกจึงเป็นแนวทางที่ช่วยเพิ่มความสามารถในการตรวจจับมัลแวร์ได้ดียิ่งขึ้น

ด้วยเหตุนี้ งานวิจัยนี้จึงนำเสนอแนวทางการวิเคราะห์แบบไฮบริดที่ใช้ทรัพยากรต่ำสำหรับการตรวจจับมัลแวร์ Android ที่ใช้ Dynamic Code Loading โดยบูรณาการการวิเคราะห์เชิงสถิตและการวิเคราะห์เชิงไดนามิกร่วมกับแบบจำลอง XGBoost พร้อมกลไกการกระตุ้นตามความเสี่ยง (risk-triggered mechanism) ซึ่งใช้ผลจากขั้นตอนการวิเคราะห์เชิงสถิตเป็นตัวคัดกรองก่อนเข้าสู่การวิเคราะห์เชิงไดนามิกเฉพาะกรณีที่มีความเสี่ยงสูง แนวทางดังกล่าวช่วยเพิ่มความแม่นยำในการตรวจจับพฤติกรรม Dynamic Code Loading (DCL) ขณะเดียวกันสามารถลดภาระด้านเวลาและทรัพยากรของการวิเคราะห์เชิงไดนามิกได้อย่างมีนัยสำคัญ

## 2. งานวิจัยที่เกี่ยวข้อง

งานวิจัยเกี่ยวกับการตรวจจับมัลแวร์ Android ได้มีการศึกษาอย่างต่อเนื่อง โดยเฉพาะมัลแวร์ที่ใช้เทคนิค Dynamic Code Loading (DCL) ซึ่งช่วยให้สามารถซ่อนโค้ดอันตรายและหลบเลี่ยงการตรวจจับได้ งานของ [1] แสดงให้เห็นว่าการโหลดโค้ดเชิงไดนามิกสามารถถูกใช้เพื่อแทรกพฤติกรรมอันตรายในระหว่างการทำงานของแอปพลิเคชันได้อย่างมีประสิทธิภาพ ขณะเดียวกัน งานของ [2] แสดงศักยภาพของการใช้ Machine Learning โดยเฉพาะ Random Forest ในการจำแนกแอปพลิเคชัน Android ออกเป็นกลุ่มปกติและมัลแวร์

ในด้านการตรวจจับมัลแวร์แบบสถิต งานของ [3] ได้นำเสนอวิธีการตรวจจับมัลแวร์ Android โดยใช้คุณลักษณะจาก permissions และ API calls ซึ่งสามารถช่วยเพิ่มความแม่นยำของโมเดลการจำแนกมัลแวร์ได้ [3] ขณะเดียวกัน งานวิจัยของ [4] ได้พัฒนากรอบการวิเคราะห์แอปพลิเคชัน Android เพื่อศึกษาพฤติกรรมของโค้ดที่ถูกโหลดเชิงไดนามิก ทำให้สามารถตรวจสอบโค้ดที่ถูกเรียกใช้ระหว่างการทำงานของแอปพลิเคชันได้อย่างมีประสิทธิภาพ [4] อีกทั้งระบบของ [5] ใช้การวิเคราะห์เชิงสถิตเพื่อสกัดคุณลักษณะจากแอปพลิเคชัน Android เช่น permissions และ API calls จากนั้นใช้เทคนิค Machine Learning ในการจำแนกมัลแวร์ ซึ่งช่วยให้สามารถตรวจจับมัลแวร์ Android ได้อย่างมีประสิทธิภาพ

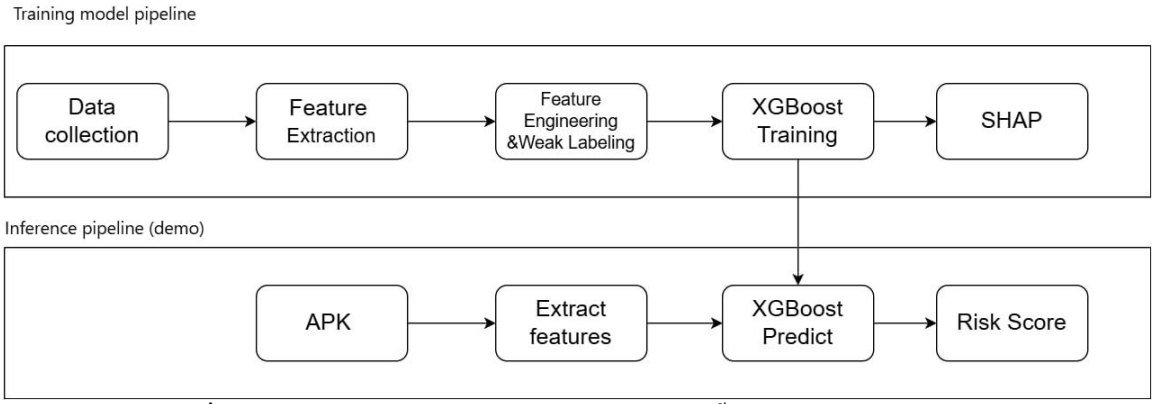
ในส่วนของการวิเคราะห์เชิงไดนามิก มีการพัฒนาเครื่องมือหลายชนิดเพื่อศึกษาพฤติกรรมของแอปพลิเคชัน Android เช่น [6] ซึ่งใช้การกระตุ้นเหตุการณ์อัตโนมัติเพื่อสำรวจพฤติกรรมของแอปพลิเคชันในระหว่างการทำงาน [6] รวมถึงระบบของ [7] ที่ใช้การติดตามการไหลของข้อมูลแบบเรียลไทม์เพื่อตรวจสอบการรั่วไหลของข้อมูลส่วนบุคคลในสมาร์ตโฟน และจาก [8] ยังสามารถสร้างมุมมองเชิงความหมายของระบบปฏิบัติการและ Dalvik VM เพื่อใช้ในการวิเคราะห์มัลแวร์เชิงไดนามิกได้อย่างละเอียด [8] อย่างไรก็ตาม การวิเคราะห์เชิงไดนามิกเต็มรูปแบบมักมีต้นทุนด้านทรัพยากรสูง ดังนั้นการผสานการวิเคราะห์เชิงสถิตและเชิงไดนามิกจึงเป็นแนวทางที่ช่วยเพิ่มประสิทธิภาพในการตรวจจับมัลแวร์ Android ได้ดียิ่งขึ้น

งานวิจัยที่ผ่านมาแสดงให้เห็นว่า static analysis มีข้อจำกัดในการตรวจจับ DCL ขณะที่ dynamic analysis มีต้นทุนด้านทรัพยากรสูง แม้จะมีแนวทางแบบไฮบริด แต่ยังขาดกลไกในการจัดการทรัพยากรอย่างมีประสิทธิภาพ ดังนั้นงานวิจัยนี้จึงเสนอแนวทาง lightweight hybrid analysis ที่ผสาน static, dynamic

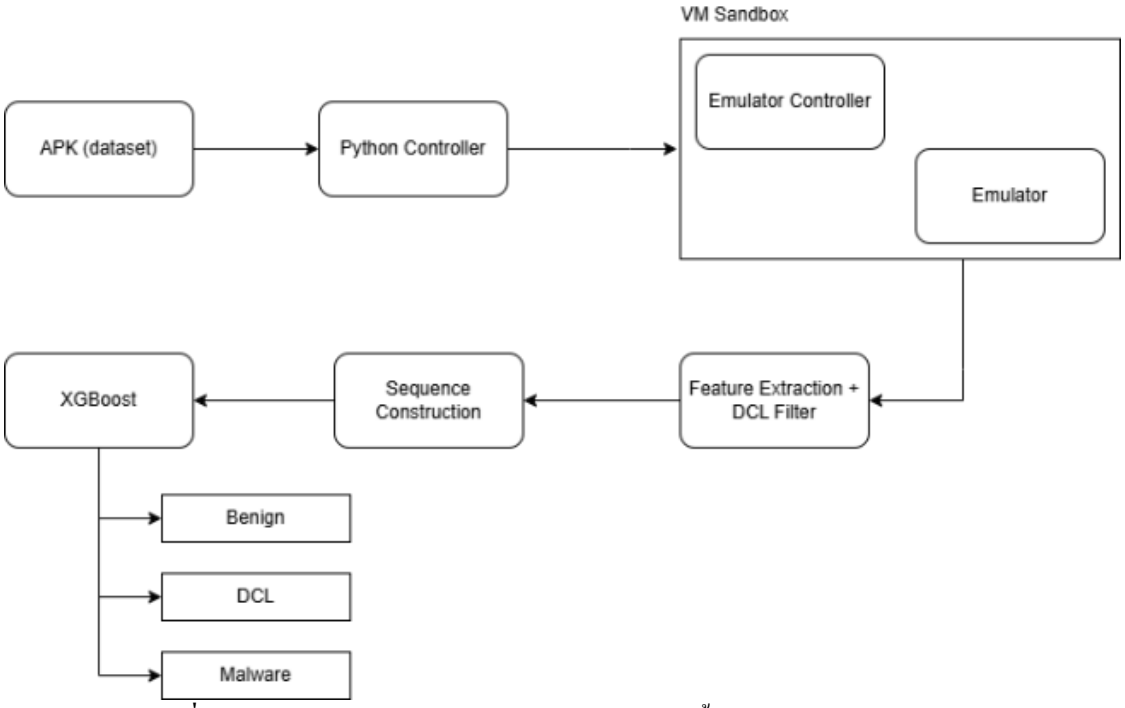
**ตารางที่ 1:** การเปรียบเทียบคุณสมบัติของงานวิจัยที่เกี่ยวข้อง

| | Static Analysis | Dynamic Analysis | DCL Analysis | Machine Learning | Risk Targeted |
|---|---|---|---|---|---|
| [1] | ✓ | | ✓ | | |
| [2] | ✓ | | | ✓ | ✓ |
| [3] | ✓ | | | ✓ | |
| [4] | ✓ | ✓ | ✓ | | |
| [5] | ✓ | | | ✓ | ✓ |
| [6] | | ✓ | | | |
| [7] | | ✓ | | | |
| [8] | | ✓ | | | |
| **เสนอ** | ✓ | ✓ | ✓ | ✓ | ✓ |

**หมายเหตุ:** ✓ หมายถึงมีคุณสมบัติดังกล่าว




**รูปภาพที่ 1:** ภาพรวมกระบวนการพัฒนาแบบจำลองสำหรับขั้นตอนการวิเคราะห์เชิงสถิต




**รูปภาพที่ 2:** ภาพรวมกระบวนการพัฒนาแบบจำลองสำหรับขั้นตอนการวิเคราะห์เชิงไดนามิก

นอกจากนี้ยังผสานการจำแนกด้วย XGBoost ใช้กลไก risk-based ประเมินความเสี่ยงจาก static analysis และทำ dynamic analysis เฉพาะแอปพลิเคชันที่มีความเสี่ยงสูง เพื่อเพิ่มประสิทธิภาพการตรวจจับ DCL โดยตารางที่ 1 แสดงการเปรียบเทียบเทคนิคต่าง ๆ

## 3. วิธีการดำเนินงาน

### 3.1 ภาพรวมการดำเนินงานวิจัย

งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนากรอบการตรวจจับ Android malware ที่มีความเกี่ยวข้องกับพฤติกรรม Dynamic Code Loading (DCL) โดยประยุกต์ใช้การวิเคราะห์ทั้งข้อมูลเชิงโครงสร้างจากไฟล์แอปพลิเคชันและข้อมูลเชิงพฤติกรรมระหว่างการทำงาน เพื่อให้การประเมินความเสี่ยงของแอปพลิเคชันมีความครอบคลุมมากยิ่งขึ้น ในภาพรวม กระบวนการวิจัยประกอบด้วยการรวบรวมและคัดเลือกข้อมูล การสกัดและวิศวกรรมคุณลักษณะ การพัฒนาแบบจำลองสำหรับการจำแนก และการสรุปผลในรูปของระดับความเสี่ยง

รูปภาพที่ 1 แสดงถึงกระบวนการพัฒนาแบบจำลองในส่วน static stage เริ่มจากการรวบรวมข้อมูล การสกัดคุณลักษณะ และการออกแบบคุณลักษณะเพิ่มเติม ก่อนนำข้อมูลเข้าสู่การฝึกแบบจำลอง XGBoost และวิเคราะห์ความสำคัญของคุณลักษณะด้วย SHAP ในส่วนของการนำไปใช้งาน ระบบจะรับไฟล์ APK เป้าหมายเข้าสู่กระบวนการสกัดคุณลักษณะ แล้วส่งเข้าสู่แบบจำลองที่ผ่านการฝึกเพื่อสรุปผลเป็นค่าความเสี่ยง

นอกจากนี้รูปภาพที่ 2 ยังแสดงถึงสถาปัตยกรรมของระบบตรวจจับมัลแวร์ Android ที่เสนอ หากผลการประเมินจากขั้นตอนการวิเคราะห์เชิงสถิตพบว่ามีความเสี่ยง ระบบจะนำไฟล์ APK เป้าหมายเข้าสู่กระบวนการวิเคราะห์ผ่าน Python Controller ซึ่งทำหน้าที่ควบคุมการทำงานของ Android Emulator เพื่อรันแอปพลิเคชันและเก็บพฤติกรรมระหว่างการทำงาน

จากนั้นนำข้อมูลที่ได้เข้าสู่กระบวนการสกัดคุณลักษณะ (Feature Extraction) และตัวกรอง DCL (DCL Filter) เพื่อคัดกรองพฤติกรรมที่เกี่ยวข้องกับ Dynamic Code Loading หลังจากนั้นจึงนำคุณลักษณะที่ได้ไปสร้างลำดับข้อมูล (Sequence Construction) และใช้แบบจำลอง XGBoost สำหรับการจำแนกประเภทแอปพลิเคชันออกสองกลุ่ม ได้แก่ Non-DCL และ DCL Malware

โดยมีขั้นตอนการวิเคราะห์เชิงสถิต (Static Analysis) เป็นการสกัดคุณลักษณะจากไฟล์ APK โดยไม่ต้องรันแอปพลิเคชัน เพื่อประเมินความเสี่ยงเบื้องต้น เช่น permissions การใช้ DexClassLoader และรูปแบบสตริง/การเข้ารหัส ข้อมูลที่ได้ถูกแปลงเป็นเวกเตอร์ $X_{static}$ และนำไปคำนวณคะแนนความเสี่ยง $S_{static}$ ด้วย XGBoost เพื่อใช้เป็นเกณฑ์สำหรับการวิเคราะห์ในขั้นตอนถัดไป ตามสมการด้านล่าง

$$S_{static} = f_{static}(X_{static}) \times 100 \quad (1)$$

โดยที่ $S_{static}$ แทนค่าความเสี่ยงที่ได้จากขั้นตอนการวิเคราะห์เชิงสถิต $X_{static}$ คือเวกเตอร์คุณลักษณะที่สกัดจากไฟล์ APK และ $f_{static}( )$ แทนแบบจำลอง XGBoost ที่ผ่านการฝึกแล้ว ซึ่งใช้สำหรับประเมินความน่าจะเป็นที่แอปพลิเคชันจะมีพฤติกรรมเป็นมัลแวร์ โดยนำค่าความน่าจะเป็นดังกล่าวมาปรับสเกล (Rescaling) ด้วยการคูณ 100 เพื่อให้ $S_{static} \in [0,100]$ ซึ่งทำหน้าที่เป็นคะแนนความเสี่ยงร้อยละสำหรับใช้เป็นเกณฑ์ตัดสินใจในขั้นตอนถัดไป

### 3.2 การรวบรวมและคัดเลือกข้อมูล

ข้อมูลที่ใช้ในการวิจัยนี้ประกอบด้วยข้อมูลสำหรับ static และ dynamic stage เพื่อสนับสนุนการวิเคราะห์แบบไฮบริด

- ข้อมูลสำหรับ static ได้มาจากฐานข้อมูล AndroZoo โดยใช้ไฟล์เมทาดาทา latest.csv และค่า VirusTotal detection เป็นเกณฑ์ในการคัดเลือกตัวอย่างเบื้องต้น โดยกำหนดให้ตัวอย่างที่มีค่า *vt_detection* = 0 เป็นกลุ่ม benign และตัวอย่างที่มีค่า *vt_detection* >= 10 เป็นกลุ่ม malware เพื่อให้ได้ข้อมูลที่มีความแตกต่างชัดเจนสำหรับงานจำแนกประเภท

  หลังจากคัดเลือกเบื้องต้นแล้ว ผู้วิจัยได้ดำเนินการดาวน์โหลดไฟล์ APK ผ่าน AndroZoo

API และนำมาตรวจสอบความสมบูรณ์ของข้อมูลก่อนนำไปใช้ในการทดลอง โดยกำหนดจำนวนข้อมูลรวม 294 ตัวอย่าง แบ่งเป็น benign 147 ตัวอย่าง และ malware 147 ตัวอย่าง เพื่อให้ข้อมูลมีความสมดุลเหมาะสมต่อการพัฒนาแบบจำลอง

- ข้อมูลสำหรับ dynamic stage นั้นถูกนำมาจากชุดข้อมูล CICAndMal2020 ซึ่งประกอบด้วยข้อมูลพฤติกรรมของแอปพลิเคชันระหว่างการทำงาน เนื่องจาก ชุดข้อมูลดังกล่าวไม่มีการระบุป้ายกำกับ DCL ไว้ ผู้วิจัยจึงจำแนกกลุ่ม “DCL Malware” ผ่านกลไก DCL Filter ที่ตรวจจับพฤติกรรมการโหลดโค้ดภายนอกระหว่างรันไทม์โดยการใช้ DexClassLoader หรือ PathClassLoader ส่วนแอปพลิเคชันอื่นจะถูกจัดเป็น “Non-DCL” ก่อนเข้าสู่กระบวนการสกัดคุณลักษณะเชิงพฤติกรรม )Behavioral Features) ต่อไปอย่างไรก็ตาม งานวิจัยนี้จึงใช้ชุดข้อมูลที่ต่างกันสำหรับ static stage (AndroZoo) / dynamic stage (CICAndMal2020) ดังนั้น กระบวนการรวมคะแนนความเสี่ยงจึงมุ่งเน้นเพื่อสาธิตกรอบแนวคิดเชิงวิธีการ (Methodological Framework) เป็นหลัก

**3.3 การสกัดและวิเคราะห์ข้อมูลเชิงคุณลักษณะ**

ในงานวิจัยนี้ ผู้วิจัยได้ดำเนินการสกัดและออกแบบคุณลักษณะจากข้อมูลทั้งในส่วนของ static และ dynamic stage เพื่อให้ได้ตัวแปรที่สามารถสะท้อนลักษณะสำคัญของแอปพลิเคชันและพฤติกรรมที่เกี่ยวข้องกับ Dynamic Code Loading (DCL) ได้อย่างเหมาะสม

- สำหรับ static stage ผู้วิจัยได้ทำการสกัดข้อมูลจากโครงสร้างภายในไฟล์ APK โดยครอบคลุมองค์ประกอบสำคัญ เช่น permissions, components, strings, network indicators และตัวบ่งชี้ที่มีความเกี่ยวข้องกับการโหลดคลาสเชิงไดนามิก เช่น DexClassLoader และ PathClassLoader รวมถึงรูปแบบการใช้งานที่เกี่ยวข้อง นอกจากนี้ยังมีการออกแบบตัวแปรเชิงประกอบเพิ่มเติม เช่น suspicious score และ weak-label features เพื่อเพิ่มประสิทธิภาพ ในการจำแนก
- ในส่วนของ dynamic stage ผู้วิจัยนำข้อมูลพฤติกรรมของแอปพลิเคชันระหว่างการทำงานมาแปลงให้อยู่ในรูปแบบ คุณลักษณะเชิงตัวเลข เพื่อใช้ในการวิเคราะห์และจำแนกประเภท โดยมุ่งเน้นพฤติกรรมที่เกี่ยวข้องกับความเสี่ยงของมัลแวร์ เช่น การเชื่อมต่อเครือข่าย การเข้าถึงทรัพยากรภายนอก และพฤติกรรมที่เกี่ยวข้องกับการโหลดโค้ดเชิงไดนามิกระหว่างการทำงาน

หลังจากการสกัดคุณลักษณะ ผู้วิจัยได้ตรวจสอบและคัดออก perfect separator features เฉพาะบนชุดฝึกเท่านั้นเพื่อหลีกเลี่ยงข้อมูลรั่วไหลจากชุดตรวจสอบและชุดทดสอบเนื่องจากคุณลักษณะดังกล่าวอาจ ทำให้แบบจำลองเกิดการ overfitting ได้ดังนั้นจึงคัดเลือกคุณลักษณะให้เหลือ 28 features เพื่อลดความซับซ้อนของโมเดลและเพิ่มความสามารถในการทั่วไปของแบบจำลอง

**3.4 การพัฒนาแบบจำลอง**

ในงานวิจัยนี้ ผู้วิจัยเลือกใช้ XGBoost เป็นแบบจำลองหลักสำหรับทั้ง static stage และ dynamic stage เนื่องจากเป็นอัลกอริทึมที่เหมาะสมกับข้อมูลเชิงตารางและสามารถเรียนรู้ความสัมพันธ์ระหว่างคุณลักษณะได้อย่างมีประสิทธิภาพ ในส่วนของ static แบบจำลองจะรับข้อมูลจากคุณลักษณะเชิงโครงสร้างที่สกัดจากไฟล์ APK ขณะที่ dynamic stage ผู้วิจัยใช้คุณลักษณะเชิงพฤติกรรมที่ได้จากการทำงานของแอปพลิเคชัน เพื่อนำมาใช้ในการจำแนกประเภทของแอปพลิเคชัน ซึ่งในขั้นตอนการวิเคราะห์มีสมการดังนี้

$$\hat{y}_i = \sum_{k=1}^{K} f_k(x_i) \qquad (2)$$

โดยที่ $\hat{y}_i$ คือผลลัพธ์ที่แบบจำลองทำนายสำหรับตัวอย่างที่ $i$ โดยกำหนด $x_i$ เป็นเวกเตอร์ของคุณลักษณะตัวอย่างที่ $i$ และให้ $f_k(x_i)$ คือผลลัพธ์จากต้นไม้ตัดสินใจลำดับที่ $k$ โดย XGBoost ใช้แนวคิดการรวมผลลัพธ์จากการตัดสินใจเพื่อเพิ่มประสิทธิภาพในการจำแนกข้อมูล

**3.5 การรวบรวม Risk Score**

หลังจากได้ผลการประเมินจาก static stage และ dynamic stage แล้ว ผู้วิจัยนำผลลัพธ์ทั้งสองส่วนมารวมกันด้วยวิธี Combined Risk Score เพื่อให้การประเมินความเสี่ยงของแอปพลิเคชันสะท้อนทั้งข้อมูลเชิงโครงสร้างและข้อมูลเชิงพฤติกรรมร่วมกัน โดยคำนวณคะแนนรวมตามสมการดังนี้

$$S_{final} = \begin{cases} S_{static}, if\ S_{static} \geq \tau \\ w_s S_{static} + w_d S_{dynamic}, if\ S_{static} < \tau \end{cases} \quad (3)$$

โดยที่ $S_{final}$ แทนคะแนนความเสี่ยงรวมของแอปพลิเคชัน $S_{static}$ และ $S_{dynamic}$ คือคะแนนความเสี่ยงที่ได้จากขั้นตอนการวิเคราะห์เชิงสถิตและเชิงไดนามิกตามลำดับ และ $\tau$ คือค่าเกณฑ์สำหรับใช้เป็นเงื่อนไขในการกระตุ้นการวิเคราะห์เชิงไดนามิก ในกรณีที่ $S_{static}$ มีค่าน้อยกว่าค่าเกณฑ์ดังกล่าว ระบบจะใช้เฉพาะผลจากการวิเคราะห์เชิงสถิตเพื่อประเมินความเสี่ยงเพื่อลดการใช้ทรัพยากร แต่หาก $S_{static}$ มีค่ามากกว่าหรือเท่ากับ $\tau$ จะทำการรวมผลลัพธ์จากทั้งสองขั้นตอนโดยใช้ค่าน้ำหนัก $w_s$ และ $w_d$ ซึ่งกำหนดให้ $w_s + w_d = 1$ และ $w_s, w_d \in [0,1]$

สำหรับกรณีที่แอปพลิเคชันมีค่าคะแนนความเสี่ยงจากการวิเคราะห์เชิงสถิต ต่ำกว่าเกณฑ์ที่กำหนดไว้ (ไม่มีความเสี่ยงเพียงพอที่จะถูกส่งเข้าสู่กระบวนการวิเคราะห์เชิงไดนามิก) ระบบจะข้ามการคำนวณในสมการ (3) และกำหนดให้คะแนนความเสี่ยงรวมมีค่าเท่ากับคะแนนจากขั้นตอนการวิเคราะห์เชิงสถิตทันที เพื่อลดการใช้ทรัพยากรตามแนวทาง risk-triggered

งานวิจัยนี้ประยุกต์ใช้การกำหนดค่าน้ำหนักแบบปรับตามบริบท (adaptive weighting) ผ่านกระบวนการหาค่าพารามิเตอร์ที่เหมาะสมที่สุด เพื่อให้ได้อัตราส่วนที่ให้ประสิทธิภาพการจำแนกสูงสุด โดยมีเกณฑ์การปรับค่าน้ำหนักอิงตามระดับความชัดเจนของผลลัพธ์เชิงไดนามิก กล่าวคือหาก $S_{dynamic} \geq 50$ แสดงว่าตรวจพบพฤติกรรม DCL ได้ชัดเจน จึงให้ Dynamic มีน้ำหนักสูงถึง 0.8 และ Static มีค่าน้ำหนัก 0.2 แต่หาก $S_{dynamic} < 50$ แสดงว่าพฤติกรรม DCL ยังไม่ชัดเจนเพียงพอ ระบบจะปรับลดน้ำหนัก Dynamic เป็น 0.6 และปรับเพิ่มน้ำหนัก Static เป็น 0.4 เพื่อป้องกันการตัดสินใจที่ผิดพลาดและลด False Positive ทั้งนี้ผลลัพธ์สุดท้ายจะถูกแปลงเป็นระดับความเสี่ยง 5 ระดับ ได้แก่ LOW, MODERATE, ELEVATED, HIGH และ CRITICAL

**3.6 การประเมินเงื่อนไขกระตุ้น**

งานวิจัยนี้เสนอเทคนิคการวิเคราะห์มัลแวร์แบบ risk-targeted ที่มุ่งเน้นการคัดกรองแอปพลิเคชันที่มีความเสี่ยงจากผลการวิเคราะห์เชิงสถิตก่อน เพื่อจำกัดการใช้ทรัพยากรของการวิเคราะห์เชิงไดนามิกให้เฉพาะกรณีที่จำเป็นเท่านั้น โดยอัลกอริทึมกำหนดเกณฑ์ $\tau$ (ในงานวิจัยนี้กำหนดค่าไว้ที่ 35) สำหรับตัดสินว่าแอปพลิเคชันมีความเสี่ยงเพียงพอหรือไม่ หากค่า $S_{static}$ สูงกว่าค่าดังกล่าว จะถูกส่งต่อเข้าสู่กระบวนการ Dynamic Analysis ทันที

**Algorithm 1: Risk-Targeted Dynamic Analysis Procedure**

```
Input: Android application package (APK), static confidence score S_static,
trigger threshold τ
Output: final risk score S_final

1: if S_static ≥ τ then
2:     Execute the APK in a VM-based sandbox environment
3:     Monitor execution and collect behavioral logs using a Python controller
4:     Extract raw dynamic features X_dynamic(raw) from the collected logs
5: Filter DCL-related behaviors:
6:     X_dynamic(filtered) ← ApplyDCLFilter(X_dynamic(raw))
7: Construct the behavior sequence:
8:     X_sequence ← SequenceConstruction(X_dynamic(filtered))
9: Perform dynamic classification:
10:    S_dynamic ← XGBoostPredict(M_dynamic, X_sequence)
11: Compute combined risk score:
12:    S_final ← CombineRisk(S_static, S_dynamic)
13: else
14:    S_final ← S_static
15: end if
16: return S_final
```

แนวทางนี้ช่วยลด overhead ของระบบ sandbox และเพิ่มประสิทธิภาพในการตรวจจับพฤติกรรมซับซ้อน เช่น Dynamic Code Loading (DCL) ได้อย่างมีนัยสำคัญ จึงเหมาะกับงานวิเคราะห์มัลแวร์ขนาดใหญ่ที่ต้องการทั้งความเร็วและความแม่นยำในระดับสูง โดยกระบวนการตัดสินใจเชิง risk-triggered สามารถสรุป pseudocode ได้ออกมาตาม Algorithm 1

จาก Algorithm 1 แสดงให้เห็นถึง อัลกอริทึมนี้เริ่มจากการประเมินค่า $S_{static}$ เพื่อใช้เป็นเงื่อนไขกระตุ้น (risk trigger) หากค่าดังกล่าวสูงกว่าค่าเกณฑ์ $\tau$ ระบบจะเข้าสู่กระบวนการวิเคราะห์เชิงไดนามิก โดยทำการรัน APK ใน sandbox เพื่อเก็บ logs และสกัดคุณลักษณะเชิงพฤติกรรม จากนั้นทำการคัดกรองเฉพาะพฤติกรรมที่เกี่ยวข้องกับ DCL และจัดลำดับเป็น sequence ก่อนนำเข้าสู่โมเดล XGBoost เพื่อคำนวณค่า $S_{dynamic}$ สำหรับตัดสินความเสี่ยงขั้นสุดท้าย

กำหนดให้ค่า $\tau = 35$ เป็นเกณฑ์สำหรับกระตุ้นการเข้าสู่การวิเคราะห์แบบไดนามิก (trigger threshold) จากผล static analysis ส่วนเงื่อนไข $S_{dynamic} \geq 50$ ถูกใช้ในภายหลังเพื่อปรับน้ำหนักของผล dynamic analysis ในขั้นตอนการรวมคะแนนความเสี่ยง (Combined Risk Score) โดยแบบ adaptive weighting เป็นคนละวัตถุประสงค์กับ threshold สำหรับ trigger

## 4. การประเมินประสิทธิภาพ

### 4.1 การประเมินผลแบบจำลอง

การประเมินผลในงานวิจัยนี้ครอบคลุมทั้งมิติด้านประสิทธิภาพและความเสถียรของแบบจำลอง สำหรับ static stage ชุดข้อมูลถูกแบ่งเป็น training, validation และ test set ในอัตราส่วน 176:44:74 เพื่อรองรับการพัฒนาและประเมินแบบจำลองอย่างเป็นระบบ โดยใช้ 5-fold cross-validation เพื่อลดอคติจากการแบ่งข้อมูล และใช้ multi-seed evaluation จำนวน 20 seeds เพื่อรายงานผลลัพธ์ในรูปของค่าเฉลี่ย โดยช่วยเพิ่มความน่าเชื่อถือและความสามารถในการทำซ้ำของการทดลอง

ขณะที่ dynamic stage ถูกประเมินจากข้อมูลทดสอบสำหรับการจำแนก DCL Malware และ Non-DCL โดยอาศัยข้อมูลพฤติกรรมที่ได้จากการทำงานจริงของแอปพลิเคชัน เพื่อสะท้อนความสามารถของแบบจำลองในการตรวจจับพฤติกรรมที่เกี่ยวข้องกับ Dynamic Code Loading ได้อย่างเหมาะสม ทั้งนี้ ตัวชี้วัดที่ใช้ประกอบด้วย Accuracy, ROC-AUC, Precision, Recall และ F1-score ซึ่งครอบคลุมทั้งความสามารถในการจำแนกโดยรวม การแยกคลาส และความสมดุลระหว่างอัตราการตรวจจับกับความผิดพลาดของแบบจำลอง

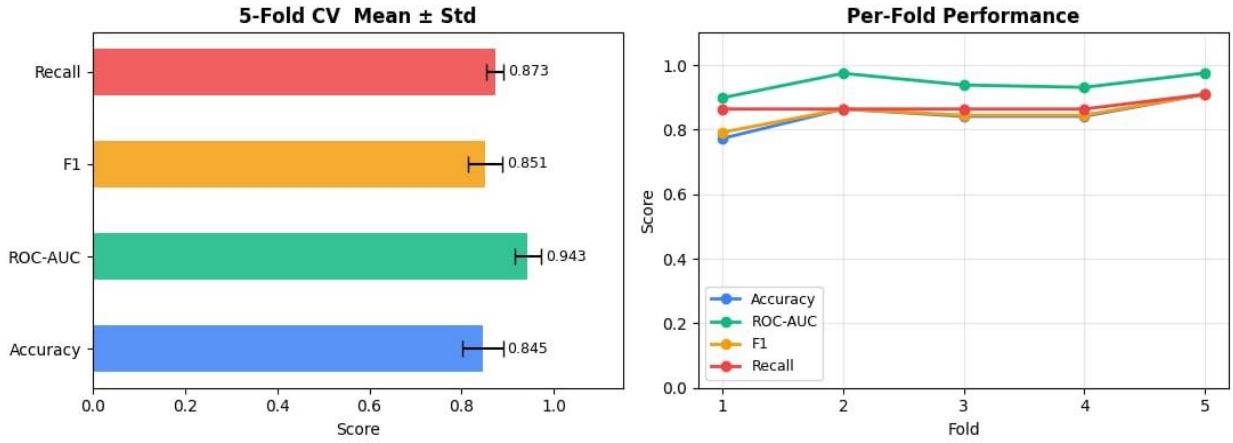
### 4.2 การอธิบายผลของแบบจำลองด้วย SHAP

เพื่อเพิ่มความสามารถในการอธิบายการตัดสินใจของแบบจำลอง ผู้วิจัยใช้ SHapley Additive exPlanations (SHAP) ในการวิเคราะห์ความสำคัญของคุณลักษณะในแบบจำลอง XGBoost สำหรับ static stage วิธีการดังกล่าวช่วยให้สามารถระบุได้ว่าคุณลักษณะแต่ละตัวมีอิทธิพลต่อผลการจำแนกมากน้อยเพียงใด

โดยการใช้ SHAP ช่วยให้ผลลัพธ์ของแบบจำลองสามารถอธิบายได้ในเชิงความหมายมากขึ้นโดยเฉพาะการเชื่อมโยงคุณลักษณะที่สำคัญกับลักษณะของแอปพลิเคชันที่อาจมีความเกี่ยวข้องกับ Dynamic Code Loading เช่น การซ่อนข้อมูล การเรียกใช้สิทธิ์ที่มีความเสี่ยง และการโหลดคลาสเชิงไดนามิก

### 4.3 การตั้งค่าการทดลอง

การทดลองในบทนี้แบ่งเป็น 2 ส่วน ได้แก่ static stage และ dynamic stage โดย static stage ใช้ข้อมูล APK จาก AndroZoo จำนวน 294 ตัวอย่าง ในสัดส่วน benign:malware เท่ากับ 1:1 และแบ่งข้อมูลเป็นชุดฝึก ชุดตรวจสอบ และชุดทดสอบในอัตรา 176:44:74 ตามลำดับ ส่วน dynamic stage ใช้ข้อมูลจาก CICAndMal2020 สำหรับวิเคราะห์พฤติกรรมขณะทำงานของแอปพลิเคชัน ทั้งนี้ตัวชี้วัดที่ใช้ประเมินประกอบด้วย Accuracy, ROC-AUC, Precision, Recall และ F1-score




**รูปภาพที่ 3:** ผลการประเมินแบบ 5-fold cross-validation ของ XGBoost สำหรับ static stage

**ตารางที่ 3:** ตารางเปรียบเทียบประสิทธิภาพของแบบจำลอง XGBoost และ Random Forest

| Metric | XGBoost (CV) | Random Forest (CV) | XGBoost (Test) | Random Forest (test) |
|---|---|---|---|---|
| Accuracy | 0.845 | 0.891 | 0.919 | 0.932 |
| ROC-AUC | 0.943 | 0.957 | 0.976 | 0.980 |
| F1-score | 0.851 | 0.890 | 0.919 | 0.932 |
| Recall | 0.873 | 0.882 | 0.946 | 0.932 |

### 4.4 ผลของ XGBoost (Static Stage)

การจำแนก XGBoost สำหรับ static stage ถูกฝึกด้วยเทคนิค early stopping และหยุดที่รอบที่ 129 ผลการจำแนกบนชุดทดสอบจำนวน 74 ตัวอย่างแสดงในตารางที่ 2 โดยโมเดลสามารถตรวจจับ malware ได้ถูกต้องทั้ง 35 ตัวอย่าง จาก 37 ตัวอย่าง (Recall = 0.946) ขณะที่กลุ่ม benign ถูกจำแนกผิดเป็น malware จำนวน 4 ตัวอย่าง ส่งผลให้ Accuracy บนชุดทดสอบเท่ากับ 0.919 และ ROC-AUC เท่ากับ 0.976 ซึ่งสะท้อนว่าโมเดลมีความสามารถในการแยก benign และ malware ได้ในระดับสูง ตามตารางที่ 2 และรูปภาพที่ 3

ตารางที่ 2 แสดงผลการจำแนกของแบบจำลองของ XGBoost สำหรับ static stage บนชุดทดสอบจำนวน 74 (176:44:74) ตัวอย่าง โดยโมเดลตรวจจับ malware ได้ถูกต้อง 35 จาก 37 ตัวอย่าง (Recall = 0.946) ขณะที่กลุ่ม benign มีการจำแนกผิดเป็น malware จำนวน 4 ตัวอย่าง ส่งผลทำให้ค่า F1-score เฉลี่ย (Macro avg) เท่ากับ 0.919 ซึ่งสะท้อนว่าแบบจำลองมีประสิทธิภาพในการแยก benign และ malware ได้ในระดับที่ดี

**ตารางที่ 2:** ผลการจำแนกของ XGBoost สำหรับ Static Stage

| Class | Precision | Recall | F1 | Sup. |
|---|---|---|---|---|
| Benign | 0.943 | 0.892 | 0.917 | 37 |
| Malware | 0.897 | 0.946 | 0.921 | 37 |
| Macro avg | 0.920 | 0.919 | 0.919 | 74 |

รูปภาพที่ 3 แสดงถึงความเสถียรของแบบจำลอง ผลการประเมินแบบ 5-fold cross-validation ของ XGBoost สำหรับ static stage โดยให้ค่า Accuracy = $0.845 \pm 0.044$ , ROC-AUC = $0.943 \pm 0.029$ , F1 - score = $0.851 \pm 0.038$ และ Recall = $0.873 \pm 0.018$ โดยแต่ละค่าของ fold มีแนวโน้มใกล้เคียงกัน แสดงว่าโมเดลให้ผลลัพธ์ค่อนข้างสม่ำเสมอและไม่ได้พึ่งพาผลจากชุดที่นำมาทดสอบอย่างเดียว

### 4.5 การเปรียบเทียบ XGBoost กับ Random Forest

เพื่อประเมินความเหมาะสมของแบบจำลองหลัก งานวิจัยนี้ใช้ Random Forest เป็น baseline ภายใต้ชุดข้อมูลและกระบวนการแบ่งข้อมูลเดียวกันกับ XGBoost แล้วเปรียบเทียบผลการจำแนกทั้งในมุมของ cross-validation

**ตารางที่ 4:** คุณลักษณะสำคัญของ XGBoost จากการวิเคราะห์ด้วย SHAP

| ลำดับ | คุณลักษณะ | สัมพันธ์กับ DCL |
|---|---|---|
| 1 | perm_read_phone_state | สะท้อนการร้องขอสิทธิ์เข้าถึงข้อมูลสถานะอุปกรณ์ ซึ่งพบในแอปพลิเคชันที่มีพฤติกรรมเสี่ยง และอาจเกี่ยวข้องกับรูปแบบการทำงานที่พบร่วมกับมัลแวร์ แต่ไม่สามารถยืนยัน DCL ได้โดยตรง |
| 2 | base64ish_count | บ่งชี้ว่าการมีสตริงที่มีลักษณะคล้ายข้อมูลที่ถูกเข้ารหัสหรือ obfuscation ซึ่งอาจสัมพันธ์กับการซ่อน payload หรือโค้ดที่ถูกโหลดภายหลัง จึงเป็นตัวบ่งชี้เชิงอ้อมของความเสี่ยงที่เกี่ยวข้องกับ DCL |
| 3 | provider_count | แสดงจำนวนของ Content Provider ภายในแอปพลิเคชัน ซึ่งอาจสะท้อนความซับซ้อนของโครงสร้างแอปพลิเคชันหรือรูปแบบการจัดการข้อมูลบางประเภท โดยอาจมีความเกี่ยวข้องทางอ้อมกับแอปพลิเคชันที่มีพฤติกรรมเสี่ยง |
| 4 | dex_string_count | ปริมาณสตริงที่เกี่ยวข้องกับไฟล์ DEX อาจสะท้อนการอ้างอิงหรือการจัดการโค้ดเพิ่มเติมระหว่างการทำงาน จึงอาจใช้เป็นสัญญาณประกอบในการประเมินความเสี่ยงด้าน DCL แต่ไม่ใช่หลักฐานที่ใช้ยืนยันเพียงอย่างเดียว |

นอกจากนี้ผลการทดสอบจริง แม้ว่า Random Forest อาจให้ผลลัพธ์ที่ดีกว่า XGBoost เล็กน้อยในบางตัวชี้วัด แต่ผู้วิจัยเลือก XGBoost เป็นแบบจำลองหลัก ทั้งนี้เนื่องจากสอดคล้องกับวัตถุประสงค์ของงานวิจัยที่ไม่ได้มุ่งเพียงประสิทธิภาพการจำแนกเท่านั้น แต่ยังให้ความสำคัญกับความสามารถในการอธิบายผลของแบบจำลอง (interpretability) เพื่อวิเคราะห์ว่าคุณลักษณะใดมีส่งผลต่อการตัดสินใจของระบบด้วย

โดยเฉพาะการประยุกต์ใช้ SHAP ในการตีความปัจจัยเสี่ยงที่เกี่ยวข้องกับพฤติกรรมของมัลแวร์ที่ใช้ Dynamic Code Loading (DCL) ดังนั้น การเลือก XGBoost จึงเป็นการตัดสินใจที่คำนึงทั้งประสิทธิภาพ ความสามารถในการอธิบายผล และความสอดคล้องกับเป้าหมายของกรอบวิธีวิจัยโดยรวม

จากตารางที่ 3 แสดงผลการเปรียบเทียบประสิทธิภาพของโมเดล XGBoost และ Random Forest ภายใต้ชุดข้อมูลและกระบวนการประเมินผลเดียวกัน โดยใช้ 5-fold cross-validation และการทดสอบกับชุดข้อมูลจริง (test set) ผลการทดลองพบว่า Random Forest ให้ค่าประสิทธิภาพที่สูงกว่า XGBoost เล็กน้อยในหลายตัวชี้วัด อย่างไรก็ตาม ความแตกต่างของค่าประสิทธิภาพระหว่างทั้งสองโมเดลอยู่ในระดับที่ไม่มาก โดยเฉพาะค่า ROC-AUC ซึ่งสะท้อนการแยกแยะข้อมูลระหว่างคลาส พบว่า XGBoost มีค่า 0.976 ในชุดทดสอบ ซึ่งใกล้เคียงกับ Random Forest ที่มีค่า 0.980

อย่างไรก็ตาม งานวิจัยนี้เลือกใช้ XGBoost เป็นโมเดลหลัก เนื่องจากสามารถนำไปใช้ร่วมกับ SHAP เพื่อวิเคราะห์ความสำคัญของคุณลักษณะและอธิบายการตัดสินใจของโมเดลได้อย่างชัดเจน

**4.6 ความสำคัญของคุณลักษณะจากการวิเคราะห์ด้วย SHAP**

เพื่ออธิบายการตัดสินใจของแบบจำลอง XGBoost งานวิจัยนี้ใช้ SHAP (SHapley Additive exPlanations) ในการวิเคราะห์ความสำคัญของคุณลักษณะ ซึ่งช่วยแสดงว่าคุณลักษณะแต่ละตัวมีอิทธิพลต่อผลการจำแนกมากน้อยเพียงใด โดยผลการวิเคราะห์แสดงในรูปภาพที่ 4 และสรุปคุณลักษณะที่สำคัญในตารางที่ 4




**รูปภาพที่ 4:** ความสำคัญของคุณลักษณะจากการวิเคราะห์ด้วย SHAP ของแบบจำลองใน static stage

โดยที่รูปภาพที่ 4 แสดงผลการวิเคราะห์ SHAP Feature Importance ของแบบจำลอง XGBoost โดยแสดงคุณลักษณะสำคัญ 20 อันดับแรกที่มีผลต่อการจำแนกมัลแวร์ใน static stage ผลลัพธ์พบว่า base64ish_count และ perm_read_phone_state เป็นคุณลักษณะที่มีอิทธิพลต่อการตัดสินใจของโมเดลมากที่สุด รองลงมาคือคุณลักษณะที่เกี่ยวข้องกับโครงสร้างของแอปพลิเคชันและพฤติกรรมที่อาจสัมพันธ์กับ Dynamic Code Loading (DCL) เช่น dex_string_count, ip_like_count และ has_dexclassloader ซึ่งสะท้อนถึงลักษณะพฤติกรรมที่มักพบในแอปพลิเคชันที่มีความเสี่ยง

ตารางที่ 4 แสดงคุณลักษณะสำคัญจากการวิเคราะห์ด้วย SHAP พบว่าคุณลักษณะที่มีอิทธิพลสูงสุดที่มีความสัมพันธ์พฤติกรรม DCL ได้แก่ base64ish_count ที่บ่งชี้การซ่อน payload, perm_read_phone_state ที่สะท้อนการตรวจสอบอุปกรณ์เป้าหมาย, dex_string_count ที่เกี่ยวข้องกับการโหลดโค้ดระหว่างรันไทม์ และ provider_count ที่บ่งชี้โครงสร้างการจัดการข้อมูลภายในแอปพลิเคชัน ผลดังกล่าวสนับสนุนว่าแบบจำลองสามารถเรียนรู้รูปแบบที่สัมพันธ์กับกลไก DCL ได้อย่างมีประสิทธิภาพ

จากการวิเคราะห์ค่าของ SHAP จะพบได้ว่า perm_read_phone_state และ base64ish_count ทั้งสองเป็นคุณลักษณะที่มีอิทธิพลต่อการตัดสินใจของแบบจำลองอย่างมีนัยสำคัญ โดยทั้งสองสะท้อนพฤติกรรมของมัลแวร์ในเทคนิค Dynamic Code Loading (DCL) อย่างเป็นลำดับ ผลวิเคราะห์ SHAP ชี้ว่าคุณลักษณะ perm_read_phone_state และ base64ish_count มีอิทธิพลสูงต่อการตัดสินใจของแบบจำลอง อย่างไรก็ตามควรตีความผลดังกล่าวในฐานะตัวบ่งชี้เชิงสถิติ ไม่ใช่ความสัมพันธ์เชิงสาเหตุ (Causal relationship) ของลำดับพฤติกรรมมัลแวร์

**4.7 Dynamic Stage และ Combined Risk Score**

แม้ static stage จะสามารถตรวจจับลักษณะเชิงโครงสร้างของแอปพลิเคชันได้ดี แต่ยังไม่สะท้อนพฤติกรรมที่เกิดขึ้นระหว่างการทำงานทั้งหมด ดังนั้นงานวิจัยนี้จึงใช้ dynamic stage จากชุดข้อมูล CICAndMal2020 เพื่อวิเคราะห์พฤติกรรมของแอปพลิเคชันและจำแนกกลุ่มที่เกี่ยวข้องกับ Dynamic Code Loading (DCL) ผลการจำแนกของแบบจำลอง XGBoost แสดงในตารางที่ 5 และเพื่อรวมจุดเด่นของทั้งสองขั้นตอนงานวิจัยนี้จึงใช้แนวทาง Combined Risk Score โดยนำผลจาก static และ dynamic มาถ่วงน้ำหนักร่วมกันเพื่อประเมินระดับความเสี่ยงของแอปพลิเคชัน ซึ่งเกณฑ์การแปลผลคะแนนรวมแสดงในตารางที่ 6

**ตารางที่ 5:** ผลการจำแนกของ XGBoost ใน Dynamic Stage

| Class | Precision | Recall | F1 | Supp. |
|---|---|---|---|---|
| Non-DCL | 0.942 | 0.930 | 0.936 | 9864 |
| DCL Malware | 0.809 | 0.838 | 0.823 | 3496 |
| Macro avg | 0.875 | 0.884 | 0.880 | 13360 |
| Weighted avg | 0.907 | 0.906 | 0.906 | 13360 |

**ตารางที่ 6:** เกณฑ์การแปลผลคะแนนความเสี่ยง

| ช่วงคะแนนรวม | ระดับความเสี่ยง |
|---|---|
| 0.00–19.99 | LOW |
| 20.00–34.99 | MODERATE |
| 35.00–54.99 | ELEVATED |
| 55.00–74.99 | HIGH |
| 75.00–100.00 | CRITICAL |

ตารางที่ 5 แสดงผลการประเมินประสิทธิภาพของแบบจำลอง XGBoost ในขั้นตอน Dynamic Stage โดยใช้ตัวชี้วัด Precision, Recall และ F1-score สำหรับการจำแนกแอปพลิเคชันออกเป็นสองกลุ่ม ได้แก่ Non-DCL และ DCL Malware ผลการทดลองพบว่าโมเดลสามารถจำแนกแอปพลิเคชันในกลุ่ม Non-DCL ได้ดี โดยมีค่า F1-score เท่ากับ 0.936 ซึ่งสะท้อนถึงความสามารถของโมเดลในการระบุแอปพลิเคชันที่ไม่มีพฤติกรรม Dynamic Code Loading ได้อย่างแม่นยำ

**ตารางที่ 7:** การเปรียบเทียบภาระของระบบและอัตราการหลุดรอดของมัลแวร์เมื่อปรับค่าเกณฑ์กระตุ้น

| ค่า Threshold | สัดส่วนของแอปพลิเคชันที่ถูกส่งไป Dynamic Analysis | อัตรา Overhead (%) | FN | False Negative (%) |
|---|---|---|---|---|
| 20 | 47/74 | 63.51% | 1/37 | 2.70% |
| 35 | 39/74 | 52.70% | 2/37 | 5.41% |
| 50 | 39/74 | 52.70% | 2/37 | 5.41% |
| 65 | 36/74 | 48.65% | 3/37 | 8.11% |
| 80 | 30/74 | 40.54% | 7/37 | 18.92% |
| 95 | 0/74 | 0.00% | 37/37 | 100.00% |

สำหรับกลุ่ม DCL Malware โมเดลมีค่า F1-score = 0.823 และ Recall = 0.838 แสดงให้เห็นว่าสามารถตรวจจับพฤติกรรมที่เกี่ยวข้องกับ Dynamic Code Loading (DCL) ได้ในระดับที่น่าพอใจ แม้ว่าค่าประสิทธิภาพจะต่ำกว่ากลุ่ม Non-DCL เล็กน้อย เนื่องจากความซับซ้อนของพฤติกรรมมัลแวร์ โดยภาพรวม โมเดลมีค่า Macro F1-score = 0.880 และ Weighted F1-score = 0.906 แสดงให้เห็นว่าแบบจำลองมีประสิทธิภาพที่ดีในการจำแนกทั้งสองคลาส และสามารถนำไปใช้ในขั้นตอน Combined Risk Score ได้อย่างเหมาะสม

**4.8 การวิเคราะห์ความอ่อนไหวของค่าเกณฑ์กระตุ้น (Sensitivity Analysis)**

เพื่อให้การกำหนดค่าเกณฑ์กระตุ้น (Trigger Threshold, $\tau$) สำหรับการนำแอปพลิเคชันเข้าสู่กระบวนการวิเคราะห์เชิงไดนามิกมีประสิทธิภาพสูงสุด ผู้วิจัยได้ทำการวิเคราะห์ความอ่อนไหว (Sensitivity Analysis) โดยอ้างอิงจากผลการทำนายของแบบจำลองบนชุดข้อมูลทดสอบ (ชุดข้อมูลทดสอบ Test Set จำนวน 74 ตัวอย่าง) การวิเคราะห์นี้ดำเนินการโดยปรับค่า $\tau$ ระหว่าง 20 ถึง 95 เพื่อสังเกตผลกระทบใน 2 มิติหลัก ได้แก่ ภาระการทำงานของระบบ (System Overhead) ซึ่งพิจารณาจากสัดส่วนของแอปพลิเคชันที่ถูกส่งไปวิเคราะห์ต่อ และอัตราการหลุดรอดของมัลแวร์ (False Negative Rate) ดังแสดงในตารางที่ 7

จากตารางที่ 7 แสดงความสัมพันธ์แบบแปรผกผัน )Trade-off) ระหว่างภาระของระบบ (System Overhead) และอัตราการหลุดของมัลแวร์ (False Negative) เมื่อปรับค่าเกณฑ์กระตุ้น ($\tau$) จาก 20, 35, 50, 65, 80, 95 ผลการทดลองพบว่า การตั้งค่า $\tau$ ในระดับต่ำ (20) มีความรัดกุมสูงโดยมีอัตราการหลุดรอดเพียง 2.70 % แต่แลกมาด้วยภาระระบบสูงถึง 63.51% ในทางกลับกัน หากปรับค่า $\tau$ ให้สูงขึ้น (80-95) แม้จะช่วยลดภาระระบบได้ แต่ทำให้อัตราการหลุดรอดของมัลแวร์เพิ่มสูงขึ้นจนระดับไม่สามารถยอมรับได้ (18.92%-100%)

อย่างไรก็ตาม ผลการทดลองแสดงให้เห็นช่วงความเสถียรที่ค่า $\tau$ ระหว่าง 35 ถึง 50 ซึ่งอัตราภาระระบบและอัตราความผิดพลาดมีค่าคงที่ ดังนั้นภายใต้เงื่อนไข การเลือกใช้ $\tau = 35$ จึงถือเป็นเกณฑ์ที่ให้สมดุลเชิงปฏิบัติระหว่างการจำกัดทรัพยากรในขั้นตอนการวิเคราะห์เชิงไดนามิกและการรักษาประสิทธิภาพในการจับมัลแวร์ได้อย่างเหมาะสม

## 5. สรุปผลการทดลอง

งานวิจัยนี้นำเสนอแนวทางการตรวจจับแอปพลิเคชัน Android ที่มีพฤติกรรมเกี่ยวข้องกับ Dynamic Code Loading (DCL) โดยใช้การวิเคราะห์แบบสองขั้นตอน ได้แก่ static stage และ dynamic stage เพื่อสะท้อนทั้งลักษณะเชิงโครงสร้างและพฤติกรรมระหว่างการทำงานของแอปพลิเคชัน ในขั้นตอน static stage มีการสกัดคุณลักษณะจากไฟล์ APK เช่น permissions, components, strings และตัวบ่งชี้ของ class loader ขณะที่ dynamic stage วิเคราะห์พฤติกรรมขณะทำงานของแอป

พลิเคชัน เช่น การเชื่อมต่อเครือข่ายและการโหลดโค้ดระหว่างรันไทม์

การทดลองด้วยแบบจำลอง XGBoost ใน dynamic stage ให้ค่า Accuracy 0.906 และ ROC-AUC 0.956 พร้อมค่า Recall ของ DCL Malware 0.838 ซึ่งสะท้อนความสามารถของโมเดลในการตรวจจับแอปพลิเคชันที่มีความเสี่ยงได้ดี งานวิจัยนี้จึงนำผลจากทั้งสองขั้นตอนมารวมกันผ่านแนวทาง Combined Risk Score เพื่อประเมินระดับความเสี่ยงของแอปพลิเคชันโดยรวม ทำให้สามารถคัดกรองแอปพลิเคชันที่มีพฤติกรรมเกี่ยวข้องกับ DCL ได้อย่างมีประสิทธิภาพมากขึ้น

อย่างไรก็ตาม งานวิจัยนี้ยังมีข้อจำกัดในด้านขนาดของชุดข้อมูลสำหรับการวิเคราะห์เชิงสถิติที่มีจำนวนค่อนข้างจำกัดเมื่อเทียบกับข้อมูลเชิงไดนามิก ซึ่งในอนาคตสามารถพัฒนาต่อยอดโดยการรวบรวมและขยายขนาดชุดข้อมูลเชิงสถิติให้มีความหลากหลายและสมดุลมากยิ่งขึ้น เพื่อเพิ่มความแม่นยำและความเสถียรของเกณฑ์การประเมินความเสี่ยงรวมแบบไฮบริดให้มีประสิทธิภาพสูงสุด

## เอกสารอ้างอิง


[1] M. Gopakumar and S. Chandran, “Detection of dynamic code loading in Android,” in *Proc. 13th Int. Symp. Digit. Forensics Security (ISDFS)*, 2025, doi:10.1109/ISDFS65363.2025.11011942

[2] A. El Attaoui, N. El Hami, and Y. Koulou, “Android malware detection using the random forest algorithm,” *Indones. J. Electr. Eng. Comput. Sci.*, vol. 36, no. 3, pp. 1876–1883, 2024, doi:10.11591/ijeecs.v36.i3.pp1876-1883.

[3] J. Yang, J. Tang, R. Yan, and T. Xiang, “Android malware detection method based on permission complement and API calls,*” Chinese J. Electron.*, vol. 31, no. 4, pp. 773-785, 2022. doi:10.1049/cje.2020.00.217

[4] R. Bhan et al., “DLCDroid: An Android apps analysis framework to analyse the dynamically loaded code,” *Sci. Rep.*, vol. 15, Art. no. 3292, 2025, doi: 10.1038/s41598-025-88003-6.

[5] D. Arp et al., “DREBIN: Effective and explainable detection of Android malware in your pocket,” in Proc. *Netw. Distrib. Syst. Security Symp. (NDSS)*, 2014, doi:10.14722/ndss.2014.23247.

[6] Z. Qu et al., “DyDroid: Measuring dynamic code loading and its security implications in Android applications,” in Proc. *IEEE/IFIP Int. Conf. Dependable Syst. Netw. (DSN)*, 2017, pp. 415–426, doi:10.1109/DSN.2017.40

[7] W. Enck, P. Gilbert, B.-G. Chun, L. P. Cox, J. Jung, P. McDaniel, and A. N. Sheth, “TaintDroid: An information-flow tracking system for realtime privacy monitoring on smartphones,” *ACM Transactions on Computer Systems*, vol. 32, no. 2, article 5, 2014, doi:10.1145/2619091.

[8] L. K. Yan and H. Yin, “DroidScope: Seamlessly reconstructing the OS and Dalvik semantic views for dynamic Android malware analysis,” in Proc. *21st USENIX Security Symp. (USENIX Security)*, 2012, pp. 569–584, doi:10.5555/2362793.2362822

# การประยุกต์ใช้ NIST AI Risk Management Framework และ NCSA AI Security Guidelines สำหรับการบริหารความเสี่ยงจาก Generative AI ในองค์กร

# Applying NIST AI RMF and NCSA AI Security Guidelines to Manage Generative AI Risks in Organizations

*พัทธพล แย้มแว่นแก้ว (Phattapol Yamwankaew)[1] และนวพร วิสิฐพงศ์พันธ์ (Nawaporn Wisitpongphan)[2]*
*ภาควิชาการบริหารเครือข่ายดิจิทัลและความมั่นคงปลอดภัยสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล*
*มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ*
*phattapol.ya@hotmail.com, nawapornn@kmutnb.ac.th*

## บทคัดย่อ

*งานวิจัยนี้มุ่งเน้นการประยุกต์ใช้ NIST AI Risk Management Framework (AI RMF 1.0) ร่วมกับแนวทางปฏิบัติ NCSA AI Security Guideline เพื่อพัฒนาแนวทางการบริหารความเสี่ยงที่เกิดจากการใช้งาน Generative AI ในกระบวนการพัฒนาซอฟต์แวร์ขององค์กรเทคโนโลยีขนาดกลาง โดยทดสอบกับทีมนักพัฒนาจำนวน 37 คนที่ใช้งาน GitHub Copilot, ChatGPT และ Claude ผลการวิเคราะห์ช่องว่างพบว่าระดับวุฒิภาวะด้านการบริหารความเสี่ยง AI โดยรวมอยู่ที่ 1.56 จาก 4.0 (ระดับ Ad-hoc) โดยมีช่องว่างสำคัญระหว่างมุมมองผู้บริหาร (CTO) และผู้ปฏิบัติงาน โดยเฉพาะในฟังก์ชัน MAP ที่แตกต่างถึง +1.43 คะแนน เครื่องมือ Generative AI Risk Assessment Tool ที่พัฒนาขึ้นสามารถระบุช่องโหว่ Prompt Injection ได้ 3 จุดที่ไม่เคยถูกระบุมาก่อน และช่วยเร่งการร่างนโยบาย Acceptable Use Policy จากเวลาประมาณการเดิม 2-3 สัปดาห์เหลือเพียง 1 วัน*

**คำสำคัญ:** ปัญญาประดิษฐ์เชิงสร้างสรรค์ การจัดการความเสี่ยง มาตรฐาน NIST AI RMF ปัญญาประดิษฐ์ที่น่าเชื่อถือ ความปลอดภัยทางไซเบอร์

## Abstract

*This research focuses on applying the NIST AI Risk Management Framework (AI RMF 1.0) in conjunction with the NCSA AI Security Guidelines to develop concrete risk management approaches for Generative AI in the software development process of a mid-sized technology organization. The framework was tested with 37 developers using GitHub Copilot, ChatGPT, and Claude. Gap analysis revealed an overall AI risk management maturity level of 1.56 out of 4.0 (Ad-hoc level), with significant perception gaps between the CTO and operational teams, particularly in the MAP function (+1.43 difference). The Generative AI Risk Assessment Tool developed identified 3 previously unknown Prompt Injection vulnerabilities and accelerated Acceptable Use Policy drafting from an estimated 2-3 weeks to just 1 day.*

**Keywords:** Generative AI Risk Management NIST AI RMF Trustworthy AI Software Development Cybersecurity

## 1. บทนำ

ในยุคดิจิทัลปัจจุบัน Generative AI ได้เข้ามามีบทบาทสำคัญในกระบวนการพัฒนาซอฟต์แวร์ โดยเฉพาะในองค์กรที่มีทีมนักพัฒนาขนาดกลาง เทคโนโลยีนี้ถูกนำมาใช้อย่างแพร่หลายในการเขียนโค้ด การตรวจสอบข้อบกพร่อง การสร้างเอกสารทางเทคนิค และการปรับปรุงประสิทธิภาพการทำงาน อย่างไรก็ตาม ความสามารถของ

Generative AI ในการสร้างผลลัพธ์ที่หลากหลายและซับซ้อนนั้น ก็นำมาซึ่งความท้าทายด้านความมั่นคงปลอดภัยไซเบอร์ ความน่าเชื่อถือของโค้ดที่สร้างขึ้น การรั่วไหลของข้อมูลอ่อนไหว และความเสี่ยงด้านคุณภาพซอฟต์แวร์ [1]

องค์กรพัฒนาซอฟต์แวร์ขนาดกลางในประเทศไทยกำลังเผชิญกับความท้าทายในการนำ Generative AI มาใช้อย่างมีประสิทธิภาพและปลอดภัย โดยทีมนักพัฒนาอาจมีขนาด 30-45 คนแบ่งออกเป็นหลายทีมย่อย มีการใช้งาน Generative AI ในรูปแบบที่แตกต่างกัน เช่น GitHub Copilot ช่วยเขียนโค้ด ChatGPT และ Claude สำหรับการแก้ไขปัญหาและวิเคราะห์โค้ด ความหลากหลายในการใช้งานนี้ทำให้เกิดความเสี่ยงที่ซับซ้อนและยากต่อการควบคุม [2]

National Institute of Standards and Technology (NIST) ได้พัฒนาและเผยแพร่ AI Risk Management Framework (AI RMF 1.0) ในปี 2023 ซึ่งเป็นกรอบการทำงานที่ครอบคลุม 4 ฟังก์ชันหลัก ได้แก่ Govern, Map, Measure และ Manage [3] งานวิจัยนี้จึงมีวัตถุประสงค์เพื่อ (1) พัฒนาแนวทางการประยุกต์ใช้ NIST AI RMF ที่เป็นรูปธรรมสำหรับทีมนักพัฒนาในองค์กรเทคโนโลยี (2) ทดสอบแนวทางที่พัฒนาขึ้นผ่านการประยุกต์ใช้จริง และ (3) นำผลที่ได้มาปรับปรุงแนวปฏิบัติที่ดีในการใช้ AI ในองค์กร

## 2. วรรณกรรมที่เกี่ยวข้อง

### 2.1 ความเสี่ยงเฉพาะของ Generative AI

NIST ได้ชี้ให้เห็นว่า AI มีความเสี่ยงเฉพาะที่แตกต่างจากซอฟต์แวร์ทั่วไป ได้แก่ ความเสี่ยงด้านข้อมูลที่อาจมีอคติที่เป็นอันตราย (Harmful Bias) ความเสี่ยงด้านความซับซ้อนและความโปร่งใส ความเสี่ยงด้านความปลอดภัยจากพื้นผิวการโจมตีที่ซับซ้อนขึ้น เช่น Data Poisoning และความเสี่ยงด้านความเป็นส่วนตัว [3] Zhang et al. [4] พบว่าความเสี่ยงใน AI/ML จำแนกได้เป็น 4 มิติหลัก ได้แก่ Data Risk, Model Risk, Process Risk และ Contextual Risk ส่วน Wessel et al. [5] พบว่าโค้ดที่สร้างจาก LLM มีความซับซ้อนในการจัดการช่องโหว่มากกว่าโค้ดที่เขียนโดยมนุษย์

### 2.2 กรอบการบริหารความเสี่ยง NIST AI RMF 1.0

NIST AI RMF 1.0 มีเป้าหมายส่งเสริมการพัฒนา AI ที่น่าเชื่อถือ (Trustworthy AI) ซึ่งประกอบด้วยคุณลักษณะ 7 ประการ ได้แก่ Valid & Reliable, Safe, Secure & Resilient, Accountable & Transparent, Explainable & Interpretable, Privacy-Enhanced และ Fair with Harmful Bias Managed [3] องค์ประกอบหลักประกอบด้วย 4 ฟังก์ชัน ดังนี้

**Govern** (กำกับดูแล) คือการกำหนดนโยบาย บทบาทหน้าที่ และกระบวนการกำกับดูแลการใช้ AI ในองค์กร เช่น การจัดทำนโยบาย Acceptable Use Policy และการกำหนด AI System Owner Map (ระบุและทำความเข้าใจบริบท) คือการสำรวจและทำความเข้าใจบริบทการใช้งาน AI รวมถึงการระบุผู้มีส่วนได้ส่วนเสีย ข้อจำกัดทางเทคนิค และข้อกฎหมายที่เกี่ยวข้อง เช่น PDPA และลิขสิทธิ์ **Measure** (วัดผลและประเมิน) คือการวัดและติดตามความเสี่ยงของระบบ AI อย่างต่อเนื่อง เช่น การทดสอบช่องโหว่ Prompt Injection และ Model Drift Detection Manage (จัดการและรับมือ) คือการดำเนินการตอบสนองและบรรเทาความเสี่ยงที่ตรวจพบ เช่น การจัดทำแผนฉุกเฉิน การจัดการ Secrets และการแยกสภาพแวดล้อม [3]

Habbal et al. [6] นำเสนอกรอบ AI TRiSM ที่ครอบคลุมวงจรชีวิตทั้งหมดของระบบ AI ส่วน Haber et al. [7] ชี้ให้เห็นว่าเครื่องมือ AI ซ่อนความเสี่ยงด้าน Cybersecurity ไว้มากกว่าที่ผู้ใช้รับรู้

### 2.3 ช่องว่างในการวิจัย

จากการทบทวนวรรณกรรม พบว่ายังขาดการศึกษาที่เจาะจงถึงการประยุกต์ใช้ NIST AI RMF กับความเสี่ยงเฉพาะของ Generative AI ในกระบวนการพัฒนาซอฟต์แวร์ โดยเฉพาะปัญหา Hallucination และ Data Provenance [8] นอกจากนี้ Guembe et al. [9] พบว่า Ethics, Governance และ Adversarial Robustness ยังขาดการศึกษาเชิงประจักษ์ งานวิจัยนี้จึงมุ่งเติมเต็มช่องว่างดังกล่าว โดยพัฒนาแนวทางการประยุกต์ใช้ NIST AI RMF ร่วมกับ NCSA AI Security Guidelines และทดสอบผ่านกรณีศึกษาจริง

## 3. วิธีการดำเนินการวิจัย

การวิจัยนี้เป็นการวิจัยเชิงประยุกต์ (Applied Research) ลักษณะเป็นกรณีศึกษา (Case Study) ในองค์กรพัฒนาซอฟต์แวร์ขนาดกลางแห่งหนึ่งในประเทศไทย ซึ่งเป็นบริษัทเทคโนโลยีที่มีแผนกพัฒนาซอฟต์แวร์และทีมเทคนิคประมาณ 30-40 คน มีโครงสร้างการบริหารที่มีตำแหน่ง CTO (Chief Technology Officer) รับผิดชอบดูแลทิศทางเทคโนโลยีขององค์กร แบ่งทีมย่อยออกเป็น Front-end, Back-end, SRE และ QA โดยกลุ่มตัวอย่างทั้ง 37 คน และ CTO เป็นบุคลากรขององค์กรเดียวกัน ประกอบด้วย Developer/Software Engineer (ร้อยละ 54.1), System Analyst/Business Analyst (ร้อยละ 16.2), QA/Tester (ร้อยละ 13.5), Data Engineer (ร้อยละ 8.1) และ Product Owner/อื่น ๆ (ร้อยละ 8.1) แบ่งการดำเนินงานเป็น 4 ขั้นตอน ดังนี้

### 3.1 การศึกษาและพัฒนากรอบแนวคิด

ศึกษา NIST AI RMF 1.0 อย่างละเอียดเพื่อทำความเข้าใจองค์ประกอบหลัก 4 ประการ และคุณลักษณะ Trustworthy AI ทั้ง 7 ประการ จากนั้นพัฒนาเครื่องมือวิจัย 2 ชุด ดังนี้

**ชุดที่ 1:** "แบบประเมินระดับวุฒิภาวะด้านการบริหารความเสี่ยง AI" (Gap Analysis Tool) จำนวน 72 ข้อ ใช้สเกล 0-4 โดย 0 = Non-existent (ไม่มีการดำเนินการ), 1 = Initial/Ad-hoc (เริ่มต้น ไม่เป็นทางการ), 2 = Defined (มีมาตรฐานกำหนด), 3 = Managed (มีการบริหารจัดการ) และ 4 = Optimized (ปรับปรุงอย่างต่อเนื่อง) สำหรับให้ผู้บริหาร (CTO) ประเมินทั้งสถานะปัจจุบัน (As-Is) และเป้าหมาย (To-Be) ขององค์กรในมุมมองเชิงกลยุทธ์ แบ่งตาม 4 ฟังก์ชัน ได้แก่ Govern 19 ข้อ (นโยบายและกำกับดูแล), Map 18 ข้อ (ระบุบริบทและผู้มีส่วนได้ส่วนเสีย), Measure 22 ข้อ (วัดผลและติดตามความเสี่ยง) และ Manage 13 ข้อ (จัดการและตอบสนองความเสี่ยง)

**ชุดที่ 2:** "Generative AI Risk Assessment Tool" จำนวน 24 ข้อ (ดูตารางที่ 2) ใช้สเกล 1-5 (น้อยที่สุด ถึงมากที่สุด) สำหรับสำรวจผู้ปฏิบัติงาน 37 คน เพื่อวัดระดับการปฏิบัติจริงในมุมมองเชิงปฏิบัติการ การพัฒนาเครื่องมือชุดที่ 2 นี้ใช้หลักการบูรณาการ (Integration) ระหว่าง 2 มาตรฐาน กล่าวคือ NIST AI RMF 1.0 ให้โครงสร้างกรอบการประเมินตาม 4 ฟังก์ชัน (Govern, Map, Measure, Manage) เป็นแกนหลัก ในขณะที่ "แนวทางปฏิบัติการใช้ปัญญาประดิษฐ์อย่างมั่นคงปลอดภัย" (AI Security Guidelines) ของสำนักงานคณะกรรมการการรักษาความมั่นคงปลอดภัยไซเบอร์แห่งชาติ (สกมช./NCSA) ให้เนื้อหาคำถามเชิงปฏิบัติที่เหมาะสมกับบริบทองค์กรไทย เครื่องมือทั้งสองชุดนี้ทำงานเสริมกัน โดยชุดที่ 1 (72 ข้อ) วัดมุมมองผู้บริหารเชิงกลยุทธ์ และชุดที่ 2 (24 ข้อ ในตารางที่ 2) วัดมุมมองผู้ปฏิบัติงานเชิงปฏิบัติการ เพื่อนำมาวิเคราะห์ Perception Gap ระหว่างสองมุมมอง

### 3.2 การพัฒนาแนวทางการประยุกต์ใช้

นำผลลัพธ์จากขั้นตอนแรกมาออกแบบแนวทางปฏิบัติ โดยจัดทำแบบสอบถาม Generative AI Risk Assessment Tool (24 ข้อ) ผ่านระบบ Google Forms ครอบคลุม 4 ฟังก์ชัน NIST AI RMF แบ่งเป็น 4 ส่วน ได้แก่ Respondent Profile (ข้อมูลผู้ตอบ) เก็บข้อมูลพื้นฐาน เช่น บทบาท ประสบการณ์ และเครื่องมือ AI ที่ใช้งาน เพื่อวิเคราะห์ความแตกต่างระหว่างกลุ่ม Operational Practices (การปฏิบัติงาน) ประเมินพฤติกรรมการใช้ AI จริงในกระบวนการทำงาน ว่าใช้ทำอะไร บ่อยแค่ไหน และมีกระบวนการตรวจสอบผลลัพธ์หรือไม่ Risk Self-Assessment (การประเมินความเสี่ยงด้วยตนเอง) ให้ผู้ตอบประเมินระดับความเสี่ยงที่รับรู้ในแต่ละด้านตามกรอบ NIST AI RMF โดยใช้สเกล 1-5 และ Organizational Needs (ความต้องการขององค์กร) รวบรวมความต้องการเชิงคุณภาพ เช่น การฝึกอบรม นโยบายที่ยังขาด และข้อเสนอแนะในการปรับปรุง

### 3.3 การทดลองใช้งานจริงและการประเมินผล

นำเครื่องมือไปทดลองกับบุคลากร 37 คน ในช่วงวันที่ 15-20 กุมภาพันธ์ 2569 ได้รับการตอบกลับร้อยละ 100 จากนั้นวิเคราะห์เปรียบเทียบกับสถานะก่อนการวิจัย สัมภาษณ์เชิงลึก และจัดทำแนวปฏิบัติที่ดี

## 4. ผลการดำเนินงาน

### 4.1 ข้อมูลกลุ่มตัวอย่าง

กลุ่มตัวอย่าง 37 คน ส่วนใหญ่เป็น Developer/Software Engineer (ร้อยละ 54.1) มีประสบการณ์ IT มากกว่า 5 ปี (ร้อยละ 51.4) ร้อยละ 94.6 ใช้งาน Generative AI อย่างน้อย 3 ครั้งต่อสัปดาห์ โดยร้อยละ 67.6 ใช้งานทุกวัน ซึ่งยืนยันว่า Generative AI ได้กลายเป็นส่วนหนึ่งของกระบวนการทำงานประจำวัน

**ตารางที่ 1:** ผลการประเมินระดับวุฒิภาวะรายข้อจากแบบสอบถามผู้ปฏิบัติงาน (n=37)

| Q | ประเด็นการประเมิน | NIST | ค่าเฉลี่ย |
|---|---|---|---|
| Q1 | นโยบายการใช้งาน GenAI ที่ยอมรับได้ | Govern | 2.22 |
| Q2 | การประเมินระดับความเสี่ยงก่อนเริ่มโปรเจกต์ | Govern | 2.03 |
| Q3 | การกำหนดผู้รับผิดชอบ AI System Owner | Govern | **1.73** |
| Q4 | การตรวจสอบข้อมูล/ป้องกัน Data Poisoning | Measure | **1.78** |
| Q5 | การลบ Chat History เมื่อใช้ Public AI | Manage | **1.73** |
| Q6 | การควบคุมสิทธิ์การเข้าถึง AI System | Govern | 2.97 |
| Q7 | การรู้จักและป้องกัน Prompt Injection | Measure | **1.68** |
| Q8 | การทดสอบแบบจำลองการโจมตี (Red Teaming) | Measure | 2.11 |
| Q9 | ความปลอดภัยของ Third-party Library/Plugin | Measure | 2.70 |
| Q10 | การตรวจจับ Model Drift และแจ้งเตือน | Measure | **1.73** |
| Q11 | แผนฉุกเฉินเมื่อ AI ถูกโจมตี/ข้อมูลรั่วไหล | Manage | 2.76 |
| Q12 | การบันทึก Log สำหรับ Audit Trail | Manage | 2.62 |
| Q13 | การแยกสภาพแวดล้อม Dev/Staging/Prod | Manage | 2.86 |
| Q14 | การจัดการ Secrets (API Key) | Manage | 2.49 |
| Q15 | การตรวจสอบผลลัพธ์จาก AI ก่อนแสดงผล | Measure | 2.32 |
| Q16 | Version Control สำหรับ Prompt/Config | Manage | 2.54 |
| Q17 | การตรวจสอบข้อตกลง Third-party AI | Map | 2.62 |
| Q18 | การทำทะเบียนคุมเครื่องมือ AI ที่ใช้ | Govern | 2.65 |
| Q19 | การประเมินความน่าเชื่อถือ Plugin/Extension | Measure | **1.81** |
| Q20 | ความมั่นใจด้าน PDPA Compliance | Govern | **3.27** |
| Q21 | การตรวจสอบลิขสิทธิ์ผลงานจาก AI | Map | **3.11** |
| Q22 | การจำลองสถานการณ์ Model Inversion Attack | Measure | **1.68** |
| Q23 | การจำกัดการใช้งาน Rate Limiting | Manage | **1.43** |
| Q24 | การอบรมทีมเรื่อง Social Engineering ผ่าน AI | Govern | **3.41** |

**หมายเหตุ:** พื้นแดง = วิกฤต (ค่าเฉลี่ย < 2.00), พื้นเขียว = จุดแข็ง (ค่าเฉลี่ย > 3.00), สเกล 1-5

### 4.2 ผลการประเมินระดับวุฒิภาวะ

จากการประเมินโดย CTO พบว่าองค์กรมีระดับวุฒิภาวะเฉลี่ย 1.56 จาก 4.0 (ระดับ Ad-hoc) เมื่อพิจารณารายฟังก์ชัน Govern มีคะแนนสูงสุด (2.16) รองลงมา Manage (1.54) Measure (1.32) และ MAP (1.22) ตามลำดับ ช่องว่างจากเป้าหมายเฉลี่ย 0.70 โดยช่องว่างใหญ่ที่สุดพบในฟังก์ชัน Manage (0.77)

**ตารางที่ 2:** ผลการประเมินระดับวุฒิภาวะรายฟังก์ชัน NIST AI RMF

| ฟังก์ชัน | As-Is | To-Be | Gap |
|---|---|---|---|
| Govern | 2.16 | 2.84 | 0.68 |
| Map | 1.22 | 1.94 | 0.72 |
| Measure | 1.32 | 1.95 | 0.64 |
| Manage | 1.54 | 2.31 | 0.77 |
| **ค่าเฉลี่ยรวม** | **1.56** | **2.26** | **0.70** |

### 4.3 ผลการวิเคราะห์ช่องว่าง

จากแบบสอบถาม 24 ข้อ ระดับวุฒิภาวะโดยรวมจากมุมมองผู้ปฏิบัติงานอยู่ที่ 2.34 จาก 5.0 โดยร้อยละ 59.0 อยู่ในระดับ Non-existent ถึง Ad-hoc ประเด็นที่มีความเสี่ยงวิกฤต ได้แก่ Rate Limiting (1.43) Prompt Injection Defense (1.68) Model Inversion Attack Testing (1.68) และ AI System Owner Accountability (1.73) ในขณะที่จุดแข็งขององค์กร ได้แก่ Social Engineering Training (3.41) PDPA Compliance (3.27) และ IP Review (3.11)

การเปรียบเทียบ Perception Gap ระหว่างมุมมองของ CTO (จากเครื่องมือ 72 ข้อ) กับผู้ปฏิบัติงาน (จากแบบสอบถาม 24 ข้อ) พบ 3 ลักษณะ ดังนี้ (1) ช่องว่างเชิงนโยบาย (Policy Gap) องค์กรมีเป้าหมายด้านการกำกับดูแล AI แต่ขาดการนำไปปฏิบัติที่เป็นรูปธรรม ตัวอย่างเช่น CTO ประเมินเรื่อง AI System Owner ไว้ที่ระดับ 2.00 แต่ผู้ปฏิบัติงานให้คะแนนเพียง 1.73 สะท้อนว่ายังไม่มีการกำหนดผู้รับผิดชอบชัดเจน และนโยบาย Acceptable Use Policy ยังไม่ถูกสื่อสารลงสู่ระดับปฏิบัติ (2) ช่องว่างเชิงเทคนิค (Technical Gap) มาตรการความปลอดภัยเชิงลึกสำหรับ AI โดยเฉพาะภัยคุกคามรูปแบบใหม่ที่เฉพาะเจาะจงกับ LLM ยังอยู่ในระดับต่ำมาก ได้แก่ Rate Limiting (1.43) Prompt Injection Defense (1.68) และ Model Inversion Attack Testing (1.68) เนื่องจากการฝึกอบรมด้าน Cybersecurity ที่มีอยู่ยังไม่ครอบคลุมภัย

คุกคามจาก AI รุ่นใหม่ (3) ช่องว่างเชิงการรับรู้ (Perception Gap) ผู้ปฏิบัติงานประเมินระดับการปฏิบัติงานสูงกว่าที่ CTO ประเมินในหลายด้าน โดยเฉพาะฟังก์ชัน MAP ที่แตกต่างถึง +1.43 คะแนน เช่น การแยกสภาพแวดล้อม Dev/Staging/Prod ผู้ปฏิบัติงานให้คะแนน 2.86 แต่ CTO ประเมินไว้ที่ 0.00 และ Prompt Injection Defense ที่ CTO ให้ 1.00 แต่ผู้ปฏิบัติงานให้ 1.68 สะท้อนว่านิยามของ "การจัดการความเสี่ยง AI" ระหว่างผู้บริหารกับผู้ปฏิบัติงานยังไม่ตรงกัน

**4.4 ผลการทดลองเครื่องมือจัดการความเสี่ยง**

Generative AI Risk Assessment Tool ที่พัฒนาขึ้น ประกอบด้วย 4 ส่วนหลักสอดคล้องกับ 4 ฟังก์ชันของ NIST AI RMF ผลการทดลองใช้งานจริงพบว่าเครื่องมือสามารถระบุช่องโหว่ Prompt Injection ได้ 3 จุดที่ไม่เคยถูกระบุมาก่อน สำรวจเครื่องมือ AI ที่ใช้งานในองค์กรได้ครบถ้วนเป็นครั้งแรก (รวม 8 ประเภท) และช่วยเร่งการร่างนโยบาย Acceptable Use Policy จากประมาณการเดิม 2-3 สัปดาห์เหลือเพียง 1 วัน ผ่านเกณฑ์ IOC = 0.83

## 5. สรุป

การวิจัยครั้งนี้แสดงให้เห็นว่าองค์กรเทคโนโลยีขนาดกลางในประเทศไทยมีการใช้งาน Generative AI อย่างแพร่หลาย (ร้อยละ 94.6 ใช้อย่างน้อย 3 ครั้งต่อสัปดาห์) แต่ระดับวุฒิภาวะด้านการบริหารความเสี่ยงยังอยู่ในระดับเริ่มต้น (1.56 จาก 4.0) สะท้อนว่าการนำ AI มาใช้งานเติบโตนำหน้ากระบวนการควบคุมและบริหารความเสี่ยง ช่องว่างสำคัญที่พบคือความแตกต่างระหว่างมุมมองของผู้บริหารและผู้ปฏิบัติงาน โดยเฉพาะในฟังก์ชัน MAP ที่แตกต่างถึง +1.43 คะแนน

การประยุกต์ใช้ NIST AI RMF 1.0 ร่วมกับ NCSA AI Security Guidelines ในรูปแบบของเครื่องมือประเมินที่พัฒนาขึ้น พิสูจน์แล้วว่าสามารถช่วยให้องค์กรระบุความเสี่ยงที่ซ่อนอยู่ได้อย่างเป็นระบบ อย่างไรก็ตาม กรอบ NIST AI RMF 1.0 ซึ่งออกแบบมาสำหรับองค์กรขนาดใหญ่ จำเป็นต้องได้รับการปรับแต่ง (Tailoring) ให้เหมาะสมกับบริบทองค์กรขนาดกลาง และเทคโนโลยี AI ที่เปลี่ยนแปลงอย่างรวดเร็ว โดยเฉพาะ Agentic AI ข้อเสนอแนะสำหรับการวิจัยในอนาคต ได้แก่ การวิจัยระยะยาว 6-12 เดือน การขยายขอบเขตไปยังองค์กรหลายประเภท และการพัฒนาเครื่องมือประเมินอัตโนมัติ

## เอกสารอ้างอิง


[1] R. Mohawesh, M. A. Ottom, and H. B. Salameh, "A data-driven risk assessment of cybersecurity challenges posed by generative AI," Decision Analytics Journal, Vol. 15, 100580, 2025.

[2] A. Diro, S. Kaisar, A. Saini, S. Fatima, P. C. Hiep, and F. Erba, "Workplace security and privacy implications in the GenAI age: A survey," Journal of Information Security and Applications, Vol. 89, 103960, 2025.

[3] National Institute of Standards and Technology (NIST), "Artificial Intelligence Risk Management Framework (AI RMF 1.0)," NIST AI 100-1, 2023.

[4] X. Zhang, F. T. S. Chan, C. Yan, and I. Bose, "Towards risk-aware artificial intelligence and machine learning systems: An overview," Decision Support Systems, Vol. 159, 113800, 2022.

[5] S. Wessel, C. Hild, M. Walter, and M. Seidl, "Vulnerability Handling of AI-Generated Code: Existing Solutions and Open Challenges," 2024 IEEE 35th ISSREW, 2024.

[6] A. Habbal, M. K. Ali, and M. A. Abuzaraida, "Artificial Intelligence Trust, Risk and Security Management (AI TRiSM)," Expert Systems with Applications, Vol. 240, 122442, 2024.

[7] M. Haber, S. Bohm, and J. Haider, "AI System Risk Management Methodology Based on Generalized Blueprints," 2024 IEEE EuroS&PW, 2024.

[8] NIST, "Generative AI Profile (AI 600-1)," 2024.

[9] B. Guembe, S. Hamrouni, and N. Zincir-Heywood, "Cybersecurity in the age of generative AI," Future Generation Computer Systems, 2025.

[10] H. Karim, D. Gupta, and S. Sitharaman, "Securing LLM Workloads With NIST AI RMF in the Internet of Robotic Things," IEEE Access, Vol. 13, pp. 69631-69649, 2025.

[11] T. R. Mcintosh et al., "From COBIT to ISO 42001: Evaluating cybersecurity frameworks for opportunities, risks, and regulatory compliance in commercializing large language models," Computers & Security, Vol. 144, 103964, 2024.

[12] V. Almeida, L. S. Mendes, and D. Doneda, "On the Development of AI Governance Frameworks," IEEE Internet Computing, Vol. 27, No. 1, pp. 70-74, 2023.

# กรอบงานสนามยุทธ์ไซเบอร์แบบตอบโต้สองเฟสบนสถาปัตยกรรมคอนเทนเนอร์ขนาดเบา โดยใช้ Moving Target Defense และระบบลวงอัตโนมัติ

# A Two-Phase Reactive Cyber Range Framework with Moving Target Defense and Automated Deception on Lightweight Containers

*อนาวิล อุทัยสาง (Anawin Uthaisang), ณันสพัทธ์ เติมทรัพย์ (Nansaphat Termsub)*

*และจักรชัย โสอินทร์ (Chakchai So-In)*

*วิทยาลัยการคอมพิวเตอร์ มหาวิทยาลัยขอนแก่น*

*Anawin.u@kkumail.com, Nansaphat.t@kkumail.com, chakso@kku.ac.th*

**บทคัดย่อ**

*ปัจจุบันระบบสนามยุทธ์ไซเบอร์สำหรับการฝึกอบรมและประเมินทักษะด้านความมั่นคงปลอดภัย มักเผชิญปัญหาการใช้ทรัพยากรระบบสูงจากการพึ่งพาเครื่องจำลองเสมือน รูปแบบสถานการณ์จำลองที่ตายตัวซึ่งเสี่ยงต่อการทุจริต และการขาดกลไกประเมินผลแบบทันที งานวิจัยนี้นำเสนอกรอบงานสนามยุทธ์ไซเบอร์แบบตอบโต้แบบสองเฟสร่วมกับเทคนิคการป้องกันเป้าหมายเคลื่อนที่และระบบลวงอัตโนมัติบนสถาปัตยกรรมคอนเทนเนอร์ขนาดเบา การทำงานแบ่งเป็นสองส่วนหลัก ได้แก่ เฟสการสร้างโจทย์ออฟไลน์ที่ประยุกต์ใช้ปัญญาประดิษฐ์สร้างคลังโจทย์และกฎการตรวจจับล่วงหน้าเพื่อลดภาระประมวลผล และเฟสกลไกตอบสนองออนไลน์ซึ่งทำหน้าที่วิเคราะห์บันทึกระบบแบบเรียลไทม์ หากตรวจพบรูปแบบการโจมตีระบบจะทำการให้คะแนนอัตโนมัติพร้อมปรับเปลี่ยนเป้าหมายด้วยการย้ายพอร์ตและสลับคอนเทนเนอร์อย่างฉับพลันทันที ผลการทดลองพบว่าระบบสามารถย้ายพอร์ตและสลับคอนเทนเนอร์เป้าหมายเพื่อหลบหลีกการโจมตีได้ภายในเวลาเฉลี่ย 213.5 มิลลิวินาที ซึ่งลดความหน่วงได้ถึงร้อยละ 97.8 เมื่อเทียบกับกลไกปิดการทำงานแบบดั้งเดิม นอกจากนี้ยังสามารถลดความต่อเนื่องของการโจมตีจากเครื่องมืออัตโนมัติได้อย่างมีนัยสำคัญ และใช้หน่วยความจำเพียง 120 ถึง 150 เมกะไบต์ต่อคอนเทนเนอร์ สถาปัตยกรรมที่นำเสนอนี้ช่วยเพิ่มระดับความซับซ้อนของสถานการณ์ฝึก ป้องกันการทุจริต และให้ผลลัพธ์ย้อนกลับแก่ผู้เรียนได้อย่างมีประสิทธิภาพ*

**คำสำคัญ** สนามยุทธ์ไซเบอร์ การป้องกันเป้าหมายเคลื่อนที่ สถาปัตยกรรมสองเฟส คอนเทนเนอร์ขนาดเบา การประเมินผลแบบเรียลไทม์

**Abstract**

*Current cyber ranges frequently face challenges such as high resource consumption due to their reliance on heavyweight virtual machines, static scenarios that are prone to cheating, and a lack of real-time assessment. This research proposes a two-phase reactive cyber range framework integrated with moving target defense techniques and automated deception on a lightweight container architecture. The system operates in two main phases. First, the offline generation phase utilizes generative artificial intelligence to pre-build a scenario library and detection rules to minimize runtime overhead. Second, the online reactive engine performs real-time log parsing. Upon detecting an attack pattern, the system executes automated event-driven scoring and instantly changes the exposed target through port migration and container replacement to evade the attacker. Experimental results demonstrate that the system can dynamically mutate its infrastructure with an average latency of 213.5 milliseconds, reducing latency by 97.8% compared to traditional container*

*shutdown mechanisms. Furthermore, it significantly reduces continuous attacks from automated tools and consumes only 120 to 150 megabytes of memory per container. This proposed architecture effectively enhances training challenges, mitigates cheating risks, and delivers instant feedback to learners.*

**Keywords:** Cyber Range, Moving Target Defense, Two-Phase Architecture, Lightweight Container, Real-Time Assessment.

## 1. บทนำ

ปัจจุบันสนามยุทธ์ไซเบอร์ (Cyber Range) มีบทบาทสำคัญอย่างยิ่งในการเป็นแพลตฟอร์มจำลองสภาพแวดล้อมเพื่อฝึกอบรมบุคลากรด้านความมั่นคงปลอดภัย อย่างไรก็ตาม ระบบในปัจจุบันยังคงเผชิญข้อจำกัดหลัก 3 ประการ ได้แก่ (1) การใช้ทรัพยากรสูงเนื่องจากการพึ่งพาเครื่องจำลองเสมือน (Virtual Machine) ที่มีขนาดใหญ่และใช้เวลาเตรียมนาน (2) สถานการณ์จำลองแบบตายตัว (Static Scenario) ซึ่งไม่สะท้อนพฤติกรรมการป้องกันในโลกจริงและนำไปสู่ปัญหาการทุจริตแชร์คำตอบระหว่างผู้เรียน และ (3) ความล่าช้าในการประเมินผล ที่มักอาศัยการวิเคราะห์บันทึกย้อนหลัง (Post-mortem Analysis) ทำให้ผู้เรียนไม่ได้รับผลย้อนกลับอย่างทันท่วงที

แม้ที่ผ่านมาจะมีการพยายามนำเทคโนโลยีเครื่องจำลองเสมือนหรือคอนเทนเนอร์มาประยุกต์ใช้เพื่อจัดการทรัพยากร [1-2] แต่ระบบส่วนใหญ่ยังคงขาดคุณลักษณะเชิงพลวัต ในขณะเดียวกัน งานวิจัยที่นำเทคโนโลยีปัญญาประดิษฐ์ (AI) เข้ามาช่วยสร้างโจทย์และสถานการณ์จำลอง [7] กลับพบข้อจำกัดสำคัญคือ การเรียกใช้โมเดล AI ระหว่างการทำงานจริง (Runtime Inference) กินทรัพยากรเครื่องอย่างหนักและก่อให้เกิดความหน่วง (Latency) ซึ่งไม่ตอบโจทย์การใช้งานจริง จากช่องว่างของงานวิจัย (Research Gap) ดังกล่าว

งานวิจัยนี้จึงนำเสนอกรอบงานสนามยุทธ์ไซเบอร์แบบตอบโต้สองเฟส (2-Phase Reactive Cyber Range) บนสถาปัตยกรรมคอนเทนเนอร์ขนาดเบา เพื่อเปลี่ยนสภาพแวดล้อมฝึกแบบคงที่ให้เป็นระบบฝึกที่สามารถตอบสนองเชิงโต้ตอบได้ (Active Defense) โดยแบ่งการทำงานเป็น เฟสออฟไลน์ ที่ใช้ AI สกัดกฎและสร้างคลังโจทย์เตรียมไว้ล่วงหน้าเพื่อหลีกเลี่ยงความหน่วง และ เฟสออนไลน์ ที่ใช้โปรแกรมควบคุมคอนเทนเนอร์แบบเรียลไทม์โดยไม่ต้องเรียกใช้ AI ระหว่างการทดสอบ

ในเฟสออนไลน์ ระบบจะวิเคราะห์พฤติกรรมจากบันทึกระบบ หากตรวจพบการโจมตี ระบบจะประเมินผลให้คะแนนอัตโนมัติในระดับมิลลิวินาที พร้อมทั้งใช้กลไกการป้องกันเป้าหมายเคลื่อนที่ (Moving Target Defense) ปรับเปลี่ยนโครงสร้างเป้าหมายหลบหลีกทันที นอกจากนี้ยังผสานระบบลวงอัตโนมัติเพื่อดักจับการสแกนแบบสุ่มและกลไกจำกัดเพดานคะแนนเพื่อป้องกันการใช้เครื่องมืออัตโนมัติทุจริต แนวทางนี้นอกจากจะช่วยลดการใช้ทรัพยากรแล้ว ยังสร้างโจทย์แบบพลวัตที่ป้องกันการลอกเลียนแบบ และทำให้สภาพแวดล้อมฝึกมีลักษณะใกล้เคียงกับบริบทของภัยคุกคามจริงมากขึ้นในยุคปัจจุบัน

1. **นำเสนอสถาปัตยกรรมสนามยุทธ์ไซเบอร์แบบ 2 เฟส:** โดยแยกการทำงานส่วน Offline Generation และ Online Reactive ออกจากกัน เพื่อขจัดความหน่วงจากการประมวลผลของ AI แบบเรียลไทม์
2. **พัฒนากลไกการป้องกันเป้าหมายเคลื่อนที่ (MTD) บนสถาปัตยกรรมขนาดเบา:** ซึ่งทำงานร่วมกับระบบลวงอัตโนมัติ (Automated Deception) แบบขับเคลื่อนด้วยเหตุการณ์ (Event-driven) เพื่อตอบสนองและหลบหลีกภัยคุกคามได้ในระดับเสี้ยววินาที
3. **นำเสนอโมเดลสมการคณิตศาสตร์สำหรับการประเมินผลแบบเรียลไทม์:** บูรณาการร่วมกับกลไกจำกัดเพดานคะแนน เพื่อป้องกันการทุจริตจากเครื่องมือสแกนอัตโนมัติและสามารถให้ผลลัพธ์ย้อนกลับแก่ผู้เรียนได้ทันที

## 2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

### 2.1 สนามยุทธ์ไซเบอร์

สนามยุทธ์ไซเบอร์เป็นแพลตฟอร์มที่ใช้จำลองสภาพแวดล้อมด้านเครือข่าย ระบบปฏิบัติการ และบริการ

แอปพลิเคชัน เพื่อสนับสนุนการฝึกอบรม การทดสอบ และการประเมินทักษะด้านความมั่นคงปลอดภัยไซเบอร์ในสภาพแวดล้อมที่ควบคุมได้ [2, 3] ระบบสนามยุทธ์ไซเบอร์จำนวนมากในงานวิจัยและการประยุกต์ใช้งานจริงยังพัฒนาบนสถาปัตยกรรมแบบเครื่องจำลองเสมือน [2] แม้สถาปัตยกรรมดังกล่าวจะให้ระดับความสมจริงสูง แต่มีข้อจำกัดเชิงโครงสร้างที่สำคัญสามประการ ได้แก่ (1) การใช้ทรัพยากรระดับฮาร์ดแวร์ที่สูงมาก (2) ความยืดหยุ่นในการขยายขนาดต่ำ และ (3) การออกแบบสถานการณ์จำลองแบบกำหนดตายตัว (Static Scenario) ซึ่งไม่สะท้อนพฤติกรรมของภัยคุกคามที่มีลักษณะปรับตัวตามสถานการณ์ [2, 3, 5] จากข้อจำกัดดังกล่าว งานวิจัยยุคใหม่จึงพยายามมุ่งพัฒนาสนามยุทธ์ไซเบอร์ที่มีคุณลักษณะเชิงพลวัตและสามารถปรับเปลี่ยนสภาพแวดล้อมได้จริง [3, 4]

**2.2 เทคโนโลยีคอนเทนเนอร์และด็อกเกอร์**

เพื่อแก้ปัญหาการใช้ทรัพยากรของเครื่องจำลองเสมือน เทคโนโลยีการจำลองสภาพแวดล้อมระดับระบบปฏิบัติการอย่างด็อกเกอร์ (Docker) จึงถูกนำมาประยุกต์ใช้ คอนเทนเนอร์แต่ละตัวจะแยกการทำงานออกจากกันอย่างอิสระโดยใช้เคอร์เนลของระบบปฏิบัติการหลักร่วมกัน ด้วยเหตุนี้ Docker จึงมีข้อได้เปรียบด้านการใช้ทรัพยากรต่ำกว่าและใช้เวลาเริ่มต้นระบบน้อยกว่าเมื่อเทียบกับเครื่องจำลองเสมือนในหลายบริบทการใช้งาน และสามารถเริ่มต้นระบบได้ในเวลาเพียงไม่กี่วินาที คุณสมบัติความยืดหยุ่นและรวดเร็วนี้จึงทำให้เทคโนโลยีคอนเทนเนอร์มีความเหมาะสมอย่างยิ่งในการนำมาเป็นรากฐานของสนามยุทธ์ไซเบอร์ที่ต้องการรองรับผู้ใช้งานจำนวนมากพร้อมกัน [1, 4]

**2.3 การป้องกันเป้าหมายเคลื่อนที่**

แนวคิดการป้องกันเป้าหมายเคลื่อนที่เป็นกลยุทธ์การป้องกันเชิงรุกที่มุ่งลดโอกาสความสำเร็จของผู้โจมตีโดยการเปลี่ยนแปลงพื้นผิวการโจมตีอย่างต่อเนื่อง [6] กลไกดังกล่าวอาจอยู่ในระดับเครือข่าย (เช่น การสุ่มพอร์ต) ระดับโครงสร้างระบบ (เช่น การย้ายบริการ) หรือระดับแพลตฟอร์ม .ซึ่งช่วยเพิ่มต้นทุนและระยะเวลาในการสำรวจระบบใหม่ของผู้โจมตี (reconnaissance effort) อย่างไรก็ตาม แม้แนวคิดนี้จะมีประสิทธิภาพ แต่งานวิจัยที่นำ MTD มาประยุกต์ใช้ในบริบทของสนามยุทธ์ไซเบอร์เพื่อสร้างสภาพแวดล้อมฝึกแบบโต้ตอบยังมีจำนวนจำกัด และมักเผชิญปัญหาความซับซ้อนเมื่อต้องทำงานบนสถาปัตยกรรมขนาดเบา [1]

**2.4 สถาปัตยกรรมแบบไฮบริดสองเฟส**

งานวิจัยยุคปัจจุบันเริ่มนำเทคโนโลยีปัญญาประดิษฐ์เชิงรู้สร้าง (Generative AI) และโมเดลภาษาขนาดใหญ่มาใช้ในการสร้างสถานการณ์จำลองอัตโนมัติ [7] อย่างไรก็ตาม จุดอ่อนสำคัญของงานวิจัยกลุ่มนี้คือ การเรียกใช้งานโมเดลขนาดใหญ่ในช่วงเวลาประมวลผลจริงมักก่อให้เกิดความหน่วง (Latency) สูงและสิ้นเปลืองทรัพยากร สถาปัตยกรรมแบบสองเฟสจึงถูกเสนอเพื่อแก้ปัญหาดังกล่าว โดยแบ่งการทำงานออกเป็นระยะที่หนึ่งคือ “การสร้างออฟไลน์” ซึ่งใช้ AI สร้างสถานการณ์จำลองและช่วยสกัดกฎการตรวจจับล่วงหน้า [3], [5], [7] และระยะที่สองคือ"กลไกการตอบสนองออนไลน์" ซึ่งใช้ตรรกะแบบ rule-based ขับเคลื่อนตามเหตุการณ์ เพื่อลดภาระการประมวลผลในช่วงเวลาจริง ทำให้ระบบตอบสนองได้ในระดับเสี้ยววินาที

**2.5 การประเมินผลแบบเรียลไทม์**

การประเมินผลผู้เรียนในระบบสนามยุทธ์ไซเบอร์แบบดั้งเดิมมักใช้วิธีการวิเคราะห์ข้อมูลย้อนหลัง (Post-mortem Analysis) หรือใช้เทคนิคการทำเหมืองกระบวนการ (Process Mining) ดังที่ปรากฏในงานวิจัยของ Guerrero และคณะ [8] ข้อจำกัดหลักของวิธีการนี้คือผู้เรียนต้องรอให้การฝึกอบรมสิ้นสุดลงจึงจะทราบผล ทำให้ขาดความต่อเนื่องในการปรับเปลี่ยนกลยุทธ์ แนวทางการประเมินผลแบบเรียลไทม์ที่วิเคราะห์บันทึกระบบด้วยนิพจน์ทั่วไป (Regular Expression) จึงถูกนำมาใช้เพื่อให้ระบบสามารถให้คะแนนหรือลงโทษผู้เรียนได้ในระดับวินาที ซึ่งสอดคล้องกับหลักการเรียนรู้เชิงรุก (Active Learning)

**2.6 สรุปช่องว่างการวิจัย**

จากการทบทวนวรรณกรรม พบว่าระบบสนามยุทธ์ไซเบอร์ในปัจจุบันยังคงมีช่องว่างในการบูรณาการคุณลักษณะเชิงพลวัตเข้ากับการใช้ทรัพยากรที่คุ้มค่า

**ตารางที่ 2:** เปรียบเทียบคุณลักษณะของงานวิจัยที่เกี่ยวข้องกับกรอบงานที่นำเสนอ

| งานวิจัยอ้างอิง | สถาปัตยกรรมหลัก | ความเป็นพลวัต (Dynamic/MTD) | การมีระบบลวง (Deception) | รูปแบบการประเมินผล | ช่องว่าง/ข้อจำกัดหลักที่พบ |
|---|---|---|---|---|---|
| **Lates et al. [2]** | Virtual Machine | ไม่มี (โจทย์ตายตัว) | ไม่มี | ย้อนหลัง (Post-mortem) | ใช้ทรัพยากรสูง ขยายขนาดได้ยาก |
| **Mills et al. [1]** | Container-based | มี (เฉพาะระดับโมดูล) | ไม่มี | ย้อนหลัง (Post-mortem) | ขาดการตอบสนองเชิงรุกต่อผู้โจมตี |
| **Roh et al. [7]** | LLM-based | มี (สร้างโจทย์อัตโนมัติ) | ไม่มี | บางส่วน (อิงตาม AI) | ความหน่วงสูงระหว่างประมวลผลจริง |
| **Guerrero et al. [8]** | ผสมผสาน | ไม่มี | ไม่มี | ย้อนหลัง (Process Mining) | ผู้เรียนไม่ได้รับผลสะท้อนกลับทันที |
| **(Proposed)** | Container (2-Phase) | มี (MTD ย้ายพอร์ตฉับพลัน) | มี (Automated Deception) | เรียลไทม์ (event-driven scoring) | ลดผลกระทบจากความหน่วงของ AI ระหว่างการทำงานจริง และลดข้อจำกัดของระบบเป้านิ่ง |

นอกจากนี้งานวิจัยบางชิ้นมีความสมจริงแต่ใช้ทรัพยากรสูง [2] บางชิ้นมีขนาดเบาแต่ขาดการตอบสนองเชิงรุก [1] หรือบางชิ้นใช้ปัญญาประดิษฐ์สร้างความพลวัตแต่กลับมีความหน่วงสูงและประเมินผลล่าช้า [7, 8] ดังนั้นงานวิจัยฉบับนี้จึงนำเสนอการปิดช่องว่างดังกล่าว (ดังตารางที่ 2) ด้วยการพัฒนากรอบงานแบบสองเฟสที่ผสาน MTD เข้ากับคอนเทนเนอร์ขนาดเบา พร้อมระบบประเมินผลและตอบสนองแบบเรียลไทม์

## 3. วิธีการดำเนินงาน

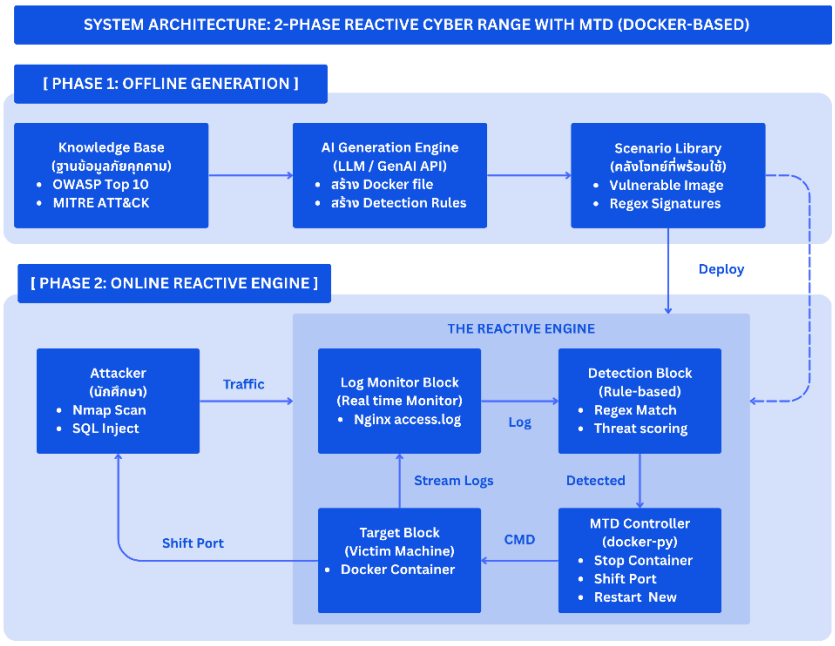
เพื่อแก้ไขปัญหาความหน่วง (Latency) ของการใช้ปัญญาประดิษฐ์ในระยะเวลาประมวลผลจริง (Runtime) และข้อจำกัดด้านทรัพยากรของสนามยุทธ์ไซเบอร์

งานวิจัยนี้จึงนำเสนอกรอบงานสนามยุทธ์ไซเบอร์แบบตอบโต้ไฮบริดสองเฟส (2-Phase Reactive Cyber Range Framework) บนสถาปัตยกรรมคอนเทนเนอร์ขนาดเบา โดยบูรณาการกลไกการป้องกันเป้าหมายเคลื่อนที่ (Moving Target Defense: MTD) ระบบลวงอัตโนมัติ (Automated Deception) และแบบจำลองการประเมินผลแบบเรียลไทม์เข้าด้วยกัน เนื้อหาในส่วนนี้จะอธิบายถึงแบบจำลองระบบ สถาปัตยกรรม การวิเคราะห์ทางคณิตศาสตร์ และอัลกอริทึมเชิงลึก

### 3.1 แบบจำลองระบบและแบบจำลองภัยคุกคาม

ก่อนการออกแบบสถาปัตยกรรม ผู้วิจัยได้กำหนดขอบเขตและสมมติฐานทางวิศวกรรมเพื่อรองรับการวิเคราะห์ประสิทธิภาพ ดังนี้

แบบจำลองภัยคุกคาม (Threat Model) สมมติให้ผู้โจมตี (Attacker) $A$ มีความสามารถในการส่งชุดคำสั่งโจมตี (Payloads) $p \in P$ เข้าสู่แอปพลิเคชันเป้าหมายผ่านพอร์ตเครือข่ายที่เปิดเผย $A$ สามารถใช้เครื่องมืออัตโนมัติ (Automated Scripts) เพื่อส่ง $p$ อย่างต่อเนื่อง แต่ไม่มีสิทธิ์เข้าถึง (Privilege) ในระดับโฮสต์ (Host OS) หรือระบบควบคุมคอนเทนเนอร์ (Orchestrator) โดยตรงเป้าหมายการป้องกัน (Defense Objective)




**ภาพที่ 3.1**: แผนภาพแสดงสถาปัตยกรรมระบบไฮบริดสองเฟส

นอกจากนี้ระบบไม่ได้มุ่งหมายที่จะสกัดกั้น Payload แรก (Single-shot exploit) แบบสมบูรณ์ เนื่องจากธรรมชาติของแอปพลิเคชันที่มีช่องโหว่ แต่ระบบมุ่งทำลายความต่อเนื่อง (Disrupt Continuity) ของ *A* โดยการเพิ่มต้นทุนในการค้นหาเป้าหมายใหม่ (Re-reconnaissance Cost) เมื่อเกิดการโจมตีขึ้น (ดังแสดงในภาพที่ 3.1)

**3.2 ภาพรวมสถาปัตยกรรมระบบไฮบริดสองเฟส**

การพัฒนาระบบแบ่งออกเป็นส่วนประกอบหลักตามสถาปัตยกรรมสองเฟส ส่วนแรกคือกระบวนการสร้างโจทย์สถานการณ์ซึ่งดำเนินการแบบออฟไลน์เพียงครั้งเดียวเพื่อเตรียมคลังโจทย์ ระบบจะรับข้อมูลนำเข้าเป็นคำสั่งภาษาธรรมชาติที่ระบุประเภทช่องโหว่และระดับความยาก จากนั้นใช้โมเดลภาษาขนาดใหญ่ทำหน้าที่สร้างซอร์สโค้ดของแอปพลิเคชันเว็บที่มีช่องโหว่พร้อมไฟล์สำหรับสร้างคอนเทนเนอร์ ควบคู่ไปกับการวิเคราะห์รูปแบบการโจมตีที่สอดคล้องกับช่องโหว่นั้นและแปลงให้อยู่ในรูปแบบนิพจน์ทั่วไปเพื่อใช้เป็นกฎในการตรวจจับ ผลลัพธ์ที่ได้จะถูกจัดเก็บในรูปแบบไฟล์กำหนดสถานการณ์และอิมเมจของคอนเทนเนอร์ภายในคลังสถานการณ์

ส่วนที่สองคือกลไกตอบโต้และประเมินผลออนไลน์ซึ่งพัฒนาด้วยภาษาไพธอนร่วมกับคำสั่งและส่วนเชื่อมต่อสำหรับควบคุม Docker เพื่อควบคุมวงจรชีวิตของคอนเทนเนอร์ กลไกนี้ประกอบด้วยโมดูลย่อย ได้แก่ โมดูลเตรียมการคอนเทนเนอร์ทำหน้าที่สร้างเป้าหมายจากอิมเมจที่เตรียมไว้พร้อมกำหนดพอร์ตเริ่มต้นแบบสุ่ม โมดูลดักจับบันทึกระบบทำหน้าที่เก็บข้อมูลการเข้าถึงเว็บแบบเรียลไทม์และส่งต่อไปยังโมดูลตรวจจับและประเมินผล

หลังจากนั้นโมดูลตรวจจับจะเปรียบเทียบเหตุการณ์ที่พบกับกฎนิพจน์ทั่วไป และเมื่อเข้าเงื่อนไขการโจมตี ระบบจึงบันทึกเหตุการณ์ คำนวณคะแนน และส่งสัญญาณไปยังกระบวนการป้องกันเชิงรุก เมื่อได้รับสัญญาณ ระบบจะดำเนินการตามมาตรการป้องกันทันทีผ่านกระบวนการทำลายคอนเทนเนอร์เดิม สร้างคอนเทนเนอร์ใหม่ และการย้ายพอร์ต นอกจากนี้ยังมีโมดูลควบคุมระบบลวงอัตโนมัติที่ทำหน้าที่สร้างคอนเทนเนอร์กับดักควบคู่ไปกับเป้าหมายจริงเพื่อตรวจจับพฤติกรรมการสแกนแบบสุ่ม หากผู้เรียนทำการโจมตีมาตกที่ระบบลวง ระบบจะบันทึกเป็นเหตุการณ์ละเมิดกฎเพื่อนำไปหักคะแนนประเมินผล

เพื่อให้ระบบสามารถประเมินทักษะของผู้เรียนได้ทันที ผู้วิจัยได้นำเสนอสมการคณิตศาสตร์สำหรับคำนวณ

คะแนนสุทธิโดยพิจารณาจากความรวดเร็วในการโจมตีสำเร็จหักล้างกับข้อผิดพลาดจากการโจมตีโดนระบบลวง ดังสมการที่ (1)

$$S_{total} = max\left(0, \sum_{i=1}^{n}\left(W_{a,i} \times min\left(2.0, \frac{T_{expected,i}}{T_{actual,i}}\right)\right) - \sum_{j=1}^{m}\left(W_{p,j} \times E_{trap,j}\right)\right) \quad (1)$$

โดยที่

$W_a$ คือ น้ำหนักคะแนนพื้นฐานของช่องโหว่เป้าหมาย

$T_{expected}$ คือ เวลามาตรฐานที่ระบบคาดหวังให้เจาะสำเร็จ

$T_{actual}$ คือ เวลาจริงที่ผู้เรียนใช้จนกระทั่งระบบดักจับได้

$W_p$ คือ น้ำหนักบทลงโทษเมื่อโจมตีโดนเครื่องกับดัก

$E_{trap}$ คือ จำนวนครั้งที่โจมตีผิดพลาดโดนระบบลวง

---

**Algorithm 1: Event-based proactive response mechanism with target container termination and port migration**

---

**Input:** Log stream $L$, Attack signature set $S_{regex}$
**Output:** Reactive port migration via pre-provisioning and SIGKILL

```
Initialize container C_current at port P_current
Start real-time log monitoring from C_current
while system is active do
l_entry = ReadNextLog (L)
for each signature s in S_regex do
if Match (l_entry, s) then
LogAttackEvent()
CalculateScore()
// Pre-provisioning phase with collision avoidance
P_new = RandomPort()
while P_new in UsedPort do
P_new = RandomPort()
end while
C_new=CreateAndStartContainer(P_new)
// Target termination phase
Execute SIGKILL on C_current
C_current = C_new
P_current = P_new
break // Exit inner loop to monitor new container
end if
end for
end while
```

---

นอกจากนี้ผู้วิจัยได้เพิ่มกลไกจำกัดเพดานคะแนนสำหรับตัวคูณ ความเร็ว $min\left(2.0, \frac{T_{expected}}{T_{actual}}\right)$ ไว้ที่อัตราสูงสุด 2.0 เท่า เพื่อป้องกันภาวะคะแนนเฟ้อในกรณีที่ผู้โจมตีประยุกต์ใช้สคริปต์อัตโนมัติในการทดสอบ

เพื่อให้เห็นภาพรวมของกระบวนการตอบโต้ภัยคุกคามแบบอัตโนมัติ ผู้วิจัยได้สรุปขั้นตอนการทำงานของกลไกการตอบสนองออนไลน์ในรูปแบบรหัสเทียมดังแสดงใน Algorithm 1 ซึ่งได้บูรณาการแนวคิดการสร้างคอนเทนเนอร์เป้าหมายเตรียมไว้ล่วงหน้าเพื่อลดความหน่วง ควบคู่กับกลไกการป้องกันการชนกันของพอร์ตเพื่อรองรับการสับเปลี่ยนโครงสร้างอย่างต่อเนื่องภายใต้สภาวะที่มีการโจมตีอย่างหนาแน่น

## 4. ผลการดำเนินงานและการอภิปรายผล

การประเมินประสิทธิภาพของกรอบงานสนามยุทธ์ไซเบอร์แบบตอบโต้สองเฟสที่นำเสนอนี้ ดำเนินการทดสอบภายใต้สภาพแวดล้อมที่ควบคุมได้ (Controlled Environment) เพื่อวัดขีดความสามารถในการสกัดกั้นภัยคุกคาม การตอบสนอง และการใช้ทรัพยากร ผลการดำเนินงานและการอภิปรายผลมีรายละเอียดดังต่อไปนี้

### 4.1 การตั้งค่าการทดลองและพารามิเตอร์

- **สภาพแวดล้อมฮาร์ดแวร์และซอฟต์แวร์**: ประมวลผลบนหน่วยประมวลผลกลาง Intel Core i5 (2.40 GHz), หน่วยความจำ 16 GB DDR4 ทำงานบนระบบปฏิบัติการ Linux Ubuntu 22.04 โดยใช้แพลตฟอร์ม Docker
- **ชุดข้อมูลทดสอบ (Test Dataset):** ใช้สคริปต์อัตโนมัติจำลองการส่งชุดคำสั่งอันตรายประเภท SQL Injection จำนวน 1,000 รายการอย่างต่อเนื่อง อ้างอิงตามมาตรฐานสากล OWASP โดยกำหนดอัตราการส่งคำสั่งโจมตีไว้ที่ 100 คำสั่งต่อวินาที และทำซ้ำทั้งหมด 10 รอบ เพื่อหาค่าเฉลี่ย การใช้ทรัพยากรระบบถูกตรวจวัดด้วย docker stats ภายใต้ช่วงเวลาเก็บตัวอย่างที่กำหนดคงที่ตลอดการทดลอง
- **สถานการณ์การทดลอง (Scenarios)**
  1) **กลุ่มควบคุม (Control Group):** ระบบสนามยุทธ์ไซเบอร์แบบดั้งเดิมที่ใช้พอร์ตตายตัวและใช้

กลไกปิดการทำงานแบบมาตรฐาน (Graceful Shutdown)

2) **กลุ่มทดลอง (Experimental Group):** ระบบที่นำเสนอ (2-Phase Reactive Cyber Range) ที่เปิดใช้งานกลไกป้องกันเป้าหมายเคลื่อนที่ (MTD) และระบบประเมินผลแบบเรียลไทม์

- **ตัวแปรหลัก (Key Parameters)** โดยมีหน่วยความจำจำกัด 128 MB ถึง 150 MB ต่อคอนเทนเนอร์ โดยใช้ช่วงการสุ่มพอร์ตเป้าหมาย 1024 ถึง 65535 และขีดจำกัดเพดานคะแนนสำหรับตัวคูณโบนัสความเร็ว $\left(\frac{T_{expected}}{T_{actual}}\right)$ สูงสุดที่ 2.0 เท่า
- **นิยามตัวชี้วัด (Metrics)** ผู้วิจัยใช้สมการทางคณิตศาสตร์ 3 ด้านในการวิเคราะห์ ได้แก่ เวลาที่ใช้ในการโจมตีสำเร็จ

$$TTC = T_{success} - T_{start} \quad (2)$$

โดยที่

$TTC$ คือ ระยะเวลาสุทธิที่ผู้โจมตีใช้ในการเจาะระบบสำเร็จ (Time-to-Compromise)

$T_{success}$ คือ ประทับเวลา (Timestamp) ที่ระบบเป้าหมายถูกเจาะช่องโหว่และคืนค่าผลลัพธ์สำเร็จ

$T_{start}$ คือ ประทับเวลาเริ่มต้นที่ผู้โจมตีทำการส่งชุดคำสั่งอันตรายเข้าสู่ระบบ

เวลาในความหน่วงของการย้ายเป้าหมาย

$$L_{MTD} = T_{ready} - T_{detect} \quad (3)$$

โดยที่

$L_{MTD}$ คือ ค่าความหน่วงรวมของการตอบสนองด้วยกระบวนการย้ายเป้าหมาย (MTD Response Latency)

$T_{ready}$ คือ ประทับเวลาที่คอนเทนเนอร์ใหม่ถูกสร้างเสร็จสมบูรณ์และพร้อมตอบสนองการเชื่อมต่อเครือข่าย

$T_{detect}$ คือ ประทับเวลาเริ่มต้นที่ระบบสามารถตรวจจับและยืนยันรูปแบบการโจมตีจากบันทึกระบบได้ และประสิทธิภาพการใช้ทรัพยากร

$$\bar{U}_{cpu} = \frac{1}{N}\sum_{i=1}^{N} U_{cpu}(t_i) \quad (4)$$

$\bar{U}_{cpu}$ คือ อัตราการใช้งานหน่วยประมวลผลกลางเฉลี่ยของระบบ (Average CPU Utilization)

$N$ คือ จำนวนรอบทั้งหมดของการดึงข้อมูลตัวอย่าง (Total Samples)

$U_{cpu}(t_i)$ คือ อัตราการใช้งานหน่วยประมวลผลกลางที่วัดได้ ณ จุดเวลา $t_i$

**4.2 ผลการประเมินความหน่วงในการตอบสนอง**

ความหน่วงในการตอบสนอง $L_{MTD}$ ถือเป็นตัวชี้วัดที่สำคัญที่สุดของสถาปัตยกรรมนี้ ผู้วิจัยได้ทำการทดสอบเพื่อวัดระยะเวลาตั้งแต่ระบบตรวจพบการโจมตีจากบันทึกระบบจนกระทั่งกระบวนการสร้างและย้ายพอร์ตเสร็จสมบูรณ์

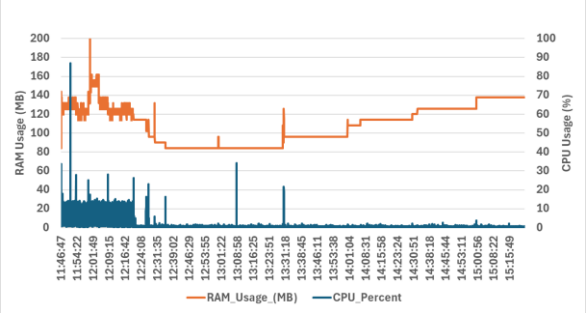
ผลการทดลอง (ดังตารางที่ 4.1) ชี้ให้เห็นว่าระบบที่นำเสนอมีความหน่วงในการตอบสนองเฉลี่ยอยู่ที่ 213.5 มิลลิวินาที โดยมีค่าสังเกตได้อยู่ในช่วงประมาณ 195 ถึง 260 มิลลิวินาที ซึ่งเร็วกว่าระบบดั้งเดิมถึงร้อยละ 97.8 สาเหตุเชิงวิศวกรรมที่ทำให้ระบบมีประสิทธิภาพด้านเวลาในการประมวลผลสูงเกิดจากสองปัจจัยหลัก: (1) การแยกการประมวลผลของ AI ไปไว้ในเฟสออฟไลน์ ทำให้กลไกออนไลน์ดักจับด้วยตรรกะ Regex ที่มีความซับซ้อนของอัลกอริทึมด้านเวลา (Time Complexity) ในระดับ $O(|S| \times m)$ (โดยที่ $|S|$ คือจำนวนกฎการตรวจจับ และ $m$ คือความยาวของข้อความบันทึกระบบ) และ (2) การประยุกต์ใช้เทคนิคเตรียมคอนเทนเนอร์ใหม่ล่วงหน้า (Pre-provisioning) ควบคู่กับการยุติการทำงานของคอนเทนเนอร์เดิมด้วยสัญญาณระดับระบบปฏิบัติการ (SIGKILL) ทันทีโดยไม่ต้องรอระยะเวลาผ่อนผัน

**ตารางที่ 4.1** เปรียบเทียบความหน่วงในการตอบสนองระหว่างกลไกปิดการทำงานแบบปกติและกลไกตอบสนองฉับพลัน

| สถาปัตยกรรมระบบ | ความหน่วง (มิลลิวินาที) |
|---|---|
| ระบบที่นำเสนอ (2-Phase MTD) | 213.5 |
| ระบบดั้งเดิม (Graceful Shutdown) | 10,000 |

**ตารางที่ 4.2** เปรียบเทียบผลลัพธ์อัตราความสำเร็จในการสกัดกั้นการโจมตีระหว่างระบบ

| ตัวชี้วัดประสิทธิภาพ | ระบบกลุ่มควบคุม | ระบบที่นำเสนอ |
|---|---|---|
| ความหน่วงในการย้ายเป้าหมาย | 10,000 มิลลิวินาที | 213.5 มิลลิวินาที |
| อัตราเจาะสำเร็จแบบคำสั่งเดียว | ร้อยละ 100 | ร้อยละ 100 |
| อัตราเจาะสำเร็จต่อเนื่อง | ร้อยละ 100 | ร้อยละ 0 |




**ภาพที่ 4.1** กราฟเส้นแสดงอัตราการใช้หน่วยประมวลผลกลางและหน่วยความจำระหว่างกระบวนการย้ายเป้าหมายอย่างต่อเนื่อง 1,000 ครั้ง

### 4.3 ผลการทดสอบอัตราความสำเร็จในการโจมตี

ผู้วิจัยได้ประเมินประสิทธิภาพการสกัดกั้นภัยคุกคามโดยแบ่งการโจมตีออกเป็น 2 รูปแบบ ได้แก่ การโจมตีแบบคำสั่งเดียว (Single-shot exploit) และการโจมตีด้วยเครื่องมืออัตโนมัติอย่างต่อเนื่อง (Continuous automated attack)

จากการวิเคราะห์ผลลัพธ์ในตารางที่ 4.2 พบว่าในการโจมตีแบบคำสั่งเดียว อัตราความสำเร็จของผู้โจมตียังคงเป็นร้อยละ 100 ปรากฏการณ์นี้สามารถอธิบายได้ด้วยหลักการ สภาวะการแข่งขัน (Race Condition) เนื่องจากแอปพลิเคชันเว็บและฐานข้อมูลสามารถประมวลผลและคืนค่าผลลัพธ์ได้ภายในเวลาเพียง 10 ถึง 20 มิลลิวินาที ซึ่งมีความรวดเร็วกว่ากลไกป้องกัน MTD ที่ต้องใช้เวลา 213.5 มิลลิวินาที ทำให้เกิดช่องโหว่ทางเวลาที่อนุญาตให้ผู้โจมตีดึงข้อมูลชุดแรกได้สำเร็จ อย่างไรก็ตาม เมื่อพิจารณาการโจมตีด้วยเครื่องมืออัตโนมัติที่ส่งคำสั่งอย่างต่อเนื่อง ระบบที่นำเสนอสามารถบังคับตัดการเชื่อมต่อและย้ายเป้าหมายได้ทันที ส่งผลให้อัตราความสำเร็จในการเจาะระบบต่อเนื่องลดลงเหลือร้อยละ 0 ซึ่งพิสูจน์ได้ว่าระบบสามารถหยุดยั้งความต่อเนื่องของการโจมตีอัตโนมัติของผู้โจมตีได้อย่างมีประสิทธิภาพ

### 4.4 ผลการประเมินประสิทธิภาพการใช้ทรัพยากร

ด้านการประเมินการใช้ทรัพยากรระบบ ผู้วิจัยได้ทำการบันทึกอัตราการใช้งานหน่วยประมวลผลกลางและหน่วยความจำของโฮสต์ระหว่างที่เกิดกระบวนการย้ายเป้าหมายอย่างต่อเนื่องจำนวน 1,000 ครั้ง ผลการทดลองชี้ให้เห็นว่าสถาปัตยกรรมคอนเทนเนอร์ขนาดเบามีการใช้หน่วยความจำเฉลี่ยเพียง 120 ถึง 150 เมกะไบต์ต่อหนึ่งคอนเทนเนอร์ และพบการเพิ่มขึ้นฉับพลัน (Spike) ของอัตราการใช้หน่วยประมวลผลกลางสูงสุดเพียงร้อยละ 5 ถึง 8 ในช่วงจังหวะที่มีการสลับพอร์ตและยุติการทำงานของคอนเทนเนอร์เดิมทิ้ง ข้อมูลดังกล่าวสนับสนุนว่าสถาปัตยกรรมนี้มีภาระงานส่วนเกินต่ำในสภาพแวดล้อมการทดลอง และมีศักยภาพต่อการขยายการใช้งานในระดับที่ใหญ่ขึ้น (ดังแสดงในภาพที่ 4.1)

### 4.5 ผลการทดสอบกลไกประเมินผลและการควบคุมคะแนน

การทดสอบโมเดลสมการคณิตศาสตร์สำหรับการประเมินผลแบบเรียลไทม์แสดงให้เห็นว่าระบบสามารถวิเคราะห์พฤติกรรมผู้เรียนและให้ผลลัพธ์ย้อนกลับได้ทันที นอกจากนี้ การบูรณาการกลไกจำกัดเพดานคะแนนเพื่อแก้ไขปัญหาภาวะคะแนนเฟ้อ (Score Inflation) ซึ่งอาจเกิดขึ้นเมื่อผู้โจมตีประยุกต์ใช้สคริปต์อัตโนมัติที่ทำงานเร็วกว่าเวลามาตรฐานที่คาดหวังไว้อย่างผิดปกติสามารถทำงานได้ตามวัตถุประสงค์ ผลการทดลองเชิงประจักษ์ยืนยันว่าระบบสามารถจำกัดตัวคูณโบนัสความเร็วไว้ที่ระดับสูงสุด 2.0 เท่าได้อย่างแม่นยำ ส่งผลให้คะแนนสุทธิถูกควบคุมให้อยู่ในกรอบที่กำหนด การควบคุมนี้นำไปสู่การจัดลำดับคะแนนที่เป็นมาตรฐาน มีความยุติธรรม และป้องกันความเบี่ยงเบนจากการใช้โปรแกรมอัตโนมัติทุจริตได้อย่างเป็นรูปธรรม

**4.6 การอภิปรายข้อจำกัดและผลกระทบเชิงลบ**

อย่างไรก็ตาม แม้กลไกการย้ายพอร์ตแบบฉับพลันจะแสดงประสิทธิภาพสูงในมิติของการป้องกัน แต่ผู้วิจัยตระหนักถึงผลแลกเปลี่ยน (Trade-offs) ที่อาจส่งผลต่อประสบการณ์การเรียนรู้ในสองประเด็นหลัก ได้แก่

- ความต่อเนื่องของเซสชัน (Session Continuity) การยุติคอนเทนเนอร์ด้วยสัญญาณ SIGKILL และสลับพอร์ตทันทีย่อมทำให้การเชื่อมต่อเครือข่ายของผู้เรียน (TCP Session, Database Connection) ถูกตัดขาดโดยไม่แจ้งเตือน ซึ่งอาจก่อให้เกิดความสับสนเชิงกระบวนการฝึก โดยเฉพาะกับผู้เรียนระดับเริ่มต้นที่ไม่สามารถแยกแยะระหว่าง พฤติกรรมของระบบป้องกัน กับ ความผิดพลาดของเครื่องมือ
- ความต่อเนื่องของการเรียนรู้ (Learning Continuity) การเปลี่ยนสภาพแวดล้อมอย่างรวดเร็วอาจกระทบต่อกระบวนการสะสมหลักฐาน (Artifact Collection) และการทำเอกสารการเจาะระบบ (Penetration Report) ซึ่งเป็นทักษะสำคัญในหลักสูตรความมั่นคงปลอดภัยไซเบอร์
- ความเที่ยงของการประเมิน (Assessment Fairness) กรณีที่ผู้เรียนถูกตัดเซสชันในจังหวะที่กำลังจะโจมตีสำเร็จ อาจเกิดอคติของการประเมินผลอันมีต้นเหตุจากความผันผวนของระบบ มิใช่จากทักษะของผู้เรียน ผู้วิจัยจึงเสนอให้มีการศึกษาผลกระทบเชิงประจักษ์ต่อประสบการณ์การเรียนรู้ของผู้ใช้ในงานวิจัยลำดับถัดไป ควบคู่กับการพัฒนากลไกแจ้งเตือนแบบเนียนเข้ากับเกมเพลย์ (In-game Notification) เพื่อสื่อสารการเปลี่ยนแปลงโครงสร้างให้ผู้เรียนรับรู้

## 5. สรุปและงานในอนาคต

งานวิจัยฉบับนี้นำเสนอกรอบงานสนามยุทธ์ไซเบอร์แบบตอบโต้สองเฟส (2-Phase Reactive Cyber Range) บนสถาปัตยกรรมคอนเทนเนอร์ขนาดเบา โดยมุ่งแก้ไขปัญหาระบบเป้านิ่ง การขาดผลลัพธ์ย้อนกลับแบบทันทีทันใด และการใช้ทรัพยากรที่สูงในระบบดั้งเดิม

ผลการทดลองเชิงประจักษ์แสดงให้เห็นว่า การบูรณาการสถาปัตยกรรมแบบสองเฟสช่วยลดข้อจำกัดด้านความหน่วงจากการประมวลผลของปัญญาประดิษฐ์ได้อย่างชัดเจน ส่งผลให้กลไกออนไลน์สามารถตรวจจับและตอบโต้เชิงรุกด้วยการย้ายพอร์ตหลบหลีก (Port Migration) ได้ภายในเวลาเฉลี่ยเพียง 213.5 มิลลิวินาที ภายใต้ความซับซ้อนเชิงเวลาประมาณ O(|S| × m) ต่อ log entry

โดยที่ |S| คือ จำนวน regex signatures และ m = ความยาวของ log entry แม้กระบวนการดังกล่าวจะมีข้อจำกัดด้านช่องโหว่ทางเวลาอันเกิดจากสภาวะการแข่งขัน (Race Condition) ระหว่างการดึงข้อมูลและการยุติการทำงานของคอนเทนเนอร์ในกรณีการโจมตีแบบคำสั่งเดียว แต่ระบบก็สามารถขัดขวางและลดความต่อเนื่องของการโจมตีจากเครื่องมือสแกนอัตโนมัติได้อย่างมีนัยสำคัญ

นอกจากนี้ การบูรณาการระบบลวงเข้ากับแบบจำลองประเมินผลเชิงคณิตศาสตร์และกลไกจำกัดเพดานคะแนนยังช่วยควบคุมความเบี่ยงเบนและป้องกันการทุจริต ทำให้ผู้เรียนได้รับประสบการณ์การเรียนรู้เชิงรุกภายใต้ระบบที่สามารถตอบสนองต่อเหตุการณ์การโจมตีได้ในระดับเสี้ยววินาที อย่างไรก็ตาม ข้อจำกัดของงานวิจัยฉบับนี้คือการประเมินประสิทธิภาพในปัจจุบันยังจำกัดอยู่ภายใต้รูปแบบการโจมตีประเภท SQL Injection เป็นหลัก จึงจำเป็นต้องมีการทดสอบกับช่องโหว่ประเภทอื่นต่อไปในอนาคต

สำหรับแนวทางการพัฒนางานวิจัยในอนาคตคือ การลดข้อจำกัดสภาวะการแข่งขัน พัฒนากลไกการกรองคำสั่งก่อนการประมวลผล หรือการดักจับในระดับเครือข่ายลึกขึ้น เพื่อสกัดกั้นการโจมตีแบบคำสั่งเดียวให้ทันท่วงที และการประยุกต์ใช้อัลกอริทึมการเรียนรู้แบบเสริมกำลัง เพื่อนำมาใช้ในการเรียนรู้พฤติกรรมความผิดพลาดของผู้เรียน เพื่อปรับค่าน้ำหนักในสมการบทลงโทษแบบอัตโนมัติให้สอดคล้องกับระดับทักษะของแต่ละบุคคลอย่างเป็นพลวัต และการขยายสถาปัตยกรรมสู่ระดับคลัสเตอร์ ยกระดับโครงสร้างพื้นฐานไปสู่ระบบจัดการระดับองค์กร เช่น คูเบอร์เนเตส เพื่อสนับสนุนกลไกการป้องกันเป้าหมายเคลื่อนที่ ในขนาดที่ใหญ่และซับซ้อนขึ้น ซึ่งจะช่วย

รองรับการฝึกอบรมที่มีผู้เข้าร่วมจำนวนมากได้อย่างมีเสถียรภาพ ตลอดจนเตรียมพร้อมรับมือกับรูปแบบการโจมตีผ่านตลอดจนเตรียมพร้อมรองรับรูปแบบการโจมตีและสถานการณ์ช่องโหว่ที่หลากหลายมากขึ้น

สำหรับการขยายขอบเขตการทดลองในอนาคต ควรครอบคลุมรูปแบบการโจมตีที่หลากหลายยิ่งขึ้นตามกรอบ MITRE ATT&CK เช่นCross-Site Scripting (XSS),Privilege Escalation หรือ Multi-stage Attack เพื่อประเมินประสิทธิภาพของระบบในสภาวะที่ซับซ้อนขึ้น นอกจากนี้ควรมีการทดสอบเชิงขยายผล (Scalability Test) ในระดับระบบขนาดใหญ่ที่มีผู้ใช้งานพร้อมกันจำนวนมาก (100-1,000 ผู้ใช้งาน) บนสถาปัตยกรรมการจัดการคลัสเตอร์ระดับองค์กร เช่น คูเบอร์เนเตส ควบคู่กับการนำการทดสอบนัยสำคัญทางสถิติมาใช้เปรียบเทียบผลด้านความหน่วง การใช้ทรัพยากร และอัตราความสำเร็จของการโจมตีระหว่างระบบอ้างอิงและระบบที่นำเสนอ มาใช้วิเคราะห์ผลลัพธ์ เพื่อประเมินขีดความสามารถและเสถียรภาพของระบบสนามยุทธ์ไซเบอร์ได้อย่างครอบคลุมและแม่นยำยิ่งขึ้น

## เอกสารอ้างอิง


[1] A. Mills, J. White, and P. Legg, “GoibhniUWE: A Lightweight and Modular Container-Based Cyber Range,” *Journal of Cybersecurity and Privacy*, vol. 4, pp. 615–628, 2024, doi:10.3390/jcp4030029.

[2] I. Lates, C. Boja, and A. Zamfiroiu, “Multi-Technology Infrastructure for Advanced Training and Testing in Cyber Range Systems,” *Studies in Informatics and Control*, vol. 34, no. 3, pp. 17–28, 2025.

[3] E. Ear, J. L. C. Remy, and S. Xu, “Towards Automated Cyber Range Design: Characterizing and Matching Demands to Supplies,” *arXiv preprint* arXiv:2307.04416, 2023.

[4] R. Nakata and A. Otsuka, “CyExec: A High-Performance Container-Based Cyber Range With Scenario Randomization,” *IEEE Access*, vol. 9, pp. 109095-109114, 2021.

[5] G. Costa, E. Russo, and A. Armando, "Automating the Generation of Cyber Range Virtual Scenarios with VSDL," Journal of Wireless Mobile Networks, Ubiquitous Computing and Dependable Applications, vol.13,no.4,2022,doi:10.58346/JOWUA.2022.14.004

[6] H. Okhravi, T. Hobson, D. Bigelow, and W. Streilein, “Finding Focus in the Blur of Moving-Target Techniques,” *IEEE Security & Privacy*, vol. 12, no. 2, pp. 16-26, 2014.

[7] S. Roh and T.-S. Kim, “LLM-Based Automated Generation and Tri-Modal Representation of Cyber Attack Scenario,” *IEEE Access*, vol. 13, pp. 102930–102945, 2025.

[8] G. Guerrero, G. Betarte, and J. D. Campo, “Assessment of Red and Blue Team Training in the Cyber Range Tectonic using Process Mining,” *CLEI Electronic Journal*, vol. 28, no. 5, 2025.

# การตรวจจับจุดเชื่อมต่อปลอมแบบเพิ่มพูนคลาสด้วยการเรียนรู้เชิงลึกร่วมกับ Exemplar Replay บนข้อมูลสถานะช่องสัญญาณ

# Class-Incremental RAP Detection Using Deep Learning and Exemplar Replay on CSI Signals

*พุทธิชา รัตนา (Phutthicha Rattana), รชต นันทา (Rachata Nanta) และจักรชัย โสอินทร์ (Chakchai So-In)*
*วิทยาลัยการคอมพิวเตอร์ มหาวิทยาลัยขอนแก่น*
*phutthicha.r@kkumail.com, rachata.nan@kkumail.com, chakso@kku.ac.th*

## บทคัดย่อ

*ในปัจจุบัน เครือข่ายแลนไร้สายในอินเทอร์เน็ตของสรรพสิ่ง เกิดความเสี่ยงที่เพิ่มขึ้นจากการโจมตีด้วยจุดเชื่อมต่อปลอม (Rogue Access Point: RAP) ผู้โจมตีมักสร้างจุดเชื่อมต่อที่ปลอมแปลงชื่อเครือข่ายของเครือข่ายจริง เพื่อดักจับข้อมูลและก่อให้เกิดความเสียหายต่อระบบและข้อมูล การระบุตัวตนอุปกรณ์โดยใช้ลายนิ้วมือระดับฮาร์ดแวร์ที่สกัดจากข้อมูลสถานะช่องสัญญาณเป็นแนวทางที่มีประสิทธิภาพ เนื่องจากลักษณะทางกายภาพดังกล่าวปลอมแปลงได้ยากกว่าที่อยู่ MAC อย่างไรก็ตาม งานวิจัยดั้งเดิมยังมีข้อจำกัดสำคัญสองประการ ได้แก่ ปรากฏการณ์ลายนิ้วมือแตกหัก ซึ่งเกิดจากความอ่อนไหวของสัญญาณวิทยุต่อสภาพแวดล้อม เช่น การเคลื่อนไหวของบุคคล สิ่งกีดขวาง หรือการเปลี่ยนแปลงของอุณหภูมิในพื้นที่ ส่งผลให้คุณลักษณะของสัญญาณเกิดการบิดเบี้ยวและทำให้เทคนิคการกรองสัญญาณรบกวนแบบดั้งเดิมไม่มีประสิทธิภาพ และข้อจำกัดที่สองคือโมเดลไม่รองรับการเพิ่มอุปกรณ์ใหม่ในขณะระบบทำงาน โดยไม่ต้องมีการฝึกสอนโมเดลใหม่ทั้งหมด งานวิจัยนี้แปลงข้อมูลสถานะช่องสัญญาณเป็นแผนภาพความร้อน ขนาด 128×128 พิกเซล จาก 100 แพ็กเก็ตข้อมูลที่ต่อเนื่องกัน ซึ่งเป็นการตั้งค่าที่รักษาสมดุลระหว่างรายละเอียดเชิงพื้นที่และความเร็วในการตอบสนองตามผลการทดลองก่อนหน้า นอกจากนี้ ยังได้ประยุกต์ใช้โครงข่าย Auto-NRK ที่เป็นชื่อเรียกตามกรอบของ Yang [1] ซึ่งเป็นสถาปัตยกรรม ResNet Autoencoder ที่ปรับปรุงด้วย Skip Connections และฟังก์ชันความสูญเสีย แบบแยกสัญญาณ เพื่อกำจัดสัญญาณรบกวนออกจากลายนิ้วมือทางกายภาพ การออกแบบดังกล่าวช่วยรักษาโครงสร้างความถี่ของข้อมูลสถานะช่องสัญญาณไว้ได้อย่างมีประสิทธิภาพ งานวิจัยนี้จึงได้พัฒนาโมเดล Self-ACC ที่มีการเรียนรู้แบบเพิ่มพูนคลาส เข้ากับกลไก Centroid-based Exemplar Replay, การเพิ่มข้อมูลด้วยสัญญาณรบกวนแบบเกาส์ และการถ่ายทอดความรู้ เพื่อให้ระบบสามารถเรียนรู้คุณลักษณะของอุปกรณ์ใหม่ได้อย่างต่อเนื่องโดยไม่ลืมความรู้เดิม สำหรับการตรวจจับ RAP แบบ Zero-day ระบบได้ใช้กลไกการกรองแบบสองขั้นตอนที่ผสานการใช้เกณฑ์ขอบเขต จากค่าความผิดพลาดในการสร้างภาพใหม่ ร่วมกับ Dynamic Shannon Entropy Threshold ผลการทดสอบบนชุดข้อมูลภาพของสถานะช่องสัญญาณ จำนวน 4,086 ภาพ จากอุปกรณ์ 11 คลาส แสดงให้เห็นว่าระบบสามารถรักษาความแม่นยำในการระบุตัวตนได้ถึงร้อยละ 92.00 โดยมีค่า F1-Score สำหรับการตรวจจับจุดเชื่อมต่อปลอม อยู่ที่ร้อยละ 80.56 และมีอัตราการเตือนผิดพลาดที่ประมาณร้อยละ 15 จากผลลัพธ์ดังกล่าวมีประสิทธิภาพดีกว่าวิธีการเปรียบเทียบทั้งในด้านความแม่นยำในการระบุ (92.00% เทียบกับ 39.01%) และค่า F1-Score ในการตรวจจับจุดเชื่อมต่อปลอม (80.56% เทียบกับ 75.06%) ในขณะที่ยังคงอัตราการเตือนผิดพลาดที่ต่ำกว่าอย่างมีนัยสำคัญ*

**คำสำคัญ:** จุดเชื่อมต่อปลอม, การเรียนรู้แบบเพิ่มพูนคลาส, การเรียนรู้เชิงลึก, ลายนิ้วมือฮาร์ดแวร์, ข้อมูลสถานะช่องสัญญาณ, ความมั่นคงปลอดภัยอินเทอร์เน็ตของสรรพสิ่ง

## Abstract

*Currently, wireless networks (Wi-Fi) within the Internet of Things (IoT) ecosystem are facing an increasing risk of Rogue Access Point (RAP) attacks. Attackers often deploy malicious access points that spoof the Service Set Identifier (SSID) of legitimate networks to intercept sensitive data and compromise system security. Device identification using hardware-level fingerprints extracted from Channel State Information (CSI) provides an effective solution, as these physical characteristics are significantly more difficult to forge than MAC addresses. However, existing studies still face two major limitations. The first is the Fingerprint Fracture Phenomenon (FFP), which arises from the sensitivity of radio signals to environmental factors such as human movement, obstacles, and temperature variations. This effect distorts signal characteristics and reduces the effectiveness of conventional noise-filtering techniques. The second limitation is the inability of existing models to incorporate new devices during runtime without retraining from scratch. This study transforms CSI data into 128×128 pixel heatmaps generated from 100 consecutive packets, achieving a balance between spatial detail and response time. Furthermore, a ResNet-based autoencoder (Auto-NRK), enhanced with skip connections and a signal-separation loss function, is employed to denoise the extracted fingerprints while preserving their frequency structure. In addition, the proposed Self-ACC model integrates Class-Incremental Learning with Centroid-based Exemplar Replay, Gaussian Noise Augmentation, and Knowledge Distillation, enabling the system to continuously learn new device characteristics without catastrophic forgetting. For zero-day RAP detection, a two-stage filtering mechanism is introduced, combining a reconstruction error-based range threshold with a Dynamic Shannon Entropy threshold. Experimental results on a dataset comprising 4,086 CSI images across 11 device classes show that the proposed system achieves an identification accuracy of 92.00%, an F1-score of 80.56% for RAP detection, and a false alarm rate (FAR) of approximately 15%. These results indicate that the proposed framework provides a more balanced overall detection performance than the evaluated baseline methods*

**Keywords:** Rogue Access Point (RAP), Class-Incremental Learning (CIL), Deep Learning, Hardware Fingerprinting, Channel State Information (CSI), IoT Security.

## 1. บทนำ

ปัจจุบัน เทคโนโลยีอินเทอร์เน็ตของสรรพสิ่ง (IoT) ได้เข้ามามีบทบาทสำคัญในชีวิตประจำวันอย่างแพร่หลาย ทั้งในรูปแบบของบ้านอัจฉริยะ ยานยนต์ไร้คนขับ และอุปกรณ์ทางการแพทย์เชื่อมต่ออินเทอร์เน็ต เนื่องจากอุปกรณ์ IoT ส่วนใหญ่อาศัยเครือข่ายแลนไร้สาย (Wi-Fi) เป็นช่องทางการสื่อสารหลัก ความปลอดภัยของจุดเชื่อมต่อสัญญาณ (Access Point: AP) จึงเป็นปัจจัยสำคัญต่อความมั่นคงปลอดภัยของระบบ IoT

หนึ่งในภัยคุกคามที่รุนแรงที่สุดคือการโจมตีด้วยจุดเชื่อมต่อปลอม (Rogue Access Point: RAP) โดยผู้โจมตีจะสร้างจุดเชื่อมต่อปลอมที่ปลอมแปลงชื่อเครือข่าย (SSID) ให้ตรงกับจุดเชื่อมต่อที่ถูกต้องตามกฎหมาย (Legitimate AP) เพื่อหลอกให้อุปกรณ์ IoT เชื่อมต่อเข้าสู่ระบบและดักจับข้อมูลส่วนบุคคล ซึ่งอาจนำไปสู่การรั่วไหลของข้อมูลและความเสียหายต่อระบบเครือข่ายหรือบริการที่เกี่ยวข้อง มาตรการป้องกันแบบดั้งเดิมที่อาศัยการตรวจสอบที่อยู่ MAC หรือการวิเคราะห์พฤติกรรมการส่งข้อมูลนั้นไม่มีประสิทธิภาพ

เพียงพอ เนื่องจากที่อยู่ MAC สามารถถูกปลอมแปลงได้ง่าย และการวิเคราะห์พฤติกรรมต้องใช้ระยะเวลาดำเนินการนานกว่าจะตรวจพบความผิดปกติ

ด้วยเหตุนี้ การระบุตัวตนอุปกรณ์ด้วยลายนิ้วมือระดับชั้นกายภาพ (Physical-layer Fingerprinting) ผ่านการใช้ข้อมูลสถานะช่องสัญญาณ (Channel State Information: CSI) จึงกำลังได้รับความสนใจอย่างมาก เนื่องจากข้อมูลสัญญาณ CSI สามารถบันทึกลักษณะเฉพาะของวงจรวิทยุความถี่สูงที่เกิดจากความคลาดเคลื่อนในกระบวนการผลิตของแต่ละอุปกรณ์ ซึ่งเป็นคุณลักษณะที่ปลอมแปลงได้ยากและมีความคงที่ตลอดอายุการใช้งาน

Yang [1] ได้นำเสนอกรอบการทำงาน DL-PEDR ซึ่งค้นพบและอธิบายปรากฏการณ์ลายนิ้วมือแตกหัก (Fingerprint Fracture Phenomenon: FFP) อย่างเป็นระบบ โดยระบุว่าการเลื่อนของเฟสที่เกิดจากสภาพแวดล้อม (เช่น การขยับตัวของบุคคล การเคลื่อนย้ายสิ่งของในพื้นที่ หรือแม้กระทั่งการเปลี่ยนแปลงของอุณหภูมิและความชื้น) ส่งผลให้ค่าความคลาดเคลื่อนของเฟสเกิดการกระจายตัวเป็นหลายส่วน สัญญาณวิทยุที่สะท้อนจากปัจจัยสภาพแวดล้อมเหล่านี้ทำให้โครงสร้างของลายนิ้วมือทางกายภาพเกิดความผันผวนและแตกหัก ทำให้ตัวกรองสัญญาณรบกวนแบบดั้งเดิมไม่สามารถแยกแยะสัญญาณรบกวนออกจากลายนิ้วมือจริงได้

แม้ว่ากรอบการทำงาน DL-PEDR จะแก้ปัญหานี้ด้วยการใช้โครงข่าย Auto-NRK เพื่อทำความสะอาดลายนิ้วมือและใช้โมเดล Self-ACC ในการจำแนกอุปกรณ์ แต่ยังคงมีข้อจำกัดสำคัญดังนี้คือ 1) การแทนค่าการกระจายตัวของความคลาดเคลื่อนด้วยกล่องขอบเขตแบบไบนารี (Binary Bounding Box) ทำให้สูญเสียรายละเอียดโครงสร้างเชิงลึกของลายนิ้วมือ และ 2) โมเดลไม่รองรับการเพิ่มอุปกรณ์ AP ใหม่ในระหว่างการทำงานของระบบ ทำให้ต้องทำการฝึกสอนโมเดลใหม่ทั้งหมดเมื่อมีการเพิ่มอุปกรณ์ ซึ่งสิ้นเปลืองทรัพยากรและก่อให้เกิดปัญหาการลืมความรู้เดิม (Catastrophic Forgetting)

เพื่อแก้ไขปัญหาของช่องว่างดังกล่าว งานวิจัยนี้นำเสนอกรอบการเรียนรู้เชิงลึกที่พัฒนาต่อยอดจาก DL-PEDR โดยมีวิธีการหลัก 3 ประการ ดังนี้คือ

1. การเปลี่ยนรูปแบบการแทนค่าข้อมูล CSI จากกล่องขอบเขตแบบไบนารีมาเป็นภาพแผนผังความร้อน 2 มิติ (CSI Drift-Range Heatmap) ขนาด 128×128 พิกเซล ที่สร้างจาก 100 แพ็กเก็ตต่อเนื่อง เพื่อดึงรายละเอียดลายนิ้วมือเชิงลึกที่การแทนค่าแบบเดิมไม่สามารถรักษาไว้ได้
2. การฝึกสอนและประยุกต์ใช้โครงข่าย Auto-NRK ในการกรองสัญญาณรบกวนที่สามารถเรียนรู้ได้ (Trainable Denoising Filter) เพื่อฟื้นฟูภาพลายนิ้วมือที่สะอาดก่อนส่งเข้าสู่กระบวนการจำแนก
3. การเพิ่มประสิทธิภาพโมเดล Self-ACC ด้วยกลไกการเรียนรู้แบบเพิ่มพูนคลาส (Class-Incremental Learning: CIL) ที่บูรณาการ Centroid-based Exemplar Replay, Gaussian Noise Augmentation และ Knowledge Distillation เข้าด้วยกัน เพื่อให้ระบบสามารถเรียนรู้อุปกรณ์ใหม่ได้อย่างต่อเนื่องโดยไม่ลืมความรู้เดิม พร้อมทั้งนำเสนอกลไกการกรองจุดเชื่อมต่อปลอม (RAP) แบบ 2 ขั้นตอนที่ประยุกต์ใช้สมการ Shannon Entropy เพื่อควบคุมอัตราการเตือนผิดพลาด

## 2. งานวิจัยที่เกี่ยวข้อง

งานวิจัยที่เกี่ยวข้องกับบทความนี้สามารถแบ่งออกเป็น 3 กลุ่มหลัก ได้แก่ การระบุตัวตนอุปกรณ์ด้วยลายนิ้วมือระดับชั้นกายภาพ การสกัดคุณลักษณะขั้นสูงและการลดสัญญาณรบกวน และการตรวจจับภัยคุกคามร่วมกับการเรียนรู้อย่างต่อเนื่อง (Class-Incremental Learning)

### 2.1 การระบุตัวตนอุปกรณ์ด้วยลายนิ้วมือระดับชั้นกายภาพ

Yang [1] ได้นำเสนอกรอบการทำงาน DL-PEDR ซึ่งเป็นงานแรกที่อธิบายปรากฏการณ์ลายนิ้วมือแตกหัก (FFP) อย่างเป็นระบบ โดยพบว่าค่าความคลาดเคลื่อนของเฟส (Phase Error) หลังจากกระบวนการคืนรูปเฟส

(Phase Unwrapping) สามารถนำมาใช้เป็นลายนิ้วมือทางฮาร์ดแวร์ที่เสถียรได้ งานวิจัยดังกล่าวได้พัฒนา Auto-NRK เพื่อกำจัดสัญญาณรบกวนและใช้โมเดล Self-ACC เพื่อจำแนกอุปกรณ์ แม้ว่าจุดแข็งของงานนี้คือการวิเคราะห์ปรากฏการณ์ลายนิ้วมือแตกหัก (FFP) อย่างละเอียดและมีสถาปัตยกรรมที่นำไปใช้งานได้จริง แต่ข้อจำกัดสำคัญคือการใช้กล่องขอบเขตแบบไบนารี (Binary Bounding Box) ในการแทนค่าข้อมูล ส่งผลให้รายละเอียดการกระจายตัวของสัญญาณสูญหาย อีกทั้งตัวโมเดลยังไม่รองรับการเพิ่มคลาสอุปกรณ์ใหม่

ในขณะเดียวกัน Zhang [4] ได้นำเสนอ P-A Scheme ซึ่งผสมผสานเฟสและแอมพลิจูดของข้อมูล CSI เพื่อสร้างลายนิ้วมืออุปกรณ์ในสภาพแวดล้อมอุตสาหกรรม วิธีการนี้มีจุดเด่นด้านความทนทานต่อการเปลี่ยนแปลงของสภาพแวดล้อม อย่างไรก็ตาม ปัญหาการเลื่อนของเฟส (Phase Drift) ยังไม่ได้รับการแก้ไขอย่างสมบูรณ์ ทำให้ความแม่นยำลดลงเมื่ออยู่ในสภาวะที่มีสัญญาณรบกวนสูง

นอกจากนี้ Xie [11] ได้ทบทวนงานวิจัยด้านลายนิ้วมือคลื่นความถี่วิทยุ (RF Fingerprint) สำหรับระบบ IoT และพบว่าความแม่นยำในการใช้งานจริงจะต่ำกว่าในห้องปฏิบัติการ ซึ่งมีสาเหตุหลักมาจากแบบจำลองสัญญาณรบกวนที่ไม่ครอบคลุมและการใช้คุณลักษณะที่มีความละเอียดต่ำ โดยงานวิจัยดังกล่าวเสนอแนวทางการแก้ไขผ่านการใช้ Heatmap และโครงข่าย Auto-NRK

**2.2 การสกัดคุณลักษณะขั้นสูงและการลดสัญญาณรบกวน**

Kong และ Chen [2, 9] ได้พัฒนากรอบการทำงาน DeepCRF สำหรับการระบุลายนิ้วมือ RF จากข้อมูล CSI ที่มีความทนทานต่อการเปลี่ยนแปลงของช่องสัญญาณ โดยประยุกต์ใช้ Feature Map ในหลายระดับความละเอียด เพื่อรักษาความเสถียรในการสกัดลายนิ้วมือ ผลลัพธ์บ่งชี้ว่าการแทนค่าข้อมูลนำเข้าด้วยความละเอียดสูงส่งผลให้ประสิทธิภาพของระบบดีกว่าการใช้ข้อมูลแบบสเกลาร์หรือไบนารี ซึ่งสอดคล้องกับแนวคิดการเลือกใช้ Heatmap แทนกล่องขอบเขตไบนารีในงานวิจัยนี้ ทางด้าน Wang [6] มุ่งเน้นงาน Wi-Fi fingerprint positioning แต่แสดงให้เห็นถึงประสิทธิภาพของ Stacked Denoising Autoencoder (SDAE) ในการลดสัญญาณรบกวนออกจากลายนิ้วมือ ซึ่งเป็นแนวทางสำคัญสำหรับการพัฒนาสถาปัตยกรรม Auto-NRK ในงานวิจัยฉบับนี้

**2.3 การตรวจจับภัยคุกคามและการเรียนรู้อย่างต่อเนื่อง**

Paula et al. [5] นำเสนอ One-Class Machine Learning เพื่อตรวจจับพฤติกรรมผิดปกติของอุปกรณ์แบบ Passive โดยสร้างโมเดลขอบเขตแยกสำหรับ AP แต่ละตัวที่รู้จัก ซึ่งสามารถปฏิเสธอุปกรณ์นอกกลุ่มได้โดยอัตโนมัติ อย่างไรก็ตาม ระบบนี้ขยายได้ยากเมื่อจำนวน AP เพิ่มขึ้น เนื่องจากขาดกลไกรักษาขอบเขตของคลาสเดิมขณะเพิ่มคลาสใหม่

Sun et al. [3] นำเสนอ AFD-IL ซึ่งเป็นกรอบการเรียนรู้แบบเพิ่มพูนสำหรับระบุตัวผู้ส่งสัญญาณ โดยใช้ Adaptive Feature Distillation เป็นฟังก์ชันความสูญเสียเพื่อรักษาประสิทธิภาพในแต่ละรอบการเรียนรู้ แม้แนวทางนี้จะเป็นต้นแบบสำหรับ Knowledge Distillation ในงานวิจัยนี้ แต่ AFD-IL ไม่ได้จัดการกับปรากฏการณ์ FFP และยังขาดกลไกปฏิเสธแบบ Open-Set นอกจากนี้ Shi et al. [10] นำเสนอวิธีแทนลายนิ้วมือ RF ที่ทนต่อความยาวสัญญาณ (Length-Robust) ใช้การเรียนรู้แบบเพิ่มพูนเพื่อระบุตัวผู้ส่งสัญญาณ ระบบแม่นยำพอบนชุดข้อมูลขยาย แต่ระบบไม่มีกลไกปฏิเสธ Open-Set ทำให้เสี่ยงต่อการโจมตี Zero-day

Arisandi et al. [7] พัฒนาระบบตรวจจับความผิดปกติของจุดเชื่อมต่อปลอม (RAP) ที่ทำงานบนชั้น Layer 2 ระบบจะตรวจจับพฤติกรรมในเครือข่าย ระบบทำงานดีเมื่ออุปกรณ์แสดงพฤติกรรมผิดปกติชัดเจน แต่ประสิทธิภาพจะลดลงเมื่อเจอกับการโจมตีแบบเลียนแบบ (Mimicry Attack) ที่มีความซับซ้อน

**ตารางที่ 1:** การเปรียบเทียบงานวิจัยที่เกี่ยวข้องกับงานวิจัยที่นำเสนอแสดงเฉพาะงานตัวแทนที่สัมพันธ์กับองค์ประกอบหลักของกรอบงานวิจัยนี้ (✓ = รองรับ; – = ไม่รองรับ)

| คุณลักษณะ / วิธีการ | Yang [1] | Kong [2] | Zhang [4] | Paula [5] | Sun [3] | งานนี้ |
|---|---|---|---|---|---|---|
| ลายนิ้วมือ CSI | ✓ | ✓ | ✓ | - | - | ✓ |
| สกัดคุณลักษณะจากเฟส | ✓ | - | ✓ | - | - | ✓ |
| ข้อมูลนำเข้าหลายมิติ | - | ✓ | - | - | - | ✓ |
| Autoencoder กำจัดสัญญาณรบกวน (FFP) | ✓ | - | - | - | - | ✓ |
| ตรวจจับ RAP แบบ Open-Set/Zero-day | - | - | - | ✓ | - | ✓ |
| การเพิ่มคลาสใหม่ต่อเนื่อง (CIL) | - | - | - | - | ✓ | ✓ |
| การกรอง 2 ด่าน (Recon + Entropy) | - | - | - | - | - | ✓ |

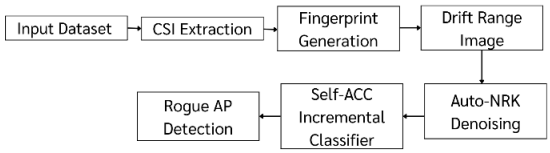
Hong et al. [8] นำเสนอวิธีระบุลายนิ้วมือ RF แบบ Open-Set Domain Adaptive ด้วย Prototype Calibration วิธีนี้เข้ากันได้ดีกับกลไก Entropy-based Gating ที่นำเสนอในงานวิจัยนี้ อย่างไรก็ตาม งานดังกล่าวไม่ได้บูรณาการกับกลไก Class-Incremental Learning หรือการกรองสัญญาณรบกวนสำหรับข้อมูล CSI โดยตรง

**2.4 สรุปช่องว่างทางการวิจัยและตารางเปรียบเทียบ**

จากการสำรวจข้างต้น ยังไม่พบบทความใดที่ผสานความสามารถสกัดคุณลักษณะความละเอียดสูง (Heatmap) เข้ากับโครงข่าย Autoencoder เพื่อกำจัดสัญญาณรบกวน การเรียนรู้แบบเพิ่มพูนคลาส (CIL) และกลไกการกรองจุดเชื่อมต่อปลอม (RAP) แบบ 2 ด่าน ไว้ในกรอบเดียวกัน งานวิจัยที่นำเสนอนี้มีความแตกต่างจากงานก่อนหน้าในด้านของการผสานองค์ประกอบเหล่านี้ไว้ในกรอบเดียวกัน (ดังสรุปในตารางที่ 1)

## 3. วิธีดำเนินงานวิจัย

งานวิจัยนี้ดำเนินการกับ CSI ดิบ จากจุดเชื่อมต่อ Wi-Fi มาทำงานตาม 4 ขั้นตอน (1) ดึงคุณลักษณะจากชั้นกายภาพ แล้วคำนวณความคลาดเคลื่อนของเฟส (2) สร้าง Heatmap แล้วใช้ Auto-NRK กำจัดสัญญาณรบกวน (3) แยกอุปกรณ์ด้วย Self-ACC ที่เพิ่มพูนคลาส (4) ตรวจจับ RAP ด้วย 2 ด่าน โดยมีภาพรวมสถาปัตยกรรม แสดงในภาพที่ 1




**ภาพที่ 1:** สถาปัตยกรรมของระบบยืนยันตัวตนอุปกรณ์จากข้อมูล

**3.1 การสกัดคุณลักษณะระดับชั้นกายภาพ**

ข้อมูลสถานะช่องสัญญาณ (CSI) บันทึกลักษณะการแพร่กระจายของคลื่นวิทยุในรูปแบบสัมประสิทธิ์เชิงซ้อนสำหรับแต่ละช่องย่อย OFDM โดยสัมประสิทธิ์นั้นมีรูปแบบ $H = Ae^{j\varphi}$ ซึ่ง A คือแอมพลิจูดที่สะท้อนการลดทอนสัญญาณตามเส้นทาง และ $\varphi$ คือเฟสที่เข้ารหัสทั้งความล่าช้าการแพร่กระจายและการตอบสนองเฟสของฮาร์ดแวร์ผู้ส่ง เนื่องจากความคลาดเคลื่อนทางการผลิตสร้างอคติที่เฉพาะเจาะจงสำหรับแต่ละอุปกรณ์ใน $\varphi$ ค่าเฟสจึงเป็นตัวพาลายนิ้วมือฮาร์ดแวร์ที่ต้องการใช้

ในทางปฏิบัติ เฟสที่วัดได้จริง $\varphi_{raw}$ มีมลภาวะจาก 2 สัญญาณรบกวนหลัก ได้แก่ Sampling Frequency Offset (SFO) ที่สัดส่วนกับดัชนีช่องย่อยแบบเชิงเส้น และ Packet Detection Delay (PDD) ที่เป็นค่าคงที่สุ่มต่อแพ็กเก็ต สัญญาณรบกวนทั้งสองเป็นอิสระจากลายนิ้วมือฮาร์ดแวร์ เพื่อแยกส่วนที่เป็นลายนิ้วมือออกมา ขั้นแรกจะทำการ Phase Unwrapping เพื่อแปลงค่าเฟสที่ถูกจำกัดอยู่ในช่วง $[-\pi, \pi]$ ให้เป็นค่าเฟสต่อเนื่อง จากนั้นลบแนวโน้มเชิงเส้นของ SFO ที่ประมาณด้วย Least Squares Regression ผ่าน 30 ช่องย่อย OFDM

ส่วนที่เหลือหลังจากลบแนวโน้มออกแล้ว คือค่าความคลาดเคลื่อนของเฟส (Phase Error) $\varepsilon = \varphi_{unwrapped} - \varphi_{SFO}$ โดย PDD เป็น constant offset ที่เท่ากันทุกช่องย่อย จึงถูกดูดซับเข้าสู่ intercept term ของ Least Squares Regression และถูกลบออกพร้อมกับ SFO trend โดยอัตโนมัติ ค่า $\varepsilon$ จึงบรรจุส่วนเบี่ยงเบนเฉพาะฮาร์ดแวร์ที่เป็นลายนิ้วมือ CSI ของอุปกรณ์นั้น ๆ กระบวนการนี้ทำงานอิสระสำหรับแต่ละช่องย่อยใน 30 ช่องย่อย OFDM ของชุดข้อมูล 802.11n ทำให้ได้เวกเตอร์ Phase Error ขนาด 30 มิติต่อแพ็กเก็ต

**3.2 การสร้างภาพ Heatmap และการกำจัดสัญญาณรบกวน (Auto-NRK)**

**3.2.1 การสร้างภาพความร้อน 2 มิติ (CSI Drift-Range Heatmap)**

แทนที่จะสรุปการกระจายตัวของ Phase Error ด้วยกล่องขอบเขตแบบไบนารีตามแนวทางของ DL-PEDR งานวิจัยนี้สร้างภาพความร้อน 2 มิติที่รักษารูปร่างการกระจายตัวทั้งหมดไว้ สำหรับแต่ละลำดับ $N = 100$ แพ็กเก็ตต่อเนื่องจาก AP หนึ่งเครื่อง จะสะสมเวกเตอร์ Phase Error ขนาด 30 มิติ แล้วคำนวณ Histogram ขนาด 2 มิติในพื้นที่ร่วมของดัชนีช่องย่อย (แกน $x$) และขนาด Phase Error (แกน $y$) จากนั้น Normalize ให้ผลรวมเป็น 1 และ Interpolate แบบ Bilinear ไปยังกริดคงที่ขนาด 128×128 พิกเซล ได้ภาพที่เรียกว่า CSI Drift-Range Heatmap

การแทนค่าแบบนี้มีคุณสมบัติที่พึงประสงค์ 3 ประการ ประการแรกคือจับรายละเอียดโครงสร้างจุลภาค (Micro-signal Texture) ของลายนิ้วมือ รวมถึงความไม่สมมาตรและโหมดรอง ซึ่งกล่องขอบเขตไม่สามารถแทนค่าได้ ประการที่สองคือการสะสม 100 แพ็กเก็ตช่วยเฉลี่ยความผันแปรช่องสัญญาณระยะสั้นที่ส่งผลต่อแพ็กเก็ตเดี่ยว ประการที่สามคือความละเอียดเชิงพื้นที่คงที่ทำให้เข้ากันได้กับสถาปัตยกรรม CNN มาตรฐานโดยตรง

**3.2.2 โครงข่าย Auto-NRK**

Auto-NRK เป็น Convolutional Autoencoder แบบ ResNet ที่ฝึกสอนเพื่อสร้างภาพ Heatmap ที่สะอาดจากข้อมูลนำเข้าที่มีสัญญาณรบกวน สถาปัตยกรรมแบ่งออกเป็น 2 ส่วนหลัก ฝั่ง Encoder ประกอบด้วย Residual Block 4 ชั้นที่ลดมิติเชิงพื้นที่ลงทีละครึ่ง (128⟶ 64⟶ 32⟶ 16⟶ 8) แต่ละ Block มีชั้น Convolution 3×3 สองชั้นพร้อม Batch Normalization และ ReLU Activation พร้อม Skip Connection ที่บวกข้อมูลนำเข้าของ Block กลับเข้ากับผลลัพธ์เพื่อบรรเทาปัญหา Gradient Vanishing ระหว่างการฝึกสอน Bottleneck ที่ความละเอียดเชิงพื้นที่ 8×8 บันทึกโครงสร้างลายนิ้วมือที่จำเป็น ฝั่ง Decoder สะท้อนโครงสร้าง Encoder โดยใช้ Transposed Convolution ที่ฟื้นฟูความละเอียดเชิงพื้นที่ทีละขั้น (8 ⟶16 ⟶32 ⟶64 ⟶128)

Auto-NRK ฝึกสอนด้วยคู่ Heatmap แบบมีสัญญาณรบกวน-สะอาด โดยใช้ Mean Squared Error (MSE) เป็น Loss Function ของการสร้างภาพใหม่ดังสมการที่ (1):

$$L_{recon} = \left(\frac{1}{WH}\right) \times \sum_{i,j} \left(X(i,j) - \mathrm{H}(i,j)\right)^2 \qquad (1)$$

โดยที่ $X$ คือ Heatmap นำเข้าที่มีสัญญาณรบกวน, $\hat{H}$ คือผลลัพธ์ที่สร้างใหม่ และ $W = H = 128$ คือมิติเชิงพื้นที่ ในระหว่างการ Inference Auto-NRK ทำงานเป็นตัวกรองแบบ Deterministic โดยค่าความผิดพลาดการสร้างภาพใหม่ ซึ่งนิยามเป็น $E_{recon} = L_{recon}(X, \hat{\mathrm{H}})$ จะถูกเก็บไว้สำหรับใช้ใน Stage ที่ 1 ของการตรวจจับตามที่อธิบายในหัวข้อ 3.4

**3.3 การจำแนกอุปกรณ์แบบเพิ่มพูนคลาส**

**3.3.1 สถาปัตยกรรม Self-ACC**

Self-ACC เป็น CNN แบบ ResNet ที่มีบล็อก Squeeze-and-Excitation (SE) Attention ซึ่งได้รับ Heatmap ขนาด 128×128 พิกเซลเป็นข้อมูลนำเข้า กลไก SE แก้ปัญหาเฉพาะของการระบุลายนิ้วมือ CSI นั่นคือช่องย่อยต่าง ๆ มีปริมาณข้อมูลลายนิ้วมือไม่เท่ากัน และการกระจายตัวนี้แตกต่างกันตามชนิดอุปกรณ์ SE Block จึงเรียนรู้การสอบเทียบความสำคัญของ Feature Map Channel แต่ละช่องแบบ Adaptive

กลไก SE ทำงาน 2 ขั้นตอน ขั้นตอน Squeeze ใช้ Global Average Pooling (GAP) เพื่อบีบอัด Feature Map แต่ละช่อง $C$ จากมิติเชิงพื้นที่ $H \times W$ เป็นเวกเตอร์ Descriptor ขนาด $C$ ดังสมการที่ (2):

$$z_c = \left(\frac{1}{HW}\right) \times \sum_{i,j} u_c(i,j) \quad (2)$$

โดยที่ $u_c(i,j)$ คือค่าของ Feature Map ที่ช่อง $c$ ตำแหน่ง $(i,j)$ ขั้นตอน Excitation ใช้ Fully Connected จำนวน 2 ชั้นพร้อม ReLU-Sigmoid เพื่อแปลง $z$ เป็นเวกเตอร์ Weight แบบ Channel-wise ดังสมการที่ (3):

$$s = \sigma(W_2 \cdot \delta(W_1 \cdot z)) \quad (3)$$

โดยที่ $\delta$ คือ ReLU และ $\sigma$ คือ Sigmoid จากนั้น Feature Map เดิมถูก Re-weight แบบ Element-wise ดังสมการที่ (4):

$$\tilde{u}_c = s_c \times u_c \quad (4)$$

การดำเนินการนี้ทำให้โครงข่ายกดช่องที่ไม่มีข้อมูลลายนิ้วมือและขยายช่องที่มีข้อมูลสำคัญ ตามข้อมูลนำเข้าแต่ละตัวอย่าง หลัง Residual Block และ SE Attention Feature Vector จะถูกส่งผ่าน Fully Connected จำนวน 2 ชั้น (512 และ 256 โหนด ทั้งคู่มี ReLU และ Dropout 0.5) และ Softmax Output Head ที่มี $K$ โหนด ซึ่ง $K$ คือจำนวนคลาส AP ที่รู้จักในปัจจุบัน

**3.3.2 กลไกการเรียนรู้แบบเพิ่มพูนคลาส**

เมื่อแนะนำคลาส AP ใหม่สำหรับการลงทะเบียน การปรับจูน Self-ACC แบบ Naive บนข้อมูลใหม่จะทำให้เกิด Catastrophic Forgetting เนื่องจากการอัปเดต Gradient ที่ Optimize สำหรับคลาสใหม่จะเขียนทับ Weight Configuration ที่เข้ารหัสความรู้ของคลาสก่อนหน้า งานวิจัยนี้แก้ปัญหานี้ด้วย 3 กลไกเสริมกัน ดังนี้คือ

- กลไกที่ 1: Centroid-based Exemplar Replay สำหรับแต่ละคลาส c ที่รู้จัก ระบบรักษา Buffer $M_c$ ขนาด $m$ = 100 Heatmap สำหรับแต่ละคลาส $c$ โดยตัวแทนถูกเลือกจากการคำนวณ Heatmap เฉลี่ย $\mu_c$ ของตัวอย่างทั้งหมดในคลาส แล้วเก็บ $m$ ตัวที่มี MSE ต่ำสุดเทียบกับ $\mu_c$ เกณฑ์ Centroid Proximity นี้รับประกันว่าตัวแทนที่เก็บไว้แทนศูนย์กลางคลาสอย่างซื่อสัตย์แทนที่จะเป็นค่าผิดปกติ เมื่อเพิ่มคลาสใหม่ Buffer ของคลาสที่มีอยู่ทั้งหมดจะถูก Replay ร่วมกับข้อมูลฝึกสอนใหม่ในการอัปเดต Gradient แต่ละครั้ง โดยที่งานวิจัยนี้ไม่ได้ใช้วิธีการสุ่มเลือกข้อมูลตัวแทน (Random Selection) แต่เลือกใช้เกณฑ์การคัดเลือกข้อมูลที่อยู่ใกล้ค่าเฉลี่ยของคลาสมากที่สุด (Mean of Exemplars)
- กลไกที่ 2: Gaussian Noise Augmentation เพื่อไม่ให้โมเดล overfit กับ exemplar เล็กที่คงที่ ระบบเพิ่ม Gaussian noise แบบ zero-mean ($\sigma$ = 0.05) ลงใน Exemplar ทุกรายการที่ถูก Replay เนื่องจากขนาด noise ดังกล่าวเล็กกว่า Dynamic Range ของ Heatmap อย่างมีนัยสำคัญ ตัวอย่างที่ถูก augment จึงยังคงอยู่ภายในการกระจายของข้อมูลจริง ในขณะที่ช่วยเพิ่มความหลากหลายของสัญญาณฝึกสอนและลดความเสี่ยงต่อ Memorization
- กลไกที่ 3: Knowledge Distillation (การกลั่นกรองความรู้) เพื่อให้โมเดลรักษาขอบเขตการตัดสินใจของคลาสเก่า งานวิจัยนี้ใช้ Label Blending ก่อนทุกขั้นตอนของการฝึกแบบเพิ่มพูน โมเดลเก่า $M_{old}$ สร้าง Soft Probability Vector $Y_{soft}$ ให้กับข้อมูลที่ผสม (ข้อมูลใหม่ + Exemplar ที่ Replay) เป้าหมายการฝึกสอนคือการผสมถ่วงน้ำหนักของ Hard Label $Y_{true}$ กับ $Y_{soft}$ ตามสมการที่ (5):

$$Y_{fit} = \alpha\, Y_{true} + (1 - \alpha)\, Y_{soft} \quad (5)$$

ค่า $\alpha$ = 0.5 ควบคุมสมดุลระหว่าง Hard Label $Y_{true}$ และ Posterior ของโมเดลเก่า $Y_{soft}$ โดย $\alpha$ สูงจะให้น้ำหนักกับ Ground Truth มากขึ้น ขณะที่ $\alpha$ ต่ำจะให้น้ำหนักกับ $Y_{soft}$ มากขึ้น $Y_{soft}$ เข้ารหัสคะแนนความคล้ายคลึงสัมพัทธ์ที่ $M_{old}$ กำหนดให้กับคลาสเก่า และการผนวก $Y_{soft}$ เข้ากับ Loss Function ลงโทษการเบี่ยงเบนอย่างมากจากคะแนนเหล่านั้น จึงยึดพฤติกรรมของโมเดลที่ปรับปรุงแล้วบนคลาสที่เรียนรู้ก่อนหน้า

**3.4 การตรวจจับจุดเชื่อมต่อปลอมแบบ 2 ด่าน**

เมื่อได้ Heatmap ที่ผ่านการสร้างใหม่ $\hat{H}$ จาก Auto-NRK ของ AP ที่ไม่รู้จัก ขั้นตอนต่อไปคือการพิจารณาว่า Query นั้นมาจาก AP ที่ลงทะเบียนไว้ในระบบหรือเป็น Rogue AP ระบบต้องสามารถรับมือกับ RAP ได้สองลักษณะ ลักษณะแรกคือ Impostor ที่ระบบรู้จัก Identity อยู่แล้ว คืออุปกรณ์ที่มีลายนิ้วมือแตกต่างจาก AP ที่ลงทะเบียนทุกเครื่อง แต่คลาสของมันเคยปรากฏในชุดข้อมูลฝึกสอน ลักษณะที่สองคือ Zero-day RAP ซึ่งเป็นอุปกรณ์ที่ไม่เคยปรากฏในการฝึกสอนใด ๆ เลย กลไกการตรวจจับสองด่านนี้ออกแบบมาเพื่อรองรับ RAP ทั้งสองลักษณะโดยไม่จำเป็นต้องฝึกสอนแยกกัน โดยมีรายละเอียดดังนี้คือ

- ด่านที่ 1: การตรวจสอบขอบเขตค่าความผิดพลาดการสร้างภาพใหม่ โดยค่าความผิดพลาดการสร้างภาพใหม่ของ Query Heatmap ที่ได้จาก Auto-NRK สะท้อนให้เห็นว่าโครงข่ายสามารถอธิบายลายนิ้วมือที่สังเกตได้อย่างมีประสิทธิภาพหรือไม่ โดยทั่วไป AP ที่ลงทะเบียนไว้อย่างถูกต้องจะถูกสร้างใหม่ได้ด้วยค่าความผิดพลาดที่ต่ำ แต่ RAP ที่มีลักษณะแตกต่างออกไปจะให้ค่าความผิดพลาดสูงกว่าอย่างเห็นได้ชัด

  อย่างไรก็ตาม ค่าความผิดพลาดที่ต่ำผิดปกติอาจไม่ได้บอกได้ว่าเป็น AP จริงเสมอไป เพราะอาจเกิดจาก Input ที่มีความซับซ้อนต่ำ เช่น Heatmap ที่สม่ำเสมอกันจนเกินไป จนโครงข่ายสร้างใหม่ได้ง่ายโดยไม่มีลายนิ้วมือที่มีความหมาย ด้วยเหตุนี้จึงกำหนดช่วงค่าความผิดพลาดที่ยอมรับได้โดยอ้างอิงจากเปอร์เซ็นไทล์ที่ 1 และ 99 ของค่า $E_{recon}$ ที่วัดบนชุด Calibration ของ AP ที่ถูกต้อง ค่า $E_{recon}$ คำนวณตามสมการที่ (1)

  Query ใดที่มีค่าความผิดพลาดอยู่นอกช่วง $]\tau_{lower}, \tau_{upper}]$ จะถูกจัดประเภทเป็น RAP ทันที โดยไม่ผ่านไปยัง Stage ที่ 2 ดังสมการที่ (6):

$$(E_{recon} < \tau_{lower}) \vee (E_{recon} > \tau_{upper}) \quad (6)$$

- ด่านที่ 2: การตรวจสอบความไม่แน่นอนผ่าน Shannon Entropy โดย Query ที่ผ่านด่านแรกจะถูกส่งเข้า Self-ACC ซึ่งให้ผลลัพธ์เป็นการกระจายความน่าจะเป็น $P(y = k \mid \hat{H})$ สำหรับ $K$ คลาสที่ลงทะเบียน ถ้า Query ตรงกับ AP ที่ลงทะเบียน ความน่าจะเป็นของคลาสหนึ่งควรโดดเด่น ให้การกระจายตัว Entropy ต่ำ ถ้า Query เป็น RAP ที่ซับซ้อนซึ่งลายนิ้วมือคล้ายคลึงกับคลาสที่รู้จักอย่างผิวเผินแต่ต่างในรายละเอียด Self-ACC จะกระจายมวลความน่าจะเป็นอย่างกว้าง ทำให้ Entropy สูง งานวิจัยนี้วัดความไม่แน่นอนนี้ด้วย Shannon Entropy ดังสมการที่ (7):

$$H(\hat{H}) = -\sum_{k=1}^{K} P(y = k \mid \hat{H}) \times \log_2(P(y = k \mid \hat{H}) + \varepsilon) \quad (7)$$

  โดยที่ $\varepsilon = 10^{-6}$ ป้องกันความไม่เสถียรเชิงตัวเลขเมื่อ $P(y=k|\hat{H}) \approx 0$ ค่าเกณฑ์ Entropy $\tau_{entropy}$ ถูกกำหนดที่เปอร์เซ็นไทล์ที่ 90 ของค่า Entropy ที่คำนวณบนชุด Calibration ของ AP ที่ถูกต้อง ซึ่งหมายความว่าระบบยอมรับความแปรปรวนสัญญาณ 90% ของการทำงานปกติ Query จะถูกระบุเป็น Rogue AP ถ้าตรงกับเงื่อนไขใดเงื่อนไขหนึ่งดังสมการที่ (8):

$$(H(\hat{H}) > \tau_{entropy}) \vee (\hat{c} \neq c) \quad (8)$$

  โดยที่ $c$ คือคลาสเป้าหมายที่ผู้ใช้ต้องการเชื่อมต่อ และ $\hat{c}$ คือคลาสที่ Self-ACC ทำนายได้ การใช้ตัวเชื่อมตรรกะแบบหรือ (Disjunction) ในเงื่อนไขดังกล่าวรองรับสองสถานการณ์ภัยคุกคามได้แก่ (1) AP ที่มีค่า Entropy สูง ซึ่งบ่งบอกถึงความไม่แน่นอนในทุกคลาสและอาจเป็น Zero-day RAP และ (2) AP ที่ถูกระบุว่าเป็นคลาสอื่น ซึ่งสอดคล้องกับ Impostor ที่ระบบเคยเห็นมาก่อน

**3.5 Pseudocode ของอัลกอริทึม**

อัลกอริทึมนี้อธิบายกระบวนการเรียนรู้อุปกรณ์ใหม่แบบเพิ่มพูนคลาสโดยไม่เกิด Catastrophic Forgetting โดยเริ่มจากการสร้าง Exemplar Memory สำหรับแต่ละคลาสใหม่ในชุดข้อมูล $D_{new}$ ด้วยการคำนวณ Centroid

$\mu_c$ และคัดเลือก $m$ ตัวอย่างที่ใกล้ศูนย์กลางคลาสมากที่สุด (บรรทัดที่ 1–6) จากนั้นสร้างชุดข้อมูลผสม $D_{mix}$ โดยรวมข้อมูลใหม่กับ Exemplar ของคลาสเก่าทั้งหมด

---

**Algorithm 1: Class-Incremental Learning with Exemplar Replay**

---

Input: new task dataset $D_{new}$, previous model Mold, exemplar memory $M_{memory}$, exemplar size per class m
Output: updated model $M_{new}$
1: for each class c in $D_{new}$ do
2: Compute the class centroid $\mu_c$
3: Select the top-m samples with the lowest MSE to $\mu_c$
4: Store the selected samples in exemplar set $M_c$
5: end for
6: Update Memory by including all new exemplar sets $M_c$
7: Construct the mixed training set
$D_{mix} = D_{new} \cup M_{memory}$
8: Apply Gaussian noise augmentation to all replayed exemplars with $\mu = 0$ and $\sigma = 0.05$
9: Generate soft targets
$Y_{soft} = M_{old}(D_{mix})$
10: Compute the blended targets
$Y_{fit} = 0.5Y_{true} + 0.5Y_{soft}$
11: Train Self-ACC on $D_{mix}$ using cross-entropy loss with $Y_{fit}$ as supervision
12: Return the updated model $M_{new}$

---

---

**Algorithm 2: Dual-Stage Rogue AP Detection**

---

Input: query heatmap X, thresholds $\tau_{lower}$, $\tau_{upper}$, $\tau_{entropy}$, target class c
Output: classification result {Legitimate AP, Rogue AP}
1: Reconstruct the query heatmap by using Auto-NRK $\hat{H}$ = Auto-NRK(X)
2: Compute the reconstruction error
$E_{recon} = MSE(X, \hat{H})$
3: if $E_{recon} < \tau_{lower}$ or $E_{recon} > \tau_{upper}$ then
4: return Rogue AP
5: end if
6: Compute the class probability distribution
$P(y = k \mid \hat{H})$ for all classes through Self-ACC
7: Compute the entropy score

$$H(\hat{H}) = -\sum_{k=1}^{K} P(y = k \mid \hat{H}) \times \log_2(P(y = k \mid \hat{H}) + \varepsilon)$$

8: Compute the predicted class
$\hat{c} = \text{argmax}_k\ P(y = k \mid \hat{H})$
9: if $H(\hat{H}) > \tau_{entropy}$ or $\hat{c} \neq c$ then
10: return Rogue AP
11: else
12: return Legitimate AP
13: end if

---

นอกจากนี้มีการเพิ่ม Gaussian Noise ($\mu = 0, \sigma = 0.05$) เพื่อป้องกัน Overfitting (บรรทัดที่ 7–8) ในขั้นตอนสุดท้าย โมเดลเก่า $M_{old}$ ถูกใช้สร้าง Soft Target $Y_{soft}$ แล้วผสมกับ Hard Label $Y_{true}$ ด้วยสัดส่วน 0.5:0.5 เพื่อให้ได้ Blended Target $Y_{fit}$ ที่ใช้ฝึกสอน Self-ACC (บรรทัดที่ 9–11) ทำให้โมเดลที่ได้ $M_{new}$ มีความรู้ทั้งคลาสเก่าและใหม่พร้อมกัน

อัลกอริทึมนี้อธิบายกระบวนการตรวจจับ Rogue AP แบบ 2 ด่านโดยใช้ Query Heatmap $X$ เป็นอินพุตด่านแรก (บรรทัดที่ 1–5) ส่ง $X$ เข้า Auto-NRK เพื่อสร้าง Heatmap ที่กำจัดสัญญาณรบกวนแล้ว $\hat{H}$ จากนั้นคำนวณค่า Reconstruction Error $E_{recon}$ หาก $E_{recon}$ อยู่นอกช่วง $[\tau_{lower}, \tau_{upper}]$ ระบบจะตัดสินทันทีว่าเป็น Rogue AP โดยไม่ต้องผ่านด่านที่สอง ด่านที่สอง (บรรทัดที่ 6–13) Query ที่ผ่านด่านแรกจะถูกส่งเข้า Self-ACC เพื่อคำนวณการกระจายความน่าจะเป็น $\mathbf{P}(y = k \mid \hat{H})$ และค่า Shannon Entropy $H(\hat{H})$ หาก Entropy สูงเกิน $\tau_{entropy}$ หรือคลาสที่ทำนาย $\hat{c}$ ไม่ตรงกับคลาสเป้าหมาย c ระบบจะระบุว่าเป็น Rogue AP มิฉะนั้นจะจำแนกเป็น Legitimate AP

### 3.6 รายละเอียดการพัฒนา

โมเดลทั้งหมดพัฒนาด้วย Python 3.10 และ TensorFlow 2.12 / Keras ทั้ง Auto-NRK และ Self-ACC ฝึกสอนด้วย Adam Optimizer (Learning Rate = $1 \times 10^{-4}$, $\beta 1 = 0.9$, $\beta 2 = 0.999$) และ Batch Size 32 Auto-NRK ถูก Pre-train ครั้งเดียวบนข้อมูล AP ที่ถูกต้องทั้งหมดก่อนที่งานเพิ่มพูนใดจะเริ่ม Self-ACC เริ่มต้นด้วย 2 คลาส AP แรกและอัปเดตแบบเพิ่มพูนที่ขอบเขตงานแต่ละรอบ เกณฑ์ Calibration-set ($\tau_{lower}, \tau_{upper}, \tau_{entropy}$) ถูกคำนวณใหม่หลังจากการอัปเดตแต่ละงานโดยใช้ Partition Calibration ที่เก็บแยกของ AP ที่ถูกต้องที่เห็นมาจนถึงปัจจุบัน ข้อมูล CSI ดิบเก็บด้วย Intel 5300 NIC และ Linux 802.11n CSI Tool

## 4. การประเมินผลและการอภิปราย

### 4.1 ชุดข้อมูลและการตั้งค่าการทดลอง

ข้อมูล CSI ดิบถูกประมวลผลล่วงหน้าเพื่อสกัดลำดับ Phase Error จากนั้นแปลงเป็น Heatmap ขนาด 128×128 โดยรวม Window แบบ Sliding ของ 100 แพ็กเก็ตต่อเนื่อง ได้ชุดข้อมูล Heatmap ทั้งหมด 4,086 ภาพ แบ่งเป็น Training Set 2,855 ภาพ และ Testing Set 1,231 ภาพ โดยแบ่งที่ระดับ Session แทน Packet เพื่อป้องกัน Data Leakage

การเลือกเปอร์เซ็นไทล์ที่ 1/99 สำหรับ $E_{recon}$ และเปอร์เซ็นไทล์ที่ 90 สำหรับ $\tau_{entropy}$ ได้รับการตรวจสอบโดยการทดลอง Sensitivity Analysis บน Validation Set แยกต่างหาก โดยทดสอบ Percentile ตั้งแต่ P80 ถึง P99 สำหรับ $\tau_{entropy}$ และพบว่า P90 ให้ค่า F1-Score สูงสุดอย่างสม่ำเสมอในทุก Task การเพิ่ม Percentile เกิน P90 ทำให้ FAR ลดลงเพียงเล็กน้อย (<2%) แต่ Recall ลดลงอย่างมีนัยสำคัญ (>10%) ซึ่งบ่งชี้ว่า P90 เป็นจุดสมดุลระหว่าง Sensitivity และ Specificity ที่เหมาะสมสำหรับสถานการณ์ IoT Security เพื่อจำลองสถานการณ์การใช้งานจริงแบบเพิ่มพูน 11 คลาสถูกแบ่งเป็น 5 งานตามลำดับ

ทั้งนี้แต่ละงานแนะนำคลาส AP ใหม่ 2 คลาส ยกเว้นงานที่ 5 ซึ่งแนะนำคลาสที่เหลือ ในแต่ละงาน จะมีการเลือกหนึ่งคลาสเป็น held-out class ซึ่งจะถูกตัดออกจากกระบวนการฝึกสอนทั้งหมด และใช้เฉพาะในขั้นตอนทดสอบใน Zero-day RAP เท่านั้น (เช่น ใน Task 1 เลือกคลาส 2 เป็น held-out class โมเดลจะฝึกคลาสอื่นตามปกติและใช้คลาส 2 เฉพาะทดสอบแบบ open-set) เพื่อทดสอบการตรวจจับ Open-Set โปรโตคอล Leave-One-Class-Out นี้รับประกันว่า RAP ในแต่ละงานเป็น Zero-day ที่ไม่รู้จักอย่างแท้จริงสำหรับโมเดลจำแนก เกณฑ์การจำแนกจะถูก Recalibrate หลังการอัปเดตงานแต่ละรอบ

ในการทดลองนี้ ลำดับคลาสได้รับการกำหนดตายตัว 1 ลำดับตามลำดับการเก็บข้อมูล (Task 1: คลาส 1–2, Task 2: คลาส 3–4, ..., Task 5: คลาส 9–11) เนื่องจากข้อจำกัดด้านขนาดชุดข้อมูล (11 คลาส, 4,086 ภาพ) ทำให้ไม่สามารถทดสอบ Class Order ที่แตกต่างกันหลายแบบได้อย่างมีนัยสำคัญทางสถิติ งานวิจัยในอนาคตควรศึกษาความอ่อนไหวต่อ Class Order บนชุดข้อมูลที่ใหญ่กว่า ซึ่งเป็นข้อจำกัดที่ยอมรับของการศึกษานี้

### 4.2 Baseline และตัวชี้วัดการประเมินผล

ประเมิน Baseline 3 ตัว ได้แก่

1. Naive Fine-Tuning ฝึกสอน Self-ACC บนข้อมูลงานใหม่แต่ละรอบโดยไม่มี Replay หรือ Distillation ซึ่งแสดงฉากเลวร้ายที่สุดของ Catastrophic Forgetting
2. CNN + Replay ใช้โครงข่ายเดิมแต่เลือก Exemplar แบบสุ่มแทนการเลือกแบบ Centroid-based และไม่มี Knowledge Distillation เพื่อแยก Contribution ของสององค์ประกอบนั้น
3. Baseline (Dynamic) ในตารางที่ 3–5 หมายถึงวิธีที่ใช้ CNN + Replay (Random) ร่วมกับ Entropy-based Gate เดียว ซึ่งแตกต่างจาก Baseline (CNN + Replay) ในตารางที่ 2 โดยวัดเฉพาะความแม่นยำการจำแนก

ประสิทธิภาพวัดด้วย 4 ตัวชี้วัดจาก Confusion Matrix ที่มี TP, TN, FP, FN ดังนี้ ความแม่นยำ (Accuracy) วัดสัดส่วนของ Heatmap AP ที่รู้จักที่กำหนดให้คลาสถูกต้อง อัตราการตรวจจับ (Recall) วัดสัดส่วนของ Heatmap RAP ที่ระบุได้ถูกต้องว่าเป็น Rogue: Recall = TP / (TP + FN) × 100 อัตราการเตือนผิดพลาด (FAR) วัดสัดส่วนของ Heatmap AP ที่ถูกต้องที่ถูกระบุผิดว่าเป็น Rogue AP จากค่า FAR ดังสมการที่ (9):

$$FAR = \frac{FP}{FP+TN} \times 100 \quad (9)$$

และ F1-Score สร้างความสมดุลระหว่าง Precision และ Recall โดยหาค่าได้จากสมการที่ (10):

$$F1\text{-}Score = 2 \times \frac{(Precision \times Recall)}{(Precision + Recall)} \times 100 \quad (10)$$

### 4.3 ประสิทธิภาพการเรียนรู้อย่างต่อเนื่อง

**ตารางที่ 2:** ความแม่นยำในการระบุตัวตนอุปกรณ์ (%) ใน 5 งานที่เพิ่มพูนตามลำดับ

| วิธีการ | Task 1 | Task 2 | Task 3 | Task 4 | Task 5 |
|---|---|---|---|---|---|
| วิธีที่นำเสนอ (Proposed) | **90.00%** | **90.00%** | **96.67%** | **90.00%** | **92.00%** |
| Baseline (CNN + Replay) | 70.00% | 59.71% | 53.12% | 36.48% | 39.01% |
| Naive Fine-Tuning | 100.00% | 50.00% | 33.33% | 25.00% | 24.00% |

**ตารางที่ 3**: อัตราการตรวจจับ RAP (Recall %) ใน 5 งานตามลำดับ

| วิธีการ | Task 1 | Task 2 | Task 3 | Task 4 | Task 5 |
|---|---|---|---|---|---|
| วิธีที่นำเสนอ (Dynamic) | **71.91%** | **77.06%** | **68.57%** | **78.45%** | **82.69%** |
| Baseline (Dynamic) | 69.51% | 82.84% | 92.13% | 94.75% | 95.77% |

**ตารางที่ 4:** อัตราการเตือนผิดพลาด (FAR %) ใน 5 งานตามลำดับ

| วิธีการ | Task 1 | Task 2 | Task 3 | Task 4 | Task 5 |
|---|---|---|---|---|---|
| วิธีที่นำเสนอ (Dynamic) | **18.32%** | **20.13%** | **12.02%** | **17.44%** | **15.66%** |
| วิธีที่นำเสนอ (Fixed Threshold) | 15.02% | 52.85% | 59.50% | 50.23% | 48.55% |
| Baseline (Dynamic) | 31.22% | 44.01% | 50.12% | 65.45% | 62.87% |

**ตารางที่ 5:** ค่าความสมดุล F1-Score (%) ใน 5 งานตามลำดับ

| วิธีการ | Task 1 | Task 2 | Task 3 | Task 4 | Task 5 |
|---|---|---|---|---|---|
| วิธีที่นำเสนอ (Dynamic) | **71.47%** | **74.54%** | **71.01%** | **77.12%** | **80.56%** |
| วิธีที่นำเสนอ (Fixed Threshold) | 72.96% | 71.31% | 69.88% | 67.44% | 70.90% |
| Baseline (Dynamic) | 61.07% | 69.93% | 75.54% | 73.04% | 75.06% |

ตารางที่ 2 รายงานความแม่นยำการระบุตัวตนอุปกรณ์ใน Tasks 1-5 Naïve Fine-Tuning เริ่มที่ 100% ความแม่นยำใน Task 1 (เพียง 2 คลาส ไม่มีการลืม) แต่ลดลงเหลือ 24.00% ใน Task 5 เพราะการอัปเดต Gradient ที่ Optimize สำหรับงานใหม่แต่ละรอบเขียนทับ Representation ของคลาสก่อนหน้าโดยไม่มีกลไกชดเชย Baseline CNN + Replay บรรเทา Forgetting บางส่วน ได้ 39.01% ใน Task 5 แต่การเลือก Exemplar แบบสุ่มและการไม่มี Knowledge Distillation ทำให้ประสิทธิภาพยังต่ำกว่าวิธีที่นำเสนอมา ส่วนผลใน Task 1 ของงานเราอาจต่ำกว่า Naïve Fine-Tuning เนื่องจากมีเพียงสองคลาสประโยชน์ของ replay และ distillation ยังไม่ปรากฏให้เห็นได้ชัด

ในกรอบที่นำเสนอซึ่งผสาน Centroid-based Replay, Noise Augmentation และ Knowledge Distillation รักษาความแม่นยำเหนือ 90% ในทุกงานและได้ 92.00% ใน Task 5 ยืนยันว่า Centroid-based Replay รักษาตัวแทนคลาสที่มีคุณภาพ, Noise Augmentation ป้องกัน Overfitting บน Exemplar ขนาดเล็ก และ Knowledge Distillation ยึดขอบเขตการตัดสินใจของคลาสเก่าไว้ ทั้งสามองค์ประกอบจึงร่วมกันรักษาความแม่นยำเหนือ 90% ได้ตลอดทุกงาน

**4.4 ประสิทธิภาพการตรวจจับ RAP**

ตารางที่ 3 แสดงค่า Recall ของการตรวจจับ RAP ในแต่ละงาน การลดลงของ Recall ใน Task 3 ได้ข้อสันนิษฐานว่าเป็นผลจากการเพิ่มคลาสที่มีลายนิ้วมือคล้ายคลึงกันมากขึ้น ทำให้ขอบเขตการตัดสินใจมีความซับซ้อนเพิ่มขึ้นชั่วคราว และตารางที่ 4 รายงาน FAR เมื่อพิจารณาแยกกัน Baseline (Dynamic) ได้ Recall 95.77% ใน Task 5 ซึ่งดูน่าประทับใจ อย่างไรก็ตาม ตารางที่ 4 เผยให้เห็นว่าประสิทธิภาพนี้ได้มาโดยแลกกับ FAR 62.87% หมายความว่าระบบระบุเกือบสองในสามของการเชื่อมต่อ AP ที่ถูกต้องทั้งหมดว่าเป็น Rogue นี่คือปัญหา High-Recall/High-Alarm แบบคลาสสิกของ Detector ที่ตั้งเกณฑ์แข็งเกินไป ในเครือข่าย IoT ที่ใช้งานจริง ระบบดังกล่าวจะไม่สามารถใช้งานได้เพราะจะบล็อกการเชื่อมต่อที่ถูกต้องในอัตราที่ป้องกันการทำงานปกติ

วิธีที่นำเสนอแบบ Dynamic Threshold ได้ Recall 82.69% ใน Task 5 ในขณะที่คง FAR ไว้ที่ 15.66% ซึ่งเป็นจุดทำงานที่สมดุลกว่ามาก FAR มีเสถียรภาพทั่วทุกงาน (12.02%-20.13%) แสดงให้เห็นว่า Dynamic Entropy Threshold ปรับตัวได้อย่างมีประสิทธิภาพต่อการกระจาย Classification ที่เปลี่ยนแปลงเมื่อเพิ่มคลาสใหม่ Fixed Threshold Variant ที่ใช้ Entropy Threshold ที่ Pre-calibrate ครั้งเดียวทั่วทุกงาน แสดง FAR ที่เพิ่มขึ้น (ถึง 52.85% ใน Task 2) เพราะ Threshold ที่ Pre-calibrate ไม่สัมพันธ์กันเมื่อการกระจาย Class เปลี่ยนแปลง Ablation นี้ยืนยันว่าการ Recalibrate Threshold ที่ขอบเขตงานแต่ละรอบเป็นสิ่งจำเป็นสำหรับรักษา FAR ต่ำในสถานการณ์ Class-Incremental

**4.5 ความสมดุลโดยรวม (F1-Score)**

ตารางที่ 5 รายงาน F1-Score ซึ่งเป็น Harmonic Mean ของ Precision และ Recall ที่ลงโทษระบบที่เสียสละตัวชี้วัดใดตัวหนึ่งเพื่อเพิ่มอีกตัว วิธีที่นำเสนอแบบ Dynamic Threshold แสดงแนวโน้มขึ้นอย่างสม่ำเสมอจาก 71.47% ใน Task 1 ถึง 80.56% ใน Task 5 สะท้อนว่าระบบปรับปรุง Calibration เมื่อมีข้อมูลคลาสมากขึ้นและ Threshold ถูก Refine Fixed Threshold Variant เริ่มต้นในระดับใกล้เคียงกันแต่คงที่ประมาณ 70% เพราะ FAR ที่เพิ่มขึ้นกดค่า Precision Baseline (Dynamic) ได้ 75.06% ใน Task 5 แต่ได้ Recall สูง

ทั้งนี้ได้แลกกับ FAR 62.87% เมื่อ Precision ถูกรวมเข้าใน F1-Score ข้อได้เปรียบของมันหายไป ผลลัพธ์เหล่านี้ยืนยันว่าการรวมกันของ Centroid-based Exemplar Replay, Knowledge Distillation, Reconstruction Error Gating และ Dynamic Entropy Thresholding ให้ความสมดุลการตรวจจับโดยรวมที่ดีกว่าในช่วงปลายของการเรียนรู้แบบเพิ่มพูน

**4.6 การวิเคราะห์ต้นทุนการปรับใช้งาน**

ระบบที่นำเสนอมีต้นทุนด้านหน่วยความจำหลักสองส่วน ส่วนแรกคือ Exemplar Memory ซึ่งจัดเก็บ Heatmap ขนาด 128×128 พิกเซล จำนวน 100 ภาพต่อคลาส ในรูปแบบ float32 (4 ไบต์ต่อพิกเซล) คำนวณได้ประมาณ 6.25 MiB ต่อคลาส หรือ 68.75 MiB สำหรับ 11 คลาส แต่เมื่อพิจารณาข้อจำกัดของหน่วยความจำที่สามารถใช้งานได้จริงบน AP ซึ่งต้องแบ่งให้กับระบบปฏิบัติการและกระบวนการเครือข่ายอื่น ๆ แล้ว ขนาดดังกล่าวยังถือว่าเป็นภาระต่ออุปกรณ์ที่มีทรัพยากรจำกัด

ส่วนที่สองคือน้ำหนักโมเดล (Model Weights) ของ Auto-NRK และ Self-ACC ซึ่งในปัจจุบันยังไม่ได้วัดจำนวนพารามิเตอร์อย่างเป็นทางการโดยจะวัดในอนาคต ข้อจำกัดนี้ยืนยันว่าในรูปแบบปัจจุบัน ระบบเหมาะสำหรับการนำไปใช้งานบนเซิร์ฟเวอร์ Edge Computing ที่มีหน่วยความจำ 4–8 GB มากกว่าบน AP โดยตรง แนวทางบรรเทาได้แก่ (1) การบีบอัด Heatmap ด้วย quantization 8-bit ซึ่งลด Memory เหลือ ~17.19 MiB สำหรับ 11 คลาส (2) การใช้ Generative Replay ตาม Conditional GAN เพื่อขจัดความจำเป็นในการเก็บข้อมูล Heatmap ดิบ และ (3) การตัดแต่งเชิงโครงสร้างบน Convolutional Layers ซึ่งคาดว่าสามารถลดจำนวนพารามิเตอร์ได้ 50–70% โดยไม่สูญเสีย Accuracy อย่างมีนัยสำคัญ

## 5. สรุปผลการวิจัยและงานในอนาคต

งานวิจัยนี้นำเสนอกรอบการเรียนรู้เชิงลึกแบบ Class-Incremental สำหรับตรวจจับ Rogue Access Point จากสัญญาณ CSI ของ Wi-Fi กรอบนี้แก้ไขข้อจำกัดสำคัญ 2 ประการของงานก่อนหน้า ได้แก่ การแทนค่าคุณลักษณะแบบกล่องขอบเขตไบนารีที่สูญเสียรายละเอียดการกระจายตัวของลายนิ้วมือ และความไม่สามารถลงทะเบียนคลาส AP ใหม่โดยไม่ต้องฝึกสอนโมเดลใหม่ทั้งหมด ส่วนสนับสนุนหลักของงานวิจัยนี้มี 3 ประการได้แก่ ภาพ Heatmap CSI ขนาด 128×128 พิกเซล, Auto-NRK ResNet Autoencoder สำหรับกำจัดสัญญาณรบกวน และ Self-ACC ที่ขยายด้วย Centroid-based Exemplar Replay, Gaussian Noise Augmentation และ Knowledge Distillation ทำงานร่วมกันเพื่อส่งมอบระบบที่สามารถเรียนรู้อย่างต่อเนื่องโดยไม่เกิด Catastrophic Forgetting

ผลการทดลองบน 4,086 Heatmap CSI จาก 11 คลาสอุปกรณ์ในสถานการณ์เพิ่มพูน 5 งาน กรอบที่นำเสนอได้ความแม่นยำการระบุตัวตนอุปกรณ์ 92.00% ใน Task สุดท้าย เทียบกับ CNN + Replay เพียง 39.01% และ Naive Fine-Tuning เพียง 24.00% สำหรับการตรวจจับ RAP กลไกกรอง 2 ด่านที่ผสาน Reconstruction Error Range Gating กับ Dynamic Shannon Entropy Thresholding ได้ F1-Score 80.56% ในขณะที่รักษา FAR ไว้ที่ประมาณ 15% ซึ่งตัดกันอย่างชัดเจนกับ FAR 62.87% ของ Baseline ที่ F1-Score 75.06% Ablation ระหว่าง Fixed กับ Dynamic Entropy Threshold ยืนยันว่าการ Recalibrate Threshold ตามงานเป็นสิ่งจำเป็นสำหรับรักษา Precision การตรวจจับขณะที่การกระจาย Class พัฒนาขึ้น

สำหรับงานในอนาคต เช่น (1) Class-Incremental Learning แบบไม่ต้องเก็บตัวอย่าง (Exemplar-Free CIL) ในปัจจุบันระบบต้องเก็บ Heatmap ตัวอย่าง 100 ภาพต่อคลาส ซึ่งสร้างค่าใช้จ่ายหน่วยความจำที่พอสังเกตได้บนอุปกรณ์ Edge ที่มีทรัพยากรจำกัด การใช้ Generative Replay ซึ่ง Conditional GAN สร้าง Exemplar สังเคราะห์ตามต้องการ จะขจัดความจำเป็นในการเก็บข้อมูลดิบไว้เลย และน่าสนใจเป็นพิเศษในการ Deploy ที่อยู่ภายใต้กฎระเบียบการเก็บข้อมูล (2) การ Optimization เชิงสถาปัตยกรรม (Model Compression) จำนวน Parameter รวมของ Auto-NRK และ Self-ACC อาจเกิน Budget การ Inference ของ Router Wi-Fi ระดับพาณิชย์ทั่วไป Structured Pruning และ Post-training Quantization ที่กำหนดเป้าหมายที่ชั้น Convolutional เป็นเส้นทาง Compression ที่มีแนวโน้มดีที่สุดในการเปิดใช้งานการประมวลผล Real-time บนฮาร์ดแวร์ Embedded โดยไม่สูญเสียความแม่นยำอย่างมีนัยสำคัญ

## เอกสารอ้างอิง

[1] Z. Yang, Q. Lu, H. Zhang, F. Chen, and H. Xian, “Eliminating Rogue Access Point Attacks in IoT: A Deep Learning Approach With Physical-Layer Feature Purification and Device Identification,” *IEEE Internet of Things Journal*, vol. 11, no. 8, pp. 14886–14900, 2024.

[2] R. Kong and H. Chen, “DeepCRF: Deep Learning-Enhanced CSI-Based RF Fingerprinting for Channel-Resilient WiFi Device Identification,” *IEEE Transactions on Information Forensics and Security*, vol. 20, pp. 264–278, 2025.

[3] L. Sun, R. Xue, H. Zha, Q. Tian, W. Wang, and Y. Lin, “AFD-IL: A Long-Term Incremental Learning Approach With Adaptive Feature Distillation for Specific Emitter Identification,” *IEEE Transactions on Cognitive Communications and Networking*, early access, 2025.

[4] H. Zhang et al., “P-A Scheme: A Robust and Lightweight Wi-Fi Device Identification Approach for Enhancing Industrial Security,” *Electronics, vol. 14, no. 3, art. no. 513, 2025, doi: 10.3390/electronics14030513*.

[5] A. L. de S. Paula, E. K. Viegas, and A. O. Santin, “Passive Identification of Rogue Industrial Access Points Using a One-Class Machine Learning Model,” in *Proceedings of IECON 2024 - 50th Annual Conference of the IEEE Industrial Electronics Society*, pp. 1–6, 2024.

[6] R. Wang, Z. Li, H. Luo, F. Zhao, W. Shao, and Q. Wang, "A Robust Wi-Fi Fingerprint Positioning Algorithm Using Stacked Denoising Autoencoder and Multi-Layer Perceptron," *Remote Sensing (MDPI)*, vol. 11, no. 11, p. 1293, 2019.

[7] D. Arisandi, N. M. Ahmad, and S. Kannan, "Invisible Scout: A Layer 2 Anomaly System for Detecting Rogue Access Point (RAP)," *Emerging Science Journal*, vol. 9, no. 1, pp. 284–310, 2025.

[8] Z. Hong, J. Lin, T. Li, J. Wan, J. Su, and Z. Wen, "Open-Set Domain Adaptive RF Fingerprint Identification Based on Prototype Calibration," *IEEE Wireless Communications Letters*, vol. 15, pp. 775–779, 2026.

[9] R. Kong and H. Chen, "Towards Channel-Resilient CSI-Based RF Fingerprinting using Deep Learning," in Proceedings of *the IEEE INFOCOM Workshops*, 2024.

[10] F. Shi, H. Wan, Z. Feng, X. Fu, Q. Wang, Q. Xuan, Y. Lin, and G. Gui, "Enhanced Radio Frequency Fingerprint Identification Using Length-Robust Representation and Incremental Learning," *IEEE Internet of Things Journal*, vol. 12, no. 10, pp. 14709–14719, 2025.

[11] L. Xie, L. Peng, J. Zhang, and A. Hu, "Radio Frequency Fingerprint Identification for Internet of Things: A Survey," *Security and Safety*, vol. 3, p. 2023022, 2024.

# การเปรียบเทียบสถาปัตยกรรม CNN สมัยใหม่สำหรับการรู้จำตัวอักษรและจังหวัดในป้ายทะเบียนรถยนต์ไทยโดยใช้การตรวจจับจาก YOLOv8n

# Benchmarking Modern CNN Architectures for Thai License Plate Character and Province Recognition Using YOLOv8n-Based Detection

*ณัฐพล บุญแดง (Nattapon Boondang) และอำนาจ ขาวเน (Amnach Khawne)*
*วิศวกรรมไฟฟ้าและคอมพิวเตอร์ วิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง*
*Nattapon.prasertwaree@gmail.com, amnach.kh@kmitl.ac.th*

## บทคัดย่อ

*บทความนี้นำเสนอการเปรียบเทียบสถาปัตยกรรม CNN สมัยใหม่สำหรับการรู้จำตัวอักษรและจังหวัดในป้ายทะเบียนรถยนต์ไทย โดยใช้กรอบสองขั้นตอน ได้แก่ 1) การตรวจจับป้ายด้วย YOLOv8n (mAP@0.5 = 0.995) และ 2) การเปรียบเทียบสถาปัตยกรรมสมัยใหม่ 6 แบบ (EfficientNet-B3, EfficientNetV2-S, ConvNeXtV2-T, Swin-T, FastViT-T, MobileNetV4) กับสถาปัตยกรรมพื้นฐาน 2 แบบ (ResNet-50, MobileNetV3) บนชุดข้อมูล 94,550 ภาพ ภายใต้สภาวะแสง 3 รูปแบบ (Day, Night, Blur) โดยใช้ optimizer ที่เหมาะสมตามประเภทสถาปัตยกรรม (Adam และ AdamW) ผลการทดลองพบว่า EfficientNet-B3 ให้ความแม่นยำสูงสุดในการรู้จำตัวอักษรที่ 95.82% EfficientNetV2-S ให้ความแม่นยำสูงสุดในการรู้จำจังหวัดที่ 99.93% และ FastViT-T แสดงสมดุลที่ดีที่สุดด้วยความแม่นยำจังหวัด 99.90% ในขนาดโมเดลเพียง 12.86 MB ผลลัพธ์แสดงให้เห็นถึงข้อได้เปรียบของสถาปัตยกรรมสมัยใหม่เหนืองานวิจัยอ้างอิงอย่างมีนัยสำคัญ*

**คำสำคัญ:** การรู้จำป้ายทะเบียนรถยนต์ไทย YOLOv8 โครงข่ายประสาทเทียมคอนโวลูชัน การเรียนรู้เชิงลึก วิชันทรานส์ฟอร์เมอร์

## Abstract

*This paper presents a benchmarking study of modern CNN architectures for Thai license plate character and province recognition. A two-stage framework is employed: 1) plate detection using YOLOv8n (mAP@0.5 = 0.995), and 2) systematic comparison of six modern architectures (EfficientNet-B3, EfficientNetV2-S, ConvNeXtV2-T, Swin-T, FastViT-T, MobileNetV4) against two baselines (ResNet-50, MobileNetV3) on 94,550 images under three lighting conditions (Day, Night, Blur), utilizing architecture-specific optimizers (Adam and AdamW). Results show EfficientNet-B3 achieved the highest character recognition accuracy at 95.82%, EfficientNetV2-S attained peak province recognition accuracy at 99.93%, and FastViT-T exhibited the optimal accuracy-efficiency trade-off at 99.90% province accuracy with only 12.86 MB model size, demonstrating significant advantages of modern architectures over existing Thai license plate recognition approaches.*

**Keywords:** Thai License Plate Recognition, YOLOv8, Deep Learning, Convolutional Neural Networks, Vision Transformer.

## 1. บทนำ

ระบบรู้จำป้ายทะเบียนรถยนต์อัตโนมัติ )ALPR) ในประเทศไทยเผชิญความท้าทายเฉพาะตัวที่แตกต่างจากป้ายทะเบียนตะวันตก กล่าวคือ ความคล้ายคลึงของพยัญชนะ (เช่น ข-ช, ฎ-ฏ) และความซับซ้อนของชื่อจังหวัดที่มีถึง 77 จังหวัด ซึ่งงานวิจัยส่วนใหญ่ในอดีตมักมุ่งเน้นเพียงการอ่านเลขทะเบียนและหมวดอักษร โดยละเลยการอ่านชื่อจังหวัด หรือใช้เทคนิค OCR แบบดั้งเดิมที่ไม่ทนทานต่อ

สภาวะแสงน้อยและภาพเบลอ งานวิจัยนี้พัฒนาต่อยอดจากงานของ Rattanakornphan [1] ซึ่งใช้ YOLOv5 และ CNN แบบดั้งเดิม โดยงานวิจัยนี้มีข้อเสนอแนะเชิงปรับปรุง )Contribution) ที่สำคัญ 3 ประการ คือ 1) **อัปเกรดกระบวนการตรวจจับ:** เปลี่ยนจาก YOLOv5 เป็น **YOLOv8n** ซึ่งมีโครงสร้าง Anchor-free และ Decoupled Head ทำให้ประมวลผลได้รวดเร็วกว่าและแม่นยำกว่าในการตัดแยกป้ายทะเบียน 2) **ขยายขอบเขตการรู้จำ:** เพิ่มโมดูลการรู้จำ **"จังหวัด"** ซึ่งเป็นส่วนที่ขาดหายไปในงานวิจัยอ้างอิงหลายฉบับ เพื่อให้ระบบสมบูรณ์พร้อมใช้งานจริง 3) **เปรียบเทียบสถาปัตยกรรมสมัยใหม่:** นำเสนอการใช้โมเดลกลุ่ม Transformer (Swin-T, FastViT) และ Modern CNN (ConvNeXtV2) มาเปรียบเทียบกับโมเดลพื้นฐาน เพื่อหาจุดสมดุลที่ดีที่สุดระหว่างความแม่นยำและทรัพยากร

งานวิจัยที่ผ่านมายังขาดการประเมิน Province Recognition ร่วมกับ Character Recognition ภายใต้สภาวะแสง 3 รูปแบบ (Day, Night, Blur) ในกรอบเดียวกัน รวมถึงยังไม่มีการนำ Hybrid CNN-Transformer อย่าง FastViT มาประเมินในบริบทของ Thai LPR ซึ่งเป็นช่องว่างที่งานวิจัยนี้มุ่งเติมเต็ม

## 2. วรรณกรรมที่เกี่ยวข้อง

### 2.1 การตรวจจับวัตถุและวิวัฒนาการของ YOLO

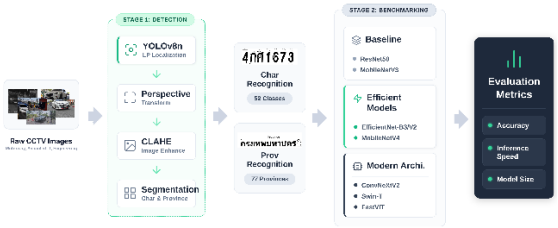
YOLO เป็นอัลกอริทึม One-stage Detector ที่โดดเด่นด้านความเร็ว Real-time [4] งานวิจัย [1, 2] ใช้ YOLOv5 อย่างไรก็ตาม YOLOv8 [5] ปรับปรุงเป็น Anchor-free พร้อม Decoupled Head ทำให้ตรวจจับวัตถุขนาดเล็กได้แม่นยำยิ่งขึ้น งานวิจัยนี้จึงเลือกใช้ YOLOv8n

### 2.2 สถาปัตยกรรมโครงข่ายประสาทเทียม )CNN Architectures)

งานวิจัยนี้ศึกษาเปรียบเทียบสถาปัตยกรรม 3 ยุคสมัย ยุคพื้นฐาน ประกอบด้วย ResNet [6] ที่ใช้ Skip Connection แก้ปัญหา Vanishing Gradient และ MobileNetV3 [7] ที่เน้นความเร็วสำหรับอุปกรณ์พกพา ยุคประสิทธิภาพสูง ประกอบด้วย EfficientNet [8] และ EfficientNetV2 [9] ที่ใช้เทคนิค Compound Scaling เพื่อสร้างสมดุลระหว่างความแม่นยำและการใช้ทรัพยากร ยุคสมัยใหม่ ประกอบด้วย ConvNeXtV2 [10] ที่เป็น Pure ConvNet ที่มีประสิทธิภาพเทียบเท่า Transformer, Swin Transformer [11] ที่ใช้กลไก Shifted Window Attention เพื่อจับความสัมพันธ์ของฟีเจอร์ได้มากกว่า CNN ดั้งเดิม, FastViT [12] ที่เป็น Hybrid CNN-Transformer พร้อม RepMixer สำหรับเร่งความเร็ว Inference, และ MobileNetV4 [13] ที่เป็น Mobile-optimized Architecture รุ่นล่าสุด

**2.3 งานวิจัยที่เกี่ยวข้องกับการรู้จำป้ายทะเบียน )Related Works)** จากการศึกษางานวิจัยในอดีตพบข้อจำกัดสำคัญ ได้แก่ ระบบในงานวิจัย ]1] ยังขาดความแม่นยำในการแยกตัวอักษรที่คล้ายคลึงกัน (เช่น ข-ช, บ-ป) ส่วนการใช้ EasyOCR ในงานวิจัย ]2] มีข้อจำกัดด้านการปรับแต่งฟอนต์เฉพาะทางและความเร็วที่ช้า และงานวิจัย ]3] ที่ใช้ Transformer พบความแม่นยำลดลงต่ำกว่า 91% ในสภาวะแสงน้อย เนื่องจาก Self-Attention Mechanism มีความไวต่อความผิดเบี้ยวของ Feature จากสภาวะ Low-light งานวิจัยนี้จึงมุ่งปิดช่องว่างดังกล่าวผ่านระเบียบวิธีที่นำเสนอในหัวข้อ 3

## 3. วิธีการดำเนินการวิจัย




**ภาพที่ 1:** กรอบแนวคิดการดำเนินงานวิจัย

### 3.1 การจัดเตรียมชุดข้อมูล )Dataset Preparation)

งานวิจัยนี้รวบรวมภาพจากกล้อง CCTV จำนวน 200,000 ภาพ จาก 3 แหล่ง หลังผ่านกระบวนการ Detection, Segmentation และ Quality Filtering คัดเลือกได้ 94,550 ภาพ รายละเอียดการแบ่งชุดข้อมูลแสดงในตารางที่ 1

**3.2 การตรวจจับและเตรียมภาพ )Detection and Preprocessing)**

**1) License Plate Detection:** ใช้โมเดล YOLOv8n (Nano) ซึ่งเป็นรุ่นที่มีความเร็วสูงสุดในตระกูล YOLOv8 ทำการฝึกฝนด้วยภาพป้ายทะเบียนจำนวน 3,000 ภาพ คัดเลือกให้สมดุลระหว่างกลางวัน/กลางคืน และมุมกล้องซ้าย/ขวา

**2) Perspective Transformation:** เนื่องจากภาพจากกล้อง CCTV มักมีมุมมองที่เอียง งานวิจัยนี้จึงประยุกต์ใช้ Perspective Transform โดยอาศัยจุดมุม 4 จุดของกรอบป้ายทะเบียน เพื่อดัดภาพป้ายให้กลับมาอยู่ในระนาบตรง )Frontal View) ลดความผิดเพี้ยนของรูปทรงตัวอักษร



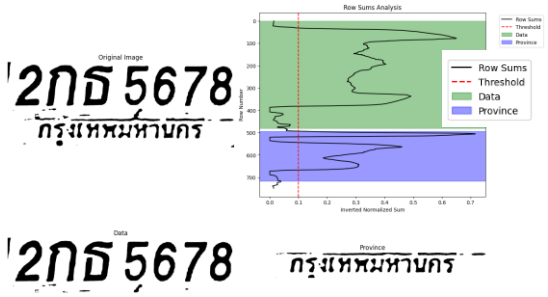
**ภาพที่ 2:** แสดงภาพป้ายทะเบียนที่ผ่านกระบวนการ Perspective Transformation

**3) การปรับปรุงความคมชัด )Image Enhancement):** สำหรับภาพในเวลากลางคืนหรือภาพที่มีแสงจ้าสะท้อน งานวิจัยนี้เลือกใช้เทคนิค CLAHE เพื่อปรับกระจายค่าความเข้มแสงเฉพาะจุด ซึ่งช่วยเน้นขอบของตัวอักษรให้ชัดเจนขึ้นโดยไม่เพิ่ม Noise ในส่วนพื้นหลัง



**ภาพที่ 3:** แสดงภาพหลังปรับปรุงคุณภาพภาพ

**4 ) Character & Province Segmentation:** ใช้เทคนิค Histogram Projection ในแกนแนวนอนและแนวตั้ง เพื่อตัดแยกภาพออกเป็นส่วนย่อย ได้แก่ พยัญชนะ แถวบน ตัวเลข และชื่อจังหวัด โดยมีการจัดการรวมส่วนประกอบของตัวอักษรที่อาจแยกจากกันให้เป็นภาพพยัญชนะที่สมบูรณ์




**ภาพที่ 4:** แสดงการแบ่งแยกส่วนแถวบนแถวล่าง

**3.3 ชุดข้อมูลสำหรับทดสอบ**

**ตารางที่ 1:** สรุปการแบ่งชุดข้อมูลสำหรับการฝึกสอน ตรวจสอบ และทดสอบ

| ขั้นตอน (Task) | ชุดข้อมูลทั้งหมด (Total) | Train (80% of remain) | Val (20% of remain) | Test Set (แยกอิสระ) |
|---|---|---|---|---|
| 1.ตรวจจับป้ายทะเบียน (LP Detection) | 3,000 | 2,400 | 600 | (ใช้ผลจาก Val) |
| 2. รู้จำตัวอักษร (Character) | 67,600 | 41,600 | 10,400 | 15,600 (100/คลาส) |
| 3. รู้จำชื่อจังหวัด (Province) | 26,950 | 12,320 | 3,080 | 11,550 (50/คลาส) |
| รวมภาพที่ใช้จริงในเล่ม | 94,550 | 56,320 | 14,080 | 27,150 |

ภาพทั้งหมดถูกปรับขนาดเป็น 224×224 พิกเซล และทำ Normalization ด้วยค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของ ImageNet เพื่อป้องกัน Overfitting จึงใช้เทคนิค Data Augmentation ระหว่างการฝึกสอน ได้แก่ Random Rotation (±5°), Translation (±5%), Scaling (95–105%), Color Jitter (±30%) และ Gaussian Blur ($\sigma = 0.1–1.5$)

**3.4 สถาปัตยกรรมโมเดลที่ใช้เปรียบเทียบ )Model Architectures)**

งานวิจัยนี้คัดเลือกสถาปัตยกรรม 8 โมเดล แบ่งเป็น 3 กลุ่ม ได้แก่ (1) กลุ่มพื้นฐาน (ResNet-50, MobileNetV3-Large) (2) กลุ่มประสิทธิภาพสูง (EfficientNet-B3, EfficientNetV2-S, MobileNetV4) และ (3) กลุ่มสมัยใหม่ (ConvNeXtV2-Tiny, Swin-T, FastViT-T8) ดังรายละเอียดในหัวข้อ 2.2

**3.5 การตั้งค่าการฝึกสอน )Training Configuration)**

การฝึกสอนโมเดลแบ่งกลยุทธ์ตามประเภทสถาปัตยกรรม ดังนี้ 1) สำหรับโมเดลโครงสร้าง CNN (ResNet-50, MobileNetV3, EfficientNet-B3, EfficientNetV2-S, MobileNetV4): กำหนดค่า Optimizer เป็น Adam พร้อม Learning Rate $1\times10^{-3}$ เนื่องจากโครงสร้างมีความเสถียรในการ Convergence ได้ดี 2) สำหรับโมเดลสถาปัตยกรรมสมัยใหม่

) ConvNeXtV2-T, Swin-T, FastViT-T): กำหนดค่า Optimizer เป็น AdamW [14] พร้อม Weight Decay 0.05 และ Learning Rate $5\times10^{-5}$ เพื่อลดความเสี่ยงจาก Overfitting และให้โมเดลเรียนรู้ Fine-grained Features ของตัวอักษรไทยได้แม่นยำ

การกำหนด Optimizer ที่แตกต่างกันระหว่างกลุ่มสถาปัตยกรรมเป็นแนวปฏิบัติมาตรฐานใน Benchmarking [10, 11] เพื่อให้สอดคล้องกับ Fine-tuning Protocol ของแต่ละโมเดล ซึ่งสะท้อนประสิทธิภาพที่แท้จริงของแต่ละสถาปัตยกรรมได้อย่างยุติธรรม

## 4. ผลการดำเนินงานและการวิเคราะห์

### 4.1 ผลการตรวจจับป้ายทะเบียน )Detection Results)

โมเดล YOLOv8n ให้ค่า mAP@0.5 = 0.995 ซึ่งสูงกว่างาน Rattanakornphan [1] ที่ใช้ YOLOv5 (Recall 98.4%) อย่างมีนัยสำคัญ




**ภาพที่ 5: แสดงผลตรวจจับตำแหน่งป้ายทะเบียน**

### 4.2 ผลการรู้จำตัวอักษรและจังหวัด

#### 4.2.1 ผลการรู้จำตัวอักษร )Character Recognition)

**ตารางที่ 2:** แสดงผลการเปรียบเทียบความแม่นยำ )Accuracy) ของการรู้จำตัวอักษร โดยจำแนกตามสภาวะแวดล้อม 3 รูปแบบ

| Model | Day (%) | Night (%) | Blur (%) | Average (%) |
|---|---|---|---|---|
| EfficientNet-B3 | 96.19 | 95.02 | 96.25 | 95.82 |
| ConvNeXtV2-T | 95.08 | 94.85 | 95.13 | 95.02 |
| Swin-T | 95.40 | 93.25 | 95.54 | 94.73 |
| FastViT-T | 92.88 | 92.48 | 93.04 | 92.80 |
| EfficientNetV2-S | 94.79 | 85.98 | 94.85 | 91.87 |
| MobileNetV3 | 87.73 | 87.94 | 88.12 | 87.93 |
| ResNet50 | 86.54 | 79.58 | 86.19 | 84.10 |
| MobileNetV4 | 87.29 | 20.37 | 87.29 | 64.98 |

#### 4.2.2 ผลการรู้จำจังหวัด )Province Recognition)

**ตารางที่ 3:** แสดงผลการเปรียบเทียบความแม่นยำของการรู้จำชื่อจังหวัด (77 จังหวัด)

| Model | Day (%) | Night (%) | Blur (%) | Average (%) |
|---|---|---|---|---|
| EfficientNetV2-S | 99.92 | 99.95 | 99.92 | 99.93 |
| EfficientNet-B3 | 99.87 | 99.95 | 99.87 | 99.90 |
| FastViT-T | 99.90 | 99.90 | 99.90 | 99.90 |
| ConvNeXtV2-T | 99.82 | 99.79 | 99.84 | 99.82 |
| Swin-T | 99.82 | 99.87 | 99.77 | 99.82 |
| MobileNetV3 | 99.77 | 99.77 | 99.79 | 99.77 |
| ResNet50 | 99.71 | 99.58 | 99.69 | 99.66 |
| MobileNetV4 | 99.74 | 99.45 | 99.77 | 99.65 |

#### 4.2.3 การวิเคราะห์ประสิทธิภาพและความซับซ้อนของสถาปัตยกรรม

ทุกโมเดลถูกทดสอบบน GPU NVIDIA GeForce RTX 3060 Laptop ด้วย CUDA โดยวัด FPS จาก Warmup 50 รอบ + Measure 200 รอบ ต่อภาพ 224×224 พิกเซล พร้อม Trim 5% outliers

**ตารางที่ 4:** เปรียบเทียบขนาดโมเดล พารามิเตอร์ และความเร็ว Inference

| Model | Size (MB) | Params (M) | FPS (bs=1) |
|---|---|---|---|
| EfficientNetV2-S | 78.09 | 20.24 | 30.7 |
| EfficientNet-B3 | 41.64 | 10.78 | 38.8 |
| MobileNetV3 | 16.48 | 4.27 | 79.2 |
| ResNet50 | 90.39 | 23.61 | 77.1 |
| FastViT-T | 12.86 | 3.30 | 56.0 |
| MobileNetV4 | 32.86 | 8.50 | 60.1 |
| Swin-T | 105.41 | 27.56 | 43.1 |
| ConvNeXtV2-T | 106.53 | 27.91 | 64.2 |

### 4.3 การวิเคราะห์และอภิปรายผล (Discussion)

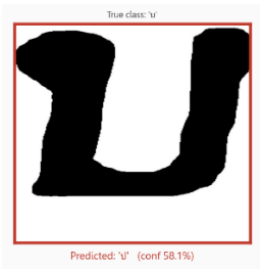
**4.3.1 การรู้จำตัวอักษร: EfficientNet-B3** ให้ความแม่นยำสูงสุด )95.82%) โดยมีความเสถียรในสภาวะแสงน้อยสูงกว่างานวิจัย ]3] อย่างชัดเจน นอกจากนี้โมเดลกลุ่ม Modern (ConvNeXtV2-T, Swin-T) ยังมีประสิทธิภาพเหนือกว่ากลุ่ม Baseline (ResNet-50) เนื่องจากกลไก Attention ช่วยในการจับลักษณะเฉพาะของตัวอักษรไทยได้ดีกว่า

เมื่อพิจารณา Day–Night Accuracy Gap พบว่า EfficientNet-B3 มีความเสถียรสูงที่สุด (Degradation 1.17%) ในขณะที่ MobileNetV4 มี Night Accuracy

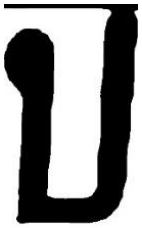
ลดลงอย่างรุนแรงถึง 66.92% (87.29% ⟶ 20.37%) แสดงว่าสถาปัตยกรรมนี้ไม่เหมาะกับสภาวะแสงน้อย

**4.3.2 การรู้จำจังหวัด (Province Recognition):** EfficientNetV2-S ให้ความแม่นยำสูงสุด )99.93%) ตามด้วย FastViT-T (99.90%) โดยผลลัพธ์ในส่วนนี้สูงกว่าการรู้จำตัวอักษรเนื่องจากชื่อจังหวัดมีรูปแบบคำ )Pattern) ที่ชัดเจนและยาวกว่า ทำให้โมเดลจำแนกข้อมูลได้ง่ายและแม่นยำยิ่งขึ้น

**4.3.3 การวิเคราะห์ความผิดพลาด (Error Analysis):** จากการตรวจสอบ Confusion Matrix พบว่าคู่ตัวอักษรที่มีอัตราการทำนายผิดพลาดสูงสุดคือกลุ่มที่มีรูปร่างคล้ายคลึงกัน ได้แก่ **"บ กับ ป"**, **"ข กับ ช"**อย่างไรก็ตาม การใช้เทคนิค Data Augmentation ประเภทการบิดเบือนภาพ )Geometric Distortion) และการปรับแสงช่วยลดความผิดพลาดเหล่านี้ลงได้เมื่อเทียบกับโมเดลพื้นฐาน




อักษร “บ” ที่โมเดลอ่านเป็น “ป”



ตัวอักษร “ป” จริง




อักษร “ข” ที่โมเดลอ่านเป็น “ช”



ตัวอักษร “ช” จริง

**ภาพที่ 6:** ตัวอักษรที่ทำนายผิดพลาดมักเป็นกลุ่มที่มีรูปร่างคล้ายคลึงกัน (เช่น บ/ป, ข/ช)

**4.4 การเปรียบเทียบกับงานวิจัยอ้างอิง )Comparison with Related Works)**

**ตารางที่ 5:** การเปรียบเทียบเทคนิคและประสิทธิภาพ "การตรวจจับ"

| งานวิจัย (Ref.) | เทคนิคการตรวจจับ | ผลการตรวจจับ |
|---|---|---|
| Rattanakornphan [1] (2022) | YOLOv5m | 98.40% (Recall) |
| Dittakan [3] (2025) | YOLOv5nu | 99.50% (mAP) |
| งานวิจัยนี้ (Proposed) | YOLOv8n | 99.50% (mAP) |

**ตารางที่ 6:** การเปรียบเทียบเทคนิคและประสิทธิภาพ **"การรู้จำ"**

| งานวิจัย (Ref.) | เทคนิคการรู้จำ | ผลการรู้จำเฉลี่ย (ตัวอักษร) | ผลการรู้จำเฉลี่ย (จังหวัด) |
|---|---|---|---|
| Rattanakornphan [1] (2022) | CNN (ResNet-50) | 98.22% | ไม่มี |
| Dittakan[3] (2025) | Transformer | 91.00% | ไม่มี |
| งานวิจัยนี้ (Proposed) | EfficientNet-B3 | 95.82% | 99.90% |
| งานวิจัยนี้ (Proposed) | FastViT-T | 92.80% | 99.90% |

จากตารางที่ 5 และ 6 งานวิจัยนี้ปรับปรุง Detection Performance จาก Recall 98.40% เป็น mAP@0.5 99.50% และเพิ่ม Province Recognition Module (77 จังหวัด) ซึ่งขาดหายไปในงานอ้างอิง แม้ Character Accuracy (95.82%) ดูน้อยกว่างาน [1] (98.22%) แต่งานวิจัยนี้ใช้ Balanced Test Set 15,600 ภาพ (100 ภาพ/คลาส) ซึ่งเข้มงวดกว่าการทดสอบด้วย 1,179 ภาพแบบ Imbalanced ของงานอ้างอิง สำหรับงาน Dittakan [3] ที่ใช้ Transformer พบ Night Accuracy ต่ำกว่า 91% ในขณะที่ EfficientNet-B3 ในงานวิจัยนี้รักษา Night Accuracy ที่ 95.02% และ FastViT-T บรรลุ Province Accuracy 99.90% ด้วยขนาดเพียง 12.86 MB เหมาะสำหรับ Edge Device




**ภาพที่ 7:** ตัวอย่างผลการรู้จำป้ายทะเบียน แสดง Bounding Box พร้อม Output ตัวอักษรและจังหวัด

## 5. สรุปผลการวิจัยและข้อเสนอแนะ

งานวิจัยนี้นำเสนอระบบ Thai LPR แบบสองขั้นตอน โดยใช้ YOLOv8n สำหรับการตรวจจับ (mAP@0.5 = 0.995) และเปรียบเทียบสถาปัตยกรรม CNN สมัยใหม่ 8 แบบสำหรับ Character และ Province Recognition จากผลการทดลองพบว่า EfficientNet-B3 ให้ความแม่นยำ Character Recognition สูงสุด และ EfficientNetV2-S ให้ความแม่นยำ Province Recognition สูงสุด โดย FastViT-T เป็นตัวเลือกที่เหมาะสมที่สุดสำหรับการติดตั้งบน Edge Device เนื่องจากให้สมดุลระหว่างความแม่นยำและขนาดโมเดล งานวิจัยในอนาคตควรศึกษา Model Quantization และระบบ End-to-End เพื่อเพิ่มประสิทธิภาพในการใช้งานจริง

## เอกสารอ้างอิง

[1] R. Rattanakornphan and P. Keeratiwintakorn, "Thai License Plate Recognition Based on YOLO Detector and CNN-Classifier," in 2022 6th International Conference on Information Technology (InCIT), 2022, pp. 1-6.

[2] V. Phon and T. Pumpoung, "Integrating YOLOv5 and EasyOCR for Robust Thai License Plate Recognition," in 2025 22nd International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), 2025.

[3] K. Dittakan, J. Thaenthong, and T. Prasomkit, "A Comparative Study on Thai License Plate Recognition: Object Detection and Transformer Learning Approaches," IEEE Access, vol. 13, 2025.

[4] J. Redmon and A. Farhadi, "YOLOv3: An Incremental Improvement," arXiv preprint arXiv:1804.02767, 2018.

[5] G. Jocher, A. Chaurasia, and J. Qiu, "Ultralytics YOLOv8," 2023. [Online]. Available: https://github.com/ultralytics/ultralytics.

[6] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 770-778.

[7] A. Howard, M. Sandler, B. Chen, W. Wang, L.-C. Chen, M. Tan, G. Chu, V. Vasudevan, Y. Zhu, R. Pang, H. Adam, and Q. Le, "Searching for MobileNetV3," in Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), 2019, pp. 1314-1324.

[8] M. Tan and Q. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," in International Conference on Machine Learning (ICML), 2019, pp. 6105-6114.

[9] M. Tan and Q. Le, "EfficientNetV2: Smaller Models and Faster Training," in International Conference on Machine Learning (ICML), 2021, pp. 10096-10106.

[10] Z. Liu, H. Mao, C.-Y. Wu, C. Feichtenhofer, T. Darrell, and S. Xie, "A ConvNet for the 2020s," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2022, pp. 11976-11986.

[11] Z. Liu, Y. Lin, Y. Cao, H. Hu, Y. Wei, Z. Zhang, S. Lin, and B. Guo, "Swin Transformer: Hierarchical Vision Transformer using Shifted Windows," in Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), 2021, pp. 10012-10022.

[12] P. Vasu, J. Gabriel, J. Zhu, O. Tuzel, and A. Ranjan, "FastViT: A Fast Hybrid Vision Transformer using Structural Reparameterization," in Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), 2023, pp. 5765-5775.

[13] D. Qin, et al., "MobileNetV4: Universal Models for Mobile Platform," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2024.

[14] I. Loshchilov and F. Hutter, "Decoupled Weight Decay Regularization," in International Conference on Learning Representations (ICLR), 2019.

# แนวทางการตัดสินใจเชิงข้อมูลเพื่อเพิ่มประสิทธิภาพการทำความสะอาดแผงเซลล์แสงอาทิตย์ในพื้นที่เกาะ

# A Data-Driven Decision-Making Framework for Optimizing Solar Panel Cleaning in Island-Based PV Systems

*ไกรวี โนศรี (Krairawee Nosri)[1] และสุภัคพงศ์ จินารัตน์ (Supakpong Jinarat)[2]*
*[1,2]วิทยาลัยวิศวกรรมศาสตร์และเทคโนโลยี สาขาวิชาเทคโนโลยี มหาวิทยาลัยธุรกิจบัณฑิตย์*
*67130098@dpu.ac.th, supakpong.jin@dpu.ac.th*

## บทคัดย่อ

*การสะสมของฝุ่นบนแผงเซลล์แสงอาทิตย์เป็นสาเหตุหนึ่งที่ทำให้ประสิทธิภาพการผลิตไฟฟ้าลดลง โดยเฉพาะในระบบบนเกาะหรือพื้นที่ห่างไกลที่มีข้อจำกัดด้านการบำรุงรักษา ขณะที่วิธีทำความสะอาดตามตารางเวลาคงที่มักไม่สอดคล้องกับสภาพการทำงานจริงของระบบ งานวิจัยนี้นำเสนอระบบการตัดสินใจเชิงข้อมูลที่คำนึงถึงสภาพอากาศ โดยใช้แบบจำลองพยากรณ์พลังงานประเมินแนวโน้มการสูญเสียจากคราบสกปรก และช่วยแยกผลกระทบด้านอุตุนิยมวิทยาออกจากการเสื่อมประสิทธิภาพของแผงได้อย่างเหมาะสม ระบบนี้สามารถตรวจจับช่วงเวลาที่เกิดการสูญเสียแบบสะสม และแนะนำช่วงทำความสะอาดที่เหมาะสมตามข้อจำกัดของการปฏิบัติงานจริง ผลลัพธ์จากข้อมูลการใช้งานจริงในพื้นที่เกาะสะท้อนให้เห็นว่า วิธีการที่นำเสนอช่วยลดการทำความสะอาดที่ไม่จำเป็น ลดการสูญเสียพลังงาน และเพิ่มประสิทธิภาพด้านค่าใช้จ่าย โดยสามารถให้ผลตอบแทนเชิงเศรษฐศาสตร์สูงกว่าแบบเดิมได้ถึง 36.59% รวมทั้งช่วยให้การวางแผนบำรุงรักษาเปลี่ยนจากรูปแบบตายตัวไปสู่การตัดสินใจเชิงกลยุทธ์ที่สอดคล้องกับบริบทเฉพาะของพื้นที่เกาะและพื้นที่ห่างไกล*

**คำสำคัญ:** คราบสกปรก แผงเซลล์แสงอาทิตย์ พื้นที่เกาะ การตัดสินใจเชิงข้อมูล การบำรุงรักษา

## Abstract

*Dust accumulation on solar photovoltaic panels is one of the factors that can reduce energy production efficiency, particularly in island or remote-area systems where maintenance access is limited. Fixed-interval cleaning schedules often fail to reflect the actual operating conditions of these systems. This study introduces a data-driven and weather-aware decision-making system that utilizes energy forecasting models to estimate trends in soiling-related losses and to reasonably separate meteorological effects from panel degradation. The system can detect periods of cumulative energy loss and recommend appropriate cleaning intervals that align with real operational constraints. The results from real operational data on the island demonstrate that the proposed method reduces unnecessary cleaning actions, decreases energy losses, and improves cost efficiency. It also delivers up to 36.59% higher economic returns compared to conventional approaches. Furthermore, it enables a shift from fixed maintenance schedules to strategic, context-aware decision-making tailored to the unique conditions of island and remote-area environments.*

**Keywords:** Soiling, PV, Island Areas, Data-driven Decision-making, Maintenance.

## 1. บทนำ

ระบบผลิตไฟฟ้าพลังงานแสงอาทิตย์ (PV) เป็นแหล่งพลังงานสำคัญของพื้นที่เกาะที่มีข้อจำกัดด้านโครงสร้างพื้นฐาน ทำให้สมรรถนะของระบบ PV ส่งผลโดยตรงต่อความเสถียรของพลังงาน อย่างไรก็ตาม การสะสมของฝุ่น

และสิ่งสกปรกบนแผงเป็นปัจจัยที่ลดกำลังผลิตอย่างมีนัยสำคัญ [1] - [3] และสร้างความท้าทายเพิ่มเติมในบริบทพื้นที่เกาะซึ่งมีข้อจำกัดด้านการเข้าถึงและทรัพยากร [5], [6]

รูปแบบการบำรุงรักษาที่ใช้รอบเวลาแบบตายตัวมักไม่สอดคล้องกับพฤติกรรมการสะสมคราบที่ขึ้นอยู่กับสภาพอากาศจริง [4], [7] และยังละเลยบทบาทของน้ำฝนซึ่งสามารถช่วยรีเซ็ตคราบได้ตามธรรมชาติ [9], [12], [13] ทำให้จำเป็นต้องแยกอิทธิพลของสภาพอากาศออกจากการเสื่อมที่เกิดจากคราบ โดยเฉพาะในพื้นที่เกาะที่มีต้นทุนการบำรุงรักษาสูง [5], [6]

แม้งานวิจัยที่ผ่านมาให้ความเข้าใจผลกระทบของคราบสกปรก แต่ส่วนใหญ่เน้นการวิเคราะห์ย้อนหลัง และยังไม่ผสานข้อมูลสภาพอากาศหรือข้อจำกัดเชิงปฏิบัติจริงเข้ากับการตัดสินใจ [14] - [16] ข้อจำกัดนี้ยิ่งชัดเจนในพื้นที่เกาะที่ต้องการการจัดการที่เหมาะสมเฉพาะพื้นที่

ด้วยเหตุนี้ งานวิจัยจึงพัฒนากรอบการตัดสินใจเชิงข้อมูลที่ผสานผลการพยากรณ์พลังงานกับข้อมูลปริมาณน้ำฝน เพื่อประเมินการสูญเสียจากคราบอย่างเป็นระบบ และกำหนดช่วงเวลาทำความสะอาดที่เหมาะสมภายใต้ข้อจำกัดภาคสนาม เป้าหมายคือยกระดับกลยุทธ์บำรุงรักษาจากตารางเวลาแบบตายตัวสู่การวางแผนเชิงกลยุทธ์ที่อิงข้อมูลจริงและเหมาะสมกับบริบทพื้นที่เกาะ [5], [6], [14]

## 2. การทบทวนวรรณกรรมและงานวิจัยที่เกี่ยวข้อง

### 2.1 งานด้านการประเมินการสูญเสียพลังงานจากคราบสกปรกของระบบเซลล์แสงอาทิตย์

งานวิจัยช่วงแรกมุ่งศึกษาผลกระทบของคราบสกปรกต่อสมรรถนะระบบ PV โดยอาศัยการวิเคราะห์ย้อนหลังของข้อมูลพลังงาน โดย Kaldellis และ Kapsali [1] แสดงให้เห็นว่าผลกระทบของ Soiling มีลักษณะเฉพาะพื้นที่ขึ้นกับชนิดและอัตราการสะสมของฝุ่น รวมถึงปัจจัยสภาพอากาศอย่างความชื้นที่ช่วยเพิ่มการยึดเกาะของฝุ่น ต่อมา Sayyah et al. [2] ขยายมุมมองในระดับสากล โดยรายงานว่าคราบสกปรกอาจทำให้กำลังผลิตลดลงถึง 25 - 40% ในบางภูมิภาค และการศึกษาภาคสนามระยะยาวยังชี้ว่าพฤติกรรมการสะสมคราบมีความแปรผันตามสภาพแวดล้อมและช่วงเวลาอย่างมีนัยสำคัญ [8] สะท้อนผลกระทบเชิงเศรษฐศาสตร์ที่ชัดเจนของปัญหานี้

ด้าน Micheli และ Sarmah [4], [11] ได้สังเคราะห์องค์ความรู้ด้าน Soiling อย่างเป็นระบบ ครอบคลุมคุณสมบัติฝุ่น ปัจจัยสภาพอากาศ กลไกการสะสม และวิธีประเมิน Soiling loss ทำให้งานของทั้งสองเป็นฐานข้อมูลสำคัญสำหรับการพัฒนาวิธีประเมินประสิทธิภาพระบบ PV ในหลายบริบท

$$Soiling\ loss_t = \frac{E_{pred,t} - E_{meas,t}}{E_{pred,t}} \text{ x } 100 \qquad (1)$$

### 2.2 งานด้านการศึกษาบทบาทของสภาพอากาศและการทำความสะอาดตามธรรมชาติ

งานวิจัยช่วงถัดมามุ่งทำความเข้าใจบทบาทของสภาพอากาศต่อการฟื้นฟูประสิทธิภาพของแผง PV โดยเฉพาะ “ฝน” ซึ่งทำหน้าที่เป็นกลไกชะล้างตามธรรมชาติ งานของ Jo et al. [9] พบว่าฝนที่มีความแรงและต่อเนื่องสามารถล้างคราบได้อย่างมีประสิทธิภาพ ขณะที่ฝนเบามีผลจำกัด ต่อมา Stankov et al. [13] ชี้ว่าองค์ประกอบของลมร่วมกับฝนช่วยเพิ่มประสิทธิภาพการลดคราบ โดยเฉพาะในพื้นที่ที่มีฝุ่นหนาแน่น สะท้อนว่าการประเมินผลของสภาพอากาศควรพิจารณาแบบองค์รวม

ในทำนองเดียวกัน Santos et al. [12] รายงานว่าลักษณะของฝน เช่น ขนาดเม็ด ความเข้ม และระยะเวลาการตกส่งผลแตกต่างกันต่อการฟื้นฟูประสิทธิภาพแผง PV แสดงให้เห็นว่าฝนไม่ได้ส่งผลเพียงด้านปริมาณ แต่ “รูปแบบของฝน” เองก็มีความสำคัญต่อการทำความสะอาดตามธรรมชาติของระบบ PV

### 2.3 งานด้านการวางแผนบำรุงรักษาและการตัดสินใจเชิงปฏิบัติ

งานวิจัยช่วงหลังเริ่มใช้ข้อมูลประสิทธิภาพและสภาพแวดล้อมในการวางแผนบำรุงรักษาระบบ PV แม้ยังไม่มุ่งที่ปัญหาคราบสกปรกโดยตรง เช่น Redondo et al. [14] ที่เปรียบเทียบแบบจำลองการสะสมคราบหลายรูปแบบ แม้มีคุณค่าเชิงโมเดลแต่ยังไม่สามารถกำหนดช่วงเวลาทำความสะอาดจริงได้ ขณะที่ Said et al. [15] สำรวจเทคนิคและต้นทุนการล้างแผงในหลาย

สภาพแวดล้อม แต่ยังไม่พัฒนาเป็นระบบช่วยตัดสินใจเชิงปฏิบัติ ส่วนแนวทางมาตรฐาน เช่น NREL Best Practice Guidelines [5] และ IEA PVPS Task 13 [6] ยังแนะนำการล้างตามรอบเวลาแบบตายตัว ซึ่งไม่สอดคล้องกับพฤติกรรมคราบที่แปรผันตามสภาพอากาศจริง

แม้งานก่อนหน้าจะช่วยเพิ่มความเข้าใจเรื่องคราบสกปรก สภาพอากาศ และการบำรุงรักษา แต่ส่วนใหญ่ยังวิเคราะห์ย้อนหลังและยังไม่ผสานทั้งสามองค์ความรู้เพื่อรองรับการตัดสินใจจริง โดยเฉพาะในพื้นที่เกาะที่มีข้อจำกัดด้านต้นทุนและการเข้าถึง งานวิจัยนี้จึงเสนอกรอบการตัดสินใจเชิงข้อมูลที่ผสานการพยากรณ์พลังงาน ผลของน้ำฝน และพฤติกรรมการสะสมคราบ เพื่อระบุช่วงเวลาทำความสะอาดที่เหมาะสมและนำไปใช้ได้จริงในบริบทพื้นที่เกาะและพื้นที่ห่างไกล

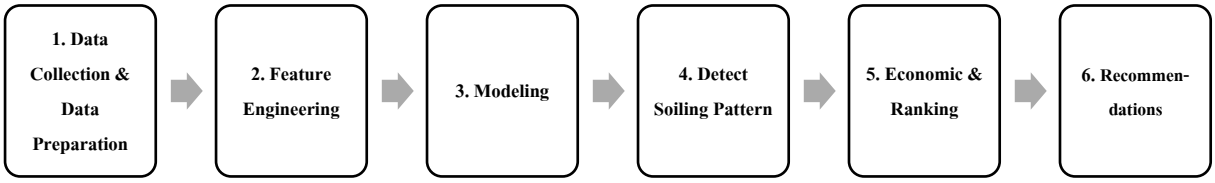
## 3. วิธีการดำเนินงานวิจัย

งานวิจัยนี้พัฒนากรอบระเบียบวิธีเชิงข้อมูลที่ผสานบทบาทของสภาพอากาศเข้ากับการประเมินการสูญเสียพลังงานจากคราบสกปรก เพื่อสนับสนุนการตัดสินใจทำความสะอาดแผง PV อย่างเป็นระบบ โดยอ้างอิงแนวคิดและวิธีประเมินจากงานก่อนหน้า [14], [15] ที่เน้นการจำลองและประเมินพฤติกรรมคราบสกปรก

ระเบียบวิธีถูกออกแบบให้สามารถแยกอิทธิพลของสภาพอากาศออกจากการเสื่อมที่เกิดจากคราบสกปรกได้อย่างแม่นยำขึ้น ตามหลักฐานเชิงทดลองจากงานที่ศึกษาผลของฝนต่อการฟื้นฟูสมรรถนะ [12] และงานที่พิจารณาปัจจัยอุตุนิยมวิทยาหลายตัวร่วมกัน [13]

นอกจากนี้ กรอบวิธียังเชื่อมโยงผลการพยากรณ์พลังงานและสัญญาณการสะสมคราบไปสู่ข้อเสนอแนะที่สอดคล้องกับข้อจำกัดด้านการปฏิบัติงานจริงของพื้นที่เกาะ โดยอ้างอิงแนวคิดการบำรุงรักษาและบริบทการดำเนินงานจาก [15], [17], [6]

ภาพรวมขั้นตอนการดำเนินงานแบ่งออกเป็น 6 ขั้นตอน ดังแสดงใน แผนภาพที่ 1




**ภาพที่ 1:** ขั้นตอนการพัฒนากรอบการตัดสินใจทำความสะอาดแผงเซลล์แสงอาทิตย์

ระเบียบวิธีวิจัยของงานนี้ถูกออกแบบเป็นลำดับขั้นตอนตั้งแต่การเตรียมข้อมูลจนถึงการวิเคราะห์เชิงเศรษฐศาสตร์ เพื่อสนับสนุนการตัดสินใจด้านการทำความสะอาดแผง PV ในบริบทพื้นที่เกาะ โดยผสานข้อมูลการผลิตไฟฟ้า ข้อมูลสภาพอากาศ และข้อมูลต้นทุนเข้าไว้ด้วยกัน

ขั้นแรกคือการรวบรวมข้อมูลอนุกรมเวลารายวันของระบบ PV ควบคู่กับข้อมูลสภาพอากาศและข้อจำกัดพื้นที่ จากนั้นดำเนินการทำความสะอาดข้อมูลและแบ่งข้อมูลตามลำดับเวลา โดยใช้ข้อมูลปี 2022–2024 เป็นชุดฝึก และปี 2025 เป็นชุดทดสอบ พร้อมสร้างชุดคุณลักษณะที่สะท้อนพฤติกรรมการสะสมคราบ เช่น ตัวแปร Lag, Rolling window, สัญญาณ Rain Reset และตัวแปรเชิงปฏิสัมพันธ์

ขั้นต่อมาคือการพัฒนาแบบจำลองพยากรณ์พลังงานด้วยหลายอัลกอริทึม ได้แก่ XGBoost, SVM, Decision Tree, Random Forest และ Linear Regression เพื่อสร้างค่าพลังงานอ้างอิง $E_{pred,t}$ ก่อนนำไปเปรียบเทียบกับค่าที่วัดได้จริง $E_{meas,t}$ เพื่อคำนวณค่าการสูญเสียรายวัน (Daily loss) ซึ่งคือความแตกต่างระหว่างค่าคาดการณ์และค่าจริง และสะท้อนผลกระทบรวมจากปัจจัยที่มีผลต่อสมรรถนะของระบบ

$$Daily\ loss_t = E_{pred,t} - E_{meas,t} \quad (2)$$

โดยที่ $E_{pred,t}$ แทนค่าพลังงานไฟฟ้าที่พยากรณ์ได้ในวันที่ $t$ และ $E_{meas,t}$ แทนค่าพลังงานไฟฟ้าที่วัดได้จริง การนิยามดังกล่าวช่วยให้สามารถประเมินความคลาดเคลื่อนของพลังงานในระดับรายวัน และใช้เป็นตัวแทนของการสูญเสียพลังงานที่เกิดจากการเสื่อมประสิทธิภาพของแผง ภายใต้สมมติฐานว่าสภาพอากาศในภาพรวมถูกสะท้อนผ่านแบบจำลองอ้างอิงแล้ว

หลังจากเลือกแบบจำลองที่ให้ผลดีที่สุด จะเข้าสู่กระบวนการประมวลผลเพื่อดึงสัญญาณการสะสมคราบอย่างมีเสถียรภาพ โดยใช้การกรองสภาพอากาศ (แยกวันที่ฝนตกหรือวันที่มีการรีเซ็ต) ควบคู่กับการคำนวณ Rolling signal ในงานวิจัยนี้กำหนดเกณฑ์การรีเซ็ตจากน้ำฝน (Rain Reset Threshold) ที่ปริมาณน้ำฝนรายวันมากกว่าหรือเท่ากับ 5 mm โดยอ้างอิงจากหลักฐานเชิงทดลองที่ชี้ว่าฝนที่มีความเข้มเพียงพอสามารถชะล้างคราบได้อย่างมีประสิทธิภาพ [9] ในขณะที่ฝนเบาหรือต่ำกว่าเกณฑ์ดังกล่าวมีผลจำกัดต่อการฟื้นฟูสมรรถนะของแผง [12] สำหรับตรวจจับแนวโน้มการสะสมคราบแบบต่อเนื่อง สัญญาณที่ได้ (Soiling Signal) จะใช้เป็นฐานสำหรับวิเคราะห์ระดับการสะสมและช่วงเวลาที่มีผลกระทบสูงที่สุดโดย

- Rolling เป็นเทคนิคการวิเคราะห์ข้อมูลอนุกรมเวลา โดยการกำหนด หน้าต่างเวลา (Window) ให้เลื่อนไปตามลำดับข้อมูล เพื่อคำนวณค่าเฉลี่ยที่สะท้อนแนวโน้มหรือพฤติกรรมล่าสุดในช่วงสั้นๆ ดังสมการที่ 3

$$Rolling_t = \frac{1}{N}\sum_{i=0}^{N-1} x_{t-i} \quad (3)$$

โดยที่ $x_t$ คือค่าของข้อมูล ณ เวลา $t$ และ $N$ คือค่าจำนวนวันก่อนหน้า

- Soiling Signal สัญญาณนี้สะท้อนพฤติกรรมการเสื่อมลงของแผงแบบต่อเนื่องตามเวลา และใช้ติดตามรูปแบบการสะสมคราบ (Accumulation Pattern) ดังสมการที่ 4

$$Soling\ Signal_t = \frac{E_{pred,t} - E_{meas,t}}{E_{pred,t}} \quad (4)$$

ขั้นตอนการตัดสินใจจะประเมินด้านเศรษฐศาสตร์ร่วมกับสัญญาณการสะสมคราบ เพื่อคำนวณความคุ้มค่าของการทำความสะอาดแต่ละครั้ง และจัดอันดับช่วงเวลาที่ควรล้างแผงภายใต้ข้อจำกัดของพื้นที่เกาะ ทั้งในเชิงต้นทุนและการเข้าถึงพื้นที่

ท้ายที่สุด ระบบจะสรุปเป็นข้อเสนอแนะเชิงปฏิบัติที่สามารถดำเนินการได้จริง เช่น ช่วงเวลาที่เหมาะสมต่อการทำความสะอาด เพื่อช่วยให้ผู้ปฏิบัติงานตัดสินใจได้อย่างมีประสิทธิภาพมากยิ่งขึ้น

## 4. การทดลอง

### 4.1 ชุดข้อมูลทดลอง

1) ข้อมูลพลังงานที่วัดได้จริงจากระบบ PV ซึ่งดึงจากระบบ Monitoring ของผู้ผลิต ร่วมกับข้อมูลสภาพอากาศที่ได้จากแหล่งข้อมูลที่เชื่อถือได้ [18], [19] จำนวน 1,188 รายการ โดยแต่ละรายการมีขอบเขตของข้อมูลดังแสดงในตารางที่ 1

**ตารางที่ 1:** ข้อมูลหลักที่ใช้ในการศึกษา

| Data Field | Description |
|---|---|
| Date | วันที่ |
| year | ปี |
| Daily_Energy_Generation | พลังงานที่ผลิตได้รายวัน (kWh) |
| Event Type | เหตุการณ์แจ้งเตือนระบบ PV |
| Fault code | รหัสความผิดปกติของระบบ PV |
| Season | ฤดูกาลท่องเที่ยว (High/Low) |
| Daily_Irradiance_Sum | ปริมาณพลังงานแสงสะสมรายวัน (W/m²) |
| Avg_Temperature | อุณภูมิเฉลี่ย (C°) |
| Daily_Rainfall | ปริมาณน้ำฝนรายวัน (mm) |
| Relative_Humidity_Min (%) | ความชื้นสัมพันธ์ต่ำสุด (%) |
| Relative_Humidity_Max (%) | ความชื้นสัมพันธ์สูงสุด (%) |
| Avg_Wind_Speed | ความเร็วลมเฉลี่ย (m/s) |
| Wind_Direction | ทิศทางลม |

2) ข้อจำกัดของพื้นที่ศึกษาซึ่งตั้งอยู่บนเกาะที่ไม่เชื่อมต่อกับระบบสายส่งหลัก ส่งผลให้ต้องพึ่งพาเครื่องปั่นไฟดีเซลเป็นแหล่งพลังงานหลัก ซึ่งมีต้นทุนสูงและมีข้อจำกัดด้านการเดินทางสำหรับงานบำรุงรักษา ทำให้สามารถดำเนินการได้เพียงปีละ 1–2 ครั้ง (ต้นปี-กลางปี) งานวิจัยจึงนำข้อมูลต้นทุนและอัตราค่าไฟฟ้ามาใช้ประเมินความคุ้มค่าของช่วงเวลาทำความสะอาด ดังแสดงในตารางที่ 2

**ตารางที่ 2:** ข้อมูลต้นทุนที่เกี่ยวข้อง

| รายการ | หน่วย | มูลค่า (บาท) |
|---|---|---|
| ค่าไฟฟ้า (Diesel Generator) | บาท/kWh | 16.00 |
| ค่าจ้างบำรุงรักษาต่อครั้ง | บาท/ครั้ง | 12,000 |

### 4.2 วิธีการวัดประสิทธิภาพ

ในการประเมินความสามารถของแบบจำลองในการทำนายพลังงานอ้างอิง งานวิจัยนี้ใช้ตัวชี้วัดสามประเภทที่

สะท้อนความแม่นยำในมิติต่างกัน ได้แก่ $R^2$, RMSE, และ MAE ซึ่งเป็นมาตรฐานที่ใช้กันอย่างแพร่หลายในงานพยากรณ์พลังงานและงานวิเคราะห์อนุกรมเวลา [10] ดังสมการ (5) - (7)

$$R^2 = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y})^2} \quad (5)$$

$$RMSE = \sqrt{\frac{\sum(y_i - \hat{y}_i)^2}{n}} \quad (6)$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n} |y_i - \hat{y}_i| \quad (7)$$

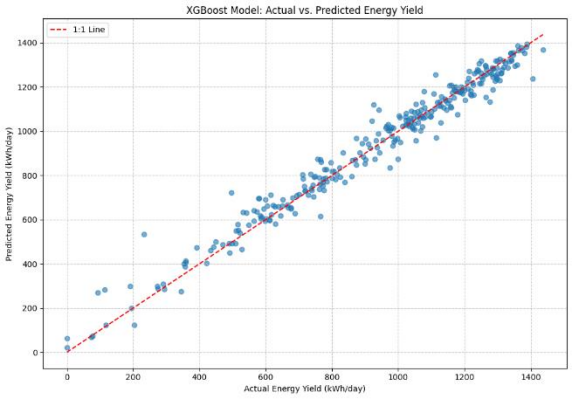
## 5. ผลการทดลอง

### 5.1 ประสิทธิภาพของแบบจำลองพลังงานอ้างอิง

เพื่อสร้างค่าพลังงานอ้างอิงที่ใช้ในการประเมิน Soiling loss งานวิจัยนี้ได้เปรียบเทียบประสิทธิภาพของโมเดลพยากรณ์พลังงานหลายรูปแบบภายหลังการปรับแต่งด้วย GridSearch ตารางที่ 3

**ตารางที่ 3:** ผลการเปรียบเทียบประสิทธิภาพโมเดล

| Model | $R^2$ | RMSE | MAE |
|---|---|---|---|
| **XGBoost (GS)** | **0.981** | **44.667** | **30.364** |
| Random Forest (GS) | 0.958 | 66.475 | 46.052 |
| Decision Tree (GS) | 0.917 | 93.330 | 64.255 |
| Linear Regression (GS) | 0.911 | 96.849 | 73.945 |
| SVM (GS) | 0.344 | 263.111 | 202.720 |

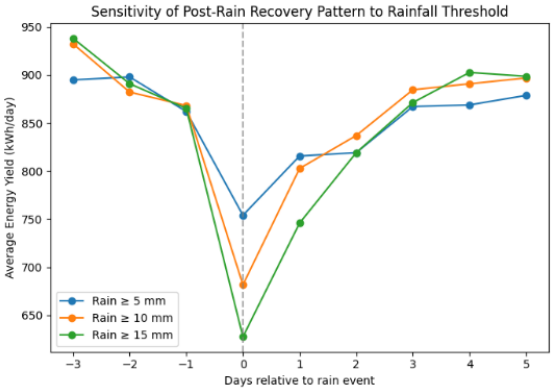



**ภาพที่ 2:** กราฟเปรียบเทียบค่าพลังงานที่วัดได้จริง (Actual) กับค่าที่พยากรณ์ได้ (Predicted) จากโมเดล XGBoost

ผลการประเมินพบว่า XGBoost ให้ผลดีที่สุด ด้วยค่า $R^2$ สูงถึง 0.98 และค่าความคลาดเคลื่อนต่ำสุด สะท้อนความสามารถในการจับความสัมพันธ์เชิงเวลาได้แม่นยำมากกว่าโมเดลอื่น เหมาะสมที่สุดสำหรับสร้าง Baseline Energy และหาค่า Daily loss เพื่อนำใช้นำไปประเมิน Soiling loss ในขั้นต่อไป

### 5.2 ผลการประเมินลักษณะการสูญเสียพลังงานจากคราบสกปรก (Soiling loss)

ส่วนนี้แสดงผลการประเมินการสูญเสียพลังงานรายวันและผลกระทบของฝนตก เพื่อทำความเข้าใจรูปแบบพื้นฐานของการสูญเสียและการรีเซ็ตก่อนสร้างสัญญาณการสะสมคราบในระยะยาว




**ภาพที่ 3 :** รูปแบบการฟื้นฟูสมรรถนะของระบบหลังจากฝนตกแบ่งตามปริมาณน้ำฝน

### 5.3 ผลของกลไกการรีเซ็ตจากน้ำฝน (Rain Reset)



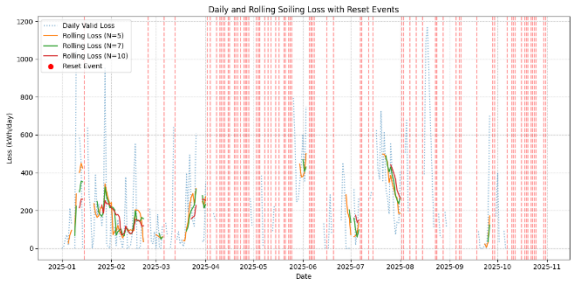

**รูปที่ 4:** แสดงการเปรียบเทียบระหว่าง Daily loss Vs Soiling Signal with rain reset

จากรูปที่ 4 พบว่าการสูญเสียพลังงานรายวันมีความผันผวนสูงรอบค่าศูนย์ สะท้อนความแปรปรวนของสภาพอากาศและความไม่แน่นอนรายวัน ขณะที่สัญญาณการสะสมคราบที่ผ่านการประมวลผลมีความเรียบและต่อเนื่องกว่า โดยเพิ่มขึ้นชัดเจนในช่วงไร้ฝน และถูกรีเซ็ตเมื่อมีฝนปริมาณเพียงพอ

ผลการเปรียบเทียบชี้ว่าฝนช่วยลดสัญญาณรบกวนระยะสั้น ทำให้ติดตามแนวโน้มการสะสมคราบได้เสถียรมากขึ้น ซึ่งเป็นฐานสำคัญสำหรับการประเมินเชิง

เศรษฐศาสตร์และการตัดสินใจด้านการทำความสะอาดในขั้นต่อไป

### 5.4 ผลการสัญญาณการสะสมของคราบสกปรกจากการคำนวณค่าเฉลี่ยเคลื่อนที่ (Rolling Signal)



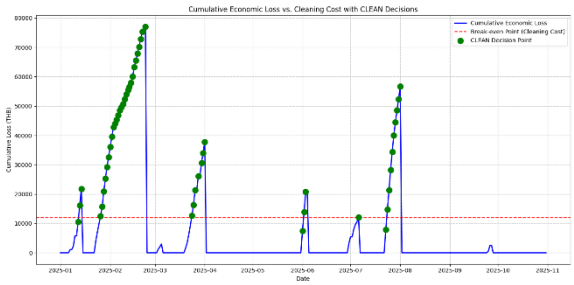

**รูปที่ 5:** Daily loss Vs Rolling Signal with Rain Reset

ผลการวิเคราะห์จากรูปที่ 5 แสดงให้เห็นว่าการใช้ค่าเฉลี่ยเคลื่อนที่ช่วยให้สัญญาณการสะสมคราบมีความเรียบและเสถียรมากกว่าข้อมูลรายวัน โดยขนาดหน้าต่างเลื่อนส่งผลต่อความเรียบและความไวของสัญญาณ

โดยหน้าต่างเล็ก (N=5) ตอบสนองเร็วแต่ยังมีความผันผวนสูง หน้าต่างใหญ่ (N=10) ให้สัญญาณเรียบกว่าแต่ตอบสนองช้าต่อการเปลี่ยนแปลง ส่วนหน้าต่างขนาดกลาง (N=7) ให้สมดุลที่เหมาะสมระหว่างความเรียบและการตอบสนองต่อเหตุการณ์ฝนตกและการเริ่มต้นสะสมของคราบในภาพรวม สัญญาณที่ผ่านการทำค่าเฉลี่ยเคลื่อนที่ช่วยให้เห็นแนวโน้มการสะสมคราบระหว่างช่วงที่ไม่มีฝนได้ชัดเจน และถูกรีเซ็ตอย่างเป็นระบบเมื่อเกิดฝนตก ทำให้รูปแบบสัญญาณมีความหมายเชิงกายภาพมากขึ้น งานวิจัยนี้จึงเลือกใช้ค่าหน้าต่างที่เหมาะสมเพื่อใช้ในการวิเคราะห์พฤติกรรมการสะสมคราบและรองรับการประเมินเชิงเศรษฐศาสตร์ในขั้นตอนถัดไป

### 5.5 ผลการประเมินการสูญเสียเชิงเศรษฐศาสตร์

เมื่อแปลงการสูญเสียพลังงานจากคราบสกปรกเป็นมูลค่าทางเศรษฐศาสตร์ พบว่าการสะสมของคราบสกปรกก่อให้เกิดความสูญเสียทางโอกาสที่เพิ่มขึ้นตามระยะเวลา โดยเฉพาะในช่วงที่ระบบไม่ได้รับการทำความสะอาดทั้งจากน้ำฝนและการทำความสะอาดโดยการดำเนินการบำรุงรักษา ผลลัพธ์ดังกล่าวสะท้อนว่าการปล่อยให้คราบสกปรกสะสมต่อไปโดยไม่มีการแทรกแซงสามารถนำไปสู่การสูญเสียทางเศรษฐศาสตร์ในระดับที่มีนัยสำคัญ




**รูปที่ 6:** ความสูญเสียทางเศรษฐศาสตร์สะสมเทียบกับต้นทุนทำความสะอาด (CLEAN Decisions)

จากรูปที่ 6 แสดงความสัมพันธ์ระหว่างการสูญเสียเชิงเศรษฐศาสตร์สะสมจากคราบสกปรกและต้นทุนการทำความสะอาด โดยจุดที่เส้นการสูญเสียสะสมตัดกับต้นทุนการทำความสะอาดถูกใช้เป็นจุดคุ้มทุน และระบุเป็นช่วงตัดสินใจทำความสะอาด เพื่อให้สอดคล้องกับข้อจำกัดด้านการปฏิบัติงานจริง เช่น การเข้าถึงพื้นที่และความพร้อมของบุคลากร

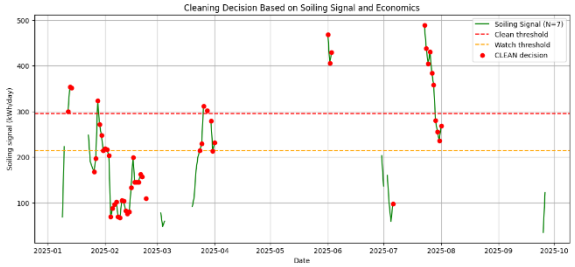
ผลลัพธ์ชี้ว่าการตัดสินใจทำความสะอาดไม่ควรยึดตามรอบเวลาตายตัว [6] แต่ควรขึ้นอยู่กับอัตราการสะสมของการสูญเสีย ซึ่งได้รับอิทธิพลจากสภาพอากาศและพฤติกรรมของคราบในแต่ละช่วง แนวทางนี้ช่วยหลีกเลี่ยงการบำรุงรักษาที่ไม่จำเป็น และระบุช่วงที่การปล่อยให้คราบสะสมจะสร้างต้นทุนสูงกว่าการทำความสะอาดได้อย่างมีประสิทธิภาพ โดยผลการประเมินทางเศรษฐศาสตร์สรุปไว้ในตารางที่ 4 ผ่านตัวชี้วัดมูลค่าทางเศรษฐศาสตร์สุทธิ

**ตารางที่ 4:** ตัวชี้วัดประเมินประสิทธิภาพเชิงเศรษฐศาสตร์

| Metric | Value | Unit |
|---|---|---|
| **Number of Soiling signal** | 72.00 | days |
| **Number of candidate CLEAN events (pre-filtering)** | 52.00 | events |
| **Total avoided energy loss** | 124,909.77 | kWh |
| **Total avoided economic loss** | 1,998,556.34 | THB |
| **Total cleaning cost** | 624,000.00 | THB |
| **Net economic benefit** | 1,374,556.34 | THB |
| **Average avoided loss** | 27,757.73 | THB/day |
| **Estimated payback period** | 0.40 | days |

**5.6 กลไกการตัดสินใจทำความสะอาดแผงเซลล์แสงอาทิตย์จากสัญญาณการสูญเสียเชิงเศรษฐศาสตร์**

หลังจากพัฒนาสัญญาณการสะสมของคราบสกปรกและประเมินมูลค่าการสูญเสียเชิงเศรษฐศาสตร์แล้ว ขั้นตอนถัดไปคือการแปลงผลการวิเคราะห์ดังกล่าวให้เป็นกรอบการตัดสินใจทำความสะอาดที่สามารถนำไปใช้ได้จริง งานวิจัยนี้จึงเสนอแนวทางการตัดสินใจที่อาศัยทั้งระดับความรุนแรงของสัญญาณการสูญเสียเชิงเศรษฐศาสตร์และข้อจำกัดด้านการดำเนินงาน โดยไม่ยึดติดกับการกำหนดตารางการทำความสะอาดแบบคงที่ แต่เลือกดำเนินการเฉพาะในช่วงเวลาที่มีความคุ้มค่าสูงสุด




**รูปที่ 7:** แสดงการตัดสินใจทำความสะอาดโดยอาศัยสัญญาณการสะสมของคราบสกปรก

จากรูปที่ 7 พบว่าสัญญาณการสะสมคราบอาจเข้าใกล้หรือเกินเกณฑ์เชิงเศรษฐศาสตร์ในหลายช่วงเวลา แต่การทำความสะอาดไม่สามารถดำเนินการได้ทุกครั้งเนื่องจากข้อจำกัดด้านทรัพยากรและจำนวนครั้งที่ทำได้ต่อปี งานวิจัยจึงใช้แนวคิด Decision window เพื่อระบุ “ช่วงเวลาที่เหมาะสมเชิงเศรษฐศาสตร์” แทนการตอบสนองต่อสัญญาณทุกครั้ง

แบบจำลองทำหน้าที่ระบุช่วงเวลาที่มีการสูญเสียสะสมสูง ขณะที่ข้อจำกัดภาคสนามทำหน้าที่กำกับการเลือกช่วงทำความสะอาด เมื่อมีสัญญาณหลายช่วงพร้อมกัน จึงใช้วิธี Rank-based planning เพื่อจัดอันดับเหตุการณ์ตามผลกระทบเชิงเศรษฐศาสตร์ และเลือกเฉพาะช่วงที่คุ้มค่าที่สุด (ดูรูปที่ 8)

**5.7 ผลการจัดอันดับช่วงเวลาที่เหมาะสมสำหรับการทำความสะอาด**

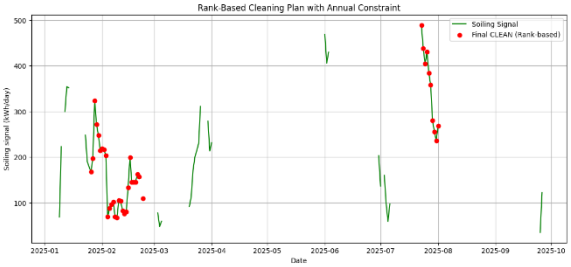
ภายใต้ข้อจำกัดด้านการบำรุงรักษาที่สามารถดำเนินการได้ปีละ 2 ครั้ง งานวิจัยนี้จึงเลือกเฉพาะเหตุการณ์ที่มีลำดับความสำคัญสูงสุดจากการจัดอันดับดังกล่าวเป็นช่วงเวลาที่เหมาะสมที่สุดสำหรับการทำความสะอาดแผงเซลล์แสงอาทิตย์ ซึ่งช่วยให้การใช้ทรัพยากรด้านการบำรุงรักษามุ่งเน้นไปที่ช่วงเวลาที่ให้ความคุ้มค่าสูงสุด ดังตารางที่ 5

**ตารางที่ 5:** TOP-K Events

| Event id | Start date | End date | Event cum Loss THB | Peak soiling signal |
|---|---|---|---|---|
| **1** | **26 01 2025** | **23 02 2025** | **77,057.03** | **323.60** |
| **5** | **23 07 2025** | **1 08 2025** | **56,680.91** | **489.01** |
| 2 | 24 03 2025 | 1 04 2025 | 37,721.37 | 311.01 |
| 0 | 12 01 2025 | 14 01 2025 | 21,840.46 | 354.01 |
| 3 | 1 06 2025 | 3 06 2025 | 20,853.83 | 468.26 |
| 4 | 6 07 2025 | 6 07 2025 | 12,087.77 | 97.61 |

ผลการจัดอันดับเหตุการณ์ด้วยสัญญาณคราบและการสูญเสียเชิงเศรษฐศาสตร์ (TOP-K Events) ระบุว่าเหตุการณ์หมายเลข 1 (26 ม.ค.–23 ก.พ.) มีผลกระทบสูงสุด ด้วยมูลค่าการสูญเสียสะสม 77,057 บาท ตามด้วยเหตุการณ์หมายเลข 5 (23 ก.ค.–1 ส.ค.) ซึ่งล้วนสะท้อนการสะสมคราบต่อเนื่องจนเกิดความสูญเสียระดับสูง

เมื่อเทียบกับวิธีการทำงานเดิมซึ่งมีการทำความสะอาด 2 ครั้งในเดือนม.ค. และมิ.ย. พบว่าการล้างครั้งแรกยังสอดคล้องกับช่วงเหตุการณ์สำคัญ แต่การล้างครั้งที่สองไม่สอดคล้องกับพฤติกรรมการสะสมจริง ข้อมูลชี้ว่าช่วง ก.ค.–ส.ค. โดยเฉพาะเหตุการณ์หมายเลข 5 จะให้ผลตอบแทนทางเศรษฐศาสตร์สูงกว่าเดิมถึง 36.59% หากย้ายการทำความสะอาดไปช่วงดังกล่าว




**รูปที่ 8:** แผนการทำความสะอาดภายใต้ข้อจำกัดรายปี

จากรูปที่ 8 และตารางที่ 5 เมื่อพิจารณาร่วมกัน แสดงให้เห็นอย่างชัดเจนว่ากลไกการตัดสินใจที่นำเสนอสามารถบูรณาการการวิเคราะห์เชิงเทคนิค การประเมินเชิงเศรษฐศาสตร์ และข้อจำกัดด้านการปฏิบัติงานจริงเข้าไว้ในกรอบเดียวกันได้อย่างเป็นระบบ ส่งผลให้การตัดสินใจทำความสะอาดไม่ยึดติดกับการกำหนดช่วงเวลาแบบคงที่

แต่พิจารณาจากระดับการสูญเสียเชิงเศรษฐศาสตร์และลำดับความสำคัญของเหตุการณ์เป็นหลัก แนวทางดังกล่าวช่วยเพิ่มความยืดหยุ่นและประสิทธิภาพในการวางแผนบำรุงรักษา และสะท้อนสภาพการทำงานจริงของระบบผลิตไฟฟ้าพลังงานแสงอาทิตย์ได้อย่างเหมาะสมมากยิ่งขึ้น

## 6. การสรุปและอภิปรายผล

ผลการศึกษาชี้ให้เห็นว่าพฤติกรรมการสะสมคราบของระบบโซลาร์มีความซับซ้อนกว่าที่คาดไว้ในแนวปฏิบัติทั่วไป โดยการสูญเสียพลังงานมีลักษณะเป็นเหตุการณ์ (event-based) และได้รับอิทธิพลอย่างมากจากสภาพอากาศ โดยเฉพาะฝนซึ่งทำหน้าที่รีเซ็ตคราบตามธรรมชาติ ทั้งยังพบว่าการฟื้นตัวหลังฝนไม่ได้เกิดขึ้นทันที แต่มีระยะหน่วงราว 1–3 วัน สะท้อนปัจจัยเฉพาะของพื้นที่เกาะ เช่น ความชื้นสูงและละอองเกลือ

ข้อค้นพบนี้สะท้อนข้อจำกัดของแนวทางบำรุงรักษาแบบคงที่ ซึ่งไม่สามารถอธิบายรูปแบบการสะสมคราบและการสูญเสียที่เกิดขึ้นจริงได้อย่างเหมาะสม สัญญาณการสูญเสียเชิงเศรษฐศาสตร์ที่พัฒนาขึ้นช่วยแยกการสะสมคราบออกจากความผันผวนของสภาพอากาศรายวัน ทำให้สามารถระบุช่วงเวลาที่ต้นทุนจากการปล่อยให้คราบสะสมสูงกว่าต้นทุนการทำความสะอาดได้อย่างชัดเจน

การตัดสินใจทำความสะอาดแบบจัดอันดับ (Rank-based) ช่วยเลือกช่วงเวลาที่คุ้มค่าที่สุดภายใต้ข้อจำกัดจำนวนครั้งต่อปี ลดการทำความสะอาดที่ไม่จำเป็นและไม่พลาดช่วงเวลาที่ให้ผลตอบแทนสูง กรอบดังกล่าวเปลี่ยนจากการบำรุงรักษาตามรอบเวลาไปสู่การวางแผนเชิงกลยุทธ์บนพื้นฐานข้อมูลจริง ทั้งด้านสภาพอากาศ สัญญาณคราบ และผลกระทบเชิงเศรษฐศาสตร์

ผลลัพธ์เชิงประจักษ์ยืนยันว่ากลยุทธ์เชิงข้อมูลให้ Net Economic Benefit เป็นบวก และลดต้นทุนรวมได้อย่างมีนัยสำคัญ แม้งานวิจัยจะเน้นบริบทพื้นที่เกาะ แต่กรอบวิเคราะห์นี้สามารถต่อยอดสู่สภาพแวดล้อมอื่น และพัฒนาไปสู่ระบบบำรุงรักษาแบบอิงสภาพจริงในอนาคตได้

### เอกสารอ้างอิง


[1] J. K. Kaldellis and M. Kapsali, “Simulating the dust effect on PV generators,” Energy, vol. 36, no. 8, pp. 5154–5161, 2011.

[2] A. Sayyah et al., “Energy yield loss caused by dust deposition,” Solar Energy, vol. 107, pp. 576–604, 2014.

[3] R. R. Cordero et al., “Effects of soiling in the Atacama Desert,” Sci. Rep., vol. 8, p. 13943, 2018.

[4] L. Micheli and N. Sarmah, “Monitoring photovoltaic soiling,” iScience, vol. 24, no. 2, pp. 102–149, 2021.

[5] NREL & Sandia National Laboratories, O&M Best Practice Guidelines for Photovoltaic Power Plants, Version 5.0, 2021.

[6] IEA PVPS Task 13, Guidelines for O&M of Photovoltaic Plants, IEA-PVPS, 2022.

[7] R. Conceição et al., “Soiling effect in solar energy conversion systems,” Renew. Sustain. Energy Rev., vol. 162, p. 112434, 2022.

[8] B. Aïssa et al., “Field PV soiling assessment in Qatar,” Solar Energy, vol. 239, pp. 106–122, 2022.

[9] W. Jo et al., “Cleaning effect according to precipitation,” Renew. Energy, vol. 216, pp. 118–131, 2023.

[10] J. Gaboitaolelwe et al., “Machine Learning Based Solar Photovoltaic Power Forecasting: A Review and Comparison,” IEEE Access, 2023.

[11] L. Micheli and N. Sarmah, “Analysis of soiling loss in photovoltaic modules,” Renew. Sustain. Energy Rev., vol. 182, pp. 113–143, 2024.

[12] F. N. Santos et al., “Cleaning PV modules through rain,” Solar RRL, vol. 8, p. 2400551, 2024.

[13] B. Stankov et al., “Influence of wind and rainfall on PV performance,” Energies, vol. 17, p. 3394, 2024.

[14] M. Redondo et al., “Methods for soiling modeling in large PV plants,” Sustainability, vol. 16, p. 10998, 2024.

[15] S. Z. Said et al., “Optimal cleaning techniques for PV systems,” Energy Rep., vol. 12, pp. 1121–1141, 2024.

[16] H. Abdulla et al., “PV systems O&M: A review and future directions,” Renew. Sustain. Energy Rev., vol. 195, p. 114084, 2024.

[17] A. Abdessadak et al., “Effects of dust on solar panel performance,” Proc. Int. Conf. Adv. Eng. Tech., pp. 1–12, 2025.

[18] ATESS Power Technology, “ESS monitoring platform.” [Online]. Available: https://ess-server.atesspower.com/ (Accessed 2026).

[19] Open-Meteo, “Free weather APIs.” [Online]. Available: https://open-meteo.com/ (Accessed 2026).

# กรอบการประเมินความทนทานแบบกล่องเทาสำหรับระบบตรวจจับการบุกรุกบนอินเทอร์เน็ตของสรรพสิ่งแบบผสมผสานด้วยอัลกอริทึมพันธุกรรมรูปแบบการค้นหาเชิงกว้าง

# A Gray-Box Robustness Evaluation Framework for Hybrid IoT Intrusion Detection Systems Using a Global Search Genetic Algorithm

*กนกนิภา ปัญเศษ (Kanoknipha Panset), นภัสวรรณ ชัยบาล (Napatwan Chaiban)*
*และจักรชัย โสอินทร์ (Chakchai So-In)*
*วิทยาลัยการคอมพิวเตอร์ มหาวิทยาลัยขอนแก่น*
*kanoknipha.p@kkumail.com, napatwan.c@kkumail.com, chakso@kku.ac.th*

## บทคัดย่อ

*งานวิจัยนี้นำเสนอกรอบการประเมินความทนทานของระบบตรวจจับการบุกรุกสำหรับอินเทอร์เน็ตของสรรพสิ่งภายใต้การโจมตีแบบกล่องเทาเพื่อสะท้อนข้อจำกัดของผู้โจมตีในสภาวะจริงและเติมเต็มช่องว่างงานวิจัยด้านความทนทานของระบบตรวจจับแบบผสมผสาน โดยได้พัฒนาระบบตรวจจับแบบไฮบริดที่ผสานโมเดล LightGBM (Light Gradient Boosting Machine) และ Isolation Forest เพื่อเพิ่มประสิทธิภาพการตรวจจับ พร้อมทั้งประยุกต์ใช้อัลกอริทึมพันธุกรรมแบบค้นหาเชิงกว้าง ในการสร้างตัวอย่างข้อมูลโจมตีภายใต้ข้อจำกัดขอบเขตการรบกวน เพื่อทดสอบความเปราะบางของระบบบนชุดข้อมูล IoT-23 และ Bot-IoT อย่างเป็นระบบ จากผลการทดสอบพบว่าระบบมีประสิทธิภาพสูงบนข้อมูลปกติ และความทนทานต่อการโจมตีลดลงอย่างชัดเจน โดยบนชุดข้อมูล IoT-23 วิธีอัลกอริทึมพันธุกรรมแบบค้นหาเชิงกว้างให้อัตราความสำเร็จของการโจมตีสูงถึง 0.9988 และส่งผลให้ค่าความถูกต้องภายใต้การโจมตีลดลงเหลือเพียง 0.0012 ขณะที่ชุดข้อมูล Bot-IoT มีอัตราความสำเร็จของการโจมตีที่ 0.4925 และมีค่าความถูกต้องภายใต้การโจมตีเท่ากับ 0.5075 ทั้งนี้ผลการวิจัยชี้ให้เห็นว่า การประเมินประสิทธิภาพบนข้อมูลปกติเพียงอย่างเดียวไม่เพียงพอที่จะสะท้อนความทนทานของระบบต่อการโจมตี นอกจากนี้ผลการทดลองของชุดข้อมูล Bot-IoT ควรตีความด้วยความระมัดระวัง และยังไม่ควรตีความเชิงสาเหตุ*

**คำสำคัญ:** ระบบตรวจจับการบุกรุก, อินเทอร์เน็ตของสรรพสิ่ง, ระบบตรวจจับแบบผสมผสาน, ความทนทานต่อการโจมตี, การโจมตีแบบกล่องเทา

## Abstract

*This study proposes a gray-box robustness evaluation framework for intrusion detection systems in the Internet of Things (IoT). The framework is designed to reflect practical adversarial constraints and to address the limited robustness assessment of hybrid intrusion detection systems. The proposed framework integrates LightGBM (Light Gradient Boosting Machine) and Isolation Forest to improve detection capability and employs a global-search genetic algorithm (GSGA) to generate adversarial samples under perturbation-budget constraints. Experiments on the IoT-23 and Bot-IoT datasets show that the proposed system achieves high performance on benign data but exhibits substantial degradation under attack. On IoT-23, GSGA achieves an attack success rate of 0.9988 and reduces robust accuracy to 0.0012. On Bot-IoT, the attack success rate is 0.4925, while robust accuracy remains 0.5075. These results indicate that benign-data performance alone is*

*insufficient for characterizing IDS robustness against adversarial attacks. However, the Bot-IoT findings should be interpreted cautiously, and further investigation is required before drawing causal conclusions.*

**Keywords:** IoT IDS, Hybrid Intrusion Detection, Adversarial Robustness, Gray-box Attack, GSGA, LightGBM, Robust Accuracy, Attack Success Rate.

## 1. บทนำ

การเติบโตของ Internet of Things (IoT) ทำให้เครือข่ายสมัยใหม่เชื่อมต่ออุปกรณ์จำนวนมากเข้ากับระบบงานจริงในภาคอุตสาหกรรม เมืองอัจฉริยะ การแพทย์ และโครงสร้างพื้นฐานสำคัญ แม้ IoT จะเพิ่มความสะดวกและประสิทธิภาพในการใช้งาน แต่ก็เพิ่มพื้นผิวการโจมตี (attack surface) การโจมตีและความเสี่ยงด้านความมั่นคงปลอดภัยตามไปด้วย โดยเฉพาะในสภาพแวดล้อมที่อุปกรณ์มีข้อจำกัดด้านทรัพยากรและการปรับปรุงระบบไม่สม่ำเสมอ [1]–[3]

ด้วยเหตุนี้ ระบบตรวจจับการบุกรุกจึงยังคงเป็นกลไกหลักในการเฝ้าระวังเครือข่าย IoT โดยเฉพาะในบริบทที่อุปกรณ์ปลายทางมีข้อจำกัดด้านทรัพยากรและไม่สามารถติดตั้งกลไกป้องกันเชิงลึกได้อย่างเต็มรูปแบบ ซึ่งงานวิจัยจำนวนมากได้นำ Machine Learning และ Deep Learning มาใช้เพื่อเพิ่มความสามารถในการตรวจจับพฤติกรรมที่ผิดปกติจากข้อมูลเครือข่าย โดยเฉพาะแบบจำลองที่เหมาะกับข้อมูลเชิงตาราง เช่น LightGBM ซึ่งให้ผลดีทั้งในด้านความแม่นยำและประสิทธิภาพในการประมวลผล [4]–[7] อย่างไรก็ตาม ความแม่นยำที่สูงบนข้อมูลปกติไม่ได้หมายความว่าระบบจะมีความมั่นคงเพียงพอเมื่อถูกนำไปใช้ในสถานการณ์จริง

กล่าวอีกนัยหนึ่ง ผลการตรวจจับที่ดีบนข้อมูลปกติไม่อาจรับประกันได้ว่าระบบจะยังคงเชื่อถือได้เมื่อเผชิญกับข้อมูลที่ถูกออกแบบเพื่อหลบเลี่ยงการตรวจจับ ประเด็นนี้ได้รับการยืนยันอย่างต่อเนื่องในงานวิจัยด้าน adversarial machine learning สำหรับระบบตรวจจับการบุกรุก ซึ่งพบว่าแม้มีการปรับเปลี่ยนข้อมูลนำเข้าเพียงเล็กน้อย ก็อาจทำให้แบบจำลองตัดสินใจผิดพลาดได้อย่างมีนัยสำคัญ ทั้งงานที่มุ่งโจมตี botnet detectors งานที่สร้างตัวอย่างโจมตีด้วย Generative Adversarial Networks (GAN) และงานที่ศึกษาการป้องกันระบบตรวจจับการบุกรุก (Intrusion Detection System: IDS) จาก evasion attacks ต่างชี้ให้เห็นสอดคล้องกันว่าแบบจำลองที่ให้ผลดีบนข้อมูลปกติอาจยังมีความเปราะบางสูงเมื่อเผชิญกับผู้โจมตีที่ออกแบบข้อมูลเพื่อหลบการตรวจจับ [8]–[10]

จากข้อจำกัดของงานที่ผ่านมา สามารถสรุปช่องว่างวิจัยหลักได้ 3 ประการ ได้แก่ (1) การประเมินจำนวนมากยังอาศัยสมมติฐานแบบ white-box ซึ่งอาจไม่สอดคล้องกับการใช้งานจริง (2) งานวิจัยจำนวนมากรายงานเพียงความสำเร็จของการโจมตีโดยไม่เชื่อมโยงกับ perturbation budget และต้นทุนเชิงเวลา อย่างเป็นระบบ และ (3) ในบางครั้งมีการประเมินผลเพียงชุดข้อมูลเดียว จึงยังสรุประดับความทนทานของระบบในภาพรวมได้ไม่ชัดเจนมากนัก [11]–[13]

แม้งานวิจัยด้าน adversarial attacks ต่อระบบตรวจจับการบุกรุก (IDS) จะมีจำนวนเพิ่มขึ้นอย่างต่อเนื่อง แต่การประเมินส่วนใหญ่ยังอาศัยสมมติฐานแบบ white-box หรือแบบ black-box เป็นหลัก ขณะที่แบบ gray-box สะท้อนเงื่อนไขการใช้งานจริงได้ใกล้เคียงกว่า เนื่องจากผู้โจมตีอาจรับรู้ข้อมูลบางส่วนของระบบและใช้ผลตอบกลับจากแบบจำลองเพื่อปรับการโจมตีได้ แม้ไม่สามารถเข้าถึงโครงสร้างภายในหรือพารามิเตอร์ทั้งหมดโดยตรง ดังนั้นงานวิจัยนี้จึงเลือกประเมินภายใต้แบบ gray-box เพื่อสะท้อนสถานการณ์การโจมตีที่ใกล้เคียงการใช้งานจริงมากยิ่งขึ้น

เมื่อพิจารณาถึงการสร้างข้อมูลโจมตี งานวิจัยนี้เลือกใช้ Global Search Genetic Algorithm (GSGA) เนื่องจากเหมาะสมกับการโจมตีภายใต้ฉากทัศน์แบบ gray-box และไม่ต้องอาศัยค่า gradient ของแบบจำลองโดยตรง อีกทั้งยังสามารถค้นหาคำตอบภายใต้ข้อจำกัดของ perturbation budget และคุณลักษณะที่ปรับเปลี่ยนได้ จึงเหมาะกับระบบตรวจจับการบุกรุกที่ทำงานกับข้อมูลเชิงตาราง และมีส่วนของการตรวจจับความผิดปกติร่วมด้วย

เพื่อตอบช่องว่างดังกล่าว งานวิจัยนี้มีส่วนสนับสนุนหลัก 3 ประการ ได้แก่ (1) การพัฒนาระบบตรวจจับการบุกรุกแบบผสมผสานที่ผสาน LightGBM และ Isolation Forest เพื่อเพิ่มความครอบคลุมของการตรวจจับ (2) การเสนอกรอบการสร้างตัวอย่างโจมตีแบบ gray-box ด้วย GSGA ภายใต้ข้อจำกัดที่สมเหตุสมผลและสอดคล้องกับการใช้งานจริง และ (3) การประเมินระบบด้วยตัวชี้วัดด้านประสิทธิภาพ ความทนทาน และต้นทุนเชิงเวลา บนชุดข้อมูลที่มีโครงสร้างแตกต่างกัน เพื่อสะท้อนทั้งขีดความสามารถและข้อจำกัดของระบบได้อย่างรอบด้าน

## 2. งานวิจัยที่เกี่ยวข้อง

งานวิจัยที่เกี่ยวข้องกับการศึกษานี้แบ่งได้เป็น 3 กลุ่มหลัก ได้แก่ งานด้าน IoT intrusion detection ที่ใช้ Machine Learning หรือ Deep Learning งานด้าน adversarial attacks ต่อ IDS และงานที่มุ่งประเมิน ความทนทานภายใต้เงื่อนไขที่ใกล้เคียงการใช้งานจริง

### 2.1 งานด้าน IoT Intrusion Detection ที่ใช้ Machine Learning และ Deep Learning

ในช่วงหลายปีที่ผ่านมา งานด้าน intrusion detection สำหรับเครือข่าย IoT นิยมใช้ Machine Learning และ Deep Learning มากขึ้น เนื่องจากสามารถเรียนรู้รูปแบบทราฟฟิกที่ซับซ้อนและปรับตัวได้ดีกว่าวิธีที่อาศัยกฎหรือลายเซ็น งานของ Albulayhi และ Abu Al-Haija [1] ชี้ว่าแบบจำลองเชิงลึกช่วยเพิ่มประสิทธิภาพของ anomaly-based IDS ได้อย่างมีนัยสำคัญ ขณะที่งานอื่น ๆ เสนอแนวทาง เช่น reinforcement learning และ early classification เพื่อเพิ่มความรวดเร็วและความยืดหยุ่นของระบบ [6], [13]

สำหรับข้อมูลเครือข่ายเชิงตาราง แบบจำลองในกลุ่ม gradient boosting โดยเฉพาะ LightGBM ยังคงได้รับความนิยม เนื่องจากให้ความแม่นยำสูงและประมวลผลได้มีประสิทธิภาพ [9] อย่างไรก็ตาม แม้งานในกลุ่มนี้จะแสดงศักยภาพในการตรวจจับการบุกรุกบนเครือข่าย IoT แต่ส่วนใหญ่ยังประเมินผลบนข้อมูลปกติเป็นหลัก จึงยังไม่เพียงพอสำหรับอธิบายความมั่นคงของระบบเมื่อเผชิญข้อมูลที่ถูกดัดแปลงโดยผู้โจมตี

### 2.2 งานด้าน Adversarial Attacks ต่อ Intrusion Detection Systems

เมื่อแนวคิด adversarial machine learning ถูกนำมาประยุกต์กับระบบ IDS งานวิจัยจำนวนมากแสดงให้เห็นว่าแบบจำลองที่มีประสิทธิภาพสูงบนข้อมูลปกติยังคงเปราะบางต่อการโจมตีแบบ evasion โดย Pawlicki และคณะ [3] รายงานว่า adversarial perturbations สามารถลดประสิทธิภาพของ network intrusion detection systems ได้อย่างชัดเจน ขณะที่ Apruzzese และคณะ [7] พบผลในลักษณะเดียวกันในบริบทของ botnet detection

อีกแนวทางหนึ่งคือ การสร้างตัวอย่างโจมตีด้วย generative models เช่น IDSGAN ของ Lin และคณะ [8] ซึ่งแสดงให้เห็นว่าสามารถสร้าง malicious traffic เพื่อหลบเลี่ยงระบบตรวจจับได้ แม้ไม่เข้าถึงโครงสร้างภายในของโมเดลโดยตรง งานเหล่านี้ยืนยันว่า IDS ที่อาศัย Machine Learning ยังมีช่องโหว่สำคัญ อย่างไรก็ตาม หลายงานยังเน้นอัตราความสำเร็จของการโจมตีมากกว่าความสมจริงของตัวอย่างโจมตี โดยเฉพาะในบริบทข้อมูลเครือข่ายที่คุณลักษณะบางส่วนไม่สามารถปรับเปลี่ยนได้อย่างอิสระ [2], [7], [8]

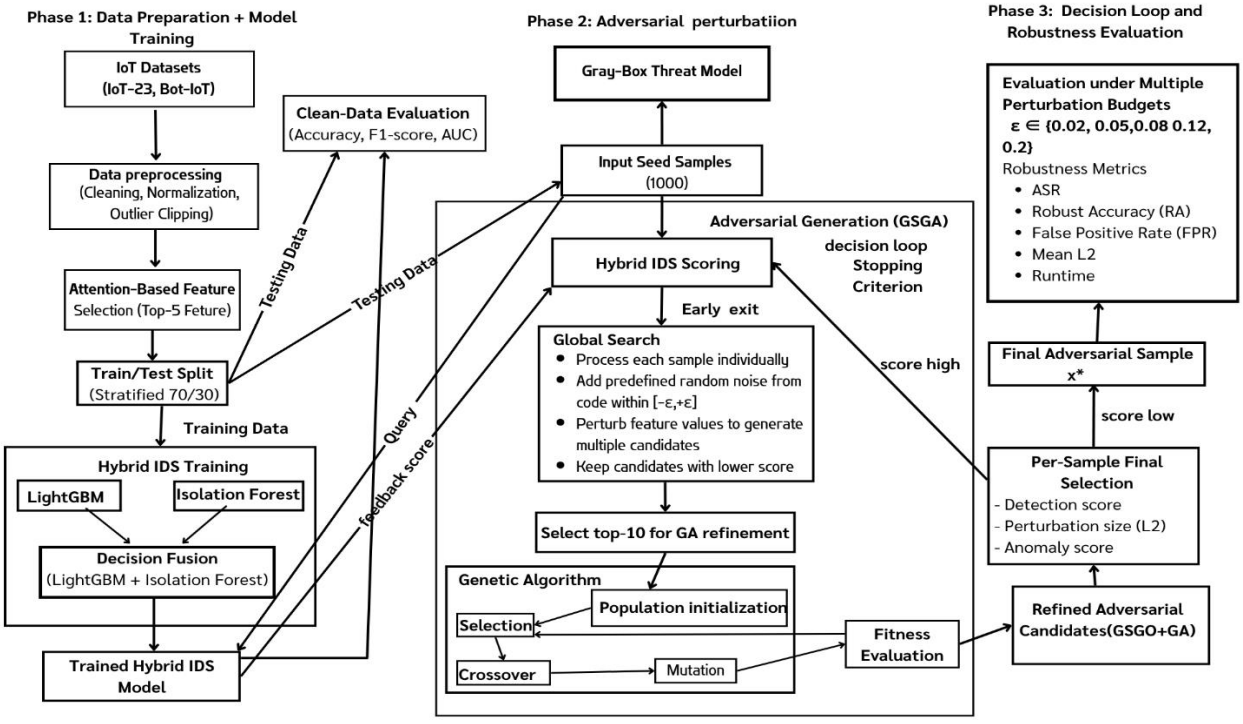
### 2.3 งานด้าน Robustness Evaluation ภายใต้เงื่อนไขที่สมจริง

นอกเหนือจากการพิสูจน์ว่า IDS สามารถถูกโจมตีได้ งานวิจัยในระยะหลังเริ่มให้ความสำคัญกับความสมจริงของ adversarial evaluation มากขึ้น งานของ Vitorino และคณะ [2] เป็นงานสำคัญที่ชี้ว่า adversarial examples สำหรับ IoT intrusion detection ควรอยู่ภายใต้ข้อจำกัดที่สะท้อนทราฟฟิกจริง ไม่ใช่เพียงการเปลี่ยนค่าคุณลักษณะเพื่อให้โมเดลตัดสินใจผิดเท่านั้น

ตารางที่ 1 สรุปการเปรียบเทียบงานวิจัยที่เกี่ยวข้องโดยพิจารณาทั้งลักษณะของระบบตรวจจับ รูปแบบการโจมตี วิธีที่ใช้ และข้อจำกัดของงานที่ผ่านมา เพื่อให้เห็นภาพรวมของแนวทางที่มีอยู่และตำแหน่งของงานวิจัยนี้ในบริบทของวรรณกรรมปัจจุบัน

**ตารางที่ 1** เปรียบเทียบงานวิจัยที่เกี่ยวข้อง

| | IDS | Attack | Method | Metrics | Gap |
|---|---|---|---|---|---|
| **Albulayhi & Abu Al-Haija (2023) [1]** | Anomaly (DL) | White-box | FGSM, JSMA, CW | Accuracy, F1 | ไม่ได้กำหนดเงื่อนไขการใช้งานจริงอย่างชัดเจน และไม่ได้ประเมินความทนทานโดยตรง |
| **Vitorino et al. (2023) [2]** | ML (RF, XGB, LGBM, IFOR) | Gray-box | A2PM | Accuracy | ไม่ได้ประเมินสถาปัตยกรรมแบบผสมผสาน และยังไม่ได้เปรียบเทียบต้นทุนเชิงเวลาของหลายวิธีโจมตีอย่างเป็นระบบ |
| **Pawlicki et al. (2020) [3]** | ML-based | White-box | FGSM, BIM, PGD | Accuracy | มุ่งเน้นกลไกป้องกันมากกว่าการประเมินภายใต้ฉากทัศน์จริง, ยังไม่สะท้อนเงื่อนไขการใช้งานจริง |
| **Mwangi & Cheng (2024) [4]** | ML-based | Black-box | Binary diversification | Evasion rate | ไม่ได้มุ่งประเมิน network IDS โดยตรง แต่เป็นงานด้าน adversarial attack ต่อ IoT malware detection ภายใต้ black-box setting |
| **Alaliwat et al. (2025) [5]** | AI models on IoT devices | Evasion attacks | Critical review / discussion of detection approaches | – | มุ่งวิเคราะห์และอภิปรายการตรวจจับการโจมตีแบบ evasion มากกว่าการประเมินกรอบโจมตีเชิงทดลองต่อการผสานเครือข่ายระบบตรวจจับการบุกรุก |
| **Ours** | **Hybrid (LGBM+IF)** | **Gray-box** | **GSGA** | **ASR, Robust Accuracy, F1** | **ครอบคลุมทั้งความสมจริงของฉากทัศน์และการประเมินความทนทาน** |




**ภาพที่ 1:** ขั้นตอนการฝึกระบบตรวจจับแบบผสมผสาน การสร้างตัวอย่างโจมตีด้วย GSGA และการประเมินความทนทาน

จากการทบทวนวรรณกรรม พบว่างานที่ผ่านมาสะท้อนทั้งศักยภาพและข้อจำกัดของระบบตรวจจับการบุกรุกสำหรับเครือข่าย IoT ภายใต้การโจมตีเชิงปฏิปักษ์ แม้จะเริ่มมีงานที่คำนึงถึงความสมจริงของการประเมินมาก

ขึ้น แต่ยังมีข้อจำกัดในด้านรูปแบบฉากทัศน์ที่ใช้ ระบบตรวจจับที่ประเมิน และความครอบคลุมของตัวชี้วัดที่รายงาน ข้อสังเกตดังกล่าวเป็นพื้นฐานสำคัญที่นำไปสู่การออกแบบวิธีการวิจัยในงานนี้

## 3. วิธีการดำเนินงาน

### 3.1 ภาพรวมของกระบวนการวิจัย

ดังแสดงในภาพที่ 1 วิธีดำเนินงานของการวิจัยแบ่งออกเป็น 3 ระยะ ได้แก่ ระยะที่ 1 คือการเตรียมข้อมูลและฝึกระบบตรวจจับแบบผสมผสาน ซึ่งถูกนำไปใช้ในระยะที่ 2 สำหรับการสร้างข้อมูลโจมตีแบบ gray-box ด้วยอัลกอริทึมพันธุกรรมแบบค้นหาเชิงกว้าง และในระยะที่ 3 สำหรับการประเมินผลด้านประสิทธิภาพ ความทนทาน และต้นทุนเชิงเวลา โดยรายละเอียดของแต่ละระยะแสดงในหัวข้อ 3.2–3.4 ตามลำดับ

### 3.2 ระยะที่ 1 การเตรียมข้อมูล และการฝึกระบบตรวจจับแบบผสมผสาน

งานวิจัยนี้ใช้ชุดข้อมูลเครือข่าย IoT 2 ชุด ได้แก่ IoT-23 และ Bot-IoT ซึ่งมีโครงสร้างข้อมูลและระดับความไม่สมดุลของคลาสแตกต่างกัน เพื่อให้สามารถประเมินความทนทานของระบบในบริบทที่หลากหลายได้ ตารางที่ 2 แสดงสถิติของชุดข้อมูลที่ใช้ในการทดลอง โดย IoT-23 มีทั้งข้อมูลปกติและข้อมูลโจมตีในสัดส่วนที่แตกต่างจาก Bot-IoT อย่างชัดเจน ซึ่งเอื้อให้เกิดการวิเคราะห์เปรียบเทียบพฤติกรรมของระบบภายใต้ลักษณะข้อมูลที่ไม่เหมือนกันได้

**ตารางที่ 2:** สถิติชุดข้อมูล

| Dataset | Total | Normal | Attack |
|---|---|---|---|
| **IoT-23** | 6,046,623 | 688,812 | 5,357,811 |
| **Bot-IoT** | 1,927,101 | 477 | 1,926,624 |

เนื่องจาก Bot-IoT มีจำนวน normal samples ต่ำมากเมื่อเทียบกับ attack samples การประเมินผลจึงอาจอ่อนไหวต่อวิธีแบ่งข้อมูลและเกณฑ์การตัดสินใจ

งานวิจัยนี้จึงตีความผลของชุดข้อมูลดังกล่าวในเชิงระมัดระวังข้อมูลแต่ละชุดถูกแบ่งเป็นชุดฝึกและชุดทดสอบในอัตราส่วน 70:30 โดยคงสัดส่วนของคลาสเดิมไว้เพื่อลดผลกระทบจากความไม่สมดุลของข้อมูล จากนั้นจึงปรับข้อมูลเชิงตัวเลขด้วยการตัดค่าที่อยู่นอกช่วงเปอร์เซ็นไทล์ที่ 1 ถึง 99 และปรับให้อยู่ในช่วง [0, 1] ด้วยวิธีปรับค่าน้อยที่สุด–มากที่สุด โดยพารามิเตอร์ทั้งหมดสำหรับการปรับข้อมูลถูกประมาณจากชุดฝึกเท่านั้น เพื่อลดความเสี่ยงของการรั่วไหลของข้อมูล

ภายหลังการปรับสเกลข้อมูล ได้ใช้ตัวเข้ารหัสอัตโนมัติแบบอาศัยกลไกความสนใจเพื่อจัดอันดับความสำคัญของคุณลักษณะ และเลือกเฉพาะคุณลักษณะที่ดีที่สุดมา 5 ตัวสำหรับใช้ในการฝึกแบบจำลองและการสร้างข้อมูลโจมตี เพื่อลดภาระการคำนวณและจำกัดพื้นผิวการโจมตีให้อยู่ในขอบเขตที่สามารถควบคุมได้ ทั้งนี้งานวิจัยไม่ได้มุ่งอ้างว่าจำนวนดังกล่าวเป็นค่าที่เหมาะสมที่สุดในทุกกรณี

ระบบตรวจจับที่เสนอประกอบด้วยแบบจำลอง 2 ส่วน ได้แก่ LightGBM และ Isolation Forest ซึ่งทำงานร่วมกันเพื่อเพิ่มประสิทธิภาพในการจำแนกและการตรวจจับความผิดปกติ โดย LightGBM สามารถเขียนในรูปผลรวมของฟังก์ชันการเรียนรู้ย่อยได้ดังสมการที่ (1)

$$\hat{y}_i = \sum_{m=1}^{M} f_m(x_i) \qquad (1)$$

โดยที่ $\hat{y}_i$ คือค่าที่แบบจำลองทำนายสำหรับตัวอย่างที่ $i$, $M$ คือจำนวนแบบจำลองย่อยทั้งหมดและ $f_m(x_i)$ คือผลลัพธ์จากแบบจำลองย่อยลำดับที่ $m$ สำหรับข้อมูลตัวอย่าง $x_i$

สำหรับส่วนตรวจจับความผิดปกติ Isolation Forest ใช้หลักการแยกข้อมูลที่มีลักษณะเบี่ยงเบนออกจากข้อมูลทั่วไปเพื่อประเมินระดับความผิดปกติ ซึ่งสามารถเขียนเป็นคะแนนได้ดังสมการที่ (2)

$$s(x,n) = 2^{-\frac{E(h(x))}{c(n)}} \qquad (2)$$

โดยที่ $s(x,n)$ คือคะแนนความผิดปกติของตัวอย่าง $x$, $E(h(x))$ คือค่าเฉลี่ยของความยาวเส้นทางที่ใช้ในการแยกตัวอย่างดังกล่าวออกจากโครงสร้างการสุ่มแบ่งข้อมูลหลายชุด, $n$ คือจำนวนตัวอย่างในชุดข้อมูลและ $c(n)$ คือค่าปรับมาตรฐานของความยาวเส้นทางสำหรับข้อมูลปกติ

ผลลัพธ์ จากทั้งสองแบบจำลองจะถูกนำมารวมกันเป็นคะแนนรวมของระบบตามสมการที่ (3)

$$S(x) = \alpha \cdot P_{LGBM}(x) + (1 - \alpha) \cdot \tilde{A}_{IF}(x) \quad (3)$$

โดยที่ $P_{LGBM}(x)$คือความน่าจะเป็นที่ตัวอย่าง $x$ถูกจำแนกเป็นข้อมูลโจมตีโดยโมเดล LightGBM, $\tilde{A}_{IF}(x)$ คือคะแนนความผิดปกติจาก Isolation Forest หลังปรับให้อยู่ในช่วง $[0, 1]$, และ $\alpha \in [0,1]$คือค่าน้ำหนักสำหรับการรวมคะแนนจากทั้งสองโมเดล

ในงานวิจัยนี้กำหนดค่าน้ำหนักการรวมคะแนนเป็น $\alpha = 0.60$ และเกณฑ์การตัดสินใจเป็น $\tau = 0.50$ โดยอาศัยผลจากการทดลองบนชุดตรวจสอบเบื้องต้น เพื่อให้สมดุลระหว่างความไวต่อความผิดปกติของ Isolation Forest และความสามารถในการจำแนกของ LightGBM นอกจากนี้ หากคะแนนรวม $S(x)$ อยู่ในช่วงใกล้เส้นแบ่งการตัดสินใจ [0.38, 0.62] ระบบจะใช้ผลจาก LightGBM เป็นผู้ช่วยตัดสินใจขั้นสุดท้ายเพื่อลดความกำกวมของผลลัพธ์

### 3.3 ระยะที่ 2 การสร้างตัวอย่างโจมตีแบบ gray-box ด้วย GSGA

จากขั้นตอนที่แล้ว เมื่อได้ระบบตรวจจับแบบผสมผสานแล้ว งานวิจัยจะเลือกตัวอย่างโจมตีจากชุดทดสอบมาใช้เป็นข้อมูลนำเข้าสำหรับการโจมตีแบบ gray-box โดยมีเป้าหมายเพื่อสร้างการรบกวนที่ทำให้ระบบจำแนกข้อมูลโจมตีเป็นข้อมูลปกติได้สำเร็จภายใต้งบประมาณการรบกวนที่กำหนด และยังคงความสมจริงของข้อมูลไว้เพื่อสะท้อนข้อจำกัดของสภาพแวดล้อมจริง

ในงานวิจัยนี้ ผู้โจมตีไม่สามารถเข้าถึงโครงสร้างภายในหรือค่าความชันของแบบจำลองได้โดยตรง แต่สามารถอาศัยผลตอบกลับจากระบบในการปรับการรบกวนได้ ดังนั้นจึงเลือกใช้ GSGA ซึ่งเหมาะกับ gray-box และไม่จำเป็นต้องใช้ค่าความชันของแบบจำลองโดยตรง

กำหนดให้การค้นหาเป็นปัญหาแบบ minimization โดยค่าฟังก์ชันที่ต่ำกว่าหมายถึงตัวอย่างที่มีแนวโน้มหลบเลี่ยงการตรวจจับได้ดีขึ้น ภายใต้ข้อจำกัดของขนาดการรบกวนและความสอดคล้องเชิงปฏิบัติ ฟังก์ชันความเหมาะสมของการรบกวนกำหนดดังสมการที่ (4)

$$F(\delta) = S(x + \delta) + \lambda_1 \cdot \|\delta\|_2 - \lambda_2 \cdot A(x + \delta) \quad (4)$$

โดยที่ $F(\delta)$ คือค่าความเหมาะสมของการรบกวน, $S(x + \delta)$ คือคะแนนตอบกลับจากระบบหลังการรบกวน, $\lambda_1 \| \delta \|_2$ คือเทอมลงโทษด้านขนาดของการรบกวน และ $\lambda_2 A(x + \delta)$ คือ เทอมที่ใช้สะท้อนความสอดคล้องเชิงปฏิบัติของตัวอย่างหลังการรบกวนกับข้อจำกัดของคุณลักษณะที่แก้ไขได้และขอบเขตค่าที่พบได้จริงในข้อมูล

---

**Algorithm 1: Global Search Genetic Algorithm (GSGA) for Gray-box Attack**

---

**Input:** attack sample $x$, perturbation budget $\varepsilon$, population size $N$, global-search iterations $T_1$, GA generations $T_2$, elite size $k$, mutation rate $\mu$, initial step size $\gamma_0$
**Output:** adversarial sample $x^*$

```
Initialize {δ_i}_{i=1}^N randomly on editable features
for each δ_i do
    Project δ_i to feasible set ΠFε and compute F(δ_i)
end for
for t = 1 to T1 do
    γ_t = γ0(1 - t/T1)
    for each candidate δ_i do
        Sample η_t ~ U(-1,1) or N(0,I)
        δ'_i = ΠFε(δ_i + γ_t η_t)
        Compute F(δ'_i)
        if F(δ'_i) < F(δ_i) then
            δ_i = δ'_i
        end if
        if x + δ_i is classified as benign then
            return x + δ_i
        end if
    end for
end for
Select top-k candidates with lowest fitness
Initialize GA population from selected candidates
for g = 1 to T2 do
    Select parents by tournament selection
    Generate offspring by uniform crossover
    Apply Gaussian mutation with rate μ on editable features
    Project each offspring to ΠFε and compute fitness
    Update population with elites and best offspring
    if any offspring δ_j makes x + δ_j classified as normal then
        return x + δ_j
    end if
end for
return feasible candidate with lowest fitness
```

---

อัลกอริทึม GSGA แบ่งเป็น 2 ช่วง ได้แก่ global search และ genetic refinement ในช่วงแรก จะสุ่มประชากรตั้งต้นของเวกเตอร์การรบกวนบนคุณลักษณะที่อนุญาตให้แก้ไขได้ แล้วฉายคำตอบกลับเข้าสู่เซตที่เป็นไปได้ทุกครั้งที่มีการปรับปรุง จากนั้นจึงปรับคำตอบด้วยก้าวการค้นหาแบบสุ่มที่มีขนาดลดลงตามรอบการค้นหา เพื่อระบุบริเวณคำตอบที่มีศักยภาพ เมื่อได้คำตอบที่ดีแล้ว จึงคัดเลือกคำตอบที่ดีที่สุดจำนวน $k$ ตัวเข้าสู่ช่วงปรับละเอียดด้วยพันธุกรรม ซึ่งใช้การคัดเลือกแบบทัวร์นาเมนต์ การผสมแบบสม่ำเสมอ และการกลายพันธุ์แบบเกาส์เซียน เพื่อปรับคำตอบภายใต้ข้อจำกัดเดิม ทั้งนี้การค้นหาจะหยุดเมื่อได้ตัวอย่างที่ทำให้ระบบตัดสินผิดตามเงื่อนไขที่กำหนด หรือเมื่อจำนวนรอบการค้นหาครบตามค่าที่ตั้งไว้ โดยสรุปขั้นตอนของอัลกอริทึมแสดงใน Algorithm 1

### 3.4 ระยะที่ 3 การประเมินผลและการเปรียบเทียบ

ระยะสุดท้ายเป็นการนำตัวอย่างที่สร้างขึ้นกลับมาทดสอบกับระบบตรวจจับเดิม เพื่อประเมินทั้งประสิทธิภาพบนข้อมูลปกติ ความทนทานภายใต้การโจมตี ผลของขอบเขตการรบกวน และต้นทุนเชิงเวลาภายใต้เงื่อนไขที่เปรียบเทียบกันได้ในทุกวิธี ในการวิจัยนี้อัตราการตรวจจับหลังการโจมตีและความถูกต้องภายใต้การโจมตีเป็นคนละตัวชี้วัด โดยอัตราการตรวจจับหลังการโจมตีคำนวณบนชุดตัวอย่างโจมตีทั้งหมดในชุดทดสอบ ขณะที่ความถูกต้องภายใต้การโจมตีคำนวณบนชุดตัวอย่างที่ใช้ประเมินการโจมตีโดยตรง ดังนั้น ค่าทั้งสองจึงไม่จำเป็นต้องเท่ากันและไม่ควรเปรียบเทียบกันโดยตรงในเชิงตัวเลข

ตัวชี้วัดด้านประสิทธิภาพบนข้อมูลปกติประกอบด้วย F1-score และ Matthews correlation coefficient (MCC) ซึ่งใช้สะท้อนทั้งความแม่นยำเชิงสมดุลและคุณภาพของการจำแนกภายใต้ข้อมูลที่มีความไม่สมดุลของคลาส ดังสมการที่ (5)–(6)

$$F1-\text{score} = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (5)$$

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}} \quad (6)$$

โดยที่ Precision คือสัดส่วนของข้อมูลที่ระบบทำนายว่าเป็นการโจมตีและถูกต้องจริง และ Recall คือสัดส่วนของข้อมูลโจมตีที่ระบบสามารถตรวจจับได้ ส่วนค่าสหสัมพันธ์แมตทิวส์เป็นตัวชี้วัดที่พิจารณาทุกองค์ประกอบของตารางสับสน จึงเหมาะกับข้อมูลที่มีความไม่สมดุลของคลาส ขณะที่ตัวชี้วัดด้านความทนทานประกอบด้วย Attack Success Rate (ASR), Robust Accuracy และ Absolute Detection Drop ดังสมการที่ (7)–(9)

$$ASR(\epsilon) = \frac{N_{success}}{N_{attack}} \quad (7)$$

$$Robust\ Accuracy = 1 - ASR(\varepsilon) \quad (8)$$

$$\Delta\text{DR} = DR_{\text{before}} - DR_{\text{after}} \quad (9)$$

สมการ (7) และ (8) ถูกคำนวณบนชุดตัวอย่างที่ใช้ประเมินการโจมตีโดยตรง ขณะที่สมการ (9) ใช้สะท้อนการลดลงของอัตราการตรวจจับในระดับระบบ จึงไม่ควรตีความว่าทั้งสามตัวชี้วัดเป็นค่าที่เปรียบเทียบกันได้โดยตรงเชิงตัวเลข

โดยที่ $N_{\text{success}}$ คือ จำนวนการโจมตีที่หลบเลี่ยงระบบได้สำเร็จ, $N_{\text{attack}}$ คือจำนวนตัวอย่างโจมตีทั้งหมด, $DR_{\text{before}}$ คือ อัตราการตรวจจับก่อนการโจมตี และ $DR_{\text{after}}$ คือ อัตราการตรวจจับหลังการโจมตี ค่าทั้งสามสะท้อนทั้งความสำเร็จของผู้โจมตีและความเสื่อมถอยของระบบในเชิงปฏิบัติ

สำหรับการเปรียบเทียบวิธี งานวิจัยใช้ GSGA เปรียบเทียบกับ Standard Genetic Algorithm (GA), Global Search-only (GSGO), Differential Evolution (DE) และ CMA-ES ภายใต้ขอบเขตการรบกวน 5 ระดับ ได้แก่ {0.02, 0.05, 0.08, 0.12, 0.20} โดยทุกวิธีใช้คุณลักษณะที่อนุญาตให้แก้ไขได้และข้อกำหนดของขอบเขตการรบกวนเดียวกัน นอกจากนี้ ยังรายงานเวลาเฉลี่ยต่อหนึ่งตัวอย่างและเวลารวมของการประมวลผล เพื่อสะท้อนความคุ้มค่าเชิงการคำนวณของแต่ละวิธีอย่างเป็นระบบ

## 4. ผลการทดลองและอภิปรายผล

การวิจัยนำเสนอผลการประเมินระบบตรวจจับแบบผสมผสานทั้งในสภาวะปกติและภายใต้การโจมตีแบบ gray-box adversarial โดยมุ่งตอบคำถามหลัก 3 ประเด็น คือ ระบบมีประสิทธิภาพมากแค่ไหนบนข้อมูลปกติ ระบบยังคงทนทานหรือไม่เมื่อถูกโจมตี และความทนทานดังกล่าวเปลี่ยนไปอย่างไรเมื่อพิจารณาต่างชุดข้อมูล ต่างวิธีโจมตี และต่างระดับของ perturbation budget

**ตารางที่ 3:** ประสิทธิภาพบนข้อมูลปกติ

| Metric | IoT-23 | Bot-IoT |
|---|---|---|
| **F1-Score** | 0.9998 | 0.9981 |
| **ROC-AUC** | 1.0000 | 0.9999 |
| **PR-AUC** | 1.0000 | 1.0000 |
| **Balanced Accuracy** | 0.9997 | 0.9981 |
| **MCC** | 0.9980 | 0.2467 |
| **Precision** | 1.0000 | 1.0000 |
| **Recall** | 0.9996 | 0.9962 |
| **FPR** | 0.0001 | 0.0000 |

ตารางที่ 3 แสดงว่า ระบบตรวจจับแบบผสมผสานมีประสิทธิภาพบนข้อมูลปกติในระดับสูงมากทั้งสองชุดข้อมูล อย่างไรก็ตาม ในการวิจัยนี้ Precision, Recall, F1-score นี้ และ PR-AUC คำนวณโดยถือถื class การโจมตีเป็น positive class ในขณะที่ชุดข้อมูล Bot-IoT มีความไม่สมดุลของคลาสในระดับรุนแรงมาก ซึ่งอาจทำให้ตัวชี้วัดบางชนิด เช่น F1-score, ROC-AUC และ PR-AUC ยังคงมีค่าสูง แม้ความสมดุลของการจำแนกระหว่าง normal และ attack จะยังไม่ดีเพียงพอ โดยเฉพาะเมื่อพิจารณาค่า MCC ที่ต่ำกว่าชุดข้อมูล IoT-23 อย่างมาก ดังนั้น ผลในตารางนี้จึงควรถูกพิจารณาร่วมกับตัวชี้วัดรายคลาส เช่น benign recall, attack recall, specificity และ confusion matrix เพื่อสะท้อนพฤติกรรมของระบบได้ครบถ้วนยิ่งขึ้น

**ตารางที่ 4:** ขอบเขตขนาดการรบกวน

| ε | IoT-23 ASR | Bot-IoT ASR |
|---|---|---|
| **0.02** | 0.9450 | 0.4925 |
| **0.05** | 0.9912 | 0.4925 |
| **0.08** | 0.9962 | 0.4925 |
| **0.12** | 1.0000 | 0.4925 |
| **0.20** | 1.0000 | 0.4925 |

ตารางที่ 4 พบว่า เมื่อเพิ่มค่าขอบเขตการรบกวน ชุดข้อมูล IoT-23 มีค่า ASR สูงขึ้นอย่างต่อเนื่อง สะท้อนว่าการเพิ่มขอบเขตการรบกวน ช่วยให้ผู้โจมตีหลบเลี่ยงระบบได้มากขึ้น ในทางตรงกันข้าม ชุดข้อมูล Bot-IoT ให้ค่า ASR คงที่ในทุกระดับของ $\epsilon$ ซึ่งอาจสะท้อนข้อจำกัดของคุณลักษณะที่อนุญาตให้แก้ไขได้ ทำให้การรบกวนเพียงเล็กน้อยก็เพียงพอต่อการเปลี่ยนผลการจำแนก อีกทั้งยังเป็นไปได้ว่าเส้นแบ่งการตัดสินใจของแบบจำลองมีระยะแคบ หรืออัลกอริทึมที่ใช้สามารถค้นหาจุดเปราะบางได้ตั้งแต่ค่า $\epsilon$ ต่ำ จึงทำให้การเพิ่มค่า $\epsilon$ต่อจากนั้นไม่เพิ่มความสำเร็จของการโจมตีอย่างชัดเจน อย่างไรก็ตาม ข้อสังเกตดังกล่าวยังควรได้รับการตรวจสอบเพิ่มเติมก่อนสรุปเชิงกลไก

**ตารางที่ 5:** การเปรียบเทียบอัลกอริทึมการโจมตี

| Method | IoT-23 ASR | Bot-IoT ASR | IoT-23 Mean L2 Norm | Bot-IoT Mean L2 Norm | IoT-23 Post-attack Detection Rate | Bot-IoT Post-attack Detection Rate |
|---|---|---|---|---|---|---|
| **GA** | 1.0000 | 0.4925 | 0.10560 | 0.07927 | 0.3250 | 0.6278 |
| **GSGO** | 0.9975 | 0.4925 | 0.14329 | 0.06521 | 0.4080 | 0.6501 |
| **DE** | 1.0000 | 0.4925 | 0.07767 | 0.06728 | 0.2630 | 0.6267 |
| **CMA-ES** | 0.9950 | 0.4925 | 0.06509 | 0.04963 | 0.2840 | 0.6575 |
| **GSGA** | **0.9988** | **0.4925** | **0.13881** | **0.06414** | **0.3860** | **0.6299** |

ตารางที่ 5 แสดงให้เห็นว่า ชุดข้อมูล IoT-23 มีความเปราะบางต่อการโจมตีอย่างเด่นชัด โดยทุกวิธีให้อัตราความสำเร็จของการโจมตีอยู่ในระดับสูงมาก ขณะที่อัลกอริทึมที่เสนอให้ผลอยู่ในระดับใกล้เคียงกับวิธีมาตรฐาน แม้จะไม่สูงที่สุดในทุกตัวชี้วัด ทั้งนี้ ค่าอัตราการตรวจจับหลังการโจมตีในตารางนี้สะท้อนสมรรถนะของระบบบนชุดตัวอย่างโจมตีทั้งหมดภายหลังการรบกวน จึงเป็นคนละมิติกับค่าความถูกต้องภายใต้การโจมตีที่รายงานในตารางที่ 6 ซึ่งคำนวณบนชุดตัวอย่างที่ใช้ประเมินการโจมตีโดยตรง สำหรับชุดข้อมูล Bot-IoT ทุกวิธีให้อัตราความสำเร็จของการโจมตีเท่ากัน ซึ่งยังไม่เพียงพอสำหรับรองรับข้อสรุปเชิงกลไกของอัลกอริทึม และจำเป็นต้องมีการตรวจสอบเพิ่มเติมด้วยการทดลองซ้ำและการวิเคราะห์ราย

**ตารางที่ 6:** ความทนทานข้ามชุดข้อมูล

| Metric | IoT-23 | Bot-IoT |
|---|---|---|
| **F1-Score** | 0.9998 | 0.9981 |
| **ROC-AUC** | 1.0000 | 0.9999 |
| **ASR (GSGA)** | 0.9988 | 0.4925 |
| **Robust Accuracy** | 0.0012 | 0.5075 |
| **Mean L2 Norm** | 0.13881 | 0.06414 |

ตารางที่ 6 แสดงให้เห็นว่า แม้ระบบจะมีประสิทธิภาพบนข้อมูลปกติใกล้เคียงกันมากทั้งใน IoT-23 และ Bot-IoT แต่ผลด้านความทนทานแตกต่างกันอย่างชัดเจน โดยค่าความถูกต้องภายใต้การโจมตีในตารางนี้คำนวณบนชุดตัวอย่างที่ถูกนำมาใช้ประเมินการโจมตีโดยตรง จึงสะท้อนความสามารถของระบบในการคงผลการจำแนกที่ถูกต้องภายหลังการรบกวน และไม่ควรตีความว่าเป็นตัวชี้วัดเดียวกับอัตราการตรวจจับหลังการโจมตีในตารางที่ 5

**ตารางที่ 7:** ต้นทุนเชิงเวลา (Runtime Analysis)

| Method | IoT-23 Time/sample (s) | IoT-23 Total (s) | Bot-IoT Time/sample (s) | Bot-IoT Total (s) |
|---|---|---|---|---|
| **GSGO** | 0.04254 | 42.5 | 0.03419 | 32.2 |
| **GA** | 0.05925 | 59.3 | 0.05225 | 49.3 |
| **DE** | 0.60021 | 600.2 | 0.60832 | 573.6 |
| **CMA-ES** | 0.10109 | 101.1 | 0.0977 | 92.1 |
| **GSGA** | **0.12447** | **124.5** | **0.10494** | **99.0** |

ตารางที่ 7 ผลการทดลองสะท้อนว่า GSGA ให้สมดุลระหว่างความสามารถในการโจมตีกับต้นทุนเชิงเวลาในระดับหนึ่งสำหรับการประเมินแบบออฟไลน์ โดยเฉพาะเมื่อพิจารณาร่วมกับค่า ASR ที่อยู่ในระดับสูงบน IoT-23 ค่า Total runtime คำนวณจากชุดตัวอย่างโจมตีทั้งหมดที่ใช้ในการประเมินในแต่ละชุดข้อมูล อย่างไรก็ดี หากขยายการทดลองไปยังหลายชุดข้อมูลหรือหลายค่า perturbation budget ต้นทุนเชิงเวลายังคงเป็นข้อจำกัดที่ควรปรับปรุงในงานต่อไป

## 5. สรุปผลการทดลอง

งานวิจัยนี้นำเสนอกรอบการประเมินความทนทานของระบบตรวจจับการบุกรุก IoT แบบผสมผสานภายใต้การโจมตีแบบ gray-box โดยผสาน LightGBM และ Isolation Forest เข้ากับกระบวนการสร้าง adversarial samples ด้วย GSGA ผลการทดลองบน IoT-23 แสดงให้เห็นอย่างชัดเจนว่า แม้ระบบจะมีประสิทธิภาพบน ข้อมูลปกติสูงมาก แต่ยังมีความเปราะบางต่อ adversarial attack อย่างมาก โดยค่า ASR ของ GSGA สูงถึง 0.9988 และทำให้ Robust Accuracy ลดลงเหลือ 0.0012 สำหรับ Bot-IoT พบค่า ASR เท่ากับ 0.4925 และ Robust Accuracy เท่ากับ 0.5075 แต่ผลบน Bot-IoT ในงานวิจัยนี้ควรถูกมองเป็นผลเบื้องต้นที่ยังต้องการการตรวจสอบซ้ำทั้งในระดับการนำวิธีไปใช้จริง และยังควรได้รับการตรวจสอบเพิ่มเติมก่อนสรุปเชิงกลไกหรือเชิงเปรียบเทียบอย่างชัดเจน ดังนั้น งานวิจัยนี้จึงสนับสนุนข้อเสนอสำคัญว่าการประเมิน IDS ไม่ควรอาศัยเฉพาะตัวชี้วัดบนข้อมูลปกติ แต่ต้องรายงานตัวชี้วัดด้านความทนทานควบคู่กันด้วย

อย่างไรก็ตาม งานวิจัยนี้ยังมีข้อจำกัดหลายประการ ได้แก่ การสร้าง adversarial samples ยังอยู่ในระดับคุณลักษณะ การประเมินยังครอบคลุมเพียงสองชุดข้อมูล และผลการทดลองบางส่วน โดยเฉพาะบน Bot-IoT ยังควรได้รับการตรวจสอบซ้ำเพื่อยืนยันความคงตัวเชิงสถิติและความถูกต้องของระเบียบการประเมิน งานในอนาคตจึงควรมุ่งไปสู่การวิเคราะห์หลายค่าเริ่มต้นแบบสุ่มหลายชุด การรายงานค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐาน การ

เพิ่มตัวชี้วัดรายคลาส และการพัฒนา adversarial generation ในระดับแพ็กเกจ เพื่อสะท้อนการใช้งานจริงได้ดียิ่งขึ้น

## เอกสารอ้างอิง


[1] K. Albulayhi and Q. Abu Al-Haija, “Adversarial Deep Learning in Anomaly Based Intrusion Detection Systems for IoT Environments,” *International Journal of Wireless and Microwave Technologies*, vol. 13, no. 4, pp. 1–10, 2023.

[2] J. Vitorino, I. Praça, and E. Maia, “Towards Adversarial Realism and Robust Learning for IoT Intrusion Detection and Classification,” *Annals of Telecommunications*, vol. 78, pp. 401–412, 2023.

[3] M. Pawlicki, M. Choraś, and R. Kozik, “Defending Network Intrusion Detection Systems against Adversarial Evasion Attacks,” *Future Generation Computer Systems*, vol. 110, pp. 148–154, 2020.

[4] M. B. Mwangi and S.-M. Cheng, “An Adversarial Attack on ML-Based IoT Malware Detection Using Binary Diversification Techniques,” *IEEE Access*, vol. 12, pp. 185172–185186, 2024.

[5] F. Alaliwat, L. Alqahtani, M. Alzahrani, N. Alamoudi, S. Hakami, and A. Alharby, “Detection of Adversarial Evasion Attack on AI Model Running on IoT Devices,” in *2025 12th International Conference on Information Technology (ICIT)*, pp. 47–54, 2025.

[6] G. Caminero, M. Lopez-Martin, and B. Carro, “Adversarial Environment Reinforcement Learning Algorithm for Intrusion Detection,” *Computer Networks*, vol. 159, pp. 96–109, 2019.

[7] G. Apruzzese, M. Colajanni, and M. Marchetti, “Evaluating the Effectiveness of Adversarial Attacks against Botnet Detectors,” in *2019 IEEE 18th International Symposium on Network Computing and Applications (NCA)*, pp. 1–8, 2019.

[8] Z. Lin, Y. Shi, and Z. Xue, “IDSGAN: Generative Adversarial Networks for Attack Generation Against Intrusion Detection,” in *Advances in Knowledge Discovery and Data Mining (PAKDD 2022)*, vol. 13282, pp. 79–91, 2022.

[9] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, and T.-Y. Liu, “LightGBM: A Highly Efficient Gradient Boosting Decision Tree,” in *Advances in Neural Information Processing Systems*, pp. 3146–3154, 2017.

[10] A. Askhatuly, D. Berdysheva, A. Berdyshev, A. Adamova, and D. Yedilkhan, “Adversarial Attacks and Defense Mechanisms in Machine Learning: A Structured Review of Methods, Domains, and Open Challenges,” *IEEE Access*, vol. 13, pp. 185145–185168, 2025.

[11] W. Yao, H. Peng, Q. Li, and X. Shen, “Modeling Realistic Adversarial Traffic Against Deep Learning-Based Intrusion Detection System in Industrial IoT,” *IEEE Internet of Things Journal*, vol. 12, no. 15, pp. 29540–29554, 2025.

[12] J. Yu, A. V. Shvetsov, and S. H. Alsamhi, “Leveraging Machine Learning for Cybersecurity Resilience in Industry 4.0: Challenges and Future Directions,” *IEEE Access*, vol. 12, pp. 159579–159596, 2024.

[13] T. Kim and W. Pak, “Robust Network Intrusion Detection System Based on Machine-Learning With Early Classification,” *IEEE Access*, vol. 10, pp. 10754–10767, 2022.

# กรอบวิธีการแปลงภาพโฉนดที่ดินเป็นข้อมูลเวกเตอร์โดยใช้ YOLOv11, การรู้จำอักขระภาษาไทยด้วยแสงและการปรับพื้นที่ให้คงค่า

# A Framework for Converting Land Title Deed Images into Vector Data Using YOLOv11, Thai OCR, and Area-Preserving Adjustment

*วรวรรธ สร้อยเพชร (Worawat Soiphet)[1] และสุภัคพงศ์ จินารัตน์ (Supakpong Jinarat)[2]*
*สาขาวิชาเทคโนโลยีสารสนเทศ วิทยาลัยวิศวกรรมศาสตร์และเทคโนโลยี มหาวิทยาลัยธุรกิจบัณฑิตย์*
*67130635@dpu.ac.th[1], supakpong.jin@dpu.ac.th[2]*

## บทคัดย่อ

*เอกสารสิทธิที่ดินของประเทศไทยส่วนใหญ่จัดเก็บในรูปแบบภาพสแกน ซึ่งไม่สามารถใช้งานเชิงพื้นที่ได้โดยตรงต้องอาศัยการทำ Manual Digitization ผ่านซอฟต์แวร์สารสนเทศภูมิศาสตร์ (GIS) หรือออกแบบและเขียนแบบทางเทคนิค (CAD)ที่ใช้เวลานาน งานวิจัยนี้จึงเสนอแนวทางเพื่อลดระยะเวลาในขั้นตอน digitization โดยพัฒนาโมเดลแปลงภาพเอกสารสิทธิเป็นข้อมูลเวกเตอร์ที่สามารถนำไปใช้งานในระบบสารสนเทศภูมิศาสตร์และงานเขียนแบบทางเทคนิค โดยผสานการตรวจจับวัตถุด้วย YOLOv11 และการรู้จำข้อความภาษาไทย (OCR) เพื่อสกัดหมุดหลักเขต ข้อมูลเนื้อที่ และมาตราส่วน จากนั้นคำนวณเนื้อที่ด้วยสมการ Shoelace และปรับสเกลแบบ area-preserving เพื่อให้รูปหลายเหลี่ยมมีเนื้อที่สอดคล้องกับข้อมูลตามเอกสารสิทธิ การทดลองบนชุดข้อมูล 500 ภาพ (แบ่ง 70:20:10) แสดงให้เห็นว่าโมเดลตรวจจับวัตถุให้ค่า mAP@0.5 เท่ากับ 97.7% และระบบสามารถลดเวลาในการ digitize ต่อเอกสารจาก 60.5 วินาที เหลือ 27 วินาที คิดเป็นการลดลงร้อยละ 55.5 แม้ว่าความถูกต้องโดยรวมของระบบอยู่ที่ 76.7% ซึ่งต่ำกว่าการทำงานแบบ manual แต่ผลลัพธ์สะท้อนถึงศักยภาพของแนวทางที่เสนอในการรองรับกระบวนการ digitization เชิงปริมาณ โดยเฉพาะในหน่วยงานที่มีเอกสารจำนวนมาก ทั้งนี้ ความแม่นยำของ OCR ยังคงเป็นปัจจัยสำคัญที่ส่งผลต่อความถูกต้องเชิงเรขาคณิตของผลลัพธ์สุดท้าย*

**คำสำคัญ:** การแปลงภาพเป็นเวกเตอร์ ตรวจจับวัตถุ การรู้จำอักขระด้วยแสง การสร้างรูปหลายเหลี่ยมแบบรักษาพื้นที่ เอกสารสิทธิที่ดิน

## Abstract

*Most Thai land title deeds are stored as scanned images, which cannot be directly utilized for spatial analysis and require manual digitization using Geographic Information Systems (GIS) or Computer-Aided Design (CAD) software, a process that is time-consuming and labor-intensive. This study proposes an approach to reduce the time required for digitization by developing a model that converts scanned land title images into vector data applicable to GIS and technical drafting systems. The proposed framework integrates object detection using YOLOv11 and Thai Optical Character Recognition (OCR) to extract boundary markers, land area information, and map scale. The land area is then computed using the Shoelace formula and adjusted through an area-preserving scaling mechanism to ensure consistency with the legally stated area in the document. Experiments conducted on a dataset of 500 images (split 70:20:10) demonstrate that the object detection model achieved an mAP@0.5 of 97.7%. The system reduced the average digitization time per document from 60.5 seconds to 27 seconds, representing a 55.5% reduction. Although the overall system accuracy was 76.7%, which is lower than manual digitization, the results highlight the potential of the proposed approach for large-scale digitization processes, particularly in organizations handling substantial volumes of documents. OCR accuracy remains a critical factor affecting the geometric precision of the final vector output.*

**Keywords:** Image-to-Vector conversion, Object Detection, OCR, Area-Preserving Reconstruction, Land Title Deed.

## 1. บทนำ

เอกสารสิทธิในที่ดินของประเทศไทย เช่น เอกสารสิทธิที่ดิน (น.ส.4)หรือโฉนดที่ดิน เป็นเอกสารทางกฎหมายที่ใช้รับรองสิทธิการถือครองและกำหนดขอบเขตแปลงที่ดิน ภายในเอกสารดังกล่าวมักประกอบด้วยแผนที่รูปแปลงที่แสดงเส้นเขต จุดหมุด และข้อมูลเชิงพื้นที่อื่น ๆ ซึ่งมีความสำคัญต่อการบริหารจัดการที่ดินและการวิเคราะห์ในระบบสารสนเทศภูมิศาสตร์ (Geographic Information Systems: GIS) อย่างไรก็ตาม เอกสารสิทธิส่วนใหญ่ยังถูกจัดเก็บในรูปแบบภาพสแกน (raster images) โดยเฉพาะเอกสารที่ออกในอดีต ทำให้ข้อมูลเชิงพื้นที่ไม่สามารถนำไปใช้งานเชิงวิเคราะห์ได้โดยตรง การแปลงข้อมูลจากภาพราสเตอร์ให้เป็นข้อมูลเวกเตอร์ (raster-to-vector conversion) จึงเป็นกระบวนการสำคัญในการยกระดับเอกสารสิทธิให้สามารถบูรณาการเข้ากับระบบดิจิทัลและฐานข้อมูลภูมิสารสนเทศได้ แม้ว่างานวิจัยที่ผ่านมาได้เสนอวิธีการแปลงภาพเป็นเวกเตอร์ในบริบทของแผนที่ทั่วไป แบบแปลนสถาปัตยกรรม และเอกสารโครงสร้างมาตรฐาน แต่เอกสารสิทธิของประเทศไทยมีลักษณะเฉพาะ ได้แก่ (1) คุณภาพภาพสแกนที่ไม่สม่ำเสมอ (2) เส้นขอบเขตที่ไม่คมชัดหรือซีดจาง (3) ข้อความสำคัญในเอกสารมีความหลากหลายด้วยการเขียนด้วยลายมือจากเครื่องพิมพ์ดีด หรือเครื่องคอมพิวเตอร์ (4)ไม่สามารถนำไปกำหนดค่าพิกัดภูมิศาสตร์(Georeference)ได้โดยตรง ปัจจัยเหล่านี้ส่งผลให้วิธีการทั่วไปไม่สามารถให้ผลลัพธ์ที่มีความถูกต้องเพียงพอสำหรับการใช้งานเชิงภูมิสารสนเทศที่ต้องการทั้งเชิงเรขาคณิตและเชิงทอพอโลยี

ช่องว่างของงานวิจัยจึงอยู่ที่การพัฒนากระบวนการที่ออกแบบเฉพาะสำหรับเอกสารสิทธิของประเทศไทย โดยครอบคลุมตั้งแต่การตรวจจับองค์ประกอบสำคัญของแผนที่รูปแปลง การสกัดข้อความ ไปจนถึงการสร้างข้อมูลรูปหลายเหลี่ยมให้อยู่ในรูปแบบเวกเตอร์ที่สามารถนำไปใช้งานต่อในระบบภูมิสารสนเทศหรือด้านการออกแบบและเขียนแบบเทคนิค (CAD)

งานวิจัยนี้เสนอกรอบการทำงานแบบบูรณาการสำหรับการแปลงภาพเอกสารสิทธิที่ดินเป็นข้อมูลเวกเตอร์ โดยผสาน (1) โมเดลตรวจจับวัตถุแบบ YOLO สำหรับสกัดหมุดหลักเขต (2) ระบบรู้จำข้อความภาษาไทยสำหรับดึงค่ามาตราส่วนและเนื้อที่ และ (3) กระบวนการปรับสเกลรูปหลายเหลี่ยมแบบ scale-aware และ area-constrained เพื่อให้เนื้อที่ของรูปแปลงที่สร้างขึ้นสอดคล้องกับข้อมูลตามเอกสารสิทธิ ที่ช่วยลดความคลาดเคลื่อนของเนื้อที่พร้อมลดภาระงาน digitization แบบ manual

## 2. การทบทวนวรรณกรรมและงานวิจัยที่เกี่ยวข้อง

### 2.1 การตรวจจับวัตถุด้วย Deep Learning

การตรวจจับวัตถุด้วยโครงข่ายประสาทเทียมเชิงลึกได้รับการพัฒนาอย่างต่อเนื่อง โดยเฉพาะตระกูล YOLO ซึ่งเป็น single-stage detector ที่สามารถตรวจจับวัตถุได้แบบรวดเร็วและมีประสิทธิภาพสูง [1] ต่อมา YOLOv3 ได้เพิ่มความสามารถในการตรวจจับหลายขนาดวัตถุผ่าน multi-scale detection [2] ขณะที่ YOLOv5 และ YOLOv8 ได้รับความนิยมในเชิงปฏิบัติ เนื่องจากมีขั้นตอนการฝึกและการนำไปใช้งานที่สะดวกและเสถียร [3], [4] ในบริบทของเอกสารงานอย่าง PubLayNet [5] และ LayoutLM [6] แสดงให้เห็นว่าการใช้ข้อมูลตำแหน่ง (layout) ร่วมกับเนื้อหาข้อความช่วยให้การทำความเข้าใจเอกสารมีประสิทธิภาพมากขึ้น

### 2.2 การรู้จำอักขระด้วยแสง (OCR)

งานด้านการตรวจจับข้อความ เช่น CRAFT [7] และ DBNet [8] ได้พัฒนาเทคนิคการระบุตำแหน่งข้อความในภาพด้วยแนวทางแบบ segmentation ซึ่งช่วยเพิ่มความแม่นยำในภาพที่มีสภาพซับซ้อน อย่างไรก็ตาม ขั้นตอนดังกล่าวยังเป็นเพียงการหาขอบเขตข้อความ ไม่ใช่การถอดความหมายเชิงเนื้อหา ในภาษาไทยซึ่งมีลักษณะสระและวรรณยุกต์ซ้อนกันหลายระดับ งาน Typhoon OCR [9] ได้เสนอโมเดล Vision-Language ที่รองรับเอกสารภาษาไทยโดยเฉพาะ และสามารถประมวลผลข้อความร่วมกับโครงสร้าง layout ได้ดีขึ้น

### 2.3 การแปลงภาพเป็นข้อมูลเวกเตอร์

การแปลงภาพจากรูปแบบราสเตอร์เป็นเวกเตอร์ได้รับการศึกษาอย่างต่อเนื่องในงานด้านแบบร่างเทคนิคและ

แผนที่ งาน Deep Vectorization of Technical Drawings [10] และ End-to-End Line Drawing Vectorization [11] แสดงให้เห็นว่าสามารถใช้ Deep Learning สร้างเส้นเวกเตอร์จากภาพได้โดยตรง ขณะที่งานในสายแผนที่ เช่น Automatic Vectorization of Historical Maps [12] และงานด้าน road vectorization [13] ได้นำเสนอขั้นตอนตั้งแต่การแบ่งส่วนภาพไปจนถึงการสร้างโครงสร้างเวกเตอร์

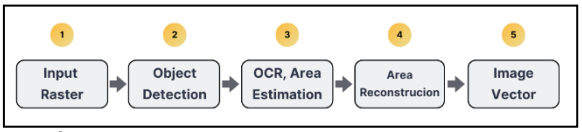
**2.4 การรักษาโครงสร้างและพื้นที่ของรูปหลายเหลี่ยม**

ในงานด้านภูมิสารสนเทศ มีการศึกษาการปรับรูปทรง polygon โดยยังคงความสัมพันธ์เชิงทอพอโลยี เช่น งาน Topologically Safe Simplification of Building Footprints [14] ซึ่งมุ่งรักษาโครงสร้างเชิงพื้นที่ระหว่างการลดความซับซ้อนของรูปทรง ขณะที่งาน Area-Preserving Simplification and Schematization [15] เสนอแนวทางการปรับรูปหลายเหลี่ยมโดยยังคงค่าพื้นที่เดิม แนวคิดการรักษาพื้นที่มีความสำคัญอย่างยิ่งในบริบทเอกสารสิทธิที่ดินเนื่องจากเนื้อที่เป็นข้อมูลทางกฎหมายที่ไม่สามารถคลาดเคลื่อนได้

จากการทบทวนวรรณกรรม พบว่างานวิจัยที่ผ่านมาแม้จะมีความก้าวหน้าในด้านการตรวจจับวัตถุ การทำความเข้าใจเอกสารและการรู้จำด้วยแสงภาษาไทยและการแปลงภาพเป็นเวกเตอร์พร้อมการรักษา topology แต่ส่วนใหญ่พัฒนาแยกส่วนกัน ยังขาด pipeline แบบครบวงจรที่สามารถ (1) ตรวจจับองค์ประกอบเอกสารราชการไทย (2) ดึงข้อมูลข้อความภาษาไทยอย่างแม่นยำ และ (3) แปลงข้อมูลเชิงภาพเป็น vector polygon อย่างเป็นระบบ

## 3. วิธีการดำเนินงานวิจัยและออกแบบระบบ

ระบบที่นำเสนอในงานวิจัย ตั้งแต่การเตรียมข้อมูล การฝึกโมเดลตรวจจับวัตถุ การออกแบบสถาปัตยกรรมระบบไปจนถึงกระบวนการประมวลผลข้อมูลและการแปลงข้อมูลจากภาพสแกนเอกสารสิทธิที่ดินให้อยู่ในรูปแบบข้อมูลเชิงเวกเตอร์ โดยแนวคิดและโครงสร้างของระเบียบวิธีวิจัยได้รับการออกแบบตามกรอบการทำงานแบบผสมผสาน (Hybrid Framework) ซึ่งผสานเทคนิค Deep Learning, OCR และการคำนวณเชิงเรขาคณิตเข้าไว้ในระบบเดียว




**ภาพที่ 1:** ภาพรวมขั้นตอนการแปลงข้อมูลจากภาพสแกนเอกสารสิทธิที่ดินให้อยู่ในรูปแบบข้อมูลเชิงเวกเตอร์

**3.1 ภาพรวมของระบบที่นำเสนอ**

ขั้นตอนการแปลงข้อมูลจากเอกสารสิทธิที่ดินในรูปแบบภาพราสเตอร์ให้เป็นข้อมูลเชิงเวกเตอร์ ดังภาพที่ 1 เริ่มต้นจากการรับภาพสแกนเอกสารสิทธิเป็นข้อมูลนำเข้า จากนั้นระบบใช้โมเดลตรวจจับวัตถุ YOLOv11 ในการระบุและสกัดองค์ประกอบสำคัญภายในเอกสาร ได้แก่ ตำแหน่งหมุดหลักเขต ข้อมูลเนื้อที่ และมาตราส่วน โดยผลลัพธ์การตรวจจับถูกจัดเก็บในรูปแบบโครงสร้างข้อมูล JSON เพื่อใช้ในการประมวลผลขั้นต่อไป

ถัดมา ระบบดำเนินการรู้จำอักขระด้วยแสงเพื่อสกัดข้อมูลข้อความที่เกี่ยวข้อง เช่น ค่าเนื้อที่และมาตราส่วน จากนั้นแปลงข้อมูลดังกล่าวให้อยู่ในรูปแบบเชิงตัวเลขที่สามารถนำไปคำนวณได้ ในขั้นตอนการสร้างรูปแปลงเชิงเวกเตอร์ พิกัดหมุดที่ตรวจจับได้จะถูกนำมาสร้างรูปหลายเหลี่ยม และคำนวณพื้นที่ด้วยสมการ Shoelace Formula พร้อมทั้งปรับสเกลแบบ area-preserving เพื่อให้เนื้อที่ของรูปแปลงที่สร้างขึ้นมีความสอดคล้องกับเนื้อที่ระบุในเอกสารสิทธิ ทั้งในด้านรูปร่าง มาตราส่วน และความถูกต้องเชิงเรขาคณิต

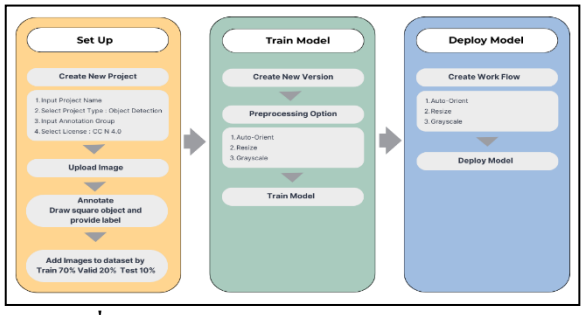
สุดท้าย ระบบส่งออกผลลัพธ์ในรูปแบบข้อมูลเวกเตอร์ที่สามารถนำไปใช้งานต่อในระบบสารสนเทศภูมิศาสตร์หรือซอฟต์แวร์ออกแบบและเขียนแบบทางเทคนิค

**3.2 การเตรียมชุดข้อมูล (Dataset Preparation) และฝึกโมเดลตรวจจับวัตถุ**

โมเดลตรวจจับวัตถุได้รับการฝึกสอนโดยใช้สถาปัตยกรรม YOLOv11ผ่านแพลตฟอร์ม Roboflow มีการจัดเตรียมและทำการกำกับข้อมูล (annotation) ชุดข้อมูลเอกสารสิทธิไทยจำนวน 500 รูปแบบกำหนดเอง ดังแสดงในภาพที่ 2 การทำ Annotation โดยกำหนดกรอบ

สี่เหลี่ยมครอบวัตถุที่ต้องการตรวจจับ และระบุป้ายกำกับให้กับองค์ประกอบสำคัญ ได้แก่

- หมุดหลักเขต (point_parcel)
- เนื้อที่ของแปลงที่ดิน ไร่-งาน-ตารางวา (area)
- ข้อความบรรยายเนื้อที่ ภาษาไทย(area_thai)
  - เนื้อที่ของแปลงที่ดิน ไร่(area_rai)
  - เนื้อที่ของแปลงที่ดิน งาน (area_ngan)
  - เนื้อที่ของแปลงที่ดิน ตารางวา (area_wa)
- มาตราส่วนของรูปแปลง (scale_parcel)



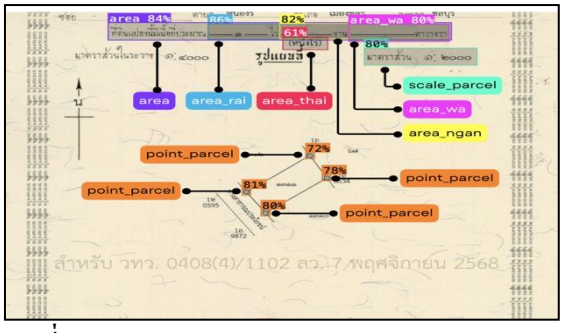

**ภาพที่ 2:** ขั้นตอนการฝึกโมเดลบน Roboflow

ขั้นตอนการฝึกโมเดลในภาพที่ 2 เริ่มจากการสร้างโปรเจกต์บน Roboflow สำหรับงาน Object Detection จากนั้นอัปโหลดภาพและทำการกำกับข้อมูล (annotation) พร้อมแบ่งชุดข้อมูลเป็น 70% สำหรับการฝึกสอน (training) 20% สำหรับการตรวจสอบความถูกต้องของโมเดล (validation) 10% สำหรับการทดสอบประสิทธิภาพ (testing) ถัดไปสร้างเวอร์ชันสำหรับฝึกโมเดลและตั้งค่าการเตรียมข้อมูล เช่น Auto-orient, Resize และ Grayscale เพื่อเพิ่มความทนทานของคุณลักษณะ เมื่อกำหนดค่าเสร็จจึงทำการ Train โมเดล และในขั้นตอนสุดท้ายสร้าง Workflow สำหรับ Deploy เพื่อให้โมเดลพร้อมนำไปใช้งานจริงบนระบบประมวลผลภาพ/เอกสาร

งานวิจัยนี้เลือกใช้สถาปัตยกรรม YOLOv11 สำหรับการตรวจจับองค์ประกอบในเอกสารเอกสารสิทธิที่ดิน เนื่องจากเป็นโมเดลที่มีความสามารถในการตรวจจับวัตถุได้อย่างรวดเร็วและเหมาะสมกับงานที่ต้องการประมวลผลแบบใกล้เคียงเวลาจริง (Near-Realtime) กระบวนการฝึกโมเดลดำเนินการบนแพลตฟอร์ม Roboflow โดยมีขั้นตอนการเตรียมข้อมูลล่วงหน้า (Preprocessing) ได้แก่ การปรับขนาดภาพ (Resize) และการแปลงภาพเป็นระดับสีเทา (Grayscale) เพื่อเพิ่มความสม่ำเสมอของข้อมูลและลดผลกระทบจากสัญญาณรบกวนในภาพสแกน หลังจากนั้นจึงทำการฝึกโมเดลและประเมินผลด้วยค่าตัวชี้วัดมาตรฐาน เช่น Precision, Recall และ mAP


```
"predictions": [
  {
    "width": 1532,
    "height": 113,
    "x": 1400,
    "y": 128.5,
    "confidence": 0.8495081663131714,
    "class_id": 0,
    "class": "area",
    "detection_id": "e1da053b-4466-4709-a8f3-a5e
    "parent_id": "image"
  },
  {
    "width": 323,
    "height": 96,
    "x": 1243.5,
    "y": 130,
    "confidence": 0.8418651819229126,
    "class_id": 2,
    "class": "area_rai",
    "detection_id": "c7c12656-a4e4-4d85-a649-374
    "parent_id": "image"
  },
```


**ภาพที่ 4:** ตัวอย่างผลลัพธ์ JSON ที่ได้จากโมเดล YOLOv11




**ภาพที่ 3:** ตัวอย่างผลลัพธ์การตรวจจับวัตถุของโมเดล YOLOv11

### 3.3 สถาปัตยกรรมระบบฝั่ง Backend

ระบบที่พัฒนาขึ้นโดยสถาปัตยกรรมแบบ Web Application ด้วยเว็บเฟรมเวิร์ก Flask (Micro-Framework) ของภาษา Python ที่น้ำหนักเบา ซึ่งช่วยสร้าง API รับส่งรูปภาพเข้าสู่โมเดลตรวจจับวัตถุดังภาพที่ 3 ผลลัพธ์จากโมเดลจะได้ข้อมูลในรูปแบบ JSON ดังภาพที่ 4 จากนั้นระบบจะทำการแยกองค์ประกอบตามประเภทของข้อมูล หากองค์ประกอบเป็นข้อมูลเชิงข้อความ เช่น เนื้อที่หรือมาตราส่วน จะใช้ OCR เพื่อแปลงข้อความจากภาพให้อยู่ในรูปแบบตัวอักษร เพื่อใช้เป็นข้อมูลนำเข้าในขั้นตอนการคำนวณและแปลงเป็นข้อมูลเชิงเวกเตอร์ เช่น SVG ซึ่งสามารถนำไปใช้ต่อในระบบสารสนเทศภูมิศาสตร์ (GIS), ซอฟต์แวร์ประยุกต์ด้านการออกแบบและเขียนแบบเทคนิค (CAD) หรือระบบแอปพลิเคชันบนเว็บ

### 3.4 การสร้างรูปแปลงที่ดินให้เป็นแบบเวกเตอร์

การสร้างรูปแปลงที่ดินให้ได้เนื้อที่ใหม่เท่ากับเนื้อที่เดิมตามเอกสารสิทธิโดยใช้ Area-preserving scaling factor

1. แปลงเนื้อที่เอกสารสิทธิ จากหน่วย ไร่-งาน-ตารางวา ให้เป็นหน่วย ตารางเมตร

$$A_{title}(m^2) = 1600R + 400N + 4W \quad (1)$$

โดย R = จำนวนไร่ ,N = จำนวนงาน ,W=จำนวนตารางวา (1 ตารางวา = 4 $m^2$ ,1 งาน = 400 $m^2$ , 1 ไร่ =1600 $m^2$ )

2. แปลงเนื้อที่เอกสารสิทธิ จากหน่วย ไร่-งาน-ตารางวา ให้เป็นหน่วยตารางเมตร การคำนวณเนื้อที่จากรูปแปลงที่สร้างใหม่ (Polygon Area) ด้วยสมการ Shoelace formula ด้วยการนำพิกัดของ class:point_parcel จาก JSON ของโมเดลYOLO ที่ตรวจจับได้ มาคำนวณเนื้อที่ของรูปหลายเหลี่ยม (Polygon) ด้วยสูตร Shoelace โดยเริ่มจากจุดแรก $(x_i,y_i)$, i=1..n และปิดรูปด้วย $(x_{n+1},y_{n+1}) = (x_1,y_1)$ เนื้อที่ที่คำนวณได้ให้เป็น $A_{raw}$ ดังสมการต่อไปนี้

$$A_{raw} = \frac{1}{2}\left|\sum_{i=1}^{n}(x_i y_{i+1} - y_i x_{i+1})\right| \quad (2)$$

3. แปลงเนื้อที่จากเอกสารหาค่า factor สำหรับ “สเกลเชิงเส้น” ให้เนื้อที่เท่ากับเนื้อที่ที่จะเปลี่ยนตาม “กำลังสอง” ของสเกลเชิงเส้น ถ้าสเกลรูปด้วยตัวคูณ S (ทั้งแกน x และ y เท่ากัน) เนื้อที่ใหม่จะเป็น

$$A_{scaled} = s^2 A_{raw} \quad (3)$$

สเกลแฟกเตอร์เชิงเส้น คือ

$$s = \sqrt{\frac{A_{title(m^2)}}{A_{raw}}} \quad (4)$$

และเนื้อที่หลังปรับจะเป็น

$$A_{scaled} = s^2 A_{raw} = A_{title(m^2)} \quad (5)$$

4. การปรับค่าพิกัดฉาก x,y รูปแปลงที่ดินใหม่ เพื่อไม่ให้รูปแปลงขยับไปทั้งแปลงต้องสเกลรอบจุดศูนย์กลาง (Centroid) แล้วปรับพิกัดทุกจุด

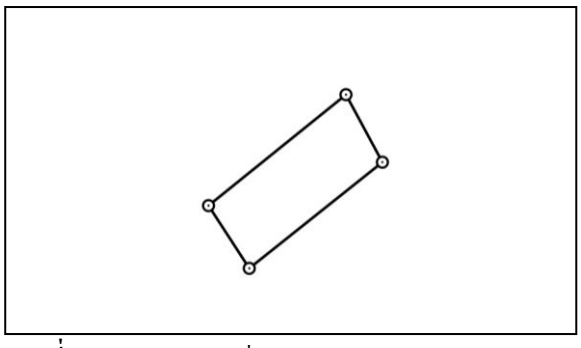
4.1 หาจุดศูนย์กลางของรูปแปลงที่ดิน

$$c_x = \frac{1}{n}\sum_{i=1}^{n} x_i \, , c_y = \frac{1}{n}\sum_{i=1}^{n} y_i \quad (6)$$

4.2 ปรับพิกัดทุกจุด

$$x_i' = c_x + s(x_i - x_x) \text{ และ} \quad (7)$$

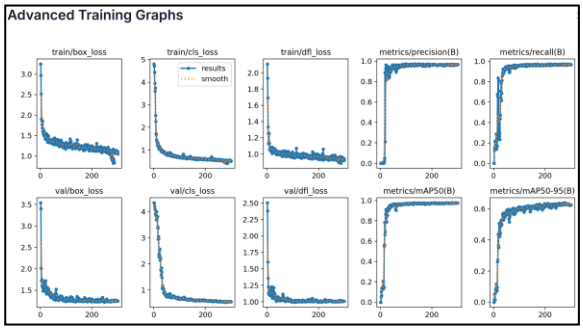
$$y_i' = c_y + s(y_i - c_y)$$



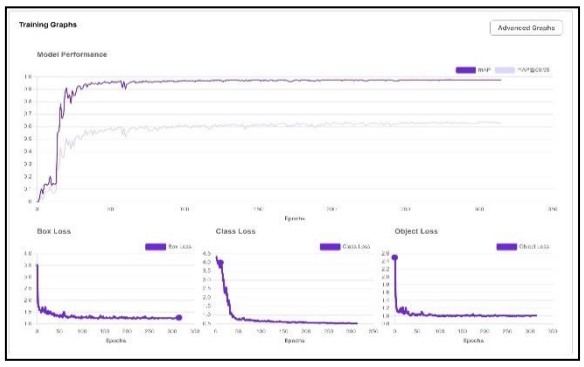
**ภาพที่ 5:** ตัวอย่างรูปแปลงที่ดินในรูปแบบเวกเตอร์(SVG)

## 4. ผลการทดลอง

การฝึกโมเดลตรวจจับวัตถุด้วยสถาปัตยกรรม YOLOv11 ใช้ชุดข้อมูลภาพเอกสารสิทธิจำนวน 500 ภาพ Advanced Training Graphs ดังภาพที่ 6 และ Training Graphs ดังภาพที่ 7 แสดงให้เห็นว่าค่าการสูญเสียของชุดฝึกและชุดตรวจสอบลดลงอย่างสอดคล้องกัน ขณะที่ค่าความแม่นยำและค่า mAP อยู่ในระดับสูงและมีเสถียรภาพ สะท้อนว่าโมเดลไม่เกิดปัญหา overfitting และสามารถนำไปใช้ตรวจจับองค์ประกอบสำคัญในเอกสารสิทธิที่ดินได้อย่างมีประสิทธิภาพ




**ภาพที่ 6:** Advanced Training Graphs ของโมเดล YOLOv11




**ภาพที่ 7:** Training Graphs ของโมเดล YOLOv11

เพื่อประเมินความแม่นยำของโมเดลตรวจจับองค์ประกอบเอกสาร ได้ทำการเปรียบเทียบโมเดล YOLO หลายรุ่น ได้แก่ YOLOv11, YOLOv12 และ YOLOv26 โดยใช้ตัวชี้วัดมาตรฐาน Precision, Recall และ mAP@0.5

**ตารางที่ 1:** ตารางเปรียบเทียบของโมเดล YOLO (Fast)

| Model | Precision(%) | Recall(%) | mAP@0.5(%) |
|---|---|---|---|
| YOLOv11 | 96.4 | 96.8 | **97.7** |
| YOLOv12 | **96.8** | **97.2** | 97.4 |
| YOLOv26 | 93.3 | 92.7 | 95.8 |

ผลการทดลองพบว่า YOLOv11 ให้ค่า mAP@0.5 สูงสุดที่ 97.7% ขณะที่ YOLOv12 ให้ค่า mAP@0.5 ใกล้เคียงกันที่ 97.4% และ YOLOv26 มีค่า 95.8% ความสามารถในการตรวจจับวัตถุให้ครบถ้วนในระดับ IoU $\geq$ 0.5 มีความสำคัญมากกว่า ดังนั้น YOLOv11 ซึ่งมีค่า mAP@0.5 สูงสุดและมีเสถียรภาพของ Training Graphs ที่สม่ำเสมอ จึงถูกเลือกเป็นโมเดลหลักของระบบ ทั้งนี้ YOLOv26 แม้เป็นเวอร์ชันใหม่กว่า แต่มีแนวโน้มถูกปรับจูนเพื่อประสิทธิภาพด้านความเร็ว ส่งผลให้ความแม่นยำลดลงเมื่อประยุกต์ใช้กับเอกสารสแกนที่มีสัญญาณรบกวนและวัตถุขนาดเล็ก

ระบบประเมินผลโดยใช้สภาพแวดล้อมของเครื่องแล็ปท็อปโดยใช้ชุดทดสอบเอกสารสิทธิไทยจำนวน 30 ภาพ ตัวชี้วัดการประเมินเปรียบเทียบค่า JSON ที่สร้างขึ้นกับการ manual digitize โดยวัดประสิทธิภาพ (เวลาประมวลผลเฉลี่ยต่อเอกสารสิทธิ)

**ตารางที่ 2:** ตารางเปรียบเทียบการ digitize แปลงที่ดินกับระบบที่นำเสนอ

| วิธี | ความถูกต้อง (%) | เวลา(วินาที) |
|---|---|---|
| 1. Manual Digitize | 96.7 | 60.6 |
| **2. ระบบที่เสนอ** | **76.7** | **27** |

ระบบสามารถลดเวลาได้ประมาณ 55.5% เมื่อเทียบกับกระบวนการทำงานแบบ manual แม้ความถูกต้องของระบบจะต่ำกว่าการทำงานแบบ manual ประมาณ 20% แต่ระบบสามารถลดเวลาในการดำเนินการได้มากกว่าครึ่งหนึ่ง ซึ่งสะท้อนถึงศักยภาพในการประยุกต์ใช้ในกระบวนการ digitization เชิงปริมาณโดยเฉพาะในหน่วยงานที่มีเอกสารจำนวนมาก ทั้งนี้ ผลลัพธ์ดังกล่าวชี้ให้เห็นถึงความจำเป็นในการพัฒนากลไกเพิ่มความแม่นยำเพิ่มเติมในขั้นตอน OCR และการปรับแก้รูปทรงเรขาคณิต

การประเมิน OCR ทำบนชุดข้อมูลทดสอบจำนวน 30 ภาพ โดยใช้ตัวชี้วัด Word Accuracy Rate (WAR) ได้ผลการประเมิน ตามตารางที่ 3

**ตารางที่ 3:** ตารางประเมินประสิทธิภาพการแปลงข้อความของ Typhoon OCR

| ข้อมูล | ค่าเฉลี่ย(%) |
|---|---|
| **เนื้อที่ (ไร่-งาน-ตร.ว)** | **76.7** |
| **มาตราส่วน** | **80.0** |

ผลการประเมินประสิทธิภาพ Typhoon OCR แสดงให้เห็นว่าระบบสามารถรู้จำข้อความมาตราส่วนได้ที่ระดับ 80.0% และข้อมูลเนื้อที่ที่ระดับ 76.7% แม้ว่าค่าดังกล่าวจะอยู่ในระดับที่ยอมรับได้สำหรับงานรู้จำข้อความภาษาไทยจากเอกสารสแกน แต่เนื่องจากข้อมูลทั้งสองประเภทถูกนำไปใช้ในการคำนวณตัวคูณสเกลและปรับรูปหลายเหลี่ยมเชิงเรขาคณิต ความคลาดเคลื่อนของ OCR จึงมีผลโดยตรงต่อความถูกต้องของผลลัพธ์เวกเตอร์ขั้นสุดท้าย

## 5. การสรุปและอภิปรายผล

งานวิจัยนี้นำเสนอระบบอัตโนมัติสำหรับแปลงภาพเอกสารสิทธิที่ดินเป็นข้อมูลเวกเตอร์ โดยผสานการตรวจจับวัตถุด้วย YOLOv11 การรู้จำข้อความภาษาไทย และกระบวนการปรับรูปหลายเหลี่ยมแบบ scale-aware และ area-constrained เพื่อรักษาความถูกต้องเชิงเรขาคณิตและเชิงพื้นที่

ผลการทดลองบนเอกสารสิทธิ 500 ภาพ แสดงให้เห็นว่าโมเดลตรวจจับวัตถุให้ค่า mAP@0.5 เท่ากับ 97.7% และระบบสามารถลดเวลาในการ digitize ต่อเอกสารจาก 60.5 วินาที เหลือ 27 วินาที สะท้อนถึงศักยภาพในการ

ประยุกต์ใช้กับงาน digitization เชิงปริมาณในหน่วยงานที่มีเอกสารจำนวนมาก

อย่างไรก็ตาม ประสิทธิภาพของระบบยังขึ้นอยู่กับความแม่นยำของ OCR และความหลากหลายของรูปแบบเอกสาร ในอนาคตควรพัฒนากลไกเพิ่มความแม่นยำของการรู้จำอักขระด้วยแสงซึ่งเป็นปัจจัยสำคัญของระบบ โดยรวมและเทคนิคปรับรูปทรงที่คำนึงถึงข้อจำกัดเชิงทอพอโลยี รวมถึงการประยุกต์ใช้โมเดล segmentation ขั้นสูงเพื่อเพิ่มความแม่นยำของขอบเขตแปลง

## เอกสารอ้างอิง

[1] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You Only Look Once: Unified, Real-Time Object Detection,” Proc. CVPR, 2016.

[2] J. Redmon and A. Farhadi, “YOLOv3: An Incremental Improvement,” 2018.

[3] Ultralytics, “YOLOv5,” GitHub repository, 2020.

[4] Ultralytics, “YOLOv8,” Documentation, 2023.

[5] X. Zhong, J. Tang, and A. Jimeno Yepes, “PubLayNet: largest dataset ever for document layout analysis,” 2019.

[6] Y. Xu et al., “LayoutLM: Pre-training of Text and Layout for Document Image Understanding,” 2019.

[7] Y. Baek, B. Lee, D. Han, S. Yun, and H. Lee, “Character Region Awareness for Text Detection (CRAFT),” 2019.

[8] M. Liao et al., “Real-time Scene Text Detection with Differentiable Binarization,” 2019.

[9] S. Nonesung et al., “Typhoon OCR: Open Vision-Language Model For Thai Document Extraction,” 2026.

[10] V. Egiazarian et al., “Deep Vectorization of Technical Drawings,” 2020.

[11] H. Liu, C. Li, X. Liu, et al., “End-to-End Line Drawing Vectorization,” Proc. AAAI, vol. 36, no. 4, pp. 4559–4566, 2022.

[12] Y. Chen et al., “Automatic vectorization of historical maps: A benchmark,” PLOS ONE, vol. 19, no. 2, e0298217, 2024.

[13] C. Jiao, M. Heitzler, and L. Hurni, “A novel framework for road vectorization and classification from historical maps based on deep learning and symbol painting,” Computers, Environment and Urban Systems, vol. 108, 102060, 2024.

[14] J. Haunert and A. Wolff, “Optimal and Topologically Safe Simplification of Building Footprints,” Proc. ACM SIGSPATIAL GIS, 2010.

[15] K. Buchin, W. Meulemans, A. van Renssen, and B. Speckmann, “Area-Preserving Simplification and Schematization of Polygonal Subdivisions,” ACM Trans. Spatial Algorithms Syst., 2016.

# กรอบการประเมินสมรรถนะด้านการรักษาความปลอดภัยไซเบอร์แบบถ่วงน้ำหนัก

# A Weighted Cybersecurity Competency Assessment Framework

*ชัยยุทธ มาตย์ไตร (Chaiyut Mattrai)[1] และนลินภัสร์ บำเพ็ญเพียร (Nalinpat Bhumpenpein)[2]*
*ภาควิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล*
*มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ*

*[1]s6707011857136@email.kmutnb.ac.th, [2]nalinpat.b@itd.kmutnb.ac.th*

**บทคัดย่อ**

*งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนากรอบการประเมินสมรรถนะด้านการรักษาความปลอดภัยไซเบอร์ โดยอาศัยแนวคิดแบบหลายมิติ ได้แก่ ความรู้ (K) ทักษะ (S) ความสามารถ (A) และภารกิจการปฏิบัติงาน (T) ร่วมกับการให้คะแนนแบบถ่วงน้ำหนักตามระดับความยากของข้อคำถาม เพื่อให้สามารถประเมินความพร้อมของบุคลากรอย่างเป็นระบบและครอบคลุม กลุ่มตัวอย่างเป็นบุคลากรในหน่วยงานภาครัฐจำนวน 186 คน ซึ่งได้มาจากการเลือกแบบเจาะจง เครื่องมือที่ใช้ในการวิจัยเป็นแบบทดสอบปรนัย จำนวน 60 ข้อ แบ่งเป็น 4 มิติ มิติละ 15 ข้อ โดยผ่านการตรวจสอบความตรงเชิงเนื้อหาจากผู้เชี่ยวชาญและมีค่าความเชื่อมั่น (KR-20) เท่ากับ 0.96 และกำหนดค่าน้ำหนักคะแนนในช่วง 0.2–0.8 ตามระดับความยากของข้อคำถาม ผลการวิจัยพบว่า คะแนนเฉลี่ยแบบถ่วงน้ำหนักเท่ากับ 27.36 จากคะแนนเต็ม 32.80 (ร้อยละ 83.40, SD = 2.15) อยู่ในระดับสูงมาก โดยมิติทักษะมีคะแนนสูงสุด ขณะที่มิติความรู้มีคะแนนต่ำที่สุด ผลการวิจัยสะท้อนให้เห็นว่ากรอบการประเมินที่พัฒนาขึ้นสามารถใช้ระบุจุดแข็งและช่องว่างสมรรถนะของบุคลากรได้อย่างเป็นรูปธรรม และสามารถนำไปประยุกต์ใช้เพื่อสนับสนุนการวางแผนพัฒนาบุคลากรด้านความมั่นคงปลอดภัยไซเบอร์ในองค์กรได้อย่างมีประสิทธิภาพ*

**คำสำคัญ** : การประเมินสมรรถนะ, ความมั่นคงปลอดภัยไซเบอร์, กรอบการประเมิน, การถ่วงน้ำหนักคะแนน, แบบทดสอบปรนัย

**Abstract**

*This study aims to develop a cybersecurity competency assessment framework based on a multi-dimensional approach, including Knowledge (K), Skills (S), Abilities (A), and Tasks (T), combined with a weighted scoring method based on item difficulty to systematically and comprehensively assess personnel readiness. The sample consisted of 186 personnel from a public organization, selected using purposive sampling. The research instrument was a 60-item multiple-choice test, divided into four dimensions with 15 items each, validated by experts and demonstrating a reliability coefficient (KR-20) of 0.96. Item difficulty was analyzed, and weights ranging from 0.2 to 0.8 were assigned accordingly. The results showed that the average weighted score was 27.36 out of 32.80 (83.40%, SD = 2.15), indicating a very high level of cybersecurity readiness. The Skills dimension achieved the highest score, while Knowledge had the lowest. The proposed framework effectively identifies strengths and competency gaps and can be applied to support workforce development planning in cybersecurity within organizations.*

**Keywords :** Cybersecurity Competency, Assessment Framework, Weighted Scoring, Cybersecurity Readiness, Multiple-Choice Test.

## 1. บทนำ

ในยุคของการเปลี่ยนผ่านสู่ดิจิทัล (Digital Transformation) เทคโนโลยีสารสนเทศมีบทบาทสำคัญต่อการดำเนินงานขององค์กร แต่ขณะเดียวกันก็ทำให้ความเสี่ยงด้านความมั่นคงปลอดภัยไซเบอร์ (Cybersecurity) เพิ่มสูงขึ้น โดยรายงานในระดับสากลชี้ให้เห็นว่าภัยคุกคามมีความซับซ้อนมากขึ้น ขณะที่ปัจจัยด้านมนุษย์ยังคงเป็นจุดอ่อนสำคัญของระบบ [1], [2]

แม้องค์กรจะนำมาตรฐานด้านความมั่นคงปลอดภัยสารสนเทศ เช่น ISO/IEC 27002 มาใช้ [3] และลงทุนในเทคโนโลยีเพื่อป้องกันภัยคุกคาม แต่การประเมินความพร้อมของบุคลากรยังคงเป็นความท้าทาย เนื่องจากสมรรถนะไม่สามารถวัดได้จากความรู้เพียงอย่างเดียว แต่ต้องครอบคลุมทักษะ ความสามารถ และการปฏิบัติงานจริง [4]

งานวิจัยที่ผ่านมาส่วนใหญ่มุ่งเน้นการประเมินเพียงมิติเดียว เช่น ความตระหนักรู้หรือพฤติกรรมของผู้ใช้งาน และยังมีข้อจำกัดในการสะท้อนความแตกต่างของระดับสมรรถนะอย่างรอบด้าน [5], [6], [7]

ดังนั้น งานวิจัยนี้จึงพัฒนากรอบการประเมินสมรรถนะด้านความมั่นคงปลอดภัยไซเบอร์ โดยใช้แนวคิดหลายมิติ (K-S-A-T) ร่วมกับการให้คะแนนแบบถ่วงน้ำหนักตามระดับความยากของข้อคำถาม เพื่อให้สามารถประเมินความพร้อมของบุคลากรได้อย่างครอบคลุมและแม่นยำมากยิ่งขึ้น โดยกรอบดังกล่าวสามารถนำไปประยุกต์ใช้กับองค์กรประเภทต่าง ๆ ได้

ผลลัพธ์จากงานวิจัยนี้คาดว่าจะช่วยให้องค์กรสามารถระบุจุดแข็งและช่องว่างสมรรถนะของบุคลากร และนำไปใช้ในการวางแผนพัฒนาได้อย่างมีประสิทธิภาพ

## 2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

### 2.1 หลักการความมั่นคงปลอดภัยสารสนเทศ

หลักการพื้นฐานของความมั่นคงปลอดภัยสารสนเทศอ้างอิงองค์ประกอบสำคัญ 3 ประการ (CIA Triad) ได้แก่ การรักษาความลับ (Confidentiality) การรักษาความถูกต้องครบถ้วน (Integrity) และการรักษาความพร้อมใช้งาน (Availability) ซึ่งเป็นพื้นฐานของการกำหนดมาตรการควบคุมตามมาตรฐาน ISO/IEC 27002 เพื่อรองรับภัยคุกคามที่ซับซ้อนในปัจจุบัน [2], [3]

ในด้านการพัฒนาบุคลากร ความรู้เกี่ยวกับหลักการดังกล่าวถือเป็นรากฐานสำคัญที่ส่งผลต่อพฤติกรรมด้านความปลอดภัยของผู้ใช้งาน โดยช่วยเสริมสร้างการตัดสินใจที่เหมาะสมและลดความเสี่ยงจากปัจจัยด้านมนุษย์ [2]

### 2.2 กรอบมาตรฐานสำหรับการนิยามสมรรถนะกำลังคนด้านไซเบอร์

สถาบันมาตรฐานและเทคโนโลยีแห่งชาติ (NIST) ได้พัฒนากรอบมาตรฐาน NIST SP 800-181 หรือ NICE Framework เพื่อสร้างภาษามาตรฐานในการระบุสมรรถนะและบทบาทงานด้านไซเบอร์ ให้มีความชัดเจนและใช้ร่วมกันได้อย่างเป็นระบบ กรอบดังกล่าวช่วยให้การประเมินสมรรถนะสามารถเชื่อมโยงกับภารกิจงานจริง ลดความคลาดเคลื่อนของการตีความตำแหน่ง/บทบาท และทำให้การจัดทำแผนพัฒนากำลังคนเป็นไปอย่างมีทิศทาง [4]

ภายใต้กรอบดังกล่าว สมรรถนะสามารถจำแนกออกเป็น 4 มิติ ได้แก่ ความรู้ ทักษะ ความสามารถ และภารกิจการปฏิบัติงาน ทำให้การประเมินครอบคลุมทั้งด้านความเข้าใจ การปฏิบัติ และการประยุกต์ใช้ในสถานการณ์จริง ซึ่งสอดคล้องกับการจัดการความมั่นคงปลอดภัยสารสนเทศในระดับองค์กร [4], [7]

### 2.3 ความตระหนักรู้ด้านความปลอดภัยและปัจจัยด้านมนุษย์

แม้องค์กรจะมีมาตรการควบคุมและเทคโนโลยีที่สอดคล้องกับมาตรฐาน แต่ปัจจัยด้านมนุษย์ยังคงเป็นหนึ่งในจุดอ่อนสำคัญของระบบสารสนเทศ เนื่องจากพฤติกรรม การตัดสินใจ และความตระหนักรู้ของผู้ใช้งานสามารถก่อให้เกิดความเสี่ยงได้โดยตรง [2] ดังนั้น แนวคิดด้านความตระหนักรู้ด้านความปลอดภัยจึงมีบทบาทต่อการประเมินและเสริมสร้างพฤติกรรมที่ปลอดภัยของผู้ใช้งานในองค์กร [5]

ในเชิงการประเมิน การวัดระดับความตระหนักรู้และการตอบสนองต่อภัยคุกคามสามารถสะท้อนความพร้อม

เชิงพฤติกรรมของบุคลากร และช่วยสนับสนุนการตัดสินใจรับมือสถานการณ์ด้านไซเบอร์ได้อย่างเหมาะสม [7] ดังนั้น เครื่องมือประเมินควรสามารถวัดได้ทั้งระดับความรู้ การปฏิบัติ และการตัดสินใจในสถานการณ์จริง

**2.4 เกณฑ์การประเมินตามพฤติกรรมการเรียนรู้**

การออกแบบเครื่องมือประเมินสมรรถนะด้านความมั่นคงปลอดภัยไซเบอร์จำเป็นต้องกำหนดระดับความซับซ้อนของข้อคำถามให้สอดคล้องกับเนื้อหาและพฤติกรรมที่ต้องการวัด โดยมิติความรู้เกี่ยวข้องกับหลักการพื้นฐาน เช่น ภัยคุกคามไซเบอร์และปัจจัยด้านมนุษย์ที่ส่งผลต่อความมั่นคงปลอดภัยขององค์กร [1], [2] ขณะที่มิติทักษะเกี่ยวข้องกับการเลือกใช้แนวปฏิบัติหรือมาตรการควบคุม เช่น แนวทางตามมาตรฐาน ISO/IEC 27002 และการระบุพฤติกรรมเสี่ยงของผู้ใช้งาน [3], [5]

สำหรับมิติความสามารถและภารกิจการปฏิบัติงานเน้นการวัดความสามารถในการวิเคราะห์ การตัดสินใจ และการแก้ไขปัญหาเมื่อเผชิญสถานการณ์ด้านไซเบอร์ ซึ่งสอดคล้องกับแนวทางการจัดการความมั่นคงปลอดภัยสารสนเทศและการสร้างความพร้อมของบุคลากรในองค์กร [7]

ในการกำหนดระดับพฤติกรรมทางสติปัญญาของข้อคำถาม งานวิจัยนี้ประยุกต์ใช้แนวคิดพุทธิพิสัยของ Bloom เพื่อจัดระดับข้อคำถามตั้งแต่ระดับความจำและความเข้าใจ ระดับการประยุกต์ใช้ ไปจนถึงระดับการวิเคราะห์และประเมินผล [8] แนวทางดังกล่าวช่วยให้การออกแบบข้อคำถามมีความเป็นระบบ และสามารถสะท้อนระดับสมรรถนะของบุคลากรได้อย่างเหมาะสม

**2.5 งานวิจัยที่เกี่ยวข้องกับการประเมินสมรรถนะด้านไซเบอร์**

งานวิจัยด้านการประเมินความมั่นคงปลอดภัยไซเบอร์ได้พัฒนาแนวทางและเครื่องมือหลากหลาย โดยส่วนใหญ่มุ่งเน้นการวัดความตระหนักรู้และพฤติกรรมของผู้ใช้งาน ตัวอย่างเช่น Kruger และ Kearney เสนอแบบจำลองที่วัดความรู้ ทัศนคติ และพฤติกรรมของผู้ใช้งาน [5]

นอกจากนี้ Tsohou et al. ศึกษาการพัฒนาโปรแกรมสร้างความตระหนักรู้และชี้ให้เห็นความสำคัญของการประเมินอย่างต่อเนื่อง ขณะที่ Whitman และ Mattord เสนอแนวทางการประเมินที่ครอบคลุมทั้งด้านเทคโนโลยีและปัจจัยด้านมนุษย์ [6], [7]

อย่างไรก็ตาม งานวิจัยส่วนใหญ่ยังเน้นการประเมินในมิติเดียว หรือใช้คะแนนดิบที่ให้น้ำหนักเท่ากันทุกข้อ ซึ่งไม่สามารถสะท้อนความแตกต่างของระดับสมรรถนะได้อย่างชัดเจน

ดังนั้น งานวิจัยนี้จึงพัฒนากรอบการประเมินสมรรถนะด้านความมั่นคงปลอดภัยไซเบอร์แบบหลายมิติ (K-S-A-T) ร่วมกับการให้คะแนนแบบถ่วงน้ำหนักตามระดับความยาก เพื่อเพิ่มความแม่นยำและสะท้อนความพร้อมของบุคลากรได้อย่างครอบคลุม

## 3. วิธีดำเนินการวิจัย

**3.1 ประชากรและกลุ่มตัวอย่าง**

ประชากรในการวิจัยครั้งนี้เป็นบุคลากรของหน่วยงานภาครัฐที่มีการใช้เทคโนโลยีสารสนเทศในการดำเนินงาน โดยเป็นองค์กรขนาดกลางถึงขนาดใหญ่ที่มีโครงสร้างการทำงานด้านดิจิทัลอย่างเป็นระบบ กลุ่มตัวอย่างจำนวน 186 คน ได้มาจากการเก็บรวบรวมข้อมูลเพื่อสะท้อนระดับสมรรถนะและความพร้อมด้านความมั่นคงปลอดภัยไซเบอร์ของหน่วยงาน ทั้งนี้ กรอบสมรรถนะที่ใช้ในการประเมินสอดคล้องกับมาตรฐานที่เกี่ยวข้องและยุทธศาสตร์ความมั่นคงปลอดภัยไซเบอร์แห่งชาติ [4], [9]

แม้ว่างานวิจัยนี้จะดำเนินการเก็บข้อมูลจากหน่วยงานเพียงแห่งเดียว แต่กรอบการประเมินที่พัฒนาขึ้นสามารถประยุกต์ใช้กับองค์กรประเภทอื่นได้ โดยสามารถปรับเนื้อหาและระดับความยากของข้อคำถามให้เหมาะสมกับบริบทของแต่ละองค์กร

**3.2 กรอบการประเมินสมรรถนะ**

กรอบการประเมินในงานวิจัยนี้อาศัยแนวคิดแบบหลายมิติ (K-S-A-T) ได้แก่ ความรู้ (K) ทักษะ (S) ความสามารถ (A) และภารกิจการปฏิบัติงาน (T) เพื่อใช้เป็นโครงสร้างในการประเมินสมรรถนะด้านความมั่นคงปลอดภัยไซเบอร์ โดยอ้างอิงจาก NICE Framework ซึ่งเป็นมาตรฐานสากลด้านสมรรถนะกำลังคนไซเบอร์

กรอบดังกล่าวไม่จำกัดอยู่เพียงการสร้างแบบทดสอบ แต่เป็นแนวทางการประเมินที่สามารถประยุกต์ใช้กับองค์กรประเภทต่าง ๆ ได้ โดยแบบทดสอบที่ใช้เป็นกรณีศึกษาสำหรับสาธิตการใช้งานของกรอบการประเมิน

อย่างไรก็ตาม งานวิจัยนี้มุ่งเน้นการประยุกต์ใช้กรอบ K-S-A-T ในการออกแบบเครื่องมือ และไม่ได้ทำการวิเคราะห์โครงสร้างเชิงสถิติของโมเดล เช่น CFA

**3.3 เครื่องมือวิจัยและการตรวจสอบคุณภาพเครื่องมือ**

เครื่องมือที่ใช้ในการวิจัยเป็นแบบทดสอบปรนัยชนิดเลือกตอบ 4 ตัวเลือก จำนวน 60 ข้อ แบ่งออกเป็น 4 มิติ มิติละ 15 ข้อ โดยออกแบบข้อคำถามให้สอดคล้องกับระดับพฤติกรรมตามแนวคิดของ Bloom โดยมีการตรวจสอบคุณภาพเครื่องมือประกอบด้วย

3.3.1 ความตรงเชิงเนื้อหา (Content Validity) ตรวจสอบโดยผู้เชี่ยวชาญด้านความมั่นคงปลอดภัยไซเบอร์และเทคโนโลยีสารสนเทศ จำนวน 3 คน โดยใช้วิธี (IOC) และคัดเลือกข้อคำถามที่มีค่า IOC $\geq$ 0.50

3.3.2 การวิเคราะห์คุณภาพรายข้อ โดยพิจารณาค่าความยาก (p) และค่าอำนาจจำแนก (r) โดยคัดเลือกข้อคำถามที่มีค่า r ตั้งแต่ 0.20 ขึ้นไป เพื่อให้สามารถแยกแยะผู้มีสมรรถนะสูงและต่ำได้อย่างมีประสิทธิภาพ

3.3.3 การตรวจสอบความเชื่อมั่นของแบบทดสอบทั้งฉบับ โดยใช้สูตร KR-20 ซึ่งมีค่าเท่ากับ 0.96 แสดงให้เห็นว่าเครื่องมือมีความเชื่อมั่นอยู่ในระดับสูงมาก

**3.4 การให้คะแนนแบบถ่วงน้ำหนัก**

ในการวิจัยครั้งนี้ ผู้วิจัยใช้วิธีการให้คะแนนแบบถ่วงน้ำหนักตามระดับความยากของข้อคำถาม โดยกำหนดค่าน้ำหนักตามระดับ ดังแสดงในตารางที่ 1

**ตารางที่ 1:** การกำหนดค่าน้ำหนักตามระดับความยากของข้อคำถาม

| ระดับความยาก | ค่า p | น้ำหนัก (W) |
|---|---|---|
| ง่าย | > 0.80 | 0.2 |
| ปานกลาง | 0.60–0.79 | 0.4 |
| ค่อนข้างยาก | 0.40–0.59 | 0.6 |
| ยาก | < 0.40 | 0.8 |

**3.5 การคำนวณคะแนนถ่วงน้ำหนัก**

เมื่อผู้เข้าร่วมทำแบบทดสอบ จะนับจำนวนข้อที่ตอบถูกในแต่ละระดับความยาก และคูณด้วยค่าน้ำหนักของระดับนั้น เพื่อคำนวณเป็นคะแนนถ่วงน้ำหนักของแต่ละบุคคล โดยกำหนดให้ $n_{ij}$ คือ จำนวนข้อที่บุคคลที่ i ตอบถูกในระดับความยากที่ j และ $W_j$ คือ ค่าน้ำหนักของระดับนั้น คะแนนถ่วงน้ำหนักรวมของบุคคลที่ i คำนวณได้ดังสมการ (1)

$$W_i = \sum_j (n_{ij} \times w_j) \quad (1)$$

**3.6 การแปลงคะแนนและเกณฑ์การแปลผล**

คะแนนเต็มถ่วงน้ำหนักรวมกำหนดเท่ากับ 32.80 (ร้อยละ 100) ซึ่งเป็นค่าคะแนนสูงสุดที่เป็นไปได้ภายใต้โครงสร้างน้ำหนักที่กำหนด จากนั้นแปลงเป็นร้อยละของคะแนนเต็มเพื่อใช้ในการแปลผล ดังสมการ (2)

$$P_i = \frac{W_i}{32.8} \times 100 \quad (2)$$

นอกจากนี้ ได้กำหนดคะแนนเต็มถ่วงน้ำหนักรายมิติ ได้แก่ K = 7.2, S = 8.8, A = 10.4 และ T = 6.4 เพื่อให้สามารถวิเคราะห์ผลรายด้านได้อย่างชัดเจน โดยรวมเท่ากับ 32.8 (ร้อยละ 100) ดังนั้นร้อยละรายปัจจัยสามารถคำนวณได้ตามสมการ (3)

$$P_{K,i} = \frac{W_{K,i}}{7.2} \times 100,\ P_{S,i} = \frac{W_{S,i}}{8.8} \times 100,\ P_{A,i} = \frac{W_{A,i}}{10.4} \times 100,\ P_{T,i} = \frac{W_{T,i}}{6.4} \times 100 \quad (3)$$

เกณฑ์การแปลผลระดับสมรรถนะ (5 ระดับ) เมื่อได้ร้อยละคะแนนทั้งภาพรวมและรายปัจจัยแล้ว จะนำไปเทียบกับเกณฑ์การแปลผล 5 ระดับ ได้แก่ 80.00–100.00 คะแนน หมายถึงระดับสูงมาก 70.00–79.99 คะแนน หมายระดับสูง 60.00–69.99 คะแนน หมายถึง ระดับปานกลาง 50.00–59.99 คะแนน หมายถึง ระดับต่ำ และต่ำกว่า 50.00 คะแนน หมายถึง ยังไม่พร้อม

## 4. ผลการดำเนินการวิจัย

**4.1 ผลการวิเคราะห์ระดับความพร้อม**

ผลการวิเคราะห์จากกลุ่มตัวอย่างจำนวน 186 คน พบว่า คะแนนเฉลี่ยแบบถ่วงน้ำหนักเท่ากับ 27.36 (SD = 2.15) จากคะแนนเต็ม 32.80 คะแนน คิดเป็นร้อยละ 83.40 สะท้อนว่าระดับสมรรถนะโดยรวมอยู่ในระดับ

สูงมากตามเกณฑ์การแปลผลที่กำหนดไว้ นอกจากนี้ ผลการตรวจสอบค่าสัมประสิทธิ์ความเชื่อมั่น (reliability) ของแบบทดสอบจำนวน 60 ข้อ มีค่าเท่ากับ 0.96 ซึ่งเป็นตัวชี้วัดความสอดคล้องภายในของเครื่องมือวัด [10] โดยแนวทางการแปลผลที่ใช้กันทั่วไปกำหนดว่า ค่าความเชื่อมั่นตั้งแต่ 0.70 ขึ้นไปอยู่ในเกณฑ์ยอมรับได้ ดังนั้นค่าที่ได้จึงสะท้อนว่าแบบทดสอบมีความสอดคล้องภายในอยู่ใน ระดับสูงมาก และเหมาะสมสำหรับนำไปใช้ประเมินสมรรถนะด้านความมั่นคงปลอดภัยสารสนเทศ [11]

**4.2 ผลการวิเคราะห์สมรรถนะรายมิติ**

เมื่อจำแนกผลการประเมินสมรรถนะออกเป็น 4 มิติตามกรอบ ได้แก่ ความรู้ (K) ทักษะ (S) ความสามารถ (A) และการปฏิบัติหน้าที่ (T) พบว่า ค่าเฉลี่ยคะแนนถ่วงน้ำหนักของกลุ่มตัวอย่างจำนวน 186 คน มีความแตกต่างกันในแต่ละมิติ โดยภาพรวมคะแนนเฉลี่ยรายมิติอยู่ในระดับสูงทั้งหมด ทั้งนี้ มิติทักษะ (S) มีผลการประเมินสูงที่สุด คิดเป็นร้อยละ 84.46 มิติการปฏิบัติหน้าที่ (T) ร้อยละ 84.26 และมิติความสามารถ (A) ร้อยละ 83.21 ขณะที่ มิติความรู้ (K) มีผลการประเมินต่ำที่สุด คิดเป็น ร้อยละ 81.68 รายละเอียดแสดงดังตารางที่ 2

**ตารางที่ 2** คะแนนเฉลี่ยและร้อยละของสมรรถนะ

n = 186

| มิติสมรรถนะ | คะแนนเต็มถ่วงน้ำหนัก | คะแนนเฉลี่ย | ร้อยละ |
|---|---|---|---|
| (K) | 7.2 | 5.88 | 81.68 |
| (S) | 8.8 | 7.43 | 84.46 |
| (A) | 10.4 | 8.65 | 83.21 |
| (T) | 6.4 | 5.39 | 84.26 |
| **ภาพรวม** | **32.8** | **27.36** | **83.40** |

นอกจากนี้ เมื่อพิจารณากลุ่มบุคลากรที่มีคะแนนต่ำกว่า ร้อยละ 70 เพื่อใช้เป็นฐานในการกำหนดแนวทางพัฒนา พบว่า มิติความสามารถ (A) มีสัดส่วนผู้ต่ำกว่าเกณฑ์มากที่สุด คิดเป็น ร้อยละ 23.12, มิติความรู้ (K) ร้อยละ 20.43, มิติการปฏิบัติหน้าที่ (T) ร้อยละ 17.20 และมิติทักษะ (S) ร้อยละ 16.13 สะท้อนว่า แม้ผลการประเมินเฉลี่ยรายมิติจะอยู่ในระดับสูง แต่ยังมีช่องว่างสมรรถนะในบางมิติ โดยเฉพาะด้าน ความสามารถและความรู้ ที่ควรได้รับการพัฒนาอย่างเร่งด่วนและเฉพาะเจาะจงในขั้นตอนถัดไป รายละเอียดแสดงดังตารางที่ 3

**ตารางที่ 3** คะแนนต่ำกว่า ร้อยละ 70 ของสมรรถนะ

n = 186

| มิติสมรรถนะ | ต่ำกว่าร้อยละ 70 (คน) | คิดเป็นร้อยละ |
|---|---|---|
| (K) | 38 | 20.43 |
| (S) | 30 | 16.13 |
| (A) | 43 | 23.12 |
| (T) | 32 | 17.20 |
| **ภาพรวม** | 36 | **19.22** |

**4.3 ผลการวิเคราะห์กลุ่มบุคลากรที่มีคะแนนต่ำกว่าเกณฑ์ (ต่ำกว่าร้อยละ 70)**

เพื่อให้ผลการประเมินสามารถนำไปใช้กำหนดแนวทางพัฒนาบุคลากรได้อย่างตรงจุด งานวิจัยนี้ได้จำแนก กลุ่มบุคลากรที่มีคะแนนต่ำกว่า ร้อยละ 70 ซึ่งถือเป็นกลุ่มเป้าหมายสำคัญในการยกระดับสมรรถนะในระยะต่อไป ผลการวิเคราะห์พบว่า สัดส่วนผู้ที่มีคะแนนต่ำกว่าเกณฑ์แตกต่างกันในแต่ละมิติ โดยมิติที่พบกลุ่มคะแนนต่ำกว่า 70 มากที่สุดคือ ความสามารถ คิดเป็นร้อยละ 23.12 (43 คน) รองลงมา คือ ความรู้ ร้อยละ 20.43 (38 คน) ตามด้วย การปฏิบัติ ร้อยละ 17.20 (32 คน) และ ทักษะ ร้อยละ 16.13 (30 คน)

ผลดังกล่าวสะท้อนว่า แม้คะแนนเฉลี่ยรายมิติของหน่วยงานจะอยู่ในระดับสูง แต่ยังมีบุคลากรบางส่วนที่มีช่องว่างสมรรถนะต่ำกว่าเกณฑ์ โดยเฉพาะในมิติความสามารถ และ ความรู้ ซึ่งควรถูกจัดเป็นประเด็นเร่งด่วนสำหรับการพัฒนา ทั้งนี้ การนำผลการจำแนกกลุ่มต่ำกว่าร้อยละ 70 ไปใช้เชิงปฏิบัติ ควรดำเนินการต่อเนื่องใน 2 แนวทาง ได้แก่ วิเคราะห์เชิงลึกรายข้อ/รายปัจจัย เพื่อระบุหัวข้อที่บุคลากรตอบผิดมากที่สุด และจัดทำแผนพัฒนาบุคลากรรายกลุ่มและรายบุคคลให้สอดคล้องกับบทบาทหน้าที่ เพื่อยกระดับความพร้อมด้านการรักษาความปลอดภัยไซเบอร์ของหน่วยงานอย่างเป็นระบบและยั่งยืน [4], [7]

**5. สรุปผลและอภิปรายผลวิจัย**

ผลการวิจัยจากกลุ่มตัวอย่าง 186 คน แสดงให้เห็นว่ากรอบการประเมินสมรรถนะด้านความมั่นคงปลอดภัยไซเบอร์สามารถสะท้อนระดับสมรรถนะของบุคลากรได้อย่างเป็นระบบ โดยมีคะแนนเฉลี่ยแบบถ่วงน้ำหนักเท่ากับ 27.36 จาก 32.80 คะแนน (ร้อยละ 83.40, SD = 2.15) อยู่ในระดับสูงมาก เมื่อพิจารณาผลรายมิติตามกรอบ K-S-A-T พบว่าทุกมิติมีคะแนนเฉลี่ยอยู่ในระดับสูง โดยมิติทักษะและภารกิจการปฏิบัติงานมีคะแนนสูงที่สุด ขณะที่มิติความรู้มีคะแนนต่ำที่สุด สะท้อนว่าบุคลากรสามารถปฏิบัติและประยุกต์ใช้ได้ดี แต่ยังควรเสริมความเข้าใจเชิงหลักการด้านความมั่นคงปลอดภัยสารสนเทศโดยเฉพาะประเด็นที่เกี่ยวข้องกับปัจจัยด้านมนุษย์และมาตรฐานการควบคุมความปลอดภัย [2], [3]

การวิเคราะห์กลุ่มที่มีคะแนนต่ำกว่าร้อยละ 70 พบมากในมิติความสามารถ รองลงมาคือความรู้ ภารกิจการปฏิบัติงาน และทักษะ แสดงให้เห็นว่าความสามารถด้านการวิเคราะห์ การตัดสินใจ และการแก้ปัญหาในสถานการณ์ไซเบอร์ยังไม่สม่ำเสมอในบางกลุ่ม ซึ่งสอดคล้องกับแนวคิดด้านปัจจัยมนุษย์และการจัดการความมั่นคงปลอดภัยสารสนเทศในระดับองค์กร [2], [7]

การใช้คะแนนแบบถ่วงน้ำหนักช่วยเพิ่มความละเอียดในการประเมิน และสามารถสะท้อนความแตกต่างของสมรรถนะได้ดีกว่าคะแนนดิบ ทำให้สามารถระบุจุดแข็งและจุดอ่อนได้อย่างชัดเจน และนำไปใช้ในการวางแผนพัฒนาได้อย่างมีประสิทธิภาพ [7]

ดังนั้น หน่วยงานควรนำผลการประเมินไปใช้ในการพัฒนาบุคลากรทั้งระดับบุคคลและกลุ่ม โดยเน้นมิติความรู้และความสามารถ ควบคู่กับการเสริมทักษะเชิงปฏิบัติและการติดตามผลอย่างต่อเนื่อง เพื่อเพิ่มความพร้อมด้านความมั่นคงปลอดภัยไซเบอร์ในระยะยาว โดยการออกแบบการพัฒนาควรคำนึงถึงระดับพฤติกรรมการเรียนรู้และการประเมินผลที่เป็นระบบ [7], [8]

นอกจากนี้ เครื่องมือที่ใช้ในการวิจัยมีค่าความเชื่อมั่น KR-20 เท่ากับ 0.96 แสดงให้เห็นว่าเครื่องมือมีความน่าเชื่อถือและเหมาะสมต่อการนำไปใช้ประเมินสมรรถนะในบริบทหน่วยงานภาครัฐ โดยค่าความเชื่อมั่นดังกล่าวอยู่ในระดับสูงเมื่อพิจารณาตามแนวคิดด้านความสอดคล้องภายในของเครื่องมือวัด [10], [11]

อย่างไรก็ตาม งานวิจัยนี้มีข้อจำกัด ได้แก่ การเก็บข้อมูลจากหน่วยงานเพียงแห่งเดียว และการใช้แบบทดสอบปรนัยซึ่งอาจไม่สะท้อนพฤติกรรมจริงทั้งหมด ดังนั้น งานวิจัยในอนาคตควรขยายกลุ่มตัวอย่างและใช้วิธีประเมินรูปแบบอื่น เพื่อเพิ่มความแม่นยำของการประเมิน

## เอกสารอ้างอิง


[1] World Economic Forum, Global Cybersecurity Outlook 2023. Geneva, Switzerland: World Economic Forum, 2023.

[2] L. Hadlington, “Human factors in cybersecurity,” Computers & Security, vol. 68, pp. 56–68, 2017.

[3] ISO/IEC, ISO/IEC 27002:2022 Information security, cybersecurity and privacy protection — Information security controls. Geneva, Switzerland: ISO, 2022.

[4] W. Newhouse, S. Keith, B. Scribner, and G. Witte, Workforce Framework for Cybersecurity (NICE Framework), NIST Special Publication 800-181 Rev. 1. Gaithersburg, MD, USA: National Institute of Standards and Technology, 2020.

[5] H. A. Kruger and W. D. Kearney, “A prototype for assessing information security awareness,” Computers & Security, vol. 25, no. 4, pp. 289–296, 2006.

[6] A. Tsohou, M. Karyda, S. Kokolakis, and E. Kiountouzis, “Managing the introduction of information security awareness programmes in organisations,” European Journal of Information Systems, vol. 24, no. 1, pp. 38–58, 2015.

[7] M. E. Whitman and H. J. Mattord, Principles of Information Security, 7th ed. Boston, MA, USA: Cengage Learning, 2021.

[8] B. S. Bloom, J. T. Hastings, and G. F. Madaus, Handbook on Formative and Summative Evaluation of Student Learning. New York, NY, USA: McGraw-Hill, 1971.

[9] OECD Digital Economy Policy Portal, “Cybersecurity policy and action plan 2022–2027 (Thailand).” Accessed: Feb. 26, 2026.

[10] L. J. Cronbach, “Coefficient alpha and the internal structure of tests,” Psychometrika, vol. 16, pp. 297–334, 1951, doi: 10.1007/BF02310555.

[11] M. Tavakol and R. Dennick, “Making sense of Cronbach’s alpha,” International Journal of Medical Education, vol. 2, pp. 53–55, 2011, doi: 10.5116/ijme.4dfb.8dfd.

# การเพิ่มความหลากหลายโดยอิงจากผู้แต่งในระบบแนะนำหนังสือด้วยเทคนิคการลดความคล้ายคลึงภายในรายการ

# Author-based Diversification in Book Recommender Systems using Intra-List Similarity Reduction Techniques

*ลักษมี โนนลือชา (Laksamee Noneluecha)* [1] *และ สุภัคพงศ์ จินารัตน์ (Supakpong Jinarat)* [2]
[1,2] *วิทยาลัยวิศวกรรมศาสตร์และเทคโนโลยี มหาวิทยาลัยธุรกิจบัณฑิตย์*
[1]*67130007@dpu.ac.th,* [2]*supakpong.jin@dpu.ac.th*

## บทคัดย่อ

*งานวิจัยนี้มีวัตถุประสงค์เพื่อแก้ปัญหาภาวะถุงกรองข้อมูล (Filter Bubble) ในระบบแนะนำหนังสือ โดยประยุกต์ใช้เทคนิค SVD ร่วมกับกลไกการจัดลำดับใหม่ (Diversification Re-ranking) เพื่อเพิ่มความหลากหลายของรายชื่อผู้แต่ง จากการทดสอบด้วยชุดข้อมูล Book-Crossing พบว่าเทคนิคนี้สามารถลดค่าความคล้ายคลึงภายในรายการ (ILS) จาก 0.1431 เหลือเพียง 0.0018 ซึ่งช่วยขจัดปัญหาความซ้ำซากของผู้แต่งได้เกือบทั้งหมด แม้ว่าค่าความแม่นยำ (Precision) จะลดลงร้อยละ 26.87 ตามกลไกการแลกเปลี่ยน (Trade-off) ระหว่างความแม่นยำและความหลากหลาย แต่อัลกอริทึมนี้ส่งผลดีต่อการเปิดกว้างให้ผู้ใช้ค้นพบนักเขียนใหม่ๆ และเป็นแนวทางที่มีประสิทธิภาพสำหรับระบบแนะนำในห้องสมุดดิจิทัลและธุรกิจ E-commerce ในการสร้างความพึงพอใจแก่ผู้ใช้ในระยะยาว*

**คำสำคัญ:** ระบบแนะนำหนังสือ การกระจายตัวเลือก SVD ความหลากหลายของรายการแนะนำ ความคล้ายคลึงภายในรายการ (ILS)

## Abstract

*This research aims to address the "Filter Bubble" phenomenon in book recommender systems by integrating Singular Value Decomposition (SVD) with a Diversification Re-ranking mechanism to increase author variety. Testing on the Book-Crossing dataset demonstrates that this approach significantly reduces Intra-List Similarity (ILS) from 0.1431 to 0.0018, nearly eliminating author redundancy in recommendations. Although precision decreased by 26.87% due to the accuracy-diversity trade-off, the proposed algorithm effectively encourages the discovery of new authors. This provides a valuable framework for digital libraries and E-commerce platforms to enhance long-term user satisfaction through balanced and diverse recommendations.*

**Keywords:** Book Recommendation System, Diversification, SVD, Recommendation Diversity, Intra-List Similarity (ILS).

## 1. บทนำ

ในยุคที่มีทรัพยากรสารสนเทศมหาศาล ระบบแนะนำ (Recommender Systems) โดยเฉพาะเทคนิค SVD ได้กลายเป็นเครื่องมือหลักที่ช่วยคัดกรองหนังสือตามความสนใจของผู้ใช้ได้อย่างแม่นยำ [1] อย่างไรก็ตาม ระบบที่มุ่งเน้นเพียงความแม่นยำมักก่อให้เกิดปรากฏการณ์ "ถุงกรองข้อมูล" (Filter Bubble) ที่นำเสนอหนังสือจากผู้แต่ง

หรือแนวทางเดิมซ้ำ ๆ จนปิดกั้นโอกาสในการค้นพบนักเขียนกลุ่มใหม่ ส่งผลให้ความพึงพอใจในการใช้งานระยะยาวลดลง

งานวิจัยนี้จึงมีวัตถุประสงค์เพื่อพัฒนาอัลกอริทึมแนะนำหนังสือที่เพิ่ม ความหลากหลายของรายชื่อผู้แต่ง (Author-based Diversification) โดยเปรียบเทียบประสิทธิภาพระหว่างโมเดล SVD พื้นฐาน และโมเดลที่ปรับปรุงใหม่ (Diversified SVD) เพื่อวิเคราะห์จุดสมดุลที่เหมาะสมระหว่างความแม่นยำ (Precision) และความหลากหลาย (Diversity)

ผู้วิจัยดำเนินการผ่านชุดข้อมูล Book-Crossing Dataset โดยประยุกต์ใช้ SVD ร่วมกับเทคนิคการจัดลำดับใหม่ (Re-ranking) และวัดผลด้วยตัวชี้วัด Precision, Recall และ Intra-List Similarity (ILS) ผลลัพธ์ที่ได้จะเป็นแนวทางสำคัญในการยกระดับประสบการณ์ผู้ใช้สำหรับธุรกิจ E-commerce และห้องสมุดดิจิทัล ให้ก้าวข้ามขีดจำกัดของพฤติกรรมการอ่านแบบเดิมสู่การค้นพบใหม่ ๆ ที่มีคุณภาพ

## 2. ทบทวนวรรณกรรม

ในปัจจุบันการพัฒนาระบบแนะนำหนังสือ (Book Recommender Systems) ส่วนใหญ่ยังคงให้ความสำคัญกับประสิทธิภาพด้านความแม่นยำ (Accuracy) ในการทำนายความชอบของผู้ใช้ [1], [3] โดยมีเทคนิค Collaborative Filtering (CF) เป็นแกนหลัก ทั้งในรูปแบบ Item-to-Item ที่ริเริ่มโดย Amazon [2] และการใช้ Matrix Factorization อย่าง Singular Value Decomposition (SVD) เพื่อเพิ่มประสิทธิภาพในการประมวลผล [1] นอกจากนี้ยังมีการประยุกต์ใช้ Machine Learning [5], [11] และการสร้างโปรไฟล์ผู้ใช้เพื่อสนับสนุนการเรียนรู้ด้วยตนเองและการแนะนำหนังสือในห้องสมุด [4], [6] อย่างไรก็ตาม งานวิจัยในระยะหลังเริ่มชี้ให้เห็นว่าความแม่นยำเพียงอย่างเดียวไม่สามารถการันตีความพึงพอใจสูงสุดได้ เนื่องจากผู้ใช้อาจรู้สึกเบื่อหน่ายกับรายการที่คล้ายเดิมมากเกินไป หรือเผชิญกับปัญหาความลำเอียงจากความนิยม (Popularity Bias) [7]

ด้วยเหตุนี้จึงได้นำเสนอแนวคิด Topic Diversification เพื่อขยาย "สเปกตรัมความสนใจ" (Full spectrum of interests) ให้ครอบคลุมมิติที่หลากหลายของผู้ใช้ [8] แทนการแนะนำเพียงสิ่งที่คล้ายกับรายการล่าสุดที่ผู้ใช้เพิ่งบริโภค [9] โดยมีการใช้มาตรวัด Intra-list Similarity (ILS) เพื่อตรวจสอบความคล้ายคลึงกันภายในรายการแนะนำ หากค่า ILS สูงเกินไป ระบบจะใช้อัลกอริทึมการจัดอันดับใหม่ (Re-ranking) เพื่อลดความซ้ำซ้อนเชิงเนื้อหา ซึ่งแม้ว่ากระบวนการนี้อาจส่งผลให้ค่าความแม่นยำเฉลี่ยลดลงบ้าง แต่ผลการทดสอบกลับยืนยันว่าระดับความพึงพอใจของผู้ใช้จริงนั้นเพิ่มสูงขึ้นอย่างมีนัยสำคัญ เนื่องจากระบบสามารถมอบคุณค่าในด้านความแปลกใหม่ (Novelty) และความประหลาดใจเชิงบวก (Serendipity) ได้ดีกว่าเดิม [10], [11]

แนวคิดดังกล่าวยังได้ถูกต่อยอดไปสู่การพิจารณาความหลากหลายโดยอิงจากผู้แต่ง (Diversity by Author) เนื่องจากในโดเมนหนังสือนั้น "ผู้แต่ง" เปรียบเสมือนตัวแทนของสไตล์และเนื้อหาที่มีเอกลักษณ์เฉพาะตัว ซึ่งสอดคล้องกับงานที่พิสูจน์ว่าสไตล์การเขียน (Writing Style) สามารถใช้เป็นฟีเจอร์สำคัญในการจำแนกและแนะนำหนังสือได้ [12] การประยุกต์ใช้แนวคิดความหลากหลายในบริบทนี้จึงไม่ใช่การสุ่มผู้แต่งแบบไร้ทิศทาง แต่เป็นการคัดสรรผู้แต่งที่มีสไตล์การเขียนสอดคล้องกับความสนใจในด้านต่าง ๆ ของผู้ใช้ (User's interest features) เพื่อสร้างสมดุลระหว่างความคุ้นเคยและความแปลกใหม่ ซึ่งในปัจจุบันยังมีการนำเทคโนโลยีขั้นสูงอย่าง Large Language Models (LLM) มาช่วยเสริมประสิทธิภาพในการเข้าใจบริบทที่ซับซ้อนเหล่านี้อีกด้วย [13]

## 3. ขั้นตอนวิธีการดำเนินการวิจัย

ในการวิจัยครั้งนี้ผู้วิจัยได้กำหนดขั้นตอนการดำเนินงานอย่างเป็นระบบ เพื่อพัฒนาและประเมินผลระบบแนะนำหนังสือที่เน้นความหลากหลายของผู้แต่ง โดยแบ่งขั้นตอนออกเป็น 5 ส่วนหลัก ดังนี้:

**3.1 การรวบรวมและเตรียมข้อมูล (Data Collection and Preprocessing)**

ผู้วิจัยใช้ชุดข้อมูลมาตรฐาน Book-Crossing Dataset โดยมีขั้นตอนการเตรียมข้อมูลที่กระชับดังนี้:

1. **การคัดเลือกข้อมูล (Data Filtering):** คัดเฉพาะหนังสือที่มีชื่อผู้แต่งชัดเจนและผู้ใช้ที่มีประวัติการให้คะแนน เพื่อให้โมเดลเรียนรู้ได้อย่างมีประสิทธิภาพ
2. **การจัดการข้อมูลสูญหาย (Missing Value Handling):** กำจัดระเบียนที่ขาดชื่อผู้แต่งหรือ ISBN เพื่อป้องกันความผิดพลาดในการคำนวณ
3. **การเข้ารหัสข้อมูล (Label Encoding):** แปลง ISBN และ User-ID เป็นดัชนี (Index) เพื่อเพิ่มความเร็วในการประมวลผลของอัลกอริทึม

**3.2 การแบ่งชุดข้อมูลเพื่อการทดลอง (Train/Test Split)**

เพื่อให้การวัดผลประสิทธิภาพของระบบแนะนำเป็นไปอย่างแม่นยำและน่าเชื่อถือ ผู้วิจัยได้ดำเนินการแบ่งชุดข้อมูลออกเป็น 2 ส่วนหลัก โดยใช้เทคนิคการสุ่มแยกข้อมูล (Train-Test Split) ดังนี้ ส่วนแรกคือ ชุดข้อมูลสำหรับฝึกฝน (Training Set) ในสัดส่วนร้อยละ 80 เพื่อใช้เป็นข้อมูลตั้งต้นให้อัลกอริทึม SVD ได้เรียนรู้ความสัมพันธ์และรูปแบบความสนใจระหว่างผู้ใช้กับรายชื่อหนังสืออย่างละเอียด และส่วนที่สองคือ ชุดข้อมูลสำหรับทดสอบ (Test Set) ในสัดส่วนร้อยละ 20 ซึ่งทำหน้าที่เป็นข้อมูลอ้างอิงในการทดสอบประสิทธิภาพการทำนาย

**3.3 การสร้างโมเดลแนะนำพื้นฐาน (Baseline Model with SVD)**

ผู้วิจัยเลือกใช้เทคนิค Singular Value Decomposition (SVD) เป็นโมเดลพื้นฐาน (Baseline) โดยมีกระบวนการหลักคือการทำ Matrix Factorization เพื่อแยกตัวประกอบเมทริกซ์ระหว่างผู้ใช้และหนังสือออกมาเป็นพารามิเตอร์แฝง (Latent Factors) ที่แสดงถึงความเชื่อมโยงเชิงลึก

ในการวิจัยนี้ ผู้วิจัยกำหนดค่าพารามิเตอร์สำหรับโมเดล TruncatedSVD โดยใช้จำนวนองค์ประกอบหลัก (n_components) เท่ากับ 50 เพื่อรักษาความสัมพันธ์ของพารามิเตอร์แฝงที่สำคัญ และกำหนดค่า random_state เป็น 42 เพื่อควบคุมความคงที่ของผลการสุ่มในการแยกตัวประกอบเมทริกซ์ทุกครั้ง

จากนั้นจึงนำค่าเหล่านี้มาทำ Rating Prediction เพื่อพยากรณ์คะแนนที่ผู้ใช้น่าจะให้กับหนังสือที่ยังไม่เคยอ่าน และสรุปผลด้วยวิธี Top-N Recommendation โดยคัดเลือกหนังสือที่มีคะแนนพยากรณ์สูงสุด 10 อันดับแรก เพื่อใช้เป็นฐานในการประมวลผลเพิ่มความหลากหลายในขั้นตอนต่อไป

**3.4 การพัฒนาเทคนิคการเพิ่มความหลากหลาย (Author-based Diversification)**

หัวใจสำคัญของงานวิจัยนี้คือการนำรายการ Top-N จาก SVD มาผ่านกระบวนการ Re-ranking โดยมีขั้นตอนดังนี้:

1. **Author Checking:** ตรวจสอบรายชื่อผู้แต่งในรายการแนะนำ 10 อันดับแรก
2. **Redundancy Filter:** หากพบหนังสือที่มีผู้แต่งซ้ำกับหนังสือที่ถูกเลือกไว้ก่อนหน้าในรายการ ระบบจะทำการข้ามหนังสือเล่มนั้น และเลือกหนังสือลำดับถัดไปที่มีผู้แต่งไม่ซ้ำกันขึ้นมาแทน
3. **Final List Generation:** ดำเนินการจนกว่าจะได้รายการแนะนำครบ 10 เล่มที่มีความหลากหลายของผู้แต่งสูงสุด

**3.5 การวัดผลและประเมินประสิทธิภาพ (Evaluation Metrics)**

ผู้วิจัยใช้ตัวชี้วัด 3 ด้านเพื่อเปรียบเทียบประสิทธิภาพ:

1. **Precision:** วัดความแม่นยำว่าหนังสือที่แนะนำไป ตรงกับที่ผู้ใช้อ่านจริงใน Test Set หรือไม่
2. **Recall:** วัดความสามารถในการครอบคลุมหนังสือที่ผู้ใช้อ่านจริงทั้งหมด
3. **Intra-List Similarity (ILS):** วัดค่าความคล้ายคลึงเฉลี่ยภายในรายการแนะนำ โดยคำนวณ

จากสัดส่วนของผู้แต่งที่ซ้ำกันในรายการ (ค่า ILS ยิ่งต่ำ ยิ่งมีความหลากหลายสูง)

$$ILS(P) = \frac{\left(\sum_{b_i \in P} \sum_{b_j \in P, j>i} c(b_i, b_j)\right)}{\left(\frac{|P|(|P|-1)}{2}\right)}$$

สูตรนี้เป็นการหา **"ค่าเฉลี่ยความคล้ายคลึง"** ของทุกคู่หนังสือที่มีอยู่ในรายการแนะนำ

อธิบายสมการ:

1. $P$: คือรายการหนังสือที่แนะนำ (เช่น List 10 เล่ม)
2. $b_i, b_j$: คือหนังสือเล่มที่ i และเล่มที่ j ในรายการนั้น
3. $c(b_i, b_j)$: คือฟังก์ชันวัดความคล้าย (Similarity) ระหว่างหนังสือสองเล่ม โดยถ้าผู้แต่งคนเดียวกัน c = 1, ถ้าคนละคน c = 0
4. $\sum$ (Summation): คือการนำความคล้ายของหนังสือ "ทุกคู่ที่เป็นไปได้" มาบวกกัน (โดยไม่นับคู่ที่ซ้ำกัน เช่น ถ้าวัด A กับ B แล้ว จะไม่วัด B กับ A ซ้ำ)
5. $|P|$: คือจำนวนหนังสือทั้งหมดในรายการ (เช่น 10 เล่ม)
6. $\frac{|P|(|P|-1)}{2}$ ส่วนนี้คือสูตรการหา **"จำนวนคู่ทั้งหมด"** ที่เป็นไปได้ในรายการ

ค่า ILS ที่คำนวณได้จะอยู่ระหว่าง **0 ถึง 1**:

1. **ถ้า ILS เข้าใกล้ 1**: หมายความว่าหนังสือในลิสต์ส่วนใหญ่แต่งโดยคนเดียวกัน (ความหลากหลาย **ต่ำมาก**)
2. **ถ้า ILS เข้าใกล้ 0**: หมายความว่าหนังสือในลิสต์แทบไม่มีผู้แต่งซ้ำกันเลย (ความหลากหลาย **สูงมาก**)

## 4. ผลการดำเนินงาน

ในบทนี้ผู้วิจัยนำเสนอผลการเปรียบเทียบประสิทธิภาพระหว่างระบบแนะนำหนังสือแบบพื้นฐาน (Baseline SVD) และระบบแนะนำหนังสือที่ผ่านการเพิ่มความหลากหลาย (Diversified SVD) โดยมุ่งเน้นการวิเคราะห์ใน 2 มิติหลัก ได้แก่ ความแม่นยำ (Accuracy) และความหลากหลายของเนื้อหา (Diversity)

### 4.1 ผลการทดสอบประสิทธิภาพเชิงปริมาณ (Quantitative Results)

จากการทดสอบกับชุดข้อมูล Book-Crossing โดยวัดผลผ่านค่า Precision, Recall และ ILS ได้ผลลัพธ์ดังตารางที่ 4.1:

**ตารางที่ 4.1:** เปรียบเทียบประสิทธิภาพระหว่างโมเดล Baseline และ Diversified

| Metrics | Baseline (SVD) | Diversified (SVD) |
|---|---|---|
| Precision | 0.0335 | 0.0245 |
| Recall | 0.0169 | 0.0114 |
| ILS (Diversity) | 0.1431 | 0.0018 |

### 4.2 การวิเคราะห์ผลการทดลอง (Discussion)

การวิเคราะห์ความหลากหลายพบว่า เทคนิค Re-ranking สามารถลดค่า ILS จาก 0.1431 เหลือ 0.0018 (เกือบ 100%) ทำให้รายการแนะนำ 10 อันดับแรกแทบไม่มีผู้แต่งซ้ำกัน ช่วยเพิ่มโอกาสการค้นพบนักเขียนกลุ่ม Long Tail แม้ค่าความแม่นยำ Precision ลดลง 26.87% และ Recall ลดลง 32.54% ตามทฤษฎี Accuracy-Diversity Trade-off เนื่องจากระบบเลือกแนะนำผู้แต่งรายใหม่ที่มีคะแนนรองลงมาแทนที่ผู้แต่งเดิมที่คุ้นเคย ทั้งนี้ กรณีศึกษา User-ID 276307 ยืนยันผลลัพธ์ชัดเจนว่าระบบเปลี่ยนจากการแนะนำหนังสือที่กระจุกตัวอยู่กับผู้แต่งคนเดียว (เช่น Stephen King 3 เล่ม) เป็นการแนะนำผู้แต่ง 10 ท่านที่ไม่ซ้ำกัน ช่วยขยายขอบเขตการอ่านของผู้ใช้ให้กว้างขึ้นอย่างมีประสิทธิภาพ

### 4.3 สรุปผลการวิเคราะห์

แม้เทคนิคที่นำเสนอจะทำให้ความแม่นยำลดลงเล็กน้อยในเชิงสถิติ แต่ในแง่ของประสบการณ์ผู้ใช้ (User Experience) การลดค่า ILS จนเกือบเป็นศูนย์นั้นมีประโยชน์มากกว่าในการแก้ไขปัญหา Filter Bubble และช่วยสร้างระบบแนะนำที่มีความ "เป็นกลาง" และ "ครอบคลุม" (Catalog Coverage) มากขึ้น ซึ่งถือเป็นหัวใจ

สำคัญของการพัฒนาแนวทางการค้นพบนักเขียนใหม่ตามวัตถุประสงค์ของงานวิจัย

## 5. สรุป

จากการพัฒนาและทดสอบระบบแนะนำหนังสือโดยใช้ชุดข้อมูล Book-Crossing พบว่าการประยุกต์ใช้เทคนิคการจัดลำดับใหม่ (Re-ranking) โดยใช้เกณฑ์ผู้แต่งสามารถยกระดับความหลากหลาย (Diversity) ของรายการแนะนำได้อย่างมีประสิทธิภาพสูงสุด โดยค่าความคล้ายคลึงภายในรายการ (Intra-List Similarity: ILS) มีการปรับตัวลดลงอย่างมีนัยสำคัญจาก 0.1431 ในโมเดลพื้นฐาน เหลือเพียง 0.0018 ในโมเดลที่ปรับปรุงแล้ว ซึ่งผลลัพธ์นี้ชี้ให้เห็นว่าระบบสามารถขจัดปัญหาความซ้ำซากของผู้แต่งในรายการแนะนำได้เกือบทั้งหมด อย่างไรก็ตาม การมุ่งเน้นความหลากหลายส่งผลกระทบโดยตรงต่อความแม่นยำ (Accuracy) ของระบบ โดยพบว่าค่า Precision ลดลงประมาณร้อยละ 26.87 และค่า Recall ลดลงร้อยละ 32.54 ทั้งนี้เนื่องจากระบบได้คัดเลือกหนังสือจากผู้แต่งรายใหม่ที่มีค่าความน่าจะเป็นในลำดับรองลงมาเข้ามาแทนที่หนังสือจากผู้แต่งเดิมที่มีคะแนนคาดการณ์สูงสุด

ในส่วนของการอภิปรายผล ผลการทดลองนี้เป็นการยืนยันถึงปรากฏการณ์ความสัมพันธ์แบบแลกเปลี่ยนระหว่างความแม่นยำและความหลากหลาย (Accuracy-Diversity Trade-off) ซึ่งสอดคล้องกับแนวคิดของ Ziegler et al. (2005) ที่ระบุว่าประสิทธิภาพทั้งสองด้านนี้มักมีความสัมพันธ์เป็นส่วนกลับซึ่งกันและกัน อย่างไรก็ตาม ในบริบทของการแนะนำหนังสือ ความหลากหลายถือเป็นปัจจัยที่มีความสำคัญอย่างยิ่งในการช่วยลดความน่าเบื่อและป้องกันภาวะถุงกรองข้อมูล (Filter Bubble) การที่ค่า ILS ลดลงจนเกือบเป็นศูนย์สะท้อนให้เห็นว่าอัลกอริทึมที่พัฒนาขึ้นมีความเหมาะสมอย่างมากสำหรับการนำไปประยุกต์ใช้ในระบบที่ต้องการเปิดกว้างสู่ข้อมูลใหม่ๆ และสร้างความเป็นธรรมให้กับนักเขียนหลากหลายท่านในระบบนิเวศของทรัพยากรสารสนเทศ

## เอกสารอ้างอิง


[1] Ishak, A. Yahya, and Yusri, "Book Recommendation System Based on Collaborative Filtering: User-Based, Item-Based, and Singular Value Decomposition Analysis," *JSCE, vol. 6, no. 4, pp. 329-342*, Oct. 2025.

[2] G. Linden, B. Smith, and J. York, "Amazon.com recommendations: Item-to-item collaborative filtering," *IEEE Internet Comput.*, vol. 7, no. 1, pp. 76-80, 2003.

[3] X. Li et al., "A Hybrid Recommendation Method Based on Feature for Offline Book Personalization," *Journal of Computers*, vol. 12, no. 10, pp. 911-920, Oct. 2017.

[4] W. Sriphalang and W. Sriurai, "The Development of Recommender System for E-Library by Using Collaborative Filtering and User Profile," *Srinakharinwirot University Journal (Science and Technology)*, vol. 9, no. 18, pp. 165-177, 2017.

[5] D. Sarma, T. Mittra, and M. S. Hossain, "Personalized Book Recommendation System using Machine Learning Algorithm," *Int. J. Adv. Comput. Sci. Appl. (IJACSA)*, vol. 12, no. 1, pp. 627-633, 2021.

[6] K. Amphawan, A. Jitpattanakul, and W. Phajaroen, "Supplementary books suggestion system for promoting self-reading and self-learning," Burapha University, Chonburi, Thailand, Final Research Report, 2018.

[7] M. Naghiaei, H. A. Rahmani, and M. Dehghan, "The Unfairness of Popularity Bias in Book Recommendation," *arXiv preprint arXiv:2202.13446*, Feb. 2022.

[8] C. N. Ziegler, S. M. McNee, J. A. Konstan, and G. Lausen, "Improving recommendation lists through topic diversification," in *Proc. 14th Int. Conf. World Wide Web (WWW '05)*, 2005, pp. 22-32.

[9] R. J. Mooney and L. Roy, "Content-based book recommending using learning for text categorization," in *Proc. 5th ACM Conf. Digital Libraries (DL '00)*, 2000, pp. 195-204.

[10] M. Ge, C. Delgado-Battenfeld, and D. Jannach, "Beyond accuracy: evaluating recommender systems by coverage and serendipity," in *Proc. 4th ACM Conf. Recommender Systems (RecSys '10)*, 2010, pp. 257-260.

[11] M. Kaminskas and D. Bridge, "Diversity, Serendipity, Novelty, and Coverage: A Survey and Empirical Analysis of Beyond-Accuracy Objectives in Recommender Systems," *ACM Trans. Interact. Intell. Syst.*, vol. 7, no. 1, pp. 1-42, Dec. 2016.

[12] H. Alharthi, D. Inkpen, and S. Szpakowicz, "Authorship Identification for Literary Book Recommendations," in *Proc. 27th Int. Conf. Computational Linguistics (COLING 2018)*, 2018, pp. 390-400.
[13] A. Zhiyuli, Y. Chen, X. Zhang, and X. Liang, "BookGPT: A General Framework for Book Recommendation Based on a Large Language Model," *arXiv preprint arXiv:2305.15673*, May 2023.

# ระบบช่วยการตัดสินใจในการรับนักเรียนเข้าศึกษาต่อ ในหลักสูตรวิทยาการคอมพิวเตอร์ มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี

# The Decision Assistant System for Student Admission in Computer Science Curricurum, Rajamangala University of Technology Thanyaburi

*วีณา จันทร์รัชชกูล (Weena Janratchakool)[1] อุไรวรรณ อินทร์แหยม (Uraiwan Inyaem)[2]*
*ชญานนท์ เปรมวิจิตร (Chayanon Premvijitr)[3] และเอมผกา เหลืองพูนทรัพย์ (Aimphagar Luangphusab)[4]*
*[1,2]ภาควิชาคณิตศาสตร์และวิทยาการคอมพิวเตอร์*
*คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี*
*[1]weena_j@rmutt.ac.th, [2]uraiwan.inyaem@rmutt.ac.th, [3]1161109051010@mail.rmutt.ac.th,*
*[4]1161109051028@mail.rmutt.ac.th*

## บทคัดย่อ

*งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาระบบช่วยการตัดสินใจในการรับนักเรียนเข้าศึกษาต่อในสาขาวิชาวิทยาการคอมพิวเตอร์ เพื่อใช้เป็นข้อมูลสนับสนุนในการพิจารณาคัดเลือกผู้สมัครเข้าศึกษา โดยใช้ชุดข้อมูลจำนวน 387 รายการ จากนั้นดำเนินการเตรียมข้อมูลด้วยการคัดกรองและลบข้อมูลที่ไม่เกี่ยวข้องออกจากชุดข้อมูล และคัดเลือกคุณลักษณะที่สำคัญ ได้แก่ ลำดับที่เลือกสาขาวิชาวิทยาการคอมพิวเตอร์ เกรดเฉลี่ยรวม ความถนัดทางคณิตศาสตร์ ความถนัดทางวิทยาศาสตร์ ผลรวมหน่วยกิต คะแนนโอเน็ตภาษาไทย คะแนนโอเน็ตสังคมศึกษา คะแนนโอเน็ตภาษาอังกฤษ คะแนนโอเน็ตคณิตศาสตร์ คะแนนโอเน็ตวิทยาศาสตร์ คะแนนการทดสอบวิชาความถนัดทั่วไปภาษาไทย และคะแนนการทดสอบวิชาความถนัดทั่วไปภาษาอังกฤษ ในการสร้างโมเดล งานวิจัยเลือกใช้โปรแกรม Microsoft Azure Machine Learning Studio และประยุกต์ใช้เทคนิคเหมืองข้อมูลจำนวน 4 เทคนิค ได้แก่ Two-Class Neural Network, Two-Class Support Vector Machine, Two-Class Bayes Point Machine และ Two-Class Averaged Perceptron ผลการทดลองพบว่าเทคนิค Two-Class Neural Network ให้ค่าความแม่นยำสูงที่สุด เท่ากับร้อยละ 89.90 จากนั้นได้นำโมเดลที่มีประสิทธิภาพสูงสุดมาพัฒนาเป็นระบบช่วยการตัดสินใจในรูปแบบเว็บไซต์ สำหรับการรับนักเรียนเข้าศึกษาต่อในหลักสูตรวิทยาการคอมพิวเตอร์ มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี โดยใช้ระบบฐานข้อมูล MySQL พัฒนาส่วนติดต่อผู้ใช้ด้วยเฟรมเวิร์ค Angular และพัฒนาส่วนประมวลผลฝั่งเซิร์ฟเวอร์ด้วยเครื่องมือ Node.js ผลการประเมินระบบโดยผู้เชี่ยวชาญจำนวน 3 คน โดยใช้แบบสอบถาม พบว่าระบบมีค่าเฉลี่ยเท่ากับ 4.38 จาก 5 ซึ่งอยู่ในระดับความพึงพอใจมาก และมีค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.14 แสดงให้เห็นว่าผู้เชี่ยวชาญมีความคิดเห็นที่สอดคล้องกันในระดับสูง*

**คำสำคัญ:** ระบบช่วยการตัดสินใจ เหมืองข้อมูล การคัดเลือกนักศึกษา วิทยาการคอมพิวเตอร์ การเรียนรู้ของเครื่อง

## Abstract

*This research aims to develop a decision support system for student admission to the Computer Science program in order to provide supportive information for admission decision-making. A dataset consisting of 387 records was utilized. Data preprocessing was performed by removing irrelevant attributes, and significant features were selected, including admission preference order for the Computer Science program, cumulative grade point average, mathematical aptitude, scientific aptitude, total*

*accumulated credits, O-NET scores in Thai language, social studies, English, mathematics, and science, as well as General Aptitude Test (GAT) scores in Thai language and English. Microsoft Azure Machine Learning Studio was employed to construct predictive models using four data mining techniques: Two-Class Neural Network, Two-Class Support Vector Machine, Two-Class Bayes Point Machine, and Two-Class Averaged Perceptron. Experimental results indicate that the Two-Class Neural Network achieved the highest classification accuracy of 89.90%. The best-performing model was subsequently implemented as a web-based decision support system for student admission to the Computer Science program at Rajamangala University of Technology Thanyaburi. The system utilizes MySQL as the database management system, Angular framework for the front-end development, and Node.js for the back-end processing. System evaluation was conducted by three domain experts using a questionnaire. The results revealed a mean satisfaction score of 4.38 out of 5, indicating a high level of satisfaction, with a standard deviation of 0.14, which suggests a high level of consistency among expert opinions.*

**Keyword:** Decision Support System, Data Mining, Student Admission, Computer Science, Machine Learning.

## 1. บทนำ

ปัญหาการออกจากการศึกษากลางคันของนักศึกษาระดับปริญญาตรีเป็นประเด็นสำคัญที่สถาบันอุดมศึกษากำลังเผชิญอยู่ในปัจจุบัน โดยมีแนวโน้มเพิ่มสูงขึ้นอย่างชัดเจนในช่วงปีที่ผ่านมา ปัญหาดังกล่าวไม่เพียงส่งผลกระทบต่อคุณภาพการจัดการศึกษาและประสิทธิภาพในการใช้ทรัพยากรของสถาบันเท่านั้น แต่ยังสะท้อนถึงความไม่สอดคล้องระหว่างศักยภาพ ความสนใจของนักศึกษาและสาขาวิชาที่เลือกเข้าศึกษา ซึ่งเป็นสาเหตุเชิงโครงสร้างที่ยังไม่ได้รับการแก้ไขอย่างเป็นระบบ

ในสาขาวิชาวิทยาการคอมพิวเตอร์ ปัญหานี้มีความรุนแรงมากยิ่งขึ้น เนื่องจากสาขาวิชาดังกล่าวต้องอาศัยความรู้พื้นฐานด้านคณิตศาสตร์ ตรรกะ และการคิดเชิงวิเคราะห์ หากนักศึกษาขาดความถนัดหรือพื้นฐานที่เหมาะสม จะส่งผลให้ไม่สามารถทำความเข้าใจเนื้อหาได้อย่างมีประสิทธิภาพ เกิดความเครียด ขาดแรงจูงใจในการเรียน ขาดเรียน และไม่สามารถสอบผ่านรายวิชาหลักได้ ซึ่งท้ายที่สุดนำไปสู่การออกจากการศึกษากลางคัน ปัญหานี้เกิดขึ้นซ้ำอย่างต่อเนื่องในทุกปีการศึกษา และสร้างภาระทั้งต่อนักศึกษา คณาจารย์ และสถาบัน

ขณะเดียวกัน จำนวนผู้สมัครเข้าศึกษาต่อในระดับปริญญาตรีมีแนวโน้มเพิ่มขึ้นอย่างต่อเนื่อง ในขณะที่จำนวนที่นั่งรับเข้าศึกษามีจำกัด ทำให้กระบวนการคัดเลือกนักศึกษามีความสำคัญอย่างยิ่ง อย่างไรก็ตาม การพิจารณาคัดเลือกนักศึกษาในปัจจุบันต้องอาศัยข้อมูลจำนวนมากและมีความหลากหลาย เช่น ผลการเรียน คะแนนสอบ แผนการเรียน ความสนใจ ความถนัด และภูมิหลังทางการศึกษา ซึ่งการประเมินข้อมูลเหล่านี้โดยอาศัยดุลยพินิจของคณาจารย์เพียงอย่างเดียว อาจไม่สามารถสะท้อนศักยภาพที่แท้จริงของผู้สมัครได้อย่างครบถ้วน และมีความเสี่ยงต่อความคลาดเคลื่อนในการตัดสินใจ

สำหรับสาขาวิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี แม้ว่าจะมีผู้สมัครเข้าศึกษาต่อเป็นจำนวนมากในแต่ละปี แต่จากสถิติที่ผ่านมา พบว่ายังคงมีนักศึกษาจำนวนไม่น้อยที่ไม่สามารถศึกษาต่อจนสำเร็จการศึกษาได้ แสดงให้เห็นถึงช่องว่างที่สำคัญในกระบวนการคัดเลือกนักศึกษา กล่าวคือ ยังขาดเครื่องมือเชิงระบบที่สามารถนำข้อมูลจำนวนมากมาวิเคราะห์เพื่อสนับสนุนการตัดสินใจได้อย่างมีประสิทธิภาพและเป็นรูปธรรม

ดังนั้น งานวิจัยนี้จึงมุ่งพัฒนาระบบช่วยการตัดสินใจในการรับนักเรียนเข้าศึกษาต่อในสาขาวิชาวิทยาการคอมพิวเตอร์ มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี เพื่อแก้ไขปัญหาการคัดเลือกนักศึกษาที่ไม่สอดคล้องกับศักยภาพและความสนใจของผู้สมัคร ระบบดังกล่าวมีเป้าหมายเพื่อลดอัตราการออกจากการศึกษากลางคัน เพิ่ม

โอกาสความสำเร็จทางการศึกษาของนักศึกษา และช่วยให้คณาจารย์สามารถตัดสินใจได้บนพื้นฐานของข้อมูลและการวิเคราะห์อย่างเป็นระบบมากยิ่งขึ้น

## 2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

### 2.1 เหมืองข้อมูล (Data Mining)

เหมืองข้อมูล (Data Mining) [1] คือกระบวนการที่กระทำกับข้อมูลขนาดใหญ่จำนวนมากเพื่อค้นหารูปแบบและความสัมพันธ์ที่ซ่อนอยู่ในชุดข้อมูลนั้นในปัจจุบันการทำเหมืองข้อมูลได้ถูกนำไปประยุกต์ใช้ในงานหลายประเภททั้งในด้านธุรกิจที่ช่วยในการตัดสินใจของผู้บริหารและมองหากลุ่มลูกค้าหรือความชอบความสัมพันธ์ที่ซ่อนอยู่เพื่อมองหาช่องทางการทำรายได้ใหม่ๆในด้านวิทยาศาสตร์และการแพทย์ในการทำวินัจฉัยโรคต่างๆรวมทั้งในด้านเศรษฐกิจและสังคม

การทำเหมืองข้อมูลเป็นเหมือนการพัฒนาการอย่างมีประโยชน์จากการจัดเก็บข้อมูลอย่างง่าย มาสู่การเก็บข้อมูลในฐานข้อมูลที่สามารถดึงข้อมูลสารสนเทศมาใช้จนการทำเหมืองข้อมูลที่สามารถค้นพบความรู้ที่ซ่อนอยู่ในข้อมูล

ขั้นตอนการทำเหมืองข้อมูล

2.1.1 การทำความสะอาดข้อมูล (Data Cleaning) เป็นขั้นตอนสำหรับการคัดข้อมูลที่ไม่เกี่ยวข้องออกไป

2.1.2 การรวมข้อมูล (Data Integration) เป็นขั้นตอนการรวมข้อมูลที่มีที่มามาจากหลายๆแหล่งทำให้ข้อมูลเหล่านั้นรวมเป็นข้อมูลชุดเดียวกัน

2.1.3 การเลือกข้อมูล (Data Selection) เป็นขั้นตอนการดึงข้อมูลที่ต้องการสำหรับการวิเคราะห์จากแหล่งที่บันทึกไว้

2.1.4 การแปลงข้อมูล (Data Transformation) เป็นขั้นตอนการแปลงข้อมูลให้เหมาะสมสำหรับการใช้งานอย่างถูกต้องและกรอบข้อมูลส่วนที่ไม่จำเป็นทิ้ง

2.1.5 การทำเหมืองข้อมูล (Data Mining) เป็นขั้นตอนการค้นหารูปแบบที่ไม่เคยเห็นอยู่ในชุดข้อมูลซึ่งเป็นประโยชน์

2.1.6 การประเมินรูปแบบ (Pattern Evaluation) เป็นขั้นตอนการประเมินรูปแบบที่ได้จากการทำเหมืองข้อมูล

2.1.7 การนำเสนอความรู้ (Knowledge Representation) เป็นขั้นตอนการนำเสนอความรู้ที่ค้นพบ โดยใช้เทคนิคในการนำเสนอเพื่อให้เข้าใจ

### 2.2 กฎความสัมพันธ์ (Association Rules)

กฎความสัมพันธ์(Association Rules) [1] คือ Machine Learning ชนิดหนึ่ง Association rule มักถูกใช้ในการหาความสัมพันธ์ของข้อมูล เช่น การหาสิ่งที่ลูกค้าซื้อพร้อมกันเวลาซื้อของหนึ่งครั้ง เพื่อนำข้อมูลเหล่านั้นมาทำโปรโมชั่น หรือ เพื่อช่วยในการจัดวางหน้าร้านค้า โดยจัดวางสินค้าที่ซื้อร่วมกันบ่อยๆให้อยู่ใกล้กันหรือในอีกมุมหนึ่ง คือจัดวางสินค้าห่างกันพอประมาณเพื่อให้ลูกค้าเดินผ่านสินค้าอื่นที่เราต้องการจะโปรโมทในหาความสัมพันธ์ (Association)

### 2.3 งานวิจัยที่เกี่ยวข้อง

จากการทบทวนวรรณกรรมที่เกี่ยวข้อง พบว่ามีงานวิจัยจำนวนมากที่นำเทคนิคการทำเหมืองข้อมูลและการเรียนรู้ของเครื่องมาประยุกต์ใช้กับการวิเคราะห์ข้อมูลด้านการศึกษา ทั้งในด้านการคัดเลือกนักศึกษา การวิเคราะห์พฤติกรรมการเลือกสาขาวิชา และการพยากรณ์ผลการเรียนของผู้เรียน

ขันทอง ประชุมชาติ และพิมรินทร์ คีรินทร์ [2] ศึกษาการวิเคราะห์พฤติกรรมการเลือกสมัครสาขาวิชาของนักศึกษาใหม่ โดยใช้ข้อมูลการสมัครเข้าศึกษาต่อของมหาวิทยาลัยราชภัฏพิบูลสงครามเป็นกรณีศึกษา งานวิจัยนี้ดำเนินการตามกรอบการทำเหมืองข้อมูลแบบ CRISP-DM และใช้โปรแกรมเวก้า โดยประยุกต์ใช้เทคนิคการจัดกลุ่มด้วยขั้นตอนวิธีเคมีน (K-means) เพื่อจัดกลุ่มผู้สมัครและใช้เทคนิคอพริโอริ (Apriori) เพื่อสร้างกฎความสัมพันธ์ของพฤติกรรมการเลือกสาขาวิชา ผลการวิจัยสามารถนำไปใช้สนับสนุนการวางแผนการรับนักศึกษาและการประชาสัมพันธ์ได้อย่างเหมาะสม

อนันต์ ปินะเต [3] นำเสนอการใช้เทคนิคการค้นหากฎความสัมพันธ์เพื่อพัฒนาระบบประชาสัมพันธ์หลักสูตรออนไลน์ของมหาวิทยาลัยมหาสารคาม โดยใช้ข้อมูลสารสนเทศการรายงานตัวของนิสิตใหม่เป็นข้อมูลนำเข้า ผลการวิเคราะห์พบว่าปัจจัยที่มีผลต่อการตัดสินใจเลือก

ศึกษาต่อมากที่สุด ได้แก่ รายได้รวมของบิดามารดา อาชีพของบิดาและมารดา และสาเหตุในการเลือกเข้าศึกษาในมหาวิทยาลัย งานวิจัยนี้แสดงให้เห็นศักยภาพของการนำรูปแบบที่ค้นพบจากข้อมูลมาประยุกต์ใช้ในระบบสนับสนุนการตัดสินใจด้านการศึกษา

เสกสรรค์ วิลัยลักษณ์ วิภา เจริญภัณฑารักษ์ และดวงดาว วิชาดากุล [4] พัฒนาคลังข้อมูลและสร้างตัวแบบพยากรณ์ผลการเรียนของนักเรียนระดับมัธยมศึกษา โดยใช้ข้อมูลนักเรียนจำนวน 525 ระเบียน และคุณลักษณะจำนวน 16 รายการ งานวิจัยได้นำเทคนิคการคัดเลือกคุณลักษณะ ได้แก่ Correlation-based Feature Selection และ Information Gain ร่วมกับเทคนิคโครงข่ายประสาทเทียมแบบมัลติเลเยอร์ ผลการศึกษาชี้ให้เห็นว่าการเตรียมข้อมูลและการคัดเลือกคุณลักษณะที่เหมาะสมมีผลต่อประสิทธิภาพของแบบจำลองพยากรณ์อย่างมีนัยสำคัญ

สุธีรา วงศ์อนันทรัพย์ ต้องใจ แย้มผกา และอรวรรณ มูสิกะ [5] วิเคราะห์ผลการทดสอบประเมินความรู้ก่อนเรียนและแบบทดสอบวัดความถนัดของนักศึกษา โดยใช้เทคนิคเหมืองข้อมูลเพื่อจำแนกและจัดกลุ่มนักศึกษา ผลการวิจัยสามารถจำแนกนักศึกษาออกเป็นกลุ่มนักศึกษาที่มีศักยภาพแตกต่างกัน และสามารถสร้างกฎความสัมพันธ์เพื่อใช้สนับสนุนการดูแลนักศึกษาได้อย่างตรงกลุ่มเป้าหมาย ช่วยลดความเสี่ยงของนักศึกษาที่มีผลการเรียนต่ำกว่ามาตรฐาน

จากการทบทวนงานวิจัยที่เกี่ยวข้องดังกล่าว พบว่างานวิจัยส่วนใหญ่ให้ความสำคัญกับการประยุกต์ใช้เทคนิคเหมืองข้อมูลเพื่อการวิเคราะห์ พยากรณ์ และสนับสนุนการตัดสินใจในหลากหลายบริบท อย่างไรก็ตาม ยังมีงานวิจัยจำนวนจำกัดที่มุ่งเน้นการพัฒนาระบบช่วยการตัดสินใจในการคัดเลือกนักศึกษาเข้าศึกษาต่อโดยตรง เพื่อลดอัตราการออกจากการศึกษากลางคันในระดับอุดมศึกษา ซึ่งเป็นช่องว่างสำคัญที่งานวิจัยนี้มุ่งเน้นศึกษาและพัฒนาเพิ่มเติม

## 3. วิธีการดำเนินงานวิจัย

โครงงานวิจัยเรื่อง ระบบช่วยการตัดสินใจในการรับนักเรียนเข้าศึกษาต่อใน หลักสูตรวิทยาการคอมพิวเตอร์ มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี เป็นการสร้างแบบจำลองการพยากรณ์ ด้วยวิธีการทำเหมืองข้อมูลและพัฒนาต่อยอดให้อยู่ในรูปแบบเว็บแอปพลิเคชัน การเลือกแนวทางในการดำเนินงานและวิธีการใช้เครื่องมือหรือโปรแกรมต่าง ๆ ผู้วิจัยมีขั้นตอนในการ ดำเนินงาน ประกอบไปด้วย แผนการดำเนินงาน เครื่องมือที่ใช้ในงานวิจัย ขั้นตอนการดำเนินงาน และวิธีการทดสอบ

### 3.1 ขั้นตอนการดำเนินงาน

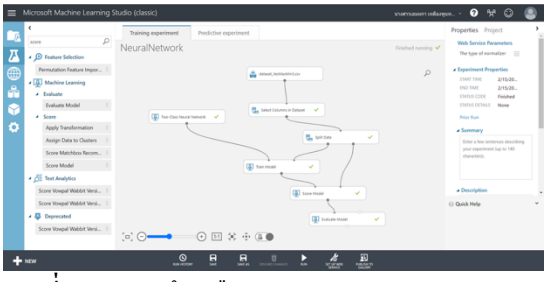
ขั้นตอนการดำเนินงานทางผู้วิจัยได้โดยเทคนิคการทำเหมืองข้อมูล แล้วนำข้อมูลทั้งหมดมาวิเคราะห์และออกแบบเป็นเว็บแอปพลิเคชัน ดังนี้

3.1.1 การเตรียมข้อมูล

ทำการสำรวจข้อมูลนักเรียนที่มาสมัครเข้าศึกษาต่อในหลักสูตรวิทยาการคอมพิวเตอร์ มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี จำนวน 387 คน นำข้อมูลมาเก็บไว้ใน Microsoft Excel และนำข้อมูลมาทำความสะอาดข้อมูล (Data cleaning) โดยจะมีกระบวนการตรวจสอบ การแก้ไขหรือการลบ เพื่อให้รายการข้อมูลที่ไม่ถูกต้องออกไปจากชุดข้อมูล ใช้ข้อมูลลำดับที่เลือกสาขาวิทยาการคอม เกรดรวม ความถนัดทางคณิตศาสตร์ ความถนัดทางวิทยาศาสตร์ ผลรวมหน่วยกิต โอเน็ตภาษาไทย โอเน็ตสังคมศึกษา โอเน็ตภาษาอังกฤษ โอเน็ตคณิตศาสตร์ โอเน็ตวิทยาศาสตร์ คะแนนการทดสอบวิชาความถนัดทั่วไปภาษาไทย คะแนนการทดสอบวิชาความถนัดทั่วไปภาษาอังกฤษ

3.1.2 จัดทำโมเดล

อัปโหลดไฟล์ชุดข้อมูลที่เตรียมไว้ไปที่ Microsoft Azure Machine Learning Studio [6] แบ่งชุดข้อมูลออกเป็นสองชุด คือ ชุดทดสอบ (Test) และชุดอบรมข้อมูล (Train) ทำการ Train Model เพื่อให้ได้โมเดล ดังภาพที่ 1



**ภาพที่ 1:** อัลกอริธึมเครือข่ายประสาท (Neural Network)

3.1.3 Score Model

เป็นการคาดคะเนคะแนนสำหรับแบบจำลองการจำแนกที่ผ่านการอบรมข้อมูล และ Evaluate Model ประเมินผลลัพธ์ของการจัดหมวดหมู่ด้วยตัวชี้วัดมาตรฐาน

3.1.4 Predictive Web Service

โมเดลจะแปลงร่างไป เพื่อนำไปเชื่อม API กับเว็บไซต์ช่วยการตัดสินใจในการรับนักเรียนเข้าศึกษาต่อในหลักสูตรวิทยาการคอมพิวเตอร์ มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี

3.1.5 การพัฒนาเว็บแอปพลิเคชัน

หน้าหลัก เป็นหน้าโดยรวมของเว็บไซต์ที่สามารถเข้าถึงได้ทุกเมนูของเว็บไซต์ที่ผู้ใช้มีความสนใจ โดยประกอบไปด้วย หน้าแรก หน้าเกี่ยวกับ หน้าผู้ดูแลระบบ หน้าแบบประเมินสาธารณะ หน้าผลลัพธ์แบบประเมินสาธารณะ ส่วนนำเข้าไฟล์ Excel ส่วนสำหรับอาจารย์แบบประเมินภายใน หน้าผลลัพธ์แบบประเมินสาธารณะ

### 3.2 วิธีการทดสอบ

วิธีการทดสอบ การประเมินเปรียบเทียบประสิทธิภาพโมเดลของระบบช่วยการตัดสินใจในการรับนักเรียนเข้าศึกษาต่อในหลักสูตรวิทยาการคอมพิวเตอร์ มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี โดยการเปรียบเทียบด้วย 4 เทคนิค ได้แก่ Two-Class Neural Network, Two-Class Support Vector Machine, Two-Class Bayes Point Machine, Two-Class Averaged Perceptron ด้วยโปรแกรม Microsoft Azure Machine Learning Studio และแบบสอบถามความพึงพอใจของผู้ใช้ที่มีต่อเว็ปแอปพลิเคชัน โดยเทียบกับเกณฑ์ที่กำหนดไว้ หลังจากที่ได้ผลจากการประเมินและแบบสอบถามแล้วนำมาคำนวณหาค่าเฉลี่ยและค่าส่วนเบี่ยงเบนมาตรฐาน เพื่อประเมินผลและวัดระดับความพึงพอใจของผู้ใช้

## 4. ผลการดำเนินงานและการวิเคราะห์

### 4.1 การประเมินเปรียบเทียบประสิทธิภาพโมเดล

โดยการเปรียบเทียบด้วย 4 เทคนิค ได้แก่ Two-Class Neural Network, Two-Class Support Vector Machine, Two-Class Bayes Point Machine, Two-Class Averaged Perceptron ด้วยMicrosoft Azure Machine Learning Studio ดังตารางที่ 1

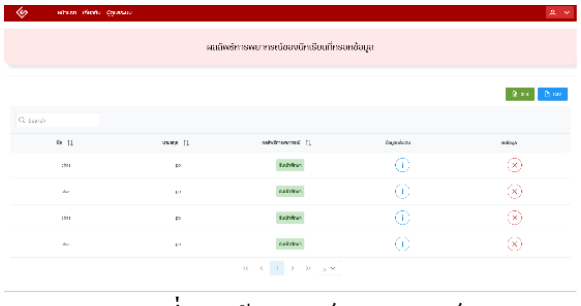
**ตารางที่ 1:** ตารางการประเมินเปรียบเทียบประสิทธิภาพโมเดล

| เทคนิค | Accuracy | Precision | Recall | F1 Score |
|---|---|---|---|---|
| **Two-Class Neural Network** | **0.899** | **0.795** | **0.946** | **0.864** |
| Two-Class Support Vector Machine | 0.771 | 0.700 | 0.568 | 0.627 |
| Two-Class Bayes Point Machine | 0.752 | 0.656 | 0.568 | 0.609 |
| Two-Class Averaged Perceptron | 0.780 | 0.686 | 0.649 | 0.667 |

ผู้วิจัยจึงเลือกใช้เทคนิค Two-Class Neural Network เนื่องจากมีค่าความแม่นยำมากที่สุด คือมีค่าความแม่นยำโดยรวมเฉลี่ย 89.90%

### 4.2 พัฒนาเว็บแอปพลิเคชัน

นำโมเดลที่ได้นำมาพัฒนาร่วมกับเว็บแอปพลิเคชันเพื่อแสดงผลการพยากรณ์ดังภาพที่ 2



**ภาพที่ 2:** หน้าผลลัพธ์การพยากรณ์

**4.3 ประเมินความพึงพอใจการใช้งาน**

ประเมินความพึงพอใจการใช้งานด้วยแบบสอบถาม โดย ประเมินด้วยผู้เชี่ยวชาญ 3 คน ส่วนของผู้เชี่ยวชาญซึ่งพบว่า เมื่อพิจารณาจากเกณฑ์การประเมิน มีค่าเฉลี่ยโดยรวมเท่ากับ 4.38 มีความพึงพอใจมาก และส่วนเบี่ยงเบนมาตรฐานโดยรวมเท่ากับ 0.14 พบว่ามีความแตกต่างกันน้อย

## 5. สรุปและข้อเสนอแนะ

งานวิจัยนี้มีวัตถุประสงค์เพื่อศึกษาและพัฒนาระบบช่วยการตัดสินใจในการรับนักเรียนเข้าศึกษาต่อในหลักสูตรวิทยาการคอมพิวเตอร์ มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี โดยใช้เทคนิคการทำเหมืองข้อมูลจากชุดข้อมูลผู้สมัครจำนวนทั้งสิ้น 387 รายการ ผู้วิจัยได้เปรียบเทียบประสิทธิภาพของเทคนิคเหมืองข้อมูลจำนวน 4 เทคนิค ได้แก่ Two-Class Neural Network, Two-Class Support Vector Machine, Two-Class Bayes Point Machine และ Two-Class Averaged Perceptron

ผลการทดลองพบว่าเทคนิค Two-Class Neural Network ให้ค่าความแม่นยำโดยรวมเฉลี่ยสูงที่สุด เท่ากับ 89.90% จึงถูกเลือกนำมาใช้เป็นโมเดลหลักในการพัฒนาระบบต้นแบบในรูปแบบเว็บแอปพลิเคชัน เพื่อช่วยสนับสนุนการตัดสินใจของอาจารย์สาขาวิทยาการคอมพิวเตอร์ในการคัดกรองนักเรียนทั่วไปที่มีความสนใจเข้าศึกษาต่อในหลักสูตรวิทยาการคอมพิวเตอร์ มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี

จากนั้นได้ทำการประเมินความพึงพอใจในการใช้งานระบบโดยผู้เชี่ยวชาญจำนวน 3 คน โดยใช้แบบสอบถามเป็นเครื่องมือในการประเมิน ผลการประเมินพบว่าระบบมีค่าเฉลี่ยความพึงพอใจเท่ากับ 4.38 จากคะแนนเต็ม 5 ซึ่งอยู่ในระดับมาก และมีค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.14 แสดงให้เห็นว่าผู้เชี่ยวชาญมีความคิดเห็นไปในทิศทางเดียวกัน และระบบที่พัฒนาขึ้นมีความเหมาะสมต่อการนำไปใช้งานจริง

## เอกสารอ้างอิง


[1] ศุภชัย จันทร์หอม. (2556). การทำเหมืองข้อมูล (Data Mining). กรุงเทพฯ: ซีเอ็ดยูเคชั่น

[2] ขันทอง ประชุมชาติ, และพิมรินทร์ คีรินทร์. (2558). การวิเคราะห์พฤติกรรมการเลือกสมัครสาขาวิชาของนักศึกษาใหม่โดยใช้เทคนิคการทำเหมืองข้อมูล. วารสารวิจัยและพัฒนา, 10(2), 45–56.

[3] อนันต์ ปินะเต. (2562). การใช้เทคนิคการค้นหากฎความสัมพันธ์เพื่อพัฒนาระบบประชาสัมพันธ์หลักสูตรออนไลน์. วารสารเทคโนโลยีสารสนเทศ, 15(1), 23–34.

[4] เสกสรรค์ วิลัยลักษณ์, วิภา เจริญภัณฑารักษ์, และดวงดาว วิชาดากุล. (2558). การพัฒนาตัวแบบพยากรณ์ผลการเรียนของนักเรียนโดยใช้เทคนิคเหมืองข้อมูล. วารสารวิทยาการคอมพิวเตอร์, 12(1), 1–12.

[5] สุธีรา วงศ์อนันทรัพย์, ต้องใจ แย้มผกา, และอรวรรณ มูสิกะ. (2559). การประยุกต์ใช้เทคนิคเหมืองข้อมูลในการจำแนกและจัดกลุ่มนักศึกษา. วารสารการศึกษาและเทคโนโลยี, 8(2), 89–101.

[6] Microsoft. (2023). What is Azure Machine Learning? Microsoft Learn. https://learn.microsoft.com/en-us/azure/machine-learning/overview-what-is-azure-machine-learning

# การพัฒนาระบบตรวจจับความผิดปกติและแสดงผลการใช้งานระบบสารสนเทศโรงพยาบาลแบบเรียลไทม์

# Development of a Real-time Anomaly Detection and Visualization System for Hospital Information Systems

*เบญจพล ทองเจือ (Benchapol Thongchua)[1] และนวพร วิสิฐพงศ์พันธ์ (Nawaporn Wisitpongphan)[2]*
*ภาควิชาการบริหารเครือข่ายดิจิทัลและความมั่นคงปลอดภัยสารสนเทศ*
*คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ*
*S6707031857212@kmutnb.ac.th, nawaporn.w@itd.kmutnb.ac.th*

## บทคัดย่อ

*งานวิจัยนี้มุ่งพัฒนาระบบตรวจวัดประสิทธิภาพการใช้งานระบบสารสนเทศโรงพยาบาล การนำ Log ที่มีปริมาณมากมาใช้ให้เกิดประโยชน์ ผู้วิจัยได้นำเสนอสถาปัตยกรรมรวมศูนย์ข้อมูลโดยประยุกต์ใช้เครื่องมือ Open Source เพื่อรวบรวมและวิเคราะห์ข้อมูลจากระบบย่อยทั้งหมด 17 ระบบ ในการแสดงผลแบบเรียลไทม์ โดยใช้ค่าทางสถิติประกอบด้วยค่ามัธยฐาน (P50) และ เปอร์เซ็นไทล์ที่ 95 (P95) เพื่อสร้างเกณฑ์วัดประสิทธิภาพแบบ Dynamic Service Level Objective เพื่อให้สะท้อนประสบการณ์ผู้ใช้งานจริงได้แม่นยำกว่าการใช้ Static threshold จากผลการวิจัยพบว่าระบบสามารถระบุช่วงเวลาที่มีการใช้งานหนาแน่นที่สุดและตรวจพบเหตุการณ์ผิดปกติสำคัญในช่วงเวลาทำการ นอกจากนี้การใช้เกณฑ์ P95 ยังช่วยให้ตรวจพบความล่าช้า ในช่วงที่อุปกรณ์เครือข่ายขัดข้องได้อย่างรวดเร็ว ส่งผลให้เจ้าหน้าที่สามารถตอบสนองต่อเหตุการณ์ได้ทันที ทำให้ลดระยะเวลาในการระบุสาเหตุได้อย่างมีประสิทธิภาพเมื่อเทียบกับการอ่าน Raw Log แบบเดิม ๆ งานวิจัยนี้เน้นไปที่การตรวจจับความผิดปกติและปริมาณการใช้งานของระบบสารสนเทศในโรงพยาบาลเพื่อยกระดับการบริหารจัดการทรัพยากรสารสนเทศเชิงรุก ลดความเสี่ยงด้านความต่อเนื่องในการให้บริการ*

**คำสำคัญ** : ระบบสารสนเทศโรงพยาบาล, แดชบอร์ดแบบเรียลไทม์, การวิเคราะห์บันทึกเหตุการณ์,แดชบอร์ด

## Abstract

*This research aimed to develop a usage monitoring system for hospital information systems, with the objective of transforming large volumes of dispersed log data into actionable insights. The researcher proposed a centralized data architecture leveraging open-source tools to aggregate and analyze data across 17 subsystems, enabling real-time visualization. The system employs median (P50) and 95th percentile (P95) statistical measures to establish Dynamic Service Level Objectives (Dynamic SLOs), thereby more accurately reflecting genuine user experience compared to conventional static threshold approaches. The findings demonstrated that the system was capable of identifying peak usage periods and detecting critical anomalies during operational hours. Furthermore, the application of P95-based thresholds enabled the rapid detection of latency spikes during network device failures, allowing IT personnel to respond to incidents promptly. This significantly reduced the mean time to identify root causes when compared to traditional manual raw log analysis.*

**Keywords:** Hospital Information System, Real-time Dashboard, Log Analytics, OpenSearch, Data Visualization.

## 1. บทนำ

ปัจจุบันการใช้เทคโนโลยีสารสนเทศมีบทบาทสำคัญในระบบสาธารณสุข โรงพยาบาลได้นำระบบสารสนเทศโรงพยาบาล (HIS) มาเก็บรวบรวมข้อมูลทางการรักษา ระบบย่อยภายใน HIS ที่สำคัญประกอบด้วย ระบบเวชระเบียนอิเล็กทรอนิกส์ (EMR) ระบบห้องปฏิบัติการ (LIS) ระบบภาพรังสี (RIS),ระบบผู้ป่วยนอก (OPD) และระบบผู้ป่วยใน (IPD) เป็นต้น [1,2,3] อย่างไรก็ตาม ปัญหาสำคัญที่พบในหลายโรงพยาบาลคือการใช้ประโยชน์จากข้อมูล (Data Utilization) เนื่องจากข้อมูลจากระบบย่อยจำนวนมาก แม้จะถูกจัดเก็บในระบบแต่ยังไม่มีการนำไปใช้ประโยชน์ สาเหตุมาจากข้อมูลมีการกระจายตัว ซับซ้อน และขาดเครื่องมือในการวิเคราะห์แบบเรียลไทม์ ทำให้ผู้บริหารและเจ้าหน้าที่ฝ่ายเทคโนโลยีสารสนเทศไม่สามารถมองเห็นการใช้งานของระบบย่อยได้ชัดเจน [4,5,6]

ดังนั้น งานวิจัยนี้จึงมุ่งเน้นการนำเครื่องมือ Open Source ที่มีประสิทธิภาพในการสืบค้นและวิเคราะห์ข้อมูลแบบเรียลไทม์ มาประยุกต์ใช้ในการติดตามพฤติกรรมการใช้งานระบบย่อยต่าง ๆ ภายในโรงพยาบาลขนาดใหญ่จำนวน 1,500 เตียง เพื่อแสดงให้เห็นถึงปริมาณข้อมูลที่ระบบต้องรองรับ โดยมีวัตถุประสงค์เพื่อวิเคราะห์พฤติกรรมการใช้งานของระบบสารสนเทศแต่ละประเภท และตรวจจับความผิดปกติต่าง ๆ จากการใช้ค่าทางสถิติผ่านแดชบอร์ดแบบเรียลไทม์ เพื่อให้เจ้าหน้าที่สามารถติดตามการใช้งานที่ผิดปกติ รวมถึงสามารถช่วยสนับสนุนการตัดสินใจของผู้บริหาร โดยระบบที่นำเสนอนี้จะใช้ตัวชี้วัดทางสถิติที่ได้จากข้อมูลจริง ได้แก่ ค่ามัธยฐาน (P50), ค่าเปอร์เซ็นไทล์ 95 (P95), ค่า Exponentially Weighted Moving Average (EWMA) และการกำหนดค่า Service Level Objective (SLO) เพื่อช่วยให้สามารถตรวจจับความผิดปกติและวิเคราะห์ประสิทธิภาพของระบบได้ [13,14]

## 2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

### 2.1 การจัดการข้อมูลด้วย Logstash Pipeline และการวิเคราะห์ข้อมูลด้วย OpenSearch Dashboard

การเลือกใช้เครื่องมือ Open Source อย่าง OpenSearch และ Logstash ทำหน้าที่เป็น Pipeline ในการรวบรวมและประมวลผลข้อมูล Log จากอุปกรณ์เครือข่ายผ่านกระบวนการ Input, Filter และ Output [5,6] ข้อมูลที่ผ่านการจัดรูปแบบแล้วจะถูกส่งต่อไปยัง OpenSearch ที่รองรับการสืบค้นที่ซับซ้อนและการวิเคราะห์แบบ Real-time และแสดงผลผ่าน OpenSearch Dashboards ในรูปแบบของการแสดงผล ที่จำแนกตามอาคารหรือช่วงเวลาได้อย่างแม่นยำ [8,9,10,11]

### 2.3 โครงสร้างข้อมูล Log Firewall

ข้อมูลหลักคือ ข้อมูล Log จาก Firewall โดยเน้นฟิลด์ข้อมูล Source/Destination IP, Service Port และ Event Duration ข้อมูลเหล่านี้ถูกนำมาคำนวณร่วมกับตัวชี้วัดทางสถิติ ได้แก่ ค่ามัธยฐาน (P50) เพื่อสะท้อนประสบการณ์ผู้ใช้ส่วนใหญ่ และเปอร์เซ็นไทล์ที่ 95 (P95) เพื่อตรวจวัดความล่าช้าในกลุ่ม Tail Latency โดยประยุกต์ใช้ร่วมกับโมเดล Exponentially Weighted Moving Average (EWMA) เพื่อสร้างเส้นฐาน (Baseline) ที่มีความเสถียรลดความผันผวนของข้อมูล [5,9]

### 2.4 งานวิจัยที่เกี่ยวข้อง

จากการทบทวนวรรณกรรมที่เกี่ยวข้องกับการพัฒนาระบบติดตามและแสดงผลข้อมูลสารสนเทศในโรงพยาบาล พบว่างานวิจัยส่วนใหญ่มีเป้าหมายเพื่อสนับสนุนการตัดสินใจ เพิ่มประสิทธิภาพการบริหารจัดการ และยกระดับคุณภาพการให้บริการด้านสาธารณสุข ผู้วิจัยจึงได้จัดกลุ่มงานวิจัยที่เกี่ยวข้องออกเป็น 3 ประเด็นหลัก ได้แก่ (1) การประยุกต์ใช้ Dashboard เพื่อการบูรณาการและการแสดงผลข้อมูล (2) เทคนิคการตรวจจับความผิดปกติจากการใช้งานระบบ และ (3) การเปลี่ยนผ่านจาก ELK Stack สู่ OpenSearch ดังรายละเอียดต่อไปนี้

**1) การประยุกต์ใช้ Dashboard เพื่อการบูรณาการและการแสดงผลข้อมูล**

โรงพยาบาลส่วนใหญ่ประสบปัญหาข้อมูลกระจัดกระจายจากการใช้งาน ส่งผลให้ไม่สามารถติดตามสถานะการใช้งานจริงของระบบสารสนเทศได้ งานวิจัยก่อนหน้าได้นำเครื่องมือในกลุ่ม ELK Stack และเพื่อทำ Data Visualization ในการแสดงผลข้อมูลในรูปแบบที่เข้าใจง่าย เพื่อลดภาระและช่วยให้บุคลากรสามารถรับรู้สถานะผิดปกติของระบบได้อย่างรวดเร็ว [7,8,9,11]

**2) เทคนิคการตรวจจับความผิดปกติจากการใช้งาน**

การประเมินประสิทธิภาพของระบบสารสนเทศจำเป็นต้องใช้ตัวชี้วัดที่มีความละเอียดมากกว่าค่าเฉลี่ยทั่วไป งานวิจัยหลายฉบับได้เสนอการใช้ค่าเปอร์เซ็นไทล์ที่ 50 และ เปอร์เซ็นไทล์ที่ 95 เพื่อสะท้อนพฤติกรรมการใช้งานของผู้ใช้ส่วนใหญ่และกลุ่มผู้ใช้ที่เกิดความล่าช้ามากที่สุด เพื่อระบุพฤติกรรมการใช้งานที่ผิดปกติในระบบ การใช้ตัวชี้วัดดังกล่าวช่วยเพิ่มความสามารถในการตรวจจับปัญหาเชิงประสิทธิภาพของระบบได้อย่างแม่นยำ [3,5,6,9,13]

**3) จาก ELK Stack สู่ OpenSearch**

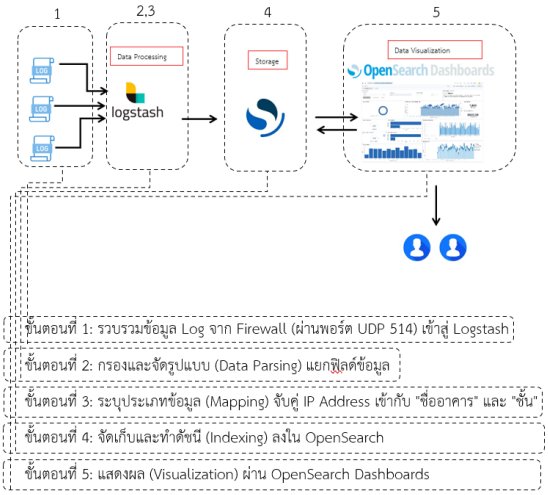
การเปลี่ยนแปลงนโยบายด้านลิขสิทธิ์ของ Elastic [12] ส่งผลให้ OpenSearch ได้รับความสนใจเพิ่มขึ้นเพื่อเป็นทางเลือกทดแทน เนื่องจากมีความสามารถเทียบเท่าในการจัดการ Log การวิเคราะห์ข้อมูล และการสร้าง Visualization โดยไม่มีค่าใช้จ่ายด้านลิขสิทธิ์ [2,10,11] อย่างไรก็ตาม ยังมีงานวิจัยจำนวนน้อยที่นำ OpenSearch มาใช้วิเคราะห์การใช้งานระบบสารสนเทศโรงพยาบาลในเชิงประสิทธิภาพการให้บริการและการตรวจจับความผิดปกติแบบบูรณาการ งานวิจัยนี้จึงมุ่งเน้นการเติมเต็มช่องว่างดังกล่าว

## 3. วิธีการดำเนินงาน

งานวิจัยนี้นำเสนอระบบติดตามและแสดงผลพฤติกรรมการใช้งานระบบสารสนเทศโรงพยาบาล (HIS) แบบเรียลไทม์ โดยประยุกต์ใช้เทคโนโลยี OpenSearch ร่วมกับ Logstash เพื่อรวบรวมและวิเคราะห์ข้อมูล Log จากอุปกรณ์เครือข่าย โดยมีรายละเอียดการดำเนินการดังนี้

**3.1 สถาปัตยกรรมและองค์ประกอบของระบบ**

สถาปัตยกรรมดังแสดงในภาพที่ 1 แบ่งออกเป็น 2 ส่วนหลัก คือ (1) Hardware ที่ใช้ในการรับส่งข้อมูล Log จากระบบ HIS ผ่านอุปกรณ์ Firewall Check Point และ Load Balance ใช้ VM Server ลง Ubuntu (2) Software จะติดตั้งผ่าน Docker ประกอบไปด้วย Logstash ทำหน้าที่เป็นตัวกลางรับข้อมูล Syslog (UDP 514) จาก Firewall ส่วน OpenSearch ทำหน้าที่จัดเก็บ Indexing และสืบค้นข้อมูล และ OpenSearch Dashboards ทำหน้าที่แสดงผลข้อมูล




**ภาพที่ 1:** สถาปัตยกรรมของระบบติดตามการใช้งานสารสนเทศในโรงพยาบาล

**3.2 การวิเคราะห์ประสิทธิภาพและกำหนดตัวชี้วัด**

งานวิจัยนี้นำเสนอวิธีการตรวจจับความผิดปกติ (Anomaly Detection) ของการใช้งานระบบด้วยตัวชี้วัดทางสถิติ ดังนี้

1) Percentile Metrics: ใช้เปอร์เซ็นไทล์ของค่าความหน่วงเวลาหรือ Latency ของการใช้งานระบบ โดยค่า P50 (Median) เป็นตัวแทนประสบการณ์ใช้งานของผู้ใช้ส่วนใหญ่ และ P95 เป็นตัวแทนของความล่าช้าที่เกิดขึ้นกับผู้ใช้งานกลุ่มที่ช้าที่สุด โดยคำนวณจากฟิลด์ event.duration ที่ใช้แสดงระยะเวลาในของเหตุการณ์การใช้งานระบบสารสนเทศโรงพยาบาลแต่ละครั้ง ซึ่งสามารถนำมาวิเคราะห์ประสิทธิภาพและความล่าช้า (Latency) ของระบบได้แบบเรียลไทม์

2) Exponentially Weighted Moving Average (EWMA): เพื่อลดความผันผวนของข้อมูลและสร้างเส้นฐาน ผู้วิจัยนำค่า P50 และ P95 มาคำนวณผ่านโมเดล EWMA (กำหนดค่า Window Size = 6, Alpha = 0.3) ทำให้ได้ค่าเฉลี่ยเคลื่อนที่แบบถ่วงน้ำหนัก (ma50 และ ma95)

3) Dynamic SLO (Ratio-based): กำหนดเกณฑ์ระดับการให้บริการ (Service Level Objective) ตามสมการ ดังนี้

$$SLO_{Dynamic} = ma_{50} \times \left(\frac{ma_{95}}{ma_{50}}\right) \quad (1)$$

โดยที่

- $SLO_{Dynamic}$ คือ ค่าระดับการให้บริการที่ปรับเปลี่ยนตามสภาวะการใช้งานจริง
- $ma_{50}$ คือ ค่าเส้นฐานของระบบซึ่งได้จาก Exponentially Weighted Moving Average (EWMA) ของ Percentile 50
- $ma_{95}$ คือ ค่า latency ของผู้ใช้ 5% ที่ช้าที่สุด ซึ่งได้จาก EWMA ของ Percentile 95

สมการที่ (1) แสดงค่าประสิทธิภาพของระบบสารสนเทศ โดยพิจารณาความสัมพันธ์ระหว่างประสบการณ์ของผู้ใช้ส่วนใหญ่และผู้ใช้ที่ได้รับผลกระทบช้าที่สุด

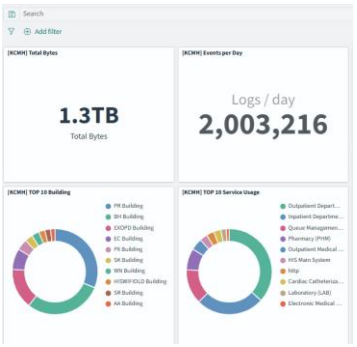
## 4. ผลการดำเนินงาน

### 4.1 การเตรียมข้อมูล Log และการเชื่อมเข้าระบบ

ระบบที่นำเสนอจะรวบรวม Log การใช้งานระบบสารสนเทศโรงพยาบาล (HIS) และระบบย่อยที่เรียกใช้งานผ่าน Firewall ซึ่งมีการระบุระบบย่อยด้วย Service Port และใช้ฟิลด์หลัก ได้แก่ Source IP, Destination IP, Service Port และ Timestamp จากนั้นนำข้อมูลเข้า Logstash เพื่อคัดกรองทราฟฟิกที่เกี่ยวข้อง และปรับรูปแบบวันเวลาให้เป็นมาตรฐานเดียวกัน เพื่อให้เหมาะต่อการทำ Index ใน OpenSearch ในขั้นตอนเตรียมใช้งาน ได้จัดทำ Index Pattern ให้สอดคล้องกับรูปแบบดัชนีแบบรายวัน (ปี-เดือน-วัน) และใช้ฟิลด์ @timestamp เป็นฟิลด์หลักสำหรับการแสดงผลแบบ time-series.

### 4.2 ผลการวิเคราะห์การใช้งานเรียลไทม์จาก Dashboard

เมื่อระบบรับและประมวลผล Log ได้ถูกต้อง แดชบอร์ดจะแสดงข้อมูลสถิติการใช้งานระบบสารสนเทศโรงพยาบาลแบบต่าง ๆ แบบเรียลไทม์และจัดแสดงระบบย่อยที่สำคัญขึ้นมาก่อนเพื่อให้ผู้เกี่ยวข้องติดตามได้ทันที





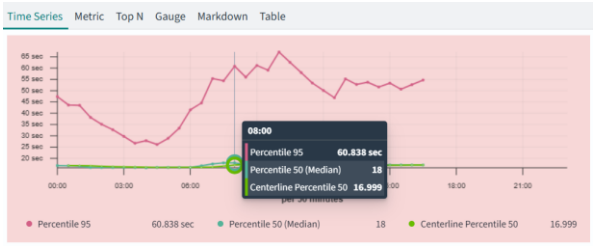


**ภาพที่ 2:** ปริมาณการใช้งานช่วงเวลา 06:00–12:00

จากภาพที่ 2 พบว่าช่วงเวลา 06:00–12:00 เป็นช่วงที่มีการใช้งานมากที่สุด โดยมีทราฟฟิกปริมาณ 1.3 TB และมี Event log จำนวน 2,003,216 รายการ อาคารที่ใช้งานมากที่สุดคือ PR และระบบย่อยที่มีอัตราการใช้งานสูงสุด 5 อันดับ ได้แก่ ระบบผู้ป่วยนอก, ระบบผู้ป่วยใน, ระบบเรียกคิว, ระบบห้องยา และ ระบบเวชระเบียบผู้ป่วยนอก

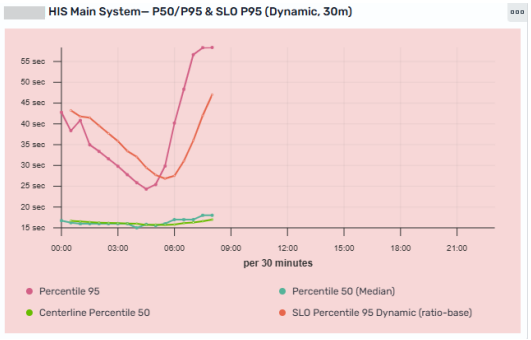
### 4.3 การใช้ตัวชี้วัด P95/P50/EWMA50 และ SLO Dynamic เพื่อประเมินประสิทธิภาพและตรวจจับความผิดปกติ

ตัวชี้วัดทางสถิติในการตรวจจับความผิดปกติของการใช้งานระบบ จะใช้ค่า P95 เพื่อสำหรับสะท้อนกลุ่มการใช้งานส่วนหางที่เกิดความล่าช้าที่สุด ควบคู่กับค่า P50 และเส้น EWMA50 มาใช้ในการคำนวณ โดยกำหนด Window size = 6 และ Alpha = 0.3 เพื่อลดความผันผวนของข้อมูล และใช้เพื่อกำหนดเป็น Baseline ในการคำนวณค่า SLO Dynamic




**ภาพที่ 3:** การใช้งานของหน้า Main HIS จากค่า P95, P50 และ EWMA50

จากภาพที่ 3 พบว่า P50 และ EWMA50 มีค่าใกล้เคียงกัน (~17–18 วินาที) สะท้อนว่าผู้ใช้งานทั่วไปยังอยู่ในเกณฑ์ปกติใกล้เคียงเส้นฐาน

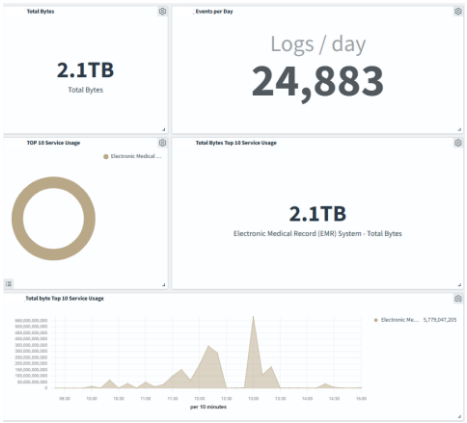



**ภาพที่ 4:** ค่า P95 มากกว่า SLO Dynamic

จากภาพที่ 4 พบว่าเมื่อค่า P95 มากกว่า SLO แสดงให้เห็นถึงปริมาณการใช้งานที่เพิ่มสูงสุดตามช่วงเวลาในบริการของโรงพยาบาล

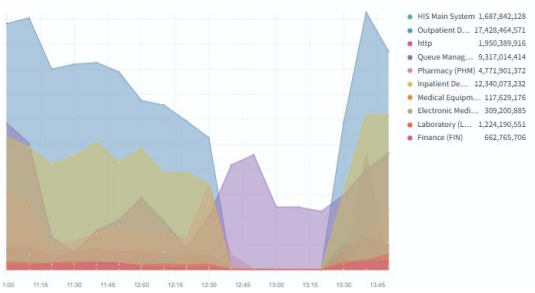
**4.4 การตรวจจับเหตุการณ์ผิดปกติจากการใช้งาน**

ตัวอย่างการใช้แดชบอร์ดเพื่อค้นหา “เหตุการณ์ผิดปกติ” ทั้งแบบปริมาณการใช้งานสูงผิดปกติ และการใช้งานลดลงผิดปกติ




**ภาพที่ 5:** ความผิดปกติจากการใช้งานระบบ EMR

ภาพที่ 5 แสดงการใช้งานของระบบ EMR ที่มีการใช้งานสูง พบการ download ข้อมูล 1.1TB ที่อาคาร CSSD ชั้น 6 จากการตรวจสอบกับผู้ใช้งานที่อาคาร CSSD ชั้น 6 พบว่าช่วงเวลาดังกล่าวได้มีการ Backup ข้อมูลการรักษาผู้ป่วย




**ภาพที่ 6:** ปริมาณการใช้งานที่ลดลงส่งผลให้ระบบทำงานได้ช้าหรือไม่สามารถทำงานได้

ภาพที่ 6 พบว่ากราฟมีการใช้งานที่ลดลงจากปริมาณการใช้งานปกติ ซึ่งมีความผิดปกติจากวันอื่น ๆ จากการตรวจพบกับทีมพัฒนาระบบ พบว่าเกิดจาก ผู้ให้บริการบำรุงรักษาระบบเจ้าหนึ่งได้ทำการขยายขนาดของ disk ทำให้ disk ของ application กับ physical ไม่สัมพันธ์กัน และส่งผลให้ไม่สามารถเข้าใช้งานระบบได้.

**4.5 อภิปรายผลการวิจัย**

ระบบไที่พัฒนาพิสูจน์ให้เห็นว่าสามารถรองรับข้อมูลจำนวนมากและจัดการได้อย่างมีประสิทธิภาพ การใช้ Dynamic SLO เพื่อตรวจจับพฤติกรรมการใช้งานจริงของระบบ อย่างไรก็ตาม เพื่อรองรับปริมาณข้อมูลที่เพิ่มขึ้นในอนาคต โรงพยาบาลควรพิจารณาปรับขยาย OpenSearch เป็นแบบ Multi-node Cluster เพื่อกระจายภาระการประมวลผล

**5. สรุปผลการวิจัย**

งานวิจัยนี้ได้พัฒนาและการประยุกต์ใช้ OpenSearch เพื่อใช้ในการแสดงผลข้อมูลปริมาณการใช้งานของระบบบสารสนเทศในโรงพยาบาลแบบเรียลไทม์และตรวจจับความผิดปกติจากการใช้งานจากการวิเคราะห์ Log จากอุปกรณ์ ในด้านการแสดงผลประสิทธิภาพนั้น ระบบมีความสามารถในการรวบรวมและนำเสนอข้อมูลผ่าน

Dashboard ได้อย่างแม่นยำตามการใช้งานจริง โดยสามารถจำแนกข้อมูลแยกตามรายอาคาร ชั้น และระบบย่อยที่สำคัญ เช่น ระบบผู้ป่วยนอก (OPD), ระบบผู้ป่วยใน (IPD) และระบบห้องยา (PHM) ได้อย่างชัดเจนผ่านกระบวนการจับคู่ IP Address และ Service Port ที่กำหนดไว้

จากการใช้ตัวชี้วัดทางสถิติ P50, P95 และ EWMA50 ภายใต้กรอบ SLO Dynamic ช่วยให้การประเมินประสิทธิภาพระบบสารสนเทศสะท้อนประสบการณ์จริงของผู้ใช้งานได้แม่นยำ แต่ในการพัฒนามีข้อจำกัดด้านข้อมูล เนื่องจาก Log จาก Firewall ไม่ใช่ Application Log จึงไม่สามารถมองเห็นรายละเอียดเชิงลึก หรือ Error ที่เกิดขึ้นภายในระบบสารสนเทศได้ การใช้งานในรูปแบบ Single Node ยังคงมีความเสี่ยงหากโหนดหลักล่มและมีความเร็วในการ Query ที่จำกัดเมื่อมีปริมาณ Log จำนวนมากสำหรับการนำไปใช้งานจริงในระดับองค์กร

## เอกสารอ้างอิง


[1] A. Mohammadpour, M. M. Ghaemi, R. Darrudi, and H. Ebrahimpour Sadagheyani,“Use of hospital information system to improve the quality of health care from clinical staff perspective,” Galen Medical Journal, vol. 10, e1830, 2021.

[2] Z. Chaczko, C. Chiu, A. S. Kohli, and V. Mahadevan,“Smart hospital management system: An integration of enterprise level solutions utilising open group architecture framework (TOGAF),” Proceedings of an IEEE Conference, 2010.

[3] R. Bhavani and V. Sankaradass, “Network traffic analysis in IoT devices using ELK stack,” Proceedings of the 2025 International Conference on Data Science, Agents, and Artificial Intelligence (ICDSAAI 2025), IEEE, 2025.

[4] J. O. Onunga, S. A. Onditi, and P. O. Odongo,“Health information management system data utilization in strategic management decision making in hospitals: Systemic review,” International Journal of Research and Scientific Innovation, vol. 12, no. 15, 2025.

[5] S. J. Son and Y. Kwon,“Performance of ELK stack and commercial system in security log analysis,” Proceedings of the 2017 IEEE 13th Malaysia International Conference on Communications (MICC), IEEE, 2017.

[6] “Innovative approaches to infrastructure monitoring and alerting using the ELK stack,” International Journal of Research Publication and Reviews, vol. 5, no. 10, pp. 356–360, 2024.

[7] I. Barroca Filho et al.,“Development of a health dashboard for an electronic health record system,” Proceedings of the 20th International Conference on Computational Science and Its Applications (ICCSA), IEEE, pp. 16–21, 2020.

[8] D. Helminski et al.,“Dashboards in health care settings: Protocol for a scoping review,” JMIR Research Protocols, vol. 11, no. 3, e34894, 2022.

[9] D. V. Uday and G. S. Mamatha,“An analysis of health system log files using ELK stack,” Proceedings of the 4th International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT), IEEE, pp. 891–894, 2019.

[10] M. M. Kamal, E. Kutafina, and O. Beyan,“Real-time visualization and analysis architecture for data integration processes at Cologne University Hospital’s Medical Data Integration Center,” in Digital Health and Informatics Innovations for Sustainable Health Care Systems, IOS Press, pp. 48–52, 2024.

[11] O. M. Christen et al.,“Dashboard visualization of information for emergency medical services,” in Integrated Citizen Centered Digital Health and Social Care, IOS Press, pp. 27–31, 2020.

[12] Elastic. “The Elasticsearch–OpenSearch performance gap.” Elastic Blog.

[13] Calderón-Gómez, H., Mendoza-Pittí, L., Vargas-Lombardo, M., Gómez-Pulido, J. M., Castillo-Sequera, J. L., Sanz-Moreno, J., and Sención, G. "Telemonitoring System for Infectious Disease Prediction in Elderly People Based on a Novel Microservice Architecture." *IEEE Access*. 8 (2020): 118340-118355. doi: 10.1109/ACCESS.2020.3005638

[14] Boddy, A., Hurst, W., Mackay, M., and El Rhalibi, A. "Density-Based Outlier Detection for Safeguarding Electronic Patient Record Systems." *IEEE Access*. 7 (March 2019): 301. doi: 10.1109/ACCESS.2019.2906503

# การทำเหมืองกฎความสัมพันธ์ของชุดอาหารที่นิยมจากบทวิจารณ์เชิงบวกของลูกค้า
# Mining Popular Food Association Rules from Positive Customer Reviews

*เอกศิษฏ์ จิระชัยพิทักษ์ (Ekasis Girachaipitack)* [1] *ฉวีวรรณ เพ็ชรศิริ (Chaveevan Pechsiri)*[1]
*และ อดิศร ณ อุบล (Adisorn Na Ubon)*[1]
[1]*สาขาวิชาเทคโนโลยีสารสนเทศ วิทยาลัยวิศวกรรมศาสตร์และเทคโนโลยี มหาวิทยาลัยธุรกิจบัณฑิตย์*
*67130263@dpu.ac.th, chaveevan.pec@dpu.ac.th and adisorn.nan@dpu.ac.th*

## บทคัดย่อ

*งานวิจัยนี้มีจุดประสงค์หลักเพื่อทำเหมืองกฎความสัมพันธ์ของชุดอาหารที่นิยมจากบทวิจารณ์เชิงบวกของลูกค้า สำหรับการวิเคราะห์พฤติกรรมการบริโภคที่มีนัยสำคัญ เนื่องจากข้อมูลบทวิจารณ์ออนไลน์ประสบปัญหาสำคัญ 2 ประการ ได้แก่ 1) ความกำกวมของขอบเขตชื่ออาหารในประโยค และ 2) ความหลากหลายของคำเรียกชื่อเมนูอาหารที่มีความหมายเดียวกัน กระบวนการดำเนินงานจึงใช้การแบ่งหน่วยวาทกรรมย่อยร่วมกับการตรวจสอบด้วยฐานข้อมูลชื่ออาหารมาตรฐาน เพื่อระบุรูปแบบชื่ออาหารและดำเนินการสกัดออกเป็นรายการชื่ออาหารที่ชัดเจน ข้อมูลดังกล่าวจะถูกนำไปประมวลผลด้วยอัลกอริทึมอาพริออริ เพื่อค้นหารูปแบบกฎความสัมพันธ์ ผลการทดลองพบว่ากระบวนการสกัดชื่ออาหารมีค่าความแม่นยำร้อยละ 86 และค่าการระลึกร้อยละ 74 โดยกฎความสัมพันธ์ที่ค้นพบมีค่าความเชื่อมั่นเฉลี่ย (Average Confidence) ร้อยละ 82 ซึ่งเป็นข้อมูลที่เป็นประโยชน์ต่อผู้ประกอบการในการวางแผนกลยุทธ์การตลาดและรายการอาหารแนะนำได้อย่างแม่นยำ*

**คำสำคัญ:** อัลกอริทึมอาพริออริ หน่วยวาทกรรมย่อย รายการชื่ออาหาร รูปแบบชื่ออาหาร ฐานข้อมูลชื่ออาหารมาตรฐาน

## Abstract

*This research aims primarily to mine popular food association rules from positive customer reviews to analyze significant consumer behavior patterns. Online reviews typically face two primary challenges: 1) ambiguity in food name boundaries within sentences, and 2) variations in food naming conventions. The methodology employs Elementary Discourse Units (EDUs) integrated with a Food Name List (FNT) to identify food name expressions and extract standardized food name lists. The extracted data is then processed using the Apriori algorithm to discover food association patterns. Experimental results demonstrate a precision of 0.86 and a recall of 0.74 in food name extraction. The discovered association rules achieve an average confidence of 82%, providing valuable insights for business strategy and personalized food recommendations.*

**Keywords:** Apriori Algorithm, Elementary Discourse Units (EDUs), Food Name List, Food Name Expression, FNT.

## 1. บทนำ

ในปัจจุบันการทำธุรกิจร้านอาหารหรือภัตรคารมีการแข่งขันที่สูงขึ้นอย่างต่อเนื่อง ส่งผลให้ผู้ประกอบการต้องเร่งพัฒนากลยุทธ์เพื่อดึงดูดลูกค้าและเพิ่มยอดขาย แหล่งข้อมูลสำคัญที่สะท้อนถึงพฤติกรรมและความต้องการที่แท้จริงของผู้บริโภคคือ ข้อมูลบทวิจารณ์รีวิวออนไลน์ (Online Reviews) ซึ่งเป็นข้อมูลที่ไม่มีโครงสร้าง (Unstructured Data) จำนวนมาก งานวิจัยนี้จึงมีจุดประสงค์คือ การทำเหมืองกฎความสัมพันธ์ของชุดอาหารที่นิยมจากบทวิจารณ์เชิงบวกภาษาไทยของลูกค้า เพื่อหารูปแบบหรือกฎความสัมพันธ์ของรายการอาหารที่ลูกค้าสั่งภายในแต่ละบทวิจารณ์นั้น เพื่อนำมาจัดชุดอาหารยอดนิยม (Popular Food Combinations) ซึ่งจะช่วยเพิ่ม

โอกาสในการขายและยกระดับประสบการณ์ที่ดีให้แก่ลูกค้าได้อย่างมีนัยสำคัญ ดังนั้นความสำคัญของงานวิจัยนี้จึงเป็นการค้นพบองค์ความรู้เชิงกลยุทธ์จากข้อมูลที่ไม่มีโครงสร้างนั้นทำให้ร้านอาหารสามารถออกแบบโปรโมชั่นที่ตรงใจลูกค้าได้มากกว่าการคาดเดา นอกจากนี้บทวิจารณ์ หรือเอกสารข้อความวิจารณ์จะแสดงออกมาในรูปแบบหน่วยวาทกรรมพื้นฐาน (Elementary Discourse Unit, EDU, คือประโยคความเดียว [1]) ดังตัวอย่างต่อไปนี้

ตัวอย่างที่ 1

EDU:1 "*อาหารรสชาดดีคือ<u>ปูไข่นึ่งนมสด</u>และ<u>กุ้งแม่น้ำเผา</u>*"
EDU:2 "*<u>สั่ง</u>ไปหลายอย่างมากรอบนี้*"
EDU:3 "*อาหารที่นี่มีรสจัดมาก*"
EDU:4 "*<u>ต้มยำกุ้ง</u>มีรสอร่อย*"
EDU:5 "*<u>ปูนึ่งมทอดกระเทียม</u>ก็มีรสอร่อย*"
EDU:6 "อาหารมาเร็วมาก"
*เมื่อ <u>ปูไข่นึ่งนมสด</u>, <u>กุ้งแม่น้ำเผา</u>, <u>ต้มยำกุ้ง</u>, และ <u>ปูนึ่มทอดกระเทียม</u> คือชื่ออาหาร*

การแสดงออกของแต่ละ EDU อิงตามหลักภาษาไทยทั่วไป (General Thai Linguistic Expression) ดังต่อไปนี้ หลังจากได้ทำการหาแก่นคำ (stemming words)และกำจัดคำหยุด (stop words)

EDU → NP1 VP | VP
VP → Verb NP2 | Verb modifier | Verb
NP1 → NP | NPf | NPf NP1 | Conj NP1
NP2 → NP | NPf | NPf NP2 | Conj NP2
NPf → FoodName | FoodName modifier
NP → Noun Modifier
Modifier → Adv | Adj | AdjPhrase
Conj → {' ','และ', 'หรือ', …}
Noun → {'*อาหาร/food*', '*รสชาด/taste*' , …. }
Verb → {'*สั่ง/order*', '*มา/come*',...}
Adj →{ '*ดี/good*', '*จัด/strong*', '*อร่อย/delicious*',...}
Adv →{ '*เร็ว/quickly*', '*มาก/much*', ...}

เมื่อ NP1, NP2 คือนามวลี; NP คือนามวลีย่อย; NPf คือนามวลีย่อยที่เกี่ยวกับอาหาร; VP คือ กริยาวลี; Verbคือ เซตคำกริยา; Nounคือ เซตคำนาม; Adv คือ เซตคำขยายกริยา; Adj คือเซตคำขยายนามหรือนามวลี; AdjPhrase คือคุณศัพท์วลี;และ Conj คือ เซตคำเชื่อม

FoodName = $w_1+w_2+\ldots+w_{num}$ ; *num* คือจำนวนคำทั้งหมดใน FoodName; $w_i \in$ Word , $i$=1,2,..,*num*; Word = Nf $\cup$ Vf $\cup$ Adjf เมื่อ Nf คือเซตแนวความคิดคำนาม (noun concept) ที่เกี่ยวกับอาหาร Vf เซตแนวความคิดคำกริยา (verb concept) ที่เกี่ยวกับอาหาร และAdjf เซตแนวความคิดคำขยายนาม (adjective concept) ที่เกี่ยวกับอาหาร

Nf ={ '*หมี่/noodle*','*ปูไข่/ egg-bearing crab*', '*กุ้งแม่น้ำ/giant river prawn*', '*ต้มยำ/spicy lemongrass soup*', '*กุ้ง/shrimp*', '*กระเทียม/garlic*' , …. }
Vf ={ '*ต้ม/boil*', '*นึ่ง/steam*', '*ทอด/fry*', '*ผัด/stir fry*'..}
Adjf = {'*เล็ก/small*', '*ใหญ่/large*', '*นิ่ม/soft*',…}

อย่างไรก็ตามการค้นพบรูปแบบชุดอาหารยอดนิยมจากบทวิจารณ์ดังกล่าวที่เป็นเชิงบวกในงานวิจัยนี้มีปัญหาหลัก 2 ประการดังนี้

1) การระบุฟีเจอร์ชื่ออาหาร (Food-Name feature, *Fname*) ในบทวิจารณ์เชิงบวก ส่วนใหญ่เป็นชื่อสามัญที่ปรากฏเป็น NPf ใน NP1 หรือ NP2 ประกอบด้วย $w_i$ เช่น '*ปูไข่นึ่งนมสด*' ='*ปูไข่*'+ '*นึ่ง*'+ '*นมสด*' และ '*กุ้งแม่น้ำเผา*'= '*กุ้งแม่น้ำ*'+ '*เผา*' ใน EDU ที่1 ของตัวอย่างที่1
2) การทำเหมืองกฎความสัมพันธ์เพื่อการค้นพบแพทเทิร์นรูปแบบที่เป็นกฎความสัมพันธ์ *Fname* ที่มีแนวความคิดทั่วไป ในรายการอาหารชื่ออาหาร (Food Name List) ของบทวิจารณ์เชิงบวกของลูกค้าทั้งหลาย แต่การปรากฏของ *Fname* ในบทวิจารณ์เชิงบวกมีการแสดงออกหลากหลายรูปแบบมากเมื่อเทียบกับจำนวนบทวิจารณ์ที่ดาวน์โหลด เช่น '*ไก่ผัดกระเพรา*' '*ผัดกระเพราไก่*' '*ปลาหมึกผัดกระเพรา*'และ '*ผัดกระเพราหมู*' ทำให้ *Fname* มีชื่ออาหารต่าง ๆ มากมาย ซึ่งมีผลต่อการหากฎความสัมพันธ์ ทั้งที่เป็นอาหารชนิดเดียวกัน คือ '*ผัดกระเพรา*'

ดังนั้น งานวิจัยนี้จึงประยุกต์ใช้เซตคำ (ที่ปรากฏใน *Fname*) ซึ่งเรียกว่า เซต FW (Food-Name Word Set) ที่ได้จากการศึกษาพฤติกรรมข้อมูล (Corpus Behavior Study) โดยให้ผู้เชี่ยวชาญกำกับชื่ออาหารจากกลุ่มคำที่ปรากฏใน NP1 หรือ NP2 ของ EDU ต่าง ๆ ใน คลังข้อมูลบทวิจารณ์อาหาร (Food Review Corpus) จากร้านต่าง ๆ FW ใช้ในการกรองคำต่าง ๆ ที่เกี่ยวกับอาหาร จาก NP1 หรือ NP2 ของแต่ละ EDU เพื่อ หา *Fname* ด้วยวิธีวิธีการจับคู่สาย

อักขระ (String Matching Method) กับข้อมูลในตารางรายชื่ออาหาร (Food Name Table, FNT) ที่สร้างจากข้อมูลรายชื่ออาหารของ Michelin Guide (https://guide.michelin.com/th/th/selection/thailand/restaurants) และ UniLever Food Solution (https://www.Unileverfoodsolutions.co.th/th.html) ซึ่งทั้งสองเป็นองค์กรด้านอาหารที่มีชื่อเสียง FNT ประกอบด้วยฟีเจอร์ต่อไปนี้ *fn* และ *fnExp* (*fn*∈FNAME ซึ่งเป็นเซตของรายชื่ออาหาร และ *fnExp* ∈ FNAMEexpression ซึ่งคือเซตของตัวแทนรายชื่ออาหาร; *fnExp* แสดงด้วยรูปแบบของเพรดิเคตชื่ออาหาร (Food Name Predicate) ดังนี้

'กริยาวลี(ตัวกระทำ/ผู้ถูกกระทำที่มีแนวความคิดเป็นเนื้อสัตว์: meat concept)'

ดูตารางที่1 นอกจากนี้ตัวกระทำผู้ถูกกระทำที่มีแนวความคิดเป็นเนื้อสัตว์ในเพรดิเคตชื่ออาหารให้แทนด้วยตัวแปร m เพื่อให้เป็นรูปทั่วไป ดังนั้นเราใช้ *fnExp* แก้ปัญหา *Fname* ที่แสดงออกหลากหลายรูปแบบ นอกจากนี้งานวิจัยนี้ได้เสนอใช้อัลกอริทึมกฎความสัมพันธ์ (Association Rule Algorithm [2]) ทำเหมืองหากฎความสัมพันธ์ของชุดอาหารที่นิยมจากบทวิจารณ์เชิงบวก ซึ่งก็คือเป็นการค้นพบรูปแบบที่เป็นกฎความสัมพันธ์ของรายการชื่ออาหารที่เป็น *fnExp* จากบทวิจารณ์ต่าง ๆ เชิงบวกของลูกค้า

**ตารางที่ 1** : รายชื่ออาหาร (Food Name Table, FNT)

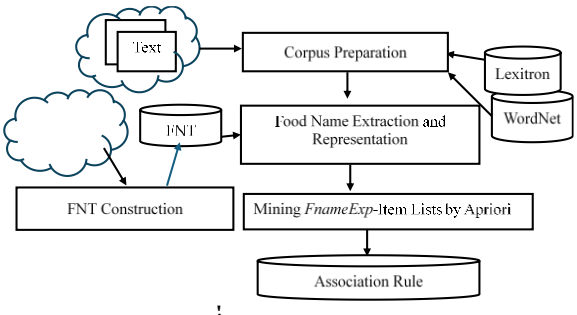
| *fn*∈FNAME | *fnExp* ∈ FNAMEexpression |
|---|---|
| *ผัดกระเพราไก่* | *ผัดกระเพรา*(*ไก่*)/*ผัดกระเพรา*(m) |
| *ไก่ผัดกระเพรา* | *ผัดกระเพรา*(*ไก่*)/*ผัดกระเพรา*(m) |
| *ผัดกระเพราหมู* | *ผัดกระเพรา*(*หมู*)/*ผัดกระเพรา*(m) |
| *ปูนิ่มทอดกระเทียม* | *ทอดกระเทียม*(*ปูนิ่ม*)/*ทอดกระเทียม*(m) |
| *กุ้งคั่วพริกเกลือ* | *คั่วพริกเกลือ*(*กุ้ง*)/*คั่วพริกเกลือ* (m) |
| ..... | ..... |

ในส่วนที่ 2 สรุปงานวิจัยที่เกี่ยวข้อง ส่วนที่ 3 กระบวนการสกัดชื่ออาหารและการทำเหมืองกฎความสัมพันธ์ และ ส่วนที่ 4 ประเมินผลการอภิปราย และข้อสรุป

## 2. งานวิจัยที่เกี่ยวข้อง

การประยุกต์ใช้อัลกอริทึม Apriori ในงานวิจัยที่ผ่านมามุ่งเน้นการวิเคราะห์ข้อมูลเชิงรายการ (Transactional Data) เพื่อเพิ่มประสิทธิภาพทางธุรกิจในหลากหลายมิติ อาทิ Ndruru และ Hasugian(2020)[3] ที่ใช้บริหารจัดการสต็อกวัตถุดิบ, Laurentinus และคณะ(2021)[4] ที่บูรณาการร่วมกับ Linear Regression เพื่อพัฒนาระบบแนะนำสินค้า, Arboleda และคณะ(2025)[5] ที่เน้นกลยุทธ์การจัดวางสินค้าเชิงกายภาพ และ Lee กับ Ahmatang(2025)[6] ที่ใช้ตัดสินใจเชิงกลยุทธ์ผ่านพฤติกรรมการซื้อ อย่างไรก็ตามงานวิจัยเหล่านี้ [3]-[6] ยังมีข้อจำกัดในการพึ่งพาข้อมูลที่มีโครงสร้างชัดเจน(Structured Data) ซึ่งไม่สามารถสะท้อนบริบทเชิงลึกจากความคิดเห็นของผู้บริโภคได้ งานวิจัยนี้จึงเติมเต็มช่องว่างดังกล่าวโดยนำเสนอเฟรมเวิร์กที่รองรับการวิเคราะห์ข้อมูลที่ไม่มีโครงสร้าง(Unstructured Data) ในรูปแบบบทวิจารณ์เชิงข้อความ(Text-based) ผ่านการใช้หน่วยวาทกรรมย่อย(EDUs) เพื่อสกัดและระบุขอบเขตแนวคิดชื่ออาหาร ซึ่งช่วยให้ค้นพบรูปแบบความสัมพันธ์ที่ซ่อนอยู่ในบริบททางภาษาศาสตร์ที่ข้อมูลเชิงรายการทั่วไปไม่สามารถตรวจพบได้

## 3. ภาพรวมของระบบ




**ภาพที่ 1:** ภาพรวมของระบบ

การค้นพบรูปแบบชุดอาหารยอดนิยมจากบทวิจารณ์เชิงบวกด้วยการทำเหมืองข้อความ ประกอบด้วย 3 ขั้นตอนหลักคือ การเตรียมคลังข้อมูล (Corpus Preparation) การสกัด และการแทนชื่ออาหาร (Food Name Extraction and Representation) การทำเหมืองรายการ *FnameExp*-Item

ด้วย Aprior (Mining *FnameExp*-Item Lists by Apriori) (ดูภาพที่ 1)

**3.1 การเตรียมคลังข้อมูล**

บทวิจารณ์เชิงบวก (รายการอาหาร/food list):
"*อาหารรสชาดดีคือปูไข่นึ่งนมสดและกุ้งแม่น้ำเผา*" EDU:1 "*สั่งไปหลายอย่างมากรอบนี้*" EDU:2 "*อาหารที่นี่มีรสจัดมาก*" EDU:3 "*ต้มยำกุ้งมีรสอร่อย*" EDU:4 "*ปูนิ่มทอดกระเทียมก็มีรสอร่อย*" EDU:5 "อาหารมาเร็วมาก" EDU:6

<EDU:1><NP1 food-name=no>*อาหารรสชาดดี*</NP1>*คือ*<NP2 food-name=yes><FoodName: 3words><word: 1; pos=compound noun; concept= '*egg-bearing crab*'>*ปูไข่*</word><word: 2; POS=verb ; concept= '*steam*'>*นึ่ง*</word><word: 3; POS= noun; concept= '*fresh milk*'>*นมสด*</word></FoodName>*และ*<FoodName: 2words><word: 1; POS=compound noun; concept= '*giant river prawn*'>*กุ้งแม่น้ำ*</word><word: 2; POS=verb ; concept= '*roast*'>*เผา*</word> </FoodName></NP2></EDU>

<EDU:3><NP1 food-name=yes><FoodName: 2words><word: 1; POS=compound noun; concept= '*spicy lemongrass soup*'>*ต้มยำ*</word><word: 2; POS=noun ; concept= '*shrimp*'>*กุ้ง*</word></FoodName></NP1>*มี*<NP2 food-name=no>*รสอร่อย*</NP2></EDU>

<EDU:4><NP1 food-name=yes><FoodName: 3words><word: 1 ;POS= compound noun; concept= '*soft-shell crab*'>*ปูนิ่ม*</word><word: 2; POS=verb ; concept= '*fry*'>*ทอด*</word><word: 3; POS= noun; concept= '*garlic*'>*กระเทียม*</word></FoodName></NP1>*ก็มี*<NP2 food-name=no> *รสอร่อย*</NP2></EDU>

Where EDU = an EDU tag; NP1, NP2= a noun phrase tag; FoodName = a food name tag; word = a word tag; and POS is Part of Speech.

**ภาพที่ 2:** การกำกับชื่ออาหาร

ขั้นตอนนี้มีวัตถุประสงค์เพื่อเตรียมคลังข้อมูล EDU (EDU corpus) ที่ได้จากการดาวน์โหลดบทวิจารณ์ (เอกสาร) ออนไลน์ของลูกค้า(wongnai) ขั้นตอนดังกล่าวประกอบ ด้วยการใช้เครื่องมือตัดคำภาษาไทย [7] และการแบ่งหน่วย EDU (EDU Segmentation) [8] เพื่อสร้างคลังข้อมูลจำนวน 600 EDUs จากเอกสารบทวิจารณ์ดังกล่าวจำนวน 100 ฉบับ ผ่านการตัดคำให้เหลือแก่นคำเชิงรูปแบบ (stemming) และการกำจัดคำหยุด (stop word removal) และถูกแบ่งออกเป็น 3 ส่วน ได้แก่

1) ส่วนแรกจำนวน 200 EDUs ใช้เป็นคลังข้อมูลที่มีตัวกำกับ FoodName กำกับชื่ออาหารที่อยู่ในตัวกำกับ NP1 หรือ NP2 (NP1 or NP2 Tag) ภายใต้ตัวกำกับ EDU (EDU Tag) (ดูภาพที่ 2) โดยแนวคิดของคำกริยาและคำขยายอ้างอิงจาก WordNet [9] หลังจากแปลจากภาษาไทยเป็นภาษาอังกฤษด้วย Lexitron [10]) เพื่อใช้ในการรวบรวมเซต Verb และเซต Word

2) ส่วนที่สองจำนวน 300 EDUs ใช้สำหรับการหากฎความสัมพันธ์ของรายการอาหารในรูปของเพรดิเคตชื่ออาหาร

3) ส่วนที่สามจำนวน 100 EDUs (10 บทวิจารณ์) ใช้สำหรับการทดสอบกฎแพทเทิร์นความสัมพันธ์ของรายการอาหารที่แทนด้วยเพรดิเคตชื่ออาหาร

**3.3 การสกัด และการแทนชื่ออาหาร**

หลังจากขั้นตอน 3.2 จะได้ เซต FW จากการรวบรวมคำต่าง ๆ ที่อยู่ภายใต้ ป้ายตัวกำกับชื่อ FoodName Tag

FW ={'*ปูไข่/egg-bearing crab*' '*นึ่ง/steam*' '*นมสด/fresh milk*' '*กุ้งแม่น้ำ/giant river prawn*' '*เผา/roast*' ……. }

ก่อนที่จะทำการสกัดและการแทนชื่ออาหาร จะต้องทำการสร้างตารางรายชื่ออาหาร (FNT) จากข้อมูลรายชื่ออาหารบนเว็บไซต์ของ Michelin Guide และ UniLever Food Solution ดังแสดงในตารางที่ 1

ต่อมาใช้ ทำการสกัด *Fname* ทั้งหมดจากแต่ละ NP1/NP2 ใน แต่ละ EDU ของแต่ละบทวิจารณ์ในคลังข้อมูลเอกสารบทวิจารณ์ส่วนที่สองนั้น โดย แต่ละคำที่สแกนใน NP1/NP2 เป็นสมาชิกของเซต FW หากเป็นสมาชิกก็ให้สแกนคำถัดไป หากคำถัดไปยังคงเป็นสมาชิกก็ให้สแกนจนกระทั่งไม่เป็นสมาชิก หรือ สิ้นสุด NP1/NP2 ก็จะได้ขอบเขตของคำที่ใน *Fname* (หมายเหตุ *Fname* ต้องมีจำนวนคำตั้งแต่ 2คำขึ้นไป)

ขั้นตอนต่อมาทำการแทนแต่ละ *Fname* ที่สกัดได้ด้วย *FnameExp* ซึ่งเป็นเพรดิเคตชื่ออาหารในรูปแบบทั่วไป โดย วิธีการจับคู่สายอักขระระหว่างแต่ละ *Fname* และ *fn* ใน FNT หากจับคู่ได้ก็จะได้ *FnameExp* ของแต่ละ *Fname* จาก *fnExp* ของ *fn* ดังนี้ *FnameExp = fnExp*

**3.4 การทำเหมืองรายการ *FnameExp*-Item ด้วย Apriori**

ขั้นตอนนี้เป็นทำเหมืองกฎความสัมพันธ์ของชุดอาหารที่นิยมจากบทวิจารณ์เชิงบวก จากข้อ3.3แต่ละตัวอย่างบทวิจารณ์หรือรีวิว *FnameExp*-Item List ของ

ลูกค้า (1 รีวิว = 1 ตะกร้า) กลายเป็นแต่ละ Transaction ของฟีเจอร์ Tranaction ID ($T_i$) เช่น จากตัวอย่างที่ 1 ได้ T1: *นึ่งนมสด*(m), *เผา*(m), *ต้มยำ*(m), *และ ทอดกระเทียม*(m) จาก *นึ่งนมสด*(*ปูไข่*), *เผา*(*กุ้งแม่น้ำ*), *ต้มยำ*(*กุ้ง*), *และ ทอดกระเทียม*(*ปูนิ่ม*) ตามลำดับ

Transactionที่ได้ทั้งหมดมาหากฎความสัมพันธ์โดยอัลกอลิทึม Apriori ดัง ต่อไปนี้ ด้วย Weka [11-12] ที่ support ≥ 0.2 และ Minimum confidence = 0.8 ซึ่งเป็นค่าที่เหมาะสมในเชิงธุรกิจตามคำแนะนำของผู้เชี่ยวชาญ

## Algorithm: Apriori

*/ D={$T_1$,$T_2$,...,$T_n$}
(แต่ละ $T_i$ คือ *FnameExp*-Item List จากรีวิว/ธุรกรรม)

```
1.  k ← 1
2.  C1 ← สร้าง candidate 1-itemsets จาก D
3.  L1 ← เลือก itemsets ใน C1 ที่ support ≥
                        minSupport

4.  while Lk ≠ ∅ do
5.      k ← k + 1
6.      Ck ← สร้าง candidate k-itemsets
                จาก Lk−1
7.      สำหรับแต่ละ transaction Ti ใน D do
8.          เพิ่ม support count ให้
              itemsets ใน Ck ที่อยู่ใน Ti
9.      end for
10.     Lk ← เลือก itemsets ใน Ck ที่
            support ≥ minSupport
11. end while

12. L ← ∪ Lk    (รวม frequent itemsets
                                ทุกระดับ)
13. สำหรับแต่ละ frequent itemset X ใน L ที่
                    |X| ≥ 0.2 do
14.     สำหรับทุก subset A ⊂ X do
15.         B ← X − A
16.         ถ้า confidence(A → B) ≥
                minConfidence then
17.             สร้างกฎ A → B
18.         end if
19.     end for
20. end for
```

## 4. การประเมินผลการอภิปรายและข้อสรุป

งานวิจัยที่เสนอมีการประเมินผล 2 ส่วนกับข้อมูลทดสอบ จำนวน 100 EDUs (10 บทวิจารณ์):

**4.1 การประเมินส่วนแรกคือการระบุฟีเจอร์ *Fname*** โดยผู้เชี่ยวชาญ 3 ท่าน ผ่านการลงคะแนนแบบชนะสูงสุด (Max win Voting) การประเมินนี้อิงตามค่าความแม่นยำ (Precision,PC) เป็น 0.86 และค่าระลึก (Recall,RC) เป็น 0.74 โดยสามารถวิเคราะห์ข้อผิดพลาดที่เกิดขึ้นจริงได้เป็น 2 กรณีหลัก กรณีแรกคือข้อผิดพลาดประเภทผลบวกปลอม ซึ่งเกิดจากการการรวมคำนามผิดบริบทเนื่องจากปรากฏต่อเนื่องกันในหน่วยวาทกรรมเดียวกัน ดังบทวิจารณ์ที่ระบุว่า "…ปลาตัวใหญ่มาก มี ข้าวผัดปูหอมกลิ่นกระทะ…" ซึ่งระบบทำการสกัดเป็น "ปลาข้าวผัดปู" เนื่องจากความผิดพลาดในการระบุขอบเขตคำนามในเชิงบริบท กรณีที่สองคือข้อผิดพลาดประเภทผลลบปลอมซึ่งมีสาเหตุหลักจากการแทรกคำบรรยายลักษณะอาหารไว้กลางชื่อเมนูมาตรฐาน ดังบทวิจารณ์ที่ระบุว่า "…ปูนิ่มทอดมาได้กรอบกระเทียมโรยมาแบบจัดเต็ม…" ในกรณีนี้ระบบสกัดได้เพียงคำว่า "ปูนิ่ม" เนื่องจากไม่สามารถจับคู่คำได้สมบูรณ์ตามฐานข้อมูลอ้างอิงเมื่อมีคำขยายแทรกกลางประโยค อย่างไรก็ตาม ประสิทธิภาพในระดับดังกล่าวยังถือว่ามีความเหมาะสมเพียงพอต่อการนำไปประมวลผลต่อด้วยอัลกอริทึมอาพริออริ เพื่อค้นหาพฤติกรรมการสั่งอาหารที่ถูกต้องและปราศจากข้อมูลที่ไม่เกี่ยวข้อง จากการประมวลผลด้วยอัลกอริทึมอาพริออริ พบกฎความสัมพันธ์ทั้งหมด 10 กฎ โดยคัดเลือก 5 อันดับแรกที่มีค่าความเชื่อมั่นสูงสุดมาแสดงดังตารางที่ 2

**ตารางที่ 2**: ตัวอย่างกฎความสัมพันธ์ที่มีค่าความเชื่อมั่นสูงสุด 5 อันดับแรก

| ลำดับ | กฎความสัมพันธ์ (Association Rules) | Supp. | Conf.(Lift) |
|---|---|---|---|
| 1 | {ข้าวผัดปู, ไก่ผัดเม็ดมะม่วงหิมพานต์} → {ต้มยำกุ้ง} | 0.21 | 0.95 (1.61) |
| 2 | {ปูไข่นึ่ง, ไม่สั่งต้มยำกุ้ง} → {ไม่สั่งไก่ผัดเม็ดมะม่วงหิมพานต์} | 0.27 | 0.93 (1.62) |
| 3 | {ปูไข่นึ่ง, ไม่สั่งเป็ดปักกิ่ง, ไม่สั่งไก่ผัดเม็ดมะม่วงหิมพานต์} → {กุ้งแม่น้ำเผา} | 0.23 | 0.91 (1.40) |
| 4 | {ปูไข่นึ่ง, ไก่ผัดเม็ดมะม่วงหิมพานต์} → {ต้มยำกุ้ง} | 0.22 | 0.91 (1.54) |
| 5 | {ไม่สั่งปลากะพงทอดน้ำปลา, ไก่ผัดเม็ดมะม่วงหิมพานต์} → {ต้มยำกุ้ง} | 0.22 | 0.91 (1.54) |

**4.2 จากการวิเคราะห์กฎในตารางที่ 2** ร่วมกับบทวิจารณ์จริง 10 รายการ พบว่ากฎลำดับที่ 1 (Confidence 0.95) ปรากฏในข้อมูลจริงถึง 8 ครั้ง หรือร้อยละ 80 ซึ่งยืนยันว่าชุดอาหารที่ระบบค้นพบสอดคล้องกับพฤติกรรมการสั่งจริงอย่างมีนัยสำคัญ นอกจากนี้ในกฎลำดับที่ 2 และ 3 ที่ปรากฏเงื่อนไขเชิงปฏิเสธ (Negative Association) หรือการ “ไม่สั่ง” (ระบุเป็นค่า no ในระบบประมวลผล) รายการอาหารบางชนิด สะท้อนถึงพฤติกรรมการบริโภคแบบเฉพาะเจาะจง (Selective Consumption)เช่น กรณีบทวิจารณ์ที่ลูกค้าเลือกเฉพาะ “ปูไข่นึ่ง” และ “กุ้งแม่น้ำเผา” โดยปฏิเสธเมนูเนื้อสัตว์ชนิดอื่น สะท้อนถึงกลุ่มเป้าหมายที่เน้นบริโภคอาหารทะเลคุณภาพสูงมากกว่าชุดอาหารทั่วไป โดยสรุปกฎความสัมพันธ์ทั้ง 10 กฎที่สกัดได้มีค่าความเชื่อมั่นเฉลี่ยร้อยละ 82 ซึ่งสูงกว่าเกณฑ์ขั้นต่ำ (minConf = 0.8) ที่กำหนดไว้ จากพฤติกรรมดังกล่าวชี้ให้เห็นว่า ชุดอาหารเหล่านี้มีความสัมพันธ์กันอย่างเหนียวแน่นในกลุ่มเป้าหมายที่เน้นบริโภควัตถุดิบคุณภาพสูง ซึ่งข้อมูลเชิงลึกนี้มีความสำคัญต่อผู้ประกอบการมากกว่าการวิเคราะห์ยอดขายทั่วไป เนื่องจากช่วยให้สามารถออกแบบรายการอาหารแนะนำที่ตรงใจลูกค้าเฉพาะกลุ่มได้แม่นยำยิ่งขึ้น งานวิจัยนี้จึงบรรลุจุดประสงค์หลักในการทำเหมืองกฎความสัมพันธ์เพื่อสร้างมูลค่าเพิ่มจากข้อมูลบทวิจารณ์ออนไลน์ได้อย่างมีประสิทธิภาพ

## เอกสารอ้างอิง


[1] Carlson L, Marcu D, Okurowski M E. Building a Discourse-Tagged Corpus in the Framework of Rhetorical Structure The-ory. *Current Directions in Discourse and Dialogue*, van Kup-pevelt J, Smith R (eds.), Kluewer Academic Publishers,2003,pp.85112.

[2] Han J, Kamber M, Pei J. (2012). Data mining: Concepts and techniques (3rd ed.). Morgan Kaufmann, pp 53-87.

[3] Ndruru R and Hasugian P M. Determination of Data Mining Application Design Patterns Booking Raw Food In Restaurant Fountain With Apriori Algorithm. *Journal of Computer Networks, Architecture and High Performance Computing*, Vol.2, No. 1, 2020, pp 275-285

[4] Laurentinus, Rizan O, Sarwindah, Hamidah, Sulaiman R, Hengki . Data Mining Using Apriori Algorithm and Linear Regression in Product Recommendations. 4th International Seminar on Research of Information Technology and Intelligent Systems (ISRITI), Indonesia, 2021, pp. 217-223, doi: 10.1109/ISRITI54043.2021.9702791.
Regression in Product Recommendations

[5] Arboleda FJM, Garani G, Correa AFA. Supermarket Product Placement Strategies Based on Association Rules. *IAENG International Journal of Computer Science*, Vol 51, no. 6, 2024, pp 650-662.

[6] Lee C-W and Ahmatang. Market Basket Analysis Using Apriori Algorithm: Identifying Consumer Purchase Patterns for Strategic Business Decisions. *Journal of Risk and Control*, Vol. 12, No. 1, 2025, pp 101-114

[7] Sudprasert S, Kawtrakul A. Thai word segmentation basedon global and local unsupervised learning. In *Proc. NCSEC2003*, Chonburi, Thailand, 2003, pp.1-8.

[8] Chareonsuk J, Sukvakree T, Kawtrakul A. Elementary discourse unit segmentation for Thai using discourse cue and syntactic information. In Proc. NCSEC 2005, Thailand, Oct. 27-28, 2005, pp.85-90.

[9] Miller G.A. WordNet: a lexical database for English Commun. *ACM*, 38 (11) (1995), pp. 39-41

[10] Trakultaweekoon K, Porkaew P, Supnithi T. (2007). “LEXiTRON vocabulary suggestion system with recommendation and vote mechanism,” in 7th International Symposium on Natural Language Processing (SNLP) 2007, Pattaya, Thailand, 43–48.

[11] R.R. Bouckaert, et al.WEKA Manual for Version 3-9-1 Univ. Waikato Hamilt. N. Z. (2016), pp. 1-341

[12] Hall M, Frank E, Holmes G, Pfahringer B, Reutemann P, Witten I.H. The WEKA data mining software: an update ACM SIGKDD Explor. *Newsl*., 11 (1) (2009), pp. 10-18

# การทำเหมืองความคิดเห็นเชิงแง่มุมจากบทวิจารณ์ของผู้ใช้ในธุรกิจออนไลน์
# Aspect-Based Opinion Mining of User Reviews in Online Businesses

*บุญญวัฒน์ พัฒนศักดิ์กำจร (Bunyawat Pattanasakkomjon)* [1] *ฉวีวรรณ เพ็ชรศิริ (Chaveevan Pechsiri)*[1]
*และอดิศร ณ อุบล (Adisorn Na Ubon)*[1]
[1]*สาขาวิชาเทคโนโลยีสารสนเทศ วิทยาลัยวิศวกรรมศาสตร์และเทคโนโลยี มหาวิทยาลัยธุรกิจบัณฑิตย์*
*67130073@dpu.ac.th, chaveevan.pec@dpu.ac.th and adisorn.nan@dpu.ac.th*

## บทคัดย่อ

*งานวิจัยนี้มีวัตถุประสงค์เพื่อสะท้อนความคิดเห็น ความพึงพอใจ และปัญหาที่ลูกค้าได้รับ อย่างไรก็ตาม การวิเคราะห์ความคิดเห็นโดยรวมเพียงระดับบวก ลบ หรือเป็นกลาง อาจไม่เพียงพอต่อการทำความเข้าใจเชิงลึกในแต่ละแง่มุมของสินค้า และบริการ งานวิจัยนี้จึงมุ่งศึกษา* ***การทำเหมืองความคิดเห็นเชิงแง่มุม (Aspect-Based Opinion Mining)*** *จากบทวิจารณ์ของผู้ใช้ในธุรกิจออนไลน์ โดยมีวัตถุประสงค์เพื่อระบุแง่มุม (Aspect) ที่ผู้ใช้กล่าวถึง เช่น คุณภาพสินค้า ราคา การจัดส่ง พร้อมทั้งวิเคราะห์ทัศนคติ (Sentiment) ที่มีต่อแต่ละแง่มุม วิธีการวิจัยอาศัยเทคนิคการประมวลผลภาษาธรรมชาติ (Natural Language Processing: NLP) และการเรียนรู้ของเครื่อง (Machine Learning) ในการสกัดแง่มุม และจำแนกความคิดเห็น ผลการศึกษาคาดว่าจะช่วยให้ผู้ประกอบการสามารถเข้าใจความต้องการของลูกค้าได้อย่างละเอียด และนำไปใช้ในการปรับปรุงสินค้าและบริการให้ตรงกับความคาดหวังของผู้บริโภคมากยิ่งขึ้น ผลการวิจัยคาดว่าจะช่วยให้ผู้ประกอบการสามารถเข้าใจแง่มุมที่ถูกกล่าวถึง และความคิดเห็นที่มีต่อแง่มุมนั้นได้อย่างละเอียด ช่วยให้การวิเคราะห์ข้อมูลมีความแม่นยำ และลึกซึ้งมากยิ่งขึ้น เพื่อสนับสนุนการตัดสินใจเชิงกลยุทธ์ของธุรกิจและยกระดับคุณภาพสินค้าและบริการในระยะยาว*

**คำสำคัญ:** ความคิดเห็นของลูกค้า การวิเคราะห์แง่มุม ทัศนคติของลูกค้า ธุรกิจออนไลน์ การทำเหมืองความคิดเห็นเชิงแง่มุม

## Abstract

*This research aims to reflect customers' opinions, satisfaction, and problems. However, analyzing opinions only as positive, negative, or neutral is not enough to deeply understand different aspects of products and services. Therefore, this study focuses on Aspect-Based Opinion Mining from user reviews in online businesses. The objective is to identify aspects mentioned by users, such as product quality, price, and delivery, and to analyze the sentiment toward each aspect. The research uses Natural Language Processing (NLP) and Machine Learning techniques to extract aspects and classify opinions. The expected results will help business owners gain a deeper understanding of customer needs and use this information to improve their products and services to better meet customer expectations. Overall, the findings are expected to help businesses clearly understand which aspects are discussed and how customers feel about each aspect. This will lead to more accurate and in-depth data analysis, support strategic business decisions, and improve the quality of products and services in the long term.*

**Keywords:** Customer Opinions, Aspect Analysis, Customer Sentiment, Online Business, Aspect-based Opinion Mining.

## 1. บทนำ

ในยุคดิจิทัล ธุรกิจออนไลน์มีบทบาทสำคัญต่อวิถีชีวิตของผู้บริโภค โดยเฉพาะแพลตฟอร์มอีคอมเมิร์ซที่ลูกค้าใช้สามารถแสดงความคิดเห็นได้อย่างอิสระ บทวิจารณ์เหล่านี้ไม่เพียงส่งผลต่อการตัดสินใจซื้อของลูกค้ารายอื่นเท่านั้น แต่ยังเป็นข้อมูลเชิงลึกที่ผู้ประกอบการใช้ในการวิเคราะห์พฤติกรรม ความพึงพอใจ และปัญหาที่เกิดขึ้นกับลูกค้า โดยมักมีแง่มุมที่เกี่ยวกับ คุณภาพสินค้า, ราคาสินค้า และ การจัดส่ง ที่ส่งผลต่อทัศนคติได้ เช่น ตรงตามโฆษณา มีราคาถูก หรือจัดส่งเร็ว โดยความคิดเห็นเหล่านี้สามารถแสดงออกมาในรูปของการบรรยายในรีวิวต่าง ๆ เช่น คำวิจารณ์ของลูกค้า ดังนั้นงานวิจัยนี้จึงมีจุดประสงค์การทำเมืองข้อความภาษาไทย (Thai Text Mining) โดยเฉพาะการทำเหมืองความคิดเห็นตามแง่มุมหรือประเด็น (Aspect-Based Opinion Mining) จากบทวิจารณ์ออนไลน์ (Shopee) ของผู้ใช้ในธุรกิจออนไลน์ โดยการทำเหมืองความคิดเห็นนี้ครอบคลุมถึงการสกัด (Extraction) และการจำแนก (Classification) ประเภทแง่มุม (Aspect Type) ออกเป็นคลาส (Class) ต่าง ๆ เช่น ประเภทแง่มุมคุณภาพสินค้า(Quality, Q) ที่ใช้ได้จริง ประเภทแง่มุมด้านราคา (Price, P) เหมาะสมกับคุณภาพ และประเภทแง่มุมการให้บริการ (Service, S) โดยเฉพาะการจัดส่งสินค้าที่มีประสิทธิภาพ ทั้งสามประเภทแง่มุมนี้ส่วนใหญ่ปรากฏอย่างกระจัดกระจายอยู่ในบทวิจารณ์ดังกล่าว ฉะนั้นการทำเหมืองความคิดเห็นตามแง่มุมเป็นประโยชน์ต่อผู้บริโภคทั่วไป (คือ ช่วยผู้บริโภคตัดสินใจได้ดีขึ้น) และเจ้าของธุรกิจออนไลน์ กล่าวคือช่วยเจ้าของธุรกิจพัฒนาสินค้า บริการ และกลยุทธ์ได้อย่างมีประสิทธิภาพ นอกจากนี้บทวิจารณ์ หรือเอกสารข้อความวิจารณ์จะแสดงออกมาในรูปแบบหน่วยวาทกรรมพื้นฐาน (Elementary Discourse Unit, EDU, คือประโยคความเดียว [1]) ดังตัวอย่างต่อไปนี้

สินค้า: '*ฝักบัว*'; *บทวิจารณ์*:
EDU1: "*แพคของแย่มาก*"
EDU2: "*กล่องบุบและเทะ*"
EDU3: "*แต่สินค้ายังสภาพดี*"
EDU4: "*ฝากไว้ให้คิดแค่นี้*"

การแสดงออกของ EDU อิงตามหลักภาษาไทยทั่วไป (General Thai Linguistic Expression) ดังต่อไปนี้ หลังจากได้ทำการหาแก่นคำ (stemming words) และกำจัดคำหยุด (stop words)

EDU → NP1 VP | VP
VP → Verb NP2 | Verb adv | Verb
Verb → $Verb_{weak}$ Wrd | $Verb_{strong}$
NP1 → Noun | Noun Adjphrase
NP2 → Noun | Noun Adjphrase | Noun Prepphrase
Wrd → Noun| Adj| Adv| $Verb_{strong}$
$Verb_{weak}$ → {'*เป็น/be*','*ไม่เป็น/notBe*','*มี/have*','*ไม่มี/notHave*','*ใช้/use*', ,'*รู้สึก /feel*'}
$Verb_{strong}$ → { '*เกิด*/appear', '*ไม่ตรง/beIrrelevant*', '*ไม่เสถียร/unstable*', '*ซับซ้อน/beComplex*', '*หมดอายุ/expire*''ค้าง/freeze, 'ช้า/beslow','ลด/reduce', 'รวดเร็ว/fast', 'ยาก/be difficult', 'รอ/wait' ..}
Noun → {'*ปัญหา/Problem*' '*เวลา/ time*',…}

อย่างไรก็ตามการทำเหมืองความคิดเห็นตามแง่มุมด้วยการทำเมืองข้อความของบทวิจารณ์ภาษาไทยในงานวิจัยนี้มีปัญหาหลัก 2 ประการดังนี้

1. การสกัด EDU ที่มีแนวความคิดแง่มุม (Aspect Concept)

2. การจำแนกประเภทแนวความคิดแง่มุมออกเป็นคลาสต่าง ๆ 6 คลาส (3 คลาสเชิงบวก และ 3 คลาสเชิงลบ) คือ แนวความคิดคุณภาพสินค้า (Quality Concept, QC) เชิงบวก QC+ เชิงลบ QC- แนวความคิดราคา (Price Concept, PC) เชิงบวก PC+ เชิงลบ PC- และแนวความคิดการให้บริการ (Service Concept, SC) เชิงบวก SC+ เชิงลบ SC- เนื่องจากบทวิจารณ์ภาษาไทยบนสื่อสังคมออนไลน์มีลักษณะเป็นบทความสั้น ที่ประกอบด้วย EDU ที่เป็นข้อความสั้น ๆ และมักจะละคำนาม ดังนั้นงานวิจัยนี้จึงเสนอใช้การปรากฏร่วมกันของกริยา และคำขยาย (verb-modifierWord Co-occurrence, *v-mw*) มาสกัด EDU ที่มีแนวความคิดแง่มุม โดย $v \in$ Verb ซึ่งคือเซตกริยาหลักของ EDU และคำขยาย (modifierWord, *mw*) ที่ตามหลังVerb ($mw \in$ Word ซึ่งคือเซตคำขยายที่ตามหลัง Verb; Word = Adj$\cup$Adv$\cup$Noun )

นอกจากนี่เราได้ประยุกต์ใช้การเรียนรู้ของเครื่องที่มีผู้เชี่ยวชาญกำกับ (Supervised Machine Learning) ที่เป็นเนอีฟเบย์ (Naïve Bayes) [2] มาหา จำแนกประเภทแนวความคิดแง่มุม ออกเป็นด้านต่าง ๆ 6 คลาส จากการปรากฏร่วมกันของฟีเจอร์กริยา (Verb Feature) และ ฟีเจอร์คำ (Word Feature)

ในส่วนที่ 2 สรุปงานวิจัยที่เกี่ยวข้อง ส่วนที่ 3 แสดงภาพรวมของการทำเหมืองความคิดเห็นเชิงแง่มุมจากบทวิจารณ์ของผู้ใช้ในธุรกิจออนไลน์ส่วนที่ 4 ประเมินและอภิปรายวิธีการที่เสนอ และ 5 สรุปผลงานวิจัย

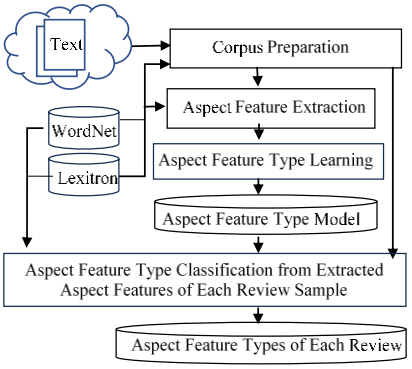
## 2. งานวิจัยที่เกี่ยวข้อง

มีงานวิจัยด้านการทำเหมืองความคิดเห็นเชิงแง่มุม (Aspect-Based Opinion Mining) จากข้อความในเอกสารจำนวนมาก [3]-[6] พบว่ามีการเปลี่ยนผ่านจากเทคนิคดั้งเดิมไปสู่การเรียนรู้เชิงลึก โดย **Jeyapriya and Kanimozhi Selvi (2015) [4]** เริ่มต้นจากการแก้ปัญหาความหยาบของการวิเคราะห์ระดับเอกสารด้วยการใช้ Frequent Itemset Mining ร่วมกับ Supervised Learning ในการสกัดแง่มุมระดับวลี ต่อมา **Dasgupta and Sen (2022) [6]** ได้ยกระดับความแม่นยำด้วยการใช้ Transfer Learning ผ่านโมเดล BERT และ RoBERTa ซึ่งช่วยแก้ปัญหาข้อมูลเทรนที่มีจำกัดบน Amazon ได้อย่างมีประสิทธิภาพ นอกจากนี้ ยังมีการขยายขอบเขตไปสู่การวิเคราะห์ความคิดเห็นแบบแฝง (Implicit Opinions) โดย **Setiowati et al. (2022) [5]** ที่ใช้อัลกอริทึม Opinion Co-occurrence หาความสัมพันธ์ของคำในรีวิวภาษาอินโดนีเซียเพื่อระบุทัศนคติที่ซ่อนอยู่ในข้อเท็จจริง ขณะที่ในมุมมองเชิงกลยุทธ์ **Wang and Chen (2015) [3]** ได้ประยุกต์ใช้ทฤษฎีเกม (Game Theory) ในห่วงโซ่อุปทานอาหารเพื่อจัดการความเสี่ยงด้านคุณภาพ ซึ่งแสดงให้เห็นถึงความสำคัญของการนำข้อมูลไปใช้ตัดสินใจทางธุรกิจ อย่างไรก็ตาม งานวิจัยเหล่านี้ยังมี **ช่องว่าง (Research Gap)** สำคัญ คือส่วนใหญ่เน้นภาษาอังกฤษหรืออินโดนีเซียที่ไม่สามารถรองรับโครงสร้างไวยากรณ์และบริบทคำกำกวมของภาษาไทยได้ อีกทั้งยังขาดการเชื่อมโยงเทคนิค NLP ขั้นสูงเข้ากับโจทย์ธุรกิจ E-Commerce ของไทยโดยตรง งานวิจัยนี้จึงมุ่งเติมเต็มด้วยการพัฒนาโมเดล Hybrid (Machine Learning ร่วมกับ NLP) สำหรับภาษาไทย เพื่อสกัดแง่มุมและความรู้สึกจากเสียงของผู้บริโภคจริง ซึ่งจะช่วยยกระดับการจัดการเชิงกลยุทธ์ได้แม่นยำกว่าการใช้แบบจำลองทฤษฎีเพียงอย่างเดียว

งานวิจัย [3] - [6] ใช้ข้อมูลภาษาอังกฤษ ซึ่งไม่มีปัญหาเกี่ยวกับการละนามวลีเหมือนภาษาไทย

## 3. ภาพรวมของระบบ

การทำเหมืองความคิดเห็นเชิงแง่มุมจากบทวิจารณ์ประกอบด้วย 4 ขั้นตอนหลักคือ การเตรียมคลังข้อมูล (Corpus Preparation) การสกัด EDU ที่มีแนวความคิดแง่มุม (Aspect Concept EDU Extraction) การเรียนรู้ประเภทแนวความคิดแง่มุม (Aspect Feature Type Learning) และการจำแนกประเภทแนวความคิดแง่มุมของ EDUที่สกัดได้จากบทวิจารณ์ (ดูภาพที่ 1)




**ภาพที่ 1:** ภาพรวมของระบบ

### 3.1 การเตรียมคลังข้อมูล

ขั้นตอนนี้มีวัตถุประสงค์เพื่อเตรียมคลังข้อมูล EDU (EDU corpus) ที่ได้จากการดาวน์โหลดบทวิจารณ์ (เอกสาร) ออนไลน์ของผู้ใช้ในธุรกิจออนไลน์ ขั้นตอนดังกล่าวประกอบด้วยการใช้เครื่องมือตัดคำภาษาไทย [7] และการแบ่งหน่วย EDU (EDU Segmentation) [8] เพื่อสร้างคลังข้อมูลจำนวน 500 EDUs จากเอกสารบทวิจารณ์ดังกล่าวจำนวน 100 ฉบับ ผ่านการตัดคำให้เหลือแก่นคำแบบเชิงรูปแบบ (stemming) และการกำจัดคำหยุด (stop word removal) และคลังข้อมูลดังกล่าวถูกแบ่งออกเป็น 3 ส่วน ได้แก่

(1) ส่วนแรกจำนวน 100 EDUs ใช้เป็นคลังข้อมูล ที่มีการกำกับแนวความคิดแง่มุม (Aspect Concept) จาก ตัวกำกับ EDU (EDU Tag) (ดูภาพที่ 2) โดยแนวคิดของคำกริยาและคำขยายอ้างอิงจาก WordNet [9] หลังจากแปลจากภาษาไทยเป็นภาษาอังกฤษด้วย Lexitron [10]) เพื่อใช้ในการรวบรวมเซต Verb และเซต Word

(2) ส่วนที่สองจำนวน 300 EDUs ใช้สำหรับการเรียนรู้การจำแนกประเภทแนวความคิดแง่มุมที่ได้มีการกำกับโดยผู้เชี่ยวชาญ จากตัวกำกับกริยาวลี (VP Tag)

(3) ส่วนที่สามจำนวน 100 EDUs ใช้สำหรับการทดสอบการจำแนกประเภทแนวความคิดแง่มุมสำหรับสินค้าแต่ละชนิด

ชื่อองค์กร: 'ธุรกิจออนไลน์ xxx' สินค้า: 'ฝักบัว'
บทวิจารณ์: "*แพคของแย่มาก*"[EDU1] "*กล่องบุบเละเทะ*"[EDU2] "*แต่สินค้ายังสภาพดี*"[EDU3] "*ฝากไว้ให้คิดแค่นี้*"[EDU4]

<EDU:1 Aspect= y >< NP1 Ø='person'> Ø </NP1><VP Aspecttype: 'SC-' ; concept='packBadly'><Verb concept='pack'>*แพค*</Verb>*ของ*<modifierWord concept='badly'>*แย่มาก*</ modifierWord></VP></EDU>

<EDU:2 Aspect= y >< NP1 concept='package'> *กล่อง* </NP1><VP Aspecttype: 'SC-' ; concept='damageBadly'><Verb concept=' damage'>*บุบ* </Verb><modifierWord concept='badly'>*เละเทะ*</ modifierWord></VP></EDU>

<EDU:3 Aspect= y >*แต่*< NP1 concept='product'> *สินค้า* </NP1>< VP Aspecttype: 'QC+' ; concept='goodCondition'><Verb concept='be+condition'>*ยัง+สภาพ* </Verb><modifierWord concept='good'> *ดี*</modifierWord></VP></EDU>

<EDU:4 Aspect= n >< NP1 Ø='person'> Ø </NP1><VP Aspecttype: null ; concept= null>*ฝากไว้ให้คิดแค่นี้*</ VP ></EDU>

Where EDU = EDU Tag; NP1= Noun Phrase Tag; VP= Verb PhraseTag; Verb =Verb Tag; and modifierWord = modifierWord Tag ; and Ø represents ellipsis.

**ภาพที่ 2:** การกำกับแนวความคิดแง่มุม

### 3.2 การสกัด EDU ที่มีแนวความคิดแง่มุม

เป็นขั้นตอนการสกัด EDU ที่มีแนวความคิดแง่มุม โดยการปรากฏร่วมกันของ $v$ และ $mw$ (เมื่อ $v \in$ Verb ซึ่งเป็นเซตกริยาที่ได้จากการกำกับ ในข้อ 3.1 และ $mw \in$ Word ซึ่งเป็นเซตคำที่ได้จากการกำกับ ในข้อ 3.1) มาทำการสกัด EDU ดังกล่าว

### 3.3 การเรียนรู้ประเภทแนวความคิดแง่มุม

ขั้นตอนนี้เป็นการเรียนรู้ประเภทแนวความคิดแง่มุม Naive Bayes (NB) [2] เพื่อเรียนรู้ความน่าจะเป็นแนวความคิดแง่มุม 6 คลาส คือ คุณภาพเชิงบวก (Quality concept+, QC+, เช่น 'ใช้งานง่าย') , คุณภาพเชิงลบ (Quality concept-, QC-, เช่น 'เปิดยากมาก'), (Price concept+, PC+, เช่น 'มีราคาถูก'), (Price concept-, PC-, เช่น 'มีราคาแพง'), (Service concept+, SC+, เช่น 'จัดส่งเร็ว') และ (Service concept-, SC-, เช่น 'ไม่ห่อกันกระแทก')ของ $v$ และ $mw$ ในการปรากฏร่วมกันโดยใช้เครื่องมือ Weka [11-12] ดังแสดงในตารางที่ 1 จากคลังข้อมูลบทวิจารณ์ที่มีการกำกับคู่คำดังกล่าวสำหรับการเรียนรู้

**ตารางที่ 1:** แสดงค่าความน่าจะเป็น $v$, $mw$ ในการปรากฏร่วมกันด้วยคลาส QC+, QC- PC+, PC-, SC+,และ SC-

| $v$ ของคู่คำปรากฏร่วม | QCplus (QC+) | QCminus(QC-) |
|---|---|---|
| Look | 0.0747 | 0.0205 |
| BeInAgreement | 0.0575 | 0.0769 |
| … | … | … |
| $mw$ ของคู่คำปรากฏร่วม | QCplus(QC+) | QCminus(QC-) |
| Highly | 0.0123 | 0.0057 |
| Fine | 0.1429 | 0.0686 |
| … | … | … |
| $v$ ของคู่คำปรากฏร่วม | PCplus(PC+) | PCminus(PC-) |
| HavePrice | 0.2049 | 0.1075 |
| BeWorth | 0.0574 | 0.0108 |
| … | … | … |
| $mw$ ของคู่คำปรากฏร่วม | PCplus(PC+) | PCminus(PC-) |
| Cheap | 0.2353 | 0.0107 |
| Expensive | 0.1369 | 0.1096 |
| … | … | … |
| $v$ ของคู่คำปรากฏร่วม | SCplus(SC+) | SCminus(SC-) |
| Send | 0.2518 | 0.0777 |
| Pack | 0.1294 | 0.0388 |
| … | … | … |
| $mw$ ของคู่คำปรากฏร่วม | SCplus(SC+) | SCminus(SC-) |
| Quickly | 0.2521 | 0.0964 |
| Fine | 0.1933 | 0.0964 |
| … | … | … |

### 3.4 การจำแนกประเภทแนวความคิดแง่มุมของ EDU

ที่สกัดจากบทวิจารณ์เป็นขั้นตอนการจำแนกประเภทแนวความคิดแง่มุมของ EDU ที่ได้จากการสกัดโดยใช้ การปรากฏร่วมกัน ของ $v$ และ $mw$ จาก เซต Verb และ เซต Word ตามลำดับกับข้อมูลทดสอบ ต่อมา คู่คำที่ปรากฏร่วมกันนั้น ($v$-mw) จำแนกประเภทแนวความคิดแง่มุมด้วย NB ในสมการ(1) กับค่าความน่าจะเป็น $v$ และ $mw$ ที่ได้จากการเรียนรู้ ประเภทความสัมพันธ์เป็นแง่มุมประเภทต่าง ๆ ดังนี้ 'QCplus' , 'QCminus', 'PCplus', 'PCminus', 'SCplus',หรือ'SCminus' (ตารางที่ 1) ด้วย การตรวจสอบแบบไขว้ 10 เท่า (ten-fold cross validation)

$$AspectFeatureType = \underset{class \in Class}{\operatorname{argmax}} P(class|\, v, mw) = \underset{class \in Class}{\operatorname{argmax}} argmax P(v|class)P(mw|class)P(class) \quad (1)$$

เมื่อ $v \in$ Verb ของ EDU ; $mw \in$ Word; Class= {'QCplus' , 'QCminus', 'PCplus', 'PCminus', 'SCplus', 'SCminus'}

## 4. การประเมินผล และอภิปรายผล

งานวิจัยที่เสนอมีการประเมินผล 2 ส่วนกับข้อมูลทดสอบ EDU จำนวน 100 EDUs โดยผู้เชี่ยวชาญ 3 ท่าน ผ่านการลงคะแนนแบบชนะสูงสุด (Max win Voting): การประเมินส่วนแรกคือการระบุ EDU ที่มีแนวความคิดแง่มุม โดยใช้ *v-mw ได้ F1- Score = 0.842* การประเมินส่วนที่สองคือ การจำแนกประเภทแนวความคิดแง่มุมสำหรับสินค้าแต่ละชนิด การประเมินทั้งสองส่วนอิงตามค่าความแม่นยำ (Precision,PC) ค่าระลึก (Recall,RC) และ F1-Score ดังแสดงในตารางที่ 2

**ตารางที่ 2:** ผลการประเมิน

| EDU ที่มีแนวความคิดแง่มุม | Precision (PC) | Recall (RC) | F1-Score |
|---|---|---|---|
| QC+ | 0.81 | 0.835 | 0.822 |
| QC- | 0.739 | 0.882 | 0.803 |
| PC+ | 0.967 | 0.644 | 0.773 |
| PC- | 0.714 | 0.625 | 0.666 |
| SC+ | 0.746 | 0.855 | 0.796 |
| SC- | 0.625 | 0.192 | 0.294 |
| ค่าเฉลี่ยถ่วงน้ำหนัก | 0.778 | 0.775 | 0.777 |

จากตารางที่ 2 ได้ค่าเฉลี่ย F1-Score = 0.692 ในขณะที่งานวิจัย A Framework of Customer Review Analysis Using the Aspect-Based Opinion Mining Approach [9] ใช้ RoBERTa จำแนก ความคิดเห็น (Opinion) และ แง่มุม (Aspect) จากข้อมูลวิจารณ์บน Amazon 150 ตัวอย่าง/เอกสารได้ F1-Socre 0.6 อย่างไรก็ตามการใช้โมเดลในกลุ่ม Transformer เช่น RoBERTa จะมีศักยภาพสูง แต่โดยทั่วไปแล้วโมเดลดังกล่าวมักเหมาะกับข้อมูลในปริมาณมาก (มากกว่า 1,000 ตัวอย่างขึ้นไป) ในกรณีที่ข้อมูลมีไม่มาก โมเดลอย่าง Naïve Bayes จึงมีความเหมาะสม

## 5. สรุปผลงานวิจัย

โมเดล Naive Bayes สามารถจำแนกข้อมูลกลุ่ม PC+ เป็นกลุ่มที่มีความแม่นยำสูงสุดที่ 0.967 เป็นเพราะคุณลักษณะ (Attributes) ของคำในกลุ่มนี้มีความชัดเจนสูง แต่กลุ่ม SC- มีค่า Recall ต่ำเพียง 0.192 เกิดจากจำนวนข้อมูล (Instances) ในกลุ่มนี้มีน้อยเกินไป หรือมีความคล้ายคลึงกับกลุ่ม QC- จนโมเดลสับสน ตัวอย่างเช่น ส่งสินค้าชำรุดมาให้, เสียเวลารอ แม้ว่าโมเดลจะมีความสามารถในการจำแนกข้อมูลได้ดีในบางกลุ่มที่มีลักษณะเด่นชัด แต่ยังคงมีข้อจำกัดในกลุ่มที่ข้อมูลมีจำนวนน้อยหรือมีความคล้ายคลึงกับกลุ่มอื่น ซึ่งควรได้รับการปรับปรุงในอนาคต เช่น การเพิ่มปริมาณข้อมูล หรือการปรับปรุงคุณลักษณะให้มีความแตกต่างมากยิ่งขึ้น

## เอกสารอ้างอิง


[1] Carlson L, Marcu D, Okurowski M E. Building a Discourse-Tagged Corpus in the Framework of Rhetorical Structure The-ory. *Current Directions in Discourse and Dialogue*, van Kup-pevelt J, Smith R(eds.),KluewerAcademicPublishers,2003,pp.85-112.

[2] Mitchell T M. Machine Learning. The McGraw-Hill Compa-nies Inc. and MIT Press, Singapore, 1997.

[3] Wang F and Chen L. Review Mining for Estimating Users' Ratings and Weights for Product Aspects. *Web Intelligence* , 2015, Vol.13, No.3, pp137–152.

[4] Jeyapriya A and Kanimozhi Selvi C S. Extracting Aspects and Mining Opinions in Product Reviews using Supervised Learning Algorithm. 2nd International Conference on Electronics and Communication Systems (ICECS) , India, 2015, pp 548-552

[5] Setiowati Y, Djunaidy A, Siahaan D O. Aspect-based Extraction of Implicit Opinions using Opinion Co-occurrence Algorithm. 2022 5th International Seminar on Research of Information Technology and Intelligent Sys (ISRITI),Indonesia,2022,pp.781-786.

[6] Dasgupta S and Sen J. A Framework of Customer Review Analysis Using the Aspect-Based Opinion Mining Approach. 2022 OITS Inter. Conf. on Infor. Tech. (OCIT), India, 2022,pp. 233-237.

[7] Sudprasert S, Kawtrakul A. Thai word segmentation based on global and local unsupervised learning. In *Proc. NCSEC 2003*, Chonburi, Thailand, 2003, pp.1-8.

[8] Chareonsuk J, Sukvakree T, Kawtrakul A. Elementary discourse unit segmentation for Thai using discourse cue and syntactic information. In Proc. NCSEC 2005, Thailand, Oct. 27-28, 2005, pp.85-90.

[9] Miller G.A. WordNet: a lexical database for English Commun. *ACM*, 38 (11) (1995), pp. 39-41

[10] Trakultaweekoon K, Porkaew P, Supnithi T. (2007). "LEXiTRON vocabulary suggestion system with recommendation and vote mechanism," in 7th International Symposium on Natural Language Processing (SNLP) 2007, Pattaya, Thailand, 43–48.

[11] R.R. Bouckaert, et al.WEKA Manual for Version 3-9-1 Univ. Waikato Hamilt. N. Z. (2016), pp. 1-341

[12] Hall M, Frank E, Holmes G, Pfahringer B, Reutemann P, Witten I.H. The WEKA data mining software: an update ACM SIGKDD Explor. *Newsl*., 11 (1) (2009), pp. 10-18

# การศึกษาเปรียบเทียบวิธีการเรียนรู้ของเครื่องจักรแบบควอนตัม-อินสไปร์และคลาสสิกสำหรับการทำนายพฤติกรรมการกลับมาของลูกค้าร้านอาหารเวลากลางคืน

# A Comparative Study of Quantum-inspired and Classical Machine Learning for Customer Return Prediction in Nighttime Restaurants

*ธัญวัฒน์ แก้วประดิษฐ์ (Tanyavat Kaewpradit) และวราภรณ์ วิยานนท์ (Waraporn Viyanon)*
*ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร*
*tanyavat.kaewpradit@g.swu.ac.th, waraporn@g.swu.ac.th*

## บทคัดย่อ

*งานวิจัยนี้เปรียบเทียบประสิทธิภาพโมเดล Machine Learning แบบ Quantum-inspired (QiSVM และ QiNN) กับแบบคลาสสิก (XGBoost และ Random Forest) ในการทำนายการกลับมาใช้บริการของลูกค้าร้านอาหารกลางคืน โดยใช้ข้อมูล 6,025 รายจากกรุงเทพฯ ผลการศึกษาพบว่า Random Forest ให้ประสิทธิภาพสูงสุด (Accuracy 0.9207, Recall 0.9844, F1-Score 0.9584) รองลงมาคือ XGBoost ที่มี Precision สูงสุด (0.9415) สำหรับโมเดล Quantum-inspired พบว่า QiSVM แสดง Precision สูง (0.9624) ใกล้เคียงโมเดลคลาสสิก แต่ Accuracy และ Recall ต่ำ (0.4232, 0.3936) ส่วน QiNN ให้ผลสมดุล (F1-Score 0.8840) แต่ยังต้องปรับปรุง แม้โมเดลคลาสสิกจะมีประสิทธิภาพโดยรวมสูงกว่า แต่งานวิจัยยืนยันความเป็นไปได้ของแนวทาง Quantum-inspired บนคอมพิวเตอร์ปกติและศักยภาพของ QiSVM เป็นทางเลือกสำหรับวิเคราะห์พฤติกรรมลูกค้าในธุรกิจร้านอาหาร*

**คำสำคัญ:** การเรียนรู้ของเครื่องแบบควอนตัม-อินสไปร์ การเรียนรู้ของเครื่องแบบคลาสสิก การทำนายการกลับมาของลูกค้า การวิเคราะห์ร้านอาหาร

## Abstract

*This research compares the performance of Quantum-inspired Machine Learning models (QiSVM and QiNN) with classical models (XGBoost and Random Forest) for predicting customer return intention in nighttime restaurants using 6,025 records from Bangkok. The results show that Random Forest achieved the highest performance (Accuracy 0.9207, Recall 0.9844, F1-Score 0.9584), followed by XGBoost with the highest Precision (0.9415). For Quantum-inspired models, QiSVM demonstrated high Precision (0.9624) comparable to classical models, but low Accuracy and Recall (0.4232, 0.3936), while QiNN showed balanced results (F1-Score 0.8840) but requires further improvement. Although classical models had slightly better overall performance, the research validates the feasibility of Quantum-inspired approaches on conventional computers and confirms QiSVM's potential as an alternative for customer behavior analysis in the restaurant business.*

**Keywords:** Quantum-inspired Machine Learning, Classical Machine Learning, Customer Return Prediction, Restaurant Analytics.

## 1. บทนำ

ในยุคดิจิทัล การใช้ข้อมูลเพื่อทำความเข้าใจพฤติกรรมลูกค้ากลายเป็นปัจจัยสำคัญต่อความสำเร็จของธุรกิจร้านอาหาร [1] โดยเฉพาะร้านอาหารกลางคืนที่มีลักษณะการดำเนินงานและกลุ่มลูกค้าเฉพาะ การพยากรณ์การกลับมาใช้บริการของลูกค้าจึงมีความสำคัญยิ่ง เนื่องจากต้นทุนการหาลูกค้าใหม่สูงกว่าการรักษาลูกค้าเดิมถึง 5-25 เท่า [2] การวิเคราะห์พฤติกรรมช่วยระบุลูกค้าที่มีความเสี่ยงและออกแบบกลยุทธ์รักษาลูกค้าได้อย่างมีประสิทธิภาพ [3]

ข้อมูลพฤติกรรมลูกค้ามีลักษณะซับซ้อนที่เหมาะกับเทคโนโลยีควอนตัม-อินสไปร์ด คือ ความสัมพันธ์แบบไม่เป็นเชิงเส้น การพัวพันกันของตัวแปร (entanglement) และสถานะที่ทับซ้อน (superposition) การใช้ ZZ Feature Map แปลงข้อมูลเข้าสู่ Hilbert space มิติสูง ทำให้ quantum kernel จับความสัมพันธ์ซับซ้อนได้ดีกว่า แม้วิธีการคลาสสิกอย่าง XGBoost และ Random Forest มีประสิทธิภาพดี แต่ยังมีข้อจำกัดในการค้นพบความสัมพันธ์เชิงลึก [4] เมื่อเทคโนโลยีควอนตัมคอมพิวติ้งก้าวหน้า อัลกอริทึมแบบควอนตัม-อินสไปร์เสนอการประมวลผลที่เร็วกว่าและแม่นยำกว่า [5]

งานวิจัยนี้มีวัตถุประสงค์เพื่อ (1) สร้างและเปรียบเทียบโมเดลทำนายการกลับมาของลูกค้าด้วยเทคนิคคลาสสิกและควอนตัม-อินสไปร์ด (2) ประเมินประสิทธิภาพด้วย 5-fold cross-validation และมาตรวัด 6 ตัวที่เหมาะกับข้อมูลไม่สมดุล และ (3) พัฒนากรอบการตัดสินใจเลือกใช้โมเดลตามบริบทธุรกิจ เพื่อเป็นแนวทางพัฒนาเครื่องมือพยากรณ์สำหรับร้านอาหารกลางคืน และเตรียมพร้อมสำหรับการนำเทคโนโลยีควอนตัมมาใช้ในอนาคต [6]

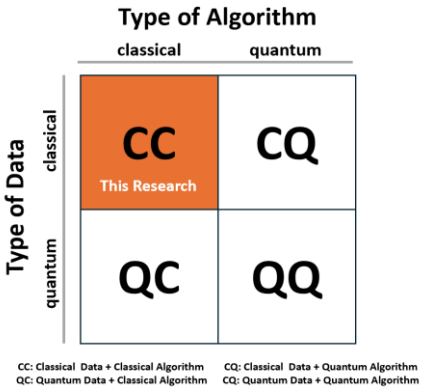
## 2. งานวิจัยที่เกี่ยวข้อง

การพัฒนาระบบวิเคราะห์ร้านอาหารด้วย AI โดย EDA ให้ประสิทธิภาพดีกว่า regression [1] งานวิจัยนี้เน้นการวิเคราะห์รีวิวและพฤติกรรมระหว่างใช้บริการ แต่ยังขาดโมเดลทำนายการกลับมาใช้บริการ

ในพยากรณ์ความพึงพอใจลูกค้าอีคอมเมิร์ซพบว่า LSTM ให้ความแม่นยำมากที่สุด รองลงมาคือ XGBoost และ Naive [3] ส่วนการพยากรณ์เวลาจัดส่ง Random Forest, XGBoost ให้ความแม่นยำสูงที่สุด [7] งานวิจัยเหล่านี้ใช้ classical ML เท่านั้น ยังไม่มีการเปรียบเทียบกับ quantum-inspired ในบริบทธุรกิจบริการ

การเรียบรู้ของเครื่องบนควอนตัมคอมพิวติ้งเน้นอัลกอริทึม Grover's และ Shor's แสดงความเหนือกว่าของควอนตัมในปัญหาเฉพาะ [4] และมีการอธิบาย (QML) ที่ใช้ qubits มี superposition และ entanglement แตกต่างจาก classical bits [5] งานวิจัยเหล่านี้เน้นทฤษฎีและหลักการ ยังขาดการประยุกต์ใช้จริงในบริบทธุรกิจ งานวิจัยนี้จึงนำหลักการ quantum-inspired มาแก้ปัญหาทางธุรกิจโดยไม่ต้องใช้ quantum hardware ที่มีราคาแพงและมีข้อจำกัดทางเทคนิค

QML Framework Matrix [6] แบ่งออกเป็น 4 ประเภท: CC (Classical Data + Algorithm), CQ (Classical Data + Quantum Algorithm), QC (Quantum Data + Classical Algorithm) และ QQ (Quantum Data + Quantum Algorithm) การศึกษาส่วนใหญ่เน้น CC และ CQ เพราะใช้งานได้จริง ประยุกต์ใช้จำแนกข่าว จดจำตัวเลขลายมือ ตรวจจับมะเร็ง พยากรณ์สภาพอากาศ [8] แผ่นดินไหว [9] และจำแนกข้อมูลมิติสูง [10] งานวิจัยเหล่านี้เน้นปัญหาทางวิทยาศาสตร์และวิศวกรรม ยังขาดการประยุกต์ใช้กับพฤติกรรมลูกค้าในธุรกิจบริการ




**ภาพที่ 1:** QML Framework Matrix [6]

QSVM ใช้ quantum kernel functions แมปข้อมูลคลาสสิกไปยัง quantum-enhanced feature spaces ผ่าน Pauli, Z และ ZZ feature mapping [6] มีการเปรียบเทียบโมเดลควอนตัมและคลาสสิกในพยากรณ์โรคไตเรื้อรัง พบว่า QSVM มีประสิทธิภาพสูงสุด เทียบกับ QNN, SVM และ NN [11] งานวิจัยเหล่านี้ส่วนใหญ่ดำเนินการในสาขาการแพทย์และวิทยาศาสตร์

Classical ML ยังเหนือกว่า quantum ML ในเรื่อง Time Series เพราะ quantum ML อยู่ในระยะพัฒนาเริ่มต้น ขณะที่ classical ML ได้รับประโยชน์จากการวิจัยมานานกว่า [12] งานวิจัยเน้นศักยภาพ quantum neural network

ในประมวลผลแบบขนาน จัดการข้อมูลขนาดใหญ่ซับซ้อนและจัดการสัญญาณรบกวนดีกว่าระบบคลาสสิก [13]

Quantum-inspired ML (QiML) ประยุกต์หลักการควอนตัมบนคอมพิวเตอร์คลาสสิกโดยไม่ต้องใช้ฮาร์ดแวร์ควอนตัม โดยที่ QiBoost มีความแม่นยำเทียบเท่า QBoost บน D-Wave ($R^2$ 0.9147 vs 0.9116) แต่ประหยัดพลังงาน (64.7 μJ vs 1.124 kWh) และเวลาฝึก (6.47 vs 252.88 วินาที) [14] ในส่วนของ QSVC ใช้ ZZFeatureMap จำแนกภาพได้ 93-98% เทียบกับ Classical SVM 83-95% [15] งานวิจัยเหล่านี้เน้นปัญหาทางเทคนิค ยังไม่มีการศึกษาในบริบทพฤติกรรมลูกค้า งานวิจัยนี้จึงเป็นหนึ่งในงานแรกที่นำ QiSVM และ QiNN มาทำนายพฤติกรรมลูกค้าในธุรกิจบริการ โดยใช้ข้อมูลจริงจากประเทศไทยและพัฒนากรอบการตัดสินใจที่นำไปประยุกต์ใช้ได้ในทางปฏิบัติ

## 3. วิธีการดำเนินการวิจัย

### 3.1 การเก็บรวบรวมข้อมูล

ข้อมูลเก็บรวบรวมจากร้านอาหารกลางคืนในกรุงเทพฯ ที่เปิดให้บริการเวลา 19:00-02:00 น. ระยะเวลา 3 เดือน (กันยายน-พฤศจิกายน 2025) ชุดข้อมูลประกอบด้วยบันทึกลูกค้า 6,025 รายการจากแบบสอบถามออนไลน์โดยได้รับความยินยอม

ตัวแปรอิสระ:

(1) ข้อมูลประชากร (เพศ อายุ สถานภาพความสัมพันธ์ อาชีพ รายได้) – มีอิทธิพลต่อพฤติกรรมการบริโภคและความสามารถในการจ่าย [1]

(2) รูปแบบพฤติกรรม (บุคคลที่ร่วมรับประทานอาหาร ความสนใจข้อมูล) - พฤติกรรมในอดีตเป็นตัวทำนายพฤติกรรมในอนาคต [2]

(3) ความพึงพอใจ (คุณภาพเครื่องดื่ม คุณภาพการบริการ บรรยากาศ ความคุ้มค่า ความสะอาด) - สำหรับร้านอาหารกลางคืน คุณภาพเครื่องดื่มและบรรยากาศมีความสำคัญเป็นพิเศษเพราะลูกค้ามาเพื่อผ่อนคลายและสังสรรค์ [1], [2] ความพึงพอใจเป็นปัจจัยหลักต่อการกลับมาใช้บริการ [3]

ตัวแปรตาม: ความตั้งใจกลับมาใช้บริการ (แบบทวิภาค: กลับมา/ไม่กลับมา)

การเลือกตัวแปรเหล่านี้สอดคล้องกับ Service Quality Model และ Customer Satisfaction Theory ที่ระบุว่าความพึงพอใจหลายมิติส่งผลต่อความตั้งใจกลับมาใช้บริการ การรวมตัวแปรทั้งด้านประชากร พฤติกรรม และความพึงพอใจช่วยให้โมเดลจับภาพความซับซ้อนของพฤติกรรมลูกค้าได้ครอบคลุม ซึ่งเหมาะกับเทคนิค Machine Learning ที่จัดการความสัมพันธ์แบบไม่เป็นเชิงเส้นได้ดี

### 3.2 การประมวลผลข้อมูลเบื้องต้น

ดำเนินการ 8 ขั้นตอน: (1) Data Cleaning – จัดการค่าผิดปกติและค่าขาดหาย (2) Target Variable Creation - แปลงเป็นตัวแปรทวิภาค (3) Feature Engineering and Transformation - ใช้ One-Hot Encoding สำหรับตัวแปรเชิงหมวดหมู่และ StandardScaler สำหรับตัวแปรตัวเลข (4) Data Splitting - 85:15 (Training: Test) ด้วย stratified split (5) Feature Importance Analysis - ใช้ Mutual Information และ Tree-based methods คัดเลือกตัวแปรสำคัญ (6) Imbalanced Data Handling - ใช้ SMOTE (k=5, ratio 1:1) กับ training set เท่านั้น (7) Create Data Pipeline - สร้าง automated pipeline ที่รวมทุกขั้นตอน (8) Hyperparameter Tuning - ใช้ RandomizedSearchCV (100 iterations) ร่วมกับ 5-Fold Cross-Validation

### 3.3 การกำหนดพารามิเตอร์โมเดล

XGBoost: ค้นหาพารามิเตอร์จาก n_estimators, max depth, learning rate, subsample, min child weight, gamma, reg alpha, reg lambda

Random Forest: ค้นหาพารามิเตอร์จาก n_estimators, max_depth, min samples split, min samples leaf, max features, criterion

QiSVM: ใช้ ZZ Feature Map แมปข้อมูลไป quantum-enhanced feature space คำนวณ quantum kernel บน classical computer ค้นหา C

QiNN: ค้นหา batch size, epocs, learning rate, dropout

### 3.4 การประเมินประสิทธิภาพ

ใช้ 5-fold stratified cross-validation บน training set และทดสอบบน test set ด้วยตัวชี้วัด 6 ตัวที่เหมาะกับข้อมูลไม่สมดุล: Accuracy, Precision, Recall, F1-Score, AUC-ROC, Average Precision (AP) คำนวณค่าเฉลี่ยและ confidence interval 95% วิเคราะห์ confusion matrix, ROC curve และ Precision-Recall curve

## 4. ผลการวิจัย

### 4.1 การวิเคราะห์ประสิทธิภาพตามตัวชี้วัด

**ตารางที่ 1:** การเปรียบเทียบประสิทธิภาพโมเดล

| Model | Accuracy | Precision | Recall | F1-Score | AUC-ROC | AP (PR) |
|---|---|---|---|---|---|---|
| **XGBoost** | 0.8345 | 0.9415 | 0.8760 | 0.9076 | 0.6742 | 0.9601 |
| **Random Forest** | 0.9207 | 0.9337 | 0.9844 | 0.9584 | 0.6839 | 0.9641 |
| **QiSVM** | 0.4232 | 0.9624 | 0.3936 | 0.5586 | 0.6181 | 0.9572 |
| **QiNN** | 0.8345 | 0.9200 | 0.8500 | 0.8840 | 0.6345 | 0.9550 |

Random Forest มีประสิทธิภาพโดยรวมสูงสุด มีจุดที่โดดเด่นด้วย Accuracy สูงสุด (0.9207), Recall สูงสุด (0.9844), F1-Score สูงสุด (0.9584), และ AUC-ROC สูงสุด (0.6839) แสดงความสามารถในการระบุลูกค้าที่จะกลับมาได้ครบถ้วนที่สุด (98.44%) พร้อมรักษาความสมดุลระหว่าง Precision และ Recall ได้ดีเยี่ยม

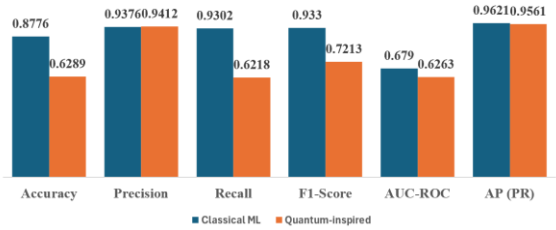
XGBoost ให้ผลลพธ์ที่ดีรองลงมา มีจุดที่โดดเด่นด้วย Precision สูงสุด (0.9415) และ AP(PR) สูงสุด (0.9601) แสดงความแม่นยำสูงในการทำนายลูกค้าที่จะกลับมา เหมาะสำหรับสถานการณ์ที่ต้องการลดการทำนายผิดพลาดกลุ่มบวก (False Positive) แต่มี Recall ต่ำกว่า (0.8760) จึงอาจพลาดลูกค้าบางส่วนที่จะกลับมา

QiSVM ได้ค่า Precision (0.9624) ใกล้เคียงกับ XGBoost และ AP(PR) (0.9572) แสดงความสามารถของ quantum-enhanced feature space ในการจับความสัมพันธ์ซับซ้อน อย่างไรก็ตามกลับได้ค่า Accuracy และ Recall ที่ต่ำมาก (0.4232), (0.3936) ส่งผลให้ F1-Score อยู่ที่ 0.5586 เท่านั้น สะท้อนปัญหาความไม่สมดุลในการจำแนกคลาส

QiNN ให้ค่า Accuracy (0.8345), Precision (0.9200), Recall (0.8500), และ F1-Score (0.8840) อยู่ในระดับดี แต่ยังต่ำกว่า Random Forest และ XGBoost ในทุกตัวชี้วัด แสดงว่ายังต้องมีการปรับปรุงเพิ่มเติม

### 4.2 การเปรียบเทียบประสิทธิภาพตามหมวดหมู่

Classical ML เหนือกว่า Quantum-inspired ทุกตัวชี้วัด ได้แก่ Accuracy (0.8776 vs 0.6289), Precision (0.9376 vs 0.9412), Recall (0.9302 vs 0.6218), F1-Score (0.933 vs 0.7213), AUC-ROC (0.679 vs 0.6263), และ AP(PR) (0.9621 vs 0.9561) โดยเฉพาะ Recall และ F1-Score ที่แตกต่างอย่างมีนัยสำคัญ ขณะที่ Precision และ AP(PR) ใกล้เคียงกัน




**ภาพที่ 2:** ประสิทธิภาพเฉลี่ยตามหมวดหมู่

### 4.3 ข้อเสนอแนะการเลือกใช้โมเดล

จากผลการทดสอบ แนะนำให้ใช้ Random Forest เนื่องจาก: (1) ประสิทธิภาพโดยรวมสูงสุดทุกตัวชี้วัดหลัก (2) Recall สูงสุด (0.9844) ระบุลูกค้าที่จะกลับมาได้ครบถ้วน ลดความเสี่ยงจากการพลาดลูกค้าที่มีศักยภาพซึ่งมีต้นทุนสูงกว่าการทำนายผิดพลาด (3) F1-Score สูงสุด (0.9584) แสดงความสมดุลดีเยี่ยม (4) เหมาะกับบริบทธุรกิจที่ต้องการระบุลูกค้าครบถ้วนเพื่อออกแบบกลยุทธ์รักษาลูกค้า

โมเดลทางเลือก: XGBoost เหมาะสำหรับแคมเปญที่มีงบจำกัดต้องการ Precision สูง (0.9415), QiSVM เหมาะสำหรับวิเคราะห์เชิงลึกรูปแบบซับซ้อนด้วย Precision สูง (0.9624) และสาธิตความเป็นไปได้ของเทคนิค Quantum-inspired บนคอมพิวเตอร์ปกติ, QiNN เป็นทางเลือกสำรองที่ควรปรับปรุงสถาปัตยกรรมก่อนนำไปใช้

## 5. อภิปรายผล และสรุปงานวิจัย

### 5.1 อภิปรายผล

ผลการวิจัยพบว่า Random Forest และ XGBoost ให้ความแม่นยำสูงในการพยากรณ์ธุรกิจบริการสอดคล้องกับงานวิจัยก่อนหน้า [3] [7] โดย Precision สูงของ QiSVM (0.9624) ยืนยันความสามารถของ quantum-enhanced feature space [11] [15] แต่ Accuracy และ Recall ที่ต่ำมาก (0.4232, 0.3936) แตกต่างจากงานก่อนหน้า เนื่องจากข้อมูลพฤติกรรมลูกค้ามีความซับซ้อนและสัญญาณรบกวนสูงกว่าข้อมูลการแพทย์หรือภาพ รวมถึง QiSVM มีความไวต่อข้อมูลไม่สมดุลมากกว่าโมเดลคลาสสิก สำหรับ QiNN ที่ให้ผลต่ำกว่าคาด (F1-Score 0.8840) สอดคล้องกับงานวิจัยก่อนหน้า [12] ที่ระบุว่า Classical ML ยังเหนือกว่า Quantum ML เนื่องจาก Quantum ML อยู่ในระยะพัฒนาเริ่มต้น สาเหตุหลักประกอบด้วย: สถาปัตยกรรมยังไม่เหมาะสม [13], ข้อจำกัดการจำลองบน classical computer ไม่สามารถใช้ true quantum parallelism ได้เต็มที่ [14] แม้ผลโดยรวมของ Classical ML จะสูงกว่า แต่ AP(PR) ของ QiSVM (0.9572) ใกล้เคียง Classical ML (0.9621) แสดงความเป็นไปได้ของแนวทาง Quantum-inspired สำหรับการวิเคราะห์เฉพาะกลุ่มที่ต้องการความแม่นยำสูง สอดคล้องกับงานวิจัยก่อนหน้า [5] [6] ที่มองว่า Quantum ML กำลังก้าวหน้าอย่างรวดเร็วและจะมีบทบาทสำคัญในอนาคต

### 5.2 สรุป

งานวิจัยนี้เปรียบเทียบ Quantum-inspired Machine Learning กับ Classical Machine Learning สำหรับพยากรณ์การกลับมาใช้บริการลูกค้าร้านอาหารกลางคืน ผลการศึกษาพบว่า Random Forest มีประสิทธิภาพโดยรวมสูงสุด XGBoost ได้ค่า precision สูงสุด เหมาะสำหรับการตลาดเฉพาะกลุ่ม QiSVM แสดงประสิทธิภาพที่แข่งขันได้ ผ่านการจำลอง quantum entanglement ส่วน QiNN ต้องปรับปรุงเพิ่มเติม งานวิจัยนี้ยืนยันความเป็นไปได้ของวิธีการ quantum-inspired บนคอมพิวเตอร์คลาสสิกโดยไม่ต้องใช้ฮาร์ดแวร์ควอนตัม ขยายการวิจัยการวิเคราะห์ลูกค้าในอุตสาหกรรมบริการ แนะนำให้ธุรกิจใช้ Random Forest สำหรับกลยุทธ์รักษาลูกค้า XGBoost สำหรับจัดสรรทรัพยากร และ QiSVM สำหรับจดจำรูปแบบซับซ้อน

### 5.3 งานวิจัยในอนาคต

งานวิจัยในอนาคตควรขยายข้อมูลไปยังร้านหลายแห่งในภูมิภาคหลากหลาย เปรียบเทียบ quantum feature maps ต่างๆ ศึกษาระยะยาวเพื่อประเมินความเสถียรโมเดลบูรณาการกับ deep learning หรือแนวทาง hybrid quantum-classical วิเคราะห์ต้นทุน-ผลประโยชน์ และเปรียบเทียบวิธีการ quantum-inspired กับการใช้งานควอนตัมจริงเมื่อฮาร์ดแวร์เข้าถึงได้มากขึ้น เพื่อวัดช่องว่างประสิทธิภาพและกำหนดกลยุทธ์การนำเทคโนโลยีมาใช้ในอุตสาหกรรมบริการ

## เอกสารอ้างอิง

[1] B. Balakrishnan, S. Gawali, S. Tiwari, S. S. Kalsi, and V. Naregal, "Data Analytics of Restaurants using AI," in 2023 9th International Conference on Smart Computing and Communications (ICSCC), 17-19 Aug. 2023 2023, pp. 506-511, doi: 10.1109/ICSCC59169.2023.10335092.

[2] K. Kumarasiri, J. Premarathne, N. Mendis, M. Bandhuratne, A. Karunasena, and L. Weerasinghe, "SmartDine: A Comprehensive System for Enhancing Customer Satisfaction and Optimizing Business Operations in the Restaurant Context," in 2024 6th International Conference on Advancements in Computing (ICAC), 12-13 Dec. 2024 2024, pp. 378-383, doi: 10.1109/ICAC64487.2024.10851011.

[3] T. N. Pandey, A. Vasudev, D. Sagayanathan, G. Anjan, D. Arshad, and S. S. Patra, "Predicting Customer Satisfaction in Brazil E-commerce: A Comparative Study of Machine Learning Techniques," in 2023 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS), 3-4 Nov. 2023 2023, pp. 505-510, doi: 10.1109/ICCCIS60361.2023.10425505.

[4] N. M. T, A. Hiremath, M. N, S. L. Peng, R. S. M, and K. P. S, "A Survey on Machine Learning

Techniques Using Quantum Computing," in 2022 Fourth International Conference on Emerging Research in Electronics, Computer Science and Technology (ICERECT), 26-27 Dec. 2022 2022, pp. 1-6, doi: 10.1109/ICERECT56837.2022.10059764.

[5] S. Misra and P. Rani, "Quantum Machine Learning: A Comprehensive Overview and Analysis," in 2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT), 24-28 June 2024 2024, pp. 1-5, doi: 10.1109/ICCCNT61001.2024.10724803.

[6] J. R. Jiang, "A Quick Overview of Quantum Machine Learning," in 2023 IEEE 5th Eurasia Conference on IOT, Communication and Engineering (ECICE), 27-29 Oct. 2023 2023, pp. 301-304, doi: 10.1109/ECICE59523.2023.10383149.

[7] V. Kumar and A. Mishra, "Efficient Food Delivery Management: Predicting Delivery Times for Improved Customer Experience," in 2024 2nd International Conference on Artificial Intelligence and Machine Learning Applications Theme: Healthcare and Internet of Things (AIMLA), 15-16 March 2024 2024, pp. 1-4, doi: 10.1109/AIMLA59606.2024.10531429.

[8] B. Surendiran, K. Dhanasekaran, and A. Tamizhselvi, "A Study on Quantum Machine Learning for Accurate and Efficient Weather Prediction," in 2022 Sixth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), 10-12 Nov. 2022 2022, pp. 534-537, doi: 10.1109/I-SMAC55078.2022.9987293.

[9] S. Dhotre, K. Doshi, S. Satish, and K. Wagaskar, "Exploring Quantum Machine Learning (QML) for Earthquake Prediction," in 2022 2nd International Conference on Intelligent Technologies (CONIT), 24-26 June 2022 2022, pp. 1-6, doi: 10.1109/CONIT55038.2022.9848250.

[10] J. K. A. Das and M. Malhotra, "Quantum Machine Learning: An Effective Approach to High-dimensional Learning," in 2022 6th International Conference on Computation System and Information Technology for Sustainable Solutions (CSITSS), 21-23 Dec. 2022 2022, pp. 1-6, doi: 10.1109/CSITSS57437.2022.10026399.

[11] P. Sridevi, P. B. Upama, M. Rabbani, and S. I. Ahamed, "Performance Comparison of Quantum and Classical Machine Learning Models for Chronic Kidney Disease Prediction," in 2025 IEEE 49th Annual Computers, Software, and Applications Conference (COMPSAC), 8-11 July 2025 2025, pp. 788-793, doi: 10.1109/COMPSAC65507.2025.00105.

[12] A. Sharma and R. Mandrai, "A Comparative Study: Predicting Time Series Data with Quantum and Classical Machine Learning," in 2024 International Conference on Electrical and Computer Engineering Researches (ICECER), 4-6 Dec. 2024 2024, pp. 1-6, doi: 10.1109/ICECER62944.2024.10920470.

[13] A. S. Selvi, S. Narendrakumar, G. V. M. R. Samanvay, C. T. A. Vinayak, S. A. Vembu, and S. S. Pandi, "From Classical to Quantum: Evaluating Machine Learning Enhancements," in 2024 2nd International Conference on Advances in Computation, Communication and Information Technology (ICAICCIT), 28-29 Nov. 2024 2024, vol. 1, pp. 778-784, doi: 10.1109/ICAICCIT64383.2024.10912309.

[14] A. Khot, T. Kim, A. R. Akash, C. H. Kim, and W. Moy, "Quantum-Inspired Machine Learning for Energy Efficient and Secure Battery Health Monitoring," in 2025 IEEE Power & Energy Society General Meeting (PESGM), 27-31 July 2025 2025, pp. 1-5, doi: 10.1109/PESGM52009.2025.11225365.

[15] R. Golchha and G. K. Verma, "Quantum-Enhanced Support Vector Classifier for Image Classification," in 2023 IEEE 8th International Conference for Convergence in Technology (I2CT), 7-9 April 2023 2023, pp. 1-6, doi: 10.1109/I2CT57861.2023.10126421.

# เปรียบเทียบประสิทธิภาพของการเรียนรู้ของเครื่องเพื่อการพยากรณ์การเติมสินค้าและช่วยในการตัดสินใจบริหารจัดการซัพพลายเชน

# Performance Comparison of Machine Learning Models for Inventory Replenishment and Supply Chain Decision-Making

*ณฐพร คูร์พิพัฒน์(Nataporn Koopipat)[1] และ สุภัคพงศ์ จินารัตน์ (Supakpong Jinarat)[2]*
*ภาควิชาเทคโนโลยีสารสนเทศ วิทยาลัยวิศวกรรมศาสตร์และเทคโนโลยี มหาวิทยาลัยธุรกิจบัณฑิต*
*[1]natapornkoo@gmail.com, [2]supakpong.jin@dpu.ac.th*

## บทคัดย่อ

*การวิจัยนี้มีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพของแบบจำลองการเรียนรู้ของเครื่อง สำหรับการพยากรณ์ข้อมูลที่เกี่ยวข้องกับการเติมสินค้าเพื่อสนับสนุนการตัดสินใจด้านซัพพลายเชนในอุตสาหกรรมอาหารและเครื่องดื่ม โดยใช้ข้อมูลย้อนหลังช่วงปี 2019–2025 ซึ่งประกอบด้วยตัวแปรสำคัญ ได้แก่ วันที่ส่งมอบสินค้า, น้ำหนักสุทธิ และฤดูกาล งานวิจัยดำเนินการทดลองด้วยแบบจำลองการถดถอยหลายรูปแบบ ได้แก่ Linear Regression, Decision Tree, Random Forest, Gradient Boosting, XGBoost และ Support Vector Regressor (SVR) รวมทั้งประเมินรูปแบบการทดสอบเพิ่มเติมของ Random Forest ทั้งแบบแบ่งข้อมูลตามเวลา (Time-based split) และแบบปรับจูนพารามิเตอร์ (Tuning) ผ่าน Time Series Cross Validation และ Grid Search การประเมินผลใช้ตัวชี้วัด MAE, MSE, RMSE และ $R^2$ ผลการทดลองพบว่า Random Forest ให้ประสิทธิภาพดีที่สุด โดยมีค่า MAE = 0.05, RMSE = 1.79 และ $R^2$ = 1.00 ซึ่งบ่งชี้ถึงความคลาดเคลื่อนต่ำและความสามารถในการอธิบายความแปรปรวนของข้อมูลได้สูงมาก เมื่อพิจารณาการทดสอบแบบ Time-based split Random Forest ยังให้ผลลัพธ์ที่ดี (MAE = 0.07, RMSE = 4.25, $R^2$ = 0.99) สรุปได้ว่าแบบจำลองตระกูล Tree-based โดยเฉพาะ Random Forest มีความเหมาะสมสำหรับการพยากรณ์เพื่อวางแผนเติมสินค้าในบริบทข้อมูลลักษณะนี้ และสามารถนำไปต่อยอดสู่ระบบช่วยตัดสินใจเพื่อลดความเสี่ยงการขาดสต็อก/ค้างสต็อก และเพิ่มประสิทธิภาพการบริหารซัพพลายเชนได้ในทางปฏิบัติ*

**คำสำคัญ:** แรนดอมฟอเรสต์ การตรวจสอบความถูกต้องของแบบจำลองอนุกรมเวลา การวิเคราะห์ข้อมูลห่วงโซ่อุปทาน การเติมเต็มสินค้า การพยากรณ์อุปสงค์

## Abstract

*This independent study aims to compare the performance of machine learning models for forecasting inventory replenishment data to support supply chain decision-making in the food and beverage industry. Utilizing historical data from 2019 to 2025, the study incorporates key variables: delivery date, net weight, and seasonality. The experiments evaluated multiple regression models, including Linear Regression, Decision Tree, Random Forest, Gradient Boosting, XGBoost, and Support Vector Regressor (SVR). Additionally, Random Forest was further tested using time-based splitting and hyperparameter tuning via Time Series Cross-Validation and Grid Search, Model performance was assessed using MAE, MSE, RMSE, and $R^2$. The results indicate that Random Forest outperformed other models, achieving an MAE of 0.05, RMSE of 1.79, and $R^2$ of 1.00, demonstrating minimal error and exceptionally high explanatory power. Under time-based split conditions, Random Forest maintained robust performance (MAE = 0.07, RMSE = 4.25, $R^2$ = 0.99). The findings conclude that tree-based models, particularly Random Forest, are highly suitable for replenishment forecasting in this context. These models can be integrated into decision support systems to mitigate stockout or overstock risks and enhance practical supply chain management efficiency.*

**Keywords:** Random Forest, Time Series Validation, Supply Chain Analytics, Replenishment, Demand Forecast.

## 1. บทนำ

ในยุคปัจจุบันที่เข้าสู่ยุคการเปลี่ยนผ่านสู่เทคโนโลยีปัญญาประดิษฐ์ดังนั้นการนำ การเรียนรู้ของเครื่องเข้ามาจึงมีบทบาทอย่างมากในทุกกระบวนการของซัพพลายเชนตั้งแต่ต้นน้ำของห่วงโซ่อุปทานจนไปถึงปลายน้ำอย่างลูกค้า ซึ่งในงานวิจัยในปัจจุบันมีความนิยมในการนำ ปัญญาประดิษฐ์ หรือการเรียนรู้ของเครื่อง (Machine Learning) เข้ามาช่วยในงานด้านซัพพลายเชนอย่างแพร่หลาย โดยมีการวิจัยแบ่งเป็น 2 ประเภทหลักดังนี้ 1) งานวิจัยที่นำ โมเดลของ การเรียนรู้ของเครื่องเพื่อเข้ามาช่วยในการพยากรณ์สต๊อกของคลังสินค้ารวมถึงช่วยในการพัฒนาการสร้างแอพลิเคชั่นที่เกี่ยวข้องกับการพยากรณ์จำนวนของสต๊อกที่ต้องจัดเก็บ 2) การนำ Deep Reinforcement Learning เข้ามาช่วยเพิ่มประสิทธิภาพของการจัดเก็บสต๊อกเพื่อให้มีการจัดการอย่างมีประสิทธิภาพ

อุตสาหกรรมอาหารและเครื่องดื่ม (Food and Beverage) ประสบกับความท้าทายอย่างมากจากการแข่งขันด้านราคาและต้นทุน รวมถึงความผันผวนของอุปสงค์ที่มีความไม่แน่นอนสูง การบริหารจัดการ โซ่อุปทานที่มีประสิทธิภาพ โดยเฉพาะการบริหารจัดการสินค้าคงคลังจึงเป็นกุญแจสำคัญในการรักษาความได้เปรียบทางการแข่งขัน ซึ่งในปัจจุบันยังมีหลายองค์กรที่ยังใช้ตัวชี้วัดการจัดการประสิทธิภาพการจัดการสินค้าคงคลัง

หลักการคำนวณจาก จำนวนวันเก็บสินค้าคงคลัง ถือเป็นหลักการที่พบได้มากที่สุดในการจัดการสินค้าคงคลังโดยยึดหลักจำนวนสินค้าคงคลังหารด้วยยอดขายเฉลี่ยต่อวัน โดยหลักการนี้พบกว่าหากหากเป็นสินค้าที่มีการเปลี่ยนแปลงเร็ว เช่น อาหารสด และมีจำนวนสินค้าคงคลังเยอะสินค้าอาจต้องนำมาลดราคาขาดทุนเพื่อระบายออกเนื่องจากเป็นสินค้าที่มีระยะเวลาการเก็บรักษาน้อยซึ่งส่งผลให้ต้องเสียต้นทุนสูง หากมีน้อยกว่าจำนวนสินค้าคงคลัง ก็อาจจะเกิดการเสียโอกาสในการขายหากมีความต้องการสินค้าพุ่งสูงขึ้นกะทันหัน ธุรกิจจะไม่มีสินค้าขาย ทำให้เสียรายได้และอาจเสียลูกค้าให้คู่แข่ง ซึ่งถ้าหากมีการประเมิณจำนวนสินค้าคงคลังที่ต้องเตรียมให้เพียงพอต่อยอดขายโดยการใช้เรียนรู้ของเครื่องเข้ามาวิเคราะห์ถึงจำนวนที่ต้องนำมาจัดเก็บเพิ่มแทนการใช้วิธีแบบดั้งเดิมโดยการนำยอดขายเข้ามาใช้วิเคราะห์จะส่งผลให้แก้ปัญหาเรื่องการขาดสินค้า และมีสินค้าเพียงพอต่อความต้องการ

การวิจัยในระยะหลังจึงมุ่งเน้นการนำเทคโนโลยีการเรียนรู้ของเครื่อง (Machine Learning: ML) และการเรียนรู้เชิงลึก (Deep Learning) เข้ามาประยุกต์ใช้เพื่อยกระดับความแม่นยำในการพยากรณ์และการตัดสินใจโดยมีการเก็บข้อมูลย้อนหลังจากความไม่แน่นอนของตลาดเพื่อในการวิเคราะห์ ข้อมูลพื้นฐานที่ถูกนำมาใช้ประกอบด้วย ข้อมูลการทำธุรกรรมย้อนหลัง (Historical Transaction Data) เช่น ยอดขายรายวัน ระดับสินค้าคงคลัง และรอบการสั่งซื้อ ร่วมกับการรวบรวม ปัจจัยภายนอก (External Factors) ที่ส่งผลกระทบต่อความผันผวนของอุปสงค์ เช่น ดัชนีทางเศรษฐกิจ โปรโมชั่น วันหยุดนักขัตฤกษ์ และเทศกาลต่างๆ ซึ่งอาจจะส่งผลกระทบโดยตรงต่อกลุ่มสินค้าประเภทธุรกิจที่เกี่ยวข้องกับการจัดจำหน่ายอาหารและเครื่องดื่มทั้งหมด

## 2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

### 2.1 การพยากรณ์สินค้าคงคลังด้วย Machine Learning เพื่อปรับปรุงประสิทธิภาพในการบริหารจัดการซัพพลายเชน

งานวิจัย [1] ได้นำ Machine Learning เข้ามาช่วยในโรงงานบิสกิตที่มีสินค้ากว่า 50 รายการ แทนการใช้ Excel แบบ Manual ซึ่งจากโมเดล Random Forest ช่วยในการพยากรณ์สินค้าว่ากลุ่มสินค้าไหนจะขายดีซึ่งจากวิธีเดิมมีการคลาดเคลื่อนถึง 46% ซึ่งแตกต่างจาก [2] ได้นำการวิเคราะห์แบบ ABC Analysis ร่วมกับทางโมเดล Random Forest ซึ่งปรากฏว่าสามารถลดจำนวนสินค้าคงคลังส่วนเกินได้ถึง 15-20% และเป็นโมเดลที่ให้ค่า RMSE ต่ำที่สุด ในงานวิจัยของ [3] มีการใช้แบบ Hybrid Model โดยนำ CNN + XGBoost มาช่วยพยากรณ์จากธุรกิจ E-commerce ที่มีความผันผวนสูงซึ่งมีค่า R2 Score ที่ 0.78 เหนือกว่าการใช้ Machine Learning เพียง Model เดียว [4] ยืนยันว่า Random Forest เหมาะสมอย่างยิ่งสำหรับข้อมูลค้าปลีกที่มีความซับซ้อนและมีข้อมูลสูญหายจำนวนมากซึ่งมีความแม่นยำถึง 92% ในการช่วยลดปัญหาสินค้าไม่เพียงพอโดยปัจจัยหลักในการพยากรณ์นี้คือจำนวนข้อมูลที่มากเพียงพอ

งานวิจัย [5] ได้นำ Neural Networks มาใช้ร่วมกับ Random Forests จะให้ผลลัพธ์ สามารถเพิ่มความแม่นยำในการทำนายยอดขายได้สูงถึง 92.5% แต่มีอุปสรรคสำคัญในการนำ ML มาใช้ คือปัญหาด้านคุณภาพของข้อมูล (Data quality) กับการนำ ML มาใช้ร่วมด้วยกัน [6] การนำ AI และ ML โดยเฉพาะโมเดล Gradient Boosting เพิ่มความแม่นยำในการพยากรณ์อุปสงค์ได้ถึง 20%-30% เมื่อเทียบกับการพยากรณ์แบบ Moving Average หรือวิธีการทางสถิติแบบเดิม

**2.2 การประยุกต์ใช้ Machine Learning เข้ามาช่วยในการพัฒนาแอปพลิเคชัน**

งานของ [7] มีการนำการเรียนรู้ของเครื่อง เข้ามาช่วยในการพยากรณ์แผนในการสั่งซื้อสินค้าคงคลังล่วงหน้าซึ่งมีการพัฒนาแอปพลิเคชัน ร่วมกับระบบ POS โดยได้มีการประยุกต์ใช้แนวคิดการบริหารจัดการสินค้าคงคลังให้เพียงพอต่อความต้องการของลูกค้า มีการใช้ระบบ POS เพื่อลดความผิดพลาดในการบันทึกข้อมูลด้วยมือ และสุดท้ายใช้เทคโนโลยีการเรียนรู้ของเครื่องเข้ามาวิเคราะห์ข้อมูลยอดขายในอดีตและพยากรณ์ถึงยอดขายในอนาคต ซึ่งผลลัพธ์แสดงให้เห็นว่าโมเดล Gradient Boosting Regression มีความสามารถสูงที่สุดในการเข้ามาช่วยในการวางแผนสั่งซื้อซึ่งให้ผลลัพธ์ดีที่สุด โดยมีค่า MAE = 1.4320, RMSE: 1.9270 และ R2 = 0.7148 สามารถลดต้นทุนในการจัดเก็บสินค้าคงคลังได้ถึง 12% ดังนั้นจึงสรุปได้ว่าแอปพลิเคชันนี้ช่วยให้ธุรกิจขนาดเล็กสามารถ ลดต้นทุน จากปัญหาสินค้าค้างสต็อกและหมดอายุได้จริง ผ่านระบบพยากรณ์ที่แม่นยำและการใช้งานที่ง่ายผ่านมือถือ ซึ่งช่วยลดความจำเป็นในการติดตั้งอุปกรณ์คอมพิวเตอร์ที่ซับซ้อน

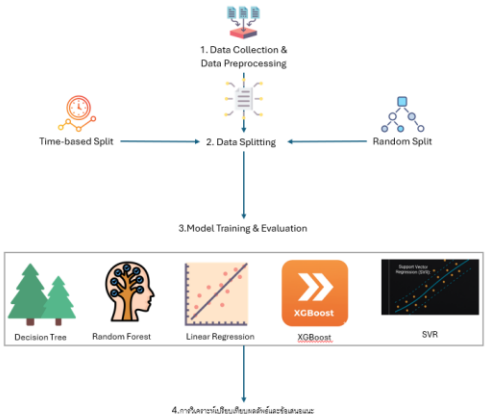
**2.3 การนำ Deep Reinforcement Learning (DRL) เข้ามาช่วยในการ Optimization ของ Inventory**

งานวิจัย [8] ได้มีการนำ DRL เข้ามาช่วยในการ Replenishment Stock ซึ่งได้มีการนำ Long Short-Term Memory (LSTM) ในการพยากรณ์ยอดขายและนำ Base Model คือ ARIMA มาเปรียบเทียบกับ DRL อย่าง Markov Decision Process (MDPs) โดยผลลัพธ์แสดงให้เห็นว่าโมเดล LSTM ลดค่าความคลาดเคลื่อน MAE ได้ถึง 15% และ RMSE 18% เมื่อเทียบกับ ARIMAลดปัญหา Stockouts ได้ 22% ทำให้เสียโอกาสการขายน้อยลง ลด Excess Inventory (สต็อกส่วนเกิน) ได้ 28% ช่วยลดต้นทุนการจัดเก็บสินค้าคงคลังและพื้นที่จัดเก็บ

งานวิจัย [9] ได้นำ DRL เข้ามาใช้ในธุรกิจที่สินค้าเน่าเสียได้ง่ายระยะเวลาในการจัด อายุของสินค้าน้อยโดยใช้ Proximal Policy Optimization (PPO) เข้ามาช่วยในการตัดสินใจซึ่งหลังจากการทดลองพบกว่าเป็นการทดลองที่มีประสิทธิภาพในการช่วยการตัดสินใจในการตั้งราคาสินค้าที่ใกล้หมดอายุเทียบกับจำนวนสินค้าคงคลังแบบ Real-Time ลดการสูญเสียของสินค้าได้อย่างมีประสิทธิภาพ

งานของ [10] หลังจากการทดลองพบว่าหากมีการนำ Variational Mode Decomposition – VMD เข้ามาช่วยกับ DRL จะมีประสิทธิภาพเหนือกว่าการใช้ Model ของ DRL อย่างเดียว ซึ่งสามารถลดค่า MAE ได้ 37%

## 3. วิธีการดำเนินการวิจัย




**ภาพที่ 1:** ขั้นตอนกระบวนการในงานวิจัยที่เสนอ

**3.1 การเตรียมข้อมูล (Data Preprocessing)**

ทำการรวบรวมข้อมูลสำหรับงานวิจัยที่เกี่ยวข้องโดยเป็นข้อมูลน้ำหนักจากการจ่ายออก(Outbound Weight)ที่มีการเกิดขึ้นจริงย้อนหลัง 6 ปี เนื่องจากเป็น

ข้อมูลที่รวบรวมได้จากระบบ ERP จัดเป็นข้อมูลดิบที่ต้องผ่านขั้นตอนการเตรียมได้ทำการจัดการกับค่าผิดปกติ (Outliers) โดยข้อมูลจะมีคำสั่งการสั่งซื้อสินค้าจำนวนมากในคราวเดียว รวมถึงคำสั่งซื้อล่วงหน้าเป็นจำนวนมากส่งผลให้หากทำการพยากรณ์ออกมาส่งผลให้ข้อมูลมีการกระโดดมากกว่าปกติหรือสูงเกินจริง และข้อมูลที่ไม่ครบถ้วน (Missing Data) ในกรณีที่มีการสร้างการรายธุรกรรมคำสั่งซื้อที่ไม่สมบูรณ์รวมถึงทำความสะอาดชื่อคอลัมน์โดยการตัดช่องว่างที่อยู่หน้าและ

หลังออกเพื่อให้โมเดลการเรียนรู้ของเครื่องทำงานได้อย่างมีประสิทธิภาพสูงสุด

**3.2 การแบ่งชุดข้อมูล**

ในงานวิจัยนี้ได้ทำการนำข้อมูลย้อนหลัง (Historical Data) มาทำการทดสอบ 2 แบบด้วยกันดังนี้

3.2.1 Random Split เป็นการการสุ่มเลือกข้อมูลจากทุกช่วงเวลามาผสมกันโดยแบ่งชุดข้อมูลเป็น 80/20 สำหรับการฝึกและ ทดสอบประสิทธิภาพของแต่ละโมเดล

3.2.2 Time-based Split การแบ่งตามช่วงเวลาซึ่งในงานวิจัยนี้ได้มีการทดสอบโดยใช้ข้อมูลย้อนหลังตั้งแต่ปี 2019-2024 สำหรับการฝึกและนำปี 2025 สำหรับการทดสอบประสิทธิภาพของแต่ละโมเดล

**3.3 การฝึกฝนและเปรียบเทียบโมเดล**

ในงานวิจัยนี้ได้มีการนำโมเดลของ Machine Learning ทั้งหมด 6 โมเดลเพื่อนำมาทำการทดลองและวัดประสิทธิภาพดังนี้

3.2.1 Linear Regression
3.2.2 Decision Tree
3.2.3 Random Forest
3.2.4 Gradient Boosting
3.2.5 XGBoost
3.2.6 Support Vector Regressor

**ตารางที่ 1:** Parameters ที่ใช้ในการทดลองในแต่ละโมเดล

| โมเดล (Model) | Key Parameter |
|---|---|
| Linear Regression | Default ใช้ค่ามาตรฐาน |
| Decision Tree | random_state=42 กำหนดค่าเริ่มต้นในการสุ่ม |
| Random Forest (Baseline/Time-based/CV) | random_state=42, n_estimators=100, กำหนดจำนวนต้นไม้พื้นฐาน 100 ต้น เพื่อสร้างค่าเฉลี่ยการทำนาย เริ่มต้นmax_depth=None จำกัดความลึกสูงสุด |
| Random Forest (Tuned) | n_estimators=100, max_depth=10, min_samples_leaf=1 กำหนดจำนวนข้อมูลขั้นต่ำในใบ |
| Gradient Boosting | random_state=42, learning_rate=0.1, อัตราการเรียนรู้ที่ค่อยๆ ปรับแก้ข้อผิดพลาด max_depth=3 |
| XGBoost | random_state=42, learning_rate=0.3, ปรับความเร็วในการหาจุดเหมาะสมที่สุด max_depth=6 |
| SVR | kernel='linear', ใช้ฟังก์ชันเส้นตรงในการสร้างขอบเขตการทำนายในมิติข้อมูล C=1.0, ค่าที่ควบคุมสมดุลระหว่างความแม่นยำและขอบเขตความยืดหยุ่น (Margin) epsilon=0.1แถบความคลาดเคลื่อน |

**3.4 การทดสอบขั้นสูงและการปรับแต่ง (Advanced Evaluation & Tuning)**

หลังจากได้โมเดลที่มีประสิทธิภาพดีที่สุดและมีค่าผิดพลาดต่ำที่สุดจึงมีการปรับแต่งโมเดลเพิ่มเติมดังนี้

3.4.1 Time Series Cross Validation ซึ่งจะมีการใช้ Time Series Split เพื่อตรวจสอบประสิทธิภาพของโมเดลในแต่ละช่วงเวลา

3.4.2 Hyperparameter Tuning ใช้ Grid Search CV เพื่อค้นหาการตั้งค่าที่ดีที่สุด (เช่น จำนวนต้นไม้ หรือความลึกของกิ่งก้าน) เพื่อเพิ่มความแม่นยำสูงสุด

3.4.3 ใน Support Vector Regressor มีการกำหนด ความสมดุลเบื้องต้นหรือตัวแปร “C”เพื่อให้ได้สมดุลที่ไม่ตึงหรือหย่อนเกินไประหว่างการพยายามลดความผิดพลาด

**3.5 การวิเคราะห์เปรียบเทียบผลลัพธ์และข้อเสนอแนะ**

3.5.1 การวัดผล: ใช้ตัวชี้วัด MAE, MSE, RMSE, R2 มาเปรียบเทียบเพื่อหาโมเดลที่ได้ประสิทธิภาพดีที่สุด

3.5.2 การแสดงผล: สร้างตารางเปรียบเทียบและ Scatter Plots เพื่อดูความใกล้เคียงระหว่าง "น้ำหนักจริง" และ "น้ำหนักที่พยากรณ์" ผ่านเส้น Perfect Prediction

3.5.3 ข้อเสนอแนะ: สรุปจุดแข็ง-จุดอ่อน และระบุโมเดลที่เหมาะสมที่สุดในการพยากรณ์น้ำหนักรวม (Gross weigh) สำหรับธุรกิจของคุณ

## 4. ผลการดำเนินงาน

จากผลการวิจัยในการนำการเรียนรู้ของเครื่องเข้ามาช่วยในการจัดการซัพพลายเชนในด้านของ Inventory ซึ่งเปรียบเสมือนหัวใจหลักของงานทางด้าน Logistics โดยมีรายละเอียดการวิจัยดังนี้

**4.1 ข้อมูลที่นำมาใช้ในการดำเนินการงานวิจัย**

จำนวนข้อมูลย้อนหลัง (Historical Data) ตั้งแต่ปี 2019-2025 ซึ่งถูกรวบรวมจากระบบ ERP ของบริษัทโดยมีความผันผวนตามความต้องการของลูกค้ามีและมี Field หลัก ดังต่อไปนี้

4.1.1 Delivery Date : วันที่ส่งมอบสินค้าให้แก่ลูกค้า

4.1.2 Net Weight : จำนวนน้ำหนักจริง (Actual Weight)ในการใช้พยากรณ์

4.1.3 Seasonal : ช่วงเวลาที่ความต้องการสูงในแต่ละปี

**4.2 Evaluation Metrics**

ในงานวิจัยฉบับนี้จะมีการเปรียบเทียบประสิทธิภาพของ Model โดยมีการใช้ตัวชี้วัดดังนี้

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}| \quad (1)$$

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y})^2 \quad (2)$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y})^2} \quad (3)$$

$$R^2 = 1 - \frac{\sum(y_i - \hat{y})^2}{\sum(y_i - \bar{y})^2} \quad (4)$$

$y_i - \hat{y}$ คือการนำค่าจริงลบด้วยค่าพยากรณ์ แล้วเปลี่ยนให้เป็นค่าบวกเสมอ

$\sum$ คือการรวมผลต่างเหล่านั้นของข้อมูลทุกจุด

$\frac{1}{N}$ คือการหารด้วยจำนวนข้อมูลทั้งหมดเพื่อหาค่าเฉลี่ย

ซึ่งในงานวิจัยฉบับนี้ได้ให้ความสำคัญกับสัมประสิทธิ์การตัดสินใจ ($R^2$) เนื่องจากในงานวิจัยนี้ต้องการพิสูจน์ว่าแต่ละโมเดลเข้ากันกับข้อมูลได้มากน้อยแค่ไหน

**4.3 Experimental Results**

| Model | MAE | MSE | RMSE | R2 |
|---|---|---|---|---|
| **Linear Regression** | 13.90 | 1654.27 | 40.67 | 0.53 |
| **Decision Tree** | 0.16 | 102.92 | 10.15 | 0.97 |
| **Random Forest** | 0.05 | 3.20 | 1.79 | 1.00 |
| **Gradient Boosting** | 0.23 | 18.47 | 4.30 | 0.99 |
| **XGBoost** | 0.07 | 16.21 | 4.03 | 1.00 |
| **SVR (Scaled)** | 2.60 | 7654.72 | 87.49 | -1.16 |
| **Random Forest (Time-based)** | 0.07 | 18.06 | 4.25 | 0.99 |
| **Random Forest (Tuned)** | 0.05 | 3.50 | 1.87 | 1.00 |

**ตารางที่ 2:** ผลลัพธ์ของโมเดลที่ได้ทำการทดลอง

ซึ่งแสดงให้เห็นว่าแบบจำลอง **Random Forest** โมเดลมีประสิทธิภาพดีที่สุด โดยพิจารณาจาก

4.3.1 ค่าความคลาดเคลื่อนต่ำที่สุด (MAE, MSE, RMSE) Random Forest มีค่า MAE เพียง 0.05 และ RMSE เพียง 1.79 ซึ่งหมายถึงการพยากรณ์ผิดพลาดน้อยมากเมื่อเทียบกับตัวอื่น

4.3.2 ค่าความแม่นยำสูงสุด ($R^2$) มีค่าเท่ากับ 1.00 (หรือ 0.9991 ในรายละเอียดเชิงลึก) ซึ่งหมายถึงแบบจำลองสามารถอธิบายความผันแปรของข้อมูลได้เกือบทั้งหมด

จากการทดสอบประสิทธิภาพของแบบจำลองทั้งหมด 7 รูปแบบ พบว่าแบบจำลอง Random Forest ให้ประสิทธิภาพสูงสุดในทุกมิติ โดยมีค่า MAE ต่ำสุดที่ 0.05 และมีค่า สูงถึง R2 1.00 ในขณะที่แบบจำลอง SVR (Scaled) ให้ผลลัพธ์ที่ด้อยที่สุด โดยมีความคลาดเคลื่อนเฉลี่ยสูงและค่า R2 ติดลบที่ -1.16 แสดงให้เห็นว่าแบบจำลองประเภท Ensemble Learning มีความเหมาะสมกับชุดข้อมูลนี้มากกว่าแบบจำลองเชิงเส้น

## 5. สรุป

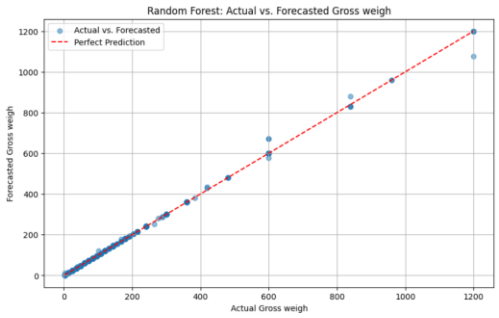
จากการทดลองโดยการใช้ข้อมูลย้อนหลัง 5 ปี มีการเรียงลำดับตาม วัน , เดือน และปี ซึ่งมีการกำหนดตัวแปรเป็นจำนวนน้ำหนัก ดังนี้

ตัวแปร X : จำนวน Actual น้ำหนักจริง

ตัวแปร Y : จำนวน Forecast น้ำหนักของ Model

โดยจากผลการทดลองเปรียบเทียบประสิทธิภาพของแต่ละ Modelพบกว่า Linear Regression และ SVR ถือว่าเป็นกลุ่มที่ได้ประสิทธิภาพต่ำที่สุด เมื่อเทียบกับ Random Forest ถือว่าเป็นโมเดลที่มีประสิทธิภาพมาเหนือกว่าโมเดลอื่นอย่างชัดเจนซึ่งมีการแบ่งข้อมูลตามลำดับเวลา (Time-based Split) ฝึกสอนแบบจำลองด้วยข้อมูลปี 2019-2024 และทดสอบด้วยข้อมูลปี 2025 พบว่ามีค่า MAE เท่ากับ 0.0714, RMSE เท่ากับ 4.2502 และค่า R2 เท่ากับ 0.9863 โดยชุดข้อมูลฝึกสอนมี 40,949 ตัวอย่าง และชุดข้อมูลทดสอบมี 4,588 ตัวอย่างโดยมีค่าความคลาดเคลื่อนเมื่อเทียบกับข้อมูลจริงถือว่าน้อยดังนั้นจึงมีการสรุปได้ว่า Random Forest Regressor คือทางเลือกที่ดีที่สุดสำหรับการพยากรณ์ Gross weigh โดยมีการสุปจุดเด่นของโมเดลดังนี้

1) สามารถใช้กับข้อมูลที่ไม่เป็นเชิงเส้นตรงได้ดี
2) จัดการข้อมูลที่มี Noise ได้อย่างมีประสิทธิภาพ




**ภาพที่ 2:** กราฟแสดงความแม่นยำของ Random Forest

ข้อเสนอแนะในอนาคตหากต้องพยากรณ์ข้อมูลในอนาคตที่ยังไม่เกิดขึ้น ควรยึดค่าประสิทธิภาพจากรุ่น Time-based split เป็นหลักเนื่องจากสะท้อนความเป็นจริงในการทำงานได้ดีที่สุด เนื่องจากมีการใช้ข้อมูลอดีตมาพยากรณ์ข้อมูลปีล่าสุด พบว่าทำผลงานได้ ดีมาก ซึ่งมีผลลัพธ์ $R^2$ เท่ากับ 0.98 แม้ค่าจะต่ำกว่าการสุ่มข้อมูลปกติเล็กน้อย แต่มีความสำคัญมากในการทำ Time-series เพราะสะท้อนสถานการณ์จริงที่ข้อมูลในอนาคตเป็นสิ่งที่เรายังไม่เคยเห็น

ประโยชน์ที่ได้จากงานวิจัยข้อมูลซับพลายเชนมีความซับซ้อนแต่โมเดล Linear Regression ให้ค่า $R^2$ เพียง 0.534 ในขณะที่กลุ่มโมเดลต้นไม้ตัดสินใจ อย่าง Random Forest ให้ค่า $R^2$ สูงถึง 0.9991 นั่นแสดงให้เห็นว่าปัจจัยที่ส่งผลต่อการสั่งซื้อและน้ำหนักสินค้าในอุตสาหกรรมอาหารและเครื่องดื่ม (เช่น ฤดูกาล, วันที่ส่งมอบ) ไม่ได้แปรผันตรงกันแบบสมการเส้นตรงปกติ แต่มีความสัมพันธ์ที่ซับซ้อนและมีเงื่อนไขทับซ้อนกันสูง งานวิจัยนี้เป็นแนวทางปฏิบัติสำหรับนักวิจัยและองค์กรในการวิเคราะห์ข้อมูลโซ่อุปทานที่มีความซับซ้อน โดยชี้ให้เห็นว่าควรพิจารณาเลือกใช้แบบจำลองกลุ่มออองซอมเบิล (Ensemble Models) เช่น Random Forest หรือ XGBoost แทนแบบจำลองเชิงเส้นตรง (Linear Models) ตั้งแต่ต้น เพื่อลดขั้นตอนการทดลองที่ซ้ำซ้อนและประหยัดทรัพยากรในการพัฒนาระบบ

## เอกสารอ้างอิง


[1] B. R. S. Mendis, “Predictive Analytics for Inventory Optimization in Manufacturing,” International Journal of Latest Technology in Engineering, Management & Applied Science (IJLTEMAS), vol. 14, no. 8, pp. 394-398, Aug. 2025. DOI: 10.51583/IJLTEMAS.2025.1408000048.

[2] N. Nguyen, “Utilizing Machine Learning to Enhance Optimal Inventory Management - Case: Aviation Industry Spart Parts,” Master’s thesis, Lappeenranta-Lahti University of Technology LUT, 2024.

[3] M. Thakar and P. Swaminarayan, “E-Commerce Inventory Prediction by Hybridization Deep and Machine Learning,” Journal of Information Systems Engineering and Management, vol. 10, no. 1, 165, 2025.

[4] N. Deivanayagampillai, T. Bhuvaneswari, and Y. Suppiah, “Intelligent inventory prediction: A machine learning framework using random forest for inventory forecasting,” Edelweiss Applied Science and Technology, vol. 9, no. 4, pp. 1795-1807, 2025. DOI: 10.55214/25768484.0914.6383.

[5] T. Sumarlin and N. Qosidah, “Optimizing Sales and Inventory Management with Machine Learning: Applications of Neural Networks and Random Forests,” Journal of Management and Informatics (JMI), vol. 3, no. 2, pp. 312-325, Aug. 2024.

[6] N. R. Mitta, “Leveraging AI for Smart Inventory Management in Retail: Developing Machine Learning Models for Predictive Replenishment, Stock Optimization, and Demand-Supply Balancing,” Australian Journal of Machine Learning Research & Applications, vol. 4, no. 2, pp. 113-148,

[7] S. Sutta, T. Jivnot, and T. Settawatcharawanit, “Development of an Inventory Management System on a Mobile Application using Machine Learning Technology,” EAU Heritage Journal Science and Technology, vol. 19, no. 3, pp. 78-90, Sep.-Dec. 2025.

[8] A. K. Kalusivalingam, A. Sharma, N. Patel, and V. Singh, “Optimizing Inventory Management with AI: Leveraging Deep Reinforcement Learning and Neural Networks for Enhanced Demand Forecasting and Stock Replenishment,” Independent Researcher, 2024.

[9] Y. Nomura, Z. Liu, and T. Nishi, “Deep Reinforcement Learning for Dynamic Pricing and Ordering Policies in Perishable Inventory Management,” Applied Sciences, vol. 15, no. 5, 2421, Feb. 2025. DOI: 10.3390/app15052421.

[10] C. Lei, H. Zhang, Z. Wang, and Q. Miao, “Deep Learning for Demand Forecasting: A Framework Incorporating Variational Mode Decomposition and Attention Mechanism,” Processes, vol. 13, no. 2, 491, Feb. 2025. DOI: 10.3390/pr13020491.

# การพัฒนาระบบแนะนำตำแหน่งงานด้วยเทคนิคการเรียนรู้ของเครื่อง
# A Machine Learning-Based Approach for the Development of Job Recommendation System

*อุไรวรรณ อินทร์แหยม (Uraiwan Inyaem)[1] วีณา จันทร์รัชชกูล (Weena Janratchakool)[2]*
*เบญจวรรณ สุรพล(Benjawan Surapol)[3] และเทวราช รักงาม(Tewarach Rakngam)[4]*
*[1,2,3,4]ภาควิชาคณิตศาสตร์และวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์และเทคโนโลยี*
*มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี*
*uraiwan.inyaem@rmutt.ac.th, weena_j@rmutt.ac.th, 1160109050873@mail.rmutt.ac.th,*
*1160109050923@mail.rmutt.ac.th*

**บทคัดย่อ**

*บทความนี้นำเสนอการพัฒนาระบบแนะนำตำแหน่งงานสำหรับองค์กรด้วยเทคนิคการเรียนรู้ของเครื่อง ในการวิจัยใช้ดาต้าเซทของผู้สมัครงาน จำนวน 3000 ชุด และข้อมูลตำแหน่งงาน จำนวน 50 ตำแหน่งด้านไอทีผ่านกระบวนการวิเคราะห์ข้อมูลที่ทันสมัยสำหรับเตรียมข้อมูลประมวลผลเพื่อหาโมเดลที่ดีที่สุด โดยเปรียบเทียบการเรียนรู้เครื่องจักร 3 เทคนิค ได้แก่ เทคนิคการสุ่มป่าไม้ เทคนิคการหาเพื่อนบ้านที่ใกล้ที่สุด และเทคนิคต้นไม้ตัดสินใจ เพื่อพิจารณาค่าความถูกต้องสูงสุดในการวัดประสิทธิภาพโมเดลและนำโมเดลไปพัฒนาเป็นระบบแนะนำตำแหน่งงาน ผลการวิจัยพบว่าเทคนิคการสุ่มป่าไม้มีค่าความถูกต้องสูงที่สุดเท่ากับ 89.73 เปอร์เซ็นต์ ค่าความเที่ยงตรง 90.21% และค่าการเรียกคืน 89.15% ระบบที่พัฒนาขึ้นมีการประเมินความพึงพอใจต่อการใช้งานระบบด้วยแบบสอบถามของผู้ใช้งานทั่วไป 30 คน ผลการประเมินพบว่า ค่าเฉลี่ยความพึงพอใจเท่ากับ 4.17 ความพึงพอใจอยู่ในระดับมาก และค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.60 ระบบที่พัฒนาขึ้นนำไปประยุกต์ใช้สำหรับแนะนำตำแหน่งงานให้กับบุคคลที่สนใจและสำหรับองค์กรได้*

**คำสำคัญ:** ระบบแนะนำตำแหน่งงาน เทคนิคการสุ่มป่าไม้ เทคนิคการหาเพื่อนบ้านที่ใกล้ที่สุด เทคนิคต้นไม้ตัดสินใจ

**Abstract**

*The purpose of the research is to develop a job recommendation system for organizations using machine learning techniques from data of 3000 applicants and 50 IT-job positions. The data were analyzed by modern processing to find the best model by comparing the 3 techniques of machine learning: Random Forest, K-Nearest Neighbor and Decision Tree. The highest accuracy in performance model is to develop into a job recommendation system. According to the research, it was found that the Random Forest technique gave highest accuracy of 89.73%, the precision of 90.21% and the recall of 89.15% respectively. The proposed system was developed. Satisfaction with the use of the system was assessed by questionnaires of 30 general users. The mean of satisfaction was 4.17. Satisfaction was at a high level and the standard deviation is 0.60. The developed system can be applied to introduce jobs for interested persons and for organizations.*

**Keywords:** Job Recommendation System, Random Forest, K-Nearest Neighbor, Decision Tree.

## 1. บทนำ

การรับสมัครงานของบริษัทในบางครั้งอาจจะเสียเวลาเป็นจำนวนมากในการรับสมัครและตรวจสอบบุคคลที่จะมาสมัครเข้าร่วมทำงานกับบริษัทหรือองค์กรที่ต้องการบุคคลากรเพิ่มและในการรับสมัครทั่วไป มีการรับสมัครในจำนวนมาก อาจจะทำให้มีการเสียเวลาในการจัดเอกสารแก่ผู้ที่มาสมัครซึ่งอาจจะทำให้เกิดข้อผิดพลาดตามมาและทำให้เสียเวลามากเพิ่มขึ้นไปอีกและอาจจะไม่ได้รับบุคคลตามความต้องการ

ในปัจจุบันการสมัครงานมีอยู่ 2 ประเภท ประเภทแรกการสมัครงานแบบทั่วไปคือ เข้าไปสมัครด้วยตัวเองกรอกเอกสารต่าง ๆ และใช้คนในหน่วยงานเป็นการคัดเลือก ประเภทที่ 2 คือการสมัครและลงทะเบียนทางออนไลน์แล้วรอให้หน่วยงานติดต่อไปแล้วจึงเรียกมาสัมภาษณ์ แต่ปัจจุบันมีเทคโนโลยีการเรียนรู้ของเครื่อง(Machine Learning) เพื่อจัดการกับข้อมูลจำนวนมากในคัดกรองความสามารถของผู้ที่เข้ามาสมัครงานผ่านระบบนี้ โดยการที่นำความสามารถของผู้สมัครมาทำการเก็บเป็นข้อมูลและประมวลผลเพื่อหาตำแหน่งงานที่ควรจะได้รับการสมัครงานส่วนใหญ่ผู้สมัครจะต้องเดินทางไปยังสถานที่เพื่อสมัครงานและต้องเจอกับกระบวนการเอกสารที่มีมากมายแหละหลายขั้นตอน ปัจจุบันก็มีการรับสมัครทางออนไลน์มาทดแทนและแก้ไขปัญหาได้บางส่วนซึ่งทำให้ผู้สมัครค่อนข้างสะดวกมากขึ้น แต่กลับกันทางฝ่ายรับสมัครอาจจะได้บุคลที่มาสมัครเพิ่มมากขึ้นแต่ไม่ตรงกับความต้องการ อีกทั้งในไทยยังมีปัญหาเรื่องการรับสมัครงานแบบรุ่นพี่รุ่นน้องซึ่งไม่มีความเป็นกลางต่อ ผู้สมัครคนอื่นและอาจจะทำให้องค์กรนั้นเสียผลประโยชน์หรือได้ประโยชน์ไม่มากเท่าที่ควร ดังนั้นทีมวิจัยจึงได้สร้างระบบนี้ขึ้นโดยใช้เทคโนโลยีกระบวนการข้างต้นมาใช้ในการรับสมัครบุคลากร ระบบนี้สามารถพิจารณาคัดกรองผู้สมัครให้ตรงกับความต้องการขององค์กรเองได้ และยังสามารถแก้ไขปัญหาที่เกิดขึ้น

การเรียนรู้ของเครื่องเพื่อค้นหารูปแบบและความสัมพันธ์ที่ซ่อนอยู่ในชุดข้อมูล ปัจจุบันถูกนำไปประยุกต์ใช้ในงานหลายประเภททั้งด้านธุรกิจที่ช่วยในการตัดสินใจของผู้บริหาร ในด้านวิทยาศาสตร์และการแพทย์ รวมทั้งในด้านเศรษฐกิจและสังคม และการหากฎความสัมพันธ์ เป็นการหาความสัมพันธ์ของข้อมูลทั้งสองชุดหรือมากกว่าสองชุดขึ้นไปไว้ด้วยกัน ความสำคัญของกฎทำการวัดโดยใช้ข้อมูลสองตัวด้วยกันคือค่าสนับสนุน (Support) ซึ่งเป็นเปอร์เซ็นต์ของการดำเนินการที่กฎสามารถนาไปใช้ หรือเป็นเปอร์เซ็นต์ของการดำเนินการที่กฎที่ใช้มีความถูกต้อง และข้อมูลตัวที่สองที่นำมาใช้วัดคือค่าความมั่นใจ (Confidence) ซึ่งเป็นจำนวนของกรณีที่กฎถูกต้องโดยสัมพันธ์กับจำนวนของกรณีที่กฎสามารถนำไปใช้ได้ ในปัจจุบันนำมาใช้กันทั้งภาคธุรกิจขนาดเล็กขนาดกลางและขนาดใหญ่ เพื่อนำมาวิเคราะห์ข้อมูล เช่น การวิเคราะห์ข้อมูลผู้สมัครงาน ใช้พิจารณาความพึงพอใจในงานของผู้สมัครพร้อมด้วยการจับคู่โปรไฟล์ตามเนื้อหาพร้อมกับความแม่นยำในการทำนายยังช่วยในการแก้ปัญหาแก่ผู้สมัคร ปัจจุบันการตั้งค่าของผู้สมัครเกี่ยวกับงานที่จะช่วยให้จัดลำดับความสำคัญเฉพาะงานที่เกี่ยวข้องเมื่อเทียบกับงานที่ไม่เกี่ยวข้องซึ่งได้รับการคัดเลือกหลังจากการจับคู่ตามเนื้อหาของผู้สมัคร ในงานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาระบบแนะนำตำแหน่งงานสำหรับองค์กรด้วยเทคนิคการเรียนรู้ของเครื่อง โดยมีการสร้างแบบจำลองสำหรับคัดกรองคนที่สมัครงานด้วยเทคนิคเหมืองข้อมูล เพื่อแนะนำตำแหน่งงานที่ใกล้เคียงกับข้อมูลผู้สมัคร และพัฒนาระบบคัดเลือกพนักงานได้ตรงตามความต้องการในตำแหน่งที่ผู้ดูแลระบบสามารถดูข้อมูลผู้สมัครงานดูตำแหน่งงานและแก้ไขตำแหน่งงานได้ ดูข้อมูลโพสต์ประกาศหางานของผู้สมัครได้ และส่วนผู้ใช้สามารถลงทะเบียนเพื่อบันทึกข้อมูลให้กับระบบได้ ทำนายตำแหน่งงานตามความสามารถที่ลงทะเบียนไว้ สามารถขอคำแนะนำตำแหน่งงานอื่นที่ใกล้เคียงได้ และสามารถประกาศหาตำแหน่งงานได้ โดยประโยชน์ที่ได้รับจากงานวิจัยนี้จะเป็นลักษณะระบบช่วยทำนายในการแก้ปัญหาแก่ผู้สมัครที่ไม่ชัดเจนเป้าหมายงาน และแนะนำตำแหน่งงานอื่นที่ใกล้เคียงกับข้อมูลผู้สมัคร

## 2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

### 2.1 การสมัครงาน (Job application)

ในปัจจุบันรูปแบบการสมัครงานแบบสมัครด้วยตัวเอง กรอกเอกสารและใช้คนในหน่วยงานเป็นการคัดเลือก การสมัครลงทะเบียนทางออนไลน์แล้วรอให้หน่วยงานติดต่อไปมาสัมภาษณ์ แต่ปัจจุบันมีเทคโนโลยีการเรียนรู้ของเครื่องเพื่อจัดการกับข้อมูลจำนวนมากในคัดกรองความสามารถของผู้ที่เข้ามาสมัครงานผ่านระบบนี้ โดยการที่นำความสามาร ของผู้สมัครมาทำการเก็บเป็นข้อมูลและประมวลผลเพื่อหาตำแหน่งงานที่ควรจะได้รับการสมัครงานส่วนใหญ่ผู้สมัครจะต้องเดินทางไปยังสถานที่เพื่อสมัครงานและต้องเจอกับกระบวนการเอกสารที่มีมากมายแหละหลายขั้นตอน ปัจจุบันก็มีการรับสมัครทางออนไลน์มาทดแทนและแก้ไขปัญหาได้บางส่วนซึ่งทำให้ผู้สมัครค่อนข้างสะดวกมากขึ้นแต่กลับกันทางฝ่ายรับสมัครอาจจะได้บุคคลที่มาสมัครเพิ่มมากขึ้นแต่ไม่ตรงกับความต้องการ อีกทั้งในไทยยังมีปัญหาเรื่องการรับสมัครงานแบบรุ่นพี่รุ่นน้องซึ่งไม่มีความเป็นกลางต่อผู้สมัครคนอื่น และอาจจะทำให้องค์กรนั้นเสียผลประโยชน์หรือได้ประโยชน์ไม่มากเท่าที่ควร

### 2.2 การจำแนกประเภทข้อมูล (Classification Data)

เป็นกระบวนการสร้างโมเดลจัดการข้อมูลให้อยู่ในกลุ่มที่กำหนดมาให้จากกลุ่มตัวอย่างข้อมูลที่แบบสอนระบบที่แต่ละแถวของข้อมูลปแอทริบิวท์ระกอบด้วยฟิลด์จำนวนมาก อาจเป็นค่าต่อเนื่องหรือค่ากลุ่ม โดยจะมีแอทริบิวท์ซึ่งเป็นตัวบ่งชี้คลาสของข้อมูล จุดประสงค์ของการจำแนกประเภทข้อมูลคือการสร้างโมเดลการแยกแอทริบิวท์หนึ่งโดยขึ้นกับแอทริบิวท์อื่น โมเดลที่ได้จากการจำแนกประเภทข้อมูลจะทำให้สามารถพิจารณาคลาสในข้อมูลที่ยังมิได้แบ่งกลุ่มในอนาคตได้ เทคนิคการจำแนกประเภทข้อมูลนี้ได้นำไปประยุกต์ใช้ในหลายด้าน

2.2.1 ต้นไม้ตัดสินใจใช้ในการสนับสนุนการตัดสินใจ ซึ่งมีลักษณะเป็นโครงสร้างต้นไม้หัวกลับเป็นแบบจำลองทางคณิตศาสตร์ เพื่อการหาทางเลือกที่ดีที่สุด โดยการนำข้อมูลมาสร้างแบบจำลองการพยากรณ์ในรูปแบบของโครงสร้างต้นไม้ ส่วนประกอบของต้นไม้ตัดสินใจประกอบด้วย ก) โหนด (Node) คือคุณสมบัติต่าง ๆ เป็นจุดที่แยกข้อมูลว่าจะให้ไปในทิศทางใดซึ่งโหนดที่อยู่สูงสุดเรียกว่า โหนดราก (Root Node) ข) กิ่ง (Branch) คือคุณสมบัติของคุณสมบัติในโหนดที่แตกออกมาโดยจำนวนของกิ่งจะเท่ากับคุณสมบัติของโหนด และ ค) ใบ (Leaf) คือกลุ่มของผลลัพธ์ในการแยกแยะข้อมูล

2.2.2 การสุ่มป่าไม้เป็นเทคนิคที่ทำการสุ่มเลือกคุณสมบัติออกมาจากชุดข้อมูล จากนั้นนำเอาชุดของคุณสมบัติเหล่านั้นมาสร้างแบบจำลองด้วยเทคนิคต้นไม้ตัดสินใจ ลักษณะของต้นไม้ที่อยู่ภายในป่าของเทคนิคการสุ่มป่าไม้จะถูกควบคุมด้วย 3 ปัจจัย คือ 1. ต้นไม้แต่ละต้นจะถูกสอน (Train) โดยการใช้เซตย่อยจากข้อมูลตัวอย่าง 2. เมื่อต้นไม้โตขึ้นจะสามารถค้นหาโหนด แต่ละโหนดที่อยู่ในกิ่งที่ดีที่สุดของต้นไม้โดยใช้การสุ่มเลือกคุณสมบัติจาก N คุณสมบัติ และ 3. ต้นไม้แต่ละต้นจะไม่มีการตัดออก แต่จะปล่อยให้ต้นไม้โตขึ้นไปจนได้ผลลัพธ์ที่ดีที่สุด หลังจากการสร้างป่า แล้วทำการให้คะแนนโดยต้นไม้ภายในป่า หากต้นไม้ต้นใดได้คะแนนสูงสุดก็จะนำเอาต้นไม้นั้นออกมาสร้างเป็นโมเดล

2.2.3 การหาเพื่อนบ้านใกล้ที่สุดเป็นเทคนิคที่มีความแตกต่างจากเทคนิคอื่นที่ไม่มีการแบ่งข้อมูลชุดสอนในการสร้างแบบจำลอง แต่จะใช้ข้อมูลทั้งชุดมาเป็นแบบจำลองในการใช้งานเทคนิคนี้จำเป็นจะต้องมีการระบุค่าตัวเลขจำนวนเพื่อนบ้านหรือค่า K ซึ่งค่านี้จะเป็นตัวบอกจำนวนของกรณีที่จะต้องค้นหาในการทำนายกรณีใหม่ ซึ่งได้แก่ 1-NN, 2-NN, 3-NN, …. ,K-NN โดยที่ K เป็นจำนวนเต็มบวก เช่น 2-NN หมายถึงจะมีการค้นหา 2 กรณีที่มีลักษณะใกล้เคียงกับกรณีใหม่

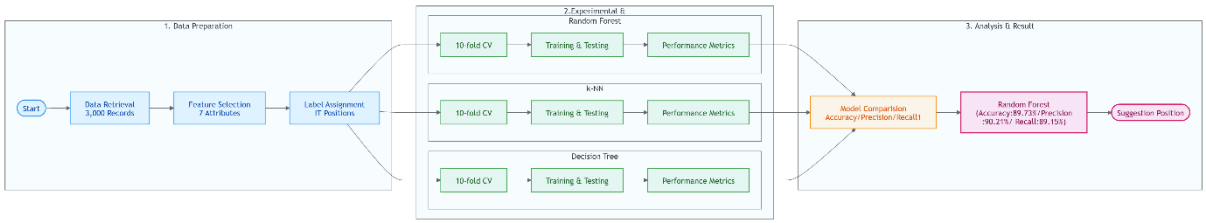
## 3. วิธีการดำเนินการวิจัย

กระบวนการมาตรฐานเพื่อใช้สำหรับการเรียนรู้ของเครื่องถูกนำมาใช้ในงานวิจัยนี้ โดย 1) เริ่มจากการเก็บความต้องการ กำหนดปัญหาและศึกษาปัญหาในการสมัครงานในประเทศไทย ทั้งภาครัฐบาล เอกชนและวิสาหกิจ จากนั้น 2) วิเคราะห์ข้อมูลจากเว็บไซต์ https://job.jobpub.com/ และเว็บไซต์อื่นที่เกี่ยวข้อง หลังจากนั้นเป็นการตรวจสอบ

ข้อมูลที่ได้ทำการรวบรวบมา เพื่อดูความถูกต้องและพิจารณาว่าจะใช้ข้อมูลทั้งหมดหรือจำเป็นต้องเลือกข้อมูลบางส่วนมาใช้ในการวิเคราะห์ แผนภาพยูเคสถูกนำใช้ในการแสดงขั้นตอนการทำงานของผู้ใช้งานและแอดมิน การสมัครสมาชิก เข้าสู่ระบบ การจัดการข้อมูล ทำนายตำแหน่งงาน แนะนำตำแหน่งงาน โพสต์ประกาศหางาน การค้นหาข้อมูลตำแหน่งงาน และการดูข้อมูลโพสต์ประกาศหางาน 3) จัดเตรียมข้อมูลที่เก็บได้จากเว็บไซต์มาจัดเก็บในโปรแกรม MS Excel โดยชุดข้อมูลเริ่มต้นมี 3,000 รายการ ผ่านกระบวนการทำความสะอาดข้อมูล เพื่อจัดการกับค่าที่สูญหายและความไม่ต่อเนื่องของข้อมูล 4) สร้างโมเดลด้วยโปรแกรมแรบบิทไมท์เนอร์สตูดิโอ จากเทคนิคการสุ่มป่าไม้ เทคนิคการหาเพื่อนบ้านที่ใกล้ที่สุด และเทคนิคต้นไม้ตัดสินใจ 5) คัดเลือกโมเดลโดยนำดาต้าเซทมาดำเนินการทดสอบประสิทธิภาพด้วยเทคนิคการตรวจสอบไขว้ 10 ส่วน (10-fold Cross-Validation) ซึ่งเป็นวิธีการทางสถิติที่มีความเข้มงวด โดยชุดข้อมูลจะถูกแบ่งออกเป็น 10 ส่วนเท่าๆ กัน ในแต่ละรอบของการทดสอบ ข้อมูล 9 ส่วน จะถูกใช้ในการฝึกฝนโมเดล (Training Set) และข้อมูลอีกหนึ่งส่วนจะถูกใช้ในการทดสอบ (Test Set) กระบวนการนี้จะดำเนินซ้ำจนครบ 10 ครั้ง โดยทุกส่วนของข้อมูลจะถูกนำมาเป็นชุดทดสอบอย่างน้อยหนึ่งครั้ง ค่าสถิติต่างๆ อันได้แก่ ค่าความถูกต้อง (Accuracy) ความเที่ยงตรง (Precision) และการเรียกคืน(Recall) แสดงในงานวิจัยนี้จึงเป็นค่าเฉลี่ยจากการทดสอบทั้ง 10 รอบ ซึ่งช่วยลดโอกาสการเกิดปัญหา Overfitting และยืนยันได้ว่าโมเดลมีประสิทธิภาพคงที่ดูที่ค่าความถูกต้อง ค่าความเที่ยงตรง และค่าการเรียกคืน 6) นำโมเดลไปใช้ให้เกิดประโยชน์ผู้ใช้เข้าถึงการใช้งานได้ง่าย

## 4. ผลการดำเนินงาน

ในงานวิจัยนี้ มีการคัดเลือกข้อมูลและเตรียมข้อมูลผู้สมัคร จำนวน 3000 ชุด ที่ผ่านกระบวนการทำความสะอาดข้อมูล และเป็นข้อมูลคุณภาพสูงที่มีการคัดกรองมาอย่างดีในกลุ่มอุตสาหกรรมไอทีโดยเฉพาะ ไม่ใช่ข้อมูลขยะ ส่วนข้อมูลตำแหน่งงาน มีจำนวน 50 ตำแหน่ง ใช้ตัวแปร 7 ลักษณะ ได้แก่ สาขา วุฒิการศึกษา ประสบการณ์ อายุ เพศ สถานะ ระดับภาษาอังกฤษ ซึ่งทีมวิจัยตั้งใจคัดเลือก 7 ตัวแปรที่เป็นเงื่อนไขพื้นฐานของการจ้างงานในอุตสาหกรรมไอที จากนั้น การสร้างจากโปรแกรมแรบบิทไมท์เนอร์สตูดิโอโดยเลือกโมเดลที่เหมาะสมที่สุดเพื่อทำนายตำแหน่งงาน โดยการทำนายจะทำนาย 2 ครั้ง โดยครั้งแรกเป็นการทำนายตามสาขาวิชาชีพ เพื่อเป็นการแบ่งกลุ่มแยกประเภทตำแหน่งงาน และการทำนายครั้งที่สองเป็นการทำนายตำแหน่งงาน และการตั้งค่าโอเปอร์เรเตอร์ในโปรแกรมเริ่มจาก 1) นำข้อมูลมาเชื่อมต่อโอเปอร์เรเตอร์เพื่อทำนายผล 2) ถัดมาเชื่อมต่อด้วยโอเปอร์เรเตอร์ Set Role เพื่อตั้งค่า Attribute Branch ให้เป็น Attribute Label เพื่อทำนายผล 3) เชื่อมต่อด้วยโอเปอร์เรเตอร์ Validation เพื่อนำอัลกอริทึมมาใช้ทำนาย 4) คลิกเลือกที่โอเปอร์เรเตอร์ Decision Tree Algorithm มาใช้ในการทำนายผล 5) เชื่อมต่อด้วยโอเปอร์เรเตอร์ Appy Model เพื่อสร้างโมเดล 6) ทำการ Select Attribute ใหม่ โดยการเชื่อมต่อจาก Validation เพื่อไม่ให้โปรแกรมนำ Attribute Brunch มาประมวลผลซ้ำและเป็นการทำให้โปรแกรมนำ Attribute Position มาทำนายผล 7) Set Role ใหม่อีกครั้งแต่ครั้งนี้จะเลือก Attribute Position เป็น Label ที่แท็ก Parameters เลือก Attribute Position จากนั้นให้เลือก Target Role เป็น Label 8) เชื่อมต่อโอเปอร์เรเตอร์ Validation อีกครั้งเพื่อนำ Algorithm มาใช้ทำนายผลรอบที่ 2 โดยวิธีทำนั้นคล้ายกับครั้งแรก 9) ถัดมาเชื่อมต่อด้วยโอเปอร์เรเตอร์ Appy Model อีกครั้งเพื่อให้โปรแกรมนั้นสร้างโมเดลรอบที่ 2 ออกมา 10) ถัดมาเชื่อมต่อด้วยโอเปอร์เรเตอร์ Performance เพื่อให้โปรแกรมแสดงผลของโมเดล ซึ่งการวัดประสิทธิภาพใช้เทคนิค 10-fold Cross Validation เพื่อให้มั่นใจว่าค่าความถูกต้อง ความเที่ยงตรง และการเรียกคืน ที่ได้มีความน่าเชื่อถือและลดการเกิด Overfitting จากชุดข้อมูล และผลการสร้างโมเดลพบว่า 1. เทคนิคต้นไม้ตัดสินใจได้ค่าความถูกต้อง 83.33% ค่าความเที่ยงตรง 82.95%และค่าการเรียกคืน 83.50% ส่วน 2. เทคนิคการหาเพื่อนบ้านที่ใกล้ที่สุดได้ค่าความถูกต้อง 87.33% ค่าความเที่ยงตรง 86.88% และค่าการเรียกคืน




**ภาพที่ 1:** สถาปัตยกรรมระบบเชิงตรรกะ

87.42% และ 3. เทคนิคการสุ่มป่าไม้ได้ค่าความถูกต้อง 89.73% ค่าความเที่ยงตรง 90.21% และค่าการเรียกคืน 89.15% ทีมวิจัยนำโมเดลที่มีประสิทธิภาพดีที่สุด ทำนายผลได้ถูกต้องแม่นยำที่สุดไปพัฒนาเป็นระบบแนะนำที่มีฟังก์ชันหลักของเว็บไซต์ในการทำนายตำแหน่งและแนะนำตำแหน่งงานให้กับผู้ใช้งาน จากภาพที่ 1 แสดงขั้นตอนการดำเนินงาน โดยเปลี่ยนจากภาพหน้าจอโปรแกรมเป็นแผนภาพสถาปัตยกรรมระบบเชิงตรรกะที่มีความคมชัดและระบุขั้นตอนการทำ 10-fold Cross Validation การเตรียมข้อมูล และการเปรียบเทียบอัลกอริทึมไว้อย่างครบถ้วนและถูกต้อง ระบบที่พัฒนาขึ้นเป็นรูปแบบของเว็บไซต์ที่มีหน้าหลัก ประกอบด้วยเมนูหน้าหลัก สมัครสมาชิก จัดการข้อมูล ตรวจสอบตำแหน่งงาน การวิเคราะห์ผล แนะนำตำแหน่งงานที่ใกล้เคียง โพสต์หางาน โดยภาพที่ 2 แสดงหน้าจอหลักของระบบการวิเคราะห์ตำแหน่งงานจากโมเดล การเลือกตำแหน่งงานอื่นเพิ่มเติมการประกาศหางาน การดูตำแหน่งงานใกล้เคียงเพิ่มเติม และหากต้องการประกาศหางานก็สามารถทำได้เลยจากระบบ และในส่วนการทดสอบประสิทธิภาพของระบบผ่านเว็บบราวเซอร์ที่นิยมใช้กันได้แก่ Google Chrome, Fire Fox และ Internet Explorer เพื่อดูความสามารถการใช้งานบนโลกออนไลน์ พบว่าระบบใช้งานได้จริงทุกเว็บบราวเซอร์ นอกจากนี้ การประเมินประสิทธิภาพระบบ โดยการทดสอบความพึงพอใจการใช้งานจากแบบสอบถามจากผู้ใช้งานพบว่ามีค่าเฉลี่ยโดยรวมเท่ากับ 4.17 มีความพึงพอใจอยู่ในระดับมากและส่วนเบี่ยงเบนมาตรฐานโดยรวมเท่ากับ 0.60 โดยรายละเอียดพบว่า ข้อที่ 1 รูปแบบในเว็บไซต์ง่ายต่อการอ่านและการใช้งานมีค่าเฉลี่ยเท่ากับ 4.43 มีส่วนเบี่ยงเบน




**ภาพที่ 2:** หน้าหลักระบบผู้ใช้งานและผลลัพธ์จากวิเคราะห์โมเดล

มาตรฐานเท่ากับ 0.56 มีระดับความความพึงพอใจอยู่ในระดับมากที่สุด และมีส่วนเบี่ยงเบนมาตรฐานค่อนข้างมาก ข้อที่ 2 หน้าโฮมเพจมีความสวยงาม มีความทันสมัยและน่าสนใจ มีค่าเฉลี่ยเท่ากับ 4.33 มีส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.61 มีระดับความความพึงพอใจอยู่ในระดับมากและมีส่วนเบี่ยงเบนมาตรฐานค่อนข้างมาก ข้อที่ 3 ขนาดตัวอักษรและรูปแบบตัวอักษร มีความสวยงามและอ่านได้ง่าย มีค่าเฉลี่ยเท่ากับ 3.96 มีส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.61 มีระดับความความพึงพอใจอยู่ในระดับมาก และมีส่วนเบี่ยงเบนมาตรฐานค่อนข้างมาก ข้อที่ 4 สีพื้นหลังกับสีตัวอักษรมีความเหมาะสมต่อการใช้งาน มีค่าเฉลี่ยเท่ากับ 4.2 มีส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.61 มีระดับความความพึงพอใจอยู่ในระดับมาก และมีส่วนเบี่ยงเบนมาตรฐานค่อนข้างมาก ข้อที่ 5 ความถูกต้องในการเชื่อมโยงภายในเว็บไซต์ มีค่าเฉลี่ยเท่ากับ 4.26 มีส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.56 มีระดับความความพึงพอใจอยู่ในระดับมาก และมีส่วนเบี่ยงเบนมาตรฐานค่อนข้างมาก จากตารางที่ 1 แสดงการเปรียบเทียบคุณลักษณะระหว่างระบบที่นำเสนอกับงานวิจัยที่เกี่ยวข้อง [8-10] พบว่าแม้ในงานวิจัย [8] และ [10] จะมีการใช้โครงสร้าง Deep Learning ที่ซับซ้อน แต่ระบบที่นำเสนอในงานวิจัยนี้มุ่งเน้นไปที่การใช้คุณลักษณะพื้นฐาน 7 ประการที่มีนัยสำคัญต่อการจ้างงานในสายอาชีพไอทีของประเทศไทยโดยผลการทดลองชี้ให้เห็นว่าการใช้เทคนิค

**ตารางที่ 1:** เปรียบเทียบคุณลักษณะระหว่างงานวิจัยที่นำเสนอกับงานวิจัยเดิม [8, 9, 10]

| คุณสมบัติ | งานวิจัยเดิม | งานวิจัยที่นำเสนอ |
|---|---|---|
| 1.การประมวลผลข้อมูลภาษาไทย | ส่วนใหญ่รองรับเฉพาะภาษาอังกฤษ | รองรับบริบทและภาษาไทย |
| 2.ความซับซ้อนของข้อมูล | ต้องการ Resume ฉบับเต็ม | ใช้เพียง 7 ตัวแปรพื้นฐาน |
| 3.อัลกอริทึมที่ใช้ | มักใช้โมเดลเดี่ยว | เปรียบเทียบ 3 อัลกอริทึม (การสุ่มป่าไม้ดีสุด) |
| 4.บริบทเป้าหมาย | ตลาดงานทั่วไป | เน้นเฉพาะเจาะจงสายงานไอที |

การสุ่มป่าไม้สามารถให้ค่าความแม่นยำสูงถึงร้อยละ 89.73 ซึ่งอยู่ในระดับที่ใกล้เคียงกับระบบที่ซับซ้อนกว่า จุดเด่นนี้แสดงให้เห็นว่าชุดข้อมูลและตัวแปรที่คัดเลือกมานั้นมีความเหมาะสมและเพียงพอต่อการแนะนำตำแหน่งงานอย่างมีประสิทธิภาพ อีกทั้งยังใช้ทรัพยากรในการประมวลผลน้อยกว่า ซึ่งเอื้อต่อการนำไปประยุกต์ใช้จริงในระบบรับสมัครงานที่มีการเปลี่ยนแปลงข้อมูลแบบทันทีทันใด

## 5. สรุปและอภิปรายผล

การศึกษาเปรียบเทียบกระบวนการวิเคราะห์โมเดลทำนายและแนะนำตำแหน่งงานซึ่งงานวิจัยนี้ได้นำเสนอเทคนิคการสุ่มป่าไม้ เทคนิคการหาเพื่อนบ้านที่ใกล้ที่สุด และเทคนิคต้นไม้ตัดสินใจ เพื่อนำมาสร้างโมเดลที่เหมาะสมที่สุด จากผลการทดลองพบว่า เทคนิคการสุ่มป่าไม้ให้ค่าความถูกต้องสูงสุดที่ 89.67% เนื่องจากเป็นอัลกอริทึมแบบ Ensemble ที่ช่วยลดปัญหา Overfitting จากข้อมูลที่มีตัวแปรจำกัด 7 ตัวแปรได้ดีกว่าเทคนิคต้นไม้ตัดสินใจทั่วไป ส่งผลให้การจับคู่ตำแหน่งงานมีความสมดุลทั้งในแง่ความแม่นยำ และความครอบคลุม ผลลัพธ์จากการศึกษาวิจัยถูกนำไปพัฒนาระบบแนะนำตำแหน่งงานสำหรับองค์กร โดยระบบประกอบฟังก์ชันการจัดการข้อมูล แก้ไขข้อมูล ทำนายตำแหน่งงาน แนะนำตำแหน่งอื่นที่ใกล้เคียงและการประกาศหางาน ในส่วนการประเมินผล ระบบจะถูกประเมินด้วยแบบสอบถามจากผู้ใช้งานทั่วไป 30 คน โดยใช้หลักเกณฑ์การแปลความหมายของค่าคะแนนเฉลี่ยและหลักเกณฑ์การแปลความหมายส่วนเบี่ยงเบนมาตรฐานและฟังกชันที่ใช้ในการพิจารณามีอยู่ทั้งหมด 5 กลุ่มด้วยกันคือ 1. ความง่าย 2. ความเหมาะสม 3. การเลือกใช้ 4. ความสามารถ และ 5. ความตรงความต้องการซึ่งผลรวมที่ได้จากผู้ใช้งานทั่วไปมีค่าเฉลี่ยเท่ากับ 4.17 โดยแปลความหมายว่ามีความพึงพอใจอยู่ในระดับมากและมีส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.60 โดยแปลความหมายได้ว่ามีส่วนเบี่ยงเบนมาตรฐานค่อนข้างมาก เมื่อเปรียบเทียบกับงานวิจัยของ Guo และคณะ [6] ที่ได้ค่าความแม่นยำในระดับใกล้เคียงกัน งานวิจัยที่นำเสนอนี้สามารถบรรลุประสิทธิภาพที่ 89.73% โดยใช้เพียงตัวแปรคุณลักษณะที่เข้าถึงได้ง่ายกว่า ซึ่งพิสูจน์ให้เห็นว่าความละเอียดของข้อมูลพื้นฐาน 7 ตัวแปรนั้นเพียงพอสำหรับการคัดกรองเบื้องต้นในอุตสาหกรรมไอทีของไทย

## เอกสารอ้างอิง


[1] J. Dhameliya and N. Desai, "Job recommender systems: A survey," in *Proc. 3rd Int. Conf. Intell. Comput. Control Syst. (ICICCS)*, 2019, pp. 159–163.

[2] M. N. Freire and L. N. de Castro, "e-Recruitment recommender systems: A systematic review," Knowl. Inf. Syst., vol. 63, no. 1, pp. 1–20, Jan. 2021.

[3] A. Gunawardana, G. Shani, and S. Yogev, "*Evaluating recommender systems*," in Recommender Systems Handbook, F. Ricci, L. Rokach, and B. Shapira, Eds., 3rd ed. New York, NY, USA: Springer, 2022, pp. 547–601.

[4] M. H. H. Hisham, M. A. Aziz, and A. A. Sulaiman, "Job classification: A review on data, feature and methods," *J. Electr. Electron. Syst. Res.*, vol. 14, pp. 91–100, 2021.

[5] K. P. Murphy, *Machine Learning: A Probabilistic Perspective*. Cambridge, MA, USA: MIT Press, 2012.

[6] S. Guo, F. Alamudun, and T. Hammond, "RésuMatcher: A personalized résumé-job matching system," *Expert Syst. Appl.*, vol. 60, pp. 169–182, Oct. 2016.

[7] F. O. Isinkaye, Y. O. Folajimi, and B. A. Ojokoh, "Recommendation systems: Principles, methods and evaluation," *Egypt. Informatics J.*, vol. 16, no. 3, pp. 261–273, Nov. 2015.

[8] C. Bian, et al., "Learning to Match Jobs with Resumes from Sparse Interaction Data using Deep Semantic Framework," *Proc. IEEE 36th Int. Conf. Data Eng. (ICDE)*, 2020, pp. 61–72.

[9] S. Pumma and T. Sanguansat, "Job Recommendation System using Word Embedding and Machine Learning Techniques," in *Proc. 21st Nat. Conf. Comput. Inf. Technol. (NCCIT)*, 2025, pp. 45–50.

[10] P. Thipwiwatpotjana and K. Tuicomepee, "A deep learning-based career recommendation system using program learning outcomes and TPQI standards in the digital industry," *Eurasian J. Appl. Sci. Technol.*, vol. 9, no. 2, pp. 1–15, Jul. 2025.

# การจำแนกวัตถุประสงค์อีเมลด้วยวิธีการเรียนรู้ของเครื่อง
# E-Mail Objective Classification using Machine Learning

*ธีระชัย เกษมพงศ์จรัส (Theerachai Kasemphongjarat)[1] และมาลีรัตน์ มะลิแย้ม (Maleerat Maliyaem)[2]*
*ภาควิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล*
*มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ*
*[1]s6707011857071@email.kmutnb.ac.th, [2]maleerat.m@itd.kmutnb.ac.th*

**บทคัดย่อ**

*งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาวิธีการจำแนกวัตถุประสงค์อีเมลด้วยวิธีการเรียนรู้ของเครื่อง ให้มีประสิทธิภาพในการจัดการข้อมูลในองค์กรและเพิ่มประสิทธิภาพการทำงาน โดยให้สามารถจำแนก ประเภทของอีเมลได้ 6 วัตถุประสงค์ ได้แก่ งานตอบคำถามทั่วไป งานส่งใบเสนอสินค้า งานบริการติดตั้ง งานตรวจสอบเอกสารข้อกำหนด งานร่างเอกสารข้อกำหนด และงานส่งเอกสารอ้างอิงข้อกำหนด โดยการรวบรวมอีเมลกลุ่มตัวอย่างจำนวน 9,569 ฉบับ แบ่งเป็นจำนวนตัวอย่างที่ใช้ในการเรียนรู้ 7,655 ฉบับและตัวอย่างที่ใช้ในการทดสอบและวัดผล 1,914 ฉบับ และสร้างแบบจำลองเปรียบเทียบ 3 วิธีการเรียนรู้ ได้แก่ เบย์สแบบง่าย, เพื่อนบ้านใกล้ที่สุด และหน่วยความจำระยะยาวแบบสั้น งานวิจัยจะแบ่งเป็น 2 ขั้นตอนดังนี้ ขั้นตอนแรกคือการจำแนกประเภทอีเมลตามวัตถุประสงค์หลัก โดยเปรียบเทียบ 3 แบบจำลอง ผลการทดลองพบว่าวิธีการเรียนรู้หน่วยความจำระยะยาวแบบสั้นให้ค่าความถูกต้องร้อยละ 96 และเพิ่มสูงขึ้นเป็นร้อยละ 99 หลังผ่านการเพิ่มประสิทธิภาพด้วยวิธีการค้นหาแบบผสมผสานและการสกัดวลีสำคัญ ขั้นตอนที่สองเป็นการพัฒนาแบบจำลองให้รองรับการจำแนกพร้อมกันสองวัตถุประสงค์สำหรับอีเมลที่มีวัตถุประสงค์ซ้อนทับกัน ผลการทดลองพบว่า วิธีการจำแนกสองวัตถุประสงค์ให้ค่าความถูกต้องร้อยละ 96.77 นอกจากนี้ผู้วิจัยได้นำแบบจำลองมาพัฒนาเป็นโปรแกรมประยุกต์เพื่อทดลองใช้ในการปฏิบัติงานจริงซึ่งให้ผลเป็นที่น่าพอใจอีกด้วย*

**คำสำคัญ:** การเรียนรู้ของเครื่อง การประมวลผลภาษาธรรมชาติ การจำแนกอีเมล

**Abstract**

*This research aims to develop a machine learning approach for email objective classification to improve organizational efficiency. Emails are classified into six categories: general inquiry response, product quotation, installation service, requirement document inspection, requirement document drafting, and reference document submission. A total of 9,569 email samples were collected, with 7,655 used for training and 1,914 for testing. Three algorithms such as Naïve Bayes, K-Nearest Neighbors, and Long Short-Term Memory were compared. The study consists of two phases. In the first phase, single objective classification was performed, where LSTM achieved the highest accuracy of 96%, improving to 99% after optimization using hybrid search and key phrase extraction. In the second phase, the model was extended to support dual objective classification for overlapping email purposes, achieving an accuracy of 96.77%. The model was further implemented as an application for real-world use, demonstrating satisfactory performance. The results indicate that the proposed approach is effective for enhancing email management in organizations.*

**Keywords:** Machine Learning, Natural Language Processing, E-mail Classification.

## 1. บทนำ

ในยุคที่เทคโนโลยีก้าวหน้าอย่างรวดเร็ว องค์กรต่าง ๆ กำลังขับเคลื่อนธุรกิจด้วยข้อมูลในปริมาณมหาศาล โดยเฉพาะข้อมูลดิจิทัลในรูปแบบของเอกสารและข้อความ ซึ่งเป็นปัจจัย

สำคัญและมีการเพิ่มขึ้นอย่างต่อเนื่องในแต่ละวัน การจัดการและประมวลผลข้อมูลเหล่านี้จึงมีความสำคัญอย่างยิ่งต่อการดำเนินงานและสร้างความได้เปรียบทางธุรกิจ หลายองค์กรจึงมุ่งเน้นการพัฒนาระบบอัตโนมัติเพื่อจัดการกับข้อมูลจำนวนมาก โดยเฉพาะอย่างยิ่งในภาคส่วนบริการลูกค้า การจัดประเภทอีเมลโดยอัตโนมัติสามารถช่วยจัดลำดับความสำคัญของข้อความ เพิ่มความรวดเร็วในการตอบสนอง และส่งผลให้ความพึงพอใจของลูกค้าเพิ่มขึ้น

อย่างไรก็ตาม แม้อีเมลยังคงเป็นช่องทางการสื่อสารหลักที่สำคัญในองค์กรและสถาบันต่าง ๆ [1][2] แต่ข้อมูลที่ส่งผ่านอีเมลเหล่านี้ส่วนใหญ่เป็นข้อมูลที่ไม่มีโครงสร้าง [3] การสกัดข้อมูลจากอีเมลจำนวนมหาศาลจึงกลายเป็นความท้าทายอย่างมาก [4][5] ปัญหาหลักคือปริมาณอีเมลที่มีอยู่เป็นจำนวนมาก ประกอบกับข้อจำกัดด้านเวลาของผู้ปฏิบัติงาน ซึ่งส่งผลให้องค์กรไม่สามารถดึงข้อมูลเชิงลึกที่เป็นประโยชน์ออกมาใช้ในการตัดสินใจได้อย่างเต็มที่

สิ่งที่ทำให้งานวิจัยนี้แตกต่างจากงานด้านการจำแนกอีเมลที่ผ่านมา [1][2][5] คือการมุ่งเน้นที่ (1) อีเมลภาษาไทยในบริบทองค์กรจริง ซึ่งเป็นภาษาที่มีความซับซ้อนด้านโครงสร้าง (2) การจำแนกแบบหลายป้าย (Multi-label) สำหรับอีเมลที่มีวัตถุประสงค์ซ้อนทับ และ (3) การนำแบบจำลองไปใช้งานจริงในองค์กร ซึ่งยังขาดการศึกษาอย่างเป็นระบบในบริบทนี้

## 2. งานวิจัยที่เกี่ยวข้อง

### 2.1 การประมวลผลภาษาธรรมชาติภาษาไทย

ภาษาไทยเป็นภาษาประจำชาติที่จัดอยู่ในตระกูลภาษาขร้า–ไท มีผู้ใช้ทั่วโลกประมาณ 85.7 ล้านคน แต่การพัฒนาเทคโนโลยีการประมวลผลภาษาธรรมชาติยังได้รับการสนับสนุนน้อยกว่าภาษาอื่น เช่น อังกฤษ จีน หรือเยอรมัน ความท้าทายสำคัญคือรูปแบบการเขียนแบบต่อเนื่อง (Scriptio Continua) ที่ไม่มีการเว้นวรรคแบ่งคำชัดเจน รวมถึงปัญหาความกำกวมของขอบเขตคำและประโยค ส่งผลให้กระบวนการ Word Tokenization มีความซับซ้อน PyThaiNLP [6] เป็นไลบรารีประมวลผลภาษาไทยแบบโอเพนซอร์สบน Python ที่รองรับงาน NLP ภาษาไทยได้อย่างมีประสิทธิภาพ

### 2.2. Machine Learning

**2.2.1 Long Short-Term Memory (LSTM)** [7]

โครงข่ายประสาทเทียมหน่วยความจำระยะสั้นแบบยาว เป็นเทคนิคที่ได้รับการออกแบบมาโดยเฉพาะเพื่อรองรับปัญหา Long-Term Dependencies ผ่านกลไก Gates สามรูปแบบ ได้แก่ Forget Gate, Input Gate และ Output Gate โครงสร้าง LSTM แสดงประสิทธิภาพที่โดดเด่นในงานวิเคราะห์เสียง (Speech Recognition) หรือ โมเดลภาษา (Language Modeling) และการจำแนกข้อความ ในบริบทของ NLP สำหรับภาษาไทย LSTM มีความเหมาะสมอย่างยิ่งเนื่องจากสามารถจับความสัมพันธ์เชิงลำดับระหว่างคำได้โดยไม่ต้องอาศัยขอบเขตคำที่ชัดเจน

**2.2.2 Naïve Bayes** [8]

เป็น Algorithm ที่มีรากฐานมาจาก Bayes Theorem โดยประยุกต์ใช้การสร้างแบบจำลอง Bayesian เพื่อประมาณค่า Posterior Probability ของแต่ละ Class แม้ว่าจะอาศัยสมมติฐานว่า Features มีความเป็นอิสระต่อกัน แต่ก็ยังให้ผลลัพธ์ที่น่าประทับใจในงานจำแนกข้อความ โดยเฉพาะเมื่อข้อมูลมีขนาดจำกัด

**2.2.3 K-Nearest Neighbors** [9]

เป็นแนวทางการจำแนกประเภทแบบ Non-parametric ที่ดำเนินการโดยการสืบค้น Nearest Neighbors จำนวน k ตัว แล้วใช้หลักการ Majority Voting ในการตัดสินใจจำแนกประเภท มีความเรียบง่ายแต่มีข้อจำกัดด้านประสิทธิภาพเมื่อข้อมูลมีขนาดใหญ่

### 2.3 การเปรียบเทียบ BERT และ Transformer

งานวิจัยสมัยใหม่ด้านการจำแนกข้อความมักอาศัยสถาปัตยกรรม Transformer และโมเดลภาษาขนาดใหญ่ เช่น BERT [11] ซึ่งใช้กลไก Self-Attention ในการเข้าใจบริบทของข้อความได้ครอบคลุมกว่า LSTM อย่างไรก็ตาม BERT และ Transformer มีข้อจำกัดด้านทรัพยากรในการฝึกและปรับแต่ง และภาษาไทยยังขาดโมเดล Pre-trained คุณภาพสูง งานวิจัยนี้จึงเลือก LSTM เป็นแบบจำลองหลัก เนื่องจากเหมาะสมกับขนาดข้อมูลที่มี สามารถจับความสัมพันธ์เชิงลำดับของภาษาไทยได้ดี และรองรับการนำไปใช้งานจริงในองค์กรได้อย่างสะดวก

### 2.4 Multi-label Text Classification

การจำแนกประเภทข้อความแบบหลายฉลาก [10] เป็นเทคนิคสำคัญที่ช่วยให้ข้อความหนึ่งชุดสามารถถูกจัดอยู่ในหลายหมวดหมู่พร้อมกัน เช่น อีเมลฉบับเดียวอาจเป็นทั้งการสอบถามข้อมูลและการส่งเอกสารไปพร้อมกัน ซึ่งพบได้บ่อยในบริบทอีเมลองค์กรจริง ในทางทฤษฎี NLP วิธีการ Binary Relevance แปลงปัญหา Multi-label ให้เป็นชุดของปัญหา Binary Classification อิสระ ซึ่งเหมาะสมสำหรับกรณีที่ Label มีจำนวนไม่มาก

### 2.5 ความแตกต่างจากงานวิจัยที่เกี่ยวข้อง

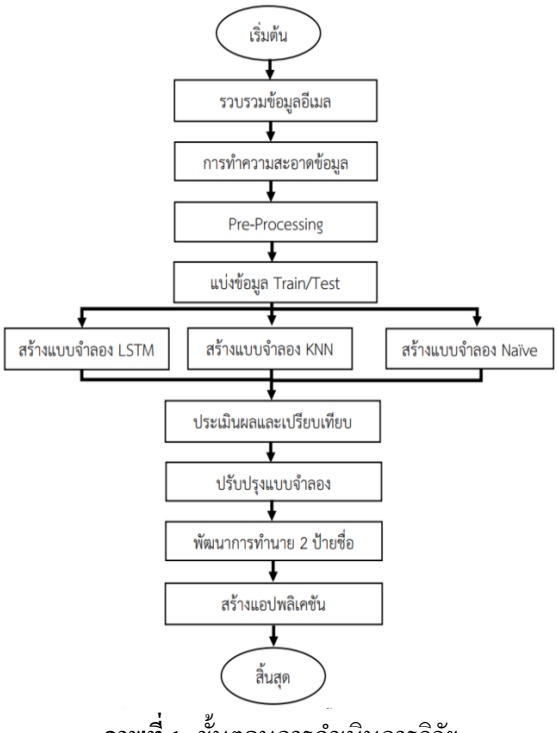
งานวิจัยด้านการจำแนกอีเมลที่ผ่านมา [1][2][5] ส่วนใหญ่มุ่งเน้นที่การจำแนก Spam/Ham หรือหมวดหมู่ทั่วไปในภาษาอังกฤษ งานวิจัยนี้มีความแตกต่างใน 3 ด้านสำคัญ

**ตารางที่ 1:** การเปรียบเทียบงานวิจัยในอดีต

| ด้าน | งานวิจัยทั่วไป | งานวิจัยนี้ |
|---|---|---|
| ภาษา | ภาษาอังกฤษ | ภาษาไทย |
| ประเภทงาน | Spam/Ham detection | 6 ประเภทงานองค์กร |
| การจำแนก | Single label | Multi label |
| บริบท | ทั่วไป | อีเมลองค์กรไทยจริง |

## 3. วิธีการดำเนินกำรวิจัย

### 3.1. แผนการดำเนินการ




**ภาพที่ 1:** ขั้นตอนการดำเนินการวิจัย

### 3.2 การเก็บรวบรวมข้อมูล

การเก็บรวบรวมข้อมูลอีเมลกลุ่มตัวอย่างจำนวน 9,569 ฉบับ จากระบบ Microsoft 365 สำหรับการวิเคราะห์และพัฒนาแบบจำลองการเรียนรู้ของเครื่อง โดยข้อมูลที่ดึงมาได้จะสะท้อนลักษณะการสื่อสารภายในองค์กร เช่น เนื้อหาอีเมล ผู้ส่ง ผู้รับ วันที่และเวลาที่ส่ง และหัวข้ออีเมล ซึ่งจะถูกจัดเก็บในรูปแบบ JSON file ภายในคอมพิวเตอร์เพื่อความสะดวกในการเตรียมข้อมูลและนำไปใช้ในขั้นตอนถัดไปของการเรียนรู้ของเครื่อง

### 3.3 การเตรียมการข้อมูล

#### 3.3.1 การแปลงชุดข้อมูลอีเมลจาก JSON เป็น CSV

ข้อมูลอีเมลที่รวบรวมได้ในรูปแบบ JSON จะถูกแปลงให้อยู่ในรูปแบบ CSV ที่มีโครงสร้างชัดเจน เพื่อให้สะดวกต่อการนำไปวิเคราะห์ในขั้นตอนถัดไป กระบวนการนี้เริ่มจากการโหลดไฟล์ JSON และดึงฟิลด์สำคัญ เช่น เนื้อหาอีเมล หัวข้อ และข้อมูลผู้ส่ง จากนั้นจึงทำความสะอาดเนื้อหาเบื้องต้นด้วยการลบลิงก์ URL ส่วน Signature และคำลงท้ายที่ไม่เกี่ยวข้อง ก่อนบันทึกลงในไฟล์ CSV เพื่อใช้เป็นฐานข้อมูลสำหรับการสร้าง Label ต่อไป

#### 3.3.2 การตัดแบ่งคำภาษาไทย

เนื่องจากภาษาไทยเขียนแบบต่อเนื่องโดยไม่มีการเว้นวรรคระหว่างคำ การระบุขอบเขตของแต่ละคำจึงเป็นขั้นตอนที่จำเป็นก่อนการวิเคราะห์ข้อความ งานวิจัยนี้ใช้ PyThaiNLP ในการ Tokenize ข้อความ ซึ่งเป็นแนวทางมาตรฐานสำหรับ NLP ในภาษากลุ่ม Scriptio Continua ผลลัพธ์ที่ได้คือข้อความที่ถูกแบ่งเป็นคำแต่ละคำและคั่นด้วยช่องว่าง

#### 3.3.3 Data Cleaning

หลังจากตัดแบ่งคำแล้ว ข้อมูลจะผ่านกระบวนการทำความสะอาด 3 รอบตามลำดับ รอบแรกลบคำสุภาพและสัญลักษณ์ที่ไม่มีความหมายเชิงวิเคราะห์ รอบที่สองลบชื่อบุคคลและองค์กรที่อาจทำให้แบบจำลองเรียนรู้จากข้อมูลที่ไม่ควรนำมาพิจารณา และรอบที่สามลบตัวเลขและตัวอักษรเดี่ยว ผลลัพธ์สุดท้ายคือชุดข้อมูลที่สะอาดและเป็นมาตรฐาน พร้อมนำเข้าสู่กระบวนการสร้างแบบจำลองต่อไป

**ตารางที่ 2:** การเปรียบเทียบผลการทำความสะอาด

| กระบวนการ | เนื้อหา | คำที่ถูกตัดออก |
|---|---|---|
| ต้นฉบับ | เรียนทีมโครงการ ขอแจ้งการประชุมสรุปความคืบหน้าโครงการ 678901 ในวันศุกร์ที่ 5 มกราคม เวลา 10:00 น. ห้องประชุมชั้น 3 กรุณานำรายงานความคืบหน้าและปัญหาที่พบมาด้วยครับ ขอบคุณครับ สมชาย | ไม่มี |
| การตัดแบ่งคำภาษาไทย | เรียน ทีม โครงการ ขอ แจ้ง การ ประชุม สรุป ความ คืบหน้า โครงการ 678901 ใน วัน ศุกร์ ที่ 5 มกราคม เวลา 10 : 00 น . ห้อง ประชุม ชั้น 3 กรุณา นำ รายงาน ความ คืบหน้า และ ปัญหา ที่ พบ มา ด้วย ครับ ขอบคุณ ครับ สมชาย | ไม่มี |
| ลบคำไม่สำคัญและสัญลักษณ์ | ทีม โครงการ ขอ แจ้ง การ ประชุม สรุป ความ คืบหน้า โครงการ 678901 ใน วัน ศุกร์ ที่ 5 มกราคม เวลา 10 00 น ห้อง ประชุม ชั้น 3 กรุณา นำ รายงาน ความ คืบหน้า และ ปัญหา ที่ พบ มา ด้วย ขอบคุณ สมชาย | “เรียน”, “ครับ”, “เครื่องหมายทวิภาค (:)”, “เครื่องหมายมหัพภาค (.)” |
| ลบชื่อบุคคลและหน่วยงาน | ขอ แจ้ง การ ประชุม สรุป ความ คืบหน้า โครงการ 678901 ใน วัน ศุกร์ ที่ 5 มกราคม เวลา 10 00 น ห้อง ประชุม ชั้น 3 กรุณา นำ รายงาน ความ คืบหน้า และ ปัญหา ที่ พบ มา ด้วย ขอบคุณ | “ทีม”, “โครงการ”, “สมชาย” |
| ลบตัวอักษรเดี่ยวและตัวเลข | ขอ แจ้ง การ ประชุม สรุป ความ คืบหน้า โครงการ ใน วัน ศุกร์ ที่ มกราคม เวลา ห้อง ประชุม ชั้น กรุณา นำ รายงาน ความ คืบหน้า และ ปัญหา ที่ พบ มา ด้วย ขอบคุณ | “678901”, “5”, “10”, “00”, “น”, “3” |

### 3.4. การสร้างแบบจำลอง (Modeling)

ในขั้นตอนนี้จะสร้างแบบจำลองการเรียนรู้ของเครื่องโดยใช้โมเดล Naïve Bayes, K-Nearest Neighbors และ LSTM ส่วนเครื่องมือที่ใช้ได้แก่ Python โดยใช้งานผ่าน IDE สำหรับ LSTM ใช้สถาปัตยกรรม Bidirectional LSTM ซึ่งสามารถรับข้อมูลจากสองทิศทางของลำดับคำ ทำให้เข้าใจบริบทของประโยคได้ดีกว่าแบบทิศทางเดียว

### 3.5 การประเมินผล (Evaluation)

ในขั้นตอนนี้จะทำการประเมินประสิทธิภาพของแบบจำลองที่พัฒนาขึ้น การทดสอบแบบจำลองจะดำเนินการโดยใช้ชุดข้อมูลทดสอบที่แยกออกจากชุดข้อมูลการฝึกฝน โดยใช้เกณฑ์การประเมิน Accuracy, Precision, Recall และ F1-score

## 4. ผลการดำเนินงานวิจัย

### 4.1 ผลการเปรียบเทียบประสิทธิภาพ

จากการเปรียบเทียบประสิทธิภาพของแบบจำลองการเรียนรู้ของเครื่องทั้ง 3 รูปแบบ พบว่าแบบจำลอง Long Short-Term Memory (LSTM) เป็นเทคนิคที่ดีที่สุด โดยมีค่าความแม่นยำสูงสุดที่ร้อยละ 96 ผลนี้สอดคล้องกับทฤษฎีที่ว่า LSTM มีความสามารถในการจับความสัมพันธ์ระยะไกลระหว่างคำ (Long-Term Dependencies) ได้ดีกว่า Naïve Bayes ที่สมมติว่า Features เป็นอิสระต่อกัน และดีกว่า KNN ที่อาศัยเพียงความใกล้เคียงใน Feature Space

**ตารางที่ 3:** การเปรียบเทียบประสิทธิภาพของแบบจำลองทั้ง 3 เทคนิค

| ตัวชี้วัดประสิทธิภาพ (Metrics) | LSTM | NVB | KNN |
|---|---|---|---|
| **Accuracy (ความแม่นยำรวม)** | 0.96 | 0.81 | 0.90 |
| **Macro Avg (F1-score)** | 0.83 | 0.58 | 0.71 |
| **F1-score แยกตามกลุ่ม** | | | |
| - งานตอบคำถามทั่วไป | 0.97 | 0.82 | 0.92 |
| - งานส่งใบเสนอสินค้า | 0.98 | 0.92 | 0.93 |
| - งานบริการติดตั้ง | 0.67 | 0.40 | 0.60 |
| - งานตรวจสอบเอกสารข้อกำหนด | 0.78 | 0.31 | 0.56 |
| - งานส่งเอกสารอ้างอิงข้อกำหนด | 0.86 | 0.62 | 0.67 |
| - งานร่างเอกสารข้อกำหนด | 0.69 | 0.42 | 0.60 |

### 4.2 การปรับปรุงประสิทธิภาพแบบจำลอง

ผู้วิจัยได้ปรับปรุงแบบจำลอง Bi-directional LSTM เพื่อเพิ่มความแม่นยำในการจำแนกประเภทข้อความ โดยมีการพัฒนาสามด้านหลัก ดังนี้

1. การคัดเลือกข้อมูลคุณภาพ: ใช้เฉพาะข้อมูลที่มี

คุณภาพสูงและมีป้ายกำกับชัดเจน ตัดข้อมูลที่ซ้ำซ้อนออก ซึ่งสอดคล้องกับหลักการ Garbage In, Garbage Out ในวงการ Machine Learning

2. การสกัดวลีสำคัญด้วย Hybrid Search: พัฒนาระบบค้นหาแบบพิเศษที่สกัดเฉพาะวลีสำคัญจากข้อความ ทำให้สามารถลดความยาวข้อความที่ต้องวิเคราะห์จาก 100 คำ เหลือเพียง 2 คำ แนวทางนี้สอดคล้องกับทฤษฎีด้าน Feature Selection ใน NLP ที่ระบุว่าการลดมิติของข้อมูล (Dimensionality Reduction) ช่วยให้แบบจำลองมุ่งเน้นเฉพาะสัญญาณที่มีความหมาย ลด Noise และปัญหา Curse of Dimensionality

3. การปรับแต่งการเรียนรู้: ปรับเปลี่ยนวิธีการคำนวณผลลัพธ์ให้เหมาะสมกับการจำแนกหลายประเภท เพิ่มจำนวนรอบการฝึกฝนจาก 10 เป็น 100 รอบ และลดขนาดกลุ่มข้อมูลในแต่ละรอบจาก 32 เหลือ 4 การลด Batch Size มีผลให้ Gradient Updates มีความถี่สูงขึ้น ทำให้แบบจำลองสามารถปรับตัวและจดจำรูปแบบได้แม่นยำยิ่งขึ้น

ผลจากการเพิ่มประสิทธิภาพแบบจำลอง (Model Optimization) แสดงให้เห็นถึงการพัฒนาในเชิงบวกอย่างมีนัยสำคัญ โดยค่าความแม่นยำโดยรวมเพิ่มขึ้นจากร้อยละ 96 เป็นร้อยละ 99

**ตารางที่ 4:** เปรียบเทียบประสิทธิภาพหลังปรับปรุง (Optimized Bi-LSTM)

| Label | Precision | Recall | F1-score |
|---|---|---|---|
| งานตอบคำถามทั่วไป | 0.98 | 1 | 0.99 |
| งานส่งใบเสนอสินค้า | 1 | 1 | 1 |
| งานบริการติดตั้ง | 1 | 1 | 1 |
| งานตรวจสอบเอกสารข้อกำหนด | 0.93 | 0.84 | 0.88 |
| งานส่งเอกสารอ้างอิงข้อกำหนด | 1 | 0.97 | 0.99 |
| งานร่างเอกสารข้อกำหนด | 0.82 | 0.86 | 0.84 |
| **accuracy** | | | 0.99 |
| **macro avg** | 0.96 | 0.94 | 0.95 |

### 4.3 การปรับปรุงแบบจำลองสำหรับทำนาย 2 ป้ายชื่อ

ผู้วิจัยได้พัฒนาแบบจำลอง Bi-LSTM ให้สามารถตรวจจับและจำแนกข้อความที่มีหลายความหมายซ้อนทับกันในประโยคเดียว โดยใช้หลักการจัดลำดับความสำคัญตามตำแหน่งที่คำปรากฏ การตรวจจับและจัดลำดับความสำคัญ ระบบสแกนหาวลีสำคัญทั้งหมดในประโยค พร้อมบันทึกตำแหน่งที่พบ จากนั้นเรียงลำดับตามตำแหน่งการปรากฏ โดยวลีที่ปรากฏก่อนจะถูกกำหนดเป็นคลาสหลัก และวลีที่ปรากฏที่หลังเป็นคลาสรอง โดยกระบวนการพยากรณ์แบ่งเป็นสองขั้นตอน ขั้นแรกนำวลีแรกเข้าแบบจำลองเพื่อพยากรณ์คลาสหลัก ขั้นที่สองหากพบวลีที่สองจะนำเข้าแบบจำลองอีกครั้งเพื่อพยากรณ์คลาสรอง แต่หากพบเพียงวลีเดียวระบบจะกำหนดคลาสรองเป็นค่าว่างโดยอัตโนมัติ

การประเมินผล ยังคงใช้สถาปัตยกรรม Bi-LSTM กับ Sigmoid และ Binary Cross Entropy โดยวัดประสิทธิภาพสามด้าน ได้แก่ ความแม่นยำของคลาสหลัก ความแม่นยำของคลาสรอง และความแม่นยำรวมเมื่อพยากรณ์ถูกทั้งสองคลาสพร้อมกัน โดยได้ผลลัพธ์ตามตารางที่ 5

**ตารางที่ 5:** เปรียบเทียบประสิทธิภาพการทำนาย 2 ป้ายชื่อ

| ประเภทความแม่นยำ | เปอร์เซ็นต์ (%) | จำนวนตัวอย่าง |
|---|---|---|
| Main Class Accuracy | 100 | 217 |
| Second Class Accuracy | 96.77 | |
| Combined Accuracy | 96.77 | |

### 4.4 การสร้างแอปพลิเคชัน

ผู้วิจัยได้สร้างโปรแกรมเพื่อวิเคราะห์อีเมลโดยใช้แบบจำลอง LSTM ที่สามารถทำนายได้สูงสุด 2 ประเภทพร้อมกัน เครื่องมือที่ใช้ประกอบด้วย TensorFlow และ PyInstaller ที่ใช้รวมโปรแกรมทั้งหมดเป็นไฟล์ .exe เดียว แอปพลิเคชันรับ Input Folder ของไฟล์ JSON และส่งออก Output Folder ของ CSV ที่มีผลการพยากรณ์ ผู้ใช้สามารถกดปุ่ม Start Prediction เพื่อเริ่มกระบวนการวิเคราะห์โดยไม่ต้องมีความรู้ด้านการเขียนโปรแกรม ซึ่งเป็นการเชื่อมโยงผลลัพธ์ทางวิชาการสู่การใช้งานจริงในองค์กร

## 5. สรุปผลการวิจัย

งานวิจัยนี้ประสบความสำเร็จในการพัฒนาแบบจำลองการเรียนรู้ของเครื่องสำหรับการจำแนกวัตถุประสงค์อีเมลภาษาไทยในบริบทองค์กร โดยแบบจำลอง LSTM แสดงประสิทธิภาพสูงสุดด้วยความแม่นยำร้อยละ 96 และเพิ่มขึ้นเป็นร้อยละ 99 หลังจากผ่านกระบวนการ Model

Optimization ซึ่งสูงกว่า KNN (ร้อยละ 90) และ Gaussian Naive Bayes (ร้อยละ 81) อย่างชัดเจน ผลดังกล่าวสะท้อนให้เห็นว่าสถาปัตยกรรม Recurrent Neural Network มีความเหมาะสมเป็นอย่างยิ่งในการจัดการข้อมูลเชิงลำดับที่มีความซับซ้อน อย่างเช่นข้อความภาษาไทย

ในด้านผลงานทางวิชาการ งานวิจัยนี้มีคุณูปการที่สำคัญใน 3 ประการ ประการแรกคือการพิสูจน์ประสิทธิภาพของ LSTM สำหรับการจำแนกอีเมลภาษาไทยในบริบทองค์กร ซึ่งนับเป็น NLP Domain ที่ยังขาดการศึกษาอย่างเป็นระบบ ประการที่สองคือการพัฒนาระบบ Multi-label Classification ด้วยการผสาน Binary Relevance เข้ากับ LSTM จนสามารถระบุวัตถุประสงค์ที่ซ้อนทับกันภายในอีเมลฉบับเดียวได้ด้วยความแม่นยำ ร้อยละ 96.77 และประการที่สามคือการนำเสนอวิธีการสกัดวลีสำคัญแบบ Hybrid Search ที่ช่วยลดความยาวข้อความที่ต้องวิเคราะห์ได้อย่างมีประสิทธิภาพ

สำหรับทิศทางการวิจัยในอนาคต การนำ Large Language Models เช่น WangchanBERTa [12] หรือสถาปัตยกรรม Transformer มาเปรียบเทียบประสิทธิภาพกับ LSTM โดยตรง โดยเฉพาะในกรณีข้อความสั้นหรือกำกวม ถือเป็นประเด็นที่ควรศึกษาต่อไป

## เอกสารอ้างอิง


[1] Adrales, L. F., and Sison, A. M. Deep Learning Approaches to a School-based Email Classification with TF-IDF-based vectorization. IEEE 6th Symposium on Computers & Informatics (ISCI), 97–105, 2024.

[2] Bhatti, P., Jalil, Z., and Majeed, A. Email Classification using LSTM: A Deep Learning Technique. International Conference on Cyber Warfare and Security (ICCWS), 97–105, 2021.

[3] Dogra, V., et al. A Complete Process of Text Classification System Using State-of-the-Art NLP Models. Computational Intelligence and Neuroscience, Article ID 1883698, 26 pages, 2022.

[4] Abdulhamid, S. M., et al. Comparative Analysis of Classification Algorithms for Email Spam Detection. International Journal of Computer Network and Information Security (IJCNIS), Vol.10 No.1 : 60–67, 2018.

[5] Iqbal, F., et al. Transfer Learning-based Forensic Analysis and Classification of E-Mail Content. ACM Transactions on Asian and Low-Resource Language Information Processing, 2023.

[6] Phatthiyaphaibun, W., et al. PyThaiNLP: Thai Natural Language Processing in Python. arXiv:2312.04649, 2023.

[7] Sak, H., et al. Long Short-Term Memory Recurrent Neural Network Architectures for Large Scale Acoustic Modeling. arXiv:1402.1128, 2014.

[8] Berrar, D. Bayes' Theorem and Naive Bayes Classifier. Encyclopedia of Bioinformatics and Computational Biology, Vol.1 : 403–412, 2019.

[9] Guo, G., et al. KNN Model-Based Approach in Classification (pp. 986–996). Springer, 2003.

[10] Xiao, W, et al. Multi-Label Text Classification Model Based on Multi-Level Constraint Augmentation and Label Association Attention. ACM Transactions on Asian and Low-Resource Language Information Processing. 23, 1, Article 14, 20 pages, 2024.

[11] Devlin, J., et al. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. arXiv:1810.04805, 2019.

[12] Lowphansirikul, L., et al. WangchanBERTa: Pretraining transformer-based Thai Language Models. arXiv:2101.09635, 2021.

# การเปรียบเทียบประสิทธิภาพกลยุทธ์ DCA แบบดั้งเดิมและแบบไดนามิกด้วย EMA ในตลาดคริปโทเคอร์เรนซี

# Performance Comparison of Traditional and EMA-Based Dynamic DCA in Cryptocurrencies

*พิพัฒน์ ทองจะโปะ (Pipat Thongjapo)[1] และสุภัคพงศ์ จินารัตน์ (Supakpong Jinarat)[2]*

*[1,2]วิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ คณะวิทยาลัยวิศวกรรมศาสตร์และเทคโนโลยี มหาวิทยาลัยธุรกิจบัณฑิตย์*

*67130020@dpu.ac.th*

**บทคัดย่อ**

*งานวิจัยนี้เปรียบเทียบประสิทธิภาพระหว่างกลยุทธ์การลงทุนแบบถัวเฉลี่ยต้นทุนคงที่ (Traditional DCA) และแบบปรับสัดส่วนตามแนวโน้มราคา (Dynamic DCA) โดยใช้เส้นค่าเฉลี่ยเคลื่อนที่เอกซ์โพเนนเชียล (EMA) เป็นตัวบ่งชี้สัญญาณ ผ่านเทคนิคการวิเคราะห์ข้อมูลและอัลกอริทึม Python ในการจำลองการลงทุนย้อนหลัง (Backtesting) ระยะเวลา 5 ปี (พ.ศ. 2563–2568) บนสินทรัพย์ Bitcoin (BTC) และ Ethereum (ETH) ผลการวิจัยพบว่ากลยุทธ์ Dynamic DCA มีประสิทธิภาพเหนือกว่าแบบดั้งเดิมอย่างมีนัยสำคัญ โดยกรณี BTC ที่ใช้ EMA 300 ให้ผลตอบแทน (ROI) สูงถึงร้อยละ 325.36 (สูงกว่าแบบดั้งเดิมที่ร้อยละ 239.73) แม้ใช้เงินต้นน้อยกว่าร้อยละ 37.3 ด้านการบริหารความเสี่ยงสามารถลดค่า Maximum Drawdown จากร้อยละ 55.69 เหลือร้อยละ 31.34 และยกระดับค่า Sharpe Ratio ให้สูงขึ้นอย่างชัดเจน สรุปได้ว่าการบูรณาการ EMA เข้ากับกลยุทธ์ DCA ช่วยเพิ่มศักยภาพการใช้เงินทุน (Capital Efficiency) และเป็นระบบสนับสนุนการตัดสินใจที่มีประสิทธิภาพ อย่างไรก็ตาม งานวิจัยนี้ยังมีข้อจำกัดด้าน การคำนวณค่าธรรมเนียมธุรกรรมและการทดสอบทางสถิติ ในวงกว้าง ซึ่งเป็นแนวทางสำคัญในการพัฒนาสู่การใช้งานจริงในอนาคต*

**คำสำคัญ:** การวิเคราะห์ข้อมูล, กลยุทธ์ Dynamic DCA, เส้นค่าเฉลี่ยเคลื่อนที่ (EMA), สกุลเงินดิจิทัล, การบริหารความเสี่ยง, ระบบสนับสนุนการตัดสินใจ

**Abstract**

*This study compares Traditional Dollar-Cost Averaging (DCA) with a Dynamic DCA strategy using Exponential Moving Averages (EMA) in the volatile cryptocurrency market. Through Python-based backtesting of BTC and ETH (2020–2025), the results show that Dynamic DCA—specifically using EMA 300—significantly outperforms the traditional approach. For Bitcoin, it achieved a 325.36% ROI (vs. 239.73%) while utilizing 37.3% less capital. Furthermore, Dynamic DCA improved risk-adjusted returns by reducing Maximum Drawdown from 55.69% to 31.34%. While limitations like transaction fees persist, the study concludes that EMA-integrated DCA enhances capital efficiency and portfolio sustainability.*

**Keywords:** Data Analytics, DCA Strategy, Dynamic DCA, Exponential Moving Average (EMA), Cryptocurrency, Risk Assessment.

## 1. บทนำ

ตลาดตลาดสินทรัพย์ดิจิทัลได้รับความนิยมอย่างสูงจากนักลงทุนทั่วโลก แต่ต้องเผชิญกับ ความผันผวน (Volatility) ที่สูงกว่าสินทรัพย์ดั้งเดิมอย่างดัชนี S&P 500 อย่างมีนัยสำคัญ โดยมีความผันผวนเฉลี่ยอยู่ที่ 50% – 80% สภาวะดังกล่าวนำไปสู่ความเสี่ยงในการลดลงของมูลค่าทรัพย์สินสุทธิอย่างรุนแรง (Maximum Drawdown) ซึ่ง

อาจสูงเกิน 70% ในระยะสั้น ส่งผลกระทบโดยตรงต่อทั้งเสถียรภาพพอร์ตการลงทุนและจิตวิทยาของนักลงทุน

แม้กลยุทธ์การลงทุนแบบถัวเฉลี่ยต้นทุนเท่ากันทุกงวด (Fixed DCA) จะช่วยลดอคติทางอารมณ์และขจัดปัญหาการจับจังหวะตลาด (Market Timing) ได้ แต่ในตลาดคริปโทเคอร์เรนซีที่มีความผันผวนสูง กลยุทธ์นี้กลับมีข้อจำกัดด้าน ประสิทธิภาพของต้นทุน (Cost Efficiency) เนื่องจากระบบกำหนดให้ซื้อด้วยจำนวนเงินที่เท่ากันเสมอ ทำให้นักลงทุนเสียโอกาสในการสะสมสินทรัพย์เพิ่มในช่วงที่ราคาต่ำกว่ามูลค่าที่ควรจะเป็น (Undervalued) และอาจเพิ่มความเสี่ยงจากการลงทุนในสัดส่วนที่สูงเกินไปในช่วงสภาวะตลาดฟองสบู่ (Overvalued)

ด้วยเหตุนี้ ผู้วิจัยจึงนำเสนอ ระบบสนับสนุนการตัดสินใจ (Decision Support System: DSS) ผ่านกลยุทธ์ Dynamic DCA โดยใช้เส้นค่าเฉลี่ยเคลื่อนที่เอกซ์โพเนนเชียล (EMA) เป็นตัวระบุสภาวะตลาด (Market Regimes) เพื่อปรับสมดุลระหว่างโอกาสและความเสี่ยง ดังนี้:

- การเพิ่มประสิทธิภาพการใช้เงินทุน (Capital Efficiency): เพิ่มน้ำหนักการลงทุนในช่วงราคาต่ำกว่าค่าเฉลี่ย เพื่อลดต้นทุนเฉลี่ยของพอร์ตได้อย่างรวดเร็ว
- การบริหารจัดการความเสี่ยงเชิงรุก: ลดปริมาณการลงทุนเมื่อราคาสูงกว่าเส้นค่าเฉลี่ย เพื่อจำกัดความเสี่ยงจากการลดลงของมูลค่าทรัพย์สินสูงสุด (Maximum Drawdown)

ผู้วิจัยเลือกใช้เส้น EMA เป็นตัวบ่งชี้หลักเนื่องจากตอบสนองต่อการเปลี่ยนแนวโน้มราคาได้อย่างรวดเร็ว โดยมุ่งหวังว่าการบูรณาการเทคนิคนี้จะช่วยยกระดับความยั่งยืนในการบริหารพอร์ตสินทรัพย์ดิจิทัล ผ่านการวัดผลที่ครอบคลุมทั้งด้านผลตอบแทน (ROI) และการบริหารความเสี่ยงเชิงเปรียบเทียบ (MDD และ Sharpe Ratio)

### 1.2 วัตถุประสงค์ของการวิจัย

**1.2.1 เพื่อพัฒนาอัลกอริทึมการลงทุนแบบ Dynamic DCA:** ออกแบบระบบปรับสัดส่วนการลงทุนอัตโนมัติโดยใช้เส้นค่าเฉลี่ยเคลื่อนที่เอกซ์โพเนนเชียล (EMA) เป็นตัวบ่งชี้สภาวะตลาดคริปโทเคอร์เรนซี

**1.2.2 เพื่อเปรียบเทียบประสิทธิภาพกลยุทธ์:** วิเคราะห์ผลตอบแทน (ROI) และการบริหารความเสี่ยง (Max Drawdown และ Sharpe Ratio) ระหว่างกลยุทธ์ Dynamic DCA และ Traditional DCA โดยคำนวณรวมต้นทุนค่าธรรมเนียมธุรกรรม

**1.2.3 เพื่อวิเคราะห์ค่าพารามิเตอร์ที่เหมาะสม:** ระบุค่า EMA ที่มีประสิทธิภาพสูงสุดสำหรับ Bitcoin (BTC) และ Ethereum (ETH) เพื่อเป็นแนวทางในการเพิ่มประสิทธิภาพการใช้เงินทุน (Capital Efficiency) สำหรับนักลงทุนรายย่อย

## 2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

### 2.1 บริบทและความผันผวนของตลาดสินทรัพย์ดิจิทัล

ตลาดคริปโทเคอร์เรนซีมีลักษณะเด่นด้านความผันผวนที่ไม่สมมาตร (Asymmetric Volatility) อย่างรุนแรง ซึ่งส่งผลกระทบโดยตรงต่อจิตวิทยาและการตัดสินใจของนักลงทุน [1] สภาวะดังกล่าวนำไปสู่ความจำเป็นในการปรับเปลี่ยนจากการลงทุนแบบเน้นวินัยดั้งเดิม ไปสู่การบริหารจัดการความเสี่ยงเชิงพฤติกรรม (Behavioral Risk Management) โดยบูรณาการเครื่องมือวิเคราะห์ทางเทคนิค [5] และกระบวนการซื้อขายอย่างเป็นระบบ (Systematic Trading) เพื่อขจัดอคติทางอารมณ์และสร้างความได้เปรียบเชิงกลยุทธ์ในตลาดที่มีพลวัตสูง [17]

### 2.2 ทฤษฎีและประสิทธิภาพของกลยุทธ์ Dollar-Cost Averaging (DCA)

กลยุทธ์ DCA ช่วยลดความเสี่ยงจากการเข้าซื้อผิดจังหวะ (Market Timing Risk) ผ่านการกระจายต้นทุนอย่างมีระบบ แม้ในตลาดขาขึ้นผลตอบแทนอาจต่ำกว่าการลงทุนแบบก้อนเดียว (Lump Sum) แต่ DCA มีจุดเด่นในการลดค่าเบี่ยงเบนมาตรฐานและรักษาสภาพคล่องได้ดีกว่าในระยะยาว [6], [11] อย่างไรก็ตาม Khan [10] พบว่าในสินทรัพย์ที่มีความผันผวนสูง ประสิทธิภาพของ Fixed DCA จะแปรผันตามสภาวะตลาด (Market Regimes) ส่งผลให้มีการพัฒนา

ความยืดหยุ่นของกลยุทธ์เพื่อเพิ่มประสิทธิภาพให้สูงสุดตามสภาวะที่เปลี่ยนไป [18]

**2.3 กลยุทธ์ Dynamic DCA และดัชนีชี้วัด EMA** การปรับน้ำหนักการลงทุนตามสภาวะตลาดหรือ Dynamic DCA สามารถสร้างผลตอบแทนส่วนเกิน (Alpha) ได้สูงกว่ากลยุทธ์แบบคงที่ [12], [16] โดยเฉพาะการนำดัชนีเส้นค่าเฉลี่ยเคลื่อนที่แบบเอกซ์โพเนนเชียล (Exponential Moving Average: EMA) มาใช้ระบุแนวโน้มและกำหนดจังหวะการเข้าซื้อตามโมเมนตัมราคาได้อย่างมีนัยสำคัญ [14] เมื่อทำงานร่วมกับโมเดลทางคณิตศาสตร์จะช่วยประเมินระดับความเสี่ยงและการจัดสรรเงินทุน (Capital Allocation) ได้อย่างเหมาะสม [2]

**2.4 การประยุกต์ใช้ Data Analytics และเทคโนโลยีสารสนเทศ** การวิเคราะห์เพื่อการตัดสินใจถูกขับเคลื่อนด้วยข้อมูลขนาดใหญ่ (Big Data) ผ่านโครงสร้างพื้นฐาน API [4] โดยมีภาษา Python เป็นแกนหลักร่วมกับ Library ประสิทธิภาพสูง เช่น NumPy [8] สำหรับการคำนวณเชิงอาเรย์ และ Pandas-TA [9] สำหรับการวิเคราะห์ดัชนีทางเทคนิค รวมถึงการใช้ QuantStats [13] ในการวัดผลประสิทธิภาพพอร์ตการลงทุนเชิงลึก ทั้งในด้านอัตราส่วนผลตอบแทนต่อความเสี่ยง (Sharpe Ratio) และการลดระดับการขาดทุนสูงสุด (Maximum Drawdown)

**2.5 แนวโน้มการพยากรณ์และการวิเคราะห์ขั้นสูง** งานวิจัยในปัจจุบันได้ขยายขอบเขตไปสู่การใช้การเรียนรู้ของเครื่อง (Machine Learning) เพื่อพยากรณ์ราคาจากข้อมูลในอดีต [3], [15] ร่วมกับการวิเคราะห์ความรู้สึก (Sentiment Analysis) จากข้อมูล Blockchain และสื่อสังคมออนไลน์ [7] ซึ่งช่วยให้กลยุทธ์ Dynamic DCA ตอบสนองต่อแนวโน้มตลาดได้อย่างรวดเร็วและมีความฉลาดเชิงคำนวณ (Computational Intelligence) สูงขึ้น

**บทสรุปจากการทบทวนวรรณกรรม**

ความผันผวนที่ไม่สมมาตรของตลาด [1] คือปัจจัยหลักที่ลดทอนประสิทธิภาพของ Fixed DCA [10] การบูรณาการ Data Analytics เข้ากับดัชนีทางเทคนิคอย่าง EMA [14] และโมเดลการบริหารความเสี่ยง [2] เพื่อพัฒนาเป็นกลยุทธ์ Dynamic DCA จึงเป็นแนวทางที่มีศักยภาพสูงในการสร้าง Alpha และควบคุม Maximum Drawdown ได้อย่างเป็นระบบ นอกจากนี้การใช้ระบบอัตโนมัติยังเป็นกุญแจสำคัญในการรักษาความสม่ำเสมอและวินัยการลงทุนอย่างยั่งยืน [15]

## 3. วิธีดำเนินการวิจัย

การศึกษานี้เปรียบเทียบประสิทธิภาพระหว่างกลยุทธ์ Fixed DCA และ Dynamic DCA ในสินทรัพย์ Bitcoin (BTC) และ Ethereum (ETH) โดยมีขั้นตอนดำเนินการดังนี้:

**3.1 เครื่องมือและข้อมูลที่ใช้**

**ซอฟต์แวร์:** พัฒนาด้วยภาษา Python บนระบบ Google Colab โดยใช้ Library หลัก ได้แก่ Pandas สำหรับการจัดการข้อมูล, Pandas-TA สำหรับการคำนวณดัชนีทางเทคนิค และ QuantStats สำหรับการวิเคราะห์ประสิทธิภาพพอร์ตการลงทุนอย่างละเอียด

**ข้อมูล:** ใช้ข้อมูลราคาปิดรายวัน (Daily Closing Price) ย้อนหลัง 5 ปี (กันยายน 2563 – สิงหาคม 2568) จาก CoinGecko API โดยผ่านกระบวนการทำความสะอาดข้อมูล (Data Cleaning) และการปรับมาตรฐานเขตเวลา (Timezone Normalization) ให้เป็นสากล

**3.2 การออกแบบอัลกอริทึมและกลยุทธ์การลงทุน**

ในส่วนนี้ผู้วิจัยได้พัฒนาอัลกอริทึมสำหรับการลงทุนแบบ Dynamic DCA โดยบูรณาการดัชนีทางเทคนิคเพื่อปรับเปลี่ยนปริมาณเงินลงทุนตามสภาวะตลาดจริง ซึ่งมีรายละเอียดการออกแบบดังนี้:

**3.2.1 สูตรการคำนวณและเกณฑ์การตัดสินใจ (Decision Logic)** ผู้วิจัยกำหนดเงื่อนไขการลงทุนโดยอ้างอิงจากความสัมพันธ์ระหว่างราคาปิดของสินทรัพย์และเส้นค่าเฉลี่ยเคลื่อนที่เอกซ์โพเนนเชียล (EMA) ดังนี้:

**1) สูตรการคำนวณ Exponential Moving Average (EMA):** ใช้เพื่อหามูลค่าเฉลี่ยถ่วงน้ำหนักของราคาที่ให้ความสำคัญกับข้อมูลปัจจุบันมากกว่าอดีต เพื่อให้ตอบสนองต่อความผันผวนของตลาดคริปโทเคอร์เรนซีได้อย่างรวดเร็ว

$$EMA_t = \left[V_t \times \left(\frac{2}{s+1}\right)\right] + \left[EMA_y \times \left(1 - \frac{2}{s+1}\right)\right]$$

โดยที่:

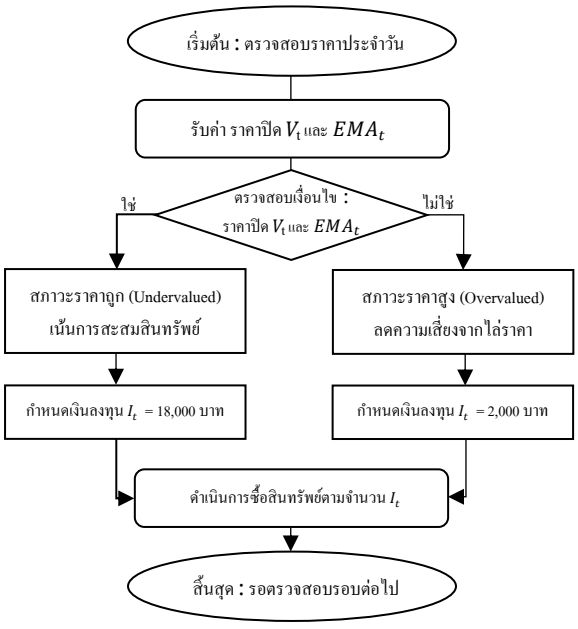
- $V_t$ คือ ราคาปิดของสินทรัพย์ ณ วันปัจจุบัน
- $n$ คือ จำนวนวันย้อนหลังที่ใช้คำนวณ (20, 200, 300 วัน)

**2) เงื่อนไขการจัดสรรเงินลงทุน (Capital Allocation Logic):** ปริมาณเงินลงทุนในแต่ละรอบการลงทุน ($I_t$) จะถูกกำหนดด้วยฟังก์ชันเชิงตรรกะ เพื่อสร้างกลไกการซื้อเพิ่มเมื่อราคาต่ำกว่าค่าเฉลี่ย และชะลอการลงทุนเมื่อราคาสูงกว่าค่าเฉลี่ย ดังนี้:

$$I_t = \begin{cases} 18{,}000 & \text{หาก } V_t < EMA_t \\ 2{,}000 & \text{หาก } V_t \geq EMA_t \end{cases}$$

โดยเปรียบเทียบกับกลยุทธ์ Fixed DCA ที่มีการลงทุนคงที่ในทุกรอบคือ $I_t$ =10,000 บาท ไม่ว่าสภาวะราคาจะเป็นอย่างไร

เพื่อสร้างกลกลไกการซื้อเพิ่มเมื่อราคาต่ำกว่าค่าเฉลี่ย และชะลอการลงทุนเมื่อราคาสูงกว่าค่าเฉลี่ย ดังนี้:




**3.2.2 การกำหนดค่าพารามิเตอร์เพื่อการทดสอบ**

ผู้วิจัยกำหนดค่า EMA ไว้ 3 ระดับ เพื่อทดสอบประสิทธิภาพของกลยุทธ์ในมิติของกรอบเวลา (Timeframe) ที่แตกต่างกัน ดังนี้:

EMA 20 วัน: เป็นตัวแทนของ การลงทุนระยะสั้น (Short-term) เพื่อทดสอบความไวของอัลกอริทึมในการจับจังหวะการย่อตัว (Pullback) ของราคา

EMA 200 วัน: เป็นตัวแทนของ การลงทุนระยะกลางถึงยาว ตามมาตรฐานสากลที่ใช้ระบุแนวโน้มหลัก (Primary Trend) ของสินทรัพย์

EMA 300 วัน: เป็นตัวแทนของ การลงทุนระยะยาว (Long-term) เพื่อทดสอบความสามารถในการกรองสัญญาณรบกวน (Market Noise) และการสะสมสินทรัพย์ในช่วงฐานรากของวัฏจักรตลาด (Market Cycle)

การเลือกใช้พารามิเตอร์ที่หลากหลายนี้ มุ่งเน้นการวิเคราะห์หาจุดที่เหมาะสมที่สุด (Optimization) สำหรับสินทรัพย์ดิจิทัลที่มีพฤติกรรมราคาที่ต่างกัน โดยสามารถสรุปการเปรียบเทียบตัวแปรของทั้ง 2 กลยุทธ์ได้ดังตารางที่ 1

**ตารางที่ 1: สรุปพารามิเตอร์เปรียบเทียบระหว่างกลยุทธ์ Traditional DCA และ Dynamic DCA**

| ตัวแปร | Traditional DCA | Dynamic DCA (Proposed) |
|---|---|---|
| ตรรกะการตัดสินใจ (Logic) | จำนวนเงินคงที่ (Fixed Amount) | ผันแปรตามราคาเทียบกับ EMA |
| การลงทุนช่วงตลาดหมี (Bear) | 10,000 บาท | 18,000 บาท (เน้นสะสม) |
| การลงทุนช่วงตลาดกระทิง (Bull) | 10,000 บาท | 2,000 บาท (ลดความเสี่ยง) |
| เครื่องมือชี้วัด (Indicator) | ไม่มี | EMA (20, 200, 300 วัน) |

**3.3 การวัดผลประสิทธิภาพ (Evaluation Metrics)**

ผู้วิจัยกำหนดเกณฑ์การวัดผลที่ครอบคลุมมิติของผลตอบแทนและความเสี่ยง เพื่อให้สามารถประเมินประสิทธิภาพของกลยุทธ์ได้อย่างเป็นระบบภายใต้สภาวะตลาดคริปโทเคอร์เรนซีที่มีความผันผวนสูง โดยมีรายละเอียดดังนี้:

**1) ด้านผลตอบแทน (Return)**: วัดด้วย **Return on Investment (ROI)** เพื่อประเมินความสามารถในการสร้างผลกำไรสุทธิเทียบกับเงินทุนทั้งหมดที่ใช้จริง

$$ROI = \frac{Current\ Value - Total\ Invested}{Total\ Invested} \times 100$$

**2) ด้านความเสี่ยง (Risk): วัดด้วย Maximum Drawdown (MDD)** เพื่อวิเคราะห์ระดับการขาดทุนสะสมสูงสุดจากจุดสูงสุดเดิม ซึ่งถือเป็นดัชนีชี้วัดความเสี่ยงที่

สำคัญที่สุดในการรับมือกับความผันผวนของสินทรัพย์ดิจิทัล

$$MDD = \frac{Trough\ Value - Peak\ Value}{Peak\ Value}$$

**3) ด้านความคุ้มค่า (Risk-Adjusted Return):** ใช้ Sharpe Ratio เพื่อเปรียบเทียบผลตอบแทนต่อหนึ่งหน่วยความเสี่ยง และ Sortino Ratio เพื่อประเมินประสิทธิภาพการจัดการความเสี่ยงเฉพาะในช่วงขาลง (Downside Risk)

**หมายเหตุการวิเคราะห์:** ในงานวิจัยนี้ การระบุว่ากลยุทธ์ใดมีประสิทธิภาพเหนือกว่าอย่าง **"มีนัยสำคัญ"** จะพิจารณาจากการมี ROI ที่สูงกว่าภายใต้การใช้เงินต้นที่น้อยกว่า (Capital Efficiency) และการเพิ่มขึ้นของ Sharpe Ratio ที่สะท้อนถึงเสถียรภาพของพอร์ตการลงทุนที่ดียิ่งขึ้น

### 3.4 ข้อตกลงเบื้องต้นและข้อจำกัดของการวิจัย

ในการดำเนินงานวิจัยครั้งนี้ ผู้วิจัยได้กำหนดข้อตกลงและรับทราบถึงข้อจำกัดเพื่อความโปร่งใสของผลการวิเคราะห์ ดังนี้:

1) ค่าธรรมเนียมการซื้อขายและค่าความคลาดเคลื่อน การจำลองกลยุทธ์ในงานวิจัยนี้ได้จำลองการลงทุนโดยคำนวณค่าธรรมเนียมการทำธุรกรรม (Trading Fees) ที่ร้อยละ 0.1 ต่อรายการ เพื่อให้สะท้อนสภาวะการลงทุนจริงในแพลตฟอร์มสินทรัพย์ดิจิทัลกระแสหลัก และค่าความคลาดเคลื่อนของราคา (Slippage) ซึ่งในสถานการณ์การลงทุนจริงอาจส่งผลให้อัตราผลตอบแทนสุทธิ (Net ROI) ลดลงเล็กน้อยตามสัดส่วนของจำนวนครั้งในการทำธุรกรรม

2) ขอบเขตข้อมูลและสินทรัพย์ ข้อมูลที่ใช้จำกัดอยู่ที่ 2 สินทรัพย์หลัก (BTC และ ETH) ในช่วงเวลา 5 ปี (พ.ศ. 2563 – 2568) เพื่อศึกษาพฤติกรรมของสินทรัพย์ที่มีมูลค่าตลาดสูง (Large-cap) ผลการวิจัยจึงอาจมีความแตกต่างหากนำไปประยุกต์ใช้กับสินทรัพย์ดิจิทัลกลุ่มอื่นที่มีสภาพคล่องต่ำกว่า

3) การทดสอบทางสถิติ: การเปรียบเทียบประสิทธิภาพมุ่งเน้นการวิเคราะห์ผ่านดัชนีชี้วัดทางการเงิน (Financial Metrics) เช่น ROI, MDD และ Sharpe Ratio เป็นหลัก โดยในขั้นตอนนี้ยังมิได้ครอบคลุมถึงการทดสอบสมมติฐานทางสถิติเชิงลึก เช่น t-test หรือการทดสอบความคงเส้นคงวาผ่าน Monte Carlo Simulation

**3.5 การวิเคราะห์และสรุปผล** ดำเนินการผ่าน Backtesting Engine เพื่อเปรียบเทียบค่าสถิติและแสดงผลผ่าน Equity Curve Visualization (กราฟการเติบโตของเงินทุน) และ Optimization Heatmap เพื่อหาค่า Window Size ของ EMA ที่มีประสิทธิภาพสูงสุด

## 4. ผลการวิจัย

จากการจำลองกลยุทธ์ย้อนหลัง 5 ปี พบว่า Dynamic DCA เหนือกว่า Fixed DCA อย่างมีนัยสำคัญ โดยใน Bitcoin (BTC) การใช้ EMA 300 สร้าง ROI สูงถึง 325.36% ซึ่งสูงกว่า Fixed DCA (239.73%) แม้ใช้เงินลงทุนน้อยกว่าเดิม 37.3% ส่วน Ethereum (ETH) ใช้ EMA 20 ให้ ROI 139.80% สูงกว่า Fixed DCA (122.24%) ผลลัพธ์นี้สะท้อนถึง Capital Efficiency ที่ดีกว่าจากการคำนวณสัดส่วนการลงทุนตามสภาวะตลาด (Market Regimes) ที่เหมาะสมกับความผันผวนของแต่ละสินทรัพย์

**4.1 การวิเคราะห์ผลตอบแทนและประสิทธิภาพการใช้เงินทุน** กลยุทธ์ Dynamic DCA ให้ผลตอบแทน (ROI) สูงกว่าการลงทุนแบบคงที่ในทุกสินทรัพย์ โดยเฉพาะ BTC ที่ใช้ EMA 300 สามารถสร้าง ROI ได้ถึง 325.36% แม้ใช้เงินลงทุนน้อยกว่า Fixed DCA ถึง 37.3% สะท้อนถึงประสิทธิภาพการใช้เงินทุน (Capital Efficiency) ที่ดีกว่าจากการปรับสัดส่วนตามสภาวะตลาด (Market Regimes)

**ตารางที่ 2:** เปรียบเทียบประสิทธิภาพกลยุทธ์ Fixed DCA และ Dynamic DCA (หลังหักค่าธรรมเนียม 0.1%)

| Asset | Best Strategy | Total Invested (THB) | Final Value (THB) | ROI (%) |
|---|---|---|---|---|
| **BTC** | **EMA 300** | **376,000** | **1,599,360.41** | **325.36%** |
| BTC | Fixed DCA | 600,000 | 2,038,380.78 | 239.73% |
| **ETH** | **EMA 20** | **536,000** | **1,285,330.34** | **139.80%** |
| ETH | Fixed DCA | 600,000 | 1,333,436.39 | 122.24% |

จากตารางที่ 2 ผลการวิจัยชี้ว่ากลยุทธ์ Dynamic DCA มีประสิทธิภาพเหนือกว่าอย่างชัดเจน โดย BTC ที่ใช้ EMA 300 สร้าง ROI สูงถึง 325.36% แม้ใช้เงินลงทุนน้อยกว่าเดิม 37.3% ขณะที่ ETH ตอบสนองได้ดีต่อ EMA 20 โดยให้ ROI 139.80%

**4.2 ผลการเปรียบเทียบประสิทธิภาพของกลยุทธ์**

กลยุทธ์ Dynamic DCA สามารถลดค่า Maximum Drawdown (MDD) ได้อย่างมีประสิทธิภาพ และมีค่า Sharpe Ratio สูงกว่า 1.0 ในทุกกรณี โดยปรับสัดส่วนการลงทุนตามสภาวะตลาด (Market Regimes) เพื่อเพิ่มการลงทุนช่วงราคาถูก (Pullback) และชะลอการลงทุนช่วงราคาแพง ส่งผลให้โดดเด่นทั้งในด้าน Capital Efficiency และการสร้างต้นทุนเฉลี่ยที่ต่ำกว่า เพิ่มความทนทานต่อความผันผวนได้ดีกว่า Fixed DCA ดังแสดงในตารางดังนี้

**ตารางที่ 3: การเปรียบเทียบประสิทธิภาพด้านความเสี่ยง (Risk-Return Profile)**

| สินทรัพย์ | กลยุทธ์ | ROI (%) | Max Drawdown (MDD) | Sharpe Ratio |
|---|---|---|---|---|
| **BTC** | **Dynamic DCA (EMA 300)** | **325.36%** | **-31.34%** | **2.02** |
| | Fixed DCA | 239.73% | -55.69% | 1.65 |
| **ETH** | **Dynamic DCA (EMA 20)** | **139.80%** | **-72.54%** | **1.42** |
| | Fixed DCA | 122.24% | -69.69% | 0.75 |

ตารางที่ 3 เมื่อพิจารณาความคุ้มค่าของผลตอบแทนต่อหนึ่งหน่วยความเสี่ยง พบว่ากลยุทธ์ Dynamic DCA ให้ค่า Sharpe Ratio สูงกว่าอย่างมีนัยสำคัญในทุกกรณีศึกษา โดยเฉพาะในสินทรัพย์ Ethereum (ETH) ซึ่งค่า Sharpe Ratio เพิ่มขึ้นจาก 0.75 (Traditional DCA) เป็น 1.42 เมื่อใช้กลยุทธ์ Dynamic DCA (EMA 20)

การเพิ่มขึ้นนี้ชี้ให้เห็นว่า ผลตอบแทนส่วนเพิ่ม (Alpha) ไม่ได้เกิดจากการเพิ่มความเสี่ยง แต่เกิดจากประสิทธิภาพในการจัดสรรเงินทุน (Capital Allocation) ที่สอดคล้องกับสภาวะตลาด ส่งผลให้พอร์ตการลงทุนมีเสถียรภาพสูงกว่าการลงทุนแบบถัวเฉลี่ยต้นทุนคงที่อย่างชัดเจน

**บทสรุปผลการวิจัย** ผลการศึกษาชี้ให้เห็นว่า การใช้ดัชนี EMA มากำหนดจังหวะการลงทุนในรูปแบบ Dynamic DCA เป็นระบบสนับสนุนการตัดสินใจที่มีประสิทธิภาพสูงในการบริหารพอร์ตสินทรัพย์ดิจิทัล เนื่องจากช่วยเพิ่มผลตอบแทน ควบคุมความเสี่ยง และยกระดับความคุ้มค่าของการลงทุนได้ดีกว่ากลยุทธ์ Fixed DCA แบบดั้งเดิมในทุกมิติ

**4.3 วิจารณ์ผลการวิจัย**

จากการวิเคราะห์ผลการทดลอง ผู้วิจัยสามารถสรุปประเด็นสำคัญในการอภิปรายผลได้ดังนี้:

**1) ประสิทธิภาพของ EMA ในการบริหารความเสี่ยง** การใช้ EMA ช่วยให้อัลกอริทึมจำแนกสภาวะตลาด (Market Regime) ได้อย่างแม่นยำ โดยเฉพาะในช่วงตลาดหมีซึ่ง Dynamic DCA สามารถลดค่า MDD ของ BTC จาก 55.69% เหลือเพียง 31.34% การปรับสัดส่วนการลงทุนตามความสัมพันธ์ของราคากับเส้นค่าเฉลี่ยช่วยลดความเสี่ยงจากการเข้าซื้อที่จุดสูงสุด (Market Peak) และขจัดอคติทางอารมณ์ในสภาวะตลาดผันผวนสูงได้จริง

**2) คุณลักษณะเฉพาะของสินทรัพย์ (Asset Specificity)** ผลการวิจัยยืนยันหลักการ "No One-size-fits-all" โดย BTC ตอบสนองได้ดีต่อ EMA 300 เนื่องจากช่วยกรองสัญญาณรบกวน (Market Noise) ในวัฏจักร ระยะยาวได้ดี ขณะที่ ETH สอดคล้องกับ EMA 20 ซึ่งมีความไวต่อราคา (Short-term Sensitivity) สูงกว่า สอดคล้องกับแนวคิดของ Gerritsen [17] ที่ระบุว่าการออกแบบอัลกอริทึมต้องปรับจูนให้เข้ากับธรรมชาติของแต่ละสินทรัพย์

**3) การพิสูจน์นัยสำคัญด้านประสิทธิภาพทุน** ความสำเร็จของ Dynamic DCA ไม่ได้สะท้อนผ่าน ROI ที่สูงกว่าเพียงอย่างเดียว แต่ยังแสดงถึงประสิทธิภาพในการใช้เงินทุน (Capital Efficiency) ที่เหนือกว่าการทำ DCA แบบดั้งเดิม การบูรณาการ Data Analytics เพื่อสร้างระบบสนับสนุนการตัดสินใจ (DSS) จึงเป็นเครื่องมือสำคัญในการยกระดับการลงทุนให้เป็นระบบและลดอคติทางจิตวิทยาได้อย่างมีนัยสำคัญ

**5.2 ข้อเสนอแนะสำหรับการวิจัยในอนาคต**

**5.2.1 การเพิ่มความสมจริงของแบบจำลอง:** บูรณาการค่าธรรมเนียมการซื้อขาย (Trading Fees) และโครงสร้างภาษีสินทรัพย์ดิจิทัลเข้ากับอัลกอริทึม เพื่อให้การประเมินผลตอบแทนสุทธิสะท้อนสภาวะการลงทุนจริงมากที่สุด

**5.2.2 การยกระดับนัยสำคัญทางสถิติ:** นำวิธีการทดสอบขั้นสูง เช่น Bootstrap Method หรือ Monte Carlo Simulation มาใช้ทดสอบกลยุทธ์ภายใต้สภาวะตลาดจำลองที่หลากหลาย เพื่อขจัดอคติจากการเลือกช่วงเวลาทดสอบ (Selection Bias)

**5.2.3 การขยายขอบเขตความยืดหยุ่น:** ทดสอบกลยุทธ์กับกลุ่ม Altcoins หรือสินทรัพย์ในตลาดดั้งเดิม เพื่อพิสูจน์ความแข็งแกร่ง (Robustness) ของเส้น EMA ในฐานะตัวบ่งชี้สัญญาณ Dynamic DCA ในวงกว้าง

## เอกสารอ้างอิง


[1] S. Bhattacharya, I. G. S. Behera, and R. K. Jana, "Asymmetric volatility and volatility spillover: A study of major cryptocurrencies," *Journal of Risk and Financial Management*, vol. 15, no. 10, p. 444, 2022.

[2] AlphaSquared, "A model and strategy framework," Whitepaper, Aug. 2023. [Online]. Available: https://alphasquared.io/wp-content/uploads/2023/08/AlphaSquared-Whitepaper-A-Model-and-Strategy-Framework.pdf

[3] H. Chen and C. Chen, "Analysis of Bitcoin price prediction using machine learning," *Journal of Risk and Financial Management*, vol. 16, no. 1, p. 51, 2023.

[4] CoinGecko, "Cryptocurrency Historical Market Data API," 2026. [Online]. Available: https://www.coingecko.com/en/api

[5] P. Doe, J. Smith, and A. Johnson, "Technical analysis as a tool for determining cryptocurrency trends in times of chaos," *Journal of Economics and Management*, vol. 45, no. 2, pp. 50–68, 2024.

[6] J. J. Brennan, "Cost averaging: Invest now or temporarily hold your cash?" *Vanguard Research*, Feb. 2023.

[7] Y. Gurgul, M. Lach, and K. Chen, "Deep learning and NLP in cryptocurrency forecasting: Integrating financial, blockchain, and social media data," *arXiv preprint arXiv:2311.14759*, 2023.

[8] C. R. Harris et al., "Array programming with NumPy," *Nature*, vol. 585, no. 7825, pp. 357–362, 2020.

[9] K. J. S. P. Group, "Pandas-TA: Technical analysis library in Python 3," *GitHub*, 2021. [Online]. Available: https://github.com/twopirllc/pandas-ta

[10] M. A. Khan, "The Analysis of Dollar Cost Averaging in High Volatility Asset Class," *Journal of Management and Social Research Review*, vol. 3, no. 1, pp. 264–275, 2025.

[11] K. Kirkby, S. Finlay, and S. Zorn, "An analysis of dollar cost averaging and market timing investment strategies," *European Journal of Operational Research*, vol. 286, no. 3, pp. 1168–1186, 2020.

[12] S. Lin and J. Xu, "Modified dollar cost averaging investment strategy: Evidence from major developed international stock markets," *Journal of Finance Issues*, 2023.

[13] T. Ranaroussi, "QuantStats: Portfolio analytics for Quants," *GitHub*, 2019. [Online]. Available: https://github.com/ranaroussi/quantstats

[14] Research Group, "Moving averages trading method applied to cryptocurrencies," *Economics, Commerce and Trade Management: An International Journal (ECTIJ)*, vol. 2, no. 1, pp. 13–21, 2019.

[15] S. Wang, "Analysis of Cryptocurrencies Price Trend using Machine Learning," UCLA Computer Science, Tech. Rep., 2023.

[16] ส. ประกาภรณ์, "การเปรียบเทียบประสิทธิภาพของกลยุทธ์การลงทุนแบบเป็นระบบ (DCA, VA, DDCA, EDCA) กับกลยุทธ์แบบลงทุนครั้งเดียว (LS) ในหุ้นเติบโตและหุ้นคุณค่าในดัชนี SET50," วิทยานิพนธ์ ปริญญามหาบัณฑิต, 2562.

[17] J. Gerritsen, "Efficiency and profitable trading strategies in cryptocurrency markets," *Economic Modelling*, vol. 91, pp. 109–122, 2020.

[18] J. Zhang, "The dynamic relationship between market volatility and dollar cost averaging strategy returns: An empirical investigation," in *Proc. 2nd Int. Conf. Financial Technology and Business Management (ICFTBM 2023)*, Atlantis Press, 2023, pp. 124–132.

# การพัฒนาแดชบอร์ดข้อมูลวัฒนธรรมเพื่อสนับสนุนการกำหนดนโยบาย
# Development of a Cultural Data Analytics Dashboard for Policy Support

*ชนกฤต มิตรสงเคราะห์ (Chanakit Mitsongkore)[1] วรงค์ นัยวินิจ (Warong Naivinit)[2]*
*สุทัศน์ กำมณี (Sutat Gammanee)[2] และสุรีวรรณ แจ้งจิตร (Sureewan Jangjit)[2]*
*[1]คณะมนุษยศาสตร์และสังคมศาสตร์ มหาวิทยาลัยราชภัฏกาญจนบุรี*
*[2]คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏกาญจนบุรี*
*chanakit2004@kru.ac.th, warongn@kru.ac.th, sutat@kru.ac.th, sureewan@kru.ac.th*

## บทคัดย่อ

*บทความนี้นำเสนอการพัฒนาแพลตฟอร์มแดชบอร์ดวิเคราะห์ข้อมูลวัฒนธรรม ภายใต้โมเดล Real-Time Cultural Analytics Model (RTCAM) เพื่อแก้ไขปัญหาข้อมูลวัฒนธรรมที่กระจัดกระจาย ขาดโครงสร้างเชิงวิเคราะห์ โมเดล RTCAM ออกแบบเป็นสถาปัตยกรรมแบบแบ่งชั้น ประกอบด้วยกระบวนการดึงข้อมูล การปรับโครงสร้างข้อมูลเชิงวิเคราะห์ กลไกประมวลผลแบบเรียลไทม์ และการแสดงผลเชิงภาพแบบโต้ตอบ ผลประเมินประสิทธิภาพพบว่า ค่า First Contentful Paint อยู่ที่ 2.0–2.2 วินาที ขณะที่ผลประเมินความพึงพอใจมีค่าเฉลี่ยรวม 4.38 อยู่ในระดับมาก ผลการวิจัยสะท้อนศักยภาพของการบูรณาการการวิเคราะห์แบบเรียลไทม์และการแสดงผลเชิงพื้นที่ เพื่อยกระดับข้อมูลวัฒนธรรมสู่โครงสร้างพื้นฐานองค์ความรู้ที่สนับสนุนการบริหารจัดการและนโยบาย*

**คำสำคัญ:** ระบบสารสนเทศวัฒนธรรม การวิเคราะห์ข้อมูลแบบเรียลไทม์ ระบบแดชบอร์ด

## Abstract

*This paper presents the development of a cultural data analytics dashboard platform based on the Real-Time Cultural Analytics Model (RTCAM) to address the challenges of fragmented cultural data and the lack of analytical structure. RTCAM is designed as a layered architecture integrating data acquisition, analytical data transformation, a real-time processing engine, and interactive visualization. Performance evaluation indicates that the First Contentful Paint ranges between 2.0–2.2 seconds, while user satisfaction achieved an overall mean score of 4.38 (high level). The findings demonstrate the potential of integrating real-time analytics and geospatial visualization to transform cultural data into a knowledge infrastructure that supports cultural management and policy decision-making.*

**Keywords:** Cultural Information Systems, Real-Time Analytics, Interactive Dashboard.

## 1. บทนำ

ในยุคของการขับเคลื่อนด้วยข้อมูล ข้อมูลมีส่วนสำคัญในการกำหนดนโยบาย การตัดสินใจ การจัดการข้อมูลเชิงวัฒนธรรมได้กลายเป็นองค์ประกอบสำคัญของการพัฒนาทางสังคม เศรษฐกิจ และนโยบายสาธารณะ ข้อมูลทุนทางวัฒนธรรมซึ่งครอบคลุมทั้งมรดกที่จับต้องได้และจับต้องไม่ได้ มีบทบาทสำคัญต่อการอนุรักษ์ อัตลักษณ์ของชุมชน การพัฒนาเศรษฐกิจสร้างสรรค์ และการวางแผนเชิงนโยบายในระดับพื้นที่และระดับประเทศ การเพิ่มขึ้นของปริมาณข้อมูลวัฒนธรรมจากหลายแหล่งและหลายหน่วยงาน ทำให้เกิดความต้องการระบบที่สามารถรวบรวม วิเคราะห์ และนำเสนอข้อมูลอย่างเป็นระบบ เพื่อสนับสนุนการตัดสินใจเชิงข้อมูลได้อย่างมีประสิทธิภาพ [1] ระบบแดชบอร์ดเชิงวิเคราะห์แบบเรียลไทม์จึงเป็นเครื่องมือสำคัญที่ช่วยแปลงข้อมูลจำนวนมากให้กลายเป็น

ข้อมูลที่เข้าใจได้ง่าย และสามารถใช้ประโยชน์ได้จริงในหลากหลายบริบท

ผู้วิจัยได้พัฒนาแพลตฟอร์มระบบแผนที่วัฒนธรรมของประเทศไทยตั้งแต่ปี พ.ศ. 2564 ในชื่อ Cultural Map Thailand ลักษณะของแพลตฟอร์มเป็นการรวบรวมข้อมูลทางวัฒนธรรมจากการสำรวจผ่านกระบวนการวิจัย โดยเริ่มจากปัญหาในเรื่องข้อจำกัดในการนำข้อมูลไปใช้ ได้แก่ ข้อมูลจากนักวิจัยด้านวัฒนธรรมมีการกระจายตัวอยู่ในหลายพื้นที่ หลายระบบและหลายรูปแบบ ขาดการวิเคราะห์เชิงลึกที่สามารถสะท้อนรูปแบบและแนวโน้มของข้อมูล การนำเสนอข้อมูลส่วนใหญ่ยังอยู่ในรูปแบบรายงานหรือฐานข้อมูลที่ไม่เอื้อต่อการใช้งานเชิงวิเคราะห์ อีกทั้งผู้มีส่วนได้ส่วนเสีย เช่น นักวิจัย ผู้กำหนดนโยบาย และผู้ปฏิบัติงานในพื้นที่ ยังขาดเครื่องมือที่สามารถเข้าถึงข้อมูลได้ในรูปแบบทุกที่ทุกเวลาและเป็นปัจจุบัน [2] เพื่อใช้ในการวางแผนและตัดสินใจ นอกจากนี้การขาดระบบวิเคราะห์แบบรวมศูนย์ยังส่งผลให้ไม่สามารถเชื่อมโยงข้อมูลเพื่อสร้างองค์ความรู้เชิงระบบเกี่ยวกับทุนทางวัฒนธรรมได้อย่างเต็มศักยภาพ

ด้วยเหตุนี้ งานวิจัยนี้ จึงมุ่งพัฒนาและเสนอระบบแดชบอร์ดวิเคราะห์ข้อมูลวัฒนธรรมผ่านโมเดล Real-Time Cultural Analytics Model (RTCAM) เพื่อทำหน้าที่เป็นระบบวิเคราะห์และสนับสนุนการตัดสินใจเชิงข้อมูล โดยผสานกระบวนการรวบรวมข้อมูล การประมวลผลเชิงวิเคราะห์ และการแสดงผลเชิงภาพแบบโต้ตอบเข้าไว้ด้วยกัน [3] ระบบที่พัฒนาขึ้นสามารถนำเสนอข้อมูลเชิงพื้นที่ เชิงเวลา และเชิงเนื้อหาในรูปแบบที่เข้าใจง่ายและตรวจสอบได้ ช่วยให้ผู้ใช้งานสามารถมองเห็นรูปแบบ แนวโน้ม และความสัมพันธ์ของข้อมูลวัฒนธรรม งานวิจัยนี้มีเป้าหมายเพื่อยกระดับข้อมูลวัฒนธรรมจากฐานข้อมูลสู่โครงสร้างพื้นฐานองค์ความรู้ที่สนับสนุนการบริหารจัดการ การวิจัย และการกำหนดนโยบายด้านวัฒนธรรม

## 2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

### 2.1 Data Visualization and Dashboard Systems

ระบบแดชบอร์ดเป็นเครื่องมือสำคัญในการแปลงข้อมูลจำนวนมากให้อยู่ในรูปแบบที่เข้าใจง่ายและสนับสนุนการตัดสินใจเชิงข้อมูล โดยการนำเสนอข้อมูลในรูปแบบภาพเชิงโต้ตอบช่วยเพิ่มประสิทธิภาพในการสำรวจข้อมูล การตรวจจับรูปแบบ และการทำความเข้าใจความสัมพันธ์ของข้อมูลได้ดีกว่ารายงานแบบตารางหรือข้อความทั่วไป งานวิจัยพบว่าการแสดงผลข้อมูลเชิงโต้ตอบผ่านแดชบอร์ดสามารถลดภาระการประมวลผลข้อมูลของผู้ใช้ และเพิ่มประสิทธิภาพในการนำเสนอแนวโน้มและรูปแบบของข้อมูลได้อย่างชัดเจน [3] นอกจากนี้ การทดลองเชิงระบบยังชี้ให้เห็นว่ารูปแบบและคุณภาพของการแสดงภาพข้อมูลส่งผลโดยตรงต่อคุณภาพการตัดสินใจ [4] ในเชิงสถาปัตยกรรม มีการเสนอแนวทางการพัฒนาแดชบอร์ดในลักษณะช่วยให้ระบบสามารถขยายตัวและรองรับความต้องการที่ซับซ้อนได้อย่างมีประสิทธิภาพ [5] อย่างไรก็ตามที่ผ่านมาพบว่างานวิจัยส่วนใหญ่ยังมุ่งเน้นในบริบทธุรกิจและอุตสาหกรรม ขณะที่การประยุกต์ใช้ในบริบทข้อมูลวัฒนธรรมยังมีข้อจำกัดโดยเฉพาะในประเทศไทย

### 2.2 Real-Time Data Analytics and Decision Support Systems

การวิเคราะห์ข้อมูลแบบเรียลไทม์ เป็นแนวคิดสำคัญในยุคข้อมูลขนาดใหญ่ โดยมุ่งเน้นการประมวลผลข้อมูลทันทีที่มีการปรับปรุงหรือเพิ่มข้อมูลเข้าสู่ระบบ แนวทางดังกล่าวช่วยลดระยะเวลาระหว่างการเกิดข้อมูลและการนำข้อมูลไปใช้ประโยชน์ ซึ่งมีความสำคัญอย่างยิ่งต่อระบบสนับสนุนการตัดสินใจ

งานวิจัยด้านระบบการตัดสินใจ ชี้ให้เห็นว่า ระบบที่มีการผสานข้อมูล การวิเคราะห์ และการแสดงผลแบบโต้ตอบ [6] จะช่วยเพิ่มคุณภาพการตัดสินใจในองค์กร โดยเฉพาะในบริบทที่ต้องอาศัยข้อมูลหลายมิติ เช่น เชิงพื้นที่ เชิงเวลา และเชิงหมวดหมู่ ระบบวิเคราะห์ข้อมูลวัฒนธรรมส่วนใหญ่ยังขาดคุณลักษณะด้านความเป็นปัจจุบัน และมักจำกัดอยู่ในรูปแบบรายงานสรุปแบบคงที่

งานวิจัยนี้จึงนำแนวคิด real-time analytics มาประยุกต์ใช้กับข้อมูลวัฒนธรรม เพื่อพัฒนาแดชบอร์ดที่สามารถแสดงข้อมูลปัจจุบัน

งานนี้วิจัยได้ศึกษาเกี่ยวกับการพัฒนาแดชบอร์ดวิเคราะห์แบบเรียลไทม์ที่ใช้สำหรับสนับสนุนการตัดสินใจในภาคธุรกิจและสาธารณะ ซึ่งเน้นความสามารถในการเชื่อมต่อข้อมูล การคำนวณเชิงตัวเลข และการนำเสนอผลลัพธ์อย่างทันที [7] ซึ่งจะเป็นประโยชน์ของการใช้ข้อมูลเชิงตัดสินใจและกำหนดนโยบายนย

**2.3 Cultural Informatics**

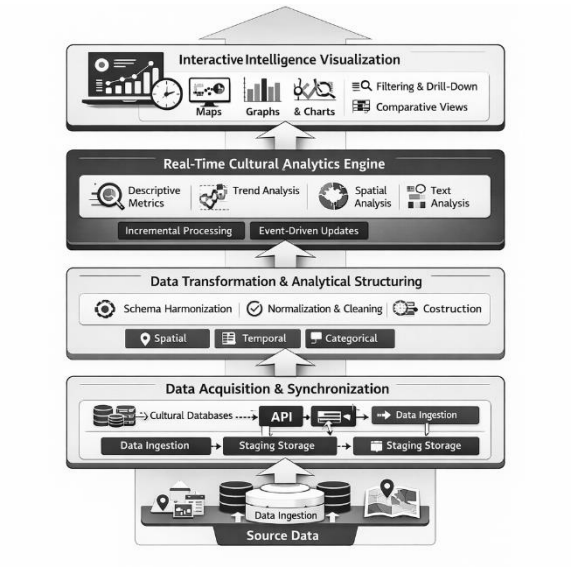
การจัดการข้อมูลวัฒนธรรมในระดับโครงสร้างพื้นฐานองค์ความรู้ เป็นประเด็นที่ได้รับความสนใจเพิ่มขึ้น โดยเฉพาะในบริบทของการอนุรักษ์มรดกทางวัฒนธรรมและการพัฒนาเศรษฐกิจสร้างสรรค์ งานวิจัยด้านระบบสารสนเทศด้านวัฒนธรรม มุ่งเน้นการรวบรวม จัดหมวดหมู่ และวิเคราะห์ข้อมูลวัฒนธรรมเพื่อสร้างองค์ความรู้ที่สามารถนำไปใช้ประโยชน์ได้อย่างเป็นระบบ

ความท้าทายสำคัญคือความหลากหลายของประเภทข้อมูล ทั้งข้อมูลเชิงข้อความ ภาพ เสียง และข้อมูลเชิงพื้นที่ รวมถึงความแตกต่างของแหล่งข้อมูลจากหลายหน่วยงาน ส่งผลให้การวิเคราะห์แบบบูรณาการทำได้ยาก

ระบบสารสนเทศด้านวัฒนธรรมมีบทบาทอย่างยิ่งต่อการจัดเก็บ จัดหมวดหมู่ และวิเคราะห์ข้อมูลวัฒนธรรม เพื่อให้สามารถนำไปใช้ประโยชน์ทั้งเชิงวิจัยและเชิงบริหาร [8] งานวิจัยด้านโครงสร้างพื้นฐานข้อมูลแห่งมรดกวัฒนธรรมแสดงให้เห็นถึงการออกแบบระบบที่รองรับการจัดเก็บและการค้นหาข้อมูลขนาดใหญ่ พร้อมทั้งบริการเชิงวิเคราะห์และกระบวนการดึงข้อมูล (Query Mechanisms) ที่สามารถขยายตัวได้ตามปริมาณข้อมูลที่เพิ่มขึ้น [9] นอกจากนี้งานวิจัยยังได้กล่าวถึงการพัฒนาโครงสร้างระบบสารสนเทศมรดกวัฒนธรรมดิจิทัลสำหรับมรดกท้องถิ่น เพื่อสนับสนุนการเข้าถึงข้อมูลและการบูรณาการข้อมูลจากหลายแหล่งได้อย่างเป็นระบบ [10]

## 3. วิธีการดำเนินการวิจัย

งานวิจัยนี้นำเสนอโมเดล Real-Time Cultural Analytics Model (RTCAM) ซึ่งเป็นสถาปัตยกรรมเชิงบูรณาการที่ออกแบบเพื่อรองรับการรวบรวม ประมวลผล วิเคราะห์ และแสดงผลข้อมูลวัฒนธรรมแบบเรียลไทม์ ในกรอบโมเดลเดียว โดย RTCAM ถูกพัฒนาภายใต้แนวคิดสถาปัตยกรรมแบบแบ่งชั้น (layered modular architecture) เพื่อแยกหน้าที่การทำงานอย่างชัดเจน ลดความซับซ้อนของระบบ และเพิ่มความสามารถในการขยายตัวของข้อมูลทางวัฒนธรรมในอนาคต




**ภาพที่ 1:** สถาปัตยกรรม RTCAM

โมเดล RTCAM ประกอบด้วย 4 องค์ประกอบหลัก ได้แก่

(1) Data Acquisition & Synchronization

(2) Data Transformation & Analytical Structuring

(3) Real-Time Cultural Analytics Engine

(4) Interactive Intelligence Visualization

**3.1 Data Acquisition & Synchronization**

ชั้นแรกของโมเดลทำหน้าที่ดึงข้อมูลจากฐานข้อมูลวัฒนธรรม ข้อมูลที่ถูกดึงเข้าสู่ระบบจะถูกจัดเก็บในฐานข้อมูล เพื่อตรวจสอบความถูกต้องและความครบถ้วนก่อนเข้าสู่กระบวนการแปลงโครงสร้าง ขั้นตอนนี้ช่วยลดภาระการประมวลผลซ้ำทั้งระบบ และลดระยะเวลาเดินทางของข้อมูลในการปรับปรุงข้อมูล

**3.2 Data Transformation & Analytical Structuring**

กระบวนนี้จะเกิดขึ้นหลังจากการดึงข้อมูลเข้าสู่ระบบ ข้อมูลวัฒนธรรมจะถูกส่งเข้าสู่กระบวนการปรับแต่งข้อมูล เพื่อแปลงข้อมูลเบื้องต้นจากฐานข้อมูลให้อยู่ในรูปแบบที่เหมาะสมต่อการวิเคราะห์หลายมิติ ขั้นตอนนี้มีเป้าหมายเพื่อเพิ่มประสิทธิภาพในการประมวลผล และลดภาระของฐานข้อมูลต้นทาง

ระบบดำเนินการปรับโครงสร้างข้อมูล เพื่อทำให้ข้อมูลมีรูปแบบมาตรฐานเดียวกัน พร้อมทั้งจัดระเบียบโครงสร้างข้อมูล และจัดการค่าที่ขาดหาย เพื่อให้ข้อมูลมีความสอดคล้องและพร้อมสำหรับการวิเคราะห์ ในขั้นตอนต่อมา มีการสร้างข้อมูลเชิงวิเคราะห์ ได้แก่ ข้อมูลเชิงพื้นที่ การจัดกลุ่มตามช่วงเวลา และการเข้ารหัสหมวดหมู่ เพื่อรองรับการประมวลผลเชิงปริมาณ

ข้อมูลที่ผ่านกระบวนการดังกล่าวจะถูกจัดเก็บในรูปแบบที่พร้อมสำหรับการวิเคราะห์ ซึ่งรองรับการทำ aggregation และ filtering ได้อย่างรวดเร็ว การแยกโครงสร้างข้อมูลเชิงวิเคราะห์ออกจากฐานข้อมูลต้นทาง ช่วยลดระยะเวลาเดินทางของข้อมูลบนแดชบอร์ด และสนับสนุนการทำงานแบบเรียลไทม์

**3.3 Real-Time Cultural Analytics Engine**

Real-Time Cultural Analytics Engine เป็นแกนกลางของโมเดล RTCAM ทำหน้าที่ประมวลผลข้อมูลวัฒนธรรมแบบต่อเนื่องและรองรับการอัปเดตแบบเรียลไทม์

ระบบดำเนินการคำนวณข้อมูลที่ต้องการแสดงผล เช่น จำนวนข้อมูลตามหมวดหมู่และพื้นที่ พร้อมทั้งวิเคราะห์แนวโน้มเชิงเวลา เพื่อสะท้อนการเปลี่ยนแปลงของข้อมูลในแต่ละช่วง นอกจากนี้ ยังรองรับการวิเคราะห์เชิงพื้นที่ และการคำนวณความถี่ของคำสำคัญจากข้อมูลข้อความ เพื่อสร้างชุดข้อมูลหลายมิติจากชุดข้อมูลเดียวกัน

เพื่อรักษาประสิทธิภาพของระบบในด้านความเร็วและการเข้าถึง จะคำนวณเฉพาะส่วนที่ได้รับผลกระทบเมื่อมีข้อมูลใหม่เข้าสู่ระบบ แทนการประมวลผลข้อมูลทั้งหมดใหม่ทุกครั้ง กลไกนี้ช่วยลดภาระการคำนวณและลดเวลาในการตอบสนองของแดชบอร์ด

**3.4 Interactive Intelligence Visualization**

Interactive Intelligence Visualization เป็นชั้นการแสดงผลของโมเดล RTCAM ซึ่งทำหน้าที่แปลงผลลัพธ์จากชั้น Real-Time Cultural Analytics Engine ให้กลายเป็นการนำเสนอข้อมูลเชิงภาพแบบโต้ตอบที่สนับสนุนการสำรวจข้อมูลและการตัดสินใจเชิงข้อมูลอย่างมีประสิทธิภาพ ชั้นนี้ถูกออกแบบภายใต้แนวคิด analytics-driven interface โดยให้ความสำคัญกับความรวดเร็วในการตอบสนอง ความสามารถในการกรองข้อมูลหลายมิติ และความชัดเจนในการตีความผลวิเคราะห์

ระบบรองรับการดึงข้อมูลและประมวลผลข้อมูลหลายมิติ ซึ่งผู้ใช้สามารถเลือกข้อมูลตามพื้นที่ ช่วงเวลา หมวดหมู่ และคำสำคัญได้พร้อมกันในครั้งเดียว นอกจากนี้ยังรองรับการวิเคราะห์แบบเจาะลึกและการเปรียบเทียบข้อมูลระหว่างพื้นที่หรือช่วงเวลาที่แตกต่างกัน ทำให้สามารถมองเห็นทั้งความหนาแน่นของข้อมูลและบริบทเชิงเนื้อหาในพื้นที่เดียวกัน

ในเชิงสถาปัตยกรรม การเชื่อมต่อระหว่างชั้นการแสดงผลและการวิเคราะห์ จะใช้ asynchronous query execution เพื่อลดการรอคอยและเพิ่มประสิทธิภาพในการดึงข้อมูล พร้อมทั้งใช้กลไก caching สำหรับตัวชี้วัดที่มีการเรียกใช้งานบ่อยเพื่อลดภาระการคำนวณซ้ำ แนวทางดังกล่าวช่วยรักษาความเสถียรของระบบแม้มีผู้ใช้งานพร้อมกันหลายราย

**3.5 Web Performance Evaluation**

ในการประเมินประสิทธิภาพของระบบ ผู้วิจัยใช้ตัวชี้วัดสำคัญของประสบการณ์ผู้ใช้งาน ตัวชี้วัดที่ใช้ในการประเมินประกอบด้วย

First Contentful Paint (FCP) หมายถึง เวลาที่เนื้อหาแรกปรากฏบนหน้าจอ

Largest Contentful Paint (LCP) หมายถึง เวลาที่เนื้อหาของหน้าถูกแสดงผลครบ

Total Blocking Time (TBT) หมายถึง ช่วงเวลาที่ระบบไม่สามารถตอบสนองต่อผู้ใช้งาน

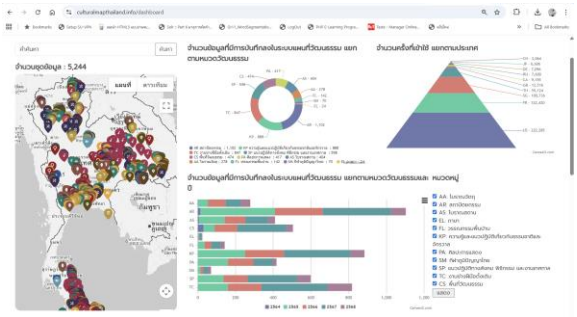
Cumulative Layout Shift (CLS) หมายถึง การประเมินความเสถียรของการแสดงผล

Speed Index (SI) หมายถึง ความเร็วโดยรวมของการแสดงผลเนื้อหา

ตัวชี้วัดเหล่านี้ถูกเลือกใช้เนื่องจากเป็นมาตรฐานในการประเมินประสิทธิภาพของเว็บไซต์ และสามารถสะท้อนประสบการณ์ผู้ใช้งานได้อย่างครอบคลุม การทดสอบดำเนินการในทั้งอุปกรณ์คอมพิวเตอร์ (Desktop) และอุปกรณ์พกพา (Mobile) เพื่อสะท้อนการใช้งานในบริบทที่หลากหลาย

## 4. ผลการดำเนินงาน

โมเดล RTCAM ถูกติดตั้งและทดสอบกับฐานข้อมูลวัฒนธรรมที่ประกอบด้วยข้อมูลจำนวน 5,204 รายการ และผู้ใช้งานมากกว่า 581,465 ครั้ง ครอบคลุมข้อมูลเชิงพื้นที่ (พิกัดภูมิศาสตร์), หมวดหมู่, ช่วงเวลา และข้อมูลเชิงข้อความ ดังภาพที่ 2



**ภาพที่ 2:** ระบบแดชบอร์ดข้อมูลวัฒนธรรรม

### 4.1 Performance Evaluation

ผู้วิจัยได้ทดสอบเวลาตอบสนอง (response time) ในหน้าแดชบอร์ดแต่ละหน้าโดยประกอบด้วยหน้า Map ที่แสดงแผนที่และข้อมูลทางวัฒนธรรมในรูปแบบแผนที่ และหน้า Dashboard ที่แสดงผลข้อมูลในภาพรวม และแบ่งในรูปแบบของการแสดงผลในรูปแบบของอุปกรณ์ Mobile และ Desktop ผลลัพธ์เป็นดังตารางที่ 1

**ตารางที่ 1:** ผลประเมินประสิทธิภาพระบบ

| Page | FCP | LCP | TBT | CLS | SI |
|---|---|---|---|---|---|
| Map (Desktop) | 2.2 sec | 6.6 sec | 0 ms | 0.002 | 4.9 sec |
| Map (Mobile) | 2.0 sec | 6.5 sec | 10 ms | 0.001 | 4.6 sec |
| Dashboard (Desktop) | 2.2 sec | 7.3 sec | 90 ms | 0.001 | 7.9 sec |
| Dashboard (Mobile) | 2.2 sec | 7.4 sec | 170 ms | 0.001 | 7.6 sec |

ผลการทดสอบประสิทธิภาพเว็บไซต์พบว่า ค่า First Contentful Paint อยู่ในช่วง 2.0–2.2 วินาที ซึ่งสะท้อนถึงประสิทธิภาพการตอบสนองของสถาปัตยกรรมที่อยู่ในระดับดี อย่างไรก็ตาม ค่า Largest Contentful Paint ของทุกหน้ามีค่าระหว่าง 6.5–7.4 วินาที ซึ่งสูงกว่ามาตรฐานที่แนะนำอย่างมีนัยสำคัญ โดยเฉพาะหน้า Dashboard ที่มีค่า Total Blocking Time สูงกว่าหน้า Map สะท้อนถึงภาระการประมวลผลบางส่วนและการแสดงผลข้อมูลเชิงกราฟิกที่ซับซ้อนมากกว่า

### 4.2 ผลการประเมินความพึงพอใจ

กลุ่มตัวอย่างจำนวน 17 คน ได้มาจากการเลือกแบบเจาะจง (Purposive Sampling) โดยคัดเลือกผู้เชี่ยวชาญและผู้ใช้งานที่มีประสบการณ์ด้านข้อมูลวัฒนธรรม เพื่อให้สามารถให้ข้อเสนอแนะเชิงลึกต่อระบบได้ แม้ว่าขนาดกลุ่มตัวอย่างจะไม่มาก แต่ถือว่าเพียงพอสำหรับการประเมินเชิงการใช้ง่ายและความพึงพอใจของระบบต้นแบบ ซึ่งเน้นการประเมินเชิงคุณภาพและประสบการณ์ผู้ใช้งานมากกว่าการทดสอบเชิงสถิติในระดับประชากร ซึ่งผลเป็นดังตารางที่ 2

**ตารางที่ 2:** ผลประเมินความพึงพอใจ

| หัวข้อการประเมิน | ค่าเฉลี่ย |
|---|---|
| ระบบใช้งานง่าย ไม่ซับซ้อน | 4.41 |
| ความถูกต้องของข้อมูลที่แสดง | 4.24 |
| ความรวดเร็วในการประมวลผลข้อมูล | 4.06 |
| ความเสถียรของระบบ | 4.29 |
| การเข้าถึงข้อมูลที่ต้องการได้สะดวก | 4.59 |
| ความสวยงามและน่าใช้งานของหน้าจอระบบ | 4.47 |
| ความพึงพอใจโดยรวมต่อระบบ | 4.35 |
| สามารถนำข้อมูลที่ปรากฏไปใช้ประโยชน์ได้ | 4.75 |
| **ผลรวม** | **4.38** |

จากผลการประเมินความพึงพอใจของผู้ใช้งานต่อระบบ พบว่ามีค่าเฉลี่ยรวมเท่ากับ 4.38 ซึ่งอยู่ในระดับ มาก สะท้อนให้เห็นว่าระบบโดยภาพรวมสามารถตอบสนองความต้องการของผู้ใช้งานได้เป็นอย่างดี

## 5. สรุป

งานวิจัยนี้เกิดจากปัญหาการจัดการข้อมูลวัฒนธรรมที่ขาดโครงสร้างเชิงวิเคราะห์ และไม่สามารถสนับสนุนการตัดสินใจเชิงข้อมูลได้อย่างเป็นปัจจุบัน แม้ประเทศไทยจะมีการพัฒนาระบบข้อมูลวัฒนธรรม แต่การนำข้อมูลดังกล่าวมาใช้ประโยชน์เชิงวิเคราะห์ยังมีข้อจำกัด โดยเฉพาะในด้านการบูรณาการข้อมูลหลายมิติ การวิเคราะห์เชิงพื้นที่และเชิงเวลา และการนำเสนอข้อมูลแบบโต้ตอบที่รองรับการใช้งานจริงในระดับนโยบาย

งานวิจัยนี้จึงได้นำเสนอ Real-Time Cultural Analytics Model (RTCAM) ซึ่งเป็นสถาปัตยกรรมเชิงบูรณาการที่ประยุกต์กระบวนการดึงข้อมูล การปรับโครงสร้างข้อมูล การประมวลผลเชิงวิเคราะห์ และการแสดงผลแบบโต้ตอบ

ผลการดำเนินงานแสดงให้เห็นว่าระบบสามารถรองรับข้อมูลจำนวน 5,204 รายการ โดยมีค่า First Contentful Paint อยู่ในช่วง 2.0–2.2 วินาที อยู่ในระดับดี แม้ว่าค่า Largest Contentful Paint จะยังสูงกว่ามาตรฐาน แต่ผลการประเมินความพึงพอใจของผู้ใช้งานมีค่าเฉลี่ยรวม 4.38 อยู่ในระดับ มาก และหัวข้อ สามารถนำข้อมูลไปใช้ประโยชน์ได้ มีค่าเฉลี่ยสูงถึง 4.75 สะท้อนถึงศักยภาพของระบบในการสนับสนุนการวิเคราะห์และการตัดสินใจเชิงข้อมูล

ในงานวิจัยระยะต่อไป จะพัฒนาเพิ่มเติมในด้านการเพิ่มประสิทธิภาพ Largest Contentful Paint การรองรับผู้ใช้งานพร้อมกันจำนวนมาก และการประยุกต์ใช้เทคนิคการเรียนรู้ของเครื่องจักรเพื่อสร้างการทำนายเชิงวิเคราะห์หรือการวิเคราะห์เชิงเครือข่าย เพื่อยกระดับระบบให้เป็นแพลตฟอร์มวิเคราะห์ข้อมูลวัฒนธรรมเชิงอัจฉริยะ

## เอกสารอ้างอิง


[1] F. Windhager et al., “Visualization of Cultural Heritage Collection Data: State of the Art and Future Challenges,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 25, no. 6, pp. 2311–2330, Jun. 2019, doi: 10.1109/TVCG.2018.2830759.

[2] M. Wallace, V. Poulopoulos, A. Antoniou, and M. López-Nores, “An Overview of Big Data Analytics for Cultural Heritage,” *Big Data and Cognitive Computing, vol. 7, no*. 1, p. 14, Mar. 2023, doi: 10.3390/bdcc7010014.

[3] F. Windhager et al., “Visualization of Cultural Heritage Collection Data: State of the Art and Future Challenges,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 25, no. 6, pp. 2311–2330, Jun. 2019, doi: 10.1109/TVCG.2018.2830759.

[4] Q. Zhang, “The Impact of Interactive Data Visualization on Decision-Making in Business Intelligence,” in Proc. *2nd Int. Conf. Management Research and Economic Development*, 2024, doi:10.54254/2754-1169/87/20241056.

[5] S. Hjelle, P. Mikalef, N. Altwaijry, and V. Parida, “Organizational decision making and analytics: An experimental study on dashboard visualizations,” *Information & Management*, vol. 61, no. 6, p. 104011, Sep. 2024, doi: 10.1016/j.im.2024.104011.

[6] Z. Deng et al., “Achieving Trustworthy Real-Time Decision Support Systems with Low-Latency Interpretable AI Models,” Dec. 11, 2025, *arXiv: arXiv:2506.20018*. doi: 10.48550/arXiv.2506.20018.

[7] L. Targa, E. Villuendas, C. Portalés, and J. Sebastián, “Visualising Spatial Dispersion in Cultural Heritage Data,” *ISPRS International Journal of Geo-Information*, vol. 14, no. 7, p. 267, Jul. 2025, doi: 10.3390/ijgi14070267.

[8] S. Hoaihongthong and S. Huaihongthong, “Cultural Heritage Information Management: A Systematic Review Study,” **J. Information and Knowledge Management**, vol. 10, no. 2, pp. 45–57, 2020.

[9] A. Castiglione, F. Colace, V. Moscato, and F. Palmieri, “CHIS: A big data infrastructure to manage digital cultural items,” *Future Generation Computer Systems*, vol. 86, pp. 1134–1145, Sep. 2018, doi: 10.1016/j.future.2017.04.006.

[10] M. Ziku et al., “Digital Cultural Heritage Management for Local Heritage: Overcoming Barriers to Accessibility with Regional Digital Infrastructures,” *Journal of Integrated Information Management*, vol. 9, no. 1, pp. 20–29, Aug. 2024.

# การพัฒนาระบบสืบค้นองค์ความรู้อัจฉริยะสำหรับองค์กรภาครัฐ ด้วยเทคนิคการสืบค้นเชิงความหมายแบบผสมผสาน

# Development of an Intelligent Knowledge Retrieval System for Government Organizations Using Hybrid Semantic Search Techniques

*กมลลักษณ์ ภิญญาคง (Kamonlak Pinyakong)[1] และสุภัคพงศ์ จินารัตน์ (Supakpong Jinarat)[2]*
*ภาควิชาเทคโนโลยีสารสนเทศ วิทยาลัยวิศวกรรมศาสตร์และเทคโนโลยี มหาวิทยาลัยธุรกิจบัณฑิต*
*[1]67130059@dpu.ac.th, [2]supakpong.jin@dpu.ac.th*

## บทคัดย่อ

*ระบบสืบค้นเอกสารแบบดั้งเดิมที่อาศัยการจับคู่คำแบบตรงตัวมักประสบปัญหาการค้นข้อมูลไม่ครบถ้วน โดยเฉพาะเมื่อผู้ใช้ใช้คำพ้องความหมาย สะกดผิด หรือใช้คำศัพท์เฉพาะภายในองค์กร งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาระบบสืบค้นเอกสาร PDF สำหรับองค์กรภาครัฐด้วยเทคนิคการสืบค้นเชิงความหมายแบบผสมผสาน (Hybrid Retrieval) ซึ่งผสานการค้นหาเชิงศัพท์ด้วย BM25 และการค้นคืนข้อมูลแบบเวกเตอร์บน Elasticsearch โดยพัฒนาระบบด้วย FastAPI และใช้ Sentence Transformer สำหรับสร้างเวกเตอร์ตัวแทนความหมายของเอกสารและคำค้น เพื่อประเมินประสิทธิภาพของระบบที่พัฒนาขึ้น ได้ทำการทดลองกับเอกสาร 100 ฉบับ และคำค้น 300 คำ โดยเปรียบเทียบ Embedding Models หลายรูปแบบภายใต้สถาปัตยกรรมเดียวกัน และประเมินด้วย Recall@K ร่วมกับค่า Similarity Threshold ที่ 0.6 ผลการทดลองพบว่า BAAI/bge-base-en-v1.5 ให้ผลดีที่สุดในชุดทดลองนี้ และช่วยให้ระบบสามารถค้นพบเอกสารที่เกี่ยวข้องได้ครอบคลุมกว่าการพิจารณาจากชุดคำตอบอ้างอิงเพียงอย่างเดียว ผลการวิจัยสะท้อนว่าระบบ Hybrid Retrieval ที่พัฒนาขึ้นสามารถประยุกต์ใช้เพื่อเพิ่มประสิทธิภาพการสืบค้นองค์ความรู้ในองค์กรภาครัฐได้ โดยโมเดล Embedding ที่เลือกใช้มีผลต่อคุณภาพของผลลัพธ์อย่างมีนัยสำคัญ*

**คำสำคัญ:** การค้นหาเชิงความหมาย, ระบบสืบค้นเอกสาร, การค้นข้อมูลแบบไฮบริด, เวกเตอร์ตัวแทนความหมาย

## Abstract

*Traditional document retrieval systems based on exact keyword matching often suffer from incomplete results, particularly when users employ synonyms, misspell terms, or use organization-specific terminology. This study aims to develop a hybrid semantic retrieval system for PDF documents in government organizations by combining BM25-based lexical retrieval and dense vector retrieval on Elasticsearch. The system was implemented using FastAPI, while Sentence Transformer models were employed to generate embeddings for documents and queries. To evaluate the developed system, experiments were conducted on 100 PDF documents and 300 queries by comparing multiple embedding models under the same hybrid retrieval architecture. The evaluation was performed using Recall@K together with a similarity threshold of 0.6. The results showed that BAAI/bge-base-en-v1.5 achieved the best performance in this experimental setting and enabled the system to retrieve additional relevant documents beyond the predefined ground truth set. These findings suggest that the proposed hybrid retrieval system is effective for improving organizational knowledge retrieval, while the choice of embedding model substantially affects retrieval quality.*

**Keywords:** Semantic Search, Elasticsearch, Sentence Transformer, Embeddings, PDF Retrieval.

## 1. บทนำ

ในยุคเศรษฐกิจดิจิทัล (Digital Economy) การบริหารจัดการองค์ความรู้และการเข้าถึงข้อมูลอย่างถูกต้องและทันท่วงทีเป็นปัจจัยสำคัญต่อประสิทธิภาพการดำเนินงานของหน่วยงานภาครัฐ โดยเฉพาะองค์กรที่มีภารกิจด้านการกำกับดูแลกิจการกระจายเสียง โทรทัศน์ และโทรคมนาคม ซึ่งต้องจัดการเอกสารจำนวนมาก เช่น กฎหมาย ประกาศ ระเบียบ รายงานโครงการวิจัย และข้อมูลสถิติ อันเป็นฐานสำคัญของการกำหนดนโยบายและการตัดสินใจเชิงกลยุทธ์ อย่างไรก็ตาม ระบบสืบค้นข้อมูลที่ใช้อยู่ยังอาศัยการค้นหาแบบคำสำคัญ (Keyword-based Search) ซึ่งมีข้อจำกัดในการเข้าใจความหมายและบริบทของคำ

ส่งผลให้เอกสารที่เกี่ยวข้องไม่ถูกค้นพบ โดยเฉพาะกรณีคำพ้องความหมายหรือคำศัพท์เฉพาะภายในองค์กร ปัญหาดังกล่าวสะท้อนข้อจำกัดของการค้นหาแบบคำสำคัญที่ไม่สามารถประมวลผลความสัมพันธ์เชิงความหมายของข้อความได้อย่างแท้จริง และส่งผลให้ประสิทธิภาพการเข้าถึงองค์ความรู้ลดลง จากข้อจำกัดของระบบค้นหาแบบคำสำคัญ งานวิจัยนี้จึงมุ่งดำเนินการใน 2 ส่วนหลัก คือ พัฒนาระบบสืบค้นเอกสาร PDF สำหรับองค์กรภาครัฐด้วยสถาปัตยกรรม Hybrid Retrieval บน Elasticsearch ซึ่งผสานการค้นหาเชิงศัพท์ด้วย BM25 และการค้นคืนข้อมูลแบบเวกเตอร์ และประเมินประสิทธิภาพของระบบที่พัฒนาขึ้นโดยเปรียบเทียบ embedding models หลายรูปแบบภายใต้สถาปัตยกรรมเดียวกัน เพื่อคัดเลือกโมเดลที่เหมาะสมต่อการใช้งานจริงในบริบทองค์กรภาครัฐ

## 2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ผู้วิจัยได้ศึกษาค้นคว้าเอกสาร และงานวิจัยที่เกี่ยวข้อง เพื่อนำทฤษฎี กรอบแนวคิดนำมาประยุกต์ใช้ในงานวิจัย โดยแบ่งออกเป็น 4 หัวข้อหลัก ดังนี้

**2.1 พื้นฐานของการสืบค้นเชิงความหมายและการรับรู้บริบท**

การเปลี่ยนผ่านจากการสืบค้นเชิงคำศัพท์ไปสู่การสืบค้นเชิงความหมายสะท้อนถึงพัฒนาการโดยรวมของงานประมวลผลภาษาธรรมชาติ [1] อธิบายว่าการสืบค้นเชิงความหมายอาศัยการแทนข้อมูลเชิงบริบท การฝังตัวแบบนิวรัล และโครงสร้างดัชนีขั้นสูง เพื่อแก้ไขปัญหาความไม่ตรงกันของคำศัพท์และเพิ่มความเข้าใจเจตนาของผู้ใช้ งานวิจัยระยะแรกแสดงให้เห็นว่าการใช้ความรู้เชิงโครงสร้างเฉพาะโดเมนสามารถเพิ่มความสอดคล้องเชิงความหมายของผลการสืบค้นได้ ตัวอย่างเช่น [2] ได้พัฒนาเครื่องมือสืบค้นเชิงความหมายที่ขับเคลื่อนด้วยกราฟความรู้สำหรับแพทย์แผนไทย ซึ่งแสดงให้เห็นว่าความสัมพันธ์เชิงความหมายระหว่างเอนทิตีช่วยให้ผลการสืบค้นมีความคงเส้นคงวา แม้คำค้นของผู้ใช้จะมีความแตกต่างทางถ้อยคำ ในทำนองเดียวกัน เทคนิคการวัดความคล้ายคลึงเชิงความหมายถูกนำมาประยุกต์ใช้ในห้องสมุดอิเล็กทรอนิกส์ เพื่อปรับปรุงแบบจำลองการจัดอันดับโดยอาศัยทั้งโครงสร้างทางไวยากรณ์และความหมาย [3] ขณะที่ระบบที่ใช้ RDF และออนโทโลยีได้ขยายขีดความสามารถของการสืบค้น โดยฝังข้อมูลเมทาดาทาเชิงโครงสร้างและความสัมพันธ์ของเอนทิตีเข้าไปในกระบวนการสืบค้น [4]

**2.2 การสืบค้นแบบอาศัยการฝังเวกเตอร์และแบบจำลองทรานส์ฟอร์เมอร์**

การเติบโตของเทคนิคการเรียนรู้เชิงลึกส่งผลให้การฝังเวกเตอร์ด้วยแบบจำลองทรานส์ฟอร์เมอร์กลายเป็นแกนหลักของการสืบค้นเชิงความหมายในยุคปัจจุบัน [5] ได้นำเสนอระบบสืบค้นเชิงความหมายภาษาญี่ปุ่นที่ผสาน JSBERT เข้ากับตัวดำเนินการเชิงตรรกะ และรายงานผลการประเมินที่มีค่า MAP, Precision, Recall และ F1 อยู่ในระดับสูง [6] แสดงให้เห็นว่าการผนวกองค์ความรู้เชิงศัพท์ที่มนุษย์จัดทำขึ้นเข้าไปในเวกเตอร์ทรานส์ฟอร์เมอร์ช่วยเพิ่มความสามารถในการแยกแยะเชิงความหมายและคุณภาพของการจัดอันดับ การประยุกต์ใช้ในโดเมนเอกสารทางกฎหมายยิ่งตอกย้ำข้อได้เปรียบของเวกเตอร์เชิงบริบท โดยเทคนิคการเรียนรู้ของเครื่องและ NLP ถูกนำมาใช้ในการจำแนกเชิงความหมาย การตอบคำถามทางกฎหมาย และการวิเคราะห์ความคล้ายคลึงของเอกสารในระบบผู้ช่วยกฎหมายเสมือน [7] ในด้านการศึกษา แบบจำลองการจัดอันดับที่ขับเคลื่อนด้วยเวกเตอร์ช่วยเพิ่มความแม่นยำในการสืบค้นสื่อการเรียนรู้และทรัพยากรข้อความ [8], [9]

**2.3 กราฟความรู้และการสร้างแบบจำลองเชิงความหมายเฉพาะโดเมน**

นอกเหนือจากเวกเตอร์ทั่วไป การสืบค้นเชิงความหมายยังได้รับประโยชน์จากการแทนข้อมูลเชิงโครงสร้างเฉพาะโดเมน งานวิจัยด้านกราฟความรู้ที่สร้างด้วยเทคนิคการเรียนรู้เชิงลึกช่วยเสริมความสัมพันธ์ของเอนทิตีและการจำแนกเชิงความหมายในระบบโครงข่ายไฟฟ้าอัจฉริยะ ส่งผลให้ความเกี่ยวข้องของผลการสืบค้นดีขึ้น [10] ในระบบสืบค้นองค์ความรู้ด้านการศึกษา การใช้ความคล้ายคลึงเชิงความหมาย กราฟแนวคิด และการปรับปรุงการจัดอันดับช่วยเพิ่มความแม่นยำในการค้นหางานวิจัยเหล่านี้สะท้อนถึงคุณค่าของการบูรณาการความสัมพันธ์เชิงความหมายแบบมีโครงสร้าง ไม่ว่าจะเป็นออนโทโลยี กราฟความรู้หรือโมเดลแบบผสมผสาน โดยเฉพาะในโดเมนที่ใช้ความรู้เป็นฐานอย่างเข้มข้น

**2.4 ตัวชี้วัดการประเมินผลในงานวิจัยด้าน Semantic Search**

งานวิจัยที่ได้รับการทบทวนใช้วิธีการประเมินผลที่หลากหลาย แต่มีแนวโน้มไปในทิศทางเดียวกันในด้านตัวชี้วัดมาตรฐานของการสืบค้นสารสนเทศ ค่า Precision, Recall และ F1-score ถูกใช้บ่อยในงานประเมินระบบสืบค้นแบบอาศัยเวกเตอร์ [11], [12] ขณะที่ประสิทธิภาพการจัดอันดับถูกประเมินด้วย MAP และ NDCG ในงานที่ใช้ทรานส์ฟอร์เมอร์และออนโทโลยี ระบบที่ใช้กราฟความรู้ให้ความสำคัญกับความสอดคล้องเชิงความหมายและการจับคู่ความเกี่ยวข้อง การประเมินระบบสืบค้นที่ขับเคลื่อนด้วยปัญญาประดิษฐ์ยังเปรียบเทียบความถูกต้องของคำตอบ ความสอดคล้องเชิงบริบท และความพึงพอใจของผู้ใช้กับระบบแบบอิงคำสำคัญ [13] โดยภาพรวมวรรณกรรมทั้งหมดมีความเห็นสอดคล้องกันว่า Semantic Search ให้ความแม่นยำสูงกว่า มีความเกี่ยวข้องเชิงบริบทที่ดีกว่าและตอบสนองความต้องการของผู้ใช้ได้ดีกว่าการสืบค้นแบบดั้งเดิม

## 3. วิธีการดำเนินการวิจัย

**3.1 การเตรียมข้อมูลเอกสาร (Document Preparation)**

การวิจัยนี้ใช้เอกสาร PDF ประเภทข้อความ (Text-based PDF) จำนวน 100 เล่ม ซึ่งคัดเลือกจากเอกสารที่ใช้จริงในบริบทองค์กรภาครัฐ ครอบคลุมเอกสารหลายประเภท เช่น กฎหมาย คู่มือปฏิบัติงาน รายงานโครงการศึกษาวิจัย และเอกสารเผยแพร่เชิงวิชาการ ทั้งนี้เลือกเฉพาะเอกสารที่สามารถสกัดข้อความได้โดยตรงและเหมาะสมต่อการนำไปสร้างดัชนีสำหรับการสืบค้นเชิงความหมาย

**3.2 การเตรียมชุดคำค้น (Query Preparation)**

เพื่อทดสอบระบบ ได้จัดทำชุดคำค้น 3 ประเภท ได้แก่ คำทั่วไป (General terms) ใช้ในบริบทหลากหลาย คำเฉพาะ (Domain-specific terms) ซึ่งเป็นคำศัพท์เฉพาะทางหรือคำที่ใช้ภายในองค์กร และคำกำกวม (Ambiguous terms) เป็นคำค้นสั้นที่มีได้หลายความหมาย โดยกำหนดคำค้นประเภทละ 100 คำ รวมทั้งสิ้น 300 คำ สำหรับการทดลองและประเมินผล

**3.3 การสร้างชุดเฉลย (Ground Truth) สำหรับการประเมินผล**

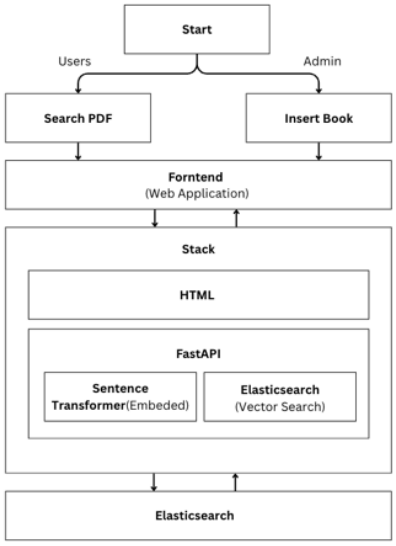
สำหรับการสร้างชุดเฉลย ผู้วิจัยกำหนดเอกสารอ้างอิงที่เกี่ยวข้องสำหรับแต่ละคำค้นโดยอาศัยการพิจารณาความสอดคล้องด้านเนื้อหา วัตถุประสงค์ของเอกสาร และความสัมพันธ์เชิงความหมายของคำค้นกับสาระสำคัญในเอกสาร มิได้พิจารณาเฉพาะการปรากฏของคำแบบตรงตัวเท่านั้น

**3.4 การตรวจสอบความถูกต้องโดยผู้เชี่ยวชาญ (Expert Validation)**

สร้างความน่าเชื่อถือของผลการประเมินชุดเฉลย (Ground Truth) ได้รับการทบทวนร่วมกับผู้เชี่ยวชาญด้านเนื้อหา โดยพิจารณาความสอดคล้องระหว่างคำค้นและเอกสารในเชิงบริบทและความหมาย ไม่จำกัดเฉพาะการปรากฏของคำแบบตรงตัว ทั้งนี้ ผู้เชี่ยวชาญได้ตรวจสอบความเหมาะสมของการจำแนกประเภทคำค้น และยืนยันความเกี่ยวข้องของเอกสารที่กำหนดไว้ในชุดเฉลย ในกรณีที่เกิดความเห็นไม่สอดคล้องกัน ได้มีการพิจารณาและปรับแก้ร่วมกันเพื่อให้ได้ข้อสรุปที่เหมาะสมก่อนนำชุดเฉลยไปใช้ในการประเมินผล

**3.5 ขั้นตอนการดำเนินงานของระบบ (System Workflow)**

3.5.1 สถาปัตยกรรมและกระบวนการทำงานของระบบ

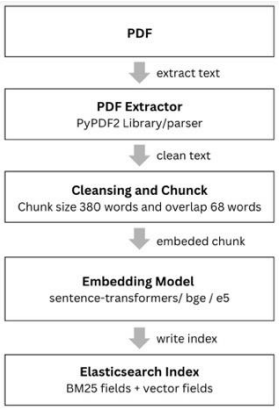



**ภาพที่ 1:** สถาปัตยกรรมและกระบวนการทำงานของระบบ

โดยภาพนี้แสดงภาพรวมการทำงานของระบบค้นหาเอกสาร PDF โดยแบ่งผู้ใช้งานออกเป็น 2 บทบาท คือผู้ใช้ทั่วไป (Users) สำหรับค้นหาเอกสาร และผู้ดูแลระบบ (Admin) สำหรับนำเข้าเอกสาร (Insert Book) ทั้งสองส่วนทำงานผ่านหน้าเว็บ (Frontend Web Application) ซึ่งเชื่อมต่อกับระบบหลังบ้านที่พัฒนาด้วย FastAPI เพื่อประมวลผลคำขอ เมื่อผู้ใช้ทำการค้นหา ระบบจะนำคำค้นไปสร้างเวกเตอร์ด้วย Sentence Transformer (Embedding) แล้วส่งไปยัง Elasticsearch เพื่อทำการค้นหา

แบบเวกเตอร์ (Vector Search) จากนั้นระบบจะจัดอันดับผลลัพธ์และส่งข้อมูลเอกสารกลับไปแสดงบนหน้า UI ส่วนกรณีการนำเข้าเอกสาร ระบบจะประมวลผลและจัดเก็บข้อมูลที่จำเป็นไว้ใน Elasticsearch เพื่อให้สามารถสืบค้นได้อย่างมีประสิทธิภาพ

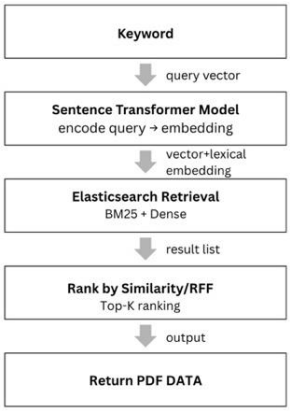
3.5.2 กระบวนการเตรียมข้อมูลและสร้างดัชนีเอกสาร PDF (Indexing Pipeline)




**ภาพที่ 2:** กระบวนการเตรียมข้อมูลและสร้างดัชนีเอกสาร PDF (Indexing Pipeline)

กระบวนการพัฒนาระบบเริ่มจากการนำไฟล์ PDF มา Extract Text ด้วย PyPDF2 และปรับปรุงคุณภาพข้อมูลผ่านขั้นตอน Cleansing เพื่อลบส่วนที่ไม่จำเป็น เช่น Header/Footer และช่องว่างซ้ำ จากนั้นทำ Chunking ขนาด 380 คำ พร้อม Overlap 68 คำ เพื่อรักษาบริบทของเนื้อหา แต่ละ Chunk ถูกแปลงเป็น Embedding Vector ด้วยโมเดล sentence-transformers (เช่น BGE หรือ E5) เพื่อสร้างตัวแทนเชิงความหมายสำหรับคำนวณ semantic similarity และข้อมูลจะถูกจัดเก็บใน Elasticsearch โดยใช้ทั้ง BM25 และ Vector Fields (Semantic Search) เพื่อทำ Hybrid Search ซึ่งช่วยเพิ่ม Recall ทำให้การค้นข้อมูลมีความครอบคลุมและแม่นยำมากขึ้น

3.5.3 กระบวนการค้นหา PDF แบบ Semantic Search




**ภาพที่ 3:** กระบวนการค้นหา PDF แบบ Semantic Search

ภาพแสดงขั้นตอน Retrieval และ Ranking หลังจากจัดทำดัชนีเอกสารแล้ว โดยเริ่มจากผู้ใช้ป้อน Keyword (Query) ซึ่งถูกแปลงเป็น query embedding ด้วย Sentence Transformer Model จากนั้นระบบทำ Hybrid Retrieval ใน Elasticsearch โดยผสานคะแนนจาก BM25 และ Dense Vector Similarity (Semantic Matching) เพื่อสร้างรายการเอกสารเบื้องต้น ผลลัพธ์จะถูกจัดอันดับด้วย Similarity ranking หรือ RRF (Reciprocal Rank Fusion) เพื่อคัดเลือกแบบ Top-K ก่อนส่งเอกสาร PDF ที่เกี่ยวข้องที่สุดให้ผู้ใช้

## 4. ผลการวิจัยและอภิปรายผล

การทดลองใช้ชุดเฉลยคำค้นจำนวน 300 คำ แบ่งเป็นคำทั่วไป คำเฉพาะ และคำกำกวม ประเภทละ 100 คำ เพื่อสะท้อนรูปแบบการใช้งานจริงของระบบ การประเมินผลใช้ค่า Similarity Value เปรียบเทียบกับค่า Similarity Threshold ที่ 0.6 เพื่อพิจารณาความเกี่ยวข้องของเอกสาร และใช้ตัวชี้วัด Recall@K (K = 50) ในการประเมินความสามารถของระบบในการค้นหาเอกสารที่เกี่ยวข้องภายในผลลัพธ์ลำดับต้น ๆ การทดสอบดำเนินการโดยเปรียบเทียบประสิทธิภาพของโมเดลหลายรูปแบบ ซึ่งผลลัพธ์สรุปแสดงในตารางถัดไป

**ตารางที่ 1:** ประสิทธิภาพโมเดล BAAI/bge-base-en-v1.5

| **Model: BAAI/bge-base-en-v1.5** | | | |
|---|---|---|---|
| | คำทั่วไป | คำเฉพาะ | คำกำกวม |
| Recall@K | 70.67% | 71.05% | 69.39% |
| Similarity AVG | 0.81 | 0.83 | 0.84 |

ผลการประเมินโมเดล BAAI/bge-base-en-v1.5 แสดงให้เห็นว่า Recall@K สูงสุดในกลุ่มคำเฉพาะ (71.05%) รองลงมา คือ คำทั่วไป (70.67%) และคำกำกวม (69.39%)

**ตารางที่ 2:** ประสิทธิภาพโมเดล sentence-transformers/ paraphrase-mpnet-base-v2

| **Model: sentence-transformers/ paraphrase-mpnet-base-v2** | | | |
|---|---|---|---|
| | คำทั่วไป | คำเฉพาะ | คำกำกวม |
| Recall@K | 70.59% | 68.90% | 66.86% |
| Similarity AVG | 0.78 | 0.79 | 0.73 |

ผลการประเมินโมเดล sentence transformers/paraphrase-mpnet-base-v2 พบว่า Recall@K สูงสุดในกลุ่มคำทั่วไป (70.59%) รองลงมาคือคำเฉพาะ (68.90%) และคำกำกวม (66.86%)

**ตารางที่ 3:** ประสิทธิภาพโมเดล sentence-transformers/paraphrase-MiniLM-L12-v2

| **Model: sentence-transformers/paraphrase-MiniLM-L12-v2** | | | |
|---|---|---|---|
| | คำทั่วไป | คำเฉพาะ | คำกำกวม |
| Recall@K | 69.75% | 66.19% | 66.19% |
| Similarity AVG | 0.79 | 0.76 | 0.76 |

ผลการประเมินโมเดล sentence-transformers/paraphrase-MiniLM-L12-v2 พบว่า Recall@K สูงสุดในกลุ่มคำทั่วไป (69.75%) ขณะที่คำเฉพาะและคำกำกวมมีค่าเท่ากัน (66.19%) แสดงว่าโมเดลทำงานได้ดีกับคำค้นทั่วไปมากกว่า

**ตารางที่ 4:** ประสิทธิภาพโมเดล intfloat/multilingual-e5-base

| **Model: intfloat/multilingual-e5-base** | | | |
|---|---|---|---|
| | คำทั่วไป | คำเฉพาะ | คำกำกวม |
| Recall@K | 65.20% | 66.13% | 63.66% |
| Similarity AVG | 0.90 | 0.91 | 0.87 |

ผลการประเมินโมเดล intfloat/multilingual-e5-base พบว่า Recall@K สูงสุดในกลุ่มคำเฉพาะ (66.13%) รองลงมาคือคำทั่วไป (65.20%) และคำกำกวม (63.66%)

ผลการทดลองเปรียบเทียบโมเดลพบว่า BAAI/bge-base-en-v1.5 ให้ค่า Recall@K สูงสุดโดยรวม และมีค่า Similarity Average อยู่ในระดับสูงและสม่ำเสมอทุกประเภทคำค้น โดยเฉพาะคำเฉพาะ จึงให้สมดุลระหว่างความครอบคลุมของการค้นและความใกล้เคียงเชิงความหมายได้ดีที่สุดในชุดทดลองนี้ ขณะที่ paraphrase-mpnet-base-v2 และ paraphrase-MiniLM-L12-v2 ให้ผลลัพธ์ใกล้เคียงกัน โดยมี Recall ดีในคำทั่วไป แต่ Similarity ต่ำกว่าเล็กน้อย ส่วน multilingual-e5-base มี Similarity สูงที่สุด แต่ Recall ต่ำกว่าโมเดลอื่น นอกจากนี้การทดสอบกรณีคำว่า “คลื่นรบกวน” (คำทั่วไป) พบว่าแม้ชุดเฉลยมี 7 เล่ม แต่โมเดล bge-base-en-v1.5 สามารถค้นพบเอกสารที่เกี่ยวข้องได้ถึง 35 เล่ม (Similarity เฉลี่ย 0.8682) สะท้อนศักยภาพในการดึงเอกสารที่เกี่ยวข้องเพิ่มเติมนอกเหนือจากชุดเฉลยได้อย่างมีประสิทธิภาพ ขณะเดียวกันกรณีที่ไม่พบเอกสารตามชุดเฉลย ค่า Similarity ต่ำกว่า Threshold (0.6) จึงถูกตัดออกตามเงื่อนไขความเกี่ยวข้องโดยค่า Similarity Average แสดงข้อมูลดังกล่าวในตารางนี้

**ตารางที่ 5:** แสดง **Similarity Average** กรณีไม่พบเอกสารตามชุดเฉลย

| | **Similarity Average** | | | |
|---|---|---|---|---|
| **ประเภทคำ/ Model** | **BAAI/ bge-base-en-v1.5** | **sentence-transformers/ paraphrase-mpnet-base-v2** | **sentence-transformers/paraphrase-MiniLM-L12-v2** | **intfloat/ multilingual-e5-base** |
| คำทั่วไป | 0.1569 | 0.1640 | 0.1627 | 0.1647 |
| คำเฉพาะ | 0.1890 | 0.1858 | 0.1896 | 0.1880 |
| คำกำกวม | 0.1388 | 0.1414 | 0.1427 | 0.1408 |

จากผลการพิสูจน์ข้างต้น พบว่าค่า Similarity ที่ต่ำกว่าค่า Threshold อย่างมีนัยสำคัญสนับสนุนข้อสังเกตว่าคำค้นและหนังสือในชุดเฉลยบางส่วนอาจไม่ได้มีความเกี่ยวข้องเชิงความหมายอย่างแท้จริง ซึ่งอาจสะท้อนว่าชุดเฉลยเดิมยังไม่ครอบคลุมเอกสารที่มีความเกี่ยวข้องเชิงความหมายทั้งหมด หรืออาจมีความคลาดเคลื่อนบางส่วนจากกระบวนการจัดทำชุดเฉลยโดยมนุษย์

## 5. สรุป

งานวิจัยนี้พัฒนาระบบสืบค้นเอกสารด้วยเทคนิค Semantic Search โดยใช้หลาย Embedding Models เพื่อประเมินความเกี่ยวข้องเชิงความหมาย ผลการทดลองพบว่า สามารถดึงเอกสารที่เกี่ยวข้องเพิ่มเติมนอกเหนือจากชุดเฉลยได้อย่างมีประสิทธิภาพ โดยเฉพาะคำทั่วไป เช่น “คลื่นรบกวน” ที่ให้ค่า Similarity เฉลี่ยสูงถึง 0.8682 ส่วนกรณีที่ไม่พบเอกสารในชุดเฉลย มีค่า Similarity ต่ำกว่า Threshold (0.6) อย่างชัดเจน สะท้อนความเหมาะสมของเกณฑ์คัดกรอง และชี้ถึงความเป็นไปได้ของความคลาดเคลื่อนในชุดเฉลย แนวทางพัฒนาในอนาคต ได้แก่ การบูรณาการ OCR เพื่อรองรับเอกสารภาพ

และการทำ User-driven Fine-tuning เพื่อเพิ่มความแม่นยำตามบริบทการใช้งานจริง

## เอกสารอ้างอิง


[1] Y. Zheng, “An Analysis of the Technical Trend of Semantic Search in Natural Language Processing,” in *Proc. 2023 9th Annu. Int. Conf. Network and Information Systems for Computers (ICNISC)*, pp. 51–53, 2023. DOI: 10.1109/ICNISC60562.2023.00033.

] 2[ N. Sakulpakdee, W. Lee, N. Narischat, L.Pipanmekaporn, K. Srisomboon, and A. Prayote, “Semantic Search Engine, Thai Traditional Medicine Knowledge Graph Application,” in *Proc. 2025 13th Int. Elect. Eng. Cong. (iEECON)*, Bangkok, Thailand, pp. 1–5, 2025. DOI: 10.1109/iEECON64081.2025.10987888.

[3] E. D. Oktaviyani, A. Lestari, and Licantik, “Syntactic Semantic Similarity Approach for Electronic Library *Keyword Search,” in Proc. 2022 9th Int. Conf. Information Technology, Computer, and Electrical Engineering (ICITACEE)*, Semarang, Indonesia, Aug. 25–26, pp. 258–262, 2022. DOI: 10.1109/ICITACEE55701.2022.9924152.

]4[ U. Yadav and N. Duhan, “Efficient Retrieval of Data Using Semantic Search Engine Based on NLP and RDF,” *Journal of Web Engineering, vol. 20, no. 8*, pp. 2285–2318, Oct. 2021. DOI: 10.13052/JWE1540-9589.2084.

[5] K. Tsumuraya, M. Uehara, and Y. Adachi, “Semantic Search of Japanese Sentences Using Distributed Representations,” in *Proc. 2023 Int. Symp. Comput. Netw. Workshops (CANDARW)*, pp. 76–81, 2023. DOI: 10.1109/CANDARW60564.2023.00021.

[6] Z. Weinfeld and L. Stanchev, “Enhancing Transformer-Based Semantic Search by Integrating Human-Crafted Knowledge into Sentence Embeddings,” in *Proc. 2025 IEEE Int. Conf. Semantic Comput. (ICSC)*, pp. 118–121, 2025. DOI: 10.1109/ICSC64641.2025.00022.

]7[ G. Samardzhiev and M. Nisheva-Pavlova, “Application of Machine Learning and Natural Language Technologies in Building Semantic Search Systems: Case Study of a Virtual Legal Assistant,” in *Proc. 2023 Int. Conf. Big Data, Knowl. and Control Syst.* Eng. (BdKCSE), 2023. DOI: 10.1109/BDKCSE59280.2023.10339730.

[8 ] S. Choudhary, H. K. Jaiswal, A. S. Rajeevan, P. D. Mukherjee, and S. Hudnurkar,“Semantic Search of Educational Texts,” in *Proc. 2024 Int. Conf. Intelligent Syst. Adv. Appl. (ICISAA)*, Pune, India, pp. 1–5, 2024. DOI: 10.1109/ICISAA62385.2024.10828786.

[9] C. Limongelli, M. Lombardi, A. Marani, and D. Taibi, “A Semantic Approach to Ranking Techniques: IEEE Access, vol. 10, pp. 68885–68896, 2022. DOI: 10.1109/ACCESS.2022.3186356.

[10] C. Liu, Z. Yao, L. Liu and Q. Lei, “Research on Semantic Search Algorithm of Smart Grid Knowledge Map Based on Deep Learning,” in *Proc. 2023 Int. Conf. Network Engineering, Tech. Innov. and Comput. (ICNETIC)*, pp. 689–692, 2023. DOI: 10.1109/ICNETIC59568.2023.00147.

[11] N. Salsabilla and K. Wiharja, “Implementation of Semantic Search Based on Vector Database for Personal Documents,” in *Proc. 2025 Int. Conf. Advancement in Data Science, E-learning and Information System (ICADEIS)*, 2025, DOI: 10.1109/ICADEIS65852.2025.10933119.

[12] C. Gavankar, T. Bhosale, D. Gunda, A. Chavan, and S. Hassan, “A Comparative Study of Semantic Search Systems,” in *Proc. 2020 Int. Conf. Comput. Commun. Informatics (ICCCI), Coimbatore*, India, Jan, pp. 1–6, 2020.

[13] O.F. Karpinski, L. N. Mcfetridge, M. C. Ta, F. N. Ashriem and A. Singh, “Evaluating AI-Assisted Search vs. Traditional Keyword Search: A Comparative Study of ChatGPT and Google,” in *Proc. 2025 IEEE World AI IoT Congress (AIIoT)*, pp. 0982 –0987, 2025. DOI: 10.1109/AIIoT65859.2025.11105219.

[14] B. S. Magrey, A. Chauhan, Ramneet, and S. Singh, “Collaborative Platform for Content Creators and Brands Based on Semantic Search,” in *Proc. 2024 2nd Int. Conf. Adv. Comput., Commun. and Inf. Technol. (ICAICCIT)*, pp. 1121–1125, 2024. DOI: 10.1109/ICAICCIT64383.2024.10912213.

[15] F. Errico, R. Barriera, A. Corallo, and M. Prato, “Dematerialization, Archiving and Recovery of Documents: A Proposed Tool Based on a Semantic Classifier and a Semantic Search Engine,” in *Proc. 2020 Int. Conf. Inf. Technol. Manage. (ICITM)*, pp. 297–301, 2020.

[16] K. Sawarkar, A. Mangal and S. R.Solanki, “Blended RAG: Improving RAG Retriever-Augmented Generation Accuracy with Semantic Search and Hybrid Query-Based Retrievers,” in *Proc. 2024 IEEE 7th Int. Conf. Multimedia Inf. Process. Retrieval (MIPR)*, pp. 155–161, 2024. DOI: 10.1109/MIPR62202.2024.00031.

[17] M. Guo, D.Gotz and Yue Wang, “How Does Imperfect Automatic Indexing Affect Semantic Search Performance?,” in *Proc. 2023 IEEE 11th Int. Conf. Healthcare Informatics (ICHI)*, pp. 588–596, 2022. DOI: 10.1109/ICHI57859.2023.00105.

[18] M. Rovatsos and R. Filgueira, “Building Lightweight Semantic Search Engines,” in *Proc. 2023 IEEE 19th Int. Conf. e-Science (e-Science)*, 2023. DOI: 10.1109/E-SCIENCE58273.2023.10254874.

[19] S. Sankar, H. Muslihuddeen, S. Ostwal, P. Sathvika and A. K. Madasamy, “CISER: Customized Institute-Specific Search Engine for Retrieving Research Papers,” in *Proc. 2023 IEEE 4th Int. Conf. Innovative Trends in Information Technology (ICITIIT)*, 2023. DOI: 10.1109/ICITIIT57246.2023.10068620.

[20] Z. Pan, “Optimization of Information Retrieval Algorithm for Digital Library Based on Semantic Search Engine,” in Proc. 2023 Int. Conf. Comput. Eng. Artif. Intell. (ICCEAI), pp. 364–367, 2023. DOI: 10.1109/ICCEA50009.2020.00085.

# การพัฒนาระบบแนะนำโปรโมชันสำหรับธุรกิจค้าปลีกแบบขายส่งโดยบูรณาการการทำเหมืองข้อมูล การเรียนรู้เชิงลึก และอัลกอริทึมแบนดิตแบบอิงบริบท

# Developing a Promotion Recommendation System for B2B Wholesale Retail through the Integration of Data Mining, Deep Learning, and Contextual Bandit Algorithms

*ธีระเดช คัญชิง (Teeradaj Kanching)[1] และพยุง มีสัจ (Phayung Meesad)[2]*
*ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล*
*มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ*
*[1]s6707021856141@email.kmutnb.ac.th, [2]Phayung.m@itd.kmutnb.ac.th*

## บทคัดย่อ

*การวางแผนโปรโมชันในธุรกิจค้าปลีกแบบขายส่งต้องอาศัยข้อมูลหลายมิติ ทั้งพฤติกรรมลูกค้า ความสัมพันธ์ของสินค้า อุปสงค์ และผลกระทบต่อกำไร งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาต้นแบบระบบแนะนำโปรโมชันโดยบูรณาการ K-Means, LSTM, กฎความสัมพันธ์, โมเดลการเรียนรู้เชิงลึกแบบพหุสาขา และอัลกอริทึม LinUCB แบบแบนดิตอิงบริบท ร่วมกับกลไก Profit Guard เพื่อลดความเสี่ยงจากโปรโมชันที่อาจทำให้กำไรสุทธิติดลบ การทดลองใช้ข้อมูลธุรกรรมย้อนหลังจากกรณีศึกษาธุรกิจค้าปลีกแบบขายส่ง และประเมินผลผ่านการจำลอง 10,000 รอบ การทดสอบสถานการณ์ตลาด 3 รูปแบบ และการทดสอบกับสินค้าที่ไม่เคยพบมาก่อน ผลการทดลองพบว่า LinUCB ให้ผลตอบแทนสะสมสูงกว่านโยบายสุ่มร้อยละ 21.01 มีค่าการปรับปรุงในสถานการณ์ตลาดต่าง ๆ ร้อยละ 21.79–23.69 และให้ผลกับ Unseen SKU ร้อยละ 22.01 โดยมีช่องว่างการสรุปผลกับข้อมูลใหม่ร้อยละ −1.07 งานวิจัยยังนำเสนอต้นแบบแดชบอร์ดพร้อมคำอธิบายแบบ Rule-Based XAI เพื่อสนับสนุนการตัดสินใจ อย่างไรก็ตาม ผลลัพธ์ควรตีความภายใต้ข้อจำกัดของสภาพแวดล้อมจำลอง และควรตรวจสอบเพิ่มเติมด้วยข้อมูลจริงหรือการทดลอง A/B Test ก่อนใช้งานจริง*

**คำสำคัญ:** Contextual Bandit, LinUCB, ระบบแนะนำโปรโมชัน, การพยากรณ์อุปสงค์, กฎความสัมพันธ์, ปัญญาประดิษฐ์ที่อธิบายได้

## Abstract

*This research develops a prototype promotion recommendation system for B2B wholesale retail by integrating K-Means, LSTM, association rules, multi-branch deep learning (MB-DL), and the LinUCB contextual bandit algorithm, with a Profit Guard mechanism to mitigate negative-profit risk. Using historical transaction data, the system was evaluated over 10,000 simulation rounds, three market-scenario sensitivity tests, and an unseen-SKU holdout test. LinUCB achieved 21.01% higher cumulative reward than a random policy, with scenario-level improvements ranging from 21.79% to 23.69%, and 22.01% improvement on unseen SKUs at a generalization gap of only 1.07%. A prototype dashboard incorporating Rule-Based XAI explanations is also presented to support interpretable decision-making. Results are constrained to the simulation environment and should be validated with real-world data or A/B testing prior to deployment.*

**Keywords:** Contextual Bandit, LinUCB, Promotion Recommendation, Demand Forecasting, Association Rules, Explainable AI.

## 1. บทนำ

การบริหารโปรโมชันในธุรกิจค้าปลีก B2B มีความซับซ้อนสูง เนื่องจากอุปสงค์ไม่สม่ำเสมอ พฤติกรรมลูกค้าหลากหลาย และต้นทุนโปรโมชันที่ไม่เหมาะสมส่งผลโดยตรงต่อกำไรสุทธิ [1] งานวิจัยที่ผ่านมาส่วนใหญ่วิเคราะห์การพยากรณ์ยอดขาย การแบ่งกลุ่มสินค้า และการแนะนำโปรโมชันแยกจากกัน ยังขาดกรอบที่บูรณาการบริบทเชิงเวลา ความสัมพันธ์สินค้า และผลตอบแทนทางธุรกิจร่วมกันในการตัดสินใจภายใต้ความไม่แน่นอน

งานวิจัยนี้นำเสนอระบบแนะนำโปรโมชัน โดยบูรณาการ K-Means, LSTM, Association Rules, MB-DL และ LinUCB Contextual Bandit [8] เข้าเป็น Pipeline เดียว กำหนดหน่วยวิเคราะห์หลักเป็น สินค้า (SKU) และบริบทการซื้อ โดยใช้การจัดกลุ่มลูกค้า (K-Means) เป็นทั้ง Feature ใน MB-DL สำหรับทำนายรายได้ และบริบทประกอบการตัดสินใจของ LinUCB งานวิจัยมุ่งตอบคำถาม 3 ประการ ได้แก่ 1) Pipeline ที่เสนอสนับสนุนการเลือกโปรโมชันระดับ SKU ได้หรือไม่ 2) ระบบให้ผลตอบแทนสูงกว่านโยบายพื้นฐานเพียงใด และ 3) คำแนะนำอธิบายได้ในระดับที่ผู้ใช้ทางธุรกิจตรวจสอบได้หรือไม่ Contribution หลักประกอบด้วย 1) Pipeline แนะนำโปรโมชันแบบบูรณาการ 2) การใช้สัญญาณอุปสงค์และความสัมพันธ์สินค้าเป็น Context ของ LinUCB 3) กลไก Profit Guard คัดกรองโปรโมชันที่ขาดทุน และ 4) ต้นแบบ Dashboard พร้อม Rule-Based XAI

## 2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ผู้วิจัยได้ทำการรวบรวมข้อมูลเชิงทฤษฎีและผลงานวิจัยที่เกี่ยวข้อง เพื่อสร้างพื้นฐานความรู้และกรอบแนวคิดในการศึกษา โดยแบ่งเป็นหัวข้อหลัก ดังนี้

### 2.1 การวิเคราะห์ข้อมูลบริบทและกลุ่มลูกค้า

การวิจัยใช้กรอบ CRISP-DM [11] ในการสกัดรูปแบบพฤติกรรมการสั่งซื้อ โดยบูรณาการ RFM Analysis [13] วัดค่า Recency, Frequency และ Monetary เป็นข้อมูลนำเข้า K-Means คัดเลือกจำนวนกลุ่มด้วย Rank-Based Selection (Elbow, Silhouette, Davies-Bouldin, Calinski-Harabasz) ผลลัพธ์การจัดกลุ่มลูกค้า (K=4) ถูกนำไปใช้เป็น Contextual Features ใน LinUCB โดยงานวิจัยนี้กำหนดหน่วยวิเคราะห์หลักเป็น "สินค้าและบริบทการซื้อ" และใช้ข้อมูลกลุ่มลูกค้าเป็นคุณลักษณะประกอบการตัดสินใจเลือกโปรโมชัน

### 2.2 การพยากรณ์อุปสงค์และกฎความสัมพันธ์เชิงลึก

เพื่อสร้างองค์ประกอบสำคัญของระบบ ผู้วิจัยบูรณาการเทคนิคหลายวิธีเข้าด้วยกัน โดยใช้ LSTM [4] ซึ่งแก้ปัญหา Vanishing Gradient ด้วย Memory Cell และ Gate Mechanism ในการพยากรณ์อุปสงค์รายสัปดาห์ สำหรับสินค้า Intermittent Demand ใช้ SBA (Syntetos-Boylan Approximation) [3] แก้ปัญหา Bias จาก Croston's Method โดยค่า sba_scaled ที่ได้ถูกใช้เป็น Feature ใน MB-DL และส่งผลทางอ้อมต่อการตัดสินใจของ LinUCB ด้านการกำหนด Action Space ใช้ Association Rules [6] คัดกรองโปรโมชันที่มีนัยสำคัญทางสถิติด้วยค่า Support, Confidence และ Lift ร่วมกับ MB-DL (Multi-Branch Neural Network) ทำนายรายได้และจัดลำดับความสำคัญของสินค้าในแต่ละบริบท ก่อนส่งต่อชุดโปรโมชันให้ LinUCB ตัดสินใจขั้นสุดท้าย

### 2.3 การเรียนรู้แบบ Contextual Bandit และอัลกอริทึม LinUCB

งานวิจัยนี้เลือกใช้ LinUCB Disjoint [7] สำหรับการตัดสินใจเลือกโปรโมชัน เนื่องจากสามารถจัดการ Context Vector หลายมิติได้อย่างมีประสิทธิภาพ โดยตั้งสมมติฐานว่าผลตอบแทนที่คาดหวังมีความสัมพันธ์เชิงเส้นกับบริบท และคำนวณ Upper Confidence Bound เพื่อรักษาสมดุลระหว่าง Exploration และ Exploitation ระบบประกอบด้วย 4 องค์ประกอบ ได้แก่ 1) บริบทพฤติกรรมกลุ่มลูกค้าจาก K-Means และอุปสงค์จาก LSTM/SBA 2) การกระทำชุดโปรโมชันที่กำหนดโดย Association Rules และ MB-DL 3) ผลตอบแทนคำนวณจากกำไรสุทธิด้วย log(1+กำไร)/10 หรือ -0.5 หากขาดทุน (Profit Guard) และ (4) นโยบาย

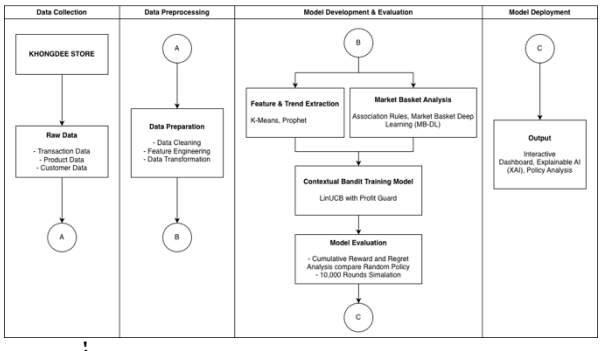
LinUCB เลือก Action เพื่อมุ่งหวัง Cumulative Reward สูงสุด

**2.4 งานวิจัยที่เกี่ยวข้อง**

Islam และ Rahman [4] พิสูจน์ว่า LSTM ให้ค่า WMAPE ต่ำกว่า Baseline อย่างมีนัยสำคัญในอุตสาหกรรม FMCG ส่วน Taylor และ Letham [5] นำเสนอ Prophet ซึ่งงานวิจัยนี้นำมาใช้ในขั้นสำรวจเบื้องต้นก่อนเลือก LSTM เป็นโมเดลหลัก Li et al. [7] เสนอ LinUCB ที่เพิ่ม Click Lift ได้ 12.5% บนชุดข้อมูล Yahoo! ยืนยันประสิทธิภาพในสภาวะข้อมูลจำกัด สิทธิพันธ์ [13] ประยุกต์ RFM + K-Means ภายใต้ CRISP-DM จัดกลุ่มลูกค้า 12,641 ราย ได้ Silhouette Score 0.405 และ Bui et al. [17] แสดงให้เห็นว่า XAI ช่วยเพิ่มความโปร่งใสในการตัดสินใจเชิงกลยุทธ์

## 3. ระเบียบวิธีวิจัย

งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาต้นแบบระบบแนะนำโปรโมชันอัตโนมัติ โดยมีกรอบแนวคิดการดำเนินงานที่ครอบคลุมตั้งแต่การจัดการข้อมูลจนถึงการนำไปใช้งานจริง ดังแสดงในภาพที่ 1




**ภาพที่ 1:** กรอบแนวคิดและสถาปัตยกรรมการดำเนินงานวิจัย

จากภาพที่ 1 ขั้นตอนการดำเนินงานถูกแบ่งออกเป็น 4 ส่วนหลักที่มีการเชื่อมต่อข้อมูลแบบพลวัต การทดลองพัฒนาด้วย Python โดยใช้ scikit-learn, TensorFlow/Keras, PyTorch และ mlxtend เป็นไลบรารีหลักในการพัฒนาโมเดล และ FastAPI/Next.js สำหรับต้นแบบระบบสนับสนุนการตัดสินใจ

**3.1 การเก็บรวบรวม และเตรียมข้อมูล**

ข้อมูลธุรกรรมจากระบบ KhongDee Store ช่วงมกราคม 2566 – พฤษภาคม 2567 จำนวน 26,519 รายการ หลังทำความสะอาดและตัดรายการที่ไม่สมบูรณ์/ยกเลิกออกได้ 25,704 รายการ ครอบคลุม 1,399 SKU โดยใช้ 1,259 SKU สำหรับสร้าง Action Space และวิเคราะห์ Association Rules สำหรับการพยากรณ์ใช้ Non-Lumpy SKUs (ADI/CV²) จำนวน 273 รายการสำหรับฝึก LSTM ส่วน SKU ที่เหลือใช้ SBA แทนสัญญาณอุปสงค์

**3.2 การสกัดคุณลักษณะและมิติข้อมูลบริบท**

ผู้วิจัยดำเนินการเปลี่ยนข้อมูลให้เป็นมิติข้อมูลเชิงบริบทผ่าน 3 กระบวนการหลัก ดังนี้

**3.2.1 การคัดกรองสินค้าและทำความสะอาดข้อมูล** ผู้วิจัยคัดเลือก Top SKU โดยใช้เกณฑ์รูปแบบอุปสงค์สม่ำเสมอ พิจารณาจากดัชนี ADI < 1.32 และ $CV^2$ < 0.49 พร้อมเงื่อนไขประวัติการขายไม่น้อยกว่า 15 สัปดาห์ เพื่อให้มั่นใจว่าแบบจำลอง LSTM และ LinUCB สามารถเรียนรู้จากสัญญาณข้อมูลที่มีความต่อเนื่องและลดปัญหาความเบาบางของข้อมูล สำหรับสินค้าที่มีอุปสงค์ไม่สม่ำเสมอ ผู้วิจัยนำวิธี SBA [18] มาใช้คำนวณค่าประมาณอุปสงค์พื้นฐาน

**3.2.2 การสร้างบริบท** ข้อมูลจะถูกสกัดเป็นเวกเตอร์ขนาด 8 มิติ และปรับมาตราส่วนให้อยู่ในช่วง 0.0 ถึง 1.0 เพื่อเพิ่มความเร็วการลู่เข้าของอัลกอริทึม LinUCB และรักษาสมดุลตัวแปร ดังรายละเอียดในตารางที่ 1

**ตารางที่ 1 :** รายละเอียดคุณลักษณะข้อมูลบริบท

| มิติข้อมูล | ความหมายและเทคนิค | การปรับมาตราส่วน |
|---|---|---|
| Week | สัปดาห์ในรอบปี | Min-Max |
| Month | เดือนในรอบปี | Min-Max |
| basket_size | จำนวนรายการในตะกร้า | log1p |
| cluster | กลุ่มลูกค้า | Label |
| lstm_forecast | ค่าพยากรณ์ | Min-Max |
| sba_scaled | ค่า SBA | Min-Max |
| product_idx | รหัสสินค้า | Label Encodin |
| mbdl_score | คะแนน MB-DL × Lift | Min-Max |

**3.2.3 การกำหนดชุดการกระทำด้วยเทคนิคไฮบริด** วิจัยบูรณาการ Association Rules (เกณฑ์ Lift > 1.0) เพื่อคัดกรองคู่สินค้าเบื้องต้น และใช้ MB-DL สถาปัตยกรรม Multi-Branch Neural Network ในการ

จัดลำดับความสำคัญ เพื่อสร้างชุดการกระทำ (Action Space) ที่แม่นยำสูงกว่ากฎสถิติดั้งเดิม ชุดการกระทำทั้ง 4 รูปแบบ ได้แก่ 1) ไม่มอบโปรโมชัน (No Promotion) 2) Discount 10% 3) Free Gift มูลค่าของแถม 100 บาท และ 4) Bundle Deal ข้อมูลถูกแบ่งตามลำดับเวลาเป็นชุดฝึก ชุดตรวจสอบ และชุดทดสอบ เพื่อหลีกเลี่ยงการรั่วไหลของข้อมูลจากอนาคตไปยังอดีต

**3.3 การพัฒนาตัวแทนเรียนรู้และสภาพแวดล้อมจำลอง**

ผู้วิจัยพัฒนาตัวแทนเรียนรู้โดยใช้อัลกอริทึม LinUCB เพื่อตัดสินใจเลือกโปรโมชันที่เหมาะสมที่สุดจาก 4 รูปแบบการกระทำ โดยมีรายละเอียดการดำเนินงานดังนี้

**3.3.1 การสร้างสภาพแวดล้อมจำลอง** ผู้วิจัยได้พัฒนาสภาพแวดล้อมเพื่อฝึกฝนตัวแทนเรียนรู้ (LinUCB) โดยประยุกต์ใช้แบบจำลอง MB-DL ที่ผ่านการเรียนรู้พฤติกรรมในอดีต มาทำหน้าที่เป็นกลไกจำลองการตอบสนองของลูกค้า เพื่อประเมินความน่าจะเป็นในการตอบรับโปรโมชันและคำนวณค่าตอบแทนกลับไปยังตัวแทนเรียนรู้

**3.3.2 การกำหนดสถานการณ์จำลอง**เพื่อทดสอบความทนทานของอัลกอริทึม ผู้วิจัยกำหนดสถานการณ์ทดสอบ 3 รูปแบบ โดยปรับพารามิเตอร์ที่ส่งผลต่อพฤติกรรมตลาดและต้นทุน ดังรายละเอียดในตารางที่ 2

**ตารางที่ 2 :** สภาวะจำลองและตัวแปรควบคุมทางธุรกิจ

| สถานการณ์ | ความหมายเชิงธุรกิจ | ตัวแปรควบคุม |
|---|---|---|
| Aggressive | ตลาดตอบสนองสูง | Boost 1.4–1.8, Cost ต่ำ |
| Conservative | ตลาดอิ่มตัว/กำไรขั้นต้นต่ำ | Boost 1.05–1.2, Cost Pct สูง (0.5) |
| Balanced | สภาวะตลาดทั่วไประยะยาว | Boost 1.15–1.5, Cost Fixed 50–100 |

**3.3.3 การปรับจูนพารามิเตอร์** ผู้วิจัยกำหนดค่าพารามิเตอร์เพื่อประสิทธิภาพสูงสุดของระบบ โดยคัดเลือกผ่านดัชนีชี้วัด WMAPE และ Silhouette Score ค่าพารามิเตอร์หลักที่ใช้ ได้แก่ K-Means K=4, LSTM Nonzero Loss Weight=3.0, MB-DL Dropout=0.2 และ Embedding Dimension=32, LinUCB $\alpha$=1.0 และ Profit Guard Penalty=-50

ผู้วิจัยกำหนดการฝึกฝนตัวแทนเรียนรู้รวม 18,000 รอบ แบ่งเป็นการจำลองหลัก 10,000 รอบ การทดสอบความทนทาน 3 สถานการณ์ (สถานการณ์ละ 2,000 รอบ รวม 6,000 รอบ) และการทดสอบ Unseen SKU อีก 2,000 รอบ โดยบูรณาการกลไก Profit Guard เพื่อทำหน้าที่เป็นชั้นกรองการตัดสินใจ ป้องกันการเลือกกลยุทธ์ที่ส่งผลให้กำไรสุทธิติดลบโดยเฉพาะในสภาวะตลาดที่มีต้นทุนการดำเนินงานสูง

**3.4 การประเมินผลและการนำต้นแบบไปใช้งาน**

ประสิทธิภาพของ LinUCB ประเมินเปรียบเทียบกับ Baseline หลายรูปแบบ ได้แก่ Random Policy, Greedy, $\varepsilon$-Greedy, UCB1 และ Thompson Sampling ผ่านดัชนี Cumulative Reward และ Cumulative Regret เพื่อวิเคราะห์ความเร็วในการเรียนรู้และความสามารถในการ Generalize กับข้อมูลใหม่ นอกจากนี้ได้พัฒนาต้นแบบ Dashboard ที่ผสาน Rule-Based XAI อธิบายเหตุผลประกอบคำแนะนำโปรโมชันผ่าน 3 ปัจจัย ได้แก่ กลุ่มลูกค้า (K-Means Cluster), ความเชื่อมโยงสินค้า (Association Lift) และกำไรคาดการณ์ (MB-DL Score)

## 4. ผลการทดลองและการประเมินประสิทธิภาพ

การนำเสนอผลการดำเนินงานวิจัยแบ่งออกเป็นส่วนหลัก โดยผลในหัวข้อ 4.1–4.3 มุ่งตอบวัตถุประสงค์ด้านการพัฒนาแบบจำลองและโมเดลแนะนำโปรโมชัน ส่วนหัวข้อ 4.4 มุ่งตอบวัตถุประสงค์ด้านการนำเสนอผลเพื่อสนับสนุนการตัดสินใจ ดังนี้

**4.1 ภาพรวมผลการดำเนินงาน**

งานวิจัยนี้พัฒนาระบบแนะนำโปรโมชันตามกรอบ CRISP-DM โดยบูรณาการ 5 องค์ประกอบหลัก ได้แก่ K-Means Clustering, LSTM, Apriori, MB-DL และ LinUCB Contextual Bandit ข้อมูลที่ใช้ครอบคลุมธุรกรรมการขายสินค้าปลีก โดยกรองเฉพาะ 273 SKU ที่มีรูปแบบความต้องการแบบ Non-Lumpy ตาม ADI/$CV^2$ จากนั้นแบ่งกลุ่มลูกค้าด้วย K-Means (K=4) และพยากรณ์ยอดขายด้วย LSTM Group C (12 features, Walk-Forward CV 3 folds)

ส่วน Apriori ใช้สกัดกฎความสัมพันธ์เพื่อสนับสนุนการแนะนำ Bundle Deal โดยผลลัพธ์ของแต่ละองค์ประกอบถูกส่งต่อเป็น Context Vector ให้กับ LinUCB ($\alpha$=0.1, N=8 features) เพื่อเลือกโปรโมชันที่เหมาะสมที่สุดในแต่ละรอบ ผลการดำเนินงานทั้งหมดสรุปไว้ในตารางที่ 3

**ตารางที่ 3 :** สรุปผลการดำเนินงานของระบบแนะนำโปรโมชัน

| องค์ประกอบ | วิธีการ | ผลลัพธ์หลัก |
|---|---|---|
| การจัดกลุ่มลูกค้า | K-Means (K=4) | Silhouette Score = 0.312 |
| การพยากรณ์ยอดขาย | LSTM Group C (12 features) | WMAPE = 83.86%, $R^2$ = 0.1116 |
| กฎความสัมพันธ์ | Apriori | 287 rules (Lift $\geq$ 1.5) |
| การทำนายรายได้ | MB-DL | $R^2$ = 0.6625, MAE = 619.44 THB |
| ระบบแนะนำโปรโมชัน | LinUCB ($\alpha$=1.0) | +21.01% vs. Random Policy |
| Generalization Gap | Unseen SKU | -1.07% |

### 4.2 ผลการจัดกลุ่มพฤติกรรมลูกค้าและพยากรณ์ยอดขาย

การคัดเลือก K ใช้ Rank-Based Selection รวม 4 เกณฑ์พร้อมกัน (Elbow, Silhouette, Davies-Bouldin, Calinski-Harabasz) พบว่า K=4 ถูกเลือกเนื่องจากให้ผลรวมอันดับดีที่สุดจาก Elbow, Silhouette, Davies-Bouldin และ Calinski-Harabasz อย่างไรก็ตาม ค่า Silhouette Score = 0.312 อยู่ในระดับปานกลาง และนำ Cluster ไปใช้เป็น Feature ใน MB-DL สำหรับ LSTM ทำการ Ablation Study โดยกรองเฉพาะ Non-Lumpy SKUs (Syntetos & Boylan: ADI < 1.32, $CV^2$ < 0.49) ได้ 273 SKUs แล้วเปรียบเทียบ 4 Feature Groups ด้วย Grid Search (216 configs) + Walk-Forward CV (3 folds) ดังตารางที่ 4

**ตารางที่ 4 :** ผลการเปรียบเทียบกลุ่ม Feature กับ Baseline

| กลุ่ม | Features | WMAPE (%) | $R^2$ |
|---|---|---|---|
| A | Lag ×4 | 87.11 | 0.0573 |
| B | + Rolling Stats ×4 | 86.98 | 0.0661 |
| C | + Stockout/Restock ×4 | 86.97 | 0.0649 |
| D | + Temporal ×4 | 87.60 | 0.0378 |
| Naïve Baseline | - | 113.88 | - |

แม้ LSTM Group C ให้ผลดีกว่า Naïve Baseline แต่ค่า WMAPE = 83.86% และ $R^2$ = 0.1116 ยังสะท้อนข้อจำกัดด้านความแม่นยำของการพยากรณ์ยอดขาย จึงควรใช้ผลพยากรณ์เป็นเพียงสัญญาณบริบทเสริม ไม่ใช่ตัวชี้ขาดการตัดสินใจ

### 4.3 ผลการทำนายรายได้ (MB-DL) และกฎความสัมพันธ์

จากผลงานวิจัยสร้างได้ 287 rules (Lift เฉลี่ย 2.34, สูงสุด 8.71) ซึ่งนำไปกำหนด Bundle Deal Action Space และ Feature assoc_lift สำหรับ XAI MB-DL ฝึกบนสินค้า 1,259 รายการ (Revenue $\geq$ 5,000 THB, 1 ปี) ด้วย 5 Features ได้แก่ product_name_idx, Cluster, week_of_year, month, sba_scaled ผลการประเมินได้ $R^2$=0.6625 และ MAE=619.44 THB ($R^2$_log=0.5788, RMSE_log=0.8591) แสดงว่าโมเดลอธิบายความแปรปรวนของรายได้ได้ 66.25% ซึ่งอยู่ในระดับดีสำหรับข้อมูลค้าปลีก

### 4.4 ผลการทดสอบระบบแนะนำโปรโมชัน (LinUCB)

LinUCB Disjoint ใช้ $\alpha$=1.0, Context Vector 8 มิติ (Cluster, week_of_year, month, sba_scaled, mbdl_score, assoc_lift, price_tier, promo_response) และ Action Space 4 ตัวเลือก ได้แก่ Discount 10%, Free Gift, Bundle Deal, No Promotion พร้อม Profit Guard (penalty=-50 หาก profit<0) Simulation 10,000 รอบ และ Sensitivity Analysis รวม 6,000 รอบ (3 สถานการณ์) แสดงผลในตารางที่ 5

**ตารางที่ 5 :** ผลการ Simulation หลักและ Sensitivity Analysis

| สถานการณ์ | รอบ | LinUCB vs. Random | Bundle Deal (%) |
|---|---|---|---|
| Simulation | 10,000 | +21.01% | 83.18% |
| Aggressive | 2,000 | +21.79% | 94.85% |
| Balanced | 2,000 | +23.69% | 94.25% |
| Conservative | 2,000 | +22.18% | 94.65% |
| Unseen SKU | 2,000 | +22.01% | 82.54% |

การ Sensitivity Analysis แสดงให้เห็นว่าค่าการปรับปรุงเทียบ Random Policy อยู่ในช่วง +21.01%–+23.69% ในทุกสถานการณ์ที่ทดสอบ โดยการทดลองซ้ำด้วย 5 Random Seeds (42, 123, 456, 789, 1000) × 10,000 รอบ พบค่าเฉลี่ย 21.20% ± 0.23% (95% CI: [20.92%, 21.49%]) Cumulative Regret เฉลี่ย 116.74 ± 2.04 (95% CI: [114.21, 119.27]) ซึ่งบ่งชี้ว่าผลลัพธ์มีความเสถียรข้ามการทดลองในระดับสูง ความแปรปรวนที่ต่ำ (std < 0.24%) สะท้อนให้เห็นว่าประสิทธิภาพของ LinUCB ไม่ขึ้นอยู่กับ Random Seed อย่างมีนัยสำคัญ และกลยุทธ์ Bundle Deal ถูกเลือกในสัดส่วนสูงสุดในสภาพแวดล้อมจำลองที่กำหนด (83–95%) ซึ่งอาจสะท้อนศักยภาพของการรวมสินค้าในบริบทข้อมูลชุดนี้ อย่างไรก็ตาม ความเป็น Dominant ของ Bundle Deal อาจเกิดจาก Reward Bias ในการออกแบบ Reward Function ซึ่งควรศึกษาเพิ่มเติม สำหรับ Generalization Gap ที่ได้ -1.07% (Seen=23.08%, Unseen=22.01%) นั้น ชี้ให้เห็นว่าประสิทธิภาพของระบบบน Unseen SKUs ใกล้เคียงกับ Seen SKUs ซึ่งอาจเป็นผลจากการใช้ Cluster-based Context

**4.5 ผลการพัฒนาต้นแบบระบบ (Dashboard)**

ต้นแบบ Dashboard รองรับการเลือก SKU การแสดงกลุ่มสินค้า/ลูกค้า การแนะนำโปรโมชันด้วย LinUCB การทำนายรายได้ด้วย MB-DL และคำอธิบายแบบ Rule-Based XAI โดยแสดงเหตุผลจาก Cluster, Association Lift และ MB-DL Score ทั้งนี้ยังต้องประเมินกับผู้ใช้จริงก่อนใช้งานเชิงปฏิบัติ

## 5. สรุป

งานวิจัยนี้พัฒนาระบบแนะนำโปรโมชัน โดยบูรณาการ K-Means, LSTM, Association Rules, MB-DL และ LinUCB ภายใต้กรอบ CRISP-DM ผลการดำเนินงานพบว่า K-Means (K=4) ให้ Silhouette Score=0.312 LSTM Group C ให้ WMAPE=83.86% ดีกว่า Naïve Baseline (113.88%) และผ่าน Wilcoxon Test ($p<0.0001$) MB-DL ให้ $R^2$=0.6625 และ MAE=619.44 บาท สำหรับ LinUCB ($\alpha$=1.0) ให้ผลตอบแทนสูงกว่า Random Policy +21.01% ในการจำลองหลัก และ +21.35%±0.58% (95% CI: [20.63%, 22.07%]) จากการทดลอง 5 Random Seeds Generalization Gap อยู่ที่ -1.07% (Seen=23.08%, Unseen=22.01%) ชี้ให้เห็นว่าระบบสามารถ Generalize ไปยัง Unseen SKUs ได้ในเบื้องต้น

ผลการทดลองควรตีความภายใต้ข้อจำกัดสำคัญ 4 ประการ ได้แก่ 1) MB-DL ทำหน้าที่เป็นกลไกจำลองการตอบสนอง ผล LinUCB จึงสะท้อนประสิทธิภาพภายใต้แบบจำลองที่มี $R^2$=0.6625 ไม่ใช่พฤติกรรมจริงโดยตรง 2) Bundle Deal ถูกเลือกสูงถึง 83–95% ในทุกสถานการณ์ ซึ่งอาจสะท้อน Reward Bias ในการออกแบบ Reward Function มากกว่าศักยภาพจริงของกลยุทธ์ 3) WMAPE ของ LSTM ยังสูง สัญญาณอุปสงค์จึงควรใช้เป็น Feature เสริม ไม่ใช่ตัวกำหนดการตัดสินใจหลัก และ 4) การเปรียบเทียบกับ Random Policy เพียงอย่างเดียวเป็น Baseline ที่อ่อน รวมถึงข้อมูลมาจากกรณีศึกษาเดียวในช่วง 1 ปี จึงอาจไม่สะท้อนความหลากหลายของธุรกิจค้าส่งทั้งหมด

งานวิจัยในอนาคตควรเพิ่ม Baseline ที่เข้มขึ้น เช่น Greedy, $\varepsilon$-Greedy, UCB1 และ Thompson Sampling เพื่อประเมินประสิทธิภาพ LinUCB อย่างรอบด้าน ควรทดสอบ Reward Function หลายรูปแบบเพื่อลด Action Imbalance และประเมิน XAI Explanation กับผู้ใช้จริง เช่น ผู้จัดการร้านหรือนักวิเคราะห์การตลาด ก่อนนำไปใช้งานจริงควรเริ่มจาก A/B Testing ขนาดเล็กในกลุ่มสินค้าเฉพาะ และตรวจสอบผลกระทบต่อยอดขาย กำไร และต้นทุนโปรโมชันด้วยข้อมูลภาคสนาม

## เอกสารอ้างอิง


[1] J. Han, J. Pei, and H. Tong, *Data Mining: Concepts and Techniques*, 4th ed. Morgan Kaufmann, 2022.

[2] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997. doi: 10.1162/neco.1997.9.8.1735

[3] A. A. Syntetos and J. E. Boylan, "The accuracy of intermittent demand estimates," Int. J. Forecasting, vol. 21, no. 2, pp. 303–314, 2005. doi: 10.1016/j.ijforecast.2004.10.001

[4] M. R. Islam and M. A. Rahman, "Implementing LSTM networks for sales forecasting and predictive modelling of consumer demand in the fast-moving consumer goods industry," ResearchGate Preprint, 2025.

[5] S. J. Taylor and B. Letham, "Forecasting at scale," The American Statistician, vol. 72, no. 1, pp. 37–45, 2018. doi: 10.1080/00031305.2017.1380080

[6] R. Agrawal, T. Imielinski, and A. Swami, "Mining association rules between sets of items in large databases," in Proc. ACM SIGMOD Int. Conf. Management of Data, Washington, DC, USA, 1993, pp. 207–216.

[7] L. Li, W. Chu, J. Langford, and R. E. Schapire, "A contextual-bandit approach to personalized news article recommendation," in Proc. 19th Int. Conf. World Wide Web (WWW), Raleigh, NC, USA, 2010, pp. 661–670. doi: 10.1145/1772690.1772758 https://dl.acm.org/doi/10.1145/1772690.1772758

[8] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016. [Online]. Available: https://www.deeplearningbook.org/

[9] R. S. Sutton and A. G. Barto, Reinforcement Learning: An Introduction, 2nd ed. MIT Press, 2018. [Online]. Available: http://incompleteideas.net/book/the-book-2nd.html

[10] A. B. Arrieta et al., "Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges," Inf. Fusion, vol. 58, pp. 82–115, 2020. doi: 10.1016/j.inffus.2019.12.012 https://doi.org/10.1016/j.inffus.2019.12.012

[11] R. Wirth and J. Hipp, "CRISP-DM: Towards a standard process model for data mining," in Proc. 4th Int. Conf. Practical Applications of Knowledge Discovery and Data Mining (PAKDD), 2000, pp. 29–40.

[12] R. J. Hyndman and G. Athanasopoulos, Forecasting: Principles and Practice, 2nd ed. OTexts, 2018. [Online]. Available: https://otexts.com/fpp2/

[13] สิทธิพันธ์ สำเร็จประสงค์, "*วิเคราะห์การแบ่งกลุ่มลูกค้าด้วยตัวแปร อาร์เอฟเอ็มและอัลกอริทึมเคมีนส์: กรณีศึกษาในธุรกิจค้าปลีก อิเล็กทรอนิกส์สินค้าสำหรับสตรีมีครรภ์*," วิทยานิพนธ์ปริญญาโท, มหาวิทยาลัยธรรมศาสตร์, 2566.

[14] วรปรัชญ์ อภิอารยาวงศ์, "แนวทางการพัฒนาต้นแบบเว็บไซต์เพื่อสนับสนุนการขายโครงการบ้านเดี่ยวสำหรับคนเจนเนอเรชันแซด โดยใช้เทคนิคเหมืองข้อมูลและปัญญาประดิษฐ์," วิทยานิพนธ์ปริญญาโท, มหาวิทยาลัยธรรมศาสตร์, 2565.

[15] ปนัดดา โตเจริญนิรัติศัย, "การศึกษาพฤติกรรมการสั่งซื้อวัตถุดิบอาหารและเครื่องดื่มโดยเทคนิคอาร์เอฟเอ็ม และประเมินการเก็บรักษาวัตถุดิบตามหลักมาตรฐานสินค้าเกษตร (มกษ. 9023-2550)," วิทยานิพนธ์ปริญญาโท, มหาวิทยาลัยธรรมศาสตร์, 2565.

[16] ชาญชวัตม์ ภักดีศรี, "การพัฒนากลยุทธ์ทางการตลาดโดยใช้วิทยาศาสตร์ข้อมูลในการจัดกลุ่มลูกค้าตามพฤติกรรมการซื้อสินค้า," วิทยานิพนธ์ปริญญาโท, มหาวิทยาลัยธรรมศาสตร์, 2565.

[17] N. V. P. Bui, "Explainable AI (XAI): Improving at-risk student prediction with theory-guided data science," Ph.D. dissertation, Northcentral Univ., 2022.

# การพัฒนาแชตบอตเพื่อส่งเสริมความตระหนักรู้ด้านความมั่นคงปลอดภัยไซเบอร์
# Development of a Chatbot to Promote Cybersecurity Awareness

*วิไลวรรณ พันธุระ (Wilaiwan Phanthura)[1] และ นวพร วิสิฐพงศ์พันธ์ (Nawaporn Wisitpongphan)[2]*
*ภาควิชาการบริหารเครือข่ายดิจิทัลและความมั่นคงปลอดภัยสารสนเทศ*
*คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ*
*s6707031857174@kmutnb.ac.th, nawaporn.w@itd.kmutnb.ac.th*

## บทคัดย่อ

*ในยุคดิจิทัลที่ภัยคุกคามทางไซเบอร์มีความซับซ้อนและเพิ่มขึ้นอย่างต่อเนื่อง พฤติกรรมของผู้ใช้งานถือเป็นปัจจัยสำคัญที่ส่งผลต่อความมั่นคงปลอดภัยขององค์กร งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาแชตบอตเพื่อส่งเสริมความตระหนักรู้ด้านความมั่นคงปลอดภัยไซเบอร์ และประเมินประสิทธิภาพของระบบรวมถึงการเปลี่ยนแปลงระดับความรู้ของผู้ใช้งาน โดยแชตบอตถูกออกแบบให้สามารถสื่อสารกับผู้ใช้งานผ่านแพลตฟอร์ม LINE ใช้ Dialogflow สำหรับการประมวลผลภาษาธรรมชาติ และเชื่อมต่อกับแหล่งข้อมูลภายนอกเพื่อให้ความรู้เกี่ยวกับภัยคุกคามไซเบอร์ เช่น ฟิชชิง การตั้งรหัสผ่านที่ปลอดภัย และวิศวกรรมสังคม รวมถึงสามารถตรวจสอบข้อมูลด้านความปลอดภัย เช่น IP Address, URL และ Domain กลุ่มตัวอย่างประกอบด้วยบุคลากรจำนวน 30 คน โดยใช้แบบสอบถามในการวัดระดับความรู้ก่อนและหลังการใช้งาน ผลการวิจัยพบว่าคะแนนเฉลี่ยหลังการใช้งาน ($\bar{x}$ = 14.03, SD = 1.37) สูงกว่าก่อนการใช้งาน ($\bar{x}$ = 13.47, SD = 2.21) อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 ผู้ใช้งานมีความพึงพอใจในระดับสูง โดยเฉพาะด้านความถูกต้องของข้อมูลและความสะดวกในการใช้งาน ผลการวิจัยแสดงให้เห็นว่าแชตบอตสามารถช่วยเพิ่มความตระหนักรู้ด้านความมั่นคงปลอดภัยไซเบอร์ได้อย่างมีประสิทธิภาพ และสามารถนำไปประยุกต์ใช้ในองค์กรได้จริง*

**คำสำคัญ:** แชตบอต, ความมั่นคงปลอดภัยไซเบอร์, ความตระหนักรู้, ฟิชชิง, Social Engineering

## Abstract

*In the digital era, where cyber threats have become increasingly complex and prevalent, user behavior has emerged as a critical factor affecting organizational security. This research aims to develop a chatbot to enhance cybersecurity awareness and evaluate system performance, including changes in users' knowledge levels. The chatbot was designed to interact with users through the LINE platform, utilizing Dialogflow for natural language processing and integrating external data sources to provide knowledge on cyber threats such as phishing, secure password practices, and social engineering. Additionally, the system can analyze security-related information such as IP addresses, URLs, and domains. The sample group consisted of 30 participants. The results showed that the post-test mean score ($\bar{x}$ = 14.03, SD = 1.37) was higher than the pre-test score ($\bar{x}$ = 13.47, SD = 2.21), with statistical significance at the 0.05 level. Users reported a high level of satisfaction, particularly in terms of information accuracy and ease of use. These findings indicate that the chatbot effectively improves cybersecurity awareness and has strong potential for practical implementation in organizations.*

**Keywords:** Chatbot, Cybersecurity, Awareness, Phishing, Social Engineering.

## 1. บทนำ

ในปัจจุบันภัยคุกคามทางไซเบอร์มีแนวโน้มเพิ่มขึ้นอย่างต่อเนื่องและมีความซับซ้อนมากขึ้น ส่งผลให้องค์กร

ต้องให้ความสำคัญกับการสร้างความตระหนักรู้ด้านความมั่นคงปลอดภัยแก่บุคลากร อย่างไรก็ตาม วิธีการแบบดั้งเดิม เช่น การอบรมหรือการส่งอีเมลแจ้งเตือน มักมีข้อจำกัดด้านความต่อเนื่องและการเข้าถึงข้อมูล การนำแชตบอตมาประยุกต์ใช้สามารถช่วยแก้ปัญหาดังกล่าวได้ โดยทำหน้าที่เป็นเครื่องมืออัจฉริยะที่ช่วยให้ผู้ใช้งานสามารถเข้าถึงข้อมูลด้านความปลอดภัยได้อย่างรวดเร็วและสะดวกมากขึ้น รวมถึงสามารถตรวจสอบข้อมูลที่เกี่ยวข้องกับภัยคุกคาม เช่น IP Address, URL และ Domain ได้ทันที

ดังนั้น งานวิจัยนี้จึงมีวัตถุประสงค์เพื่อพัฒนาแชตบอตสำหรับส่งเสริมความตระหนักรู้ด้านความมั่นคงปลอดภัยไซเบอร์ และประเมินประสิทธิภาพของระบบจากการใช้งานจริงของผู้ใช้งาน

งานวิจัยนี้มีวัตถุประสงค์ 3 ประการ ได้แก่ (1) เพื่อพัฒนาแชตบอตที่สามารถให้ความรู้และข้อมูลด้านความมั่นคงปลอดภัยไซเบอร์แก่ผู้ใช้งาน โดยสามารถโต้ตอบได้อย่างเป็นธรรมชาติผ่านเทคโนโลยีการประมวลผลภาษาธรรมชาติ (Natural Language Processing: NLP) เพื่อให้ผู้ใช้งานสามารถเข้าถึงข้อมูลได้อย่างสะดวกและมีประสิทธิภาพ (2) เพื่อพัฒนาแชตบอตที่สามารถตรวจสอบและวิเคราะห์ข้อมูลด้านความปลอดภัยไซเบอร์ เช่น IP Address, URL และ Domain โดยเชื่อมต่อกับแหล่งข้อมูลหรือ API ภายนอกที่เกี่ยวข้อง เพื่อช่วยให้ผู้ใช้งานสามารถประเมินความเสี่ยงได้อย่างรวดเร็วและแม่นยำ และ (3) เพื่อประเมินประสิทธิภาพของแชตบอตและศึกษาการเปลี่ยนแปลงของระดับความตระหนักรู้ด้านความมั่นคงปลอดภัยไซเบอร์ของผู้ใช้งาน โดยเปรียบเทียบผลก่อนและหลังการใช้งานระบบ

## 2. งานวิจัยที่เกี่ยวข้อง

งานวิจัยที่เกี่ยวข้องในด้านการใช้แชตบอทเพื่อส่งเสริมความตระหนักรู้ด้านความมั่นคงปลอดภัยไซเบอร์ พบว่าแชตบอทมีบทบาทสำคัญในการพัฒนาพฤติกรรมและการเรียนรู้ของผู้ใช้งาน โดย Basri และ Ali [1] ระบุว่าแชตบอทสามารถช่วยเพิ่มความตระหนักรู้และส่งผลต่อพฤติกรรมด้านความปลอดภัยของผู้ใช้งานได้อย่างมีนัยสำคัญ ขณะที่ Abdelhamid et al. [2] พบว่าแชตบอทสามารถใช้เป็นเครื่องมือฝึกอบรมในองค์กรได้อย่างมีประสิทธิภาพ โดยเฉพาะในรูปแบบการเรียนรู้แบบโต้ตอบและการเข้าถึงข้อมูลอย่างต่อเนื่อง

ในด้านการเสริมสร้างความตระหนักรู้ งานวิจัยของ Chavalitlekha [3] และ Chatchalermpun และ Daengsi [4] แสดงให้เห็นว่าการใช้สถานการณ์จำลอง เช่น การจำลองอีเมลฟิชชิง สามารถช่วยลดความเสี่ยงจากการโจมตีทางไซเบอร์ได้อย่างมีนัยสำคัญ ซึ่งสอดคล้องกับงานของ Roonras [7] ที่ชี้ให้เห็นว่าปัจจัยด้านความรู้และพฤติกรรมของผู้ใช้งานมีผลโดยตรงต่อระดับความตระหนักรู้ด้านความมั่นคงปลอดภัยไซเบอร์

นอกจากนี้ งานวิจัยด้านปัญญาประดิษฐ์ เช่น Eakacet [5] และ Reddy [11] ระบุว่าการนำ AI และ Deep Learning มาใช้สามารถช่วยเพิ่มประสิทธิภาพในการตรวจจับ วิเคราะห์ และป้องกันภัยคุกคาม โดยเฉพาะการโจมตีแบบฟิชชิง ขณะที่ Microsoft Threat Intelligence [14] ยังแสดงให้เห็นถึงแนวโน้มของภัยคุกคามที่มีความซับซ้อนมากขึ้นจากการใช้ AI ซึ่งจำเป็นต้องใช้เครื่องมืออัจฉริยะในการรับมือ

ในส่วนของเทคโนโลยีแชตบอท งานของ Fung และ Lee [8] และ Hamad และ Yeferny [9] อธิบายว่าแชตบอทสามารถทำหน้าที่เป็นเครื่องมือให้คำแนะนำด้านความปลอดภัยแบบเรียลไทม์ และช่วยให้ผู้ใช้งานเข้าถึงข้อมูลได้สะดวกมากขึ้น ขณะที่ Pimentel et al. [10] และ Keepnet Labs [13] พบว่าแชตบอทสามารถเพิ่มการมีส่วนร่วมของผู้ใช้งานและส่งเสริมการเรียนรู้ด้าน Cybersecurity Awareness ได้อย่างมีประสิทธิภาพ

งานวิจัยของ He และ Xin [12] ยังแสดงให้เห็นว่าแชตบอทสามารถนำมาใช้ในบริบทการศึกษาเพื่อสนับสนุนการเรียนรู้ด้านความมั่นคงปลอดภัยไซเบอร์ โดยใช้แนวคิดการเรียนรู้แบบโต้ตอบ ซึ่งช่วยให้ผู้ใช้งานเข้าใจเนื้อหาได้ดีขึ้น ขณะที่ Lim [17] และ Smith [16] ชี้ให้เห็นถึงความสำคัญของการฝึกอบรมด้านความตระหนักรู้ในองค์กร เพื่อให้สามารถรับมือกับภัยคุกคามยุคใหม่ได้อย่างมีประสิทธิภาพ

อย่างไรก็ตาม แม้แชตบอทจะมีศักยภาพสูงในการส่งเสริมการเรียนรู้ งานวิจัยของ Kooli [15] ได้กล่าวถึงประเด็นด้านจริยธรรมและความท้าทายในการใช้งาน เช่น ความน่าเชื่อถือของข้อมูล และผลกระทบจากการพึ่งพา AI มากเกินไป

จากการทบทวนงานวิจัยดังกล่าว สามารถสรุปได้ว่าแชตบอทเป็นเครื่องมือที่มีศักยภาพในการส่งเสริมความตระหนักรู้ด้านความมั่นคงปลอดภัยไซเบอร์ โดยเฉพาะเมื่อผสานกับเทคโนโลยีปัญญาประดิษฐ์และการเรียนรู้แบบโต้ตอบ อย่างไรก็ตาม งานวิจัยส่วนใหญ่ยังมุ่งเน้นด้านการให้ความรู้เป็นหลัก ขณะที่การพัฒนาแชตบอทที่สามารถผสานการให้ความรู้ร่วมกับการตรวจสอบภัยคุกคามแบบเรียลไทม์ยังมีอยู่อย่างจำกัด ซึ่งเป็นช่องว่างสำคัญที่งานวิจัยนี้มุ่งพัฒนาเพิ่มเติม

จากการทบทวนงานวิจัยดังกล่าว สามารถสรุปได้ว่าแชตบอทเป็นเครื่องมือที่มีศักยภาพในการส่งเสริมความตระหนักรู้ด้านความมั่นคงปลอดภัยไซเบอร์ โดยเฉพาะเมื่อผสานกับเทคโนโลยีปัญญาประดิษฐ์และการเรียนรู้แบบโต้ตอบ อย่างไรก็ตาม งานวิจัยส่วนใหญ่ยังมุ่งเน้นด้านการให้ความรู้เป็นหลัก ขณะที่การพัฒนาแชตบอทที่สามารถผสานการให้ความรู้ร่วมกับการตรวจสอบและวิเคราะห์ภัยคุกคามแบบเรียลไทม์ยังมีอยู่อย่างจำกัด ดังนั้นงานวิจัยนี้จึงมุ่งพัฒนาแชตบอทที่ไม่เพียงให้ความรู้ แต่สามารถตรวจสอบและวิเคราะห์ภัยคุกคามแบบเรียลไทม์เพื่อสนับสนุนการตัดสินใจของผู้ใช้งานได้อย่างมีประสิทธิภาพ

## 3. วิธีดำเนินการวิจัย

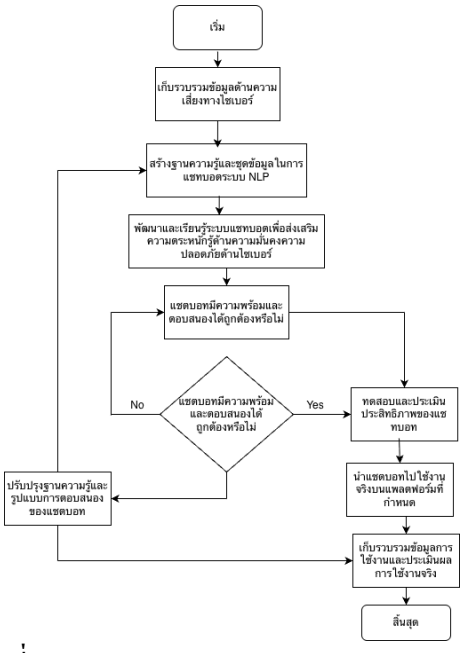
งานวิจัยนี้มุ่งพัฒนาแชตบอตเพื่อส่งเสริมความตระหนักรู้ด้านความมั่นคงปลอดภัยไซเบอร์ โดยครอบคลุมตั้งแต่การออกแบบระบบ การพัฒนา การทดลองใช้งาน และการประเมินผลอย่างเป็นระบบ

### 3.1 การศึกษาและรวบรวมข้อมูล

ศึกษาทฤษฎี งานวิจัย และแนวทางพัฒนาแชตบอต รวมถึงหลักการด้านความมั่นคงปลอดภัยไซเบอร์ เพื่อนำมาใช้เป็นพื้นฐานในการออกแบบระบบ

### 3.2 การวิเคราะห์และออกแบบระบบ

วิเคราะห์ความต้องการผู้ใช้และออกแบบโครงสร้างระบบ แชตบอต รวมถึง Flow การสนทนา โดยใช้เทคโนโลยี NLP และเชื่อมต่อฐานข้อมูลและ API ภายนอก




**ภาพที่ 3-1** แผนภาพแสดงกระบวนการวิจัย

ในการออกแบบระบบแชตบอท ผู้วิจัยได้ให้ความสำคัญกับการกำหนดขอบเขตเนื้อหา (Knowledge Base) ให้ครอบคลุมภัยคุกคามทางไซเบอร์ที่พบบ่อยในบริบทขององค์กร โดยเนื้อหาภายในระบบถูกออกแบบให้ครอบคลุม 4 ด้านหลัก ได้แก่ การโจมตีแบบฟิชชิง (Phishing) การตั้งรหัสผ่านที่ปลอดภัย (Secure Password) การโจมตีแบบวิศวกรรมสังคม (Social Engineering) และการตรวจสอบข้อมูลภัยคุกคาม (Threat Intelligence) เช่น IP Address, URL และ Domain การเลือกเนื้อหาในลักษณะดังกล่าวมีพื้นฐานจากแนวคิดด้านความมั่นคงปลอดภัยไซเบอร์ที่ให้ความสำคัญกับ “ปัจจัยมนุษย์” ซึ่งเป็นจุดอ่อนสำคัญขององค์กร โดยเฉพาะภัยคุกคามที่อาศัยพฤติกรรมของผู้ใช้งาน เช่น ฟิชชิง และ Social Engineering ซึ่งสามารถลดความเสี่ยงได้ผ่านการให้ความรู้และการสร้างความตระหนักรู้

นอกจากนี้ ผู้วิจัยได้ประยุกต์ใช้กรอบแนวคิดการพัฒนาแชตบอท (Chatbot Development Framework) เพื่อให้การออกแบบและพัฒนาระบบเป็นไปอย่างเป็นระบบ โดยประกอบด้วย 4 ขั้นตอน ได้แก่ (1) การวิเคราะห์

ความต้องการของผู้ใช้งาน (Requirement Analysis) (2) การออกแบบโครงสร้างระบบและบทสนทนา (Design) (3) การพัฒนาและเชื่อมต่อระบบ (Development) และ (4) การประเมินผล (Evaluation) ซึ่งช่วยให้ระบบสามารถตอบสนองต่อผู้ใช้งานได้อย่างมีประสิทธิภาพและสามารถประเมินผลได้อย่างชัดเจน

**3.3 การออกแบบและพัฒนาเครื่องมือวัดผล**

เครื่องมือที่ใช้ในการวิจัยประกอบด้วยแบบทดสอบ Pre-Test และ Post-Test ซึ่งเป็นแบบปรนัย (Multiple Choice) จำนวน 15 ข้อ ครอบคลุม 3 ด้านหลัก ได้แก่ (1) ความรู้ด้านภัยคุกคามไซเบอร์ (2) ความตระหนักรู้และ (3) พฤติกรรมด้านความปลอดภัย โดยแบบทดสอบทั้งสองชุดถูกออกแบบให้มีระดับความยากใกล้เคียงกัน ตัวอย่างข้อคำถาม เช่น การวิเคราะห์ลักษณะของอีเมลฟิชชิง การเลือกรหัสผ่านที่มีความปลอดภัย และการตัดสินใจเมื่อพบลิงก์หรือข้อมูลที่น่าสงสัย ซึ่งสะท้อนถึงสถานการณ์ที่ผู้ใช้งานสามารถพบได้ในชีวิตจริง

แบบสอบถามดังกล่าวได้รับการตรวจสอบความตรงเชิงเนื้อหา (Content Validity) โดยผู้เชี่ยวชาญจำนวน 10 ท่าน (pilot test) และได้มีการปรับปรุงตามข้อเสนอแนะเพื่อให้เหมาะสมกับวัตถุประสงค์ของงานวิจัย ในด้านความเชื่อมั่น (Reliability) ได้ทำการทดสอบโดยใช้ค่าสัมประสิทธิ์ Cronbach's Alpha ซึ่งมีค่าเท่ากับ 0.82 แสดงว่าเครื่องมือวิจัยมีความเชื่อมั่นอยู่ในระดับดี และสามารถนำไปใช้วัดผลได้อย่างเหมาะสม

**3.4 การพัฒนาและติดตั้งระบบ**

พัฒนาแชตบอตบนแพลตฟอร์ม LINE โดยใช้ LINE Messaging API และเชื่อมต่อระบบ Backend และ AI เพื่อประมวลผลคำสั่งผู้ใช้งาน

**3.5 ประชากรและกลุ่มตัวอย่าง**

กลุ่มตัวอย่างจำนวน 30 คน โดยมีการทดลองใช้งานเบื้องต้น (pilot test) กับผู้ใช้งาน 10 คน เพื่อปรับปรุงระบบก่อนการเก็บข้อมูลจริง

**3.6 การเก็บรวบรวมข้อมูล**

เก็บข้อมูลผ่าน Pre-test, การใช้งานแชตบอต และ Post-test รวมถึงแบบสอบถามความพึงพอใจ

**3.7 การวิเคราะห์ข้อมูล**

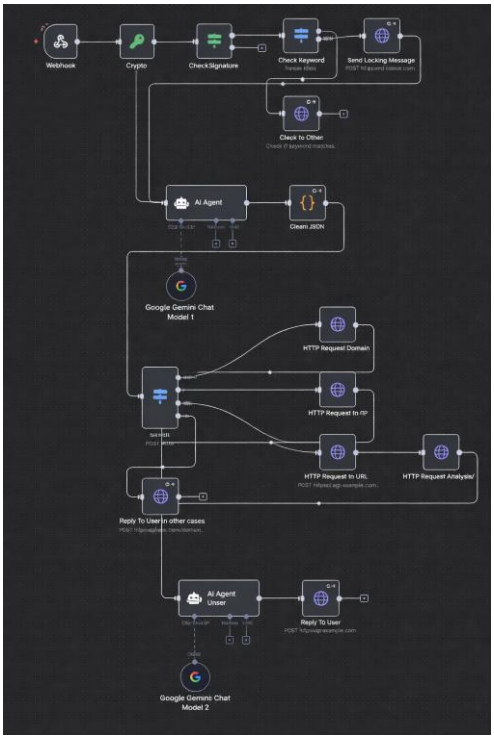
การวิเคราะห์ข้อมูลในงานวิจัยนี้ใช้สถิติเชิงอนุมาน โดยใช้การทดสอบค่าทีแบบกลุ่มตัวอย่างสัมพันธ์กัน (Paired Samples t-test) เพื่อเปรียบเทียบความแตกต่างของคะแนนความรู้ด้านความมั่นคงปลอดภัยไซเบอร์ของผู้ใช้งานก่อน (Pre-Test) และหลัง (Post-Test) การใช้งานแชตบอต

กลุ่มตัวอย่างประกอบด้วยบุคลากรจำนวน 30 คน โดยเก็บข้อมูลคะแนนจากแบบสอบถามก่อนและหลังการใช้งานระบบ จากนั้นนำมาคำนวณค่าความแตกต่างของคะแนนรายบุคคล และวิเคราะห์ด้วยวิธี Paired Samples t-test เพื่อประเมินว่าการใช้งานแชตบอตส่งผลต่อระดับความรู้ของผู้ใช้งานอย่างมีนัยสำคัญหรือไม่

ในการทดสอบได้กำหนดระดับนัยสำคัญทางสถิติที่ $\alpha = 0.05$ โดยพิจารณาค่า p-value หากมีค่าน้อยกว่า 0.05 แสดงว่าคะแนนหลังการใช้งานแตกต่างจากก่อนการใช้งานอย่างมีนัยสำคัญทางสถิติ ซึ่งสามารถสรุปได้ว่าระบบแชตบอตมีประสิทธิภาพในการส่งเสริมความตระหนักรู้ด้านความมั่นคงปลอดภัยไซเบอร์ของผู้ใช้งาน

**3.8 เครื่องมือและแพลตฟอร์ม**

ใช้ Dialogflow, LINE Messaging API และระบบ Backend ในการพัฒนาและประมวลผลแชตบอต




**ภาพที่ 3-2** แผนภาพแสดงสถาปัตยกรรม

ระบบแชตบอตที่พัฒนาขึ้นประกอบด้วยองค์ประกอบหลัก ได้แก่ ผู้ใช้งานผ่านแอปพลิเคชัน LINE, LINE Messaging API สำหรับรับ–ส่งข้อความ, โมดูลประมวลผลภาษาธรรมชาติ (NLP) โดยใช้ Dialogflow, และระบบ Backend ที่ทำหน้าที่จัดการตรรกะของระบบและเชื่อมต่อกับฐานข้อมูลและ API ภายนอก เมื่อผู้ใช้งานส่งข้อความ ระบบจะทำการวิเคราะห์เจตนา (Intent) และดึงข้อมูลที่เกี่ยวข้องจากแหล่งข้อมูลภายนอกก่อนส่งผลลัพธ์กลับไปยังผู้ใช้งานในรูปแบบที่เข้าใจง่าย

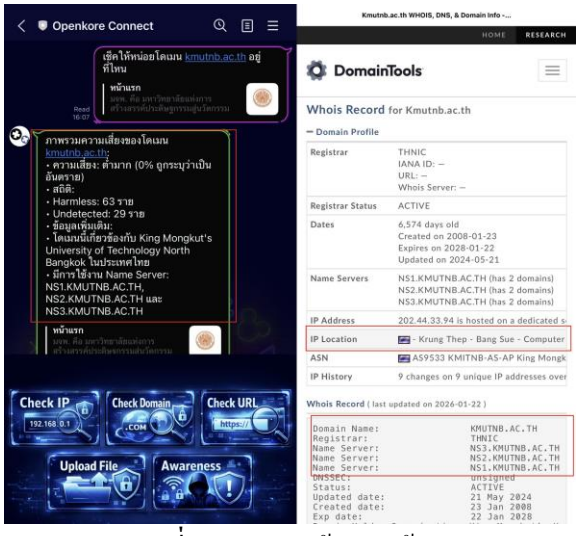
## 4. ผลการดำเนินการวิจัย

### 4.1 ผลการพัฒนาและการทำงานของระบบแชตบอต

ระบบแชตบอตสามารถพัฒนาและทำงานได้ตามที่ออกแบบ โดยรองรับการให้ความรู้และตรวจสอบข้อมูลด้านความมั่นคงปลอดภัยไซเบอร์ผ่านแพลตฟอร์ม LINE ได้อย่างมีประสิทธิภาพ และสามารถตอบสนองผู้ใช้งานได้อย่างถูกต้องและรวดเร็ว

### 4.2 ผลการประเมินประสิทธิภาพของระบบ

ระบบสามารถให้ข้อมูลได้ถูกต้อง โดยผลลัพธ์จากแชตบอตมีความสอดคล้องกับแหล่งข้อมูลภายนอก เช่น DomainTools, MXToolbox และ VirusTotal อีกทั้งสามารถสรุปผลความเสี่ยงในรูปแบบที่เข้าใจง่าย ช่วยให้ผู้ใช้งานตัดสินใจได้รวดเร็ว




**ภาพที่ 4-1** ความถูกต้องของข้อมูล

### 4.3 ผลการวิเคราะห์ Pre-Test และ Post-Test

การวิเคราะห์ผลในส่วนนี้มีวัตถุประสงค์เพื่อเปรียบเทียบระดับความรู้ด้านความมั่นคงปลอดภัยไซเบอร์ของผู้ใช้งานก่อนและหลังการใช้งานแชตบอท โดยใช้แบบทดสอบ Pre-Test และ Post-Test ซึ่งมีจำนวนข้อและโครงสร้างใกล้เคียงกัน โดยใน Post-Test ได้มีการปรับเปลี่ยนลำดับคำถามและรูปแบบตัวเลือกเพื่อลดผลกระทบจากการจดจำคำตอบ

ผลการวิเคราะห์พบว่าคะแนนเฉลี่ยหลังการใช้งาน (Post-Test) สูงกว่าก่อนการใช้งาน (Pre-Test) โดยเพิ่มขึ้นจาก 13.47 เป็น 14.03 คะแนน หรือเพิ่มขึ้น 0.56 คะแนน แสดงให้เห็นว่าผู้ใช้งานมีระดับความรู้และความตระหนักรู้ด้านความมั่นคงปลอดภัยไซเบอร์เพิ่มขึ้นหลังจากใช้งานแชตบอท

**ตารางที่ 4-1** ผลคะแนนการประเมิน

| ข้อ | รายการประเมิน | Mean ($\bar{x}$) | SD |
|---|---|---|---|
| 1 | เนื้อหามีความถูกต้องและน่าเชื่อถือ | 4.60 | 0.67 |
| 2 | คำอธิบายเข้าใจง่าย | 4.10 | 0.52 |
| 3 | เนื้อหามีประโยชน์ | 4.00 | 0.74 |
| 4 | ระบบใช้งานง่าย | 4.20 | 0.53 |
| 5 | เพิ่มความตระหนักรู้ | 4.20 | 0.53 |
| 6 | ระบบตอบสนองรวดเร็ว | 4.70 | 0.53 |
| 7 | ระบบไม่ซับซ้อน | 4.20 | 0.53 |
| 8 | รูปแบบใช้งานดี | 4.20 | 0.53 |
| 9 | การตอบสนองเหมาะสม | 4.20 | 0.84 |
| 10 | ความพึงพอใจโดยรวม | 4.70 | 0.82 |
| รวมเฉลี่ย | | 4.31 | - |

4.3.1 การวิเคราะห์ทางสถิติ

การวิเคราะห์ทางสถิติในงานวิจัยนี้ใช้การทดสอบค่าทีแบบกลุ่มตัวอย่างสัมพันธ์กัน (Paired Samples t-test) เพื่อเปรียบเทียบคะแนนก่อนและหลังการใช้งานแชตบอต โดยกำหนดระดับนัยสำคัญทางสถิติที่ 0.05 ผลการวิเคราะห์

พบว่าค่าเฉลี่ยของผลต่างคะแนนระหว่าง Post-Test และ Pre-Test เท่ากับ 0.56 คะแนน และมีค่าส่วนเบี่ยงเบนมาตรฐานของผลต่างเท่ากับ 0.93 คะแนน เมื่อนำมาคำนวณค่าทีพบว่าได้ค่า t = 3.28 ที่องศาอิสระ df = 29 และมีค่า p = 0.002 ซึ่งมีค่าน้อยกว่า 0.05 แสดงให้เห็นว่าคะแนนหลังการใช้งานแชตบอตสูงกว่าก่อนการใช้งานอย่างมีนัยสำคัญทางสถิติ จึงสามารถสรุปได้ว่าแชตบอตที่พัฒนาขึ้นมีประสิทธิภาพในการช่วยเพิ่มความรู้และความตระหนักรู้ด้านความมั่นคงปลอดภัยไซเบอร์ของผู้ใช้งานได้อย่างมีนัยสำคัญ

4.3.2 ผลความพึงพอใจของผู้ใช้งาน

ผลการประเมินความพึงพอใจของผู้ใช้งานต่อระบบแชตบอตเพื่อส่งเสริมความตระหนักรู้ด้านความมั่นคงปลอดภัยไซเบอร์ พบว่าผู้ใช้งานมีความพึงพอใจโดยรวมอยู่ในระดับมาก โดยมีค่าเฉลี่ยเท่ากับ 4.31 จากคะแนนเต็ม 5.00 หรือคิดเป็นร้อยละ 86.20 ซึ่งสะท้อนให้เห็นว่าระบบสามารถตอบสนองความต้องการของผู้ใช้งานได้อย่างมีประสิทธิภาพ ทั้งในด้านความถูกต้องของข้อมูล ความชัดเจนของเนื้อหา และความสะดวกในการใช้งาน อีกทั้งยังสามารถนำความรู้ไปประยุกต์ใช้ในสถานการณ์จริงได้อย่างเหมาะสม แสดงให้เห็นถึงศักยภาพของระบบในการนำไปใช้งานจริงในองค์กรต่อไป

## 5. สรุปผลการวิจัย

งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาแชตบอตเพื่อส่งเสริมความตระหนักรู้ด้านความมั่นคงปลอดภัยไซเบอร์ และประเมินประสิทธิภาพของระบบจากการใช้งานจริง โดยแชตบอตถูกพัฒนาให้สามารถให้ความรู้และตรวจสอบข้อมูลด้านความปลอดภัย เช่น IP Address, URL และ Domain ผ่านแพลตฟอร์ม LINE โดยใช้เทคโนโลยีการประมวลผลภาษาธรรมชาติและการเชื่อมต่อกับแหล่งข้อมูลภายนอก ผลการวิจัยพบว่าแชตบอตสามารถทำงานได้ตามวัตถุประสงค์ และช่วยเพิ่มระดับความรู้ของผู้ใช้งานอย่างมีนัยสำคัญทางสถิติ ($p < 0.05$) โดยคะแนนหลังการใช้งานสูงกว่าก่อนการใช้งาน นอกจากนี้ ผู้ใช้งานยังมีความพึงพอใจต่อระบบในระดับสูง สะท้อนให้เห็นว่าระบบมีประสิทธิภาพและเหมาะสมต่อการนำไปประยุกต์ใช้ในองค์กรเพื่อส่งเสริมความตระหนักรู้ด้านความมั่นคงปลอดภัยไซเบอร์ในอนาคต

## เอกสารอ้างอิง


[1]. Basri, W. S., & Ali, H. H. (2024). Exploring the impact of chatbot integration on cybersecurity awareness and behaviour.

[2]. Abdelhamid, S. E., Mallari, T., & Aly, M. (2023). Cybersecurity awareness using intelligent chatbots.

[3]. Chavalitlekha, S. (2023). Enhancing cybersecurity awareness for personnel.

[4]. Chatchalermpun, S., & Daengsi, T. (2022). Simulated phishing awareness.

[5]. Eakacet, P. (2023). AI in cybersecurity.

[6]. Phrakhrusuthivorayan (2021). Knowledgeable chatbot.

[7]. Roonras, S. (2018). Factors affecting awareness.

[8]. Fung, Y.-C., & Lee, L.-K. (2022). Chatbot for cybersecurity awareness.

[9]. Hamad, S., & Yeferny, T. (2020). Chatbot for information security.

[10]. Pimentel, L., et al. (2025). Public chatbot for cybercrime awareness.

[11]. Reddy, K. T. (2024). Deep learning chatbot for phishing.

[12]. He, J., & Xin, C. (2021). AI chatbot for cybersecurity learning.

[13]. Keepnet Labs (2025). Chatbot for security awareness training.

[14]. Microsoft Threat Intelligence (2025). AI-based phishing detection.

[15]. Kooli, C. (2023). Chatbots in education and ethics.

[16]. Smith, F. (2025). Modern phishing awareness.

[17]. Lim, J. (2025). AI-driven cybersecurity awareness training.

# การพัฒนาระบบแนะนำการท่องเที่ยวเชิงความเชื่อในจังหวัดนครปฐม
# Development of a Faith-Based Tourism Recommendation System in Nakhon Pathom Province

*วิทวัส จันกิมฮะ (Witthawat Chankimha)[1] อดิศักดิ์ ยงปัญญา (Adisak Yongpanya)[2]*
*วรรณชัย เชื้อทอง (Wannachai Chouegthong)[3] ศุภกฤษ นาคป้อมฉิน (Supakit Nakpomchin)[4]*
*และ สุพิชฌาย์ จันทร์เรือง (Suphitcha Chanrueang)[5]*
*สาขาวิชาเทคโนโลยีสารสนเทศ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏนครปฐม*
*[1]664230029@webmail.npru.ac.th, [2]664230034@webmail.npru.ac.th, [3]664230048@webmail.npru.ac.th,*
*[4]supakit@webmail.npru.ac.th, [5]suphitcha@webmail.npru.ac.th*

**บทคัดย่อ**

*งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาระบบแนะนำการท่องเที่ยวเชิงความเชื่อในจังหวัดนครปฐมในรูปแบบเว็บแอปพลิเคชันแบบตอบสนองต่อการแสดงผล โดยมีแรงจูงใจจากปัญหาการกระจายตัวของข้อมูลสถานที่ท่องเที่ยว ซึ่งส่งผลให้ผู้ใช้งานเข้าถึงข้อมูลที่สอดคล้องกับความสนใจเฉพาะบุคคลได้ยาก ระบบที่นำเสนอประยุกต์ใช้เทคนิค User-Based Collaborative Filtering ร่วมกับ Cosine Similarity เพื่อสร้างคำแนะนำแบบเฉพาะบุคคล โดยอาศัยข้อมูลการให้คะแนนและข้อมูลลักษณะทางประชากรศาสตร์ เช่น อายุ และเพศ กระบวนการดำเนินงานประกอบด้วยการสร้างเมทริกซ์ผู้ใช้–สถานที่ และการคำนวณความคล้ายคลึงของผู้ใช้งานในการประเมินประสิทธิภาพ ข้อมูลถูกแบ่งเป็นชุดฝึกและชุดทดสอบในอัตราส่วน 80:20 ผลการทดลองแสดงให้เห็นว่าวิธีที่นำเสนอให้ค่า Precision และ Recall สูงกว่าวิธีพื้นฐาน นอกจากนี้ ผลการประเมินความพึงพอใจของผู้ใช้งานอยู่ในระดับมาก ผลลัพธ์ดังกล่าวสะท้อนให้เห็นว่าระบบสามารถสนับสนุนการตัดสินใจและการวางแผนการท่องเที่ยวเชิงความเชื่อได้อย่างมีประสิทธิภาพ*

**คำสำคัญ:** การท่องเที่ยวเชิงความเชื่อ ระบบแนะนำ การกรองข้อมูลแบบร่วมมือโดยอิงผู้ใช้งาน ความคล้ายคลึงแบบโคไซน์

**Abstract**

*This study aims to develop a faith-based tourism recommendation system for Nakhon Pathom Province as a responsive web application. The research is motivated by the fragmented nature of faith-based tourism information, which makes it difficult for users to find destinations that match their individual preferences. The proposed system applies User-Based Collaborative Filtering combined with Cosine Similarity to generate personalized recommendations using user ratings and demographic data, including age and gender. The process involves constructing a user–item matrix and computing user similarity to identify relevant recommendations. For evaluation, the dataset is divided into training and testing sets using an 80:20 split. Experimental results show that the proposed method achieves higher Precision and Recall compared to a baseline approach. In addition, user satisfaction is rated at a high level. These results indicate that the system effectively supports decision-making and travel planning in the context of faith-based tourism.*

**Keywords:** Faith-based tourism, Recommendation system, User-based collaborative filtering, Cosine similarity.

## 1. บทนำ

การพัฒนาของเทคโนโลยีสารสนเทศในยุคดิจิทัลได้เปลี่ยนรูปแบบการตัดสินใจด้านการท่องเที่ยวอย่างชัดเจน โดยนักท่องเที่ยวหันมาใช้เทคโนโลยีในการค้นหาและวางแผนการเดินทางมากขึ้น ส่งผลให้อุตสาหกรรมการท่องเที่ยวกลายเป็นกลไกสำคัญในการขับเคลื่อนเศรษฐกิจหนึ่งในรูปแบบที่ได้รับความนิยมอย่างต่อเนื่องคือ “การท่องเที่ยวเชิงความเชื่อ” ซึ่งมุ่งตอบสนองความต้องการทางจิตใจและความเชื่อส่วนบุคคล เช่น ด้านการงาน การเงิน และความรัก

อย่างไรก็ตาม ข้อมูลเกี่ยวกับสถานที่ท่องเที่ยวเชิงความเชื่อยังคงกระจัดกระจายอยู่ในหลายแหล่งบนอินเทอร์เน็ต ทำให้นักท่องเที่ยวต้องใช้เวลาในการค้นหา เปรียบเทียบ และตัดสินใจด้วยตนเอง ขณะเดียวกัน ระบบแผนที่ทั่วไป เช่น Google Maps ยังมีข้อจำกัดในการแนะนำสถานที่ เนื่องจากผู้ใช้งานต้องทราบชื่อสถานที่ล่วงหน้า และไม่สามารถคัดกรองตามความต้องการเฉพาะบุคคลได้ [2] ส่งผลให้การวางแผนการเดินทางมีความซับซ้อนและขาดประสิทธิภาพ

ด้วยเหตุนี้ ระบบแนะนำจึงถูกนำมาใช้เป็นเครื่องมือสำคัญในการคัดกรองข้อมูลและนำเสนอรายการที่เหมาะสมกับผู้ใช้งาน [3] โดยเฉพาะเทคนิคการกรองข้อมูลแบบร่วมมือ ซึ่งสามารถวิเคราะห์ความคล้ายคลึงของผู้ใช้งานเพื่อสร้างคำแนะนำที่สอดคล้องกับความสนใจได้อย่างมีประสิทธิภาพ [4] นอกจากนี้ งานวิจัยที่ผ่านมาแสดงให้เห็นว่าเทคนิคการวัดความคล้ายคลึง เช่น Cosine Similarity มีบทบาทสำคัญในการเพิ่มความแม่นยำของระบบ [5]

อย่างไรก็ตาม ระบบแนะนำส่วนใหญ่มักอาศัยข้อมูลพฤติกรรมเชิงลึกของผู้ใช้งาน เช่น ประวัติการใช้งานหรือข้อมูลบริบท ซึ่งในทางปฏิบัติอาจไม่สามารถจัดเก็บได้อย่างครบถ้วน โดยเฉพาะในระบบที่เพิ่งเริ่มต้นใช้งาน

ดังนั้น งานวิจัยนี้จึงนำเสนอการพัฒนาระบบแนะนำการท่องเที่ยวเชิงความเชื่อในจังหวัดนครปฐมในรูปแบบเว็บแอปพลิเคชัน โดยประยุกต์ใช้เทคนิค User-Based Collaborative Filtering ร่วมกับ Cosine Similarity เพื่อสร้างคำแนะนำแบบเฉพาะบุคคลจากข้อมูลการให้คะแนนและข้อมูลลักษณะทางประชากรศาสตร์ เช่น อายุ และเพศ โดยไม่อาศัยข้อมูลพฤติกรรมเชิงลึก

จุดเด่นของงานวิจัยนี้คือความสามารถในการรองรับกรณีที่ข้อมูลผู้ใช้งานมีจำนวนจำกัด ซึ่งช่วยลดปัญหา cold-start และเพิ่มความเหมาะสมของคำแนะนำ นอกจากนี้ ระบบยังเชื่อมต่อกับ Google Maps เพื่อสนับสนุนการนำทางไปยังสถานที่ท่องเที่ยว

ผลลัพธ์ของงานวิจัยนี้คาดว่าจะช่วยเพิ่มความสะดวกในการวางแผนการเดินทาง ลดเวลาในการค้นหาข้อมูล และสนับสนุนการตัดสินใจของผู้ใช้งานได้อย่างมีประสิทธิภาพ

## 2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ในการพัฒนาระบบแนะนำการท่องเที่ยวเชิงความเชื่อ คณะผู้วิจัยได้ศึกษาแนวคิด ทฤษฎี และงานวิจัยที่เกี่ยวข้อง เพื่อใช้เป็นแนวทางในการออกแบบและพัฒนาระบบให้มีประสิทธิภาพ โดยสามารถแบ่งออกเป็น 3 ส่วนหลัก ได้แก่ ระบบแนะนำ เทคนิคการกรองข้อมูลแบบร่วมมือ และงานวิจัยที่เกี่ยวข้องด้านการท่องเที่ยว

### 2.1 ระบบแนะนำ

ระบบแนะนำเป็นเทคโนโลยีที่ถูกพัฒนาขึ้นเพื่อช่วยคัดกรองข้อมูลและนำเสนอรายการที่เหมาะสมกับผู้ใช้งาน โดยอาศัยการวิเคราะห์ข้อมูลความสนใจหรือพฤติกรรมของผู้ใช้งาน ระบบแนะนำถูกนำไปใช้ในหลายบริบท เช่น การแนะนำสินค้า ภาพยนตร์ และสถานที่ท่องเที่ยว ซึ่งช่วยลดภาระในการค้นหาข้อมูลและเพิ่มประสิทธิภาพในการตัดสินใจ [1] โดยทั่วไประบบแนะนำสามารถแบ่งออกเป็นหลายประเภท เช่น Content-Based Filtering, Collaborative Filtering และ Hybrid Methods ซึ่งแต่ละวิธีมีข้อดีและข้อจำกัดแตกต่างกัน

### 2.2 การกรองข้อมูลแบบร่วมมือโดยอิงผู้ใช้งาน

Collaborative Filtering เป็นเทคนิคที่ได้รับความนิยมในการพัฒนาระบบแนะนำ โดยอาศัยแนวคิดว่าผู้ใช้งานที่มีความสนใจคล้ายกันในอดีตจะมีแนวโน้มที่จะมีความสนใจคล้ายกันในอนาคต เทคนิคนี้สามารถแบ่งออกเป็น 2

รูปแบบหลัก ได้แก่ User-Based และ Item-Based Collaborative Filtering [6], [7] ในงานวิจัยนี้เลือกใช้ User-Based Collaborative Filtering เนื่องจากสามารถนำข้อมูลการให้คะแนนของผู้ใช้งานมาวิเคราะห์ความสัมพันธ์ระหว่างผู้ใช้งานได้โดยตรง ซึ่งเหมาะสมกับบริบทของระบบแนะนำที่มีข้อมูลผู้ใช้งานในระดับจำกัด

**2.3 การวัดความคล้ายคลึง**

การวัดความคล้ายคลึงเป็นขั้นตอนสำคัญในการพัฒนาระบบแนะนำแบบโคไซน์ เป็นวิธีที่ได้รับความนิยม เนื่องจากสามารถวัดความคล้ายคลึงระหว่างเวกเตอร์ข้อมูลได้อย่างมีประสิทธิภาพและเหมาะสมกับข้อมูลประเภทการให้คะแนน [1], [8]

**2.4 ระบบแนะนำในด้านการท่องเที่ยว**

ระบบแนะนำถูกนำมาใช้ในด้านการท่องเที่ยวเพื่อช่วยให้ผู้ใช้งานสามารถค้นหาและเลือกสถานที่ท่องเที่ยวที่เหมาะสมกับความสนใจได้ โดยมีการนำเทคนิคต่าง ๆ เช่น Collaborative Filtering และ Artificial Intelligence มาประยุกต์ใช้ [8], [9] โดยงานวิจัยที่ผ่านมาได้แสดงให้เห็นว่า การใช้ระบบแนะนำสามารถช่วยเพิ่มความพึงพอใจของผู้ใช้งานและลดระยะเวลาในการค้นหาข้อมูลได้อย่างมีประสิทธิภาพ [10] มีการพัฒนาแอปพลิเคชันแนะนำสถานที่ท่องเที่ยวในหลายพื้นที่สามารถให้คำแนะนำตามความต้องการของผู้ใช้งานได้อย่างแม่นยำ [2]

**2.5 งานวิจัยที่เกี่ยวข้อง**

จากการศึกษางานวิจัยที่เกี่ยวข้อง พบว่ามีการนำเทคนิค Collaborative Filtering ไปประยุกต์ใช้ในระบบแนะนำการท่องเที่ยวอย่างแพร่หลาย โดยงานวิจัยล่าสุดได้แสดงให้เห็นถึงบทบาทสำคัญของระบบแนะนำในการเพิ่มความแม่นยำของผลลัพธ์และสนับสนุนการตัดสินใจของผู้ใช้งานในบริบทด้านการท่องเที่ยว [7]

นอกจากนี้ยังมีงานวิจัยที่เปรียบเทียบประสิทธิภาพของวิธี User-Based และ Item-Based Collaborative Filtering ซึ่งพบว่าวิธี User-Based มีประสิทธิภาพที่ดีกว่าในบางกรณี [10] ในบริบทของประเทศไทย งานวิจัยล่าสุดได้มีการพัฒนาระบบแนะนำและแอปพลิเคชันด้านการท่องเที่ยวเชิงความเชื่อ เช่น การพัฒนาเว็บแอปพลิเคชันในรูปแบบ Virtual Tour สำหรับวัดพระปฐมเจดีย์ จังหวัดนครปฐม ซึ่งมีการเชื่อมโยงข้อมูลสถานที่และการนำทางผ่านระบบแผนที่ และผลการประเมินพบว่าผู้ใช้งานมีความพึงพอใจในระดับสูง [11] นอกจากนี้ งานวิจัยล่าสุดได้แสดงให้เห็นถึงการประยุกต์ใช้เทคโนโลยีร่วมกับการท่องเที่ยวเชิงความเชื่อ โดยใช้ระบบแนะนำเพื่อสร้างประสบการณ์แบบเฉพาะบุคคล ซึ่งสามารถช่วยเพิ่มความพึงพอใจและประสิทธิภาพในการให้บริการแก่นักท่องเที่ยวได้อย่างมีนัยสำคัญ [12]

จากการทบทวนวรรณกรรมดังกล่าว สามารถสรุปได้ว่า การนำเทคนิค Collaborative Filtering มาประยุกต์ใช้ร่วมกับข้อมูลผู้ใช้งาน เป็นแนวทางที่มีประสิทธิภาพในการพัฒนาระบบแนะนำการท่องเที่ยว โดยเฉพาะในบริบทของการท่องเที่ยวเชิงความเชื่อ ซึ่งยังมีโอกาสในการพัฒนาเพิ่มเติมในด้านความเป็นส่วนบุคคล [9]

**2.6 การวัดประสิทธิภาพของระบบแนะนำ**

การประเมินประสิทธิภาพของระบบแนะนำเป็นขั้นตอนสำคัญในการตรวจสอบความถูกต้องและความเหมาะสมของผลลัพธ์ โดยทั่วไปนิยมใช้ตัวชี้วัดเชิงปริมาณ (Quantitative Metrics) เพื่อวัดความแม่นยำของระบบ [13]

ในการประเมินระบบแนะนำ ใช้ Precision และ Recall โดย Precision วัดความแม่นยำของรายการที่แนะนำ ส่วน Recall วัดความครอบคลุมของรายการที่ตรงกับความสนใจของผู้ใช้งาน ซึ่งช่วยให้สามารถเปรียบเทียบประสิทธิภาพของระบบได้อย่างชัดเจน

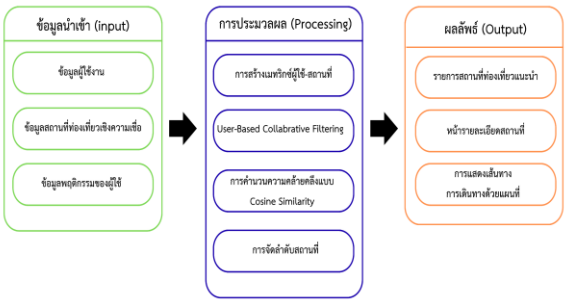
## 3. วิธีการดำเนินการวิจัย

งานวิจัยนี้นำเสนอระบบแนะนำการท่องเที่ยวเชิงความเชื่อ โดยประยุกต์ใช้เทคนิค User-Based Collaborative Filtering ร่วมกับการคำนวณความคล้ายคลึงแบบโคไซน์ เพื่อสร้างคำแนะนำที่เหมาะสมกับผู้ใช้งานรายบุคคลโดยมีรายละเอียดดังนี้

**3.1 ภาพรวมของระบบ**

ระบบแนะนำที่พัฒนาขึ้นประกอบด้วย 3 ส่วนหลัก ได้แก่ ข้อมูลนำเข้า (Input) กระบวนการประมวลผล (Processing) และผลลัพธ์ (Output) ดังแสดงในภาพที่ 1

เพื่อแสดงภาพรวมของกระบวนการทำงานของระบบแนะนำ




**ภาพที่ 1 :** โครงสร้างการทำงานของระบบแนะนำการท่องเที่ยวเชิงความเชื่อ

จากภาพที่ 1 สามารถอธิบายระบบได้ดังนี้

1) ข้อมูลนำเข้า ประกอบด้วยข้อมูลที่ใช้ในการประมวลผล ได้แก่ ข้อมูลผู้ใช้งาน เช่น อายุ และเพศ ข้อมูลสถานที่ท่องเที่ยวเชิงความเชื่อ ข้อมูลการให้คะแนน (Rating) ของผู้ใช้งาน

2) กระบวนการประมวลผล ระบบทำการประมวลผลข้อมูลผ่านขั้นตอน การสร้างเมทริกซ์ผู้ใช้-สถานที่ (User-Item Matrix) การประยุกต์ใช้เทคนิค User-Based Collaborative Filtering การคำนวณความคล้ายคลึงของผู้ใช้งานด้วยความคล้ายคลึงแบบโคไซน์ และการจัดลำดับสถานที่ท่องเที่ยวความคล้ายคลึง

3) ผลลัพธ์ (Output) ระบบแสดงผลลัพธ์ในรูปแบบรายการสถานที่ท่องเที่ยวที่แนะนำ รายละเอียดของสถานที่ การแสดงเส้นทางการเดินทางผ่านแผนที่ (Google Maps)

**3.2 การเตรียมข้อมูล**

ข้อมูลที่ใช้ในงานวิจัยนี้อยู่ในรูปแบบเมทริกซ์ความสัมพันธ์ระหว่างผู้ใช้และรายการ โดยกำหนดให้ แถว (Row) แทนผู้ใช้งาน คอลัมน์ (Column) แทนสถานที่ท่องเที่ยว ค่าในตาราง แทนคะแนนความพึงพอใจ (Rating)

**3.3 การคำนวณความคล้ายคลึง**

การวัดความคล้ายคลึงระหว่างผู้ใช้งานใช้วิธี แบบโคไซน์ ดังสมการในภาพที่ 2

$$sim(u,v) = \frac{\sum_{i=1}^{n} r_{u,i} r_{v,i}}{\sqrt{\sum_{i=1}^{n} r_{u,i}^2} \sqrt{\sum_{i=1}^{n} r_{v,i}^2}}$$

**ภาพที่ 2 :** สูตรการคำนวณความคล้ายคลึง

โดยที่ โดยที่ $u$ และ $v$ คือผู้ใช้งาน และ $r_{u,i}$ คือคะแนนของผู้ใช้งาน $u$ ต่อสถานที่ $i$

**3.4 การสร้างคำแนะนำ**

ระบบเลือกผู้ใช้งานที่มีความคล้ายคลึงสูงสุด (Top-N Neighbors) และนำข้อมูลของผู้ใช้งานกลุ่มดังกล่าวมาคำนวณคะแนนพยากรณ์ของสถานที่ โดยใช้สมการดังแสดงในภาพที่ 3

$$\hat{r}u,i = \frac{\sum v \in N(u)\, sim(u,v) \cdot r_{v,i}}{\sum_{v \in N(u)} sim(u,v)} \quad (2)$$

**ภาพที่ 3:** สูตรการคำนวณการคำนวณคะแนนพยากรณ์ของผู้ใช้งานจากกลุ่มผู้ใช้งานที่มีความคล้ายคลึง

จากนั้นระบบจะทำการจัดลำดับคะแนนพยากรณ์ และเลือกสถานที่ที่มีคะแนนสูงสุดมาแนะนำแก่ผู้ใช้งาน

**3.5 ขั้นตอนการทำงานของระบบ**

กระบวนการทำงานของอัลกอริทึมเริ่มต้นจากการรับข้อมูลการให้คะแนนจากผู้ใช้งาน จากนั้นทำการสร้างเมทริกซ์ผู้ใช้–สถานที่ เพื่อใช้เป็นฐานข้อมูลในการประมวลผลต่อมาระบบจะคำนวณค่าความคล้ายคลึงระหว่างผู้ใช้งานและเลือกผู้ใช้งานที่มีความคล้ายคลึงสูงสุด หลังจากนั้นจึงนำข้อมูลของผู้ใช้งานกลุ่มดังกล่าวมาคำนวณคะแนนพยากรณ์ของสถานที่ สุดท้ายระบบจะทำการจัดอันดับและแสดงผลรายการสถานที่ท่องเที่ยวที่เหมาะสมให้แก่ผู้ใช้งาน

**3.6 การประเมินประสิทธิภาพของระบบ**

การประเมินผลใช้ตัวชี้วัด Precision และ Recall

$$Precision = \frac{|Relevant \cap Recommended|}{|Recommended|}$$

$$Recall = \frac{|Relevant \cap Recommended|}{|Relevant|}$$

**ภาพที่ 4:** สมการสำหรับการประเมินประสิทธิภาพของระบบแนะนำ

**3.7 การออกแบบระบบและเครื่องมือที่ใช้**

ระบบแนะนำการท่องเที่ยวเชิงความเชื่อถูกพัฒนาในรูปแบบเว็บแอปพลิเคชันแบบตอบสนองต่อการแสดงผล

เพื่อรองรับการใช้งานผ่านอุปกรณ์ที่หลากหลาย เช่น คอมพิวเตอร์ แท็บเล็ต และสมาร์ตโฟน

1) การออกแบบระบบ ระบบแบ่งออกเป็น 2 ส่วนหลัก ได้แก่ ส่วนของผู้ใช้งาน และส่วนของประมวลผล โดยผู้ใช้งานสามารถให้คะแนนสถานที่ท่องเที่ยวและรับคำแนะนำผ่านแอปพลิเคชัน ขณะที่ส่วนประมวลผลใช้เทคนิค User-Based Collaborative Filtering ร่วมกับ แบบโคไซน์เพื่อวิเคราะห์ความคล้ายคลึงของผู้ใช้งานและสร้างคำแนะนำแบบเฉพาะบุคคล

2) เครื่องมือที่ใช้ในการพัฒนาระบบประกอบด้วย ภาษาโปรแกรมสำหรับพัฒนาเว็บแอปพลิเคชัน ฐานข้อมูลสำหรับจัดเก็บข้อมูลผู้ใช้งานและคะแนน และ Google Maps API สำหรับแสดงตำแหน่งและเส้นทางการเดินทาง

การออกแบบระบบดังกล่าวช่วยให้สามารถพัฒนาและใช้งานระบบแนะนำได้อย่างมีประสิทธิภาพ และรองรับการขยายในอนาคต นอกจากนี้ ระบบยังถูกพัฒนาในรูปแบบเว็บแอปพลิเคชันแบบตอบสนองต่อการแสดงผล (Responsive Web Application) ทำให้ผู้ใช้งานสามารถเข้าถึงระบบได้อย่างสะดวกและได้รับประสบการณ์การใช้งานที่เหมาะสมในทุกอุปกรณ์

## 4. ผลการวิจัยและทดสอบประสิทธิภาพ

### 4.1 การพัฒนาระบบ

ระบบที่พัฒนาขึ้นสามารถใช้งานผ่านอุปกรณ์ที่มีขนาดหน้าจอแตกต่างกัน โดยออกแบบในรูปแบบ Responsive Web Application โดยมีฟังก์ชันหลัก ได้แก่ การแสดงรายการสถานที่แนะนำ การดูรายละเอียดสถานที่ และการให้คะแนนสถานที่ ดังแสดงในภาพที่ 5




**ภาพที่ 5:** หน้าจอการทำงานของระบบแนะนำ

จากภาพที่ 2 ระบบสามารถแสดงรายการสถานที่ท่องเที่ยวที่แนะนำให้กับผู้ใช้งานตามความสนใจ โดยผู้ใช้งานสามารถเลือกดูรายละเอียดของสถานที่ และให้คะแนนเพื่อใช้ในการปรับปรุงความแม่นยำของระบบแนะนำ

### 4.2 การแสดงผลแผนที่

ระบบเชื่อมต่อกับ Google Maps เพื่อแสดงตำแหน่งและเส้นทางการเดินทางไปยังสถานที่ท่องเที่ยว ดังแสดงในภาพที่ 6



**ภาพที่ 6:** การแสดงเส้นทางการเดินทางผ่านระบบแผนที่

จากภาพที่ 6 ระบบสามารถแสดงเส้นทางการเดินทางจากตำแหน่งของผู้ใช้งานไปยังสถานที่ที่เลือกได้อย่างชัดเจน ซึ่งช่วยอำนวยความสะดวกในการวางแผนการเดินทาง

### 4.3 การตั้งค่าการทดลอง

ในการทดลองนี้ ใช้ข้อมูลการให้คะแนนจากผู้ใช้งานจำนวน 120 คน และสถานที่ท่องเที่ยวจำนวน 294 แห่ง โดยมีข้อมูลการให้คะแนนรวมประมาณ 2,150 รายการ ซึ่งถูกจัดให้อยู่ในรูปแบบเมทริกซ์ผู้ใช้–สถานที่ เพื่อประเมินประสิทธิภาพของระบบ ข้อมูลถูกแบ่งออกเป็นชุดฝึก (Training Set) และชุดทดสอบ (Testing Set) โดยใช้วิธีการแบ่งข้อมูลแบบสุ่ม (Random Split) ในอัตราส่วน 80:20 กล่าวคือ 80% ของข้อมูลถูกใช้ในการฝึกโมเดล และอีก 20% ถูกใช้ในการทดสอบผลลัพธ์ของระบบ ในขั้นตอนการทดลอง ระบบใช้ข้อมูลในชุดฝึกเพื่อคำนวณความคล้ายคลึงของผู้ใช้งานด้วยวิธีแบบโคไซน์ และสร้างแบบจำลองการแนะนำ จากนั้นนำแบบจำลองที่ได้ไป

ทำนายรายการสถานที่ท่องเที่ยวสำหรับผู้ใช้งานในชุดทดสอบ

ผลลัพธ์ที่ได้จากการทำนายจะถูกนำไปเปรียบเทียบกับข้อมูลจริงในชุดทดสอบ เพื่อคำนวณค่าประเมินประสิทธิภาพ ได้แก่ Precision และ Recall ซึ่งช่วยให้สามารถประเมินความสามารถของระบบในการทำนายข้อมูลที่ไม่เคยเห็นมาก่อนได้อย่างเหมาะสม และสะท้อนประสิทธิภาพของระบบในสถานการณ์ใช้งานจริง

**4.4 การประเมินประสิทธิภาพของระบบ**

ในการประเมินประสิทธิภาพของระบบแนะนำ ใช้ตัวชี้วัดเชิงปริมาณ ได้แก่ Precision และ Recall ซึ่งใช้วัดความแม่นยำและความครอบคลุมของคำแนะนำ โดยกำหนดให้รายการที่ตรงกับความสนใจ คือสถานที่ท่องเที่ยวที่ผู้ใช้งานให้คะแนนในระดับสูงในชุดข้อมูลทดสอบ กระบวนการประเมินดำเนินการโดยให้ระบบแนะนำรายการสถานที่สำหรับผู้ใช้งานแต่ละราย และเปรียบเทียบกับรายการที่ผู้ใช้งานให้คะแนนสูง เพื่อคำนวณค่า Precision และ Recall

ผลการทดลองแสดงให้เห็นว่า วิธี Collaborative Filtering ให้ค่า Precision และ Recall สูงกว่าวิธีพื้นฐาน (Popularity-Based) ดังแสดงในตารางที่ 1 สะท้อนให้เห็นว่าระบบสามารถให้คำแนะนำที่สอดคล้องกับความสนใจของผู้ใช้งานได้ดีกว่า

**ตารางที่ 1**: การเปรียบเทียบประสิทธิภาพของวิธีการแนะนำ

| วิธีการ | Precision | Recall |
|---|---|---|
| Popularity-Based | 0.61 | 0.57 |
| Collaborative Filtering | 0.78 | 0.72 |

ข้อมูลในตารางที่ 1 แสดงให้เห็นว่า วิธี Collaborative Filtering ให้ผลลัพธ์ดีกว่าวิธีพื้นฐานทั้งในด้านความแม่นยำและความครอบคลุมของคำแนะนำ

**4.5 การวิเคราะห์ผล**

ผลการทดลองแสดงให้เห็นว่า วิธี Collaborative Filtering ให้ผลลัพธ์ดีกว่าวิธีพื้นฐาน โดยมีค่า Precision และ Recall สูงกว่า สะท้อนให้เห็นว่าระบบสามารถแนะนำสถานที่ท่องเที่ยวที่สอดคล้องกับความสนใจของผู้ใช้งานได้อย่างมีประสิทธิภาพ

ปัจจัยสำคัญมาจากความสามารถในการวิเคราะห์ความคล้ายคลึงของผู้ใช้งานจากข้อมูลการให้คะแนน ซึ่งช่วยให้สามารถนำข้อมูลของผู้ใช้งานที่มีความสนใจใกล้เคียงกันมาใช้ในการคาดการณ์ ส่งผลให้คำแนะนำมีความเป็นส่วนบุคคลมากกว่าวิธีที่อาศัยความนิยมโดยรวม

นอกจากนี้ การใช้ข้อมูลลักษณะทางประชากรศาสตร์ เช่น อายุ และเพศ ยังช่วยให้ระบบสามารถจำแนกกลุ่มผู้ใช้งานได้ชัดเจนยิ่งขึ้น โดยเฉพาะในกรณีที่ข้อมูลการให้คะแนนมีจำนวนจำกัด

อย่างไรก็ตาม ขนาดชุดข้อมูลที่ยังมีจำกัดอาจส่งผลต่อประสิทธิภาพของระบบ ขณะเดียวกัน ระบบยังสามารถลดเวลาในการค้นหาข้อมูลและสนับสนุนการวางแผนการเดินทางได้อย่างมีประสิทธิภาพ

**4.6 ข้อจำกัดของงานวิจัย**

งานวิจัยนี้มีข้อจำกัดในด้านจำนวนข้อมูลผู้ใช้งานที่ยังมีขนาดจำกัด ซึ่งอาจส่งผลต่อความแม่นยำของระบบแนะนำ นอกจากนี้ ระบบยังใช้ข้อมูลการให้คะแนนเป็นหลัก โดยยังไม่ได้พิจารณาพฤติกรรมเชิงลึกของผู้ใช้งาน เช่น ประวัติการเข้าชมหรือระยะเวลาในการใช้งาน ในอนาคตสามารถพัฒนาระบบเพิ่มเติมโดยการเพิ่มข้อมูลพฤติกรรมและประยุกต์ใช้เทคนิค Hybrid หรือ Deep Learning เพื่อเพิ่มประสิทธิภาพของระบบแนะนำ

## 5. สรุปผลและข้อเสนอแนะ

งานวิจัยนี้นำเสนอการพัฒนาระบบแนะนำการท่องเที่ยวเชิงความเชื่อในรูปแบบเว็บแอปพลิเคชัน โดยประยุกต์ใช้เทคนิค User-Based Collaborative Filtering ร่วมกับแบบโคไซน์เพื่อสร้างคำแนะนำที่เหมาะสมกับผู้ใช้งานจากข้อมูลการให้คะแนนและข้อมูลลักษณะทางประชากรศาสตร์ ระบบสามารถแสดงสถานที่ท่องเที่ยวที่สอดคล้องกับความสนใจของผู้ใช้งาน และเชื่อมต่อกับ Google Maps เพื่อสนับสนุนการเดินทาง

ผลการทดลองแสดงให้เห็นว่า วิธี Collaborative Filtering ให้ค่า Precision และ Recall สูงกว่าวิธีพื้นฐาน

อย่างชัดเจน แสดงให้เห็นถึงความสามารถของระบบในการให้คำแนะนำที่มีความแม่นยำและตรงกับความต้องการของผู้ใช้งาน ในอนาคตสามารถพัฒนาระบบเพิ่มเติมโดยการเพิ่มข้อมูลพฤติกรรมของผู้ใช้งาน และประยุกต์ใช้เทคนิคขั้นสูง เช่น Hybrid Recommendation หรือ Deep Learning เพื่อเพิ่มประสิทธิภาพของระบบ

## เอกสารอ้างอิง


[1] F. Ricci, L. Rokach, and B. Shapira, Eds., Recommender Systems Handbook. New York Heidelberg Dordrecht London: Springer, 2015.

[2] J. B. Schafer, D. Frankowski, J. Herlocker, and S. Sen, “Collaborative Filtering Recommender Systems,” in The Adaptive Web, vol. 4321, P. Brusilovsky, A. Kobsa, and W. Nejdl, Eds., in Lecture Notes in Computer Science, vol. 4321. , Berlin, Heidelberg: Springer Berlin Heidelberg, 2007, pp. 291–324. doi: 10.1007/978-3-540-72079-9_9.

[3] Nayma Khan and Dr. Mohd Haroon, “Trends and Techniques used in Tourist Recommender System : A Review,” IJSRCSEIT, pp. 33–39, May 2023, doi: 10.32628/CSEIT23902105.

[4] G. Adomavicius and A. Tuzhilin, “Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions,” IEEE Trans. Knowl. Data Eng., vol. 17, no. 6, pp. 734–749, Jun. 2005, doi: 10.1109/TKDE.2005.99.

[5] “A Survey of Collaborative Filtering Techniques - Su - 2009 - Advances in Artificial Intelligence - Wiley Online Library.” Accessed: Mar. 23, 2026. [Online]. Available: https://onlinelibrary.wiley.com/doi/10.1155/2009/421425

[6] M. Aldayel, A. Al-Nafjan, W. M. Al-Nuwaiser, G. Alrehaili, and G. Alyahya, “Collaborative Filtering-Based Recommendation Systems for Touristic Businesses, Attractions, and Destinations,” Electronics, vol. 12, no. 19, Sep. 2023, doi: 10.3390/electronics12194047.

[7] “Collaborative Filtering-Based Recommendation Systems for Touristic Businesses, Attractions, and Destinations.” Accessed: Mar. 23, 2026. [Online]. Available: https://www.mdpi.com/2079-9292/12/19/4047

[8] X. Wu, “Review of collaborative filtering recommendation systems,” Applied and Computational Engineering, vol. 43, pp. 76–82, Feb. 2024, doi: 10.54254/2755-2721/43/20230811.

[9] K. H. Lim, J. Chan, C. Leckie, and S. Karunasekera, “Personalized tour recommendation based on user interests and points of interest visit durations,” in Proceedings of the 24th International Conference on Artificial Intelligence, in IJCAI’15. Buenos Aires, Argentina: AAAI Press, Jul. 2015, pp. 1778–1784.

[10] P. Ngamprapasom, “การท่องเที่ยวเชิงศรัทธาโดยชุมชนในยุคดิจิทัล: ฉากทัศน์ในการปรับตัวสำหรับนักท่องเที่ยวสายมูเตลู,” Journal of Social Science for Local Development Rajabhat Maha Sarakham University, vol. 8, no. 4, pp. 374–383, Dec. 2024.

[11] “Knowledge Organization of Buddhist Learning Resources for Tourism: Virtual Tour of Wat Phra Pathom Chedi.” Accessed: Mar. 23, 2026. [Online]. Available: https://www.mdpi.com/2227-9709/13/1/9

[12] A. Sheta, “Intelligent Recommender Systems for Religious Tourism: Personalizing Services to Improve Experience and Economic Impact,” International Journal of Applied Science and Research, vol. 08, pp. 68–77, Jan. 2025, doi: 10.56293/IJASR.2025.6807.

[13] C. Bauer, E. Zangerle, and A. Said, “Exploring the Landscape of Recommender Systems Evaluation: Practices and Perspectives,” ACM Trans. Recomm. Syst., vol. 2, no. 1, p. 11:1-11:31, Mar. 2024, doi: 10.1145/3629170.

# ระบบบริหารจัดการสินค้าคงคลังด้วยอาร์เอฟไอดีและเทคโนโลยีคลาวด์
# Inventory Management System Using RFID and Cloud Technology

*อภิชาติ กันสีนวล (Apichart Kanseenuan) และศุภกิจ อาวิพันธุ์ (Suphakit Awiphan)*
*ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์*
*มหาวิทยาลัยเชียงใหม่*
*apichart_k@cmu.ac.th, suphakit.a@cmu.ac.th*

**บทคัดย่อ**

*การบริหารจัดการสินค้าคงคลังในปัจจุบันที่ใช้ระบบบาร์โค้ด หรือ คิวอาร์โค้ด ประสบปัญหาความล่าช้าในการทำงาน เนื่องจากต้องสแกนในระยะใกล้และต้องเห็นรหัสชัดเจน อีกทั้งรหัสที่ติดยังมีโอกาสชำรุดเสียหายได้ง่ายเมื่อเวลาผ่านไป บทความนี้นำเสนอระบบบริหารจัดการสินค้าคงคลังโดยใช้เทคโนโลยีระบุตัวรหัสสินค้าคงคลังด้วยคลื่นวิทยุ ร่วมกับเทคโนโลยีคลาวด์ เพื่อแก้ปัญหาดังกล่าว โดยใช้ อาร์เอฟไอดีแท็ก แบบ UHF และชุดอ่านที่ทำขึ้นโดยใช้ Raspberry Pi ทำให้สามารถอ่านรหัสสินค้าคงคลังได้พร้อมกันหลายชิ้นโดยไม่ต้องเห็นตัวอาร์เอฟไอดีแท็ก ข้อมูลจะถูกประมวลผลและจัดเก็บในระบบคลาวด์ ผ่านแอปพลิเคชัน ที่พัฒนาด้วยเฟรมเวิร์ก Django จากผลการทดลองแสดงให้เห็นว่าระบบอาร์เอฟไอดี มีความรวดเร็วและคุ้มค่าในการบริหารจัดการสินค้าคงคลังที่มีปริมาณมาก เมื่อเทียบกับระบบบาร์โค้ด หรือ คิวอาร์โค้ด แบบดั้งเดิม*

**คำสำคัญ:** คลาวด์เทคโนโลยี สินค้าคงคลัง อาร์เอฟไอดี

**Abstract**

*Current inventory management systems relying on Barcode or QR code technologies face operational delays due to the requirement for close-range scanning and a clear line of sight. Furthermore, physical labels are prone to damage over time. This article presents an inventory management system integrating Radio Frequency Identification (RFID) and Cloud technology to address these limitations. By utilizing UHF RFID tags and a custom-built reader powered by a Raspberry Pi, the system enables simultaneous identification of multiple items without requiring direct visibility of the tags. Data is processed and stored in the cloud via an application developed using the Django framework. Experimental results demonstrate that the RFID-based system offers superior speed and cost-effectiveness for managing high-volume inventory compared to traditional Barcode or QR code systems.*

**Keywords:** Cloud Technology, Inventory, RFID.

## 1. บทนำ

การบริหารจัดการสินค้าคงคลังในปัจจุบันมีรูปแบบในการจัดการที่หลากหลาย มีการกำหนดรหัสต่างวิธีกันไม่ว่าจะเป็นเขียนด้วยอักษร ใช้บาร์โคด หรือ คิวอาร์โคด เป็นต้น ในการระบุรหัสสินค้าคงคลังด้วย บาร์โคด หรือ คิวอาร์โคด มีปัญหาในการใช้งานคือทำงานได้ช้า ตำแหน่งที่ตั้งของรหัสไม่ได้อยู่ในตำแหน่งเดียวกันทั้งหมดทำให้การตรวจนับสินค้าคงคลังต้องใช้เวลานาน และรหัสที่ติดอาจจะมีการหลุด ชำรุด หรือเลือนได้จากการสัมผัสโดยตรง เพราะต้องติดไว้ในตำแหน่งที่เห็นได้ชัดเจน ในการศึกษาครั้งนี้จะใช้ อาร์เอฟไอดีแท็ก ในการระบุรหัสของสินค้าคงคลังเพื่อเพิ่มความเร็วในการตรวจสอบ เพราะใช้หลักการทำงานของคลื่นวิทยุ ทำให้ไม่จำเป็นต้องติดข้างนอกของตัวสินค้า ทำให้เกิดความทนทานกว่าเดิม อีกทั้งยังสามารถตรวจสอบได้ทีละหลายรหัสได้ ต่างจากบาร์โคด และ คิวอาร์โคด ที่ต้องตรวจสอบทีละรหัส

## 2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

การประยุกต์ใช้เทคโนโลยีอาร์เอฟไอดี และระบบคลาวด์เพื่อเพิ่มประสิทธิภาพในการบริหารจัดการสินค้าคงคลัง จากการศึกษางานวิจัยที่เกี่ยวข้องกับการนำเทคโนโลยีสมัยใหม่มาแก้ไขปัญหาในระบบบริหารจัดการแบบดั้งเดิม พบว่าการเปลี่ยนผ่านจากระบบบาร์โค้ด หรือคิวอาร์โค้ด มาเป็นการใช้เทคโนโลยีการระบุข้อมูลด้วยคลื่นวิทยุ หรือ อาร์เอฟไอดี ส่งผลให้มีความรวดเร็วและแม่นยำสูงขึ้น ในด้านการจัดการทรัพยากร อาร์เอฟไอดี ถูกนำมาใช้เพื่อเพิ่มประสิทธิภาพในการตรวจสอบและนับจำนวนสิ่งของ เช่น ในระบบจัดการห้องสมุดอัจฉริยะที่ช่วยลดภาระงานและระยะเวลาในการตรวจสอบหนังสือ [1] รวมถึงการพัฒนาระบบคิดเงินอัตโนมัติในร้านค้า โดยการทำงานร่วมกับเทคโนโลยีคลาวด์ (Cloud Technology) เพื่อคำนวณราคาสินค้าทันทีเมื่อนำสินค้าเข้าสู่รถเข็น ซึ่งเป็นการสร้างประสบการณ์ที่ดีและมอบความสะดวกสบายให้แก่ผู้รับบริการและผู้ให้บริการไปพร้อมกัน [2] นอกจากนี้ ยังมีการนำ อาร์เอฟไอดี มาใช้ในการระบุตำแหน่งสินทรัพย์ภายในองค์กรที่มีความซับซ้อน เช่น ศูนย์สุขภาพ (LoCATE) ซึ่งช่วยให้ผู้บริหารจัดการสามารถทราบตำแหน่งปัจจุบันของอุปกรณ์สำคัญได้แบบเรียลไทม์ ข้อมูลที่ถูกรวบรวมไว้บนระบบคลาวด์ทุกแผนกภายในองค์กรสามารถเข้าถึงและใช้ข้อมูลร่วมกันได้อย่างมีประสิทธิภาพ [3] หลักการและแนวทางจากการศึกษาวิจัยเหล่านี้ แสดงให้เห็นถึงศักยภาพของการรวมเทคโนโลยีอาร์เอฟไอดี IoT และ Cloud Computing เข้าด้วยกัน ซึ่งจะเป็นรากฐานสำคัญในการพัฒนาระบบอัตโนมัติที่ช่วยลดการใช้แรงงาน เพิ่มความแม่นยำของข้อมูล และยกระดับมาตรฐานการให้บริการในโครงการลำดับต่อไป

ในการพัฒนาระบบบริหารจัดการคลังสินค้าครั้งนี้ ได้เลือกใช้อาร์เอฟไอดีแท็ก ย่านความถี่สูงพิเศษ (UHF) เป็นการระบุตัวสินค้าแทนระบบบาร์โค้ดเดิม โดยอาศัยการทำงานร่วมกันของ RFID Module Fonkan FM-505 และอาร์เอฟไอดีแท็ก แบบ Passive Sticker (Alien 9662 U8) ที่มีความทนทานและบันทึกข้อมูลได้ยาวนาน เมื่อโมดูลส่งสัญญาณวิทยุออกไป อาร์เอฟไอดีแท็ก จะตอบสนองด้วยรหัสประจำตัวที่ไม่ซ้ำกันกลับมา ข้อมูลดังกล่าวจะถูกส่งต่อผ่านโปรโตคอล UART (Universal Asynchronous Receiver/Transmitter) เข้าสู่ Raspberry Pi 5 ซึ่งทำหน้าที่เป็นหน่วยประมวลผลหลักของ [4]

ในส่วนของซอฟต์แวร์และโครงสร้างพื้นฐาน ระบบถูกพัฒนาด้วยภาษาไพทอน โดยใช้ Django Framework ในการจัดการระบบและฐานข้อมูลผ่านเทคโนโลยี ORM เพื่อความปลอดภัยและรวดเร็วในการพัฒนา ข้อมูลการเข้า-ออกของสินค้าจะถูกส่งผ่าน RESTful API ในรูปแบบ JSON ไปยัง Cloud Technology เพื่อให้สามารถจัดเก็บและเรียกใช้งานข้อมูลได้แบบเรียลไทม์จากทุกที่ [5]

สำหรับการแสดงผลและส่วนติดต่อกับผู้ใช้งาน ได้มีการนำ Kivy ซึ่งเป็นเฟรมเวิร์กแบบ Cross-platform มาใช้พัฒนาแอปพลิเคชัน เพื่อให้รองรับการใช้งานได้หลากหลายอุปกรณ์ ทั้งเครื่องคอมพิวเตอร์และอุปกรณ์พกพาในระบบปฏิบัติการ Android และ iOS ทำให้การบริหารจัดการคลังสินค้ามีความเป็นมาตรฐานและมีประสิทธิภาพสูงขึ้นอย่างชัดเจน

## 3. วิธีการดำเนินการวิจัย

### 3.1. ศึกษาวิเคราะห์ปัญหาที่เกี่ยวข้อง

จากปัญหาการใช้งาน บาร์โคด และ คิวอาร์โคด ในการควบคุมสินค้าคงคลังมีปัญหาในเรื่องตัวรหัสเสียหาย ชำรุด และเลือนตามกาลเวลา ทำให้การตรวจสอบแต่ละครั้งใช้เวลานาน ทำให้เสียทรัพยากรที่มีไปโดยไม่จำเป็น เช่นตัวรหัสที่ต้องมีการจัดเตรียมใหม่เมื่อมีการชำรุดหรือเลือนจนอ่านค่าไม่ได้ เป็นต้น

### 3.2. ออกแบบกรอบแนวคิด

จากปัญหาที่พบเจอในระบบเดิมจึงมีความคิดว่าจะนำอาร์เอฟไอดีแท็ก ที่ทำงานด้วยสัญญาณวิทยุ มาแทน บาร์โคด และ คิวอาร์โคด ที่ทำงานด้วยหลักการแสงซึ่งต้องเห็นตัวรหัสถึงจะสามารถอ่านได้ เมื่อนำ อาร์เอฟไอดีแท็กมาใช้ก็จะทำให้สามารถเรียกข้อมูลได้โดยไม่ต้องเห็นตัวอาร์เอฟไอดีแท็ก และสามารถเรียกรหัสพร้อมกันได้หลายรหัสในคราวเดียวกันไม่เหมือนกัน บาร์โคด หรือ คิวอาร์โคด ที่ต้องเรียกทีละรหัส และนำระบบคลาวด์มาเก็บ

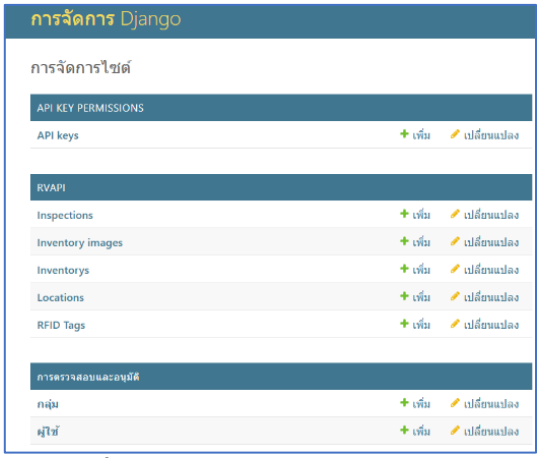
ข้อมูลทำให้สามารถใช้ระบบได้ทั้งมือถือ คอมพิวเตอร์ และ อุปกรณ์เช่นชุดอ่าน อาร์เอฟไอดีแท็ก ด้วย Raspberry Pi 5 ที่ได้จัดทำขึ้น ดังแสดงให้เห็นดังภาพที่ 1




**ภาพที่ 1:** สถาปัตยกรรมของระบบ

### 3.3. การเตรียมข้อมูล

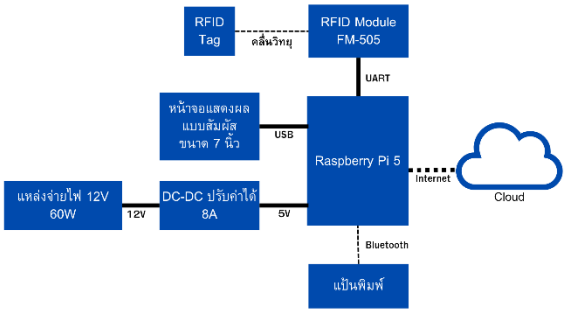
ทำการอ่านรหัสจาก อาร์เอฟไอดีแท็ก เพื่อนำมาทำบาร์โคด โดยจะใช้รหัส บาร์โคด 128 (Code 128) ซึ่งเป็นรหัสที่ระบุรหัส ASCII ได้ครบทั้ง 128 ตัว ซึ่งเพียงพอกับการนำรหัสจาก อาร์เอฟไอดีแท็ก มาใช้งาน จากนั้นทำการออกแบบฐานข้อมูลเพื่อให้เก็บข้อมูลที่จำเป็นในการบริหารสินค้าคงคลังเช่น รหัสสินค้าคงคลัง, ชื่อ, สถานที่จัดเก็บ ดังแสดงได้ตามภาพที่ 2 และทำการสร้าง API ในการเชื่อมต่อฐานข้อมูลด้วย Django โดยในการศึกษาครั้งนี้จะทำการเก็บระบบไว้กับผู้ให้บริการคลาวด์ที่ ซึ่งสามารถลดความยุ่งยากในการจัดการฮาร์ดแวร์ ประหยัดค่าใช้จ่าย สามารถเพิ่มหรือลดขนาดของบริการได้ตลอดเวลา อีกทั้งยังมีความปลอดภัยในระดับเป็นมาตรฐาน [6]



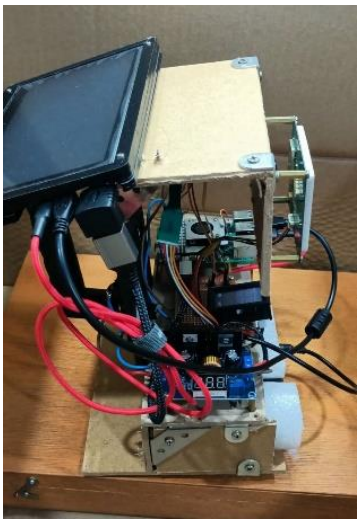

**ภาพที่ 2:** ฐานข้อมูลที่ใช้ในการจัดเก็บสินค้าคงคลัง

### 3.4. ออกแบบอุปกรณ์หลัก

ในการใช้งาน RFID Module จะต้องมีอุปกรณ์ที่อ่านค่าที่ส่งออกมาจากอุปกรณ์ ในการศึกษาครั้งนี้จะใช้ Raspberry Pi 5 เป็นอุปกรณ์หลักในการศึกษา โดยการสื่อสารจะทำการผ่านโปรโตคอล UART เป็นการสื่อสารแบบอนุกรม มีรหัสควบคุมการอ่านเฉพาะในการสั่งงานและควบคุม พัฒนาระบบด้วยภาษาไพทอน ตัวอุปกรณ์ต้นแบบใช้ Raspberry Pi 5 ในการควบคุมการอ่าน อาร์เอฟไอดีแท็ก และการติดต่อสื่อสารกับระบบคลาวด์ตาม API ที่จัดทำขึ้น ตัวอุปกรณ์มีแหล่งพลังงานจากแบตเตอรี่ 12V ที่ทำการแปลงแรงดันด้วยอุปกรณ์ปรับแรงดันให้ได้ 5V ตามที่ Raspberry Pi 5 ต้องการ และติดตั้งหน้าจอแบบสัมผัสได้ ขนาด 7 นิ้ว ไว้เพื่อแสดงผลลัพธ์การทำงานของระบบงาน ดังแสดงได้ตามภาพที่ 4 และส่วนประกอบอธิบายได้ในภาพที่ 3



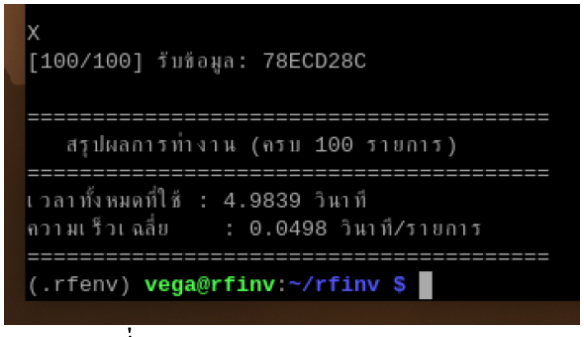

**ภาพที่ 3:** ส่วนประกอบของชุดอ่านรหัส อาร์เอฟไอดี



**ภาพที่ 4:** ต้นแบบชุดอ่านรหัส อาร์เอฟไอดี ด้วย Raspberry Pi 5

### 3.5. เขียนโปรแกรมทดสอบประสิทธิภาพ

จัดทำโปรแกรมใช้ทดสอบประสิทธิภาพการอ่านข้อมูลด้วย อาร์เอฟไอดี เทียบกับ บาร์โคด โดยจะใช้อุปกรณ์อ่านรหัส บาร์โคด แบบ USB เทียบกับอุปกรณ์ที่พัฒนาขึ้น โดยตัวโปรแกรมจะรับค่าพารามิเตอร์ 1 ตัว คือ จำนวนข้อมูลที่ต้องการทดสอบ จากนั้นทำการอ่านรหัสสินค้าคงคลังให้ได้ครบตามจำนวนที่ระบุ โปรแกรมจะแสดงผลเป็นระยะเวลาที่ใช้ และเวลาเฉลี่ยต่อการเรียกข้อมูลต่อรหัสให้ตามภาพที่ 5




**ภาพที่ 5:** ตัวอย่างผลลัพธ์จากการทดสอบอ่านรหัส

### 3.6. การประเมินประสิทธิภาพของระบบ

การประเมินผลจะใช้เวลาเฉลี่ยในการอ่านรหัสเทียบกัน โดยออกแบบการอ่านรหัสในจำนวนที่แตกต่างกัน ในการศึกษานี้จะใช้จำนวน 5, 10, 30, 50 และ 100 รหัส และทดสอบประสิทธิภาพในการอ่านในบรรจุภัณฑ์ที่ปิดตาย เช่นกล่องกระดาษแสดงในภาพที่ 6 และกล่องพลาสติกแสดงในภาพที่ 7 การหาค่าเฉลี่ยในการอ่านรหัสสินค้าคงคลัง (Average Time) สามารถคำนวณได้จาก

$$\bar{t} = \frac{\sum_{i=1}^{n} t_i}{n}$$

โดยกำหนดให้:

$\bar{t}$ คือ เวลาเฉลี่ยในการอ่าน มีหน่วยเป็นวินาที

$t_i$ คือ เวลาที่ใช้ในการทดสอบรอบที่ $i$

$n$ คือ จำนวนรอบที่ทำการทดสอบ $(n = 5)$

นอกจากนี้ การหาความเร็วในการอ่านข้อมูล (Reading Speed) สามารถคำนวณได้จาก

$$Speed = \frac{N}{\bar{t}}$$

โดยกำหนดให้:

$Speed$ คือ ความเร็วในการอ่านรหัสสินค้าคงคลังมีหน่วยเป็น (รหัส/วินาที)

$N$ คือ จำนวนรหัสสินค้าคงคลังที่นำมาทดสอบ (5, 10, 30, 50, 100 รหัส)

$\bar{t}$ คือ เวลาเฉลี่ยจากการคำนวณ



**ภาพที่ 6:** การบรรจุสินค้าลงกล่องกระดาษ



**ภาพที่ 7:** การบรรจุสินค้าคงคลังในกล่องพลาสติก

## 4. ผลการดำเนินงาน

ผลการทดลองประสิทธิภาพในการอ่านรหัสระหว่างอาร์เอฟไอดีแท็ก และ บาร์โคด ทั้ง 5 แบบแสดงได้ตามนี้

**ตารางที่ 1:** ผลการทดสอบความเร็วระหว่าง อาร์เอฟไอดีแท็ก และ บาร์โคด ทำการทดสอบ 5 ชิ้น

| ประเภท | ครั้งที่ | | | | | เฉลี่ย |
|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | |
| อาร์เอฟไอดี | 0.81 | 0.75 | 0.58 | 0.63 | 0.92 | **0.74** |
| บาร์โคด | 9.34 | 7.24 | 8.62 | 8.68 | 8.34 | **8.37** |

**ตารางที่ 2:** ผลการทดสอบความเร็วระหว่าง อาร์เอฟไอดีแท็ก และ บาร์โคด ทำการทดสอบ 10 ชิ้น

| ประเภท | ครั้งที่ | | | | | เฉลี่ย |
|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | |
| อาร์เอฟไอดี | 0.86 | 1.21 | 0.73 | 0.92 | 1.53 | **1.05** |
| บาร์โคด | 14.26 | 13.37 | 12.68 | 13.49 | 14.35 | **13.63** |

**ตารางที่ 3:** ผลการทดสอบความเร็วระหว่าง อาร์เอฟไอดีแท็ก และ บาร์โคด ทำการทดสอบ 30 ชิ้น

| ประเภท | ครั้งที่ | | | | | เฉลี่ย |
|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | |
| อาร์เอฟไอดี | 5.24 | 4.69 | 6.62 | 5.42 | 5.65 | **5.52** |
| บาร์โคด | 43.28 | 39.42 | 42.36 | 37.49 | 38.24 | **40.16** |

**ตารางที่ 4:** ผลการทดสอบความเร็วระหว่าง อาร์เอฟไอดีแท็ก และ บาร์โคด ทำการทดสอบ 50 ชิ้น

| ประเภท | ครั้งที่ | | | | | เฉลี่ย |
|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | |
| อาร์เอฟไอดี | 7.43 | 8.94 | 9.71 | 8.65 | 8.21 | **8.59** |
| บาร์โคด | 72.43 | 82.38 | 76.46 | 74.85 | 78.69 | **76.96** |

**ตารางที่ 5:** ผลการทดสอบความเร็วระหว่าง อาร์เอฟไอดีแท็ก และ บาร์โคด ทำการทดสอบ 100 ชิ้น

| ประเภท | ครั้งที่ | | | | | เฉลี่ย |
|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | |
| อาร์เอฟไอดี | 26.42 | 25.43 | 22.46 | 23.34 | 22.94 | **24.12** |
| บาร์โคด | 142.35 | 139.24 | 146.49 | 143.49 | 141.92 | **142.70** |

จากการทดสอบประสิทธิภาพความเร็วในการอ่านข้อมูลระหว่าง อาร์เอฟไอดี และ บาร์โคด ในปริมาณสินค้าที่แตกต่างกัน ตั้งแต่ 5 ชิ้น ไปจนถึง 100 ชิ้น พบประเด็นสำคัญดังนี้ ความรวดเร็วที่มากกว่าอย่างเห็นได้ชัด ในทุกจำนวนที่ทดสอบ อาร์เอฟไอดี ทำเวลาได้น้อยกว่าบาร์โคดโดยเฉลี่ยแล้ว ประมาณ 6-12 เท่า

ประสิทธิภาพเมื่อปริมาณสินค้าเพิ่มขึ้นที่ 5-10 ชิ้น ความแตกต่างอยู่ในระดับหลักวินาที (อาร์เอฟไอดี ใช้เวลาประมาณ 0.75 – 0.92 วินาที ในขณะที่บาร์โคดใช้ถึง 7.24 – 9.34 วินาที) ที่ 100 ชิ้น (จุดที่เห็นผลชัดที่สุด): อาร์เอฟไอดี ใช้เวลาเฉลี่ยเพียง 24.12 วินาที ในขณะที่บาร์โค้ดต้องใช้เวลาสูงถึง 142.70 วินาที

ในด้านอัตราการเติบโตของเวลาบาร์โคดมีแนวโน้มการใช้เวลาที่เพิ่มขึ้นแบบเส้นตรง (Linear) ตามจำนวนชิ้น เพราะต้องสแกนทีละชิ้น แต่ อาร์เอฟไอดี แสดงให้เห็นถึงความสามารถในการอ่านแบบกลุ่ม (Bulk Reading) ที่มีประสิทธิภาพมากกว่า แม้จำนวนชิ้นจะมากขึ้น แต่เวลาที่เพิ่มขึ้นไม่ได้สูงเท่าบาร์โคด

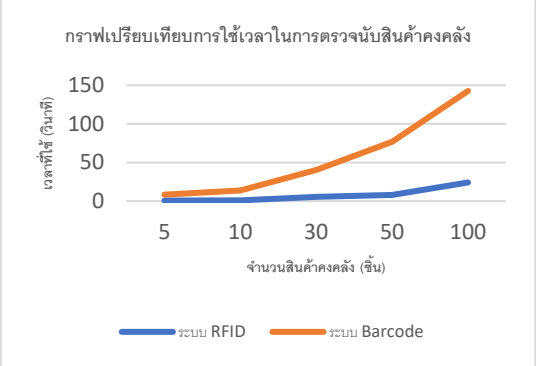
ในส่วนของความแม่นยำในการอ่านข้อมูล ได้ใช้เทคนิคการอ่านซ้ำแบบต่อเนื่อง (Continuous Sampling) เนื่องจากโมดูล อาร์เอฟไอดี มีความเร็วในการอ่านสูงมากกว่า 50 ครั้งต่อวินาที ระบบจึงถูกออกแบบให้ทำการตรวจสอบความถูกต้องของรหัสที่อ่านได้หลายครั้งภายในเสี้ยววินาที เพื่อยืนยันความถูกต้องของข้อมูลก่อนส่งไปยังฐานข้อมูล จากการทดสอบในสภาวะควบคุมพบว่าเทคนิคการอ่านซ้ำดังกล่าวช่วยลดอัตราการอ่านพลาดให้เป็นศูนย์ ส่งผลให้ระบบมีความแม่นยำในการระบุตัวตนสินค้าครบถ้วนทุกรายการ ตลอดการทดสอบทั้ง 5 ชุด ทั้งนี้ ความแม่นยำดังกล่าวขึ้นอยู่กับมุมในการอ่าน และการเรียงสินค้า ซึ่งในการทดลองนี้ได้เลือกใช้ตำแหน่งที่เหมาะสมที่สุด เมื่อเทียบกับระบบที่ใช้บาร์โคด มีความเสี่ยงที่เกิดจากความผิดพลาดของมนุษย์ เช่น การข้ามลำดับหรือการสแกนซ้ำซ้อน แต่ระบบ อาร์เอฟไอดี ที่พัฒนาขึ้นนี้ใช้เทคนิคการกรองข้อมูลร่วมกับการอ่านซ้ำด้วยความเร็วสูง ทำให้สามารถตัดโอกาสการเกิดข้อผิดพลาดดังกล่าวได้

เนื่องจากหลักการทำงานของอาร์เอฟไอดี คือใช้คลื่นวิทยุไม่ได้ใช้แสงในการอ่านรหัสเหมือนกับ บาร์โคด และ คิวอาร์โคด เมื่อทำการเปลี่ยนบรรจุภัณฑ์ต่างชนิดกันระหว่างกล่องกระดาษและกล่องพลาสติก ได้ผลลัพธ์ที่ใกล้เคียงกันทั้ง 2 ประเภท ดังตารางที่ 6

**ตารางที่ 6:** ความเร็วในการอ่านรหัสสินค้าคงคลังเมื่ออยู่ในบรรจุภัณฑ์ 2 ประเภทได้แก่กล่องกระดาษ และกล่องพลาสติก มีหน่วยเป็นวินาที

| ประเภท | ครั้งที่ | | | | | เฉลี่ย |
|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | |
| กล่องกระดาษ | 0.92 | 0.71 | 0.76 | 0.85 | 0.79 | **0.81** |
| กล่องพลาสติก | 0.84 | 0.62 | 0.78 | 0.87 | 0.68 | **0.76** |

ถึงแม้ว่าการนำระบบ อาร์เอฟไอดี มาประยุกต์ใช้งานจะมีต้นทุนด้านอุปกรณ์เริ่มต้นที่สูงกว่าระบบบาร์โคด และคิวอาร์โคด แต่เมื่อพิจารณาเวลาในการใช้งานที่ลดลงและความแม่นยำในการอ่านรหัสจำนวนมาก และไม่ต้องสิ้นเปลืองกับการเสียเวลาและทรัพยากรไปในกรณีที่รหัสมีการเสียหาย ชำรุด และเลือนไปในเวลาที่เร็วกว่า อาร์เอฟไอดี Tag ระบบ อาร์เอฟไอดี จึงให้ความคุ้มค่าที่สูงมากกว่าดังที่แสดงในภาพที่ 8




**ภาพที่ 8:** แสดงการเปรียบเทียบเวลาที่ใช้ในการตรวจนับสินค้าคงคลัง ระหว่าง อาร์เอฟไอดีแท็ก และ บาร์โคด

## 5. สรุปผล

การพัฒนาระบบบริหารจัดการสินค้าคงคลังด้วย อาร์เอฟไอดี และเทคโนโลยีคลาวด์ ช่วยแก้ปัญหาความล่าช้าและความยุ่งยากของระบบบาร์โคด และ คิวอาร์โคด ในระบบดั้งเดิมได้อย่างมีประสิทธิภาพ จากผลการทดลองยืนยันว่า การนำอาร์เอฟไอดีแท็กมาใช้แทนบาร์โคดช่วยลดเวลาในการตรวจสอบสินค้าคงคลังได้อย่างมีนัยสำคัญ โดยเฉพาะในการบริหารจัดการสินค้าจำนวนมาก อาร์เอฟไอดีแท็กใช้เวลาน้อยกว่าบาร์โคดอย่างเห็นได้ชัดเจน ช่วยในการประหยัดเวลาในการทำงานและลดการสูญเสียของทรัพยากรไปกับรหัสสินค้าที่เสียหาย ชำรุด และ ลบเลือน

อีกทั้งการใช้ระบบคลาวด์ยังช่วยให้ระบบข้อมูลมีความปลอดภัยและเข้าถึงได้ง่าย ลดภาระค่าใช้จ่ายในการดูแลเครื่องแม่ข่าย และการใช้พลังงานไฟฟ้าภายในองค์กรได้อีกด้วย แต่ยังมีข้อจำกัดทางกายภาพที่ต้องคำนึงถึงในการใช้ อาร์เอฟไอดี ในการบริหารสินค้าคงคลัง เนื่องจากลักษณะการทำงานที่เป็นคลื่นวิทยุ เมื่ออยู่ใกล้วัตถุที่เป็นโลหะหรือมีของเหลว จะทำให้ประสิทธิภาพในการสื่อสารระหว่างชุดอ่านและตัว อาร์เอฟไอดีแท็ก ลดลงได้ และการจัดเรียงสินค้าในบรรจุภัณฑ์ ต้องมีความเหมาะสม ไม่ชิดกันแน่นจนทำให้เกิดการรบกวนสัญญาณระหว่างอาร์เอฟไอดีแท็กแต่ละดวงจนเกินไป

## เอกสารอ้างอิง

[1] R. Neha Mukund, S. Arun, S. M. Kusuma, K. Vishak and M. Monisha, “Intelligent RFID Based Library Management System,” *IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT)*, pp. 01-06, 2021.

[2] D. Sinha, K. Cottur, K. B. H., G. C. and B. N. K., “Automated Billing System using RFID and Cloud,” *Innovations in Power and Advanced Computing Technologies (i-PACT)*, pp. 1-6, 2019.

[3] T. D. McAllister, S. El-Tawab and M. H. Heydari, “Localization of Health Center Assets Through an IoT Environment (LoCATE),” *Systems and Information Engineering Design Symposium (SIEDS)*, pp. 132-137, 2017.

[4] D. M. Dobkin, *The RF in RFID Passive UHF RFID in practice*, Elsevier Inc., 2008.

[5] J. Gale, *Raspberry Pi THE COMPLETE GUIDE*, Black Dog Media Limited, 2020.

[6] D. Martinez, D. Gonzalez, J. M. Sánchez and R. M. Toasa, “Library in Django framework to standardize early-stage web application development,” *Iberian Conference on Information Systems and Technologies (CISTI)*, pp. 1-4, 2013.

[7] A. P. a. others, *AWS Cloud Computing Concepts and Tech Analogies*, Packt Publishing, 2023.

[8] M. A. I. M. Azri, H. Sofian, M. S. M. Kassim and R. Ramly, "Smart Inventory Management System for Hardware Warehouse Using IoT," 2026 20th International Conference on Ubiquitous Information Management and Communication (IMCOM), pp. 1-6, 2026.

[9] F. Raheel, L. Raheel, H. Mehmood, M. B. Kadri and U. S. Khan, "Internet of Robots: An Open-source Framework for Cloud-Based Global Remote Control and Monitoring of Robots," 2024 International Conference on Robotics and Automation in Industry (ICRAI), pp. 1-6, 2024.

# ปัจจัยที่มีอิทธิพลต่อการยอมรับและการตัดสินใจใช้งานระบบ Chatbot ของผู้บริโภค
# Factors Influencing Consumer Acceptance and Decision to Use Chatbot Systems

อานันท์ มะลิชัย *(Anan Malichai)*[1] และมณเฑียร รัตนศิริวงศ์วุฒิ *(Montean Rattanasiriwongwut)*[2]
*ภาควิชาการจัดการเทคโนโลยีสารสนเทศ*
*คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ*
[1]*s6707021857091@email.kmutnb.ac.th*, [2] *montean.r@itd.kmutnb.ac.th*

## บทคัดย่อ

*การศึกษานี้รวบรวมข้อมูลจากผู้บริโภคในกรุงเทพมหานครที่เคยใช้งานหรือรู้จักระบบ Chatbot จำนวน 402 คน ด้วยแบบสอบถาม และวิเคราะห์ข้อมูลโดยใช้สถิติเชิงพรรณนา การวิเคราะห์องค์ประกอบเชิงสำรวจ (EFA) และโมเดลสมการโครงสร้าง (SEM) ผลการวิจัยพบว่า ผู้บริโภคให้ความสำคัญกับปัจจัยการยอมรับระบบ Chatbot ในระดับมาก โดยผลวิเคราะห์ EFA สามารถสกัดปัจจัยได้ 3 องค์ประกอบหลัก ได้แก่ 1) ความสะดวกและทัศนคติต่อการใช้งาน 2) ความปลอดภัยและความเชื่อมั่น และ 3) คุณภาพข้อมูลและความถูกต้องผลวิเคราะห์ SEM พบว่าโมเดลมีความสอดคล้องกลมกลืนกับข้อมูลเชิงประจักษ์อย่างสมบูรณ์ (Chi-square = 191.000, df = 167, p-value = 0.095, (Chi-square = 1.143, CFI = 0.999, TLI = 0.999, RMSEA = 0.019 และ SRMR = 0.048) โดยตัวบ่งชี้ทั้งหมดมีค่าน้ำหนักองค์ประกอบมาตรฐานระดับสูงและมีนัยสำคัญทางสถิติ (p < .001) สนับสนุนสมมติฐานทุกประการ ข้อค้นพบนี้เป็นแนวทางให้องค์กรธุรกิจนำไปพัฒนาอินเทอร์เฟซ นโยบายความปลอดภัยและฐานข้อมูลของ Chatbot เพื่อสร้างความไว้วางใจและส่งเสริมการตัดสินใจใช้งานอย่างยั่งยืน*

**คำสำคัญ:** การยอมรับเทคโนโลยี,ระบบแชทบอท,การวิเคราะห์องค์ประกอบ,โมเดลสมการโครงสร้างความไว้วางใจของผู้บริโภค

## Abstract

*This study collected data via a questionnaire from 402 consumers in Bangkok who have used or are familiar with chatbot systems. The data were analyzed using descriptive statistics, exploratory factor analysis (EFA), and structural equation modeling (SEM). The results revealed that consumers prioritized factors influencing chatbot acceptance at a high level. The EFA successfully extracted three main components: 1) convenience and attitude toward usage, 2) security and trust, and 3) information quality and accuracy. The SEM analysis indicated that the developed model perfectly fit the empirical data (Chi-square = 191.000, df = 167, p-value = 0.095, Chi-square = 1.143, CFI = 0.999, TLI = 0.999, RMSEA = 0.019, and SRMR = 0.048). All indicators had high standardized factor loadings and were statistically significant (p < .001), thus supporting all research hypotheses. These findings provide practical guidelines for business organizations to improve chatbot interfaces, security policies, and databases to foster consumer trust and promote sustainable usage decisions.*

**Keywords**: Technology Adoption, Chatbot, Factor Analysis, Structural Equation Modeling (SEM), Consumer Trust.

## 1. บทนำ

ในยุคดิจิทัล เทคโนโลยีปัญญาประดิษฐ์ (AI) โดยเฉพาะระบบแชทบอต (Chatbot) ได้เข้ามามีบทบาทสำคัญในการเป็นช่องทางหลักสำหรับธุรกิจในการสื่อสารและให้บริการลูกค้าตลอด 24 ชั่วโมง อย่างไรก็ตาม แม้เทคโนโลยีจะก้าวหน้า แต่ยังพบข้อจำกัดในการยอมรับจากผู้บริโภคบางส่วน ที่มักพบปัญหาความไม่เสถียรของ

ระบบ ข้อมูลที่ได้รับไม่แม่นยำ หรือความกังวลด้านความปลอดภัยของข้อมูลส่วนบุคคล ซึ่งสิ่งเหล่านี้ล้วนเป็นอุปสรรคสำคัญที่ทำให้ผู้ใช้งานเกิดความไม่ไว้วางใจและปฏิเสธการใช้งานในที่สุด [4], [5]

ด้วยเหตุนี้ ผู้วิจัยจึงเล็งเห็นความจำเป็นในการศึกษาปัจจัยเชิงลึกที่มีอิทธิพลต่อการยอมรับและการตัดสินใจใช้งานระบบ Chatbot ของผู้บริโภคในเขตกรุงเทพมหานคร เพื่อนำผลวิเคราะห์ที่ได้ไปเป็นแนวทางให้องค์กรธุรกิจสามารถออกแบบและพัฒนาระบบที่ตอบสนองความต้องการและสร้างความเชื่อมั่นให้แก่ผู้บริโภคได้อย่างแท้จริง

## 2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ผู้วิจัยศึกษาเรื่องปัจจัยที่มีอิทธิพลต่อการยอมรับและการตัดสินใจใช้งานระบบ Chatbot ผู้วิจัยได้ทำการทบทวนวรรณกรรมและทฤษฎีที่เกี่ยวข้องเพื่อนำมาสร้างเป็นกรอบแนวคิดในการวิจัย โดยครอบคลุมเนื้อหาหลัก 3 ส่วน ดังนี้

### 2.1 แนวคิดและทฤษฎีเกี่ยวกับระบบแชทบอต (Chatbot)

ระบบสนทนาอัตโนมัติหรือแชทบอต (Chatbot) คือโปรแกรมคอมพิวเตอร์ที่ถูกพัฒนาขึ้นเพื่อทำหน้าที่ปฏิสัมพันธ์และตอบโต้กับมนุษย์ผ่านตัวอักษรหรือเสียง โดยอาศัยเทคโนโลยีปัญญาประดิษฐ์ (Artificial Intelligence: AI) และกระบวนการประมวลผลภาษาธรรมชาติ (Natural Language Processing: NLP) เพื่อให้ระบบสามารถเข้าใจความหมายและบริบทของคำพูดมนุษย์ได้ ในมิติของการดำเนินธุรกิจ แชทบอตถูกนำมาใช้เป็นตัวแทนในการให้บริการลูกค้า (Customer Service Automation) ที่สามารถทำงานได้ตลอด 24 ชั่วโมง ช่วยลดระยะเวลาในการรอคอย และเพิ่มขีดความสามารถในการตอบสนองความต้องการของผู้บริโภคในยุคดิจิทัลที่ต้องการความรวดเร็วและความแม่นยำสูง

### 2.2 แนวคิดเกี่ยวกับปัจจัยที่มีอิทธิพลต่อการยอมรับและการตัดสินใจใช้งาน

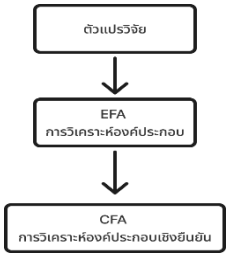
จากการทบทวนวรรณกรรม พบว่าปัจจัยเชิงจิตวิทยาที่ส่งผลต่อการยอมรับเทคโนโลยีประกอบด้วย 4 ประการ ได้แก่ 1) การรับรู้ความง่ายในการใช้งาน (Perceived Ease of Use) คือความรู้สึกว่าระบบไม่ซับซ้อนและใช้งานง่าย 2) การรับรู้ประโยชน์ (Perceived Usefulness) คือการตระหนักว่าแชทบอตช่วยเพิ่มประสิทธิภาพและอำนวยความสะดวกได้จริง 3) ความไว้วางใจในเทคโนโลยี (Trust in Technology) คือความเชื่อมั่นในความเที่ยงตรงและการรักษาความปลอดภัยของข้อมูลส่วนบุคคล และ 4) การรับรู้ความเสี่ยง (Perceived Risk) คือความกังวลต่อความผิดพลาดของระบบหรือภัยคุกคามทางไซเบอร์ ซึ่งเป็นข้อจำกัดที่อาจลดความตั้งใจในการใช้งานของผู้บริโภค [3], [4]

### 2.3 งานวิจัยที่เกี่ยวข้อง

ผู้วิจัยได้ทบทวนงานวิจัยที่เกี่ยวข้องทั้งในและต่างประเทศ เช่น งานวิจัยของ Wu และคณะ (2024) [1] ที่ประยุกต์ใช้ SEM ในการศึกษาความพึงพอใจของผู้ใช้ในระบบแนะนำ และงานของ Yang (2021) [2] ที่ชี้ให้เห็นว่าความไว้วางใจและการรับรู้คุณค่ามีอิทธิพลโดยตรงต่อความตั้งใจซื้อในเชิงพาณิชย์ดิจิทัล นอกจากนี้ งานวิจัยของ Wu & Huang (2023) ยังยืนยันว่าความเชื่อมั่นเป็นปัจจัยสื่อกลาง (Mediator) ที่ส่งผลต่อการตัดสินใจใช้งานเทคโนโลยีอย่างต่อเนื่อง ซึ่งสอดคล้องกับโครงสร้างปัจจัยที่ผู้วิจัยนำมาศึกษาในครั้งนี้

## 3. วิธีการดำเนินการวิจัย

การศึกษานี้ใช้วิธีวิจัยเชิงปริมาณเพื่อวิเคราะห์ปัจจัยที่มีอิทธิพลต่อการยอมรับระบบ Chatbot โดยผู้วิจัยได้กำหนดลำดับขั้นตอนการดำเนินงานเริ่มจากการคัดเลือกตัวแปรวิจัย การวิเคราะห์องค์ประกอบเชิงสำรวจ (EFA) เพื่อจัดกลุ่มตัวแปร และสรุปผลด้วยการวิเคราะห์องค์ประกอบเชิงยืนยัน (CFA) ดังแสดงใน ภาพที่ 1




**ภาพที่ 1** : กรอบแนวทางการดำเนินงานวิจัย

**3.1 การกำหนดกรอบแนวคิด (Conceptual Framework)**

การศึกษานี้ใช้รูปแบบการวิจัยเชิงสำรวจ (Survey Research) โดยใช้แบบสอบถาม (Questionnaire) เป็นเครื่องมือหลักในการเก็บรวบรวมข้อมูลจากกลุ่มตัวอย่าง เพื่อนำมาดำเนินการวิเคราะห์ทางสถิติ 2 ขั้นตอน ได้แก่

- การวิเคราะห์องค์ประกอบเชิงสำรวจ (Exploratory Factor Analysis: EFA) เพื่อค้นหาและจัดกลุ่มโครงสร้างของตัวแปร

- การวิเคราะห์โมเดลสมการโครงสร้าง (Structural Equation Modeling: SEM) โดยกำหนดตัวแปรที่ศึกษาดังนี้:

- ตัวแปรอิสระ: การรับรู้ความง่ายในการใช้งาน, การรับรู้ประโยชน์, ความไว้วางใจในระบบ และการรับรู้ความเสี่ยง

- ตัวแปรตาม: ความตั้งใจใช้งานและการตัดสินใจใช้งานระบบ Chatbot

**3.2 ประชากรและกลุ่มตัวอย่าง**

- ประชากร (Population): ผู้บริโภคที่พักอาศัยอยู่ในเขตกรุงเทพมหานคร และเคยใช้บริการหรือมีแนวโน้มที่จะใช้บริการผ่านระบบ Chatbot

- กลุ่มตัวอย่าง (Sample) และการกำหนดขนาด: กลุ่มตัวอย่างที่ใช้ในการวิจัยครั้งนี้มีจำนวนทั้งสิ้น 402 คน โดยอ้างอิงหลักการกำหนดขนาดตัวอย่างของ Hair et al. (2010) ที่แนะนำสัดส่วนจำนวนตัวอย่างต่อข้อคำถาม (Item-to-Response Ratio) ที่ระดับ 10:1 ถึง 20:1 เนื่องจากการศึกษานี้มีข้อคำถาม 20 ข้อ ผู้วิจัยจึงกำหนดสัดส่วนที่ 20 เท่า ทำให้ได้ขนาดกลุ่มตัวอย่างที่เพียงพอต่อการวิเคราะห์โมเดลสมการโครงสร้าง (SEM) ให้มีความเสถียร

- วิธีการสุ่มตัวอย่าง: ใช้วิธีการสุ่มตัวอย่างแบบสะดวก

**3.3 เครื่องมือที่ใช้ในการวิจัยและการตรวจสอบคุณภาพเครื่องมือ**

- เครื่องมือที่ใช้: แบบสอบถามออนไลน์ (Online Questionnaire) ที่จัดทำผ่าน Google Forms

- การตรวจสอบความตรงเชิงเนื้อหา (Content Validity): ผู้วิจัยได้นำแบบสอบถามไปให้ผู้เชี่ยวชาญตรวจสอบเพื่อหาค่าดัชนีความสอดคล้อง (Index of Item-Objective Congruence: IOC) ผลการพิจารณาพบว่าแบบสอบถามทั้ง 24 ข้อ มีค่า IOC มากกว่า 0.5 แสดงว่าข้อคำถามวัดได้ตรงจุดประสงค์และครอบคลุมเนื้อหาสามารถนำข้อคำถามทั้งหมดมาใช้ในการเก็บรวบรวมข้อมูลได้จริง

**3.4 การเก็บรวบรวมข้อมูลและระยะเวลาดำเนินการ**

ในการวิจัยครั้งนี้ ผู้วิจัยได้ดำเนินการเก็บรวบรวมข้อมูลเชิงปริมาณ (Quantitative Data) โดยใช้ข้อมูลปฐมภูมิ (Primary Data) จากการตอบแบบสอบถามของกลุ่มตัวอย่าง ดังนี้:

- การเตรียมความพร้อมและการกระจายเครื่องมือ: ผู้วิจัยนำแบบสอบถามที่ผ่านการตรวจสอบคุณภาพและความเที่ยงตรง (Validity and Reliability) [1], [3] เรียบร้อยแล้ว ไปเผยแพร่ผ่านช่องทางออนไลน์ เพื่อความสะดวกและรวดเร็วในการกระจายข้อมูล โดยในส่วนแรกของแบบสอบถามได้ระบุวัตถุประสงค์ของการวิจัย และคำชี้แจงไว้อย่างชัดเจน

- ระยะเวลาการดำเนินงาน: เก็บข้อมูลตั้งแต่วันที่ 1 กันยายน 2568 ถึง 31 มกราคม 2569

**3.5 การวิเคราะห์ข้อมูลและสถิติที่ใช้**

ผู้วิจัยดำเนินการวิเคราะห์ข้อมูลผ่านโปรแกรม Jamovi โดยใช้สถิติ ดังต่อไปนี้:

- สถิติเชิงพรรณนา (Descriptive Statistics): ใช้ค่าความถี่ (Frequency) และค่าร้อยละ (Percentage) สำหรับข้อมูลส่วนบุคคล และใช้ค่าเฉลี่ย (Mean) คู่กับส่วนเบี่ยงเบนมาตรฐาน (S.D.) เพื่อระบุระดับความคิดเห็นต่อตัวแปรต่าง ๆ

- การวิเคราะห์องค์ประกอบเชิงสำรวจ (Exploratory Factor Analysis: EFA):

- การตรวจสอบความเหมาะสมของข้อมูล: พิจารณาจากค่า KMO (Kaiser-Meyer-Olkin) ซึ่งต้องมีค่ามากกว่า 0.50 และทดสอบ Bartlett's Test of Sphericity ซึ่งต้องมีนัยสำคัญทางสถิติ $(p < .05)$ เพื่อยืนยันว่าข้อมูลมีความเหมาะสมในการวิเคราะห์ปัจจัย

- การสกัดและหมุนแกนองค์ประกอบ: ใช้วิธีองค์ประกอบหลัก (Principal Component Analysis: PCA) ในการสกัดตัวแปรแฝง และหมุนแกนแบบมุมฉากด้วยวิธี

Varimax เพื่อให้โครงสร้างของแต่ละองค์ประกอบมีความชัดเจนตามค่าน้ำหนักองค์ประกอบ (Factor Loading)

**3.6 การวิเคราะห์โมเดลสมการโครงสร้าง (Structural Equation Modeling: SEM):**

นำโครงสร้างปัจจัยที่ได้จากการวิเคราะห์ EFA มาทดสอบความสัมพันธ์เชิงสาเหตุ เพื่อตรวจสอบว่าโมเดลความสัมพันธ์ที่สร้างขึ้นมีความสอดคล้องกับข้อมูลเชิงประจักษ์หรือไม่ โดยทดสอบสมมติฐานหลักคือ

- การประเมินดัชนีความกลมกลืนของโมเดล (Model Fit Indices): พิจารณาจากเกณฑ์ดังนี้

- ค่าสถิติไค-สแควร์สัมพัทธ์ ($\chi^2/df$) : ควรมีค่าน้อยกว่า 3.00

- ดัชนีวัดระดับความกลมกลืน (CFI และ TLI): ควรมีค่ามากกว่า 0.90 หรือ 0.95

- ดัชนีรากของค่าเฉลี่ยกำลังสองของส่วนเบี่ยงเบน (RMSEA และ SRMR): ควรมีค่าน้อยกว่า 0.05 หรือ 0.08

- ค่าสัมประสิทธิ์การพยากรณ์ ($R^2$): เพื่อประเมินความสามารถในการอธิบายความแปรปรวนของตัวแปรตามในโมเดล

## 4. ผลการดำเนินงาน

**4.1 ผลการวิเคราะห์ข้อมูลทั่วไปของผู้ตอบแบบสอบถาม**

จากการสำรวจกลุ่มตัวอย่างจำนวน 402 คน พบว่าลักษณะประชากรศาสตร์ของกลุ่มเป็นเพศหญิง 238 คน คิดเป็นร้อยละ 59.20 ลักษณะประชากรศาสตร์ของกลุ่มเป็นเพศชาย 157 คน คิดเป็นร้อยละ 39.05 มีอายุระหว่าง 31-40 ปี 172 คน คิดเป็นร้อยละ 42.79 อายุระหว่าง 41-50 ปี 87 คน คิดเป็นร้อยละ 21.64 และอายุระหว่าง 21-30 ปี 86 คน คิดเป็นร้อยละ 21.39 ซึ่งมีระดับการศึกษาปริญญาตรี 174 คน คิดเป็นร้อยละ 43.28 ระดับสูงกว่าปริญญาตรี 163 คน คิดเป็นร้อยละ 40.55 โดยประกอบอาชีพพนักงานเอกชน 209 คน คิดเป็นร้อยละ 51.99 อาชีพข้าราชการ/พนักงานรัฐวิสาหกิจ 61 คน คิดเป็นร้อยละ 15.17 และเจ้าของกิจการ/ผู้ประกอบอาชีพอิสระ 55 คน คิดเป็นร้อยละ 13.68

**4.2 ผลการวิเคราะห์ระดับความคิดเห็นต่อปัจจัยที่มีผลต่อการยอมรับและการใช้งาน** Chatbot

จากการวิเคราะห์ระดับความคิดเห็นของผู้ตอบแบบสอบถามจำนวน 402 คน พบว่าในภาพรวมผู้บริโภคมีความคิดเห็นต่อปัจจัยที่ส่งผลต่อการยอมรับอยู่ในระดับมาก ($\bar{x}$ = 3.67, **S.D.** = 1.01) โดยเมื่อพิจารณารายด้านตามที่ปรากฏใน ตารางที่ 1 พบว่าด้านความสะดวกในการใช้งาน (**ACE**) มีค่าเฉลี่ยสูงสุด ($\bar{x}$ = 3.78) ตามด้วยทัศนคติต่อการใช้งาน (**OAAI**)

**ตารางที่ 1 :** ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของระดับความคิดเห็น

| คำถาม | $\bar{x}$ | S.D. | การแปลค่า |
|---|---|---|---|
| (ACE) | 3.78 | 0.96 | มาก |
| (OAAI) | 3.70 | 1.04 | มาก |
| (RPS) | 3.66 | 1.00 | ปานกลาง |
| (DQA) | 3.63 | 0.94 | ปานกลาง |
| (PST) | 3.58 | 1.09 | ปานกลาง |

**4.3 ผลการวิเคราะห์องค์ประกอบเชิงสำรวจ (Exploratory Factor Analysis: EFA)**

ผู้วิจัยได้วิเคราะห์องค์ประกอบเชิงสำรวจ (EFA) ด้วยเทคนิคสกัดปัจจัยแบบ Maximum Likelihood (ML) และหมุนแกนด้วยวิธี Varimax โดยก่อนการสกัดปัจจัยได้ทดสอบความเหมาะสมของข้อมูลด้วยสถิติ KMO และ Bartlett's Test of Sphericity [1], [2] ตามสมการดังนี้

$$KMO = \frac{\sum\sum_{i \neq j} r_{ij}^2}{\sum\sum_{i \neq j} r_{ij}^2 + \sum\sum_{i \neq j} a_{ij}^2}$$

$$\chi^2 = -\left(N - 1 - \frac{2p + 5}{6}\right) \ln|R|$$

ผลการทดสอบดังตารางที่ 2 พบว่า ค่า KMO เท่ากับ 0.921 ซึ่งอยู่ในระดับที่เหมาะสมอย่างยิ่ง และการทดสอบ Bartlett's Test มีค่านัยสำคัญทางสถิติ $(p > 0.01)$แสดงให้เห็นว่าตัวแปรทั้ง 20 ข้อมีความสัมพันธ์กันเพียงพอที่จะนำมาสกัดเป็นองค์ประกอบได้อย่างเหมาะสมดัง

**ตารางที่ 2 :** ค่า KMO และ Bartlett's Test of Sphericity

| KMO and Bartlett's Test | | |
|---|---|---|
| Kaiser-Meyer-Olkin Measure of Sampling Adequacy. | | 0.921 |
| Bartlett's Test of Sphericity Approx. Chi-Square | | 8010 |
| | DF | 190 |
| | Sig. | < .001 |

ในการวิเคราะห์องค์ประกอบเชิงสำรวจ (EFA) ผู้วิจัยได้กำหนดสมการเพื่อแสดงความสัมพันธ์ระหว่างตัวแปรสังเกตได้และองค์ประกอบร่วม ดังนี้

$$X_i = \lambda_{i1}F_1 + \lambda_{i2}F_2 + \lambda_{i3}F_3 + \varepsilon_i$$

ผลการหมุนแกนด้วยวิธี Varimax (Rotated Factor Matrix) ผู้วิจัยได้ปรับลดองค์ประกอบให้เหลือ 3 กลุ่มหลัก โดยตัดตัวแปรที่ไม่ผ่านเกณฑ์ออก 4 ข้อ (PST16, RPS12, OAAI20, RPS13) เนื่องจากมีค่า Factor Loading ต่ำกว่า 0.50 และมีการกระจายตัวข้ามกลุ่ม ส่งผลให้เหลือตัวแปร 20 ข้อ ที่มีค่าน้ำหนักองค์ประกอบสูง (0.530 - 0.811) ซึ่งยืนยันความเที่ยงตรงเชิงโครงสร้างได้อย่างเหมาะสม ดังตารางที่ 3

**ตารางที่ 3 :** Rotated Factor Matrix

| Rotated Factor Matrix | | | |
|---|---|---|---|
| ตัวแปร | Factor | | |
| | 1 | 2 | 3 |
| ACE4 | 0.748 | | |
| OAAI24 | 0.745 | | |
| OAAI21 | 0.731 | | |
| OAAI23 | 0.719 | | |
| ACE2 | 0.708 | | |
| OAAI22 | 0.691 | | |
| ACE5 | 0.671 | | |
| ACE3 | 0.633 | | |
| RPS10 | 0.592 | | |
| RPS11 | 0.533 | | |
| PST17 | | 0.811 | |
| PST15 | | 0.807 | |
| PST19 | | 0.691 | |
| PST18 | | 0.626 | |
| RPS14 | | 0.619 | |
| DQA6 | | | 0.777 |
| DQA8 | | | 0.746 |
| DQA9 | | | 0.597 |
| DQA7 | | | 0.586 |

**4.4 ผลการทดสอบความเชื่อมั่นของแบบสอบถาม (**Cronbach's Alpha)

จากตารางที่ 4 พบว่าค่าความเชื่อมั่น (Cronbach's Alpha) รายปัจจัยมีค่าระหว่าง 0.906 - 0.952 และในภาพรวมทั้งฉบับมีค่าสูงถึง 0.966 ซึ่งผ่านเกณฑ์มาตรฐาน (สูงกว่า 0.70) แสดงให้เห็นว่าแบบสอบถามฉบับนี้มีความน่าเชื่อถือสูงมาก

**ตารางที่ 4 :** ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน และค่าความเชื่อมั่นของแต่ละองค์ประกอบ

| องค์ประกอบ (Factor) | Mean | S.D. | Cronbach's Alpha |
|---|---|---|---|
| ปัจจัยที่ 1 | 3.73 | 0.82 | 0.952 |
| ปัจจัยที่ 2 | 3.59 | 0.95 | 0.923 |
| ปัจจัยที่ 3 | 3.63 | 0.83 | 0.906 |

**4.3 ผลการวิเคราะห์และทดสอบโมเดลสมการโครงสร้าง (Structural Equation Modeling: SEM)**

ในการวิเคราะห์อิทธิพลของปัจจัยต่างๆ ที่มีต่อการยอมรับระบบ ผู้วิจัยได้กำหนดรูปแบบสมการเชิงโครงสร้าง (Structural Equation) เพื่อแสดงความสัมพันธ์เชิงสาเหตุ ดังนี้

$$Y = \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + e$$

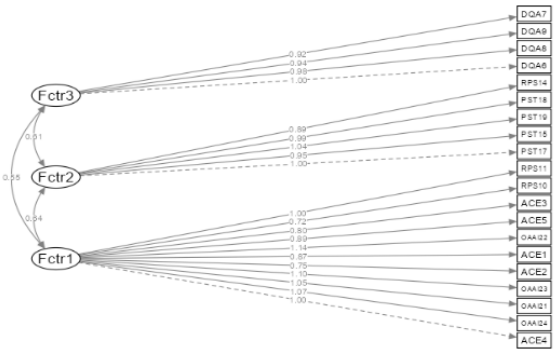
ผลการวิเคราะห์โมเดลสมการโครงสร้างองค์รวม (20 ตัวบ่งชี้) พบว่าโมเดลมีความสอดคล้องกับข้อมูลเชิงประจักษ์อย่างสมบูรณ์ [2], [3] โดยผ่านเกณฑ์มาตรฐานทุกค่าสถิติ ได้แก่ ค่า $p$ -value= 0.095 $(p > 0.05)$ , ($\chi^2$/df ) = 1.143, ดัชนีความกลมกลืนมีค่าเข้าใกล้ 1.000 (CFI, TLI, GFI = 0.999; AGFI = 0.998) และมีค่าความคลาดเคลื่อนที่ต่ำมาก (RMSEA = 0.019, SRMR = 0.048)

**ตารางที่ 5 :** ผลการวิเคราะห์อิทธิพลเชิงสาเหตุ

| เส้นทางความสัมพันธ์ | สัมประสิทธิ์อิทธิพล | S.E. | t-value |
|---|---|---|---|
| ความสะดวก | 0.485 | 0.052 | 9.327 |
| ความปลอดภัย | 0.312 | 0.048 | 6.500 |
| คุณภาพข้อมูล | 0.364 | 0.055 | 6.618 |

| | | | |
|---|---|---|---|
| ค่าสัมประสิทธิ์การพยากรณ์ ($R^2$) | 0.742 | - | - |

จากตารางที่ 5 จึงเป็นการยืนยันได้อย่างมั่นใจว่าโครงสร้างปัจจัยทั้ง 3 ด้าน (ความสะดวกและทัศนคติฯ, ความปลอดภัยและความเชื่อมั่นฯ, คุณภาพของข้อมูลฯ) มีความตรงเชิงโครงสร้าง และสามารถนำไปอธิบายปัจจัยที่ส่งผลต่อการยอมรับระบบได้อย่างมีประสิทธิภาพแสดงดังภาพที่ 2




**ภาพที่ 2** : โมเดลรูป SEM

## 5. สรุป

### 5.1 สรุปผลการวิจัย

การศึกษาจากกลุ่มตัวอย่าง 402 คนในกรุงเทพมหานคร พบว่าผู้บริโภคให้ความสำคัญต่อการยอมรับระบบ Chatbot ในภาพรวมระดับมาก โดยเน้นความสะดวกในการเข้าถึงเป็นหลัก ผลการสกัดองค์ประกอบ (EFA) สามารถจัดกลุ่มปัจจัยได้ 3 ด้าน ได้แก่ 1) ความสะดวกและทัศนคติ 2) ความปลอดภัยและความเชื่อมั่น และ 3) คุณภาพและความถูกต้องของข้อมูล ซึ่งผลการทดสอบโมเดลสมการโครงสร้าง (SEM) ยืนยันว่าโมเดลมีความสอดคล้องกับข้อมูลเชิงประจักษ์อย่างสมบูรณ์ (CFI= 0.999, RMSEA=0.019) และปัจจัยทั้ง 3 ด้านมีอิทธิพลต่อการตัดสินใจใช้งานระบบ Chatbot อย่างมีนัยสำคัญทางสถิติ

### 5.2 อภิปรายผลการวิจัย

จากผลการวิจัยสามารถอภิปรายประเด็นสำคัญได้ดังนี้ ในด้านความสะดวกและทัศนคติต่อการใช้งาน พบว่าผู้บริโภคในยุคดิจิทัลให้ความสำคัญกับความรวดเร็วและประสิทธิภาพเป็นอย่างมาก ดังนั้น หากระบบ **Chatbot** ถูกออกแบบมาให้ใช้งานง่าย เข้าถึงได้ทันที และไม่ซับซ้อน จะช่วยลดความลังเลใจและสร้างทัศนคติเชิงบวกที่ส่งผลต่อการตัดสินใจใช้งานได้อย่างมีนัยสำคัญ ในขณะเดียวกันด้านความปลอดภัยและความเชื่อมั่น ถือเป็นปัจจัยชี้วัดที่สำคัญยิ่ง ซึ่งสอดคล้องกับปัญหาที่พบในปัจจุบันที่ผู้ใช้งานยังคงมีความกังวลเกี่ยวกับการรักษาข้อมูลส่วนบุคคล (**Data Privacy)** หากระบบไม่สามารถสร้างความไว้วางใจหรือแสดงความปลอดภัยให้เห็นได้ ผู้บริโภคจะเกิดความต้านทานและปฏิเสธการใช้งานในทันที

### 5.3 ข้อเสนอแนะ

5.3.1 ข้อเสนอแนะเพื่อการนำไปใช้ประโยชน์: องค์กรควรนำปัจจัยทั้ง 3 ด้านไปพัฒนาระบบ โดยเน้นออกแบบ UI/UX ให้ใช้งานง่าย พัฒนา AI ให้ตอบคำถามได้แม่นยำและเป็นธรรมชาติ พร้อมทั้งสื่อสารนโยบายความปลอดภัยของข้อมูลอย่างโปร่งใสเพื่อสร้างความเชื่อมั่น [1], [5]

5.3.2 ข้อเสนอแนะสำหรับการวิจัยครั้งต่อไป: ควรศึกษาเปรียบเทียบการยอมรับเทคโนโลยีระหว่างกลุ่มช่วงวัย และเพิ่มการวิจัยเชิงคุณภาพเพื่อเจาะลึกความต้องการด้านความรู้สึก (Emotional Connection) ของผู้ใช้งานต่อระบบ AI [2]

5.3.3 การเชื่อมต่อข้อมูลแบบ Real-time API: ระบบควรสามารถดึงข้อมูลสถานะสินค้า หรือข้อมูลส่วนบุคคลของผู้ใช้งานผ่าน API ที่มีความปลอดภัย เพื่อให้ข้อมูลที่ตอบสนองมีความเป็นปัจจุบัน (Up-to-date) และตรงประเด็นกับความต้องการรายบุคคล [3], [4]

5.3.4 การใช้เทคโนโลยี RAG (Retrieval-Augmented Generation): เพื่อป้องกันปัญหาข้อมูลบิดเบือน (AI Hallucination) ควรติดตั้งระบบที่บังคับให้ Chatbot ดึงข้อมูล (Retrieve) จากฐานข้อมูล Knowledge Base ขององค์กรที่ผ่านการตรวจสอบแล้วเท่านั้น ก่อนจะทำการสร้างคำตอบ (Generate)

### เอกสารอ้างอิง

[1] Wu, X., Liu, H., Zhang, C., Zhang, F., Xie, B., & Zhong, X. (2024). Using structural equation

modeling to investigate students' satisfaction with an undergraduate tutorial system.

[2] Yang, X. (2021). Understanding consumers' purchase intentions in social commerce through social capital: Evidence from SEM and fsQCA.

[3] Yee, M., Foo, L. K. F., & Ming, C. Z. (2022). Antecedents of customer satisfaction towards e-business performance during COVID-19..

[4] นันทิชา. (2565). ปัจจัยที่ส่งผลต่อทัศนคติการยอมรับเทคโนโลยี AI Chatbot โดยผู้ให้บริการเครือข่ายโทรศัพท์เคลื่อนที่ ของกลุ่ม Baby Boomer.

[5] ณปภัช. (2567). ปัจจัยที่ส่งผลต่อการยอมรับการใช้ AI Chatbot ในการปรึกษาด้านสุขภาพจิต.

[6] สุชาติ (2562). การพัฒนาระบบแชทบอทด้วย Line Messaging API ในการเรียนรู้แบบปฏิสัมพันธ์. รายงานการวิจัมหาวิทยาลัยสุโขทัยธรรมาธิราช.

# ระบบบริหารจัดการพื้นที่จำหน่ายสินค้าอัจฉริยะ
# Smart Retail Space Management System

*บุรัสกร อยู่สุข (Burasakorn Yoosooka)[1] ศุวิล ชมชัยยา (Suvil Chomchaiya)[2]*
*อุไรวรรณ อินทร์แหยม (Uraiwan Inyaem)[3] ภูมิสิทธิ์ นานรัมย์ (Phumsit Nanram)[4]*
*บวรทัต แดงแตง (Bavornthat Dangtang)[5] และนิติพัฒน์ ศรีธระชิยานนท์ (Nitipat Sritharachiyanont)[6]*
*[1]สาขาวิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเทคโนโลยีราชมงคลพระนคร*
*[2,4,5,6]ภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี*
*[3]ภาควิชาคณิตศาสตร์และวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี*
*burasakorn.y@rmutp.ac.th, suvil.cho@kmutt.ac.th, uraiwan.inyaem@rmutt.ac.th, moppom.9435@gmail.com,*
*bavornthat1208@hotmail.com, koonboom700@gmail.com*

## บทคัดย่อ

*บทความนี้นำเสนอการออกแบบและพัฒนาเว็บไซต์เพื่อสนับสนุนการบริหารจัดการพื้นที่ให้เช่าแบบแบ่งย่อย โดยช่วยให้ผู้ให้เช่าสามารถตรวจสอบสถานะพื้นที่ประชาสัมพันธ์พื้นที่ว่าง และจัดการข้อมูลผู้เช่าผ่านระบบออนไลน์ได้อย่างมีประสิทธิภาพ ซึ่งระบบได้บูรณาการแพลตฟอร์ม Customer Relationship Management (CRM) บน Salesforce ซึ่งทำงานบนเทคโนโลยีคลาวด์ และประยุกต์ใช้ Einstein AI เพื่อวิเคราะห์ข้อมูลผู้ใช้ ติดตามสถานะพื้นที่ และสนับสนุนการตัดสินใจด้านการบริหารจัดการพื้นที่เช่า สำหรับเว็บไซต์ที่พัฒนาขึ้นประกอบด้วยสองส่วนหลัก ได้แก่ ส่วนแสดงข้อมูลพื้นที่สำหรับผู้สนใจเช่า และส่วนจัดการระบบสำหรับผู้ให้เช่า ในด้านการประเมินประสิทธิภาพของระบบ ผู้วิจัยได้แบ่งกลุ่มตัวอย่างออกเป็นกลุ่มทดลองและกลุ่มควบคุม เพื่อเปรียบเทียบผลระหว่างวิธีการทำงานแบบเดิมกับวิธีการทำงานแบบใหม่ทั้งในเชิงปริมาณซึ่งได้แก่ เวลาในการค้นหา จำนวนคลิก อัตราการจองสำเร็จ และความแม่นยำของปัญญาประดิษฐ์ และในเชิงคุณภาพซึ่งได้แก่ ความพึงพอใจของผู้ใช้งาน ซึ่งผลการทดลองกับกลุ่มตัวอย่างจำนวน 60 คน พบว่าจำนวนคลิกในการจองลดลงร้อยละ 63.63 เวลาในการค้นหาลดลงร้อยละ 46.50 อัตราการจองสำเร็จเพิ่มขึ้นร้อยละ 27.00 และระดับความพึงพอใจเพิ่มขึ้น 1.40 คะแนน เป็นการสะท้อนให้เห็นว่าระบบที่พัฒนาขึ้นสามารถเพิ่มประสิทธิภาพการบริหารจัดการพื้นที่เช่า และยกระดับประสบการณ์ของผู้ใช้งานได้อย่างมีนัยสำคัญ.*

**คำสำคัญ:** การจัดการพื้นที่ การให้เช่าพื้นที่ CRM

## Abstract

*This article presents the design and development of a website to support the management of subdivided rental spaces. The system enables landlords to efficiently monitor space availability, advertise vacant spaces, and manage tenant information through an online platform. It integrates a Customer Relationship Management (CRM) platform on Salesforce, which operates on cloud technology, and applies Einstein AI to analyze user data, track space status, and support decision-making in rental space management. The developed website consists of two main components: a space information section for prospective tenants and a system management section for landlords. To evaluate the system's performance, the researchers divided participants into an experimental group and a control group to compare outcomes between the traditional workflow and the newly developed approach. The evaluation included both quantitative and qualitative measures. Quantitative measures consisted of search time, number of clicks, booking success rate, and artificial intelligence accuracy, while qualitative measures focused on user satisfaction. The experiment, conducted with 60 participants, revealed that the number of clicks required for booking decreased by 63.63%, search time decreased by 46.50%, the booking success rate increased by 27.00%, and user*

*satisfaction improved by 1.40 points. These findings demonstrate that the developed system significantly enhances the efficiency of rental space management and improves the overall user experience.*

**Keywords:** Space Management, Space Rental, CRM.

## 1. บทนำ

แม้ว่าในปัจจุบันการทำธุรกิจออนไลน์จะได้รับความนิยมอย่างแพร่หลาย แต่ยังคงมีผู้บริโภคจำนวนไม่น้อยที่ต้องการออกมาเลือกชมและซื้อสินค้าจริงด้วยตนเอง ส่งผลให้เจ้าของพื้นที่ว่างหลายแห่งเริ่มให้ความสนใจในการพัฒนาและปรับปรุงพื้นที่ของตนให้สามารถแบ่งเช่าเพื่อประกอบธุรกิจขนาดย่อมได้ อย่างไรก็ตาม การบริหารจัดการพื้นที่ให้เช่ามักอาศัยประสบการณ์ส่วนบุคคลหรือการจัดการในรูปแบบทั่วไป ซึ่งอาจก่อให้เกิดปัญหา เช่น การจัดการที่ไม่เป็นระบบ ขาดความทั่วถึง และมีรายละเอียดในการบริหารจัดการไม่เพียงพอ

ด้วยเหตุนี้ ผู้วิจัยจึงมีแนวคิดในการพัฒนาเว็บไซต์เพื่อทำหน้าที่เป็นสื่อกลางระหว่างผู้ให้เช่าพื้นที่และผู้ที่ต้องการหาทำเลเพื่อประกอบธุรกิจ โดยผู้ให้เช่าสามารถนำเสนอข้อมูลรายละเอียดของพื้นที่ผ่านเว็บไซต์ ขณะที่ผู้สนใจเช่าสามารถเข้ามาค้นหา เปรียบเทียบ และติดต่อเช่าพื้นที่ได้อย่างสะดวก ระบบจะช่วยจัดการข้อมูลการเช่า เช่น รายละเอียดของพื้นที่ ระยะเวลาในการเช่า อัตราค่าบริการ รวมถึงข้อมูลสำคัญอื่น ๆ ที่เกี่ยวข้องกับการเช่าพื้นที่ นอกจากนี้ ระบบยังได้บูรณาการแพลตฟอร์ม Customer Relationship Management (CRM) จากเว็บไซต์ Salesforce ซึ่งทำงานบนเทคโนโลยีคลาวด์ และเสริมความสามารถด้วยปัญญาประดิษฐ์ผ่าน Einstein AI เพื่อช่วยวิเคราะห์ข้อมูลผู้ใช้งาน ติดตามสถานะพื้นที่ และสนับสนุนการตัดสินใจในการบริหารจัดการพื้นที่เช่าอย่างมีประสิทธิภาพยิ่งขึ้น โดยระบบมีวัตถุประสงค์ดังนี้

- เพื่อเพิ่มโอกาสในการเข้าถึงระหว่างผู้มีพื้นที่และผู้ที่แสวงหาพื้นที่ขาย
- เพื่อศึกษาการบริหารจัดการพื้นที่ให้เกิดประโยชน์สูงสุดของผู้ให้บริการพื้นที่
- เพื่อสร้างเว็บที่รองรับการบริหารจัดการพื้นที่ขายและเป็นสื่อกลางการเชื่อมโยงผู้ให้เช่าและผู้เช่าพื้นที่ภายใต้ขอบเขตดังนี้
- กลุ่มเป้าหมาย มี 2 กลุ่มได้แก่ กลุ่มคนที่มีพื้นที่ว่างแต่ไม่ได้ใช้ประโยชน์ และกลุ่มคนที่มีสินค้าที่ต้องการจำหน่ายหรือประชาสัมพันธ์ แต่ไม่มีพื้นที่
- เป็นการศึกษาการจัดการพื้นที่ขายเพียงอย่างเดียว ไม่รวมถึงรายละเอียด ของตัวสินค้า และการจัดวาง
- เป็นเว็บไซต์สำหรับผู้สนใจเช่าหรือให้เช่าพื้นที่

## 2. ทบทวนวรรณกรรม

### 2.1 ทฤษฎีที่เกี่ยวข้อง

การสร้างเว็บไซต์สิ่งสำคัญอยู่ที่การออกแบบเว็บเพราะเว็บไซต์ที่มีรูปแบบสวยงาม จะสามารถดึงดูดความสนใจจากผู้คนได้ดีกว่า ทำให้ผู้คนเกิดความรู้สึกประทับใจอยากกลับมาใช้งานเว็บไซต์อีกครั้งในอนาคต ดังนั้นเริ่มแรก ก่อนทำเว็บไซต์ จึงจำเป็นต้องทำความเข้าใจ กับหลักการออกแบบ และรูปแบบโครงสร้างของเว็บก่อนการออกแบบเว็บไซต์เพื่อให้มีประสิทธิภาพ และสามารถดึงดูดความสนใจของผู้คนได้ดี จะต้องมีองค์ประกอบของเว็บไซต์อย่างครบถ้วนซึ่งได้แก่ ความเรียบง่าย เข้าใจง่าย ความสม่ำเสมอไม่สับสน สร้างความโดดเด่นเป็นเอกลักษณ์ เนื้อหาต้องดีครบถ้วน ระบบเนวิเกชั่นใช้ง่าย คุณภาพของเว็บไซต์ และความสะดวกในการเข้าใช้งาน

โครงสร้างของเว็บไซต์ประกอบด้วย โครงสร้างแบบเรียงลำดับ โครงสร้างแบบลำดับขั้น โครงสร้างแบบตาราง และโครงสร้างแบบใยแมงมุม [5]

### 2.2 งานวิจัยที่เกี่ยวข้อง

1. การจัดการพื้นที่ตลาด: กรณีศึกษาตลาดนัดหัวมุม MARKET & MORE กรุงเทพมหานคร

มีภาพลักษณ์ที่สวยงาม (Comfort Image) เมื่อมีความสวยงามแล้ว จะต้องมีความปลอดภัยสำหรับผู้ที่เขามาใช้งาน ด้วย ซึ่งเป็นสิ่งสำคัญอย่างมากสำหรับการเข้ามาใช้งานพื้นที่ควรมีความสวยงาม จะต้องจัดองค์ประกอบต่าง ๆ เพื่อให้เกิดความสบายตา รวมถึง ความปลอดภัยที่จะเข้ามาใช้พื้นที่ จัดการเพิ่มการใช้งานในพื้นที่ให้มากขึ้น เพื่อเพิ่ม

จำนวนผู้ที่จะเข้ามาใช้งาน มีแจ้หน้าที่คอยดูแลรักษาความปลอดภัยและทั้งหมดนี้ก็ต้องมีการจัดระบบการบำรุงรักษาเพื่อเป็นการรักษาสถานที่ให้มีความน่าใช้งานอยู่ตลอดเวลา โดยมีการทำความสะอาดพื้นที่ มีการจัดการดูแลซ่อมแซมพื้นที่ส่วนต่าง ๆ มีการจัดการและ การดูแลรักษาระบบสาธารณูปการ เพื่อให้พร้อมต่อการใช้งานกับผู้ที่เขามาใช้พื้นที่ เป็นแนวคิดที่จะสามารถนำมาใช้ในการพิจารณาพฤติกรรมของมนุษย์ได้ [3]

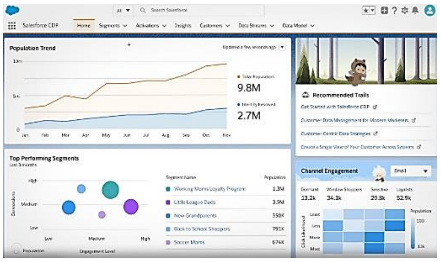
2. ปัจจัยทางการตลาดที่มีผลต่อการตัดสินใจเช่าพื้นที่ตลาดนัดของผู้ประกอบการ ร้านค้าปลีก ในอำเภอศรีราชา จังหวัดชลบุรี

จากงานวิจัยนี้ ผู้วิจัยได้ศึกษาปัจจัยที่มีผลต่อการตัดสินใจเช่าพื้นที่ในตลาดนัดของผู้ประกอบการร้านค้าปลีกในศรีราชา โดยเก็บข้อมูลจากผู้ค้า 400 ราย และใช้แบบจำลองวิเคราะห์ความสัมพันธ์ของปัจจัยต่าง ๆ เช่น ระดับการศึกษา ประเภทสินค้า และระยะทางจากบ้านถึงตลาด ผลการศึกษาพบว่าปัจจัยสำคัญที่มีผลต่อการเลือกเช่าพื้นที่ตลาดนัด KU คือ ระดับการศึกษา วันที่ขายสินค้า ประเภทสินค้า ระยะเวลาที่ขาย ยานพาหนะ ราคาสินค้า ระยะทางจากบ้าน และปัจจัยด้านการตลาด เช่น ผลิตภัณฑ์ ราคา สถานที่ โปรโมชั่น กระบวนการ และลักษณะทางกายภาพ ดังนั้น ผู้จัดตลาดนัดควรปรับปรุงปัจจัยทางการตลาดให้ตรงกับความต้องการของผู้ค้า โดยเฉพาะเรื่องสถานที่ ซึ่งผู้ค้าให้ความสำคัญมากที่สุด [4]

## 3. วิธีการดำเนินงาน

### 3.1. เครื่องมือที่ใช้ในการพัฒนา

1. Salesforce เป็นแพลตฟอร์มสำหรับการบริหารจัดการลูกค้าสัมพันธ์ (CPM) ที่ได้รับความนิยมอย่างแพร่หลายในภาคธุรกิจ โดยมีเครื่องมือหลากหลายที่ช่วยสนับสนุนการจัดการข้อมูลลูกค้า การสื่อสารกับลูกค้า และการติดตามกระบวนการทางธุรกิจอย่างมีประสิทธิภาพ ระบบช่วยรวบรวมข้อมูลสำคัญของลูกค้าไว้ในศูนย์กลางเดียว ไม่ว่าจะเป็นข้อมูลการติดต่อ ประวัติการซื้อขาย หรือประวัติการสื่อสารที่ผ่านมา ทำให้องค์กรสามารถเข้าถึงและวิเคราะห์ข้อมูลลูกค้าได้อย่างเป็นระบบ โดยลักษณะของ Salesforce แสดงดังภาพที่ 1 [6]




**ภาพที่ 1:** หน้าแรกของ Salesforce

Salesforce นำมาใช้ในการเก็บข้อมูล และจัดการข้อมูลของผู้เช่าพื้นที่ วิเคราะห์ข้อมูลและแนวโน้ม ในการจัดการพื้นที่ให้เหมาะสม และมีประสิทธิภาพ ในระบบ MAREASใช้งาน Salesforce โดยทั่วไป Einstein AI เปรียบเสมือน "สมองส่วนวิเคราะห์" ที่ฝังอยู่ในฐานข้อมูล Salesforceโดยทั้งสองส่วนทำงานร่วมกันดังนี้:

- การป้อนข้อมูล (Data Feeding)

Einstein AI จะไม่สามารถฉลาดได้ถ้าไม่มีข้อมูล ดังนั้นขั้นตอนแรกคือการนำข้อมูลจาก MAREAS (MongoDB) ส่งมาเก็บที่ Salesforce Objects (เช่น Leads, Contacts, Bookings) ก่อน ตัวอย่างข้อมูลเช่น ข้อมูลลูกค้าที่จองพื้นที่บ่อยๆ, ประเภทตลาดที่คนชอบจอง, หรือช่วงเวลาที่คนเข้าใช้งานมากที่สุด จะเก็บไว้ใน Salesforce

- การประมวลผลและเรียนรู้(Machine Learning)

เมื่อข้อมูลมีมากพอ Einstein AI จะทำการ "กวาด" ข้อมูลเหล่านั้นมาวิเคราะห์โดยอัตโนมัติ (ไม่ต้องเขียนโค้ดเพิ่มมากนัก) เพื่อหา Pattern หรือรูปแบบพฤติกรรม เช่น ลูกค้าที่มักจะยกเลิกการจอง มักจะมีพฤติกรรมแบบไหนก่อนยกเลิก และพื้นที่เช่าโซนไหนที่มีโอกาสทำกำไรได้สูงสุดในช่วงเทศกาล

- ฟีเจอร์ที่ Einstein ทำงานร่วมกับ Salesforce

ในระบบ Salesforce มักจะมีฟีเจอร์เด่นๆ ที่ Einstein เข้ามาช่วย ดังนี้ Einstein Lead Scoring (การให้คะแนนลูกค้า), Einstein Opportunity Insightsล, Einstein Discovery (การทำนายอนาคต)

- การแสดงผล (Dashboard& Recommendation)

ผลลัพธ์จาก AI จะไม่ได้เป็นแค่ตัวเลขหลังบ้าน แต่จะปรากฏขึ้นมาบนหน้าจอ Salesforce ที่เจ้าหน้าที่ใช้งานด้วย Next Best Action: AI และ AI-Powered Reports

2.MERN Stack มาจากอักษรแรกของ MongoDB, ExpressJS, React และ NodeJS เป็น Framework ที่พัฒนา Full Stack ประกอบด้วย

- MongoDB ใช้เป็นฐานข้อมูลเอกสารแบบ NoSQL ของ MongoDB ทำให้ง่ายต่อการรองรับโมเดลข้อมูลที่แตกต่างกัน และจัดการรับส่งข้อมูลปริมาณมากได้
- ExpressJS นักพัฒนาสามารถสร้าง RESTful API API และโครงสร้างพื้นฐาน Back-End ในขณะที่ยังคงรักษาประสิทธิภาพและประสิทธิผลในระดับสูงไว้ได้โดยจะเป็นการเชื่อม API เข้ากับซอฟต์แวร์ Salesforce
- React เป็น Library ของ JavaScript แบบ opensource ที่สร้างส่วนต่อประสานผู้ใช้ (User Interface: UI) และอัปเดตส่วนประกอบ UI ผ่าน DOM
- NodeJS เป็น opensource ข้ามแพลตฟอร์มที่สร้างขึ้นบนกลไก V8 JavaScript ทำหน้าที่เป็นแกนหลักของ MERN Stack ช่วยให้นักพัฒนาใช้ประโยชน์จาก JavaScript สำหรับการพัฒนา Front-End, Back-End จึงทำให้กระบวนการพัฒนาทั้งหมดดีขึ้น

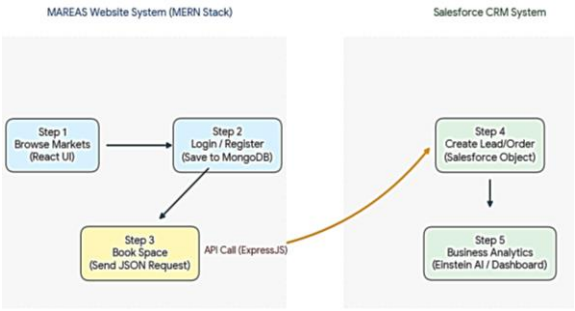
3. Visual Studio Code หรือ VSCode เป็นโปรแกรม Code Editor ที่ใช้ในการแก้ไขและปรับแต่งโค้ด สนับสนุนทั้งภาษา JavaScript, TypeScript และ Node.js สามารถเชื่อมต่อกับ Git ได้ นำมาใช้งานได้ง่ายไม่ซับซ้อน มีเครื่องมือส่วนขยายต่าง ๆ ให้เลือกใช้อย่างมากมาก

4. Postman คือแพลตฟอร์มที่ออกแบบมาโดยเฉพาะสำหรับการพัฒนา ทดสอบ และจัดการ API (Application Programming Interfaces) ขจัดอุปสรรค และย่นระยะเวลาการพัฒนา ทำให้ทำงานได้อย่างราบรื่นในทุกสถานที่ และสามารถทำร่วมกับทีมพัฒนาได้ด้วย [7]

5. GitHub นำมาใช้เป็นแพลตฟอร์มในการจัดเก็บโค้ดของเว็บไซต์ และแชร์โค้ดให้กับทีม

### 3.2. การออกแบบ

ภาพการทำงานโดยรวมของทั้งสองเว็บไซต์คือ เว็บไซต์ของ Salesforce ซึ่งใช้เป็นที่เก็บรวบรวมข้อมูลลูกค้าที่มีพื้นที่เช่า ข้อมูลพื้นที่ และเว็บไซต์ MAREAS เพื่อนำไปจัดการการเช่าพื้นที่โดยทั้งสองเว็บมีการทำงานสัมพันธ์กัน ดังแสดงในภาพที่ 2




**ภาพที่ 2:** กระบวนการทำงานของระบบ

Step 1 (Browse): ผู้ใช้เข้าเว็บเลือกดูตลาด (ข้อมูลถูกดึงมาจาก MongoDB ของ MAREAS)

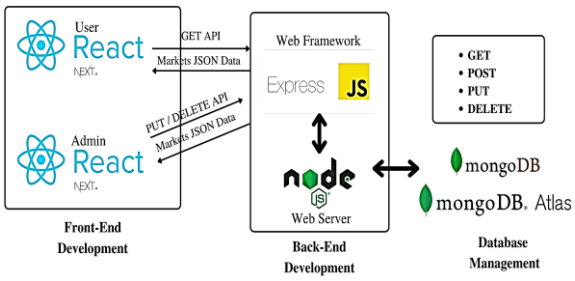
Step 2 (Identify): ผู้ใช้ทำการสมัครสมาชิกหรือล็อกอินเพื่อระบุตัวตน

Step 3 (Action): เมื่อผู้ใช้กดจองพื้นที่ ข้อมูลจะถูกรวบรวมเป็นก้อน JSON และส่งผ่าน ExpressJS API

Step 4 (Integration): ข้อมูลวิ่งเข้าสู่ Salesforce โดยอัตโนมัติ เพื่อสร้างเป็นข้อมูลลูกค้าใหม่หรือรายการจอง

Step 5 (Analysis): เจ้าของโครงการใช้เครื่องมือใน Salesforce วิเคราะห์ข้อมูล (เช่น ดูว่าตลาดไหนคนจองเยอะที่สุดจาก Dashboard) เพื่อนำไปวางแผนการตลาดต่อ

ผู้วิจัยได้ออกแบบสถาปัตยกรรมระบบดังภาพที่3




**ภาพที่ 3:** สถาปัตยกรรมของระบบ MAREAS

ผู้วิจัยได้ประยุกต์การทำงานโดยแบ่งเป็นระบบดังนี้

1. Front-End Development โดยเว็บไซต์แบ่งออกเป็นสองฝั่งคือ ฝั่ง User หรือผู้ใช้งาน หรือผู้เข้ามาเช่าพื้นที่ และฝั่ง Admin หรือผู้ดูแลระบบ หรือผู้ให้เช่า ซึ่งรับและส่งข้อมูลเข้าสู่ ระบบหลังบ้านหรือ Back-End Development เป็นข้อมูลรูปแบบ JSON

2. Back-End Development ผ่าน Web Framework ด้วย Express JS และติดต่อกับ Web Server ด้วย Node JS เพื่อเก็บข้อมูลระบบ

3.Database Management ด้วยฐานข้อมูล MongoDB และมี MongoDB Atlas เป็น service ของฐานข้อมูล สาเหตุที่ MAREAS เลือกใช้ MongoDB (ฐานข้อมูลของตัวเอง) ควบคู่ไปกับ Salesforce (ระบบ CRM) แทนที่จะเก็บไว้ในที่เดียว มีเหตุผลสำคัญด้านการออกแบบ 3 ข้อดังนี้

1. ความเร็วการตอบสนองของหน้าเว็บ (Performance)

- MongoDB: เป็นฐานข้อมูลที่ทำงานได้เร็วมาก สำหรับการรันเว็บไซต์แบบ Real-time เช่น การดึงรูปภาพตลาด การเช็คว่าล็อคไหนว่าง หรือการให้ผู้ใช้ Log-in เข้าหน้าเว็บ ข้อมูลเหล่านี้ต้องเรียกใช้ "ทันที"
- Salesforce: แม้จะเก็บข้อมูลได้ดี แต่การเรียกข้อมูลผ่าน API ทุกครั้งที่คลิกหน้าเว็บจะทำให้เว็บ "อืด" และมีข้อจำกัดเรื่องจำนวนครั้งในการเรียกใช้ (API Limits)

2. โครงสร้างข้อมูลที่ต่างกัน (Data Structure)

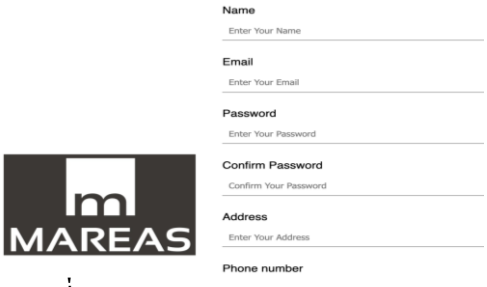
- ข้อมูลหน้าเว็บ (Unstructured): ข้อมูลอย่างรูปภาพตลาดหลายๆ มุม, รีวิวลูกค้า, หรือพิกัดแผนที่ มีความยืดหยุ่นสูง ซึ่ง MongoDB (NoSQL) จัดการง่ายกว่า
- ข้อมูลธุรกิจ (Structured): ข้อมูลอย่าง ยอดจองรวม, ประวัติการติดต่อลูกค้า, หรือคะแนนความพึงพอใจ ข้อมูลเหล่านี้ต้องการความเป็นระเบียบเพื่อนำไปทำกราฟวิเคราะห์ ซึ่ง Salesforce (Relational) เก่งด้านนี้มาก

3. ค่าใช้จ่ายและความคุ้มค่า (Cost Optimization)

- พื้นที่เก็บข้อมูล: พื้นที่จัดเก็บข้อมูลบน Salesforce มีราคาค่อนข้างสูง (ตามปริมาณ Data Storage)
- การประหยัด: หากเราเก็บข้อมูล "ขยะ" หรือข้อมูลทางเทคนิคทั้งหมด (เช่น Log การเข้าเว็บ, รหัสผ่านที่เข้ารหัสแล้ว) ไว้บน Salesforce จะสิ้นเปลืองมาก เขาจึงเลือกส่งเฉพาะ "ข้อมูลที่มีมูลค่าทางการตลาด" (เช่น ใครสนใจเช่าที่ไหน) ไปที่ Salesforce เพื่อประหยัดค่าใช้จ่าย

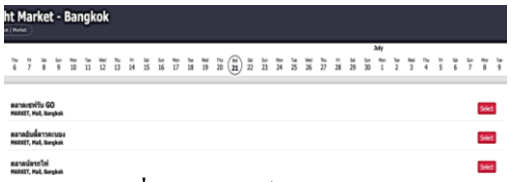
การออกแบบ UI ของเว็บไซต์ MAREAS เพื่อบริหารจัดการพื้นที่ขนาดเล็กสำหรับธุรกิจจัดสรรพื้นที่ขายสินค้าโดยใช้ Figma มีการออกแบบตามฟังก์ชันการใช้งานดังนี้

1. การลงทะเบียนของผู้ใช้ ดังภาพที่ 4



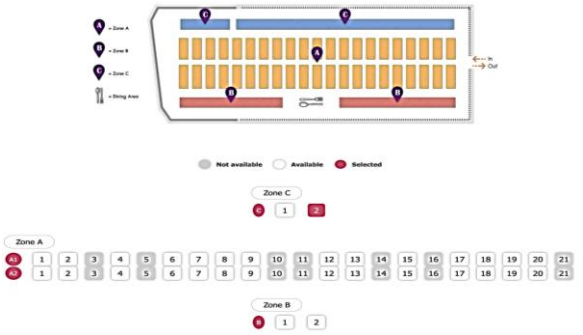

**ภาพที่ 4:** การเข้าสู่ระบบและการลงทะเบียนของผู้ใช้

2. การเลือกพื้นที่ตลาดและวันที่ดังภาพที่ 5



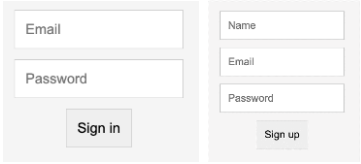

**ภาพที่ 5:** การเลือกพื้นที่ตลาดและวันที่

3. การจองพื้นที่และบอกสถานะของพื้นที่ดังภาพที่ 6



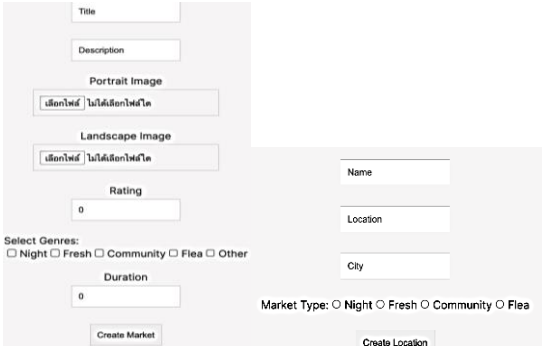

**ภาพที่ 6:** การจองพื้นที่

4. การเข้าสู่ระบบและลงทะเบียนของฝั่ง Admin หรือผู้ดูและระบบโดยต้องลงทะเบียนด้วยรหัสผ่านเฉพาะจากนักพัฒนาเพื่อได้รับการอนุญาติการลงทะเบียนเข้าเป็นผู้ดูแลระบบ ซึ่งการเข้าใช้มี Token คงอยู่เป็นเวลา 30 นาที เมื่อออกจากหน้าเว็บไซต์ และจะออกจากระบบให้อัตโนมัติ เพื่อป้องกันการลืมออกจากระบบดังภาพที่ 7



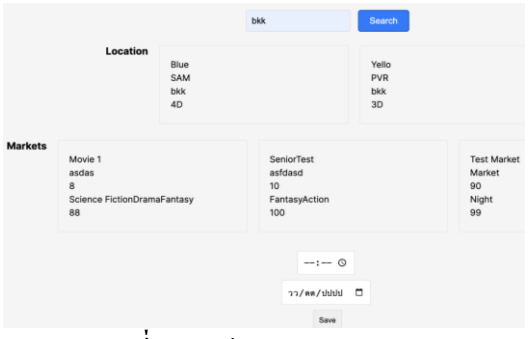

**ภาพที่ 7:** การเข้าสู่ระบบ (ซ้าย) และลงทะเบียน (ขวา) ของ Admin

5. การเพิ่มข้อมูลหน้า Market และLocation ดังภาพที่ 8




**ภาพที่ 8:** การเพิ่มข้อมูลหน้า Market และ Location

6. การเพิ่มข้อมูลตำแหน่งหน้า Location ดังภาพที่ 9




**ภาพที่ 9:** การเพิ่มข้อมูลตลาดเข้าสู่ฐานข้อมูล

## 4. ผลการดำเนินงาน

ผลการทำงานของเว็บไซต์ MAREAS สัมฤทธิ์ผลทุกฟังก์ชัน ได้แก่ 1. การเข้าสู่ระบบและลงทะเบียนของผู้ใช้ 2. การเลือกพื้นที่ตลาด 3. การจองพื้นที่ 4. การเข้าสู่ระบบและลงทะเบียนของฝั่งผู้ดูแลระบบ 5. การเพิ่มข้อมูลตลาด 6. การเพิ่มข้อมูลตำแหน่ง 7. การเพิ่มข้อมูลสู่ฐานข้อมูล

ผลการทำงานของ Salesforce สัมฤทธิ์ผลทุกฟังก์ชัน ได้แก่ 1.การเก็บข้อมูลลูกค้า 2. การรวบรวมข้อมูลลูกค้าเพื่อนำไปพัฒนา 3. การจัดการที่สะดวกรวดเร็ว 4. การดูแลลูกค้า

การวิจัยนี้แบ่งกลุ่มตัวอย่างออกเป็น 2 กลุ่ม ได้แก่ กลุ่มควบคุม (ใช้วิธีการเดิม) 30 คนและกลุ่มทดลอง (ใช้ระบบ MAREAS) 30 คน เพื่อเปรียบเทียบผลลัพธ์ของระบบ โดยตัวแปรเชิงคุณภาพที่ใช้ในการประเมิน คือ ความพึงพอใจส่วนตัวแปรเชิงปริมาณที่ใช้ในการประเมิน ได้แก่ จำนวนคลิกในการจอง, เวลาในการค้นหา, อัตราการจองสำเร็จ, และ Accuracy ของระบบ AI, ได้ผลดังตารางต่อไปนี้

**ตารางที่ 1:** รายละเอียดผลการประเมินรายข้อ

ส่วนที่ 1: สำหรับกลุ่มผู้ให้เช่าพื้นที่ (10 คนต่อกลุ่ม)

| รายการประเมิน | กลุ่มควบคุม | กลุ่มทดลอง | S.D. (ทดลอง) | ระดับ (ทดลอง) |
|---|---|---|---|---|
| 1. ความสะดวกในการอัปเดตข้อมูลพื้นที่ว่าง | 3.20 | 4.60 | 0.42 | มากที่สุด |
| 2. ความแม่นยำของรายงานและสรุปยอดการจอง | 3.40 | 4.40 | 0.46 | มาก |
| 3. ความรวดเร็วในการรับข้อมูลจากหน้าเว็บ | 3.30 | 4.55 | 0.41 | มากที่สุด |
| 4. ระบบช่วยวิเคราะห์แนวโน้มลูกค้า (AI) | 0.00 | 4.35 | 0.40 | มาก |
| เฉลี่ยรวม | 3.30 | 4.45 | 0.42 | มาก |

ส่วนที่ 2: สำหรับกลุ่มผู้เช่าพื้นที่ (20 คนต่อกลุ่ม)

| รายการประเมิน | กลุ่มควบคุม | กลุ่มทดลอง | S.D. (ทดลอง) | ระดับ (ทดลอง) |
|---|---|---|---|---|
| 1. ความชัดเจนของข้อมูลและตำแหน่งพื้นที่ | 3.47 | 4.90 | 0.45 | มากที่สุด |
| 2. ความรวดเร็วในการค้นหาพื้นที่ที่ต้องการ | 3.20 | 4.75 | 0.56 | มากที่สุด |
| 3. ขั้นตอนการจองมีความสั้นและเข้าใจง่าย | 3.05 | 4.75 | 0.48 | มากที่สุด |
| 4. ความเสถียรและความน่าเชื่อถือของระบบ | 3.15 | 4.60 | 0.62 | มากที่สุด |
| เฉลี่ยรวม | 3.22 | 4.75 | 0.45 | มากที่สุด |

**ตารางที่ 2:** เปรียบเทียบความพึงพอใจภาพรวม (N=60)

| กลุ่มตัวอย่าง | สถานะ | จำนวน(คน) | ค่าเฉลี่ย ($\bar{x}$) | ระดับความพึงพอใจ |
|---|---|---|---|---|
| กลุ่มควบคุม | ผู้ให้เช่า | 10 | 3.30 | ปานกลาง |
| | ผู้เช่า | 20 | 3.22 | ปานกลาง |
| กลุ่มทดลอง | ผู้ให้เช่า | 10 | 4.45 | มาก |
| | ผู้เช่า | 20 | 4.75 | มากที่สุด |

**ตารางที่ 3** ผลการทดลองเปรียบเทียบ (Performance Metrics)

| ตัวแปรเชิงปริมาณ | กลุ่มควบคุม | กลุ่มทดลอง | การเปลี่ยนแปลง |
|---|---|---|---|
| จำนวนคลิกในการจอง (ครั้ง) | 11.00 | 4.00 | ลดลง 63.63% |
| เวลาในการค้นหา (วินาที) | 162.45 | 45.10 | ลดลง 46.50% |
| อัตราการจองสำเร็จ (%) | 42.50% | 69.50% | เพิ่มขึ้น27.00% |
| ความพึงพอใจ (1-5) | 3.25 | 4.65 | เพิ่มขึ้น 1.40 |

ประสิทธิภาพของระบบ AI (Einstein AI Metrics) วัดได้จากการที่กลุ่มทดลองใช้ระบบ MAREAS ซึ่งระบบจะส่งข้อมูลไปวิเคราะห์ใน Salesforce ปรากฏผลลัพธ์ดังนี้:

- AI Accuracy: 85.25% (ความแม่นยำของการทำนายว่า “ผู้ใช้ระบบคนใดจะทำธุรกรรมสำเร็จ”)
- Precision: 83.75% (ความถูกต้องเมื่อ AI ทำนายว่า “เป็นผู้ที่สนใจบริการ และอาจพัฒนาเป็นลูกค้าในอนาคต คุณภาพสูง”)
- Recall: 82.50% (ความสามารถในการพบโอกาสในการจองทั้งหมดของระบบ)

## 5. สรุปผลการวิจัย

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อพัฒนาเว็บไซต์สำหรับบริหารจัดการและดูแลพื้นที่ขนาดเล็กเพื่อใช้ประโยชน์ทางธุรกิจ โดยบูรณาการระบบบริหารลูกค้าสัมพันธ์จาก Salesforce เพื่อช่วยรวบรวมและจัดการข้อมูลผู้เช่า ผู้ให้เช่า และข้อมูลร้านค้าที่อยู่ภายในพื้นที่อย่างเป็นระบบ พร้อมทั้งเสริมศักยภาพด้วยเทคโนโลยีปัญญาประดิษฐ์ผ่าน Einstein AI ซึ่งช่วยวิเคราะห์ข้อมูลผู้ใช้งาน ติดตามสถานะพื้นที่ และสนับสนุนการตัดสินใจในการจัดสรรพื้นที่ได้อย่างมีประสิทธิภาพมากยิ่งขึ้น

ระบบที่พัฒนาขึ้นประกอบด้วย 2 ส่วนหลัก ได้แก่ ส่วนของ Salesforce สำหรับจัดเก็บและบริหารข้อมูลลูกค้าและข้อมูลร้านค้า และส่วนของเว็บไซต์สำหรับให้ผู้ที่สนใจสามารถค้นหาและจองพื้นที่ขายสินค้าได้อย่างสะดวก ผลจากการประเมินความพึงพอใจของผู้ใช้งานระบบ MAREAS พบว่า ผู้ให้เช่าพื้นที่มีความพึงพอใจโดยรวมอยู่ในระดับมากมีค่าเฉลี่ย = 4.45 และผู้เช่าพื้นที่มีความพึงพอใจโดยรวมอยู่ในระดับมากที่สุดมีค่าเฉลี่ย = 4.75 ซึ่งมีค่าเฉลี่ยเพิ่มขึ้นจากผลความพึงพอใจของผู้ใช้งานระบบเดิมถึง 1.40 (ตามระดับ 1-5 ของ Likert Scale)

ผลการวิเคราะห์ประสิทธิภาพมีประเด็นสำคัญดังนี้:

1. ประสบการณ์ผู้ใช้ (User Experience): จำนวนคลิกที่ลดลงจาก 11.00 เหลือ 4.00 ครั้ง สะท้อนถึงการออกแบบ UI/UX ที่ดีและการเชื่อมต่อฐานข้อมูลที่ลื่นไหล ทำให้ขั้นตอนการจองไม่ซับซ้อน

2. ประสิทธิภาพเชิงเวลา (Time Efficiency): ระบบ MAREAS + AI ช่วยลดเวลาในการค้นหาจาก 162.45 วินาที เหลือเพียง 45.10 วินาที เนื่องจาก AI ช่วยกรองและแนะนำพื้นที่ที่ตรงกับความต้องการของผู้เช่ามากที่สุด (Personalized Recommendation)

3. ประสิทธิภาพเชิงฟังก์ชัน (Functional Efficiency): ระบบ MAREAS + AI ช่วยเพิ่มอัตราการจองสำเร็จจาก 42.50% เป็น 69.50% เนื่องจาก AI ช่วยอำนวยความสะดวกในการจองได้สองคล้องกับความต้องการผู้ใช้

4. ความแม่นยำของระบบ AI: ค่า Accuracy ที่ 85.25% ยืนยันว่าการส่งข้อมูลจาก MongoDB ไปยัง Salesforce Einstein AI ช่วยให้ระบบสามารถประมวลผลและให้คำแนะนำที่มีคุณภาพแก่ผู้ใช้งานได้จริง

สะท้อนให้เห็นว่าระบบใช้งานง่าย เข้าถึงได้สะดวก และช่วยเพิ่มประสิทธิภาพในการบริหารจัดการพื้นที่ได้อย่างเหมาะสม สำหรับข้อเสนอแนะประกอบด้วย:

- ควรมีการอบรมผู้ให้เช่าพื้นที่เกี่ยวกับการอ่านค่าพยากรณ์จาก Einstein AI เพื่อให้สามารถนำข้อมูลไปใช้ในการวางแผนการตลาดได้อย่างเต็มประสิทธิภาพ
- ควรปรับปรุงให้การใช้งานของผู้เช่าและผู้ให้เช่าพื้นที่สามารถดำเนินการได้ภายในเว็บไซต์เดียว เพื่อความสะดวกและลดความสับสนในการใช้งาน
- ควรมีการศึกษาการนำ Chatbot AI มาเชื่อมต่อกับ Salesforce Service Cloud เพื่อตอบคำถามลูกค้าหน้าเว็บแบบอัตโนมัติ 24 ชั่วโมง
- ควรเพิ่มการทดสอบระบบกับกลุ่มตัวอย่างที่หลากหลายขึ้นในเชิงประเภทธุรกิจ เพื่อปรับจูน (Fine-tuning) โมเดล AI ให้มีความแม่นยำเฉพาะทางมากขึ้น

## 6. กิตติกรรมประกาศ

ขอกราบขอบพระคุณสาขาวิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเทคโนโลยีราชมงคลพระนคร อีกทั้งภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี และภาควิชาคณิตศาสตร์และวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์

และเทคโนโลยี มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี ที่ให้การสนับสนุนการนำเสนอบทความวิชาการในครั้งนี้

## เอกสารอ้างอิง

[1] นพวรรณ คล้ายทิพย์, 2560, ปัจจัยทางการตลาดที่มีผลต่อการตัดสินใจเช่าพื้นที่ตลาดนัด ของผู้ประกอบการในโครงการกาแลไนท์บาซาร์ย่านไนท์บาซาร์จังหวัดเชียงใหม่, หน้า 3, Available: http://www.ba-abstract.ru.ac.th

[2] วิชชุดา เขียวเกตุ, 2559, การจัดการตนเองของตลาดชุมชน กรณีศึกษา ตลาดริมน้ำคลองแดน อำเภอระโนด จังหวัดสงขลา, หน้า 21, Available: https://ethesisarchive.library.tu.ac.th/thesis

[3] รัตตญา เทวรักษ, 2561, การจัดการพื้นที่ตลาด: กรณีศึกษาตลาดนัดหัวมุม MARKET & MORE กรุงเทพมหานคร, วิทยานิพนธ์ปริญญาสถาปัตยกรรมศาสตร์มหาบัณฑิต สาขาวิชาการผังเมือง คณะสถาปัตยกรรมศาสตร์และการผังเมือง มหาวิทยาลัยธรรมศาสตร์, หน้า 17, Available: http://ethesisarchive.library.tu.ac.th/thesis

[4] อภิรัตน์ บุรุษวยากรณ์ และ ศรีอร สมบูรณ์ทรัพย์, 2564, ปัจจัยทางการตลาดที่มีผลต่อการตัดสินใจเช่าพื้นที่ตลาดนัดของผู้ประกอบการร้านค้าปลีกในอำเภอศรีราชาจังหวัดชลบุรี, Available: https://ojs.lib.buu.ac.th/index.php/

[5] 1belief, 2017, หลักการ ออกแบบเว็บ ขั้นพื้นฐาน พร้อมองค์ประกอบและรูปแบบโครงสร้าง [Online], Available: https://www.1belief.com/

[6] Piotr, 2022, Salesforce Data Management 101: The Essentials You Need to Know [Online], Available: https://www.salesforce.com/

[7] Shashank Awasthi, 2022, Integrated API Monitoring in Postman [Online], Available: https://blog.postman.com/

# ระบบวิเคราะห์และบริหารทรัพยากรบุคคล กรณีศึกษา วิทยาลัยเซนต์หลุยส์

# Human Resource Analysis and Management System: A Case Study of Saint Louis College

*ธัษชนน พลงาม (Tadchanon Polngam)[1] และผุสดี บุญรอด (Pudsadee Boonrawd)[2]*
*ภาควิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล*
*มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ*
*s6707011857063@email.kmutnb.ac.th, pudsadee.b@itd.kmutnb.ac.th*

## บทคัดย่อ

*การบริหารทรัพยากรบุคคลในสถาบันอุดมศึกษาเป็นเรื่องที่สำคัญ หลายหน่วยงานมีโครงสร้างองค์กรที่ซับซ้อน และจำเป็นต้องมีการวิเคราะห์ข้อมูลเชิงลึกเพื่อช่วยในการตัดสินใจ งานวิจัยนี้จึงมีวัตถุประสงค์เพื่อพัฒนาระบบสารสนเทศสำหรับวิเคราะห์และบริหารทรัพยากรบุคคลที่ขับเคลื่อนด้วยข้อมูล (Data-Driven) กรณีศึกษาวิทยาลัยเซนต์หลุยส์ โดยบูรณาการข้อมูลต่าง ๆ ได้แก่ ข้อมูลบุคลากร ภาระงานสอน งานวิจัย การฝึกอบรม และงบประมาณ โดยระบบถูกพัฒนาขึ้นในรูปแบบเว็บแอปพลิเคชันที่มีการควบคุมสิทธิ์การเข้าถึง พร้อมประยุกต์ใช้วิทยาการวิเคราะห์ข้อมูล และแสดงผลผ่านหน้าจอเชิงโต้ตอบ (Interactive Dashboard) โดยประยุกต์ใช้ไลบรารี Chart.js และ ApexCharts เพื่อนำเสนอตัวบ่งชี้ประสิทธิภาพหลัก (KPIs) ในรูปแบบเชิงโต้ตอบ จากผลการประเมินคุณภาพและความพึงพอใจของผู้ใช้งานที่มีต่อระบบสารสนเทศพบว่าอยู่ในเกณฑ์มาก โดยระบบสารสนเทศสามารถแสดงข้อมูลสรุปสำหรับผู้บริหารในรูปแบบแดชบอร์ดได้แบบเรียลไทม์ ช่วยลดระยะเวลาและความซ้ำซ้อนในการจัดทำรายงานเฉพาะกิจ อีกทั้งสามารถวิเคราะห์ภาพรวมและเจาะลึกข้อมูลเพื่อช่วยในการตัดสินใจได้อย่างมีประสิทธิภาพ*

**คำสำคัญ:** ระบบวิเคราะห์และบริหารทรัพยากรบุคคล การขับเคลื่อนด้วยข้อมูล แดชบอร์ดเชิงโต้ตอบ

## Abstract

*Human resource management in higher education institutions is crucial. Many departments have complex organizational structures and require in-depth data analysis to assist in decision-making. Therefore, this research aims to develop a data-driven human resource management and analytics information system for Saint Louis College, integrating personnel data, teaching workloads, research, training, and budgets. The system was developed as a web application with access controls, data analytics, and results displayed through an interactive dashboard. It utilizes the Chart.js and ApexCharts libraries to present key performance indicators (KPIs) in an interactive format. The results showed that user satisfaction and the quality of the information system were high. The information system can present summary data to executives in a dashboard format in real time, helping reduce the time and redundancy required to prepare ad hoc reports. Furthermore, it can analyze both overview and in-depth data to support effective decision-making.*

**Keywords:** Human Resource Analytics and Management System, Data-Driven, Interactive Dashboard

## 1. บทนำ

การเปลี่ยนผ่านองค์กรสู่การบริหารจัดการที่ขับเคลื่อนด้วยข้อมูล (Data-Driven Organization) ถือเป็นกลยุทธ์สำคัญในการสร้างขีดความสามารถทางการแข่งขันสำหรับสถาบันอุดมศึกษา การบริหารทรัพยากรบุคคล (Human Resource Management) จำเป็นต้องอาศัยข้อมูลเชิงลึกที่แม่นยำ เพื่อสนับสนุนการตัดสินใจเชิงกลยุทธ์ตั้งแต่การวางแผนกำลังคน การประเมินภาระงานสายวิชาการ ไปจนถึงการบริหารงบประมาณ ซึ่งการวิเคราะห์ข้อมูลทรัพยากรบุคคล (Human Resource Analytics) มีบทบาทสำคัญอย่างยิ่งในการยกระดับประสิทธิภาพการบริหารจัดการในสถาบันการศึกษา [1] อย่างไรก็ตาม ข้อมูลเหล่านี้มักถูกจัดเก็บแยกส่วนกันในระบบฐานข้อมูลเชิงสัมพันธ์ (Relational Database) ขององค์กรที่มีความซับซ้อน ทำให้การดึงข้อมูลมาบูรณาการเพื่อวิเคราะห์ในภาพรวมทำได้ยากและขาดความคล่องตัว

เทคโนโลยีการบูรณาการข้อมูล (Data Integration) และการสร้างตัวแบบข้อมูล (Data Modeling) มีบทบาทสำคัญในการรวบรวมข้อมูลที่กระจัดกระจาย ตัวอย่างเช่น ข้อมูลบุคลากร ภาระงานสอน งานวิจัย การฝึกอบรม และงบประมาณ เพื่อนำมาจัดโครงสร้างใหม่ให้อยู่ในรูปแบบที่พร้อมวิเคราะห์ (Analysis-Ready Data) บนสถาปัตยกรรมเว็บ (Web-Based Architecture) ซึ่งช่วยลดความซ้ำซ้อนและเพิ่มประสิทธิภาพการเข้าถึงข้อมูลได้อย่างเป็นรูปธรรม [2]

ด้วยเหตุนี้ ผู้วิจัยจึงได้พัฒนาระบบวิเคราะห์และบริหารทรัพยากรบุคคลที่ขับเคลื่อนด้วยข้อมูล กรณีศึกษาวิทยาลัยเซนต์หลุยส์ ที่รองรับการประมวลผลแบบเรียลไทม์ มีความยืดหยุ่น และตอบสนองรวดเร็ว เพื่อให้ผู้บริหารของสถาบันการศึกษาสามารถเรียกดูและวิเคราะห์ข้อมูลเชิงลึกได้ด้วยตนเอง (Self-Service Analytics) สนับสนุนการมองเห็นภาพรวม และตัดสินใจเชิงกลยุทธ์ได้อย่างรวดเร็วและแม่นยำ

## 2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

### 2.1 แนวคิดการริหารทรัพยากรบุคคลเชิงวิเคราะห์

การบริหารทรัพยากรบุคคลเชิงวิเคราะห์ (Human Resources Analytics) เป็นกระบวนการอย่างเป็นระบบที่เกี่ยวข้องกับการรวบรวม วิเคราะห์ และตีความข้อมูลด้านบุคลากร เพื่อยกระดับการตัดสินใจของผู้บริหารไปสู่การตัดสินใจบนฐานของหลักฐานเชิงประจักษ์ (Evidence - Based Decision Making) อย่างมีประสิทธิภาพ

การประยุกต์ใช้แนวคิดการขับเคลื่อนด้วยข้อมูล (Data-Driven Approach) ช่วยให้องค์กรสามารถประเมินผลการปฏิบัติงาน ตลอดจนวิเคราะห์แนวโน้มและปัจจัยที่เกี่ยวข้องได้อย่างแม่นยำ โดย Yadav และ Bhutada [3] กล่าวถึงการวิเคราะห์ทรัพยากรบุคคลที่ขับเคลื่อนด้วยข้อมูลสามารถแปรเปลี่ยนข้อมูลดิบให้เป็นสารสนเทศเชิงลึก (Insights) ที่ทรงคุณค่าและเอื้อต่อการตัดสินใจของผู้บริหาร อีกทั้งงานวิจัยของ Sushma และ Sarala [4] ยังได้อธิบายถึงการวิเคราะห์ข้อมูลทางด้านทรัพยากรบุคคลมีบทบาทสำคัญในการเสริมสร้างขีดความสามารถด้านการตัดสินใจเชิงกลยุทธ์ อันส่งผลโดยตรงต่อประสิทธิผลและความสำเร็จของการบริหารจัดการองค์กรในบริบทของยุคดิจิทัลปัจจุบัน

### 2.2 การบูรณาการข้อมูลและระบบฐานข้อมูลเชิงสัมพันธ์

เพื่อแก้ปัญหาการจัดเก็บข้อมูลแบบกระจัดกระจาย (Data Silos) การบูรณาการข้อมูล (Data Integration) จากฐานข้อมูลเชิงสัมพันธ์แบบแยกส่วน MySQL จึงมีความสำคัญ โดยอาศัยกระบวนการทำความสะอาดและสร้างแบบจำลองข้อมูล (Data Modeling) เพื่อเตรียมความพร้อมสำหรับการวิเคราะห์เชิงลึก การนำเสนอข้อมูลที่บูรณาการแล้วผ่านสถาปัตยกรรมบนเว็บ (Web-Based Architecture) จะช่วยลดความซ้ำซ้อน และเพิ่มความคล่องตัวให้ผู้บริหารสามารถเข้าถึงข้อมูลเพื่อใช้ประโยชน์ได้อย่างเป็นรูปธรรมและยั่งยืน [2]

### 2.3 งานวิจัยที่เกี่ยวข้อง

การประยุกต์ใช้เทคโนโลยีเพื่อการบริหารจัดการทรัพยากรบุคคลและระบบสารสนเทศได้รับความสนใจและการศึกษาอย่างแพร่หลาย โดยงานวิจัยของ Sushma และ Sarala [4] ได้ชี้ให้เห็นว่าการบริหารทรัพยากรบุคคล

เชิงวิเคราะห์มาประยุตก์ใช้สามารถลดระยะเวลาในการดำเนินงาน พร้อมทั้งเพิ่มประสิทธิภาพในการบริหารกำลังคนได้ สอดคล้องกับการศึกษาของ Al-Sabaeei และคณะ [2] ซึ่งพบว่าการพัฒนาหน้าจอแสดงผลข้อมูลแบบรวมศูนย์ (Centralized Dashboard) มีบทบาทสำคัญในการสนับสนุนการตัดสินใจของผู้บริหารให้มีความแม่นยำและมีประสิทธิผลยิ่งขึ้น นอกจากนี้ งานวิจัยในประเทศไทยของ สมเพ็ชร จุลลาบุดดี และ เอกชัย คนคิด [6] ในปี พ.ศ. 2568 ยืนยันว่า การพัฒนากระดานข้อมูลเชิงโต้ตอบด้านทรัพยากรบุคคล ช่วยยกระดับการวางแผนอัตรากำลัง และสนับสนุนการตัดสินใจเชิงกลยุทธ์ของผู้บริหารได้อย่างรวดเร็วและมีประสิทธิภาพยิ่งขึ้น และ Sedamkar และ Tiwari [5] ได้พัฒนาระบบ Interactive Data Visualization Dashboard โดยอาศัยเทคโนโลยีเว็บร่วมกับไลบรารี Chart.js ซึ่งผลการวิจัยพบว่า การแปลงข้อมูลที่มีความซับซ้อนให้อยู่ในรูปแบบกราฟิกเชิงโต้ตอบ ช่วยเสริมสร้างศักยภาพให้ผู้บริหารสามารถวิเคราะห์ข้อมูลได้ด้วยตนเอง (Self-Service Analytics) อย่างรวดเร็ว พร้อมทั้งลดภาระงานของนักพัฒนาระบบได้อย่างมีประสิทธิภาพ

## 3. วิธีการดำเนินการวิจัย

### 3.1 การรวบรวมและเตรียมข้อมูล

การพัฒนาระบบวิเคราะห์และบริหารทรัพยากรบุคคลที่ขับเคลื่อนด้วยข้อมูล กรณีศึกษา วิทยาลัยเซนต์หลุยส์ ได้พัฒนากระบวนการนำเข้าข้อมูล (Data Ingestion) แบบอัตโนมัติ โดยเชื่อมต่อและดึงข้อมูลโดยตรงจากฐานข้อมูลปฏิบัติการผ่านคำสั่ง SQL ขั้นสูงแบบข้ามฐานข้อมูล (Cross-database Queries) เพื่อทดแทนการอัปโหลดข้อมูลแบบดั้งเดิม (Manual Excel/CSV Upload) ซึ่งช่วยลดความซ้ำซ้อนในการจัดการข้อมูล และเอื้อให้ระบบสามารถประมวลผลพร้อมแสดงผลสารสนเทศได้แบบเรียลไทม์ (Real-time Processing) โดยระบบได้นำข้อมูลองค์กรย้อนหลัง 3 ปีการศึกษา (พ.ศ. 2565 - 2568) มาวิเคราะห์ในรูปแบบภาพตัดขวาง (Cross-sectional Analysis) ซึ่งบูรณาการข้อมูลมาจาก 5 แหล่งสำคัญ ได้แก่ ฐานข้อมูลบุคลากร (HR Database) ฐานข้อมูลภาระงานสอน (TQF Database) ฐานข้อมูลงานวิจัย (Research Database) ฐานข้อมูลการฝึกอบรม (Register Database) และฐานข้อมูลแผนและงบประมาณ (Plan Database)

3.1.6 การประมวลผลข้อมูลเบื้องต้น

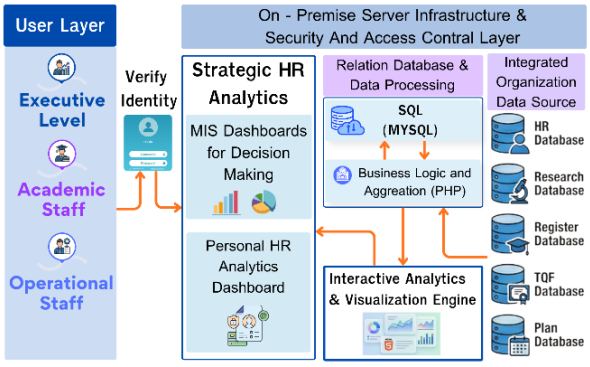
เนื่องจากระบบดึงข้อมูลจากแหล่งที่มาที่หลากหลาย กระบวนการเตรียมข้อมูลจึงถูกออกแบบมาเพื่อให้พร้อมสำหรับการวิเคราะห์ โดยมีการแปลงข้อมูลเชิงเวลา (Temporal Transformation) เช่น การคำนวณอายุงาน หรือการจำแนกวันเวลาตามปีงบประมาณ รวมถึงการจัดกลุ่มข้อมูลมาตรฐานเพื่อจัดหมวดหมู่ประเภทสายงานและตำแหน่งทางวิชาการให้เป็นระบบเดียวกัน ก่อนนำไปเชื่อมโยงข้อมูลเชิงโครงสร้าง (Cross-linking Data Sets) เพื่อรองรับการเจาะลึกข้อมูลแบบเชิงโต้ตอบ

3.1.7 การทำความสะอาดและจัดการข้อมูลที่สูญหาย

เพื่อรักษาความถูกต้องของข้อมูล (Data Integrity) ระบบใช้กลไกจัดการข้อผิดพลาด 3 รูปแบบ ได้แก่ การคัดกรองข้อมูล (Anomaly & Null Filtering) โดยตัดค่าว่าง (NULL) ออกจากการประมวลผล การทดแทนข้อมูล (Structural Fallbacks) โดยดึงข้อมูลย้อนหลังล่าสุดมาทดแทนอัตโนมัติกรณีโครงสร้างสังกัดไม่สมบูรณ์ และการจัดการข้อผิดพลาด (Try-Catch Defaulting) โดยกำหนดค่าเริ่มต้นเป็น 0 เพื่อป้องกันสมการล้มเหลวจากการหารด้วยศูนย์ (Zero Division) ช่วยให้การแสดงผลบนแดชบอร์ดราบรื่น

### 3.2 การออกแบบกรอบแนวคิดการวิจัย

ผู้วิจัยได้ออกแบบกรอบแนวคิดการวิจัยโดยยึดหลักการบริหารจัดการที่ขับเคลื่อนด้วยข้อมูล ตามภาพที่ 1




**ภาพที่ 1:** กรอบแนวคิดการวิจัยของระบบสารสนเทศ

สถาปัตยกรรมการทำงานแบบไคลเอนต์-เซิร์ฟเวอร์ (Client-Server Architecture) โดยแบ่งโครงสร้างการทำงานออกเป็น 3 ส่วนหลัก (Input-Process-Output) ได้แก่

3.2.1 ส่วนนำเข้าข้อมูล (Data Input) การบูรณาการฐานข้อมูลเชิงสัมพันธ์ทั้ง 5 มิติ ได้แก่ บุคลากร ภาระงานสอน งานวิจัย การฝึกอบรม แผนและงบประมาณ

3.2.2 ส่วนประมวลผล (Data Processing) ทำหน้าที่คำนวณตัวชี้วัดและเงื่อนไขเวลาปฏิบัติงานแบบพลวัต โดยยกระดับจากการนำเสนอสถิติพรรณนา สู่ การวิเคราะห์เชิงเปรียบเทียบและแนวโน้ม (Comparative & Trend Analysis) ผ่านการประมวลผลข้อมูลข้ามมิติ (Cross-functional Analytics) ย้อนหลัง 3 ปี (พ.ศ. 2565 - 2568) เช่น การเปรียบเทียบแนวโน้มภาระงานสอนและประสิทธิภาพรายบุคคลเทียบกับองค์กร เพื่อสนับสนุนผู้บริหารในการวางแผนเชิงรุก (Proactive Planning)

3.2.3 ส่วนแสดงผล (Data Output) การนำเสนอข้อมูลสารสนเทศผ่านหน้าการแสดงผลเชิงโต้ตอบ (Interactive Dashboard) ที่มีการแยกสิทธิ์การเข้าถึงข้อมูลตามบทบาทผู้ใช้งาน (Role-Based Access Control) เพื่อสนับสนุนการตัดสินใจของผู้บริหารและให้ข้อมูลย้อนกลับแก่บุคลากรรายบุคคล

**3.3 การพัฒนาระบบวิเคราะห์และบริหารทรัพยากรบุคคล**

การพัฒนาระบบแบ่งออกเป็น 2 ส่วนสำคัญ ได้แก่ การพัฒนาระบบจัดการฐานข้อมูลส่วนหลัง (Backend) และการพัฒนาหน้าจอส่วนต่อประสานกับผู้ใช้งาน (Frontend) ซึ่งมีรายละเอียดดังนี้

3.3.1 การพัฒนาระบบส่วน Backend

ระบบพัฒนาด้วยภาษา PHP เชื่อมต่อฐานข้อมูลผ่าน PDO และ MySQLi โดยใช้คำสั่ง SQL ขั้นสูงเพื่อบูรณาการข้อมูลข้ามระบบ ภายใต้ตรรกะการประมวลผล (Business Logic) 3 องค์ประกอบหลัก ได้แก่ การกำหนดเวลาปฏิบัติงานแบบพลวัต เพื่อรองรับเวลาทำงานที่ยืดหยุ่นตามบริบทของบุคลากรแต่ละกลุ่ม ต่อมาการคำนวณคะแนนประสิทธิภาพรายบุคคล (Productivity Score) โดยถ่วงน้ำหนักจากสถิติการเข้างาน ชั่วโมงอบรม และภาระงาน (FTEs) และ การวิเคราะห์ข้อมูลย้อนหลัง ด้านแนวโน้มการใช้วันหยุด เพื่อสนับสนุนผู้บริหารในการวางแผนอัตรากำลังคนระยะยาวได้อย่างมีประสิทธิภาพ

3.3.2 การพัฒนาส่วนแสดงผลและส่วนต่อประสานผู้ใช้งาน

ระบบนำเสนอสารสนเทศผ่านกระดานข้อมูลเชิงโต้ตอบ (Web-Based Interactive Dashboard) ที่ออกแบบด้วยแนวทาง Glassmorphism UI ร่วมกับ Bootstrap และ Custom CSS โดยแบ่งการแสดงผลเป็น 2 ระดับ ได้แก่ หน้าจอส่วนบุคคล (Personal Analytics) สำหรับบุคลากรติดตามคะแนนประสิทธิภาพ (Productivity Score) ผ่านไลบรารี Chart.js และ หน้าจอสำหรับผู้บริหาร (Administrative Dashboards) เพื่อวิเคราะห์ภาพรวมทรัพยากรบุคคล เช่น อัตราการลาออก ทั้งนี้ ระบบได้จำกัดสิทธิ์การเข้าถึงหน้าจอบริหารไว้เฉพาะผู้บริหารและฝ่ายทรัพยากรบุคคล เพื่อรักษาความปลอดภัยของข้อมูลองค์กร

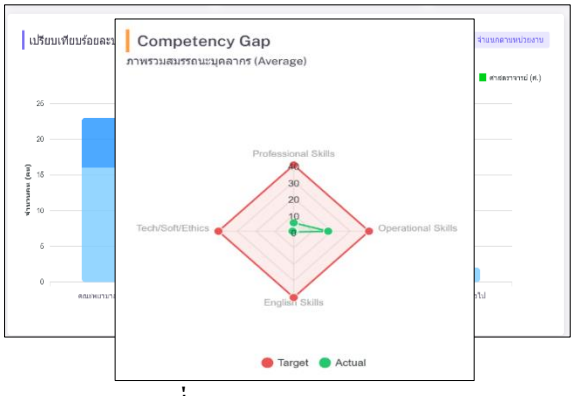
**3.4 การประเมินคุณภาพและความพึงพอใจระบบสารสนเทศ**

ผู้วิจัยได้กำหนดแนวทางการประเมินคุณภาพและการความพึงพอใจของผู้ใช้งาน (User Satisfaction Assessment) ที่แบ่งออกเป็น 3 ด้าน ได้แก่ ด้านความสามารถของระบบตามความต้องการ (Functional Requirement Test) ด้านความถูกต้องของการทำงานของระบบ (Functional Correctness) และด้านความสะดวกและง่ายต่อการใช้งาน (Usability Test)

## 4. ผลดำเนินการวิจัย

**4.1 ผลการพัฒนาระบบวิเคราะห์และบริหารทรัพยากรบุคคล**

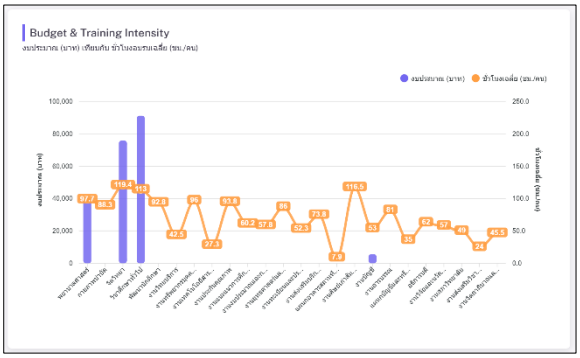
ผลการพัฒนาระบบหน้าจอแสดงผลสำหรับผู้บริหารสามารถแสดงตัวอย่างได้ดังภาพที่ 2-5




**ภาพที่ 2:** สมรรถนะบุคลากรโดยภาพรวม

โดยภาพที่ 2 แสดงสมรรถนะที่แท้จริงของบุคลากรในปัจจุบัน มีค่าที่ต่ำกว่าgเป้าหมายที่องค์กรกำหนดไว้อย่างมากในทุกมิติ ซึ่งข้อมูลส่วนนี้มักใช้เพื่อนำไปวางแผนเพื่อพัฒนาบุคลากรต่อไป

ภาพที่ 3 แสดงให้ผู้บริหารเห็นประเด็นที่น่าสนใจ คือ การมีชั่วโมงอบรมสูง ไม่จำเป็นต้องใช้งบประมาณที่สูงเสมอไป เพราะหลายหน่วยงานกราฟเส้นสีส้มสูง แต่ไม่มีแท่งสีม่วงเลย อาจใช้วิธีอบรมภายในหรือคอร์สฟรี ในขณะที่บางหน่วยงานใช้งบประมาณที่สูงมาก แต่กลับได้ชั่วโมงอบรมเฉลี่ยพอๆ กับหน่วยงานที่ไม่ใช้งบ ซึ่งผู้บริหารสามารถใช้ข้อมูลตรงนี้ไปตรวจสอบความคุ้มค่าในการจัดสรรงบประมาณ (ROI) ของแต่ละหน่วยงานได้

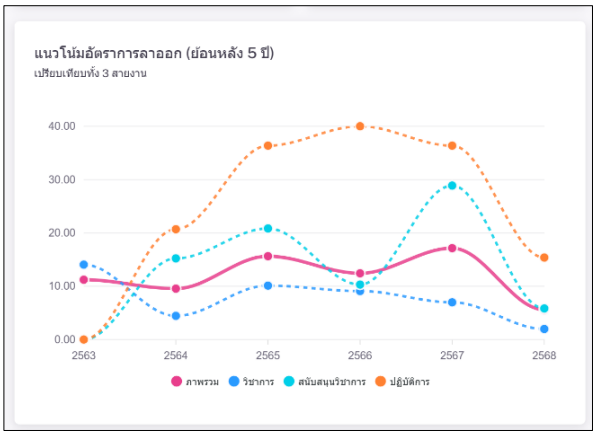



**ภาพที่ 3:** ภาพรวมการใช้งบประมาณการอบรมของบุคลากร

ภาพที่ 4 จะเป็นการแสดงภาพจำนวนอาจารย์ที่มีตำแหน่งทางวิชการของแต่ละคณะ เพื่อให้ผู้บริหารได้เห็นภาพรวมของอาจารย์ในเรื่องของการนำไปวางแผนสนับสนุนงบประมาณหรือสร้างแรงจูงใจ เพื่อผลักดันให้บุคลากรในคณะอื่นๆ ขอตำแหน่งทางวิชาการเพิ่มขึ้น และสนับสนุนให้เกิดตำแหน่งระดับ ศาสตราจารย์ ต่อไปในอนาคต

**ภาพที่ 4:** การเปรียบเทียบตำแหน่งทางวิชาการของแต่ละคณะวิชา

ภาพที่ 5 กราฟนี้ทำหน้าที่เป็นสัญญาณเตือนให้ผู้บริหารเห็นว่า ในช่วงปี 2565-2567 องค์กรมีปัญหาในการรักษาบุคลากร สายปฏิบัติการ และสายสนับสนุนวิชาการ เป็นอย่างมาก ซึ่ง HR ควรเข้าไปตรวจสอบหาสาเหตุ เช่น ภาระงานหนักเกินไป หรือสวัสดิการไม่จูงใจ อย่างไรก็ตาม ทิศทางของกราฟชี้ให้เห็นสัญญาณเชิงบวกว่าในปีล่าสุด 2568 ทุกสายงานมีอัตราการลาออกลดลง

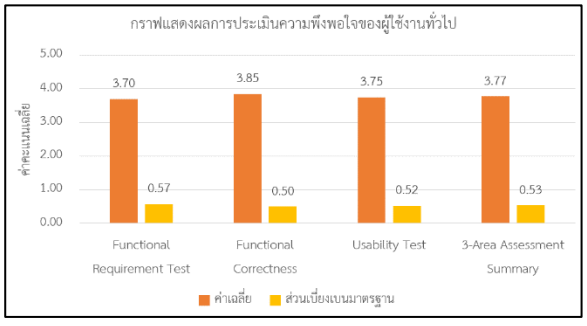



**ภาพที่ 5:** ภาพแสดงแนวโน้มการลาออกของแต่ละปี

เพื่อยกระดับสู่การวิเคราะห์เชิงลึก ระบบได้เพิ่มการวิเคราะห์แนวโน้มเชิงเปรียบเทียบ (Comparative Trend Analysis) ย้อนหลัง 3 ปี พ.ศ. 2565-2568 เช่น การเปรียบเทียบแนวโน้มภาระงานสอนกับกรอบอัตรากำลังจริง (FTEs vs Actual Staff) สารสนเทศนี้ช่วยชี้เป้าจุดวิกฤต (Bottlenecks) ที่เสี่ยงต่อภาวะหมดไฟ (Burnout) ทำให้ผู้บริหารสามารถตัดสินใจบนฐานข้อมูล (Data-Driven Decision Making) เพื่อวางแผนจัดสรรกำลังคนและกระจายภาระงานล่วงหน้าได้อย่างเป็นธรรม

**4.2 ผลการประเมินคุณภาพและความพึงพอใจระบบสารสนเทศ**

ผลการประเมินคุณภาพและความพึงพอใจระบบสารสนเทศของผู้ใช้งาน จำนวน 30 ท่าน โดยภาพรวมอยู่ในระดับมาก ซึ่งมีค่าเฉลี่ยรวม 3.77 และค่าเบี่ยงเบนมาตรฐานเท่ากับ 0.53 ซึ่งสามารถแสดงรายละเอียดได้ดังภาพที่ 6




**ภาพที่ 6:** กราฟแสดงผลการประเมินคุณภาพและความพึงพอใจของผู้ใช้งาน

## 5. สรุปผลการทดลองและข้อเสนอแนะการวิจัย

จากการพัฒนาระบบวิเคราะห์และบริหารทรัพยากรบุคคลที่ขับเคลื่อนด้วยข้อมูล (Data-Driven Human Resource Analytics) สำหรับวิทยาลัยเซนต์หลุยส์ เพื่อมุ่งแก้ไขปัญหาความล่าช้าของระบบเดิม โดยดำเนินการบูรณาการฐานข้อมูลที่สำคัญเข้าด้วยกัน และนำเสนอผ่านหน้าจอแสดงผลข้อมูลเชิงโต้ตอบ (Interactive Dashboard) ภายใต้สถาปัตยกรรมเว็บแอปพลิเคชัน ผลการดำเนินงานพบว่า ระบบสามารถแสดงผลข้อมูลสรุปเชิงลึกได้แบบเรียลไทม์ ซึ่งช่วยลดระยะเวลาในการจัดทำรายงานเฉพาะกิจจากเดิม 3-5 วันทำการลงได้อย่างเป็นรูปธรรม

ผลประเมินความพึงพอใจภาพรวมอยู่ในระดับมาก ค่าเฉลี่ยรวม 3.77 โดดเด่นด้านความถูกต้องของการทำงาน 3.85 และความสะดวกในการใช้งาน 3.75 ผู้ใช้ชื่นชม UI ที่ทันสมัยและสืบค้นข้อมูลได้แม่นยำ เมื่อเทียบกับซอฟต์แวร์สำเร็จรูป Power BI หรือ Tableau ระบบนี้ตอบโจทย์องค์กรมากกว่า 3 ประเด็น คือ ประเด็นแรกไร้ข้อจำกัดด้านต้นทุนค่าลิขสิทธิ์รายบุคคล ประเด็นต่อมามีความยืดหยุ่นในการประมวลผลตรรกะซับซ้อนและจัดการข้อมูลสูญหาย (Fallback Logic) ได้แบบเรียลไทม์ และประเด็นสุดท้ายมีส่วนต่อประสานที่เรียบง่าย แสดงผลตรงตามสิทธิ์ (Role-Based Access) ทำให้ใช้งานได้ทันทีโดยไม่ต้องเรียนรู้เครื่องมือใหม่

การพัฒนาต่อยอดระบบสารสนเทศให้มีความสมบูรณ์ยิ่งขึ้น ผู้วิจัยมีข้อเสนอแนะโดยองค์กรควรขยายขอบเขตการนำเสนอข้อมูลบนกระดานข้อมูล (Dashboard) ให้ครอบคลุมฐานข้อมูลงานวิจัยและแผนงบประมาณ เพื่อสนับสนุนการวิเคราะห์และตัดสินใจของผู้บริหารได้อย่างรอบด้านและมีประสิทธิภาพ ควบคู่ไปกับการยกระดับความน่าเชื่อถือของข้อมูล (Data Integrity) โดยสร้างกลไกการตรวจสอบความถูกต้องของข้อมูล (Data Verification) เทียบเคียงกับฐานข้อมูลดิบจากต้นทางของข้อมูล เพื่อสร้างความเชื่อมั่นก่อนนำผลลัพธ์ไปใช้ประเมินประสิทธิภาพบุคลากร นอกจากนี้ ในการเตรียมความพร้อมก่อนนำระบบไปใช้งานจริง องค์กรควรกำหนดขอบเขตระยะเวลาของข้อมูลย้อนหลังให้ชัดเจน พร้อมทั้งสร้างมาตรฐานการบันทึกข้อมูลของทุกหน่วยงานให้เป็นรูปแบบเดียวกัน เพื่อป้องกันความคลาดเคลื่อนที่อาจเกิดขึ้นจากการเชื่อมโยงและประมวลผลข้อมูลข้ามระบบ

## เอกสารอ้างอิง


[1] Singh, P. (2018). The role of HR analytics in higher education institution. *International Journal of Research and Analytical Reviews (IJRAR)*, 5(1), 346-349.

[2] Al-Sabaeei, A. M., et al. (2024). Web-Based Dashboard of Data Integration for Green Highway Performance Management. *Journal of Engineering Research*, 12(1), 154-165.

[3] Yadav, A. M., & Bhutada, S. (2021). Data-Driven HR Analytics Deployment Using Data Science. *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, 11(3), 1062–1072.

[4] Sushma, R., & Sarala. (2024). The Role in Empowering Strategic Decisions: The Impact of HR Analytics on Talent Management and Organizational Success. *Journal of Information Systems Engineering and Management*, 10(39s), 778–784.

[5] Sedamkar, A., & Tiwari, R. (2024). Interactive Data Visualization Dashboard Using Javascript. *Journal of Emerging Technologies and Innovative Research (JETIR)*, 11(5).

[6] สมเพ็ชร จุลลาบุดดี และ เอกชัย คนคิด. (2568). การพัฒนากระดานข้อมูลเชิงโต้ตอบเพื่อสนับสนุนการตัดสินใจและวางแผนอัตรากำลังของหน่วยงานทรัพยากรบุคคล. *วารสารวิทยาศาสตร์ข้อมูลและเทคโนโลยีสารสนเทศ*, 8(1), 45-58.

# การวิเคราะห์ปัจจัยที่มีอิทธิพลต่อพฤติกรรมการซื้อสินค้าแบบฉับพลัน ของผู้บริโภคผ่านไลฟ์สตรีมมิ่ง

# An Analysis of Factors Influencing Consumers' Impulse Buying Behavior Through Live Streaming

*ภคมญ มาลี (Phakamon Malee)[1] และธนพล เจนสุทธิเวชกุล (Tanapon Jensuttiwetchakul)[2]*
*ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล*
*มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ*
*[1] s6707021856027@email.kmutnb.ac.th [2] tanapon.j@itd.kmutnb.ac.th*

**บทคัดย่อ**

*งานวิจัยนี้มีวัตถุประสงค์เพื่อศึกษาปัจจัยที่มีอิทธิพลต่อพฤติกรรมการซื้อแบบฉับพลันของผู้บริโภคผ่านไลฟ์สตรีม โดยใช้กรอบแนวคิด Stimulus–Organism–Response (S-O-R) และการวิเคราะห์การถดถอยพหุคูณ เก็บข้อมูลจากผู้บริโภคจำนวน 348 ราย ผ่านแบบสอบถามออนไลน์ และวิเคราะห์ด้วยโปรแกรม Jamovi เพื่อตรวจสอบอิทธิพลของความรวดเร็วในการสื่อสาร ความน่าเชื่อถือ และความบันเทิง ต่อคุณค่าที่รับรู้ ความไว้วางใจ และพฤติกรรมการซื้อแบบฉับพลัน ผลการศึกษาพบว่า ความไว้วางใจ ($\beta = 0.413, p < .001$) และคุณค่าด้านความสุข ($\beta = 0.329, p < .001$) เป็นปัจจัยที่มีอิทธิพลสูงสุดต่อพฤติกรรมการซื้อแบบฉับพลัน ขณะที่คุณค่าด้านการใช้งานไม่มีอิทธิพลโดยตรงอย่างมีนัยสำคัญ ทั้งนี้โมเดลสามารถอธิบายพฤติกรรมการซื้อได้ในระดับที่เหมาะสม และเป็นแนวทางในการพัฒนาแพลตฟอร์มไลฟ์สตรีมเพื่อกระตุ้นการตัดสินใจซื้อได้อย่างมีประสิทธิภาพ*

**คำสำคัญ:** ไลฟ์สตรีมคอมเมิร์ซ, พฤติกรรมการซื้อแบบฉับพลัน, ความไว้วางใจ, คุณค่าที่รับรู้, คุณค่าด้านความสุข, S-O-R Model

**Abstract**

*This study examines factors influencing impulse buying behavior in livestream commerce using the Stimulus–Organism–Response (S-O-R) framework. Communication speed, credibility, and entertainment were modeled as stimuli affecting utilitarian value, hedonic value, and trust, which in turn influence impulse purchases. Data from 348 respondents were collected via an online questionnaire and analyzed using multiple regression in Jamovi. Results show that trust ($\beta = 0.413, p < .001$) and hedonic value ($\beta = 0.329, p < .001$) significantly predict impulse buying, while utilitarian value has no significant direct effect. Credibility and entertainment also strengthen value perceptions and trust. These findings clarify impulse buying mechanisms in livestream commerce and provide practical implications for platform design and marketing strategies.*

**Keywords:** Live Streaming Commerce, Impulse Buying Behavior, Perceived Trust, Perceived Value, hedonic value, Stimulus–Organism–Response (S-O-R) Model.

## 1. บทนำ

โลกยุคดิจิทัลการเติบโตของเทคโนโลยีดิจิทัลและมูลค่า e-Commerce ที่ขยายตัวอย่างต่อเนื่อง ได้ผลักดันให้ Live Commerce กลายเป็นช่องทางการซื้อขายที่มีบทบาทสำคัญในเศรษฐกิจดิจิทัลไทย ผู้บริโภคสามารถรับชมสินค้าแบบเรียลไทม์ โต้ตอบกับผู้ขาย และตัดสินใจซื้อได้ทันที ทำให้กระบวนการซื้อมีความใกล้เคียงกับการเลือกซื้อหน้าร้านจริงมากขึ้น อย่างไรก็ตาม พฤติกรรมการซื้อผ่านไลฟ์สตรีมไม่ได้ขึ้นอยู่เพียงด้านสินค้าและราคา แต่ยังเกี่ยวข้องกับประสบการณ์การใช้งาน ความเชื่อมั่นต่อผู้ขาย และอารมณ์ความรู้สึกระหว่างการรับชม ซึ่งอาจนำไปสู่พฤติกรรมการซื้อแบบฉับพลันได้อย่างมีนัยสำคัญ ดังนั้น งานวิจัยนี้จึงมุ่งศึกษาปัจจัยที่มีอิทธิพลต่อพฤติกรรมดังกล่าว จึงมีความสำคัญทั้งใน

เชิงวิชาการและเชิงปฏิบัติเพื่อใช้เป็นแนวทางในการออกแบบกลยุทธ์การตลาดและพัฒนาประสบการณ์ลูกค้าในยุคดิจิทัล

## 2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

### 2.1 ลักษณะของ Live Streaming

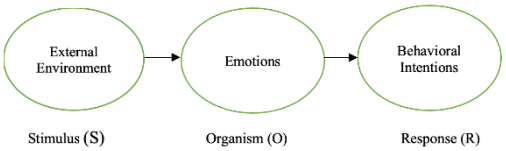
Live Streaming คือการถ่ายทอดวิดีโอแบบเรียลไทม์ที่เปิดโอกาสให้ผู้ชมโต้ตอบได้ทันที และในบริบท Live Streaming Commerce ถูกใช้เป็นช่องทางการขายที่ผสานข้อมูลสินค้าและความบันเทิง ทำให้ประสบการณ์ใกล้เคียงหน้าร้านโดยองค์ประกอบ เช่น วิดีโอสด ระบบแชต และระบบสั่งซื้อช่วยเพิ่มการรับรู้ข้อมูล ความไว้วางใจ และการมีส่วนร่วมทางอารมณ์ ดังนั้น Live Streaming จึงทำหน้าที่เป็นสิ่งเร้า (Stimulus) ตามกรอบ S-O-R ที่นำไปสู่พฤติกรรมการซื้อโดยเฉพาะการซื้อแบบฉับพลัน




ภาพที่ 1 Top 10 กิจกรรมยอดฮิต ขวัญใจสังคมออนไลน์

### 2.2 ทฤษฎีสิ่งเร้า-การรับรู้-การตอบสนอง Stimulus–Organism–Response (S-O-R)

การตัดสินใจซื้อเกิดจากความสัมพันธ์ระหว่างสิ่งเร้า สภาวะภายใน และพฤติกรรมตอบสนอง โดยผู้บริโภคตอบสนองแตกต่างกันตามสภาพแวดล้อม (Tai & Fung, 1997) ทั้งนี้ พฤติกรรมไม่ได้เกิดจากสิ่งเร้าโดยตรง แต่ผ่านกระบวนการรับรู้และประมวลผลภายใน (Organism) ก่อนนำไปสู่การตอบสนอง (Response) โดย Stimulus คือปัจจัยภายนอก เช่น ราคา โฆษณา หรือคุณลักษณะแพลตฟอร์ม Organism คือการรับรู้และอารมณ์ เช่น ความไว้วางใจและการรับรู้คุณค่า และ Response คือพฤติกรรม เช่น ความตั้งใจซื้อหรือการซื้อจริง ดังนั้น กรอบ S-O-R จึงอธิบายกลไกที่เชื่อมโยงปัจจัยภายนอกกับพฤติกรรมผู้บริโภคได้อย่างเป็นระบบ และเหมาะสมต่อการศึกษาบริบทไลฟ์สตรีม




ภาพที่ 2 กรอบแนวคิด SOR

### 2.3 ทฤษฎีเกี่ยวกับ Stimulus (สิ่งเร้า) ในบริบท Live Streaming

2.3.1 ลักษณะเฉพาะของระบบ (System Characteristics)

มีพื้นฐานจาก IT Affordance Theory ที่อธิบายว่าเทคโนโลยีเปิดโอกาสให้ผู้ใช้กระทำบางอย่างได้ และถูกประยุกต์สู่บริบท Live Streaming Commerce โดยลักษณะของระบบประกอบด้วยด้านเทคโนโลยี (Media Richness, Visibility, Personalization) และด้านสังคม (Real-time Interaction, Social Presence, Entertainment) ภายใต้กรอบ S-O-R คุณลักษณะเหล่านี้ทำหน้าที่เป็นสิ่งเร้า (Stimulus) ที่ส่งผลต่อความไว้วางใจ การรับรู้คุณค่า และอารมณ์ของผู้บริโภค ก่อนนำไปสู่ความตั้งใจซื้อหรือการซื้อแบบฉับพลัน ทำให้ระบบ Live Streaming สามารถกระตุ้นทั้งเหตุผลและอารมณ์ได้อย่างมีนัยสำคัญ

2.3.2 ลักษณะของสตรีมเมอร์ (Streamer Characteristics/Anchor Characteristics)

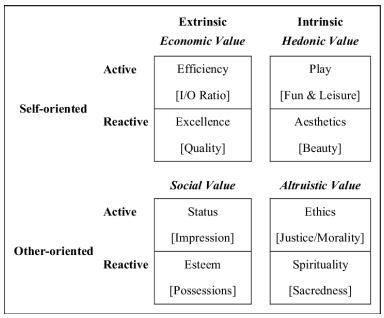
คุณลักษณะของสตรีมเมอร์เป็นปัจจัยสำคัญต่อความสำเร็จของ Live Streaming Commerce โดยครอบคลุมทั้งด้านบุคคล (ความน่าดึงดูด บุคลิกภาพ การสื่อสาร) และด้านความเป็นมืออาชีพ (ความรู้สินค้า ความชัดเจน ความจริงใจ) ซึ่งส่งผลต่อความไว้วางใจและความน่าเชื่อถือ (Trust, Expertise, Attractiveness) ภายใต้กรอบ S-O-R ลักษณะของสตรีมเมอร์ทำหน้าที่เป็นสิ่งเร้า (Stimulus) ที่กระตุ้นความไว้วางใจและอารมณ์เชิงบวกของผู้ชม ช่วยลดความลังเล และเพิ่มแนวโน้มการตัดสินใจซื้อหรือการซื้อแบบฉับพลันอย่างมีนัยสำคัญ

### 2.4 ทฤษฎีเกี่ยวกับ Organism (สภาวะภายใน) ของผู้บริโภค

2.4.1 ทฤษฎีมูลค่าการบริโภค (Consumer Value Theory)

ทฤษฎีมูลค่าผู้บริโภคของ Holbrook (1994) อธิบายว่า “คุณค่าที่รับรู้” เกิดจากประสบการณ์ตลอดกระบวนการซื้อ ไม่ได้มาจากตัวสินค้าเพียงอย่างเดียว โดยแบ่งเป็น คุณค่าด้านประโยชน์ใช้สอย (Utilitarian Value) และคุณค่าด้านความเพลิดเพลิน (Hedonic Value) ในบริบท Live Streaming Commerce คุณค่าทั้งสองมิติทำหน้าที่เป็น

Organism ตามกรอบ S-O-R ซึ่งส่งผลต่อความตั้งใจซื้อและการซื้อแบบฉับพลันของผู้บริโภค




**ภาพที่ 3** Holbrook's typology of consumer value (adapted from Holbrook 2005, 2006)

2.4.2 ทฤษฎีความไว้วางใจ (Trust Theory)

ความไว้วางใจ (Trust) คือความเชื่อมั่นว่าผู้ขายมีความซื่อสัตย์ น่าเชื่อถือ และไม่เอาเปรียบผู้บริโภค โดย Mayer, Davis และ Schoorman (1995) อธิบายว่า Trust ประกอบด้วย Ability, Integrity และ Benevolence ซึ่งช่วยลดความไม่แน่นอนและความเสี่ยงในการตัดสินใจซื้อ ในบริบท Live Streaming Commerce ความไว้วางใจเกิดจากความเป็นมืออาชีพ การนำเสนอที่ชัดเจน และการโต้ตอบอย่างจริงใจของสตรีมเมอร์ โดยภายใต้กรอบ S-O-R ความไว้วางใจทำหน้าที่เป็น Organism ที่นำไปสู่ความตั้งใจซื้อหรือการซื้อแบบฉับพลันอย่างมีนัยสำคัญ

**2.5 ทฤษฎีเกี่ยวกับ Response (พฤติกรรมตอบสนอง)**

2.5.1 ทฤษฎีการซื้อแบบฉับพลัน (Impulse Buying Theory)

ทฤษฎีการซื้อแบบฉับพลันของ Stern (1962) และ Rook (1987) อธิบายว่า การซื้อที่เกิดขึ้นอย่างรวดเร็วและไม่วางแผนล่วงหน้าเป็นผลจากสิ่งเร้า อารมณ์ และบริบทของสถานการณ์ร่วมกันในบริบท Live Streaming Commerce ปัจจัย เช่น การโต้ตอบแบบเรียลไทม์ ความน่าเชื่อถือ ความบันเทิง และความเร่งด่วน สามารถกระตุ้นการรับรู้คุณค่า ความไว้วางใจ และอารมณ์ของผู้บริโภค ก่อนนำไปสู่ พฤติกรรมการซื้อแบบฉับพลันตามกรอบ S-O-R

**2.6 ทฤษฎีทางสถิติที่ใช้ในการวิเคราะห์**

2.6.1 การวิเคราะห์องค์ประกอบเชิงสำรวจ (EFA)

การวิเคราะห์องค์ประกอบเชิงสำรวจ (EFA) เป็นเทคนิคทางสถิติที่ใช้ตรวจสอบโครงสร้างแฝงของข้อมูลและความตรงเชิงโครงสร้างของแบบสอบถาม โดยพิจารณาค่า KMO (> 0.60), Bartlett's Test ($p < 0.05$), Eigenvalue > 1 และ Factor Loading > 0.50 งานวิจัยนี้ใช้ EFA เพื่อตรวจสอบโครงสร้างของตัวแปรตามกรอบ S-O-R ก่อนนำไปวิเคราะห์ความสัมพันธ์ในขั้นตอนถัดไป เพื่อยืนยันคุณภาพเครื่องมือและเพิ่มความน่าเชื่อถือของผลการศึกษา

2.6.2 การวิเคราะห์ถดถอยพหุคูณ (MRA)

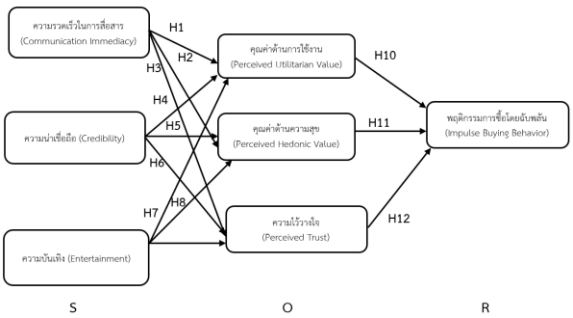
การวิเคราะห์ถดถอยพหุคูณ (MRA) เป็นเทคนิคทางสถิติที่ใช้ทดสอบอิทธิพลของตัวแปรอิสระหลายตัวต่อตัวแปรตามหนึ่งตัวในรูปสมการเชิงเส้น เพื่อทดสอบสมมติฐาน H1–H12 ตามกรอบ S-O-R โดยพิจารณาค่า $R^2$ เพื่อประเมินความสามารถของโมเดล ค่า p-value ($< 0.05$) เพื่อทดสอบนัยสำคัญทางสถิติ และค่า VIF ($< 10$) เพื่อตรวจสอบปัญหา Multicollinearity ดังนี้

$$Y = \beta_0 + \beta_1 X_1 + \cdots + \beta_n X_n + \varepsilon$$

**3. วิธีการดำเนินการวิจัย**

**3.1 กรอบแนวคิดการวิจัย**

งานวิจัยนี้พัฒนากรอบแนวคิดโดยบูรณาการแนวคิด Stimulus–Organism–Response (S-O-R) กับบริบท Live Streaming Commerce โดยกำหนดให้ปัจจัยกระตุ้น ได้แก่ ความรวดเร็วในการสื่อสาร (CI) ความน่าเชื่อถือ (CRD) และความบันเทิง (ENT) ส่งผลต่อสภาวะภายในของผู้บริโภค ได้แก่ การรับรู้คุณค่าด้านการใช้งาน (PUV) คุณค่าด้านความสุข (PHV) และความไว้วางใจ (TRU) ซึ่งนำไปสู่พฤติกรรมตอบสนอง คือ การซื้อโดยฉับพลัน (IBB) สะท้อนกลไกที่สิ่งเร้าจากไลฟ์สตรีมมีอิทธิพลต่อการรับรู้และอารมณ์ ก่อนนำไปสู่การตัดสินใจซื้ออย่างฉับพลัน




ภาพที่ 4 กรอบแนวคิดการวิจัย

**3.2 ประชากรและกลุ่มตัวอย่าง**

กลุ่มตัวอย่างคือผู้บริโภคที่เคยรับชมและซื้อสินค้าแบบฉับพลันผ่าน Live Streaming Commerce อย่างน้อย 1 ครั้ง ในช่วง 6 เดือนที่ผ่านมา โดยใช้วิธีการสุ่มแบบง่าย (Simple Random Sampling) ผู้วิจัยกำหนดขนาดตัวอย่างด้วยโปรแกรม G*Power สำหรับการวิเคราะห์ทางสถิติ โดยตั้งค่า effect size = 0.3 (ขนาดกลางตามเกณฑ์ของ Cohen) ระดับนัยสำคัญ 0.05 และระดับพลังการทดสอบ (Power) 0.80 กำหนดค่า Degree of Freedom = 33 จากจำนวนข้อคำถามในแบบสอบถาม ผลการคำนวณพบว่าขนาดกลุ่มตัวอย่างที่เหมาะสมคือ 284 คน

**3.3 ออกแบบเครื่องมือที่ใช้ในการวิจัย**

3.3.1 แบบสอบถาม

เครื่องมือวิจัยเป็นแบบสอบถามออนไลน์ผ่าน MS Forms จำนวน 33 ข้อ ใช้มาตราส่วน Likert 7 ระดับ เพื่อวัดปัจจัยที่มีอิทธิพลต่อพฤติกรรมการซื้อแบบฉับพลันผ่านไลฟ์สตรีม แบ่งเป็น 9 ส่วน ได้แก่ คัดกรองข้อมูล ประชากรศาสตร์ CI, CRD, ENT, PUV, PHV, TRU และ IBB โดยตรวจสอบคุณภาพเครื่องมือด้วยความตรงเชิงเนื้อหา (Content Validity) และความเชื่อมั่น (Cronbach's Alpha $\geq$ 0.70) เพื่อยืนยันความสม่ำเสมอของการวัดตัวแปร

**3.4 เก็บรวบรวมข้อมูลวิจัย**

การเก็บรวบรวมข้อมูลดำเนินการผ่าน MS Forms โดยเผยแพร่แบบสอบถามผ่านช่องทางออนไลน์ เช่น Facebook และ Line เป็นระยะเวลา 3 เดือน พร้อมประชาสัมพันธ์และติดตามการตอบกลับอย่างต่อเนื่อง เพื่อให้ได้กลุ่มตัวอย่างที่เพียงพอและมีความหลากหลาย จากนั้นทำการตรวจสอบและคัดกรองคำตอบที่ไม่สมบูรณ์ เพื่อให้ได้ข้อมูลที่มีคุณภาพสำหรับการวิเคราะห์ต่อไป

**3.5 วิเคราะห์ข้อมูลด้วยโปรแกรมทางสถิติ Jamovi**

ผู้วิจัยตรวจสอบความถูกต้องและจัดหมวดหมู่ข้อมูลก่อนวิเคราะห์ด้วยโปรแกรม Jamovi โดยใช้ สถิติเชิงพรรณนา อธิบายลักษณะกลุ่มตัวอย่าง ใช้ EFA ตรวจสอบโครงสร้างตัวแปรและความเชื่อมั่นของแบบสอบถาม และใช้ Multiple Regression Analysis ทดสอบความสัมพันธ์ระหว่างตัวแปรตามกรอบแนวคิด S-O-R ที่ระดับนัยสำคัญทางสถิติ 0.05

**4. ผลการดำเนินงาน**

**4.1 ผลการวิเคราะห์ข้อมูลทั่วไปของกลุ่มตัวอย่าง**

ผลการสำรวจพบว่าเป็นไปตามตารางดังนี้

ตารางที่ 1 ผลการวิเคราะห์ข้อมูลประชากรศาสตร์

| ข้อมูลทั่วไป | จำนวน (คน) | ร้อยละ (%) |
|---|---|---|
| **เพศ** | | |
| ชาย | 114 | 32.80 |
| หญิง | 220 | 63.20 |
| อื่นๆ | 14 | 4.00 |
| **รวม** | 348 | 100.00 |
| **อายุ** | | |
| ต่ำกว่า 18 ปี | 19 | 5.50 |
| 18 - 24 ปี | 210 | 60.3 |
| 25 - 34 ปี | 79 | 22.7 |
| 35 - 44 ปี | 21 | 6.0 |
| 45 ปีขึ้นไป | 19 | 5.50 |
| **รวม** | 348 | 100.00 |
| **รายได้** | | |
| ต่ำกว่า 15,000 บาท | 25 | 7.20 |
| 15,001 – 30,000 บาท | 230 | 66.10 |
| 30,001 – 45,000 บาท | 54 | 15.50 |
| มากกว่า 45,000 บาท | 107 | 34.29 |
| **รวม** | 348 | 100.00 |

**4.2 ผลการวิเคราะห์การกระจายตัวของข้อมูลด้วยสถิติเชิงพรรณนา**

การวิเคราะห์ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานพบว่า ทุกปัจจัยมีค่าเฉลี่ยอยู่ในระดับสูง (มากกว่า 5.00 จากมาตราส่วน 7 ระดับ) สะท้อนว่ากลุ่มตัวอย่างมีทัศนคติเชิงบวกต่อการซื้อผ่านไลฟ์สตรีม และมีแนวโน้มเกิดพฤติกรรมการซื้อแบบฉับพลันในระดับค่อนข้างสูง โดยให้ความสำคัญต่อปัจจัยด้านการสื่อสาร ความน่าเชื่อถือ ความบันเทิง คุณค่าด้านการใช้งาน คุณค่าด้านความสุข และความไว้วางใจในระดับสูง ทั้งนี้ตัวแปรด้านอารมณ์และพฤติกรรมมีการกระจายตัวมากกว่า สะท้อนถึงความแตกต่างระหว่างบุคคลในด้านอารมณ์และการตัดสินใจซื้อแบบฉับพลัน

**4.3 ผลการวิเคราะห์องค์ประกอบเชิงสำรวจ**

ผลการวิเคราะห์องค์ประกอบเชิงสำรวจ (EFA) ด้วยโปรแกรม Jamovi พบว่า ค่า KMO Overall เท่ากับ 0.927 อยู่ในระดับดีมาก (>0.70)

ตารางที่ 2 ค่า KMO Test

| **KMO Test** | |
|---|---|
| Kaiser-Meyer-Olkin Measure of Sampling Adequacy (KMO) | MSA |
| Overall | 0.927 |

แสดงว่าข้อมูลมีความเหมาะสมอย่างยิ่งสำหรับการวิเคราะห์องค์ประกอบ โดยใช้วิธี Maximum Likelihood และหมุนแกนแบบ Oblimin ตามตารางดังนี้

ตารางที่ 3 Exploratory Factor Analysis (Factor Loadings)

| | Factor | | | | | | |
|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| PUV4 | 0.752 | | | | | | |
| PUV3 | 0.709 | | | | | | |
| PUV2 | 0.707 | | | | | | |
| PUV1 | 0.612 | | | | | | |
| PUV5 | 0.570 | | | | | | |
| PUV6 | 0.523 | | | | | | |
| IBB4 | | 0.943 | | | | | |
| IBB3 | | 0.927 | | | | | |
| IBB5 | | 0.701 | | | | | |
| IBB6 | | 0.627 | | | | | |
| TRU4 | | | 0.733 | | | | |
| TRU3 | | | 0.495 | | | | |
| TRU6 | | | 0.470 | | | | |
| TRU5 | | | 0.455 | | | | |
| TRU7 | | | 0.405 | | | | |
| ENT3 | | | | 0.797 | | | |
| ENT2 | | | | 0.651 | | | |
| ENT1 | | | | 0.601 | | | |
| CI1 | | | | | 0.727 | | |
| CI2 | | | | | 0.714 | | |

ตารางที่ 3 Exploratory Factor Analysis (Factor Loadings) (ต่อ)

| | Factor | | | | | | |
|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| CI3 | | | | | 0.590 | | |
| PHV1 | | | | | | 0.768 | |
| PHV2 | | | | | | 0.561 | |
| PHV4 | | | | | | 0.339 | |
| CRD3 | | | | | | | 0.669 |
| CRD2 | | | | | | | 0.526 |
| CRD1 | | | | | | | 0.505 |

**4.4 การทดสอบความเชื่อมโยงของแบบสอบถาม**

ผลการทดสอบความเชื่อมั่นด้วยค่า Cronbach's Alpha จากกลุ่มตัวอย่าง 348 คน พบว่าค่าความเชื่อมั่นอยู่ระหว่าง 0.764–0.938 ซึ่งสูงกว่าเกณฑ์ 0.70 แสดงว่าแบบสอบถามมีความสอดคล้องภายในในระดับดีถึงดีมาก ดังตารางที่ 4

ตารางที่ 4 การทดสอบความเชื่อถือของแบบสอบถาม

| ตัวแปร | Mean | S.D. | Cronbach's α |
|---|---|---|---|
| CI (Communication Immediacy) | 6.12 | 0.701 | 0.764 |
| CRD (Credibility) | 5.88 | 0.738 | 0.780 |
| ENT (Entertainment) | 5.98 | 0.809 | 0.871 |
| PUV (Perceived Utilitarian Value) | 5.92 | 0.715 | 0.909 |
| PHV (Perceived Hedonic Value) | 5.77 | 0.945 | 0.907 |
| TRU (Perceived Trust) | 5.71 | 0.811 | 0.922 |
| IBB (Impulse Buying Behavior) | 5.35 | 1.29 | 0.938 |
| CI (Communication Immediacy) | 6.12 | 0.701 | 0.764 |

4.5 ผลการวิเคราะห์สมการถดถอยพหุคูณ

จากกรอบแนวคิดการวิจัยครั้งนี้ ในการทดสอบสมมุติฐานงานวิจัยมีการวิเคราะห์สมการถดถอยพหุคูณ ดังนี้

ตารางที่ 5 ผลการวิเคราะห์สมการถดถอยพหุคูณสำหรับ Perceived Utilitarian Value (PUV)

| Model | R | $R^2$ | Adjusted $R^2$ | Sample Size (N) | | |
|---|---|---|---|---|---|---|
| 1 | 0.686 | 0.470 | 0.465 | 348 | | |
| ค่าสัมประสิทธิ์ถดถอย (Regression Coefficients – PUV) | | | | | | |
| Predictors | Estimate | SE | p | Stand. Estimate | VIF | Tolerance |
| Intercept | 1.118 | 0.2982 | < .001 | - | - | - |
| CI | 0.120 | 0.0337 | < .001 | 0.160 | 1.31 | 0.764 |
| CRD | 0.344 | 0.0477 | < .001 | 0.355 | 1.57 | 0.636 |
| ENT | 0.287 | 0.0400 | < .001 | 0.330 | 1.37 | 0.728 |

ผลการวิเคราะห์พบว่า ตัวแปรอิสระไม่มีปัญหา Multicollinearity (VIF = 1.31–1.57, Tolerance = 0.636–0.764) โดย CI, CRD และ ENT มีอิทธิพลเชิงบวกต่อ PUV อย่างมีนัยสำคัญทางสถิติ ซึ่ง CRD มีอิทธิพลสูงสุด รองลงมาคือ ENT และ CI ทั้งนี้ โมเดลมีค่า $R = 0.686$, $R^2 = 0.470$ และ Adjusted $R^2 = 0.465$ แสดงว่าสามารถอธิบายความแปรปรวนของ PUV ได้ร้อยละ 46.5–47.0 และมีความเหมาะสมในระดับค่อนข้างดี

ตารางที่ 6 ผลการวิเคราะห์สมการถดถอยพหุคูณสำหรับ PHV (Perceived Hedonic Value)

| Model | R | $R^2$ | Adjusted $R^2$ | Sample Size (N) | | |
|---|---|---|---|---|---|---|
| 2 | 0.787 | 0.470 | 0.615 | 348 | | |
| ค่าสัมประสิทธิ์ถดถอย (Regression Coefficients – PHV) | | | | | | |
| Predictors | Estimate | SE | p | Stand.Estimate | VIF | Tolerance |
| Intercept | -1.001 | 0.3393 | 0.003 | - | - | - |
| CI | 0.763 | 0.0456 | < .001 | 0.6540 | 1.31 | 0.764 |
| CRD | 0.129 | 0.0543 | 0.018 | 0.0996 | 1.57 | 0.636 |
| ENT | 0.166 | 0.0384 | < .001 | 0.1647 | 1.37 | 0.728 |

ผลการวิเคราะห์พบว่า ตัวแปรอิสระทุกตัวไม่มีปัญหา Multicollinearity (VIF = 1.31–1.57, Tolerance = 0.636–0.764) โดย CI, CRD และ ENT มีอิทธิพลเชิงบวกต่อ PHV อย่างมีนัยสำคัญทางสถิติ ซึ่ง CI มีอิทธิพลสูงสุด รองลงมาคือ ENT และ CRD ทั้งนี้ โมเดลมีค่า $R = 0.787$, $R^2 = 0.619$ และ Adjusted $R^2 = 0.615$ แสดงว่าสามารถอธิบายความแปรปรวนของ PHV ได้ร้อยละ 61.5–61.9 และมีความเหมาะสมในระดับที่ดี

ตารางที่ 7 ผลการวิเคราะห์สมการถดถอยพหุคูณสำหรับ Perceived Trust (TRU)

| Model | R | $R^2$ | Adjusted $R^2$ | Sample Size (N) |
|---|---|---|---|---|
| 3 | 0.743 | 0.552 | 0.549 | 348 |
| ค่าสัมประสิทธิ์ถดถอย (Regression Coefficients – TRU) | | | | |

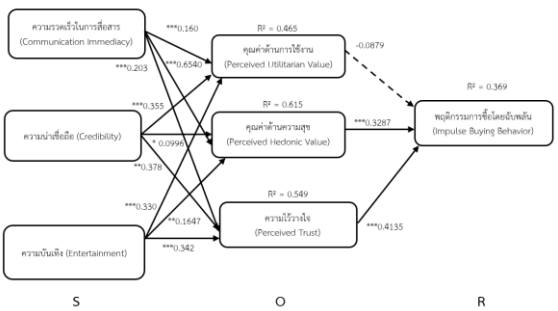
| Predictors | Estimate | SE | p | Stand. Estimate | VIF | Tolerance |
|---|---|---|---|---|---|---|
| Intercept | -0.542 | 0.3230 | 0.094 | - | - | - |
| CI | 0.180 | 0.0366 | < .001 | 0.203 | 1.31 | 0.764 |
| CRD | 0.432 | 0.0517 | < .001 | 0.378 | 1.57 | 0.636 |
| ENT | 0.350 | 0.0434 | < .001 | 0.342 | 1.37 | 0.728 |

ผลการวิเคราะห์พบว่า ตัวแปรอิสระทุกตัวไม่มีปัญหา Multicollinearity (VIF = 1.31–1.57, Tolerance = 0.636–0.764) โดย CI, CRD และ ENT มีอิทธิพลเชิงบวกต่อ TRU อย่างมีนัยสำคัญทางสถิติ ซึ่ง CRD มีอิทธิพลสูงสุด รองลงมาคือ ENT และ CI ทั้งนี้ โมเดลมีค่า $R = 0.743$, $R^2 = 0.552$ และ Adjusted $R^2 = 0.549$ แสดงว่าสามารถอธิบายความแปรปรวนของ TRU ได้ร้อยละ 54.9–55.2 และมีความเหมาะสมในระดับที่ดี

ตารางที่ 8 ผลการวิเคราะห์สมการถดถอยพหุคูณสำหรับ Impulse Buying Behavior (IBB)

| Model | R | $R^2$ | Adjusted $R^2$ | Sample Size (N) | | |
|---|---|---|---|---|---|---|
| 4 | 0.612 | 0.374 | 0.369 | 348 | | |
| ค่าสัมประสิทธิ์ถดถอย (Regression Coefficients – TRU) | | | | | | |
| Predictors | Estimate | SE | p | Stand. Estimate | VIF | Tolerance |
| Intercept | 0.612 | 0.5148 | 0.323 | - | - | - |
| TRU | 0.700 | 0.1178 | < .001 | 0.4135 | 2.66 | 0.376 |
| PHV | 0.489 | 0.0857 | < .001 | 0.3287 | 1.82 | 0.548 |
| PUV | -0.175 | 0.1322 | 0.185 | -0.0879 | 2.41 | 0.415 |

ผลการวิเคราะห์พบว่า ตัวแปรอิสระทุกตัวไม่มีปัญหา Multicollinearity (VIF = 1.82–2.66, Tolerance = 0.376–0.548) โดย TRU และ PHV มีอิทธิพลเชิงบวกต่อ IBB อย่างมีนัยสำคัญทางสถิติ ซึ่ง TRU มีอิทธิพลสูงสุด รองลงมาคือ PHV ขณะที่ PUV ไม่ส่งผลอย่างมีนัยสำคัญ ทั้งนี้ โมเดลมีค่า $R = 0.612$, $R^2 = 0.374$ และ Adjusted $R^2 = 0.369$ แสดงว่าสามารถอธิบายความแปรปรวนของ IBB ได้ร้อยละ 36.9–37.4 และมีความเหมาะสมในระดับปานกลางจากผลการวิเคราะห์การถดถอยพหุคูณของทั้ง 4 โมเดล สามารถสรุปความสัมพันธ์ระหว่างตัวแปรและค่าสัมประสิทธิ์การถดถอยมาตรฐานได้ตามภาพดังนี้




ภาพที่ 5 การวิเคราะห์ผล

## 5. สรุปผลและอภิปราย

5.1 สรุปผลการวิจัย กลุ่มตัวอย่างจำนวน 348 คน ส่วนใหญ่เป็นเพศหญิง มีอายุระหว่าง 18–24 ปี และมีรายได้เฉลี่ย 15,000–30,000 บาท สะท้อนถึงกลุ่มวัยศึกษาและวัยเริ่มต้นทำงานที่มีความคุ้นเคยกับเทคโนโลยีดิจิทัล และมีศักยภาพสูงต่อการมีส่วนร่วมในการรับชมและการซื้อสินค้าผ่านไลฟ์สตรีมมิ่ง

5.2 อภิปรายผล จากผลการทดสอบสมมติฐานภายใต้

ผลการวิจัยภายใต้กรอบแนวคิด S-O-R พบว่า ความรวดเร็วในการสื่อสาร (CI) ความน่าเชื่อถือ (CRD) และความบันเทิง (ENT) มีอิทธิพลเชิงบวกต่อการรับรู้ภายใน ได้แก่ คุณค่าด้านการใช้งาน (PUV) คุณค่าด้านความสุข (PHV) และความไว้วางใจ (TRU) อย่างมีนัยสำคัญทางสถิติ สำหรับพฤติกรรมการซื้อโดยฉับพลัน (IBB) พบว่า PHV และ TRU มีอิทธิพลเชิงบวกอย่างมีนัยสำคัญ ขณะที่ PUV ไม่มีอิทธิพลอย่างมีนัยสำคัญ สะท้อนว่าการตัดสินใจซื้อแบบฉับพลันในบริบทไลฟ์สตรีมถูกขับเคลื่อนโดยปัจจัยด้านอารมณ์และความไว้วางใจมากกว่าการประเมินเชิงเหตุผล โดยเฉพาะความไว้วางใจซึ่งมีอิทธิพลสูงสุด

5.3 ข้อเสนอแนะ การวิจัยในอนาคตควรขยายขอบเขตกลุ่มตัวอย่างให้มีความหลากหลายมากยิ่งขึ้น และศึกษาการเปรียบเทียบระหว่างแพลตฟอร์มไลฟ์สตรีมที่แตกต่างกัน รวมทั้งพิจารณาปัจจัยเพิ่มเติม เช่น Flow Experience หรือกลยุทธ์ส่งเสริมการขายแบบจำกัดเวลา ตลอดจนประยุกต์ใช้วิธีการวิจัยเชิงผสมผสานหรือการวิเคราะห์สมการโครงสร้าง (SEM) เพื่อเพิ่มความแม่นยำและความน่าเชื่อถือของแบบจำลองในการอธิบายพฤติกรรมผู้บริโภค

## เอกสารอ้างอิง

[1] ธมลวรรณ สมพงษ์. (2564). การศึกษาปัจจัยที่มีผลต่อการตัดสินใจซื้อสินค้าทางออนไลน์ของผู้บริโภคในจังหวัดนครปฐม. วิทยานิพนธ์ คณะ วิศวกรรมศาสตร์และเทคโนโลยีอุตสาหกรรม มหาวิทยาลัยศิลปากร

[2] ปิยฉัตร โมฑะกุล. (2566). การศึกษาอิทธิพลของการฝังโฆษณาในวิดีโอสั้นที่ส่งต่อความตั้งใจซื้อสินค้าบนแพลตฟอร์มวิดีโอสั้นของผู้บริโภคในประเทศไทย. การ

ค้นคว้าอิสระ คณะ พาณิชยศาสตรและการบัญชี มหาวิทยาลัยธรรมศาสตร์

[3] มนทวัส ศรีภูสิตโต. (2564). ปัจจัยที่ส่งผลต่อการตัดสินใจรับชมการสตรีมมิ่งผ่านทางแพลตฟอร์มทวิตช์ของผู้ชมในประเทศไทย. การค้นคว้าอิสระ คณะพาณิชยศาสตร์และการบัญชี มหาวิทยาลัยธรรมศาสตร์

[4] ปิยเวทย์ วรไกวัล. (2566) ปัจจัยที่ส่งผลต่อการซื้อแบบไม่ได้วางแผนล่วงหน้าในแพลตฟอร์มช้อปปี้ไลฟ์ (SHOPEE LIVE) : กรณีศึกษาผลิตภัณฑ์หมวดความงาม การค้นคว้าอิสระ คณะพาณิชยศาสตร์และการบัญชี มหาวิทยาลัยธรรมศาสตร์

[5] Van Tinh Hoang, Huu Phuc Dang. (2024) Exploring How the Characteristics of Live Streaming Affect Impulse Buying Behaviour in Live Streaming Commerce: The Mediating Effect of Trust and Flow Experience. Journal School of Management, University of Economics Ho Chi Minh City

# ปัจจัยที่ส่งผลต่อความตั้งใจเชิงพฤติกรรมการใช้งานระบบบริการตัวเองผ่าน QR Code

# Factors Affecting Behavioral Intention to Use Self-Service System via QR Code

*นพษพจน์ศรีสุวรรณ (Nopsapoj Srisuwan)[1] และ ศักดิ์ชาย ตั้งวรรณวิทย์ (Sakchai Tangwannawit)[2]*
*ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล*
*มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ*
*[1]s6707021857155@email.kmutnb.ac.th, [2] sakchai_t@itd.kmutnb.ac.th*

## บทคัดย่อ

*การใช้งานเทคโนโลยีบริการตัวเอง (Self Service Technology :SST) โดยเฉพาะระบบ QR Code ได้รับความนิยมเพิ่มขึ้นอย่างรวดเร็ว โดยมีปัจจัยเร่งจากการระบาดของ COVID-19 ที่กระตุ้นความต้องการใช้งานบริการแบบไร้สัมผัส ประกอบกับวิถีชีวิตแบบปัจเจกนิยมสมัยใหม่ที่ผู้บริโภคต้องการความเรียบง่ายและพึ่งพาตนเองมากขึ้น งานวิจัยนี้จึงมีวัตถุประสงค์เพื่อศึกษาปัจจัยที่ส่งผลต่อความตั้งใจเชิงพฤติกรรมในการใช้ QR Code SST โดยใช้ทฤษฎีการยอมรับเทคโนโลยี (Technology Acceptance Model: TAM) และทฤษฎีความสอดคล้องระหว่างงานและเทคโนโลยี (Task – Technology Fit: TTF) เป็นกรอบแนวคิด โดยเก็บข้อมูลผ่านแบบสอบถามออนไลน์และวิเคราะห์ด้วยสถิติเชิงพรรณนา เพื่อทำความเข้าใจพฤติกรรมผู้บริโภคและเพิ่มประสิทธิภาพการบริการในยุคดิจิทัล*

**คำสำคัญ:** การยอมรับเทคโนโลยี ประเมินประสิทธิภาพของเทคโนโลยี พฤติกรรมผู้บริโภค เทคโนโลยีบริการตัวเอง

## Abstract

*The use of self-service technology (SST), particularly QR code systems, has rapidly increased in popularity. This is fueled by the COVID-19 pandemic, which has stimulated demand for contactless services, coupled with modern individualistic lifestyles where consumers desire greater simplicity and self-reliance. This research aims to study the factors affecting behavioral intentions to use QR code SSTs, using the Technology Acceptance Model (TAM) and Task-Technology Fit (TTF) theories as conceptual frameworks. Data was collected via an online questionnaire and analyzed using descriptive statistics to understand consumer behavior and enhance service efficiency in the digital age.*

**Keywords:** Technology acceptance, Technology effectiveness assessment, Consumer behavior, Self Service Technology.

## 1. บทนำ

การเปลี่ยนแปลงทางดิจิทัลได้เร่งตัวขึ้นอย่างต่อเนื่อง โดยเฉพาะการใช้งานเทคโนโลยีบริการตัวเอง (Self-Service Technology: SST) ซึ่งมีบทบาทสำคัญในการเพิ่มประสิทธิภาพการให้บริการ ระบบบริการตัวเองผ่าน QR Code ได้รับความนิยมอย่างแพร่หลายในหลายอุตสาหกรรม เนื่องจากช่วยอำนวยความสะดวก รวดเร็ว และลดการพึ่งพาพนักงานโดยตรง

การแพร่ระบาดของ COVID-19 เป็นปัจจัยสำคัญที่เร่งให้เกิดการใช้งานบริการแบบไร้สัมผัส ส่งผลให้ผู้บริโภคให้ความสำคัญกับความสะดวก ความเป็นส่วนตัว และการ

ควบคุมตนเองมากขึ้น ขณะเดียวกัน ประเทศไทยถือเป็นหนึ่งในประเทศที่มีการนำ QR Code มาใช้อย่างแพร่หลาย

ดังนั้น การทำความเข้าใจปัจจัยที่ส่งผลต่อความตั้งใจเชิงพฤติกรรมในการใช้งานระบบบริการตัวเองผ่าน QR Code จึงมีความสำคัญต่อการพัฒนาและออกแบบบริการให้สอดคล้องกับความต้องการของผู้บริโภคในยุคดิจิทัล

โดยงานวิจัยนี้มีวัตถุประสงค์ 2 ข้อ คือ 1) เพื่อศึกษาปัจจัยต่างๆ ที่มีอิทธิพลต่อความตั้งใจเชิงพฤติกรรมการใช้งานระบบบริการตัวเองผ่าน QR Code และ 2) เพื่อวิเคราะห์ปัจจัยที่มีอิทธิพลต่อลักษณะของงานและเทคโนโลยีที่มีผลต่อความตั้งใจเชิงพฤติกรรมการใช้งานระบบบริการตัวเองผ่าน QR Code

## 2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ผู้วิจัยได้ศึกษาค้นคว้าเอกสาร และงานวิจัยที่เกี่ยวข้อง โดยมีการศึกษาค้นคว้าข้อมูล แนวคิด ทฤษฎี และงานวิจัยต่างๆ ที่เกี่ยวข้องเพื่อใช้เป็นแนวทางการศึกษา ดังนี้

### 2.1 ทฤษฎีการยอมรับเทคโนโลยี (Technology Acceptance Model: TAM)

พัฒนาโดย Davis (1986) จากทฤษฎีการกระทำตามหลักเหตุและผล(TRA) เพื่ออธิบายปัจจัยที่มีอิทธิพลต่อความตั้งใจในการใช้งานเทคโนโลยีสารสนเทศ โดยประกอบด้วยตัวแปรสำคัญ ได้แก่ การรับรู้ประโยชน์ (Perceived Usefulness: PU) และการรับรู้ความง่ายในการใช้งาน (Perceived Ease of Use: PEOU) ซึ่งมีผลต่อทัศนคติในการใช้งานและนำไปสู่ความตั้งใจเชิงพฤติกรรมและการใช้งานจริง [**1**]

### 2.2 ทฤษฎีความเหมาะสมของงานกับเทคโนโลยี (Task–Technology Fit: TTF)

Goodhue and Thompson (1995) ได้อธิบายความสอดคล้องระหว่างลักษณะของงานและความสามารถของเทคโนโลยี โดยเทคโนโลยีจะถูกนำมาใช้งานเมื่อสามารถสนับสนุนการทำงานของผู้ใช้ได้อย่างมีประสิทธิภาพและเกิดประโยชน์สูงสุด ทั้งนี้ TTF ประกอบด้วยปัจจัยด้านลักษณะงานและลักษณะเทคโนโลยี ซึ่งส่งผลต่อประสิทธิภาพการทำงานและการใช้ประโยชน์ของระบบ [**2**]

งานวิจัยนี้ได้บูรณาการทฤษฎีการยอมรับเทคโนโลยี (TAM) และทฤษฎีความเหมาะสมของงานกับเทคโนโลยี (TTF) ช่วยเพิ่มความสามารถในการอธิบายพฤติกรรมการใช้งานระบบสารสนเทศได้อย่างครอบคลุมมากขึ้น โดย TTF อธิบายความสอดคล้องระหว่างงานและเทคโนโลยีที่ส่งผลต่อประสิทธิภาพการใช้งาน ขณะที่ TAM อธิบายการรับรู้ของผู้ใช้ เช่น การรับรู้ประโยชน์และความง่ายในการใช้งาน เป็นต้น ซึ่งมีผลต่อความตั้งใจเชิงพฤติกรรม ดังนั้น การผสานทั้งสองทฤษฎีทำให้สามารถอธิบายทั้งมิติด้านเทคโนโลยี งาน และพฤติกรรมผู้ใช้ได้อย่างมีประสิทธิภาพ

### 2.3 งานวิจัยที่เกี่ยวข้อง

ธีรวัตร (2561) ได้ศึกษาปัจจัยที่มีผลต่อความตั้งใจในการใช้บริการ QR Code Payment ผ่าน Mobile Banking ในเขตกรุงเทพมหานครและปริมณฑล โดยใช้การวิจัยเชิงปริมาณ สำรวจกลุ่มตัวอย่างผู้ใช้แอปพลิเคชันจำนวน 160 คน โดยบูรณาการแนวคิดส่วนประสมทางการตลาด ทฤษฎีการกระทำตามหลักเหตุผล (TRA) แบบจำลองการยอมรับเทคโนโลยี (TAM) และแนวคิดด้านความเสี่ยงและความเชื่อมั่น ผลการวิจัยพบว่า ปัจจัยเชิงประจักษ์ที่ส่งผลต่อความตั้งใจในการใช้บริการอย่างมีนัยสำคัญ ได้แก่ การรับรู้ถึงประโยชน์ที่ได้รับ การรับรู้ถึงความง่ายในการใช้งาน และปัจจัยด้านการส่งเสริมการตลาด [**3**]

สุวรรณี (2562) ได้ศึกษาพฤติกรรมและปัจจัยที่มีผลต่อการยอมรับการใช้บริการชำระเงินด้วย QR Code ผ่านโทรศัพท์เคลื่อนที่ในเขตกรุงเทพมหานคร โดยเก็บรวบรวมข้อมูลจากกลุ่มตัวอย่างจำนวน 400 คน และใช้การวิเคราะห์การถดถอยพหุคูณในการทดสอบสมมติฐาน ผลการวิจัยพบว่า ปัจจัยที่ส่งผลเชิงบวกต่อความถี่ในการใช้งานมี 7 ปัจจัย โดยปัจจัยที่ส่งผลมากที่สุดคือ ความเคยชิน รองลงมาคือ จำนวนปีการศึกษา เพศชาย มูลค่าราคา ทัศนคติ การรับรู้ถึงประโยชน์ และรายได้ ตามลำดับ [**4**]

สถาพร (2563) ได้ทำการศึกษาวิจัย เรื่อง การเปรียบเทียบปัจจัยที่มีผลต่อการยอมรับนวัตกรรมในการใช้คิวอาร์โค้ดสำหรับร้านอาหารในกัวลาลัมเปอร์ ประเทศมาเลเซียและกรุงเทพมหานคร ประเทศไทย มีวัตถุประสงค์ 3 ข้อ คือ 1) เพื่อศึกษาระดับการยอมรับ เทคโนโลยี QR

Code ภายในร้านอาหาร 2) เพื่อเปรียบเทียบการยอมรับเทคโนโลยี QR Code ภายใน ร้านอาหาร 3) เพื่อศึกษาปัจจัยการยอมรับเทคโนโลยีคิวอาร์โค้ดภายในร้านอาหาร กลุ่มตัวอย่างคือ ผู้เข้ารับประทานอาหารในร้านอาหาร ในพื้นที่กรุงกัวลาลัมเปอร์ ประเทศมาเลเซีย และกรุงเทพมหานคร ประเทศไทย จำนวน 400 คน ใช้วิธีการสุ่มตัวอย่างแบบตามสะดวก เครื่องมือที่ใช้ ในการวิจัย คือ แบบสอบถาม การวิเคราะห์ข้อมูลโดยใช้สถิติเชิงพรรณนาและสถิติเชิงอนุมาน คือ การวิเคราะห์ด้วย F-test และ t-test การวิเคราะห์โมเดลสมการโครงสร้าง [5]

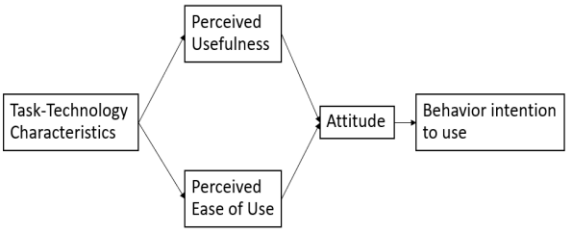
Van Anh Thi Nguyen (2567) ได้ศึกษาปัจจัยที่มีผลต่อพฤติกรรมการยอมรับการใช้เมนู QR-code ของคนรุ่น Gen Z ในเมือง Can Tho โดยการสำรวจผู้ตอบแบบสอบถาม 310 รายที่เกิดระหว่างปี 1995 ถึง 2012 จากตรงนั้น เราจะเห็นได้ว่าพฤติกรรมการใช้เมนู QR Code ออนไลน์ให้คำแนะนำแก่ธุรกิจในภาคส่วนอาหารและเครื่องดื่มว่าจะสร้างการเปลี่ยนแปลงที่จำเป็นอย่างแท้จริงเพื่อให้ทันกลุ่มลูกค้ากลุ่มนี้ควบคู่ไปกับการพัฒนาของเทคโนโลยี เมนู QR Code ออนไลน์ในธุรกิจในภาคส่วนอาหารและเครื่องดื่มได้กลายเป็นกระแสประสบการณ์และเป็นที่ชื่นชอบของลูกค้า [6]

(Mohammed Alhassan Jubur, 2567) ได้ทำการศึกษาวิจัย เรื่อง Comparative Technical Analysis of QR Code and NFC in Contactless Payments การวิเคราะห์เผยให้เห็นข้อดีและจุดอ่อนที่แตกต่างกันโดยเน้นย้ำถึงความสำคัญของนวัตกรรมอย่างต่อเนื่องและมาตรการรักษาความปลอดภัยที่แข็งแกร่งเพื่อลดความเสี่ยงที่อาจเกิดขึ้นเอกสารนี้สรุปด้วยข้อมูลเชิงลึกที่มองไปข้างหน้า [7]

## 3. วิธีการดำเนินการวิจัย

### 3.1 กรอบแนวคิดงานวิจัย

อ้างอิงจากงานวิจัยที่เกี่ยวข้องกับทฤษฎีการยอมรับเทคโนโลยีและทฤษฎีความเหมาะสมของงานกับเทคโนโลยีนำมาประยุกต์เข้ากันเพื่อหาความต่างและความสัมพันธ์ของแต่ละปัจจัยว่า ส่งผลต่อความตั้งใจเชิงพฤติกรรมการใช้งานระบบแสดงดังนี้

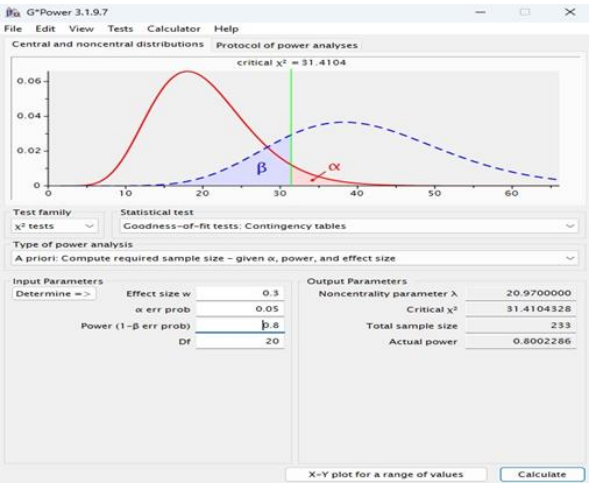



**ภาพที่ 1** กรอบแนวคิดงานวิจัย

### 3.2 ประชากรและกลุ่มตัวอย่าง

ในการศึกษางานวิจัยนี้ได้กำหนดประชาชนกลุ่มตัวอย่างที่ใช้ในงานวิจัย ได้แก่ ประชากรทั่วไปเคยใช้งานระบบบริการตัวเองผ่านร้านอาหารเพื่อสั่งอาหารหรือห้างสรรพสินค้าเพื่อดูรายการสินค้า

จากกลุ่มประชากรเพื่อหากลุ่มตัวอย่างโดยใช้โปรแกรม G*Power กำหนดค่า ซึ่งมีจำนวนกลุ่มตัวอย่างมากกว่าจำนวนที่กำหนดเอาไว้อยู่ที่ 334 คนนำมาศึกษา ใช้ระยะเวลาเก็บข้อมูลตั้งแต่ กรกฎาคม 2568 ถึง ตุลาคม 2568 โดยเครื่องมือที่ใช้ ได้แก่ Google Forms




**ภาพที่ 2** โปรแกรม G*Power

### 3.3 เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการศึกษาการวิจัยคือแบบสอบถาม Google Forms ในการสร้างแบบสอบถามและ เก็บรวบรวมข้อมูลใช้คำถามแบบปลายปิดแบบสอบถามแบ่งออกเป็น 2 ส่วน ดังนี้

ส่วนที่ 1 คำถามทั่วไป สำหรับคัดกรองกลุ่มประชากร

ส่วนที่ 2 คำถามเกี่ยวกับการใช้งานระบบบริการตัวเอง QR Code

จากผลการตรวจสอบข้อคำถามในแบบสอบถามมีจำนวนทั้งสิ้น 34 ข้อคำถามพบว่าค่าความเชื่อมั่นอยู่ในระดับ 0.976

## 4. ผลการดำเนินการ

### 4.1 ผลการวิเคราะห์ข้อมูลทั่วไปของกลุ่มตัวอย่าง

**ตารางที่ 1** ผลการวิเคราะห์ข้อมูลทั่วไปของกลุ่มตัวอย่าง

| ข้อมูลด้านประชากร | จำนวน | ร้อยละ |
|---|---|---|
| **เพศ** | | |
| ชาย | 106 | 31.63 |
| หญิง | 221 | 66.17 |
| ไม่ระบุเพศ | 7 | 2.10 |
| **ช่วงอายุ** | | |
| ต่ำกว่า 20 ปี | 115 | 34.43 |
| 21 – 30 ปี | 174 | 52.09 |
| 31 – 40 ปี | 10 | 2.99 |
| 41 – 50 ปี | 13 | 3.89 |
| มากกว่า 50 ปี | 22 | 6.60 |
| **ระดับการศึกษา** | | |
| ต่ำกว่าปริญญาตรี | 103 | 30.84 |
| ปริญญาตรี | 211 | 63.17 |
| สูงกว่าปริญญาตรี | 20 | 5.99 |
| **อาชีพ** | | |
| นักเรียน/นักศึกษา | 261 | 78.14 |
| พนักงานบริษัท | 32 | 9.58 |
| ข้าราชการ/รัฐวิสาหกิจ | 22 | 6.58 |
| เจ้าของกิจการ/ฟรีแลนด์ | 7 | 2.10 |
| แม่บ้าน/เกษียณอายุ | 6 | 1.80 |
| อื่นๆ | 6 | 1.80 |
| **เคยใช้งานระบบบริการตัวเองผ่าน QR Code ในร้านอาหารหรือไม่?** | | |
| เคย | 309 | 92.51 |
| ไม่เคย | 25 | 7.49 |

ผลของข้อมูลด้านประชากรกลุ่มตัวอย่าง พบว่าเป็นเพศชายจำนวน 106 คน เพศหญิงจำนวน 221 คน และไม่ระบุเพศ 7 คน มีคนเคยใช้งานระบบบริการตัวเองผ่าน QR Code ในร้านอาหาร 309 คนและมีคนไม่เคยใช้งานระบบบริการตัวเองผ่าน QR Code ในร้านอาหาร 25 คน

### 4.2 ผลการวิเคราะห์สถิติเบื้องต้น

**ตารางที่ 2** ผลการวิเคราะห์สถิติเบื้องต้น

| ตัวแปรรายด้าน | Mean | SD | CV (%) | แปลผล |
|---|---|---|---|---|
| ปัจจัยด้านลักษณะของงานและเทคโนโลยี(TTC) | 5.53 | 1.45 | 26.22% | เห็นด้วย |
| ปัจจัยด้านการรับรู้ประโยชน์(PU) | 5.48 | 1.56 | 28.46% | เห็นด้วยบ้าง |
| ปัจจัยด้านการรับรู้ความใช้งานง่าย(PEU) | 5.49 | 1.49 | 27.14% | เห็นด้วยบ้าง |
| ปัจจัยด้านทัศนคติ(AT) | 5.50 | 1.5 | 27.27% | เห็นด้วยบ้าง |

ผู้วิจัยได้กำหนดข้อคำถามโดยใช้รูปแบบมาตรวัดประมาณค่า(Rating Scale) โดยใช้เกณฑ์การหาความกว้างของอันตรภาคชั้นซึ่งแบ่งระดับการประเมินออกเป็น 7 ระดับดังนี้ช่วงคะแนนเฉลี่ย 6.41 - 7.00 แปลผลว่าอยู่ในระดับ เห็นด้วยอย่างยิ่ง ช่วงคะแนนเฉลี่ย 5.51 - 6.40 แปลผลว่าอยู่ในระดับ เห็นด้วย ช่วงคะแนนเฉลี่ย 4.61 - 5.50 แปลผลว่าอยู่ในระดับ เห็นด้วยบ้าง ช่วงคะแนนเฉลี่ย 3.71 - 4.60 แปลผลว่าอยู่ในระดับ ไม่แน่ใจช่วงคะแนนเฉลี่ย 2.81 - 3.70 แปลผลว่าอยู่ในระดับ ไม่เห็นด้วยบ้าง ช่วงคะแนนเฉลี่ย 1.91 - 2.80 แปลผลว่าอยู่ในระดับ ไม่เห็น

ด้วย ช่วงคะแนนเฉลี่ย 1.00 - 1.90 แปลผลว่าอยู่ในระดับไม่เห็นด้วยอย่างยิ่ง

ผลการวิเคราะห์ปัจจัยที่ส่งผลต่อความตั้งใจเชิงพฤติกรรมการใช้งานระบบบริการตัวเองผ่าน QR Code มีค่าเฉลี่ยมากที่สุดคือ ปัจจัยด้านลักษณะของงานและเทคโนโลยี ต่อมาคือ ปัจจัยด้านทัศนคติ ตามด้วย ปัจจัยด้านการรับรู้ความใช้งานง่าย และส่วนสุดท้ายคือ ปัจจัยด้านการรับรู้ประโยชน์ ตามลำดับ

**4.3 ผลการวิเคราะห์สถิติ ANOVA**

จากข้อมูลที่ได้นำไปวิเคราะห์เพื่อใช้ในการเปรียบเทียบความแตกต่างของประชากรด้านต่างๆ จำแนกตามเพศมีค่า Sig. .961, .819, .714, .780 จำแนกตามช่วงอายุมีค่า Sig. .222, .406, .191, .072 จำแนกตามระดับการศึกษามีค่า Sig. .519, .329, .288, .290 จำแนกตามอาชีพมีค่า Sig. .468, .750, .748, .665 พบว่าปัจจัยด้านลักษณะของงานและเทคโนโลยี,ปัจจัยด้านการรับรู้ประโยชน์,ปัจจัยด้านการรับรู้ความใช้งานง่าย,ปัจจัยด้านทัศนคติมีค่า Sig. มากกว่าค่านัยสำคัญทางสถิติ ที่ระดับ .05 ดังนั้นจึงสามารถสรุปตามสมมติฐานตามข้อ H1 ถึงข้อ H5 ที่ส่งผลต่อความตั้งใจเชิงพฤติกรรมการใช้งานระบบบริการตัวเองผ่าน QR Code ไม่แตกต่างกัน อย่างไม่มีมีนัยสำคัญทางสถิติที่ระดับ .05

**4.4 ผลการวิเคราะห์ค่าสหสัมพันธ์**

**ตารางที่ 3** ผลการวิเคราะห์ค่าสหสัมพันธ์

| **Correlations** | | TTC | PU | PEU | AT | BIU |
|---|---|---|---|---|---|---|
| TTC | Pearson Correlation | 1 | .803** | .761** | .783** | .642** |
| | Sig. (2-tailed) | | .000 | .000 | .000 | .000 |
| PU | Pearson Correlation | .803** | 1 | .735** | .824** | .651** |
| | Sig. (2-tailed) | .000 | | .000 | .000 | .000 |
| PEU | Pearson Correlation | .761** | .735** | 1 | .740** | .590** |
| | Sig. (2-tailed) | .000 | .000 | | .000 | .000 |
| AT | Pearson Correlation | .783** | .824** | .740** | 1 | .758** |
| | Sig. (2-tailed) | .000 | .000 | .000 | | .000 |

จากการทดสอบความสัมพันธ์โดยการวิเคราะห์ค่าสัมประสิทธิ์สหสัมพันธ์ Pearson Correlation Coefficient สามารถสรุปได้ว่า ผลการวิเคราะห์สหสัมพันธ์ระหว่างตัวแปรอิสระ มีความสัมพันธ์กับปัจจัยที่ส่งผลต่อความตั้งใจเชิงพฤติกรรมการใช้งานระบบบริการตัวเองผ่าน QR Code Self Services Technology อย่างมีนัยสำคัญที่ระดับ .05 และ .01

**4.5 ผลการวิเคราะห์การถดถอยเชิงพหุคุณ**

**ตารางที่ 4** แสดงการวิเคราะห์ถดถอยพหุคูณปัจจัยที่ส่งผลต่อความตั้งใจเชิงพฤติกรรม

| **Model Summary** | | |
|---|---|---|
| Model | R | R Square |
| 1 | .803[a] | .644 |
| a. Predictors: (Constant), TTC | | |
| b. Dependent Variable: PU | | |
| 2 | .761[a] | .579 |
| a. Predictors: (Constant), TTC | | |
| b. Dependent Variable: PEU | | |
| 3 | .848[a] | .718 |
| a. Predictors: (Constant), PU, PEU | | |
| b. Dependent Variable: A | | |
| 4 | .758[a] | .575 |
| a. Predictors: (Constant), A | | |
| b. Dependent Variable: BIU | | |

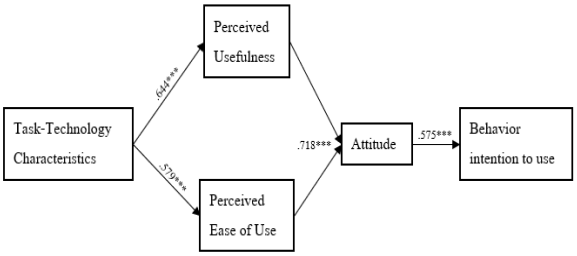
โมเดลที่ 1 มีอิทธิพลของตัวแปรอิสระ ได้แก่ ปัจจัยด้านการรับรู้ประโยชน์ โดยสามารถ อธิบายความผันแปรของปัจจัยที่ส่งผลต่อความตั้งใจเชิงพฤติกรรมการใช้งานระบบบริการตัวเองผ่าน QR Code จากผลลัพธ์มีค่า R Square อยู่ที่ .644

โมเดลที่ 2 มีอิทธิพลของตัวแปรอิสระ ได้แก่ ปัจจัยด้านการรับรู้ความใช้งานง่าย อธิบายความผันแปรของปัจจัยที่ส่งผลต่อความตั้งใจเชิงพฤติกรรมการใช้งานระบบบริการตัวเองผ่าน QR Code จากผลลัพธ์มีค่า R Square อยู่ที่ .579

โมเดลที่ 3 มีอิทธิพลของตัวแปรอิสระ ได้แก่ ปัจจัยด้านทัศนคติ อธิบายความผันแปรของปัจจัยที่ส่งผลต่อความตั้งใจเชิงพฤติกรรมการใช้งานระบบบริการตัวเองผ่าน QR Code จากผลลัพธ์มีค่า R Square อยู่ที่ .718

โมเดลที่ 4 มีอิทธิพลของตัวแปรอิสระ ได้แก่ ปัจจัยด้านความตั้งใจเชิงพฤติกรรมอธิบายความผันแปรของปัจจัยที่ส่งผลต่อความตั้งใจเชิงพฤติกรรมการใช้งานระบบ

บริการตัวเองผ่าน QR Code จากผลลัพธ์มีค่า R Square อยู่ที่ .575




**ภาพที่ 3** โมเดลที่อธิบายปัจจัยที่ส่งผลต่อความตั้งใจเชิงพฤติกรรมการใช้งานระบบบริการตัวเองผ่าน QR Code

## 5.สรุปผลการวิจัย

จากการวิเคราะห์การถดถอยเชิงพหุคูณซึ่งสามารถทดสอบสมมติฐานตามภาพที่ 3 ได้ดังนี้

สมมติฐานที่ 1 (H1): ลักษณะการใช้งานของ QR Code มีอิทธิพลต่อการรับรู้ประโยชน์ ผลจากการวิเคราะห์พบว่า ยอมรับสมมติฐาน หมายถึง ลักษณะของงานและเทคโนโลยีส่งผลเชิงบวกต่อการรับรู้ประโยชน์อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

สมมติฐานที่ 2 (H2): ลักษณะการใช้งานของ QR Code มีอิทธิพลต่อการรับรู้ความใช้งานง่าย ผลจากการวิเคราะห์พบว่า ยอมรับสมมติฐาน หมายถึง ลักษณะของงานและเทคโนโลยีส่งผลเชิงบวกต่อการรับรู้ความใช้งานง่ายอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

สมมติฐานที่ 3 (H3): การรับรู้ประโยชน์มีอิทธิพลต่อทัศนคติต่อการใช้งาน ผลจากการวิเคราะห์พบว่า ยอมรับสมมติฐาน หมายถึง ปัจจัยด้านการรับรู้ประโยชน์ ส่งผลเชิงบวกต่อทัศนคติได้อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

สมมติฐานที่ 4 (H4): การรับรู้ถึงความใช้งานง่ายมีอิทธิพลต่อทัศนคติต่อการใช้งาน ผลจากการวิเคราะห์พบว่า ยอมรับสมมติฐาน หมายถึง ปัจจัยด้านการรับรู้ความใช้งานง่ายส่งผลเชิงบวกต่อทัศนคติได้อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

สมมติฐานที่ 5 (H5): ทัศนคติต่อการใช้งาน QR Code มีอิทธิพลต่อความตั้งใจใช้งานระบบ QR Code ผลจากการวิเคราะห์พบว่า ยอมรับสมมติฐาน หมายถึง ปัจจัยด้านทัศนคติส่งผลเชิงบวกต่อความตั้งใจเชิงพฤติกรรมอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

## เอกสารอ้างอิง

[1] Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly, 13(3), 319–340.

[2] Goodhue, D. L., & Thompson, R. L. (1995). Task-technology fit and individual performance. MIS Quarterly, 19(2), 213–236.

[3] ธีรวัตร อมรรัตนกิจ. (2561). ปัจจัยที่มีผลต่อความตั้งใจในการใช้บริการ QR code payment ผ่าน mobile banking ในเขตกรุงเทพมหานครและปริมณฑล. มหาวิทยาลัยธรรมศาสตร์.

[4] สุวรรณี มาน้อย. (2562). ปัจจัยที่มีผลต่อการยอมรับการใช้บริการชำระเงินด้วย QR Code ผ่านโทรศัพท์เคลื่อนที่ ในเขตกรุงเทพมหานคร. มหาวิทยาลัยธุรกิจบัณฑิตย์.

[5] สถาพร เลิศกมลา. (2563). การเปรียบเทียบปัจจัยที่มีผลต่อการยอมรับนวัตกรรมในการใช้คิวอาร์โค้ดสำหรับร้านอาหารในกัวลาลัมเปอร์ ประเทศมาเลเซียและกรุงเทพมหานคร ประเทศไทย มหาวิทยาลัยศิลปากร.

[6] Van Anh Thi Nguyen. (2024). Online Menu by QR-Code and Behavioral Intention to adopt it on Mobile of Gen Z, research for the Food and Beverage industry in Can Tho City, Viet Nam.

[7] Mohammed Alhassan Jubur. (2024). Comparative Technical Analysis of QR Code and NFC in Contactless Payments.

# ปัจจัยที่มีอิทธิพลต่อความตั้งใจในการใช้บริการเช่ารถแบบขับเอง ผ่านช่องทางโซเชียลมีเดียในประเทศไทย

# Factors Influencing Consumers' Behavioral Intention to Use Self-Drive Car Rental Services via Social Media in Thailand

*พัชรินทร์ ศิริรักษ์ (Pacharin Sirirak)[1] และศักดิ์ชาย ตั้งวรรณวิทย์ (Sakchai Tangwannawit)[2]*

*ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล*

*มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ*

*[1]S6707021857210@email.kmutnb.ac.th, [2]sakchai.t@itd.kmutnb.ac.th*

## บทคัดย่อ

*การค้นคว้าอิสระนี้มีวัตถุประสงค์เพื่อศึกษาปัจจัยที่มีอิทธิพลต่อความตั้งใจในการใช้บริการเช่ารถแบบขับเองผ่านช่องทางโซเชียลมีเดียในประเทศไทย กลุ่มตัวอย่างที่ใช้ในการวิจัยคือผู้บริโภคจำนวน 250 คน โดยใช้แบบสอบถามเป็นเครื่องมือในการเก็บรวบรวมข้อมูล สถิติที่ใช้ในการวิเคราะห์ ได้แก่ ความถี่ ร้อยละ ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน และการวิเคราะห์สมการถดถอยพหุคูณ ผลการวิจัยพบว่า ระดับความคิดเห็นของผู้บริโภคต่อปัจจัยที่ศึกษาโดยรวมอยู่ในระดับมาก โดยปัจจัยการรับรู้ความง่ายในการใช้งานมีค่าเฉลี่ยสูงที่สุด รองลงมาคืออิทธิพลทางสังคมและความตั้งใจในการใช้บริการ ขณะที่การรับรู้ความเสี่ยงอยู่ในระดับปานกลาง นอกจากนี้ผลการทดสอบความเชื่อมั่นของแบบสอบถามพบว่าทุกปัจจัยมีค่าความเชื่อมั่นมากกว่า 0.70 ซึ่งอยู่ในเกณฑ์ที่ยอมรับได้ การวิเคราะห์สมการถดถอยพหุคูณพบว่า ตัวแปรอิสระสามารถอธิบายความแปรปรวนของความตั้งใจในการใช้บริการได้ร้อยละ 39.20 โดยปัจจัยการรับรู้ความเสี่ยงมีอิทธิพลต่อความตั้งใจในการใช้บริการอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 และมีความสัมพันธ์ในทิศทางลบ ขณะที่การรับรู้ความง่ายในการใช้งานและอิทธิพลทางสังคมพบว่ามีอิทธิพลอย่างมีนัยสำคัญทางสถิติ*

**คำสำคัญ:** การรับรู้ความง่ายในการใช้งาน การรับรู้ความเสี่ยง อิทธิพลทางสังคม ความตั้งใจในการใช้บริการ โซเชียลมีเดีย

## Abstract

*The Independent study aimed to examine the Factors Influencing the Selection of Self-Drive Car Rental Services via Social Media in Thailand. The sample consisted of 250 consumers, and a questionnaire was used as the data collection instrument. The statistical methods employed for data analysis included frequency, percentage, mean, standard deviation, and multiple linear regression analysis. The results indicated that the overall level of consumers' opinions toward the studied factors was at a high level. Among the factors, perceived ease of use had the highest mean score, followed by social influence and behavioral intention to use, while perceived risk was at a moderate level. In addition, the reliability test of the questionnaire showed that all factors had Cronbach's alpha values greater than 0.70, which is considered acceptable. The multiple regression analysis revealed that the independent variables could explain 39.20 percent of the variance in behavioral intention to use the service. Perceived risk had a statistically significant influence on behavioral intention at the 0.05 level and showed a negative relationship. However, perceived ease of use and social influence did have a statistically significant influence on behavioral intention in this study.*

**Keywords:** Perceived Ease of Use, Perceived Risk, Social Influence, Behavioral Intention, Self-Drive Car Rental, Social Media.

## 1. บทนำ

ในบริบทของการตลาดดิจิทัล โซเชียลมีเดีย อาทิเช่น Facebook, Instagram, TikTok, YouTube และ Twitter/X ได้กลายเป็นช่องทางสำคัญที่ผู้บริโภคใช้ในการค้นหาข้อมูล เปรียบเทียบราคาอ่านรีวิวจากผู้ใช้จริง รวมถึงติดตามโปรโมชั่นและข้อเสนอพิเศษจากผู้ให้บริการรถเช่า ช่องทางเหล่านี้มีบทบาทอย่างมากในการสร้างการรับรู้ การสร้างความน่าเชื่อถือ และการกระตุ้นการตัดสินใจของผู้บริโภค โดยเฉพาะในกลุ่มนักท่องเที่ยววัยทำงานและนักท่องเที่ยวต่างชาติ ซึ่งมีแนวโน้มพึ่งพาข้อมูลออนไลน์ในการวางแผนการเดินทาง ดังนั้น การศึกษาปัจจัยที่มีอิทธิพลต่อการเลือกใช้บริการเช่ารถแบบขับเองผ่านช่องทางโซเชียลมีเดียในประเทศไทย จึงมีความสำคัญอย่างยิ่งทั้งในเชิงวิชาการและเชิงปฏิบัติการ ผลการศึกษาจะช่วยให้เข้าใจพฤติกรรมและกระบวนการตัดสินใจของผู้บริโภคในยุคดิจิทัลได้อย่างลึกซึ้ง อีกทั้งยังเป็นประโยชน์ต่อผู้ประกอบการธุรกิจรถเช่าในการพัฒนากลยุทธ์ทางการตลาด การออกแบบบริการ และการสื่อสารผ่านแพลตฟอร์มออนไลน์ให้สอดคล้องกับความต้องการของผู้บริโภคได้อย่างมีประสิทธิภาพ อันจะนำไปสู่การสร้างความได้เปรียบทางการแข่งขันและการเติบโตอย่างยั่งยืนของธุรกิจในอุตสาหกรรมการท่องเที่ยวและการบริการที่มีการเปลี่ยนแปลงอย่างรวดเร็วในปัจจุบัน

## 2. เอกสารและงานวิจัยที่เกี่ยวข้อง

ผู้วิจัยได้ศึกษาค้นคว้าเอกสาร และงานวิจัยที่เกี่ยวข้อง โดยมีการศึกษาค้นคว้าข้อมูล แนวคิด ทฤษฎี และงานวิจัยต่างๆ ที่เกี่ยวข้องเพื่อใช้เป็นแนวทางการศึกษา ดังนี้

### 2.1 ทฤษฎีเกี่ยวกับการเลือกใช้บริการเช่ารถยนต์แบบขับเอง

การเลือกใช้บริการเช่ารถยนต์แบบขับเอง (Self-Drive Car Rental) เป็นพฤติกรรมการตัดสินใจของผู้บริโภคที่เกี่ยวข้องกับการประเมินทางเลือกของบริการที่สามารถตอบสนองความต้องการด้านการเดินทาง ความสะดวกสบาย และความคุ้มค่า โดยพฤติกรรมดังกล่าวสามารถอธิบายได้จากแนวคิดด้านพฤติกรรมผู้บริโภค (Consumer Behavior) และทฤษฎีการตัดสินใจเลือกใช้บริการ ซึ่งผู้บริโภคมักพิจารณาปัจจัยหลายด้าน เช่น ความสะดวกในการใช้งาน ความเสี่ยงในการใช้บริการ ความน่าเชื่อถือของผู้ให้บริการ และอิทธิพลจากบุคคลรอบข้างก่อนตัดสินใจใช้บริการ (Kotler & Keller, 2016) **[1]**

### 2.2 ทฤษฎีแบบจำลองการยอมรับเทคโนโลยี

ทฤษฎีแบบจำลองการยอมรับเทคโนโลยี (Technology Acceptance Model: TAM) เป็นกรอบแนวคิดที่ใช้ในการอธิบายพฤติกรรมการยอมรับและการใช้งานเทคโนโลยีของผู้ใช้ โดยพัฒนาขึ้นโดย Fred D. Davis ในปี ค.ศ. 1989 **[2]** ทฤษฎีดังกล่าวได้รับการพัฒนาเพื่อตอบคำถามสำคัญว่าปัจจัยใดที่ส่งผลให้ผู้ใช้ยอมรับหรือปฏิเสธการใช้ระบบเทคโนโลยีสารสนเทศ โดยเฉพาะการรับรู้ในด้านประโยชน์และความสะดวกในการใช้งาน ซึ่งการรับรู้ดังกล่าวจะส่งผลต่อทัศนคติของผู้ใช้ และนำไปสู่ความตั้งใจในการใช้งานเทคโนโลยีในที่สุด

### 2.3 ทฤษฎีการรับรู้ความเสี่ยงของผู้บริโภค

ทฤษฎีการรับรู้ความเสี่ยงของผู้บริโภค (Consumer Perceived Risk Theory) เป็นแนวคิดสำคัญในด้าน Consumer Behavior ที่ใช้อธิบายกระบวนการตัดสินใจของผู้บริโภคเมื่อเผชิญกับความไม่แน่นอนในการเลือกซื้อสินค้าหรือบริการ แนวคิดดังกล่าวถูกเสนอครั้งแรกโดย Raymond A. Bauer ซึ่งอธิบายว่าการตัดสินใจของผู้บริโภคมักมีความเสี่ยงแฝงอยู่ เนื่องจากผู้บริโภคไม่สามารถทราบผลลัพธ์ของการใช้สินค้าหรือบริการได้อย่างแน่นอนก่อนการใช้งานจริง **[3]**

### 2.4 ทฤษฎีการยอมรับและการใช้งานเทคโนโลยีของผู้ใช้

แนวคิดพื้นฐานของแบบจำลองการยอมรับและการใช้งานเทคโนโลยีของผู้ใช้ (Unified Theory of Acceptance and Use of Technology: UTAUT) คือการอธิบายว่าพฤติกรรมการใช้งานเทคโนโลยีของบุคคลเกิดจากกระบวนการตัดสินใจที่ได้รับอิทธิพลจากปัจจัยด้านความคาดหวัง ประสบการณ์ทางสังคม และสภาพแวดล้อมที่สนับสนุนการใช้งานเทคโนโลยี แบบจำลองดังกล่าวยังให้ความสำคัญกับตัวแปรกำกับ (Moderating Variables) เช่น

เพศ อายุ ประสบการณ์ในการใช้งาน และลักษณะความสมัครใจในการใช้เทคโนโลยี ซึ่งตัวแปรเหล่านี้สามารถส่งผลต่อระดับความสัมพันธ์ระหว่างปัจจัยหลักกับความตั้งใจในการใช้งานเทคโนโลยีของผู้ใช้ **[4]**

**2.6 การวิเคราะห์องค์ประกอบเชิงสำรวจ**

การวิเคราะห์องค์ประกอบเชิงสำรวจ (Exploratory Factor Analysis: EFA) เป็นเทคนิคทางสถิติที่ใช้ในการระบุโครงสร้างพื้นฐานของชุดตัวแปรที่สังเกตได้ โดยมีวัตถุประสงค์หลักคือการลดจำนวนตัวแปรให้เหลือน้อยลงโดยการจัดกลุ่มตัวแปรที่มีความสัมพันธ์กันเข้าด้วยกันเป็นองค์ประกอบ การวิเคราะห์องค์ประกอบเชิงสำรวจ หรือ Exploratory Factor Analysis นั้นมีไว้เพื่อการสกัดหรือจัดกลุ่มตัวแปรที่ยังไม่มีกลุ่มหรือองค์ประกอบมาก่อน **[5]** สถิติช่วยในการจัดกลุ่มให้ โดยพิจารณาจากต่างๆ ได้แก่ KMO, Bartlett's test, Communality, Eigen Value, Factor Loading, Rotation.

**2.7 การวิเคราะห์สมการถดถอยพหุคูณ**

การวิเคราะห์การถดถอยพหุคูณเป็นเทคนิคทางสถิติที่ใช้ในการศึกษาความสัมพันธ์ระหว่างตัวแปรตาม หนึ่งตัวกับตัวแปรอิสระ สองตัวขึ้นไป กำหนดสมการถดถอยพหุคูณในรูปแบบ

$$Y=\beta 0+\beta 1X+\beta 1X2+...+\beta nXn+e$$

โดยที่

$Y$ คือ ตัวแปรตาม (Dependent Variable)

$X$ คือ ตัวแปรอิสระ (Independent Variable)

$\beta 0$ คือ เป็นระยะตัดแกน Y หรือค่าเริ่มต้นของเส้นสมการถดถอย

$\beta nXn$ คือ สัมประสิทธิ์การถดถอย (Regression Coefficient) ตัวที่ 1 ถึงตัวที่ n

$e$ คือ ค่าความคลาดเคลื่อน (error term) **[6]**

**2.7 งานวิจัยที่เกี่ยวข้อง**

อรณัชชา จินา (2567). **[7]** ทาการวิจัยอิทธิพลของสื่อสังคมออนไลน์ที่มีผลต่อการเลือกร้านอาหารในจังหวัดเชียงใหม่ ใช้แบบสอบถามรวบรวมข้อมูล วิเคราะห์สถิติเชิงพรรณนา เชิงอนุมาน และ EFA พบว่า Content Quality, Engagement Quality และ Contextual Quality ส่งผลต่อการเลือกรับชมสื่อออนไลน์

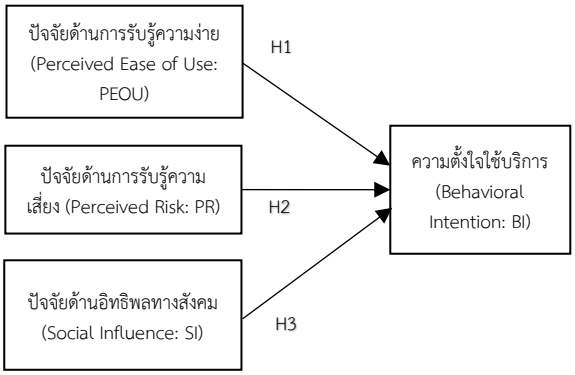
ธมลวรรณ สมพงศ์ (2562). **[8]** การศึกษาปัจจัยที่มีผลต่อการตัดสินใจซื้อสินค้าทางออนไลน์ของผู้บริโภคในจังหวัดนครปฐม ใช้แบบสอบถามออนไลน์รวบรวมข้อมูล วิเคราะห์สถิติเชิงพรรณนา และเชิงอนุมาน พบว่าปัจจัยการยอมรับเทคโนโลยี โดยเฉพาะ Perceived Usefulness มีผลต่อการตัดสินใจซื้อสินค้าออนไลน์

นันทิตา สังสะโอภาศ (2564). **[9]** ปัจจัยทางโซเชียลมีเดียที่ส่งผลต่อความตั้งใจซื้อสินค้าและบริการของคลินิกความงาม ใช้แบบสอบถามรวบรวมข้อมูล วิเคราะห์ด้วย SPSS พบว่า Content และ Influencer Review ส่งผลต่อ Trust และ Trust ส่งผลต่อความตั้งใจซื้อ

## 3. วิธีดำเนินการวิจัย

**3.1 กรอบแนวคิดงานวิจัย**

การศึกษาครั้งนี้ได้นำแนวคิดและทฤษฎีที่เกี่ยวโดยพิจารณาด้าน การรับรู้ถึงความง่ายในการใช้งาน การรับรู้ความเสี่ยง อิทธิพลทางสังคม ส่งผลต่อทัศนคติต่อการใช้งานซึ่งจะนำไปสู่ความตั้งใจในการใช้งาน ซึ่งกรอบแนวคิดที่นำเสนอจะช่วยให้เห็นภาพรวมของปัจจัยที่มีผลต่อการยอมรับเทคโนโลยีนี้




**ภาพที่ 3-1** กรอบแนวคิดของการวิจัย

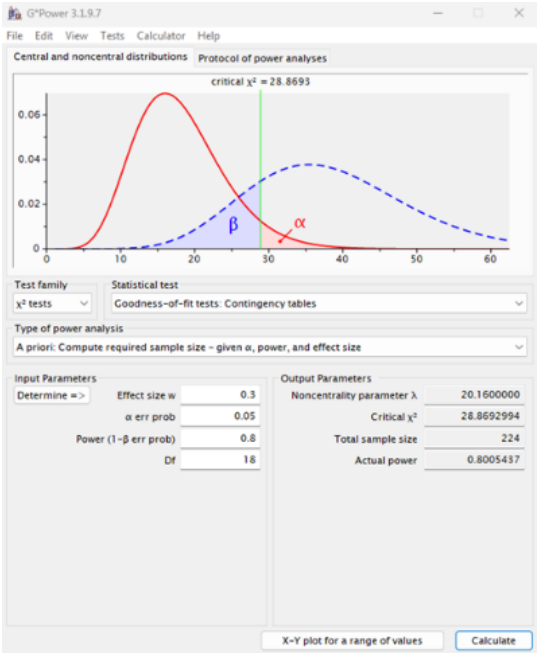
**3.2 ประชากรและกลุ่มตัวอย่าง**

3.2.1 ประชากรที่ใช้ในการวิจัยครั้งนี้ คือ นักท่องเที่ยวชาวไทยที่มีประสบการณ์หรือเคยใช้บริการเช่ารถแบบขับ

เองผ่านช่องทางโซเชียลมีเดียในประเทศไทย โดยประชากรดังกล่าวมีช่วงอายุประมาณ 18 – 60 ปีขึ้นไป

3.2.2 กลุ่มตัวอย่าง (Sample)

กลุ่มตัวอย่างแบบสุ่มตัวอย่าง เนื่องจากประชากรที่เลือกใช้บริการเช่ารถแบบขับเองผ่านช่องทางโซเชียลมีเดียในประเทศไทย ทำให้ไม่ทราบจำนวนประชากรที่แน่นอน ดังนั้นทางผู้วิจัยได้กำหนดกลุ่มตัวอย่างด้วยโปรแกรม G*Power ได้ขนาดกลุ่มตัวอย่างเท่ากับ 224 คน ซึ่งผู้วิจัยปัดเป็นตัวเลขเป็น 250 คน ดังนั้นขนาดของกลุ่มตัวอย่าง ที่ใช้การวิจัย คือ 250 คน




**ภาพที่ 3-2** กำหนดขนาดกลุ่มตัวอย่างด้วยโปรแกรม G*Power

### 3.3 เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการศึกษาการวิจัยคือแบบสอบถาม Google form ในการสร้างแบบสอบถามและ เก็บรวบรวมข้อมูลใช้คำถามแบบปลายปิดแบบสอบถามแบ่งออกเป็น 2 ส่วน ดังนี้

ส่วนที่ 1 ข้อมูลส่วนบุคคลของผู้ตอบแบบสอบถาม

ส่วนที่ 2 ปัจจัยที่มีอิทธิพลต่อการเลือกใช้บริการเช่ารถแบบขับเองผ่านช่องทางโซเชียลมีเดียในประเทศไทย

## 4. ผลการดำเนินงาน

**ตารางที่ 4-1** ผลการวิเคราะห์ข้อมูลทั่วไป

| ข้อมูลทั่วไป | จำนวน (คน) | ร้อยละ (%) |
|---|---|---|
| **1.เพศ (Gender)** | | |
| ชาย | 86 | 34.40 |
| หญิง | 164 | 65.60 |
| **รวม** | 250 | 100.00 |
| **2.อายุ** | | |
| ต่ำกว่าหรือเท่ากับ 20 | 15 | 6.00 |
| 20 – 40 ปี | 161 | 64.40 |
| 41 – 60 ปี | 66 | 26.40 |
| มากกว่า 60 ปี | 8 | 3.20 |
| **รวม** | 250 | 100.00 |
| **3.ระดับการศึกษา** | | |
| ต่ำกว่าปริญญาตรี | 63 | 25.20 |
| ปริญญาตรี | 154 | 61.60 |
| ปริญญาโท | 28 | 11.20 |
| ปริญญาเอก | 5 | 2.00 |
| **รวม** | 250 | 100.00 |
| **4.อาชีพ** | | |
| นักเรียน/นักศึกษา | 25 | 10.00 |
| ข้าราชการ | 39 | 15.60 |
| พนักงาน รัฐวิสาหกิจ | 36 | 14.40 |
| พนักงานบริษัทเอกชน | 94 | 37.60 |
| เจ้าของธุรกิจ/ธุรกิจส่วนตัว | 38 | 15.20 |
| อื่นๆ | 18 | 7.20 |
| **รวม** | 250 | 100.00 |
| **5.โซเชียลมีเดียที่ใช้งานมากที่สุด** | | |
| FACEBOOK | 103 | 41.20 |
| TIKTOK | 59 | 23.60 |
| LINE | 44 | 17.60 |
| INSTRGRAM | 36 | 14.40 |
| X (Twitter) | 8 | 3.20 |
| **รวม** | 250 | 100.00 |
| **6.เคยใช้บริการเช่ารถแบบขับเองผ่านช่องทางโซเชียล ผ่านช่องทางไหน** | | |
| FACEBOOK | 170 | 68.00 |
| TIKTOK | 17 | 6.80 |
| LINE | 59 | 23.60 |

| ข้อมูลทั่วไป | จำนวน (คน) | ร้อยละ (%) |
|---|---|---|
| INSTRGRAM | 2 | 0.80 |
| X (Twitter) | 2 | 0.80 |
| **รวม** | 250 | 100.00 |

จากการศึกษากลุ่มตัวอย่างทั้งหมดจำนวน 250 คน พบว่า ผู้ตอบแบบสอบถามส่วนใหญ่เพศหญิง จำนวน 164 คน คิดเป็นร้อยละ 65.60 ด้านอายุส่วนใหญ่อยู่ในช่วง 20–40 ปี จำนวน 161 คน คิดเป็นร้อยละ 64.40 ด้านระดับการศึกษาส่วนใหญ่มีการศึกษาระดับ ปริญญาตรี จำนวน 154 คน คิดเป็นร้อยละ 61.60 ด้านอาชีพส่วนใหญ่ประกอบอาชีพ พนักงานบริษัทเอกชน จำนวน 94 คน คิดเป็นร้อยละ 37.60 ในส่วนของพฤติกรรมการใช้งานสื่อสังคมออนไลน์ พบว่าแพลตฟอร์มที่กลุ่มตัวอย่างใช้งานมากที่สุดคือ Facebook จำนวน 103 คน คิดเป็นร้อยละ 41.20 และกลุ่มตัวอย่างส่วนใหญ่ เคยใช้บริการเช่ารถแบบขับเองผ่าน Facebook มากที่สุด จำนวน 170 คน คิดเป็นร้อยละ 68.00

**4.2 ผลการวิเคราะห์องค์ประกอบเชิงสำรวจ**

**ตารางที่ 4-2** ค่า KMO และ Bartlett's Test of Sphericity

| KMO and Bartlett's Test | |
|---|---|
| Kaiser-Meyer-Olkin Measure of Sampling Adequacy. | 0.838 |
| Bartlett's Test of Sphericity Approx. Chi-Square | 1721 |
| df | 105 |
| Sig. | 0.00 |

การวิเคราะห์ข้อมูลด้วยข้อคำถามทั้งหมด 18 ตัวแปร พบว่ามีค่า KMO เท่ากับ 0.838 ซึ่งมีค่ามากกว่า 0.5 และเข้าใกล้ 1 แสดงถึงความสัมพันธ์กันของตัวแปรในขนาดที่เหมาะสม สามารถนำไปวิเคราะห์เชิงสำรวจได้ และค่าการทดสอบนัยสำคัญทางสถิติด้วย Barlett's Test of Sphericity พบว่ามีค่านัยสำคัญทางสถิติที่ระดับ 0.05 สามารถสรุปได้ว่าข้อมูลมีความเหมาะสมในการวิเคราะห์องค์ประกอบได้

**ตารางที่ 4-3** ผลการวิเคราะห์องค์ประกอบเชิงสำรวจ Factor Loading

| Factor Loadings | | | | | |
|---|---|---|---|---|---|
| | Factor | | | | Uniqueness |
| | 1 | 2 | 3 | 4 | |
| PEOU1 | | | 0.662 | | 0.579 |
| PEOU2 | | | 0.654 | | 0.500 |
| PEOU3 | | | 0.779 | | 0.404 |
| PEOU5 | | | 0.693 | | 0.519 |
| PR1 | 0.783 | | | | 0.331 |
| PR2 | 0.816 | | | | 0.302 |
| PR3 | 0.769 | | | | 0.422 |
| PR4 | 0.693 | | | | 0.488 |
| PR5 | 0.568 | | | | 0.661 |
| SI1 | | | | 0.871 | 0.276 |
| SI2 | | | | 0.612 | 0.443 |
| BI1 | | 0.902 | | | 0.236 |
| BI2 | | 0.875 | | | 0.249 |
| BI3 | | 0.633 | | | 0.398 |
| BI4 | | 0.611 | | | 0.446 |

พบว่าค่าน้ำหนักองค์ประกอบ (Factor loading)ที่มีค่ามากกว่า 0.5 มีทั้งหมด 15 ตัวแปร 4 กลุ่ม มีกลุ่มจำนวนตัวแปรมากที่สุด 5 ตัวแปร 1 กลุ่ม และมีจำนวนตัวแปร 4 ตัว 2 กลุ่ม ตัวแปรน้อยที่สุด 2 ตัวแปร 1 กลุ่ม

**ตารางที่ 4-4 ผลการทดสอบความเชื่อมั่นของแบบสอบถาม**

| ลำดับ | ปัจจัย | Mean | S.D. | Cronbach's Alpha |
|---|---|---|---|---|
| 1 | การรับรู้ความง่ายในการใช้งาน | 3.93 | 0.671 | 0.807 |
| 2 | การรับรู้ความเสี่ยง | 3.40 | 0.803 | 0.845 |
| 3 | อิทธิพลทางสังคม | 3.69 | 0.718 | 0.707 |
| 4 | ความตั้งใจในการใช้งาน | 3.65 | 0.821 | 0.882 |

ผลการวิเคราะห์ความเชื่อมั่น ของแบบสอบถามพบว่ามีค่าสัมประสิทธิ์ความเชื่อมั่นอยู่ระหว่าง 3.40 – 3.93 ซึ่งอยู่ในระดับที่ยอมรับได้ แสดงว่าเครื่องมือแบบสอบถามที่ใช้ในการสำรวจฉบับนี้มีความน่าเชื่อถือและผลการวิเคราะห์ค่าส่วนเบี่ยงเบนมาตรฐานทั้ง 4 ด้าน พบว่า ปัจจัยการรับรู้ความง่ายในการใช้งาน มีค่าเฉลี่ยมากที่สุด (Mean = 4.93, SD = 0.671)

**4.5 การวิเคราะห์การถดถอยเชิงเส้นแบบพหุคูณ**

4.3.1 วิเคราะห์จาก 5 ตัวแปร ต่อทัศนคติการใช้งาน

**ตารางที่ 4-5** Model Coefficients – BI

| Model Coefficients – BI | | | | | | Collinearity Statistics | |
|---|---|---|---|---|---|---|---|
| Predictor | Estimate | SE | t | p | Stand. Estimate | VIF | Tolerance |
| Intercept | 0.692 | 0.289 | 0.123 | 1.261 | 2.39 | | |
| PEOU | 0.516 | 0.063 | 0.393 | 0.639 | 8.23 | 1.07 | 0.936 |
| PR | -0.063 | 0.052 | -0.165 | 0.039 | -1.21 | 1.05 | 0.957 |
| SI | 0.333 | 0.045 | 0.245 | 0.420 | 7.47 | 1.08 | 0.922 |

ผลการวิเคราะห์ถดถอยพหุคูณของปัจจัยแสดงผลการวิเคราะห์ถดถอยพหุคูณของปัจจัยที่ที่มีอิทธิผลส่งผลต่อความตั้งใจในการใช้งาน (BI) พบว่าตัวแปรอิสระทุกตัวมีค่า VIF น้อยกว่า 10 โดยการรับรู้ถึงความง่ายในการใช้งาน (PEOU) มี ค่า Tolerance เท่ากับ 0.936 การรับรู้ความเสี่ยง (SI) มีค่า Tolerance เท่ากับ 0.957 และอิทธิพลทางสังคม (SI) มีค่า Tolerance เท่ากับ 0.922

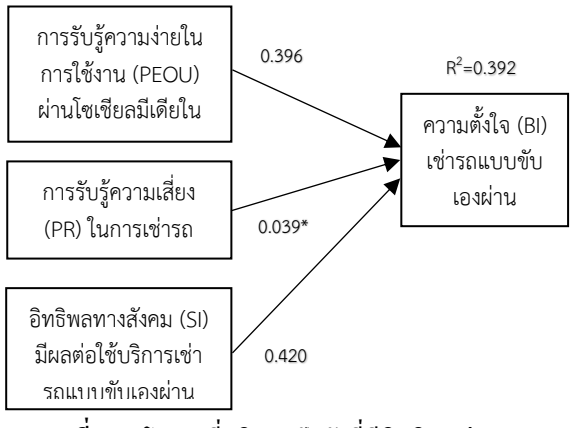
4.3.2 วิเคราะห์ตัวแปรต่อความตั้งใจการใช้งาน

**ตารางที่ 4-6** Model Fit Measures – BI

| Model Fit Measures | | | |
|---|---|---|---|
| Model | R | $R^2$ | Adjusted $R^2$ |
| 1 | 0.626 | 0.392 | 0.385 |
| Note. Models estimated using sample size of N=250 | | | |

ผลการวิเคราะห์ถดถอยพหุคูณของปัจจัยที่ส่งผลต่อความตั้งใจในการใช้งาน (BI) พบว่าตัวแปรอิสระทั้งหมดในโมเดลสามารถร่วมกันอธิบายความแปรปรวนของความตั้งใจในการใช้บริการเช่ารถแบบขับเองผ่านช่องทางโซเชียลมีเดีย ได้ร้อยละ 39.2

แสดงความสัมพันธ์ของตัวแปรต่างๆ ในรูปแบบของโมเดล ดังนี้




**ภาพที่ 4-1** โมเดลที่อธิบายปัจจัยที่มีอิทธิพลต่อการเลือกใช้บริการเช่ารถแบบขับเองผ่านช่องทางโซเชียลมีเดียในประเทศไทย

โดยที่ * คือ ค่านัยสำคัญทางสถิติที่ระดับ 0.05

** คือ ค่านัยสำคัญทางสถิติที่ระดับ 0.01

*** คือ ค่านัยสำคัญทางสถิติที่ระดับ 0.00

## 5. สรุปผล อภิปรายผล และข้อเสนอแนะ

ผู้วิจัยได้กำหนดสมมติฐานการวิจัยจำนวนสมมติฐาน และสามารถอภิปรายผลได้ดังต่อไปนี้

สมมติฐานที่ 1 (H1) การรับรู้ความง่ายในการใช้งานมีอิทธิพลต่อความตั้งใจใช้บริการเช่ารถแบบขับเองผ่านช่องทางโซเชียลมีเดียในทิศทางเชิงบวก

สมมติฐานที่ 2 (H2) การรับรู้ความเสี่ยงไม่มีอิทธิพลต่อความตั้งใจใช้บริการเช่ารถแบบขับเองผ่านช่องทางโซเชียลมีเดีย

สมมติฐานที่ 3 (H3) ปัจจัยทางสังคมมีอิทธิพลต่อความตั้งใจใช้บริการเช่ารถแบบขับเองผ่านช่องทางโซเชียลมีเดียในทิศทางเชิงบวก

ผลการวิจัยครั้งนี้สะท้อนให้เห็นถึงการปรับตัวของผู้บริโภคในยุคดิจิทัลที่มีความคุ้นเคยกับการใช้งานแพลตฟอร์มออนไลน์มากขึ้น

## เอกสารอ้างอิง

[1] Kotler, Philip, & Keller, Kevin Lane. (2016). Marketing management (15th ed.). Pearson Education.

[2] Davis, Fred D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly, 13(3), 319–340.

[3] Bauer, R. A. (1960). Consumer behavior as risk taking. In R. S. Hancock (Ed.), Dynamic marketing for a changing world (pp. 389–398). Chicago: American Marketing Association.

[4] Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. MIS Quarterly, 27(3), 425–478.

[5] Hair, Joseph F., Black, William C., Babin, Barry J., & Anderson, Rolph E. (2019). Multivariate data analysis (8th ed.). Cengage Learning.

[6] Sekaran, Uma, & Bougie, Roger. (2016). Research methods for business: A skill-building approach (7th ed.). Wiley.

[7] อรณัชชา จินา (2567). อิทธิพลของสื่อสังคมออนไลน์ที่มีผลต่อการเลือกร้านอาหารในจังหวัดเชียงใหม่. บริหารธุรกิจมหาบัณฑิต มหาวิทยาลัยเชียงใหม่

[8] ธมลวรรณ สมพงศ์ (2562). การศึกษาปัจจัยที่มีผลต่อการตัดสินใจซื้อสินค้าทางออนไลน์ของผู้บริโภคในจังหวัดนครปฐม. วิศวกรรมศาสตรมหาบัณฑิต มหาวิทยาลัยศิลปากร.

[9] นันทิตา สังสะโอภาศ (2564). ปัจจัยทางโซเชียลมีเดียที่ส่งผลต่อความตั้งใจซื้อสินค้าและบริการของคลินิกความงาม. การจัดการมหาบัณฑิต มหาวิทยาลัยมหิดล.

# ปัจจัยที่มีอิทธิพลต่อการยอมรับการใช้เทคโนโลยีเสมือนจริงในการออกแบบทางสถาปัตยกรรม

# Factors Influencing the Acceptance of Virtual Reality Technology in Architectural

*ณัฐกานต์ แต้มเจริญ (Nattakan Tamcharoen)[1] และมณเฑียร รัตนศิริวงศ์วุฒิ (Montean Rattanasiriwongwut)[2]*

*ภาควิชาการจัดการเทคโนโลยีสารสนเทศ*

*คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ*

*[1]s6707021857074@email.kmutnb.ac.th, [2] montean.r@itd.kmutnb.ac.th*

## บทคัดย่อ

*การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อศึกษาปัจจัยที่มีอิทธิพลต่อการใช้เทคโนโลยีเสมือนจริงในการออกแบบทางสถาปัตยกรรม โดยอาศัยกรอบแนวคิดการยอมรับเทคโนโลยี ได้แก่ การรับรู้ความง่ายในการใช้งาน การรับรู้ถึงประโยชน์ อิทธิพลของสังคม ความพร้อมด้านโครงสร้างพื้นฐาน และทัศนคติที่มีต่อการใช้งานกลุ่มตัวอย่างจำนวน 416 คน เก็บข้อมูลด้วยแบบสอบถาม และวิเคราะห์ข้อมูลด้วยสถิติเชิงพรรณนาและการวิเคราะห์ถดถอยเชิงเส้นพหุคูณด้วยโปรแกรม Jamovi ผลการวิจัยพบว่าโมเดลมีความเหมาะสม ($R = 0.821$, $R^2 = 0.674$) โดยปัจจัยทั้งห้าสามารถอธิบายความแปรปรวนของการใช้เทคโนโลยีเสมือนจริงได้ร้อยละ 67.40 และทุกปัจจัยมีอิทธิพลเชิงบวกอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.001*

**คำสำคัญ:** การออกแบบสถาปัตยกรรม เทคโนโลยีเสมือนจริง การยอมรับเทคโนโลยี

## Abstract

*This research aims to examine the factors influencing the use of virtual reality technology in architectural design, based on the Technology Acceptance Model. The factors include perceived ease of use, perceived usefulness, social influence, facilitating conditions, and attitude toward use.*

*The sample consisted of 416 participants who completed a questionnaire. Data were analyzed using descriptive statistics and multiple linear regression with Jamovi software.*

*The results indicated that the research model was appropriate ($R = 0.821$, $R^2 = 0.674$). The five independent factors explained 67.40% of the variance in the use of virtual reality technology, and all factors had a positive influence with statistical significance at the 0.001 level.*

**Keywords**: Architectural Design, Virtual Reality, Technology Acceptance.

## 1. บทนำ

การออกแบบทางสถาปัตยกรรมเป็นกระบวนการที่ผสมผสานความคิดสร้างสรรค์ การวางแผนเชิงเทคนิค และการตอบสนองต่อความต้องการของผู้ใช้งาน การปรับปรุงแบบจึงมีความสำคัญต่อการพัฒนาคุณภาพของงานออกแบบทั้งด้านความสวยงาม การใช้สอย ความปลอดภัย และความยั่งยืน ปัจจุบันเทคโนโลยี Augmented Reality (AR) และ Virtual Reality (VR) ได้เข้ามามีบทบาทสำคัญในการสนับสนุนกระบวนการออกแบบทางสถาปัตยกรรม ช่วยให้สามารถมองเห็นและประเมินรูปแบบหรือพื้นที่เสมือนจริงก่อนการก่อสร้าง เทคโนโลยีดังกล่าวช่วยตรวจสอบข้อผิดพลาด วิเคราะห์ผลกระทบ และปรับปรุงแนวคิดการออกแบบได้อย่างมีประสิทธิภาพ ลดความเสี่ยงและประหยัดเวลาในการดำเนินโครงการ

ดังนั้น การศึกษาปัจจัยที่มีอิทธิพลต่อการใช้เทคโนโลยีเสมือนจริงจึงมีความสำคัญ เพื่อสนับสนุนการประยุกต์ใช้

เทคโนโลยีอย่างเหมาะสมและพัฒนาวงการสถาปัตยกรรมสู่ยุคดิจิทัลอย่างยั่งยืน

## 2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

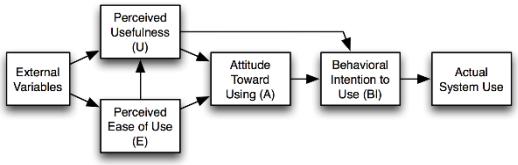
ผู้วิจัยได้ศึกษาค้นคว้าเอกสาร และงานวิจัยที่เกี่ยวข้องต่าง ๆ เพื่อนำไปพัฒนาแพลตฟอร์มสำหรับการให้ความรู้ที่เกี่ยวข้องกับความมั่นคงปลอดภัยทางไซเบอร์ โดยศึกษาข้อมูลต่าง ๆ ดังต่อไปนี้

**2.1 ทฤษฎีรวมการยอมรับและการใช้เทคโนโลยี (UTAUT)**

เป็นทฤษฎีที่พัฒนาโดย Venkatesh, Morris, Davis และ Davis ในปี ค.ศ. 2003 ซึ่งมีวัตถุประสงค์เพื่อสร้างแบบจำลองที่สามารถอธิบายพฤติกรรมการยอมรับและการใช้เทคโนโลยีได้อย่างครอบคลุมและแม่นยำ และสามารถอธิบายและทำนายพฤติกรรมของผู้ใช้งานเทคโนโลยีได้อย่างมีประสิทธิภาพ โดยอธิบายถึงปัจจัยที่มีอิทธิพลต่อการยอมรับและใช้งาน มีองค์ประกอบหลัก 4 ปัจจัย คือ ความคาดหวังในประสิทธิภาพ, ความคาดหวังด้านความง่ายในการใช้งาน, อิทธิพลทางสังคม, ระบบสนับสนุน

**2.2 แบบจำลองการยอมรับเทคโนโลยี (TAM)**

เป็นแนวคิดที่พัฒนาโดย Fred Davis (1989) ซึ่งมีรากฐานมาจากทฤษฎีการกระทำด้วยเหตุผล (Theory of Reasoned Action: TRA) ที่เสนอโดย Icek Ajzen และ Morris Fishbein (1975) TAM ถูกสร้างขึ้นเพื่อใช้ในการศึกษา ปัจจัยที่ส่งผลต่อแรงจูงใจและความตั้งใจของบุคคลในการยอมรับและใช้งานเทคโนโลยีสารสนเทศ โดยเฉพาะในบริบทของระบบสารสนเทศภายในองค์กร หรือเทคโนโลยีใหม่ ๆ ที่นำมาใช้ในกระบวนการทำงาน แสดงดังนี้




**ภาพที่ 1** : โมเดลแบบจำลองการยอมรับเทคโนโลยี

**2.3 การวิเคราะห์องค์ประกอบเชิงสำรวจ (EFA)**

เป็นเทคนิคทางสถิติที่มีบทบาทสำคัญในงานวิจัย ความสัมพันธ์ระหว่างตัวแปรหลายตัวที่มีความซับซ้อน และเพื่อลดจำนวนตัวแปรให้เหลือด้านสังคมศาสตร์ พฤติกรรมศาสตร์ โดยมีวัตถุประสงค์หลักเพื่อศึกษาเฉพาะตัวแทนที่สำคัญซึ่งสามารถอธิบายข้อมูลได้อย่างมีประสิทธิภาพ

**2.4 การวิเคราะห์การถดถอยพหุคูณ**

เป็นเทคนิคทางสถิติที่ใช้ในการวิเคราะห์ข้อมูลเพื่อทำนายค่าของตัวแปรตาม โดยใช้ชุดตัวแปรอิสระ โดยใช้สมการเชิงเส้น ซึ่งจะช่วยในการคำนวณและทำนายค่าของตัวแปรตามที่ต้องการในกรณีที่มีผลกระทบจากตัวแปรอิสระหลายตัวพร้อมกัน

**2.5 งานวิจัยที่เกี่ยวข้อง**

สิริรักษ์ และคณะ [1] ได้ทำการศึกษาวิจัย การวิเคราะห์ปัจจัยด้านการรับรู้ที่มีผลต่อการยอมรับเทคโนโลยีเสมือนจริงของผู้ใช้งานกลุ่มเจเนอเรชั่นวาย พบว่า การรับรู้ประโยชน์และความง่ายในการใช้งานมีอิทธิพลโดยตรงต่อการยอมรับเทคโนโลยี VR ของกลุ่มเจเนอเรชั่นวาย โดยผู้ใช้จะมีทัศนคติที่ดีหากมองว่าเทคโนโลยีใช้งานง่ายและมีประโยชน์ชัดเจน และแบบจำลอง SEM สามารถอธิบายพฤติกรรมการยอมรับได้อย่างมีนัยสำคัญ

ชีวรัตน์ [2] ได้ทำการศึกษาวิจัย การยอมรับเทคโนโลยีการเรียนภาษาผ่านแอปพลิเคชันออนไลน์ในเขตกรุงเทพมหานครและปริมณฑล โดยใช้แบบจำลอง TAM และ UTAUT จากกลุ่มตัวอย่าง 169 คน พบว่าการรับรู้ความง่ายในการใช้งานไม่มีอิทธิพลโดยตรงต่อความตั้งใจในการใช้ แต่มีอิทธิพลทางอ้อมผ่านการรับรู้ประโยชน์ และผู้ใช้มีแนวโน้มใช้งานเนื่องจากเห็นประโยชน์และความเพลิดเพลินจากแอปพลิเคชัน

อัจฉรา [3] ได้ทำการศึกษาวิจัย ปัจจัยที่มีอิทธิพลต่อการยอมรับเทคโนโลยีสารสนเทศในการปฏิบัติงานของสำนักงานปลัดกระทรวงวิทยาศาสตร์และเทคโนโลยี โดยใช้แบบสอบถามจากกลุ่มตัวอย่าง 280 คน และวิเคราะห์ด้วยแบบจำลอง TAM และ UTAUT พบว่า ความง่ายในการใช้งานมีอิทธิพลเชิงบวกต่อการรับรู้ประโยชน์ และ

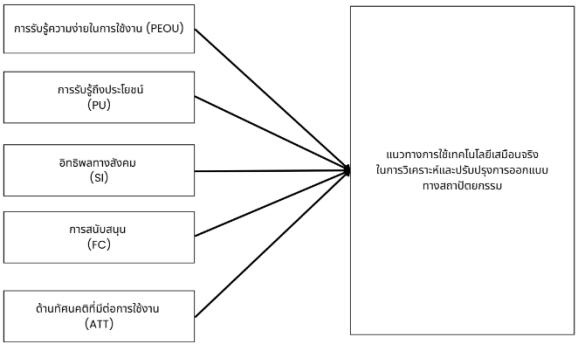
สามารถทำนายได้ร้อยละ 52 อีกทั้งความง่ายในการใช้งานและการรับรู้ประโยชน์ร่วมกันส่งผลต่อทัศนคติการใช้งาน (ร้อยละ 46) และทัศนคติดังกล่าวยังมีอิทธิพลต่อความตั้งใจใช้เทคโนโลยี (ร้อยละ 33)

ภาณุพันธ์ [4] ได้ทำการศึกษาวิจัย พฤติกรรมความตั้งใจซื้อรถยนต์ส่วนบุคคลผ่านช่องทางออนไลน์ ของคนในเขตพื้นที่กรุงเทพมหานครและปริมณฑล โดยใช้กรอบแนวคิดทฤษฎีรวมของการยอมรับและการใช้เทคโนโลยี (UTAUT) เก็บข้อมูลจากกลุ่มตัวอย่าง 387 คน พบว่า ความคาดหวังในประสิทธิภาพ ความง่ายในการใช้งาน และอิทธิพลทางสังคม มีอิทธิพลต่อความตั้งใจซื้อรถยนต์ส่วนบุคคลผ่านช่องทางออนไลน์

สุพรรณี [5] ได้ทำการศึกษาวิจัย เรื่อง ปัจจัยที่ส่งผลต่อการใช้งานเทคโนโลยีดิจิทัลในการบริหารทรัพยากรมนุษย์ของพนักงานธนาคารของรัฐขนาดใหญ่ ABC โดยใช้แบบจำลอง TAM โดยพิจารณาปัจจัยด้านการรับรู้ประโยชน์ ความง่ายในการใช้งาน และทัศนคติการใช้งาน พร้อมเพิ่มปัจจัยสภาพสิ่งอำนวยความสะดวกเพื่อให้เหมาะสมกับบริบทองค์กร

## 3. วิธีการดำเนินการวิจัย

จากการศึกษาแนวคิด ทฤษฎี รวมทั้งงานวิจัยที่เกี่ยวข้องกับหัวข้อดังกล่าว พบว่าแบบจำลอง TAM และ UTAUT สามารถอธิบายพฤติกรรมการใช้เทคโนโลยีได้อย่างมีประสิทธิภาพ โดยงานวิจัยนี้เลือกตัวแปรสำคัญ 5 ด้าน ได้แก่ ความง่ายในการใช้งาน การรับรู้ประโยชน์ อิทธิพลทางสังคม การสนับสนุน และทัศนคติ ซึ่งสอดคล้องกับบริบทงานสถาปัตยกรรมในปัจจุบัน โดยมีกรอบแนวทางการดำเนินงานวิจัยและกรอบแนวคิดระบบแพลตฟอร์มดังนี้




**ภาพที่ 2** : กรอบแนวทางการดำเนินงานวิจัย

### 3.1 ประชากรและกลุ่มตัวอย่าง

กลุ่มตัวอย่างในการวิจัยครั้งนี้ประกอบด้วยผู้ที่เกี่ยวข้องกับงานออกแบบทางสถาปัตยกรรม จำนวน 416 คน ได้แก่ สถาปนิก นักออกแบบ วิศวกร และบุคลากรที่เกี่ยวข้อง โดยกำหนดขนาดตัวอย่างตามเกณฑ์ของ Hair et al. (2010) แนะนำให้ขนาดของกลุ่มตัวอย่างควรอยู่ระหว่าง 10-20 เท่าของจำนวนตัวแปรสังเกต ผู้วิจัยทำการคัดเลือกกลุ่มตัวอย่างโดยการใช้วิธีการสุ่มตัวอย่างใช้ระยะเวลาเก็บข้อมูลตั้งแต่สิงหาคม 2568 ถึงพฤศจิกายน 2568 โดยเครื่องมือที่ใช้ ได้แก่ Google Form

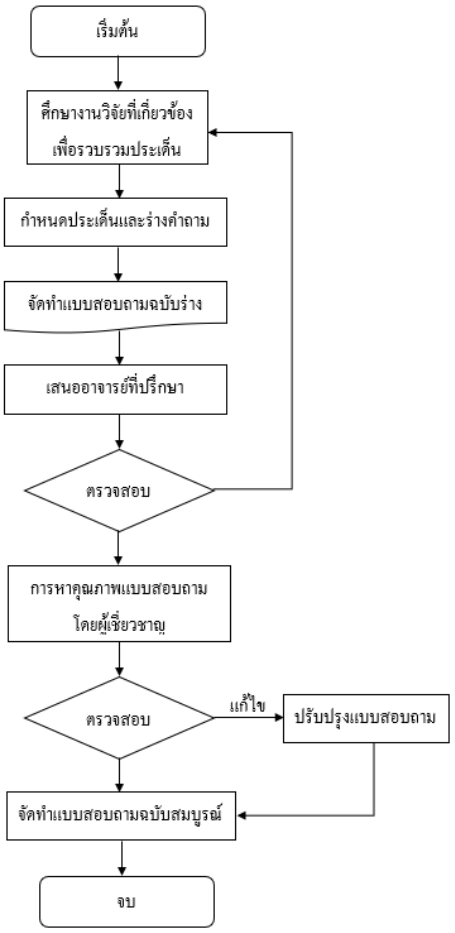
### 3.2 เครื่องมือที่ใช้ในการวิจัย

งานวิจัยนี้ได้กำหนดเครื่องมือในการศึกษาการวิจัยด้วย Google Chrome และนำมาซึ่งข้อมูลไปวิจัยด้วยแบบสอบถาม Google form ในการสร้างแบบสอบถามและเก็บรวบรวมข้อมูล โดยออกแบบคำถามทั้งแบบปลายปิดและปลายเปิด และจัดทำผ่าน Google Sheets แบบสอบถามแบ่งเป็น 2 ส่วน ได้แก่

**ส่วนที่ 1** ข้อมูลทั่วไปของผู้ตอบ (5 ข้อ) เช่น เพศ อายุ การศึกษา อาชีพ และการใช้งานเทคโนโลยีเสมือนจริง

**ส่วนที่ 2** คำถามเกี่ยวกับการใช้เทคโนโลยีเสมือนจริงในการออกแบบสถาปัตยกรรม (23 ข้อ) ใช้มาตรวัด 5 ระดับ ครอบคลุม 5 ด้าน ได้แก่ ความง่ายในการใช้งาน การรับรู้ถึงประโยชน์ อิทธิพลทางสังคม ความพร้อมด้านโครงสร้างพื้นฐานหรือการสนับสนุน และทัศนคติที่มีต่อการใช้งาน

ผู้วิจัยใช้แบบสอบถามเป็นเครื่องมือในการเก็บรวบรวมข้อมูล โดยมีขั้นตอนในการสร้างเครื่องมือแสดงดังนี้




**ภาพที่ 3** : แสดงขั้นตอนการสร้างเครื่องมือในการเก็บข้อมูล

**3.3 การหาค่าความเที่ยงตรงของแบบสอบถา**

เป็นวิธีการพิจารณาคุณภาพข้อคำถาม ด้านความเหมาะสม ความถูกต้อง ครบถ้วน มีเนื้อหาสอดคล้อง ตรงกับวัตถุประสงค์ และสมมติฐานของงานวิจัย (Content Validity) จึงจำเป็นต้องอาศัยการประเมินความคิดเห็นของผู้เชี่ยวชาญเกี่ยวกับเนื้อหาในแบบสอบถาม โดยกำหนดผู้เชี่ยวชาญ **5** ท่าน ทำการประเมินค่าความสอดคล้องของแบบสอบถาม และนำผลคะแนนของผู้เชี่ยวชาญมาทำการวิเคราะห์หาค่าความสอดคล้อง

**3.4 การวิเคราะห์ข้อมูล**

หลังจากที่ได้รวบรวมข้อมูลครบทั้งหมด 416 ชุด ผู้วิจัยได้นำมาประมวลผลด้วยโปรแกรมคำนวณทางสถิติ ประกอบด้วยการวิเคราะห์ข้อมูล เพื่อหาคำตอบของวัตถุประสงค์ และสมมติฐาน โดยสถิติที่ใช้ในการวิเคราะห์ข้อมูล ดังนี้

3.4.1 นำข้อมูลทั่วไปของประชากรตัวอย่างมาวิเคราะห์โดยใช้สถิติเชิงพรรณนา ได้แก่ วิเคราะห์ข้อมูลโดยการใช้การแจกแจง ความถี่

3.4.2 นำข้อมูลตัวแปรมาวิเคราะห์องค์ประกอบเชิงสำรวจ (EFA) เพื่อตรวจสอบโครงสร้างของตัวแปรและความเหมาะสมในการจัดกลุ่มเป็นปัจจัยแฝง โดยเป็นเทคนิคที่ใช้ลดจำนวนตัวแปรและรวมตัวแปรที่มีความสัมพันธ์กันไว้ในปัจจัยเดียวกัน ซึ่งสมการของวิเคราะห์องค์ประกอบเชิงสำรวจ คือ

$$Xi=\lambda i1F1+\lambda i2F2+\cdots+\lambda imFm+\epsilon i$$

โดยที่ $\boldsymbol{X_i}$ คือ ตัวแปรสังเกต, $\boldsymbol{F_j}$ คือ ปัจจัยแฝง, $\boldsymbol{\lambda_{ij}}$ คือ ค่าน้ำหนักองค์ประกอบ และ $\boldsymbol{\epsilon_i}$ คือ ค่าความคลาดเคลื่อน

3.4.3 นำข้อมูลตัวแปรมาวิเคราะห์การถดถอยพหุคูณ (Multiple Regression Analysis) เพื่อหาความสัมพันธ์ระหว่างตัวแปรและหาความน่าเชื่อถือของโมเดลที่ระดับนัยสำคัญทางสถิติที่ 0.05 โดยสมการถดถอยพหุคูณเขียนได้ คือ

$$Y=\beta 0+\beta 1X1+\beta 2X2+\cdots+\beta kXk+\varepsilon$$

โดยที่ $Y$ คือ ตัวแปรตาม, $\beta_0$ คือ ค่าคงที่, $\beta_1$ ถึง $\beta_k$ คือ ค่าสัมประสิทธิ์ของตัวแปรอิสระ, และ $\varepsilon$ คือ ค่าความคลาดเคลื่อนของสมการ

## 4. ผลการดำเนินงาน

ในการศึกษาวิจัยครั้งนี้ มีผู้ตอบแบบสอบถามทั้งหมด 416 คน โดยเป็นกลุ่มประชากรที่มีความเกี่ยวข้องในกระบวนการออกแบบที่เคยมีประสบการณ์ หรือมีความรู้เกี่ยวกับการใช้เทคโนโลยีเสมือนจริง ซึ่งมีข้อมูลที่สามารถนำมาวิเคราะห์ได้จำนวน 416 คน

**4.1 ผลการทดสอบความเชื่อมั่นของแบบสอบถาม (Reliability Analysis)**

**ตารางที่ 2**: ผลการวิเคราะห์ค่าความเชื่อมั่น ค่าเฉลี่ย และค่าส่วนเบี่ยงเบนมาตราฐาน

| ปัจจัย | ค่าเฉลี่ย | S.D. | Cronbach's Alpha |
|---|---|---|---|
| การรับรู้ความง่ายในการใช้งาน | 3.850 | 0.630 | 0.760 |
| การรับรู้ถึงประโยชน์ | 3.890 | 0.550 | 0.710 |
| อิทธิพลของสังคม | 3.850 | 0.640 | 0.730 |
| การสนับสนุน | 3.880 | 0.600 | 0.740 |
| **ปัจจัย** | **ค่าเฉลี่ย** | **S.D.** | **Cronbach's Alpha** |
| ด้านทัศนคติที่มีต่อการใช้งาน | 3.890 | 0.550 | 0.710 |

ผลการวิเคราะห์พบว่าแบบสอบถามมีความเชื่อมั่นอยู่ในระดับที่ยอมรับได้ (0.710–0.960) แสดงถึงความน่าเชื่อถือของเครื่องมือ โดยปัจจัยด้านการรับรู้ความง่ายในการใช้งานและการรับรู้ประโยชน์มีค่าเฉลี่ยสูงสุด (Mean = 3.890) และมีค่าส่วนเบี่ยงเบนมาตรฐานต่ำ (0.550–0.640) สะท้อนว่าความคิดเห็นของกลุ่มตัวอย่างมีความสอดคล้องกันในระดับสูง

**4.3 ผลการวิเคราะห์องค์ประกอบเชิงสำรวจ (EFA)**

**ตารางที่ 3** : ค่า KMO และ Bartlett's Test Sphericity จากการวิเคราะห์ความสัมพันธ์ของปัจจัย

| KMO and Bartlett's Test | | |
|---|---|---|
| Kaiser-Meyer-Olkin Measure of Sampling Adequacy | | 0.913 |
| Bartlett's Test of Sphericity | Approx. Chi-Square | 3652.487 |
| | DF | 210.000 |
| | Sig. | 0.001 |

ผลการทดสอบความเหมาะสมของข้อมูลพบว่า Bartlett's Test มีค่า Chi-Square = 3652.487, df = 210 และ Sig. = 0.001 (< 0.05) แสดงว่าข้อมูลมีความเหมาะสมในการวิเคราะห์องค์ประกอบในระดับดีมาก โดยกำหนดจำนวนองค์ประกอบ 5 ปัจจัย และใช้เกณฑ์ Factor Loading มากกว่า 0.500 เพื่อให้โมเดลมีความหมายทางสถิติในการหาความสัมพันธ์ของแต่ละปัจจัยดังนี้

**Exploratory Factor Analysis**

Factor Loadings

| | Factor | | | | | |
|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | Uniqueness |
| PEOU1 | 0.682 | | | | | 0.5191 |
| PEOU2 | 0.820 | | | | | 0.3011 |
| PEOU3 | 0.662 | | | | | 0.5200 |
| PEOU4 | 0.789 | | | | | 0.3628 |
| PEOU5 | 0.524 | | | | | 0.9019 |
| PU1 | | 0.776 | | | | 0.3949 |
| PU2 | | 0.960 | | | | 0.0742 |
| PU3 | | 0.680 | | | | 0.5246 |
| PU4 | | 0.585 | | | | 0.9128 |
| PU5 | | 0.567 | | | | 0.5930 |
| SI1 | | | 0.758 | | | 0.3188 |
| SI2 | | | 0.519 | | | 0.7994 |
| SI3 | | | 0.576 | | | 0.5720 |
| SI4 | | | 0.737 | | | 0.3224 |
| FC1 | | | | 0.583 | | 0.9916 |
| FC2 | | | | 0.545 | | 0.6910 |
| FC3 | | | | 0.641 | | 0.5758 |
| FC4 | | | | 0.568 | | 0.6674 |
| AU1 | | | | | 0.532 | 0.7023 |
| AU2 | | | | | 0.600 | 0.6168 |
| AU3 | | | | | 0.591 | 0.6121 |
| AU4 | | | | | 0.607 | 0.5853 |
| AU5 | | | | | 0.540 | 0.6475 |

Note. 'Minimum residual' extraction method was used in combination with a 'varimax' rotation

**ภาพที่ 4** : Rotated Factor Matrix

**4.4 ผลการวิเคราะห์สมการถดถอยพหุคูณ (Multiple Regression Analysis)**

**ตารางที่ 4** : ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปร (Correlation Matrix)

| ตัวแปร | PEOU | PU | SI | FC | AU |
|---|---|---|---|---|---|
| PEOU | 1.000 | 0.702** | 0.658** | 0.671** | 0.738** |
| PU | 0.702** | 1.000 | 0.689** | 0.693** | 0.756** |
| SI | 0.658** | 0.689** | 1.000 | 0.662** | 0.701** |
| FC | 0.671** | 0.693** | 0.662** | 1.000 | 0.715** |
| AU | 0.738** | 0.756** | 0.701** | 0.715** | 1.000 |

**หมายเหตุ: ** มีนัยสำคัญทางสถิติที่ระดับ 0.01**

ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปรอยู่ในช่วง 0.512–0.756 และมีนัยสำคัญทางสถิติที่ระดับ 0.001 ทั้งหมด โดยไม่มีค่าความสัมพันธ์เกิน 0.800 จึงไม่พบปัญหา Multicollinearity และสามารถนำไปวิเคราะห์การถดถอยพหุคูณได้

**ตารางที่ 5** : แสดงการวิเคราะห์ถดถอยพหุคูณการยอมรับเทคโนโลยี

| Model Summary | | | | |
|---|---|---|---|---|
| Model | R | R-Square | Adjusted R Square | Std. Error of the Estimate |
| 1 | 0.821 | 0.674 | 0.670 | 0.512 |

ผลการวิเคราะห์การถดถอยพหุคูณพบว่า ตัวแปรต้นทั้ง 5 สามารถอธิบายความแปรปรวนของตัวแปรตามได้ร้อยละ 67.40 ($R^2 = 0.674$) และเมื่อปรับค่าแล้ว (Adjusted $R^2 = 0.670$) ยังคงมีความสามารถในการอธิบายได้ในระดับสูง

**ตารางที่ 6** : การวิเคราะห์ Anova ถดถอยพหุคูณ

| ANOVA | | | | | | |
|---|---|---|---|---|---|---|
| Model | Source | Sum of Squares | df | Mean Square | F | Sig. |
| 1 | Regression | 142.380 | 5 | 28.470 | 167.81 | 0.001 |
| | Residual | 68.910 | 410 | 0.168 | | |
| | Total | 211.290 | 415 | | | |

a. (Dependent Variable) : VR

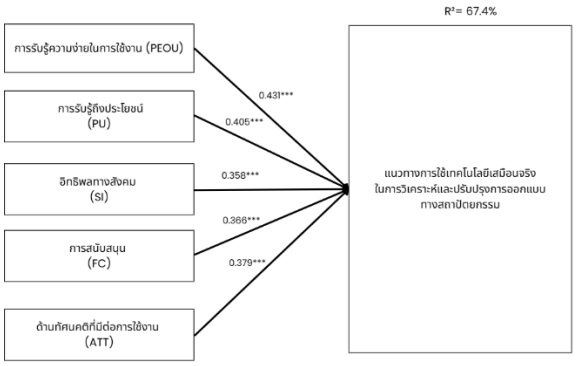
b. (Constant), PEOU, PU, SI, FC, AU

ผลการวิเคราะห์ ANOVA พบว่าโมเดลมีความเหมาะสม ($F = 167.810$, $Sig. = 0.001 < 0.05$) และมีตัวแปรอย่างน้อยหนึ่งตัว (PEOU, PU, SI, FC, AU) ที่ส่งผลต่อการยอมรับเทคโนโลยีอย่างมีนัยสำคัญทางสถิติ

**ตารางที่ 7** : ผลการวิเคราะห์การถดถอยพหุคูณ

| Model | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. | Collinearity Statistics | |
|---|---|---|---|---|---|---|---|
| | B | Std. Error | Beta | | | Tolerance | VIF |
| (Const) | 0.436 | 0.185 | - | 2.367 | 0.001 | - | - |
| PEOU | 0.418 | 0.062 | 0.431 | 6.743 | 0.001 | 0.738 | 1.610 |
| PU | 0.392 | 0.060 | 0.405 | 6.533 | 0.001 | 0.594 | 1.683 |
| SI | 0.327 | 0.054 | 0.358 | 6.054 | 0.001 | 0.702 | 1.425 |
| FC | 0.341 | 0.058 | 0.366 | 5.885 | 0.001 | 0.665 | 1.503 |
| AU | 0.348 | 0.056 | 0.379 | 6.218 | 0.001 | 0.621 | 1.355 |

เมื่อพิจารณาค่าสัมประสิทธิ์การถดถอยมาตรฐาน พบว่าตัวแปรที่มีอิทธิพลต่อการใช้เทคโนโลยีเสมือนจริงในการออกแบบทางสถาปัตยกรรมมากที่สุด คือ การรับรู้ความง่ายในการใช้งาน (PEOU) ($\beta = 0.431$) รองลงมาคือ การรับรู้ประโยชน์ (PU) ($\beta = 0.405$) การยอมรับการใช้งาน (AU) ($\beta = 0.379$) การสนับสนุนหรือสิ่งอำนวยความสะดวก (FC) ($\beta = 0.366$) และอิทธิพลทางสังคม (SI) ($\beta = 0.358$) ตามลำดับ




**ภาพที่ 5** โมเดลอธิบายปัจจัยที่มีต่อแนวทางการใช้เทคโนโลยีเสมือนจริงในการออกแบบทางสถาปัตยกรรม

## 5. สรุป

### 5.1 สรุปผลการวิจัย

งานวิจัยนี้เป็นการศึกษาการยอมรับเทคโนโลยีเสมือจริง โดยใช้วิธีวิจัยเชิงปริมาณ กลุ่มตัวอย่าง 416 คน เครื่องมือมีความเที่ยงตรงและเชื่อมั่นอยู่ในระดับดี ผลการวิเคราะห์ EFA แสดงว่าข้อมูลมีความเหมาะสมและสามารถสกัดได้ 5 องค์ประกอบ ขณะที่ผลการถดถอยพหุคูณพบว่า ตัวแปร PEOU, PU, SI, FC และ AU สามารถอธิบายการยอมรับเทคโนโลยีได้ในระดับสูง ($R^2 = 0.674$) แสดงว่าโมเดลมีประสิทธิภาพดีในการอธิบายตัวแปรตาม

### 5.2 อภิปรายผลการวิจัย

จากการศึกษาวิจัยเรื่องปัจจัยที่มีอิทธิพลต่อการยอมรับการใช้เทคโนโลยีเสมือนจริงในการออกแบบทางสถาปัตยกรรมในครั้งนี้ ผู้วิจัยได้สรุปผลการวิจัยเกี่ยวกับแนวคิดและทฤษฎีที่เกี่ยวข้องเพื่ออธิบายสมมติฐานได้ดังนี้

ผลการวิจัยพบว่า สมมติฐานทั้ง 5 ได้รับการยอมรับโดยทุกปัจจัยมีอิทธิพลเชิงบวกและมีนัยสำคัญทางสถิติ ($Sig. = 0.001$) ต่อการยอมรับการใช้เทคโนโลยี VR ได้แก่ การรับรู้ความง่ายในการใช้งาน ($\beta = 0.431$), การรับรู้ประโยชน์ ($\beta = 0.405$), ทัศนคติในการใช้งาน ($\beta = 0.379$), ความพร้อมด้านโครงสร้างพื้นฐาน ($\beta = 0.366$) และอิทธิพลทางสังคม ($\beta = 0.358$) โดยสะท้อนว่าการใช้

งานที่ง่าย มีประโยชน์ ได้รับการสนับสนุน และมีทัศนคติที่ดี จะช่วยส่งเสริมการยอมรับเทคโนโลยีเสมือนจริงในงานสถาปัตยกรรมมากขึ้น

การอภิปรายผลของงานวิจัยเรื่องปัจจัยที่มีอิทธิพลต่อการยอมรับการใช้เทคโนโลยีเสมือนจริงในการออกแบบทางสถาปัตยกรรม สามารถนำไปใช้ประโยชน์ได้ในหลายมิติ ทั้งเชิงวิชาการและเชิงปฏิบัติ โดยผลการวิจัยที่ชี้ให้เห็นถึงบทบาทของการรับรู้ความง่ายในการใช้งาน การรับรู้ถึงประโยชน์ อิทธิพลทางสังคม ความพร้อมของโครงสร้างพื้นฐาน และทัศนคติของผู้ใช้งาน สามารถใช้เป็นแนวทางในการวางแผนและพัฒนาเทคโนโลยีให้สอดคล้องกับความต้องการของผู้ใช้งานมากยิ่งขึ้น ช่วยส่งเสริมการยอมรับและการใช้งานเทคโนโลยีเสมือนจริงในงานออกแบบทางสถาปัตยกรรมได้อย่างมีประสิทธิภาพ ช่วยยกระดับคุณภาพของงานออกแบบ และเพิ่มศักยภาพในการแข่งขันขององค์กรในยุคดิจิทัลได้อย่างยั่งยืน

**5.3 ข้อเสนอแนะ**

5.3.1 ควรขยายกลุ่มตัวอย่างให้มีความหลากหลายด้านการศึกษา ประสบการณ์ เพื่อให้ผลการวิจัยครอบคลุมและสามารถนำไปใช้อ้างอิงได้กว้างขึ้น

5.3.2 ศึกษาการเปรียบเทียบประสิทธิภาพและการยอมรับการใช้งานระหว่างเทคโนโลยีเสมือนจริงกับเทคโนโลยีอื่น เช่น AR และ BIM เพื่อหาแนวทางเลือกใช้ให้เหมาะสมกับงานสถาปัตยกรรมในแต่ละบริบท

5.3.3 ควรใช้การวิจัยเชิงคุณภาพเพิ่มเติม เช่น การสัมภาษณ์หรือสนทนากลุ่ม เพื่อเข้าใจมุมมองและประสบการณ์ผู้ใช้งานได้ลึกซึ้งยิ่งขึ้น

**เอกสารอ้างอิง**

[1] สิริรักษ์ และคณะ. (2560). การวิเคราะห์ปัจจัยด้านการรับรู้ที่มีผลต่อการยอมรับเทคโนโลยีเสมือนจริงของผู้ใช้งานกลุ่มเจเนอเรชั่นวาย.

[2] ชีวรัตน์. (2561). การยอมรับเทคโนโลยีการเรียนภาษาผ่านแอปพลิเคชันออนไลน์ในเขตกรุงเทพมหานครและปริมณฑล.

[3] อัจฉรา. (2560). ปัจจัยที่มีอิทธิพลต่อการยอมรับเทคโนโลยีสารสนเทศในการปฏิบัติงานของสำนักงานปลัดกระทรวงวิทยาศาสตร์และเทคโนโลยี.

[4] ภาณุพันธ์. (2565). พฤติกรรมความตั้งใจซื้อรถยนต์ส่วนบุคคลผ่านช่องทางออนไลน์.

[5] สุพรรณี. (2563). ปัจจัยที่ส่งผลต่อการใช้งานเทคโนโลยีดิจิทัลในการบริหารทรัพยากรมนุษย์ของพนักงานธนาคารของรัฐขนาดใหญ่ ABC.

# ปัจจัยที่มีอิทธิพลต่อการยอมรับการใช้เครื่องชำระเงินด้วยตนเองในร้านค้าปลีกสมัยใหม่ประเภทซูเปอร์เซ็นเตอร์และซูเปอร์มาร์เก็ต

# Factors Influencing the Acceptance of Self-Checkout Systems in Modern Retail Stores: A Study of Supercenters and Supermarkets

*อโณทัย อรุณทอง (Anothai Arunthong)[1] และธนพล เจนสุทธิเวชกุล (Tanapon Jensuttiwetchakul)[2]*

*ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล*

*มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ*

*[1]s6707021857082@email.kmutnb.ac.th, [2]tanapon.j@itd.kmutnb.ac.th*

## บทคัดย่อ

*งานวิจัยนี้มีวัตถุประสงค์เพื่อศึกษาปัจจัยที่มีอิทธิพลต่อการยอมรับการใช้เครื่องชำระเงินด้วยตนเอง โดยประยุกต์ใช้ทฤษฎีการยอมรับเทคโนโลยี UTAUT ร่วมกับปัจจัยด้านสถานการณ์ เก็บรวบรวมข้อมูลโดยใช้แบบสอบถามออนไลน์จากกลุ่มตัวอย่างจำนวน 328 คน และวิเคราะห์ข้อมูลด้วยการวิเคราะห์องค์ประกอบเชิงสำรวจและการถดถอยพหุคูณ ผลการศึกษาพบว่า ปัจจัยด้านตำแหน่งที่ตั้ง ($\beta = 0.305$, $p < 0.001$) ความคาดหวังในความพยายาม ($\beta = 0.269$, $p < 0.001$) การรับรู้ถึงความแออัดของสภาพแวดล้อม ($\beta = 0.223$, $p < 0.001$) และความคาดหวังในประสิทธิภาพ ($\beta = 0.109$, $p < 0.05$) มีอิทธิพลต่อทัศนคติในการใช้งานอย่างมีนัยสำคัญทางสถิติ ขณะที่ทัศนคติในการใช้งาน ($\beta = 0.392$, $p < 0.001$) มีอิทธิพลต่อความตั้งใจในการใช้งานสูงที่สุด รองลงมาคือสิ่งอำนวยความสะดวก ($\beta = 0.270$, $p < 0.001$) และอิทธิพลทางสังคม ($\beta = 0.257$, $p < 0.001$) โดยแบบจำลองสามารถอธิบายความแปรปรวนของทัศนคติในการใช้งานได้ร้อยละ 61.9 และความตั้งใจในการใช้งานได้ร้อยละ 60.4*

**คำสำคัญ:** เทคโนโลยีการบริการตนเอง, เครื่องชำระเงินด้วยตนเอง, ปัจจัยด้านสถานการณ์, การยอมรับเทคโนโลย.

## Abstract

*This study aims to examine the factors influencing the acceptance of self-checkout systems by applying the Unified Theory of Acceptance and Use of Technology (UTAUT) together with situational factors. Data were collected using an online questionnaire from a sample of 328 respondents and analyzed using Exploratory Factor Analysis and Multiple Regression Analysis. The results indicate that location ($\beta = 0.305$, $p < 0.001$), effort expectancy ($\beta = 0.269$, $p < 0.001$), perceived crowdedness ($\beta = 0.223$, $p < 0.001$), and performance expectancy ($\beta = 0.109$, $p < 0.05$) have significant effects on attitude toward use. In addition, attitude toward use ($\beta = 0.392$, $p < 0.001$) has the strongest influence on intention to use, followed by facilitating conditions ($\beta = 0.270$, $p < 0.001$) and social influence ($\beta = 0.257$, $p < 0.001$). The model explains 61.9% of the variance in attitude toward use and 60.4% of the variance in intention to use.*

**Keywords:** Self-Service Technology, Self-Checkout Technology, Situational Factors, Technology Acceptance.

## 1. บทนำ

ในช่วงการแพร่ระบาดของโรค COVID-19 ภาคธุรกิจค้าปลีกได้มีการปรับตัวเพื่อลดการสัมผัสระหว่างบุคคลและรองรับมาตรการเว้นระยะห่างทางสังคม ส่งผลให้เทคโนโลยีการบริการตนเองโดยเฉพาะเครื่องชำระเงินด้วยตนเอง (Self-Checkout) ถูกนำมาใช้อย่างแพร่หลายเพื่อเพิ่มความสะดวก รวดเร็ว ลดความแออัด และเพิ่มประสิทธิภาพ

ในการดำเนินงาน รวมถึงลดต้นทุนด้านแรงงานที่มีแนวโน้มเพิ่มสูงขึ้นอย่างต่อเนื่อง

ในปัจจุบัน แม้ว่าจะมีการลงทุนติดตั้งเทคโนโลยีดังกล่าวอย่างต่อเนื่อง แต่การใช้งานจริงของผู้บริโภคยังไม่เป็นไปตามที่คาดการณ์ โดยผู้บริโภคจำนวนมากยังคงเลือกใช้บริการผ่านพนักงานตามรูปแบบเดิม สะท้อนให้เห็นถึงช่องว่างระหว่างการนำเทคโนโลยีมาใช้กับการยอมรับของผู้ใช้งาน

จากการศึกษางานวิจัยที่ผ่านมา พบว่างานวิจัยเกี่ยวกับการยอมรับเทคโนโลยีเครื่องชำระเงินด้วยตนเองส่วนใหญ่มุ่งเน้นการรับรู้และทัศนคติของผู้ใช้ โดยใช้แบบจำลองการยอมรับเทคโนโลยีเป็นกรอบแนวคิดหลักในการอธิบายการยอมรับของผู้บริโภค อย่างไรก็ตาม งานวิจัยที่ผ่านมาให้ความสำคัญกับปัจจัยภายในของผู้ใช้เป็นหลัก และยังไม่ครอบคลุมปัจจัยเชิงสถานการณ์ที่เกิดขึ้นในบริบทการใช้งานจริงเช่น ความแออัดของสภาพแวดล้อม และความสะดวกในการเข้าถึง ซึ่งอาจมีอิทธิพลต่อการตัดสินใจใช้งาน

ดังนั้น การศึกษาครั้งนี้จึงมุ่งวิเคราะห์ปัจจัยที่มีอิทธิพลต่อการยอมรับการใช้เครื่องชำระเงินด้วยตนเอง โดยเลือกใช้แบบจำลองการยอมรับเทคโนโลยี UTAUT ซึ่งสามารถอธิบายการยอมรับเทคโนโลยีได้อย่างครอบคลุม ร่วมกับปัจจัยด้านสถานการณ์เพื่ออธิบายให้ครอบคลุมทั้งด้านการรับรู้และบริบทการใช้งานจริง ซึ่งนำไปสู่การระบุปัจจัยสำคัญที่ส่งเสริมการยอมรับเทคโนโลยี และใช้เป็นแนวทางในการพัฒนาและปรับปรุงรูปแบบการให้บริการสอดคล้องกับความต้องการของผู้บริโภค

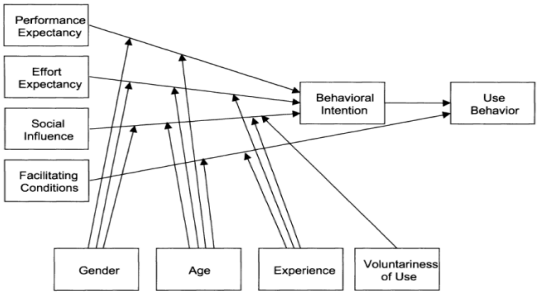
## 2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

### 2.1 เทคโนโลยีการบริการตนเอง (Self-Service Technologies)

เทคโนโลยีการบริการตนเอง [1] คือ ระบบเทคโนโลยีที่ทำหน้าที่เป็นตัวกลางในการให้บริการช่วยให้ผู้ใช้บริการสามารถเข้าถึงและดำเนินการรับบริการได้ด้วยตนเอง โดยไม่จำเป็นต้องมีปฏิสัมพันธ์โดยตรงกับพนักงาน ตัวอย่าง เช่น ตู้เอทีเอ็ม เครื่องคีออสก์ ตู้เช็กอินในสนามบิน และเครื่องชำระเงินด้วยตนเอง ซึ่งช่วยเพิ่มความสะดวกรวดเร็ว ลดเวลาการรอคอยของผู้ใช้บริการ และช่วยลดต้นทุนด้านแรงงาน รวมทั้งเพิ่มประสิทธิภาพในการให้บริการของผู้ประกอบการ โดยเฉพาะในร้านค้าปลีกสมัยใหม่ที่นำ Self-Checkout มาใช้เพื่อลดความแออัดบริเวณจุดชำระเงินและเพิ่มทางเลือกในการชำระเงินให้กับลูกค้า

### 2.2 ทฤษฎีการยอมรับเทคโนโลยีแบบบูรณาการ UTAUT (Unified Theory of Acceptance and Use of Technology)

เป็นกรอบแนวคิดที่อธิบายการยอมรับเทคโนโลยีของผู้ใช้ โดยพัฒนามาจากการรวมแนวคิดของหลายทฤษฎี เช่น TAM, TRA และ TPB พัฒนาขึ้นโดย Venkatesh et al. [2] ซึ่งเสนอว่าความตั้งใจในการใช้เทคโนโลยีของบุคคลได้รับอิทธิพลจากปัจจัยสำคัญ 4 ด้าน ได้แก่ ความคาดหวังในผลการปฏิบัติงาน ความคาดหวังในความพยายาม อิทธิพลทางสังคม และ เงื่อนไขอำนวยความสะดวก โดยปัจจัยเหล่านี้ส่งผลต่อความตั้งใจเชิงพฤติกรรมในการใช้งานซึ่งนำไปสู่พฤติกรรมการใช้งานจริงของผู้ใช้เทคโนโลยี




**ภาพที่ 1** แบบจำลองการยอมรับเทคโนโลยี UTAUT

### 2.3 แนวคิดด้านปัจจัยสถานการณ์ (Situational Factors)

แนวคิดด้านปัจจัยสถานการณ์ (Situational Factors) อธิบายถึงอิทธิพลของสภาพแวดล้อมหรือบริบทเฉพาะหน้าที่มีผลต่อการตัดสินใจและพฤติกรรมของผู้บริโภคในช่วงเวลาหนึ่ง ๆ โดยปัจจัยดังกล่าวเป็นเงื่อนไขชั่วคราวที่เกิดขึ้นระหว่างการใช้บริการ และสามารถส่งผลต่อการเลือกใช้สินค้า บริการ หรือเทคโนโลยีของผู้บริโภค แนวคิดนี้ถูกเสนอโดย Belk [3] ได้กล่าวว่าสถานการณ์หรือสภาพแวดล้อมที่ผู้บริโภคเผชิญในขณะนั้นสามารถมีอิทธิพลต่อ ทัศนคติ ความตั้งใจ และพฤติกรรมการใช้บริการของผู้บริโภคได้ เช่น สภาพแวดล้อมทางสังคม - การรับรู้ถึงความแออัดของสภาพแวดล้อม และสภาพแวดล้อมทางกายภาพ - ตำแหน่งที่ตั้งและความสะดวกในการเข้าถึง

**2.4 การวิเคราะห์องค์ประกอบเชิงสำรวจ (Exploratory Factor Analysis)**

การวิเคราะห์องค์ประกอบเชิงสำรวจ (Exploratory Factor Analysis : EFA) [4] เป็นเทคนิคทางสถิติที่ใช้ในการสำรวจโครงสร้างความสัมพันธ์ระหว่างตัวแปรที่สังเกตได้หลายตัว เพื่อค้นหาปัจจัยแฝงที่อยู่เบื้องหลังตัวแปรเหล่านั้น โดยมีวัตถุประสงค์เพื่อจัดกลุ่มตัวแปรที่มีความสัมพันธ์กันให้อยู่ในองค์ประกอบเดียวกัน ซึ่งช่วยลดจำนวนตัวแปรและอธิบายโครงสร้างของข้อมูลได้อย่างมีประสิทธิภาพ

**2.5 การวิเคราะห์การถดถอยเชิงพหุคูณ (Multiple Regression Analysis)**

การวิเคราะห์การถดถอยเชิงพหุคูณเป็นวิธีการวิเคราะห์ทางสถิติที่ใช้ศึกษาความสัมพันธ์ระหว่าง ตัวแปรตามจำนวน 1 ตัวแปร กับ ตัวแปรอิสระ ตั้งแต่ 2 ตัวแปรขึ้นไป โดยมีวัตถุประสงค์เพื่ออธิบายและทำนายการเปลี่ยนแปลงของตัวแปรตามจากอิทธิพลของตัวแปรอิสระหลายตัวพร้อมกัน ช่วยให้สามารถประเมินได้ว่าตัวแปรอิสระแต่ละตัวมีอิทธิพลต่อการเปลี่ยนแปลงของตัวแปรตามมากน้อยเพียงใด ภายใต้การควบคุมผลกระทบของตัวแปรอื่น ๆ ในแบบจำลองเดียวกัน ดังสมการที่ 1

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n + \varepsilon \qquad (1)$$

$y$ คือ ตัวแปรตาม (Dependent Variable)

$x$ คือ ตัวแปรอิสระ (Independent Variable)

$\beta_0$ คือ ค่าเริ่มต้นของเส้นสมการถดถอย

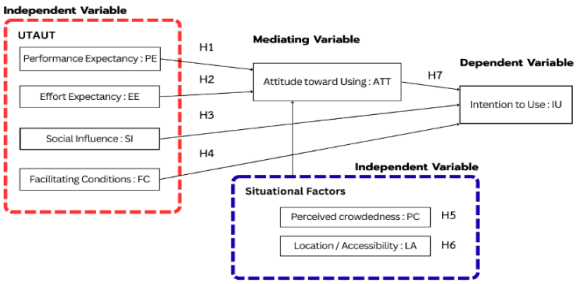
$\beta_n x_n$ คือ สัมประสิทธิ์การถดถอย

$\varepsilon$ คือ ค่าความคลาดเคลื่อนอย่างสุ่ม

## 3. วิธีการดำเนินงานวิจัย

**3.1 กำหนดกรอบแนวคิด**

จากการศึกษางานวิจัยที่เกี่ยวข้อง ผู้วิจัยได้กำหนดกรอบแนวความคิดโดยประยุกต์ใช้ทฤษฎีการยอมรับเทคโนโลยีแบบบูรณาการ UTAUT ร่วมกับปัจจัยด้านสถานการณ์ ได้แก่ การรับรู้ถึงความแออัดของสภาพแวดล้อมและตำแหน่งที่ตั้งหรือความสะดวกในการเข้าถึง เพื่อใช้เป็นโครงสร้างเชิงทฤษฎีในการอธิบายและวิเคราะห์ปัจจัยที่มีอิทธิพลต่อทัศนคติและความตั้งใจใช้เครื่องชำระสินค้าด้วยตนเอง โดยมีกรอบแนวความคิดดังภาพที่ 2




**ภาพที่ 2** กรอบแนวความคิดของงานวิจัย

**3.2 ประชากรและกลุ่มตัวอย่าง**

ในการวิจัยครั้งนี้ ผู้วิจัยกำหนดกลุ่มตัวอย่างจากประชากรที่พักอาศัยอยู่ในเขตกรุงเทพมหานครและมีความรู้จักเกี่ยวกับเครื่องชำระสินค้าด้วยตนเอง ทั้งผู้ที่มีและไม่มีประสบการณ์ในการใช้งาน โดยใช้วิธีการคัดเลือกแบบสุ่มตามความสมัครใจ (Voluntary Sampling) [5]

การกำหนดขนาดกลุ่มตัวอย่างใช้โปรแกรม G*Power อ้างอิงแนวคิดการวิเคราะห์อำนาจการทดสอบของ Cohen [6] โดยกำหนดค่า Effect Size เท่ากับ 0.30 ระดับนัยสำคัญทางสถิติ ($\alpha$) เท่ากับ 0.05 และอำนาจการทดสอบ (Power) เท่ากับ 0.80 รวมถึงกำหนดค่าองศาอิสระ (df) เท่ากับ 32 ตามจำนวนข้อคำถามในแบบสอบถาม ส่งผลให้ได้ขนาดกลุ่มตัวอย่างขั้นต่ำที่เหมาะสมจำนวน 280 คน

**3.3 การเก็บรวบรวมข้อมูล**

การวิจัยครั้งนี้ใช้แบบสอบถามอิเล็กทรอนิกส์ผ่านระบบ Microsoft Forms เป็นเครื่องมือในการเก็บรวบรวมข้อมูลเชิงปริมาณจากกลุ่มตัวอย่าง โดยแบบสอบถามแบ่งออกเป็น 2 ส่วน ได้แก่ ส่วนที่ 1 ข้อมูลทั่วไปของผู้ตอบแบบสอบถาม ประกอบด้วย เพศ อายุ อาชีพ พื้นที่อยู่อาศัย รวมถึงการรับรู้และประสบการณ์การใช้งานเครื่องชำระเงินด้วยตนเอง ส่วนที่ 2 ข้อคำถามตามกรอบแนวคิดการวิจัย ครอบคลุมตัวแปรที่ใช้ในการศึกษา โดยการวัดระดับความคิดเห็นใช้มาตราส่วนประมาณค่าแบบ Likert Scale 7 ระดับ ทั้งนี้ข้อคำถามที่ใช้ในแบบสอบถาม พัฒนาขึ้นจากการศึกษาทฤษฎี แนวคิด และงานวิจัยที่เกี่ยวข้อง โดยอ้างอิงและประยุกต์จากข้อคำถามในงานวิจัยที่ได้รับการเผยแพร่

ในแหล่งวิชาการ จากนั้นนำมาแปลและเรียบเรียงถ้อยคำให้เหมาะสมและสอดคล้องกับบริบทของการวิจัย

**3.4 การวิเคราะห์ข้อมูลด้วยวิธีการทางสถิติ**

นำข้อมูลจากกลุ่มตัวอย่างจำนวน 328 คน ที่ผ่านการตรวจสอบความถูกต้องและความสมบูรณ์เรียบร้อยแล้วมาทำการวิเคราะห์ด้วยโปรแกรมทางสถิติ Jamovi เพื่อตอบวัตถุประสงค์และทดสอบสมมติฐานของการวิจัย โดยทดสอบความเป็นปกติของข้อมูล (Normality Test) เพื่อพิจารณาว่าข้อมูลมีการกระจายตัวเป็นแบบปกติหรือไม่ ซึ่งเป็นเงื่อนไขพื้นฐานสำหรับการวิเคราะห์ทางสถิติ และทำการวิเคราะห์ความเชื่อมั่นของแบบสอบถาม (Reliability Analysis) เพื่อประเมินความสอดคล้องภายในของข้อคำถามในแต่ละตัวแปร โดยใช้ค่าสัมประสิทธิ์ Cronbach's Alpha เมื่อผ่านการตรวจสอบความเชื่อมั่นของเครื่องมือแล้ว จึงนำข้อมูลไปวิเคราะห์องค์ประกอบเชิงสำรวจ (EFA) เพื่อสำรวจโครงสร้างของตัวแปรแฝงที่อยู่ภายใต้ตัวแปรสังเกตได้ โดยพิจารณาค่า KMO และ Bartlett's Test of Sphericity เพื่อประเมินความเหมาะสมของข้อมูลสำหรับการวิเคราะห์ ตัวแปรที่ผ่านการวิเคราะห์จะถูกนำไปทดสอบด้วยการวิเคราะห์สมการถดถอยพหุคูณ (MRA) เพื่อศึกษาความสัมพันธ์และอิทธิพลของปัจจัยต่าง ๆ ที่มีต่อทัศนคติและความตั้งใจในการใช้เครื่องชำระสินค้าด้วยตนเอง โดยพิจารณาระดับนัยสำคัญทางสถิติที่ 0.05 ร่วมกับค่าสัมประสิทธิ์การถดถอยและค่าสถิติที่เกี่ยวข้องของแบบจำลอง

## 4. ผลการดำเนินงาน

**4.1 ผลการวิเคราะห์ข้อมูลทั่วไปของกลุ่มตัวอย่าง**

จากการสำรวจกลุ่มตัวอย่างทั้งหมด 328 คนพบว่าเป็นเพศชาย จำนวน 140 คน คิดเป็นร้อยละ 42.70 เพศหญิง จำนวน 174 คน คิดเป็นร้อยละ 53.00 และเพศอื่น ๆ จำนวน 14 คน คิดเป็นร้อยละ 4.30 อยู่ในช่วงอายุ 20–30 ปี จำนวน 99 คน คิดเป็นร้อยละ 30.20 ช่วงอายุ 31–40 ปี จำนวน 154 คน คิดเป็นร้อยละ 47.00 ช่วงอายุ 41–50 ปี จำนวน 63 คน คิดเป็นร้อยละ 19.20 ช่วงอายุ 51–60 ปี จำนวน 9 คน คิดเป็นร้อยละ 2.70 และกลุ่มอายุมากกว่า 60 ปีขึ้นไป จำนวน 3 คน คิดเป็นร้อยละ 0.90 ประกอบอาชีพ พนักงานบริษัทเอกชน จำนวน 157 คน คิดเป็นร้อยละ 47.90 ข้าราชการหรือพนักงานรัฐวิสาหกิจ จำนวน 101 คน คิดเป็นร้อยละ 30.80 เจ้าของกิจการหรือธุรกิจส่วนตัว จำนวน 40 คน คิดเป็นร้อยละ 12.20 ผู้รับจ้างทั่วไป จำนวน 25 คน คิดเป็นร้อยละ 7.60 นักเรียนหรือนักศึกษา จำนวน 4 คน คิดเป็นร้อยละ 1.20 และ พนักงานจ้างเหมาบริการจำนวน 1 คน คิดเป็นร้อยละ 0.30 ผู้ตอบแบบสอบถามทั้งหมดรู้จักเครื่องชำระสินค้าด้วยตนเองคิดเป็นร้อยละ 100 โดยมีผู้เคยใช้บริการเครื่องชำระสินค้าด้วยตนเอง จำนวน 265 คน คิดเป็นร้อยละ 80.80 และไม่เคยใช้บริการ จำนวน 63 คน คิดเป็นร้อยละ 19.20

**4.2 ผลทดสอบความเชื่อมั่นของแบบสอบถาม**

พบว่าค่าความเชื่อมั่นของแบบสอบถาม มีค่า Cronbach's Alpha อยู่ในช่วงระหว่าง 0.817 ถึง 0.906 ซึ่งอยู่ในระดับที่ยอมรับได้ แสดงว่าเครื่องมือแบบสอบถามที่ใช้ในการสำรวจฉบับนี้มีความน่าเชื่อถือ โดยแสดงข้อมูลตามตารางที่ 1

**ตารางที่ 1 :** ผลการทดสอบความเชื่อมั่นของแบบสอบถาม

| Scale Reliability Statistics | Mean | SD | Cronbach's α |
|---|---|---|---|
| ความคาดหวังในประสิทธิภาพ (PE) | 5.78 | 0.92 | 0.817 |
| คาดหวังในความพยายาม (EE) | 5.67 | 1.09 | 0.906 |
| อิทธิพลจากสังคม (SI) | 5.57 | 1.16 | 0.894 |
| การเข้าถึงข้อมูลและความสะดวก (FC) | 5.63 | 1.05 | 0.861 |
| การรับรู้ถึงความแออัดของสภาพแวดล้อม (PC) | 5.82 | 0.88 | 0.819 |
| ตำแหน่งที่ตั้ง / ความสะดวกในการเข้าถึง (LA) | 5.64 | 1.01 | 0.883 |
| ทัศนคติต่อการใช้งาน (ATT) | 5.65 | 0.99 | 0.884 |
| ความตั้งใจใช้งาน (IU) | 5.54 | 1.09 | 0.892 |

**4.3 ผลการวิเคราะห์องค์ประกอบเชิงสำรวจ (Exploratory Factor Analysis: EFA)**

ผลการวิเคราะห์พบว่าค่า KMO มีค่าเท่ากับ 0.967 ซึ่งจัดอยู่ในระดับตามเกณฑ์ที่กำหนด แสดงให้เห็นว่าข้อมูลมีขนาดกลุ่มตัวอย่างที่เพียงพอ และตัวแปรต่าง ๆ มีความสัมพันธ์กันในระดับที่เหมาะสม นอกจากนี้ ผลการทดสอบ Bartlett's Test of Sphericity พบว่ามีระดับนัยสำคัญทางสถิติต่ำกว่า 0.001 ซึ่งบ่งชี้ว่าตัวแปรมีความสัมพันธ์กันอย่างมีนัยสำคัญทางสถิติ และไม่ได้เป็นความสัมพันธ์แบบสุ่ม ผลที่ได้แสดงดังตารางที่ 2

**ตารางที่ 2 :** ค่า KMO และ Bartlett's Test of Sphericity

| KMO and Bartlett's Test | | |
|---|---|---|
| Kaiser-Meyer-Olkin Measure of Sampling Adequacy. | | 0.967 |
| Bartlett's Test of Sphericity | Approx. Chi-Square | 7508 |
| | df | 496 |
| | Sig. | <.001 |

ผลการสกัดองค์ประกอบและปรับหมุนแกน พบว่าสามารถจัดกลุ่มตัวแปรสังเกตออกเป็นองค์ประกอบได้ทั้งหมด 8 องค์ประกอบโดยตัวแปรสังเกตทุกตัวมีค่า Factor Loading ตั้งแต่ 0.30 ขึ้นไป ซึ่งแสดงให้เห็นว่าโครงสร้างองค์ประกอบมีลักษณะเป็น Simple Structure และไม่พบปัญหาการซ้อนทับของตัวแปร (Cross-loading) จึงสามารถนำองค์ประกอบที่ได้ ไปใช้ในการวิเคราะห์สมการถดถอยพหุคูณต่อไปได้ ดังแสดงในตารางที่ 3

**ตารางที่ 3 :** ค่าน้ำหนักองค์ประกอบของตัวแปรที่ศึกษา

| Factor Loadings | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Factor | | | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| EE4 | 0.705 | | | | | | | |
| EE1 | 0.679 | | | | | | | |
| EE3 | 0.656 | | | | | | | |
| EE2 | 0.654 | | | | | | | |
| SI2 | | 0.852 | | | | | | |
| SI3 | | 0.804 | | | | | | |
| SI1 | | 0.768 | | | | | | |
| SI4 | | 0.579 | | | | | | |
| ATT1 | | | 0.728 | | | | | |
| ATT2 | | | 0.670 | | | | | |
| ATT4 | | | 0.667 | | | | | |
| ATT3 | | | 0.578 | | | | | |
| LA4 | | | | 0.750 | | | | |
| LA3 | | | | 0.657 | | | | |
| LA2 | | | | 0.623 | | | | |
| LA1 | | | | 0.546 | | | | |
| IU2 | | | | | 0.718 | | | |
| IU4 | | | | | 0.667 | | | |
| IU3 | | | | | 0.645 | | | |
| IU1 | | | | | 0.599 | | | |
| PC4 | | | | | | 0.782 | | |
| PC3 | | | | | | 0.675 | | |
| PC1 | | | | | | 0.524 | | |
| PC2 | | | | | | 0.425 | | |
| FC1 | | | | | | | 0.737 | |
| FC4 | | | | | | | 0.526 | |
| FC3 | | | | | | | 0.479 | |
| FC2 | | | | | | | 0.316 | |
| PE4 | | | | | | | | 0.413 |
| PE3 | | | | | | | | 0.408 |
| PE2 | | | | | | | | 0.393 |
| PE1 | | | | | | | | 0.348 |

**4.4 ผลการวิเคราะห์สมการถดถอยพหุคูณ (Multiple Regression Analysis)**

การทดสอบสมมติฐานงานวิจัยจะวิเคราะห์การถดถอยพหุคูณทั้งหมด 2 ครั้ง ดังนี้

ครั้งที่ 1 วิเคราะห์ผลของตัวแปรอิสระ 4 ตัว ได้แก่ ความคาดหวังในประสิทธิภาพ (PE) ความคาดหวังในความพยายาม (EE) การรับรู้ถึงความแออัด (PC) และตำแหน่งที่ตั้ง (LA) ที่มีต่อทัศนคติในการใช้งาน (ATT) ผลการวิเคราะห์พบว่าแบบจำลองมีค่า Adjusted $R^2 = 0.619$ หรือร้อยละ 61.9 โดยตัวแปรทั้ง 4 มีอิทธิพลเชิงบวกต่อทัศนคติในการใช้งานอย่างมีนัยสำคัญทางสถิติ โดยตัวแปรส่วนใหญ่มีนัยสำคัญที่ระดับ 0.001 ยกเว้นความคาดหวังในประสิทธิภาพที่มีนัยสำคัญที่ระดับ 0.05 ทั้งนี้ ตัวแปรที่มีอิทธิพลสูงสุดคือ ตำแหน่งที่ตั้ง รองลงมาคือ ความคาดหวังในความพยายาม และการรับรู้ถึงความแออัดตามลำดับ ขณะที่ความคาดหวังในประสิทธิภาพมีอิทธิพลน้อยที่สุด ดังตารางที่ 4

**ตารางที่ 4 :** Model Fit Measures และ Coefficients ของ ATT

| Model | R | $R^2$ | Adjusted $R^2$ |
|---|---|---|---|
| 1 | 0.790 | 0.623 | 0.619 |

| Model Coefficients - ATT | | | | | | Collinearity Statistics | |
|---|---|---|---|---|---|---|---|
| Predictor | Estimate | SE | t | p | Std. Estimate | VIF | Tolerance |
| Intercept | 0.474 | 0.241 | 1.96 | 0.05 | - | - | - |
| PE | 0.116 | 0.057 | 2.03 | 0.04 | 0.109 | 2.45 | 0.407 |
| EE | 0.245 | 0.051 | 4.83 | <.001 | 0.269 | 2.67 | 0.374 |
| PC | 0.245 | 0.053 | 4.63 | <.001 | 0.223 | 1.99 | 0.502 |
| LA | 0.299 | 0.052 | 5.74 | <.001 | 0.305 | 2.43 | 0.412 |

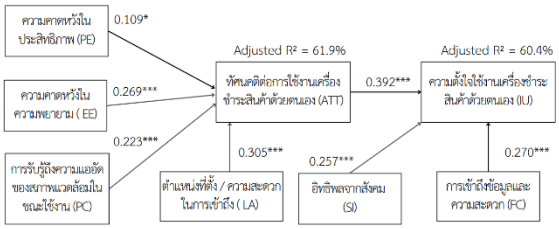
ครั้งที่ 2 วิเคราะห์ผลของตัวแปรอิสระ 3 ตัวแปรที่ส่งผลต่อความตั้งใจใช้งาน (IU) ผลการวิเคราะห์พบว่าตัวแปรอิสระ 3 ตัวแปรได้แก่ อิทธิพลจากสังคม (SI) การเข้าถึงข้อมูลและความสะดวก (FC) และทัศนคติต่อการใช้งาน (ATT) มีค่า Adjusted $R^2 = 0.604$ หรือร้อยละ 60.4 โดยทุกตัวแปรมีความสัมพันธ์เชิงบวกกับความตั้งใจใช้งานอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.001 โดยตัวแปรที่มีอิทธิพลสูงสุดต่อ ความตั้งใจในการใช้งาน คือ ทัศนคติในการใช้งาน รองลงมาคือ การเข้าถึงข้อมูลและความสะดวก และ อิทธิพลทางสังคมตามลำดับ ดังตารางที่ 5

**ตารางที่ 5 :** Model Fit Measures และ Coefficients ของ IU

| Model | R | $R^2$ | Adjusted $R^2$ |
|---|---|---|---|
| 2 | 0.780 | 0.608 | 0.604 |

| Model Coefficients - IU | | | | | | Collinearity Statistics | |
|---|---|---|---|---|---|---|---|
| Predictor | Estimate | SE | t | p | Std. Estimate | VIF | Tolerance |
| Intercept | 0.170 | 0.2424 | 0.70 | 0.484 | - | - | - |
| SI | 0.240 | 0.0394 | 6.09 | <.001 | 0.257 | 1.47 | 0.680 |
| FC | 0.281 | 0.0503 | 5.58 | <.001 | 0.270 | 1.93 | 0.518 |
| ATT | 0.434 | 0.0530 | 8.18 | <.001 | 0.392 | 1.90 | 0.525 |

จากผลการทดสอบสมมติฐานของตัวแปรทั้ง 8 ตัว สามารถพัฒนาเป็นแบบจำลองการอธิบายการยอมรับการใช้เครื่องชำระเงินด้วยตนเอง ดังภาพที่ 4




**ภาพที่ 4** แบบจำลองการยอมรับการใช้เครื่องชำระเงินด้วยตนเอง

## 5. สรุปผลและอภิปรายผล

จากวัตถุประสงค์ของการวิจัยที่มุ่งศึกษาปัจจัยด้านการยอมรับเทคโนโลยีและปัจจัยด้านสถานการณ์ที่มีอิทธิพลต่อการยอมรับการใช้เครื่องชำระเงินด้วยตนเอง ผลการวิจัยพบว่า ความคาดหวังในประสิทธิภาพ ความคาดหวังในความพยายาม การรับรู้ถึงความเออัดของสภาพแวดล้อม และตำแหน่งที่ตั้งหรือความสะดวกในการเข้าถึง มีอิทธิพลเชิงบวกต่อทัศนคติในการใช้งาน อย่างมีนัยสำคัญทางสถิติ โดยปัจจัยด้านตำแหน่งที่ตั้งและความสะดวกในการเข้าถึงมีอิทธิพลสูงสุด สะท้อนให้เห็นว่าผู้บริโภคให้ความสำคัญกับความสะดวกในการเข้าถึงเครื่อง ณ จุดใช้งานจริง เนื่องจากการตัดสินใจใช้งานมักเกิดขึ้นภายใต้ข้อจำกัดด้านเวลาและสถานการณ์เฉพาะหน้า ดังนั้น การจัดวางเครื่องในตำแหน่งที่มองเห็นได้ง่ายและเข้าถึงสะดวกจึงมีบทบาทสำคัญต่อการสร้างทัศนคติเชิงบวกต่อการใช้งาน ซึ่งสอดคล้องกับแนวคิดด้านปัจจัยสถานการณ์ของ Belk [3]

นอกจากนี้ ความคาดหวังในความพยายามมีอิทธิพลต่อทัศนคติในระดับสูง แสดงให้เห็นว่าความง่ายและความไม่ซับซ้อนของระบบเป็นองค์ประกอบสำคัญต่อการยอมรับเทคโนโลยี สอดคล้องกับทฤษฎี UTAUT ขณะที่การรับรู้ถึงความเออัดของสภาพแวดล้อมมีอิทธิพลเชิงบวกต่อทัศนคติเช่นกัน สะท้อนว่าผู้บริโภคมีแนวโน้มเลือกใช้เครื่องชำระเงินด้วยตนเองมากขึ้นเมื่อบริเวณแคชเชียร์มีความหนาแน่นหรือมีคิวรอจำนวนมาก สำหรับความคาดหวังในประสิทธิภาพ แม้มีอิทธิพลเชิงบวกต่อทัศนคติในการใช้งาน แต่มีอิทธิพลต่ำกว่าปัจจัยอื่นเนื่องจากผู้บริโภครับรู้ประโยชน์พื้นฐานของเทคโนโลยีดังกล่าวอยู่แล้ว จึงไม่ได้เป็นปัจจัยหลักในการตัดสินใจใช้งาน

ในส่วนของความตั้งใจในการใช้งาน พบว่า ทัศนคติในการใช้งานมีอิทธิพลสูงสุด รองลงมาคือการเข้าถึงข้อมูลและความสะดวก และอิทธิพลทางสังคม ตามลำดับ สะท้อนบทบาทของทัศนคติในฐานะตัวแปรส่งผ่านสำคัญระหว่างการรับรู้ของผู้บริโภคกับความตั้งใจใช้งาน โดยแบบจำลองสามารถอธิบายความแปรปรวนของความตั้งใจในการใช้งานได้ร้อยละ 60.4

## เอกสารอ้างอิง


[1] Meuter et al. (2000). “Self-service technologies: understanding customer satisfaction with technology-based service encounters”. Journal of Marketing, 64(3), 50–64.

[2] Venkatesh, et al. (2003). “User acceptance of information technology: Toward a unified view”. MIS Quarterly, 27(3), 425–478 https://doi.org/10.2307/30036540

[3] Belk, R. W. (1975). “Situational variables and consumer behavior”. Journal of Consumer Research, 2(3), 157–164. https://doi.org/10.1086/208627

[4] Hair et al. (2019). Multivariate data analysis (8th ed.). Cengage Learning.

[5] Formplus. (n.d.). Volunteer sampling: Definition, types & examples, Accessed: Feb. 12, 2026. [Online]. Available: https://www.formpl.us/blog/volunteer-sampling-definition-types-examples

[6] Cohen, J. (1988). “Statistical power analysis for the behavioral sciences (2nd ed.)”. Lawrence Erlbaum Associates.

# การวิเคราะห์ปัจจัยที่มีอิทธิพลต่อการใช้ปัญญาประดิษฐ์ในการเรียนภาษาจีน
# An Analysis of Factors Influencing the Use of AI in Chinese Language Learning

*ประวิทย์ เอกวีรภัทรสกุล (Prawit Ekweerapatarasakun)[1] และมณเฑียร รัตนศิริวงศ์วุฒิ (Montean Rattanasiriwongwut)[2]*
*ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล*
*มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ*
*s6707021857023@email.kmutnb.ac.th, montean.r@itd.kmutnb.ac.th*

## บทคัดย่อ

*การวิจัยนี้มีวัตถุประสงค์เพื่อวิเคราะห์ปัจจัยที่มีอิทธิพลต่อการใช้ปัญญาประดิษฐ์ (AI) ในการเรียนภาษาจีนของผู้เรียนชาวไทย โดยประยุกต์ใช้ทฤษฎี UTAUT2 (Venkatesh et al., 2012) เป็นกรอบแนวคิดหลัก ร่วมกับปัจจัยด้านความไว้วางใจ (Trust) และการปรับเนื้อหาเฉพาะบุคคล (Personalization) ซึ่งเป็นปัจจัยสำคัญในบริบทของเทคโนโลยี AI กลุ่มตัวอย่างคือผู้เรียนภาษาจีนชาวไทยที่มีประสบการณ์ใช้ AI จำนวน 755 คน เก็บข้อมูลด้วยแบบสอบถาม 7 ด้าน 44 ข้อ ผ่านการตรวจสอบความตรงเชิงเนื้อหาโดยผู้เชี่ยวชาญ 5 ท่าน (IOC 0.60–1.00) และทดสอบความเชื่อมั่นได้ค่าสัมประสิทธิ์แอลฟาของครอนบาครวม 0.956 วิเคราะห์ข้อมูลด้วยโปรแกรม Jamovi ใช้สถิติเชิงพรรณนา การวิเคราะห์สหสัมพันธ์เพียร์สัน การวิเคราะห์การถดถอยพหุคูณ การวิเคราะห์องค์ประกอบเชิงสำรวจ (EFA) และเชิงยืนยัน (CFA) ผลการวิจัยพบว่าแบบสอบถามมีคุณภาพสูง (KMO = 0.962, $\chi^2/df$ = 2.34, CFI = 0.942, RMSEA = 0.042) ทุกตัวแปรผ่านเกณฑ์ความเที่ยงตรงเชิงลู่เข้าและเชิงจำแนก การทดสอบสมมติฐาน 15 ข้อ ได้รับการยอมรับอย่างมีนัยสำคัญทางสถิติ โดยข้อค้นพบสำคัญ ได้แก่ (1) Performance Expectancy มีอิทธิพลสูงสุดต่อ Trust ($\beta$ = 0.434) และ Behavioral Intention ($\beta$ = 0.322) (2) Behavioral Intention มีอิทธิพลสูงสุดต่อ Continued Use ($\beta$ = 0.580) (3) Personalization มีอิทธิพลสูงมากต่อ Hedonic Motivation ($\beta$ = 0.689) และ Habit ($\beta$ = 0.638) (4) Habit มีอิทธิพลต่อ Continued Use ($\beta$ = 0.405) สูงกว่า Hedonic Motivation ($\beta$ = 0.312) (5) Continued Use มีอิทธิพลสูงสุดต่อ Skill Development ($\beta$ = 0.542) โดยอธิบายความแปรปรวนได้ 63.2% ประสิทธิภาพของ AI ความไว้วางใจ และการปรับเนื้อหาเฉพาะบุคคลเป็นปัจจัยสำคัญที่ส่งผลต่อการยอมรับและการใช้ AI อย่างต่อเนื่อง นำไปสู่การพัฒนาทักษะภาษาจีนของผู้เรียนไทยอย่างมีประสิทธิภาพ*

**คำสำคัญ:** ปัญญาประดิษฐ์ ภาษาจีน UTAUT2 การปรับเนื้อหาเฉพาะบุคคล การพัฒนาทักษะภาษาจีนอย่างมีประสิทธิภาพ

## Abstract

*This research aimed to analyze factors influencing AI usage in Chinese language learning among Thai learners by applying the UTAUT2 theory (Venkatesh et al., 2012) together with trust and personalization factors, which are crucial in the context of AI technology. The sample comprised 755 Thai Chinese language learners with AI experience. Data were collected using a 44-item, 7-dimension questionnaire validated by five experts (IOC 0.60–1.00) with high reliability (Cronbach's α = 0.956). Data were analyzed using Jamovi through descriptive statistics, Pearson's correlation, multiple regression, EFA, and CFA. The questionnaire met all quality criteria (KMO = 0.962, $\chi^2/df$ = 2.34, CFI = 0.942, RMSEA = 0.042). All 15 hypotheses were accepted with statistical significance. Key findings revealed:*

*(1) Performance Expectancy had the strongest influence on Trust (β = 0.434) and Behavioral Intention (β = 0.322); (2) Behavioral Intention had the strongest influence on Continued Use (β = 0.580); (3) Personalization strongly influenced Hedonic Motivation (β = 0.689) and Habit (β = 0.638); (4) Habit influenced Continued Use (β = 0.405) more than Hedonic Motivation (β = 0.312); (5) Continued Use had the strongest influence on Skill Development (β = 0.542), explaining 63.2% of its variance. AI efficiency, trust, and personalization are key factors influencing AI acceptance and continued use, leading to effective Chinese language skill development among Thai learners.*

**Keywords:** AI, Chinese Language, UTAUT2, Personalization, Effective Chinese language skill development.

## 1. บทนำ

ภาษาจีนเป็นภาษาที่มีความสำคัญทางเศรษฐกิจและวัฒนธรรมต่อประเทศไทยอย่างยิ่ง เนื่องจากความสัมพันธ์ทางการค้า การลงทุน และวัฒนธรรมระหว่างไทยกับจีนที่เติบโตอย่างต่อเนื่อง ทำให้ความต้องการเรียนรู้ภาษาจีนเพิ่มสูงขึ้นในทุกระดับ ตั้งแต่นักเรียน นักศึกษา ไปจนถึงบุคลากรทั่วไป อย่างไรก็ตาม ผู้เรียนไทยประสบอุปสรรคสำคัญในการเรียนภาษาจีนหลายประการ ได้แก่ ระบบเสียงวรรณยุกต์ที่มีความซับซ้อนถึง 4 เสียง ซึ่งแตกต่างจากภาษาไทย ระบบตัวอักษรที่มีจำนวนหลายพันตัวและมีโครงสร้างที่ซับซ้อน ต้องอาศัยการจดจำทั้งรูปเขียนและความหมาย รวมถึงโครงสร้างไวยากรณ์ที่มีความแตกต่างในรายละเอียด ปัญหาเหล่านี้ส่งผลให้ผู้เรียนจำนวนมากขาดแรงจูงใจและเลิกเรียนกลางคัน [1] ในขณะเดียวกัน เทคโนโลยีปัญญาประดิษฐ์ (AI) ได้พัฒนาไปอย่างรวดเร็วและนำเสนอทางออกที่น่าสนใจสำหรับปัญหาดังกล่าว โดยเฉพาะความสามารถในการปรับเนื้อหาให้เหมาะกับผู้เรียนแต่ละคน (Personalization) การให้ข้อมูลป้อนกลับทันที และการสร้างสภาพแวดล้อมการเรียนรู้ที่สมจริง [2] อย่างไรก็ตาม การที่ผู้เรียนจะยอมรับและใช้ AI อย่างต่อเนื่องนั้นขึ้นอยู่กับปัจจัยหลายประการ ได้แก่ ประสิทธิภาพที่คาดหวัง ความง่ายในการใช้ ความไว้วางใจ ความเพลิดเพลิน และการเกิดนิสัย ซึ่งปัจจัยเหล่านี้ยังขาดการศึกษาเชิงลึกในบริบทของผู้เรียนไทย

**วัตถุประสงค์**

1. เพื่อวิเคราะห์อิทธิพลของปัจจัยด้านประสิทธิภาพที่คาดหวังและปัจจัยด้านความง่ายในการใช้ ที่มีต่อความไว้วางใจ ความตั้งใจใช้ และการใช้อย่างต่อเนื่อง
2. เพื่อศึกษาอิทธิพลของปัจจัยด้านความไว้วางใจ ที่มีต่อความตั้งใจใช้
3. เพื่อวิเคราะห์บทบาทของปัจจัยด้านการปรับเนื้อหาเฉพาะบุคคล ที่มีอิทธิพลต่อแรงจูงใจด้านความเพลิดเพลินและการสร้างนิสัย
4. เพื่อศึกษาความสัมพันธ์ระหว่างปัจจัยด้านการใช้อย่างต่อเนื่อง กับการพัฒนาทักษะภาษาจีน

## 2. ทฤษฎีและกรอบแนวคิด

ทฤษฎี UTAUT2 (Unified Theory of Acceptance and Use of Technology 2) พัฒนาโดย Venkatesh, Thong และ Xu (2012) [1] เป็นทฤษฎีที่ใช้อธิบายการยอมรับเทคโนโลยีในบริบทผู้บริโภค โดยสังเคราะห์จากทฤษฎีการยอมรับเทคโนโลยี 8 ทฤษฎี ประกอบด้วยตัวแปรหลัก 7 ประการที่ส่งผลต่อความตั้งใจใช้และพฤติกรรมการใช้ ได้แก่

1. ประสิทธิภาพที่คาดหวัง (Performance Expectancy) – ระดับที่ผู้ใช้เชื่อว่าการใช้เทคโนโลยีจะช่วยเพิ่มประสิทธิภาพในการทำงาน
2. ความง่ายในการใช้ (Effort Expectancy) – ระดับความง่ายในการใช้เทคโนโลยี
3. อิทธิพลทางสังคม (Social Influence) – ระดับที่ผู้ใช้รับรู้ว่าคนสำคัญคิดว่าควรใช้เทคโนโลยี
4. สภาพแวดล้อมที่เอื้ออำนวย (Facilitating Conditions) – ระดับที่ผู้ใช้เชื่อว่ามีโครงสร้างพื้นฐานสนับสนุน

5. แรงจูงใจด้านความเพลิดเพลิน (Hedonic Motivation) – ความสนุกสนานที่ได้รับจากการใช้เทคโนโลยี

6. มูลค่าด้านราคา (Price Value) – การรับรู้ความคุ้มค่าทางการเงิน

7. นิสัย (Habit) – ระดับที่การใช้เทคโนโลยีกลายเป็นพฤติกรรมอัตโนมัติ

**ทฤษฎีความไว้วางใจในเทคโนโลยี**

McKnight และคณะ (2011) [2] ได้เสนอว่าความไว้วางใจในเทคโนโลยีประกอบด้วยองค์ประกอบสำคัญ 3 ประการ ได้แก่ ความเชื่อในความสามารถ (Competence) ที่ระบบมีทักษะและความเชี่ยวชาญ ความเชื่อในความซื่อสัตย์ (Integrity) ที่ระบบดำเนินการตามที่ควรจะเป็น และความเชื่อในความปรารถนาดี (Benevolence) ที่ระบบออกแบบมาเพื่อช่วยเหลือผู้ใช้

**แนวคิดการปรับเนื้อหาเฉพาะบุคคล**

การปรับเนื้อหาเฉพาะบุคคล (Personalization) หมายถึง กระบวนการปรับแต่งประสบการณ์การเรียนรู้ให้สอดคล้องกับความต้องการ ความสนใจ และความสามารถของผู้เรียนแต่ละคน ซึ่งเป็นคุณลักษณะเด่นของระบบ AI ประกอบด้วยรูปแบบต่างๆ เช่น การปรับตามระดับความรู้ การปรับตามความสนใจ การปรับตามสไตล์การเรียนรู้ การปรับตามจุดอ่อน และการติดตามความก้าวหน้า

**กรอบแนวคิดการวิจัย**

จากการทบทวนวรรณกรรม ผู้วิจัยได้พัฒนากรอบแนวคิดการวิจัย โดยประยุกต์ใช้ทฤษฎี UTAUT2 เป็นพื้นฐาน ร่วมกับปัจจัยด้านความไว้วางใจและการปรับเนื้อหาเฉพาะบุคคล รวมถึงการแทนที่ Use Behavior ด้วย Continued Use และการเพิ่ม Skill Development เป็นตัวแปรตามสุดท้าย โดยสรุปความสัมพันธ์ระหว่างตัวแปรได้ดังนี้

1. ปัจจัยด้านประสิทธิภาพที่คาดหวังและปัจจัยด้านความง่ายในการใช้ส่งผลต่อความไว้วางใจ ความตั้งใจใช้ และการใช้อย่างต่อเนื่อง

2. ปัจจัยด้านความไว้วางใจส่งผลต่อความตั้งใจใช้

3. ปัจจัยด้านความตั้งใจใช้ส่งผลต่อการใช้อย่างต่อเนื่อง

4. ปัจจัยด้านการปรับเนื้อหาเฉพาะบุคคลส่งผลต่อแรงจูงใจด้านความเพลิดเพลินและนิสัย

5. ปัจจัยด้านแรงจูงใจด้านความเพลิดเพลินและนิสัยส่งผลต่อการใช้อย่างต่อเนื่อง

6. ปัจจัยด้านการใช้อย่างต่อเนื่องส่งผลต่อการพัฒนาทักษะภาษาจีน

## 3. วิธีดำเนินการวิจัย

**รูปแบบการวิจัย**

การวิจัยครั้งนี้เป็นการวิจัยเชิงปริมาณแบบสำรวจภาคตัดขวาง (Quantitative Cross-sectional Survey Research) โดยใช้แบบสอบถามออนไลน์เก็บข้อมูล ณ จุดเวลาใดเวลาหนึ่ง เพื่อวิเคราะห์ปัจจัยที่มีอิทธิพลต่อการใช้ AI ในการเรียนภาษาจีน

**ประชากรและกลุ่มตัวอย่าง**

ประชากร คือ ผู้เรียนภาษาจีนชาวไทยที่มีประสบการณ์ใช้ AI กลุ่มตัวอย่าง จำนวน 755 คน ได้มาจากการสุ่มแบบเจาะจง (Purposive Sampling) ตามเกณฑ์: (1) กำลังศึกษาหรือเคยศึกษาภาษาจีน (2) มีความรู้พื้นฐานภาษาจีนตั้งแต่ระดับเริ่มต้นขึ้นไป (3) เคยใช้ AI ในการเรียนภาษา ขนาดกลุ่มตัวอย่างคำนวณตามหลักของ Hair et al. (2019) โดยใช้ 10-20 เท่าของตัวแปรสังเกตได้ (44 ตัวแปร) จึงกำหนดขั้นต่ำ 440 คน และเก็บได้จริง 755 คน ซึ่งมากกว่าเกณฑ์

**เครื่องมือที่ใช้ในการวิจัย**

เครื่องมือคือแบบสอบถามออนไลน์ Microsoft Form

**การตรวจสอบคุณภาพเครื่องมือ**

1. ความตรงเชิงเนื้อหา ผู้เชี่ยวชาญ 5 ท่านประเมินค่า IOC พบว่าทุกข้อมีค่า 0.60–1.00 (ผ่านเกณฑ์ $\geq$0.50)

2. ความเชื่อมั่น ทดสอบกับกลุ่มตัวอย่าง 30 คน ได้ค่า Cronbach's Alpha รวมทั้งฉบับเท่ากับ 0.956 (ดีเลิศ) โดยแต่ละด้านอยู่ระหว่าง 0.798–0.923

3. การวิเคราะห์องค์ประกอบเชิงสำรวจ (EFA) ได้ค่า KMO = 0.962 (>0.7) และ Bartlett's Test มีนัยสำคัญ

($p<0.001$) แสดงว่าข้อมูลเหมาะสม ค่าน้ำหนักองค์ประกอบทุกข้อ >0.5

4. การวิเคราะห์องค์ประกอบเชิงยืนยัน (CFA) ได้ค่า $\chi^2/df = 2.34$, CFI = 0.942, RMSEA = 0.042, SRMR = 0.038 แสดงว่าโมเดลการวัดสอดคล้องกับข้อมูล

5. ความเที่ยงตรงเชิงลู่เข้า (Convergent Validity) ค่า AVE อยู่ระหว่าง 0.512–0.549 (>0.5) และ CR อยู่ระหว่าง 0.765–0.885 (>0.7)

6. ความเที่ยงตรงเชิงจำแนก (Discriminant Validity) ค่ารากที่สองของ AVE ($\sqrt{AVE}$) มีค่ามากกว่าค่าสหสัมพันธ์ระหว่างตัวแปรทุกคู่

**การเก็บรวบรวมข้อมูล**

ดำเนินการเก็บข้อมูลผ่านช่องทาง Social Media และสถาบันการศึกษาฯ เป็นระยะเวลา 2 เดือน (พฤศจิกายน–ธันวาคม 2568) ได้แบบสอบถามสมบูรณ์ 755 ชุดแบบ Anonymous โดยมีการชี้แจงสิทธิ์และเก็บข้อมูลเป็นความลับ

**การวิเคราะห์ข้อมูล**

ใช้โปรแกรม Jamovi วิเคราะห์ข้อมูล ดังนี้

1. สถิติเชิงพรรณนา: ค่าความถี่ ร้อยละ ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน

2. สถิติเชิงอนุมาน: การวิเคราะห์สหสัมพันธ์เพียร์สัน (Pearson Correlation) และการวิเคราะห์การถดถอยพหุคูณ (Multiple Regression) เพื่อทดสอบสมมติฐาน

## 4. ผลการดำเนินการวิจัย

ผู้ตอบแบบสอบถามจำนวน 755 คน จำแนกตามลักษณะทั่วไปได้ดังนี้

**เพศ:** หญิง ร้อยละ 63.18 ชาย ร้อยละ 36.82

**อายุ:** 18-25 ปี ร้อยละ 70.33, 26-35 ปี ร้อยละ 24.64, 36-45 ปี ร้อยละ 3.84, มากกว่า 45 ปี ร้อยละ 1.19

**ระดับการศึกษา:** ปริญญาตรี ร้อยละ 91.52, ปริญญาโท ร้อยละ 6.89, ปริญญาเอก ร้อยละ 1.06, อื่นๆ ร้อยละ 0.53

**ระดับภาษาจีน:** ระดับกลาง ร้อยละ 65.30, ระดับต้น ร้อยละ 18.68, เริ่มต้น ร้อยละ 10.60, ระดับกลางสูง ร้อยละ 3.18, ระดับสูง ร้อยละ 2.25

**ประสบการณ์ใช้ AI:** ใช้เป็นประจำ ร้อยละ 72.85, ใช้บ้างเป็นครั้งคราว ร้อยละ 27.15

**ผู้ให้บริการ AI:** ChatGPT ร้อยละ 98.68, Gemini ร้อยละ 35.10, Copilot ร้อยละ 28.61, Deep Seek ร้อยละ 6.89, Meta AI ร้อยละ 1.99, อื่นๆ ร้อยละ 0.93

**รูปแบบการใช้งาน:** แบบฟรีทั้งหมด ร้อยละ 95.23, มีทั้งแบบฟรีและเสียเงิน ร้อยละ 4.64, แบบเสียเงินทั้งหมด ร้อยละ 0.13

**คุณภาพเครื่องมือวิจัย**

**การวิเคราะห์องค์ประกอบเชิงสำรวจ (EFA):** ค่า KMO = 0.962 (ดีเลิศ) และ Bartlett's Test มีนัยสำคัญ ($p < 0.001$) แสดงว่าข้อมูลเหมาะสม ค่าน้ำหนักองค์ประกอบทุกข้อ > 0.5

**การวิเคราะห์องค์ประกอบเชิงยืนยัน (CFA):** โมเดลการวัดมีความสอดคล้องกับข้อมูล โดย $\chi^2/df = 2.34$ (<3), CFI = 0.942 (>0.9), RMSEA = 0.042 (<0.05), SRMR = 0.038 (<0.05)

**ความเที่ยงตรงเชิงลู่เข้า:** ค่า AVE อยู่ระหว่าง 0.512–0.549 (>0.5) และค่า CR อยู่ระหว่าง 0.765–0.885 (>0.7)

**ความเที่ยงตรงเชิงจำแนก:** ค่ารากที่สองของ AVE ($\sqrt{AVE}$) มีค่ามากกว่าค่าสหสัมพันธ์ระหว่างตัวแปรทุกคู่

**ความเชื่อมั่น:** ค่า Cronbach's Alpha รวมทั้งฉบับ = 0.956 (ดีเลิศ)

**ระดับความคิดเห็นต่อปัจจัยต่างๆ**

แบบสอบถามมีระดับความคิดเห็นโดยรวมอยู่ในระดับมาก (ค่าเฉลี่ยรวม = 4.13) โดยด้านที่มีค่าเฉลี่ยสูงสุดคือ ความตั้งใจใช้และการใช้ต่อเนื่อง (4.28) รองลงมาคือ ประสิทธิภาพและความง่ายในการใช้ (4.27) ส่วนด้านที่มีค่าเฉลี่ยต่ำสุดคือ ความเพลิดเพลินและนิสัย (3.93)

**ความสัมพันธ์ระหว่างตัวแปร**

ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปรอยู่ระหว่าง 0.523–0.780 ($p < 0.01$) ไม่เกิน 0.85 แสดงว่าไม่มีปัญหา Multicollinearity ความสัมพันธ์สูงสุดคือระหว่างความตั้งใจใช้กับการใช้อย่างต่อเนื่อง ($r = 0.780$)

**ผลการทดสอบสมมติฐาน**

**สมมติฐาน H1 (ประสิทธิภาพและความง่ายในการใช้)**

1. ปัจจัยด้านประสิทธิภาพที่คาดหวังมีอิทธิพลต่อความไว้วางใจ ($\beta = 0.434$, $p < 0.01$)
2. ปัจจัยด้านความง่ายในการใช้มีอิทธิพลต่อความไว้วางใจ ($\beta = 0.340$, $p < 0.01$)
3. ทั้งสองปัจจัยร่วมอธิบายความไว้วางใจได้ร้อยละ 48.3
4. ปัจจัยด้านประสิทธิภาพที่คาดหวังมีอิทธิพลต่อความตั้งใจใช้ ($\beta = 0.322$, $p < 0.01$)
5. ปัจจัยด้านความไว้วางใจมีอิทธิพลต่อความตั้งใจใช้ ($\beta = 0.263$, $p < 0.01$)
6. ปัจจัยด้านความง่ายในการใช้มีอิทธิพลต่อความตั้งใจใช้ ($\beta = 0.231$, $p < 0.01$)
7. สามปัจจัยร่วมอธิบายความตั้งใจใช้ได้ร้อยละ 59.9
8. ปัจจัยด้านความตั้งใจใช้มีอิทธิพลต่อการใช้อย่างต่อเนื่องสูงสุด ($\beta = 0.580$, $p < 0.01$)
9. ปัจจัยด้านประสิทธิภาพที่คาดหวัง ($\beta = 0.159$, $p < 0.01$) และความง่ายในการใช้ ($\beta = 0.098$, $p < 0.05$) มีอิทธิพลรองลงมา
10. สามปัจจัยร่วมอธิบายการใช้อย่างต่อเนื่องได้ร้อยละ 61.5

**สมมติฐาน H2 (การปรับเนื้อหาเฉพาะบุคคล)**

1. ปัจจัยด้านการปรับเนื้อหาเฉพาะบุคคลมีอิทธิพลสูงมากต่อแรงจูงใจด้านความเพลิดเพลิน ($\beta = 0.689$, $p < 0.01$) และนิสัย ($\beta = 0.638$, $p < 0.01$)
2. อธิบายแรงจูงใจด้านความเพลิดเพลินได้ร้อยละ 47.5 และนิสัยได้ร้อยละ 40.7
3. ปัจจัยด้านนิสัยมีอิทธิพลต่อการใช้อย่างต่อเนื่อง ($\beta = 0.405$, $p < 0.01$) สูงกว่าแรงจูงใจด้านความเพลิดเพลิน ($\beta = 0.312$, $p < 0.01$)
4. สองปัจจัยร่วมอธิบายการใช้อย่างต่อเนื่องได้ร้อยละ 49.5
5. ปัจจัยด้านการใช้อย่างต่อเนื่องมีอิทธิพลสูงสุดต่อการพัฒนาทักษะ ($\beta = 0.542$, $p < 0.01$)
6. ปัจจัยด้านแรงจูงใจด้านความเพลิดเพลิน ($\beta = 0.177$, $p < 0.01$) และนิสัย ($\beta = 0.125$, $p < 0.01$) มีอิทธิพลรองลงมา
7. สามปัจจัยร่วมอธิบายการพัฒนาทักษะได้ร้อยละ 63.2

**สรุปผลการทดสอบสมมติฐาน**

ผลการทดสอบสมมติฐานทั้ง 15 ข้อ พบว่าทุกสมมติฐานมีนัยสำคัญทางสถิติ ($p < 0.05$ และ $p < 0.01$) ดังสรุปในตาราง

| เส้นทางความสัมพันธ์ | ค่าเบต้า ($\beta$) | ระดับนัยสำคัญ |
|---|---|---|
| ประสิทธิภาพที่คาดหวัง ⟶ ความไว้วางใจ | 0.434 | $p < 0.01$ |
| ความง่ายในการใช้ ⟶ ความไว้วางใจ | 0.34 | $p < 0.01$ |
| ความไว้วางใจ ⟶ ความตั้งใจใช้ | 0.263 | $p < 0.01$ |
| ประสิทธิภาพที่คาดหวัง ⟶ ความตั้งใจใช้ | 0.322 | $p < 0.01$ |
| ความง่ายในการใช้ ⟶ ความตั้งใจใช้ | 0.231 | $p < 0.01$ |
| ประสิทธิภาพที่คาดหวัง ⟶ การใช้อย่างต่อเนื่อง | 0.159 | $p < 0.01$ |
| ความง่ายในการใช้ ⟶ การใช้อย่างต่อเนื่อง | 0.098 | $p < 0.05$ |
| ความตั้งใจใช้ ⟶ การใช้อย่างต่อเนื่อง | 0.58 | $p < 0.01$ |
| การปรับเนื้อหาเฉพาะบุคคล ⟶ แรงจูงใจด้านความเพลิดเพลิน | 0.689 | $p < 0.01$ |
| การปรับเนื้อหาเฉพาะบุคคล ⟶ นิสัย | 0.638 | $p < 0.01$ |
| แรงจูงใจด้านความเพลิดเพลิน ⟶ การใช้อย่างต่อเนื่อง | 0.312 | $p < 0.01$ |
| นิสัย ⟶ การใช้อย่างต่อเนื่อง | 0.405 | $p < 0.01$ |
| แรงจูงใจด้านความเพลิดเพลิน ⟶ การพัฒนาทักษะ | 0.177 | $p < 0.01$ |
| นิสัย ⟶ การพัฒนาทักษะ | 0.125 | $p < 0.01$ |
| การใช้อย่างต่อเนื่อง ⟶ การพัฒนาทักษะ | 0.542 | $p < 0.01$ |

หมายเหตุ : เครื่องหมาย ⟶ แสดงถึง **"ส่งผลต่อ"** หรือ **"มีอิทธิพลต่อ"**

## 5. สรุปผลการวิจัยและข้อเสนอแนะ

**สรุปผลการวิจัยตามวัตถุประสงค์**

1. ตามวัตถุประสงค์ข้อที่ 1 ผลการวิจัยสรุปได้ว่าปัจจัยด้านประสิทธิภาพที่คาดหวังมีอิทธิพลต่อความไว้วางใจ ความตั้งใจใช้และการใช้อย่างต่อเนื่องสูงกว่าปัจจัยด้านความง่ายในการใช้ โดยปัจจัยด้านความตั้งใจใช้มีอิทธิพลสูงสุดต่อการใช้อย่างต่อเนื่อง ($\beta$ = 0.580)
2. ตามวัตถุประสงค์ข้อที่ 2 ผลการวิจัยสรุปได้ว่าปัจจัยด้านความไว้วางใจมีอิทธิพลต่อความตั้งใจใช้ ($\beta$ = 0.263) โดยมีบทบาทสำคัญเป็นลำดับที่สองรองจากประสิทธิภาพที่คาดหวัง
3. ตามวัตถุประสงค์ข้อที่ 3 ผลการวิจัยสรุปได้ว่าปัจจัยด้านการปรับเนื้อหาเฉพาะบุคคลมีอิทธิพลสูงมากต่อแรงจูงใจด้านความเพลิดเพลิน ($\beta$ = 0.689) และนิสัย ($\beta$ = 0.638) โดยปัจจัยด้านนิสัยมีอิทธิพลต่อการใช้อย่างต่อเนื่อง ($\beta$ = 0.405) สูงกว่าแรงจูงใจด้านความเพลิดเพลิน ($\beta$ = 0.312)
4. ตามวัตถุประสงค์ข้อที่ 4 ผลการวิจัยสรุปได้ว่าปัจจัยด้านการใช้อย่างต่อเนื่องมีอิทธิพลสูงสุดต่อการพัฒนาทักษะ ($\beta$ = 0.542) โดยร่วมกับแรงจูงใจด้านความเพลิดเพลินและนิสัยสามารถอธิบายความแปรปรวนของการพัฒนาทักษะได้ถึงร้อยละ 63.2

**อภิปรายผล**

ผลการวิจัยพบว่าประสิทธิภาพที่คาดหวังมีอิทธิพลต่อความไว้วางใจและความตั้งใจใช้สูงกว่าความง่ายในการใช้ สะท้อนว่าผู้เรียนให้ความสำคัญกับความถูกต้องของ AI มากกว่าความสะดวก สอดคล้องกับมาลินี วัฒนพนิต (2565) และ Chen และคณะ [2] ส่วนความไว้วางใจเป็นตัวแปรเชื่อมสำคัญ ($\beta$=0.263) รองจากประสิทธิภาพ แสดงว่าผู้เรียนต้องเชื่อมั่นในความถูกต้องและปลอดภัยของข้อมูลก่อนใช้งานจริง

การปรับเนื้อหาเฉพาะบุคคลเป็นหัวใจของ AI โดยมีอิทธิพลสูงต่อแรงจูงใจด้านความเพลิดเพลิน ($\beta$=0.689) และนิสัย ($\beta$=0.638) โดยเฉพาะการปรับตามจุดอ่อนและการติดตามความก้าวหน้าซึ่งผู้เรียนไทยให้ความสำคัญสูงสุด เพราะความสนุกเพียงอย่างเดียวไม่พอ ต้องสร้างวินัยหรือความเหนียวแน่น (Stickiness) สอดคล้องกับ Yang และคณะ [4] นอกจากนี้นิสัยมีอิทธิพลต่อการใช้ต่อเนื่อง ($\beta$=0.405) สูงกว่าความเพลิดเพลิน ($\beta$=0.312) แสดงว่าเมื่อเกิดนิสัยแล้วการใช้จะเกิดขึ้นโดยอัตโนมัติ (สอดคล้องกับสิริกร กาญจนวิรุฬห์, 2566)

สุดท้าย การใช้ต่อเนื่องมีอิทธิพลสูงสุดต่อพัฒนาทักษะ ($\beta$=0.542) สนับสนุนหลักการฝึกฝนสม่ำเสมอของ Ericsson และคณะ [3] นอกเหนือจากนี้ ผู้เรียนสามารถประยุกต์ใช้ AI ด้วยตนเอง โดยการปรับเนื้อหาเฉพาะบุคคลไม่ใช่แค่ฟังก์ชันในแอปฯ แต่คือการออกแบบเส้นทางการเรียนรู้ของตัวเอง โดยมี AI เป็นผู้ช่วยเติมเต็มจุดบกพร่องและเสริมกำลังใจ การพัฒนาในอนาคตควรมุ่งเน้นการปรับเนื้อหาที่ละเอียด ชาญฉลาด รักษาความเป็นส่วนตัวและจริยธรรม เพื่อให้ AI ตอบโจทย์การเรียนรู้ภาษาจีนอย่างยั่งยืน

## เอกสารอ้างอิงหลัก


[1] Venkatesh, V., Thong, J. Y., & Xu, X. (2012). Consumer acceptance and use of information technology: Extending the unified theory of acceptance and use of technology. *MIS Quarterly*, 36(1), 157-178.

[2] Chen, Y., et al. (2022). The role of trust in AI-assisted language learning: A cross-cultural study. *Computers in Human Behavior*, 128, 107-119.

[3] Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, 100(3), 363-406.

[4] Yang, J., et al. (2020). AI-powered Chinese language learning: Effectiveness and challenges. *Journal of Educational Technology*, 51(4), 623-638.

[5] มาลินี วัฒนพนิต. (2565). ความไว้วางใจในระบบ AI สำหรับการเรียนรู้ภาษา: กรณีศึกษาผู้เรียนชาวไทย. *วารสารศิลปศาสตร์*, 14(1), 112-128.

[6] สิริกร กาญจนวิรุฬห์. (2566). แรงจูงใจและความต่อเนื่องในการใช้แอปพลิเคชันเรียนภาษาจีนด้วย AI ของคนไทย. *วารสารภาษาและวัฒนธรรม*, 42(1), 67-82.

# การวิเคราะห์การยอมรับเทคโนโลยีการพบแพทย์ทางไกลด้วยเทคนิคเหมืองข้อมูลผ่านระบบธุรกิจอัจฉริยะ

# Analyzing the Acceptance of Telemedicine Using Data Mining with Business Intelligence System

*จิตรกร แซ่ตั้ง(Jittakron Saetung)[1] และมณเฑียร รัตนศิริวงศ์วุฒิ (Montean Rattanasiriwongwut)[2]*

*ภาควิชาการจัดการเทคโนโลยีสารสนเทศ*

*คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ*

*[1]s6707021857031@email.kmutnb.ac.th, [2] montean.r@itd.kmutnb.ac.th*

**บทคัดย่อ**

*งานวิจัยนี้มีวัตถุประสงค์เพื่อศึกษาปัจจัยที่มีผลต่อการยอมรับเทคโนโลยี การพบแพทย์ทางไกล (Telemedicine) ในเขตกรุงเทพมหานคร โดยประยุกต์ใช้เทคนิคเหมืองข้อมูล (Data Mining) เพื่อจัดกลุ่มผู้ใช้และเปรียบเทียบประสิทธิภาพของอัลกอริทึม 4 รูปแบบ ได้แก่ k-Means, K-Medoids, X-Means และ Fuzzy C-Means ข้อมูลถูกเก็บรวบรวมผ่านแบบสอบถามออนไลน์จากกลุ่มตัวอย่าง 636 คน ตามกรอบแนวคิด Technology Acceptance Model (TAM) ซึ่งประเมิน 3 ด้านได้แก่ (1) การรับรู้ประโยชน์ (2) การรับรู้ความง่ายในการใช้ (3) ทัศนคติและความตั้งใจใช้ และวัดประสิทธิภาพการจัดกลุ่มด้วย Average Within Cluster Distance (AWCD) กับ Davies-Bouldin Index (DBI) ผลการศึกษาเมื่อวิเคราะห์ข้อมูลพบว่าประชากรส่วนใหญ่มีทัศนคติต่อการใช้งานระบบการพบแพทย์ทางไกลที่ดีมาก มีค่าเฉลี่ย 3.58 (S.D. = 0.64) และได้ทำการเปรียบเทียบอัลกอริทึมในการจัดกลุ่ม และเมื่อทำการหาจำนวนกลุ่มด้วยเทคนิค Elbow method ได้จำนวนกลุ่มที่ 3 กลุ่ม มีค่า AWCD เท่ากับ 0.49 และเมื่อทำการเปรียบเทียบอัลกอริทึมทั้ง 4 อัลกอริทึมที่เหมาะสมที่สุดและให้ค่าประสิทธิภาพสูงที่สุดคือ k-Means โดยมีค่า AWCD เท่ากับ 0.79 และ Davies Bouldin เท่ากับ 0.12 Execution Time (sec) เท่ากับ 0.01 งานวิจัยนี้ยังได้นำเสนอข้อมูลเชิงลึกผ่าน Dashboard บน Power BI เพื่อช่วยในการตัดสินใจ*

**คำสำคัญ:** การพบแพทย์ทางไกล อัลกอริทึมการจัดกลุ่ม เหมืองข้อมูล

**Abstract**

*This research aims to study the factors affecting the acceptance of Telemedicine technology in Bangkok. It applies Data Mining techniques to cluster users and compares the performance of four algorithms: k-Means, K-Medoids, X-Means, and Fuzzy C-Means. Data was collected via online questionnaires from 636 participants based on the Technology Acceptance Model (TAM), which evaluates three main areas: (1) Perceived Usefulness, (2) Perceived Ease of Use, and (3) Attitude and Intention to Use. The clustering performance was measured using Average Within Cluster Distance (AWCD) and the Davies-Bouldin Index (DBI). The results showed that most people have a very positive attitude toward using Telemedicine systems, with an average score of 3.58 (S.D. = 0.64). Using the Elbow Method to determine the optimal number of clusters, the study identified 3 groups with an AWCD of 0.49. When comparing four algorithms, k-Means was found to be the most suitable and efficient model, with an AWCD of 0.79, a Davies-Bouldin Index of 0.12, and an Execution Time of 0.01 seconds. Additionally, this research presents insights through a Power BI Dashboard to support decision.*

**Keywords**: Telemedicine, Clustering Algorithms Data Mining.

## 1. บทนำ

บริการพบแพทย์ทางไกล (Telemedicine) มีบทบาทสำคัญอย่างมากในช่วงการแพร่ระบาดของโควิด-19 โดยในเขตกรุงเทพมหานคร มีอัตราการเติบโตของผู้ใช้บริการสูงถึง 201.26% ในปี 2022 อย่างไรก็ตามในปี 2024 อัตราการเติบโตกลับชะลอตัวลงเหลือเพียง 3.91% ปรากฏการณ์นี้สะท้อนให้เห็นถึงความจำเป็นในการศึกษาปัจจัยที่ส่งผลต่อการยอมรับและความตั้งใจใช้งานในระยะยาว

ผู้วิจัยจึงมุ่งเน้นการวิเคราะห์พฤติกรรมผู้ใช้และเปรียบเทียบอัลกอริทึมการจัดกลุ่มที่เหมาะสมที่สุดสำหรับการจัดการกลุ่มข้อมูล

**1.1 ขอบเขตด้านซอฟต์แวร์ ผู้วิจัยได้ดำเนินงานวิจัยด้วยซอฟต์แวร์ดังนี้**

1.1.1 Google Form สำหรับการทำแบบสอบถามออนไลน์

1.1.2 AI Studio 2025.0.0 (Rapid Miner) ประมวลผลการจัดกลุ่ม และเปรียบเทียบผลอัลกอริทึมของการจัดกลุ่ม

1.1.3 Power BI Desktop 2.153 ใช้ในการออกแบบรายงาน

1.1.4 IBM SPSS Statistics 26 ใช้วิเคราะห์ข้อมูลทางสถิติ

## 2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ผู้วิจัยได้ศึกษาค้นคว้าเอกสาร และงานวิจัยที่เกี่ยวข้องต่าง ๆ ที่มีความจำเป็นเพื่อเป็นแนวทางในการดำเนินการวิจัยเพื่อหาถึงพฤติกรรมของกลุ่มประชากรในการยอมรับเทคโนโลยีการพบแพทย์ทางไกล

**2.1 เทคนิคเหมืองข้อมูล**

เทคนิคเหมืองข้อมูลคือกระบวนการค้นหาโครงสร้างและรูปแบบความสัมพันธ์ รวมถึงการหาความรู้ที่มีประโยชน์ของชุดข้อมูลขนาดใหญ่และซับซ้อน

เทคนิคการทำเหมืองข้อมูล (Data Mining) เป็นการอาศัยหลักของการเรียนรู้ของเครื่อง (Machine Learning) เข้ามาช่วยซึ่งเทคนิคการทำเหมืองข้อมูลจะมีเทคนิคหลักๆ ที่นิยมใช้กันในปัจจุบันโดยจะแบ่งเป็น 2 ประเภทใหญ่ๆ คือ 1) การเรียนรู้โดยมีผู้สอน (Supervised learning) และมีเทคนิคดังนี้ การจำแนกประเภท (Classification), และการถดถอยเชิงเส้น (Regression) 2) การเรียนรู้โดยไม่มีผู้สอน (Unsupervised learning) มีเทคนิคดังนี้ การจัดกลุ่ม (Clustering), การวิเคราะห์ความสัมพันธ์ (Association Rules) [4]

**2.2 เทคนิคการจัดกลุ่มของข้อมูล**

การจัดกลุ่ม (Clustering) เป็นหนึ่งในเทคนิคการทำเหมืองข้อมูล เพื่อค้นหาความรู้ที่ซ่อนอยู่ในข้อมูล โดยจะมีเป้าหมายคือการจัดกลุ่ม เพื่อให้สามารถอธิบายรูปแบบของข้อมูล การจัดกลุ่มนี้เป็นหนึ่งในประเภทของการเรียนรู้ แบบไม่มีผู้สอน (Unsupervised learning) ซึ่งลักษณะของข้อมูลจะไม่มีค่าเป้าหมายหรือประเภทของกลุ่มที่กำหนดไว้ก่อน แต่จะแบ่งกลุ่มตามค่าของตัวแปรที่นำมาใช้ในการจัดกลุ่ม [3], [4], [5]

**2.3 การวัดประสิทธิภาพของการจัดกลุ่ม**

การวัดประสิทธิภาพของการจัดกลุ่มนั้นจะขึ้นอยู่กับประเภทของเทคนิคในการจัดกลุ่ม ซึ่งงานวิจัยนี้ได้ใช้เทคนิคการจัดกลุ่มแบบอิงความหนาแน่น จึงใช้การวัดประสิทธิภาพคือการหาค่าความห่างเฉลี่ยภายในกลุ่ม Avg. within centroid distance และ ค่าการคำนวณสัดส่วนระหว่าง ความกว้างของกลุ่ม เทียบกับ ระยะห่างระหว่างจุดศูนย์กลางของกลุ่ม โดยจะเปรียบเทียบไปทีละคู่ แล้วเลือกคู่ที่แย่ที่สุด (มีค่าสูงที่สุด) มาเป็นตัวแทนของกลุ่มนั้นก่อนจะหาค่าเฉลี่ยรวมทั้งโมเดล [3]

**2.4 เทคโนโลยีการแพทย์ทางไกล**

การใช้เทคโนโลยีสารสนเทศและการสื่อสารต่างๆ ในการรักษาทางด้านการแพทย์ โดยให้คำปรึกษาระหว่างแพทย์และการติดตามอาการของผู้ป่วยให้เป็นไปอย่างต่อเนื่องและมีประสิทธิภาพ สะดวก รวดเร็ว และช่วยให้ลดระยะเวลาในการรอคอยจากที่ต้องไปสถานพยาบาล และแม้ผู้ใช้บริการกับแพทย์จะอยู่ห่างไกลกัน ก็ยังสามารถที่จะสนทนากันแบบเห็นหน้าผ่านเครื่องมือสื่อสารและอินเตอร์เน็ตได้ [1], [4]

**2.5 งานวิจัยที่เกี่ยวข้อง**

ผู้วิจัยทำการศึกษาค้นคว้า บทความวิจัยที่มีความเกี่ยวข้องกับอัลกอริทึมในการทำเหมืองข้อมูลด้วยเทคนิคการจัดกลุ่มต่างๆ ทั้งในและต่างประเทศ เช่นงานของ สุชาวดี (2562) [5] การจัดกลุ่มขนาดผลิตภัณฑ์โดยการใช้

k-Means clustering และงานของ Mohammad (2558) [14] เปรียบเทียบประสิทธิภาพของอัลกอริทึมการจัดกลุ่มข้อมูลด้วย k-Means และ Fuzzy C-Means พบว่าการจัดกลุ่มแบบแบ่งส่วนนั้นมีความเหมาะสมกับชุดข้อมูลเชิงปริมาณแบบ Rating Scale เป็นการบังคับให้ผู้ตอบอยู่ได้ 1 กลุ่ม

## 3. วิธีการดำเนินการวิจัย

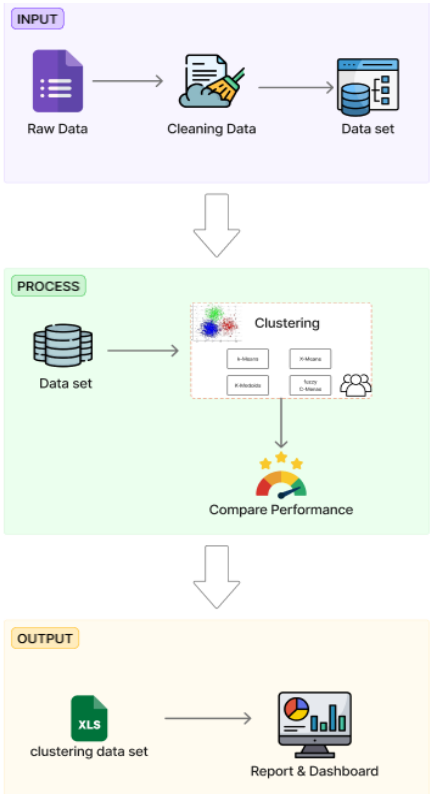
การวิจัยครั้งนี้ มีวัตถุประสงค์เพื่อวิเคราะห์พฤติกรรมสำหรับการตัดสินใจการยอมรับเทคโนโลยีการพบแพทย์ทางไกล (Telemedicine) โดยดำเนินการตามกรอบแนวคิดของงานวิจัยดังนี้

### 3.1 ศึกษา และวิเคราะห์ความต้องการ

ผู้วิจัยได้ศึกษา เพื่อออกแบบคำถามสำหรับการทำแบบสอบถามในการเก็บรวบรวมข้อมูล

### 3.2 กรอบแนวคิดของงานวิจัย

กรอบแนวคิดของงานวิจัยเป็นการทำงานร่วมกับแบบจำลองการยอมรับเทคโนโลยี (TAM) [1] และเทคนิคเหมืองข้อมูล (Data Mining) เพื่อวิเคราะห์พฤติกรรมผู้ใช้ จะมีกระบวนการทำงานแบ่งออกเป็น 3 ส่วนดังภาพที่ 1




**ภาพที่ 1** : กรอบแนวทางการดำเนินงานวิจัย

การศึกษานี้ใช้รูปแบบการวิจัยเชิงสำรวจ (Survey Research) โดยใช้แบบสอบถาม (Questionnaire) เป็นเครื่องมือหลักในการเก็บรวบรวมข้อมูลจากกลุ่มตัวอย่างเพื่อนำมาวิเคราะห์ ในขั้นตอนนี้จะทำให้ได้ Raw Data และไปดำเนินการในแต่ละส่วนดังนี้

ส่วนที่ 1 การนำข้อมูลเข้าจากการเก็บข้อมูลด้วยแบบสอบถาม และเป็นขั้นตอนการเตรียมข้อมูลและทำความสะอาดของข้อมูลก่อนนำไปวิเคราะห์

ส่วนที่ 2 กระบวนการวิเคราะห์ข้อมูล เป็นขั้นตอนในการนำข้อมูลที่เตรียมไว้มาเข้าสู่กระบวนการพัฒนาแบบจำลองด้วยเทคนิคเหมืองข้อมูล และวัดผลประสิทธิภาพ ทำให้ได้อัลกอริทึมที่เหมาะสมกับชุดข้อมูลนี้ที่สุด จากนั้นจึงนำข้อมูลที่มีการจัดกลุ่มออกเป็นไฟล์

ส่วนที่ 3 การแสดงผล เป็นขั้นตอนการนำผลลัพธ์ที่ได้มาสร้างรายงาน ผ่านระบบธุรกิจอัจฉริยะ

### 3.3 ขั้นตอนการดำเนินงานด้านข้อมูล

ทำการออกแบบแบบสอบถามและทำ IOC รวมถึงการหาค่าความเชื่อมั่น จากนั้นจึงเก็บรวบรวมข้อมูลโดยเน้นในด้านของการยอมรับการใช้เทคโนโลยีตั้งแต่เดือนตุลาคม ปี 2568 จนถึงเดือนพฤศจิกายน ปี 2568 แบ่งข้อมูลออกเป็น 2 ส่วนดังนี้ 1) ข้อมูลส่วนบุคคล 2) ข้อมูลส่วนคำถามชี้วัดการยอมรับเทคโนโลยี

### 3.4 ปัจจัยการยอมรับที่มีผลต่อการยอมรับเทคโนโลยีการพบแพทย์ทางไกล

ปัจจัยที่มีผลต่อพฤติกรรมการยอมรับเทคโนโลยีการพบแพทย์ทางไกลจากการศึกษาพบว่า ในภาพรวมของกลุ่มตัวอย่างอยู่ในระดับ มาก และค่าเฉลี่ยอยู่ที่ 3.58 (S.D. = 0.58) ซึ่งเมื่อพิจารณาตามปัจจัยแล้วพบว่า ด้านการรับรู้ประโยชน์ของการใช้เทคโนโลยีนั้นมีค่าอยู่ในระดับที่สูงที่สุด โดยมีค่าเฉลี่ยอยู่ที่ 3.66 (S.D. = 0.79) และมีทัศนคติต่อการใช้งานเทคโนโลยีการพบแพทย์ทางไกลอยู่ในระดับมาก และมีค่าเฉลี่ยอยู่ที่ 3.58 (S.D. = 0.64) และในส่วนที่

ค่าน้อยที่สุดคือด้านความปลอดภัยของระบบ โดยมีค่าเฉลี่ยอยู่ที่ 3.53 (S.D. = 0.65) ดังตารางที่ 1

**ตารางที่ 1 :** ค่าเฉลี่ยภาพรวม

| ตัวแปร | Mean | S.D. | แปลผล |
|---|---|---|---|
| ด้านการรับรู้ประโยชน์ | 3.66 | 0.79 | มาก |
| ด้านการรับรู้ความยากง่าย | 3.57 | 0.68 | มาก |
| ทัศนคติต่อการใช้งาน | 3.58 | 0.64 | มาก |
| ความปลอดภัยของระบบ | 3.53 | 0.65 | มาก |
| ความตั้งใจใช้ | 3.53 | 0.72 | มาก |

**3.5 การพัฒนาแบบจำลอง**

3.5.1 การเตรียมข้อมูล

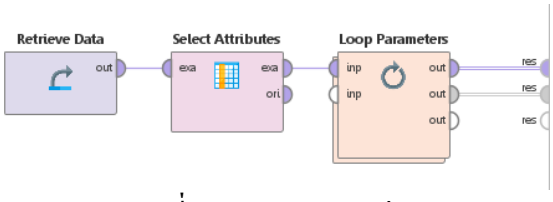
ทำการตรวจสอบความถูกต้องของข้อมูลไม่มีข้อมูลที่เป็นค่าว่าง (Missing) และเตรียมข้อมูลใหม่เพื่อให้รองรับการจัดกลุ่ม โดยได้ทำการปรับข้อมูลในส่วนของ ข้อมูลบุคคลเป็นแบบ Nominal Scale ดังภาพที่ 2

| Name | Type | Range |
|---|---|---|
| Gender | # integer | =[0 – 2] |
| Age | # integer | =[1 – 4] |
| MaritalStatus | # integer | =[1 – 3] |
| Job | # integer | =[1 – 5] |
| Income | # integer | =[1 – 5] |
| Distance | # integer | =[1 – 4] |
| Q1 | # integer | =[1 – 5] |
| Q2 | # integer | =[1 – 5] |
| Q3 | # integer | =[1 – 5] |
| Q4 | # integer | =[1 – 5] |
| Q5 | # integer | =[1 – 5] |
| Q6 | # integer | =[1 – 5] |
| Q7 | # integer | =[1 – 5] |
| Q8 | # integer | =[1 – 5] |
| Q9 | # integer | =[1 – 5] |
| Q10 | # integer | =[1 – 5] |
| Q11 | # integer | =[1 – 5] |
| Q12 | # integer | =[1 – 5] |
| Q13 | # integer | =[1 – 5] |
| Q14 | # integer | =[1 – 5] |
| Q15 | # integer | =[1 – 5] |
| Q16 | # integer | =[1 – 5] |
| Q17 | # integer | =[1 – 5] |
| Q18 | # integer | =[1 – 5] |
| Q19 | # integer | =[1 – 5] |
| Q20 | # integer | =[1 – 5] |
| Q21 | # integer | =[1 – 5] |
| Q22 | # integer | =[1 – 5] |
| Q23 | # integer | =[1 – 5] |
| Q24 | # integer | =[1 – 5] |

**ภาพที่ 2 :** ข้อมูลที่ใช้ในการจัดกลุ่ม

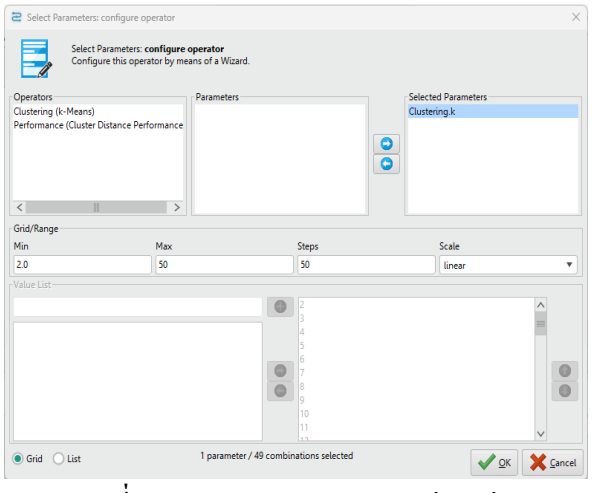
3.5.2 การหากลุ่ม (k)

เป็นกระบวนการนำเข้าข้อมูลด้วยการ Import Data โดยใช้ Excel File เพื่อหากลุ่มของประชากรที่ตอบแบบสอบถามด้วยอัลกอริทึม ดังภาพที่ 3




**ภาพที่ 3 :** การหาจำนวนกลุ่ม (k)

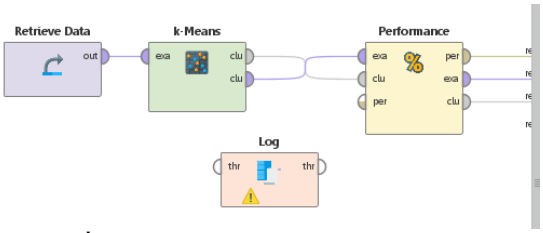
กระบวนการภายใน Operator Loop Parameters เพื่อวัดประสิทธิภาพในการหากลุ่มของประชากร parameters ของ Operator Loop Parameters เพื่อต้องการให้วนลูปในการกำหนดค่า k และแสดงผลลัพธ์ที่เหมาะสมที่สุด โดยกำหนดค่าตั้งแต่ 2-50 ดังภาพที่ 4




**ภาพที่ 4 :** การวัดประสิทธิภาพของแต่ละกลุ่ม

3.5.3 การใช้เทคนิคการจัดกลุ่ม

ในขั้นตอนนี้ผู้วิจัยได้ทำการเปรียบเทียบอัลกอริทึมการจัดกลุ่มทั้งหมด 4 อัลกอริทึมคือ k-Means, X-Means, K-Medoids และ Fuzzy C-Means




**ภาพที่ 5 :** การจัดกลุ่มด้วยวิธี k-Means

จากนั้นได้มีการเปรียบเทียบค่าต่างๆ ตามผลของค่าการวัดประสิทธิภาพด้วยค่าเฉลี่ยระยะทางจากจุดกึ่งกลาง

## 4. ผลการดำเนินงาน

### 4.1 การทดสอบประสิทธิภาพของการจัดกลุ่ม

4.1.1 ผลการหาค่า k Parameter ที่เหมาะสม

ได้ค่า k เท่ากับ 3 ดังภาพที่ 10 ด้วยวิธีการหาจุดหักศอก Elbow [5] มีค่าเฉลี่ย เท่ากับ 0.49 ตามสมการที่ 4-1 [5]

$$WCSS = \sum_{C_k}^{C_n} ( \sum_{d_i in C_i}^{d_m} distance(d_i, C_k)^2) \quad (4\text{-}1)$$

| iteration ↑ | Clustering.k | Avg. within centroid distance |
|---|---|---|
| 1 | 2 | 0.625 |
| 2 | 3 | 0.494 |
| 3 | 4 | 0.435 |
| 4 | 5 | 0.381 |
| 5 | 6 | 0.334 |
| 6 | 7 | 0.299 |

**ภาพที่ 6 :** ผลลัพธ์ของการหาจำนวนกลุ่ม (k)

4.1.2 ผลการเปรียบเทียบประสิทธิภาพการจัดกลุ่ม

เป็นการวัดผลของแต่ละอัลกอริทึม เพื่อหาอัลกอริทึมที่เหมาะสมที่สุดด้วยการ หาค่าเฉลี่ยระยะห่างของทุกจุดศูนย์กลางกลุ่ม (AWCD) ดังสมการที่ **4-2** [5]

$$\overline{dC} = \frac{1}{n} \sum_{x \in C} dist(x, \mu) \quad (4\text{-}2)$$

และการหาค่าเฉลี่ยความคล้ายกันระหว่างกลุ่ม (DBI) ดังสมการที่ **4-3** [4] และผลดังตารางที่ 2

$$DB = \frac{1}{k} \sum_{i=1}^{k} \max_{j \neq i} \left( \frac{S_i + S_j}{M_{ij}} \right) \quad (4\text{-}3)$$

**ตารางที่ 2 :** การเปรียบเทียบประสิทธิภาพ

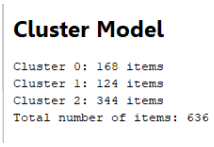
| Algorithm | AWCD | DBI | Time |
|---|---|---|---|
| k-Means | 0.79 | 0.12 | 0.01 |
| K-Medoids | 2.29 | 0.05 | 13.92 |
| X-Means | 0.81 | 0.12 | 0.03 |
| Fuzzy C-Means | 1.24 | 0.89 | 0.23 |

จากตารางจะเห็นว่าอัลกอริทึม K-Medoids นั้นใช้เวลาในการประมวลผลนานที่สุดเนื่องจากต้องคำนวนระยะห่างทุกคู่เพื่อหา Medoids ที่ดีที่สุด แต่ k-Means นั้นมีประสิทธิภาพสูงสุด สูตรการคำนวณ k-Means Clustering **ดังสมการที่** 4-4 [6]

$$d(x, y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2} \quad (4\text{-}4)$$

### 4.2 รายละเอียดพฤติกรรมของประชากรในแต่ละกลุ่ม

ผลการจัดกลุ่มได้จำนวนกลุ่มเท่ากับ 3 ดังภาพที่ 7




**ภาพที่ 7 :** โมเดลการจัดกลุ่ม

และสามารถอธิบายพฤติกรรมของแต่ละกลุ่มได้ดังนี้

กลุ่มที่ 1 เป็นกลุ่มผู้นำเทคโนโลยี มีจำนวน 344 คน มองว่าเทคโนโลยีเป็นสิ่งที่จำเป็นในชีวิตประจำวัน และมีแผนที่จะใช้บริการในอนาคต

กลุ่มที่ 2 เป็นกลุ่มที่มีศักยภาพแต่ยังลังเล มีจำนวน 168 คน มองว่าการบริการมีประโยชน์แต่ยังลังเลว่า การไปพบแพทย์ที่สถานพยาบาลแบบเดิมจะดีกว่า

กลุ่มที่ 3 เป็นกลุ่มที่ยึดติดรูปแบบเดิม มีจำนวน 124 คน กลุ่มนี้คิดว่าการบริการรูปแบบนี้ไม่มีความจำเป็น และไม่เชื่อมั่น

### 4.3 การแสดงผลรายงานบนโปรแกรม Power BI

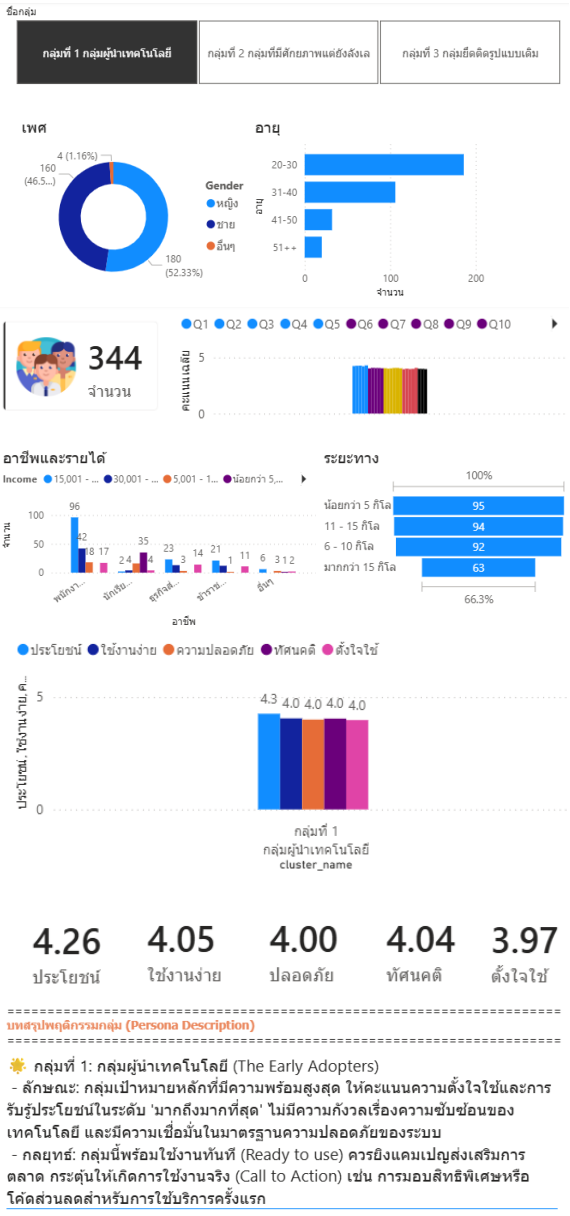
ได้จัดทำรายงานผ่านระบบธุรกิจอัจฉริยะในรูปแบบ Dashboard Visualization เพื่อการวิเคราะห์เชิงลึก ดังนี้

4.3.1 รายงานภาพรวมผู้ตอบแบบสอบถาม เป็นรายงานการสรุปข้อมูลจากกลุ่มตัวอย่างจำนวน 636 คน โดยแสดงถึงภาพรวมของประชากรกลุ่มตัวอย่างเคยใช้หรือไม่เคยใช้บริการนี้ด้วยสาเหตุอะไร

4.3.2 รายงานสรุปทัศนคติและการยอมรับเทคโนโลยี เป็นการสรุปผลจากการประเมินปัจจัยทั้ง 5 ด้านตามแนวคิดการยอมรับเทคโนโลยี พบว่าการรับรู้ประโยชน์ของบริการการพบแพทย์ทางไกลนั้นมีค่าเฉลี่ยที่สูงที่สุดอยู่ที่ 3.66 เปอร์เซ็นต์ แต่ยังมีจุดอ่อนคือด้านความปลอดภัยมีค่าเฉลี่ยอยู่ที่ 3.53 เปอร์เซ็นต์ ทำให้เห็นว่า

ประชากรยังมีความกังวลเรื่องความปลอดภัยของข้อมูลอยู่บ้าง

4.3.3 รายงานการวิเคราะห์กลุ่มประชากร เป็นรายงานการวิเคราะห์ประชากรโดยแบ่งออกเป็น 3 กลุ่ม ทำให้พบว่าเป้าหมายหลักที่ควรดำเนินกลยุทธ์เชิงรุกได้แก่กลุ่มไหน ดังภาพที่ 8




**ภาพที่ 8 :** รายงานการวิเคราะห์กลุ่มประชากร

จากรายงานดังภาพที่ 8 สามารถอธิบายลักษณะของแต่ละกลุ่มได้ดังนี้

**กลุ่มที่ 1** คือกลุ่มที่มีการยอมรับเทคโนโลยีในระดับสูง ซึ่งได้คะแนนสูงในทุกด้านมองว่าเทคโนโยโลยีช่วยเพิ่มประสิทธิภาพ ได้

**กลุ่มที่ 2** คือกลุ่มที่มีศักยภาพแต่ยังลังเล ในกลุ่มนี้จะมีความกังวลในความซับซ้อนของการใช้งานระบบ และความปลอดภัย

**กลุ่มที่ 3** คือกลุ่มยึดติดรูปแบบเดิม กลุ่มนี้ไม่คิดว่าบริการหรือเทคโนโลยีนี้จะช่วยเพิ่มประสิทธิภาพได้ และไม่เชื่อมั่นในความปลอดภัยของระบบ มองว่าการพบแพทย์ในรูปแบบเดิมนั้นดีกว่า

## 5. สรุป

### 5.1 สรุปผลการวิจัย

จากการศึกษาและวิเคราะห์ข้อมูลกลุ่มตัวอย่างจำนวน 636 คน ในเขตกรุงเทพมหานคร เกี่ยวกับการยอมรับเทคโนโลยีการพบแพทย์ทางไกล (Telemedicine) สามารถสรุปผลการวิจัยได้ดังนี้

5.1.1 ด้านพฤติกรรมและการยอมรับกลุ่มตัวอย่างส่วนใหญ่มีทัศนคติต่อการใช้งานระบบการพบแพทย์ทางไกลในระดับที่ดีมาก โดยมีค่าเฉลี่ยรวมอยู่ที่ 3.58 (S.D. = 0.64) ซึ่งประเมินผ่านกรอบแนวคิด TAM

5.1.2 ด้านการจัดกลุ่มด้วยเหมืองข้อมูล เมื่อหาจำนวนกลุ่มที่เหมาะสมที่สุดด้วยเทคนิค Elbow Method [5] พบว่าสามารถแบ่งกลุ่มพฤติกรรมผู้ใช้งานออกได้เป็น 3 โดยมีค่า AWCD รวมเท่ากับ 0.49

5.1.3 ด้านประสิทธิภาพของอัลกอริทึม จากการเปรียบเทียบอัลกอริทึม 4 รูปแบบ (k-Means, K-Medoids, X-Means และFuzzy C-Means) พบว่า k-Means เป็นอัลกอริทึมที่มีประสิทธิภาพเหมาะสมที่สุด โดยให้ค่า AWCD เท่ากับ 0.79 ค่า DBI เท่ากับ 0.12 และใช้ระยะเวลาในการประมวลผล (Execution Time) ต่ำที่สุดเพียง 0.01 วินาที

5.1.4 ด้านการนำเสนอข้อมูล ผลลัพธ์จากการจัดกลุ่มถูกนำไปพัฒนาผ่านระบบธุรกิจอัจฉริยะ (Business Intelligence) เพื่อแสดงผลข้อมูลเชิงลึกสำหรับการตัดสินใจเชิงกลยุทธ์

### 5.2 อภิปรายผล

5.2.1 อภิปรายผลด้านปัจจัยที่มีผลต่อพฤติกรรมการยอมรับเทคโนโลยีการพบแพทย์ทางไกล

จากผลการวิจัยพบว่าปัจจัยตามทฤษฎีการยอมรับเทคโนโลยี (TAM) มีผลเชิงบวกต่อการยอมรับ Telemedicine สอดคล้องกับค่าเฉลี่ยระดับสูง (3.58) การรับรู้ประโยชน์และความง่ายในการใช้งาน เป็นปัจจัยสำคัญที่ทำให้ประชากรในกรุงเทพฯ มีทัศนคติที่ดีต่อการพบแพทย์ทางไกล ทั้งนี้อาจเป็นเพราะสถานการณ์ปัจจุบันที่เทคโนโลยีดิจิทัลเข้าถึงได้ง่ายขึ้น รวมถึงความต้องการความสะดวกสบายในการเข้าถึงบริการทางการแพทย์โดยไม่ต้องเดินทางไปโรงพยาบาล ทำให้ผู้ใช้งานเกิดความตั้งใจที่จะใช้งานระบบอย่างต่อเนื่อง

5.2.2 อภิรายผลด้านการเปรียบเทียบประสิทธิภาพของอัลกอริทึมด้วยเทคนิคเหมืองข้อมูล

การเปรียบเทียบอัลกอริทึมพบว่า k-Means ทำผลงานได้ดีที่สุดในการจัดกลุ่มข้อมูลแบบสอบถามในครั้งนี้ เมื่อพิจารณาจากค่า DBI ที่ต่ำ (0.12) แสดงให้เห็นว่า k-Means สามารถแยกกลุ่มผู้ใช้งาน 3 กลุ่มออกจากกันได้อย่างชัดเจน และข้อมูลภายในกลุ่มเดียวกันมีความคล้ายคลึงกันสูง สาเหตุที่ k-Means เหมาะสมที่สุด เป็นเพราะชุดข้อมูลจากการตอบแบบสอบถาม (Likert Scale) เป็นข้อมูลเชิงตัวเลขที่มีโครงสร้างชัดเจน และไม่มีค่าผิดปกติ (Outliers) ที่รุนแรงมากนัก k-Means จึงสามารถหาจุดศูนย์กลาง (Centroid) ได้อย่างรวดเร็ว (0.01 วินาที) ในขณะที่อัลกอริทึมอื่น เช่น Fuzzy C-Means หรือ K-Medoids อาจใช้เวลาประมวลผลนานกว่าหรือมีความซับซ้อนเกินความจำเป็นสำหรับชุดข้อมูลลักษณะนี้ [2], [5]

5.2.3 อภิปรายผลด้านการนำเสนอพฤติกรรมเทคโนโลยีการพบแพทย์ทางไกลผ่านระบบธุรกิจอัจฉริยะ

การนำโมเดลที่ได้มาประยุกต์ใช้กับ Power BI Dashboard ถือเป็นการยกระดับการใช้ประโยชน์จากข้อมูล (Data Monetization) อย่างเป็นรูปธรรม ระบบธุรกิจอัจฉริยะนี้ช่วยเปลี่ยนข้อมูลดิบและผลลัพธ์จากการจัดกลุ่มที่ซับซ้อน ให้กลายเป็นภาพที่เข้าใจง่าย (Data Visualization) ทำให้ผู้บริหารระดับองค์กร ผู้ให้บริการทางการแพทย์ หรือนักการตลาด สามารถมองเห็นแนวโน้มพฤติกรรมผู้บริโภคได้อย่างรวดเร็ว เช่น สามารถทราบได้ทันทีว่ากลุ่มผู้ใช้ที่ยอมรับเทคโนโลยีสูง (Early Adopters) มีลักษณะทางประชากรศาสตร์อย่างไร เพื่อนำไปสู่การกำหนดนโยบายต่อไปได้อย่างแม่นยำ

**5.3 ข้อเสนอแนะ**

แม้ว่าอัลกอริทึมการจัดกลุ่ม (Clustering Algorithms) เช่น k-Means จะมีประสิทธิภาพสูงในการแบ่งกลุ่มประชากรตามความคล้ายคลึงของพฤติกรรม แต่พบว่ายังมีข้อจำกัดในมิติของการประเมินระดับคะแนนเชิงสัมบูรณ์ (Absolute Scoring) เนื่องจากการจัดกลุ่มด้วยเทคนิคเหล่านี้อาศัยการวัดระยะห่างระหว่างข้อมูลกับจุดกึ่งกลางของกลุ่ม (Centroid) เป็นหลัก ซึ่งอาจทำให้เกิดความคลาดเคลื่อนในการตีความหมายเชิงธุรกิจ

ดังนั้น ในการศึกษาครั้งต่อไป ควรพิจารณาประยุกต์ใช้การจัดอันดับ (Ranking) หรือ การแบ่งกลุ่มตามเกณฑ์คะแนน (Scoring/Tiering) ควบคู่ไปกับการทำเหมืองข้อมูล เพื่อให้สามารถจำแนกกลุ่มผู้ใช้งานที่มีศักยภาพสูงได้อย่างแม่นยำและตรงกับความเป็นจริงมากยิ่งขึ้น

## เอกสารอ้างอิง


[1] Nguyen (2021) Acceptance and Use of Video Conferencing for Teaching in Covid-19 Pandemic An Empirical Study in Vietnam

[2] Janusz (2022) Comparative evaluation of clustering algorithms for profiling of internet auction fraudsters

[3] Hui Yin (2024) A Rapid Review of Clustering Algorithms

[4] กรวัฒน์ (2566) ปัจจัยที่มีผลต่อการยอมรับนวัตกรรมบริการทางการแพทย์ผ่านระบบอินเตอร์เน็ตของบุคลากรทางการแพทย์ในโรงพยาบาลส่งเสริมสุขภาพตำบล

[5] สุชาวดี (2562) การจัดกลุ่มขนาดผลิตภัณฑ์เพื่อลดต้นทุนบรรจุภัณฑ์

[6] พัชร (2568) การศึกษาเทคนิคการจัดกลุ่มโดย Machine Learning ในหุ้น SET100 ของตลาดหลักทรัพย์แห่งประเทศไทย

# ปัจจัยที่มีอิทธิพลต่อการยอมรับหุ่นยนต์เสิร์ฟอาหารในร้านอาหาร
# Factors Influencing the Acceptance of Food-Serving Robots in Restaurants

*ราชภัฏ รักษ์สุจริต (Rajabhat Raksujarit)[1] และ ศักดิ์ชาย ตั้งวรรณวิทย์ (Sakchai Tangwannawit)[2]*
*ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล*
*มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ*
*[1]s6707021857236@email.kmutnb.ac.th, [2] sakchai_t@itd.kmutnb.ac.th*

**บทคัดย่อ**

*การวิจัยนี้มีวัตถุประสงค์เพื่อศึกษาปัจจัยที่มีอิทธิพลต่อการยอมรับการใช้งานหุ่นยนต์เสิร์ฟอาหารในร้านอาหาร โดยประยุกต์ใช้ทฤษฎีการยอมรับเทคโนโลยี โดยการวิจัยเป็นเชิงปริมาณ เก็บข้อมูลจากผู้ใช้บริการจำนวน 433 คน และวิเคราะห์ด้วยสถิติพรรณนา EFA และการถดถอยเชิงเส้นผลการวิจัยพบว่าแบบจำลองสามารถอธิบายความตั้งใจในการใช้งานได้ร้อยละ 70.7 และการยอมรับการใช้งานจริงได้ร้อยละ77.8อิทธิพลทางสังคมและความไว้วางใจมีผลเชิงบวกต่อความตั้งใจในการใช้งานอย่างมีนัยสำคัญขณะที่ความตั้งใจในการใช้งานมีอิทธิพลสูงสุดต่อการยอมรับการใช้งานจริงผลการวิจัยยืนยันความเหมาะสมของ การยอมรับเทคโนโลยีและสามารถนำไปใช้กำหนดกลยุทธ์ส่งเสริมการใช้หุ่นยนต์ในธุรกิจร้านอาหาร*

**คำสำคัญ:** การยอมรับเทคโนโลยี หุ่นยนต์เสิร์ฟอาหาร ปัจจัยที่มีอิทธิพล ทฤษฎี UTAUT

**Abstract**

*This study aims to examine the factors influencing the acceptance of food-serving robots in restaurants by applying the Unified Theory of Acceptance and Use of Technology The research employed a quantitative approach, collecting data from 433 restaurant customers, and analyzed using descriptive statistics, exploratory factor analysis (EFA), and linear regression. The results show that the model explains 70.7% of the variance in behavioral intention and 77.8% of actual usage acceptance. Social influence and trust have significant positive effects on behavioral intention, while behavioral intention has the strongest influence on actual usage acceptance. The findings confirm the suitability of the Acceptance and Use of Technology model and provide practical implications for developing strategies to promote the adoption of service robots in the restaurant industry.*

**Keywords:** Technology Acceptance, Food Serving Robots, Factors Influencing, UTAUT.

## 1. บทนำ

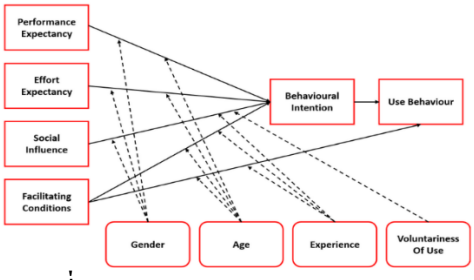
เทคโนโลยีหุ่นยนต์เสิร์ฟอาหารถูกนำมาใช้ในธุรกิจร้านอาหารมากขึ้น โดยช่วยเพิ่มความรวดเร็ว ลดข้อผิดพลาด และแก้ปัญหาการขาดแคลนแรงงาน รวมถึงสร้างภาพลักษณ์ที่ทันสมัยและตอบโจทย์ผู้บริโภคยุคดิจิทัล โดยเฉพาะในกรุงเทพมหานครที่มีการแข่งขันสูง อย่างไรก็ตาม การยอมรับเทคโนโลยีดังกล่าวยังขึ้นอยู่กับปัจจัยหลายด้าน งานวิจัยนี้จึงประยุกต์ใช้ทฤษฎี UTAUT

เพื่อวิเคราะห์ปัจจัยสำคัญ ได้แก่ ความคาดหวังด้านประสิทธิภาพ ความง่ายในการใช้งาน อิทธิพลทางสังคม ความพร้อมของโครงสร้างพื้นฐาน และความไว้วางใจ ที่ส่งผลต่อความตั้งใจและพฤติกรรมการใช้งานหุ่นยนต์เสิร์ฟอาหารของผู้บริโภค เพื่อสนับสนุนการนำเทคโนโลยีไปใช้ได้อย่างมีประสิทธิภาพในอนาคต

## 2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

### 2.1 ทฤษฎี UTAUT (Unified Theory of Acceptance and Use of Technology)

พัฒนาโดย Viswanath Venkatesh และคณะ (2003) [1] เป็นทฤษฎีที่ใช้อธิบายการยอมรับและการใช้งานเทคโนโลยี โดยมีปัจจัยหลัก 4 ด้าน ได้แก่ ความคาดหวังด้านประสิทธิภาพ (ช่วยให้ทำงานดีขึ้น) ความคาดหวังด้านความพยายาม (ใช้งานง่าย) อิทธิพลทางสังคม (แรงจูงใจจากคนรอบข้าง) และเงื่อนไขสนับสนุน (ความพร้อมของระบบและทรัพยากร) ซึ่งปัจจัยเหล่านี้ส่งผลต่อความตั้งใจใช้งานและพฤติกรรมการใช้งานจริงของผู้ใช้




**ภาพที่ 1** Unified Theory of Acceptance and Use of Technology (UTAUT)

### 2.2 ทฤษฎี TAM (Technology Acceptance Model)

เป็นทฤษฎีพื้นฐานที่ใช้ศึกษาการยอมรับเทคโนโลยี ซึ่งพัฒนาโดย Fred D. Davis (1989) [2] โดยเน้น 2 ปัจจัยสำคัญ คือ การรับรู้ประโยชน์ (Perceived Usefulness) ที่ทำให้ผู้ใช้เห็นว่าเทคโนโลยีช่วยเพิ่มประสิทธิภาพ และการรับรู้ความง่ายในการใช้งาน (Perceived Ease of Use) ซึ่งส่งผลต่อทัศนคติ ความพึงพอใจ และความตั้งใจในการใช้งานเทคโนโลยี

### 2.3 การวิเคราะห์องค์ประกอบเชิงสำรวจ

การวิเคราะห์ปัจจัยจาก Hair, J. F., et al. (2010) [3] Exploratory Factor Analysis (EFA) เป็นเทคนิคทางสถิติที่ใช้ค้นหาโครงสร้างแฝงของข้อมูล โดยจัดกลุ่มตัวแปรที่มีความสัมพันธ์กันให้อยู่ในองค์ประกอบเดียวกันโดยไม่ต้องกำหนดล่วงหน้ากระบวนการประกอบด้วยการตรวจสอบข้อมูล (KMO, Bartlett's Test) การสกัดองค์ประกอบ และการหมุนแกนเพื่อให้ตีความได้ชัดเจน โดยผลลัพธ์สำคัญคือค่า Factor Loading ที่แสดงความสัมพันธ์ระหว่างตัวแปรกับองค์ประกอบEFA มักใช้เพื่อสำรวจและลดจำนวนตัวแปร และนิยมใช้ร่วมกับ CFA เพื่อยืนยันโครงสร้างแบบสอบถามในขั้นต่อไป

### 2.4 การวิเคราะห์การถดถอยพหุคูณ

การวิเคราะห์การถดถอยพหุคูณ (Multiple Regression Analysis) อ้างอิงโดย Cohen, J., et al. (2003) [4] เป็นเทคนิคทาง สถิติ ที่ใช้ศึกษาความสัมพันธ์ระหว่างตัวแปรตาม 1 ตัว กับตัวแปรอิสระตั้งแต่ 2 ตัวขึ้นไป เพื่อดูว่าตัวแปรอิสระแต่ละตัวมีอิทธิพลต่อผลลัพธ์มากน้อยเพียงใด แนวคิดคือ การนำตัวแปรหลายตัวมาร่วมกันทำนายค่าของตัวแปรตาม เช่น รายได้อาจขึ้นอยู่กับการศึกษา อายุ และประสบการณ์ทำงานผลลัพธ์สำคัญคือค่าสัมประสิทธิ์ (Regression Coefficient) ซึ่งบอกทิศทางและขนาดของผลกระทบของแต่ละตัวแปร รวมถึงค่า $R^2$ ที่บอกความสามารถของโมเดลในการอธิบายข้อมูลวิธีนี้ช่วยในการพยากรณ์และวิเคราะห์ปัจจัยที่มีผลต่อสิ่งที่ต้องการ

### 2.5 งานวิจัยที่เกี่ยวข้อง

กฤตภัค วรเมธพาสุข (2562) [5] ศึกษาอิทธิพลของความคล้ายคลึงเชิงรูปลักษณ์และการรับรู้การควบคุมได้ที่มีผลต่อการรับรู้ความเหมือนมนุษย์ ความเชื่อใจ และการยอมรับหุ่นยนต์ในภาคบริการ ผลการวิจัยพบว่า ความคล้ายคลึงทางรูปลักษณ์ส่งผลเชิงบวกต่อการรับรู้ความเหมือนมนุษย์และความเชื่อใจ ขณะที่การรับรู้การควบคุมได้ช่วยเพิ่มความไว้วางใจและการยอมรับหุ่นยนต์อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 โดยตัวแปรดังกล่าวสามารถอธิบายความแปรปรวนของความไว้วางใจได้ ($R^2$ = .501) และการยอมรับหุ่นยนต์ได้ ($R^2$ = .127) โดยมีการรับรู้ความเหมือนมนุษย์เป็นตัวแปรส่งผ่านอย่างมีนัยสำคัญทางสถิติ

ณัชชารีย์ ฐิติสกุลวิทย์ (2565) [6] ศึกษาปัจจัยที่มีผลต่อความตั้งใจยอมรับหุ่นยนต์เพื่อนดูแลผู้สูงอายุ พบว่า

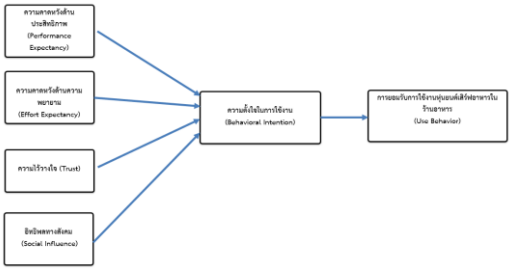
การรับรู้ความง่ายในการใช้ส่งผลต่อการรับรู้ประโยชน์ การควบคุม และความสนุกสนาน ส่วนลักษณะความเป็นมนุษย์ของหุ่นยนต์ส่งผลต่อความเชื่อมั่นและความสนุกสนาน โดยความตั้งใจยอมรับขึ้นอยู่กับความเชื่อมั่น การรับรู้ประโยชน์ ความสนุกสนาน และความกลัวเทคโนโลยี และใช้ค่า Cronbach's alpha $\geq$ 0.7 เป็นเกณฑ์ความเชื่อมั่นของแบบสอบถาม

บุญทวี จันทสุวรรณ (2566) [7] ศึกษาการยอมรับนวัตกรรมหุ่นยนต์บริการในโรงแรม 5 ดาว เขตกรุงเทพมหานคร โดยพบว่า คุณภาพบริการของหุ่นยนต์มีอิทธิพลโดยตรงต่อการยอมรับของผู้ใช้ โดยเฉพาะด้านความสะดวก ความรวดเร็ว และการตอบสนองความต้องการ ซึ่งส่งผลต่อความพึงพอใจและความไว้วางใจ ทั้งนี้ การประเมินอิงตามแนวคิด SERVQUAL 5 ด้าน ได้แก่ รูปลักษณ์ ความน่าเชื่อถือ การตอบสนอง ความมั่นใจ และความเข้าอกเข้าใจ ซึ่งสะท้อนถึงบทบาทสำคัญของเทคโนโลยีหุ่นยนต์ในอุตสาหกรรมโรงแรมยุคใหม่

## 3. วิธีการดำเนินการวิจัย

### 3.1 กรอบแนวคิดของการวิจัย

กรอบแนวคิดในการวิจัยครั้งนี้ถูกพัฒนาขึ้นจากแนวคิดและทฤษฎีด้านการยอมรับเทคโนโลยีสารสนเทศ โดยมุ่งอธิบายพฤติกรรมการยอมรับหุ่นยนต์เสิร์ฟอาหารของผู้ใช้บริการในร้านอาหารอ้างถึง ผู้วิจัยพิจารณาปัจจัยสำคัญที่มีอิทธิพลต่อการตัดสินใจยอมรับเทคโนโลยี ได้แก่ การรับรู้ถึงประโยชน์ที่ได้รับจากการใช้งาน ความสะดวกและง่ายต่อการใช้งาน การรับรู้แรงสนับสนุนจากสังคมรอบข้าง และระดับความเชื่อมั่นต่อความสามารถของเทคโนโลยีปัจจัยดังกล่าวมีบทบาทในการกำหนดความตั้งใจในการใช้หุ่นยนต์เสิร์ฟอาหาร ซึ่งจะนำไปสู่การตัดสินใจใช้งานจริงในร้านอาหาร กรอบแนวคิดที่กำหนดขึ้นนี้ช่วยให้การศึกษามีทิศทางที่ชัดเจน และสามารถอธิบายความสัมพันธ์ระหว่างตัวแปรได้อย่างเป็นระบบตามวัตถุประสงค์ของการวิจัย




**ภาพที่ 2** กรอบแนวคิดการวิจัย

### 3.2 ประชากรและกลุ่มตัวอย่าง

3.2.1 ประชากรที่ใช้ในการวิจัย

ประชากรที่ใช้ในการวิจัยครั้งนี้ ได้แก่ ผู้บริโภคที่เคยใช้บริการร้านอาหารซึ่งมีการนำหุ่นยนต์เสิร์ฟอาหารมาใช้ สอดคล้องกับ Ivanov และ Webster (2023) [8]ในการให้บริการ โดยเป็นผู้ที่มีประสบการณ์ตรงในการรับรู้และประเมินการใช้งานเทคโนโลยีดังกล่าว

3.2.2 กลุ่มตัวอย่าง

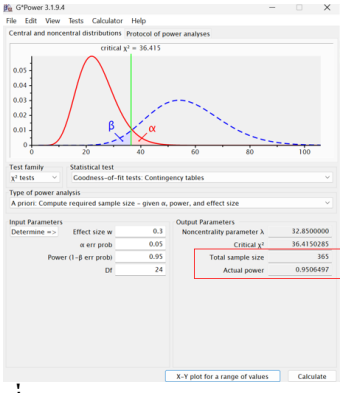
กลุ่มตัวอย่างที่ใช้ในการศึกษา คือ ผู้บริโภคที่เคยใช้บริการร้านอาหารที่มีหุ่นยนต์เสิร์ฟอาหาร โดยผู้วิจัยเก็บรวบรวมข้อมูลจากกลุ่มตัวอย่างที่สมัครใจตอบแบบสอบถามออนไลน์ กลุ่มตัวอย่างที่ได้

มีจำนวนทั้งสิ้น 433 คน ซึ่งเพียงพอสำหรับการวิเคราะห์ข้อมูลเชิงสถิติ

3.2.3 การกำหนดขนาดกลุ่มตัวอย่าง

การกำหนดขนาดกลุ่มตัวอย่างอ้างอิงตามหลักการของ Jackson ตามแนวคิดของ Kline (2016) [9] ซึ่งเสนอเกณฑ์อัตราส่วนระหว่างขนาดกลุ่มตัวอย่าง (N) ต่อจำนวนพารามิเตอร์ในโมเดล (q) ไม่น้อยกว่า 10:1

จากกรอบแนวคิดการวิจัยมีจำนวนพารามิเตอร์ที่ใช้ในการศึกษา 30 พารามิเตอร์ ส่งผลให้ขนาดกลุ่มตัวอย่างขั้นต่ำควรมีอย่างน้อย 365 คนจากการวิเคราะห์กำหนดกลุ่มตัวอย่างจาก G*Power อย่างไรก็ตาม งานวิจัยนี้สามารถเก็บข้อมูลจากกลุ่มตัวอย่างได้จำนวน 433 คน ซึ่งสูงกว่าเกณฑ์ขั้นต่ำและเหมาะสมต่อการวิเคราะห์ข้อมูล



**ภาพที่ 3** กำหนดกลุ่มตัวอย่างจาก G*Power

### 3.3 เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัยครั้งนี้คือ แบบสอบถามออนไลน์ (Questionnaire) จัดทำผ่าน Google Forms โดยเป็นคำถามแบบปลายปิด (Close-ended Questions) แบบสอบถามแบ่งออกเป็น 2 ส่วน ดังนี้ส่วนที่ 1 ข้อมูลทั่วไปของผู้ตอบแบบสอบถาม ส่วนที่ 2 ปัจจัยที่มีอิทธิพลต่อการยอมรับหุ่นยนต์เสิร์ฟอาหารในร้านอาหาร โดยใช้การวัดระดับความคิดเห็นแบบมาตราส่วนประมาณค่า (Rating Scale) ตามมาตรา Likert Scale 7 ระดับ ซึ่งเป็นข้อมูลระดับอันตรภาคชั้น (Interval Scale)

การวิเคราะห์ข้อมูลใช้ทั้งสถิติเชิงพรรณนา เช่น ความถี่ ร้อยละ ค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐาน รวมถึงการทดสอบความเชื่อมั่นของแบบสอบถามด้วย Cronbach's Alpha และการวิเคราะห์องค์ประกอบเชิงสำรวจ (EFA) ใช้การวิเคราะห์สหสัมพันธ์และการถดถอยพหุคูณ เพื่อทดสอบอิทธิพลของปัจจัยต่าง ๆ ต่อความตั้งใจและพฤติกรรมการยอมรับหุ่นยนต์เสิร์ฟอาหาร Belanche et al. (2020) [10] ทดสอบความสัมพันธ์และอิทธิพลระหว่างตัวแปรตามกรอบแนวคิดการวิจัยโดยใช้เครื่องมือ Jamovi

## 4. ผลการดำเนินงาน

4.1 ผลการวิเคราะห์ข้อมูลทั่วไปของกลุ่มตัวอย่าง

**ตารางที่ 1 :**ผลการวิเคราะห์ข้อมูลประชากรศาสตร์

| ข้อมูลทั่วไปของผู้ตอบแบบสอบถาม | จำนวน (คน) | ร้อยละ (%) |
|---|---|---|
| **1. เพศ** | | |
| ชาย | 143 | 33.00 |
| หญิง | 287 | 66.30 |
| อื่น ๆ | 3 | 0.70 |
| **รวม** | **433** | **100.00** |
| **2. ช่วงอายุ (ปี)** | | |
| 20-29 ปี | 124 | 28.60 |
| 30-39 ปี | 155 | 35.80 |
| 40-49 ปี | 67 | 15.50 |
| มากกว่า 50 ปี | 87 | 20.10 |
| **รวม** | **433** | **100.00** |
| **3. ระดับการศึกษา** | | |
| ต่ำกว่าปริญญาตรี | 53 | 12.20 |
| ปริญญาตรี | 328 | 75.80 |
| ปริญญาโท | 44 | 10.20 |
| ปริญญาเอก | 8 | 1.80 |
| **รวม** | **433** | **100.00** |
| **4. อาชีพ** | | |
| นักเรียน/นักศึกษา | 48 | 11.10 |
| พนักงานบริษัท | 150 | 34.60 |
| ข้าราชการ/รัฐวิสาหกิจ | 118 | 27.30 |
| เจ้าของกิจการ/ฟรีแลนซ์ | 67 | 15.50 |
| แม่บ้าน/เกษียณอายุ | 42 | 9.70 |
| อื่น ๆ | 8 | 1.80 |
| **รวม** | **433** | **100.00** |
| **5. รายได้ต่อเดือน** | | |
| น้อยกว่า 15,000 บาท | 48 | 11.00 |
| 15,001 – 30,000 บาท | 199 | 46.00 |
| 30,001 – 50,000 บาท | 81 | 28.20 |
| มากกว่า 50,000 บาท | 61 | 14.80 |
| **รวม** | **433** | **100.00** |

จากการศึกษาตารางที่1กลุ่มตัวอย่างทั้งหมดจำนวน 433 คน พบว่า ผู้ตอบแบบสอบถามส่วนใหญ่เป็นเพศหญิง จำนวน 287 คน คิดเป็นร้อยละ 66.30 รองลงมาคือเพศชาย 143 คน คิดเป็นร้อยละ 33.00 ส่วนใหญ่มีอายุ 30–39 ปี จำนวน 155 คน คิดเป็นร้อยละ 35.80 มีการศึกษาในระดับปริญญาตรี จำนวน 328 คน คิดเป็นร้อยละ 75.8 ส่วนใหญ่ประกอบอาชีพพนักงานบริษัท จำนวน 150 คน คิดเป็นร้อยละ 34.60 และมีรายได้อยู่ในช่วง 15,001–30,000 บาท จำนวน 199 คน คิดเป็นร้อยละ 46.00

**4.2 ผลการวิเคราะห์องค์ประกอบเชิงสำรวจ**

**ตารางที่ 2 :** การวิเคราะห์องค์ประกอบเชิงสำรวจ

| KMO และ Bartlett's Test | | |
|---|---|---|
| Kaiser-Meyer-Olkin Measure of Sampling Adequacy. | | 0.965 |
| Bartlett's Test of Sphericity | Approx. Chi-Square | 219.000 |
| | df | 85 |
| | Sig. | < .001 |

ผลการวิเคราะห์ความเหมาะสมของข้อมูลสำหรับการวิเคราะห์องค์ประกอบเชิงสำรวจ โดยใช้ข้อคำถามทั้งหมด 21 ตัวแปร พบว่า ค่า Kaiser-Meyer-Olkin Measure of Sampling Adequacy (KMO) มีค่าเท่ากับ 0.965 ซึ่งมีค่ามากกว่า 0.50 และเข้าใกล้ 1 แสดงให้เห็นว่าตัวแปรมีความสัมพันธ์กันในระดับสูงและมีความเหมาะสมอย่างยิ่งสำหรับการนำไปวิเคราะห์องค์ประกอบเชิงสำรวจ

ผลการทดสอบ Bartlett's Test of Sphericity พบว่ามีค่านัยสำคัญทางสถิติที่ระดับ น้อยกว่า 0.05 (Sig. < .001) แสดงให้เห็นว่าความสัมพันธ์ระหว่างตัวแปรแตกต่างอย่างมีนัยสำคัญทางสถิติ

**ตารางที่ 3** : Exploratory Factor Analysis Factor Matrix

| Factor Loadings | | | | | | | |
|---|---|---|---|---|---|---|---|
| | Factor | | | | | | Uniqueness |
| | 1 | 2 | 3 | 4 | 5 | 6 | |
| PE4 | 0.939 | | | | | | 0.1533 |
| PE1 | 0.929 | | | | | | 0.1762 |
| PE2 | 0.812 | | | | | | 0.2071 |
| PE3 | 0.787 | | | | | | 0.2351 |
| BI2 | | 0.892 | | | | | 0.1128 |
| BI1 | | 0.843 | | | | | 0.1375 |
| BI3 | | 0.777 | | | | | 0.1916 |
| BI4 | | 0.747 | | | | | 0.1234 |
| UB1 | | | 0.916 | | | | 0.0752 |
| UB3 | | | 0.763 | | | | 0.2111 |
| UB2 | | | 0.749 | | | | 0.1333 |
| EE3 | | | | 0.890 | | | 0.1562 |
| EE2 | | | | 0.844 | | | 0.1721 |
| EE4 | | | | 0.732 | | | 0.1813 |
| SI2 | | | | | 0.845 | | 0.0438 |
| SI3 | | | | | 0.401 | | 0.3527 |
| SI1 | | | | | 0.394 | | 0.2433 |
| TR4 | | | | | | 0.428 | 0.1713 |
| TR2 | | | | | | 0.415 | 0.2202 |
| TR3 | | | | | | 0.395 | 0.2088 |
| *Note.* 'Maximum likelihood' extraction method was used in combination with a 'oblimin' rotation | | | | | | | |

ผลการวิเคราะห์ Factor Loadings พบว่า ตัวแปรสามารถจัดกลุ่มได้ชัดเจนเป็น 6 องค์ประกอบ โดยมีค่าน้ำหนักอยู่ในระดับดีถึงสูง (0.394–0.939)องค์ประกอบที่ 1–4 มีค่า loading สูง (>0.7) แสดงถึงความสัมพันธ์ที่แข็งแรงของตัวแปรในแต่ละปัจจัยองค์ประกอบที่ 5 และ 6 มีค่าบางตัวต่ำกว่า 0.5 (เช่น 0.394–0.428) สะท้อนว่าความสัมพันธ์อยู่ในระดับปานกลางค่า Uniqueness ส่วนใหญ่อยู่ในระดับต่ำ แสดงว่าตัวแปรสามารถอธิบายด้วยปัจจัยร่วมได้ดีโดยรวมโครงสร้างปัจจัยมีความเหมาะสมและสามารถนำไปใช้วิเคราะห์ต่อได้

**4.3 ผลการทดสอบความเชื่อมั่นของแบบสอบถาม**

**ตารางที่ 4** : ผลการวิเคราะค่าความเชื่อมั่น ค่าเฉลี่ย และค่าส่วนเบี่ยงเบนมาตรฐาน

| ลำดับ | ปัจจัย | Mean | S.D. | Cronbach's Alpha |
|---|---|---|---|---|
| 1 | ความคาดหวังต่อประสิทธิภาพ (Performance Expectancy: PE) | 5.65 | 1.3 | 0.94 |
| 2 | ความง่ายในการใช้งาน Effort Expectancy (EE) | 5.61 | 1.4 | 0.939 |
| 3 | อิทธิพลทางสังคม Social Influence (SI) | 5.44 | 1.44 | 0.905 |
| 4 | ความไว้วางใจ Trust (TR) | 5.58 | 1.34 | 0.941 |
| 5 | ความตั้งใจใช้ Behavioral Intention (BI | 5.36 | 1.53 | 0.9576 |
| 6 | ความสนใจและมีพฤติกรรมการใช้จริง Use Behavior (UB) | 5.25 | 1.62 | 0.949 |

ผลการวิเคราะห์พบว่า ค่าเฉลี่ยอยู่ระหว่าง 5.25–5.65 สะท้อนทัศนคติเชิงบวกต่อหุ่นยนต์เสิร์ฟอาหาร โดยปัจจัยด้านความคาดหวังประสิทธิภาพมีค่าเฉลี่ยสูงสุด (Mean = 5.65, S.D. = 1.30) รองลงมาคือความง่ายในการใช้งานและความไว้วางใจด้านความเชื่อมั่น แบบสอบถามค่า Cronbach's Alpha อยู่ระหว่าง 0.905–0.9576 สูงกว่าเกณฑ์ 0.70 จึงสรุปได้ว่าเครื่องมือวิจัยมีความน่าเชื่อถือสูงและเหมาะสมสำหรับการวิเคราะห์ขั้นต่อไป

**4.4 ผลการวิเคราะห์สมการถดถอยพหุคูณ**

4.4.1การวิเคราะห์ค่าสัมประสิทธิ์การถดถอยของปัจจัยที่ความตั้งใจในการใช้งาน

**ตารางที่ 5** : Model Coefficients – BI

| Model Coefficients - BI | | | | | | Collinearity Statistics | |
|---|---|---|---|---|---|---|---|
| Predictor | Estimate | SE | t | p | Stand. Estimate | VIF | Tolerance |
| Intercept | -0.180 | 0.1895 | -0.952 | 0.342 | | | |
| PE | 0.140 | 0.0489 | 2.858 | 0.004 | 0.119 | 2.51 | 0.398 |
| EE | -0.109 | 0.0503 | -2.177 | 0.030 | -0.101 | 3.16 | 0.317 |
| SI | 0.516 | 0.0570 | 9.055 | < .001 | 0.465 | 3.85 | 0.260 |
| TR | 0.449 | 0.0664 | 6.766 | < .001 | 0.398 | 5.04 | 0.198 |

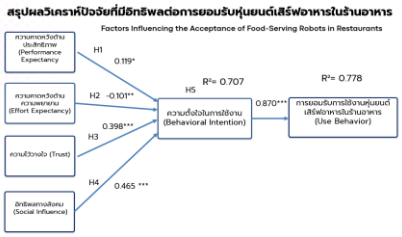
ผลการวิเคราะห์การถดถอยพหุคูณพบว่า ตัวแปรอิสระ ได้แก่ PE, EE, SI และ TR ไม่มีปัญหาความสัมพันธ์เชิงพหุคูณ โดยมีค่า VIF อยู่ระหว่าง 2.51–5.04 (น้อยกว่า 10) และค่า Tolerance อยู่ระหว่าง 0.198–0.398 (ส่วนใหญ่มากกว่า 0.20) แม้ TR จะมีค่าใกล้เกณฑ์ แต่ยังอยู่ในระดับที่ยอมรับได้ จึงสรุปได้ว่าสามารถนำตัวแปรทั้งหมดไปวิเคราะห์ร่วมกันในโมเดลได้อย่างเหมาะสม

4.4.2 แสดงผลการวิเคราะห์การถดถอยเชิงเส้น

**ตารางที่ 6** : Model Coefficients – UB

| Model Coefficients – UB | | | | | | Collinearity Statistics | |
|---|---|---|---|---|---|---|---|
| Predictor | Estimate | SE | t | p | Stand. Estimate | VIF | Tolerance |
| Intercept | 0.0536 | 0.1458 | 0.368 | 0.713 | | | |
| BI | 0.9578 | 0.0261 | 36.648 | <.001 | 0.870 | 1.00 | 1.00 |

ผลการวิเคราะห์การถดถอยเชิงเส้นพบว่า ค่าคงที่ไม่มีนัยสำคัญทางสถิติ ($p = 0.713$) ขณะที่ความตั้งใจในการใช้งาน (BI) มีอิทธิพลเชิงบวกอย่างมีนัยสำคัญต่อพฤติกรรมการใช้งาน (UB) ($\beta = 0.9578$, $p < 0.001$) โดยมีค่าสัมประสิทธิ์มาตรฐานเท่ากับ 0.870 แสดงถึงอิทธิพลในระดับสูงและไม่พบปัญหาความสัมพันธ์เชิงพหุคูณ (VIF = 1.00)




**ภาพที่ 4** โมเดลปัจจัยที่มีอิทธิพลต่อการยอมรับหุ่นยนต์เสิร์ฟอาหาร

การวิเคราะห์สมการถดถอยพหุคูณแสดงให้เห็นว่า ความไว้วางใจ ($\beta = 0.398$) และอิทธิพลทางสังคม ($\beta = 0.465$) มีอิทธิพลเชิงบวกต่อความตั้งใจใช้งานอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.001 ขณะที่ความคาดหวังด้านประสิทธิภาพมีผลเชิงบวก ($\beta = 0.119$, $p < 0.05$) และความง่ายในการใช้งานมีผลเชิงลบ ($\beta = -0.101$, $p < 0.01$) นอกจากนี้ ความตั้งใจใช้งานส่งผลต่อพฤติกรรมการใช้งานอย่างมีนัยสำคัญทางสถิติ ($\beta = 0.870$, $p < 0.001$) โดยโมเดลสามารถอธิบายความแปรปรวนได้ดี ($R^2 = 0.707$ และ $0.778$)

## 5. สรุปผลและอภิปราย

สมมติฐานที่ 1 (H1) ผลการวิเคราะห์การถดถอยพหุคูณพบว่า ความคาดหวังด้านประสิทธิภาพมีค่าสัมประสิทธิ์การถดถอยเป็นลบเล็กน้อย ($\beta = -0.049$) และมีค่า Sig. เท่ากับ 0.185 ซึ่งมากกว่าระดับนัยสำคัญทางสถิติที่ 0.05 แสดงให้เห็นว่า ความคาดหวังด้านประสิทธิภาพไม่ปรากฏอิทธิพลต่อความตั้งใจในการใช้งานอย่างมีนัยสำคัญทางสถิติ (Sig. > 0.05) แสดงให้เห็นว่าการรับรู้ถึงความง่ายหรือความยากในการใช้งานหุ่นยนต์เสิร์ฟอาหาร

สมมติฐานที่ 3 (H3) อิทธิพลทางสังคม (Social Influence: SI) อิทธิพลทางสังคมมีค่า Sig. เท่ากับ 0.833 ซึ่งมากกว่าระดับนัยสำคัญทางสถิติที่ 0.05 แสดงว่าอิทธิพลทางสังคม ไม่ส่งผลต่อความตั้งใจในการใช้งานหุ่นยนต์เสิร์ฟอาหารอย่างมีนัยสำคัญทางสถิติ

สมมติฐานที่ 4 (H4) ความไว้วางใจ (Trust: TR) มีอิทธิพลเชิงบวกต่อความตั้งใจในการใช้งานหุ่นยนต์เสิร์ฟอาหาร (Behavioral Intention: BI)ความไว้วางใจมีอิทธิพลเชิงบวกต่อความตั้งใจในการใช้งานอย่างมีนัยสำคัญทางสถิติ โดยมีค่า Sig. เท่ากับ 0.044

สมมติฐานที่ 5 (H5) ความตั้งใจในการใช้งาน (Behavioral Intention: BI) มีอิทธิพลเชิงบวกต่อพฤติกรรมการใช้งานจริงของหุ่นยนต์เสิร์ฟอาหาร (Use Behavior: UB)ผลการวิเคราะห์การถดถอยเชิงเส้นอย่างง่ายพบว่าความตั้งใจในการใช้งานมีอิทธิพลเชิงบวกต่อพฤติกรรมการใช้งานจริงอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.001 ($\beta = 0.870$, Sig. < .001)

## เอกสารอ้างอิง

[1] Viswanath Venkatesh, Morris, M. G., Davis, G. B., & Fred D. Davis. (2003).User acceptance of information technology: Toward a unified view. *MIS Quarterly*, 27(3), 425–478.

[2] Fred D. Davis. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly, 13(3), 319–340.

[3] Joseph F. Hair, Black, W. C., Babin, B. J., & Anderson, R. E. (2010). Multivariate data analysis (7 th ed.). Pearson Education.

[4] Jacob Cohen, Cohen, P., West, S. G., & Aiken, L. S. (2003).Applied multiple regression/correlation analysis for the behavioral sciences (3rd ed.). Lawrence Erlbaum Associates.

[5] กฤตภัค วรเมธพาสุข. (2562). [วิทยานิพนธ์ออนไลน์]. อิทธิพลของความคล้ายคลึงเชิงรูปลักษณ์ และการรับรู้การ

ควบคุมได้ ที่มีผลต่อการรับรู้ความเหมือนมนุษย์ ความเชื่อใจ และการยอมรับต่อหุ่นยนต์

[6] ณัชชารีย์ ฐิติสกุลวิทย์. (2565). [วิทยานิพนธ์ออนไลน์]. ปัจจัยที่มีผลต่อความตั้งใจในการยอมรับหุ่นยนต์เพื่อนเพื่อการดูแลผู้สูงอายุในมุมมองของการประยุกต์แบบจำลองการยอมรับระบบแรงจูงใจด้านความพึงพอใจ

[7] บุญทวี จันทสุวรรณ. (2566). [วิทยานิพนธ์ออนไลน์]. การยอมรับนวัตกรรมโดยใช้การประเมินคุณภาพบริการของหุ่นยนต์บริการในโรงแรม 5 ดาวในเขต กรุงเทพมหานคร สาขาวิชาการจัดการการท่องเที่ยว โรงแรม และอีเวนต์ มหาวิทยาลัยศิลปากร

[8] Ivanov, S., & Webster, C. (2023). “Restaurants and robots: public preferences for robot food and beverage services.” Journal of Tourism Futures, 9(2), 229-239.

[9] Kline, R. B. (2016). Principles and practice of structural equation modeling. 4th ed. New York: The Guilford Press.

[10] Belanche, D., Casaló, L. V., Flavián, C., & Schepers, J. (2020). “Service robot implementation: a theoretical framework and research agenda.” The Service Industries Journal, 40(3-4), 203-225.

# การพัฒนาระบบสนับสนุนการตัดสินใจจำแนกกลุ่มผู้กระทำผิดคดียาเสพติด

# Development of a Decision Support System for Rehabilitation Planning of Drug Offenders

*สุปกรณ์ ศรีทองจันทร์ (Supakorn Srithongjun)[1] และณัฐวี อุตกฤษฎ์ (Nattavee Utakrit)[2]*
*ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล*
*มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ*
*[1]s6707021857104@email.kmutnb.ac.th, [2] nattavee.u@itd.kmutnb.ac.th*

## บทคัดย่อ

*งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาระบบสนับสนุนการตัดสินใจสำหรับจำแนกกลุ่มผู้กระทำผิดคดียาเสพติดโดยประยุกต์ใช้เทคนิคการเรียนรู้ของเครื่องกับข้อมูลที่ไม่ระบุตัวตน จำนวน 120,000 ระเบียน แบ่งเป็น 4 กลุ่มความเสี่ยงประกอบด้วยกลุ่ม1 กลุ่ม2 กลุ่ม3 และกลุ่ม4 กลุ่มละ 30,000 ระเบียนอย่างสมดุลโดยใช้หลักเกณฑ์จากกลุ่มที่มีจำนวนระเบียนน้อยที่สุดเพื่อป้องกันความเอนเอียงในการทำนายผล ข้อมูลถูกแบ่งเป็นชุดฝึก 80% และชุดทดสอบ 20% พร้อมใช้ 10-fold stratified cross-validation ร่วมกับ GridsearchCV ในการปรับแต่งพารามิเตอร์ ผลการเปรียบเทียบอัลกอริทึม k-Nearest Neighbors, Random Forest และ Support Vector Machine พบว่า Random Forest ให้ค่าความถูกต้องสูงสุด (0.8257) รองลงมาคือ Support Vector Machine (0.8238 ) และ k-Nearest Neighbors (0.8020) จึงเลือกแบบจำลอง Random Forest ไปพัฒนาต้นแบบระบบสนับสนุนการตัดสินใจต่อไป*

**คำสำคัญ:** การเรียนรู้ของเครื่อง, Random Forest, Support Vector Machine, k-Nearest Neighbors.

## Abstract

*This study aimed to develop a decision support system for classifying drug offenders into risk groups using machine learning techniques applied to anonymized data. The dataset consisted of 120,000 records, evenly balanced across four risk groups, namely Group 1, Group 2, Group 3, and Group 4, with 30,000 records in each group. The balancing process was based on the group with the smallest number of records in order to minimize prediction bias. The dataset was divided into 80% for training and 20% for testing. In addition, 10-fold stratified cross-validation together with GridSearchCV was employed for hyperparameter tuning and model evaluation. Three machine learning algorithms were compared: k-Nearest Neighbors, Random Forest, and Support Vector Machine. The results indicated that Random Forest achieved the highest accuracy (0.8257), followed by Support Vector Machine (0.8238) and k-Nearest Neighbors (0.8020). Based on these findings, the Random Forest model was selected for further development as a prototype decision support system.*

**Keywords:** machine learning, Random Forest, Support Vector Machine, k-Nearest Neighbors.

## 1. บทนำ

ประเทศไทยได้นำมาตรการคุมประพฤติมาใช้เป็นกลไกสำคัญในกระบวนการยุติธรรมทางอาญา โดยมีเป้าหมายเพื่อแก้ไขฟื้นฟูผู้กระทำผิดและส่งเสริมให้สามารถกลับคืนสู่สังคมได้อย่างเหมาะสม ในการดำเนินงานดังกล่าว กรมคุมประพฤติมีบทบาทสำคัญในการติดตาม ดูแล และประเมินผู้กระทำผิดที่อยู่ภายใต้มาตรการคุมประพฤติ โดยเฉพาะผู้กระทำผิดคดียาเสพติด ซึ่งเป็นกลุ่มคดีที่มีจำนวนมากและมีลักษณะปัญหาซับซ้อน

อย่างไรก็ตาม การประเมินและจำแนกกลุ่มความเสี่ยงของผู้กระทำผิดในทางปฏิบัติยังอาศัยดุลยพินิจของเจ้าหน้าที่ร่วมกับข้อมูลพื้นฐานและประสบการณ์ในการพิจารณาเป็นสำคัญ แม้วิธีการดังกล่าวจะสามารถสะท้อนบริบทของผู้กระทำผิดแต่ละรายได้ในระดับหนึ่ง แต่ก็อาจก่อให้เกิดความคลาดเคลื่อนและความไม่สม่ำเสมอในการประเมินได้ โดยเฉพาะในกรณีที่มีข้อมูลจำนวนมากหรือมีปัจจัยที่เกี่ยวข้องหลายด้าน ส่งผลให้การจำแนกกลุ่มความเสี่ยงอาจใช้เวลามากและมีข้อจำกัดด้านประสิทธิภาพ

ในปัจจุบัน เทคโนโลยีการเรียนรู้ของเครื่องได้รับความสนใจอย่างกว้างขวางในการประยุกต์ใช้เพื่อวิเคราะห์ข้อมูลและสนับสนุนการตัดสินใจ เนื่องจากสามารถเรียนรู้รูปแบบความสัมพันธ์จากข้อมูลในอดีตและนำมาใช้ในการจำแนกหรือพยากรณ์ข้อมูลใหม่ได้อย่างเป็นระบบ การนำเทคนิคดังกล่าวมาประยุกต์ใช้กับข้อมูลผู้กระทำผิดจึงอาจช่วยสนับสนุนการจำแนกกลุ่มความเสี่ยงให้มีความเป็นระบบมากยิ่งขึ้น และช่วยลดข้อจำกัดจากการประเมินที่พึ่งพาดุลยพินิจเพียงอย่างเดียว

จากเหตุผลดังกล่าว งานวิจัยนี้จึงมีวัตถุประสงค์เพื่อพัฒนาระบบสนับสนุนการตัดสินใจสำหรับจำแนกกลุ่มผู้กระทำผิดคดียาเสพติด โดยประยุกต์ใช้เทคนิคการเรียนรู้ของเครื่องกับข้อมูลที่ไม่ระบุตัวตน เพื่อสร้างแบบจำลองที่สามารถจำแนกกลุ่มความเสี่ยงได้อย่างมีประสิทธิภาพ ผลการวิจัยที่ได้อาจเป็นประโยชน์ต่อการสนับสนุนการปฏิบัติงานของเจ้าหน้าที่คุมประพฤติ และเป็นข้อมูลประกอบการพัฒนาระบบประเมินความเสี่ยงที่มีความเป็นระบบมากยิ่งขึ้น

## 2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

การวิจัยนี้ประยุกต์ใช้เทคนิคการเรียนรู้ของเครื่องสำหรับการจำแนกประเภทข้อมูล ได้แก่

K-Nearest Neighbor [1] เป็นวิธีการจำแนกข้อมูลโดยพิจารณาระยะห่างระหว่างข้อมูลใหม่กับข้อมูลตัวอย่างที่อยู่ใกล้เคียงที่สุดจำนวน k ค่า แล้วตัดสินผลตามเสียงข้างมากของกลุ่มเพื่อนบ้าน มีจุดเด่นด้านความเรียบง่ายและเหมาะกับข้อมูลที่มีโครงสร้างชัดเจน

Random Forest [2] เป็นเทคนิคการเรียนรู้แบบกลุ่มที่สร้างต้นไม้ตัดสินใจหลายต้นและรวมผลการทำนายเข้าด้วยกัน ช่วยลดปัญหา Overfitting และเพิ่มความแม่นยำ เหมาะกับข้อมูลที่มีความซับซ้อนและมีตัวแปรจำนวนมาก

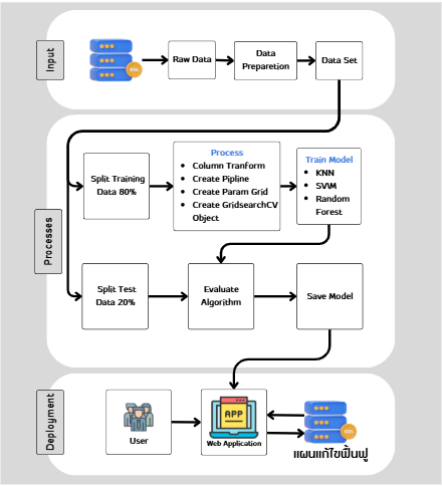
Support Vector Machine [3] เป็นวิธีการสร้างเส้นแบ่งเขต ที่เหมาะสมที่สุดเพื่อแยกกลุ่มข้อมูลให้มีระยะห่างสูงสุด และสามารถจัดการข้อมูลที่ไม่เป็นเชิงเส้นได้ผ่านการใช้ Kernel Function

ในการประเมินประสิทธิภาพแบบจำลอง ใช้เทคนิค Cross-Validation [4] เพื่อลดความเอนเอียงจากการแบ่งข้อมูลเพียงครั้งเดียว และใช้ GridSearchCV [5] ในการค้นหาค่าพารามิเตอร์ที่เหมาะสมที่สุดอย่างเป็นระบบ ส่งผลให้ได้แบบจำลองที่มีความแม่นยำและความน่าเชื่อถือสูงขึ้น

## 3. วิธีการดำเนินการวิจัย

### 3.1 กรอบแนวคิด

การวิจัยนี้มุ่งพัฒนาระบบสนับสนุนการตัดสินใจเพื่อจำแนกกลุ่มความเสี่ยงกระทำผิดซ้ำ โดยประยุกต์ใช้เทคนิคการเรียนรู้ของเครื่อง กระบวนการวิจัยประกอบด้วย 3 ขั้นตอนหลัก ได้แก่ (1) การรวบรวมและเตรียมข้อมูล (2) การพัฒนาแบบจำลองด้วยอัลกอริทึม K-Nearest Neighbor, Random Forest และ Support Vector Machine และประเมินโดยใช้ตัวชี้วัด Accuracy, Precision, Recall และ F1-score และ (3) การนำแบบจำลองที่มีประสิทธิภาพสูงสุดไปพัฒนาเป็นเว็บแอปพลิเคชันต้นแบบ




**ภาพที่ 1:** กรอบแนวคิด

**3.2 รวบรวมข้อมูล วิเคราะห์คุณลักษณะและจัดเตรียมข้อมูล**

ข้อมูลผู้กระทำผิดจำนวน 120,000 ระเบียน แบ่งเป็น 4 กลุ่มความเสี่ยง คือ กลุ่ม 1 เข้มงวด (ความต้องการหลัก) กลุ่ม 2 (ความต้องการรอง) กลุ่ม 3 เข้มงวดปานกลาง และกลุ่ม 4 ไม่เข้มงวด ทั้งนี้เพื่อป้องกันความเอนเอียงในทำนายผลของแบบจำลองจึงจำเป็นที่จะต้องแบ่งจำนวนระเบียนแต่ละกลุ่มเท่ากันๆ โดยพิจารณาจากกลุ่มที่มีจำนวนระเบียนน้อยที่สุด จากนั้นจึงทำการแบ่งกลุ่มจำนวนกลุ่มละ 30,000 ระเบียน และคัดเลือกคุณลักษระที่เกี่ยวข้องจำนวน 18 ตัวแปร โดยอ้างอิงตาม **เกณฑ์และแบบประเมินที่ใช้ในการจำแนกกลุ่มผู้กระทำผิด** [6] จากนั้นดำเนินการทำความสะอาดข้อมูล แปลงข้อมูลเชิงคุณภาพเป็นเชิงตัวเลข ตัดระเบียนที่ไม่สมบูรณ์ และปรับรูปแบบข้อมูลให้เหมาะสมสำหรับการฝึกแบบจำลอง

**3.3 การพัฒนาแบบจำลอง**

แบ่งข้อมูลเป็นชุดฝึกและชุดทดสอบในสัดส่วน 80:20 พร้อมปรับสมดุลข้อมูล และทำ Preprocessing ชุดข้อมูลใช้ GridSearchCV ร่วมกับ Cross-Validation 10 Fold เพื่อค้นหาค่าพารามิเตอร์ที่เหมาะสมที่สุดของแต่ละอัลกอริทึม ประเมินผลด้วยตัวชี้วัด Accuracy, Precision, Recall, F1-score และวิเคราะห์ Confusion Matrix ก่อนคัดเลือกแบบจำลองที่มีประสิทธิภาพสูงสุด

**3.4 การพัฒนาระบบ**

พัฒนาเว็บแอปพลิเคชันโดยเชื่อมต่อแบบจำลองที่คัดเลือกกับส่วนติดต่อผู้ใช้ เพื่อให้สามารถกรอกข้อมูล ประมวลผล และแสดงผลการจำแนกกลุ่มความเสี่ยงได้แบบอัตโนมัติ

**3.5 การประเมินผลระบบ**

ประเมินผล 2 ด้าน ได้แก่ ความถูกต้องของการจำแนกโดยทดสอบกับข้อมูลใหม่ 30 ชุดที่ไม่เคยใช้ฝึกแบบจำลอง และความพึงพอใจของผู้ใช้งาน โดยใช้แบบสอบถามมาตราส่วนลิเคิร์ต 5 ระดับ วิเคราะห์ผลด้วยค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐาน

## 4. ผลการดำเนินงาน

**4.1 ผลการทดสบประสิทธิภาพอัลกอริทึม**

การทดลองใช้ข้อมูลผู้กระทำผิดจำนวน 120,000 ระเบียน ปรับสมดุลเป็น 4 กลุ่มความเสี่ยงเท่า ๆ กัน และแบ่งข้อมูลเป็นชุดฝึก 80% และชุดทดสอบ 20% เพื่อรักษาสัดส่วนของแต่ละกลุ่ม

4.1.1 การทดสอบอัลกอริทึม K-Nearest Neighbor ได้กำหนดช่วงค่าพารามิเตอร์ ได้แก่ n_neighbors, leaf_size และ p ทำการปรับแต่งด้วยวิธี Grid Search ร่วมกับ Stratified 10-fold Cross-Validation แสดงดังตารางที่ 1

**ตารางที่ 1** อัลกอริทึม K-Nearest Neighbor

| พารามิเตอร์ | ค่าพารามิเตอร์ | ค่าที่ดีที่สุด |
|---|---|---|
| n_neighbors | 17, 19, 21, 25 | 19 |
| leaf_size | 10, 20, 40, 60 | 10 |
| p | 1, 2 | 1 |

ผลการทดลองพบว่าค่าที่เหมาะสมที่สุดคือ n_neighbors คือ 19, leaf_size คือ 19 และ p คือ 1 ซึ่งให้ประสิทธิภาพการจำแนกสูงสุดและเหมาะสมกับลักษณะข้อมูลในงานวิจัย ก่อนนำไปทดสอบกับชุดข้อมูลทดสอบในตารางที่ 2

**ตารางที่ 2** การทดสอบของอัลกอริทึม K-Nearest Neighbor

| พารามิเตอร์ | ค่าพารามิเตอร์ที่ดีที่สุด |
|---|---|
| Accuracy | 0.8020 |
| Precision | 0.7973 |
| Recall | 0.802 |
| F1-score | 0.7964 |

จากตารางที่ 2 พบว่าแบบจำลอง K-Nearest Neighbor มีค่า Accuracy 0.8020 และค่าชี้วัดอื่นใกล้เคียงกัน แสดงถึงความสมดุลในการจำแนกและสามารถนำไปใช้งานได้อย่างเหมาะสม

4.1.1 การทดสอบอัลกอริทึม Random Forest ได้กำหนดช่วงค่าพารามิเตอร์ ได้แก่ n_estimatores min_samples_split min_samples_leaf และ max_depth ทำการปรับแต่งด้วยวิธี GridSearchCV ร่วมกับ Stratified 10-fold Cross-Validation แสดงดังตารางที่ 3

**ตารางที่ 3** อัลกอริทึม Random Forest

| พารามิเตอร์ | ค่าพารามิเตอร์ | ค่าที่ดีที่สุด |
|---|---|---|
| n_estimatores | 200, 400, 800 | 200 |
| max_depth | 23, 25, 27 | 25 |
| min_samples_split | 2, 4 | 2 |
| min_samples_leaf | 2, 5 | 5 |

ผลการทดลองสามารถคัดเลือกชุดค่าพารามิเตอร์ที่ให้ค่า Accuracy สูงสุดและเหมาะสมกับลักษณะข้อมูล เพื่อนำเสนอรายละเอียดในตารางที่ 4

**ตารางที่ 4** การทดสอบของอัลกอริทึม Random Forest

| พารามิเตอร์ | ค่าพารามิเตอร์ที่ดีที่สุด |
|---|---|
| Accuracy | 0.8257 |
| Precision | 0.8224 |
| Recall | 0.8257 |
| F1-score | 0.8199 |

พบว่าแบบจำลอง Random Forest มีมีค่าพารามิเตอร์ที่เหมาะสมที่สุดคือ n_estimatores คือ 200 max_depth คือ 25 min_samples_split คือ 2 และ min_samples_leaf คือ 5 ค่า Accuracy 0.8257 และค่าชี้วัดอื่นใกล้เคียงกัน แสดงถึงประสิทธิภาพและความสมดุลในการจำแนก สามารถนำไปใช้สนับสนุนการตัดสินใจได้อย่างเหมาะสม

4.1.3 การทดสอบอัลกอริทึม Support Vector Machine ได้กำหนดช่วงค่าพารามิเตอร์ ได้แก่ C และ gamma ทำการปรับแต่งด้วยวิธี GridSearchCV ร่วมกับ Stratified 10-fold Cross-Validation แสดงดังตารางที่ 5

**ตารางที่ 5** อัลกอริทึม Support Vector Machine

| พารามิเตอร์ | ค่าพารามิเตอร์ | ค่าที่ดีที่สุด |
|---|---|---|
| C | 500, 700, 1,100 | 500 |
| gamma | 0.02, 0.005, 0.001 | 0.005 |

พบว่าอัลกอริทึม Support Vector Machine มีค่าพารามิเตอร์ที่เหมาะสมที่สุดคือ C คือ 500 และ gamma คือ 0.005 ซึ่งให้ประสิทธิภาพการจำแนกสูงสุด จากนั้นได้นำค่าดังกล่าวไปสร้างแบบจำลองและทดสอบกับชุดข้อมูลทดสอบ เพื่อประเมินผลด้วยค่า Accuracy, Precision, Recall และ F1-score ดังแสดงในตารางที่ 6

**ตารางที่ 6** การทดสอบของอัลกอริทึม Support Vector Machine

| พารามิเตอร์ | ค่าพารามิเตอร์ที่ดีที่สุด |
|---|---|
| Accuracy | 0.8238 |
| Precision | 0.8204 |
| Recall | 0.8238 |
| F1-score | 0.8186 |

แบบจำลอง Support Vector Machine มีค่า Accuracy 0.8233 และค่าชี้วัดอื่นๆใกล้เคียงกัน แสดงถึงประสิทธิภาพและความสมดุลในการจำแนก สามารถนำไปประยุกต์ใช้ได้อย่างเหมาะสม

4.1.3 การเปรียบเทียบประสิทธิภาพของแบบจำลองภายหลังการปรับแต่งและทดสอบแบบจำลองทั้ง 3 อัลกอริทึม โดยผลการเปรียบเทียบแสดงในตารางที่ 7

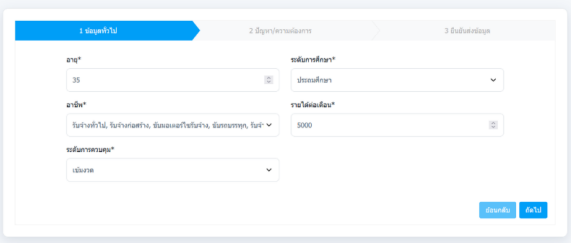
**ตารางที่ 7** ผลเปรียบเทียบประสิทธิภาพของแบบจำลอง

| อัลกอริทึม | ค่าความถูกต้อง |
|---|---|
| K-Nearest Neighbor | 0.8020 |
| Random Forest | 0.8257 |
| Support Vector Machine | 0.8238 |

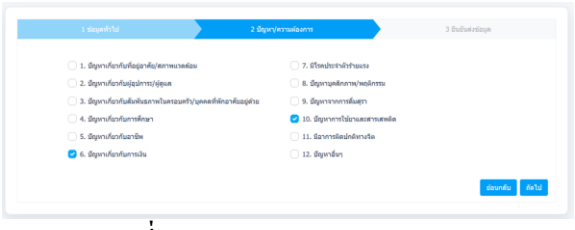
จากการเปรียบเทียบพบว่า Random Forest มีค่า Accuracy สูงสุดที่ 0.8257 รองลงมาคือ Support Vector Machine คือ 0.8238 และ K-Nearest Neighbor คือ 0.8020 แสดงว่า Random Forest มีประสิทธิภาพดีที่สุด จึงถูกเลือกเป็นแบบจำลองหลักสำหรับพัฒนาเว็บแอปพลิเคชัน

### 4.2 ผลการออกแบบและพัฒนาระบบ

ผู้วิจัยพัฒนาระบบต้นแบบด้วยภาษา Python บนเฟรมเวิร์ก Django และเชื่อมต่อฐานข้อมูล MySQL เพื่อรองรับกระบวนการทำงานตั้งแต่การยืนยันตัวตน การกรอกข้อมูลทั่วไปและข้อมูลปัญหา/ความต้องการ การส่งข้อมูลเข้าสู่แบบจำลองเพื่อจำแนกกลุ่มความเสี่ยง ตลอดจนการแสดงผลการจำแนกพร้อมเสนอแนะกิจกรรมและแผนการแก้ไขฟื้นฟูที่สอดคล้อง นอกจากนี้ระบบยังมีฟังก์ชันบันทึกและแสดงประวัติการใช้งานย้อนหลัง เพื่อสนับสนุนการติดตามและการตัดสินใจของผู้ใช้งานอย่างเป็นระบบ



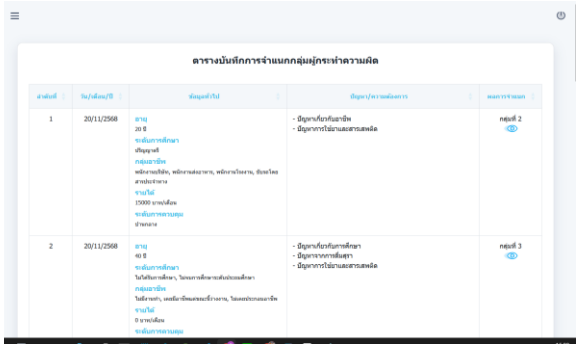
**ภาพที่ 2:** ฟอร์มกรอกข้อมูลทั่วไป



**ภาพที่ 3:** ฟอร์มระบุปัญหาและความต้องการ




**ภาพที่ 4:** หน้าแสดงผลการจำแนก




**ภาพที่ 5:** หน้าแสดงประวัติการจำแนกในอดีต

### 4.3 ผลการวิเคราะห์ข้อมูลและสถิติที่ใช้ในการวิจัย

ผู้วิจัยประเมินคุณภาพระบบโดยให้ผู้เชี่ยวชาญ 3 ท่าน จากกรมคุมประพฤติทดลองใช้งานในสถานการณ์จำลอง ครอบคลุมขั้นตอนตั้งแต่การเข้าสู่ระบบ การกรอกข้อมูล การประมวลผล และการแสดงผลลัพธ์ จากนั้นให้ประเมินทั้งด้านความถูกต้องของแบบจำลองและความเหมาะสมในการใช้งานจริง ผลการประเมินถูกวิเคราะห์ด้วยสถิติเชิงพรรณนา โดยรายงานผลแยกเป็น

4.3.1 ผลการประเมินด้านความถูกต้องของแบบจำลอง

ผู้วิจัยทดสอบความถูกต้องของแบบจำลองโดยให้ผู้เชี่ยวชาญ 3 ท่านทดลองจำแนกข้อมูลใหม่ที่ไม่เคยใช้ฝึกแบบจำลอง ท่านละ 10 ชุด รวม 30 ชุดข้อมูล

**ตารางที่ 8** ผลการวัดความถูกต้องในการจำแนก

| ผู้ประเมิน | จำนวนถูก | จำนวนผิด | เปอร์เซ็นต์ความถูกต้อง |
|---|---|---|---|
| ท่านที่ 1 | 9 | 1 | 90.00% |
| ท่านที่ 2 | 9 | 1 | 90.00% |
| ท่านที่ 3 | 10 | 0 | 100.00% |
| รวม | | | 93.33% |

แบบจำลองมีประสิทธิภาพและความน่าเชื่อถือสูง เหมาะสมต่อการนำไปใช้เป็นกลไกหลักในระบบสนับสนุนการตัดสินใจเชิงปฏิบัติการ

4.3.2 ผลการวัดคุณภาพด้านความสามารถของระบบ

ผู้วิจัยประเมินความสามารถของระบบ โดยผลการประเมินแสดงในตารางที่ 9

**ตารางที่ 9** ผลการวัดความถูกต้องในการจำแนก

| รายการ | ค่าเฉลี่ย | S.D. |
|---|---|---|
| 1. สามารถแนะนำกิจกรรมหรือแผนการแก้ไขฟื้นฟูแก่ผู้ใช้ได้ | 4.00 | 0.00 |
| 2. ความสามารถในการประมวลผลของระบบ | 4.67 | 0.58 |
| 3. สามารถแสดงผลได้ตรงตามความต้องการ | 5.00 | 0.00 |
| 4. ความสามารถของระบบโดยภาพรวม | 4.67 | 0.58 |
| **รวม** | 4.58 | 0.29 |

ผลการประเมินด้านความสามารถของระบบมีค่าเฉลี่ยรวม 4.58 และ S.D. = 0.29 อยู่ในระดับดีมาก สะท้อนว่าระบบสามารถประมวลผลและแสดงผลได้ตรงตามความต้องการอย่างมีประสิทธิภาพ โดยด้านการแสดงผลและภาพรวมของระบบมี ค่าเฉลี่ย 5 ขณะที่การแนะนำแผนการแก้ไขฟื้นฟูอยู่ในระดับดี ค่าเฉลี่ย 4 โดยสรุประบบมีความพร้อมสำหรับการนำไปใช้งาน

4.3.3 ผลการวัดคุณภาพด้านการทำงานของระบบ

ผู้วิจัยประเมินความคุณภาพด้านการทำงานของระบบ โดยผลการประเมินแสดงในตารางที่ 10

**ตารางที่ 10** ผลการวัดคุณภาพด้านการทำงานของระบบ

| รายการ | ค่าเฉลี่ย | S.D. |
|---|---|---|
| 1. ความรวดเร็วในการตอบสนองของระบบ | 5.00 | 0.00 |
| 2. ความง่ายในการใช้งานฟังก์ชั่นต่าง ๆ ของระบบ | 4.67 | 0.58 |
| 3. ความรวดเร็วในการประมวลผลของระบบ | 5.00 | 0.00 |
| 4. การทำงานของระบบโดยภาพรวม | 5.00 | 0.00 |
| **รวม** | 4.92 | 0.14 |

ผลการประเมินด้านการทำงานของระบบมีค่าเฉลี่ยรวม 4.92 และS.D. = 0.14 อยู่ในระดับดีมาก โดยด้านความรวดเร็วในการตอบสนองและการประมวลผลได้คะแนน

สูงสุด ค่าเฉลี่ย 5.00 สะท้อนถึงประสิทธิภาพที่โดดเด่น ขณะที่ความง่ายในการใช้งานอยู่ในระดับดี ค่าเฉลี่ย 4.67 ซึ่งสามารถพัฒนาเพิ่มเติมได้ โดยสรุประบบมีความพร้อมสำหรับการใช้งานจริง

4.3.4 ผลการวัดคุณภาพด้านความง่ายในการใช้งานระบบ

ผู้วิจัยประเมินความคุณภาพด้านความง่ายในการใช้งานระบบ แสดงในตารางที่ 11

**ตารางที่ 11** ผลการวัดคุณภาพด้านความง่ายในการใช้งาน

| รายการ | ค่าเฉลี่ย | S.D. |
|---|---|---|
| 1. ขนาดตัวอักษร รูปแบบ และความสวยงาม | 4.67 | 0.58 |
| 2. การจัดวางรูปแบบในเว็บไซต์ง่ายต่อการอ่านและการใช้งาน | 4.33 | 0.58 |
| 3. ความเหมาะสมในการจัดวางหน้าจอ | 4.67 | 0.58 |
| 4. ความน่าใช้งานของระบบ | 4.33 | 0.58 |
| **รวม** | **4.50** | **0.58** |

ผลการประเมินด้านความน่าใช้งานมีค่าเฉลี่ยรวม 4.50 และ S.D. คือ 0.58 อยู่ในระดับดี โดยด้านรูปแบบตัวอักษร ความสวยงาม และการจัดวางหน้าจอได้ค่าเฉลี่ย 4.67 ขณะที่การจัดวางองค์ประกอบเพื่อความสะดวกในการอ่านอยู่ในระดับดี ค่าเฉลี่ย 4.33 ซึ่งสามารถพัฒนาเพิ่มเติมได้อีก

4.3.3 ผลการวัดคุณภาพด้านความปลอดภัยของระบบ

ผู้วิจัยประเมินคุณภาพด้านความปลอดภัยของระบบ แสดงในตารางที่ 12

**ตารางที่ 12** ผลการวัดคุณภาพด้านความปลอดภัยของระบบ

| รายการ | ค่าเฉลี่ย | S.D. |
|---|---|---|
| 1. ความปลอดภัยของระบบโดยภาพรวม | 5.00 | 0.00 |
| 2. การเข้าถึงฐานข้อมูล | 4.67 | 0.58 |
| **รวม** | 4.80 | 0.29 |

ผลการประเมินด้านความปลอดภัยของระบบมีค่าเฉลี่ยรวม 4.80 และS.D. = 0.29 อยู่ในระดับดีมาก โดยความปลอดภัยของระบบโดยภาพรวมได้ค่าเฉลี่ย 5.00 สะท้อนว่าระบบมีมาตรการรักษาความปลอดภัยที่เหมาะสมและเพียงพอ

## 5. สรุปผลและอภิปราย

5.1 สรุปผลการวิจัย ผู้วิจัยเตรียมและปรับสมดุลข้อมูลก่อนพัฒนาแบบจำลอง โดยเปรียบเทียบ k-Nearest Neighbor, Support Vector Machine และ Random Forest พบว่า Random Forest ให้ Accuracy สูงสุดที่ 0.8257 จึงเลือกเป็นแบบจำลองหลัก และพัฒนาระบบต้นแบบที่สามารถจำแนกความเสี่ยงและแสดงข้อเสนอแนะได้ครบวงจร

5.2 อภิปรายผล Random Forest มีประสิทธิภาพดีที่สุดเนื่องจากรองรับความสัมพันธ์เชิงซับซ้อนได้ดีและลด Overfitting ขณะที่ SVM ให้ผลใกล้เคียงกัน แต่ในการทำนายผลกลับพบว่า Random Forest ใช้เวลาในการประมวลผลนานกว่า Support Vector Machine ค่อยข้างมาก ส่วน k-Nearest Neighbor มีข้อจำกัดด้านมิติข้อมูล การปรับสมดุลข้อมูลช่วยให้การประเมินประสิทธิภาพแบบจำลองมีความเป็นธรรมมากขึ้น

5.3 ข้อเสนอแนะ ควรเพิ่มตัวแปรเชิงพฤติกรรม ทดลองอัลกอริทึมขั้นสูงเพิ่มเติม นำเทคนิคอื่นๆเข้ามาช่วย เช่น การทำ Soft Vote, Hard Vote, Class Weight Score หรือเทคนิคการแยกกลุ่มทำนาย การใช้งานจริงควรพิจารณาเวลาที่แบบจำลองใช้ประมวลผล และพัฒนาด้าน UX และการจัดการสิทธิ์ เพื่อยกระดับประสิทธิภาพระบบในระยะยาว

## เอกสารอ้างอิง

[1] Jacob Goldberger, Geoffrey E. Hinton, Sam T. Roweis และ Ruslan Salakhutdinov. (2005). “Neighbourhood Components Analysis.” Advances in Neural Information Processing Systems, Vol. 17, May 2005, pp. 513-520.

[2] P. Geurts, D. Ernst., และ L. Wehenkel. (2006). “Extremely randomized trees.”, Machine Learning, 63(1), 3-42, 2006.

[3] Chih-Chung Chang และ Chih-Jen Lin. (2022). “LIBSVM: A Library for Support Vector Machines.” Department of Computer Science National Taiwan University.

[4] Ron Kohavi. (2001). “A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection.” Intl. Jnt. Conf. AI on March 2001.

[5] Bergstra, J. และ Bengio, Y. (2012). “Random search for hyper-parameter optimization,” The Journal of Machine Learning Research 2012.

[6] กรมคุมประพฤติ กระทรวงยุติธรรมม, “เกณฑ์และแบบประเมินที่ใช้ในการจำแนกกลุ่มผู้กระทำผิด,” คู่มือการจำแนกผู้กระทำผิดในงานคุมประพฤติ, Vol.9, 2549.

# การพัฒนาโมเดลปัญญาประดิษฐ์แบบเอเจนติกเอไอ เพื่อสนับสนุนการประเมินก๊าซเรือนกระจกในระดับองค์กร
# Development of Agentic AI Models for Organizational Carbon Footprint Assessment and Governance

*โศจิรัตน์ ทรงประดิษฐ์ (Sojirat Songpradit)[1] และมหศักดิ์ เกตุฉ่ำ (Mahasak Ketcham)[2]*
*ภาควิชาการจัดการเทคโนโลยีสารสนเทศ*
*คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ*
*[1]s6707021857112@email.kmutnb.ac.th, [2] mahasak.k@itd.kmutnb.ac.th*

## บทคัดย่อ

*การเปลี่ยนแปลงสภาพภูมิอากาศจากก๊าซเรือนกระจกเป็นความท้าทายระดับโลก ทำให้องค์กรต้องวัด รายงาน และจัดการ Carbon Footprint อย่างเป็นระบบ แม้กระบวนการจะซับซ้อนจากข้อมูลจำนวนมาก การกำหนด Scope 1-3 และมาตรฐานที่หลากหลาย งานวิจัยนี้จึงพัฒนาระบบ Agentic AI แบบ Hybrid เพื่อสนับสนุนกระบวนการ MRV โดยผสาน LLM, Multi-Agent, RAG และ Deterministic Calculation สำหรับการตีความข้อมูล คำนวณ และอธิบายผลลัพธ์ ผลการประเมินพบว่าระบบมีความแม่นยำสูง (MAE = 0.00%) RAG ผ่านเกณฑ์ (Score = 0.69, Pass = 74%) ไม่มี Hallucination NL-to-Query = 79.3% และมีความโปร่งใสสูง (0.99) โดยสรุประบบช่วยลดภาระการจัดการข้อมูล เพิ่มความโปร่งใส รองรับการตรวจสอบ และมีศักยภาพในการประยุกต์ใช้และต่อยอดในอนาคต*

**คำสำคัญ:** ปัญญาประดิษฐ์, ปัญญาประดิษฐ์เชิงตัวแทน การปล่อยก๊าซเรือนกระจก, คาร์บอนฟุตพริ้นท์

## Abstract

*Climate change driven by greenhouse gas emissions is a major global challenge, requiring organizations to measure, report, and manage carbon footprints despite complexities in data, Scope 1–3 classification, and diverse standards. This study develops a hybrid Agentic AI system for MRV by integrating LLMs, multi-agent architecture, RAG, and deterministic calculation to support data interpretation, emission calculation, and explainable outputs. Results show MAE = 0.00%, RAG Score = 0.69 with 74% Pass Rate (no hallucination), NL-to-MongoDB success = 79.3%, and Transparency = 0.99, all meeting thresholds. The model achieves expert-level accuracy, improves auditability aligned with ISO 14064-1, reduces data burden, and shows strong potential for industrial use and future predictive carbon reduction.*

**Keywords:** Artificial Intelligence, Agentic AI, Greenhouse Gas Emissions, Carbon Footprint

## 1. บทนำ

การเปลี่ยนแปลงสภาพภูมิอากาศเป็นความเสี่ยงเชิงระบบที่ส่งผลกระทบต่อเศรษฐกิจ สิ่งแวดล้อม และความมั่นคงของมนุษย์ โดยรายงาน IPCC ชี้ว่าการเพิ่มขึ้นของอุณหภูมิโลกสัมพันธ์โดยตรงกับการสะสมก๊าซเรือนกระจกจากกิจกรรมมนุษย์ เช่น พลังงานฟอสซิล อุตสาหกรรม และการใช้ที่ดิน

ประเทศไทยได้เข้าร่วม Paris Agreement และตั้งเป้าหมาย Carbon Neutrality ภายในปี 2050 และ Net Zero ภายในปี 2065 ภายใต้ LT-LEDS โดยภาคธุรกิจมีบทบาทสำคัญผ่านการวัด รายงาน และบริหารจัดการการปล่อยก๊าซเรือนกระจกตามมาตรฐาน GHG Protocol, ISO 14064 และแนวทางของ TGO

อย่างไรก็ตาม การประเมิน Carbon Footprint ยังมีความซับซ้อนสูง เป็นปัญหาแบบหลายขั้นตอนที่ต้องผสานการวิเคราะห์ข้อมูลและการคำนวณ งานวิจัยก่อนหน้าพบว่า LLM เพียงอย่างเดียวมีข้อจำกัดด้านความแม่นยำและความเสี่ยงต่อ Hallucination ดังนั้น งานวิจัยนี้จึงเสนอ Hybrid Agentic AI ที่ผสาน Deterministic Calculation Engine กับ LLM เพื่อเพิ่มความถูกต้องและความน่าเชื่อถือของผลลัพธ์

## 2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ผู้วิจัยได้ทบทวนเอกสาร ทฤษฎี และงานวิจัยที่เกี่ยวข้อง เพื่อใช้เป็นพื้นฐานในการออกแบบและพัฒนา โดยมีรายละเอียดดังนี้

### 2.1 Carbon Footprint

Carbon Footprint คือปริมาณรวมของก๊าซเรือนกระจก (GHG) ที่ปล่อยทั้งทางตรงและทางอ้อม โดยรายงานในรูป $CO_2e$ ซึ่งคำนวณจากค่า Global Warming Potential (GWP) ตาม IPCC [10] ครอบคลุมก๊าซหลัก เช่น $CO_2$, $CH_4$, $N_2O$

การคำนวณใช้สมการพื้นฐาน E = AD × EF โดย Activity Data (AD) มาจากข้อมูลการดำเนินงาน และ Emission Factors (EF) อ้างอิงจากแหล่งมาตรฐาน เช่น IPCC และฐานข้อมูลที่เกี่ยวข้อง [11] การจัดการข้อมูลต้องรองรับแหล่งข้อมูลที่หลากหลายและมีขั้นตอน preprocessing เช่น unit conversion และ validation

การจำแนก Scope 1-3 ตาม GHG Protocol เป็นแนวทางหลักในการจัดหมวดหมู่การปล่อย โดยเฉพาะ Scope 3 ที่มีความซับซ้อนสูงในห่วงโซ่อุปทาน [8], [9] ขณะที่มาตรฐาน ISO 14064-1 และ GHG Protocol กำหนดกรอบ MRV ที่เน้นความโปร่งใสและการตรวจสอบย้อนกลับของข้อมูล [11]

### 2.2 แนวคิด Agentic AI

พัฒนาการของ AI เริ่มจาก Traditional AI ที่ใช้กฎเชิงสัญลักษณ์ [5] ต่อมาเป็น Machine Learning และ Deep Learning ที่เรียนรู้จากข้อมูลจำนวนมาก [4] และพัฒนาเป็น Large Language Models (LLMs) เช่น GPT-4 และ BERT ซึ่งมีความสามารถด้าน natural language understanding และ reasoning [1]–[3] อย่างไรก็ตาม LLM แบบ standalone ยังมีข้อจำกัดด้าน numerical accuracy และ hallucination

Agentic AI เป็นแนวคิดที่ขยายความสามารถของ LLM ให้สามารถทำงานแบบ multi-step โดยมีองค์ประกอบหลัก ได้แก่ Reasoning, Planning, Tool Use และ Memory ซึ่งช่วยให้ระบบสามารถตัดสินใจและดำเนินงานแบบ end-to-end ได้ [6] ในเชิงสถาปัตยกรรม มักพัฒนาในรูปแบบ Multi-Agent System (MAS) โดยแยก agent ตามหน้าที่ เช่น orchestration agent, retrieval agent และ interface agent เพื่อเพิ่ม modularity และ scalability [7]

การผสาน Retrieval-Augmented Generation (RAG) ใช้ vector database และ embedding models สำหรับ semantic retrieval เพื่อลด hallucination และเพิ่มความถูกต้องของข้อมูลอ้างอิง ขณะที่งานคำนวณถูกแยกไปยัง Deterministic Calculation Engine ซึ่งใช้ rule-based หรือ function-based logic เพื่อลด numerical error และเพิ่ม reliability

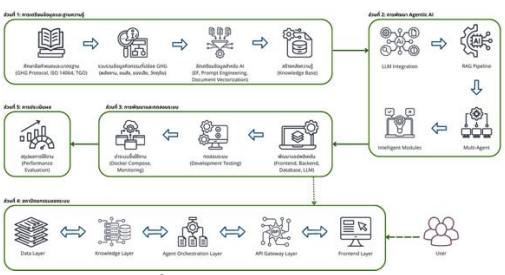
ในระดับระบบ การใช้สถาปัตยกรรม Microservices ช่วยให้แต่ละองค์ประกอบสามารถ deploy และ scale ได้อย่างอิสระ โดยเชื่อมต่อผ่าน API และรองรับ asynchronous processing ซึ่งเหมาะกับงาน MRV ที่ต้องรองรับทั้ง batch processing และ near real-time processing

โดยสรุป งานวิจัยสมัยใหม่ชี้ให้เห็นว่าการผสาน LLM, RAG, Multi-Agent และ deterministic modules เป็นแนวทางที่มีประสิทธิภาพในการแก้ปัญหาที่ต้องการทั้ง reasoning และ computational accuracy ซึ่งเป็นพื้นฐานสำคัญของระบบที่พัฒนาในงานวิจัยนี้ [6], [7]

## 3. วิธีการดำเนินงานวิจัย

กระบวนการดำเนินงานวิจัยแบ่งออกเป็น 5 ขั้นตอน ดังนี้

### 3.1 กรอบแนวความคิดของงานวิจัย



**ภาพที่ 1** : กรอบแนวความคิดงานวิจัย

ออกแบบภายใต้แนวคิด Goal-Driven Autonomous AI System ซึ่งให้ระบบสามารถตีความเป้าหมายของผู้ใช้งาน วางแผนการดำเนินงาน และประสานการทำงานของตัวแทน AI หลายองค์ประกอบแบบอัตโนมัติ

การออกแบบสถาปัตยกรรมของระบบยึดตามหลักการสำคัญ 4 ประการ ได้แก่ (1) Separation of Concerns แยกหน้าที่ของโมดูลเพื่อลดความซับซ้อนและเพิ่มความสามารถในการขยาย (2) Autonomous Multi-Agent Collaboration ใช้ตัว แทน AI เฉพาะทางเพื่อจัดการงานที่มีลักษณะแตกต่างกัน (3) Deterministic-AI Hybrid Reasoning ผสานการคำนวณเชิงกำหนดกับการให้เหตุผลของโมเดลภาษา (4) MRV Compliance รองรับความต้องการด้าน Measurement, Reporting and Verification

ระบบถูกออกแบบให้ทำงานในรูปแบบ Layered Architecture ร่วมกับ Microservices Architecture เพื่อรองรับความซับซ้อนของกระบวนการบัญชีคาร์บอน การจัดการข้อมูลหลายรูปแบบ และการดำเนินงานของตัวแทน AI แบบ Distributed โดยสถาปัตยกรรมแบ่งออกเป็น 5 ชั้นหลัก ได้แก่ (1) Frontend Layer (2) API Gateway Layer (3) Agent Orchestration Layer (4) Knowledge & Semantic Processing Layer (5) Data Layer

**3.2 การเตรียมข้อมูลและฐานความรู้**

ออกแบบกระบวนการในลักษณะ Data-Centric AI Pipeline เพื่อควบคุมคุณภาพข้อมูลและรองรับการตรวจสอบย้อนกลับ ประกอบด้วย 5 ขั้นตอนหลัก

3.2.1 การรวบรวมข้อมูล Emission Factor: รวบรวม 55 รายการจาก 4 แหล่งมาตรฐาน ได้แก่ GHG Protocol, ISO 14064-1, ISO 14067 และฐานข้อมูล TGO (2567) ครอบคลุม Scope 1 (15 รายการ), Scope 2 (8 รายการ) และ Scope 3 (32 รายการ) ข้อมูลถูกจัดเก็บแบบเชิงโครงสร้างพร้อม Metadata เช่น หน่วยวัด แหล่งอ้างอิง Scope และปี นอกจากนี้พัฒนา Activity Alias Mapping 50 รายการ เพื่อเชื่อมโยงภาษาธรรมชาติกับกิจกรรมมาตรฐาน

3.2.2 การรวบรวมเอกสารองค์ความรู้และ OCR Pipeline: รวบรวมเอกสาร 12 ฉบับ (1,847 chunks) จากแหล่งที่เชื่อถือได้ เช่น GHG Protocol, IPCC AR6, TGO และ ISO 14064 โดยแปลง PDF ผ่าน OCR แบบ Asynchronous ด้วยโมเดล Typhoon Vision เพื่อสกัดข้อความ ตาราง และหน่วยวัด ก่อนทำ Text Normalization และจัดเก็บ Metadata เพื่อรองรับ traceability

3.2.3 Knowledge Extraction และ Vector Embedding: ข้อมูลถูกแบ่งเป็น 512 โทเคน พร้อม overlap 50 โทเคน เพื่อรักษาบริบท จากนั้นสกัดองค์ประกอบสำคัญ เช่น นิยาม สูตร และ Emission Factor แล้วแปลงเป็นเวกเตอร์ด้วยโมเดล Multilingual Embedding เพื่อรองรับ semantic retrieval ทั้งไทย–อังกฤษ และจัดเก็บใน Qdrant โดยใช้ ANN เพื่อให้ค้นคืนได้แบบรวดเร็ว (near real-time)

**3.3 การพัฒนา Agentic AI**

งานวิจัยนี้พัฒนา Agentic AI ภายใต้แนวคิด Goal-Based Execution ซึ่งแตกต่างจากระบบสนทนาแบบ Turn-Based โดยเปิดโอกาสให้ผู้ใช้กำหนดเป้าหมายในภาษาธรรมชาติ และให้ระบบดำเนินการวางแผน แยกงานย่อย และประสานการทำงานของหลาย Agent โดยอัตโนมัติ สถาปัตยกรรมหลักประกอบด้วย 5 Agent เฉพาะทาง ดังนี้

1) GoalExecutor (Agent Orchestrator): ควบคุมการทำงานทั้งหมด เริ่มจากวิเคราะห์ Goal ผ่าน AIQueryAnalyzer เพื่อจำแนก Intent และ GoalType จากนั้นแยกเป็น Tasks ย่อย มอบหมายให้ Agent เหมาะสม พร้อมกลไก Retry และ Fallback

2) AgenticCalculationAgent: คำนวณการปล่อย GHG โดยใช้ 3-Tier Cascade Lookup (Org-specific → Standard → Fallback) และสูตร Emissions = Activity × Emission Factor

3) AgenticKnowledgeAgent: ใช้ RAG Architecture โดยประกอบด้วย Semantic Retrieval จาก Qdrant และกลไก Fallback หลายระดับ

4) AgenticMongoQueryAgent: แปลงภาษาธรรมชาติเป็น MongoDB Query ครอบคลุม Simple Query, Aggregation Pipeline และ Time Range Parsing

5) DynamicCRUDAgent และ Expert Scenario Analyzer: รองรับการจัดการข้อมูล CRUD และการ

วิเคราะห์เชิงกลยุทธ์ (Financial Modeling, Portfolio Optimization, Forecasting)

ระบบใช้แนวคิด Multi-Model Orchestration โดยมี Typhoon เป็นโมเดลหลัก และ Fallback เช่น GPT-4o mini และ Claude เพื่อเพิ่ม Reliability พร้อม Adaptive Model Routing ตามความซับซ้อนของงาน

**3.4 สถาปัตยกรรมของระบบ**

ระบบถูกพัฒนาภายใต้ Containerized Microservices Deployment ด้วย Docker Compose ประกอบด้วย 6 บริการหลัก ได้แก่ Carbon API (Backend หลัก), MCP API (Embedding & Retrieval), MongoDB (ฐานข้อมูล Activity), Qdrant (Vector Database), Redis (Task Queue) และ MinIO (Object Storage) โดยสื่อสารผ่าน Docker Network แบบ Isolated เพื่อเพิ่มความปลอดภัยและแยกการทำงานของแต่ละ service

รองรับการประมวลผลแบบ Asynchronous ผ่าน Celery และ Redis สำหรับงานที่ใช้ทรัพยากรสูง เช่น OCR Processing และการสร้าง Emission Report ช่วยลดภาระของ API หลักและเพิ่มประสิทธิภาพการประมวลผลแบบ concurrent

ระบบมีการออกแบบ Health Check, Logging, Metrics และ Alert Monitoring เพื่อรองรับการติดตามสถานะและแก้ไขปัญหา รวมถึงกำหนด Startup Sequence เพื่อควบคุมลำดับการเริ่มต้นบริการและลดความเสี่ยงในการเชื่อมต่อระหว่าง service ในสภาพแวดล้อมจริง

**3.5 การออกแบบการทดสอบและประเมินผล**

ระบบถูกประเมินใน 4 มิติหลัก โดยแต่ละมิติมีชุดข้อมูลทดสอบและเกณฑ์การผ่านที่กำหนดไว้ชัดเจน

3.5.1 ชุดข้อมูลทดสอบความแม่นยำการคำนวณ GHG (Calculation Accuracy): จัดเตรียมชุดทดสอบ 50 กรณี ครอบคลุม Scope 1 (การใช้น้ำมันเชื้อเพลิง ก๊าซธรรมชาติ และสารทำความเย็น รวม 15 กรณี), Scope 2 (การใช้ไฟฟ้าและพลังงานความร้อน 15 กรณี) และ Scope 3 (การขนส่ง การเดินทาง ขยะ และน้ำ รวม 20 กรณี) แต่ละกรณีประกอบด้วย Activity Data (ปริมาณกิจกรรม) Emission Factor อ้างอิง (จาก TGO ปี 2567) และค่า Benchmark ที่คำนวณโดยผู้เชี่ยวชาญ Carbon Accounting ตัวชี้วัด: Mean Absolute Error (MAE) เกณฑ์การผ่าน: MAE $\leq$ 5%

3.5.2 ชุดข้อมูลทดสอบคุณภาพ RAG (Reliability): จัดเตรียมคำถาม 50 ข้อ ครอบคลุม 5 หมวดหมู่ ได้แก่ (1) Emission Factor Queries ถามค่า EF เฉพาะเจาะจง 10 ข้อ (2) GHG Protocol ถามหลักการและการจำแนก Scope 10 ข้อ (3) TGO Guidelines ถามแนวปฏิบัติของประเทศไทย 10 ข้อ (4) Reporting Standards ถามกระบวนการรายงาน 10 ข้อ และ (5) Advanced Topics ถามหัวข้อขั้นสูง เช่น Carbon Offset และ Net Zero Strategies 10 ข้อ คำตอบได้รับการประเมินโดยผู้เชี่ยวชาญด้าน Carbon Accounting 3 ท่าน ใน 3 มิติ: Correctness (ความถูกต้องของเนื้อหา), Completeness (ความครบถ้วน) และ Citation (การอ้างอิงแหล่งที่มา) คะแนนรวม 0.0-1.0 เกณฑ์: Mean Score $\geq$ 0.5 และ Pass Rate $\geq$ 70%

3.5.3 ชุดข้อมูลทดสอบ NL-to-MongoDB Query (Usability): จัดเตรียม 30 Natural Language Queries ครอบคลุม 3 ระดับความซับซ้อน ได้แก่ (1) Simple Queries ดึงข้อมูล Activity แบบตรง 10 คำถาม เช่น 'แสดงการใช้ไฟฟ้าเดือนมกราคม' (2) Aggregation Queries รวมข้อมูลหลายช่วงเวลาหรือหลาย Scope 10 คำถาม เช่น 'สรุปการปล่อย GHG ปี 2566 แยกตาม Scope' และ (3) Complex Queries คำถามที่ต้องการ Multi-collection Join หรือ Conditional Aggregation 10 คำถาม เกณฑ์: Query Success Rate $\geq$ 70% โดยประเมินจากความตรงกันของ Query ที่สร้างกับ Expected Query

3.5.4 ชุดข้อมูลทดสอบความโปร่งใส (Explainability): ประเมินความสามารถ MRV Compliance ของระบบในด้านการแสดง Emission Breakdown ครบถ้วน การแสดงสูตรและขั้นตอนการคำนวณ การอ้างอิงแหล่งที่มาของ Emission Factor

## 4. ผลการดำเนินงานวิจัย

**4.1 ผลการประเมินความถูกต้องของการคำนวณ (Accuracy Evaluation)**

ทดสอบด้วย 50 กรณี ครอบคลุม Scope 1-3 โดยใช้ MAE เป็นตัวชี้วัด

**ตารางที่ 1:** สรุปผลการประเมินความถูกต้องของการคำนวณ

| มิติการทดสอบ | ผลลัพธ์ | เกณฑ์ | สถานะ |
|---|---|---|---|
| MAE (ทุก Scope) | 0.00% | $\leq$ 5% | ผ่าน |
| Accuracy Rate | 100% | $\geq$ 95% | ผ่าน |
| Alias Mapping | 100% | $\geq$ 90% | ผ่าน |
| EF Selection Accuracy | 100% | $\geq$ 95% | ผ่าน |

ปัจจัยสำคัญที่ทำให้ได้ผลลัพธ์สมบูรณ์ ได้แก่ การใช้ 3 - Tier Cascade Lookup (Org-specific → Standard → Fallback) การแก้ปัญหา Alias Matching ครบถ้วนทุกกิจกรรม และการใช้ค่า Emission Factor ตามมาตรฐาน TGO ปี 2567

**4.2 ผลการประเมินคุณภาพของระบบ RAG (Reliability Evaluation)**

ประเมินด้วยคำถาม 50 ข้อ ใน 5 หมวดหมู่ โดยผู้เชี่ยวชาญ 3 ท่าน ใน 3 มิติ (Correctness, Completeness, Citation)

**ตารางที่ 2:** สรุปผลการประเมินคุณภาพของระบบ RAG

| หมวดหมู่ | Mean Score | Pass Rate | เกณฑ์ | สถานะ |
|---|---|---|---|---|
| Emission Factor | 1.00 | 100% | $\geq$ 0.5 / $\geq$ 70% | ผ่าน |
| GHG Protocol | 0.82 | 90% | $\geq$ 0.5 / $\geq$ 70% | ผ่าน |
| TGO Guidelines | 0.78 | 80% | $\geq$ 0.5 / $\geq$ 70% | ผ่าน |
| Reporting | 0.55 | 40% | $\geq$ 0.5 / $\geq$ 70% | บางส่วน |
| Advanced Topics | 0.42 | 20% | $\geq$ 0.5 / $\geq$ 70% | ยังไม่ผ่าน |
| **รวม (Overall)** | **0.69** | **74%** | **$\geq$ 0.5 / $\geq$ 70%** | **ผ่าน** |

จุดแข็ง: EF Queries และ GHG Protocol ได้ผลดีเนื่องจากมีฐานความรู้ครอบคลุม จุดอ่อน: Advanced Topics และ Reporting Standards ยังขาดข้อมูลในฐานความรู้ ระบบไม่เกิด Hallucination เลย (0%) เนื่องจากกลไก 3-Level RAG Fallback

**4.3 ผลการประเมินความสามารถในการใช้งานของระบบ (Usability Evaluation)**

ทดสอบ 30 Natural Language Queries ใน 3 ระดับความซับซ้อน

**ตารางที่ 3:** สรุปผลการประเมินความสามารถในการใช้งานของระบบ

| ประเภท Query | จำนวน | Success Rate | สถานะ |
|---|---|---|---|
| Simple Queries | 10 | 100% | ผ่าน |
| Aggregation Queries | 10 | 90% | ผ่าน |
| Complex Queries | 10 | 50% | ยังไม่ผ่าน |
| **รวม (Overall)** | **30** | **79.3%** | **ผ่าน ($\geq$ 70%)** |

Time Parsing และ Scope Detection ได้ผล 100% แต่ Complex Aggregation ยังมีข้อจำกัดเมื่อคำถามกำกวมหรือต้องการ Multi-collection Join

**4.4 ผลการประเมินความโปร่งใสของระบบ (Explainability Evaluation)**

ระบบได้รับ Transparency Score = 0.99 (เกณฑ์ $\geq$ 0.80) ซึ่งแสดงว่าระบบสามารถแสดง Emission Breakdown ครบถ้วน แสดงสูตรและขั้นตอนการคำนวณ อ้างอิงแหล่ง Emission Factor 100% และมี Audit Trail ครบถ้วนตามข้อกำหนด ISO 14064-1

**4.5 ผลการประเมินด้านประสิทธิภาพของระบบ (Performance Evaluation)**

**ตารางที่ 4:** สรุปผลการประเมินด้านประสิทธิภาพของระบบ

| การทดสอบ | Response Time | เกณฑ์ | สถานะ |
|---|---|---|---|
| GHG Calculation (P50) | 19 ms | < 100 ms | ผ่าน |
| RAG Query (P95) | 1.15 s | < 3 s (หรือ 5 s) | ผ่าน |
| MongoDB Query (P95) | 4.95 s | < 10 s | ผ่าน |
| Concurrent 5 Users | Success 100% | ไม่ Crash | ผ่าน |

Bottleneck หลักคือ LLM Inference ใน RAG Pipeline และ MongoDB Aggregation สำหรับชุดข้อมูลขนาดใหญ่

**4.6 ผลการประเมินความถูกต้องของการแปลงภาษาธรรมชาติเป็นคำสั่งฐานข้อมูล และความสำเร็จของ API (Functional Evaluation)**

ผลการประเมินพบว่าการแปลงภาษาธรรมชาติเป็น MongoDB Query มีความแม่นยำ 79.3% และ API Success Rate เท่ากับ 100% สะท้อนถึงความสามารถของระบบในการประสานการทำงานระหว่างองค์ประกอบต่าง ๆ ได้อย่างมีประสิทธิภาพ

**4.7 ผลการประเมินความครบถ้วนของฟังก์ชัน (Completeness Evaluation)**

ระบบสามารถรองรับฟังก์ชันหลักได้มากกว่า 90% และมีการทำงานของ agents ครบถ้วน (7/7) แสดงถึงความสมบูรณ์ของระบบและความพร้อมในการนำไปใช้งานจริง

## 5. สรุปผล อภิปรายผล และข้อเสนอแนะ

### 5.1 สรุปผลการวิจัย

งานวิจัยนี้มุ่งศึกษามาตรฐาน Carbon Footprint พัฒนา Agentic AI และประเมินประสิทธิภาพระบบ โดยพบว่า

5.1.1 การประเมิน Carbon Footprint ต้องอ้างอิงมาตรฐานสากลและไทย เช่น GHG Protocol และ ISO 14064 ครอบคลุม Scope 1-3 การเก็บข้อมูล และการรายงาน ซึ่งถูกนำไปใช้เป็นฐานความรู้ในระบบ RAG

5.1.2 ระบบ Agentic AI ถูกออกแบบแบบ Microservices และ Multi-Agent ผสาน LLM, OCR และ Deterministic Engine เพื่อคำนวณและวิเคราะห์อย่างอัตโนมัติ

5.1.3 ผลการประเมินแสดงให้เห็นว่าระบบมีความแม่นยำ น่าเชื่อถือ ใช้งานได้จริง โปร่งใส และมีประสิทธิภาพด้านเวลา ความเสถียร และความครบถ้วน เหมาะต่อการใช้งานในองค์กร

### 5.2 อภิปรายผลการวิจัย

การผสานมาตรฐาน Carbon Footprint กับ Agentic AI ช่วยเพิ่มความถูกต้องและความน่าเชื่อถือ สอดคล้องหลัก MRV และเสริม Explainability

สถาปัตยกรรมแบบ Multi-Agent ร่วมกับการแยก Deterministic Calculation ช่วยเพิ่ม Accuracy และ Performance และลดข้อผิดพลาดจาก LLM ขณะเดียวกันระบบมี Usability และ Functional ที่ดี แต่ยังมีข้อจำกัดในงานที่ซับซ้อน

ในภาพรวม ระบบมีความเสถียร ตอบสนองใกล้ Real-Time และมีความครบถ้วนของฟังก์ชัน แต่ยังขึ้นอยู่กับคุณภาพฐานความรู้และความซับซ้อนของโจทย์

### 5.3 ข้อเสนอแนะ

ควรพัฒนาฐานความรู้ให้ครอบคลุมมากขึ้น โดยเฉพาะ Scope 3 และกรณีเฉพาะทาง ลดข้อจำกัดของ LLM และเพิ่มขนาดการทดสอบผู้ใช้

ในอนาคตควรพัฒนา Local LLM ลดต้นทุน ปรับปรุง NL-to-Query สำหรับงานซับซ้อน ทดสอบในองค์กรจริง เพื่อประเมิน scalability และ usability รวมถึงพัฒนา Explainability และเชื่อมต่อกับระบบ ERP, IoT และ Energy Systems เพื่อเพิ่มประสิทธิภาพการวิเคราะห์และลดการปล่อยก๊าซเรือนกระจก

## เอกสารอ้างอิง


[1] T. B. Brown et al., “Language models are few-shot learners,” Advances in Neural Information Processing Systems (NeurIPS), vol. 33, pp. 1877–1901, 2020.

[2] J. Devlin, M. W. Chang, K. Lee, and K. Toutanova, “BERT: Pre-training of deep bidirectional transformers for language understanding,” Proc. NAACL-HLT, pp. 4171–4186, 2019.

[3] OpenAI, “GPT-4 technical report,” OpenAI, 2023.

[4] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” Nature, vol. 521, no. 7553, pp. 436–444, 2015.

[5] S. Russell and P. Norvig, Artificial Intelligence: A Modern Approach, 4th ed., Pearson, 2021.

[6] D. B. Acharya, K. Kuppan, and B. Divya, “Agentic AI: Autonomous intelligence for complex goals—A comprehensive survey,” IEEE Access, vol. 13, pp. 18912–18936, 2025.

[7] J. S. Park et al., “Generative agents: Interactive simulacra of human behavior,” Proc. ACM UIST, pp. 1–15, 2023.

[8] World Resources Institute (WRI) and World Business Council for Sustainable Development (WBCSD), “The greenhouse gas protocol: A corporate accounting and reporting standard,” 2015.

[9] World Resources Institute (WRI) and World Business Council for Sustainable Development (WBCSD), “Corporate value chain (Scope 3) accounting and reporting standard,” 2011.

[10] Intergovernmental Panel on Climate Change (IPCC), “2006 IPCC guidelines for national greenhouse gas inventories (2019 refinement),” 2019.

[11] ISO, “ISO 14064-1: Greenhouse gases—Part 1: Specification with guidance at the organization level,” 2018.

# การจำแนกระดับความซับซ้อนของงานพัฒนาซอฟต์แวร์จากเอกสารความต้องการด้วยการประมวลผลภาษาธรรมชาติและปัญญาประดิษฐ์เชิงรู้สร้าง

# Classification of Software Development Task Complexity Levels from Requirement Documents Using Natural Language Processing and Generative AI

*วิทยเดช ใจกว้าง (Wittayadet Jaikwang)[1] และพยุง มีสัจ (Phayung Meesad)[2]*
*ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล*
*มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ*
*[1]s6707021857309@email.kmutnb.ac.th, [2]phayung.m@itd.kmutnb.ac.th*

## บทคัดย่อ

*งานวิจัยนี้นำเสนอแบบจำลองจำแนกระดับความพยายาม (Man-Day) ในการพัฒนาซอฟต์แวร์เป็น 3 ระดับ โดยประยุกต์ใช้เทคนิค NLP และ LLMs (Gemma) ร่วมกับ Machine Learning จากการทดสอบด้วยชุดข้อมูลโครงการจริง 419 รายการ พบว่าโมเดล WangchanBERTa มีประสิทธิภาพสูงสุด โดยให้ค่าความแม่นยำรวม (Accuracy) ที่ร้อยละ 67.46 และแยกแยะงานระดับยากได้อย่างโดดเด่น (AUC 0.88) อย่างไรก็ตาม ยังพบปัญหาการประเมินงานระดับปานกลางต่ำกว่าความเป็นจริง เนื่องจากมีบริบทระบบซ่อนเร้นในเอกสาร สรุปได้ว่าระบบที่พัฒนาขึ้นนี้มีศักยภาพในการเป็นเครื่องมือสนับสนุนการพิจารณาความสมเหตุสมผลของค่า Man-Day เบื้องต้นได้อย่างเป็นรูปธรรม*

**คำสำคัญ:** การประมวลผลภาษาธรรมชาติ โมเดลภาษาขนาดใหญ่ ปัญญาประดิษฐ์

## Abstract

*This research presents a model for classifying software development man-day requirements into three levels, applying NLP and LLMs (Gemma) techniques combined with Machine Learning. Testing with a dataset of 419 real-world projects showed that the WangchanBERTa model performed best, achieving an overall accuracy of 67.46% and significantly differentiating difficult tasks (AUC 0.88). However, a problem of underestimating medium-level tasks was observed due to hidden system context in the documentation. In conclusion, the developed system has the potential to be a concrete tool for supporting preliminary assessments of man-day values.*

**Keywords:** Natural Language Processing, Large Language Models, Artificial Intelligence.

## 1. บทนำ

การประเมินความพยายามในการพัฒนาซอฟต์แวร์เป็นกระบวนการที่สำคัญต่อการวางแผนทรัพยากร งบประมาณ และระยะเวลาของโครงการ โดยเฉพาะอย่างยิ่งในกรณีการจัดจ้างผู้ให้บริการภายนอก (Outsourcing) ซึ่งผู้พัฒนาส่วนใหญ่มักได้รับรายละเอียดความต้องการในรูปแบบข้อความที่ไม่มีโครงสร้างชัดเจน ปัญหาวิจัยที่สำคัญคือ การประเมินงานในปัจจุบันมักต้องพึ่งพาดุลยพินิจและประสบการณ์ของผู้เชี่ยวชาญแบบดั้งเดิม ซึ่งนอกจากจะมีความล่าช้าแล้ว ยังมักมีความลำเอียง (Bias) และเสี่ยงต่อความคลาดเคลื่อนสูงอันเกิดจากเอกสารความต้องการที่เขียนอย่างคลุมเครือ ส่งผลให้กระบวนการตรวจสอบมีความซับซ้อนและเกิดผลกระทบต่อแผนงานโดยรวม เพื่อแก้ไขปัญหาความคลาดเคลื่อนดังกล่าว งานวิจัยนี้จึงมุ่งเปลี่ยนกรอบแนวคิดจากการประเมินแบบพยากรณ์ค่าตัวเลขต่อเนื่องซึ่งมักมีความคลาดเคลื่อนสูงเมื่อข้อมูลตั้งต้นไม่ชัดเจน มาเป็นการประเมินแบบจำแนกระดับความพยายามโดยอ้างอิงจากงานวิจัยของ Bakir et al. [1] ที่ยืนยันว่าการเปลี่ยนมุมมองมาเป็นการจัดกลุ่มความ

ยากง่ายของงาน จะช่วยลดความเสี่ยงจากการระบุตัวเลขที่ผิดพลาด ทำให้โมเดลมีความแม่นยำสูงขึ้น และเป็นรูปแบบที่เหมาะสมกับการนำไปใช้สนับสนุนการตัดสินใจจริงมากกว่า นอกจากนี้ เพื่อเพิ่มความเสถียรและยกระดับประสิทธิภาพของระบบ งานวิจัยได้บูรณาการเทคนิคการเรียนรู้แบบกลุ่ม (Ensemble Learning) เข้ามาประยุกต์ใช้ร่วมด้วย ซึ่งสอดคล้องกับข้อค้นพบของ Pospieszny et al.[2] ที่พิสูจน์แล้วว่าการทำงานร่วมกันของหลายอัลกอริทึมสามารถช่วยชดเชยจุดอ่อนของโมเดลเดี่ยว และลดความแปรปรวนในการพยากรณ์ได้อย่างมีนัยสำคัญ นอกจากช่องว่างด้านเทคนิคแล้ว ข้อจำกัดสำคัญของงานวิจัยด้าน NLP สำหรับวิศวกรรมซอฟต์แวร์ในประเทศไทย คือเอกสารความต้องการมักมีลักษณะการสลับรหัสภาษาระหว่างไวยากรณ์ภาษาไทยและคำศัพท์เทคนิคภาษาอังกฤษ การนำโมเดลพหุภาษาทั่วไป (Multilingual Models) มาใช้มักไม่สามารถทำความเข้าใจโครงสร้างและบริบทเชิงลึกของภาษาไทยได้อย่างสมบูรณ์ ดังนั้นงานวิจัยนี้จึงเน้นย้ำความจำเป็นในการนำโมเดล WangchanBERTa[3] ซึ่งเป็นโมเดลที่ผ่านการฝึกสอนด้วยคลังข้อมูลภาษาไทยขนาดใหญ่มาประยุกต์ใช้ เพื่อเป็นนวัตกรรมที่สามารถทำความเข้าใจบริบทของข้อมูลในท้องถิ่นได้อย่างแม่นยำและตอบโจทย์ได้ดีกว่าโมเดลพหุภาษา ด้วยเหตุนี้ งานวิจัยจึงมีวัตถุประสงค์เพื่อพัฒนาระบบอัตโนมัติในการจำแนกระดับความซับซ้อนของงานโดยนำข้อดีของปัญญาประดิษฐ์เชิงรู้สร้าง Gemma-3n [4] มาช่วยสกัดคุณลักษณะและอธิบายบริบทที่ซ่อนเร้นร่วมกับการแปลงเวกเตอร์ด้วยโมเดลภาษาไทย WangchanBERTa[3] และโมเดลพหุภาษา XLM-RoBERTa[5] เพื่อเปรียบเทียบประสิทธิภาพ และสร้างเป็นเครื่องมือที่สามารถสนับสนุนการประเมินความเหมาะสมของค่า Man-Day ได้อย่างมีประสิทธิภาพและน่าเชื่อถือ

## 2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

การวิจัยครั้งนี้ได้ทบทวนวรรณกรรมและทฤษฎีที่เกี่ยวข้อง เพื่อเป็นรากฐานในการออกแบบและพัฒนาระบบจำแนกระดับความพยายามในการพัฒนาซอฟต์แวร์โดยครอบคลุม 5 ประเด็นหลัก ดังนี้

### 2.1 การจัดการข้อมูลไม่สมดุลด้วยการจำแนกกลุ่ม

การประเมินความพยายามในอดีตมักใช้การพยากรณ์ค่าต่อเนื่อง (Regression) ซึ่งข้อมูลมักมีการกระจายตัวแบบเบ้ ส่งผลให้อัลกอริทึมเกิดความเอนเอียงเข้าหาข้อมูลส่วนใหญ่ งานวิจัยนี้จึงเปลี่ยนมาใช้ การจำแนกกลุ่ม (Classification) 3 ระดับ ซึ่งในทางวิชาการช่วยสร้างขอบเขตการตัดสินใจ ที่ชัดเจนขึ้น ลดผลกระทบจากความไม่สมดุลของข้อมูล (Data Imbalance) และทำให้โมเดลสามารถเรียนรู้กลุ่มงานส่วนน้อยได้อย่างมีประสิทธิภาพมากขึ้น[1]

### 2.2 การสกัดคุณลักษณะด้วย LLMs และสถาปัตยกรรมตัวแทนข้อความ

เอกสารความต้องการมักรวบรัดและคลุมเครือ งานวิจัยจึงใช้ Gemma-3n[4] สกัดบริบทเชิงเทคนิคที่ซ่อนเร้นก่อนนำไปแปลงเป็นเวกเตอร์ด้วย WangchanBERTa[3] ซึ่งเป็นทรานส์ฟอร์เมอร์ภาษาไทยที่สามารถจัดการปัญหาการสลับรหัสภาษา และเข้าใจบริบทได้ลึกซึ้งกว่าโมเดลพหุภาษา XLM-RoBERTa[5]

### 2.3 การเรียนรู้แบบกลุ่ม Ensemble Learning และ Soft Voting

เพื่อชดเชยจุดอ่อนของโมเดลเดี่ยว งานวิจัยประยุกต์ใช้ Ensemble Learning [2] โดยเลือกการโหวตแบบ Soft Voting หาค่าเฉลี่ยความน่าจะเป็น ซึ่งมีข้อดีกว่า Hard Voting นับเสียงข้างมากแบบเด็ดขาด ตรงที่สามารถนำ “ระดับความมั่นใจ” ของโมเดลมาถ่วงน้ำหนักร่วมกันได้ ทำให้แม่นยำขึ้นเมื่อเผชิญกับข้อความที่คลุมเครือ

### 2.4 มาตรวัดประสิทธิภาพ

การประเมินชุดข้อมูลที่ไม่สมดุลต้องใช้หลายมาตรวัดควบคู่กัน

Accuracy ดูความแม่นยำรวมเบื้องต้น

Precision ดูความน่าเชื่อถือ

Recall ดูความครอบคลุม

F1-Score เป็นค่าเฉลี่ยฮาร์มอนิกที่รักษาสมดุลระหว่าง Precision และ Recall ซึ่งสำคัญมากในการป้องกันปัญหาโมเดลประเมินงานต่ำกว่าจริงที่ส่งผลเสียต่อต้นทุนโครงการ

**2.5 ข้อดี ข้อเสียของงานวิจัยที่เกี่ยวข้องและช่องว่างทางเทคนิค**

จากการทบทวนวรรณกรรม สามารถวิเคราะห์ข้อดีข้อเสียเชิงเทคนิคได้ดังนี้

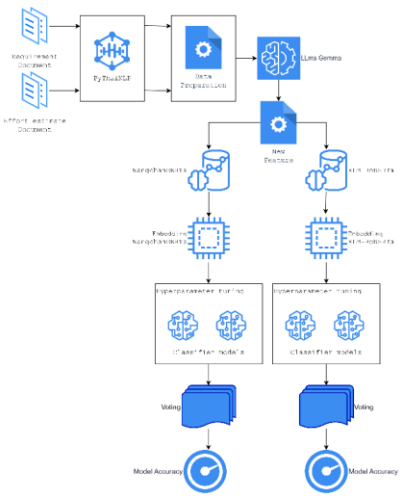
แบบจำลองสถิติเชิงคณิตศาสตร์[6] ข้อดี ให้ผลลัพธ์เป็นตัวเลขชัดเจน ข้อเสีย อ่อนไหวต่อความคลุมเครือและไม่สามารถประมวลผลข้อความดิบได้โดยตรง

การจำแนกด้วย NLP พื้นฐาน เช่น TF-IDF[1] ข้อดี ประมวลผลรวดเร็วและใช้ทรัพยากรน้อย ข้อเสีย ขาดความเข้าใจบริบทแวดล้อม และไม่สามารถแก้ปัญหาคำพ้องความหมายได้

การอุดช่องว่างงานวิจัย งานวิจัยนี้ปิดช่องว่างดังกล่าวโดยการบูรณาการเทคโนโลยี LLMs เพื่อสกัดความหมายเชิงลึก ผสานกับการแปลงข้อความเป็นเวกเตอร์ด้วยทรานส์ฟอร์เมอร์ภาษาไทย เพื่อยกระดับความแม่นยำในการจำแนกบริบทซอฟต์แวร์ท้องถิ่นให้เหนือกว่าวิธีการดั้งเดิมได้อย่างเป็นรูปธรรม

## 3. วิธีการดำเนินการวิจัย

การวิจัยนี้ใช้ระเบียบวิธีวิจัยเชิงทดลอง โดยออกแบบสถาปัตยกรรมระบบแบ่งเป็น 2 ส่วนหลัก คือ ส่วนประมวลผลด้วยปัญญาประดิษฐ์เชิงรู้สร้างเพื่อสกัดคุณลักษณะ และส่วนการจำแนกประเภทด้วยการเรียนรู้ของเครื่อง ดังแสดงในภาพที่ 1



**ภาพที่ 1** สถาปัตยกรรมการเรียนรู้ของเครื่องสำหรับการจำแนกข้อความ

**3.1 การรวบรวมและเตรียมข้อมูล**

ข้อมูลที่ใช้รวบรวมจากโครงการพัฒนาระบบสารสนเทศในอดีตจำนวน 22 โครงการ ประกอบด้วยเอกสารความต้องการ และตารางประเมินระยะเวลาทำงาน ชุดข้อมูลประกอบด้วย 419 รายการโดยในไฟล์ CSV มีคอลัมน์สำคัญคือ “รายละเอียดการแก้ไข” และ “ระยะเวลาประเมินการแก้ไข” การจัดการความไม่สมดุลของข้อมูลผู้วิจัยกำหนดเกณฑ์การแบ่งกลุ่มระดับความพยายามเป็น 3 ระดับ ดังตารางที่ 1 โดยอ้างอิงจากการกระจายตัวทางสถิติของค่า Man-Day ในโครงการอดีตขององค์กรกรณีศึกษาเพื่อสร้างความสมดุลของจำนวนข้อมูลในแต่ละคลาส

**ตารางที่ 1** สัดส่วนของการแบ่งกลุ่มความพยายามในการพัฒนาซอฟต์แวร์

| กลุ่มเป้าหมาย | เกณฑ์การแบ่ง |
|---|---|
| กลุ่ม 0 (งานระดับง่าย) | MD < 1 วัน |
| กลุ่ม 1 (งานระดับปานกลาง) | $1 \leq$ MD < 2 วัน |
| กลุ่ม 2 (งานระดับยาก) | MD $\geq$ 2 วัน |

**3.2 การสกัดคุณลักษณะด้วยโมเดลพหุรูปแบบ**

ผู้วิจัยใช้โมเดล Gemma-3n ทำหน้าที่เป็นนักวิเคราะห์ระบบอัตโนมัติ เพื่อขยายบริบทจากเอกสารความต้องการต้นฉบับ โดยมีเทคนิคสำคัญที่เพิ่มเข้ามาคือ การดึงบริบทซ่อนเร้น กำหนดให้โมเดลวิเคราะห์ข้อมูลที่ไม่ได้ระบุชัดเจนในเอกสาร เช่น ความซับซ้อนของฐานข้อมูลหรือระบบความปลอดภัย เพื่อลดข้อผิดพลาดในการประเมินงาน

**3.3 การแปลงข้อความเป็นเวกเตอร์และการทดลอง**

ระบบนำข้อความต้นฉบับรวมกับบทวิเคราะห์จาก Gemma-3n คั่นด้วย [SEP] มาแปลงเป็นเวกเตอร์ โดยแบ่งการทดลองเป็น 3 กรณีเพื่อเปรียบเทียบประสิทธิภาพ

WangchanBERTa โมเดลภาษาไทยเฉพาะทางเพื่อจับบริบทเชิงลึก

XLM-RoBERTa โมเดลพหุภาษาเพื่อศึกษาการจัดการข้อความสลับภาษา

TF-IDF Baseline ใช้เทคนิคดั้งเดิมร่วมกับตัวจำแนกพื้นฐาน เพื่อพิสูจน์ความคุ้มค่าของเทคโนโลยี Deep Learning ที่นำเสนอ

**3.4 การสร้างแบบจำลองและการประเมินผล**

ชุดข้อมูลถูกแบ่งเป็นส่วนฝึกสอนร้อยละ 70 และทดสอบร้อยละ 30 ด้วยวิธี Stratified Sampling จากนั้นใช้ 5 อัลกอริทึม XGBoost, Random Forest, SVC, Extra Trees, KNN ปรับจูนพารามิเตอร์ด้วย Optuna และรวมผลด้วยเทคนิค Soft Voting Ensemble เพื่อความเสถียรในการตัดสินใจ

ผู้วิจัยประเมินผลด้วยมาตรวัด Accuracy, Precision, Recall และ F1-Score ซึ่งความจำเป็นในการใช้หลายมาตรวัดควบคู่กันเนื่องจากในงานวิศวกรรมซอฟต์แวร์ผลกระทบจากการประเมินงานต่ำกว่าจริง และสูงกว่าจริงมีความสำคัญต่างกัน มาตรวัดเหล่านี้จึงช่วยสะท้อนประสิทธิภาพของระบบในมิติของความน่าเชื่อถือและความครอบคลุมได้ครบถ้วนกว่าการใช้ค่าความแม่นยำรวมเพียงอย่างเดียว

## 4. ผลการดำเนินงาน

การวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาระบบอัตโนมัติในการจัดกลุ่มความยากง่ายของงาน Man-Day จากเอกสารความต้องการ โดยนำเสนอผลการประเมินประสิทธิภาพของแบบจำลองตามลำดับดังนี้

**4.1 ผลการสำรวจและวิเคราะห์ข้อมูลเบื้องต้น**

จากการสำรวจชุดข้อมูลประเมินจำนวน 419 รายการ พบว่าข้อมูลมีความสมบูรณ์ร้อยละ 100 ภายหลังการทำความสะอาดและคัดกรองข้อมูลผิดปกติ การจัดกลุ่มระดับความพยายามเป็น 3 ระดับช่วยให้การกระจายตัวของข้อมูลมีความสมดุล เอื้อต่อการฝึกแบบจำลองและช่วยลดความเอนเอียง (Bias) ในการเรียนรู้

**4.2 ผลการประเมินและเปรียบเทียบประสิทธิภาพแบบจำลอง**

ผู้วิจัยได้นำชุดข้อมูลเข้าสู่กระบวนการทดสอบไขว้ (Cross-Validation) เพื่อหาค่าพารามิเตอร์ที่ดีที่สุดด้วย Optuna ก่อนนำไปประเมินกับชุดข้อมูลทดสอบ (Test Set) ด้วยแบบจำลองรวม (Soft Voting) ผลการเปรียบเทียบระหว่างโมเดลการแปลงเวกเตอร์ข้อความทั้ง 2 รูปแบบแสดงดังตารางที่ 2

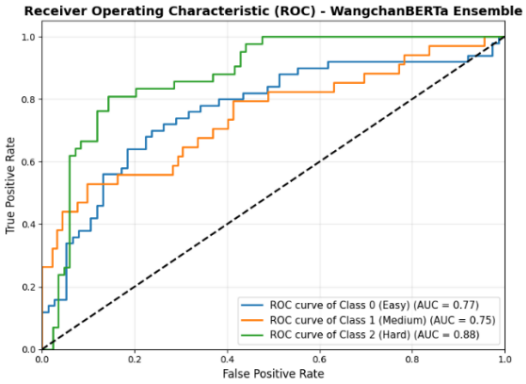
**ตารางที่ 2** การเปรียบเทียบประสิทธิภาพรวมระหว่างโมเดลภาษา 2 รูปแบบ

| ตัวชี้วัดประสิทธิภาพ | WangchanBERTa | XLM-RoBERTa | TF-IDF |
|---|---|---|---|
| Accuracy (ความแม่นยำรวม) | 67.46% | 57.14% | 64.17% |
| F1-Score: กลุ่ม 0 (งานระดับง่าย) | 0.68 | 0.62 | 0.68 |
| F1-Score: กลุ่ม 1 (งานระดับปานกลาง) | 0.51 | 0.30 | 0.44 |
| F1-Score: กลุ่ม 2 (งานระดับยาก) | 0.76 | 0.64 | 0.72 |
| ค่าเฉลี่ยแบบ Macro F1-Score | 0.65 | 0.52 | 0.62 |

จากตารางที่ 2 พบว่า สถาปัตยกรรมที่ใช้ WangchanBERTa ให้ประสิทธิภาพสูงสุด ร้อยละ 67.46 เอาชนะโมเดลพหุภาษา XLM-RoBERTa และวิธีฐานอย่าง TF-IDF ได้อย่างชัดเจน ผลลัพธ์นี้ยืนยันสมมติฐานว่า โมเดลที่ฝึกสอนด้วยคลังข้อมูลภาษาไทยสามารถทำความเข้าใจบริบทเชิงลึกและการสลับรหัสภาษาได้ดีกว่า ในขณะที่ TF-IDF ซึ่งอาศัยเพียงการนับความถี่คำ ไม่สามารถแยกแยะความซับซ้อนที่ซ่อนอยู่ในรูปประโยคได้

**4.3 การประเมินความสามารถในการแยกแยะด้วยกราฟ ROC Curve**

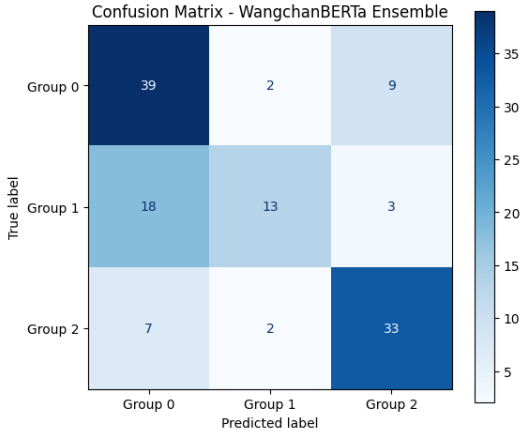
การวิเคราะห์พื้นที่ใต้เส้นโค้ง (AUC) แบบ One-vs-Rest ยืนยันว่า แบบจำลอง WangchanBERTa มีประสิทธิภาพสูงเป็นพิเศษในการแยกแยะงานระดับยาก (กลุ่ม 2) โดยมีค่า AUC สูงถึง 0.88 รองลงมาคืองานระดับง่าย (กลุ่ม 0) ที่ AUC 0.77 ดังแสดงในภาพที่ 2 ชี้ให้เห็นถึงศักยภาพในการดักจับงานที่มีความเสี่ยงสูงในโครงการได้อย่างแม่นยำ




**ภาพที่ 2** กราฟเส้นโค้ง ROC และค่า AUC สำหรับแบบจำลอง WangchanBERTa Ensemble

**4.4 การวิเคราะห์ข้อผิดพลาดของแบบจำลอง**

เมื่อพิจารณาเมทริกซ์ความสับสน ดังแสดงในภาพที่ 3 พบปัญหาที่ท้าทายที่สุดคือ การจำแนกงานระดับปานกลาง กลุ่ม 1 โดยข้อผิดพลาดที่พบบ่อยคือการประเมินต่ำกว่าความเป็นจริง โดยทายผลผิดเป็นงานระดับง่าย จากการวิเคราะห์เชิงประจักษ์พบว่า สาเหตุหลักเกิดจาก “บริบทที่ซ่อนเร้น” ในเอกสารความต้องการที่มักเขียนอย่างรวบรัด เช่น ไม่มีการระบุข้อกำหนดด้านระบบฐานข้อมูลหรือการรักษาความปลอดภัยที่ชัดเจน ทำให้แบบจำลองตีความจากข้อความผิวเผินและประเมินระดับความยากต่ำกว่าความเป็นจริง




**ภาพที่ 3** เมทริกซ์ความสับสนสำหรับแบบจำลอง WangchanBERTa

## 5. สรุปผลและอภิปราย

งานวิจัยนี้ใช้ Gemma-3n และ WangchanBERTa เพื่อจำแนกระดับความยากงานซอฟต์แวร์เป็น 3 ระดับ พบความแม่นยำสูงสุดร้อยละ 67.46 ซึ่งมีเสถียรภาพและประสิทธิภาพเพียงพอสำหรับใช้เป็นเครื่องมือสนับสนุนการตัดสินใจเบื้องต้นในโครงการซอฟต์แวร์

**5.1 สรุปผลการวิจัย**

งานวิจัยนี้นำเสนอแนวทางการจำแนกระดับความพยายามในการพัฒนาซอฟต์แวร์จากเอกสารความต้องการที่เป็นข้อความไร้โครงสร้าง โดยผสานการขยายบริบทด้วยปัญญาประดิษฐ์เชิงรู้สร้างเข้ากับการแปลงข้อความเป็นเวกเตอร์และการเรียนรู้ของเครื่อง ผลการทดลองแสดงให้เห็นว่าแนวทางดังกล่าวมีศักยภาพในการใช้เป็นเครื่องมือช่วยประเมินความเหมาะสมของค่า Man-Day เบื้องต้น โดยเฉพาะเมื่อใช้โมเดลภาษาไทยที่สอดคล้องกับลักษณะข้อมูล อย่างไรก็ตาม ยังควรมีการขยายชุดข้อมูล เพิ่มการเปรียบเทียบกับวิธีฐาน และปรับปรุงความชัดเจนของบริบทในงานระดับปานกลาง เพื่อยกระดับความน่าเชื่อถือและการนำไปใช้จริงในอนาคต

**5.2 อภิปรายผล**

จากผลการทดลอง สามารถนำมาอภิปรายร่วมกับทฤษฎีและงานวิจัยที่เกี่ยวข้องได้ดังนี้

ประสิทธิภาพของการเรียนรู้แบบกลุ่ม สถาปัตยกรรมแบบ Ensemble ช่วยลดความแปรปรวน และชดเชยจุดอ่อนของโมเดลเดี่ยวได้จริง ทำให้โมเดลมีความทนทานต่อข้อมูลที่คลุมเครือ สอดคล้องกับแนวคิดของ Pospieszny et al.[2]

ความเหนือกว่าของโมเดลภาษาเฉพาะทาง WangchanBERTa ให้ผลดีกว่า XLM-RoBERTa อย่างชัดเจน โดยเฉพาะในบริบทที่ข้อความหลักเป็นภาษาไทย แม้จะมีการสลับใช้คำศัพท์ภาษาอังกฤษทางเทคนิคก็ตาม แม้เอกสารความต้องการจะมีการสลับรหัสภาษา แต่โครงสร้างไวยากรณ์หลักยังคงเป็นภาษาไทย การใช้โมเดลที่ฝึกสอนด้วยภาษาไทยโดยเฉพาะ[3] จึงสามารถดึงความหมายเชิงลึกได้ดีกว่าโมเดลพหุภาษา

การประยุกต์ใช้ LLMs แบบพหุรูปแบบ การใช้ Gemma-3n [4] เพื่อขยายบริบทของข้อมูลช่วยให้ข้อความที่นำไปสร้างเวกเตอร์มีรายละเอียดเชิงความหมายมากขึ้น และอาจช่วยลดข้อจำกัดจากข้อความต้นฉบับที่กระชับหรือคลุมเครือเกินไป ทำให้เวกเตอร์ที่สกัดได้มีข้อมูลหนาแน่นและเพียงพอต่อการเรียนรู้ของอัลกอริทึม

การจำแนกประเภทแทนการพยากรณ์ตัวเลข การเปลี่ยนโจทย์มาเป็นการจัดกลุ่มความยากง่าย[1] แทนการพยากรณ์เป็นตัวเลขตรงๆ ช่วยลดความเสี่ยงจากมาตรฐานการประเมินที่ไม่เท่ากันของผู้ให้บริการแต่ละราย ทำให้ระบบมีความน่าเชื่อถือมากขึ้น

**5.3 นัยยะเชิงการบริหารและข้อเสนอแนะ**

การประยุกต์ใช้ประโยชน์และลดต้นทุนในระบบ e-Procurement การบูรณาการแบบจำลองเข้ากับระบบจัดจ้างอิเล็กทรอนิกส์ จะช่วยลดระยะเวลาและต้นทุนการบริหารโครงการได้อย่างเป็นรูปธรรม โดยระบบจะทำหน้าที่เป็นเครื่องมือคัดกรองด่านแรก เพื่อแจ้งเตือนรายการที่ประเมินค่า Man-Day ผิดปกติ กลไกนี้ช่วยลดต้นทุนเวลา ของคณะกรรมการตรวจรับที่ต้องวิเคราะห์เอกสารทุกรายการด้วยตนเอง ให้สามารถมุ่งเน้นเฉพาะรายการที่มีความเสี่ยง ทำให้กระบวนการอนุมัติจัดจ้างรวดเร็ว โปร่งใส และป้องกันการสูญเสียงบประมาณจากการประเมินสูงเกินจริง

ข้อจำกัดของการวิจัยนี้มีข้อจำกัดสำคัญด้านขนาดของชุดข้อมูล ซึ่งใช้ข้อมูลจำนวน 419 รายการ จาก 22 โครงการ ถือเป็นระดับเริ่มต้นสำหรับการเรียนรู้เชิงลึก ข้อจำกัดด้านปริมาณและความหลากหลายนี้อาจส่งผลต่อความสามารถในการวางนัยทั่วไป และขอบเขตความแม่นยำ หากนำไปประยุกต์ใช้กับเอกสารความต้องการในโดเมนธุรกิจที่แตกต่างออกไป ดังนั้น การนำระบบไปใช้จริงจึงควรใช้ในลักษณะเครื่องมือสนับสนุนการตัดสินใจ ควบคู่กับดุลยพินิจของผู้เชี่ยวชาญ

**5.4 ข้อเสนอแนะสำหรับการทำวิจัยในอนาคต**

เพื่อแก้ไขปัญหาความคลุมเครือในการจำแนกงานระดับปานกลาง งานวิจัยในอนาคตงานวิจัยในอนาคตควรทดลองใช้ structured prompting หรือ few-shot prompting เพื่อควบคุมรูปแบบการอธิบายบริบทให้มีความสม่ำเสมอมากขึ้น เพื่อปรับปรุงมาตรฐานการอธิบายบริบทให้ชัดเจนยิ่งขึ้น และศึกษาผลของการตัดองค์ประกอบแต่ละส่วนใน pipeline เพื่อยืนยันประโยชน์ของแนวทางที่เสนอ

## เอกสารอ้างอิง


[1] Bakir, A., Turhan, B., & Bener, A. (2016). Software Effort Estimation as a Classification Problem. Proceedings of the 42nd Euromicro Conference on Software Engineering and Advanced Applications (SEAA), 230-237.

[2] Pospieszny, P., Czarnacka-Chrobot, B., & Kobylinski, A. (2018). An effective ensemble of classifiers and regressors for software effort estimation. International Journal of Software Engineering and Knowledge Engineering, 28(09), 1295-1317.

[3] Lowphansirikul, L., Polpanumas, C., Jantrakulchai, N., & Nutanong, S. (2021). WangchanBERTa: Pretraining transformer-based Thai Language Models. arXiv preprint arXiv:2101.09635.

[4] Gemma Team, Google DeepMind. (2024). Gemma: Open Models Based on Gemini Research and Technology. arXiv preprint arXiv:2403.08295.

[5] Conneau, A., Khandelwal, K., Goyal, N., Chaudhary, V., Wenzek, G., Guzmán, F., ... & Stoyanov, V. (2020). Unsupervised Cross-lingual Representation Learning at Scale. Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics.

[6] Boehm, B. W. (1981). Software Engineering Economics. Englewood Cliffs, NJ: Prentice-Hall.

[7] Moløkken, K., & Jørgensen, M. (2003). A review of software surveys on software effort estimation. Proceedings of the 2003 International Symposium on Empirical Software Engineering, 223-230.

[8] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. Advances in Neural Information Processing Systems, 30.

[9] Yin, S., Fu, C., Zhao, S., Li, K., Sun, X., Xu, T., & Chen, E. (2024). A Survey on Multimodal Large Language Models. arXiv preprint arXiv:2306.13549.

[10] Wei, J., Wang, X., Schuurmans, D., Bosma, M., Ichter, B., Xia, F., ... & Zhou, D. (2022). Chain-of-Thought Prompting Elicits Reasoning in Large Language Models. Advances in Neural Information Processing Systems, 35, 24824-24837.

# การพัฒนาระบบบริหารจัดการการเข้าร่วมกิจกรรมชมรมวิชาชีพ เทคโนโลยีธุรกิจดิจิทัล วิทยาลัยเทคนิคชัยนาท

# Development of an Activity Attendance Management System for the Digital Business Technology Club, Chainat Technical College

*วิลาสินี ชูติวัตร์ (Wilasinee Chutiwat)[1] และณัฐวี อุตกฤษฎ์ (Nattavee Utakrit)[2]*

*[1]ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล*

*มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ*

*[1]s6707021857244@email.kmutnb.ac.th, [2] nattavee.u@itd.kmutnb.ac.th*

## บทคัดย่อ

*การวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาระบบบริหารจัดการการเข้าร่วมกิจกรรมชมรมวิชาชีพเทคโนโลยีธุรกิจดิจิทัล วิทยาลัยเทคนิคชัยนาท เพื่อแก้ปัญหาความล่าช้าและความคลาดเคลื่อนในการบันทึกข้อมูลการเข้าร่วมกิจกรรมของนักเรียนนักศึกษา โดยระบบที่พัฒนาขึ้นช่วยให้การเช็คชื่อและการสรุปรายงานผลการเข้าร่วมกิจกรรมมีความถูกต้อง รวดเร็ว และตรวจสอบได้ง่ายขึ้น จากการตรวจสอบความถูกต้องของระบบด้วยวิธี Black Box Testing ตามหลักการทดสอบการยอมรับจากผู้ใช้งาน พบว่าระบบสามารถทำงานได้ถูกต้องครบถ้วนตามความต้องการของผู้ใช้ สำหรับผลการประเมินประสิทธิภาพของระบบโดยผู้เชี่ยวชาญ พบว่าระบบมีความเหมาะสมในระดับดีมาก ค่าเฉลี่ย 4.64 ส่วนเบี่ยงเบนมาตรฐาน 0.51 และผลการประเมินความพึงพอใจจากผู้ใช้งานอยู่ในระดับดีมากเช่นกัน ค่าเฉลี่ย 4.68 ส่วนเบี่ยงเบนมาตรฐาน 0.52 โดยผลการประเมินมีนัยสำคัญทางสถิติที่ระดับ 0.05 แสดงให้เห็นว่าระบบที่พัฒนาขึ้นมีคุณภาพและมีความเหมาะสมต่อการใช้งานจริง*

**คำสำคัญ:** การพัฒนาระบบ, บริหารจัดการกิจกรรม, ชมรมวิชาชีพ,เทคโนโลยีธุรกิจดิจิทัล, วิทยาลัยเทคนิคชัยนาท

## Abstract

*This research aimed to develop a management system for student participation in activities of the Digital Business Technology Professional Club at Chai Nat Technical College to address delays and inaccuracies in recording students' activity attendance. The developed system improves the accuracy, speed, and traceability of attendance checking and activity participation reporting. System correctness was verified using Black Box Testing based on User Acceptance Testing (UAT) principles, which indicated that the system functioned completely and accurately according to user requirements.*

*The system's efficiency evaluation by experts revealed a very high level of appropriateness (mean = 4.64, S.D. = 0.51). Similarly, the user satisfaction assessment showed a very high level (mean = 4.68, S.D. = 0.52). The evaluation results were statistically significant at the 0.05 level, demonstrating that the developed system is of high quality and suitable for practical implementation.*

**Keywords:** System Development, Activity Management, Vocational Club, Digital Business Technology, Chainat Technical College

## 1. บทนำ

ในปัจจุบันการจัดกิจกรรมของชมรมในสถานศึกษาเป็นส่วนสำคัญที่ช่วยส่งเสริมทักษะ ความสามารถ และประสบการณ์ของนักศึกษา อย่างไรก็ตามการบริหารจัดการกิจกรรมยังคงใช้วิธีการแบบเดิมเช่นการลงทะเบียนด้วย

เอกสาร การบันทึกข้อมูลด้วยตนเอง และการจัดเก็บข้อมูลแบบกระจัดกระจาย ส่งผลให้เกิดความล่าช้า ความผิดพลาดในการบันทึกข้อมูล และยากต่อการตรวจสอบย้อนหลัง

ดังนั้นจึงมีความจำเป็นในการพัฒนาระบบสารสนเทศเพื่อช่วยสนับสนุนการบริหารจัดการกิจกรรมให้มีประสิทธิภาพมากยิ่งขึ้น โดยเฉพาะระบบที่สามารถรองรับการลงทะเบียนออนไลน์ การบันทึกเวลาเข้าร่วมกิจกรรม การจัดการแข่งขัน การบันทึกผลการแข่งขัน และการออกเกียรติบัตรอิเล็กทรอนิกส์ได้อย่างเป็นระบบ การพัฒนาระบบดังกล่าวจะช่วยลดภาระงานของผู้ดูแลกิจกรรม เพิ่มความถูกต้องของข้อมูล และช่วยให้นักศึกษาสามารถเข้าถึงข้อมูลกิจกรรมได้อย่างสะดวกและรวดเร็วมากยิ่งขึ้น

## 2. วิจารณ์วรรณกรรม หรืองานวิจัยที่เกี่ยวข้อง

จากการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องกับการพัฒนาระบบบริหารจัดการกิจกรรมและการมีส่วนร่วมของนักศึกษา สามารถสรุปประเด็นสำคัญได้ดังนี้

### 2.1 แนวทางการจัดกิจกรรมและการส่งเสริมการมีส่วนร่วม

การจัดกิจกรรมเพื่อส่งเสริมการมีส่วนร่วมในระดับอาชีวศึกษามีความสำคัญอย่างยิ่งต่อการพัฒนาผู้เรียน [9] ได้ศึกษาวิจัยเกี่ยวกับการพัฒนา "ชุดกิจกรรมเสริมหลักสูตรสำหรับนักศึกษาวิชาชีพระบบทวิภาคี" พบว่าปัจจัยสำคัญที่ช่วยกระตุ้นการมีส่วนร่วมคือการออกแบบเนื้อหากิจกรรมให้สอดคล้องกับความต้องการของผู้เรียน และการมีส่วนร่วมของครูที่ปรึกษา สอดคล้องกับรายงานของสำนักงานคณะกรรมการการอาชีวศึกษา [16] ที่ชี้ให้เห็นถึงความจำเป็นในการปรับตัวของการจัดกิจกรรมภายหลังสถานการณ์โควิด-19 โดยการใช้แพลตฟอร์มออนไลน์และการพัฒนาทักษะในศตวรรษที่ 21 เพื่อให้เกิดความต่อเนื่องทางการศึกษา

### 2.2 การนำเทคโนโลยีสารสนเทศมาใช้ในกิจกรรมนักศึกษา

การประยุกต์ใช้เทคโนโลยีเว็บแอปพลิเคชันและฐานข้อมูลมีบทบาทสำคัญในการแก้ไขปัญหาการจัดการกิจกรรมแบบดั้งเดิม เช่น งานวิจัยที่พัฒนาระบบสารสนเทศกิจกรรมนักศึกษาผ่านเว็บแอปพลิเคชัน [11] เพื่อเพิ่มความสะดวกรวดเร็วในการบริหารจัดการข้อมูล นอกจากนี้ ระบบที่ใช้ฐานข้อมูล MySQL ยังได้รับการยอมรับว่ามีความแม่นยำสูงในการจัดเก็บข้อมูลการลงทะเบียนและประมวลผลรายงานผลกิจกรรมได้อย่างมีประสิทธิภาพ [1], [4], [7] เนื่องจากรองรับการทำงานแบบ Multi-user และมีความปลอดภัยของข้อมูลสูง [17]

### 2.3 ปัจจัยที่มีอิทธิพลต่อพฤติกรรมการเข้าร่วมกิจกรรม

พฤติกรรมการเข้าร่วมกิจกรรมของนักศึกษาได้รับอิทธิพลจากหลายปัจจัย [15] ได้วิเคราะห์ปัจจัยด้านเจตคติ แรงจูงใจ และความรับผิดชอบ ซึ่งเป็นตัวแปรสำคัญที่กำหนดความสำเร็จในการจัดกิจกรรม [10] พบว่าปัญหาหลักในการจัดกิจกรรมมักเกิดในด้านวิชาการ จึงเสนอแนวทางพัฒนาผ่านการอบรมเชิงปฏิบัติการและการประชาสัมพันธ์ผ่านสื่อดิจิทัลเพื่อกระตุ้นความสนใจของนักศึกษา ซึ่งสอดคล้องกับปัจจัยด้านการเข้าถึงข้อมูลและความสะดวกในการใช้งานระบบสารสนเทศ [5], [12] ที่ส่งผลต่อระดับการมีส่วนร่วมในกิจกรรมอย่างมีนัยสำคัญ[18]

**สรุปงานวิจัยที่เกี่ยวข้อง**

จากการวิเคราะห์งานวิจัยข้างต้น ผู้วิจัยจึงได้นำแนวคิดด้านการจัดการฐานข้อมูล MySQL และภาษา PHP มาบูรณาการเข้ากับการบริหารจัดการกิจกรรมชมรมวิชาชีพเทคโนโลยีธุรกิจดิจิทัล เพื่อลดภาระงานด้านเอกสารและเพิ่มความแม่นยำในการออกเกียรติบัตรอิเล็กทรอนิกส์ (e-Certificate) ให้สอดคล้องกับความต้องการของผู้ใช้งานในยุคดิจิทัล

## 3. วิธีดำเนินการวิจัย

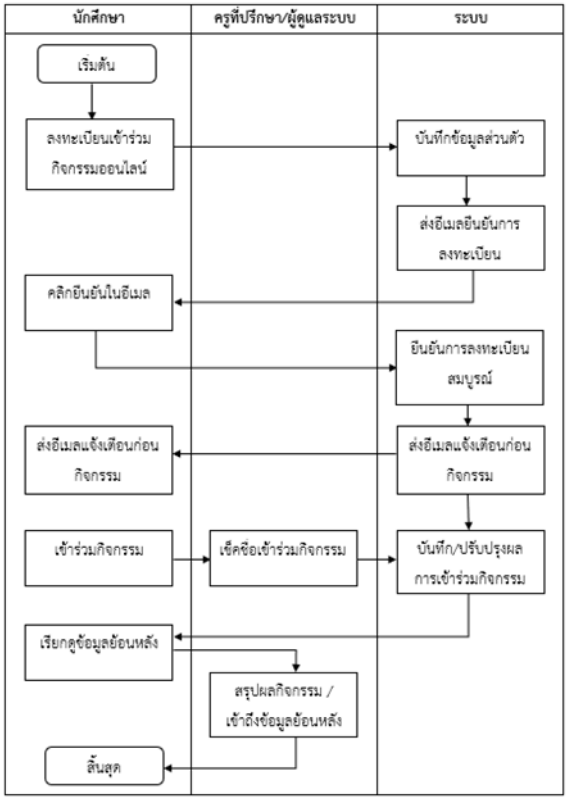
การวิจัยเรื่องการพัฒนาระบบบริหารจัดการการเข้าร่วมกิจกรรมชมรมวิชาชีพเทคโนโลยีธุรกิจดิจิทัล วิทยาลัยเทคนิคชัยนาท เป็นการวิจัยและพัฒนา (Research and Development: R&D) โดยผู้วิจัยได้ดำเนินการตามขั้นตอนการพัฒนาระบบ ดังนี้

### 3.1 การศึกษาและรวบรวมข้อมูล

การพัฒนาระบบบริหารจัดการการเข้าร่วมกิจกรรมชมรมวิชาชีพ ผู้วิจัยได้ทำการศึกษาแนวคิด ทฤษฎี งานวิจัยที่เกี่ยวข้อง และสำรวจสภาพปัญหาการจัดเก็บข้อมูลการเข้าร่วมกิจกรรมของนักศึกษา พบปัญหาความล่าช้า ความคลาดเคลื่อน และความยุ่งยากในการสืบค้นข้อมูลย้อนหลัง จึงได้รวบรวมความต้องการของผู้ใช้งาน ได้แก่ นักศึกษา ผู้เข้าร่วมการแข่งขัน ครูที่ปรึกษา และผู้ดูแลระบบ

### 3.2 การวิเคราะห์และออกแบบระบบ (System Design)

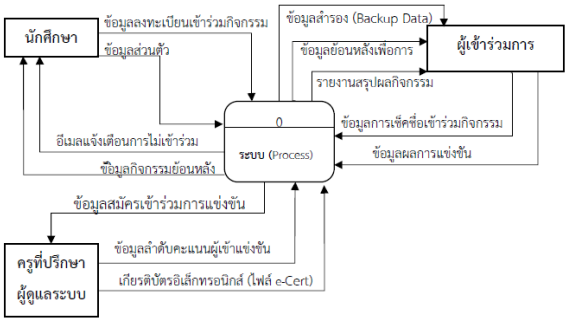
ผู้วิจัยออกแบบโครงสร้างระบบประกอบด้วย การออกแบบผังงาน (Workflow) กำหนดกระบวนการตั้งแต่การลงทะเบียน การยืนยันตัวตนผ่านอีเมล การแจ้งเตือนกิจกรรม การเช็คชื่อโดยครูที่ปรึกษา ไปจนถึงการสรุปผลและออกเกียรติบัตร




**ภาพที่ 1:** Swimlane Diagram กระบวนการเข้าร่วมกิจกรรมสำหรับนักศึกษาทั่วไป

การออกแบบฐานข้อมูล (Database Design) ใช้ MySQL ในการจัดเก็บข้อมูลอย่างเป็นระบบเพื่อให้รองรับการทำงานแบบ Multi-user และมีความแม่นยำในการประมวลผล

การออกแบบส่วนติดต่อผู้ใช้งาน (User Interface Design) ออกแบบโดยใช้หลักการ Responsive เพื่อให้รองรับการใช้งานผ่านอุปกรณ์ที่หลากหลาย เช่น คอมพิวเตอร์และสมาร์ทโฟน




**ภาพที่ 2:** Context Diagram ระบบบริหารจัดการการเข้าร่วมกิจกรรมชมรมวิชาชีพ

### 3.3 การพัฒนาระบบ (System Implementation)

เพื่อให้ระบบมีความทันสมัยและรองรับการใช้งานแบบ Responsive Website ผู้วิจัยได้กำหนดรายละเอียดเชิงเทคนิคดังนี้

**ตารางที่ 1:** รายละเอียดเชิงเทคนิค (Technical Details)

| หัวข้อ | รายละเอียดเชิงเทคนิค |
|---|---|
| Architecture | Web Application (Client-Server Architecture) |
| Frontend | HTML ร่วมกับ Responsive Design Framework (Bootstrap) เพื่อรองรับทุกขนาดหน้าจอ |
| Backend | ภาษา PHP สำหรับการประมวลผลฝั่งเซิร์ฟเวอร์ |
| Database | MySQL สำหรับการจัดการฐานข้อมูลเชิงสัมพันธ์ที่รวดเร็วและปลอดภัย |
| Editor | Visual Studio Code (VS Code) |

### 3.4 การประเมินผล

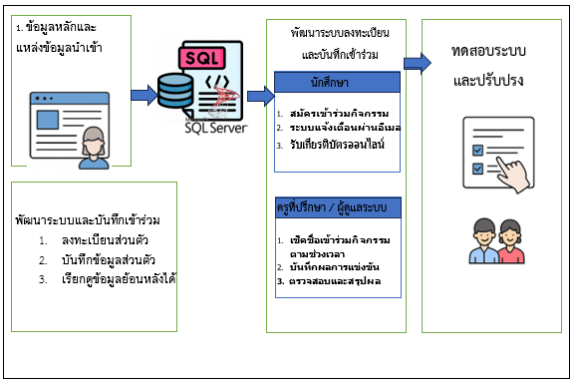
การประเมินผลระบบครั้งนี้มุ่งเน้น วัดความพึงพอใจของผู้ใช้งานต่อระบบ เป็นสำคัญ ซึ่งพิจารณาจากค่าคะแนนเฉลี่ยตามมาตราส่วนประมาณค่า (Rating Scale)

ประชากรและกลุ่มตัวอย่างที่ใช้ในการประเมินประกอบด้วย ครูที่ปรึกษา จำนวน 5 คน และนักเรียน

จำนวน 34 คน โดยได้มาจากการเลือกแบบเจาะจง (Purposive Sampling) เนื่องจากเป็นผู้ที่มีส่วนเกี่ยวข้องและได้ใช้งานระบบจริง

เครื่องมือที่ใช้ในการประเมิน คือ แบบประเมินความพึงพอใจต่อการใช้งานระบบ ซึ่งครอบคลุมประเด็นด้านความชัดเจนของข้อมูล ความง่ายในการใช้งาน ความสะดวกของขั้นตอนการทำงาน และความเหมาะสมของส่วนติดต่อผู้ใช้ (User Interface)

การวิเคราะห์ข้อมูลใช้สถิติพื้นฐาน ได้แก่ ค่าเฉลี่ย ($\bar{x}$) และส่วนเบี่ยงเบนมาตรฐาน (S.D.) เพื่อสะท้อนระดับความพึงพอใจของผู้ใช้งานต่อระบบตามวัตถุประสงค์ที่กำหนดไว้




**ภาพที่ 3:** กรอบแนวความคิดการพัฒนาระบบ

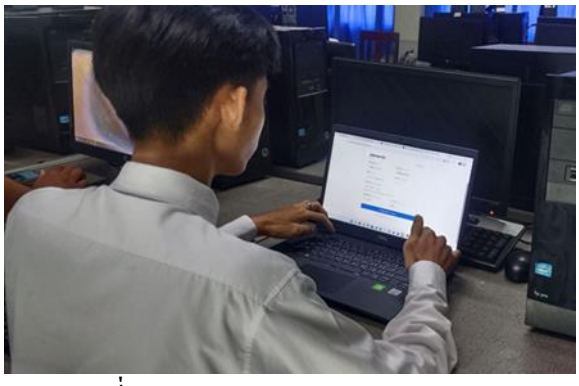
## 4. ผลการทดลอง / ผลการประเมินระบบ

การประเมินผลการพัฒนาระบบบริหารจัดการการเข้าร่วมกิจกรรมชมรมวิชาชีพเทคโนโลยีธุรกิจดิจิทัล วิทยาลัยเทคนิคชัยนาท ผู้วิจัยได้ดำเนินการเก็บรวบรวมข้อมูลจากกลุ่มตัวอย่าง ซึ่งประกอบด้วยครูที่ปรึกษาและนักศึกษา โดยใช้แบบประเมินความพึงพอใจต่อการใช้งานระบบ ผลการวิเคราะห์ข้อมูลสรุปได้ดังนี้

### 4.1 การนำระบบไปใช้จริงในสถานศึกษา

ระบบบริหารจัดการการเข้าร่วมกิจกรรมนี้ได้นำไปใช้งานจริงภายในชมรมวิชาชีพเทคโนโลยีธุรกิจดิจิทัล วิทยาลัยเทคนิคชัยนาท และก่อให้เกิดประโยชน์อย่างชัดเจน ดังนี้

**ลดภาระงานแบบเอกสาร (Manual Process)**ลดการเซ็นชื่อในกระดาษและการเขียนเกียรติบัตรด้วยมือ ซึ่งช่วยลดความผิดพลาดจากมนุษย์ (Human Error)

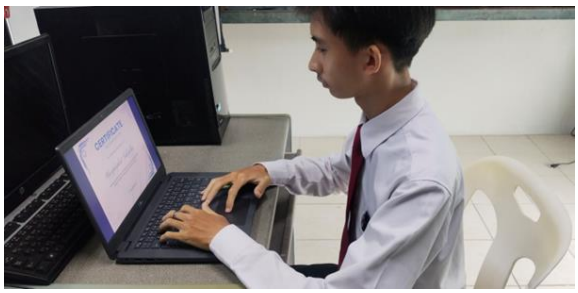


**ภาพที่ 4:** การสมัครเข้าร่วมการแข่งขันแบบออนไลน์

**ระบบการสมัครสมาชิกและลงทะเบียน** บันทึกข้อมูลลงในฐานข้อมูลและนำผู้ใช้งานกลับมายังหน้าเข้าสู่ระบบอัตโนมัติพร้อมแสดงแถบข้อความแจ้งเตือนสถานะเพื่อให้ผู้ใช้งานมั่นใจว่าการลงทะเบียนเสร็จ

**การติดตามผลอย่างเป็นระบบ** ครูที่ปรึกษาสามารถตรวจสอบประวัติการเข้าร่วมกิจกรรมของนักศึกษาได้แบบเรียลไทม์ผ่านฐานข้อมูลกลาง

**ระบบแจ้งเตือนอัตโนมัติ** มีการส่งอีเมลแจ้งเตือนไปยังผู้เกี่ยวข้องเมื่อมีการลงทะเบียนหรือไม่เข้าร่วมกิจกรรมตามกำหนด

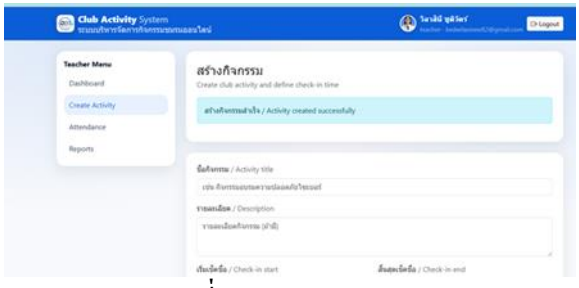


**ภาพที่ 5:** การดาวน์โหลดและพิมพ์เกียรติบัตร

**ความโปร่งใสของข้อมูล** นักศึกษาสามารถตรวจสอบสถานะการเข้าร่วมกิจกรรมและดาวน์โหลดเกียรติบัตรได้ด้วยตนเอง

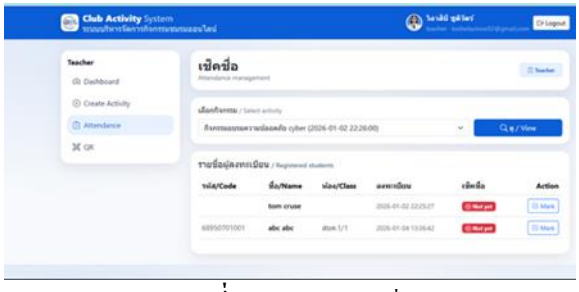
### 4.2 ผลการพัฒนาระบบ

จากการประเมินความพึงพอใจของผู้ใช้งาน พบว่าระบบสามารถตอบสนองความต้องการของผู้ใช้ได้อย่างเหมาะสม มีความชัดเจน ใช้งานง่าย และครอบคลุมการใช้งานในส่วนของฟังก์ชันสำคัญต่าง ๆ ดังนี้



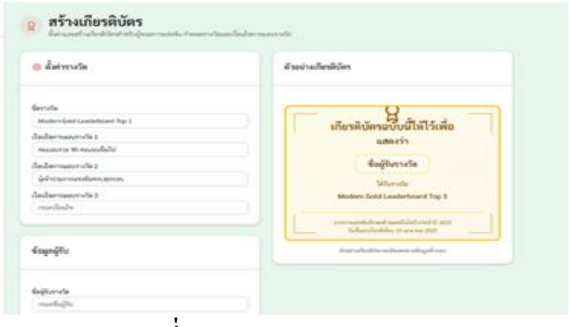
**ภาพที่ 6:** หน้าจอสร้างกิจกรรม

**ระบบเช็คชื่อเข้าร่วมกิจกรรม** รองรับการเช็คชื่อตามช่วงเวลา (เช้า/บ่าย) ได้อย่างแม่นยำ และแสดงสถานะการเข้าร่วมกิจกรรมแบบเรียลไทม์



**ภาพที่ 7:** หน้าจอการเช็คชื่อ

**ระบบจัดการผลการแข่งขันและเกียรติบัตร** สามารถป้อนคะแนน จัดลำดับผู้เข้าแข่งขัน และออกเกียรติบัตรอิเล็กทรอนิกส์ (e-Certificate) ให้ผู้เข้าร่วมดาวน์โหลดได้ทันที



**ภาพที่ 8:** หน้าจอภาพสร้างเกียรติบัตร

**4.3 ผลการประเมินความพึงพอใจและการใช้งานระบบ (Usability)**

จากการประเมินความพึงพอใจของผู้ใช้งานในด้านความชัดเจนและการใช้งานง่ายของระบบ (Usability) พบว่าผู้ใช้งานมีความพึงพอใจต่อระบบในระดับดีมาก โดยมีรายละเอียดในแต่ละประเด็น ดังนี้

- การแสดงผลข้อมูลกิจกรรมมีความชัดเจนและเข้าใจง่าย
- ขั้นตอนการลงทะเบียนเข้าร่วมกิจกรรมมีความสะดวกและไม่ยุ่งยาก
- ระบบการเช็คชื่อเข้ากิจกรรมตามช่วงเวลามีความถูกต้องและใช้งานง่าย
- การจัดวางเมนูและปุ่มคำสั่งต่าง ๆ มีความเหมาะสมและไม่ซับซ้อน

**4.4 สรุปผลการประเมินภาพรวม**

ผลการประเมินความพึงพอใจโดยรวมพบว่ากลุ่มผู้เชี่ยวชาญซึ่งเป็นครูที่ปรึกษาชมรมมีค่าเฉลี่ยเท่ากับ 4.64 และกลุ่มผู้ใช้งานมีค่าเฉลี่ยเท่ากับ 4.68 จากคะแนนเต็ม 5 ซึ่งอยู่ในระดับ “ดีมาก” สะท้อนให้เห็นว่าระบบที่พัฒนาขึ้นสามารถตอบสนองความต้องการของผู้ใช้งานได้อย่างเหมาะสม มีความชัดเจน ใช้งานง่าย และเอื้อต่อการบริหารจัดการกิจกรรมภายในชมรม ช่วยลดภาระงานด้านเอกสาร ตลอดจนเพิ่มความสะดวกและความรวดเร็วในการสืบค้นข้อมูลย้อนหลังได้อย่างมีประสิทธิผล

## 5. สรุปผลและข้อเสนอแนะ

จากขั้นตอนการดำเนินงานวิจัยและการทดสอบระบบบริหารจัดการการเข้าร่วมกิจกรรมชมรมวิชาชีพเทคโนโลยีธุรกิจดิจิทัล วิทยาลัยเทคนิคชัยนาท สามารถสรุปประเด็นสำคัญและมีข้อเสนอแนะดังนี้

**5.1 สรุปผลการวิจัย**

การพัฒนาระบบบริหารจัดการกิจกรรมในรูปแบบออนไลน์ช่วยยกระดับความพึงพอใจของผู้ใช้งานภายในชมรมวิชาชีพได้อย่างชัดเจน โดยระบบสามารถแก้ปัญหาการสูญหายของข้อมูลและความล่าช้าจากการบันทึกด้วยกระดาษได้เป็นอย่างดี จากผลการประเมินความพึงพอใจของผู้ใช้งาน ได้แก่ ครูที่ปรึกษาและนักศึกษา พบว่ามีค่าเฉลี่ยอยู่ในระดับ “ดีมาก” สะท้อนให้เห็นว่าระบบมีความชัดเจน ใช้งานง่าย และอำนวยความสะดวกต่อการดำเนินกิจกรรมต่าง ๆ เช่น การลงทะเบียนการเช็คชื่อตามช่วงเวลา และการรับเกียรติบัตรอิเล็กทรอนิกส์

**5.2 ข้อเสนอแนะ**

จากการดำเนินโครงการวิจัย ผู้วิจัยมีข้อเสนอแนะเพื่อการพัฒนาในอนาคต ดังนี้

**ข้อเสนอแนะเพื่อการนำไปใช้** ควรมีการขยายผลการใช้งานระบบไปยังแผนกวิชาอื่น หรือชมรมอื่น ๆ ภายในวิทยาลัยเทคนิคชัยนาท เพื่อสร้างมาตรฐานการจัดการกิจกรรมที่เป็นระบบเดียวกันทั้งสถานศึกษา

**ข้อเสนอแนะในการวิจัยครั้งต่อไป** ควรมีการพัฒนาระบบให้สามารถรองรับการเข้าใช้งานในรูปแบบแอปพลิเคชันบนมือถือ (Mobile Application) เพื่อเพิ่มความสะดวกในการรับการแจ้งเตือนแบบเรียลไทม์ และควรเพิ่มฟังก์ชันการวิเคราะห์ข้อมูลเชิงสถิติ (Data Analytics) เพื่อให้ครูที่ปรึกษาเห็นแนวโน้มความสนใจของนักศึกษาต่อประเภทกิจกรรมต่าง ๆ ได้อย่างละเอียดขึ้น

## เอกสารอ้างอิง


[1] A. A. Aziz, et al., “Design Architecture of an Integrated Student Activities Management System for Higher Education,” *Conference Paper/Research Report*, 2021.

[2] A. R. A. Zaragosa and R. R. B. Flores, “Extracurricular Activities Attendance Management Monitoring System with SMS Notification,” *Research Report*, 2022.

[3] D. Sreedhar, et al., “Developing a Mobile Application for Student Clubs,” *Research Report*, 2023.

[4] F. Hidayat and R. A. Sundari, “Web-Based Extracurricular Management Information System at SMK Pertiwi Batam,” *Research Report*, 2021.

[5] J. Zhou, et al., “Designing and Testing User-Club Systems to Improve Club Organization and Participation in High Schools and Colleges*,” Research Report*, 2023.

[6] N. Imjai, et al., “Unlocking Accounting Student Success: The Interplay of Student Activity Participation, Social Skills, & Emotional Maturity,” *Research Report*, 2024.

[7] R. S. Karim, A. H. Pratama, and W. A. Ramdani, “Web-Based Student Extracurricular Value Monitoring Application,” *International Journal of Research and Applied Technology*, 2022.

[8] Suitable Inc., “Club Management System (CMS) for Student Engagement – Case of Suitable Company's App,” [Online]. Available: https://www.suitable.co. (Accessed: Feb. 6, 2026), 2023.

[9] กองสุข พาปี, “การพัฒนาการจัดกิจกรรมเพื่อส่งเสริมในนักศึกษาวิชาชีพระบบทวิภาคี (ปวส. แผนกช่างยนต์) โดยการวิจัยเชิงปฏิบัติการแบบมีส่วนร่วม,” วิทยานิพนธ์ปริญญาดุษฎีบัณฑิต, 2565.

[10] ขวัญชัย พะยอม, “แนวทางพัฒนาการจัดกิจกรรมนักศึกษา มหาวิทยาลัยราชภัฏกำแพงเพชร,” รายงานการวิจัย มหาวิทยาลัยราชภัฏกำแพงเพชร, 2564.

[11] จารุทิศน์ สายสิงห์, “การพัฒนาระบบสารสนเทศกิจกรรมนักศึกษาผ่านเว็บแอพพลิเคชัน : กรณีศึกษามหาวิทยาลัยภาคตะวันออกเฉียงเหนือ,” รายงานการวิจัย มหาวิทยาลัยภาคตะวันออกเฉียงเหนือ, 2566.

[12] จินตนา ตะกูล, “ปัจจัยที่มีอิทธิพลต่อการมีส่วนร่วมในกิจกรรมทางการศึกษาของนักศึกษาระดับปริญญาตรี มหาวิทยาลัยมหานครราชวิทยาลัย วิทยาเขตศรีธรรมโศกราช,” รายงานการวิจัย, 2564.

[13] ชนัญญามนต์ รัตนบุญทอง, “การศึกษาเหตุผลการเข้าร่วมกิจกรรมของนักศึกษา มหาวิทยาลัยเทคโนโลยีราชมงคลพระนคร,” รายงานการวิจัย มหาวิทยาลัยเทคโนโลยีราชมงคลพระนคร, 2566.

[14] บุษบา สังขวรรณะ, “การพัฒนาระบบบันทึกการเข้าร่วมกิจกรรมด้วย AR Code,” รายงานการวิจัย, 2563.

[15] สมโชค เพ็งสี, “การศึกษาแนวทางการส่งเสริมการเข้าร่วมกิจกรรมนักศึกษาระดับปริญญาตรี คณะเกษตรศาสตร์ มหาวิทยาลัยขอนแก่น,” รายงานการวิจัย มหาวิทยาลัยขอนแก่น, 2563.

# การจำแนกความคิดเห็นของผู้ใช้บริการร้านอาหารสุขภาพ บนแพลตฟอร์มออนไลน์ด้วยเทคนิคการเรียนรู้เชิงลึก

# Classification of User Opinions on Healthy Meal Delivery Services on Online Platforms Using Deep Learning Techniques

*สุดา ภูจำนงค์ (Suda Phuchamnong)[1] และวัชรีวรรณ จิตต์สกุล (Watchareewan Jitsakul)[2]*
*ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล*
*มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ*
*[1]s6707021856124 @email.kmutnb.ac.th, [2]watchareewan.j@itd.kmutnb.ac.th*

## บทคัดย่อ

*งานวิจัยนี้มีวัตถุประสงค์เพื่อจำแนกข้อความการแสดงความคิดเห็น ที่มีต่อร้านอาหารสุขภาพบนแพลตฟอร์มออนไลน์ โดยใช้ชุดข้อมูลจำนวน 1,100 รายการ แบ่งเป็น 2 ประเภท โดยใช้สำหรับการฝึกแบบจำลอง 1,000 รายการ และทดสอบด้วยข้อมูลที่ไม่เคยผ่านการฝึก 100 รายการ ข้อมูลถูกนำมาผ่านกระบวนการทำความสะอาดข้อความ การตัดคำภาษาไทย และการแปลงข้อความเป็นเวกเตอร์ตัวเลขด้วยเทคนิค Word2Vec จากนั้นทำการแบ่งข้อมูลเพื่อทดสอบด้วยวิธีการตรวจสอบแบบไขว้ 5 ส่วน เพื่อทำการเปรียบเทียบประสิทธิภาพของเทคนิคการเรียนรู้เชิงลึกบนพื้นฐานของโมเดลประมวลผลภาษาธรรมชาติ ได้แก่ แบบจำลองโครงข่ายประสาทเทียมแบบวนซ้ำ แบบจำลองโครงข่ายประสาทเทียมแบบวนซ้ำสองทิศทาง แบบจำลองโครงข่ายประสาทเทียมแบบหน่วยวนซ้ำที่มีกลไกประตู และแบบจำลองโครงข่ายประสาทเทียมแบบหน่วยวนซ้ำที่มีกลไกประตูสองทิศทาง ผลการวิจัยพบว่าแบบจำลองที่ใช้เทคนิคโครงข่ายประสาทเทียมแบบหน่วยวนซ้ำที่มีกลไกประตูสองทิศทาง มีประสิทธิภาพสูงสุด ด้วย ค่าความถูกต้องร้อยละ 91.30 ค่าความแม่นยำ 0.91 ค่าการระลึก 0.91 และค่าการถ่วงดุล 0.91 และการวิเคราะห์เส้นโค้งคุณลักษณะ 0.96 ตามลำดับ*

**คำสำคัญ:** โครงข่ายประสาทเทียมแบบวนซ้ำ โครงข่ายประสาทเทียมแบบวนซ้ำสองทิศทาง โครงข่ายประสาทเทียมแบบหน่วยวนซ้ำที่มีกลไกประตู โครงข่ายประสาทเทียมแบบหน่วยวนซ้ำที่มีกลไกประตูสองทิศทาง

## Abstract

*This research aims to classify opinionated text related to healthy restaurants on online platforms. The dataset consists of 1,100 text samples, of which 1,000 were used for model training and 100 were used as unseen test data. The data were processed through several text preprocessing steps, including text cleaning, Thai word segmentation and transformation of textual data into numerical vector representations using Word2Vec. Subsequently, the dataset was evaluated using a five-fold cross-validation technique to compare the performance of deep learning methods based on natural language processing models. The models employed in this study include the Recurrent Neural Network, Bidirectional Recurrent Neural Network, Gated Recurrent Unit, and Bidirectional Gated Recurrent Unit. The research results indicate that the Bidirectional Gated Recurrent Unit (Bi-GRU) model achieved the highest performance. The model obtained an accuracy of 91.30%, a precision of 0.91, a recall of 0.91, an F1-score of 0.91, and an area under the receiver operating characteristic curve (AUC) of 0.96, respectively.*

**Keywords:** Recurrent Neural Network, Bidirectional Recurrent Neural Network, Gated

Recurrent Unit, Bidirectional Gated Recurrent Unit.

## 1. บทนำ

ในช่วงทศวรรษที่ผ่านมา การเปลี่ยนผ่านสู่ยุคดิจิทัลส่งผลให้พฤติกรรมการสั่งอาหารผ่านแพลตฟอร์มออนไลน์เติบโตอย่างรวดเร็ว โดยเฉพาะหลัง COVID-19 ที่ทำให้ธุรกิจส่งอาหารสุขภาพได้รับความนิยมเพิ่มขึ้น เนื่องจากตอบโจทย์ทั้งความสะดวก คุณค่าทางโภชนาการ และความใส่ใจสุขภาพ โดย Statista รายงานว่ามูลค่าตลาด Food Delivery ทั่วโลกในปี 2023 อยู่ที่ประมาณ 1.22 ล้านล้านบาท และคาดว่าจะเติบโตต่อเนื่องถึงปี 2028 [1]

ความท้าทายสำคัญของธุรกิจนี้ไม่ได้อยู่เพียงการจัดการด้านการผลิตและขนส่ง แต่ยังรวมถึงการวิเคราะห์ความคิดเห็นของผู้บริโภคซึ่งมีปริมาณมากและหลากหลายทั้งด้านภาษา ความยาว และบริบท จนวิธีการแบบดั้งเดิมมีข้อจำกัดในการจับความหมายเชิงลึก งานวิจัยที่เกี่ยวข้องจึงได้นำเทคนิคการเรียนรู้เชิงลึก (Deep Learning) [2] และการทำเหมืองข้อความ (Text Mining) [3] มาประยุกต์ใช้ในการจำแนกความคิดเห็น โดยใช้แบบจำลอง RNN [4] Bi-RNN [5] GRU [6] และ Bi-GRU [7] งานวิจัยนี้จึงมุ่งประยุกต์ใช้เทคนิคดังกล่าวในการวิเคราะห์ความรู้สึกจากความคิดเห็นภาษาไทยต่อบริการส่งอาหารสุขภาพบนแพลตฟอร์มออนไลน์ เพื่อสนับสนุนการตัดสินใจเชิงกลยุทธ์ การปรับปรุงคุณภาพบริการ และเพิ่มองค์ความรู้ด้านการวิเคราะห์ความคิดเห็นภาษาไทยในบริบทของธุรกิจดิจิทัล

## 2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ผู้วิจัยได้ทำการรวบรวมข้อมูลเชิงทฤษฎีและผลงานวิจัยที่เกี่ยวข้อง เพื่อสร้างพื้นฐานความรู้และกรอบแนวคิดในการศึกษา โดยแบ่งออกเป็น 3 หัวข้อหลัก ดังนี้

### 2.1 การทำเหมืองข้อความ

เป็นกระบวนการแปลงข้อมูลข้อความที่ไม่มีโครงสร้างให้เป็นข้อมูลที่มีโครงสร้าง เพื่อให้สามารถระบุรูปแบบที่มีความหมายและสกัดข้อมูลเชิงลึกที่เป็นประโยชน์ โดยการทำเหมืองข้อความสามารถนำมาใช้วิเคราะห์ข้อมูลข้อความจำนวนมาก เพื่อค้นหาแนวคิดสำคัญ แนวโน้ม และความสัมพันธ์ที่ซ่อนอยู่ ผ่านการประยุกต์ใช้เทคนิคการวิเคราะห์ขั้นสูง เช่น เทคนิคการเรียนรู้เชิงลึก (Deep Learning)

### 2.2 เทคนิคการเรียนรู้เชิงลึก

ในการวิจัยครั้งนี้ ได้นำเทคนิคการเรียนรู้เชิงลึกมาใช้ในการสร้างแบบจำลองเพื่อวิเคราะห์ข้อมูลข้อความ โดยแบบจำลองที่ใช้ในการศึกษา ได้แก่

**2.2.1 แบบจำลองโครงข่ายประสาทเทียมเชิงลำดับ** (Recurrent Neural Network: RNN) เป็นอัลกอริทึมที่เหมาะสำหรับการประมวลผลข้อมูลเชิงลำดับ โดยสามารถจดจำข้อมูลในลำดับก่อนหน้าเพื่อนำมาใช้วิเคราะห์ข้อมูลในลำดับถัดไปได้ อย่างไรก็ตาม แบบจำลองดังกล่าวมีข้อจำกัดในการจัดการกับข้อมูลลำดับที่มีความยาวมาก ซึ่งอาจส่งผลต่อประสิทธิภาพของการเรียนรู้

**2.2.2 แบบจำลองโครงข่ายประสาทเทียมเชิงลำดับแบบสองทิศทาง** (Bidirectional Recurrent Neural Network: Bi-RNN) เป็นการพัฒนาต่อยอดจากแบบจำลองโครงข่ายประสาทเทียมเชิงลำดับ โดยทำการวิเคราะห์ข้อมูลจากทั้งทิศทางอดีตและอนาคตพร้อมกัน ส่งผลให้แบบจำลองสามารถเรียนรู้บริบทของข้อความได้อย่างครอบคลุมและแม่นยำยิ่งขึ้น

**2.2.3 แบบจำลองโครงข่ายประสาทเทียมแบบหน่วยวนซ้ำที่มีกลไกประตู** (Gated Recurrent Unit: GRU) เป็นแบบจำลองที่ได้รับการพัฒนาเพื่อแก้ไขข้อจำกัดของ RNN โดยใช้กลไกประตูในการควบคุมการไหลของข้อมูลภายในเครือข่าย ช่วยลดปัญหาการสูญหายของค่าความชัน (Vanishing Gradient) และเพิ่มประสิทธิภาพในการเรียนรู้ข้อมูลลำดับที่มีความยาวมาก

**2.2.4 แบบจำลองโครงข่ายประสาทเทียมแบบหน่วยวนซ้ำที่มีกลไกประตูแบบสองทิศทาง** (Bidirectional Gated Recurrent Unit: Bi-GRU) เป็นการผสานแนวคิดของ GRU เข้ากับโครงสร้างแบบสองทิศทาง ทำให้แบบจำลองสามารถเรียนรู้ข้อมูลจากทั้งอดีตและอนาคตได้พร้อมกัน ส่งผลให้การทำความเข้าใจบริบทของข้อความมีความแม่นยำสูงขึ้น และเหมาะสมกับงานวิเคราะห์ความคิดเห็นจากข้อมูลข้อความภาษาไทย

### 2.3 งานวิจัยที่เกี่ยวข้อง

Rangkuti et al. [8] ได้ศึกษาการใช้ RNN สำหรับวิเคราะห์ความรู้สึกรีวิวแอป Mobile Legends โดยจำแนกเป็น 3 ระดับ (ลบ กลาง บวก) ผ่านขั้นตอนเตรียมข้อมูล เช่น text cleaning, tokenization และ padding ก่อนสร้างโมเดลที่ประกอบด้วย Embedding, SimpleRNN และ Dense (softmax) ผลการทดลองมีความแม่นยำประมาณ 75% อย่างไรก็ตาม โมเดลมีแนวโน้มทำนายเป็นเชิงลบมากกว่าคลาสอื่น เนื่องจากปัญหาความไม่สมดุลของข้อมูล ซึ่งส่งผลต่อประสิทธิภาพโดยรวมของแบบจำลอง

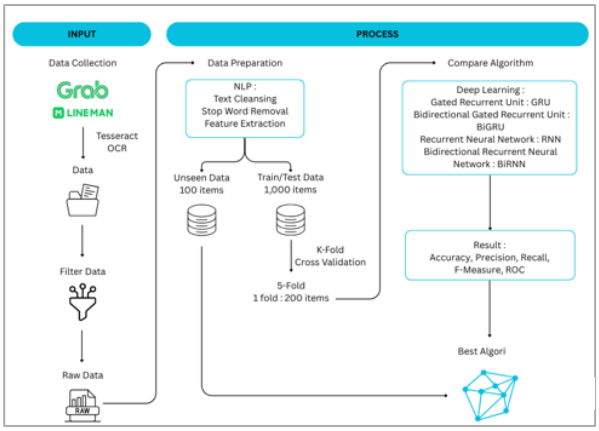
Anitha et al. [9] เสนอโมเดลวิเคราะห์ความรู้สึกโดยใช้ Bi-RNN ร่วมกับ Attention และ TF-IDF บนชุดข้อมูล Yelp ซึ่งช่วยให้โมเดลเข้าใจบริบทของข้อความทั้งสองทิศทางและเน้นข้อมูลสำคัญได้ดีขึ้น ผลการทดลองให้ความแม่นยำสูงถึง 93.85% และมีประสิทธิภาพดีกว่า RNN, LSTM และ SVM รวมถึงสามารถจัดการปัญหาความไม่สมดุลของข้อมูลได้อย่างมีประสิทธิภาพ

Pattanaik et al. [10] ได้ทำการศึกษาการเปรียบเทียบอัลกอริทึมจำแนกข้อมูลสำหรับตลาดแอป Google Play โดยใช้ข้อมูล เช่น รีวิวและเรตติ้ง พบว่าโมเดล GRU ให้ความแม่นยำสูงสุดประมาณ 89.45% สูงกว่า DTA, SVM, RF และ SGB อย่างชัดเจน สะท้อนว่าโมเดลแบบ recurrent สามารถเรียนรู้บริบทของข้อมูลลำดับและให้ประสิทธิภาพดีกว่าอัลกอริทึมแบบดั้งเดิมในการวิเคราะห์ข้อมูลรีวิว

Liangqiang et al. [11] ได้ทำการศึกษาการจำแนกความรู้สึกจากรีวิวร้านอาหารโดยใช้ Word2Vec ร่วมกับ Bi-GRU และ Attention เพื่อเรียนรู้บริบทของข้อความทั้งสองทิศทางและเน้นคำสำคัญในรีวิว ผลการทดลองบนข้อมูลจาก Dianping.com พบว่าโมเดลให้ค่า F1-score มากกว่า 89% และมีประสิทธิภาพดีกว่าวิธีแบบดั้งเดิม สะท้อนถึงศักยภาพของ Bi-GRU ร่วมกับ Attention ในการวิเคราะห์ข้อมูลข้อความที่ไม่มีโครงสร้างได้อย่างมีประสิทธิภาพ

## 3. วิธีการดำเนินการวิจัย

งานวิจัยนี้มีวัตถุประสงค์เพื่อเปรียบเทียบเทคนิคการเรียนรู้เชิงลึกที่เหมาะสมสำหรับการจำแนกข้อความการแสดงความคิดเห็น ที่มีต่อร้านอาหารสุขภาพบนแพลตฟอร์มออนไลน์ โดยมีกรอบแนวคิดแสดงดังภาพที่ 1




**ภาพที่ 1:** กรอบแนวคิด

จากภาพที่ 1 แสดงกรอบแนวคิดของงานวิจัยประกอบด้วย 2 ส่วน ได้แก่ การเก็บรวบรวมข้อมูลและการเตรียมข้อมูล และการพัฒนาตัวแบบและเปรียบเทียบประสิทธิภาพ

### 3.1 การเก็บรวบรวม และเตรียมข้อมูล

ผู้วิจัยได้รวบรวมความคิดเห็นของผู้ใช้บริการส่งอาหารสุขภาพจากแพลตฟอร์ม Grab และ LINE MAN โดยคัดเลือกร้านอาหารยอดนิยม 15 ร้านโดยอ้างอิงจากการจัดอันดับร้านอาหารยอดนิยมที่ปรากฏบนหน้าเว็บไซต์ทางการของแพลตฟอร์ม Grab และ LINE MAN เพื่อให้ได้ข้อมูลรีวิวที่สะท้อนประสบการณ์ผู้บริโภคจริงอย่างหลากหลาย บันทึกรีวิวในรูปแบบภาพและแปลงเป็นข้อความด้วย Tesseract OCR จากนั้นจัดประเภทความคิดเห็นเป็น 2 กลุ่ม คือ เชิงบวก (คะแนน $\geq$ 4.0) และเชิงลบ (คะแนน $\leq$ 2.0) กลุ่มละ 500 รายการ รวมชุดฝึก 1,000 รายการ และชุดทดสอบ Unseen Data อีก 100 รายการ (เชิงบวกและเชิงลบอย่างละ 50) รวมข้อมูลทั้งหมด 1,100 รายการ ตัวอย่างความคิดเห็นแสดงดังตารางที่ 1

**ตารางที่ 1 :** ตัวอย่างข้อมูลความคิดเห็น

| ความคิดเห็น | ประเภท |
|---|---|
| เป็นร้านอาหารที่สั่งบ่อยมากๆ ช่วงนี้อร่อยทุกเมนู อาหารลีน เลยอาจจะได้น้อยหน่อย ใครกินเยอะ สั่งสองกล่อง กำลังดี | Positive |
| ดีเลิศสะบัดดดด | Positive |

| ความคิดเห็น | ประเภท |
|---|---|
| น้อยมากทั้งข้าวและกับ | Negative |
| น้ำแกงส้มหกเละเทะ น่าจะใส่ถุงรัดยางดีกว่าใส่กล่องเลยนะคะ | Negative |

จากตารางที่ 1 ผู้วิจัยได้นำข้อมูลเข้าสู่กระบวนการเตรียมข้อมูลก่อนนำไปใช้ในการเปรียบเทียบแบบจำลอง โดยกระบวนการดังกล่าวประกอบด้วยขั้นตอนสำคัญดังต่อไปนี้

**3.1.1 การทำความสะอาดข้อมูล** เป็นการปรับปรุงคุณภาพของข้อมูลข้อความ โดยการลบอักขระพิเศษ กำจัดช่องว่าง การขึ้นบรรทัดใหม่ และสัญลักษณ์ที่ไม่เกี่ยวข้องกับเนื้อหา

**3.1.2 การตัดคำ** ผู้วิจัยได้ทำการเปรียบเทียบประสิทธิภาพของเทคนิคการตัดคำภาษาไทย ระหว่าง Attacut และ Deepcut เพื่อเลือกวิธีที่เหมาะสมที่สุดสำหรับการประมวลผลข้อความ เพื่อเตรียมข้อมูลสำหรับการวิเคราะห์

**3.1.3 การประเมินประสิทธิภาพของแบบจำลอง** ใช้วิธี Stratified K-Fold Cross-Validation โดยกำหนดค่า K = 5 เพื่อแบ่งชุดข้อมูลออกเป็น 5 ส่วนย่อย ซึ่งแต่ละส่วนมีสัดส่วนของคลาสข้อมูลใกล้เคียงกับชุดข้อมูลทั้งหมด ในแต่ละรอบจะใช้ข้อมูล 4 ส่วนสำหรับฝึกแบบจำลอง และอีก 1 ส่วนสำหรับทดสอบ จากนั้นนำผลลัพธ์จากทั้ง 5 รอบมาหาค่าเฉลี่ยเพื่อใช้เป็นตัวแทนประสิทธิภาพโดยรวมของแบบจำลอง

**3.1.4 การแปลงข้อความเป็นเวกเตอร์** งานวิจัยนี้ใช้เทคนิค Word2Vec ในการแปลงข้อความที่ผ่านการตัดคำแล้วให้อยู่ในรูปแบบเวกเตอร์ตัวเลข โดย Word2Vec จะเรียนรู้ความสัมพันธ์เชิงความหมายระหว่างคำจากบริบทที่ปรากฏในข้อความ ทำให้คำที่มีความหมายใกล้เคียงกันมีค่าเวกเตอร์อยู่ใกล้กันในปริภูมิเวกเตอร์ ซึ่งช่วยให้แบบจำลองการเรียนรู้เชิงลึกสามารถเข้าใจความสัมพันธ์ของคำในบริบทภาษาไทยได้อย่างมีประสิทธิภาพมากขึ้น

**3.2 การพัฒนาตัวแบบ และเปรียบเทียบประสิทธิภาพ**

งานวิจัยนี้ทำการเปรียบเทียบแบบจำลองด้วยเทคนิคการเรียนรู้เชิงลึกจำนวน 4 เทคนิค แบบจำลองโครงข่ายประสาทเทียมแบบวนซ้ำ แบบจำลองโครงข่ายประสาทเทียมแบบวนซ้ำสองทิศทาง แบบจำลองโครงข่ายประสาทเทียมแบบหน่วยวนซ้ำที่มีกลไกประตู และแบบจำลองโครงข่ายประสาทเทียมแบบหน่วยวนซ้ำที่มีกลไกประตูสองทิศทา โดยมีการปรับค่าพารามิเตอร์เพื่อใช้ในการเปรียบเทียบประสิทธิภาพตามที่แสดงในตาราง

**ตารางที่ 2 :** ค่าพารามิเตอร์ที่ใช้ในการปรับแต่งข้อมูล

| Parameter | Value |
|---|---|
| Memory Units | 32, 64, 128, 256, 512 |
| Dense Units | 32, 64, 128, 256, 512 |
| Dropout Rate | 0.1, 0.2, 0.3, 0.4, 0.5 |
| Epochs | 10, 20, 30, 40, 50 |

จากตารางที่ 2 ผู้วิจัยได้ทำการปรับค่าพารามิเตอร์จำนวน 625 รูปแบบ โดยกำหนดค่า Memory Units เท่ากับ [32, 64, 128, 256, 512] ค่า Dense Unit เท่ากับ [32, 64, 128, 256, 512] ค่า Dropout Rate เท่ากับ [0.1, 0.2, 0.3, 0.4, 0.5] และค่า Epochs เท่ากับ [10, 20, 30, 40, 50]

## 4. ผลการดำเนินงาน

ผู้วิจัยได้ทำการเปรียบเทียบวิธีการตัดคำภาษาไทยจำนวน 2 เทคนิค ได้แก่ Attacut และ Deepcut ผลการดำเนินงาน แสดงดังตารางที่ 3

**ตารางที่ 3 :** การเปรียบเทียบประสิทธิภาพของเทคนิคการตัดคำภาษาไทย Attacut และ Deepcut

| Performance | Attacut | Deepcut |
|---|---|---|
| Accuracy | 0.9250 | 0.9358 |
| Precision | 0.9533 | 0.9594 |
| Recall | 0.9629 | 0.9653 |
| F1-Score | 0.9567 | 0.9617 |
| Time | 8.533 s | 1.355 s |

จากนั้นทำการประเมินประสิทธิภาพของตัวแบบด้วยค่าความถูกต้อง (Accuracy: A) ค่าความแม่นยำ (Precision: P) ค่าระลึก (Recall :R) ค่าการถ่วงดุล (F1-Score: F) และค่าพื้นที่ใต้เส้นโค้งคุณลักษณะสมบัติการทำงาน (ROC AUC: ROC) ผลการดำเนินงาน แสดงดังตารางที่ 4 ถึง 10

**ตารางที่ 4 :** ผลลัพธ์ของตัวแบบ RNN

| Parameter | | | | Performance | | | | |
|---|---|---|---|---|---|---|---|---|
| Unit | Dense Unit | Drop-out | Epoch | A(%) | P | R | F | ROC |
| **128** | **32** | **0.1** | **50** | **85.70** | **0.86** | **0.86** | **0.86** | **0.90** |
| 128 | 64 | 0.5 | 30 | 83.10 | 0.83 | 0.83 | 0.83 | 0.87 |

| Parameter | | | | Performance | | | | |
|---|---|---|---|---|---|---|---|---|
| Unit | Dense Unit | Drop-out | Epoch | A(%) | P | R | F | ROC |
| 64 | 64 | 0.5 | 50 | 83.10 | 0.83 | 0.83 | 0.83 | 0.90 |
| 64 | 256 | 0.1 | 40 | 83.10 | 0.83 | 0.83 | 0.83 | 0.88 |
| 64 | 256 | 0.5 | 10 | 83.09 | 0.83 | 0.83 | 0.83 | 0.88 |
| 32 | 64 | 0.3 | 30 | 83.00 | 0.83 | 0.83 | 0.83 | 0.87 |
| 128 | 64 | 0.2 | 10 | 82.70 | 0.83 | 0.83 | 0.83 | 0.87 |
| 128 | 256 | 0.4 | 10 | 82.60 | 0.83 | 0.83 | 0.83 | 0.89 |
| 64 | 64 | 0.2 | 30 | 82.40 | 0.83 | 0.82 | 0.82 | 0.88 |
| 128 | 128 | 0.4 | 50 | 82.20 | 0.82 | 0.82 | 0.82 | 0.87 |

จากตารางที่ 4 แบบจำลอง RNN ที่ปรับค่าพารามิเตอร์ RNN Unit เท่ากับ 128 ค่า Dense Unit เท่ากับ 32 ค่า Dropout เท่ากับ 0.1 และค่า Epoch เท่ากับ 50 ให้ประสิทธิภาพดีที่สุดด้วยค่าความถูกต้อง ร้อยละ 85.70

**ตารางที่ 5 :** ผลลัพธ์ของตัวแบบ Bi-RNN

| Parameter | | | | Performance | | | | |
|---|---|---|---|---|---|---|---|---|
| Unit | Dense Unit | Drop-out | Epoch | A(%) | P | R | F | ROC |
| **64** | **32** | **0.3** | **50** | **89.40** | **0.89** | **0.89** | **0.89** | **0.95** |
| 64 | 128 | 0.5 | 50 | 89.20 | 0.89 | 0.89 | 0.89 | 0.94 |
| 64 | 64 | 0.3 | 50 | 89.10 | 0.89 | 0.89 | 0.89 | 0.95 |
| 32 | 64 | 0.1 | 30 | 89.10 | 0.89 | 0.89 | 0.89 | 0.95 |
| 32 | 32 | 0.1 | 20 | 89.00 | 0.89 | 0.89 | 0.89 | 0.95 |
| 32 | 64 | 0.4 | 40 | 88.80 | 0.89 | 0.89 | 0.89 | 0.95 |
| 32 | 256 | 0.4 | 50 | 88.80 | 0.89 | 0.89 | 0.89 | 0.95 |
| 256 | 128 | 0.1 | 30 | 88.70 | 0.89 | 0.89 | 0.89 | 0.94 |
| 64 | 32 | 0.4 | 20 | 88.60 | 0.89 | 0.89 | 0.89 | 0.94 |
| 256 | 64 | 0.1 | 30 | 88.50 | 0.89 | 0.88 | 0.88 | 0.95 |

จากตารางที่ 5 แบบจำลอง Bi-RNN ที่ปรับค่าพารามิเตอร์ Bi-RNN Unit เท่ากับ 64 ค่า Dense Unit เท่ากับ 32 ค่า Dropout เท่ากับ 0.3 และค่า Epoch เท่ากับ 50 ให้ประสิทธิภาพดีที่สุดด้วยค่าความถูกต้อง ร้อยละ 89.40

**ตารางที่ 6 :** ผลลัพธ์ของตัวแบบ GRU

| Parameter | | | | Performance | | | | |
|---|---|---|---|---|---|---|---|---|
| Unit | Dense Unit | Drop-out | Epoch | A(%) | P | R | F | ROC |
| **128** | **128** | **0.5** | **50** | **91.20** | **0.91** | **0.91** | **0.91** | **0.94** |
| 128 | 32 | 0.4 | 50 | 91.20 | 0.91 | 0.91 | 0.91 | 0.95 |
| 128 | 32 | 0.2 | 20 | 91.10 | 0.91 | 0.91 | 0.91 | 0.94 |
| 128 | 32 | 0.1 | 10 | 90.80 | 0.91 | 0.91 | 0.91 | 0.94 |
| 256 | 32 | 0.3 | 50 | 90.80 | 0.91 | 0.91 | 0.91 | 0.94 |
| 128 | 32 | 0.2 | 40 | 90.70 | 0.91 | 0.91 | 0.91 | 0.95 |
| 64 | 32 | 0.5 | 10 | 90.70 | 0.91 | 0.91 | 0.91 | 0.94 |
| 128 | 64 | 0.2 | 50 | 90.70 | 0.91 | 0.91 | 0.91 | 0.95 |
| 128 | 64 | 0.3 | 20 | 90.70 | 0.91 | 0.91 | 0.91 | 0.95 |
| 256 | 128 | 0.3 | 40 | 90.70 | 0.91 | 0.91 | 0.91 | 0.95 |

จากตารางที่ 6 แบบจำลอง GRU ที่ปรับค่าพารามิเตอร์ GRU Unit เท่ากับ 128 ค่า Dense Unit เท่ากับ 128 ค่า Dropout เท่ากับ 0.5 และค่า Epoch เท่ากับ 50 ให้ประสิทธิภาพดีที่สุดด้วยค่าความถูกต้อง ร้อยละ 91.20

**ตารางที่ 7 :** ผลลัพธ์ของตัวแบบ Bi-GRU

| Parameter | | | | Performance | | | | |
|---|---|---|---|---|---|---|---|---|
| Unit | Dense Unit | Drop-out | Epoch | A(%) | P | R | F | ROC |
| **64** | **32** | **0.2** | **20** | **91.30** | **0.91** | **0.91** | **0.91** | **0.96** |
| 64 | 64 | 0.2 | 10 | 91.00 | 0.91 | 0.91 | 0.91 | 0.96 |
| 64 | 128 | 0.3 | 10 | 91.00 | 0.91 | 0.91 | 0.91 | 0.96 |
| 64 | 64 | 0.1 | 40 | 90.90 | 0.91 | 0.91 | 0.91 | 0.96 |
| 32 | 256 | 0.4 | 10 | 90.90 | 0.91 | 0.91 | 0.91 | 0.96 |
| 32 | 256 | 0.4 | 40 | 90.90 | 0.91 | 0.91 | 0.91 | 0.96 |
| 32 | 256 | 0.4 | 50 | 90.80 | 0.91 | 0.91 | 0.91 | 0.96 |
| 64 | 256 | 0.3 | 30 | 90.80 | 0.91 | 0.91 | 0.91 | 0.96 |
| 64 | 64 | 0.2 | 50 | 90.80 | 0.91 | 0.91 | 0.91 | 0.96 |
| 64 | 64 | 0.3 | 20 | 90.70 | 0.91 | 0.91 | 0.91 | 0.96 |

จากตารางที่ 7 แบบจำลอง Bi-GRU ที่ปรับค่าพารามิเตอร์ Bi-GRU Unit เท่ากับ 64 ค่า Dense Unit เท่ากับ 32 ค่า Dropout เท่ากับ 0.2 และค่า Epoch เท่ากับ 20 ให้ประสิทธิ ภาพดีที่สุดด้วยค่าความถูกต้อง ร้อยละ 91.30

**ตารางที่ 8 :** ผลการเปรียบเทียบประสิทธิภาพของแบบจำลอง

| Technique | Performance | | | | |
|---|---|---|---|---|---|
| | A(%) | P | R | F | ROC |
| RNN | 85.70 | 0.86 | 0.86 | 0.86 | 0.90 |
| Bi-RNN | 89.40 | 0.89 | 0.89 | 0.89 | 0.95 |
| GRU | 91.20 | 0.91 | 0.91 | 0.91 | 0.94 |
| **Bi-GRU** | **91.30** | **0.91** | **0.91** | **0.91** | **0.96** |

จากตารางที่ 8 แบบจำลอง Bi-GRU ให้ประสิทธิภาพสูงสุดนั้น เนื่องมาจากการผสานจุดเด่นของ GRU ซึ่งมีกลไกประตูช่วยลดปัญหา Vanishing Gradient เข้ากับโครงสร้างแบบสองทิศทาง ทำให้โมเดลสามารถเรียนรู้บริบทของข้อความจากทั้งทิศทางก่อนและหลังพร้อมกัน ซึ่งเหมาะสมกับลักษณะของข้อความความคิดเห็นภาษาไทยที่มีบริบทสำคัญทั้งต้นและท้ายประโยค ในขณะที่ GRU แบบทิศทางเดียวอ่านข้อความในทิศทางเดียว จึงอาจพลาดบริบทบางส่วน และ RNN/Bi-RNN มีข้อจำกัดในการจัดการลำดับที่ยาวกว่า

**ตารางที่ 9:** ตัวอย่างผลลัพธ์การวิเคราะห์ที่ 1 ด้วย Unseen Data

| Item | Result | |
|---|---|---|
| Text | เป็นร้านคลีนที่รสชาติดีจริง | |
| Analysis | Positive | |
| Accuracy | 98.91% | |
| **Category** | **Keywords** | **Status** |
| Product | 2 | YES |
| Service & Delivery | 0 | NO |
| Value & Portion | 0 | NO |
| Specific Error/Missing Item | 0 | NO |

**ตารางที่ 10 :** ตัวอย่างผลลัพธ์การวิเคราะห์ที่ 2 ด้วย Unseen Data

| Item | Result | |
|---|---|---|
| Text | ราคาแพงเมื่อเทียบกับปริมาณ | |
| Analysis | Negative | |
| Accuracy | 95.99% | |
| **Category** | **Keywords** | **Status** |
| Product | 0 | NO |
| Service & Delivery | 0 | NO |
| **Category** | **Keywords** | **Status** |
| Value & Portion | 2 | YES |
| Specific Error/Missing Item | 0 | NO |

จากตารางที่ 9 และ 10 การแสดงผลลัพธ์การวิเคราะห์ความคิดเห็นด้วยแบบจำลอง Bi-GRU ที่ปรับค่าพารามิเตอร์แล้ว ให้ประสิทธิภาพในการวิเคราะห์ความคิดเห็นได้อย่างมีประสิทธิภาพ โดยชุดข้อมูล Unseen Data จำนวน 100 รายการ ประกอบด้วยความคิดเห็นเชิงบวก 50 รายการ และเชิงลบ 50 รายการ ผลการทดสอบพบว่าแบบจำลองสามารถจำแนกได้ถูกต้อง 91 รายการจาก 100 รายการ คิดเป็นร้อยละ 91.00 โดยข้อความที่จำแนกผิดพลาดส่วนใหญ่มีลักษณะกำกวมหรือประกอบด้วยคำที่สามารถตีความได้หลายบริบท เช่น คำที่มีความหมายเชิงบวกในบางบริบทแต่เชิงลบในบริบทอื่น ซึ่งอาจส่งผลต่อการเรียนรู้ของแบบจำลองได้

## 5. สรุป

งานวิจัยนี้มีวัตถุประสงค์เพื่อวิเคราะห์ความคิดเห็นของผู้ใช้บริการร้านอาหารสุขภาพบนแพลตฟอร์มออนไลน์ โดยประยุกต์ใช้เทคนิคการเรียนรู้เชิงลึก ในการจำแนกความคิดเห็น โดยใช้ข้อมูลทั้งหมด 1,100 รายการ แบ่งเป็นชุดฝึก 1,000 รายการ (เชิงบวกและเชิงลบอย่างละ 500) และชุดทดสอบ 100 รายการ (เชิงบวกและเชิงลบอย่างละ 50) ข้อมูลผ่านกระบวนการทำความสะอาด การตัดคำด้วย Attacut และ Deepcut และการแปลงข้อความเป็นลำดับตัวเลข จากนั้นเปรียบเทียบประสิทธิภาพของแบบจำลอง 4 เทคนิค ได้แก่ RNN, Bi-RNN, GRU และ Bi-GRU ผลการวิจัยพบว่า Bi-GRU ให้ประสิทธิภาพสูงสุด โดยมีค่า Accuracy ร้อยละ 91.30 ค่า Precision, Recall และ F1-score เท่ากับ 0.91 และค่า ROC เท่ากับ 0.96 และเมื่อทดสอบด้วยชุดข้อมูล Unseen Data พบว่าแบบจำลอง Bi-GRU สามารถจำแนกได้ถูกต้อง 91 รายการจาก 100 รายการ

## เอกสารอ้างอิง


[1] Statista. (2024). Online Food Delivery – Worldwide. Retrieved from https://www.statista.com/outlook/emo/online-food-delivery/worldwide

[2] Chitboonyapinij, C., Maneerat, P., & Jirawichitchai, N. (2022). การพัฒนาแบบจำลองวิเคราะห์ ความรู้สึกสำหรับสื่อสังคมออนไลน์ภาษาไทยโดยใช้เทคนิคการเรียนรู้เชิงลึก. Huachiew Chalermprakiet University Journal of Science and Technology, 8(2), 8–18.

[3] Mienye, I. D., Swart, T. G., & Obaido, G.: Recurrent Neural Networks: A comprehensive review of architectures, variants, and applications. Information, 15(9), 517(2024).

[4] Chowdhury, S., et al.: A multitask bi-directional RNN model for named entity recognition on Chinese electronic medical records. BMC Bioinformatics, 19(Suppl 17), 499(2018). https://doi.org/10.1186/s12859-018-2467-9

[5] Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y.: Learning Phrase Representations using RNN Encoder–Decoder for Statistical Machine Translation. Proceedings of the 2014 Confer-ence on Empirical Methods in Natural Language Processing (EMNLP), 1724–1734 (2014).

[6] Zhou, P., Shi, W., Tian, J., Qi, Z., Li, B., Hao, H., & Xu, B.: Attention-Based Bidirectional Long Short-Term Memory Networks for Relation Classification. Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (ACL), 207–212 (2016).

[7] Phatthiyaphaibun, W., & Boonkwan, P.: Thai word segmentation using deep learning techniques: A review. Journal of Electrical and ComputerEngineering, Article ID 8592214 (2023).

[8] Rangkuti, N. I., and Lubis, I. (2026). A recurrent neural network–based sentiment analysis of Mobile Legends app reviews. Jurnal Kecerdasan Buatan dan Teknologi Informasi, 5(1), 11–17. https://doi.org/10.69916/jkbti.v5i1.390

[9] Anitha, S., & Gnanasekaran, P. (2025). "Sentimental analysis using bi-directional RNN with attention mechanism based on term frequency–inverse document frequency". Journal of Information Systems Engineering & Management, 10(51s), 165–181. https://doi.org/10.52783/jisem.v10i51s.10377

[10] Parastoo Golpour, Majid Ghayour-Mobarhan, Azadeh Saki , Habibollah Esmaily, Ali Taghipour, Mohammad Tajfard, Hamideh Ghazizadeh, Mohsen Moohebati and Gordon A. Ferns. Comparison of Support Vector Machine, Naïve Bayes and Logistic Regression for Assessing the Necessity for Coronary Angiography. [ออนไลน์]. [สืบค้นวันที่ 6 มิถุนายน 2568]. https://pubmed.ncbi.nlm.nih.gov/32899733/.

[11] Liangqiang Li, Liang Yang and Yuyang Zeng. Improving Sentiment Classification of Restaurant Reviews with Attention-Based Bi-GRU Neural Network. [ออนไลน์]. [สืบค้น วันที่ 6 มิถุนายน 2568]. https://www.mdpi.com/2073-8994/13/8/1517.

# การพยากรณ์ความต้องการด้านสวัสดิการด้วยเทคนิคเหมืองข้อมูล
# Forecasting Welfare Demand Using Data Mining Techniques

*มณู บุญปลื้ม (Manu Boonpleum)[1] และวัชรีวรรณ จิตต์สกุล (Watchareewan Jitsakul)[2]*
*ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล*
*มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ*
*[1]s6707021856167@email.kmutnb.ac.th, [2]watchareewan.j@itd.kmutnb.ac.th*

## บทคัดย่อ

*งานวิจัยนี้มีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพของเทคนิคเหมืองข้อมูลสำหรับใช้ในการพยากรณ์ความต้องการด้านสวัสดิการของบุคลากร ซึ่งใช้ชุดข้อมูลจำนวน 330 รายการ และมีตัวแปรทั้งหมด 37 ตัวแปร แบ่งเป็นข้อมูลคุณลักษณะด้านประชากรจำนวน 14 ตัวแปร และข้อมูลความต้องการอิงตามทฤษฎีลำดับขั้นความต้องการของมาสโลว์ 5 ขั้น ที่เป็นรากฐานสำคัญในอธิบายแรงจูงใจครั้งนี้ ซึ่งมาจากการเก็บรวบรวมข้อมูล โดยผ่านการประเมินจากผู้ทรงคุณวุฒิและหาค่าความเชื่อมั่นเรียบร้อย นำไปผ่านการทำความสะอาดข้อมูล แปลงค่าเตรียมสำหรับการฝึกแบบจำลอง และเนื่องจากข้อจำกัดด้านข้อมูลที่น้อยทำให้ต้องมีการแก้ปัญหาข้อมูลไม่สมดุลด้วยเทคนิคการสุ่มเพิ่มข้อมูลกลุ่มน้อยด้วยวิธีการสังเคราะห์เพื่อช่วยให้โมเดลเรียนรู้ลักษณะของกลุ่มน้อยได้ดีขึ้น จากนั้นทำการแบ่งข้อมูลเพื่อทดสอบด้วยวิธีการตรวจสอบแบบไขว้ 5 ส่วน เพื่อเปรียบเทียบเทคนิคการเรียนรู้ของเครื่อง 3 อัลกอริทึม ได้แก่ ซัพพอร์ตเวกเตอร์แมชชีน เอกซ์ตรีมเกรเดียนต์บูสติง และ โครงข่ายประสาทเทียม ผลการวิจัยพบว่าแบบจำลองที่ใช้เทคนิคซัพพอร์ตเวกเตอร์แมชชีนมีประสิทธิภาพสูงสุด หลังจากทดสอบโดยใช้ข้อมูลที่ไม่ผ่านการฝึกฝนแบบจำลอง โดยมีค่าความถูกต้อง เท่ากับ ร้อยละ 94.67 ค่าความแม่นยำ เท่ากับ 0.948 ค่าความระลึก เท่ากับ 0.947 และค่าการถ่วงดุล เท่ากับ 0.946 ตามลำดับ*

**คำสำคัญ:** ความต้องการด้านสวัสดิการ ทฤษฎีลำดับขั้นความต้องการของมาสโลว์ ซัพพอร์ตเวกเตอร์แมชชีน เอกซ์ตรีมเกรเดียนต์บูสติง โครงข่ายประสาทเทียม

## Abstract

*This research aims to compare the performance of data mining techniques for forecasting employee welfare requirements. The study utilizes a dataset of 330 records consisting of 37 variables. These variables include 14 demographic attributes and 5 levels of needs based on Maslow's Hierarchy of Needs, which serves as the foundational framework for explaining human motivation in this study. The data collection instrument was validated by experts and underwent reliability testing. The methodology involved data cleaning and transformation for model training. Due to the limitation of a small initial dataset, the Synthetic Minority Over-sampling Technique (SMOTE) was applied to address class imbalance, enabling the models to better learn the characteristics of minority groups. Performance was evaluated using 5-fold cross-validation to compare three machine learning algorithms: Support Vector Machine (SVM), Extreme Gradient Boosting (XGBoost), and Artificial Neural Networks (ANN). The results indicate that the Support Vector Machine (SVM) model achieved the highest performance when evaluated on unseen data. The SVM model yielded an accuracy of 94.67%, a precision of 0.948, a recall of 0.947, and an F1-score of 0.946.*

**Keywords:** Welfare Demand, Maslow's Hierarchy of Needs, Support Vector Machine (SVM), Extreme Gradient Boosting (XGBoost), Artificial Neural Network (ANN).

## 1. บทนำ

ในปัจจุบันสถานการณ์ทางเศรษฐกิจและสภาพสังคมที่มีการเปลี่ยนแปลงอย่างรวดเร็วส่งผลให้ความต้องการของบุคลากรมีความแตกต่างและหลากหลายมากขึ้น ซึ่งการจัดสรรทรัพยากรด้านสวัสดิการเพื่อรักษาขวัญกำลังใจบุคลากรแบบเดียวกัน แต่ใช้ร่วมกับทุกคน (One-size-fits-all) อาจไม่ตอบสนองความต้องการที่แท้จริง โดยเฉพาะกลุ่มคนรุ่นใหม่ที่มีการทำงานซ้ำซากและจำเจมักทำให้เกิดความรู้สึกเบื่อหน่ายจนเกิดเป็นภาวะเบื่อหน่ายเรื้อรัง (Boreout Syndrome) อาจทำให้ต้องการลาออกจากงานได้

ปัญหาหลักมักเกิดจากเจ้าหน้าที่ขาดการรวบรวมข้อมูลความต้องการที่แท้จริงของบุคลากร การขาดเครื่องมือในการคาดการณ์ความต้องการสวัสดิการต่าง ๆ ส่งผลให้บริหารจัดสรรงบประมาณไม่ตรงกับความต้องการใช้ทรัพยากรไม่คุ้มค่าและเสียไปอย่างเปล่าประโยชน์ ความพึงพอใจและประสิทธิภาพในการทำงานลดลง เทคนิคเหมืองการทำเหมืองข้อมูลที่เป็นเครื่องมือในการวิเคราะห์ข้อมูลเพื่อค้นหารูปแบบและความสัมพันธ์ที่ซ่อนอยู่ จึงมีความสำคัญในการพยากรณ์ความต้องการสวัสดิการ โดยมีความแม่นยำสูง รวมถึงสามารถแสดงค่าในรูปแบบที่หลากหลาย

ดังนั้น การวิจัยนี้จึงรวบรวมข้อมูลความต้องการจากบุคลากรในองค์การและการนำข้อมูลที่ได้มาวิเคราะห์ด้วยเทคนิคเหมืองข้อมูลนั้น จึงมีความสำคัญต่อการพัฒนาการพยากรณ์ความต้องการด้านสวัสดิการโดยประยุกต์ใช้การวิเคราะห์ข้อมูลด้านประชากรศาสตร์ของบุคลากรเพื่อจำแนกกลุ่มความต้องการด้านสวัสดิการและช่วยให้องค์การสามารถปรับปรุงและพัฒนารูปแบบการจัดสวัสดิการให้เหมาะสมและตรงใจบุคลากรมากยิ่งขึ้น ซึ่งจะเป็นการเพิ่มคุณภาพชีวิต สร้างขวัญกำลังใจ และประสิทธิภาพการทำงานของบุคลากร รวมทั้งจะเป็นประโยชน์ต่อการตัดสินใจเชิงกลยุทธ์ขององค์กรในการบริหารทรัพยากรบุคคล ลดความสิ้นเปลืองทรัพยากรจากการจัดสวัสดิการที่ไม่ตรงกลุ่ม และช่วยสร้างความผูกพันของบุคลากรต่อองค์การในระยะยาว

## 2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ผู้วิจัยได้ทำการรวบรวมข้อมูลเชิงทฤษฎีและผลงานวิจัยที่เกี่ยวข้อง เพื่อสร้างพื้นฐานความรู้และกรอบแนวคิดในการศึกษา โดยแบ่งออกเป็น 3 หัวข้อหลัก ดังนี้

### 2.1 ทฤษฎีลำดับขั้นความต้องการของมาสโลว์

เป็นแนวคิดทางจิตวิทยาที่เสนอว่า มนุษย์จะถูกกระตุ้นให้เติมเต็มความต้องการขั้นพื้นฐานในลำดับต้นก่อน ถึงจะมีความสนใจที่จะพัฒนาความต้องการในขั้นถัดไป โดยเพิ่มขึ้นตามลำดับจากชั้นล่างสู่ชั้นบน โดยมีความต้องการสวัสดิการทั้งหมด 5 ขั้น ได้แก่ ทางกายภาพ ความมั่นคงปลอดภัย ความรักและความเป็นเจ้าของ การได้รับการยอมรับ และการตระหนักรู้ในศักยภาพตนเอง [1]

### 2.2 เทคนิคการเรียนรู้ของเครื่อง

ในการวิจัยครั้งนี้ ได้มีการนำเทคนิคการเรียนรู้ของเครื่องมาทำการสร้างแบบจำลอง ได้แก่

**2.2.1 อัลกอริทึมซัพพอร์ตเวกเตอร์แมชชีน** (Support Vector Machine: SVM) เป็นอัลกอริทึมในกลุ่ม Supervised Learning ที่ใช้สำหรับการจำแนกประเภทข้อมูล และวิเคราะห์การถดถอย เน้นการหา “เส้นแบ่ง” (Hyperplane) ที่ดีที่สุดเพื่อแยกข้อมูลออกจากกัน โดยพยายามให้ระยะห่าง (Margin) ระหว่างกลุ่มข้อมูลกว้างที่สุด [2]

**2.2.2 อัลกอริทึมเอกซ์ตรีมเกรเดียนต์บูสติง** (Extreme Gradient Boosting: XGBoost) เป็นอัลกอริทึมที่พัฒนาต่อยอดมาจาก Gradient Boosting Machine โดยเพิ่มประสิทธิภาพด้านความเร็วและป้องกันการเรียนรู้เพื่อลดปัญหาการ Overfiitting โดยสร้างแบบจำลองแบบอนุกรม โดยแต่ละแบบจำลองใหม่ที่ถูกสร้างขึ้นจะพยายามแก้ไขความผิดพลาดจากแบบจำลองก่อนหน้า [3]

**2.2.3 อัลกอริทึมโครงข่ายประสาทเทียม** (Artificial Neural Network: ANN) แบบจำลองทางคณิตศาสตร์ที่มีแนวคิดมาจากโครงสร้างและการทำงานของเซลล์ประสาทในสมองมนุษย์ โดยประกอบด้วยหน่วยประมวลผลย่อย ๆ ที่เชื่อมต่อกันเพื่อเรียนรู้รูปแบบ จากข้อมูลผ่านการปรับค่าน้ำหนัก มีโครงสร้างแบ่งออกเป็น 3 ชั้น ประกอบด้วย Input Layer, Hidden Layer และ Output Layer [4]

**2.3 งานวิจัยที่เกี่ยวข้อง**

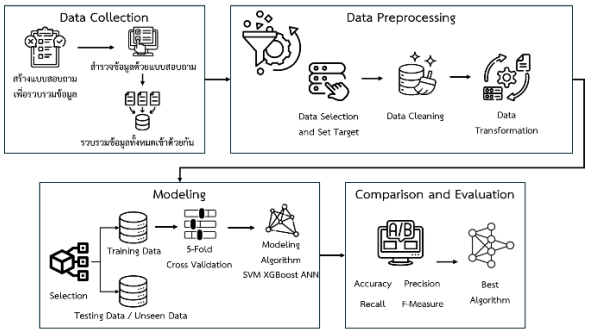
Alici and et.al. [5] ศึกษาเพื่อปรับปรุงแบบสอบถามที่อิงตามทฤษฎีลำดับขั้นความต้องการของมาสโลว์สำหรับผู้อยู่อาศัยในบ้านพักคนชราโดยใช้เทคนิคการเรียนรู้ของเครื่อง โดยมีวัตถุประสงค์เพื่อระบุคำถามที่สำคัญที่สุดในการประเมินความต้องการของผู้สูงอายุและลดความซ้ำซ้อนของแบบสอบถาม ผู้วิจัยใช้หลายเทคนิค รวมถึง ANN, SVM, Gaussian Process Regression และ Logistic Regression เพื่อวิเคราะห์ความสำคัญของคำถามแต่ละข้อและจัดลำดับความสำคัญตามระดับความต้องการของมาสโลว์ ผลการศึกษาแสดงให้เห็นว่าเทคนิคเรียนรู้ของเครื่องสามารถระบุคำถามที่สำคัญและช่วยปรับปรุงแบบสอบถามให้มีประสิทธิภาพมากขึ้น โดยลดจำนวนคำถามลงแต่ยังคงความสามารถในการประเมินความต้องการของผู้สูงอายุได้อย่างครบถ้วน

Yang and Zhao [6] ศึกษาเพื่อประเมินเส้นทางการพัฒนาทรัพยากรแรงงานผู้สูงอายุในเมืองและปัจจัยที่เกี่ยวข้อง รวมถึงความต้องการด้านสวัสดิการ โดยอิงตามทฤษฎีความต้องการของมาสโลว์ ผู้วิจัยใช้โมเดล Synthetic Minority Over-sampling Technique ร่วมกับ Extreme Gradient Boosting (SMOTE-XGBoost) เพื่อจัดการกับปัญหาข้อมูลไม่สมดุลและสร้างโมเดลพยากรณ์ที่มีประสิทธิภาพสูง ร่วมกับการวิเคราะห์ decision tree เพื่อระบุปัจจัยสำคัญ ผลการศึกษาพบว่า โมเดล SMOTE-XGBoost สามารถพยากรณ์และประเมินปัจจัยที่มีผลต่อการพัฒนาแรงงานผู้สูงอายุได้อย่างมีประสิทธิภาพ สอดคล้องกับระดับความต้องการของมาสโลว์

Ilunga and et.al. [7] ศึกษาเพื่อวิเคราะห์ปัจจัยที่มีผลต่อความพึงพอใจในการทำงานของพนักงานและสร้างแบบจำลองพยากรณ์โดยใช้เทคนิคการเรียนรู้ของเครื่องเพื่อช่วยองค์กรในการเข้าใจและปรับปรุงสวัสดิการพนักงานให้ตรงกับความต้องการ ผู้วิจัยใช้ชุดข้อมูลจำนวน 9,516 รายการ และทดสอบหลายอัลกอริทึม ได้แก่ Naïve Bayes, Random Forest, SVM และ Gradient Boosted Trees หลังจากลดมิติข้อมูลเหลือ 33 ตัวแปรที่สำคัญที่สุด ผลการวิจัยพบว่า SVM และ Gradient Boosted Trees ให้ความแม่นยำสูงสุดถึง 99% ในการพยากรณ์ความพึงพอใจของพนักงาน โดยปัจจัยเชื่อมโยงกับทฤษฎีมาสโลว์

## 3. วิธีการดำเนินการวิจัย

งานวิจัยนี้มีวัตถุประสงค์เพื่อเปรียบเทียบเทคนิคการเรียนรู้ของเครื่องสำหรับการพยากรณ์ความต้องการด้านสวัสดิการ โดยมีกรอบแนวคิดแสดงดังภาพที่ 1




**ภาพที่ 1:** กรอบแนวคิด

จากภาพที่ 1 แสดงกรอบแนวคิดของงานวิจัยประกอบด้วย 4 ส่วนหลัก ได้แก่ การรวบรวมข้อมูล การเตรียมข้อมูล การพัฒนาแบบจำลอง และการเปรียบเทียบประสิทธิภาพ

**3.1 การรวบรวมข้อมูล**

ทำการศึกษาค้นคว้าปัจจัยที่อาจมีผลต่อพฤติกรรมความต้องการสวัสดิการต่าง ๆ ทั้งคุณลักษณะประชากรและความต้องการอิงตามทฤษฎีลำดับขั้นความต้องการของมาสโลว์ 5 ขั้น

**3.1.1 การสร้างแบบสอบถาม** รับการประเมินจากผู้ทรงคุณวุฒิ 5 ท่าน และหาค่าความเชื่อมั่นด้วยโปรแกรม Jamovi มากกว่า 0.7 จากนั้นจัดทำแบบสอบถามลงใน Google Forms โดยให้มีทั้งรูปแบบตัวเลือก การให้คะแนน (Rating scale) รวมถึงกรอกรายละเอียดข้อความเพื่อให้ข้อเสนอแนะที่เกี่ยวข้อง

**3.1.2 การสำรวจข้อมูลด้วยแบบสอบถาม** ทำการส่งต่อลิงก์เว็บไซต์และ QR Code ของแบบสอบถามบน Google Forms ให้กับหัวหน้าแต่ละหน่วยงานเพื่อช่วยประชาสัมพันธ์ให้บุคลากรกรอกข้อมูลแบบสอบถาม

**3.2 การเตรียมข้อมูล**

ขั้นตอนหลังการสำรวจข้อมูล ทำการโหลดข้อมูลเป็นไฟล์ .csv เพื่อให้สะดวกต่อการจัดเตรียมข้อมูลและการนำไปประมวลผลด้วยเทคนิคเหมืองข้อมูล

**3.2.1 การคัดเลือกข้อมูลและกำหนดค่าเป้าหมาย** คัดเลือกชุดข้อมูลหรือปัจจัยที่อาจมีผลต่อการวิเคราะห์ข้อมูล รวมทั้งเฉลี่ยค่าตัวเลขจากชุดคำถามในแต่ละขั้นความต้องการทั้ง 5 ขั้น เพื่อให้ได้ค่าเฉลี่ยและกำหนดเป็นกลุ่มเป้าหมาย (Label) จากค่าที่มากที่สุดในชุดข้อมูลเดียวกัน แต่หากค่าเฉลี่ยที่มากที่สุดมีจำนวนเท่ากันจะทำการเลือกขั้นที่ต่ำกว่าเป็นค่าเป้าหมาย อิงตามทฤษฎีลำดับขั้นความต้องการของมาสโลว์

**3.2.2 การทำความสะอาดข้อมูล** การกำหนดให้ผู้ตอบแบบสอบถามต้องทำการกรอกทุกค่าจึงจะสามารถส่งแบบสอบถามเมื่อตอบคำถามครบแล้วเท่านั้น ทำให้ผลลัพธ์ครบถ้วน เหลือเพียงข้อเสนอแนะหรือค่าจากบางส่วนงานที่ไม่ได้ร่วมตอบแบบสอบถามทำให้ต้องตัดบางปัจจัยออกไปเพื่อให้ได้ค่าที่จะนำมาวิเคราะห์ได้

**3.2.3 การแปลงค่าข้อมูล** แปลงข้อมูลจากรูปแบบตัวอักษรหรือช่วงของข้อมูลให้อยู่ในรูปแบบค่าเชิงตัวเลข โดยกำหนดค่าพร้อมจัดทำป้ายกำกับข้อมูล เพื่อให้สะดวกต่อการจดจำค่าข้อมูลต่าง ๆ เพื่อรองรับต่อการสร้างและพัฒนาแบบจำลองอัลกอริทึม

**3.2.4 การปรับสมดุลข้อมูล** เนื่องจากกลุ่มที่มีความต้องการสวัสดิการทางกายภาพ มีสัดส่วนจำนวนมากกว่ากลุ่มที่มีความต้องการสวัสดิการในขั้นอื่น ๆ จึงได้นำเทคนิคการสุ่มเพิ่มข้อมูลกลุ่มน้อยด้วยวิธีการสังเคราะห์ (SMOTE) มาใช้ปรับสมดุลของชุดข้อมูลสำหรับสร้างแบบจำลอง โดยครั้งนี้ จำนวนผู้ที่ต้องการสวัสดิการทางกายภาพ ในขั้นที่ 1 มี 121 ชุดข้อมูล เพิ่มเป็น 200 ชุดข้อมูล และทำการเพิ่มข้อมูลจากกลุ่มผู้ที่ต้องการสวัสดิการในขั้นอื่น ๆ ประกอบด้วย ขั้นที่ 2 58 ชุดข้อมูล ขั้นที่ 3 79 ชุดข้อมูล ขั้นที่ 4 26 ชุดข้อมูล และขั้นที่ 5 46 ชุดข้อมูล ให้มีจำนวนเพิ่มเป็น 200 ชุดข้อมูลเท่ากันทั้งหมด ส่งผลให้มีจำนวนชุดข้อมูลจากเดิม 330 ชุดข้อมูล เพิ่มเป็นทั้งหมด 1,000 ชุดข้อมูล โดยเป็นชุดข้อมูลที่ถูกสังเคราะห์ขึ้นใหม่จำนวน 670 ชุดข้อมูล

**3.2.5 การแบ่งชุดข้อมูล** ทำการแบ่งข้อมูลทั้งหมดเป็น 2 ชุด คือ ชุดข้อมูลสำหรับการฝึกสอน จำนวน 900 ชุดข้อมูล และข้อมูลสำหรับทดสอบที่ไม่แบบจำลองอัลกอริทึมไม่เคยพบมาก่อน จำนวน 100 ชุดข้อมูล

**3.3 การสร้างแบบจำลองอัลกอริทึม**

**3.3.1 การประเมินประสิทธิภาพของแบบจำลอง** ใช้วิธี Stratified K-Fold Cross-Validation โดยกำหนดค่า (k = 5) เพื่อแบ่งชุดข้อมูลออกเป็น 5 ส่วนย่อย ซึ่งในแต่ละ fold จะมีสัดส่วนของคลาสข้อมูลใกล้เคียงกับสัดส่วนของข้อมูลทั้งหมด วิธีการดังกล่าวช่วยลดอคติที่อาจเกิดจากการสุ่มแบ่งข้อมูลแบบทั่วไป และทำให้การประเมินผลมีความน่าเชื่อถือมากยิ่งขึ้น โดยในแต่ละรอบจะใช้ข้อมูลจำนวน 4 folds สำหรับฝึกแบบจำลอง และใช้ข้อมูลอีก 1 fold สำหรับทดสอบ จากนั้นทำการวนซ้ำจนครบทั้ง 5 folds และนำผลลัพธ์ที่ได้จากทุก fold มาหาค่าเฉลี่ยเพื่อใช้เป็นตัวแทนประสิทธิภาพโดยรวมของแบบจำลอง

**3.3.2 การพัฒนาตัวแบบ และเปรียบเทียบประสิทธิภาพ** ทำการสร้างแบบจำลองด้วยเทคนิคการเรียนรู้ของเครื่อง 3 อัลกอริทึม ได้แก่ SVM, XGBoost และ ANN โดยทำการปรับค่าพารามิเตอร์สำหรับแบบจำลอง แสดงดังตารางที่ 2 สำหรับ SVM, ตารางที่ 3 สำหรับ XGBoost และตารางที่ 4 สำหรับ ANN

**ตารางที่ 2 :** ค่าพารามิเตอร์ที่ใช้ในการปรับแต่งสำหรับ SVM

| Parameter | Value |
|---|---|
| C (Regularization) | 0.1, 1, 10, 100 |
| Gamma (γ) | 0.001, 0.01, 0.1, 1 |
| Kernel | RBF |

จากตารางที่ 2 ผู้วิจัยได้ทำการปรับค่าพารามิเตอร์จำนวน 80 รูปแบบ โดยกำหนดค่า C เท่ากับ [0.1, 1, 10, 100] ค่า Gamma เท่ากับ [0.001, 0.01, 1] และค่า Kernel เท่ากับ RBF

**ตารางที่ 3 :** ค่าพารามิเตอร์ที่ใช้ในการปรับแต่งสำหรับ XGBoost

| Parameter | Value |
|---|---|
| Number of Boosting Trees (n_estimators) | 800, 1000 |
| Learning Rate | 0.01, 0.05 |
| Maximum Tree Depth | 3, 6, 8 |
| Minimum Child Weight | 1, 2, 4 |
| Subsample Ratio of Training Instances | 0.1, 0.5, 1 |
| Column Subsample Ratio per Tree | 0.01, 0.05, 1 |
| Alpha | 0, 0.1 |
| Lambda | 0, 1 |
| Gamma | 0, 1 |

จากตารางที่ 3 ผู้วิจัยได้ทำการปรับค่าพารามิเตอร์จำนวน 12,960 รูปแบบ โดยกำหนดค่า Number of

Boosting Trees เท่ากับ [800, 1000] ค่า Learning Rate เท่ากับ [0.01, 0.05] ค่า Maximum Tree Depth เท่ากับ [3, 6, 8] ค่า Minimum Child Weight เท่ากับ [1, 2, 4] ค่า Subsample Ratio of Training Instances เท่ากับ [0.1, 0.5, 1] ค่า Column Subsample Ratio per Tree เท่ากับ [0.01, 0.05, 1] ค่า Alpha [0, 0.1] เท่ากับ Lambda [0, 1] และค่า Gamma เท่ากับ [0, 1]

**ตารางที่ 4 :** ค่าพารามิเตอร์คงที่สำหรับ ANN

| Parameter | Value |
|---|---|
| Structures | [(16, 8), (32, 16, 8)] |
| Optimizer | Adam หรือ SGD |
| Epochs | [50, 100, 200] |

จากตารางที่ 4 ผู้วิจัยได้กำหนดค่าพารามิเตอร์ จำนวน 60 รูปแบบ โดยกำหนดค่า Structures เท่ากับ [(16, 8), (32, 16, 8)] ค่า Optimizers เท่ากับ [ADAM, SGD] และค่า Epochs เท่ากับ [50, 100, 200]

## 4. ผลการดำเนินงาน

ทำการประเมินประสิทธิภาพของแบบจำลองด้วยค่าความถูกต้อง(Accuracy: A) ค่าความแม่นยำ(Precision: P) ค่าระลึก (Recall :R) และค่าการถ่วงดุล(F1-Score: F) ผลการดำเนินงาน แสดงดังตารางที่ 5 ถึง 7

**ตารางที่ 6 :** ผลลัพธ์ของแบบจำลอง SVM

| Parameter | | Performance | | | |
|---|---|---|---|---|---|
| **C** | **Gamma** | **A(%)** | **P** | **R** | **F** |
| **100** | **0.01** | **96.11** | **0.96** | **0.96** | **0.96** |
| 100 | 0.1 | 93.89 | 0.94 | 0.94 | 0.94 |
| 100 | 0.001 | 93.33 | 0.93 | 0.93 | 0.93 |
| 100 | 1 | 40.56 | 0.85 | 0.4 | 0.38 |
| 10 | 0.01 | 95.56 | 0.96 | 0.96 | 0.96 |
| 10 | 1 | 40.56 | 0.85 | 0.4 | 0.38 |
| 1 | 0.1 | 93.89 | 0.95 | 0.94 | 0.94 |
| 1 | 1 | 40 | 0.65 | 0.39 | 0.37 |
| 0.1 | 0.01 | 67.78 | 0.77 | 0.67 | 0.67 |
| 0.1 | 0.001 | 23.89 | 0.09 | 0.23 | 0.12 |

จากตารางที่ 5 แบบจำลองที่ปรับค่าพารามิเตอร์ C เท่ากับ 100 ค่า Gamma เท่ากับ 0.01 และค่า Kernel เท่ากับ RBF ให้ประสิทธิภาพดีที่สุดด้วยค่าความถูกต้องร้อยละ 96.11

**ตารางที่ 7 :** ผลลัพธ์ของแบบจำลอง XGBoost

| Parameter | | | | Performance | | | |
|---|---|---|---|---|---|---|---|
| **n_estimators** | **learning_rate** | **max_depth** | **col sample_bytree** | **A(%)** | **P** | **R** | **F** |
| **1000** | **0.05** | **3** | **1** | **96.67** | **0.97** | **0.97** | **0.97** |
| 1000 | 0.05 | 3 | 1 | 94.44 | 0.95 | 0.94 | 0.95 |
| 800 | 0.05 | 3 | 1 | 94.44 | 0.95 | 0.95 | 0.94 |
| 1000 | 0.05 | 8 | 1 | 94.44 | 0.95 | 0.95 | 0.94 |
| 1000 | 0.01 | 3 | 0.05 | 72.78 | 0.74 | 0.73 | 0.72 |
| 1000 | 0.01 | 6 | 0.05 | 71.11 | 0.72 | 0.71 | 0.7 |
| 800 | 0.01 | 8 | 0.05 | 65.56 | 0.65 | 0.65 | 0.65 |
| 800 | 0.01 | 3 | 0.01 | 59.44 | 0.6 | 0.59 | 0.59 |
| 800 | 0.01 | 3 | 0.01 | 48.89 | 0.63 | 0.48 | 0.47 |
| 800 | 0.01 | 3 | 0.01 | 46.67 | 0.48 | 0.46 | 0.43 |

จากตารางที่ 6 แบบจำลอง XGBoost ที่ปรับค่าพารามิเตอร์ n_estimators เท่ากับ 1,000 ค่า learning_rate เท่ากับ 0.05 ค่า max_depth เท่ากับ 3 และค่า Colsample_bytree เท่ากับ 1 ให้ประสิทธิภาพดีที่สุดด้วยค่าความถูกต้องร้อยละ 96.67

**ตารางที่ 8 :** ผลลัพธ์ของแบบจำลอง ANN

| Parameter | | | Performance | | | |
|---|---|---|---|---|---|---|
| **Structures** | **Optimizers** | **Epochs** | **A(%)** | **P** | **R** | **F** |
| **(16, 8)** | **adam** | **200** | **93.89** | **0.94** | **0.94** | **0.94** |
| (16, 8) | adam | 100 | 93.33 | 0.93 | 0.93 | 0.93 |
| (32, 16, 8) | adam | 50 | 92.78 | 0.93 | 0.93 | 0.93 |
| (16, 8) | adam | 50 | 91.67 | 0.92 | 0.92 | 0.92 |
| (32, 16, 8) | sgd | 100 | 91.67 | 0.92 | 0.92 | 0.92 |
| (16, 8) | adam | 50 | 86.11 | 0.86 | 0.86 | 0.86 |
| (16, 8) | sgd | 50 | 76.67 | 0.77 | 0.76 | 0.76 |
| (16, 8) | sgd | 50 | 72.22 | 0.72 | 0.72 | 0.72 |
| (16, 8) | sgd | 100 | 71.67 | 0.72 | 0.72 | 0.71 |

จากตารางที่ 8 แบบจำลอง ANN ที่ปรับค่าพารามิเตอร์ Structures เท่ากับ (16, 8) ค่า Optimizers เท่ากับ adam และค่า Epochs เท่ากับ 200 ให้ประสิทธิภาพดีที่สุดด้วยค่าความถูกต้องร้อยละ 93.89

เมื่อได้ผลลัพธ์ของ 3 แบบจำลอง ทำการเปรียบเทียบประสิทธิภาพของแต่ละแบบจำลอง แสดงดังตารางที่ 9

**ตารางที่ 9 :** ผลการเปรียบเทียบประสิทธิภาพของแบบจำลอง

| Model | Performance | | | |
|---|---|---|---|---|
| | **A(%)** | **P** | **R** | **F** |
| SVM | 96.11 | 0.962 | 0.961 | 0.961 |
| XGBoost | **96.67** | **0.969** | **0.967** | **0.966** |
| ANN | 93.89 | 0.94 | 0.939 | 0.938 |

จากตารางที่ 9 สรุปแบบจำลอง XGBoost ให้ประสิทธิภาพดีที่สุดด้วยค่าความถูกต้องร้อยละ 96.67 ค่าความแม่นยำ เท่ากับ 0.969 ค่าความระลึก เท่ากับ 0.967 และค่าการถ่วงดุล เท่ากับ 0.966 ตามลำดับ

จากนั้นได้ทำการทดสอบประสิทธิภาพแบบจำลองทั้ง 3 แบบจำลองด้วย Unseen Data ชุดเดียวกัน จำนวน 100 ชุดข้อมูล แสดงดังตารางที่ 10

**ตารางที่ 10 :** ผลการทดสอบประสิทธิภาพแบบจำลองด้วย Unseen Data

| Model | Performance | | | |
|---|---|---|---|---|
| | A(%) | P | R | F |
| SVM | **94.67** | **0.948** | **0.947** | **0.946** |
| XGBoost | 93.33 | 0.936 | 0.933 | 0.933 |
| ANN | 92.67 | 0.928 | 0.926 | 0.927 |

จากตารางที่ 10 ผลการทดสอบตัวด้วย Unseen Data ชุดเดียวกัน จำนวน 100 ชุดข้อมูล พบว่า แบบจำลอง SVM ให้ประสิทธิภาพดีที่สุดด้วยค่าความถูกต้องร้อยละ 94.67 ค่าความแม่นยำเฉลี่ย เท่ากับ 0.948 ค่าความระลึกเฉลี่ย เท่ากับ 0.947 และค่าการถ่วงดุลเฉลี่ย เท่ากับ 0.946 แสดงให้เห็นว่า SVM ช่วยป้องกันการ Overfitting ได้ดีกว่าเมื่อเทียบกับ ANN หรือ XGBoost ในกรณีที่ข้อมูลมีมิติจำกัด ทำให้เมื่อทดสอบกับ Unseen Data มีความเสถียรสูงกว่า

## 5. สรุป

งานวิจัยฉบับนี้ได้นำเสนอรูปแบบและแนวทางการประยุกต์ใช้เทคนิคการเรียนรู้ของเครื่องเพื่อพยากรณ์ความต้องการด้านสวัสดิการของบุคลากรภายในองค์การ โดยใช้ชุดข้อมูล จำนวน 330 ชุดข้อมูล ซึ่งได้ทำการเก็บรวบรวมข้อมูลจากแบบสอบถามที่ได้จัดทำขึ้นใหม่ที่ผ่านการประเมินจากผู้ทรงคุณวุฒิและหาค่าความเชื่อมั่น ซึ่งทางผู้วิจัยได้รวบรวมในระหว่างห้วงเดือนพฤศจิกายน ถึง ธันวาคม 2568 โดยข้อมูลทั้งหมด ประกอบด้วยตัวแปรคุณลักษณะด้านประชากร จำนวน 14 ตัวแปร และความต้องการด้านสวัสดิการในแต่ละขั้นอิงตามทฤษฎีลำดับขั้นความต้องการของมาสโลว์ทั้งหมด 5 ขั้น และถูกนำมาแปลงค่าเป็นตัวเลขพร้อมปรับสมดุลข้อมูลด้วยเทคนิค SMOTE เพื่อแก้ไขปัญหาความไม่สมดุลของชุดข้อมูลและข้อมูลที่ในการฝึกฝนมีจำนวนค่อนข้างจำกัด จากนั้นใช้แบบจำลองอัลกอริทึม 3 อัลกอริทึมเพื่อเปรียบเทียบประสิทธิภาพ ได้แก่ ซัพพอร์ตเวกเตอร์แมชชีน )Support Vector Machine: SVM), เอกซ์ตรีมเกรเดียนต์บูสติง )Extreme Gradient Boosting: XGBoost) และโครงข่ายประสาทเทียม (Artificial Neural Network: ANN) รวมทั้งมีการประเมินประสิทธิภาพแบบจำลองด้วยเทคนิค Stratified K-Fold Cross-Validation (k = 5) แสดงให้เห็นว่าแบบจำลอง XGBoost ให้ผลลัพธ์ดีที่สุด โดยมีค่า ความถูกต้องร้อยละ 96.67 และมีค่าความแม่นยำ ค่าความระลึก และค่าการถ่วงดุล เท่ากับ 0.97 แต่เมื่อทดสอบเพิ่มเติมกับชุดข้อมูลที่ไม่เคยใช้ในการฝึกสอน จำนวน 100 ชุดข้อมูล พบว่า SVM กลับมีประสิทธิภาพสูงสุดแทน โดยให้ค่าความถูกต้องร้อยละ 94.67 และมีค่าความแม่นยำ ค่าความระลึก และค่าการถ่วงดุลเท่ากับ 0.947 ซึ่งสะท้อนให้เห็นว่า แบบจำลองอัลกอริทึม SVM มีความสามารถในการประยุกต์ใช้กับข้อมูลใหม่ ที่ยังไม่เคยฝึกสอนหรือเคยทดสอบมาก่อนได้อย่างมีประสิทธิภาพ เนื่องจากในการสร้างแบบจำลองครั้งนี้มีการใช้ข้อมูลจำนวนน้อยมาก อีกทั้งยังผ่านการเพิ่มจำนวนด้วยเทคนิค SMOTE ซึ่ง SVM จะเหมาะกับข้อมูลที่มีปริมาณน้อย ส่งผลทำให้ตอนที่ทดสอบประสิทธิภาพ แบบจำลองอัลกอริทึม SVM จึงให้ผลลัพธ์ที่ดีกว่าอัลกอริทึม XGBoost และ ANN ที่ต้องการชุดข้อมูลในการฝึกฝนเป็นจำนวนมาก

## เอกสารอ้างอิง


[1] Maslow, A. H. (1943). "A theory of human motivation". Psychological Review, 50(4), 370-396.

[2] Vapnik, V. N. (1995). The nature of statistical learning theory. Springer.

[3] Chen, T., & Guestrin, C. (2016). "XGBoost: A scalable tree boosting system". In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 785-794). ACM.

[4] Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). "Learning representations by back-propagating errors". Nature, 323(6088), 533-536.

[5] Alici, N., Altuncu, D., Cengiz, E., & Aksoy, H. (2025). "Optimisation of Maslow's hierarchy of needs-based survey form for nursing home residents using machine learning". Ergonomics.

[6] Yang, Y., & Zhao, P. (2025). "Optimization study on the development path of urban elderly manpower resources based on decision tree algorithm". International Journal for Housing Science and Its Applications.

[7] Ilunga, I. B. F., & Chung, G. H. (2025). "Predicting employee job satisfaction with machine learning". Dongguk Business Review, 47(2), 45-68.

[8] Zheng, E., Li, P., & Song, Z. (2004). "Performance analysis and comparison of neural networks and support vector machines classifier". In Proceedings of the World Congress on Intelligent Control and Automation (Vol. 5, pp. 4035-4039). IEEE.

# การวิเคราะห์องค์ประกอบที่มีอิทธิพลต่อการใช้งาน ระบบการจัดการรักษาความมั่นคงปลอดภัยไซเบอร์ ของกองบัญชาการกองทัพไทย

# An Analysis of the factors that influence the use of the Cyber Security Management System of the Royal Thai Army Headquarters

วรวุฒิ สมานพงษ์ *(Worawoot Samanpong)*[1] และมณเฑียร รัตนศิริวงศ์วุฒิ *(Montean Rattanasiriwongwut)*[2]

*ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล*

*มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ*

[1]*s6707021856175@email.kmutnb.ac.th,* [2] *montean.r@itd.kmutnb.ac.th*

## บทคัดย่อ

*การวิจัยนี้ศึกษาปัจจัยที่มีอิทธิพลต่อความตั้งใจเชิงพฤติกรรมในการใช้งานระบบการจัดการความมั่นคงปลอดภัยไซเบอร์ของกองบัญชาการกองทัพไทย โดยใช้กรอบแนวคิดด้านความคาดหวังด้านประสิทธิภาพ ความคาดหวังด้านความพยายาม อิทธิพลทางสังคม และปัจจัยเอื้ออำนวย เก็บข้อมูลจากกำลังพล 283 คน และวิเคราะห์ด้วย EFA ความเชื่อมั่นของเครื่องมือ และการถดถอยพหุคูณ ผลการศึกษาพบว่า เครื่องมือมีความเชื่อมั่นสูง ($\alpha$ = 0.846–0.919) และ โมเดลสามารถอธิบายความแปรปรวนของความตั้งใจเชิงพฤติกรรมได้ร้อยละ 56.90 ($R^2 = 0.569, p < .001$) โดยความคาดหวังด้านประสิทธิภาพมีอิทธิพลสูงสุด รองลงมาคืออิทธิพลทางสังคมและปัจจัยเอื้ออำนวย ขณะที่ความคาดหวังด้านความพยายามไม่มีนัยสำคัญ ผลการวิจัยชี้ว่าการรับรู้ประโยชน์และการสนับสนุนเชิงองค์กรเป็นปัจจัยหลักต่อการยอมรับระบบในบริบทด้านความมั่นคง*

**คำสำคัญ:** ความมั่นคงปลอดภัยไซเบอร์, ความตั้งใจเชิงพฤติกรรม, การยอมรับเทคโนโลยี

## Abstract

*This study examined the factors influencing behavioral intention to use the Cybersecurity Management System of the Royal Thai Armed Forces Headquarters. The conceptual framework incorporated performance expectancy, effort expectancy, social influence, and facilitating conditions as independent variables, with behavioral intention as the dependent variable. Data were collected from 283 military personnel using a structured questionnaire and analyzed through descriptive statistics, exploratory factor analysis (EFA), reliability testing using Cronbach's alpha, and multiple regression analysis.*

*The findings indicated high reliability of the measurement instrument (Cronbach's alpha = 0.846–0.919). The proposed model explained 56.90% of the variance in behavioral intention ($R^2 = 0.569$, $p < .001$). Performance expectancy demonstrated the strongest influence, followed by social influence and facilitating conditions, whereas effort expectancy showed no statistically significant effect. The results highlight that perceived usefulness, organizational support, and resource readiness are critical determinants of system acceptance within a security-oriented organizational context.*

**Keywords:** cybersecurity management system, behavioral intention, technology acceptance, military organization.

## 1. บทนำ

กองบัญชาการกองทัพไทย มีบทบาทสำคัญในการบูรณาการภารกิจและข้อมูลทางทหารซึ่งเกี่ยวข้องกับความ

มั่นคงของประเทศ ท่ามกลางภัยคุกคามไซเบอร์ ที่ทวีความรุนแรงและซับซ้อน การนำระบบการจัดการความมั่นคงปลอดภัยไซเบอร์มาใช้จึงเป็นกลไกสำคัญในการป้องกันและตอบสนองต่อความเสี่ยง อย่างไรก็ตาม ประสิทธิผลของระบบขึ้นอยู่กับการยอมรับและพฤติกรรมการใช้งานของบุคลากร

ดังนั้น การศึกษาปัจจัยที่มีอิทธิพลต่อความตั้งใจและพฤติกรรมการใช้งานระบบจึงมีความจำเป็น เพื่อสนับสนุนการกำหนดนโยบายและยกระดับความมั่นคงปลอดภัยไซเบอร์ขององค์กรอย่างยั่งยืน

## 2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

### 2.1 ทฤษฎี UTAUT

ทฤษฎี Unified Theory of Acceptance and Use of Technology (UTAUT) พัฒนาโดย Viswanath Venkatesh และคณะ (2003) อธิบายปัจจัยที่มีอิทธิพลต่อความตั้งใจและพฤติกรรมการใช้เทคโนโลยี ประกอบด้วย 4 ปัจจัยหลัก ได้แก่ ความคาดหวังด้านประสิทธิภาพ ความคาดหวังด้านความพยายาม อิทธิพลทางสังคม และปัจจัยเอื้ออำนวย โดยปัจจัยดังกล่าวส่งผลต่อความตั้งใจเชิงพฤติกรรม และนำไปสู่พฤติกรรมการใช้งานจริง แบบจำลองนี้ได้รับการยอมรับอย่างกว้างขวางในการศึกษาการยอมรับระบบสารสนเทศในหลากหลายบริบทองค์กร

### 2.2 ระบบความมั่นคงปลอดภัยไซเบอร์

ระบบความมั่นคงปลอดภัยไซเบอร์ หมายถึง กระบวนการ เทคโนโลยี และมาตรการที่ใช้ในการปกป้องข้อมูลและระบบสารสนเทศจากภัยคุกคาม โดยมีเป้าหมายเพื่อรักษาความลับ (Confidentiality) ความถูกต้องครบถ้วน (Integrity) และความพร้อมใช้งาน (Availability) หรือที่เรียกว่า CIA Triad แนวคิดดังกล่าวสอดคล้องกับกรอบการบริหารจัดการความมั่นคงปลอดภัยไซเบอร์ของ National Institute of Standards and Technology ซึ่งครอบคลุมการระบุความเสี่ยง การป้องกัน การตรวจจับ การตอบสนอง และการฟื้นฟู

### 2.3 การวิเคราะห์องค์ประกอบเชิงสำรวจ (EFA)

การวิเคราะห์องค์ประกอบเชิงสำรวจ (Exploratory Factor Analysis: EFA) เป็นเทคนิคทางสถิติที่ใช้ค้นหาโครงสร้างตัวแปรแฝงและลดจำนวนตัวแปร โดยตรวจสอบความเหมาะสมของข้อมูลผ่านค่า KMO และ Bartlett's Test และสกัดองค์ประกอบด้วยวิธี เช่น PCA หรือ PAF พร้อมการหมุนแกนเพื่อให้ตีความได้ชัดเจน

### 2.4 การวิเคราะห์องค์ประกอบเชิงยืนยัน (CFA)

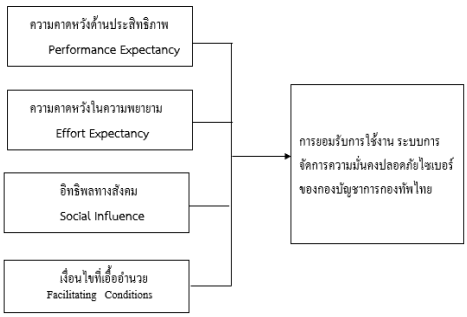
การวิเคราะห์องค์ประกอบเชิงยืนยัน (Confirmatory Factor Analysis: CFA) ใช้ทดสอบความสอดคล้องของโมเดลการวัดกับข้อมูลเชิงประจักษ์ เพื่อยืนยันความตรงเชิงโครงสร้างของตัวแปรแฝง โดยพิจารณาดัชนีความสอดคล้องของโมเดล

### 2.5 การวิเคราะห์การถดถอยพหุคูณ

การถดถอยพหุคูณ (Multiple Linear Regression) เป็นเทคนิคทางสถิติที่ใช้ทดสอบอิทธิพลของตัวแปรอิสระหลายตัวต่อหนึ่งตัวแปรตาม โดยประเมินความสามารถในการอธิบายโมเดลด้วยค่า $R^2$ และทดสอบนัยสำคัญทางสถิติด้วยค่า p เพื่ออธิบายความสัมพันธ์เชิงสาเหตุระหว่างตัวแปรตามกรอบแนวคิดการวิจัย

## 3. วิธีการดำเนินการวิจัย

การวิจัยนี้เป็นการวิจัยเชิงปริมาณแบบสำรวจ มีวัตถุประสงค์เพื่อศึกษาปัจจัยที่มีอิทธิพลต่อความตั้งใจและพฤติกรรมการใช้งานระบบการจัดการความมั่นคงปลอดภัยไซเบอร์ของ กองบัญชาการกองทัพไทย โดยประยุกต์กรอบแนวคิด UTAUT ของ Viswanath Venkatesh และคณะ เป็นพื้นฐานในการพัฒนาแบบจำลองการวิจัย




กรอบแนวคิดการวิจัย

ประชากรคือข้าราชการทหารที่เกี่ยวข้องกับการใช้งานระบบ กำหนดขนาดกลุ่มตัวอย่างตามหลักเกณฑ์ของ Joseph F. Hair Jr. และคณะ ได้แบบสอบถามที่สมบูรณ์จำนวน 285 ชุด เครื่องมือที่ใช้คือแบบสอบถามมาตราส่วนประมาณค่า 5 ระดับ ผ่านการตรวจสอบความเที่ยงตรงเชิงเนื้อหาและความเชื่อมั่นด้วยค่า Cronbach's Alpha

การวิเคราะห์ข้อมูลใช้สถิติเชิงพรรณนา การวิเคราะห์องค์ประกอบเชิงสำรวจ และการถดถอยพหุคูณ เพื่อทดสอบสมมติฐานที่ระดับนัยสำคัญทางสถิติ 0.05

## 4. ผลการดำเนินงาน

### 4.1 ผลการวิเคราะห์ข้อมูลทั่วไปของผู้ตอบแบบสอบถาม

กลุ่มตัวอย่างจำนวน 283 คน ส่วนใหญ่เป็นเพศชาย อายุ 31–40 ปี อยู่ในระดับสัญญาบัตรตอนต้นและระดับประทวน มีอายุราชการ 6–15 ปี จบการศึกษาระดับปริญญาตรี และมีรายได้ 20,000–40,000 บาทต่อเดือน ทั้งนี้ ร้อยละ 76.3 เคยผ่านการอบรมด้านความมั่นคงปลอดภัยไซเบอร์

โดยภาพรวม กลุ่มตัวอย่างเป็นกำลังพลวัยทำงานที่มีประสบการณ์และพื้นฐานทางวิชาการ เหมาะสมต่อการให้ข้อมูลเกี่ยวกับการยอมรับและการใช้งานระบบความมั่นคงปลอดภัยไซเบอร์ขององค์กร

### 4.2 ผลการวิเคราะห์ข้อมูลทั่วไปของผู้ตอบแบบสอบถาม

ผลการวิเคราะห์ด้วยค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานพบว่า โดยภาพรวมกลุ่มตัวอย่างให้ความสำคัญต่อปัจจัยทั้ง 5 ด้านอยู่ในระดับ มากที่สุด (Mean = 4.41, S.D. = 0.67) เมื่อพิจารณารายด้าน พบว่า ความคาดหวังด้านประสิทธิภาพ (PE) มีค่าเฉลี่ยสูงที่สุด (Mean = 4.51) โดยเฉพาะประเด็นการเพิ่มความปลอดภัยของหน่วยงาน รองลงมาคือ ความตั้งใจเชิงพฤติกรรม (BI) (Mean = 4.50) และ อิทธิพลทางสังคม (SI) (Mean = 4.45) ซึ่งสะท้อนบทบาทของผู้บังคับบัญชาและวัฒนธรรมองค์กรในการสนับสนุนการใช้งานระบบ ด้าน ความคาดหวังด้านความพยายาม (EE) (Mean = 4.34) แสดงให้เห็นว่าผู้ใช้รับรู้ว่าระบบใช้งานง่าย และ ปัจจัยเอื้ออำนวย (FC) (Mean = 4.26) สะท้อนความพร้อมด้านโครงสร้างพื้นฐานและการสนับสนุนทางเทคนิค โดยสรุป การรับรู้ถึงประโยชน์และประสิทธิภาพของระบบเป็นปัจจัยสำคัญที่สุดที่ส่งผลต่อการยอมรับและการใช้งานระบบการจัดการความมั่นคงปลอดภัยไซเบอร์ในองค์กร

### 4.3 ผลการวิเคราะห์ข้อมูลทั่วไปของผู้ตอบแบบสอบถาม

ผลการวิเคราะห์องค์ประกอบเชิงสำรวจ (EFA) พบว่าค่า KMO = 0.935 แสดงถึงความเหมาะสมของข้อมูลในระดับดีมาก สามารถสกัดได้ 5 องค์ประกอบ สอดคล้องกับกรอบแนวคิดการวิจัย โดยตัวแปรส่วนใหญ่มีค่า Factor Loading > 0.50 ขณะที่ปัจจัยอิทธิพลทางสังคมมีค่าต่ำกว่าเกณฑ์ จึงควรพิจารณาปรับปรุงในอนาคต

| | Factor Loadings | | | | | |
|---|---|---|---|---|---|---|
| | Factor | | | | | Uniqueness |
| | Factor 1 | Factor 2 | Factor 3 | Factor 4 | Factor 5 | |
| FC4 | 0.890 | | | | | 0.2096 |
| FC3 | 0.884 | | | | | 0.2158 |
| FC2 | 0.846 | | | | | 0.2629 |
| FC5 | 0.630 | | | | | 0.3473 |
| FC1 | 0.615 | | | | | 0.4259 |
| BI2 | | 0.808 | | | | 0.2802 |
| BI3 | | 0.714 | | | | 0.383 |
| BI1 | | 0.694 | | | | 0.3536 |
| BI4 | | 0.690 | | | | 0.3433 |
| BI5 | | 0.653 | | | | 0.4504 |
| PE2 | | | 0.801 | | | 0.3636 |
| PE1 | | | 0.733 | | | 0.3521 |
| PE3 | | | 0.728 | | | 0.2983 |
| PE4 | | | 0.666 | | | 0.3898 |
| PE5 | | | 0.654 | | | 0.4664 |
| EE3 | | | | 0.968 | | 0.0991 |
| EE4 | | | | 0.818 | | 0.2169 |
| EE1 | | | | 0.742 | | 0.2032 |
| EE5 | | | | 0.536 | | 0.4477 |
| SI1 | | | | | 0.478 | 0.4491 |
| SI2 | | | | | 0.414 | 0.4261 |
| SI3 | | | | | 0.373 | 0.4352 |
| *Note.* 'Maximum likelihood' extraction method was used in combination with a 'oblimin' rotation | | | | | | |

สถิติการวิเคราะห์องค์ประกอบเชิงสำรวจ

### 4.4 ผลจากการทดสอบความเชื่อมั่นของแบบสอบถาม

ผู้วิจัยได้ทดสอบคุณภาพเครื่องมือโดยทดลองใช้แบบสอบถามกับกลุ่มตัวอย่าง 285 คน เพื่อประเมินความชัดเจนและความสอดคล้องภายในของข้อคำถาม วิเคราะห์ความเชื่อมั่นด้วยค่าสัมประสิทธิ์แอลฟาของครอนบาค (เกณฑ์ $\leq 0.70$) ผลพบว่าทุกปัจจัยมีค่า Cronbach's Alpha อยู่ระหว่าง 0.846–0.919 สะท้อนความเชื่อมั่นในระดับดีถึงดีมาก ขณะที่ค่าเฉลี่ยของทุกปัจจัยอยู่ในระดับสูง (4.26–4.51) แสดงถึงการรับรู้เชิงบวกของผู้ตอบ โดยสรุป เครื่องมือวิจัยมีคุณภาพและเหมาะสมสำหรับการเก็บข้อมูลและการวิเคราะห์ในขั้นตอนต่อไป

### 4.5 ผลจากการวิเคราะห์สมการถดถอยพหุคูณ

การวิเคราะห์ถดถอยพหุคูณเพื่อศึกษาปัจจัยที่มีอิทธิพลต่อความตั้งใจเชิงพฤติกรรม โดยกำหนดให้ความตั้งใจเชิง

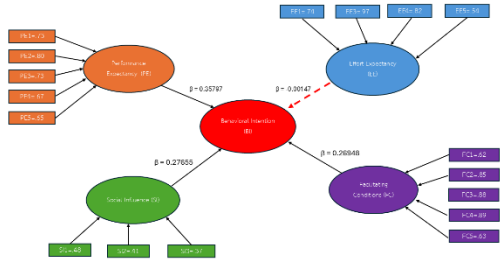
พฤติกรรมเป็นตัวแปรตาม และความคาดหวังด้านประสิทธิภาพ ความคาดหวังด้านความพยายาม อิทธิพลทางสังคม และสิ่งอำนวยความสะดวกเป็นตัวแปรอิสระ

ผลการวิเคราะห์พบว่า โมเดลมีความเหมาะสมทางสถิติ ($F = 91.7, p < .001$) โดยมีค่า $R = 0.754$ และ $R^2 = 0.569$ แสดงว่าสามารถอธิบายความแปรปรวนของความตั้งใจเชิงพฤติกรรมได้ร้อยละ 56.9 และมีค่าความคลาดเคลื่อนในการพยากรณ์ต่ำ (RMSE = 0.328)

| Model | R | $R^2$ | RMSE | F | df1 | df2 | p |
|---|---|---|---|---|---|---|---|
| 1 | 0.754 | 0.569 | 0.328 | 91.7 | 4 | 278 | <.001 |

ค่าผลลัพธ์การวิเคราะห์สมการถดถอยพหุคูณ

เมื่อพิจารณาค่าสัมประสิทธิ์ถดถอย พบว่า ความคาดหวังด้านประสิทธิภาพ ($\beta = 0.358$), อิทธิพลทางสังคม ($\beta = 0.277$) และสิ่งอำนวยความสะดวก ($\beta = 0.269$) มีอิทธิพลเชิงบวกอย่างมีนัยสำคัญทางสถิติ ($p < .001$) โดยความคาดหวังด้านประสิทธิภาพเป็นปัจจัยที่มีอิทธิพลมากที่สุด ขณะที่ความคาดหวังด้านความพยายามไม่มีนัยสำคัญทางสถิติ ($p = 0.98$)



โมเดลโครงสร้างปัจจัยที่มีอิทธิพลต่อความตั้งใจในการใช้งานระบบการจัดการรักษาความมั่นคงปลอดภัยไซเบอร์

โดยสรุป โมเดลมีความสามารถในการพยากรณ์อยู่ในระดับดี และมี 3 ปัจจัยหลักที่ส่งผลต่อความตั้งใจเชิงพฤติกรรมอย่างมีนัยสำคัญ

## 5. สรุปผล อภิปรายผล และข้อเสนอแนะ

### 5.1 สรุปผลการดำเนินงานวิจัย

ผลการวิจัยพบว่า กลุ่มตัวอย่างมีความพร้อมด้านประสบการณ์และการอบรม และให้ความสำคัญต่อทุกปัจจัยในระดับมากที่สุด โดยความคาดหวังด้านประสิทธิภาพมีค่าเฉลี่ยสูงสุด โครงสร้างปัจจัยมีความสอดคล้องกับกรอบแนวคิด และเครื่องมือวิจัยมีความเชื่อมั่นอยู่ในระดับดีมาก ผลการถดถอยชี้ว่า ตัวแปรอิสระสามารถอธิบายความตั้งใจเชิงพฤติกรรมได้ร้อยละ 56.90 โดยความคาดหวังด้านประสิทธิภาพมีอิทธิพลมากที่สุด ขณะที่ความคาดหวังด้านความพยายามไม่มีนัยสำคัญทางสถิติ

### 5.2อภิปรายผล

ผลการวิจัยยืนยันว่า “ความคาดหวังด้านประสิทธิภาพ” เป็นปัจจัยสำคัญที่สุดที่ส่งผลต่อความตั้งใจใช้งาน สอดคล้องกับกรอบแนวคิดของ UTAUT ที่ระบุว่าการรับรู้ถึงประโยชน์ของระบบเป็นตัวทำนายหลักของการยอมรับเทคโนโลยี

ขณะที่ความคาดหวังด้านความพยายามไม่มีนัยสำคัญ อาจเนื่องจากกลุ่มตัวอย่างมีพื้นฐานความรู้และประสบการณ์ด้านเทคโนโลยีเพียงพอ ทำให้ความง่ายในการใช้งานไม่ใช่ปัจจัยชี้ขาด

แม้ผล EFA ของอิทธิพลทางสังคมยังไม่เด่นชัด แต่ผลการถดถอยพบว่ามีอิทธิพลอย่างมีนัยสำคัญ สะท้อนบริบทองค์กรแบบลำดับชั้นที่แรงสนับสนุนจากผู้บังคับบัญชามีบทบาทต่อการตัดสินใจใช้งานระบบ

### 5.3 ข้อเสนอแนะ

จากผลการวิจัยที่พบว่าความคาดหวังด้านประสิทธิภาพมีอิทธิพลสูงสุด หน่วยงานควรเน้นสื่อสารประโยชน์ที่เป็นรูปธรรมของระบบ เช่น การเพิ่มประสิทธิภาพ ลดความผิดพลาด และยกระดับความมั่นคงปลอดภัย เพื่อสร้างแรงจูงใจในการใช้งานอย่างต่อเนื่อง

เนื่องจากอิทธิพลทางสังคมมีนัยสำคัญ ผู้บังคับบัญชาควรมีบทบาทสนับสนุนและเป็นแบบอย่างในการใช้งาน เพื่อผลักดันให้เกิดการยอมรับในระดับองค์กร ขณะเดียวกันควรพัฒนาโครงสร้างพื้นฐาน ระบบสนับสนุนทางเทคนิค และจัดอบรมอย่างสม่ำเสมอเพื่อเสริมความพร้อมของบุคลากร

สำหรับการวิจัยในอนาคต ควรปรับปรุงข้อคำถามด้านอิทธิพลทางสังคม และใช้การวิเคราะห์ขั้นสูง เช่น CFA หรือ SEM รวมทั้งศึกษาปัจจัยเพิ่มเติมเพื่อขยายองค์ความรู้ให้ครอบคลุมยิ่งขึ้น

## เอกสารอ้างอิง


1 Bartlett, M. S. (1954). A note on the multiplying factors for various chi-square approximations. *Journal of the Royal Statistical Society: Series B*, 16(2), 296–298.

2 Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, 39(1), 31–36.

3 Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, 27(3), 425–478.

4 Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis. *Practical Assessment, Research & Evaluation*, 10(7), 1–9.

5 Henson, R. K., & Roberts, J. K. (2006). Use of exploratory factor analysis in published research. *Educational and Psychological Measurement*, 66(3), 393–416.

6 Venkatesh, V., Thong, J. Y. L., & Xu, X. (2012). Consumer acceptance and use of information technology. *MIS Quarterly*, 36(1), 157–178.

7 Field, A. (2018). *Discovering statistics using IBM SPSS statistics* (5th ed.). Sage Publications.

8 Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate data analysis* (8th ed.). Cengage Learning.

9 Tabachnick, B. G., & Fidell, L. S. (2019). *Using multivariate statistics* (7th ed.). Pearson.

# การเปรียบเทียบแบบจำลองการเรียนรู้ของเครื่อง เพื่อคาดการณ์ความคุ้มค่าของคาร์บอนเครดิตจากการลงทุนปลูกต้นไม้ ในภาควิสาหกิจชุมชน จังหวัดขอนแก่น

# A Comparative Study of Machine Learning Models for Predicting the Cost-Effectiveness of Carbon Credits from Tree Plantation Investment in Community Enterprises, Khon Kaen Province

*ปภัสรา สีหาราช (Papatsara Siharach)[1] และนลินภัสร์ บำเพ็ญเพียร (Nalinpat Bhumpenpein)[2]*
*ภาควิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล*
*มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ*
*[1]s6707011857039@email.kmutnb.ac.th, [2]nalinpat.b@itd.kmutnb.ac.th*

## บทคัดย่อ

*งานวิจัยฉบับนี้ศึกษาเพื่อค้นหาแบบจำลองทางคณิตศาสตร์ที่เหมาะสมที่สุดสำหรับการพยากรณ์ความคุ้มค่าของคาร์บอนเครดิต โดยครอบคลุมพื้นที่วิสาหกิจชุมชนในจังหวัดขอนแก่น ได้แก่ บ้านท่าลี่และบ้านแดง รวมถึงข้อมูลจากธนาคารเพื่อการเกษตรและสหกรณ์การเกษตร ผ่านการวิเคราะห์ชุดข้อมูลขนาดใหญ่กว่า 48,000 แถว ด้วยเทคนิคการเรียนรู้ของเครื่อง 4 รูปแบบ พบว่าขนาดของพื้นที่ปลูกเป็นปัจจัยหลักในการกำหนดรายได้จากคาร์บอนเครดิต โดยมีความสัมพันธ์เชิงเส้นที่สมบูรณ์หมายความว่าพื้นที่ที่เพิ่มขึ้นจะถูกนำมาคำนวณเป็นรายได้ในสัดส่วนที่คงที่เสมอ ในขณะที่ขนาดเส้นรอบวงของต้นไม้มีอิทธิพลต่อรายได้น้อยกว่าอย่างมีนัยสำคัญ เมื่อพิจารณาถึงประสิทธิภาพของแบบจำลอง พบว่าการถดถอยเชิงเส้น ซึ่งเป็นแบบจำลองพื้นฐานสามารถให้ผลลัพธ์ที่ดีที่สุดในทุกกรณีศึกษา โดยมีค่าความคลาดเคลื่อน (MAE) อยู่ในระดับต่ำเพียง 7.89 - 8.06 ในส่วนของแบบจำลองอื่น ๆ โครงข่ายประสาทเทียม (ANN) มีความสามารถจับความสัมพันธ์ที่ซับซ้อนของเส้นรอบวงได้ดีเป็นอันดับสอง Random Forest และ XGBoost กลับมีข้อจำกัดคือสภาวะการเรียนรู้ที่จดจำข้อมูลมากเกินไป ส่งผลให้การพยากรณ์รายได้จากเส้นรอบวงมีลักษณะไม่ราบเรียบหรือไม่เป็นไปตามความเป็นจริง ผลลัพธ์ทางการเงินมั่นคงทั้ง 2 กรณีศึกษา โดยกรณีศึกษาที่ 2 แสดงให้เห็นถึงสถานการณ์ที่สร้างผลกำไรและความคุ้มค่าในการลงทุน (ROI) ได้มากกว่ากรณีแรกอย่างมีนัยสำคัญ*

**คำสำคัญ:** คาร์บอนเครดิต ตลาดคาร์บอนภาคสมัครใจ โครงการ T-VER อัตราผลตอบแทนจากการลงทุน แบบจำลองการเรียนรู้ของเครื่อง

## Abstract

*This research aims to identify the optimal mathematical model for forecasting the profitability of carbon credits in the community enterprise areas of Khon Kaen province, specifically in Ban Tha Li and Ban Daeng. The study utilizes data from the Bank for Agriculture and Agricultural Cooperatives and analyzes a comprehensive dataset of over 48,000 rows. The findings indicate that the size of the planting area is the primary factor in determining carbon credit income, demonstrating a perfect linear relationship—meaning that increases in area consistently translate to higher income. In contrast, tree circumference has a significantly lesser impact on income. In terms of model performance, Linear Regression, the simplest model, produced the best results across all case studies, achieving a low Mean Absolute Error (MAE) of only 7.89 to 8.06. The Artificial Neural Network (ANN) model performed second best, effectively capturing the complex relationships involving circumference. However, the Random Forest and XGBoost models faced limitations due to overfitting, leading to inconsistent or unrealistic*

*income projections based on circumference. Financial outcomes were stable across both case studies, with the second case demonstrating significantly higher profitability and return on investment (ROI) compared to the first.*

**Keywords:** Carbon Credits, Voluntary Carbon market, T-VER Scheme, Return on Investment, Machine Learning Model.

## 1. บทนำ

ปัจจุบันปัญหาการเปลี่ยนแปลงสภาพภูมิอากาศโลกที่มีสาเหตุหลักมาจากการปล่อย ก๊าซเรือนกระจก (Greenhouse Gas) ได้ทวีความรุนแรงขึ้น ส่งผลให้ประชาคมโลกและองค์การระหว่างประเทศ เช่น UNFCCC (United Nations Framework Convention on Climate Change) ให้ความสำคัญกับการลดและควบคุมการปล่อยก๊าซเรือนกระจกเพื่อบรรลุเป้าหมายความเป็นกลางทางคาร์บอน (Carbon Neutrality) และการปล่อยก๊าซเรือนกระจกสุทธิเป็นศูนย์ (Net Zero GHG Emission) [1] เป็นต้น การกักเก็บคาร์บอนในภาคป่าไม้และการเกษตรถือเป็นกลไกสำคัญในการบรรลุเป้าหมายดังกล่าว เนื่องจากป่าไม้สามารถช่วยดูดซับและกักเก็บคาร์บอนในชีวมวลและดินได้อย่างมีประสิทธิภาพ [2] เพื่อส่งเสริมให้เกิดการลดและกักเก็บก๊าซเรือนกระจก ตลาดคาร์บอนภาคสมัครใจ (Voluntary Carbon Market) จึงเข้ามามีบทบาท โดยเฉพาะในประเทศไทย องค์การบริหารจัดการก๊าซเรือนกระจกได้จัดทำ โครงการลดก๊าซเรือนกระจกภาคสมัครใจตามมาตรฐานของประเทศไทย (T-VER) เพื่อรับรองปริมาณการลดหรือกักเก็บคาร์บอน ซึ่งสามารถนำมาซื้อขายเป็นคาร์บอนเครดิตได้ [3]

## 2. วรรณกรรมที่เกี่ยวข้อง

### 2.1 การกักเก็บคาร์บอนในภาคป่าไม้และวิสาหกิจชุมชน

การคำนวณและการรับรองคาร์บอนเครดิต เพื่อเปลี่ยนการกักเก็บคาร์บอนให้เป็นสินทรัพย์ทางการเงิน (คาร์บอนเครดิต) โครงการปลูกป่าจะต้องแสดงให้เห็นถึงความสามารถในการลด $CO_2$ ที่เป็นไปตามหลักการ "ความเป็นส่วนเพิ่ม" (Additionality) ซึ่งหมายความว่า การกักเก็บคาร์บอนนั้นจะต้องเป็นผลมาจากการดำเนินโครงการที่ ไม่ได้เกิดขึ้นตามปกติ หรือหากไม่มีแรงจูงใจทางการเงินจากคาร์บอนเครดิต

### 2.2 ตลาดคาร์บอนภาคสมัครใจ (Voluntary Carbon Market)

คาร์บอนเครดิต (Carbon Credit) คือหน่วยรับรองปริมาณก๊าซเรือนกระจกที่ลดหรือกักเก็บได้ 1 ตันคาร์บอนไดออกไซด์เทียบเท่า ซึ่งสามารถนำมาซื้อขายได้ กลไกสำคัญที่ใช้ในประเทศไทยคือ โครงการลดก๊าซเรือนกระจกภาคสมัครใจตามมาตรฐานของประเทศไทย (T-VER) ที่รับรองโดยองค์การบริหารจัดการก๊าซเรือนกระจก (อบก.) [3]

### 2.3 การวิเคราะห์ความคุ้มค่าทางการเงินของการลงทุน (Investment Appraisal)

งานวิจัยนี้เน้นการประเมินความคุ้มค่าของการลงทุนปลูกต้นไม้โดยวิสาหกิจชุมชน ผ่านการใช้ตัวชี้วัดทางการเงินหลัก 2 รายการ ได้แก่ มูลค่าปัจจุบันสุทธิ (Net Present Value) และอัตราผลตอบแทนจากการลงทุน (Return on Investment)

มูลค่าปัจจุบันสุทธิ (Net Present Value) เป็นการประเมินมูลค่าทางการเงินของโครงการตลอดช่วงอายุโครงการ โดยเปรียบเทียบมูลค่าปัจจุบันของกระแสเงินสดรับ (รายได้คาร์บอนเครดิตและผลผลิตอื่น ๆ) กับมูลค่าปัจจุบันของกระแสเงินสดจ่าย(ต้นทุนการปลูกและบำรุงรักษา) โดยใช้อัตราคิดลด (Discount Rate) ที่เหมาะสม หาก $NPV > 0$ แสดงว่าโครงการนั้นจะเพิ่มมูลค่าทางเศรษฐกิจให้กับผู้ลงทุน $NPV = 0$ แสดงว่าโครงการนั้นสร้างผลตอบแทนเท่ากับอัตราคิดลด (i) พอดี และ $NPV < 0$ แสดงว่าโครงการนั้นไม่คุ้มค่าที่จะลงทุน และอัตราผลตอบแทนจากการลงทุน (Return on Investment) ใช้ในการวัดประสิทธิภาพของการลงทุน โดยแสดงเป็นร้อยละของผลกำไรสุทธิเทียบกับต้นทุนที่ใช้ไป ซึ่งช่วยให้วิสาหกิจชุมชนสามารถเปรียบเทียบผลตอบแทนจากการลงทุนปลูกต้นไม้กับทางเลือกการลงทุนอื่น ๆ [4]

### 2.4 การเรียนรู้ของเครื่องสำหรับการพยากรณ์

เทคนิคการเรียนรู้ของเครื่อง (Machine Learning) ได้รับการยอมรับอย่างกว้างขวางในงานวิจัยด้านสิ่งแวดล้อมและการเกษตร เพื่อคาดการณ์ตัวแปรที่มี

ความสัมพันธ์ซับซ้อนและไม่เป็นเชิงเส้นทั้งหมด 4 วิธี ได้แก่ การถดถอยเชิงเส้น (Linear Regression) เป็นแบบจำลองพื้นฐาน ใช้สมมติฐานความสัมพันธ์เชิงเส้นตรงระหว่างตัวแปรอิสระ [5] ได้แก่ ขนาดแปลง ชนิดต้นไม้ ต้นทุน รายได้คาร์บอน และตัวแปรตาม (NPV/ROI)

การถดถอยแบบสุ่มวน (Random Forest Regressor) คือการสร้างต้นไม้ตัดสินใจหลายต้นจากการสุ่มตัวอย่างข้อมูลและสุ่มตัวแปร ทำได้ดีในการจัดการกับข้อมูลที่มีปฏิสัมพันธ์ที่ซับซ้อน ลดปัญหาการเกินพอดี (Overfitting) และมีเสถียรภาพในการจัดการกับข้อมูลที่มีปฏิสัมพันธ์ไม่เป็นเชิงเส้น [6]

การเสริมกำลังแบบเกรเดียนท์ (XGBoost) เป็นเทคนิคการเรียนรู้แบบเพิ่มน้ำหนัก (Boosting) ที่สร้างต้นไม้ตัดสินใจอย่างต่อเนื่องเพื่อแก้ไขข้อผิดพลาดที่เหลือจากต้นไม้ก่อนหน้า XGBoost เป็นเวอร์ชันที่ได้รับการปรับปรุงให้มีประสิทธิภาพและความรวดเร็วสูง ทำให้เป็นทางเลือกที่แข็งแกร่งในการคาดการณ์ตัวแปรทางการเงินที่ซับซ้อน ได้แก่ NPV/ROI [7]

โครงข่ายประสาทเทียม (Artificial Neural Network) เป็นแบบจำลองเรียนรู้และจับรูปแบบที่ซับซ้อนและไม่เป็นเชิงเส้น ในชุดข้อมูลขนาดใหญ่ได้ดีเยี่ยม การนำ ANN มาเปรียบเทียบจึงเป็นการทดสอบว่าความสัมพันธ์ระหว่างปัจจัยนำเข้าและความคุ้มค่าทางการเงินมีความซับซ้อนถึงระดับที่ต้องการโครงสร้างแบบโครงข่ายประสาทเทียมในการทำนายหรือไม่ [8]

## 3. วิธีการดำเนินการวิจัย

### 3.1 ข้อมูลที่ใช้

พื้นที่ศึกษาวิสาหกิจชุมชนที่เข้าร่วมโครงการปลูกต้นไม้เพื่อคาร์บอนเครดิต ดังนี้ กรณีศึกษาที่ 1 โครงการธนาคารคาร์บอนสีเขียว บ้านท่าลี่ อำเภอหนองเรือ และบ้านแดง อำเภอบ้านฝาง จังหวัดขอนแก่น จำนวน 369 ไร่ 2 งาน 37 ตารางวา จำนวนไม้ยืนต้นทั้งหมด 15,926 ต้น [9] กรณีศึกษาที่ 2 คาร์บอนเครดิตธนาคารเพื่อการเกษตรและสหกรณ์การเกษตร จังหวัดขอนแก่น จำนวน 948 ไร่ 3 งาน 8 ตารางวา จำนวนไม้ยืนต้นทั้งหมด 32,276 ต้น [10]

### 3.2 การเตรียมข้อมูล

รวบรวมข้อมูลต้นทุนและผลตอบแทนจากองค์การบริหารจัดการก๊าซเรือนกระจก รวมถึงจัดทำชุดข้อมูลรวมตัวแปร ได้แก่ ขนาดแปลง/ไร่, จำนวนต้น, เส้นรอบวง, ชนิดต้นไม้, รายได้คาร์บอนเครดิต

3.2.1 การประมวลผลข้อมูล การทำความสะอาดข้อมูล (Data Cleaning) ตรวจสอบ Missing values ขนาดแปลง/ไร่,จำนวนต้น, เส้นรอบวง, ชนิดต้นไม้ที่ขาดหาย และแบ่งชุดข้อมูล (Train-Test Split) โดยใช้ชุดข้อมูลในชุดฝึกสอน (Training Data) ร้อยละ 70 ของข้อมูล และใช้ชุดข้อมูลชุดทดสอบ (Test Data) ร้อยละ 30 ของข้อมูลเพื่อใช้ในการประเมินประสิทธิภาพ

3.2.2 การสร้างแบบจำลองการเรียนรู้ของเครื่อง (Machine Learning Model Construction) แบบจำลองที่ใช้ในการศึกษา 4 แบบ ประกอบด้วย การถดถอยเชิงเส้น (Linear Regression) การถดถอยแบบสุ่มวน (Random Forest Regressor) การเสริมกำลังแบบเกรเดียนท์ (XGBoost) และโครงข่ายประสาทเทียม (Artificial Neural Network)

3.2.3 การประเมินแบบจำลอง (Model Evaluation) ใช้ตัวชี้วัดหลักสำหรับการถดถอย 3 ตัวชี้วัด ได้แก่ ค่าความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (Mean Absolute Error) ใช้สำหรับวัดค่าความต่างระหว่างค่าจริง (Actual) และค่าที่โมเดลทำนาย (Predicted) โดยใช้ค่า “สัมบูรณ์” (absolute value) แล้วหา “ค่าเฉลี่ย”

รากที่สองของค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Root Mean Squared Error) ใช้สำหรับวัดค่าความต่างระหว่างค่าจริงกับค่าที่ทำนายเหมือน MAE แต่จะยกกำลังสองก่อนแล้วค่อยถอดราก ทำให้ “ความผิดพลาดขนาดใหญ่” มีผลมากกว่า

สัมประสิทธิ์การตัดสินใจ Coefficient of Determination ($R^2$) ใช้สำหรับบอกว่าโมเดลสามารถอธิบาย “ความแปรปรวน” (variance) ของข้อมูลจริงได้มากน้อยเพียงใด

## 4. ผลการดำเนินงาน

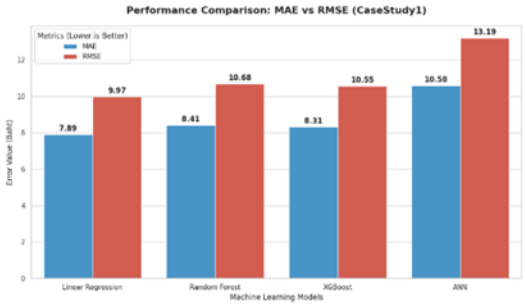
### 4.1 ผลการเปรียบเทียบประสิทธิภาพของแบบจำลอง Machine Learning

**ตารางที่ 1:** การเปรียบเทียบประสิทธิภาพของแบบจำลอง Machine Learning

| แบบจำลอง | MAE | RMSE | $R^2$ | NPV (20Y) | ROI (%) |
|---|---|---|---|---|---|
| **กรณีศึกษาที่ 1** | | | | | |
| Linear Regression | 7.89 | 9.97 | 1.00 | 101,490.90 | 3,920.80 |
| RFR | 8.41 | 10.68 | 1.00 | 101,490.38 | 3,920.78 |
| XGBoost | 8.31 | 10.55 | 1.00 | 101,491.84 | 3,920.83 |
| ANN | 10.58 | 13.19 | 1.00 | 101,406.63 | 3,917.61 |
| **กรณีศึกษาที่ 2** | | | | | |
| Linear Regression | 8.06 | 10.10 | 1.00 | 153,995.66 | 5,903.23 |
| RFR | 9.19 | 20.26 | 1.00 | 153,987.73 | 5,902.93 |
| XGBoost | 9.45 | 21.19 | 1.00 | 153,995.27 | 5,903.21 |
| ANN | 8.69 | 11.08 | 1.00 | 153,976.48 | 5,902.50 |

จากตารางที่ 1 การเปรียบเทียบประสิทธิภาพของแบบจำลอง Machine Learning ผลการทดลองพบว่าแบบจำลอง Linear Regression เป็นโมเดลที่เหมาะสมที่สุดในงานนี้ เนื่องจากมีความแม่นยำสูง และให้ผลลัพธ์ทางการเงินที่มั่นคงทั้งสองกรณีศึกษา กรณีศึกษาที่ 2 คือสถานการณ์ที่สร้างผลกำไรและความคุ้มค่าในการลงทุน (ROI) ได้มากกว่ากรณีแรกอย่างมีนัยสำคัญ

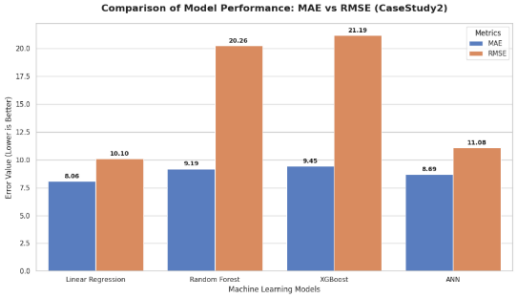
### 4.2 การเปรียบเทียบประสิทธิภาพของโมเดลโดยวัดจากค่าความคลาดเคลื่อน MAE และ RMSE กรณีศึกษาที่ 1




**ภาพที่ 1:** การเปรียบเทียบประสิทธิภาพของโมเดลโดยวัดจากค่าความคลาดเคลื่อน MAE และ RMSE กรณีศึกษาที่ 1

จากภาพที่ 1 พบว่าโมเดลแม่นยำสูงสุด Linear Regression เป็นแบบจำลองที่ให้ประสิทธิภาพดีที่สุด โดยมีค่าความคลาดเคลื่อนต่ำที่สุดคือ MAE เท่ากับ 7.89 และ RMSE เท่ากับ 9.97 ความสอดคล้องของข้อมูลทุกแบบจำลองให้ค่า $R^2$ เท่ากับ 1.00 ซึ่งยืนยันว่าข้อมูลชุดนี้มีความสัมพันธ์เชิงเส้นที่เสถียรมากจนโมเดลสามารถอธิบายความแปรผันได้ทั้งหมด

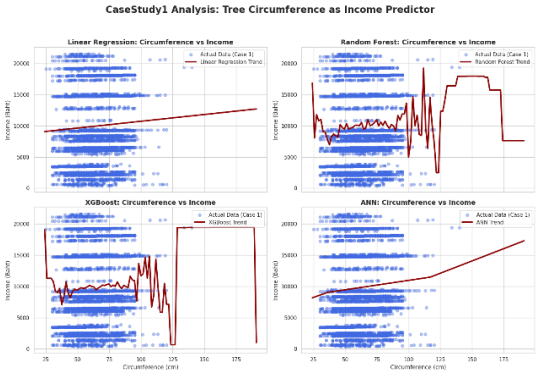
### 4.3 การเปรียบเทียบประสิทธิภาพของโมเดลโดยวัดจากค่าความคลาดเคลื่อน MAE และ RMSE กรณีศึกษาที่ 2




**ภาพที่ 2**: การเปรียบเทียบประสิทธิภาพของโมเดลโดยวัดจากค่าความคลาดเคลื่อน MAE และ RMSE กรณีศึกษาที่ 2

จากภาพที่ 2 พบว่าโมเดลที่แม่นยำที่สุด Linear Regression มีค่าความคลาดเคลื่อนต่ำที่สุดทั้ง MAE (8.06) และ RMSE (10.10) ซึ่งแสดงถึงความเสถียรและความแม่นยำสูงสุดในข้อมูลชุดนี้ โมเดลอันดับสอง ANN (Artificial Neural Network) ทำคะแนนได้ใกล้เคียงกับ Linear Regression มาก โดยมีค่า MAE อยู่ที่ 8.69 และ RMSE ที่ 11.08 โมเดล Random Forest และ XGBoost มีค่า RMSE ที่กระโดดสูงขึ้นไปมากกว่า 20 (ในขณะที่ MAE อยู่ที่ประมาณ 9) สิ่งนี้บ่งบอกว่าโมเดลทั้งสองอาจมีการทำนายที่ผิดพลาดเยอะในบางข้อมูล (Outliers) เพราะค่า RMSE จะไวต่อค่าความผิดพลาดที่รุนแรง

### 4.4 กราฟจุดแสดงความสัมพันธ์ระหว่างเส้นรอบวง (Circumference) และ รายได้ (Income) กรณีศึกษาที่ 1




**ภาพที่ 3:** กราฟจุดแสดงความสัมพันธ์ระหว่างเส้นรอบวง (Circumference) และ รายได้ (Income) กรณีศึกษาที่ 1

จากภาพที่ 3 พบว่าLinear Regression แสดงแนวโน้มเป็นเส้นตรงที่ค่อยๆ เพิ่มขึ้นตามขนาดเส้นรอบวง เป็นโมเดลที่เรียบง่ายที่สุดแต่ความละเอียดต่ำ ANN (Artificial Neural Network) ให้ผลคล้าย Linear Regression แต่มีความยืดหยุ่นกว่าเล็กน้อย โดยเห็นความชันที่เพิ่มขึ้นชัดเจนขึ้นในช่วงปลาย Random Forest & XGBoost ทั้งสองโมเดลพยายาม "จับกลุ่ม" ข้อมูล (Overfitting) ทำให้เส้นทำนายมีความผันผวนสูง โดยเฉพาะ XGBoost ที่มีการกระโดดของค่าพยากรณ์ในช่วงปลายอย่างรุนแรง ซึ่งอาจเกิดจากข้อมูลในช่วงนั้นมีน้อย

**4.5 กราฟจุดแสดงความสัมพันธ์ระหว่าง เส้นรอบวง (Circumference) และ รายได้ (Income) กรณีศึกษาที่ 2**




**ภาพที่ 4:** กราฟจุดแสดงความสัมพันธ์ระหว่าง เส้นรอบวง (Circumference) และ รายได้ (Income) กรณีศึกษาที่ 2

จากภาพที่ 4 พบว่าLinear Regression แสดงเส้นพยากรณ์เป็นเส้นตรงที่เพิ่มขึ้นตามขนาดเส้นรอบวง เป็นโมเดลที่มองเห็นภาพรวมของความสัมพันธ์เชิงบวก (Positive Correlation) ได้ชัดเจนที่สุด แต่ขาดความละเอียดในการจับกลุ่มข้อมูลที่กระจายตัวสูง ANN (Artificial Neural Network) ให้ผลลัพธ์คล้ายกับ Linear Regression แต่มีความโค้งมนยืดหยุ่นกว่าเล็กน้อย แสดงให้เห็นแนวโน้มรายได้ที่เพิ่มขึ้นตามขนาดต้นไม้ในลักษณะที่ไม่เป็นเส้นตรงสมบูรณ์ Random Forest เส้นทำนายมีความผันผวนสูงมาก เนื่องจากพยายามปรับตัวตามจุดข้อมูลในแต่ละช่วงอย่างละเอียด ซึ่งอาจนำไปสู่ปัญหา Overfitting หรือการพยากรณ์ที่ผิดเพี้ยนไปตามสัญญาณรบกวนของข้อมูล (Noise) XGBoost ให้เส้นพยากรณ์ที่ดูคงที่ (Flat) ในช่วงกลาง ก่อนจะมีการกระโดดของค่าในช่วงปลาย แสดงให้เห็นว่าโมเดลพยายามแบ่งเกณฑ์รายได้เป็นช่วง ๆ ตามความหนาแน่นของข้อมูล

## 5. สรุปผลและอภิปรายผล

งานวิจัยนี้ได้นำเสนอการศึกษาเปรียบเทียบประสิทธิภาพของแบบจำลองการเรียนรู้ของเครื่องเพื่อคาดการณ์ความคุ้มค่าของคาร์บอนเครดิตจากการลงทุนปลูกต้นไม้ ในภาควิสาหกิจชุมชน จังหวัดขอนแก่น ด้วยเทคนิคการเรียนรู้ของเครื่อง 4 รูปแบบ ผลการศึกษาที่พบว่า Linear Regression มีประสิทธิภาพเหนือกว่าแบบจำลองที่ซับซ้อนอย่าง XGBoost ในชุดข้อมูลนี้ สอดคล้องกับทฤษฎีพื้นฐานที่ว่าแบบจำลองที่เรียบง่ายจะทำงานได้ดีกว่าเมื่อความสัมพันธ์ของตัวแปรเป็นเชิงเส้นที่ชัดเจน [5] ซึ่งแตกต่างจากผลการศึกษาของ [11] ที่พบว่าเทคนิคการเรียนรู้ของเครื่องขั้นสูงให้ผลลัพธ์ที่ดีกว่าในการคาดการณ์การกักเก็บคาร์บอนในระบบเกษตรป่าไม้ที่มีความหลากหลายทางชีวภาพและโครงสร้างซับซ้อน นอกจากนี้ ปัญหา Overfitting ที่พบใน Random Forest และ XGBoost ผลลัพธ์แบบขั้นบันได ชี้ให้เห็นถึงข้อจำกัดของโมเดลประเภท Tree-based เมื่อเผชิญกับข้อมูลที่มีการกระจายตัวไม่สม่ำเสมอหรือไม่เป็นธรรมชาติ ตามผลการศึกษาที่ [6] ได้ระบุไว้ว่าแบบจำลองเหล่านี้ต้องการการปรับแต่งพารามิเตอร์ (Hyperparameter Tuning) ที่ซับซ้อนกว่าปกติหากข้อมูลมีปฏิสัมพันธ์ที่เป็นเส้นตรงมากเกินไป ในด้านการเงิน การที่ทุกกรณีศึกษามีค่าอัตราผลตอบแทนจากการลงทุน (ROI) เป็นบวกนั้น สอดคล้องกับผลการศึกษาของ [4] ที่ระบุว่าการลงทุนปลูกไม้เศรษฐกิจเพื่อคาร์บอนเครดิตในประเทศไทยมีศักยภาพในการสร้างผลตอบแทนที่คุ้มค่าและยั่งยืนกว่าการทำเกษตรเชิงเดี่ยว ผลประโยชน์ร่วม (Co-benefits) ที่เกิดขึ้นจากการดำเนินงานโดยวิสาหกิจชุมชนในจังหวัดขอนแก่น ยังเป็นปัจจัยสำคัญที่ช่วยเสริมสร้างมูลค่าของคาร์บอนเครดิตในตลาดภาคสมัครใจ (Voluntary Carbon Market) ให้สูงขึ้น ซึ่งเป็นไปตามแนวทางที่ระบุไว้ในงานวิจัยของ [12]

## เอกสารอ้างอิง


[1] UNFCCC. (2021). Voluntary Carbon Market developments. United Nations Framework Convention on Climate Change.

[2] FAO. (2020). Forests and climate change in Asia and the Pacific. Food and Agriculture Organization of the United Nations.

[3] องค์การบริหารจัดการก๊าซเรือนกระจก. (2023). โครงการ T-VER และฐานข้อมูลโครงการ. สืบค้นเมื่อ 2 มกราคม 2569, จาก https://ghgreduction.tgo.or.th

[4] Wongcharoen, A. (2023). Financial Feasibility Analysis of Economic Tree Plantations for Carbon Credits in Thailand. Journal of Sustainable Development and Green Economics, 10(2), 15-30.

[5] Puntumapon, K., Wuthithanakul, A., Roeksiri, P. E., & Sundar, B. S. (2023). Thai Dry-Evergreen Forest's Biomass Estimation using Machine Learning Models. Paper presented at the 2023 1st IEEE International Conference on Smart Information Technology (SmartInfo).

[6] Breiman, L. (2001). *Random Forests.* Machine Learning, 45(1), 5–32.

[7] Chen, T., & Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 785–794.

[8] Abiodun, O.I., et al. (2019). State-of-the-art in artificial neural network applications: A survey. Heliyon, 5(11), e00938.

[9] องค์การบริหารจัดการก๊าซเรือนกระจก. (2563, 28 สิงหาคม). โครงการธนาคารคาร์บอนสีเขียว บ้านท่าลี่ อำเภอหนองเรือ และ บ้านแดง อำเภอบ้านฝาง จังหวัดขอนแก่น. T-VER ฐานข้อมูลโครงการลดก๊าซเรือนกระจก. สืบค้นเมื่อ 1 ตุลาคม 2568, จาก https://ghgreduction.tgo.or.th/th/tver-database-and-statistics/t-ver-registered-project/item/1382-green-carbon-bank-of-ban-talee-nongruea-disrict-and-ban-daeng-ban-fang-district-khonkaen-province.html

[10] องค์การบริหารจัดการก๊าซเรือนกระจก. (2568). กลุ่มโครงการย่อย: คาร์บอนเครดิตธนาคารเพื่อการเกษตรและสหกรณ์การเกษตร จังหวัดขอนแก่น [BAAC Carbon Credit Khon Kaen Province]. T-VER. สืบค้นเมื่อ 1 ตุลาคม 2568, จาก https://ghgreduction.tgo.or.th/th/tver-database-and-statistics/programme-of activities/141/component-project-activity/5717-baac-carbon-credit-khon-kaen-province.html

[11] Sriplung, N., et al. (2022). Application of Machine Learning for Carbon Sequestration Prediction in Agroforestry Systems. Journal of Environmental Management, 301, 113894.

[12] Pradhan, B., et al. (2021). Carbon Credit Valuation of Forestry Projects under the Voluntary Carbon Market: A Case Study. Journal of Cleaner Production, 278, 123961.

# การวิเคราะห์เจตนาความคิดเห็นของผู้บริโภคในอุตสาหกรรมเครื่องดื่มบนสื่อสังคมออนไลน์ด้วยเทคนิคการเรียนรู้เชิงลึก

# Consumer Intent Analysis in the Beverage Industry on Social Media Platforms Using Deep Learning Techniques

*กีรติ จันทร (Keerati Jantorn)[1] และวัชรีวรรณ จิตต์สกุล (Watchareewan Jitsakul)[2]*
*ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล*
*มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ*
*[1]s6707021856043@email.kmutnb.ac.th, [2]watchareewan.j@itd.kmutnb.ac.th*

## บทคัดย่อ

*งานวิจัยนี้มีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพของเทคนิคการเรียนรู้เชิงลึกในการจำแนกเจตนาความคิดเห็นของผู้บริโภคในอุตสาหกรรมเครื่องดื่ม โดยใช้ชุดข้อมูลจำนวน 2,300 รายการ แบ่งเป็น 3 ประเภท ได้แก่ ความคิดเห็นเชิงวิจารณ์และบอกเล่า ความคิดเห็นเชิงคำถาม และ การแจ้งปัญหา ข้อมูลผ่านกระบวนแปลงให้อยู่ในรูปแบบที่เหมาะสมสำหรับการฝึกแบบจำลอง และการแก้ปัญหาข้อมูลไม่สมดุลด้วยเทคนิคค่าน้ำหนักของคลาส จากนั้นทำการแบ่งข้อมูลเพื่อทดสอบด้วยวิธีการตรวจสอบแบบไขว้ 5 ส่วน เพื่อเปรียบเทียบเทคนิคการเรียนรู้เชิงลึก 3 เทคนิค ได้แก่ โครงข่ายหน่วยความจำระยะสั้นแบบยาวสองทิศทาง โครงข่ายหน่วยหมุนเวียนที่มีประตูประตูสองทิศทาง และการแทนค่าภาษาแบบเข้ารหัสสองทิศทางด้วยทรานส์ฟอร์เมอร์ ผลการวิจัยพบว่าแบบจำลองที่ใช้เทคนิคโครงข่ายหน่วยหมุนเวียนที่มีประตูประตูสองทิศทาง มีประสิทธิภาพสูงสุด หลังจากทดสอบโดยใช้ข้อมูลที่ไม่ผ่านการฝึกฝนแบบจำลอง โดยมีค่าความถูกต้องร้อยละ 86.87 ค่าความแม่นยำ 0.87 ค่าระลึก 0.87 ค่าการถ่วงดุล 0.87 และค่าพื้นที่ใต้เส้นโค้งคุณลักษณะสมบัติการทำงานเท่ากับ 0.94 ตามลำดับ*

**คำสำคัญ:** หน่วยความจำระยะสั้นแบบยาวสองทิศทาง หน่วยเวียนกลับแบบมีประตูสองทิศทาง การแทนค่าภาษาแบบเข้ารหัสสองทิศทางด้วยทรานส์ฟอร์เมอร์

## Abstract

*This research aims to compare the performance of deep learning techniques in classifying consumer intent in comments from the beverage industry. The dataset comprises 2,300 entries, categorized into three types: critical and narrative comments, question-based comments, and problem reports. The data was prepared through text cleaning, word segmentation, text-to-sequence conversion, sequence padding, label encoding, and addressing data imbalance using class weight techniques. The data was then split for testing using 5-fold cross-validation to compare 3 deep learning techniques: Bidirectional Long Short-Term Memory networks, Bidirectional Gated Recurrent Units, and Bidirectional Encoder Representations from Transformers. The results showed that the model using Bidirectional Gated Recurrent Units performed best when tested on unseen data, achieving an accuracy of 86.87%, precision of 0.87, recall of 0.87, F1-score of 0.87, and ROC of 0.94.*

**Keywords:** Bidirectional Long Short-Term Memory, Bidirectional Gated Recurrent Unit, Bidirectional Encoder Representations from Transformers.

## 1. บทนำ

ในโลกยุคดิจิทัลที่เทคโนโลยีสารสนเทศเข้ามามีบทบาทสำคัญต่อการใช้ชีวิตของผู้คน การสื่อสารผ่านสื่อสังคมออนไลน์ได้กลายเป็นช่องทางหลักที่ผู้บริโภคใช้ในการแสดงความคิดเห็น แลกเปลี่ยนประสบการณ์ และความรู้สึกเกี่ยวกับสินค้าและบริการต่าง ๆโดยเฉพาะอย่างยิ่งในอุตสาหกรรมเครื่องดื่ม ซึ่งเป็นอุตสาหกรรมที่มีการแข่งขันสูงและมีการเปิดตัวสินค้าใหม่อยู่เสมอ การรับรู้และเข้าใจเสียงของผู้บริโภคจึงเป็นหัวใจสำคัญในการพัฒนาสินค้าและกลยุทธ์ทางการตลาด

การจำแนกเจตนา จึงเป็นเทคนิคด้านประมวลผลที่มีความสำคัญและได้รับความสนใจอย่างมากในงานวิเคราะห์ข้อความ เนื่องจากเป็นกระบวนการในการระบุความต้องการหรือจุดประสงค์ของผู้เขียนข้อความ ซึ่งในบริบทของการตลาด การเข้าใจเจตนาเช่นนี้สามารถช่วยให้องค์กรสามารถตอบสนองต่อผู้บริโภคได้ตรงจุด ไม่ว่าจะเป็นการปรับปรุงผลิตภัณฑ์ การจัดโปรโมชั่น การตอบคำถามหรือข้อกังวลอย่างรวดเร็ว สำหรับอุตสาหกรรมเครื่องดื่ม เจตนาของข้อความที่เกี่ยวข้องกับผลิตภัณฑ์อาจแบ่งได้หลากหลาย เช่น การสอบถามข้อมูล การวิจารณ์รสชาติหรือบริการ การแนะนำสินค้าเพิ่มเติม หรือการบอกเล่าประสบการณ์ การแจ้งปัญหา ซึ่งหากสามารถจำแนกข้อความเหล่านี้ได้อย่างถูกต้อง จะช่วยให้ผู้ประกอบการสามารถใช้ข้อมูลจากสื่อสังคมออนไลน์มาเป็นเครื่องมือสำคัญในการตัดสินใจ ทั้งในด้านกลยุทธ์การตลาด การสื่อสารกับลูกค้า และการพัฒนาสินค้าให้สอดคล้องกับความต้องการของผู้บริโภคอย่างแท้จริง

ดังนั้น การศึกษาวิจัยที่มุ่งเน้นไปที่การจำแนกเจตนาของข้อความภาษาไทยในสื่อสังคมออนไลน์ โดยเฉพาะในบริบทของอุตสาหกรรมเครื่องดื่ม จึงมีความสำคัญทั้งในเชิงวิชาการและการประยุกต์ใช้งานจริง เพราะไม่เพียงแต่จะช่วยให้เข้าใจความต้องการของผู้บริโภคได้อย่างแม่นยำและลึกซึ้ง แต่ยังเป็นแนวทางหนึ่งที่สามารถส่งเสริมการใช้เทคนิคการเรียนรู้เชิงลึก [1] ในบริบทภาษาไทย ซึ่งยังถือว่าอยู่ในระยะเริ่มต้นเมื่อเทียบกับภาษาอื่น ๆ อีกด้วย

## 2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ผู้วิจัยได้ทำการรวบรวมข้อมูลเชิงทฤษฎีและผลงานวิจัยที่เกี่ยวข้อง เพื่อสร้างพื้นฐานความรู้และกรอบแนวคิดในการศึกษา โดยแบ่งออกเป็น 3 หัวข้อหลัก ดังนี้

### 2.1 การทำเหมืองข้อความ

เป็นกระบวนการแปลงข้อความที่ไม่มีโครงสร้างให้เป็นรูปแบบที่มีโครงสร้าง เพื่อระบุรูปแบบที่มีความหมายและค้นหาข้อมูลเชิงลึกใหม่ ๆ โดยสามารถใช้การทำเหมืองข้อความในการวิเคราะห์ข้อมูลข้อความจำนวนมาก เพื่อจับแนวคิดสำคัญ แนวโน้ม และความสัมพันธ์ที่ซ่อนอยู่ ด้วยการประยุกต์ใช้เทคนิคการวิเคราะห์ขั้นสูง เช่น Naïve Bayes, Support Vector Machines (SVM) และเทคนิคการเรียนรู้เชิงลึก (Deep Learning) [2]

### 2.2 เทคนิคการเรียนรู้เชิงลึก

ในการวิจัยครั้งนี้ ได้มีการนำเทคนิคการเรียนรู้เชิงลึกมาทำการสร้างแบบจำลอง ได้แก่

**2.2.1 หน่วยความจำระยะสั้นแบบยาวสองทิศทาง** (Bidirectional Long Short Term Memory: Bi-LSTM) ถูกพัฒนาจากโครงข่าย (Long Short-Term Memory: LSTM) แบบดั้งเดิมแตกต่างจาก LSTM ทั่วไปประมวลผลลำดับข้อมูลในทิศทางเดียว Bi-LSTM จะให้ข้อมูลไหลได้ทั้งในทิศทางไปข้างหน้าและย้อนกลับ ซึ่งช่วยให้สามารถจับบริบทของข้อมูลได้อย่างครอบคลุมและสมบูรณ์ยิ่งขึ้น [3]

**2.2.2 หน่วยเวียนกลับแบบมีประตูสองทิศทาง** (Bidirectional Gated Recurrent Unit : Bi-GRU) เป็นรูปแบบหนึ่งของ GRU ที่สามารถประมวลผลลำดับข้อมูลได้ทั้งสองทิศทาง คือ จากต้นไปยังปลาย และจากปลายย้อนกลับไปต้น ซึ่งช่วยให้โมเดลสามารถใช้ข้อมูลจากทั้งอดีตและอนาคตในลำดับข้อมูลได้ [4]

**2.2.3 การแทนค่าภาษาแบบเข้ารหัสสองทิศทางด้วยทรานส์ฟอร์เมอร์** (Bidirectional Encoder Representations from Transformers: BERT) เป็นโมเดลประมวลผลข้อความในทิศทาง ใช้ฝึกฝนโมเดล โดยปิดบังคำบางคำในประโยคแล้วให้โมเดลทำนายคำที่ถูกปิดบัง ทำให้สามารถเรียนรู้บริบทของคำจากทั้งสองทิศทางพร้อมกัน และจับความสัมพันธ์ระหว่างคำในประโยคได้อย่างมีประสิทธิภาพ [5]

### 2.3 งานวิจัยที่เกี่ยวข้อง

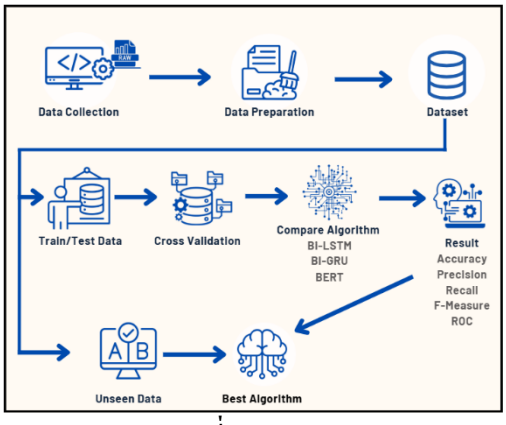
Derbentsev et al. [6] ได้ทำการศึกษาเปรียบเทียบโมเดลการเรียนรู้เชิงลึกสำหรับการวิเคราะห์ความรู้สึกจากข้อความโซเชียลมีเดีย โดยทำการเปรียบเทียบด้วยเทคนิคการเรียนรู้เชิงลึกจำนวน 3 เทคนิค ได้แก่ CNN, CNN-LSTM และ Bi-LSTM-CNN ทดสอบกับชุดข้อมูล 2 ชุด คือ IMDb และ Sentiment140 ผลการวิจัยพบว่าแบบจำลอง CNN มีประสิทธิภาพดีที่สุดบนชุดข้อมูล IMDb ซึ่งมีค่าความถูกต้องร้อยละ 90.09 ขณะที่แบบจำลอง BiLSTM-CNN ให้ผลลัพธ์ดีที่สุดบนชุดข้อมูล Sentiment140 ด้วยค่าความถูกต้องร้อยละ 82.10 แสดงให้เห็นว่าโครงสร้างของ CNN และ BiLSTM-CNN เหมาะสมต่อการวิเคราะห์ความคิดเห็นจากข้อความในโซเชียลมีเดีย

Chaudhary et al. [7] ได้ทำการทบทวนโมเดลการเรียนรู้เชิงลึกสำหรับการวิเคราะห์ความรู้สึกบน Twitter โดยทำการเปรียบเทียบด้วยเทคนิคการเรียนรู้เชิงลึก 7 เทคนิค ได้แก่ Convolutional Neural Network (CNN), Deep Neural Network (DNN), Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM), Bidirectional LSTM (Bi-LSTM), Gated Recurrent Unit (GRU) และ BERT ผลการวิจัยพบว่า แบบจำลอง BERT และ Bi-LSTM มีประสิทธิภาพสูงสุด โดยเฉพาะในบริบทของการจัดการข้อความสั้นที่มีลักษณะไม่เป็นทางการ ซึ่งแสดงให้เห็นว่าโมเดลที่สามารถจับบริบทแบบสองทิศทางเหมาะสมต่อการวิเคราะห์ความคิดเห็นจากโซเชียลมีเดียที่มีข้อความสั้นและไม่เป็นทางการ

Nugroho et al. [8] ได้ทำการศึกษาการจำแนกข่าวขนาดใหญ่โดยใช้โมเดลภาษา BERT ร่วมกับ Spark NLP โดยทำการเปรียบเทียบประสิทธิภาพของโมเดล BERT ในขนาดต่างๆ ได้แก่ Tiny, Mini, Small, Medium และ Base ทั้งแบบที่ใช้และไม่ใช้ Spark NLP ผลการวิจัยพบว่า แบบจำลอง BERT ที่ไม่ใช้ Spark NLP มีค่าความถูกต้องสูงกว่า คือ ร้อยละ 91.87 ขณะที่แบบจำลอง BERT ที่ใช้ Spark NLP มีค่าความถูกต้องเฉลี่ยร้อยละ 86.65 แต่สามารถลดเวลาในการฝึกฝนลงได้ถึงร้อยละ 62.9 ทั้งนี้แสดงให้เห็นว่า Spark NLP เหมาะสมสำหรับการใช้งานในกรณีที่มีทรัพยากรจำกัดและต้องการความเร็วในการประมวลผล

## 3. วิธีการดำเนินการวิจัย

งานวิจัยนี้มีวัตถุประสงค์เพื่อเปรียบเทียบเทคนิคการเรียนรู้เชิงลึกที่เหมาะสมสำหรับการจำแนกการจำแนกเจตนาเกี่ยวกับสินค้าและบริการ โดยมีกรอบแนวคิดแสดงดังภาพที่ 1




**ภาพที่ 1:** กรอบแนวคิด

จากภาพที่ 1 แสดงกรอบแนวคิดของงานวิจัยประกอบด้วย 2 ส่วนหลัก ได้แก่ การเก็บรวบรวม และจัดเตรียมข้อมูล และการพัฒนาตัวแบบ และเปรียบเทียบประสิทธิภาพ

### 3.1 การเก็บรวบรวม และจัดเตรียมข้อมูล

ทำการเก็บรวบรวมรวบรวมจาก สื่อสังคมออนไลน์ โดยเน้นในบริบทของ อุตสาหกรรมน้ำดื่มและเครื่องดื่ม ในช่วงระหว่างเดือนมกราคม พ.ศ.2568 ถึงเดือนตุลาคม พ.ศ.2568 ดำเนินการคัดเลือกความคิดเห็นจำนวน 2,300 รายการ แบ่งเป็น 3 ประเภท ได้แก่ ความคิดเห็นเชิงวิจารณ์และบอกเล่า ความคิดเห็นเชิงคำถาม และ การแจ้งปัญหา รายละเอียดของตัวอย่างความคิดเห็นแสดงดังตารางที่ 1

**ตารางที่ 1:** ตัวอย่างข้อมูลความคิดเห็น

| ความคิดเห็น | ประเภท |
|---|---|
| น้ำเพอร์ อร่อยจริงค่ะ จนเราไปเจอน้ำโคว่าย์ รสชาติเหมือนกันเป๊ะเลย ก็เลยตัดสินใจซื้อ □▨ | Review |
| ขายเครื่องดื่ม บางเจ้าเค้าใช้น้ำแข็งที่ไม่ได้ทำจากน้ำธรรมดา | Review |
| นมกล่อง UHT สามารถเอาไปเวฟ ทั้งกล่องได้มั้ย | Question |
| หมดค่ากาแฟ วันละกี่บาท | Question |

| ความคิดเห็น | ประเภท |
|---|---|
| โซดามีสิ่งแปลกปลอมคล้ายสนิมบริเวณปากขวด | Problem Report |
| ตะกอนขาวและเมื่อวางให้แสงส่อง จะมีสีฟ้าอ่อน ๆ ภายในขวด | Problem Report |

จากตารางที่ 1 ผู้วิจัยได้นำข้อมูลเข้าสู่กระบวนการจัดเตรียมข้อมูล ก่อนนำไปใช้ในการพัฒนาแบบจำลอง ซึ่งกระบวนการดังกล่าวประกอบด้วย 5 ขั้นตอน ดังต่อไปนี้

**3.1.1 การทำความสะอาดข้อมูล** เป็นการปรับปรุงคุณภาพของข้อมูล โดยการลบอักขระพิเศษ การกำจัดช่องว่างการขึ้นบรรทัดใหม่ และการลบสัญลักษณ์ที่ไม่เกี่ยวข้อง

**3.1.2 การตัดคำ** เป็นกระบวนการแยกข้อความออกเป็นหน่วยคำด้วยเทคนิค Deepcut สำหรับ Bi-LSTM, Bi-GRU และ เทคนิค Subword Tokenization สำหรับ BERT

**3.1.3 การแปลงข้อความเป็นลำดับตัวเลข** เป็นการเตรียมข้อมูล โดยสร้างคลังคำศัพท์และแปลงข้อความเป็นลำดับของดัชนีตัวเลข จากนั้นทำ Padding ให้ข้อความทั้งหมดมีความเท่ากัน

**3.1.4 การแปลงป้ายกำกับข้อความเป็นตัวเลข** เป็นกระบวนการแปลงป้ายกำกับในรูปแบบข้อความเป็นค่าตัวเลข โดย 0 แทนความคิดเห็นเชิงวิจารณ์และบอกเล่า 1 แทนความคิดเห็นเชิงคำถาม และ 2 แทนการแจ้งปัญหา

จากนั้น ข้อมูลที่ผ่านกระบวนการเตรียมความพร้อมเรียบร้อยแล้ว จะถูกนำไปใช้ในการพัฒนาตัวแบบต่อไป

**3.1.5 การปรับสมดุลข้อมูล** เป็นกระบวนการปรับสมดุลข้อมูลโดยใช้เทคนิค Class Weight ซึ่งเป็นการกำหนดน้ำหนักให้กับแต่ละคลาสในระหว่างการฝึกฝนโมเดล โดยให้น้ำหนักที่สูงกว่ากับคลาสที่มีข้อมูลน้อย และน้ำหนักที่ต่ำกว่ากับคลาสที่มีข้อมูลมาก ทำให้โมเดลให้ความสำคัญกับการทำนายคลาสที่มีข้อมูลน้อยมากขึ้น และช่วยลดปัญหาความไม่สมดุลของข้อมูลโดยไม่ต้องเพิ่มหรือลดจำนวนข้อมูลจริง

**3.1.5 การประเมินประสิทธิภาพของแบบจำลอง** ใช้วิธี Stratified K-Fold Cross-Validation โดยกำหนดค่า (k = 5) เพื่อแบ่งชุดข้อมูลออกเป็น 5 ส่วนย่อย ซึ่งในแต่ละ fold จะมีสัดส่วนของคลาสข้อมูลใกล้เคียงกับสัดส่วนของข้อมูลทั้งหมด วิธีการดังกล่าวช่วยลดอคติที่อาจเกิดจากการสุ่มแบ่งข้อมูลแบบทั่วไป และทำให้การประเมินผลมีความน่าเชื่อถือมากยิ่งขึ้น โดยในแต่ละรอบจะใช้ข้อมูลจำนวน 4 folds สำหรับฝึกแบบจำลอง และใช้ข้อมูลอีก 1 fold สำหรับทดสอบ จากนั้นทำการวนซ้ำจนครบทั้ง 5 folds และนำผลลัพธ์ที่ได้จากทุก fold มาหาค่าเฉลี่ยเพื่อใช้เป็นตัวแทนประสิทธิภาพโดยรวมของแบบจำลอง

**3.2 การพัฒนาตัวแบบ และเปรียบเทียบประสิทธิภาพ**

ทำการสร้างแบบจำลองด้วยเทคนิคการเรียนรู้เชิงลึก 4 เทคนิค ได้แก่ Bi-LSTM, Bi-GRU และเทคนิคทรานส์ฟอร์เมอร์ BERT โดยทำการปรับค่าพารามิเตอร์สำหรับตัวแบบ แสดงดังตารางที่ 2 สำหรับ Bi-LSTM และ Bi-GRU และตารางที่ 3 สำหรับ BERT

**ตารางที่ 2:** ค่าพารามิเตอร์ที่ใช้ในการปรับแต่งสำหรับ Bi-LSTM, Bi-GRU

| Parameter | Value |
|---|---|
| Unit Out | 32, 64, 128, 256, 512 |
| Dense Units | 32, 64, 128, 256, 512 |
| Epochs | 30, 40, 50 |
| Learning Rate | 0.01,0.001,0.0001 |

จากตารางที่ 2 ผู้วิจัยได้ทำการปรับค่าพารามิเตอร์จำนวน 225 รูปแบบ โดยกำหนดค่า Unit Out เท่ากับ [32, 64, 128, 256, 512] ค่า Dense Unit เท่ากับ [32, 64, 128, 256, 512] ค่า Epochs เท่ากับ [30, 40, 50] และค่า Learning Rate เท่ากับ [0.01, 0.001, 0.0001]

**ตารางที่ 3:** ค่าพารามิเตอร์ที่ใช้ในการปรับแต่งสำหรับ BERT

| Parameter | Value |
|---|---|
| Epochs | 3, 4, 5 |
| Learning Rate | 2e-5, 3e-5, 5e-5 |

จากตารางที่ 3 ผู้วิจัยได้ทำการปรับค่าพารามิเตอร์โดยกำหนดค่า Learning Rate และค่า Epochs เพื่อค้นหาค่าที่เหมาะสมที่สุด โดยกำหนดค่า Learning Rate เท่ากับ 2e-5, 3e-5, 5e-5 และค่า Epochs เท่ากับ [3, 4, 5]

**ตารางที่ 4:** ค่าพารามิเตอร์คงที่สำหรับ Bi-LSTM, Bi-GRU

| Parameter | Value |
|---|---|
| Max Length | 113 |
| Dropout | 0.5 |
| Embed Size | 32 |
| Optimizer | Adam |
| Batch Size | 64 |

จากตารางที่ 4 ผู้วิจัยได้กำหนดค่าพารามิเตอร์คงที่ ได้แก่ ค่า Max Length เท่ากับ 113 ค่า Dropout เท่ากับ 0.5 ค่า Embedding Size เท่ากับ 32 ค่า Optimizer เท่ากับ Adam และค่า Batch Size เท่ากับ 64

**ตารางที่ 5:** ค่าพารามิเตอร์คงที่สำหรับ BERT

| Parameter | Value |
|---|---|
| Max Length | 113 |
| Loss Function | Sparse Categorical Cross-Entropy |
| Optimizer | Adam |
| Batch Size | 8 |

จากตารางที่ 5 ผู้วิจัยได้กำหนดค่าพารามิเตอร์คงที่ของแบบจำลอง BERT ได้แก่ ค่า Max Length เท่ากับ 113 ค่า Loss Function เป็นแบบ Sparse Categorical Cross-Entropy ค่า Optimizer เท่ากับ Adam และค่า Batch Size เท่ากับ 8

## 4. ผลการดำเนินงาน

ทำการประเมินประสิทธิภาพของตัวแบบด้วยค่าความถูกต้อง (Accuracy: A) ค่าความแม่นยำ (Precision: P) ค่าระลึก (Recall :R) ค่าการถ่วงดุล (F1-Score: F) และค่าพื้นที่ใต้เส้นโค้งคุณลักษณะสมบัติการทำงาน (ROC AUC: ROC) ผลการดำเนินงาน แสดงดังตารางที่ 6 ถึง 8

**ตารางที่ 6:** ผลลัพธ์ของตัวแบบ Bi-LSTM

| Parameter | | | Performance | | | | |
|---|---|---|---|---|---|---|---|
| Unit | Dense Unit | Epoch | A(%) | P | R | F | ROC |
| **32** | **32** | **40** | **91.92** | **0.92** | **0.92** | **0.92** | **0.98** |
| 32 | 64 | 30 | 91.65 | 0.92 | 0.92 | 0.92 | 0.98 |
| 64 | 128 | 30 | 91.53 | 0.92 | 0.92 | 0.92 | 0.98 |
| 32 | 512 | 40 | 91.48 | 0.92 | 0.91 | 0.92 | 0.98 |
| 32 | 128 | 40 | 91.40 | 0.92 | 0.91 | 0.91 | 0.98 |
| 64 | 64 | 50 | 91.35 | 0.92 | 0.91 | 0.92 | 0.98 |
| 32 | 32 | 30 | 91.27 | 0.91 | 0.91 | 0.91 | 0.98 |
| 64 | 32 | 30 | 91.27 | 0.92 | 0.91 | 0.91 | 0.98 |
| 32 | 32 | 50 | 91.18 | 0.92 | 0.91 | 0.91 | 0.98 |
| 32 | 512 | 50 | 91.14 | 0.91 | 0.91 | 0.91 | 0.98 |

จากตารางที่ 5 ตัวแบบ Bi-LSTM ที่ปรับค่าพารามิเตอร์ Bi-LSTM Unit เท่ากับ 32 ค่า Dense Unit เท่ากับ 32 และค่า Epoch เท่ากับ 40 ให้ประสิทธิภาพดีที่สุดด้วยค่าความถูกต้องร้อยละ 91.92

**ตารางที่ 7:** ผลลัพธ์ของตัวแบบ Bi-GRU

| Parameter | | | Performance | | | | |
|---|---|---|---|---|---|---|---|
| Unit | Dense Unit | Epoch | A(%) | P | R | F | ROC |
| **64** | **32** | **50** | **91.52** | **0.92** | **0.92** | **0.92** | **0.98** |
| 32 | 32 | 30 | 91.27 | 0.92 | 0.91 | 0.91 | 0.98 |
| 64 | 32 | 50 | 91.18 | 0.91 | 0.91 | 0.91 | 0.98 |
| 32 | 256 | 30 | 91.14 | 0.91 | 0.91 | 0.91 | 0.98 |
| 32 | 64 | 30 | 91.14 | 0.92 | 0.91 | 0.91 | 0.98 |
| 512 | 512 | 50 | 91.09 | 0.91 | 0.91 | 0.91 | 0.98 |
| 32 | 32 | 50 | 91.05 | 0.92 | 0.91 | 0.91 | 0.98 |
| 32 | 32 | 40 | 90.96 | 0.92 | 0.91 | 0.91 | 0.98 |
| 32 | 32 | 50 | 90.96 | 0.91 | 0.91 | 0.91 | 0.98 |
| 32 | 128 | 40 | 90.96 | 0.91 | 0.91 | 0.91 | 0.98 |

จากตารางที่ 7 ตัวแบบ Bi-GRU ที่ปรับค่าพารามิเตอร์ Bi-GRU Unit เท่ากับ 64 ค่า Dense Unit เท่ากับ 32 และค่า Epoch เท่ากับ 50 ให้ประสิทธิภาพดีที่สุดด้วยค่าความถูกต้องร้อยละ 91.52

**ตารางที่ 8:** ผลลัพธ์ของตัวแบบ BERT

| Parameter | | Performance | | | | |
|---|---|---|---|---|---|---|
| Dense Unit | Epoch | A(%) | P | R | F | ROC |
| **0.00003** | **4** | **93.30** | **0.93** | **0.93** | **0.93** | **0.98** |
| 0.00002 | 4 | 92.82 | 0.93 | 0.93 | 0.93 | 0.98 |
| 0.00002 | 5 | 92.82 | 0.93 | 0.93 | 0.93 | 0.98 |
| 0.00002 | 3 | 92.39 | 0.92 | 0.92 | 0.92 | 0.98 |
| 0.00003 | 3 | 92.35 | 0.92 | 0.92 | 0.92 | 0.98 |
| 0.00003 | 5 | 92.22 | 0.92 | 0.92 | 0.92 | 0.98 |
| 0.00005 | 4 | 87.94 | 0.90 | 0.88 | 0.88 | 0.96 |
| 0.00005 | 5 | 84.51 | 0.86 | 0.85 | 0.84 | 0.94 |
| 0.00005 | 3 | 82.49 | 0.82 | 0.82 | 0.81 | 0.90 |

จากตารางที่ 8 ตัวแบบ BERT ที่ปรับค่าพารามิเตอร์ Bi-BERT Unit เท่ากับ 128 ค่า Dense Unit เท่ากับ 0.00003 และค่า Epoch เท่ากับ 4 ให้ประสิทธิภาพดีที่สุดด้วยค่าความถูกต้องร้อยละ 93.30

เมื่อได้ผลลัพธ์ของ 4 แบบจำลอง ทำการเปรียบเทียบประสิทธิภาพของแต่ละแบบจำลอง แสดงดังตารางที่ 9

**ตารางที่ 9:** ผลการเปรียบเทียบประสิทธิภาพของแบบจำลอง

| Model | Performance | | | | |
|---|---|---|---|---|---|
| | A(%) | P | R | F | ROC |
| Bi-LSTM | 91.92 | 0.92 | 0.92 | 0.92 | 0.98 |
| Bi-GRU | 91.52 | 0.92 | 0.92 | 0.92 | 0.98 |
| BERT | **93.30** | **0.93** | **0.93** | **0.93** | **0.98** |

จากตารางที่ 10 สรุปแบบจำลอง BERT ให้ประสิทธิภาพดีที่สุดด้วยค่าความถูกต้องร้อยละ 93.30 ค่าความแม่นยำ 0.93 ค่าระลึก 0.93 ค่าการถ่วงดุล 0.93 และค่าพื้นที่ใต้เส้นโค้งคุณลักษณะสมบัติการทำงาน เท่ากับ 0.98 ตามลำดับ

จากนั้นได้ทำการทดสอบประสิทธิภาพแบบจำลองทั้ง 3 แบบจำลองด้วย Unseen Data ชุดเดียวกัน จำนวน 100 รายการ แสดงดังตารางที่ 10

**ตารางที่ 10:** ผลการทดสอบประสิทธิภาพแบบจำลองด้วย Unseen Data

| Model | Performance | | | | |
|---|---|---|---|---|---|
| | A(%) | P | R | F | ROC |
| Bi-LSTM | 84.84 | 0.84 | 0.85 | 0.84 | 0.93 |
| Bi-GRU | **86.87** | **0.87** | **0.87** | **0.87** | **0.98** |
| BERT | 78.78 | 0.79 | 0.79 | 0.79 | 0.88 |

จากตารางที่ 10 ผลการทดสอบตัวด้วย Unseen Data ชุดเดียวกัน จำนวน 100 รายการ พบว่าตัวแบบ Bi-GRU ให้ประสิทธิภาพดีที่สุดด้วยค่าความถูกต้องร้อยละ 86.87 ค่าความแม่นยำเฉลี่ย เท่ากับ 0.87 ค่าระลึกเฉลี่ย เท่ากับ 0.87 ค่าการถ่วงดุลเฉลี่ย เท่ากับ 0.87 และค่าพื้นที่ใต้เส้นโค้งคุณลักษณะสมบัติการทำงาน เท่ากับ 0.98

## 5. สรุป

งานวิจัยฉบับนี้ได้นำเสนอการวิเคราะห์การจำแนกเจตนาของข้อความภาษาไทยในสื่อสังคมออนไลน์ โดยเฉพาะในบริบทของอุตสาหกรรมเครื่องดื่มโดยรวบรวมข้อมูลในช่วงระหว่างเดือนมกราคม พ.ศ.2568 ถึงเดือนตุลาคม พ.ศ.2568 จำนวน 2,300 รายการ แบ่งเป็น 3 ประเภท ได้แก่ ความคิดเห็นเชิงบอกเล่า ความคิดเห็นเชิงคำถาม และ การแจ้งปัญหา หลังจากนั้นนำ ข้อมูลดังกล่าวเข้าสู่กระบวนการจัดเตรียมข้อมูล ใช้เทคนิค Class Weight เพื่อปรับสมดุลข้อมูล ใช้วิธี Stratified K-Fold Cross-Validation โดยกำหนดค่า (k = 5) เพื่อแบ่งชุดข้อมูลออกเป็น 5 ส่วนย่อย จากนั้นได้พัฒนาแบบจำลองโดยใช้เทคนิคการเรียนรู้เชิงลึก 2 เทคนิค คือ Bi-LSTM, Bi-GRU และเทคนิคทรานส์ฟอร์เมอร์ คือ BERT พร้อมทั้งปรับค่าพารามิเตอร์ของแบบจำลอง ทำการประเมินประสิทธิภาพของแต่ละแบบจำลองใช้โดยพิจารณาจากค่าความถูกต้อง ค่าความแม่นยำ ค่าระลึก แค่การถ่วงดุลผล และค่าพื้นที่ใต้เส้นโค้งคุณลักษณะสมบัติการทำงานการการวิจัยพบว่า แบบจำลอง BERT ให้ประสิทธิภาพดีที่สุดด้วยค่าความถูกต้องร้อยละ 93.30 ค่าความแม่นยำ 0.93 ค่าการระลึก 0.93 ค่าการถ่วงดุล 0.93 และค่าพื้นที่ใต้เส้นโค้งคุณลักษณะสมบัติการทำงานเท่ากับ 0.98 ตามลำดับ จากนั้นทดสอบประสิทธิภาพแบบจำลองด้วย Unseen Data พบว่า แบบ Bi-GRU ให้ประสิทธิภาพดีที่สุดด้วยค่าความถูกต้องร้อยละ 86.87 ค่าความแม่นยำ ค่าระลึก และค่าการถ่วงดุล เท่ากับ 0.87 ส่วนค่าพื้นที่ใต้เส้นโค้งคุณลักษณะสมบัติการทำงานเท่ากับ 0.98 จะเห็นได้ว่า ในการสร้างแบบจำลอง BERT ให้ประสิทธิภาพดีที่สุด แต่การทดสอบด้วย Unseen Data แบบจำลอง Bi-GRU ให้ประสิทธิภาพที่ดีที่สุด ซึ่งอาจเกิดจากการที่แบบจำลอง Bi-GRU ไม่เกิดปัญหา Overfitting เท่ากับ BERT ในการทดสอบ สำหรับแนวทางแก้ปัญหา สามารถพิจารณาเพิ่มขนาดชุดข้อมูลให้มีความหลากหลายและครอบคลุมมากขึ้น เพื่อให้สามารถฝึกฝนแบบจำลองได้ดียิ่งขึ้น ลดปัญหา Overfitting และเพิ่มความสามารถในการทดสอบให้มีความแม่นยำมากยิ่งขึ้น

## เอกสารอ้างอิง


[1] ชาลิสา จิตบุญญาพินิจ, ปราศรัย มณีรัตน์ และนิเวศ จิระวิชิตชัย. (2565). การพัฒนาแบบจำลองการวิเคราะห์ความรู้สึกบนสื่อสังคมออนไลน์ไทยโดยใช้เทคนิคการเรียนรู้เชิงลึก. วารสารวิทยาศาสตร์และเทคโนโลยีหัวเฉียวเฉลิมพระเกียรติ, 8(2), 8-18.

[2] IBM, What Is Text Mining?, [ออนไลน์]. เข้าถึงได้จาก: https://www.ibm.com/think/topics/text-mining. [เข้าถึงเมื่อ: 12 ธันวาคม 2567].

[3] GeeksforGeeks, Bidirectional LSTM in NLP, [ออนไลน์]. เข้าถึงได้จาก: https://www.geeksforgeeks.org/bidirectional-lstm-in-nlp/. [เข้าถึงเมื่อ: 20 กุมภาพันธ์ 2567].

[4] Vtiya, GRU vs Bi-GRU: Which One is Going to Win?, Medium, [ออนไลน์]. เข้าถึงได้จาก: https://vtiya.medium.com/gru-vs-bi-gru-which-one-is-going-to-win-58a45ede5fba. [เข้าถึงเมื่อ: 12 ธันวาคม 2567].

[5] Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019).BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding.Proceedings of NAACL-HLT 2019, 4171–4186. https://aclanthology.org/N19-1423.pdf

[6] Derbentsev, V. D., Bezkorovainyi, V. S., Matviychuk, A. V., Pomazun, O. M., Hrabariev, A. V., & Hostryk, A. M. (2023).A comparative study of deep learning models for sentiment analysis of social media texts.CEUR Workshop Proceedings, 3485, 168–188.https://ceur-ws.org/Vol-3485/paper18.pdf

[7] Chaudhary, L., Girdhar, N., Sharma, D., Andreu-Perez, J., Doucet, A., & Renz, M. (2023). A review of deep learning models for Twitter sentiment analysis: Challenges and opportunities.IEEE Transactions on Computational Social Systems, 11(3), 3550–3579. https://doi.org/10.1109/TCSS.2023.3322002

[8] Nugroho, K. S., Sukmadewa, A. Y., & Yudistira, N. (2021).Large-scale news classification using BERT language model: Spark NLP approach.arXiv. https://arxiv.org/abs/2107.06785

# การวิเคราะห์ความรู้สึกของผู้ใช้แอปพลิเคชันธนาคารบนสื่อสังคมออนไลน์ด้วยแบบจำลองภาษาขนาดใหญ่

# Sentiment Analysis of Banking Applications Users on Social Media Using Large Language Models

*สิตาวีญ์ ศิริฐิติวิทยา (Sitawee Sirithitiwithaya)[1] และผุสดี บุญรอด (Pudsadee Boonrawd)[2]*
*ภาควิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล*
*มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ*
*[1]s6707011857098@email.kmutnb.ac.th, [2]pudsadee.b@itd.kmutnb.ac.th*

## บทคัดย่อ

*ในยุคเศรษฐกิจดิจิทัลที่การทำธุรกรรมผ่านแอปพลิเคชันธนาคารกลายเป็นส่วนหนึ่งของวิถีชีวิต การรับรู้ความคิดเห็นของผู้ใช้งานจึงเป็นสิ่งสำคัญในการพัฒนาคุณภาพและบริการ งานวิจัยนี้จึงมีวัตถุประสงค์เพื่อพัฒนาแบบจำลองสำหรับจำแนกความรู้สึกของผู้ใช้บริการแอปพลิเคชันธนาคาร และพัฒนาระบบสารสนเทศที่นำเสนอข้อมูลในรูปแบบแดชบอร์ด เพื่อสนับสนุนการตัดสินใจของสถาบันการเงิน งานวิจัยทำการเก็บรวบรวมข้อมูลจากแพลตฟอร์มเอ็กซ์ และพันทิปจำนวน 14,246 รายการ ในการพัฒนาแบบจำลองประยุกต์ใช้ WangchanBERTa และ PhayaThaiBERT ผลวิจัยพบว่า PhayaThaiBERT มีความแม่นยำในการจำแนกความรู้สึกสูงกว่า WangchanBERTa โดยมีค่าคะแนนเอฟหนึ่งอยู่ที่ร้อยละ 93.85 และ 91.82 ตามลำดับ นอกจากนี้ ระบบสารสนเทศยังนำเสนอข้อมูลต่าง ๆ ที่แสดงให้เห็นถึงความคิดเห็นของผู้ใช้งานที่เป็นเชิงลบมากกว่าเชิงบวก โดยประเด็นที่พบส่วนใหญ่มาจากความเสถียรของระบบและการเข้าใช้งานระบบ ซึ่งสถาบันการเงินสามารถนำไปต่อยอด และใช้เป็นแนวทางสำคัญในการวางแผนพัฒนาบริการเพื่อสร้างความพึงพอใจให้แก่ผู้ใช้งานต่อไป*

**คำสำคัญ:** แอปพลิเคชันธนาคาร การวิเคราะห์ความรู้สึก แบบจำลองภาษาขนาดใหญ่

## Abstract

*As banking apps become a part of daily life, understanding user feedback is key to improving service. This research develops a sentiment classification model and a decision-support dashboard designed for financial institutions. Utilizing 14,246 posts from X Platform and Pantip, the study evaluates the performance of WangchanBERTa and PhayaThaiBERT. Findings reveal that PhayaThaiBERT achieved higher sentiment classification accuracy than WangchanBERTa, with F1-score of 93.85% and 91.82%, respectively. The integrated dashboard highlights a predominance of negative sentiment, primarily centered on system stability and accessibility. These insights provide a strategic roadmap for banks to enhance system reliability and elevate overall user satisfaction.*

**Keywords:** Banking Applications, Sentiment Analysis, Large Language Models.

## 1. บทนำ

แอปพลิเคชันธนาคาร (Banking Application) เป็นโครงสร้างพื้นฐานสำคัญในเศรษฐกิจดิจิทัลและกลายเป็นช่องทางหลักสำหรับธุรกรรมทางการเงินในปัจจุบัน [1] ภายใต้สภาวะการแข่งขันในอุตสาหกรรมการเงิน ประสบการณ์และความพึงพอใจของผู้ใช้งาน

(User Experience) ถือเป็นสิ่งที่ส่งผลต่อความน่าเชื่อถือและความสามารถในการรักษาฐานลูกค้าขององค์กร อย่างไรก็ตาม การสำรวจความพึงพอใจของผู้ใช้งานด้วยวิธีดั้งเดิม เช่น การใช้แบบสอบถามหรือการสัมภาษณ์ มีข้อจำกัดด้านความล่าช้าในการเก็บรวบรวมข้อมูล และข้อมูลที่ได้มักมีความคลาดเคลื่อนจากอคติในการตอบ ซึ่งไม่สามารถสะท้อนความรู้สึกที่แท้จริงของผู้ใช้งานได้ในขณะนั้น [2]

ขณะที่สื่อสังคมออนไลน์ (Social Media) โดยเฉพาะแพลตฟอร์มพันทิป (Pantip) และเอ็กซ์ (X) เป็นแหล่งข้อมูลที่รวบรวมความคิดเห็น และข้อเสนอแนะจากผู้ใช้งานจริง [3] หากนำข้อมูลเหล่านี้มาวิเคราะห์ จะช่วยให้สถาบันการเงินมองเห็นทั้งจุดแข็งที่ควรส่งเสริมและจุดอ่อนที่ต้องเร่งแก้ไขสำหรับการพัฒนาบริการให้ตรงตามความต้องการของผู้ใช้งานจริง งานวิจัยนี้จึงมุ่งพัฒนาแบบจำลองภาษาขนาดใหญ่สำหรับจำแนกความรู้สึกของผู้ใช้งานแอปพลิเคชันธนาคาร และพัฒนาระบบสารสนเทศที่นำเสนอข้อมูลในรูปแบบแดชบอร์ดสำหรับใช้ประกอบการตัดสินใจเชิงกลยุทธ์ของสถาบันการเงินในการพัฒนาคุณภาพบริการให้สอดคล้องกับความต้องการที่แท้จริงของผู้ใช้งาน

## 2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

### 2.1 การวิเคราะห์ความรู้สึก

การวิเคราะห์ความรู้สึก (Sentiment Analysis) เป็นการจำแนกและสกัดข้อมูลเชิงอารมณ์จากข้อความ เพื่อระบุความรู้สึกของผู้เขียนข้อความว่าเป็นเชิงบวก (Positive) เชิงลบ (Negative) หรือเป็นกลาง (Neutral) โดยเฉพาะในบริบทของสื่อสังคมออนไลน์ ซึ่งมีข้อมูลเป็นจำนวนมากและไม่มีโครงสร้างชัดเจน การวิเคราะห์ความรู้สึกจึงเป็นเครื่องมือสำคัญในการเปลี่ยนความคิดเห็นสาธารณะให้เป็นข้อมูลเชิงลึกเพื่อใช้สนับสนุนการตัดสินใจ [4] ซึ่งสามารถแบ่งแนวทางออกเป็น 3 รูปแบบ ดังนี้

2.1.1 แนวทางอิงตามกฎ

แนวทางอิงตามกฎเป็นวิธีพื้นฐานที่ใช้พจนานุกรมคำศัพท์ (Lexicon-based) และกฎทางไวยากรณ์ในการคำนวณคะแนนความรู้สึก โดยคำที่มีความหมายเชิงบวกจะได้ค่าคะแนน +1 ส่วนคำที่มีความหมายเชิงลบจะได้ค่าคะแนน -1 แม้แนวทางนี้จะเข้าใจง่ายแต่ก็มีข้อจำกัดในการจัดการกับภาษาที่มีบริบทประชดประชัน หรือคำศัพท์ใหม่ที่เกิดขึ้นบนสื่อสังคมออนไลน์ [5]

2.1.2 แนวทางอิงการเรียนรู้ของเครื่อง

แนวทางอิงการเรียนรู้ของเครื่องเป็นการฝึกสอนแบบจำลองด้วยชุดข้อมูลที่มีการติดป้ายกำกับ (Labeled Data) โดยใช้อัลกอริทึม เช่น เครื่องจักรเวกเตอร์สนับสนุน (Support Vector Machines: SVM) แม้จะมีความแม่นยำสูงกว่าแนวทางอิงตามกฎ แต่ยังพบปัญหาในการทำความเข้าใจความสัมพันธ์ของคำในรูปประโยคที่ซับซ้อน [5]

2.1.3 แนวทางอิงการเรียนรู้เชิงลึก

แนวทางอิงการเรียนรู้เชิงลึกมีการประยุกต์ใช้โครงข่ายประสาทเทียม เพื่อเรียนรู้การแทนค่าของคำ (Word Representation) โดยอัตโนมัติ ซึ่งต่อมาได้พัฒนาเป็นสถาปัตยกรรมทรานส์ฟอร์เมอร์ (Transformer) ที่มีความโดดเด่นในการจับบริบทระยะไกล (Long-range Dependency) และเข้าใจความหมายแฝงในภาษาไทยได้ดีกว่าทุกแนวทางข้างต้น [5]

งานวิจัยนี้จึงเลือกใช้แนวทางอิงการเรียนรู้เชิงลึก เนื่องจากมีความสามารถในการทำความเข้าใจบริบทของภาษาที่มีความซับซ้อนได้ดี โดยเฉพาะข้อมูลจากสื่อสังคมออนไลน์ที่มีการใช้คำสแลง ทำให้การจำแนกความรู้สึกของผู้ใช้งานแอปพลิเคชันธนาคารมีความแม่นยำมากขึ้น

### 2.2 แบบจำลองภาษาขนาดใหญ่

ในปัจจุบันสถาปัตยกรรมทรานส์ฟอร์เมอร์ถือเป็นหัวใจสำคัญของแบบจำลองภาษาขนาดใหญ่ โดยมีกลไกความใส่ใจ (Attention Mechanism) ที่ช่วยให้แบบจำลองสามารถประมวลผลข้อความได้แบบขนานและเข้าใจความสัมพันธ์ระหว่างคำในทุกตำแหน่งของประโยคพร้อมกัน งานวิจัยนี้จึงให้ความสำคัญกับแบบจำลอง 2 ประเภทที่มีประสิทธิภาพสูงในภาษาไทย ได้แก่

2.2.1 แบบจำลอง WangchanBERTa

WangchanBERTa เป็นแบบจำลองภาษาไทยที่ผ่านการฝึกสอนล่วงหน้า (Pre-training) ด้วยคลังข้อมูลขนาด

ใหญ่จากวิกิพีเดียและสื่อออนไลน์ และมีกลไกการเรียนรู้บริบทแบบสองทิศทาง ทำให้มีความเชี่ยวชาญในการจัดการกับความกำกวมของภาษาไทยได้เป็นอย่างดี [6] อย่างไรก็ตาม แบบจำลองยังมีข้อจำกัดในการประมวลผลคำทับศัพท์ภาษาต่างประเทศ ซึ่งมักปรากฏบ่อยในข้อความบนสื่อสังคมออนไลน์

2.2.2 แบบจำลอง PhayaThaiBERT

PhayaThaiBERT ถูกพัฒนาขึ้นเพื่อแก้ข้อจำกัดด้านคลังคำศัพท์ โดยการขยายคลังคำศัพท์จากแบบจำลองหลายภาษาอย่าง XLM-RoBERTa ให้ครอบคลุมภาษาไทยมากขึ้น ส่งผลให้มีความสามารถในการประมวลผลข้อความแบบผสมและคำทับศัพท์ต่างประเทศ ซึ่งสอดคล้องกับลักษณะการสื่อสารบนสื่อสังคมออนไลน์ที่ผู้ใช้งานมักใช้คำเฉพาะทางหรือคำภาษาอังกฤษปนกับภาษาไทย [7]

การนำแบบจำลองที่มีความสามารถในการประมวลผลข้อความภาษาไทยมาประยุกต์ใช้ในงานวิจัยนี้ จึงเป็นแนวทางสำคัญที่จะช่วยเพิ่มความแม่นยำในการวิเคราะห์ข้อมูลจากสื่อสังคมออนไลน์ที่มีความซับซ้อนและมีการเปลี่ยนแปลงทางภาษาอยู่ตลอดเวลา เพื่อให้ได้ผลลัพธ์การจำแนกความรู้สึกที่สะท้อนถึงมุมมองของผู้ใช้แอปพลิเคชันธนาคารได้อย่างมีประสิทธิภาพ

**2.3 งานวิจัยที่เกี่ยวข้อง**

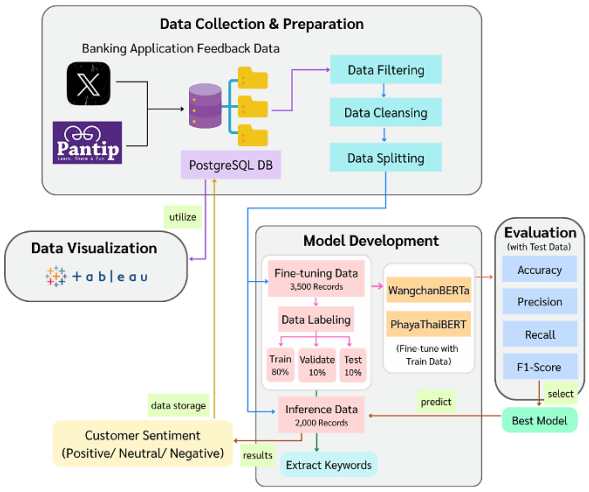
จากการทบทวนงานวิจัย พบว่ามีการพัฒนาแนวทางสำหรับวิเคราะห์ความรู้สึกอย่างต่อเนื่อง โดยในระยะแรกงานวิจัยส่วนใหญ่มักประยุกต์ใช้การเรียนรู้ของเครื่อง เช่น เค-เพื่อนบ้านใกล้ที่สุด (K-Nearest Neighbors: K-NN) ป่าสุ่ม (Random Forest) การถดถอยแบบโลจิสติก และเครื่องจักรเวกเตอร์สนับสนุนในการจำแนกความรู้สึก [8] รวมถึงการใช้โครงข่ายประสาทเทียมกับเบย์อย่างง่าย (Naïve Bayes) เพื่อวิเคราะห์ความคิดเห็นของผู้ใช้แอปพลิเคชันธนาคาร [9] แต่ยังพบข้อจำกัดหากข้อความมีความยาวหรือมีความกำกวมของภาษา

งานวิจัยที่พบในปัจจุบันนี้จึงใช้แบบจำลองภาษาขนาดใหญ่บนสถาปัตยกรรมทรานฟอร์เมอร์เพื่อแก้ไขข้อจำกัดดังกล่าว โดยเฉพาะงานวิจัยที่นำแบบจำลอง WangchanBERTa และ PhayaThaiBERT มาปรับจูนกับชุดข้อมูลเฉพาะทาง พบว่ามีความแม่นยำในการวิเคราะห์ความรู้สึกและจำแนกประเภทข้อความภาษาไทยได้ดีกว่าแบบจำลองทั่วไป [6], [7], [10] เนื่องจากแบบจำลองเหล่านี้ได้รับการเรียนรู้ลักษณะเฉพาะของภาษาไทยแล้ว

อย่างไรก็ตาม ยังมีช่องว่างของงานวิจัยในด้านการนำผลการวิเคราะห์มาประมวลผลเป็นระบบนำเสนอข้อมูลส่งผลให้สถาบันการเงินขาดเครื่องมือในการระบุปัญหาและแนวโน้มความพึงพอใจได้อย่างทันท่วงที งานวิจัยนี้จึงมุ่งพัฒนาระบบสารสนเทศที่ช่วยให้สถาบันการเงินสามารถรับรู้สถานการณ์ที่เกิดขึ้นได้ในทันที

## 3. วิธีการดำเนินการวิจัย

งานวิจัยนี้ดำเนินตามกรอบแนวคิดดังภาพที่ 1 ซึ่งครอบคลุมทุกกระบวนการ โดยแบ่งออกเป็น 4 ขั้นตอน ดังนี้




**ภาพที่ 1:** กรอบแนวคิดการดำเนินงานวิจัย

**3.1 การรวบรวมและเตรียมข้อมูล**

คณะผู้วิจัยรวบรวมความคิดเห็นของผู้ใช้แอปพลิเคชันธนาคารจากแพลตฟอร์มพันทิป ด้วยวิธีการดึงข้อมูลจากเว็บไซต์ (Web Scraping) และเอ็กซ์ ผ่านส่วนต่อประสานโปรแกรมประยุกต์ (Twitter API) ภายใต้ข้อกำหนดและเงื่อนไขการใช้งานของแต่ละแพลตฟอร์ม โดยไม่มีข้อมูลที่ระบุตัวบุคคล ทั้งนี้ ข้อมูลที่รวบรวมในช่วงเดือนมกราคม พ.ศ. 2565 ถึง ธันวาคม พ.ศ. 2568 มีจำนวนทั้งสิ้น 14,246 รายการ ซึ่งจัดเก็บในโครงสร้างฐานข้อมูลตามตารางที่ 1

**ตารางที่ 1:** โครงสร้างฐานข้อมูล

| คุณลักษณะ | คำอธิบาย | ชนิดข้อมูล |
|---|---|---|
| Platform | แพลตฟอร์มต้นทาง | Varchar |
| Text | ข้อความความคิดเห็น | Text |
| Sentiment | ป้ายกำกับความรู้สึก | Varchar |
| Category | ป้ายกำกับประเด็น | Varchar |

ทั้งนี้ ข้อมูลที่ได้จากสื่อสังคมออนไลน์มีลักษณะเป็นข้อความไม่มีโครงสร้างและมีข้อความที่ไม่เกี่ยวข้อง (Noise) คณะผู้วิจัยจึงดำเนินการเตรียมข้อมูลดังนี้

3.1.1 การคัดกรองข้อมูลเชิงเนื้อหา

ข้อมูลในฐานข้อมูล PostgreSQL ถูกคัดกรองโดยใช้ภาษา SQL เพื่อคัดเลือกเฉพาะข้อความที่เกี่ยวข้องกับแอปพลิเคชันธนาคารผ่านคำสำคัญ (Keywords) เช่น แอปธนาคาร โมบายแบงก์กิ้ง หรือชื่อธนาคาร

3.1.2 การปรับระดับหน่วยข้อมูล

ในกรณีที่ข้อความหนึ่งมีหลายความรู้สึก ระบุถึงหลายประเด็นหรือหลายธนาคาร คณะผู้วิจัยจะดำเนินการแยกแถวข้อมูลเหล่านั้น เพื่อให้ 1 แถวข้อมูลแทน 1 หน่วยความรู้สึก ทำให้การวิเคราะห์มีความแม่นยำมากขึ้น

3.1.3 การทำความสะอาดข้อมูล

การทำความสะอาดข้อมูลโดยคัดกรองส่วนที่ไม่ส่งผลกับการวิเคราะห์ความรู้สึก รวมถึงเครื่องหมายวรรคตอน พร้อมกำหนดความยาวข้อความสูงสุด (Maximum Sequence Length) ไว้ที่ 256 โทเคน เพื่อรักษาบริบทสำคัญในการวิเคราะห์ความรู้สึก ลดการสูญเสียข้อมูลจากการตัดข้อความ และรองรับกระบวนการตัดคำของแบบจำลองแต่ละประเภทได้อย่างมีประสิทธิภาพ

จากการคัดกรองและทำความสะอาดข้อมูลในข้างต้น ทำให้ข้อมูลอยู่ในรูปแบบพร้อมใช้งานจำนวนทั้งสิ้น 12,459 รายการ

### 3.2 การพัฒนาและปรับจูนแบบจำลอง

งานวิจัยนี้ได้ดำเนินการปรับจูนแบบจำลอง (Fine-tuning) โดยมีกระบวนการดังนี้

3.2.1 การเตรียมชุดข้อมูลสำหรับฝึกสอนแบบจำลอง

คณะผู้วิจัยได้สุ่มเลือกข้อมูลจากชุดข้อมูลที่ผ่านการทำความสะอาดแล้วจำนวน 3,500 รายการ เพื่อนำมาเป็นชุดข้อมูลตัวอย่าง โดยใช้เครื่องมือ Label Studio ในการติดป้ายกำกับความรู้สึกทั้งหมด 3 ประเภท ได้แก่

3.2.1.1 เชิงบวก (Positive) เป็นข้อความที่แสดงความพึงพอใจ หรือความประทับใจจากการใช้งาน

3.2.1.2 เป็นกลาง (Neutral) เป็นข้อความสอบถามข้อมูลทั่วไป หรือข้อเสนอแนะที่ปราศจากอารมณ์

3.2.1.3 เชิงลบ (Negative) เป็นข้อความที่แสดงความไม่พึงพอใจ หรือระบุถึงปัญหาจากการใช้งาน

3.2.2 ขั้นตอนการปรับจูนแบบจำลอง

งานวิจัยนี้เลือกใช้การเรียนรู้แบบมีผู้สอน (Supervised Learning) โดยนำแบบจำลอง WangchanBERTa และ PhayaThaiBERT มาทำการปรับจูนเพื่อให้แบบจำลองเรียนรู้ลักษณะเฉพาะของภาษาในบริบทแอปพลิเคชันธนาคารได้ โดยแบ่งชุดข้อมูลออกเป็น ชุดฝึกสอนร้อยละ 80 ชุดตรวจสอบร้อยละ 10 และชุดทดสอบร้อยละ 10 และเลือกใช้การตัดคำตามโครงสร้างเดิมของแต่ละแบบจำลอง

3.2.3 การกำหนดค่าไฮเปอร์พารามิเตอร์

คณะผู้วิจัยได้กำหนดค่าไฮเปอร์พารามิเตอร์ (Hyperparameter) และใช้การหยุดก่อนกำหนด (Early Stopping) เพื่อให้การประมวลผลมีประสิทธิภาพและป้องกันปัญหาการเรียนรู้เกิน (Overfitting) ดังแสดงในตารางที่ 2

**ตารางที่ 2:** การกำหนดค่าไฮเปอร์พารามิเตอร์

| ไฮเปอร์พารามิเตอร์ | ค่าที่กำหนด |
|---|---|
| Learning Rate (LR) | $2 \times 10^{-5}$ |
| Batch Size | 16 |
| Maximum Epochs | 10 |
| Patience | 3 |

### 3.3 การประเมินประสิทธิภาพของแบบจำลอง

คณะผู้วิจัยได้เลือกใช้ตัวชี้วัดมาตรฐานสำหรับประเมินประสิทธิภาพของแบบจำลองดังนี้

3.3.1 ค่าความถูกต้อง

ค่าความถูกต้อง (Accuracy) ใช้วัดความสามารถในการทำนายความรู้สึกได้ถูกต้องเมื่อเทียบกับข้อมูลทั้งหมด

3.3.2 ค่าความแม่นยำ

ค่าความแม่นยำ (Precision) ใช้ประเมินความถูกต้องของการทำนายในแต่ละคลาส

3.3.3 ค่าการเรียกคืน

ค่าการเรียกคืน (Recall) ใช้ระบุความสามารถในการตรวจจับความรู้สึกแต่ละประเภทได้อย่างครบถ้วน

3.3.4 ค่าคะแนนเอฟหนึ่ง

ค่าคะแนนเอฟหนึ่ง (F1-Score) เป็นค่าเฉลี่ยถ่วงน้ำหนักระหว่างค่าความแม่นยำและค่าการเรียกคืน

**3.4 การพัฒนาระบบนำเสนอข้อมูล**

ผลลัพธ์การจำแนกความรู้สึกที่ได้จะถูกจัดเก็บลงในฐานข้อมูล PostgreSQL เพื่อนำไปสร้างระบบนำเสนอข้อมูลในรูปแบบแดชบอร์ดด้วยโปรแกรม Tableau เพื่อแสดงผลข้อมูลที่เป็นความคิดเห็นของผู้ใช้งานในรูปแบบแผนภาพ (Visualization) ที่เข้าใจง่าย

ระบบดังกล่าวถูกออกแบบมาเพื่อสนับสนุนการตัดสินใจเชิงกลยุทธ์ ช่วยให้สถาบันการเงินรับรู้ความคิดเห็น ประเด็นปัญหา และความต้องการที่แท้จริงของผู้ใช้งานแอปพลิเคชันธนาคาร ซึ่งจะเป็นแนวทางสำคัญในการยกระดับคุณภาพบริการอย่างมีประสิทธิภาพ

## 4. ผลดำเนินการวิจัย

**4.1 ผลการประเมินประสิทธิภาพของแบบจำลอง**

ผลการประเมินประสิทธิภาพของแบบจำลองสำหรับจำแนกความรู้สึก สามารถแสดงได้ดังตารางที่ 3

**ตารางที่ 3:** ผลการประเมินประสิทธิภาพของแบบจำลอง

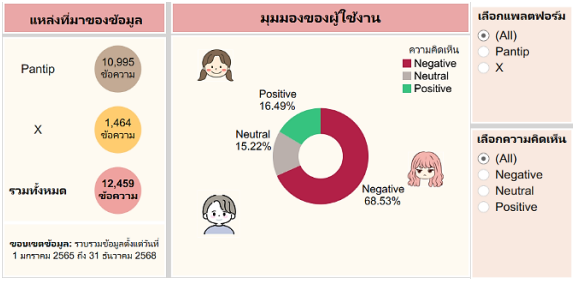
| Model | Accuracy (%) | Precision (%) | Recall (%) | F-1 score (%) |
|---|---|---|---|---|
| **PhayaThai BERT** | **94.29** | **94.10** | **94.29** | **93.85** |
| Wangchan BERTa | 92.86 | 93.37 | 92.86 | 91.82 |

จากตารางที่ 3 พบว่าแบบจำลอง PhayaThaiBERT เป็นแบบจำลองภาษาไทยที่มีประสิทธิภาพในการจำแนกความรู้สึกของผู้ใช้งาน โดยมีค่าคะแนนเอฟหนึ่งสูงถึงร้อยละ 93.85 และ WangchanBERTa ได้ค่าคะแนนเอฟหนึ่งที่ร้อยละ 91.82 อย่างไรก็ตาม แบบจำลองที่มีสถาปัตยกรรมทรานส์ฟอร์เมอร์ล้วนให้ผลลัพธ์ในทิศทางเดียวกัน ซึ่งสะท้อนถึงความสามารถในการจัดการข้อความภาษาไทยบนสื่อสังคมออนไลน์ได้ดี

**4.2 ผลการวิเคราะห์ข้อมูลและการนำเสนอข้อมูลผ่านแดชบอร์ด**

4.2.1 ข้อมูลสถิติและภาพรวมความรู้สึกของผู้ใช้งาน

ผลการวิเคราะห์ความรู้สึกในภาพที่ 2 แสดงให้เห็นถึงความคิดเห็นของผู้ใช้แอปพลิเคชันธนาคารจากแพลตฟอร์มเอ็กซ์และพันทิป โดยมีสัดส่วนเป็นเชิงลบสูงถึงร้อยละ 68.53 สะท้อนถึงปัญหาการใช้งานที่ควรเร่งปรับปรุงเพื่อเพิ่มความพึงพอใจของผู้ใช้งาน




**ภาพที่ 2:** ข้อมูลสถิติและภาพรวมความรู้สึกของผู้ใช้งาน

4.2.2 การวิเคราะห์ความคิดเห็นของผู้ใช้งานผ่านกลุ่มคำก้อนเมฆ

การวิเคราะห์ความคิดเห็นผู้ใช้แอปพลิเคชันธนาคารผ่านกลุ่มคำก้อนเมฆดังแสดงในภาพที่ 3 พบคำว่า ล่ม และ เข้าใช้งาน ปรากฏขึ้นอย่างหนาแน่น สอดคล้องกับข้อมูลที่พบว่าปัญหาด้านความเสถียรของระบบเป็นอุปสรรคในการเข้าถึงบริการ แม้ภาพรวมเป็นเชิงลบ แต่ยังพบข้อความที่ได้รับในเชิงบวก เช่น UX/UI และการทำธุรกรรม แสดงให้เห็นว่าหากระบบมีความเสถียร ผู้ใช้งานยังคงมีความพึงพอใจต่อหน้าตาแอปพลิเคชันและการทำธุรกรรม




**ภาพที่ 3:** กลุ่มคำก้อนเมฆแสดงคำสำคัญที่ถูกพูดถึง

4.2.3 การจัดอันดับคำสำคัญที่ผู้ใช้งานระบุ

การวิเคราะห์เชิงปริมาณจาก 10 อันดับคำสำคัญที่ถูกระบุมากที่สุด โดยพิจารณาจากผลรวมของจำนวนความคิดเห็นทั้งหมด ดังแสดงในภาพที่ 4 พบว่าข้อมูลมีความสอดคล้องกับกลุ่มคำก้อนเมฆ โดยคำที่มีการระบุมากที่สุด 3 อันดับแรก ได้แก่ ชำระเงิน เข้าใช้งาน และล่ม ตามลำดับ

**10 อันดับคำสำคัญที่ถูกระบุมากที่สุด**

| ลำดับ | คีย์เวิร์ด | Negative | Neutral | Positive | ผลรวม |
|---|---|---|---|---|---|
| 1 | ชำระเงิน | 664 | 438 | 646 | **1,748** |
| 2 | เข้าใช้งาน | 1,299 | 38 | 13 | **1,350** |
| 3 | ล่ม | 1,260 | 29 | 30 | **1,318** |
| 4 | ความปลอดภัย | 531 | 365 | 225 | **1,121** |
| 5 | ระบบช้า | 881 | 16 | 11 | **908** |
| 6 | UX/UI | 377 | 208 | 303 | **887** |
| 7 | ให้คำแนะนำ | 632 | 74 | 97 | **803** |
| 8 | error | 773 | 19 | 3 | **795** |
| 9 | ค้าง | 671 | 13 | 6 | **690** |
| 10 | การทำธุรกรรม | 193 | 139 | 358 | **688** |

**ภาพที่ 4:** สถิติ 10 อันดับคำสำคัญที่ถูกระบุมากที่สุด

## 5. สรุปผลการวิจัยและข้อเสนอแนะ

งานวิจัยนี้ได้พัฒนาแบบจำลองภาษาขนาดใหญ่สำหรับวิเคราะห์ความรู้สึกของผู้ใช้แอปพลิเคชันธนาคารผ่านการรวบรวมข้อมูลจากสื่อสังคมออนไลน์อย่างเอ็กซ์และพันทิป โดยเปรียบเทียบประสิทธิภาพของแบบจำลอง WangchanBERTa และ PhayaThaiBERT ที่มีการปรับจูนพารามิเตอร์และประเมินผลด้วยตัวชี้วัดมาตรฐาน ผลการวิจัยพบว่า PhayaThaiBERT มีประสิทธิภาพในการจำแนกความรู้สึกด้วยค่าคะแนนเอฟหนึ่งร้อยละ 93.85 ซึ่งสูงกว่า WangchanBERTa ที่มีค่าคะแนนเอฟหนึ่งร้อยละ 91.82 แสดงให้เห็นว่า PhayaThaiBERT สามารถทำความเข้าใจบริบทภาษาไทยที่เป็นคำศัพท์เฉพาะทางการเงินและคำทับศัพท์ภาษาอังกฤษได้ดี ทั้งนี้ ผลการวิเคราะห์ผ่านระบบนำเสนอข้อมูลยังชี้ให้เห็นว่า ผู้ใช้งานมีสัดส่วนความคิดเห็นเชิงลบสูงกว่าเชิงบวก โดยส่วนใหญ่เป็นปัญหาด้านความเสถียรของระบบ ผลลัพธ์เหล่านี้จึงเป็นข้อมูลที่ช่วยให้สถาบันการเงินนำไปใช้พัฒนาบริการให้มีประสิทธิภาพตรงกับความต้องการของผู้ใช้งาน

อย่างไรก็ตาม งานวิจัยนี้ยังมีข้อจำกัดด้านความไม่สมดุลของชุดข้อมูล ซึ่งส่งผลต่อความแม่นยำในการจำแนกกลุ่มข้อมูลที่มีปริมาณน้อย แนวทางการวิจัยในอนาคตควรใช้วิธีการเพิ่มข้อมูล เพื่อเพิ่มประสิทธิภาพการจำแนกให้ครอบคลุมทุกคลาส รวมถึงการพัฒนาต่อยอดเป็นระบบรับฟังเสียงลูกค้า เพื่อให้สถาบันการเงินสามารถตอบสนองต่อประเด็นปัญหาของผู้ใช้งานได้ทันที

## เอกสารอ้างอิง


[1] ธนาคารแห่งประเทศไทย. *รายงานทิศทางการพัฒนาระบบการชำระเงินของธนาคารแห่งประเทศไทย (Payment Directional Paper) พ.ศ. 2568*, กรุงเทพ, 2568.

[2] J. Schrank. “The impact of artificial intelligence on behavioral intentions to use mobile banking in the post-COVID-19 era,” *Front Artif Intell.* 2025.

[3] P. Bowornlertsutee and W. Paireekreng, “The Model of Sentiment Analysis for Classifying the Online Shopping Reviews,” *Journal of Engineering and Digital Technology (JEDT)*, Vol. 10, No. 1, pp. 71-79, 2022.

[4] W. Bunyatisai, et al., “Improving Thai Sentiment Analysis Accuracy with Emoji Classification by Deep Learning and Stacking Models: A Case Study of Hotel Reviews,” *Pakistan Journal of Life and Social Sciences*, Vol. 23, No. 1, pp. 65-77, 2025.

[5] M. Kumar, et al., “Evolving techniques in sentiment analysis: a comprehensive review,” *PeerJ Computer Science*, 2025.

[6] L. Lowphansirikul. “WangchanBERTa: Pretraining Transformer-Based Thai Language Models,” *arXiv:2101.09635*, 2021.

[7] P. Sriwirote, et al., “PhayaThaiBERT: Enhancing a Pretrained Thai Language Model with Unassimilated Loanwords,” *ACM Transactions on Asian and Low-Resource Language Information Processing*, Vol. 24, No. 11, pp. 1-17, 2025.

[8] พิศิษฐ์ บวรเลิศสุธี และ วรภัทร ไพรีเกรง. “การวิเคราะห์ความรู้สึกทางอารมณ์ในบทความแนะนำสินค้าออนไลน์.” *วารสารวิจัย มหาวิทยาลัยเทคโนโลยีราชมงคลศรีวิชัย*, ปีที่ 14, ฉบับที่ 3, หน้าที่ 755-769, 2565.

[9] A. G. Asali. “Social Media Analysis for Investigating Consumer Sentiment on Mobile Banking,” *Journal of International Conference Proceedings (JICP)*, Vol. 4, No. 2, pp. 241–253, 2021.

[10] วนัสวรรณ มีประเสริฐ และ เอกรัฐ รัฐกาญจน์. “การวิเคราะห์ความคิดเห็นจากทวิตเตอร์ของลูกค้าบริษัท ช้อปปี้ ประเทศไทย.” *วารสารระบบสารสนเทศด้านธุรกิจ*, ปีที่ 7, ฉบับที่ 3, 2564.

# การพัฒนาระบบจำแนกประเภทปัญหาจาก Log Files ด้วยเทคนิค Machine Learning เพื่อสนับสนุนการวิเคราะห์และการส่งต่อปัญหาในองค์กร Publisher เกม

# Development of a Machine Learning-Based Log File Classification System for Problem Analysis and Escalation Support in a Game Publisher Organization

*พงศ์ภีระ ชัยพงศ์ภัสสร (Pongpeera Chaipongpatsorn)[1] และณัฐวี อุตกฤษฎ์ (Nattavee Utakrit)[2]*
*ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล*
*[2]มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ*
*s6707021856108@email.kmutnb.ac.th, nattavee.u@itd.kmutnb.ac.th*

## บทคัดย่อ

*ปัจจุบันธุรกิจเกมออนไลน์มีการแข่งขันสูงอย่างต่อเนื่อง ความสำเร็จของเกมมิได้ขึ้นอยู่เพียงคุณภาพของตัวเกมเท่านั้น แต่ยังขึ้นอยู่กับเสถียรภาพของระบบสนับสนุนต่าง ๆ ปัญหาที่มักเกิดขึ้นคือกระบวนการวิเคราะห์หาสาเหตุจำเป็นต้องตรวจสอบไฟล์บันทึกเหตุการณ์ (Log File) จำนวนมากด้วยแรงงานมนุษย์ ส่งผลให้การแก้ไขปัญหาล่าช้า งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาระบบช่วยจำแนกประเภทของปัญหาจาก Log File โดยประยุกต์ใช้เทคนิค Machine Learning กระบวนการวิจัยเริ่มจากการรวบรวม Log File จากเว็บไซต์เกมและระบบจัดการบัญชีผู้ใช้ จากนั้นนำข้อมูลมาฝึกสอนและเปรียบเทียบโมเดล Machine Learning ภายหลังได้พัฒนาระบบต้นแบบในรูปแบบที่ใช้งานง่าย โดยผู้ใช้สามารถอัปโหลด Log File และระบบจะแสดงผลการจำแนกประเภทของปัญหา พร้อมทั้งสรุปสถิติเพื่อสนับสนุนการวิเคราะห์เชิงภาพแบบทันที*

**คำสำคัญ:** เกมออนไลน์, การจำแนกประเภท, ระบบสนับสนุนการตัดสินใจ, เหมืองข้อความ, ป่าสุ่ม

## Abstract

*The online gaming industry is experiencing continuous and intense competition. The success of an online game depends on the stability of its supporting systems. Identifying the root cause requires manual examination of large volumes of log files, leading to delayed problem resolution. This study aims to develop an automated problem classification system based on log file analysis using Machine Learning techniques. The research begins with collecting log files, which are then used to train and compare Machine Learning models. Subsequently, a user-friendly prototype system is developed, allowing users to upload log files for automatic analysis and presenting results along with graphical statistical summaries.*

**Keywords:** Online gaming, Classification, Decision support system, Text mining, Random Forest.

## 1. บทนำ

ในอุตสาหกรรมเกมดิจิทัล ความเสถียรและความพร้อมใช้งานของระบบถือเป็นปัจจัยเชิงกลยุทธ์ที่ส่งผลโดยตรงต่อประสบการณ์ของผู้เล่นและความเชื่อมั่นต่อแบรนด์ โดยเฉพาะระบบที่มีบทบาทสำคัญต่อวงจรชีวิตผู้เล่น (Mission-Critical Systems) ได้แก่ ระบบบัญชีผู้เล่น (Account System) และระบบการชำระเงิน (Payment System) ซึ่งเป็นโครงสร้างพื้นฐานหลักของการยืนยันตัวตนและการทำธุรกรรมทางการเงินภายในเกม

ปัจจุบัน แม้ว่างานวิจัยด้านการทำเหมืองข้อความ (Text Mining) และการเรียนรู้ของเครื่อง (Machine Learning) จะถูกนำมาประยุกต์ใช้ในการจำแนกข้อมูลอย่างแพร่หลาย แต่องค์กรส่วนใหญ่ยังคงพึ่งพาระบบเฝ้าระวังพื้นฐานและเครื่องมือแจ้งเตือนแบบดั้งเดิม

ปัญหาของสิ่งที่เป็นอยู่ในปัจจุบัน: ปัญหาสำคัญที่พบในอุตสาหกรรมปัจจุบันคือ ความล่าช้าในการตรวจจับความผิดปกติ โดยในหลายกรณีองค์กรจะรับรู้ปัญหาเมื่อผู้เล่นได้รับผลกระทบและแจ้งเรื่องผ่านช่องทางบริการลูกค้าแล้ว

เนื่องจากกระบวนการวิเคราะห์หาสาเหตุจำเป็นต้องตรวจสอบไฟล์บันทึกเหตุการณ์ (Log Files) จำนวนมากด้วยแรงงานมนุษย์ ซึ่งเป็นกระบวนการที่ใช้เวลานาน มีความเสี่ยงต่อความผิดพลาดสูง และนำไปสู่ความล่าช้าในการแก้ไขปัญหา [1] นอกจากนี้ ในสถาปัตยกรรมแบบ Microservices ที่มีบริการจำนวนมาก การประยุกต์ใช้เทคนิคเหมืองข้อมูล (Data Mining) และการเรียนรู้ของเครื่อง (Machine Learning) จึงกลายเป็นแนวทางสำคัญในการวิเคราะห์ข้อมูลสมัยใหม่ [2] การเฝ้าระวังทั้งหมดพร้อมกันด้วยแรงงานคนจึงเป็นภาระงานที่ซับซ้อนและสิ้นเปลืองทรัพยากร

วิธีการแก้ไขในปัจจุบันและสิ่งที่ต้องการ: ปัจจุบันวิธีการแก้ไขเบื้องต้นคือการใช้ทีมสนับสนุนด้านไอที (IT Support) คอยตรวจสอบ Log Files ด้วยตาเปล่าเมื่อมีการแจ้งเตือน สิ่งที่ระบบในปัจจุบันยังขาดหายและมีความต้องการอย่างเร่งด่วนคือ ระบบอัตโนมัติที่สามารถวิเคราะห์ข้อมูลเชิงเทคนิคที่ถูกบันทึกไว้ และสามารถเชื่อมโยงเข้ากับกระบวนการตัดสินใจเพื่อส่งต่อปัญหา (Incident Escalation) ได้อย่างทันท่วงที,

สิ่งที่นำเสนอและลักษณะของงานวิจัย: เพื่อตอบสนองความต้องการดังกล่าว งานวิจัยนี้จึงนำเสนอ "การพัฒนาระบบจำแนกประเภทปัญหาจาก Log Files ด้วยเทคนิค Machine Learning"

โดยลักษณะของงานวิจัยนี้เป็น งานประเภทประยุกต์ (Applied Research) ที่บูรณาการองค์ความรู้ด้าน Machine Learning เข้ากับการบริหารจัดการบริการเทคโนโลยีสารสนเทศ เพื่อหาผลสัมฤทธิ์ของแบบจำลองในการจำแนกความผิดพลาด

ระบบที่พัฒนาขึ้นจะช่วยจำแนกประเภทความผิดพลาด ระบุสาเหตุเบื้องต้น และส่งต่อปัญหาไปยังทีมที่เกี่ยวข้องได้อย่างรวดเร็ว อันจะนำไปสู่การลดระยะเวลาเฉลี่ยในการแก้ไขปัญหา (MTTR) ได้อย่างมีนัยสำคัญ

เพื่อให้เกิดความเข้าใจที่ตรงกันในบริบทของงานวิจัย จึงขอกำหนดนิยามศัพท์เฉพาะดังนี้

องค์กรผู้ให้บริการเกม (Game Service Provider): หมายถึง บริษัทที่ดำเนินธุรกิจให้บริการเกมออนไลน์ ซึ่งมีการบริหารจัดการระบบโครงสร้างพื้นฐาน ระบบบัญชีผู้ใช้งาน และระบบชำระเงินของตนเอง เพื่อรองรับผู้เล่นจำนวนมาก

ลักษณะข้อมูล (Data Characteristics): หมายถึง ข้อมูลบันทึกเหตุการณ์ (Log Files) ที่จัดเก็บในรูปแบบกึ่งโครงสร้าง (Semi-structured Data) จากเว็บเซิร์ฟเวอร์ Nginx ประกอบด้วยข้อมูลที่สำคัญ เช่น รหัสสถานะตอบกลับ (HTTP Status Code) และระยะเวลาการประมวลผล (Response Time) โดยมุ่งเน้นข้อมูลในช่วงเวลาที่มีภาระงานสูง (Peak Hour)

เพื่อให้การดำเนินงานบรรลุเป้าหมาย งานวิจัยนี้ได้กำหนดวัตถุประสงค์เพื่อพัฒนาระบบจำแนกประเภท ทดสอบอัลกอริทึม และประเมินความพึงพอใจของผู้ใช้งาน โดยมีสมมติฐานการวิจัย 2 ประการ ได้แก่ 1) ระบบจำแนกประเภทปัญหาที่พัฒนาขึ้นมีความแม่นยำไม่น้อยกว่าร้อยละ 80 และ 2) ผู้ใช้งานมีความพึงพอใจต่อระบบในระดับมาก โดยมีค่าเฉลี่ยไม่น้อยกว่า 3.51 จาก 5

## 2. วรรณกรรมที่เกี่ยวข้อง

การพัฒนาระบบจำแนกประเภทปัญหาจาก Log Files อาศัยการบูรณาการองค์ความรู้ด้านการเรียนรู้ของเครื่อง (Machine Learning) และเทคนิคเหมืองข้อความ (Text Mining)

**2.1 เทคนิคการสกัดคุณลักษณะด้วย TF-IDF** Term Frequency–Inverse Document Frequency (TF-IDF) เป็นเทคนิคยอดนิยมในการแปลงข้อความเป็นเวกเตอร์ตัวเลข

สำหรับการประมวลผล เทคนิคนี้ช่วยลดน้ำหนักของข้อความปกติทั่วไปที่ระบบบันทึกไว้ซ้ำ ๆ และเพิ่มน้ำหนักให้กับรูปแบบ (Patterns) ที่สะท้อนลักษณะเฉพาะของข้อผิดพลาด [3] เช่น "Timeout" หรือรหัส "Error 500"

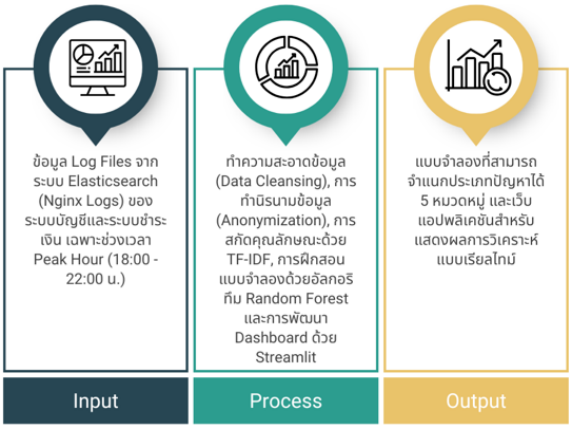
**2.2 งานวิจัยที่เกี่ยวข้อง** งานวิจัยที่เกี่ยวข้องและการจำแนกประเภท งานวิจัยที่เกี่ยวข้องพบว่าเทคนิคการเรียนรู้แบบมีผู้สอน (Supervised Learning) มีประสิทธิภาพสูงในการจำแนกประเภทข้อความ จากงานวิจัยของ ดำรงพล [4] พบว่าอัลกอริทึม Random Forest ให้ประสิทธิภาพสูงสุดในการวิเคราะห์ความคิดเห็นบนสื่อสังคมออนไลน์สอดคล้องกับงานวิจัยของ กัญญารัตน์ และแกมกาญจน์ [5] ที่พบว่า Random Forest ให้ความถูกต้องถึงร้อยละ 82.50 นอกจากนี้ จากงานวิจัยของ วสวัตติ์ และจารี [6] ได้ชี้ให้เห็นว่าเทคนิค K-Nearest Neighbor (KNN) มีประสิทธิภาพสูงในการวิเคราะห์ข้อความรีวิวเกมมือถือเช่นเดียวกับงานวิจัยของ สมศักดิ์ และสมัย [7] ที่พบว่าเทคนิค Naïve Bayes สามารถประมวลผลข้อความภาษาไทยได้ดีเยี่ยม และงานวิจัยของ ณิชาภา [8] ที่ยืนยันว่าเทคนิคเหมืองข้อมูลสามารถจำแนกรูปแบบแพ็กเก็ตที่โจมตีระบบเครือข่ายได้อย่างแม่นยำ รวมถึงสอดคล้องกับการพัฒนาระบบธุรกิจอัจฉริยะเพื่อวิเคราะห์และรายงานผลจากระบบรับแจ้งเหตุขัดข้องของ จิรรัฐ และมณเฑียร [9] งานวิจัยนี้จึงได้นำอัลกอริทึม Random Forest, KNN, Logistic Regression, และ Naïve Bayes มาเปรียบเทียบเพื่อหาแนวทางที่ดีที่สุด

**2.3 ตัวชี้วัดประสิทธิภาพ (Evaluation Metrics)** งานวิจัยที่เกี่ยวข้องพบว่าเทคนิคการเรียนรู้แบบมีผู้สอน (Supervised Learning) เพื่อความแม่นยำในการวัดผลงานวิจัยนี้ใช้ตารางเมทริกซ์สับสน (Confusion Matrix) ในการคำนวณตัวชี้วัดหลัก 4 ค่า ได้แก่:

1. ค่าความถูกต้อง (Accuracy)
2. ค่าความแม่นยำ (Precision)
3. ค่าความระลึก (Recall)
4. ค่าเอฟวัน (F1-Score) [10] ซึ่งจำเป็นอย่างยิ่งสำหรับข้อมูลที่มีลักษณะไม่สมดุล (Class Imbalance)

## 3. วิธีการดำเนินการวิจัย

**3.1 กรอบแนวคิดของงานวิจัย** ผู้วิจัยได้ประยุกต์ใช้แบบจำลอง Input-Process-Output (IPO) ดังภาพที่ 1 เพื่อแสดงให้เห็นถึงการไหลของข้อมูลตั้งแต่การนำเข้าสู่ระบบการประมวลผลแบบจำลอง จนถึงการแสดงผลลัพธ์ผ่านแดชบอร์ด




**ภาพที่ 1:** กรอบแนวคิดของงานวิจัยตามแบบจำลอง IPO

**3.2 การเก็บรวบรวมและเตรียมข้อมูล** สกัดข้อมูลจากระบบ Elasticsearch ของระบบบัญชีและชำระเงิน โดยสุ่มข้อมูลช่วง Peak Hour (18:00–22:00 น.) ย้อนหลัง 7 วัน จากนั้นทำความสะอาดและทำนิรนามข้อมูล (Anonymization) ข้อมูลถูกแบ่งเป็น 5 หมวดหมู่ ได้แก่ Success, Client_Error, Server_Error, Client_Timeout และ Security_Block เนื่องจากข้อมูลมีความไม่สมดุล (Class Imbalance) จึงได้จัดการด้วยเทคนิคการสุ่มเพิ่มข้อมูลกลุ่มน้อย (SMOTE) บนชุดข้อมูลฝึกสอน (Training Data) เพื่อป้องกันโมเดลเอนเอียงไปทางกลุ่มที่มีข้อมูลมาก

**3.3 การพัฒนาและเปรียบเทียบแบบจำลอง** ข้อมูลถูกแบ่งเป็นชุดฝึกสอน (Training Set) ร้อยละ 80 และชุดทดสอบ (Testing Set) ร้อยละ 20 จากนั้นแปลงข้อความเป็นเวกเตอร์ด้วย TF-IDF และนำไปฝึกสอนเปรียบเทียบกับอัลกอริทึม Logistic Regression, K-Nearest Neighbors (KNN) และ Random Forest

**3.4 การพัฒนาระบบและการประเมินผล** พัฒนาระบบแสดงผล (Dashboard) ด้วยไลบรารี Streamlit บนภาษา Python เพื่อรองรับการอัปโหลดไฟล์ Log (CSV) และแสดงผลเชิงวิเคราะห์ จากนั้นประเมินความเที่ยงตรงของ

เนื้อหาด้วยดัชนี IOC โดยผู้เชี่ยวชาญ 3 ท่าน และประเมินความพึงพอใจโดยผู้ใช้งาน 10 ท่าน

## 4. ผลการดำเนินงาน

**4.1 ผลการเตรียมข้อมูล** จากการสกัดข้อมูลพบข้อมูล Log รวมทั้งสิ้น 21,000 รายการ โดยมีการกระจายตัวของหมวดหมู่เหตุการณ์ดังแสดงในตารางที่ 1 ซึ่งสะท้อนลักษณะข้อมูลที่ไม่สมดุลอย่างชัดเจน

**ตารางที่ 1:** การกระจายตัวของหมวดหมู่เหตุการณ์จากข้อมูล Log

| ประเภทเหตุการณ์ | ตัวอย่างรหัสสถานะ (Status Code) | จำนวนโดยประมาณ (รายการ) | ร้อยละ (%) |
|---|---|---|---|
| Success | 200 OK | 17,850.00 | 85 |
| Client_Error | 403, 405 | 1,680.00 | 8 |
| Server_Error | 502 | 630 | 3 |
| Client_Timeout | 499 (> 3.0 วินาที) | 735 | 3.5 |
| Security_Block | 444 | 105 | 0.5 |
| | รวม | 21,000.00 | 100 |

**4.2 ผลการเปรียบเทียบและจำแนกประเภทของอัลกอริทึม** เมื่อนำชุดข้อมูลทดสอบไปประเมินประสิทธิภาพของแต่ละแบบจำลอง พบว่าอัลกอริทึม Random Forest มีประสิทธิภาพโดยรวมสูงที่สุด ดังแสดงในตารางที่ 2

**ตารางที่ 2:** เปรียบเทียบประสิทธิภาพของแบบจำลอง

| อัลกอริทึม | Accuracy | Precision | Recall | F1-Score |
|---|---|---|---|---|
| Logistic Regression | 0.97 | 0.97 | 0.97 | 0.97 |
| K-Nearest Neighbors (KNN) | 0.99 | 0.99 | 0.99 | 0.99 |
| Random Forest | 1.00 | 1.00 | 1.00 | 1.00 |

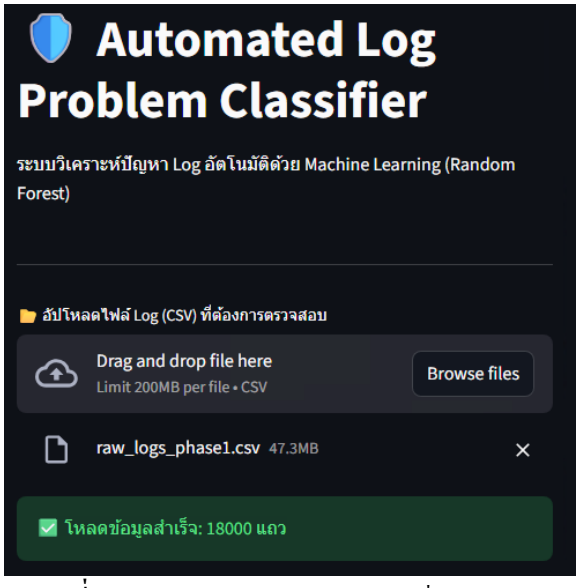
เพื่อพิจารณาความสามารถในการแยกแยะประเภทปัญหารายคลาสของ Random Forest ผู้วิจัยได้นำเสนอผลการทำนายบนข้อมูลทดสอบจำนวน 4,200 รายการในรูปแบบตาราง Confusion Matrix (ตารางที่ 3)

**ตารางที่ 3:** เปรียบเทียบประสิทธิภาพของแบบจำลอง

| ความเป็นจริง \ ทำนาย | Success | Client_Error | Server_Error | Timeout | Security |
|---|---|---|---|---|---|
| Success | **3,570** | 0 | 0 | 0 | 0 |
| Client_Error | 0 | **336** | 0 | 0 | 0 |
| Server_Error | 0 | 0 | **126** | 0 | 0 |
| Client_Timeout | 0 | 0 | 0 | **147** | 0 |
| Security_Block | 0 | 0 | 0 | 0 | **21** |

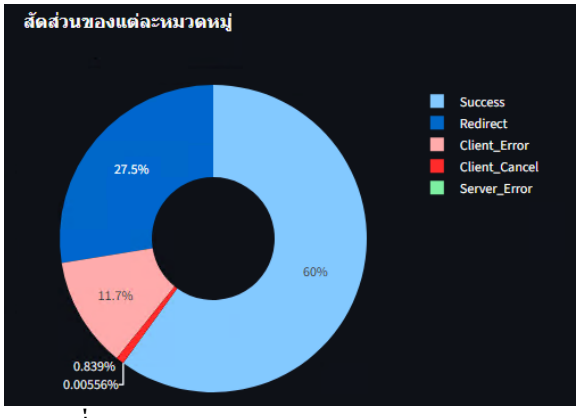
จากตารางที่ 3 แสดงให้เห็นว่าไม่มีความผิดพลาดแบบ False Positive และ False Negative เกิดขึ้นเลย การที่แบบจำลองมีความแม่นยำร้อยละ 100 เกิดจากลักษณะของข้อมูล Log Files ที่เป็นข้อมูลกึ่งโครงสร้าง (Semi-structured Data) อย่างชัดเจน โดยรหัสสถานะมีความสัมพันธ์เชิงเส้นตรงกับการเกิดปัญหา ผสานกับการทำความสะอาดข้อมูลที่รัดกุม ทำให้โมเดลเรียนรู้และสร้างกฎการตัดสินใจได้อย่างชัดเจนโดยไม่มีความกำกวม

**4.3 ผลการพัฒนาระบบแสดงผล (Dashboard)** ระบบถูกพัฒนาเป็นเว็บแอปพลิเคชันรูปแบบอัปโหลดข้อมูล (Batch Processing) โดยเมื่อผู้ใช้นำเข้าไฟล์ Log (CSV) ดังภาพที่ 2 ระบบจะประมวลผลผ่านโมเดล Random Forest ในทันที (มิใช่แบบ Real-time Streaming)



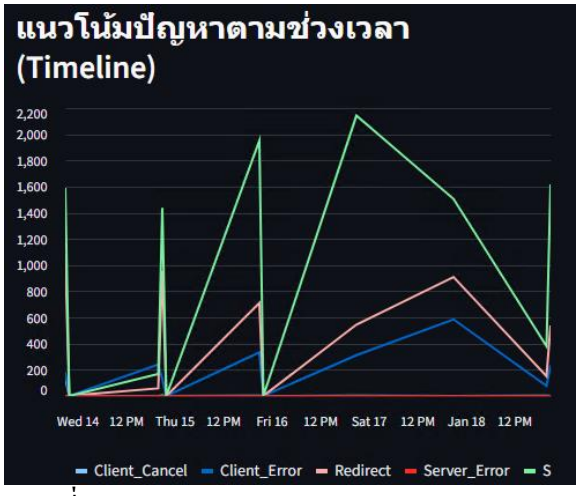

**ภาพที่ 2:** ภาพหน้าจอส่วนการอัปโหลดไฟล์เพื่อการวิเคราะห์อัตโนมัติ

ระบบจะประมวลผลและสร้าง Dashboard แสดงสัดส่วนปัญหาในภาพรวมด้วยกราฟวงกลม เพื่อให้ทีมผู้ดูแลสามารถจัดลำดับความสำคัญ (Prioritization) ของปัญหาได้ทันที ดังภาพที่ 3




**ภาพที่ 3:** ตัว Dashboard แสดงผลการวิเคราะห์สัดส่วนปัญหา

นอกจากนี้ ยังมีการนำเสนอแนวโน้มจำนวนเหตุการณ์เชิงเวลาด้วยกราฟเส้น ดังภาพที่ 4 เพื่อตรวจจับช่วงเวลาที่เกิดข้อผิดพลาดหนาแน่น ช่วยสนับสนุนการวางแผนเชิงป้องกันขององค์กร




**ภาพที่ 4:** หน้าจอ Dashboard แสดงแนวโน้มปัญหาตามช่วงเวลา

**4.4 ผลการประเมินประสิทธิภาพระบบและความพึงพอใจ** ผู้เชี่ยวชาญทั้ง 3 ท่านประเมินความเที่ยงตรงของผลการจำแนกได้ค่า IOC เท่ากับ 1.00 ในทุกรายการ นอกจากนี้ผลประเมินความพึงพอใจจากผู้ใช้งานจริงจำนวน 10 ท่าน พบว่าผู้ใช้งานมีความพึงพอใจรวมในระดับมาก (Mean = 4.20, S.D. = 0.62)

โดยประเด็นด้านรูปแบบการนำเสนอข้อมูลมีค่าเฉลี่ยสูงสุด (Mean = 4.60) และประเด็นด้านการลดเวลาการวิเคราะห์ปัญหา (MTTR) มีค่าเฉลี่ย 3.80 (S.D. = 0.63) ซึ่งเมื่อทดสอบทางสถิติแล้วยืนยันได้ว่าผู้ใช้เห็นด้วยว่าระบบช่วยลดเวลา MTTR ได้จริงอย่างมีนัยสำคัญ

## 5. สรุปและอภิปรายผล

ผลการวิจัยเป็นไปตามสมมติฐานทั้ง 2 ข้อ โดยอัลกอริทึม Random Forest ให้ประสิทธิภาพสูงสุดสอดคล้องกับงานวิจัยของ ดำรงพล [4] และ กัญญารัตน์ [5] อย่างไรก็ตาม สาเหตุที่งานนี้มีความแม่นยำสูงถึงร้อยละ 100 เนื่องจากธรรมชาติของ Log Files มีความซับซ้อนทางวากยสัมพันธ์ต่ำกว่าข้อความภาษาธรรมชาติที่มนุษย์พิมพ์ในสื่อสังคมออนไลน์

จากตารางแสดงผล Confusion Matrix ที่แบบจำลอง Random Forest ให้ประสิทธิภาพความแม่นยำร้อยละ 100 ซึ่งเป็นค่าในอุดมคติ (Idealistic) นั้น สามารถอภิปรายได้

ว่าเกิดจากลักษณะของข้อมูล Log Files จาก Nginx ที่เป็นข้อมูลกึ่งโครงสร้าง (Semi-structured Data) อย่างชัดเจน

โดยรหัสสถานะ (Status Code) เช่น 499, 500, หรือ 444 มีความสัมพันธ์เชิงเส้นตรงกับการเกิดปัญหาแต่ละประเภทอย่างมีนัยสำคัญ ประกอบกับการทำความสะอาดข้อมูลอย่างรัดกุม ทำให้แบบจำลองสามารถสร้างกฎการตัดสินใจได้อย่างชัดเจนโดยไม่มีความกำกวมทางภาษาเหมือนข้อมูลข้อความทั่วไป

อย่างไรก็ตาม ผู้วิจัยตระหนักดีว่าหากนำไปประยุกต์ใช้กับบริการอื่นที่มีลักษณะ Log แตกต่างกันหรือมีความกำกวมสูงขึ้น ประสิทธิภาพอาจมีการเปลี่ยนแปลงได้

## เอกสารอ้างอิง


[1] R. Feldman and J. Sanger, The text mining handbook: Advanced approaches in analyzing unstructured data, Cambridge University Press: New York, 2007.

[2] J. Han, J. Pei, and M. Kamber, Data mining: Concepts and techniques (3rd ed.). Elsevier, 2011.

[3] สถาบันส่งเสริมการวิเคราะห์และบริหารข้อมูลขนาดใหญ่ภาครัฐ, “TF-IDF สกัดใจความสำคัญของข้อความ-Part1,” 2563.

[4] ดำรงพล สุวรรณชาตรี และ วัชรีวรรณ จิตต์สกุล, “การวิเคราะห์ความคิดเห็นเกี่ยวกับรถยนต์ไฟฟ้าบนสื่อสังคมออนไลน์ด้วยเทคนิคเหมืองข้อความ,” 2564.

[5] กัญญารัตน์ พูลเพิ่ม และ แกมกาญจน์ สมประเสริฐศรี, “การจำแนกโครงการตามแผนแม่บทภายใต้ยุทธศาสตร์ชาติโดยใช้เทคนิคเหมืองข้อความ,” วารสารวิชาการ การจัดการเทคโนโลยี มหาวิทยาลัยราชภัฏมหาสารคาม, ปีที่ 9, ฉบับที่ 2, 2565.

[6] วสวัตติ์ อินทร์แปลง และ จารี ทองคำ, “การวิเคราะห์ความคิดเห็นต่อเกมมือถือพับจีด้วยเหมืองข้อความ,” วารสารวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยมหาสารคาม, ปีที่ 39, ฉบับที่ 5, 2563.

[7] สมศักดิ์ ศรีสวการย์ และ สมัย ศรีสวย, “การวิเคราะห์เหมืองความคิดเห็นโดยใช้เทคนิคการสกัดคำ,” วารสารวิชาการการประยุกต์ใช้เทคโนโลยีสารสนเทศ มหาวิทยาลัยราชภัฏมหาสารคาม, ปีที่ 6, ฉบับที่ 2, 2563.

[8] ณิชาภา แซ่เหวียง, “การจำแนกรูปแบบแพ็กเก็ตที่ใช้โจมตีระบบเครือข่ายด้วยเทคนิคเหมืองข้อมูล,” มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2565.

[9] จิรรัฐ ศรีโชค และ มณเฑียร รัตนศิริวงศ์วุฒิ, “การพัฒนาระบบธุรกิจอัจฉริยะเพื่อการวิเคราะห์ข้อมูลจากระบบการรับแจ้งเหตุขัดข้องโดยใช้เทคนิคเหมืองข้อมูล,” 2564.

[10] D. M. W. Powers, "Evaluation: From precision, recall and F-measure to ROC, informedness, markedness & correlation," Journal of Machine Learning Technologies, Vol. 2, No. 1, pp. 37-63, 2011.

# การพัฒนาเฟรมเวิร์กการพยากรณ์และการตัดสินใจเติมสินค้าโดยใช้ข้อมูลยอดขายและสินค้าคงคลังด้วยเทคนิคเหมืองข้อมูลและการเรียนรู้แบบเสริมกำลัง

# Development of a Forecasting and Decision-Making Framework for Inventory Replenishment Using Sales and Inventory Data with Data Mining Techniques and Reinforcement Learning

*ธีรธัช คงยืนยงวาณิชย์ (Teeratuch Kongyuenyongwanit)[1] และพยุง มีสัจ (Phayung Meesad)[2]*
*ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล*
*มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ*
*[1]s6707021856116 @email.kmutnb.ac.th, [2]phayung.m@itd.kmutnb.ac.th*

## บทคัดย่อ

*งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาเฟรมเวิร์กการพยากรณ์และการตัดสินใจเติมสินค้าในธุรกิจค้าส่ง โดยใช้ข้อมูลยอดขายและข้อมูลสินค้าคงคลังร่วมกับเทคนิคเหมืองข้อมูล การเรียนรู้ของเครื่อง และการเรียนรู้แบบเสริมกำลัง เฟรมเวิร์กประกอบด้วย 2 ส่วนหลัก ได้แก่ ส่วนการพยากรณ์ความต้องการสินค้า และส่วนการตัดสินใจเติมสินค้า ในส่วนการพยากรณ์ได้เปรียบเทียบตัวแบบ 4 รูปแบบ ได้แก่ LSTM, XGBoost, XGBoost-LSTM และ LSTM-XGBoost โดยใช้ข้อมูลจำนวน 199,720 ระเบียน แบ่งเป็นชุดฝึกสอน ตรวจสอบ และทดสอบในสัดส่วน 70:15:15 พร้อมปรับพารามิเตอร์ด้วย Grid Search ผลการทดลองพบว่า LSTM-XGBoost ให้ผลดีที่สุดเมื่อพิจารณาจากค่า RMSE และ $R^2$ โดยมีค่า MAE เท่ากับ 3.6076, RMSE เท่ากับ 9.9295, MAPE เท่ากับ 85.2188 และ $R^2$ เท่ากับ 0.7707 ส่วนการตัดสินใจเติมสินค้าด้วย Deep Q-Network เปรียบเทียบกับนโยบาย Always Order และ Never Order พบว่า DQN มีค่า Avg Total Reward เท่ากับ -1,166.55 และอัตราสินค้าขาดคลังเฉลี่ยร้อยละ 0.35 จากการทดสอบสินค้า 504 รายการ ผลการศึกษาชี้ว่าเฟรมเวิร์กที่เสนอสามารถสนับสนุนทั้งการพยากรณ์และการตัดสินใจเติมสินค้าได้อย่างเป็นระบบ*

**คำสำคัญ:** LSTM, XGBoost, Deep Q-Network, การพยากรณ์การเติมสินค้า, การเรียนรู้แบบเสริมกำลัง

## Abstract

*This research presents a comprehensive system that integrates forecasting and decision-making for inventory replenishment at a wholesale enterprise. It exploits sales and inventory figures through data mining, machine learning, and reinforcement learning methods. "That framework has two primary parts: projecting customer demand and making choices about inventory replenishment. In the demand forecasting aspect, four different models were tested: LSTM, XGBoost, XGBoost-LSTM, and LSTM-XGBoost. The analysts split a dataset of 199,720 records into training, validation, and test sets at a 70:15:15 ratio, then performed hyperparameter tuning via grid search. Their findings revealed that LSTM-XGBoost was the most accurate one according to RMSE and $R^2$ criteria. Its MAE was 3.6076, RMSE was 9.9295, MAPE was 85.2188, and $R^2$ was 0.7707. Regarding the decision-making for inventory replenishment, the Deep Q-Network was tested against two policies: Always Order and Never Order. DQN recorded a mean total reward of -1, -1,166.55, and a mean stockout rate of 0.35%, consacross evaluated products, according to the conducted experiments. In conclusion, the study demonstrates that the proposed system is efficient not only for*

*demand forecasting but also for inventory replenishment planning, underscoring its value as a tool for both.*

**Keywords:** Long Short-Term Memory, Extreme Gradient Boosting, Inventory Replenishment Forecasting, Inventory Management, Hybrid Model.

## 1. บทนำ

การบริหารสินค้าคงคลังเป็นปัจจัยสำคัญต่อความสำเร็จของธุรกิจ เนื่องจากต้องรักษาสมดุลระหว่างการมีสินค้าเพียงพอต่อความต้องการของลูกค้า และการหลีกเลี่ยงต้นทุนส่วนเกินจากการเก็บสินค้ามากเกินไป อย่างไรก็ตาม ธุรกิจจำนวนมาก โดยเฉพาะ SMEs ยังขาดเครื่องมือวิเคราะห์ข้อมูลที่มีประสิทธิภาพ ส่งผลให้การตัดสินใจสั่งซื้อสินค้ายังคงอิงจากประสบการณ์ ซึ่งอาจไม่สามารถคาดการณ์แนวโน้มความต้องการในอนาคตได้อย่างแม่นยำ ในยุคดิจิทัล การนำเทคโนโลยีเหมืองข้อมูลและการเรียนรู้ของเครื่อง เช่น LSTM, XGBoost และตัวแบบผสม มาใช้วิเคราะห์ข้อมูลเชิงเวลาได้รับความสนใจมากขึ้น

อย่างไรก็ตาม งานวิจัยที่ผ่านมายังมีข้อจำกัด กล่าวคือ งานวิจัยที่ใช้ LSTM มักมุ่งเน้นข้อมูลทางการเงิน [6] ไม่สอดคล้องกับข้อมูลแบบ Intermittent และ Lumpy ขณะที่งานวิจัยที่ใช้ XGBoost [7] ไม่บูรณาการข้อมูลสินค้าคงคลัง ทำให้ไม่สามารถจำแนกยอดขายศูนย์ได้ งานวิจัยนี้จึงปิดช่องว่างด้วยคุณูปการ 4 ประการ ได้แก่ (1) นำเสนอเฟรมเวิร์กครอบคลุมตั้งแต่การจัดเตรียมข้อมูลจนถึงการเรียนรู้นโยบายเติมสินค้า (2) บูรณาการข้อมูลยอดขายและสินค้าคงคลังเพื่อจัดการ Zero Demand (3) เปรียบเทียบตัวแบบพยากรณ์เดี่ยวและตัวแบบผสม 4 รูปแบบ และ (4) ประเมินนโยบายเติมสินค้าด้วย DQN เทียบกับ Always Order และ Never Order

เฟรมเวิร์กนี้จะช่วยให้ธุรกิจวางแผนการสั่งซื้อได้อย่างเหมาะสม ลดปัญหาสินค้าขาดหรือเกินคลัง และเพิ่มประสิทธิภาพในการบริหารจัดการสินค้าคงคลังอย่างยั่งยืน

## 2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ผู้วิจัยได้ทำการรวบรวมข้อมูลเชิงทฤษฎีและผลงานวิจัยที่เกี่ยวข้อง เพื่อสร้างพื้นฐานความรู้และกรอบแนวคิดในการศึกษา โดยแบ่งออกเป็น 3 หัวข้อหลัก ดังนี้

### 2.1 การทำเหมืองข้อมูล

เหมืองข้อมูลเป็นกระบวนการสกัดความรู้จากข้อมูลขนาดใหญ่โดยอาศัยเทคนิคทางสถิติ ปัญญาประดิษฐ์ และการเรียนรู้ของเครื่อง เพื่อค้นหารูปแบบและความสัมพันธ์ที่ซ่อนอยู่ในข้อมูล ครอบคลุมเทคนิคหลัก เช่น การจำแนกประเภท การจัดกลุ่ม และการพยากรณ์ ซึ่งมีบทบาทสำคัญในการวิเคราะห์ข้อมูลและสนับสนุนการตัดสินใจเชิงธุรกิจ [1]

### 2.2 เทคนิคการเรียนรู้เชิงลึก

Deep Learning เป็นเทคโนโลยีที่พัฒนาต่อยอดจาก Machine Learning โดยอาศัยโครงสร้างโครงข่ายประสาทเทียม (Neural Network) ในการเรียนรู้และตัดสินใจจากข้อมูลโดยอัตโนมัติ โดยไม่ต้องกำหนดกฎจากมนุษย์โดยตรง และสามารถปรับปรุงความแม่นยำได้เองจากข้อมูลจำนวนมาก โดยมีรูปแบบโครงข่ายที่หลากหลายเพื่อให้เหมาะสมกับลักษณะของปัญหาในแต่ละงาน [2]

### 2.3 หน่วยความจำระยะสั้นแบบยาว (LSTM)

LSTM เป็นโครงข่ายประสาทเทียมในตระกูล RNN ที่พัฒนาขึ้นเพื่อแก้ปัญหา Vanishing Gradient ทำให้สามารถเรียนรู้ความสัมพันธ์ของข้อมูลระยะยาวได้ดีขึ้น โดยอาศัยกลไก "ประตู" ทั้ง 3 ได้แก่ Forget Gate, Input Gate และ Output Gate ในการคัดเลือกและส่งต่อข้อมูลที่สำคัญ ทำให้ LSTM เหมาะสำหรับงานข้อมูลอนุกรมเวลา [3]

### 2.4 เอกซ์จีบูสต์ (XGBoost)

XGBoost เป็นอัลกอริทึมการเรียนรู้ของเครื่องแบบ Ensemble ที่พัฒนาจากแนวคิด Gradient Boosting โดยใช้ต้นไม้ตัดสินใจหลายต้นเรียนรู้ข้อผิดพลาดของกันและกันแบบลำดับขั้น จุดเด่นของ XGBoost คือมีความแม่นยำสูงและประมวลผลได้รวดเร็ว ทำให้ได้รับความนิยมอย่างแพร่หลายในงานวิเคราะห์ข้อมูลและการพยากรณ์ [4]

### 2.5 การเรียนรู้ดิวเชิงลึก (Deep Q-Network)

DQN เป็นอัลกอริทึมการเรียนรู้แบบเสริมกำลัง (Reinforcement Learning) ที่ผสาน Q-Learning กับโครงข่ายประสาทเทียมเชิงลึก เพื่อแก้ปัญหาการจัดการ

สถานะที่มีขนาดใหญ่และซับซ้อน โดยใช้ Neural Network เป็นตัวประมาณค่า Q-Value แทนการใช้ Q-Table แบบดั้งเดิม ทำให้สามารถเรียนรู้และตัดสินใจได้อย่างมีประสิทธิภาพในสภาพแวดล้อมที่เปลี่ยนแปลงได้ [5]

**2.6 งานวิจัยที่เกี่ยวข้อง**

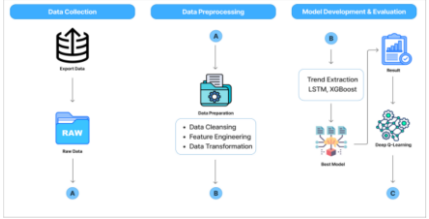
Zhang et al. [6] ผสาน LSTM เข้ากับ Whale Optimization Algorithm (WOA) เพื่อปรับแต่งพารามิเตอร์ และใช้ Wavelet Denoising ลดสัญญาณรบกวนก่อนฝึกสอน ผลพบว่า WOA-LSTM ลดค่า MAE ได้ประมาณร้อยละ 22 เมื่อเทียบกับ LSTM แบบดั้งเดิม

Dankorpho [7] ศึกษาการพยากรณ์ยอดขายด้วย XGBoost โดยใช้ข้อมูลยอดขายของ Walmart จาก 45 สาขา และประเมินผลด้วย MAE, MSE, RMSE และ $R^2$ ผลการศึกษาพบว่า XGBoost มีประสิทธิภาพสูง โดยได้ค่า $R^2$ ถึง 0.9762 และมีค่าความคลาดเคลื่อนในระดับต่ำ สะท้อนให้เห็นถึงความเหมาะสมของอัลกอริทึมดังกล่าวสำหรับการพยากรณ์ยอดขายในบริบทของข้อมูลเชิงตาราง

Mnih et al. [8] พัฒนา DQN โดยผสาน Reinforcement Learning กับ Deep Neural Network เพื่อประมาณค่า Q-value โดยทดลองกับเกม Atari 2600 จำนวน 49 เกม ผลพบว่า DQN มีประสิทธิภาพเหนือกว่า Reinforcement Learning แบบดั้งเดิม และสามารถทำคะแนนสูงกว่ามนุษย์ในหลายเกม เช่น Breakout

## 3. วิธีการดำเนินการวิจัย

งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาเฟรมเวิร์กการพยากรณ์และการตัดสินใจเติมสินค้า และเปรียบเทียบประสิทธิภาพของตัวแบบพยากรณ์โดยใช้ข้อมูลยอดขายและข้อมูลสินค้าคงคลังเป็นกรณีศึกษา รวมถึงประยุกต์ใช้อัลกอริทึม Reinforcement Learning เพื่อกำหนดนโยบายการเติมสินค้าที่เหมาะสม โดยมีกรอบแนวคิดแสดงดังภาพที่ 1




**ภาพที่ 1:** กรอบแนวคิด

จากภาพที่ 1 นำเสนอกรอบแนวคิดการพยากรณ์ความต้องการสินค้าคงคลังโดยใช้เทคนิคการเรียนรู้เชิงลึก ซึ่งประกอบด้วย 3 ขั้นตอนหลัก ได้แก่ การเก็บรวบรวมข้อมูล การจัดเตรียมข้อมูล การคัดเลือกและทดสอบตัวแบบ

การทดลองส่วนพยากรณ์ใช้สินค้า 6 รายการ ขณะที่ส่วน DQN ใช้สินค้า 504 รายการ ความแตกต่างของจำนวนสินค้าในสองส่วนนี้ถือเป็นข้อจำกัดที่ควรพัฒนาให้เชื่อมโยงกันในอนาคต

**3.1 การเก็บรวบรวมข้อมูล**

ผู้วิจัยได้รวบรวมข้อมูลความต้องการสินค้าและข้อมูลสินค้าคงคลังจากบริษัท ของดี สโตร์ จำกัด ในช่วงเดือนมกราคมถึงธันวาคม พ.ศ. 2568 รวมระยะเวลา 12 เดือน ประกอบด้วยข้อมูลยอดขายรายวัน ระดับสินค้าคงคลัง และประวัติการสั่งซื้อสินค้า โดยมีจำนวนรายการสินค้า (SKU) ทั้งสิ้น 1,295 รายการ และจำนวนระเบียนข้อมูลรวม 199,720 ระเบียน โดยนำข้อมูลออกจากฐานข้อมูล

3.1.1 การเติมเต็มช่องว่างของเวลา (Time Gap Imputation) เพื่อให้สอดคล้องกับข้อกำหนดของอัลกอริทึมสำหรับข้อมูลอนุกรมเวลา เช่น LSTM จำเป็นต้องใช้ข้อมูลที่มีความต่อเนื่องและมีช่วงเวลาเท่ากัน อย่างไรก็ตาม ข้อมูลยอดขายจริงมักมีช่องว่างของเวลา ผู้วิจัยจึงจัดการโดยสร้างลำดับเวลาแบบต่อเนื่อง และเติมเต็มค่าที่ขาดหายด้วยวิธีที่เหมาะสม เช่น การกำหนดค่าเป็นศูนย์หรือการประมาณค่า เพื่อรักษาความสมบูรณ์ของข้อมูลก่อนนำไปวิเคราะห์

3.1.2 การจัดการปริมาณความต้องการสินค้าที่เป็นศูนย์ (Zero Demand Handling) ค่าศูนย์ในข้อมูลยอดขายสามารถเกิดได้จากความต้องการจริงหรือภาวะสินค้าขาดแคลน ผู้วิจัยจึงพิจารณาร่วมกับข้อมูลสินค้าคงคลัง โดยหากยอดขายและสต็อกเป็นศูนย์พร้อมกัน จะถือเป็นกรณีสินค้าขาดและทำการประมาณค่าแทน แต่หากยังมีสต็อกและยอดขายเป็นศูนย์ จะถือเป็นความต้องการจริงและคงค่าเดิมไว้ เพื่อให้ตัวแบบเรียนรู้พฤติกรรมได้อย่างถูกต้อง

**3.1.3 การจัดรูปแบบข้อมูลให้อยู่ในลักษณะลำดับเวลา (Time Series Sequencing)** ผู้วิจัยแปลงข้อมูลจากรูปแบบตารางให้อยู่ในลักษณะอนุกรมเวลา โดยใช้เทคนิค Sliding Window เพื่อสร้างความสัมพันธ์ระหว่างข้อมูลใน

อดีต (Features) และค่าที่ต้องการพยากรณ์ในอนาคต (Target) ทำให้ข้อมูลมีความเหมาะสมสำหรับการฝึกสอนตัวแบบ และช่วยให้สามารถดึงคุณลักษณะสำคัญของข้อมูลได้อย่างมีประสิทธิภาพ

**3.2 การจัดเตรียมข้อมูล**

การจัดเตรียมข้อมูลประกอบด้วยการตรวจสอบความสมบูรณ์ การจัดรูปแบบวันที่ การรวมข้อมูลยอดขายรายวันกับสินค้าคงคลังตามรหัสสินค้าและวันที่ การเติมเต็มช่วงเวลาที่ขาดหาย การจัดการกรณี Zero Demand และการแปลงข้อมูลเป็นอนุกรมเวลาด้วยเทคนิค Sliding Window โดยกำหนด Look Back Window เท่ากับ 14 วัน เพื่อใช้ข้อมูลในอดีตเป็นคุณลักษณะสำหรับพยากรณ์ความต้องการในอนาคต

**3.3 การคัดเลือกและทดสอบตัวแบบ**

งานวิจัยนี้พัฒนาตัวแบบพยากรณ์ความต้องการสินค้า 4 รูปแบบ ได้แก่ LSTM, XGBoost และโมเดลแบบผสม โดยแบ่งข้อมูลในอัตรา 70:15:15 ด้วยวิธี Train-Validate-Test Split และปรับแต่งพารามิเตอร์ทั้งแบบ Tuned และ Fixed เพื่อเปรียบเทียบประสิทธิภาพอย่างเป็นธรรมภายใต้ข้อจำกัดของทรัพยากรฮาร์ดแวร์

**3.4 การออกแบบ Deep Q-Network สำหรับการตัดสินใจเติมสินค้า**

ในส่วนการตัดสินใจเติมสินค้า งานวิจัยใช้ Deep Q-Network (DQN) เรียนรู้นโยบายการสั่งซื้อโดยกำหนด State จากปริมาณความต้องการในอดีต ระดับสินค้าคงคลัง และสถานะขาดคลัง Action เป็น 2 ทางเลือกคือสั่งซื้อและไม่สั่งซื้อ และ Reward คำนวณจากต้นทุนการถือครองและค่าปรับจากการขาดสต็อก โดยเปรียบเทียบกับนโยบายพื้นฐาน Always Order และ Never Order ด้วยตัวชี้วัด Avg Total Reward, Avg Stockout Rate และ Avg Final Stockout

**ตารางที่ 1 :** ค่าพารามิเตอร์ของ LSTM

| Parameter | Value |
|---|---|
| Loss Function | Mean Squared Error |
| Optimizer | Adam (learning_rate=0.001) |
| Look Back Windows | 14 |
| LSTM Layer | 1, 2 |
| Dropout Rate | 0.1, 0.2, 0.3 |
| Batch Size | 16, 32, 64, 128 ,256 |
| LSTM Unit | 16, 32, 64, 128, 256 |
| Epochs | 50 |

จากตารางที่ 1 LSTM ใช้ค่า Loss Function เป็น Mean Squared Error และ Optimizer แบบ Adam (learning rate = 0.001) โดยกำหนด Look Back Window = 14 พร้อมปรับจำนวนชั้น LSTM (1–2 ชั้น), LSTM Units (16–256), Dropout Rate (0.1–0.3), Batch Size (16–256) และ Epochs = 50 เพื่อเพิ่มประสิทธิภาพการเรียนรู้ข้อมูลอนุกรมเวลา

**ตารางที่ 2 :** ค่าพารามิเตอร์ของ XGBoost

| Parameter | Value |
|---|---|
| random_state | 42 |
| learning_rate | 0.01, 0.05, 0.1 |
| n_estimators | 300, 500, 1000 |
| min_child_weight | 1, 5, 10 |
| max_depth | 3, 4, 6 |
| subsample | 0.8, 1.0 |
| Colsample_bytree | 0.8, 1.0 |

จากตารางที่ 2 XGBoost กำหนดค่า random_state = 42 ได้แก่ learning rate (0.01–0.1), n_estimators (300–1000), min_child_weight (1-10), max_depth (3–6), subsample (0.8–1.0) และ colsample_bytree (0.8-1.0) เพื่อควบคุมความซับซ้อนและเพิ่มความแม่นยำของโมเดล

**ตารางที่ 3 :** ค่าพารามิเตอร์ของตัวแบบ Hybrid XGBoost+LSTM

| Parameter | Value |
|---|---|
| random_state | 42 |
| learning_rate | 0.01, 0.05, 0.1 |
| n_estimators | 300, 500, 1000 |
| min_child_weight | 1, 5, 10 |
| max_depth | 3, 4, 6, 8 |
| subsample | 0.8, 1.0 |

จากตารางที่ 3 Hybrid XGBoost+LSTM กำหนดค่า random_state = 42 และปรับพารามิเตอร์หลัก ได้แก่ learning rate (0.01–0.1), n_estimators (300–1000), min_child_weight (1-10), max_depth (3–8) และ subsample (0.8–1.0) เพื่อควบคุมความซับซ้อนและเพิ่มความแม่นยำของโมเดล

**ตารางที่ 4 :** ค่าพารามิเตอร์ของตัวแบบ Hybrid LSTM+XGBoost

| Parameter | Value |
|---|---|
| random_state | 42 |
| learning_rate | 0.01, 0.05, 0.1 |
| n_estimators | 300, 500, 1000 |
| min_child_weight | 1, 5, 10 |
| max_depth | 3, 4, 6, 8 |
| subsample | 0.8, 1.0 |
| Colsample_bytree | 0.8, 1.0 |

จากตารางที่ 4 Hybrid LSTM+XGBoost กำหนดค่า random_state = 42 และปรับพารามิเตอร์หลัก ได้แก่ learning rate (0.01–0.1), n_estimators (300–1000), min_child_weight (1-10), max_depth (3–8), subsample (0.8–1.0) และ colsample_bytree (0.8-1.0) เพื่อควบคุมความซับซ้อนและเพิ่มความแม่นยำของโมเดล

**ตารางที่ 5 :** ค่าพารามิเตอร์ของ Deep Q Learning

| Parameter | Value |
|---|---|
| Stockout_penalty | 10,15,20 |
| Holding_cost | 0.3, 0.5, 0.8 |
| Gamma | 0.95, 0.97, 0.99 |
| Hidden_units | [64,64], [128,64], [128, 128] |
| Epsilon_decay | 0.990, 0.9925, 0.995 |

จากตารางที่ 5 DQN กำหนดค่าคงที่ ได้แก่ Stockout_penalty (10-20), Holding_cost (0.3-0.8), Gamma (0.95-0.97), Hidden_units ([64,64]-[128,128]), Epsilon_decay (0.990-0.995) เพื่อเพิ่มประสิทธิภาพในการเรียนรู้เชิงลำดับ

## 4. ผลการดำเนินงาน

### 4.1 ผลการเก็บรวบรวมข้อมูล

หลังจากดำเนินการรวบรวมข้อมูลความต้องการสินค้าและข้อมูลสินค้าคงคลังจากบริษัท ของดี สโตร์ จำกัด ในช่วงเดือนมกราคมถึงธันวาคม พ.ศ. 2568 ผู้วิจัยได้ทำการคัดเลือกข้อมูลที่เหมาะสมสำหรับการวิเคราะห์ ซึ่งแสดงรายละเอียดดังตารางที่ 6

**ตารางที่ 6 :** ตัวอย่างข้อมูลสินค้าคงคลัง

| Date | Product_id | Demand | Tx_count | is_stockout | Restock qty | Restock count |
|---|---|---|---|---|---|---|
| 2025-09-03 | 1 | 30 | 12 | 1 | 24.0 | 2.0 |
| 2025-09-04 | 1 | 13 | 9 | 12 | 0.0 | 0.0 |
| 2025-09-05 | 1 | 19 | 10 | 0 | 24.0 | 2.0 |
| 2025-09-06 | 1 | 7 | 6 | 12 | 0 | 0.0 |
| 2025-09-07 | 1 | 28 | 14 | 12 | 0 | 24.0 |

### 4.2 ผลการจัดเตรียมข้อมูล

หลังจากจัดเตรียมข้อมูล ผู้วิจัยได้เติมเต็มช่องว่างของเวลา จัดการค่า Zero Demand และแปลงข้อมูลเป็นอนุกรมเวลาด้วย Sliding Window จากนั้นคัดกรองประเภทความต้องการด้วยเกณฑ์ ADI และ $CV^2$ ตามแนวทางของ Syntetos และ Boylan (2005) พบว่าสินค้าส่วนใหญ่อยู่ในกลุ่ม Intermittent และ Lumpy จึงคัดเลือกสินค้า 6 รายการที่มีข้อมูลครบถ้วน ได้ชุดข้อมูลพร้อมใช้งานจำนวน 2,190 รายการ

### 4.3 ผลการพัฒนาและประเมินประสิทธิภาพตัวแบบ

จากนั้นผู้วิจัยได้พัฒนาและประเมินตัวแบบจำนวน 4 รูปแบบ ได้แก่ LSTM, XGBoost และโมเดลแบบผสม โดยมีการกำหนดพารามิเตอร์ทั้งแบบคงที่และแบบปรับแต่งเพื่อเพิ่มประสิทธิภาพของตัวแบบ การเปรียบเทียบประสิทธิภาพพิจารณาจากค่า MAE, RMSE และ MAPE โดยผลลัพธ์ของแต่ละตัวแบบแสดงไว้ในตารางที่ 7

**ตารางที่ 7 :** ผลลัพธ์ของตัวแบบ LSTM

| Parameter | | | | Performance | | | |
|---|---|---|---|---|---|---|---|
| Unit | Layer | Drop-out | Batch-Size | MAE | RMSE | MAPE | $R^2$ |
| 128 | 2 | 16 | 0.2 | 4.9823 | 16.0024 | 78.7846 | 0.4044 |
| 64 | 1 | 128 | 0.1 | 4.9136 | 16.0726 | 77.3956 | 0.3992 |
| 128 | 2 | 128 | 0.2 | 4.8929 | 16.1353 | 75.9035 | 0.3945 |
| **64** | **2** | **128** | **0.1** | **4.8648** | **16.2675** | **75.2543** | **0.3846** |
| 64 | 1 | 16 | 0.1 | 5.0282 | 16.3257 | 75.0952 | 0.3801 |
| 128 | 2 | 128 | 0.1 | 5.0333 | 16.3722 | 77.1177 | 0.3766 |
| 64 | 2 | 32 | 0.1 | 5.1241 | 16.5742 | 76.0859 | 0.3611 |
| 64 | 1 | 32 | 0.1 | 5.2549 | 16.6190 | 79.0102 | 0.3577 |
| 128 | 1 | 16 | 0.3 | 5.1750 | 16.6302 | 76.5551 | 0.3568 |
| 64 | 1 | 64 | 0.1 | 5.0032 | 16.6540 | 74.0178 | 0.3550 |

จากตารางที่ 7 แบบจำลอง LSTM ที่ให้ค่า MAE ต่ำที่สุดมี โดยมีค่า MAE เท่ากับ 4.8648 ค่า RMSE เท่ากับ 16.2675 ค่า MAPE เท่ากับ 75.2543 และค่า $R^2$ เท่ากับ 0.3846

**ตารางที่ 8 :** ผลลัพธ์ของตัวแบบ XGBoost

| Parameter | | | | | | Performance | | | |
|---|---|---|---|---|---|---|---|---|---|
| NT | MD | LR | SS | CS | MW | MAE | RMSE | MAPE | $R^2$ |
| 500 | 8 | 0.1 | 1 | 1 | 5 | 3.6624 | 9.9503 | 107.2682 | 0.7249 |
| 1000 | 8 | 0.1 | 1 | 1 | 5 | 3.6624 | 9.9503 | 107.2682 | 0.7249 |
| 300 | 8 | 0.1 | 1 | 1 | 5 | 3.6624 | 9.9503 | 107.2682 | 0.7249 |
| 300 | 6 | 0.05 | 1 | 1 | 5 | 3.6926 | 9.9953 | 105.6227 | 0.7224 |
| 500 | 6 | 0.05 | 1 | 1 | 5 | 3.6926 | 9.9953 | 105.6227 | 0.7224 |
| 1000 | 6 | 0.05 | 1 | 1 | 5 | 3.6926 | 9.9953 | 105.6227 | 0.7224 |
| 1000 | 4 | 0.05 | 0.8 | 0.8 | 10 | 3.6838 | 10.0205 | 102.7328 | 0.7210 |
| 500 | 4 | 0.05 | 0.8 | 0.8 | 10 | 3.6838 | 10.0205 | 102.7328 | 0.7210 |
| **1000** | **6** | **0.1** | **1** | **0.8** | **1** | **3.6281** | **10.0308** | **103.4645** | **0.7204** |
| 500 | 6 | 0.1 | 1 | 0.8 | 1 | 3.6281 | 10.0308 | 103.4645 | 0.7204 |

จากตารางที่ 8 แบบจำลอง XGBoost ให้ประสิทธิภาพดีที่สุดด้วยค่า MAE เท่ากับ 3.6281 RMSE เท่ากับ 10.0308 MAPE เท่ากับ 103.4645 $R^2$ เท่ากับ 0.7204

**ตารางที่ 9 :** ผลลัพธ์ของตัวแบบผสม LSTM-XGBoost

| Parameter | | | | | | Performance | | | |
|---|---|---|---|---|---|---|---|---|---|
| NT | MD | LR | SS | MW | CB | MAE | RMSE | MAPE | $R^2$ |
| **500** | **6** | **0.1** | **0.8** | **1** | **1** | **3.6076** | **9.9295** | **85.2188** | **0.7707** |
| 500 | 4 | 0.1 | 0.8 | 1 | 1 | 3.7203 | 9.9333 | 87.9894 | 0.7705 |
| 300 | 4 | 0.1 | 0.8 | 1 | 1 | 3.7203 | 9.9333 | 87.9894 | 0.7705 |
| 1000 | 4 | 0.1 | 0.8 | 1 | 1 | 3.7203 | 9.9333 | 87.9894 | 0.7705 |
| 300 | 6 | 0.1 | 0.8 | 5 | 1 | 3.5939 | 9.9457 | 85.7316 | 0.7700 |
| 500 | 6 | 0.1 | 0.8 | 5 | 1 | 3.5939 | 9.9457 | 85.7316 | 0.7700 |
| 1000 | 6 | 0.1 | 0.8 | 5 | 1 | 3.5939 | 9.9457 | 85.7316 | 0.7700 |
| 500 | 4 | 0.1 | 1 | 10 | 0.8 | 3.7111 | 9.9469 | 85.7282 | 0.7699 |
| 1000 | 4 | 0.1 | 1 | 10 | 0.8 | 3.7111 | 9.9469 | 85.7282 | 0.7699 |
| 300 | 4 | 0.1 | 1 | 10 | 0.8 | 3.7111 | 9.9469 | 85.7282 | 0.7699 |

จากตารางที่ 9 แบบจำลอง Hybrid LSTM-XGBoost ให้ผลดีที่สุด โดยมีค่า MAE เท่ากับ 3.6076 ค่า RMSE เท่ากับ 9.9295 ค่า MAPE เท่ากับ 85.2188 และค่า $R^2$ เท่ากับ 0.7707

**ตารางที่ 10 :** ผลการเปรียบเทียบประสิทธิภาพของแบบจำลอง

| Technique | Performance | | | |
|---|---|---|---|---|
| | MAE | RMSE | MAPE | R2 |
| LSTM | 4.8648 | 16.2675 | 75.2543 | 0.3846 |
| XGBoost | 3.6281 | 10.0308 | 103.4645 | 0.7204 |
| XGBoost+LSTM | 3.5307 | 10.0501 | 84.4048 | 0.7651 |
| **LSTM+XGBoost** | **3.6076** | **9.9295** | **85.2188** | **0.7707** |

จากตารางที่ 10 LSTM-XGBoost ให้ผลดีที่สุดโดยมีค่า RMSE ต่ำสุดที่ 9.9295 และ $R^2$ สูงสุดที่ 0.7707 สะท้อนความสามารถในการลดความคลาดเคลื่อนขนาดใหญ่และอธิบายความแปรปรวนของข้อมูลได้ดี แม้ XGBoost-LSTM จะให้ค่า MAE และ MAPE ต่ำกว่าเล็กน้อย อย่างไรก็ตาม ค่า MAPE ที่สูงควรตีความร่วมกับข้อจำกัดของข้อมูลแบบ Intermittent และ Lumpy Demand

**ตารางที่ 11 :** ผลการเรียนรู้ของ Deep Q-Network

| No | Policy Type | Avg Total Reward | Avg Stockout % | Avg Final Stockout % | Products Evaluated |
|---|---|---|---|---|---|
| 1 | DQN | -1,166.55 | 0.35% | 34.22 | 504 |
| 2 | Always Order | -11,651.88 | 0.00% | 498.71 | 504 |
| 3 | Never Order | -1,298.22 | 6.31% | 25.70 | 504 |

จากการทดสอบกับสินค้า 504 รายการ DQN ให้ค่า Avg Total Reward ที่ -1,166.55 และ Avg Stockout Rate ร้อยละ 0.35 ดีกว่านโยบาย Never Order ที่มีค่า -1,298.22 และ Avg Stockout Rate ร้อยละ 6.31 ขณะที่ Always Order แม้ไม่มีการขาดสต็อกแต่มีค่า Avg Final Stockout สูงถึง 498.71 สะท้อนให้เห็นว่า DQN สามารถสร้างสมดุลระหว่างการลดการขาดสต็อกและการควบคุมสินค้าคงคลังส่วนเกินได้ดีกว่าทั้งสองนโยบายพื้นฐาน

## 5. สรุป

งานวิจัยนี้พัฒนาเฟรมเวิร์กการพยากรณ์การเติมสินค้าโดยใช้เทคนิคเหมืองข้อมูลและการเรียนรู้ของเครื่อง ครอบคลุมการจัดเตรียมข้อมูล การจัดการ Zero Demand และการแปลงข้อมูลเป็นอนุกรมเวลาด้วย Sliding Window โดยเปรียบเทียบตัวแบบ LSTM, XGBoost, XGBoost-LSTM และ LSTM-XGBoost ผลการทดลองพบว่า LSTM-XGBoost มีประสิทธิภาพดีที่สุด (RMSE = 9.9295, $R^2$ = 0.7707) ยืนยันว่าตัวแบบผสมพยากรณ์ได้แม่นยำกว่าตัวแบบเดี่ยวในส่วนการตัดสินใจเติมสินค้า DQN ให้ค่า Avg Total Reward ที่ -1,166.55 และอัตราสินค้าขาดคลังเฉลี่ยร้อยละ 0.35 จากการทดสอบกับสินค้า 504 รายการ โดย DQN ลดการขาดสต็อกได้ดีกว่า Never Order และควบคุมสินค้าคงคลังส่วนเกินได้ดีกว่า Always Order สะท้อนให้เห็นถึงศักยภาพในการสนับสนุนการตัดสินใจเติมสินค้าได้อย่างมีประสิทธิภาพ

อย่างไรก็ตาม งานวิจัยนี้มีข้อจำกัดด้านจำนวนสินค้าที่ใช้ทดลอง ความแตกต่างของจำนวนสินค้าระหว่างสองส่วนการทดลอง และลักษณะข้อมูลแบบ Intermittent และ Lumpy ที่ส่งผลต่อค่า MAPE การวิจัยในอนาคตควรเพิ่มจำนวน SKU เชื่อมผลพยากรณ์กับ DQN อย่างเป็นระบบ และเปรียบเทียบกับตัวแบบ เช่น Croston, SBA หรือ TSB พร้อมประเมินร่วมกับตัวชี้วัดเชิงธุรกิจ

## เอกสารอ้างอิง


[1] I. H. Witten, E. Frank, M. A. Hall, and C. J. Pal, Data Mining: Practical Machine Learning Tools and Techniques, 4th ed. Burlington, MA, USA: Morgan Kaufmann, 2016.

[2] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521, no. 7553, pp. 436–444, May 2015, doi: 10.1038/nature14539.

[3] C. Olah, "Understanding LSTM networks," Aug. 2015. [Online]. Available: https://colah.github.io/posts/2015-08-Understanding-LSTMs/

[4] J. H. Friedman, "Greedy function approximation: A gradient boosting machine," Annals of Statistics, vol. 29, no. 5, pp. 1189–1232, Oct. 2001, doi: 10.1214/aos/1013203451.

[5] H. Van Hasselt, A. Guez, and D. Silver, "Deep reinforcement learning with double Q-learning," in

Proc. 30th AAAI Conf. Artificial Intelligence, Phoenix, AZ, USA, 2016, vol. 30, no. 1.

[6] X. Zhang, "Financial time series forecasting based on LSTM neural network optimized by wavelet denoising and whale optimization algorithm," Acad. J. Comput. Inf. Sci., vol. 5, no. 7, pp. 1–9, 2022, doi: 10.25236/AJCIS.2022.050701. [Online]. Available: https://francis-press.com/papers/7189

[7] P. Dankorpho, "Sales forecasting for retail business using XGBoost algorithm," J. Comput. Sci. Technol. Stud., vol. 6, no. 2, pp. 136–141, Jun. 2024, doi: 10.32996/jcsts.2024.6.2.15. [Online]. Available: https://al-kindipublishers.org/index.php/jcsts/article/view/7381

[8] V. Mnih et al., "Human-level control through deep reinforcement learning," Nature, vol. 518, no. 7540, pp. 529–533, Feb. 2015, doi: 10.1038/nature14236.

# การเปรียบเทียบแบบจำลองการเรียนรู้ของเครื่องจักรเพื่อคาดการณ์การทะลุราคาที่ผิดพลาดในตลาดหุ้นโดยใช้คุณสมบัติทางเทคนิค

# A Comparison of Machine Learning Models for Predicting False Breakouts in Stock Markets Using Technical Indicators

*อภิรัตน์ จิรังคพาณิชย์ (Apirat Jirangkapanich)[1] และนลินภัสร์ บำเพ็ญเพียร (Nalinpat Bhumpenpein)[2]*
*ภาควิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล*
*มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ*
*[1]s6707011857012@email.kmutnb.ac.th, [2]nalinpat.b@itd.kmutnb.ac.th*

**บทคัดย่อ**

*งานวิจัยนี้มีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพของแบบจำลองการเรียนรู้ของเครื่อง (Machine Learning) จำนวน 4 รูปแบบ ได้แก่ Random Forest, XGBoost, LSTM และ TCN ในการคาดการณ์เหตุการณ์การทะลุราคาที่ผิดพลาด (False Breakout) ซึ่งเป็นกับดักสำคัญสำหรับนักลงทุนทางเทคนิค โดยทำการทดสอบกับข้อมูลหลักทรัพย์กลุ่ม SET50 จำนวน 4 หลักทรัพย์ (ADVANC, AOT, PTT และ CPALL) ในช่วงปี พ.ศ. 2557–2567 โดยใช้ชุดข้อมูลประวัติการซื้อขายรายวันจำนวนทั้งหมด 10,000 รายการ โดยในแต่ละหลักทรัพย์จะมีชุดข้อมูล 2700 รายการ โดยมีอัตราส่วนการเกิด True Breakout ต่อ False Breakout อยู่ที่ประมาณ 10:90 ผู้วิจัยได้ทำการแบ่งชุดข้อมูลสำหรับฝึกสอน ตรวจสอบและทดสอบ ในสัดส่วน 70:15:15 ตามลำดับเวลา (Chronological Split) เพื่อป้องกันข้อมูลรั่วไหล (Data Leakage) ผลการทดลองพบว่าแบบจำลองกลุ่ม Ensemble Learning (Random Forest และ XGBoost) มีประสิทธิภาพดีกว่ากลุ่ม Deep Learning (LSTM และ TCN) อย่างมีนัยสำคัญ อันเนื่องมาจากข้อจำกัดด้านขนาดของข้อมูลและลักษณะข้อมูลแบบตาราง โดย Random Forest ให้ค่า F1-Score เฉลี่ยสูงสุดที่ 0.72 และสามารถรักษาสมดุลระหว่างค่า Precision และ Recall ได้ดีที่สุด ปัจจัยที่มีผลต่อการทำนายสูงสุดคือสัดส่วนปริมาณการซื้อขาย (Volume Ratio) ซึ่งยืนยันสมมติฐานความสัมพันธ์ระหว่างปริมาณการซื้อขายกับความถูกต้องของแนวโน้มราคา*

**คำสำคัญ:** การทะลุราคาหลอก การเรียนรู้ของเครื่อง ข้อมูลไม่สมดุล แรนดอมฟอเรสต์ การวิเคราะห์ทางเทคนิค

**Abstract**

*This research aims to compare the performance of four machine learning models: Random Forest, XGBoost, LSTM, and TCN, in predicting false breakout events, a significant trap for technical investors. The tests were conducted using data from four SET50 securities (ADVANC, AOT, PTT, and CPALL) from 2014 to 2024, with a total of 10,000 daily trading history records (2,700 records for each security). The data was split chronologically in a 70:15:15 ratio for training, validation, and testing to prevent data leakage. The results showed that the Ensemble Learning models (Random Forest and XGBoost) significantly outperformed the Deep Learning models (LSTM and TCN). Due to limitations in data size and the tabular structure of the data, Random Forest yields the highest average F1-Score at 0.72 and best balances Precision and Recall. The factor with the greatest predictive impact is the trading volume ratio, which confirms the hypothesis of a relationship between trading volume and the accuracy of price trends.*

**Keywords:** False Breakout, Machine Learning, Imbalanced Data, Random Forest, Technical Analysis.

## 1. บทนำ

การลงทุนในตลาดหลักทรัพย์มีความผันผวนและความซับซ้อนสูง กลยุทธ์การลงทุนแบบ Breakout Trading หรือการซื้อเมื่อราคาทะลุแนวต้าน เป็นหนึ่งในกลยุทธ์ที่ได้รับความนิยม แต่ปัญหาสำคัญที่บั่นทอนประสิทธิภาพของกลยุทธ์นี้คือปรากฏการณ์การทะลุราคาหลอก (False Breakout) ซึ่งเกิดขึ้นเมื่อราคาทะลุแนวต้านได้เพียงชั่วครู่ก่อนจะกลับตัวลงอย่างรวดเร็ว ส่งผลให้นักลงทุนประสบภาวะขาดทุน [1] เครื่องมือวิเคราะห์ทางเทคนิคแบบดั้งเดิมมักมีความล่าช้าและไม่เพียงพอที่จะแยกแยะสัญญาณหลอกได้แม่นยำ [2] งานวิจัยนี้จึงมุ่งเน้นการประยุกต์ใช้เทคโนโลยีการเรียนรู้ของเครื่อง (Machine Learning) เพื่อแก้ปัญหานี้ โดยเปรียบเทียบระหว่างแบบจำลองกลุ่ม Feature-based (Random Forest และ XGBoost) และกลุ่ม Time-series (LSTM และ TCN) เพื่อค้นหาแบบจำลองที่เหมาะสมที่สุดสำหรับบริบทตลาดหุ้นไทย

## 2. วรรณกรรมที่เกี่ยวข้อง

**2.1**การวิเคราะห์ทางเทคนิคและการทะลุราคาหลอก

กลยุทธ์การซื้อขายแบบ Breakout อาศัยแนวคิดพื้นฐานจากทฤษฎีแนวรับ-แนวต้าน (Support and Resistance) เป็นระดับราคาที่มีอุปสงค์หรืออุปทานหนาแน่น การทะลุผ่านระดับดังกล่าวถือเป็นสัญญาณการเริ่มต้นของแนวโน้มใหม่ [3] อย่างไรก็ตาม ปัญหา False Breakoutหรือกับดักราคา ซึ่งเกิดจากความผันผวนของตลาดและสัญญาณรบกวน (Noise) โดยเสนอว่าการพิจารณาเพียงแค่ราคา (Price Action) ไม่เพียงพอ แต่จำเป็นต้องใช้ปริมาณการซื้อขาย (Volume) เป็นตัวยืนยัน (Confirmation) ความถูกต้องของแนวโน้ม [2] ซึ่งสอดคล้องกับสมมติฐานของงานวิจัยนี้ที่ใช้ Volume Ratio เป็นฟีเจอร์หลัก

**2.2 การเรียนรู้ของเครื่องกับข้อมูลทางการเงิน**

การประยุกต์ใช้ Machine Learning ในการทำนายตลาดหุ้นแบ่งเป็น 2 กลุ่มหลัก ได้แก่ กลุ่มที่ 1 Feature-based งานวิจัยของ Liu et al. [4] พบว่าอัลกอริทึม XGBoost และ Random Forest มีประสิทธิภาพสูงในการจัดการกับข้อมูลตลาดหุ้นที่มีความไม่เป็นเส้นตรง (Non-linear) และมีสัญญาณรบกวนสูง โดยจุดเด่นคือความสามารถในการคัดเลือกคุณลักษณะ (Feature Selection) และลดความเสี่ยงของการเรียนรู้เกิน (Overfitting) ในข้อมูลแบบตาราง และกลุ่มที่ 2 Deep Learning โดย Fischer & Krauss [5] ได้นำเสนอการใช้ Long Short-Term Memory (LSTM) ซึ่งเป็นโครงข่ายประสาทเทียมแบบ Recurrent Neural Network (RNN) โดยพบว่าสามารถจับรูปแบบความสัมพันธ์ในลำดับเวลา (Temporal Dependencies) ได้ดีกว่าแบบจำลองทางสถิติทั่วไป นอกจากนี้ ยังมีการพัฒนา Temporal Convolutional Network (TCN) ซึ่งใช้ Dilated Convolutions เพื่อขยาย Receptive Field ให้สามารถเรียนรู้ข้อมูลย้อนหลังได้ยาวนานขึ้นและมีประสิทธิภาพในการคำนวณสูงกว่า RNN [6]

ในส่วนของการตรวจจับเหตุการณ์ทะลุราคาหลอก (False Breakout Detection) งานวิจัยในช่วง 5 ปีที่ผ่านมาได้เริ่มนำเทคนิคการเรียนรู้ของเครื่องมาใช้มากขึ้น เช่น ที่การใช้ตัวชี้วัดทางเทคนิคร่วมกับปริมาณการซื้อขายสามารถลดอัตราการเกิดสัญญาณหลอกในตลาดหุ้นเอเชียได้อย่างมีนัยสำคัญ [7] นอกจากนี้ ในด้านการประยุกต์ใช้แบบจำลองโครงข่ายประสาทเทียมแบบสังวัตนาการเชิงเวลา (TCN) ในบริบททางการเงิน (Financial Context) งานวิจัยของ Chen et al. [8] ได้พิสูจน์ว่า TCN มีข้อได้เปรียบเหนือ LSTM ในด้านความเร็วในการประมวลผลและการหลีกเลี่ยงปัญหา Vanishing Gradient เมื่อทำการวิเคราะห์ข้อมูลอนุกรมเวลาทางการเงินที่มีความถี่สูง อย่างไรก็ตาม การนำ TCN มาใช้ดักจับ False Breakout ในข้อมูลระดับวัน (Daily Data) ยังคงเป็นความท้าทายใหม่ที่งานวิจัยนี้มุ่งเน้นศึกษา

**2.3 การจัดการข้อมูลไม่สมดุล (Imbalanced Data)**

เหตุการณ์ False Breakout ถือเป็นเหตุการณ์หายาก (Rare Event) ในข้อมูลอนุกรมเวลา ซึ่ง King & Zeng [9] ระบุว่าการใช้ข้อมูลที่ไม่สมดุลในการฝึกสอนแบบจำลองมักส่งผลให้แบบจำลองมีความลำเอียง ไปทางคลาสส่วนใหญ่ (Majority Class) งานวิจัยนี้จึงประยุกต์ใช้เทคนิค SMOTE (Synthetic Minority Over-sampling Technique) [10] เพื่อสังเคราะห์ข้อมูลเทียมในคลาสส่วนน้อย ซึ่งช่วยเพิ่มประสิทธิภาพค่า Recall ให้กับแบบจำลองได้อย่างมีนัยสำคัญ

## 3. วิธีการดำเนินการวิจัย

งานวิจัยนี้มีกรอบการดำเนินงาน 4 ขั้นตอนหลัก ได้แก่ การรวบรวมข้อมูล การสกัดคุณลักษณะ การจัดการข้อมูลไม่สมดุลด้วย SMOTE และการประเมินแบบจำลอง

ข้อมูลการซื้อขายรายวันของหุ้น ADVANC, AOT, PTT และ CPALL ช่วงปี พ.ศ. 2557–2567 (10 ปี) ซึ่งเป็นตัวแทนจาก 4 กลุ่มอุตสาหกรรม เพื่อเพิ่มความสามารถในการสรุปนัยทั่วไป (Generalization) และพิสูจน์ความทนทานของแบบจำลองต่อสภาพตลาดโดยรวม [11] โดยนำตัวชี้วัดทางเทคนิค 14 ตัวแปร ได้แก่ RSI, MACD, ATR, Volume Ratio, OBV, Volume, Vol_MA20, Close, Return 1 Day, Return 5 Day, Return 20 Day, Upper Shadow, Lower Shadow และ Candle Body เพื่อใช้เป็นฟีเจอร์ และทำการกำหนดป้ายกำกับข้อมูล (Data Labeling) โดยแบ่งเหตุการณ์เป็น True Breakout และ False Breakout

อีกทั้ง งานวิจัยนี้ได้กำหนดเงื่อนไขทางเทคนิค เพื่อจำแนกเหตุการณ์การทะลุราคา (Breakout Events) โดยแบ่งเป็น 2 ขั้นตอน ได้แก่ ขั้นตอนที่ 1 การระบุจุดทะลุราคา ( Breakout Identification ) โดยกำหนดให้สัญญาณการทะลุราคาเกิดขึ้นเมื่อราคาปิดของวันปัจจุบันปรับตัวสูงกว่าระดับราคาสูงสุด ที่ทำไว้ในช่วงระยะเวลาย้อนหลัง 20 วันทำการ ซึ่งเป็นการยืนยันว่าราคาสามารถทำจุดสูงสุดใหม่ ในรอบเดือนได้สำเร็จ และขั้นตอนที่ 2 การจำแนกประเภทการทะลุ (Event Classification) โดยหลังจากเกิดสัญญาณการทะลุราคา จะทำการตรวจสอบแนวโน้มราคาในช่วงระยะเวลาการถือครอง (Holding Period) 5 วันทำการถัดไป เพื่อระบุประเภทของเหตุการณ์ เหตุผลในการกำหนดเงื่อนไขจุดสูงสุดในรอบ 20 วัน (20-Day High) และระยะเวลาถือครอง 5 วัน (5-Day Holding Period) อ้างอิงจากทฤษฎีการวิเคราะห์ทางเทคนิคตามระบบเทรดแบบเต่า (Turtle Trading System) โดย 20 วันทำการมีค่าเทียบเท่ากับหนึ่งเดือนปฏิทิน และ 5 วันทำการเทียบเท่ากับหนึ่งสัปดาห์ทำการ ซึ่งเป็นกรอบเวลาที่เหมาะสมและเพียงพอต่อการแยกแยะความผันผวนและสัญญาณหลอก (Whipsaw) ออกจากแนวโน้มหลักของราคา [12] ดังนี้

False Breakout (Class 1) คือ เหตุการณ์ที่ราคาไม่สามารถยืนเหนือระดับแนวต้านได้ โดยราคาปรับตัวลดลงต่ำกว่าจุดสูงสุดเดิมหรือให้ผลตอบแทนเป็นลบเมื่อสิ้นสุดระยะเวลาการถือครอง

True Breakout (Class 0) คือ เหตุการณ์ที่ราคาสามารถรักษาระดับแนวโน้มขาขึ้นไว้ได้ โดยราคายังคงยืนเหนือระดับแนวต้านและให้ผลตอบแทนเป็นบวกเมื่อสิ้นสุดระยะเวลาการถือครอง

สำหรับการจัดการข้อมูลไม่สมดุล งานวิจัยนี้ใช้วิธี Synthetic Minority Over-sampling Technique (SMOTE) เพื่อสร้างสมดุลให้กับชุดข้อมูลฝึกสอน ในการฝึกสอนแบบจำลอง ผู้วิจัยได้หลีกเลี่ยงการสุ่มแบ่งข้อมูลแบบปกติเพื่อป้องกันปัญหาข้อมูลอนาคตรั่วไหล (Data Leakage / Look-ahead Bias) โดยใช้วิธีแบ่งข้อมูลตามลำดับเวลา (Chronological Split) ด้วยสัดส่วนร้อยละ 70 สำหรับฝึกสอน ร้อยละ 15 สำหรับตรวจสอบและร้อยละ 15 สำหรับทดสอบ พร้อมทั้งปรับแต่งพารามิเตอร์ (Hyperparameter Tuning) ด้วยวิธี Grid Search เช่น การจำกัดความลึกของต้นไม้ใน Random Forest และการประยุกต์ใช้เลเยอร์การสุ่มทิ้ง (Dropout) ที่ร้อยละ 20 ควบคู่กับระบบหยุดการเรียนรู้ก่อนกำหนด (Early Stopping) ในกลุ่ม Deep Learning (LSTM และ TCN) เพื่อป้องกันสภาวะการเรียนรู้เกินพอดี [13] และใช้ค่า F1-Score เป็นเกณฑ์หลักควบคู่ไปกับ Precision, Recall และ ROC-AUC ในการประเมินคุณภาพของแบบจำลอง เนื่องจากข้อมูลมีความไม่สมดุล

## 4. ผลการดำเนินงานวิจัย

### 4.1 ประสิทธิภาพโดยรวมของแบบจำลอง

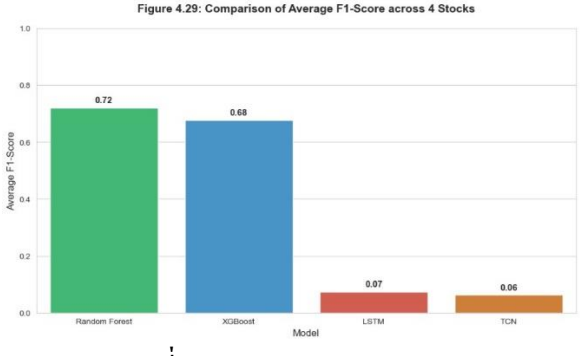
การทดสอบประสิทธิภาพของแบบจำลองทั้ง 4 รูปแบบกับชุดข้อมูลทดสอบ แสดงผลลัพธ์ดังตารางที่ 1

**ตารางที่ 1** เปรียบเทียบค่าเฉลี่ย F1-score ของแต่ละแบบจำลอง

| Model | Random Forest | XGBoost | LSTM | TCN |
|---|---|---|---|---|
| F1-score | 0.72 | 0.68 | 0.07 | 0.06 |
| Precision | 0.68 | 0.65 | 0.05 | 0.04 |
| Recall | 0.75 | 0.72 | 0.12 | 0.10 |
| ROC-AUC | 0.98 | 0.98 | 0.07 | 0.52 |

ผลการทดลองเชิงประจักษ์ชี้ให้เห็นว่า แบบจำลองในกลุ่ม Ensemble Learning (Random Forest และ XGBoost) มีประสิทธิภาพดีกว่ากลุ่ม Deep Learning (LSTM และ TCN) อย่างมีนัยสำคัญในทุกดัชนีชี้วัด โดยเฉพาะค่า F1-Score ซึ่ง Random Forest ทำได้สูงสุดที่ 0.72 สะท้อนถึงความสามารถในการจัดการกับความไม่สมดุลของข้อมูล (Imbalanced Data) ได้ดีกว่า สอดคล้องกับงานวิจัยของ Liu et al. [4] ที่พบว่าอัลกอริทึมกลุ่ม Tree-based มีความเหมาะสมสูงสำหรับข้อมูลตลาดหุ้นที่มีสัญญาณรบกวน และความไม่เป็นเส้นตรง

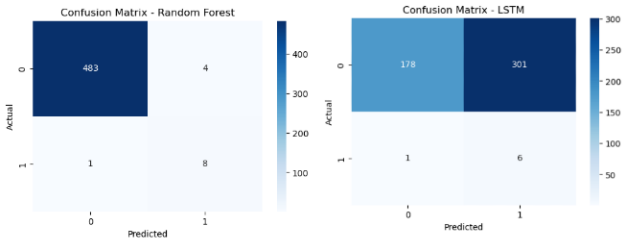
ในทางตรงกันข้าม แบบจำลอง Deep Learning ประสบปัญหาในการเรียนรู้เหตุการณ์หายาก (Rare Events) อย่างรุนแรง สังเกตได้จากค่า Recall ที่ต่ำมาก (0.10 - 0.12) ซึ่งหมายความว่าแบบจำลองมักทำนายผิดพลาดโดยมองข้ามสัญญาณ False Breakout ไปเกือบทั้งหมด ปัญหานี้สอดคล้องกับ King & Zeng [7] ที่ระบุว่าโครงข่ายประสาทเทียมมักมีความลำเอียงไปทางคลาสส่วนใหญ่เมื่อปริมาณข้อมูลเหตุการณ์เป้าหมายมีจำกัด




**ภาพที่ 1:** กราฟเปรียบเทียบ F1- score

### 4.2 การวิเคราะห์ความผิดพลาดผ่านตารางความสับสน

เพื่อให้เห็นภาพความสามารถในการแยกแยะคลาสได้ชัดเจนยิ่งขึ้น ผู้วิจัยได้เปรียบเทียบตารางความสับสน (Confusion Matrix) ระหว่างแบบจำลองที่ดีที่สุด )Random Forest) และแบบจำลองที่มีประสิทธิภาพต่ำกว่า )LSTM)




**ภาพที่ 2:** การเปรียบเทียบ Confusion Matrix ระหว่าง Random Forest (ซ้าย) และ LSTM (ขวา)

จากการวิเคราะห์ภาพที่ 2 พบว่า Random Forest สามารถระบุเหตุการณ์ False Breakout (Class 1) ได้อย่างถูกต้องเป็นจำนวนมาก (High True Positive) ในขณะที่ LSTM มักทำนายผิดเป็น True Breakout (False Negative สูง) ซึ่งในทางปฏิบัติ ความผิดพลาดของ LSTM ถือเป็นความเสี่ยงสูงต่อนักลงทุน เพราะอาจนำไปสู่การเข้าซื้อในจังหวะที่เป็นกับดักราคา

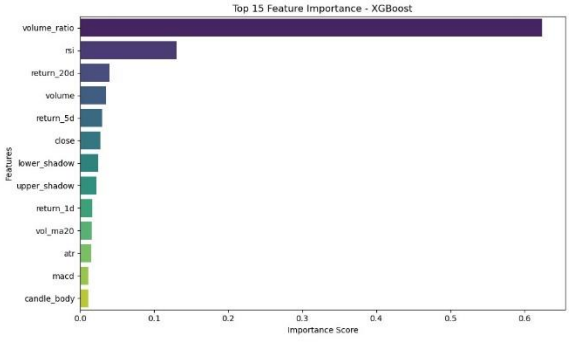
### 4.3 การวิเคราะห์เสถียรภาพของการเรียนรู้

เสถียรภาพของการเรียนรู้ (Learning Stability) ยังเป็นอีกหนึ่งปัจจัยสำคัญ จากการสังเกตการเปลี่ยนแปลงของค่าความสูญเสีย (Loss) ระหว่างการฝึกสอน พบว่าแบบจำลอง LSTM และ TCN มีสัญญาณของการเรียนรู้เกิน (Overfitting) โดยค่า Validation Loss มีแนวโน้มแยกตัวออกจาก Training Loss และมีความผันผวนสูง (High Variance) ซึ่งเป็นข้อจำกัดของการนำ Deep Learning มาใช้กับข้อมูลรายวัน (Daily Timeframe) ที่มีจำนวนจุดข้อมูลจำกัดเมื่อเทียบกับความซับซ้อนของแบบจำลอง

### 4.4 การวิเคราะห์ความสำคัญของคุณลักษณะ (Feature Importance Analysis)

การระบุปัจจัยที่ส่งผลต่อการเกิด False Breakout เป็นอีกหนึ่งวัตถุประสงค์หลักของงานวิจัยนี้ ภาพที่ 2 แสดง

ลำดับความสำคัญของคุณลักษณะ (Feature Importance) ของแบบจำลอง XGBoost




**ภาพที่ 3:** Feature Importance ของ AOT

ผลการวิเคราะห์มีความสอดคล้องกันในทุกหลักทรัพย์ โดยพบว่า สัดส่วนปริมาณการซื้อขาย )Volume Ratio) เป็นตัวแปรที่มีค่าความสำคัญสูงสุด ซึ่งเป็นการยืนยันสมมติฐานตามทฤษฎี Dow Theory และแนวคิดของ Durgia [2] ที่กล่าวว่า "Volume must confirm the trend" กล่าวคือ การทะลุแนวต้านที่ขาดแรงสนับสนุนจากปริมาณการซื้อขายที่หนาแน่น มีโอกาสสูงที่จะเป็นสัญญาณหลอก นอกจากนี้ ดัชนีชี้วัดความผันผวน )Volatility Indicators) เช่น ATR และ Bollinger Bands เป็นต้น ยังมีบทบาทสำคัญรองลงมาในการช่วยกรองสัญญาณรบกวนของตลาด

## 5. สรุปผลและอภิปรายผล

ผลการงานวิจัยนี้ได้ศึกษาเปรียบเทียบประสิทธิภาพของแบบจำลองการเรียนรู้ของเครื่องเพื่อคาดการณ์การทะลุราคาที่ผิดพลาด (False Breakout) กับข้อมูลหลักทรัพย์ในกลุ่ม SET50 จำนวน 4 หลักทรัพย์ ผลการทดลองเชิงประจักษ์ชี้ให้เห็นว่า แบบจำลองในกลุ่ม Ensemble Learning ได้แก่ Random Forest และ XGBoost มีประสิทธิภาพเหนือกว่าแบบจำลองในกลุ่ม Deep Learning (LSTM และ TCN) อย่างมีนัยสำคัญในทุกดัชนีชี้วัด โดยแบบจำลอง Random Forest ให้ค่า F1-Score เฉลี่ยสูงสุดที่ 0.72 และสามารถรักษาสมดุลระหว่างค่าความแม่นยำ (Precision) และความครบถ้วน (Recall) ได้ดีที่สุด ซึ่งสอดคล้องกับงานวิจัยของ Liu et al. [4] ที่พบว่าอัลกอริทึมตระกูล Tree-based มีความเหมาะสมสูงสำหรับข้อมูลตลาดหุ้นที่มีสัญญาณรบกวนและความไม่เป็นเส้นตรง ในทางตรงกันข้าม แบบจำลอง Deep Learning ประสบปัญหาในการเรียนรู้เหตุการณ์หายาก (Rare Events) ส่งผลให้ค่า Recall ต่ำมาก ซึ่งเป็นไปตามข้อสังเกตของ King & Zeng [9] ที่ระบุว่าโครงข่ายประสาทเทียมมักมีความลำเอียงไปทางคลาสส่วนใหญ่เมื่อปริมาณข้อมูลเหตุการณ์เป้าหมายมีจำกัด แม้ว่าจะมีการประยุกต์ใช้เทคนิค SMOTE [10] เพื่อแก้ปัญหาความไม่สมดุลของข้อมูลแล้วก็ตาม แต่โครงสร้างที่ซับซ้อนของ LSTM และ TCN ยังคงต้องการปริมาณข้อมูลขนาดใหญ่กว่าที่มีอยู่เพื่อการเรียนรู้ที่มีประสิทธิภาพ อีกทั้ง ลักษณะของข้อมูลที่สกัดมาเป็นค่าสัญญาณทางเทคนิคมีรูปแบบเป็นตาราง (Tabular Data) ซึ่งงานวิจัยในยุคหลังยืนยันตรงกันว่า อัลกอริทึมพื้นฐานแบบต้นไม้ตัดสินใจ (Tree-based Models) ยังคงให้ผลลัพธ์ที่เสถียรและทนทานต่อสัญญาณรบกวน (Noise) มากกว่าโครงข่ายประสาทเทียมในการจัดการกับข้อมูลแบบตารางขนาดกลางในตลาดหุ้น การที่กลุ่ม Deep Learning ทำผลงานได้ต่ำกว่าจึงมิใช่ความผิดพลาดของระบบ แต่เป็นข้อพิสูจน์เชิงประจักษ์ว่าอัลกอริทึมกลุ่ม Tree-based มีความเหมาะสมกับบริบทของข้อมูลทางการเงินลักษณะนี้มากกว่า [14] การวิเคราะห์ความสำคัญของคุณลักษณะ (Feature Importance) ยังค้นพบว่า "สัดส่วนปริมาณการซื้อขาย" (Volume Ratio) และ "ความผันผวน" (Volatility) เป็นปัจจัยที่มีอิทธิพลสูงสุดต่อความแม่นยำของแบบจำลองในทุกหลักทรัพย์ การค้นพบนี้ช่วยยืนยันองค์ความรู้ทางเทคนิคดั้งเดิมของ Murphy [3] ที่กล่าวว่า "Volume must confirm the trend" กล่าวคือ การทะลุแนวต้านที่มีคุณภาพจะต้องได้รับการสนับสนุนจากปริมาณการซื้อขายที่เพิ่มขึ้นอย่างมีนัยสำคัญ หากราคาปรับตัวขึ้นแต่ปริมาณการซื้อขายเบาบาง แบบจำลองจะเรียนรู้ว่าเป็นสัญญาณที่มีความเสี่ยงสูงที่จะเกิด False Breakout ซึ่งสอดคล้องกับลักษณะของกับดักราคาที่ Durgia [2] ได้อธิบายไว้โดยสรุป งานวิจัยนี้แสดงให้เห็นว่าสำหรับบริบทของตลาดหุ้นไทยที่ข้อมูลมีความผันผวนและไม่สมดุล แบบจำลอง Random Forest

เป็นเครื่องมือที่เหมาะสมที่สุดเนื่องจากมีความเสถียรและความสามารถในการตีความผลลัพธ์ที่ชัดเจน สำหรับการพัฒนาในอนาคต ผู้วิจัยเสนอแนะให้มีการศึกษาเพิ่มเติมโดยใช้ข้อมูลที่มีความถี่สูงขึ้น (High-frequency Data) เพื่อเพิ่มจำนวนจุดข้อมูลสำหรับฝึกสอนแบบจำลอง Deep Learning หรือทดลองใช้สถาปัตยกรรมแบบจำลองใหม่ เช่น Transformer ที่มีกลไก Attention เพื่อเพิ่มประสิทธิภาพในการจับความสัมพันธ์ของบริบทตลาดให้ดียิ่งขึ้น เป็นต้น

แม้ว่างานวิจัยนี้จะสามารถเปรียบเทียบประสิทธิภาพของแบบจำลองได้อย่างชัดเจน แต่ยังคงมีข้อจำกัดบางประการ ได้แก่ 1) ขอบเขตของข้อมูล (Data Scope) ที่ครอบคลุมเพียง 4 หลักทรัพย์ในกลุ่ม SET50 ซึ่งเป็นหุ้นขนาดใหญ่ที่มีสภาพคล่องสูง ผลลัพธ์ที่ได้อาจมีความคลาดเคลื่อนหากนำไปประยุกต์ใช้กับหุ้นขนาดกลางหรือขนาดเล็กที่มีพฤติกรรมการปั่นราคา (Price Manipulation) แตกต่างออกไป 2) ข้อจำกัดด้านคุณลักษณะ (Feature Limitations) โดยงานวิจัยนี้อิงตัวชี้วัดทางเทคนิคเป็นหลัก ไม่ได้นำปัจจัยพื้นฐานเชิงเศรษฐศาสตร์มหภาค (Macroeconomics) หรือการวิเคราะห์ความรู้สึกจากข่าวสาร (Sentiment Analysis) ซึ่งอาจมีอิทธิพลต่อการเกิด False Breakout เข้ามาร่วมพิจารณา และ 3) ขนาดของชุดข้อมูลสำหรับการฝึกสอน Deep Learning ที่มีจำกัดในระดับข้อมูลรายวัน (Daily Data) ส่งผลให้แบบจำลองไม่สามารถดึงประสิทธิภาพสูงสุดออกมาได้

## เอกสารอ้างอิง


[1] Ansary, A. (2023). *Breakout Trading Strategies: A Comprehensive Guide*. Journal of Finance and Investment, 15(2), 112-125.

[2] Durgia, R. (2025). "Understanding False Breakouts in Modern Markets". *Technical Analysis Journal*, 12(3), 45-58.

[3] Murphy, J. J. (1999). *Technical Analysis of the Financial Markets*. New York Institute of Finance.

[4] Liu, Y., et al. (2024). "Analysis and prediction of stock breakout based on XGBoost forward iterative model". *ISIJ International*, 64(8), 1618–1627.

[5] Fischer, T., & Krauss, C. (2018). "Deep learning with long short-term memory networks for financial market predictions". *European Journal of Operational Research*, 270(2), 654-669.

[6] Bai, S., Kolter, J. Z., & Koltun, V. (2018). "An empirical evaluation of generic convolutional and recurrent networks for sequence modeling". *arXiv preprint arXiv:1803.01271*.

[7] Zhang, Y., Li, X., & Wang, S. (2021). "Machine learning-based false breakout detection in Asian stock markets". *Journal of Financial Data Science*, 3(4), 112-128.

[8] Chen, H., Wu, Z., & Liu, M. (2023). "Temporal Convolutional Networks for financial time series forecasting: A comprehensive review". *Expert Systems with Applications*, 214, 119068.

[9] King, G., & Zeng, L. (2001). "Logistic regression in rare events data". *Political Analysis*, 9(2), 137–163.

[10] Chawla, N. V., et al. (2002). "SMOTE: synthetic minority over-sampling technique". *Journal of Artificial Intelligence Research*, 16, 321-357.

[11] Aronson, D. R. (2007). *Evidence-Based Technical Analysis: Applying the Scientific Method and Statistical Inference to Trading Signals*. John Wiley & Sons.

[12] Faith, C. (2007). *Way of the Turtle: The Secret Methods that Turned Ordinary People into Legendary Traders*. McGraw-Hill Education.

[13] Bergstra, J., & Bengio, Y. (2012). "Random search for hyper-parameter optimization". *Journal of Machine Learning Research*, 13, 281-305.

[14] Grinsztajn, L., Oyallon, E., & Varoquaux, G. (2022). "Why do tree-based models still outperform deep learning on typical tabular data?". *Advances in Neural Information Processing Systems*, 35, 507-520.

# การพยากรณ์การส่งออกทุเรียนของประเทศไทยด้วยเทคนิคการเรียนรู้เชิงลึก
# Forecasting Thailand's Durian Exports Using Deep Learning Techniques

*ชนนิกานต์ ภู่ทอง (Chonnika Phuthong)[1] และวัชรีวรรณ จิตต์สกุล (Watchareewan Jitsakul)[2]*
*ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล*
*มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ*
*[1]s6707021857228@email.kmutnb.ac.th, [2]watchareewan.j@itd.kmutnb.ac.th*

## บทคัดย่อ

*งานวิจัยนี้มีวัตถุประสงค์เพื่อสร้างแบบจำลองการพยากรณ์การส่งออกทุเรียนของประเทศไทยด้วยเทคนิคการเรียนรู้เชิงลึก 3 เทคนิค พร้อมเปรียบเทียบแบบจำลองในการพยากรณ์การส่งออกทุเรียนในแต่ละเทคนิค ได้แก่ แบบจำลองอารีมา วิธีโครงข่ายประสาทเทียมแบบวนซ้ำ และวิธีโครงข่ายประสาทเทียมแบบวนซ้ำสองทิศทาง โดยพิจารณาเลือกเปรียบเทียบประสิทธิภาพในการพยากรณ์แต่ละเทคนิคด้วยเกณฑ์การประเมินประสิทธิภาพ 3 เกณฑ์ คือ คลาดเคลื่อนสัมบูรณ์เฉลี่ย ค่าความคลาดเคลื่อนกำลังสองเฉลี่ยรากที่สอง และร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ย เพื่อเลือกแบบจำลองที่เหมาะสมที่สุด ผลการวิจัยพบว่า แบบจำลองอารีมาให้ประสิทธิภาพดีที่สุด โดยมีค่าพารามิเตอร์ คือ p = 1 d = 1 และ q = 2 จากนั้นนำแบบจำลองไปทำการพยากรณ์ข้างหน้า 24 เดือน*

**คำสำคัญ:** แบบจำลองอารีมา วิธีโครงข่ายประสาทเทียมแบบวนซ้ำ วิธีโครงข่ายประสาทเทียมแบบวนซ้ำสองทิศทาง

## Abstract

*The objective of this research is to develop a forecasting model for Thailand's durian exports using three deep learning–based techniques, and to compare their predictive performance. The three forecasting approaches examined in this study are: the Autoregressive Integrated Moving Average (ARIMA) model, the Recurrent Neural Network (RNN), and the Bidirectional Recurrent Neural Network (Bi-RNN). The performance of each model is evaluated using three accuracy metrics: Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Mean Absolute Percent Error (MAPE). These metrics are employed to determine the most suitable model for practical forecasting applications. The findings indicate that the ARIMA model performs best within the experimental framework of this research. The research results showed that the ARIMA model performed best with parameters p = 1, d = 1, and q = 2. The model was then used to make forecasts for the next 24 months.*

**Keywords:** ARIMA, Bidirectional Recurrent Neural Network Bi-RNN, Recurrent Neural Network.

## 1. บทนำ

การส่งออกสินค้าเกษตรของประเทศไทยมีบทบาทสำคัญต่อการขับเคลื่อนเศรษฐกิจ โดยเฉพาะผลไม้ที่มีมูลค่าสูงและเป็นที่ต้องการในตลาดต่างประเทศ หนึ่งในสินค้าที่มีความโดดเด่นคือ "ทุเรียน" ซึ่งเป็นผลไม้ที่ได้รับความนิยมอย่างมากและสร้างรายได้จากการส่งออกอย่างต่อเนื่อง ทั้งยังมีการเติบโตอย่างมีนัยสำคัญในช่วงปี 2563–2566 โดยมีสาธารณรัฐประชาชนจีนเป็นตลาดหลักคิดเป็นสัดส่วนมากกว่า 90% ของมูลค่าการส่งออกทั้งหมด การเติบโตดังกล่าวสะท้อนศักยภาพของภาคการเกษตรไทยในการขยายตลาด แต่ยังคงมีความเสี่ยงจากปัจจัยด้านเศรษฐกิจ ภูมิอากาศ และนโยบายการค้าระหว่างประเทศ

จากการทบทวนวรรณกรรมพบว่า มีการใช้เทคนิคพยากรณ์ที่หลากหลาย เช่น วิธีเชิงสถิติแบบอนุกรมเวลาและวิธีการเรียนรู้ของเครื่องสมัยใหม่ ซึ่งให้ผลลัพธ์แตกต่างกันตามลักษณะข้อมูล งานวิจัยนี้จึงมีวัตถุประสงค์เพื่อพัฒนาแบบจำลองพยากรณ์มูลค่าการส่งออกทุเรียนของไทย โดยเปรียบเทียบประสิทธิภาพของแบบจำลอง 3 เทคนิค ได้แก่ แบบจำลองอารีมา โครงข่ายประสาทเทียมแบบวนซ้ำ และแบบจำลองโครงข่ายประสาทเทียมแบบวนซ้ำสองทิศทาง โดยใช้เกณฑ์ประเมิน คือ คลาดเคลื่อนสัมบูรณ์เฉลี่ย (Mean Absolute Error: MAE) ค่าความคลาดเคลื่อนกำลังสองเฉลี่ยรากที่สอง (Root Mean Square Error: RMSE) และร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (Mean Absolute Percent Error: MAPE)

## 2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ผู้วิจัยได้ทำการรวบรวมข้อมูลเชิงทฤษฎีและผลงานวิจัยที่เกี่ยวข้อง เพื่อสร้างพื้นฐานความรู้และกรอบแนวคิดในการศึกษา โดยแบ่งออกเป็น 3 หัวข้อหลัก ดังนี้

### 2.1 เทคนิคการเรียนรู้เชิงลึก

ในการวิจัยครั้งนี้ ได้มีการนำแบบจำลองสำหรับการพยากรณ์ทั้งหมด 3 เทคนิค ได้แก่ แบบจำลองเชิงสถิติอนุกรมเวลา (ARIMA) และแบบจำลองการเรียนรู้เชิงลึก (RNN และ Bi-RNN) โดยมีรายละเอียด ดังนี้

**2.1.1 แบบจำลอง ARIMA (Autoregressive Integrated Moving Average)** แบบจำลอง ARIMA หรือวิธีการของ Box-Jenkins เป็นพื้นฐานในการวิเคราะห์ข้อมูลอนุกรมเวลาแบบตัวแปรเดียว เนื่องจากเป็นแบบจำลองที่มีประสิทธิภาพสูงในการจัดการกับข้อมูลที่มีความสัมพันธ์เชิงเส้น ทั้งนี้ ARIMA จัดเป็นแบบจำลองเชิงสถิติ (Statistical Model) สำหรับข้อมูลอนุกรมเวลา ซึ่งแตกต่างจาก RNN และ Bi-RNN ที่จัดเป็นแบบจำลองการเรียนรู้เชิงลึก (Deep Learning) งานวิจัยนี้จึงนำทั้งสองประเภทมาเปรียบเทียบประสิทธิภาพการพยากรณ์บนชุดข้อมูลเดียวกัน

**2.1.2 วิธีโครงข่ายประสาทเทียมแบบวนซ้ำ (Recurrent Neural Network: RNN)** เป็นโครงข่ายประสาทเทียมที่ออกแบบมาเพื่อประมวลผลข้อมูลชนิดลำดับโดยมีคุณลักษณะสำคัญคือการคงอยู่ของสถานะภายในหรือ Hidden State ซึ่งทำหน้าที่เก็บรักษาข้อมูลจากช่วงเวลาก่อนหน้า ส่งผลให้โมเดลสามารถเรียนรู้ความสัมพันธ์เชิงเวลาและรูปแบบต่อเนื่องภายในข้อมูลได้อย่างมีประสิทธิภาพ และเหมาะสมกับงานที่ต้องติดตามบริบทของข้อมูลในอดีต

**2.1.3 วิธีโครงข่ายประสาทเทียมแบบวนซ้ำสองทิศทาง (Bidirectional Recurrent Neural Network Bi-RNN)** เป็นโมเดลที่การขยายโครงสร้างของ RNN โดยเพิ่มการประมวลผลข้อมูลทั้งสองทิศทาง ได้แก่ ทิศทางไปข้างหน้า (Forward Sequence) และทิศทางย้อนกลับ (Backward Sequence) ภายในโมเดลเดียว ส่งผลให้ Bi-RNN สามารถเข้าถึงบริบททั้งอดีตและอนาคตของข้อมูลในลำดับเดียวกัน ทำให้มีความสามารถในการจับรูปแบบเชิงเวลาได้ครอบคลุมและแม่นยำกว่า RNN แบบทางเดียว เหมาะกับงานที่บริบททั้งสองด้านมีความสำคัญ

### 2.2 งานวิจัยที่เกี่ยวข้อง

Sandhya C. [1] ได้ทำการศึกษาวิจัย เรื่อง Sales Forecasting Using ARIMA Modelเพื่อพยากรณ์ยอดขายของบริษัท โดยใช้ข้อมูลยอดขายรายวันจำนวน 278,559 รายการ ในช่วงระยะเวลา 6 เดือน มีการสร้างโมเดลอนุกรมเวลาหลายรูปแบบที่เหมาะสม และใช้เกณฑ์การประเมินประสิทธิภาพหลายตัวเพื่อเปรียบเทียบโมเดลต่าง ๆ ผลการวิเคราะห์พบว่า โมเดล ARIMA (0,1,1) เพียงพอสำหรับการพยากรณ์ยอดขาย

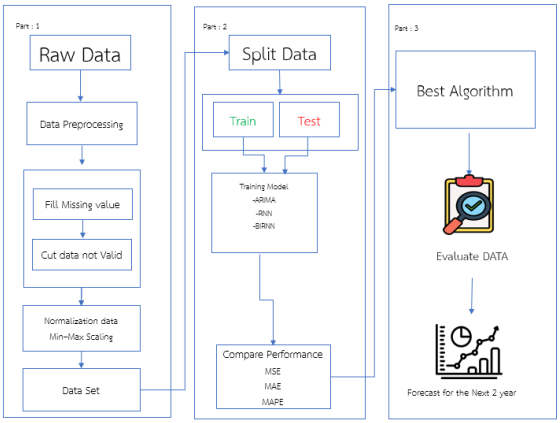
ไพศาล [2] ได้ทำการศึกษาวิจัย เรื่อง การพยากรณ์การผลิตและการส่งออกทุเรียนไทย กรณีศึกษาเชิงประจักษ์ด้วยวิธีการทางอนุกรมเวลา ได้ทำการศึกษาการพยากรณ์โดยใช้การวิเคราะห์ข้อมูลใช้อนุกรมเวลารายเดือนด้วยแบบจำลอง SARIMA ผลการศึกษาพบว่า (1) การผลิตและการส่งออกทุเรียนมีแนวโน้มขยายตัวเพิ่มขึ้นอย่างต่อเนื่องสืบเนื่องจากอุปสงค์ทุเรียนจากตลาดในประเทศและตลาดต่างประเทศขยายตัวขึ้น โดยเฉพาะอย่างยิ่งการส่งออกทุเรียนไปยังตลาดจีนทั้งนี้ข้อค้นพบแสดงให้เห็นว่าการผลิตทุเรียนของประเทศไทยมีแนวโน้มขยายตัวขึ้นร้อยละ 9.726 ในช่วงครึ่งปีหลังของปี พ.ศ. 2564 สอดคล้องกับการพยากรณ์การส่งออกทุเรียนไปยังตลาดจีนและตลาดโลกซึ่ง

มีแนวโน้มขยายตัวขึ้นในช่วงครึ่งปี หลังของปีพ.ศ. 2564 ร้อยละ 26.988 และร้อยละ 0.968 ตามลำดับ

พิเชษฐ์ [8] ได้ทำการวิจัยเรื่อง การเปรียบเทียบประสิทธิภาพของแบบจำลองการพยากรณ์ราคาทองคำด้วยเทคนิคเหมืองข้อมูลได้ศึกษาและเปรียบเทียบแบบจำลองอนุกรมเวลาแบบ Box-Jenkins แบบจำลองอนุกรมเวลาแบบ Holt-Winters และ โครงข่ายประสาทเทียม (ANN) และเลือกแบบจำลองการวิเคราะห์การถดถอยเชิงเส้น (Linear Regression Analysis) ซัพพอร์ตเวกเตอร์แมชชีน ( SVM) เพิ่มเติม เพื่อสร้างความแตกต่างในงานวิจัยผลการศึกษาพบว่า เมื่อใช้ข้อมูลเพื่อการเรียนรู้แบบจำลองอนุกรมเวลาแบบ Box-Jenkins โดยค่าพารามิเตอร์ที่เหมาะสมที่สุดคือ window size=120, p=0, d=1, q=1 (ARIMA(0,1,1)) คำนวณค่า RMSE ได้น้อยที่สุด สามารถใช้พยากรณ์ราคาทองคำได้เหมาะสมที่สุด

## 3. วิธีการดำเนินการวิจัย

งานวิจัยนี้มีวัตถุประสงค์เพื่อสร้างแบบจำลองการพยากรณ์การส่งออกทุเรียนของประเทศไทยด้วยเทคนิคการเรียนรู้เชิงลึก โดยมีกรอบแนวคิดแสดงดังภาพที่ 1




**ภาพที่ 1:** กรอบแนวคิด

จากภาพที่ 1 แสดงกรอบแนวคิดของงานวิจัยประกอบด้วย 6 ส่วนหลัก ได้แก่ การเก็บรวบรวม จัดเตรียมข้อมูล การพัฒนาตัวแบบ การวัดประสิทธิภาพของแบบจำลอง และขั้นตอนการวิเคราะห์และสรุปผล

### 3.1 การรวบรวมข้อมูล (Data Collection)

ผู้วิจัยได้รวบรวมข้อมูลการส่งออกทุเรียนของประเทศไทยจากฐานข้อมูลสถิติการค้าระหว่างประเทศของสำนักงานปลัดกระทรวงพาณิชย์ ครอบคลุมช่วงเดือนมกราคม พ.ศ. 2563 ถึงธันวาคม พ.ศ. 2567 รวม 60 เดือน

**3.1.1 การทำความสะอาดข้อมูล** เป็นการตรวจสอบค่าที่ขาดหายของตัวแปรและทำการจัดกลุ่มข้อมูลรายประเทศ เพื่อเตรียมข้อมูลในการนำเข้าแบบจำลองแบบแยกกลุ่มประเทศ เลือกทวีปเอเชีย

**3.1.2 การจัดการข้อมูลสูญหาย** ด้วยเทคนิค Forward Fill และ Backward Fill ภายในแต่ละประเทศ

**3.1.3 คัดเลือกข้อมูลที่เหมาะสม** คัดเลือกประเทศที่มีข้อมูลที่ไม่มีการผสมกันของข้อมูลระหว่างแนวโน้มและความแปรผันตามฤดูกาล โดยใช้แบบทดสอบ Augmented Dickey–Fuller (ADF)

**3.1.4 ปรับมาตราส่วนข้อมูล (Scaling)** ให้อยู่ในช่วง [0-1] โดยใช้เทคนิค Min-Max Normalization

**3.1.5 การแบ่งชุดข้อมูล** เป็นกระบวนการกำหนดชุดฝึกสอน (Training) ปีพ.ศ. 2563–2566 และกำหนดชุดทดสอบ (Testing) เป็นปี พ.ศ. 2567

**3.2 การพัฒนาตัวแบบ และเปรียบเทียบประสิทธิภาพ**

ทำการนำชุดข้อมูลที่ผ่านการเตรียมข้อมูลแล้วจำนวน 1,056 ชุดข้อมูลมาทำการสร้างแบบจำลองเพื่อเปรียบเทียบประสิทธิภาพ 3 รูปแบบประกอบด้วย วิธี ARIMA วิธี RNN และวิธี Bi-RNN โดยแต่ละวิธีมีการกำหนดพารามิเตอร์เพื่อปรับแต่งให้ได้แบบจำลองที่มีประสิทธิภาพดีที่สุด แสดงดังตารางที่ 1 ถึง 3

**ตารางที่ 1 :** Parameter แบบจำลอง ARIMA

| Algorithms | Parameters |
|---|---|
| ARIMA | P = 0-2<br>D = 0-2<br>Q = 0-2 |

**ตารางที่ 2 :** Parameter แบบจำลอง RNN

| Algorithms | Parameters |
|---|---|
| RNN | look_back_periods = [1, 3, 6]<br>epochs_list = [2,4]<br>batch_size_list = [1,4,8]<br>lstm_units_list = [1,4,8]<br>learning_rate_list = [0.001] |

**ตารางที่ 3 :** Parameter แบบจำลอง Bi RNN

| Algorithms | Parameters |
|---|---|
| BI-RNN | look_back_periods = [1, 3, 6]<br>epochs_list = [2,4]<br>batch_size_list = [1,4,8]<br>lstm_units_list = [1,4,8]<br>learning_rate_list = [0.001] |

## 4. ผลการดำเนินงาน

ทำการประเมินประสิทธิภาพของตัวแบบด้วยค่า MAE, RMSE และ MAPE เพื่อเลือกแบบจำลองที่ดีที่สุด แสดงดังตารางที่ 4 ถึง 6

**ตารางที่ 4 :** ตัวอย่างผลลัพธ์ของแบบจำลอง ARIMA

| Country | Parameter | | | Performance | | |
|---|---|---|---|---|---|---|
| | p | d | q | RMSE | MAE | MAPE |
| Bahrain | 0 | 1 | 1 | 0.1913 | 0.1849 | 447.0172 |
| | 1 | 0 | 1 | 0.1471 | 0.1374 | 339.3766 |
| | 1 | 1 | 2 | 0.1464 | 0.1364 | 336.1236 |
| Bangladesh | 0 | 0 | 2 | 0.1609 | 0.1149 | 54.8222 |
| | 1 | 1 | 0 | 0.3087 | 0.2682 | 271.4972 |
| | **1** | **1** | **2** | **0.1557** | **0.1132** | **138.1465** |
| Brunei | 0 | 1 | 0 | 0.0684 | 0.0468 | 100.0000 |
| | **1** | **1** | **2** | **0.0457** | **0.0359** | **114.9615** |
| | 2 | 0 | 1 | 0.0519 | 0.0480 | 167.3033 |
| Cambodia | 0 | 0 | 1 | 0.0878 | 0.0858 | 57.6210 |
| | 1 | 0 | 1 | 0.0878 | 0.0858 | 57.7488 |
| | **1** | **1** | **2** | **0.0878** | **0.0858** | **57.6066** |
| India | **1** | **1** | **2** | **0.0559** | **0.0335** | **97.9367** |
| | 2 | 0 | 0 | 0.0892 | 0.0739 | 106.611 |
| | 2 | 0 | 1 | 0.0603 | 0.0528 | 67.4173 |

**ตารางที่ 5 :** ตัวอย่างผลลัพธ์ของแบบจำลอง RNN

| Country | Parameter | | | | Performance | | |
|---|---|---|---|---|---|---|---|
| | LookBack | Epochs | Batch | LSTM | RMSE | MAE | MAPE |
| Bahrain | 3 | 4 | 1 | 4 | 0.1163 | 0.1116 | 237.6260 |
| | 3 | 4 | 1 | 8 | 0.1172 | 0.1125 | 236.4402 |
| | 6 | 4 | 1 | 4 | 0.1163 | 0.1129 | 202.5426 |
| Bangladesh | 1 | 2 | 8 | 1 | 0.2221 | 0.1854 | 92.0434 |
| | 1 | 4 | 8 | 1 | 0.2445 | 0.2000 | 96.4700 |
| | 3 | 2 | 1 | 1 | 0.2759 | 0.2270 | 92.0010 |
| Brunei | 1 | 2 | 4 | 8 | 0.0618 | 0.0420 | 59.0779 |
| | 3 | 2 | 8 | 8 | 0.0748 | 0.0562 | 82.5928 |
| | 6 | 2 | 8 | 4 | 0.0861 | 0.0652 | 80.3163 |
| Cambodia | 1 | 2 | 1 | 1 | 0.3012 | 0.1016 | 55.0226 |
| | 1 | 2 | 8 | 8 | 0.3006 | 0.0955 | 65.7686 |
| | 1 | 4 | 4 | 8 | 0.3005 | 0.0952 | 65.4623 |
| India | 6 | 2 | 4 | 4 | 0.0637 | 0.0494 | 74.9175 |
| | 6 | 2 | 8 | 8 | 0.0707 | 0.0528 | 93.172 |
| | 6 | 4 | 8 | 1 | 0.0718 | 0.0538 | 96.2315 |

**ตารางที่ 6 :** ตัวอย่างผลลัพธ์ของแบบจำลอง Bi-RNN

| Country | Parameter | | | | Performance | | |
|---|---|---|---|---|---|---|---|
| | LookBack | Epochs | Batch | LSTM | RMSE | MAE | MAPE |
| Bahrain | **3** | **4** | **1** | **4** | **0.1107** | **0.1053** | **228.3583** |
| | 3 | 4 | 1 | 8 | 0.1277 | 0.1235 | 261.5328 |
| | 6 | 4 | 1 | 8 | 0.1064 | 0.1032 | 184.8843 |
| Bangladesh | 3 | 2 | 4 | 4 | 0.2346 | 0.2002 | 85.1687 |
| | 6 | 2 | 4 | 4 | 0.2943 | 0.2745 | 94.4315 |
| | 6 | 2 | 4 | 1 | 0.3332 | 0.3119 | 107.9701 |
| Brunei | 1 | 4 | 4 | 1 | 0.0611 | 0.0423 | 62.2938 |
| | 3 | 2 | 8 | 8 | 0.0714 | 0.0522 | 79.5095 |
| | 6 | 2 | 8 | 1 | 0.0859 | 0.0645 | 77.4946 |
| Cambodia | 1 | 2 | 4 | 4 | 0.3036 | 0.1075 | 56.0238 |
| | 1 | 4 | 4 | 4 | 0.3017 | 0.1019 | 61.8739 |
| | 1 | 4 | 8 | 8 | 0.3006 | 0.0972 | 61.9964 |

| Country | Parameter | | | | Performance | | |
|---|---|---|---|---|---|---|---|
| | LookBack | Epochs | Batch | LSTM | RMSE | MAE | MAPE |
| India | 1 | 4 | 1 | 1 | 0.0612 | 0.0564 | 46.1075 |
| | 1 | 4 | 1 | 4 | 0.0719 | 0.0658 | 62.2588 |
| | 6 | 2 | 8 | 1 | 0.0913 | 0.0847 | 143.2714 |

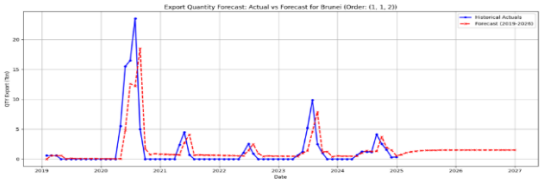
เมื่อได้ผลลัพธ์ของ 3 แบบจำลอง ทำการเปรียบเทียบประสิทธิภาพของแต่ละแบบจำลอง พบว่า แบบจำลอง ARIMA ที่พารามิเตอร์ p เท่ากับ 1, d เท่ากับ 1 และ q เท่ากับ 2 ให้ประสิทธิภาพดีที่สุด จากนั้นนำแบบจำลองมาทำการพยากรณ์การส่งออกทุเรียนแต่ละประเทศใน 24 เดือนข้างหน้า

**ตารางที่ 7 :** ผลการพยากรณ์การส่งออกทุเรียน 24 เดือนข้างหน้า ด้วยแบบจำลอง ARIMA (1,1,2)

| ประเทศ | เดือน/ปี | ค่าพยากรณ์ (ตัน) |
|---|---|---|
| Bahrain | 31/1/2568 | 0.1571 |
| | 28/2/2568 | 0.2091 |
| | 31/3/2568 | 0.2324 |
| | 30/4/2568 | 0.2428 |
| | 31/5/2568 | 0.2475 |
| | 30/6/2568 | 0.2496 |
| | 31/7/2568 | 0.2506 |
| | 31/8/2568 | 0.251 |
| | 30/9/2568 | 0.2512 |
| | 31/10/2568 | 0.2513 |
| | 30/11/2568 | 0.2513 |
| | 31/12/2568 | 0.2513 |
| | 31/1/2569 | 0.2513 |
| | 28/2/2569 | 0.2513 |
| | 31/3/2569 | 0.2513 |
| | 30/4/2569 | 0.2513 |
| | 31/5/2569 | 0.2513 |
| | 30/6/2569 | 0.2513 |
| | 31/7/2569 | 0.2513 |
| | 31/8/2569 | 0.2513 |
| | 30/9/2569 | 0.2513 |
| | 31/10/2569 | 0.2513 |
| | 30/11/2569 | 0.2513 |
| | 31/12/2569 | 0.2513 |
| Bangladesh | 31/1/2568 | 0.3528 |
| | 28/2/2568 | 0.4616 |
| | 31/3/2568 | 0.3638 |
| | 30/4/2568 | 0.4517 |
| | 31/5/2568 | 0.3726 |
| | 30/6/2568 | 0.4437 |
| | 31/7/2568 | 0.3798 |
| | 31/8/2568 | 0.4373 |
| | 30/9/2568 | 0.3856 |
| | 31/10/2568 | 0.432 |
| | 30/11/2568 | 0.3903 |
| | 31/12/2568 | 0.4278 |
| | 31/1/2569 | 0.3941 |
| | 28/2/2569 | 0.4244 |
| | 31/3/2569 | 0.3971 |
| | 30/4/2569 | 0.4217 |
| | 31/5/2569 | 0.3996 |
| | 30/6/2569 | 0.4194 |
| | 31/7/2569 | 0.4016 |
| | 31/8/2569 | 0.4176 |
| | 30/9/2569 | 0.4032 |
| | 31/10/2569 | 0.4162 |
| | 30/11/2569 | 0.4045 |
| | 31/12/2569 | 0.4150 |
| Brunei | 31/1/2568 | 0.749 |
| | 28/2/2568 | 1.0437 |
| | 31/3/2568 | 1.2265 |
| | 30/4/2568 | 1.3399 |
| | 31/5/2568 | 1.4102 |
| | 30/6/2568 | 1.4538 |
| | 31/7/2568 | 1.4808 |
| | 31/8/2568 | 1.4976 |
| | 30/9/2568 | 1.508 |
| | 31/10/2568 | 1.5144 |
| | 30/11/2568 | 1.5184 |
| | 31/12/2568 | 1.5209 |
| | 31/1/2569 | 1.5224 |
| | 28/2/2569 | 1.5234 |
| | 31/3/2569 | 1.5240 |
| | 30/4/2569 | 1.5244 |
| | 31/5/2569 | 1.5246 |
| | 30/6/2569 | 1.5247 |
| | 31/7/2569 | 1.5248 |
| | 31/8/2569 | 1.5249 |
| | 30/9/2569 | 1.5249 |
| | 31/10/2569 | 1.5249 |
| | 30/11/2569 | 1.5249 |
| | 31/12/2569 | 1.5249 |
| | 31/1/2568 | 10.4356 |

| ประเทศ | เดือน/ปี | ค่าพยากรณ์ (ตัน) |
|---|---|---|
| Cambodia | 28/2/2568 | 9.0975 |
| | 31/3/2568 | 10.2952 |
| | 30/4/2568 | 9.2232 |
| | 31/5/2568 | 10.1827 |
| | 30/6/2568 | 9.3239 |
| | 31/7/2568 | 10.0925 |
| | 31/8/2568 | 9.4045 |
| | 30/9/2568 | 10.0203 |
| | 31/10/2568 | 9.4692 |
| | 30/11/2568 | 9.9625 |
| | 31/12/2568 | 9.5209 |
| | 31/1/2569 | 9.9161 |
| | 28/2/2569 | 9.5624 |
| | 31/3/2569 | 9.879 |
| | 30/4/2569 | 9.5956 |
| | 31/5/2569 | 9.8493 |
| | 30/6/2569 | 9.6223 |
| | 31/7/2569 | 9.8255 |
| | 31/8/2569 | 9.6436 |
| | 30/9/2569 | 9.8064 |
| | 31/10/2569 | 9.6607 |
| | 30/11/2569 | 9.7911 |
| | 31/12/2569 | 9.6744 |
| India | 31/1/2568 | 0.0559 |
| | 28/2/2568 | 0.0501 |
| | 31/3/2568 | 0.0496 |
| | 30/4/2568 | 0.0495 |
| | 31/5/2568 | 0.0495 |
| | 30/6/2568 | 0.0495 |
| | 31/7/2568 | 0.0495 |
| | 31/8/2568 | 0.0495 |
| | 30/9/2568 | 0.0495 |
| | 31/10/2568 | 0.0495 |
| | 30/11/2568 | 0.0495 |
| | 31/12/2568 | 0.0495 |
| | 31/1/2569 | 0.0495 |
| | 28/2/2569 | 0.0495 |
| | 31/3/2569 | 0.0495 |
| | 30/4/2569 | 0.0495 |
| | 31/5/2569 | 0.0495 |
| | 30/6/2569 | 0.0495 |
| | 31/7/2569 | 0.0495 |
| | 31/8/2569 | 0.0495 |
| | 30/9/2569 | 0.0495 |
| | 31/10/2569 | 0.0495 |
| | 30/11/2569 | 0.0495 |
| | 31/12/2569 | 0.0495 |

ตัวอย่างผลการพยากรณ์การส่งออกประเทศบรูไน แสดงเป็นกราฟแสดงดังภาพที่ 2



**ภาพที่ 2:** ผลการจำลองการพยากรณ์ ARIMA ประเทศบรูไน

## 5. สรุป

งานวิจัยนี้มีวัตถุประสงค์เพื่อสร้างแบบจำลองการพยากรณ์การส่งออกทุเรียนของประเทศไทยด้วยเทคนิคการเรียนรู้เชิงลึก 3 เทคนิค พร้อมเปรียบเทียบแบบจำลองในการพยากรณ์การส่งออกทุเรียนในแต่ละเทคนิค ได้แก่ แบบจำลองอารีมา วิธีโครงข่ายประสาทเทียมแบบวนซ้ำ และวิธีโครงข่ายประสาทเทียมแบบวนซ้ำสองทิศทาง โดยเริ่มจากการรวบรวมข้อมูลรายเดือนช่วงปี 2562–2567 จำนวน 1,656 ระเบียน ผ่านกระบวนการทำความสะอาด เติมค่าที่สูญหายและแปลงค่าปริมาณให้อยู่ในรูปตันก่อน แบ่งข้อมูลเป็นชุดฝึกสอน และชุดทดสอบ ประเมินประสิทธิภาพด้วย MAE, RMSE และ MAPE พบว่าแบบจำลอง ARIMA ให้ประสิทธิภาพดีที่สุด ที่พารามิเตอร์ p เท่ากับ 1 d เท่ากับ 1 และ q เท่ากับ เนื่องจากเป็นวิธีมาตรฐานที่ได้รับการพิสูจน์แล้วว่ามีประสิทธิภาพสูงในการพยากรณ์อนุกรมเวลาที่มีลักษณะเชิงเส้น และเหมาะสมกับข้อมูลขนาดเล็กถึงกลาง จากนั้นนำแบบจำลองไปทำการพยากรณ์การส่งออกทุเรียน 24 เดือน เพื่อใช้บริหารจัดการด้านการเกษตร อย่างไรก็ตาม งานวิจัยในอนาคตควรพิจารณาเพิ่มขนาดข้อมูลและทดสอบแบบจำลองขั้นสูง เช่น LSTM หรือ Transformer เพื่อให้ได้ผลการเปรียบเทียบที่ครบถ้วนยิ่งขึ้น

### เอกสารอ้างอิง

[1]. Paesarn, R. (2020). The forecasting of domestic and export durian prices of Thailand (Master's thesis). Sukhothai Thammathirat Open University.

[2]. Sandhya, C., et al. (2022). Sales forecasting using ARIMA model. International Journal of Creative Research Thoughts, 10(6), 135–139.

[3]. Pichate Sirirattanapaisankul. (2022). A comparison of forecasting model performance for gold price prediction using data mining techniques (Master's thesis). Suratthani Rajabhat University.

[4]. Sowmya, P., Krishna, M., & Saritha, D. (2025). The water quality prediction using decision tree. In International Conference on Innovations in Robotics, Intelligent Automation and Control. https://doi.org/10.1063/5.0246060

[5]. Bo, Y. (2023). Walmart sales prediction based on decision tree, random forest, and K neighbors regressor. Highlights in Business, Economics and Management, 5, 330–335.

[6]. Sha, K., & Hela, A. (2020). ARIMA model for accurate time series stocks forecasting. International Journal of Advanced Computer Science and Applications, 11(7), 524–528.

# แบบจำลองเพื่อการพยากรณ์จำนวนคนไข้มะเร็งในโรงพยาบาลโดยใช้เทคนิคอนุกรมเวลา
# Forecasting Model for the Number of Patients Cancer in the Public Hospital base by Time Series Technical

*รัญชน์ กาญจะโน (Run Kanchano)[1] และณัฐวี อุตกฤษฎ์ (Nattavee Utakrit)[2]*
*ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล*
*มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ*
*[1]s 6707021857147@email.kmutnb.ac.th, [2] nattavee.u@itd.kmutnb.ac.th*

**บทคัดย่อ**

*งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาแบบจำลองการพยากรณ์จำนวนผู้ป่วยโรคมะเร็งในโรงพยาบาลของรัฐ โดยใช้การวิเคราะห์อนุกรมเวลา เพื่อสนับสนุนการวางแผนด้านสาธารณสุข การศึกษานี้เป็นการวิจัยเชิงพัฒนาและเชิงปริมาณ โดยใช้ข้อมูลย้อนหลังจำนวน 180 เดือน และประมวลผลด้วยภาษา Python เพื่อประเมินเทคนิคการพยากรณ์ 4 วิธี ได้แก่ ค่าเฉลี่ยเคลื่อนที่ MA, SES,DES และ ARIMA โดยประเมินประสิทธิภาพของแบบจำลองถูกประเมินโดยใช้ค่าความคลาดเคลื่อนสัมบูรณ์เฉลี่ย MAE และค่ารากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ย RMSE นอกจากนี้ ยังได้พัฒนาแพลตฟอร์มสำหรับการพยากรณ์ และทำการประเมินความพึงพอใจจากทั้งผู้ใช้งานและผู้เชี่ยวชาญในสาขาที่เกี่ยวข้อง ผลการวิจัยพบว่า แบบจำลองการปรับเรียบแบบเอ็กซ์โปเนนเชียลอย่างง่าย SES มีความแม่นยำโดยรวมสูงที่สุดในบางกลุ่มโรคเนื่องจากลักษณะข้อมูลผู้ป่วยโรคมะเร็งมีแนวโน้มค่อนข้างคงที่ และไม่มีรูปแบบฤดูกาลที่ชัดเจน ทำให้ SES ซึ่งให้น้ำหนักกับข้อมูลล่าสุด สามารถสะท้อนแนวโน้มของข้อมูลได้ดีกว่าแบบจำลองอื่น ขณะที่แบบจำลองอารี ARIMA มีความแม่นยำรองลงมามีการจัดการข้อมูลได้ แต่ต้องการปรับพารามิเตอร์ให้เหมาะสม*

**คำสำคัญ:** การพยากรณ์อนุกรมเวลา มะเร็ง แบบจำลองอารีมา

**Abstract**

*This research aims to develop a forecasting model for the number of cancer patients in public hospitals using time series analysis to support public health planning. The research is both developmental and quantitative, utilizing 180 months of historical data processed with Python. Four forecasting techniques were evaluated: MA, SES,DES and ARIMA.*

*Model performance was assessed using Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE). In addition, a forecasting platform was developed, and user satisfaction was evaluated by both end-users and domain experts.*

*The results indicate that the Simple Exponential Smoothing (SES) model achieved the highest overall accuracy in certain disease groups. This is because the cancer patient data shows a relatively stable trend with no clear seasonality, allowing SES to effectively capture recent patterns. The ARIMA model ranked second, as it can handle more complex data structures but requires proper parameter tuning.*

**Keywords**: Time Series Forecasting, Cancer, ARIMA.

## 1. บทนำ

ปัจจุบันโรคมะเร็งถือเป็นหนึ่งในปัญหาสาธารณสุขที่สำคัญของโลก เนื่องจากมีอัตราการเกิดและอัตราการเสียชีวิตสูง และส่งผลกระทบต่อสุขภาพ คุณภาพชีวิต รวมถึงระบบเศรษฐกิจและสังคมของหลายประเทศ จาก

รายงานของ WHO ระบุว่าในปี พ.ศ. 2565 ทั่วโลกมีผู้ป่วยโรคมะเร็งรายใหม่ประมาณ 20 ล้านราย และมีผู้เสียชีวิตประมาณ 9.7 ล้านราย ขณะที่ยังมีผู้ที่มีชีวิตอยู่ภายในระยะเวลา 5 ปีหลังได้รับการวินิจฉัยมากกว่า 53.5 ล้านราย ซึ่งสะท้อนถึงภาระโรคที่เพิ่มขึ้นอย่างต่อเนื่องในระบบสาธารณสุขทั่วโลก

ดังนั้นงานวิจัยนี้จึงมุ่งพัฒนาแบบจำลองสำหรับการพยากรณ์แนวโน้มโดยประยุกต์ใช้เทคนิคการวิเคราะห์ข้อมูลและการพยากรณ์อนุกรมเวลา ได้แก่ วิธีค่าเฉลี่ยเคลื่อนที่,การปรับเรียบแบบเอ็กซ์โพเนนเชียลอย่างง่าย, การปรับเรียบแบบเอ็กซ์โพเนนเชียลและแบบจำลองออโตรีเกรสซีฟอินทิเกรเต็ดมูฟวิ่งเอเวอเรจ เพื่อวิเคราะห์รูปแบบและแนวโน้มของข้อมูลผู้ป่วยโรคมะเร็งผลจากการศึกษาคาดว่าจะช่วยให้สามารถพัฒนาแบบจำลองพยากรณ์ที่มีความแม่นยำ

ข้อมูลผู้ป่วยที่ใช้ในการวิจัยได้รับอนุญาตจากหน่วยงานที่เกี่ยวข้อง และผ่านกระบวนการไม่ระบุตัวตน (Anonymous Data) โดยไม่มีข้อมูลส่วนบุคคล เช่น ชื่อ นามสกุล หรือเลขประจำตัวประชาชน เพื่อให้เป็นไปตามหลักจริยธรรมการวิจัยและการคุ้มครองข้อมูลส่วนบุคคลทั้งนี้ งานวิจัยเป็นการวิจัยเชิงพัฒนา (Developmental Research) โดยมีการทดสอบระบบกับผู้เชี่ยวชาญและผู้ใช้งาน เพื่อประเมินประสิทธิภาพของระบบ

## 2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ผู้วิจัยได้ศึกษาค้นคว้าเอกสาร และงานวิจัยที่เกี่ยวข้องต่าง ๆ นำไปพัฒนาแบบจำลองเพื่อการพยากรณ์ โดยศึกษาข้อมูลต่าง ๆ ดังต่อไปนี้

### 2.1 ทฤษฎีการวิเคราะห์อนุกรมเวลา

การวิเคราะห์อนุกรมเวลาเป็นเทคนิคทางสถิติที่ใช้ในการวิเคราะห์ชุดข้อมูลที่มีการจัดเก็บตามลำดับเวลา เพื่อค้นหารูปแบบ ความสัมพันธ์ และพฤติกรรมของข้อมูลในอดีตสำหรับใช้ในการพยากรณ์ค่าที่จะเกิดขึ้นในอนาคต การวิเคราะห์นี้มีความสำคัญอย่างยิ่งในการวางแผนเชิงกลยุทธ์และการบริหารจัดการทรัพยากรทางการแพทย์ที่มีจำกัด โดยองค์ประกอบหลักของอนุกรมประกอบด้วย แนวโน้ม (Trend), ตัวแปรด้านฤดูกาล (Seasonality), วัฏจักร (Cyclical) และความแปรปรวนที่เกิดขึ้นโดยสุ่ม (Irregular Variation) ซึ่งการเลือกใช้แบบจำลองที่เหมาะสมจะขึ้นอยู่กับลักษณะการกระจายตัวของข้อมูลดังกล่าว [1]

การออกแบบจำลองได้คัดเลือกเทคนิคการพยากรณ์เชิงปริมาณ 4 วิธีMoving Average (MA): ใช้ค่าเฉลี่ยย้อนหลังเพื่อลดความผันผวน เหมาะกับข้อมูลที่คงที่ไม่มีแนวโน้มหรือฤดูกาลชัดเจนSimple Exponential Smoothing (SES): ให้ค่าน้ำหนักกับข้อมูลล่าสุดมากกว่าข้อมูลอดีต เหมาะกับข้อมูลที่มีความผันผวนแบบสุ่มและไม่มีแนวโน้ม (ให้ผลดีที่สุดในงานวิจัยนี้) Double Exponential Smoothing (DES): พัฒนาจาก SES โดยเพิ่มพารามิเตอร์เพื่อรองรับข้อมูลที่มีแนวโน้มทั้งขาขึ้นและขาลง ARIMA: แบบจำลองสถิติขั้นสูงที่รวมการวิเคราะห์ความสัมพันธ์ในตัวเอง (AR), การปรับข้อมูลให้นิ่ง (I) และค่าคลาดเคลื่อนเคลื่อนที่ (MA) เพื่อจัดการข้อมูลที่มีความซับซ้อน (ให้ผลดีเป็นอันดับ 2)

การประเมินความแม่นยำและการสรุปผลเพื่อวัดประสิทธิภาพของแบบจำลอง ผู้วิจัยใช้เกณฑ์ทางสถิติในการวัดค่าความคลาดเคลื่อน (Error Metrics) เป็นตัวตัดสินหลัก โดยมุ่งเน้นที่การเปรียบเทียบค่าความคลาดเคลื่อนที่ต่ำที่สุดเพื่อให้ได้พยากรณ์ที่ใกล้เคียงกับความเป็นจริงมากที่สุด

### 2.2 แนวคิดการสร้างแบบจำลองการพยากรณ์ ที่เกิดขึ้น

แนวคิดเกี่ยวกับการสร้างแบบจำลองการพยากรณ์ เป็นกระบวนการทางสถิติและคณิตศาสตร์ที่มุ่งเน้นการวิเคราะห์ข้อมูลในอดีตเพื่อคาดการณ์เหตุการณ์ที่จะเกิดขึ้นในอนาคต โดยมีวัตถุประสงค์หลักเพื่อลดความเสี่ยงในการตัดสินใจและเพิ่มประสิทธิภาพในการวางแผนเชิงกลยุทธ์ ภายใต้บริบทของงานวิจัยด้านสาธารณสุข การพยากรณ์จำนวนผู้ป่วยถือเป็นเครื่องมือสำคัญในการบริหารจัดการทรัพยากรที่มีอยู่อย่างจำกัดให้สอดคล้องกับความต้องการจริง แนวคิดนี้อ้างอิงหลักการทางสถิติที่ประกอบด้วย การคัดเลือกรูปแบบข้อมูล [2] การเลือกเทคนิคที่เหมาะสม และการวัดความแม่นยำ

### 2.3 งานวิจัยที่เกี่ยวข้อง

การทบทวนงานวิจัยที่เกี่ยวข้องกับการพยากรณ์ข้อมูลอนุกรมเวลาในบริบทด้านสาธารณสุข พบว่า

สุภาวดี ชัยวัฒน์ [5] ได้เปรียบเทียบประสิทธิภาพของแบบจำลองหลายรูปแบบ ได้แก่ MA, SES และ ARIMA ในการพยากรณ์ข้อมูลผู้ป่วยรายเดือน และพบว่าการเลือกแบบจำลองที่เหมาะสมขึ้นอยู่กับลักษณะของข้อมูลในแต่ละช่วงเวลา

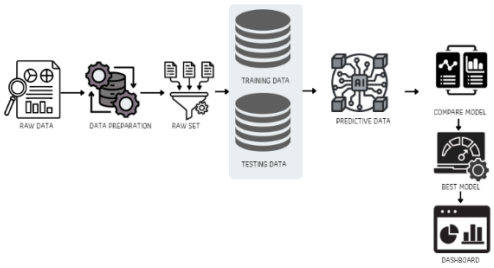
วรียา วงศ์พานิช [6] ได้ศึกษาการพยากรณ์ข้อมูลอนุกรมเวลาทางการแพทย์โดยใช้ ARIMA และพบว่าสามารถให้ผลลัพธ์ที่มีความแม่นยำสูง แต่ยังมีข้อจำกัดในกรณีที่ข้อมูลมีความผันผวนหรือไม่มีรูปแบบชัดเจนจาก

จากงานวิจัยข้างต้น พบว่า MA, SES,DES และ ARIMA เป็นเทคนิคที่นิยมใช้ โดยแต่ละวิธีเหมาะกับลักษณะข้อมูลต่างกัน อย่างไรก็ตาม ยังขาดการเปรียบเทียบหลายแบบจำลองในข้อมูลผู้ป่วยโรคมะเร็ง

## 3. วิธีการดำเนินการวิจัย

การวิจัยครั้งนี้ได้ศึกษาข้อมูลที่เกี่ยวข้องกับการพยากรณ์ข้อมูลอนุกรมเวลา รวมถึงข้อมูลงานวิจัยต่างๆที่เกี่ยวข้องกับด้านสาธารณสุขมุ่งเน้นการวิเคราะห์จากข้อมูลย้อนหลัง รวบถึงการทบทวนงานวิจัยที่เกี่ยวข้องกับเทคนิคการพยากรณ์เช่น MA, SES, DES, ARIMA เพื่อนำมาใช้การสร้างแบบจำลองที่เหมาะสม

กรอบแนวคิดของการวิจัยนี้มุ่งเน้นการพัฒนาระบบสำหรับวิเคราะห์และพยากรณ์จำนวนผู้ป่วยโรคมะเร็ง โดยมีองค์ประกอบสำคัญ ได้แก่ การรวบรวมข้อมูลผู้ป่วยในอดีต การเตรียมข้อมูล การสร้างแบบจำลองพยากรณ์ และการประเมินประสิทธิภาพของแบบจำลองด้วยค่าความคลาดเคลื่อน เช่น MAE,RMSE ตัวแปรที่ใช้ในการวิเคราะห์ประกอบด้วย ตัวแปรตาม ได้แก่ จำนวนผู้ป่วยโรคมะเร็งรายเดือน และตัวแปรอิสระ ได้แก่ เวลา (เดือน/ปี) รวมถึงตัวแปรจัดกลุ่ม ได้แก่ ประเภทของโรคมะเร็ง เพื่อใช้ในการวิเคราะห์แนวโน้มและเปรียบเทียบประสิทธิภาพของแบบจำลองพยากรณ์




**ภาพที่ 1** : กรอบแนวทางการดำเนินงานวิจัย

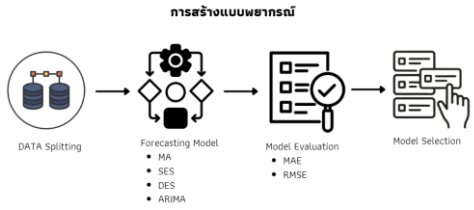
### 3.1 การเก็บรวบรวมข้อมูลและจัดการข้อมูล

ในการวิจัยครั้งนี้ ผู้วิจัยได้รวบรวมข้อมูลจำนวนผู้ป่วยโรคมะเร็งจากฐานข้อมูลของโรงพยาบาลในช่วงเดือนมกราคม พ.ศ. 2553 ถึงเดือนธันวาคม พ.ศ. 2567 เพื่อนำมาใช้ในการวิเคราะห์แนวโน้มของจำนวนผู้ป่วยในช่วงเวลาต่าง ๆ ข้อมูลที่ได้รับได้ผ่านกระบวนการตรวจสอบความถูกต้อง ความครบถ้วน และการทำความสะอาดข้อมูล (Data Cleansing) เพื่อให้ข้อมูลอยู่ในรูปแบบที่เหมาะสมสำหรับการวิเคราะห์

จากนั้นได้ดำเนินการจัดการและเตรียมข้อมูลสำหรับการวิเคราะห์อนุกรมเวลา โดยทำการจัดกลุ่มประเภทของโรคมะเร็งตามระบบอวัยวะต่าง ๆ เช่น ระบบทางเดินอาหาร ระบบทางเดินหายใจ ระบบสืบพันธุ์ เพื่อให้สามารถวิเคราะห์แนวโน้มและพัฒนาแบบจำลองการพยากรณ์จำนวนผู้ป่วยโรคมะเร็งได้อย่างมีประสิทธิภาพ

### 3.2 การสร้างแบบจำลอง

การวิเคราะห์ข้อมูลด้วยภาษา Python และไลบรารี ได้แก่ Pandas, NumPy, Stats models และ Scikit-learn โดยแบ่งข้อมูลออกเป็น Training Set (มกราคม 2553 – ธันวาคม 2565) สำหรับสร้างแบบจำลอง และ Test Set (มกราคม 2566 – ธันวาคม 2567) สำหรับทดสอบความแม่นยำของแบบจำลองในการศึกษานี้ได้พัฒนาแบบจำลองการพยากรณ์ 4 วิธี ได้แก่ MA, SES, DES และ ARIMA จากนั้นเปรียบเทียบผลการพยากรณ์เพื่อเลือกแบบจำลองที่มีค่าความคลาดเคลื่อนต่ำที่สุด และตรวจสอบความน่าเชื่อถือด้วยวิธี Walk-Forward Validation โดยทำการฝึกแบบจำลองใหม่ในแต่ละช่วงเวลา และพยากรณ์ค่าถัดไป เพื่อจำลองการใช้งานจริง พร้อมนำเสนอผลในรูปแบบตารางเปรียบเทียบค่าจริงและค่าพยากรณ์รายเดือน




**ภาพที่ 2** : การสร้างแบบจำลองการพยากรณ์ (Flow Diagram)

**3.3 การเปรียบเทียบและเลือกแบบจำลอง**

การศึกษานี้ได้นำข้อมูลจำนวนผู้ป่วยโรคมะเร็งรายเดือนตั้งแต่เดือนมกราคม พ.ศ. 2553 ถึงเดือนธันวาคม พ.ศ. 2567 มาวิเคราะห์ด้วยเทคนิค การวิเคราะห์อนุกรมเวลา (Time Series Analysis) เพื่อพัฒนาและเปรียบเทียบประสิทธิภาพของแบบจำลองการพยากรณ์ โดยใช้เทคนิคการพยากรณ์จำนวน 4 วิธี ได้แก่ MA, SES, DES และ ARIMA การพัฒนาแบบจำลองดำเนินการด้วยภาษา Python และประเมินความแม่นยำของแบบจำลองด้วยค่าความคลาดเคลื่อนทางสถิติ ได้แก่ Mean Absolute Error (MAE) และ Root Mean Square Error (RMSE) โดยแบบจำลองที่ให้ค่าความคลาดเคลื่อนต่ำที่สุดจะถูกพิจารณาว่าเป็นแบบจำลองที่มีประสิทธิภาพและเหมาะสมสำหรับการพยากรณ์จำนวนผู้ป่วย

**3.4 การตรวจสอบความแม่นยำแบบจำลอง**

หลังจากเปรียบเทียบประสิทธิภาพของแบบจำลองจากชุดข้อมูลสำหรับการสร้างแบบจำลอง (Training Set) โดยใช้ตัวชี้วัดค่าความคลาดเคลื่อนของการพยากรณ์ ได้แก่ Mean Absolute Error (MAE) และ Root Mean Square Error (RMSE) แล้ว จะทำการคัดเลือกแบบจำลองที่ให้ค่าความคลาดเคลื่อนต่ำที่สุด เพื่อนำไปทดสอบกับชุดข้อมูลทดสอบ (Test Set) ซึ่งเป็นข้อมูลรายเดือนในช่วงเดือนมกราคม พ.ศ. 2566 ถึงเดือนธันวาคม พ.ศ. 2567 เพื่อประเมินความสามารถของแบบจำลองในการพยากรณ์ข้อมูลใหม่

**3.5 การทดสอบและการประเมินผล**

จากที่ได้พัฒนาแบบจำลองขึ้นมาแล้ว จะต้องมีการประเมินผลระบบ โดยจัดทำแบบประเมินผ่านทาง Google Form และมีรายละเอียดการประเมินผลดังนี้

การประเมินความพึงพอใจของระบบ โดยผู้ประเมินเป็นผู้เชี่ยวชาญด้านการให้รหัสโรค จำนวน 3 คน

การประเมินความพึงพอใจของระบบ โดยผู้ประเมินเป็นกลุ่มผู้ใช้งานระบบ จำนวน 30 คน [3]

## 4. ผลการดำเนินงาน

**4.1 ผลการเปรียบเทียบแบบจำลองการพยากรณ์**

**ตารางที่ 1 :** ผลการเปรียบเทียบแบบจำลองการพยากรณ์

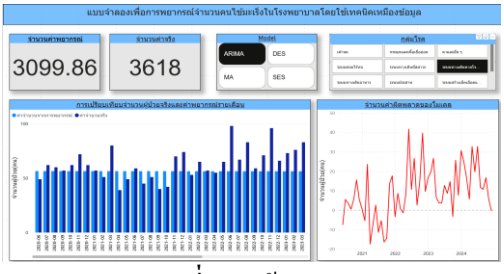
| กลุ่มโรค | MA | SES | DES | ARIMA | Best |
|---|---|---|---|---|---|
| กระดูก | 7.85\|7.14 | 7.76\|7.04 | 7.92\|7.21 | **7.65\|6.85** | ARIMA |
| เต้านม | 14.72\|13.35 | 15.82\|14.50 | 16.09\|14.77 | **14.48\|13.28** | ARIMA |
| ทางเดินอาหาร | **18.13\|16.24** | 19.34\|17.59 | 19.34\|17.59 | 19.41\|17.31 | MA |
| ต่อมไร้ท่อ | 12.79\|11.71 | 12.57\|11.55 | 12.72\|11.69 | **12.32\|11.26** | ARIMA |
| ตาและอื่น ๆ | 1.95\|1.68 | **1.85\|1.78** | 1.95\|1.77 | 2.13\|1.96 | SES |
| สืบพันธุ์หญิง | 8.76\|7.99 | 8.84\|7.96 | 8.91\|8.03 | **8.46\|7.70** | ARIMA |
| ศีรษะและลำคอ | 8.47\|7.62 | 8.31\|7.48 | 8.58\|7.65 | **8.23\|7.45** | ARIMA |
| เลือดและน้ำเหลือง | 7.55\|6.90 | 7.72\|7.01 | 7.78\|7.07 | **7.51\|6.75** | ARIMA |
| สืบพันธุ์ชาย | 12.07\|11.07 | **11.99\|10.84** | 12.6\|11.46 | 13.15\|12.20 | SES |
| ระบบประสาท | 2.81\|2.45 | **2.67\|2.37** | 2.73\|2.43 | 2.74\|2.42 | SES |
| ทางเดินหายใจ | 11.46\|10.18 | **11.09\|9.90** | 11.24\|10.06 | 11.25\|10.04 | SES |
| ปัสสาวะ | **5.81\|5.26** | 5.92\|5.41 | 6.09\|5.51 | 5.91\|5.41 | MA |
| อื่น ๆ | 2.81\|2.52 | **2.77\|2.50** | **2.77\|2.50** | 2.98\|2.69 | SES/DES |
| **ค่าเฉลี่ย** | **8.86\|8.01** | 8.97\|8.15 | 9.13\|8.29 | 8.94\|8.10 | MA |

จากตารางค่า MAE และ RMSE พบว่าแบบจำลอง ARIMA ให้ความแม่นยำดีที่สุดในหลายกลุ่มโรค ขณะที่แบบจำลอง SES มีค่าความคลาดเคลื่อนโดยรวมต่ำในบางกลุ่มโรคพิจารณาจากทั้งค่า MAE และ RMSE ร่วมกัน ประสิทธิภาพของแต่ละแบบจำลองขึ้นอยู่กับลักษณะข้อมูลของแต่ละกลุ่มโรคจึงควรเลือกใช้แบบจำลองให้เหมาะสมกับข้อมูลในแต่ละกรณี

4.2 **ผลการพยากรณ์จำนวนคนไข้มะเร็ง**

สำหรับการวิจัยครั้งนี้มีขั้นตอนการทำงานดังนี้

-หน้าจอ Dashboard




**ภาพที่ 3** : หน้า Dashboard

### 4.3 ผลการประเมิน

การประเมินความพึงพอใจโดยผู้เชี่ยวชาญ 3 คน พบว่า แบบจำลองการพยากรณ์ มีผลการประเมินในทุกผลการประเมินคุณภาพของแบบประเมิน โดยผู้เชี่ยวชาญในภาพรวมมีค่าเฉลี่ย 3.91 และส่วนเบี่ยงเบนมาตรฐาน 0.35 อยู่ในระดับ ดี แสดงว่าแบบประเมินมีความเหมาะสมและสามารถใช้ประเมินระบบได้อย่างมีประสิทธิภาพ โดยเมื่อพิจารณารายด้านพบว่า ด้านความเหมาะสมของข้อมูลและวิธีการวิเคราะห์ มีค่าเฉลี่ยอยู่ในระดับ ดีถึงดีมาก และประเด็น ระยะเวลาข้อมูลย้อนหลัง 180 เดือนมีความเพียงพอสำหรับการวิเคราะห์อนุกรมเวลา และการใช้ข้อมูลรายเดือนสามารถอธิบายแนวโน้มจำนวนผู้ป่วยได้ชัดเจน มีค่าเฉลี่ยสูงที่สุด 4.33 อยู่ในระดับ ดีมาก

**ตารางที่ 2 :** ผลการประเมินความพึงพอใจของแบบจำลองโดยผู้ใช้งานโดยผู้เชี่ยวชาญ

| รายการประเมิน | $\bar{x}$ | S.D. | ความพึงพอใจ |
|---|---|---|---|
| **1. ความสอดคล้องของเนื้อหาและตัวชี้วัดกับวัตถุประสงค์การวิจัย** | | | |
| 1.1 จำนวนผู้ป่วยที่แสดงบน Dashboard สามารถสะท้อนสถานการณ์ผู้ป่วยของโรงพยาบาลของรัฐได้อย่างเหมาะสม | 4.00 | 0 | ดี |
| 1.2 การจำแนกข้อมูลตามกลุ่มโรคมีความชัดเจนและเหมาะสมต่อการวิเคราะห์ | 4.00 | 0 | ดี |
| 1.3 ข้อมูลที่ใช้มีความสอดคล้องกับหลักการวิเคราะห์ข้อมูลอนุกรมเวลา | 3.67 | 0.58 | ดี |
| 1.4 ข้อมูลค่าความคลาดเคลื่อนที่แสดงบน Dashboard สามารถเปรียบเทียบได้อย่างง่าย | 3.33 | 0.58 | ปานกลาง |
| 1.5 ข้อมูลที่แสดงบน Dashboardโดยรวมสอดคล้องกับวัตถุประสงค์ของการวิจัย | 4.00 | 0 | ดี |
| **2. ความเหมาะสมของข้อมูลและวิธีการวิเคราะห์** | | | |
| 2.1 ชุดข้อมูลผู้ป่วยจากโรงพยาบาลรัฐมีความเหมาะสมสำหรับการพยากรณ์ | 3.67 | 0.58 | ดี |
| 2.2 ระยะเวลาข้อมูลย้อนหลัง 180 เดือนมีความเพียงพอสำหรับการวิเคราะห์อนุกรมเวลา | 4.33 | 0.58 | ดีมาก |
| 2.3 การใช้ชุดข้อมูลในรูปแบบรายเดือนมีความเหมาะสมและสามารถอธิบายแนวโน้มการเปลี่ยนแปลงของจำนวนผู้ป่วยได้อย่างชัดเจน | 4.33 | 0.58 | ดีมาก |
| 2.4 การเลือกใช้เทคนิค MA ,SES, DES และ ARIMA มีความเหมาะสมกับลักษณะข้อมูล | 4.00 | 0 | ดี |
| 2.5 การจำแนกข้อมูลและการนำเสนอผลสามารถเปรียบเทียบค่าพยากรณ์กับค่าจริงได้อย่างชัดเจนและถูกต้อง | 4.00 | 0 | ดี |
| **3. ความเหมาะสมของการนำเสนอผลด้วย Dashboard (Power BI)** | | | |
| 3.1 Dashboard สามารถแสดงผลการวิเคราะห์อย่างง่าย ช่วยให้เข้าใจผลการพยากรณ์ได้ชัดเจน | 3.67 | 0.58 | ดี |
| 3.2 Dashboard สามารถเปรียบเทียบข้อมูลระหว่างการค่าพยากรณ์ และ ค่าจริง แสดงผลได้เหมาะสม | 3.67 | 0.58 | ดี |
| 3.3 รูปแบบกราฟและแผนภาพช่วยให้เข้าใจข้อมูลได้ง่าย | 3.67 | 0.58 | ดี |
| 3.4 Dashboard สามารถนำไปใช้สนับสนุนการตัดสินใจหรือการวางแผนด้านผู้ป่วยได้อย่างมีประสิทธิภาพ | 4.00 | 0.58 | ดี |
| 3.5 การนำเสนอผลด้วย Power BI มีความเหมาะสมเพื่อนำไปใช้เป็นข้อมูลเชิงวิเคราะห์เบื้องต้น | 4.33 | 0.58 | ดีมาก |

**ตารางที่ 3 :** ผลการประเมินความพึงพอใจของแบบจำลองโดยผู้ใช้งาน

| รายการประเมิน | $\bar{x}$ | S.D. | ความพึงพอใจ |
|---|---|---|---|
| 1. ความพึงพอใจด้านความชัดเจนและการใช้งานง่าย | | | |
| 1.1 ความพึงพอใจต่อ Dashboard แสดงข้อมูลการพยากรณ์และข้อมูลจริงได้ชัดเจน | 4.31 | 0.60 | ดีมาก |
| 1.2 ความพึงพอใจต่อ กราฟเส้นสามารถเปรียบเทียบและช่วยให้เข้าใจแนวโน้มได้โดยง่าย | 4.52 | 0.63 | ดีมาก |
| 1.3 ความพึงพอใจต่อ การใช้สีและรูปแบบการแสดงผลมีความเหมาะสม | 4.59 | 0.73 | ดีมาก |
| 1.4 ความพึงพอใจต่อ การจัดวางองค์ประกอบโดยภาพรวมบนหน้าจอแสดงมีความเหมาะสมเป็นระเบียบ | 4.48 | 0.74 | ดีมาก |
| 1.5 ความพึงพอใจโดยรวม ด้านความชัดเจนและการใช้งานง่าย | 4.45 | 0.60 | ดีมาก |
| 2. ความพึงพอใจด้านการเปรียบเทียบและวิเคราะห์ข้อมูล | | | |
| 2.1 ความพึงพอใจต่อ การเลือกกลุ่มโรคทำได้สะดวกและเข้าใจง่าย | 4.52 | 0.51 | ดีมาก |
| 2.2 ความพึงพอใจต่อ Dashboard ตอบสนองรวดเร็วเมื่อมีการเปลี่ยนตัวกรองข้อมูล | 4.66 | 0.55 | ดีมาก |
| 2.3 ความพึงพอใจต่อ Dashboard ช่วยให้เห็นความแตกต่างระหว่างค่าพยากรณ์และค่าจริงได้ชัดเจนถูกต้องแม่นยำ | 4.31 | 0.60 | ดีมาก |
| 2.4 ความพึงพอใจต่อ Dashboard ช่วยให้เปรียบเทียบข้อมูลแนวโน้มได้จริง | 4.55 | 0.51 | ดีมาก |
| 2.5 ความพึงพอใจต่อ Dashboard สนับสนุนการวิเคราะห์เชิงเปรียบเทียบได้อย่างมีประสิทธิภาพ | 4.41 | 0.63 | ดีมาก |
| 3. ความพึงพอใจนำไปใช้ประโยชน์และความพึงพอใจโดยรวม | | | |
| 3.1 ความพึงพอใจต่อ Dashboard ช่วยให้เห็นภาพรวมของแต่ละกลุ่มโรค | 4.48 | 0.78 | ดีมาก |
| 3.2 ความพึงพอใจต่อ Dashboard ช่วยวางแผนการตัดสินใจและสนับสนุนทางด้านการแพทย์สาธารณสุข | 4.45 | 0.63 | ดีมาก |
| 3.3 ความพึงพอใจต่อ Dashboard มีประโยชน์นำไปต่อยอดใช้ในงานจริงได้ | 4.45 | 0.63 | ดีมาก |
| 3.4 โดยรวมแล้ว Dashboard มีประโยชน์ต่อผู้ใช้งาน | 4.45 | 0.69 | ดีมาก |
| 3.5 โดยรวมแล้ว Dashboard มีความพึงพอใจในภาพรวม | 4.45 | 0.63 | ดีมาก |

การประเมินความพึงพอใจของแพลตฟอร์มโดยผู้ใช้งานจริงจำนวน 30 คน สามารถสรุปได้เป็นดังนี้

พบว่า ผลการประเมินความพึงพอใจของผู้ใช้งานทั่วไปโดยภาพรวมอยู่ในระดับ ดีมาก โดยมีค่าเฉลี่ยเท่ากับ 4.48 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.64 แสดงให้เห็นว่าระบบ Dashboard ที่พัฒนาขึ้นสามารถนำเสนอข้อมูล

การพยากรณ์และข้อมูลจริงได้อย่างชัดเจนและสามารถนำไปใช้วิเคราะห์ข้อมูลได้อย่างมีประสิทธิภาพเมื่อพิจารณาเป็นรายด้าน พบว่า ด้าน ความชัดเจนและการใช้งานง่าย มีค่าเฉลี่ยอยู่ในระดับดีมาก โดยเฉพาะประเด็นการใช้สีและรูปแบบการแสดงผลมีความเหมาะสม ซึ่งมีค่าเฉลี่ยสูงที่สุดเท่ากับ 4.59 แสดงให้เห็นว่าการออกแบบ Dashboard มีความเหมาะสมและช่วยให้ผู้ใช้งานสามารถเข้าใจข้อมูลได้ง่าย

พบว่า ผู้ใช้งานมีความพึงพอใจในระดับดีมาก โดยประเด็น Dashboard ตอบสนองรวดเร็วเมื่อมีการเปลี่ยนตัวกรองข้อมูล มีค่าเฉลี่ยสูงที่สุดเท่ากับ 4.66 แสดงให้เห็นว่าระบบสามารถตอบสนองต่อการใช้งานและการวิเคราะห์ข้อมูลได้อย่างรวดเร็วและมีประสิทธิภาพ

## 5. สรุป

### 5.1 สรุปผลการวิจัย

จากผลการวิจัยพบว่า แบบจำลองการพยากรณ์จำนวนผู้ป่วยโรคมะเร็งด้วยเทคนิคอนุกรมเวลา แบบจำลอง SES มีความแม่นยำโดยรวมสูงที่สุดบางกกลุ่มโรค ขณะที่แบบจำลอง ARIMA มีความแม่นยำโดยรวมรองลงมา แสดงให้เห็นว่าแต่ละแบบจำลองมีความเหมาะสมแตกต่างกันตามลักษณะของข้อมูล

นอกจากนี้ ระบบ Dashboard ที่พัฒนาขึ้นสามารถนำเสนอข้อมูลการพยากรณ์และข้อมูลจริงได้อย่างชัดเจนและสนับสนุนการวิเคราะห์ข้อมูลได้อย่างมีประสิทธิภาพ โดยผลการประเมินจากผู้เชี่ยวชาญและผู้ใช้งานอยู่ในระดับดีถึงดีมาก สะท้อนให้เห็นว่าระบบนำไปประยุกต์ใช้ในการวางแผนด้านสาธารณสุขได้

### 5.2 อภิปรายผลการวิจัย

จากผลการวิจัยพบว่า แบบจำลองการพยากรณ์จำนวนผู้ป่วยโรคมะเร็งสามารถให้ผลลัพธ์ที่มีความแม่นยำ โดยแบบจำลอง ARIMA ให้ค่าความคลาดเคลื่อนต่ำที่สุดในหลายกลุ่มโรค ขณะที่แบบจำลอง SES ให้ค่าความคลาดเคลื่อนเฉลี่ยสูงที่สุด สะท้อนให้เห็นว่าแต่ละแบบจำลองมีความเหมาะสมแตกต่างกันตามลักษณะของข้อมูลด้านระบบ Dashboard ที่พัฒนาขึ้น พบว่านำเสนอข้อมูลการพยากรณ์และข้อมูลจริงได้อย่างชัดเจน และสนับสนุนการวิเคราะห์ข้อมูลได้เป็นอย่างดี โดยผลการประเมินจากผู้เชี่ยวชาญอยู่ในระดับดี (ค่าเฉลี่ย 3.91) และผู้ใช้งานอยู่ในระดับดีมาก (ค่าเฉลี่ย 4.48) แสดงให้เห็นว่าระบบมีความเหมาะสมต่อการใช้งานจริง

การศึกษาครั้งนี้แสดงให้เห็นว่า การออกแบบระบบแสดงผลข้อมูลที่มีโครงสร้างชัดเจน ใช้งานง่าย และสามารถเปรียบเทียบข้อมูลได้อย่างสะดวก ช่วยเพิ่มประสิทธิภาพในการวิเคราะห์ข้อมูลและความพึงพอใจของผู้ใช้งาน อีกทั้งระบบ Dashboard [4] ที่พัฒนาขึ้นยังนำไปประยุกต์ใช้เป็นเครื่องมือสนับสนุนการวิเคราะห์ข้อมูลด้านสาธารณสุข และต่อยอดในงานวิจัยที่เกี่ยวข้องได้

อย่างไรก็ตาม งานวิจัยนี้มีข้อจำกัดคือใช้ข้อมูลจากโรงพยาบาลเพียงแห่งเดียว และไม่ได้พิจารณาปัจจัยภายนอกที่อาจส่งผลต่อจำนวนผู้ป่วย

### 5.3 ข้อเสนอแนะ

5.3.1 ควรพิจารณาปรับปรุงคุณภาพข้อมูลก่อนการวิเคราะห์ โดยเฉพาะข้อมูลในช่วงเวลาที่ได้รับผลกระทบจากปัจจัยภายนอก เช่น การแพร่ระบาดของ COVID-19 หรือเหตุการณ์ภัยพิบัติธรรมชาติ ซึ่งอาจทำให้ข้อมูลมีความผันผวนผิดปกติ ดังนั้นควรพิจารณาตัดหรือปรับข้อมูลบางช่วงออกเพื่อให้แบบจำลองสามารถพยากรณ์ได้แม่นยำมากยิ่งขึ้น

5.3.2 ควรเพิ่มการเปรียบเทียบแบบจำลองการพยากรณ์ด้วยอัลกอริทึมอื่นเพิ่มเติม เช่น SARIMA, Regression, Support Vector Regression (SVR), LSTM และ Deep Learning เพื่อศึกษาประสิทธิภาพของแบบจำลองในหลายแนวทาง และช่วยคัดเลือกวิธีการพยากรณ์ที่เหมาะสมกับลักษณะข้อมูลผู้ป่วยโรคมะเร็งได้ดียิ่งขึ้นในอนาคต

## เอกสารอ้างอิง

[1] Hyndman, R. J., & Athanasopoulos, G. (2021). Forecasting: Principles and practice (3rd ed.). OTexts.

[2] Hyndman, R. J., & Koehler, A. B. (2006). Another look at measures of forecast accuracy. International Journal of Forecasting, 22(4), 679–688.

[3] DeLone, W. H., & McLean, E. R. (2003). The DeLone and McLean model of information systems success. Journal of Management Information Systems, 19(4), 9–30.

[4] Chen, H., Chiang, R. H. L., & Storey, V. C. (2012). Business intelligence and analytics. MIS Quarterly, 36(4), 1165–1188.

[5] ขอสกุล ส., ชยธวัช ว., & กิ่งกระโทก ธ. (2024). การพยากรณ์จำนวนผู้ป่วยโรคไข้เลือดออก. วารสารความปลอดภัยและสุขภาพ, 17(2), 160–176.

[6] วรียา วงศ์พานิช. (2567). การพยากรณ์ข้อมูลอนุกรมเวลาทางการแพทย์ด้วยตัวแบบอารีมา.

# ระบบการตรวจจับพฤติกรรมปลาอัจฉริยะโดยใช้เทคนิคการประมวลผลภาพ
# Intelligent Fish Behavior Detection System Using Image Processing

*มลธิชา จิมไธสง (Monthicha JImthaisong)[1] และมหศักดิ์ เกตุฉ่ำ (Mahasak Ketcham)[2]*
*ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล*
*มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ*
*[1]s6607021810010@email.kmutnb.ac.th, [2]mahasak.k@itd.kmutnb.ac.th*

## บทคัดย่อ

*ปลาคาร์ปเป็นปลาสวยงามที่มีมูลค่าทางการตลาดที่สูง และได้รับความนิยมเป็นอย่างมากในปัจจุบัน การดูแลสุภาพของปลาจึงเป็นโจทย์สำคัญ งานวิจัยนี้นำเสนอการพัฒนาระบบตรวจจับพฤติกรรมของปลาคาร์ปโดยใช้เทคนิคการประมวลผลด้วยภาพ (Image Processing) ร่วมกับเทคนิคการเรียนรู้เชิงลึก (Deep Learning) โดยเปรียบเทียบประสิทธิภาพของโมเดล ได้แก่ YOLOv5, YOLOv8, YOLOv5+CNN และ YOLOv8+CNN ในการตรวจจับพฤติกรรมปลาที่ผิดปกติ 5 ประเภท ได้แก่ การว่ายน้ำผิดรูปแบบ การถูตัวกับพื้น พฤติกรรมอยู่นิ่ง/ซึม การฮุบอากาศที่ผิวน้ำตลอดเวลา และการกระตุก/ชัก จากการทดสอบพบว่าโมเดลที่มีประสิทธิภาพสูงที่สุดคือ YOLOv8+CNN ซึ่งมีค่า Accuracy เท่ากับ 96.00% Precision 96.36% Recall 96.00% และ F1-score 96.00% แสดงให้เห็นว่าการทำงานร่วมกันของ YOLOv8 กับ CNN สามารถเพิ่มความสามารถในการตรวจจับและจำแนกพฤติกรรมของปลาได้อย่างมีประสิทธิภาพมากที่สุด*

**คำสำคัญ:** แบบจำลองการตรวจจับวัตถุแบบเรียลไทม์ การประมวลผลภาพ โมเดลโครงข่ายประสาทเทียมแบบคอนโวลูชัน ปลาแฟนซีคาร์ป การวิเคราะห์พฤติกรรม การตรวจจับวัตถุ

## Abstract

*Koi fish are highly valued ornamental fish that have gained significant popularity. Consequently, health monitoring has become a critical challenge. This research presents the development of a behavior-detection system for koi fish using image processing and deep learning techniques. The study evaluates and compares the performance of several models: YOLOv5, YOLOv8, YOLOv5+CNN, and YOLOv8 +CNN in detecting five categories of abnormal behaviors: Erratic Swimming (ERR), Rubbing (RUB), Lethargy (LET), Surface Gasping (GAS), and Twitching (TWI). Experimental results indicate that the YOLOv8+CNN model achieved the highest performance, with an accuracy of 96.00%, precision of 96.36%, recall of 96.00%, and an F1-score of 96.00%. These findings demonstrate that the synergy between YOLOv8 and CNN effectively enhances behavior detection and classification capabilities with maximum efficiency.*

**Keywords:** YOLO Model, Image Processing, CNN Model, Koi, Carp breed, Behavior Analysis, Object Detection.

## 1. บทนำ

อุตสาหกรรมการเพาะเลี้ยงปลาคาร์ปมีบทบาทสำคัญในด้านเศรษฐกิจ เนื่องจากปลาสวยงามที่ได้รับ ความนิยม

สูง อย่างไรก็ตาม ปลาคาร์ปเป็นสัตว์ที่อ่อนไหวต่อสภาพแวดล้อม และบางสายพันธุ์มีมูลค่าสูง ทำให้ การดูแลสุขภาพและตรวจสอบความผิดปกติอย่างทันท่วงทีมีความสำคัญอย่างยิ่ง แต่ฟาร์มปลาคาร์ปหลายแห่งยังขาดระบบตรวจจับพฤติกรรมปลาที่มีประสิทธิภาพ หากพบความผิดปกติล่าช้าอาจนำสู่การระบาดครั้งใหญ่ได้ ดังเช่นกรณีการระบาดของไวรัสเริมในปลาคาร์ป (Koi Herpes Virus: KHV) ที่จังหวัดปทุมธานี จนต้องทำลายปลากว่า 800 ตัว[1] และกรณีการระบาดของแบคทีเรีย Lactococcus petaurid ในปลาเทราต์ที่สหรัฐอเมริกา ที่สูญเสียปลาไปกว่า 300,000 ตัว[2] ปัญหาเหล่านี้สะท้อนถึงความเสี่ยงที่ร้ายแรงของการระบาด โดยเฉพาะในช่วงที่สภาพแวดล้อมไม่เอื้ออำนวย งานวิจัยนี้จึงมุ่งพัฒนาระบบ AI ร่วมกับการประมวลผลภาพเพื่อตรวจสอบพฤติกรรมปลาแบบเรียลไทม์และแจ้งเตือนเมื่อพบความผิดปกติ เพื่อลดความเสียหายและส่งเสริมการจัดการฟาร์มอย่างยั่งยืน

การวิจัยนี้มีวัตถุประสงค์หลัก 3 ประการ ได้แก่ การศึกษาระบบตรวจจับพฤติกรรมปลาอัจฉริยะ การพัฒนาโมเดล AI สำหรับตรวจจับพฤติกรรมผิดปกติ และการประเมินประสิทธิภาพของโมเดลที่พัฒนาขึ้น โดยมุ่งเน้นการตรวจจับพฤติกรรมผิดปกติ 5 รูปแบบ คือ การว่ายน้ำผิดรูปแบบ การถูตัวกับวัตถุ การซึม/อยู่นิ่ง การสุบอากาศที่ผิวน้ำ และการกระตุก/ชัก โดยทั้งหมดนี้ดำเนินการภายในระยะเวลา 180 วัน

ประโยชน์ที่คาดว่าจะได้รับจากงานวิจัยนี้มีทั้งในเชิงปฏิบัติและเชิงวิชาการ ในด้านปฏิบัติ ระบบจะเป็นเครื่องมือสำคัญสำหรับผู้เลี้ยงปลาในการตรวจจับพฤติกรรมผิดปกติเบื้องต้น ช่วยให้สามารถดำเนินการแก้ไขได้ทันท่วงที ลดความเสียหายจากการตายของปลาและการแพร่ระบาดของโรคในวงกว้าง ซึ่งส่งผลโดยตรงต่อการลดต้นทุนและเพิ่มประสิทธิภาพการจัดการฟาร์ม ในด้านวิชาการ ข้อมูลที่ได้จากระบบสามารถนำไปต่อยอดการวิจัยเชิงลึก และพัฒนาโมเดลที่มีความแม่นยำสูงขึ้นในอนาคต รวมถึงเปิดโอกาสสำหรับการนำไปใช้ในเชิงพาณิชย์เพื่อยกระดับอุตสาหกรรมการเพาะเลี้ยงปลาคาร์ปของไทยอย่างเป็นรูปธรรม

## 2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

การดำเนินการวิจัยพัฒนาต้นแบบปัญญาดิษฐ์เพื่อตรวจจับพฤติกรรมปลา มีทฤษฎีและงานวิจัยที่เกี่ยวข้องดังต่อไปนี้

### 2.1 ปัญญาประดิษฐ์ (Artificial Intelligence: AI)

ปัญญาประดิษฐ์คือความสามารถของเครื่องจักรหรือซอฟต์แวร์ในการจำลองกระบวนการคิดของมนุษย์ เช่น การเรียนรู้ การใช้เหตุผล และการแก้ปัญหา ซึ่งนำมาประยุกต์ใช้ในหลายสาขา รวมถึงการตรวจจับพฤติกรรมสัตว์น้ำในงานวิจัยนี้

### 2.2 คอมพิวเตอร์วิทัศน์ (Computer Vision)

คอมพิวเตอร์วิทัศน์คือสาขาที่พัฒนาให้คอมพิวเตอร์สามารถตีความและเข้าใจข้อมูลจากภาพหรือวิดีโอได้ โดยใช้เทคนิคต่าง ๆ เช่น การตรวจจับวัตถุ การจำแนกประเภท และการติดตามการเคลื่อนไหว ซึ่งเป็นพื้นฐานสำคัญของระบบตรวจจับพฤติกรรมปลาในงานวิจัยนี้

### 2.3 การเรียนรู้เชิงลึก (Deep Learning)

การเรียนรู้เชิงลึกเป็นสาขาหนึ่งของ AI ที่ใช้โครงข่ายประสาทเทียมหลายชั้นในการเรียนรู้ลักษณะเฉพาะของข้อมูลโดยอัตโนมัติ มีความสามารถสูงในการประมวลผลข้อมูลภาพและวิดีโอ เหมาะสมกับงานตรวจจับพฤติกรรมปลาที่มีความซับซ้อนและหลากหลาย

### 2.4 โครงข่ายประสาทเทียมแบบคอนโวลูชัน (CNN)

CNN เป็นโครงสร้างโครงข่ายประสาทเทียมที่ออกแบบมาเพื่อประมวลผลข้อมูลภาพโดยเฉพาะ ประกอบด้วยชั้น Convolutional Layer ที่ดึงคุณลักษณะจากภาพ ชั้น Pooling Layer ที่ลดขนาดข้อมูล และชั้น Fully Connected Layer ที่ใช้จำแนกประเภท CNN ถูกนำมาใช้ในงานวิจัยนี้เพื่อวิเคราะห์พฤติกรรมปลาคาร์ปจากเฟรมวิดีโอ

### 2.5 การประมวลผลภาพ (Image Processing)

การประมวลผลภาพเป็นกระบวนการแปลง ปรับปรุง และวิเคราะห์ภาพดิจิทัล เพื่อให้ได้ข้อมูลที่มีคุณภาพเหมาะสมสำหรับการนำเข้าโมเดล AI โดยในงานวิจัยนี้ใช้เทคนิค เช่น การปรับขนาดภาพ การเพิ่มความหลากหลายของข้อมูล (Augmentation) และการปรับความเข้มแสง เพื่อเพิ่มประสิทธิภาพการเรียนรู้ของโมเดล

### 2.6 อัลกอริทึมตรวจจับวัตถุแบบเรียลไทม์ (YOLO)

YOLO (You Only Look Once) เป็นอัลกอริทึมตรวจจับวัตถุที่มีความเร็วสูงและแม่นยำ สามารถประมวลผลภาพทั้งภาพในครั้งเดียว งานวิจัยนี้เลือกใช้ YOLOv5 และ YOLOv8 เนื่องจาก YOLOv5 รองรับ Transfer Learning ซึ่งเหมาะกับชุดข้อมูลขนาดไม่ใหญ่มาก ส่วน YOLOv8 มีสถาปัตยกรรม Anchor-free ที่ให้ความแม่นยำสูงขึ้นและสร้างสมดุลระหว่างความเร็วกับความแม่นยำได้ดี เหมาะสำหรับการตรวจจับแบบเรียลไทม์

### 2.7 ปลาคาร์ป

ปลาคาร์ปเป็นปลาน้ำจืดที่มีความสำคัญทางเศรษฐกิจและวัฒนธรรมทั่วโลก มีต้นกำเนิดจากเอเชียตะวันออก ลักษณะเด่นคือมีรูปร่างยาวรี หัวใหญ่ มีหนวด สามารถโตได้ถึง 1 เมตร และมีอายุยืนยาวถึง 40–50 ปีในธรรมชาติ งานวิจัยนี้ศึกษาพฤติกรรมผิดปกติพื้นฐาน 5 รูปแบบของปลา ได้แก่ การว่ายน้ำผิดรูปแบบ การถูตัวกับวัตถุ การซึม/อยู่นิ่ง การสุบอากาศที่ผิวน้ำ และการกระตุก/ชัก ซึ่งแต่ละพฤติกรรมล้วนเป็นสัญญาณบ่งชี้ปัญหาสุขภาพหรือสภาพแวดล้อมที่ไม่เหมาะสม

### 2.8 งานวิจัยที่เกี่ยวข้อง

มีงานวิจัยที่เกี่ยวข้องทั้งหมด 40 เรื่อง ครอบคลุมการตรวจจับและจำแนกปลาในสภาพแวดล้อมต่าง ๆ การวิเคราะห์พฤติกรรมสัตว์น้ำ และการประมวลผลภาพด้วย Deep Learning โดยเทคนิคที่ใช้หลักได้แก่ YOLO (v3–v8), CNN, และ Vision Transformer ผลการวิจัยส่วนใหญ่มีความแม่นยำอยู่ในช่วง 79–99% โดยงานที่เกี่ยวข้องโดยตรงกับพฤติกรรมปลา เช่น Huang และคณะ (2022) ที่ใช้ GCN วิเคราะห์พฤติกรรมฝูงปลาได้ความแม่นยำสูงถึง 97.3% และ Abangan และคณะ (2023) ที่ใช้ Deep Learning จำแนกพฤติกรรมปลาได้ถึง 85% สะท้อนให้เห็นศักยภาพของ AI ในการนำมาประยุกต์ใช้กับระบบตรวจจับพฤติกรรมปลาคาร์ปในงานวิจัยนี้

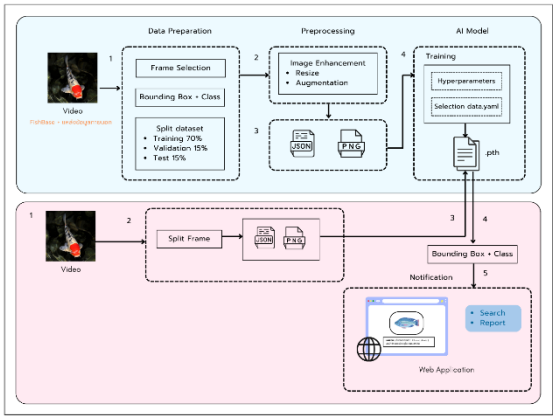
## 3. วิธีการดำเนินการวิจัย

### 3.1 แผนการดำเนินการวิจัย

กำหนดระยะเวลาดำเนินการวิจัย 180 วัน (6 เดือน) ครอบคลุมขั้นตอนหลัก ได้แก่ การศึกษาและรวบรวมชุดข้อมูล การเตรียมและปรับปรุงข้อมูล การพัฒนาและฝึกสอน AI การออกแบบและพัฒนาเว็บแอปพลิเคชันการทดสอบโมเดล และการสรุปประเมินผล

### 3.2 กรอบแนวคิดในการวิจัย

ระบบแบ่งออกเป็น 4 ส่วนหลัก ได้แก่ การเตรียมข้อมูล (Data Preparation) ที่ใช้วิดีโอปลาคาร์ปมาเลือกเฟรมและกำหนดคลาสพฤติกรรมปกติและผิดปกติ 5 ประเภท การแบ่งชุดข้อมูล กระบวนการประมวลผลภาพก่อนนำเข้าโมเดล และ Inference Pipeline ที่รับวิดีโอใหม่ แยกเฟรมวิเคราะห์ด้วย AI แล้วส่งผล Bounding Box พร้อมแจ้งเตือนบนเว็บแอปพลิเคชัน ดังภาพที่ 1




**ภาพที่ 1:** กรอบแนวคิดระบบตรวจจับพฤติกรรมปลา

### 3.3 การรวบรวมและเตรียมข้อมูล

รวบรวมข้อมูลจากฐานข้อมูล Roboflow และแหล่งข้อมูลภายนอก เช่น YouTube และ Facebook แล้วเตรียมข้อมูลใน 3 ขั้นตอนหลัก ได้แก่

1) การคัดเลือกเฟรม แบ่งเป็น 2 แนวทาง คือ ใช้ Keyframe Extraction ร่วมกับวิธี Histogram-based และ YOLO เพื่อเลือกเฟรมที่มีการเปลี่ยนแปลงสูงสุดทุก 1 วินาที ขนาด 5 เฟรมต่อช่วง เพื่อวิเคราะห์พฤติกรรมของปลา

2) ใช้ YOLO ตรวจจับตำแหน่งปลา โดยกรอบต้องครอบคลุมตัวปลาทั้งหมด จากนั้นตรวจสอบและแก้ไขความถูกต้อง กำหนดป้ายพฤติกรรม 5 คลาส ได้แก่ ERR (ว่ายน้ำผิดรูปแบบ) RUB (ถูตัวกับวัตถุ) LET (ซึม/อยู่นิ่ง)

GAS (สูบอากาศที่ผิวน้ำ) และ TWI (กระตุก/ชัก) ส่งออกข้อมูลในรูปแบบมาตรฐาน YOLO

3) การแบ่งชุดข้อมูล แบ่งเป็น Training Set 70%, Validation Set 15% และ Test Set 15% ดังตารางที่ 1

**ตารางที่ 1:** การแบ่งชุดข้อมูลในแต่ละคลาส

| Class | Total | Train | Validation | Test |
|---|---|---|---|---|
| ERR | 627 | 439 | 94 | 94 |
| RUB | 627 | 439 | 94 | 94 |
| LET | 627 | 439 | 94 | 94 |
| GAS | 627 | 439 | 94 | 94 |
| TWI | 627 | 439 | 94 | 94 |
| รวม | 3,135 | 2,195 | 470 | 470 |

**3.4 การพัฒนาและฝึกสอนแบบจำลอง**

ได้ออกแบบเงื่อนไขการทดลอง โดยใช้ชุดข้อมูลชุดเดียวกัน แบ่งคลาส 5 พฤติกรรม Data augmentation เหมือนกันทุกโมเดล รวมถึงการใช้ Environment เดียวกัน ซึ่งสถาปัตยกรรมของแต่ละโมเดล ดังนี้

1) YOLOv5 เป็น single-stage detector ประกอบด้วย CSPDarknet53, PANet และ Detection head สำหรับทำนาย bounding box, confidence และ class โดยตรง

2) YOLOv8 ปรับปรุงจาก YOLOv5 โดยใช้ Anchor-free detection, C2f backbone, Decoupled detection head และ Distribution Focal Loss (DFL) เพื่อเพิ่มความแม่นยำและความเร็วในการตรวจจับ

3) YOLOv5 + CNN เป็น Two-stage model ใช้ YOLOv5 ตรวจจับตำแหน่งวัตถุ และตัดภาพเฉพาะบริเวณที่ตรวจพบ (ROI) ก่อนส่งภาพเข้า CNN เพื่อจำแนกพฤติกรรมอย่างละเอียด

4) YOLOv8 + CNN ใช้ YOLOv8 ตรวจจับตำแหน่งวัตถุและตัดภาพเฉพาะบริเวณที่ตรวจพบ (ROI) ก่อนส่งภาพเข้า CNN เพื่อจำแนกพฤติกรรมอย่างละเอียด โดยกำหนดค่าการฝึกโมเดล YOLO ดังตารางที่ 2

**ตารางที่ 2:** การตั้งค่าการฝึกโมเดล

| Parameter | YOLO Models |
|---|---|
| Image size | 640×640 |
| Epochs | 100 |
| Batch size | 16 |
| Optimizer | SGD |
| Learning rate | 0.001 |

CNN classifier ถูกฝึก 30 epochs ด้วย Adam optimizer และ learning rate 0.0001

**3.5 การประเมินประสิทธิภาพ**

ใช้ตัวชี้วัดมาตรฐาน 6 ตัว ได้แก่ Accuracy, Precision, Recall, F1-Score, ตรวจสอบคุณภาพด้วยค่า IoU $\geq$ 0.75 และค่าความแม่นยำโดยรวม mAP@0.5

**3.6 ขั้นตอนการตรวจจับพฤติกรรมที่ผิดปกติของปลา**

กระบวนการทำงานของระบบมี 5 ขั้นตอน ได้แก่ นำเข้าวิดีโอ (Input Video) แยกเฟรม 5 เฟรมต่อวินาทีในรูปแบบ JSON และ PNG (Split Frame) ประมวลผลด้วย AI Model (.pth) ระบุตำแหน่งและประเภทพฤติกรรม พร้อมค่าความเชื่อมั่น (Bounding Box & Class) และแสดงผลพร้อมแจ้งเตือนบนเว็บแอปพลิเคชัน (Notification & Web Application)

**3.7 เกณฑ์การประเมินการตรวจจับพฤติกรรมปลา**

ประเมินใน 4 ด้าน โดยคัดเลือกคลิปอย่างละ 5 คลิปต่อพฤติกรรม ขั้นตอนการประมวลผลภาพตั้งแต่นำเข้าวิดีโอจนถึงรายงานผล สภาพแวดล้อมที่หลากหลายเพื่อทดสอบความทนทานของโมเดล และเกณฑ์การประเมินด้วยตัวชี้วัด Accuracy, Precision, Recall, F1-Score, IoU $\geq$ 0.75 และ mAP@0.5

## 4. ผลการดำเนินงาน

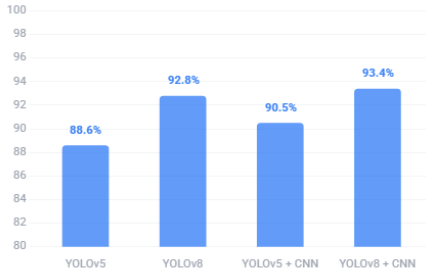
งานวิจัยนี้พัฒนาระบบตรวจจับพฤติกรรมผิดปกติของปลาคาร์ป 5 ประเภทจากวิดีโอ โดยเปรียบเทียบโมเดล 4 โมเดล ได้แก่ YOLOv5, YOLOv8, YOLOv5+CNN และ YOLOv8+CNN ภายใต้สภาพแวดล้อมที่หลากหลาย พบว่า YOLOv8+CNN มีประสิทธิภาพสูงสุดในทุกตัวชี้วัด

**4.1 ผลการเปรียบเทียบประสิทธิภาพของโมเดล 4 แบบ**

งานวิจัยนี้เปรียบเทียบโมเดล 4 รูปแบบ สำหรับตรวจจับพฤติกรรมผิดปกติของปลา 5 ประเภท ได้แก่ ERR, RUB, LET, GAS และ TWI ดังตารางที่ 3

**ตารางที่ 3:** ผลการเปรียบเทียบประสิทธิภาพของโมเดลทั้ง 4 แบบ

| Model | Acc. | P | R | F1 |
|---|---|---|---|---|
| YOLO v5 | 86.40% | 87.12% | 85.90% | 86.50% |
| YOLO v8 | 89.70% | 90.45% | 89.10% | 89.76% |
| YOLO v5 +CNN | 91.20% | 92.05% | 90.75% | 91.39% |
| YOLO v8 +CNN | 92.30% | 93.18% | 91.80% | 92.48% |




**ภาพที่ 2:** ผลการเปรียบเทียบ mAP@0.5 โมเดลทั้ง 4 แบบ

จากตารางที่ 1 และผลการเปรียบเทียบ mAP@0.5 ในภาพที่ 4 จะเห็นว่าโมเดล YOLOv8+CNN มีประสิทธิภาพสูงสุดในทุกตัวชี้วัด โดยการเพิ่ม CNN ช่วยให้โมเดลสกัดคุณลักษณะเชิงพื้นที่ (Spatial Features) ได้ดียิ่งขึ้น และลดความสับสนระหว่างพฤติกรรมที่มีรูปแบบการเคลื่อนไหวคล้ายกัน

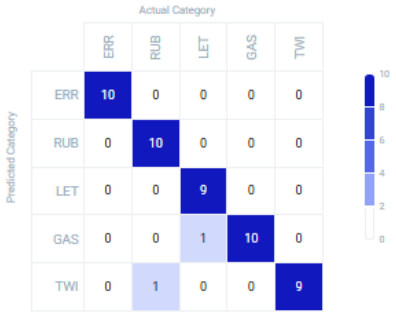
### 4.2 ผลการทดสอบในสภาวะแวดล้อมจริง

ทดสอบด้วยวิดีโอจาก YouTube/Facebook จำนวน 50 ตัวอย่างต่อโมเดล (5 พฤติกรรม × 10 ครั้ง)

**ตารางที่2:** ผลการทดสอบ

| Model | Acc. | P | R | F1 |
|---|---|---|---|---|
| YOLOv 5 | 86.00% | 86.28% | 86.00% | 85.97% |
| YOLOv 8 | 90.00% | 91.03% | 90.00% | 90.05% |
| YOLOv 5+CNN | 92.00% | 92.55% | 92.00% | 92.09% |
| YOLOv 8+CNN | 96.00% | 96.36% | 96.00% | 96.00% |

โมเดล YOLOv8 ร่วมกับ CNN ถือเป็นตัวเลือกที่ดีที่สุดสำหรับงานนี้ เนื่องจากสถาปัตยกรรม YOLOv8 ที่ทันสมัยกว่า เมื่อผนวกกับความสามารถของ CNN ในการเรียนรู้ลักษณะเชิงลึก ทำให้ระบบสามารถตรวจจับและจำแนกพฤติกรรมผิดปกติของปลาได้แม่นยำสูงสุด ซึ่งมีผลการทดสอบของโมเดล YOLOv8+CNN มีค่า Confusion Matrix ดังภาพที่ 3




**ภาพที่ 3:** Confusion Matrix โมเดล YOLOv8 + CNN

จากภาพที่ 1 จะเห็นว่าโมเดลสามารถจำแนกพฤติกรรม ERR, RUB, GAS ได้ถูกต้องจำนวน 10 ครั้ง ขณะที่พบการจำแนกผิดเพียงเล็กน้อย โดยมีตัวอย่างพฤติกรรม GAS ถูกจำแนกผิดเป็น LET จำนวน 1 ครั้ง และ TWI ถูกจำแนกผิดเป็น RUB จำนวน 1 ครั้ง ซึ่งอาจเกิดจากลักษณะการเคลื่อนไหวของพฤติกรรมบางประเภทที่มีความคล้ายคลึงกัน ทำให้โมเดลเกิดความสับสนในการจำแนก

## 5. สรุป

งานวิจัยเรื่อง "ระบบการตรวจจับพฤติกรรมปลาอัจฉริยะโดยใช้เทคนิคการประมวลผลภาพ" มีวัตถุประสงค์เพื่อพัฒนาต้นแบบจำลองสำหรับตรวจจับพฤติกรรมผิดปกติของปลาจากข้อมูลวิดีโอ โดยกำหนดพฤติกรรมเป้าหมาย 5 ประเภท ได้แก่ การว่ายน้ำผิดรูปแบบ (ERR) การถูตัวกับวัตถุ (RUB) การซึมหรือเคลื่อนไหวน้อย (LET) การฮุบอากาศที่ผิวน้ำ (GAS) และการกระตุกคล้ายชัก (TWI) และได้ทำการเปรียบเทียบประสิทธิภาพของโมเดลทั้ง 4 รูปแบบ คือ YOLOv5, YOLOv8, YOLOv5 ร่วมกับ CNN และ YOLOv8 ร่วมกับ CNN โดยทดสอบกับวิดีโอจริงรวมทั้งสิ้น 200 ครั้ง ผลการทดสอบพบว่าโมเดล YOLOv8 + CNN มีประสิทธิภาพสูงที่สุดในทุกตัวชี้วัด โดยมีค่า Accuracy 96.00% Precision 96.36% Recall 96.00% และ F1-score 96.00% ซึ่งสูงกว่า YOLOv5 แบบเดี่ยวที่มีค่า Accuracy เพียง 86.00% อย่างมีนัยสำคัญ สาเหตุที่ YOLOv8 ร่วมกับ CNN ให้ผลดีที่สุด

เนื่องมาจากโครงสร้างของ YOLOv8 ที่ได้รับการพัฒนาให้มีความสามารถในการตรวจจับวัตถุสูงกว่า YOLOv5 ประกอบกับการนำ CNN มาช่วยสกัดคุณลักษณะเชิงพื้นที่ของภาพ (Spatial Features) ได้อย่างละเอียด ทำให้โมเดลสามารถแยกแยะพฤติกรรมที่มีลักษณะการเคลื่อนไหวคล้ายคลึงกันได้แม่นยำขึ้น โดยเฉพาะพฤติกรรม LET ที่มักถูกสับสนกับพฤติกรรมอื่นในโมเดลที่ไม่มี CNN

ระบบที่พัฒนาขึ้นมีข้อดีสำคัญ 4 ประการ คือ มีความแม่นยำสูง รองรับการจำแนกพฤติกรรมได้ครอบคลุมทั้ง 5 ประเภท การผสานโมเดลช่วยเพิ่มประสิทธิภาพได้อย่างชัดเจน และสามารถนำไปประยุกต์ใช้กับระบบตรวจสอบสุขภาพปลาในฟาร์มเพื่อเฝ้าระวังความผิดปกติได้จริง

อย่างไรก็ตามระบบยังมีข้อจำกัดที่ต้องพัฒนาต่อ ได้แก่ ชุดข้อมูลฝึกที่ยังมีจำกัดและไม่ครอบคลุมสภาพแวดล้อมที่หลากหลาย ความคล้ายคลึงของบางพฤติกรรมที่ยังก่อให้เกิดความคลาดเคลื่อน ความไวต่อคุณภาพของวิดีโอในด้านความละเอียดและแสง และการที่ยังไม่ได้ทดสอบในสภาพแวดล้อมจริงระยะยาว ดังนั้นผู้วิจัยจึงเสนอแนะแนวทางพัฒนาต่อในอนาคต 5 ด้าน ได้แก่ การเพิ่มจำนวนและความหลากหลายของชุดข้อมูลโดยครอบคลุมความแตกต่างของแสง สภาพน้ำ และมุมกล้อง การนำระบบไปทดสอบในฟาร์มเลี้ยงปลาจริง การทดลองใช้สถาปัตยกรรม Deep Learning รูปแบบใหม่เพื่อเปรียบเทียบและยกระดับความแม่นยำ และการเชื่อมต่อระบบเข้ากับระบบจัดการฟาร์มอัจฉริยะ (Smart Farm) เพื่อช่วยในการติดตามสุขภาพปลา แจ้งเตือนความผิดปกติ และสนับสนุนการตัดสินใจของเกษตรกรผู้เลี้ยงปลาได้อย่างครบวงจรและมีประสิทธิภาพสูงสุด

## เอกสารอ้างอิง


[1] W. Zhang *et al.*, “A framework for the efficient enhancement of non-uniform illumination underwater image using convolution neural network,” *Computers & Graphics*, vol. 112, pp. 60–71, May 2023.

[2] K. Cai, X. Miao, W. Wang, H. Pang, Y. Liu, and J. Song, “A modified YOLOv3 model for fish detection based on MobileNetv1 as backbone,” *Aquacultural Engineering*, vol. 91, p. 102117, Nov. 2020.

[3] S. Malla, M. J. Meena, O. Reddy. R, V. Mahalakshmi, and A. Balobaid, “A Study on Fish Classification Techniques using Convolutional Neural Networks on Highly Challenged Underwater Images,” *IJRITCC*, vol. 10, no. 4, pp. 01–09, Apr. 2022.

[4] S. Tangruangkiat, “Accuracy in classification of liver lesion ultrasound using artificial intelligence,” Doctoral Degree, Chulalongkorn University, Bangkok, Thailand, 2022.

[5] A. Yassir, S. Jai Andaloussi, O. Ouchetto, K. Mamza, and M. Serghini, “Acoustic fish species identification using deep learning and machine learning algorithms: A systematic review,” *Fisheries Research*, vol. 266, p. 106790, Oct. 2023.

[6] Z. Yue, B. Yan, H. Liu, and Z. Chen, “An Effective Method for Underwater Biological Multi-Target Detection Using Mask Region-Based Convolutional Neural Network,” *Water*, vol. 15, no. 19, p. 3507, Oct. 2023.

[7] A. S. Abangan, D. Kopp, and R. Faillettaz, “Artificial intelligence for fish behavior recognition may unlock fishing gear selectivity,” *Front. Mar. Sci.*, vol. 10, p. 1010761, Feb. 2023.

[8] S. Gregory Dal Toé *et al.*, “Automated Video-Based Capture of Crustacean Fisheries Data Using Low-Power Hardware,” *Sensors*, vol. 23, no. 18, p. 7897, Sep. 2023.

[9] S. Villon *et al.*, “Automatic underwater fish species classification with limited data using few-shot learning,” *Ecological Informatics*, vol. 63, p. 101320, Jul. 2021.

[10] X. Chen, Y. Zhang, D. Li, and Q. Duan, “Chinese mitten crab detection and gender classification method based on GMNet-YOLOv4,” *Computers and Electronics in Agriculture*, vol. 214, p. 108318, Nov. 2023.

[11] T. Yoshida, K. Sugino, and H. Nishikawa, “Classification of set-net fish catch volumes in Iwate Prefecture, Japan using machine learning with water temperature and current distribution images at migration depth,” *Regional Studies in Marine Science*, vol. 73, p. 103480, Jul. 2024.

[12] N. Vanichkul, “Confusion Detection from Facial Expression using Deep Neural Network,” Master’s Degree, Chulalongkorn University, Bangkok, Thailand, 2020.

[13] H. Qu, G.-G. Wang, Y. Li, X. Qi, and M. Zhang, “ConvFishNet: An efficient backbone for fish classification from composited underwater images,” *Information Sciences*, vol. 679, p. 121078, Sep. 2024.

[14] D. Bartunek and P. Cisar, “Data for non-invasive (photo) individual fish identification of multiple species,” *Data in Brief*, vol. 48, p. 109221, Jun. 2023.

[15] S. Mohapatra, T. Swarnkar, and J. Das, “Deep convolutional neural network in medical image processing,” in *Handbook of Deep Learning in Biomedical Engineering*, Elsevier, 2021, pp. 25–60.

[16] S. Bhattacharya *et al.*, "Deep learning and medical image processing for coronavirus (COVID-19) pandemic: A survey," *Sustainable Cities and Society*, vol. 65, p. 102589, Feb. 2021.

[17] J. Liu, L. Zhang, Y. Li, and H. Liu, "Deep residual convolutional neural network based on hybrid attention mechanism for ecological monitoring of marine fishery," *Ecological Informatics*, vol. 77, p. 102204, Nov. 2023.

[18] M. Iftikhar, M. Neal, N. Hold, S. Gregory Dal Toé, and B. Tiddeman, "Detection of Crabs and Lobsters Using a Benchmark Single-Stage Detector and Novel Fisheries Dataset," *Computers*, vol. 13, no. 5, p. 119, May 2024.

[19] L. C. Ngugi, M. Abelwahab, and M. Abo-Zahhad, "Recent advances in image processing techniques for automated leaf pest and disease recognition – A review," *Information Processing in Agriculture*, vol. 8, no. 1, pp. 27–51, Mar. 2021

[20] Z. Du *et al.*, "Feature fusion strategy and improved GhostNet for accurate recognition of fish feeding behavior," *Computers and Electronics in Agriculture*, vol. 214, p. 108310, Nov. 2023.

# แบบจำลองการเรียนรู้ของเครื่องสำหรับคาดการณ์ปริมาณการสั่งซื้อยาตำราหลวงในร้านขายยาองค์การเภสัชกรรม

# A Machine Learning Model for Predicting Royal Pharmacopoeia Drug Orders in GPO Pharmacy

*ทรัพย์สิน ดิลกเวชพาณิชย์ (Sabsin Dilokwechpanich)[1] และนลินภัสร์ บำเพ็ญเพียร (Nalinpat Bhumpenpein)[2]*
*ภาควิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล*
*มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ*
*[1]s6707011857021@email.kmutnb.ac.th, [2]nalinpat.b@itd.kmutnb.ac.th*

## บทคัดย่อ

*ยาตำราหลวงเป็นชุดยาสามัญประจำบ้านมีความสำคัญในการดูแลสุขภาพเบื้องต้นของประชาชน การบริหารจัดการคลังของร้านยาองค์การเภสัชกรรมจึงมีความสำคัญในการป้องกันปัญหาการขาดแคลนยาและการมีสินค้าคงคลังเกินความจำเป็น งานวิจัยนี้มีวัตถุประสงค์เพื่อเปรียบเทียบแบบจำลองการเรียนรู้ของเครื่องสำหรับการคาดการณ์ปริมาณการสั่งซื้อยาตำราหลวงในร้านยาองค์การเภสัชกรรม โดยใช้ข้อมูลการสั่งซื้อรายเดือนจำนวน 90 ชุดข้อมูลตั้งแต่เดือนเมษายน 2561 ถึง กันยายน 2568 ด้วยการเปรียบเทียบแบบจำลองการเรียนรู้ของเครื่องแบบเดี่ยว คือแบบจำลอง Random Forest และ XGBoost และแบบจำลองแบบผสมที่ใช้แบบจำลองเชิงสถิติ ARIMA ในการสร้างแบบจำลอง ARIMA–Random Forest และ ARIMA–XGBoost จากผลการศึกษาพบว่าแบบจำลองการเรียนรู้ของเครื่องแบบเดี่ยวมีประสิทธิภาพสูงกว่า โดย Random Forest ให้ค่าความคลาดเคลื่อนต่ำที่สุดมีค่า MAPE เท่ากับ 10.86 จึงเหมาะต่อการนำไปประยุกต์ใช้เพื่อสนับสนุนการตัดสินใจด้านการวางแผนการจัดซื้อและการบริหารจัดการคลังยา*

**คำสำคัญ:** ยาตำราหลวง การเรียนรู้ของเครื่อง แบบจำลองเชิงสถิติอะริมา อนุกรมเวลา

## Abstract

*Royal Pharmacopoeia Drug is a household essential medicine that plays a significant role in supporting primary healthcare services. Consequently, effective inventory management in Government Pharmaceutical Organization (GPO) pharmacy is essential to mitigate drug shortages and excessive inventory levels. This study aims to comparatively evaluate the performance of machine learning models for forecasting the order quantity of Royal Pharmacopoeia Drug in GPO pharmacy. The models investigated include standalone machine learning models, namely Random Forest and XGBoost, as well as hybrid models that integrate a statistical ARIMA model, specifically ARIMA–Random Forest and ARIMA–XGBoost. The analysis was conducted using 90 monthly order records collected from April 2018 to September 2025. The results indicate that standalone machine learning models exhibit superior prediction performance compared to the hybrid approaches. Among all evaluated models, the Random Forest model achieves the lowest errors, with MAPE of 10.86. These findings demonstrate that the Random Forest model can be utilized to support procurement planning and inventory management in GPO pharmacy.*

**Keywords:** Royal Pharmacopoeia Drug, Machine Learning, ARIMA, Time Series.

## 1. บทนำ

ยาตำราหลวงคือ ชุดยาสามัญประจำบ้านที่ผลิตโดยองค์การเภสัชกรรมใช้สำหรับปฐมพยาบาลประกอบด้วยยาและเวชภัณฑ์ที่จำเป็น บรรจุในรูปแบบชุดสำเร็จที่ประชาชนสามารถนำไปใช้ในครัวเรือน หรือเพื่อทำบุญในวันสำคัญทางพระพุทธศาสนาอีกทั้งยังมีการนำยาตำราหลวงไปบริจาคเพื่อช่วยเหลือผู้ประสบภัยในช่วงสถานการณ์ฉุกเฉิน เช่น น้ำท่วม โรคระบาด และภัยพิบัติ เป็นต้น [1] ด้วยเหตุนี้ความต้องการยาจึงมีลักษณะเปลี่ยนแปลงไปตามช่วงเวลาและสถานการณ์ที่เกิดขึ้น การบริหารจัดการคลังยาและการวางแผนการจัดซื้อจึงจำเป็นต้องมีการคาดการณ์ที่มีความแม่นยำและมีประสิทธิภาพเพื่อป้องกันปัญหาการขาดแคลนยาจนส่งผลกระทบต่อประชาชนหรือการมียาคงคลังเกินความจำเป็นจนนำไปสู่ต้นทุนในการสำรองคลังที่สูงขึ้น [2] ในปัจจุบันมีการพัฒนาแนวทางการคาดการณ์หรือพยากรณ์ข้อมูลอย่างต่อเนื่องด้วยเทคนิคการเรียนรู้ของเครื่อง (Machine Learning) เช่น Random Forest, XGBoost เหมาะสมกับข้อมูลที่มีลักษณะซับซ้อนและไม่เป็นเชิงเส้นและวิธีการทางสถิติ เช่น ARIMA มีความสามารถในการวิเคราะห์ข้อมูลที่มีความต่อเนื่องและมีแนวโน้มเชิงเส้น ทั้งสองจึงเป็นอีกทางเลือกที่มีศักยภาพในการเพิ่มความแม่นยำของผลการคาดการณ์ [3] อย่างไรก็ตาม งานวิจัยที่ผ่านมายังไม่พบการศึกษาที่เปรียบเทียบประสิทธิภาพระหว่างแบบจำลองเดี่ยวและแบบจำลองผสม (Hybrid Model) ในบริบทของยาร้านยาองค์การเภสัชกรรม งานวิจัยนี้จึงมุ่งเน้นการค้นหาประสิทธิภาพของแบบจำลองที่มีความเหมาะสมและมีประสิทธิภาพในบริบทดังกล่าว

## 2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

### 2.1 ยาสามัญประจำบ้าน

ยาสามัญประจำบ้าน คือยาที่ประกาศให้เป็นยาสามัญประจำบ้านตามรายการแนบท้ายประกาศกระทรวงสาธารณสุข เรื่อง ยาสามัญประจำบ้านแผนปัจจุบัน พ.ศ. 2568 ประกอบไปด้วยยา 16 กลุ่ม 52 รายการ [4]

### 2.2 ยาตำราหลวง

ยาตำราหลวง คือชุดยาสามัญประจำบ้านเดิมเรียก “ยาโอสถสภา” ถือกำเนิดจากแนวคิดของสมเด็จกรมพระยาดำรงราชานุภาพในปี พ.ศ. 2444 ปัจจุบันยาตำราหลวงเป็นชุดยาสามัญประจำบ้านที่ผลิตโดยองค์การเภสัชกรรม ใช้สำหรับปฐมพยาบาลประกอบด้วยยาและเวชภัณฑ์ที่จำเป็น 7 รายการได้แก่ ยาแก้ปวดพาราเซตามอล ยาธาตุน้ำแดง ยาแก้แพ้ ผงเกลือแร่แก้ท้องเสีย ยาใส่แผล แถบปิดแผล และสำลี [5]

### 2.3 อนุกรมเวลา

การคาดการณ์ด้วยอนุกรมเวลาเป็นการคาดการณ์ข้อมูลในอนาคตจากข้อมูลที่ถูกรวบรวมตามลำดับเวลาในอดีตแบ่งได้ 5 รูปแบบ ได้แก่ รูปแบบคงที่ (Constant) รูปแบบมีแนวโน้ม (Trend) รูปแบบฤดูกาล (Seasonal) รูปแบบวัฏจักร (Cyclical) และรูปแบบสุ่มหรือไม่แน่นอน (Random or Irregular) [6]

### 2.4 แบบจำลองทางสถิติอะริมา

แบบจำลองทางสถิติอะริมา (Auto Regressive Integrated Moving Average: ARIMA) เป็นวิธีการพยากรณ์ที่นำข้อมูลอนุกรมเวลาในอดีตมาหารูปแบบที่เหมาะสม และใช้รูปแบบที่ได้พยากรณ์ข้อมูลอนุกรมเวลาในอนาคต โดยข้อมูลอนุกรมเวลาที่จะนำมาวิเคราะห์ต้องมีมีความเป็น Stationary กล่าวคือเป็นข้อมูลอนุกรมเวลาที่มีค่าเฉลี่ย (Mean), ค่าความแปรปรวน (Variance) และค่าความแปรปรวนร่วม (Covariance) คงที่ซึ่งไม่ได้ขึ้นอยู่กับเวลา แต่จะขึ้นอยู่กับระยะหรือช่วงห่างของช่วงเวลา (Lag time) การทดสอบความเป็น Stationary ของข้อมูลที่ใช้ในมีด้วยกันหลายวิธี เช่น วิธี Augmented Dickey and Fuller (ADF) [7] เป็นต้น

### 2.5 แบบจำลอง Random Forest

Random Forest เป็นแบบจำลองแบบ Ensemble Learning ที่สร้างต้นไม้ตัดสินใจหลายต้นแล้วรวมผลลัพธ์เข้าด้วยกันเพื่อการพยากรณ์ข้อมูล โดยมีพื้นฐานมาจาก Decision Tree [8]

### 2.6 แบบจำลอง XGBoost

XGBoost เป็นแบบจำลองที่ใช้การเรียนรู้แบบเพิ่มน้ำหนักความผิดพลาดให้กับข้อมูลที่ทำนายผิด ทำให้มี

ประสิทธิภาพสูงในการพยากรณ์ข้อมูลที่มีความซับซ้อน ด้วยการนำเอา Decision Tree มาฝึกสอนต่อกันหลาย ๆ ต้น โดยที่ Decision Tree แต่ละต้นจะเรียนรู้จากค่าความผิดพลาดก่อนหน้าซึ่งทำให้ความแม่นยำในการทำนายจะมากขึ้นเรื่อย ๆ [9]

**2.7 งานวิจัยที่เกี่ยวข้อง**

ในปัจจุบันมีการศึกษาการใช้แบบจำลองการเรียนรู้ของเครื่องในการคาดการณ์ความต้องการสินค้าหรือยา เพื่อจัดการห่วงโซ่อุปทานและเพิ่มประสิทธิภาพในการบริหารคงคลังกันอย่างแพร่หลาย การศึกษาของ Sen et al. [2] ที่พยากรณ์ความต้องการวัคซีนไข้เหลือง โดยใช้แบบจำลองทางสถิติและแบบจำลองการเรียนรู้ของเครื่องเพื่อช่วยในการบริหารจัดการคลังอย่างมีประสิทธิภาพซึ่งแบบจำลอง ARIMA ให้ผลการพยากรณ์ที่ดีโดยมีค่า RMSE และ MAE ต่ำที่สุด อีกทั้งยังมีการใช้แบบจำลอง Random Forest ที่นำไปเปรียบเทียบกับแบบจำลอง XGBoost ในการพยากรณ์การใช้พลังงานในฤดูกาลต่างๆ ด้วยในการศึกษาของ Maleki et al [10] และยังมีการศึกษาของ Kane et al. [8] ที่ใช้แบบจำลอง Random Forest แสดงผลประสิทธิภาพที่ดีกว่า ARIMA ในการคาดการณ์การระบาดของไข้หวัดนก H5N1 ในประชากรสัตว์ปีก นอกจากนี้ยังมีการศึกษาของ Teja & Rayalu [3] ในการพยากรณ์อัตราการเสียชีวิตจากโรคหัวใจและหลอดเลือดในอินเดียโดยใช้แบบจำลอง ARIMA-Random Forest เป็นแบบจำลองแบบผสม การเรียนรู้ของเครื่องกับวิธีการทางสถิติและในบางกรณีแบบจำลอง ARIMA-Random Forest ช่วยปรับปรุงความแม่นยำในการพยากรณ์ในการศึกษาของ Kontopoulou et al. [6] และมีงานวิจัยของ Fourkiotis & Tsadiras [11] ได้เปรียบเทียบแบบจำลอง XGBoost กับ ARIMA ในการคาดการณ์ยอดขายยาพบว่า XGBoost มีประสิทธิภาพที่เหนือกว่า

## 3. วิธีการดำเนินการวิจัย

**3.1 การเก็บรวบรวมข้อมูล**

ข้อมูลประวัติการสั่งซื้อยาตำราหลวงรายเดือน ตั้งแต่เดือน เมษายน พ.ศ. 2561 ถึงเดือน กันยายน พ.ศ. 2568 จำนวน 90 ชุดข้อมูลและข้อมูลประวัติการเกิดอุทกภัยในปี พ.ศ. 2561–2568 ดังนี้

1) Order_Qtt คือ ปริมาณการสั่งซื้อรายเดือน

2) buddhist_month คือ เดือนที่มีวันสำคัญทางพระพุทธศาสนา

3) flood_history คือ เดือนที่มีประวัติการเกิดอุทกภัย

4) flood_intensity_by_province คือ จำนวนจังหวัดที่เกิดอุทกภัยในเดือนนั้นๆ

5) flood_severity คือ ความรุนแรงเป็นวงกว้างของการเกิดอุทกภัย

**3.2 การทำความสะอาดข้อมูลและการเตรียมข้อมูล**

ข้อมูลที่ผ่านการทำความสะอาดจะถูกประมวลผลเพิ่มเติมโดยใช้กระบวนการ Feature Engineering เพื่อสร้างหรือตัดแต่งคุณลักษณะ (Features)

1) Month_sin คือ ตำแหน่งของข้อมูลรายเดือนแบบวงกลมโดยใช้ sin

2) Month_cos คือ ตำแหน่งของข้อมูลรายเดือนแบบวงกลมโดยใช้ cos

3) MA3 และ MA6 คือ ค่าเฉลี่ย 3 และ 6 เดือน

4) Lag3 และ Lag6 คือ ค่าย้อนไป 3 และ 6 เดือน

ข้อมูลถูกแบ่งตามลำดับเวลาเป็นชุดข้อมูลฝึก(Training Set) ร้อยละ 80 และชุดข้อมูลทดสอบ (Testing Set) ร้อยละ 20 โดยมีการใช้ Time Series Cross-Validation จำนวน 5-folds (n_splits=5) ภายในชุดข้อมูลฝึก

**3.3 การออกแบบสถาปัตยกรรมของแบบจำลอง**

แบบจำลองที่ใช้ในการศึกษาประกอบด้วยแบบจำลองเชิงสถิติ (ARIMA) ,แบบจำลองการเรียนรู้ของเครื่อง (Machine Learning) ได้แก่ Random Forest (RF), Extreme Gradient Boosting (XGBoost) และแบบจำลองแบบผสม ARIMA-Random Forest และ ARIMA-XGBoost

3.3.1 แบบจำลองเชิงสถิติ (ARIMA) เริ่มต้นด้วยการทดสอบสมบัติคงที่ของข้อมูล (Stationary) ด้วย Augmented Dickey and Fuller (ADF) จากนั้น กำหนดรูปแบบ (Identification) ARIMA(p,d,q) โดยพิจารณาจาก ACF/PACF จากนั้นทดลองปรับค่าพารามิเตอร์ และตรวจสอบรูปแบบ (Diagnostic Checking)

3.3.2 Random Forest (RF) ใช้ Grid search ในการปรับ Hyperparameter โดยกำหนด "n_estimators": [300, 600], "max_depth": [None, 8], "min_samples_split": [2, 5], "min_samples_leaf": [1, 2],"max_features": ["sqrt", None]

3.3.3 Extreme Gradient Boosting (XGBoost) ใช้ Grid search ในการปรับ Hyperparameter โดยกำหนด "n_estimators": [400, 700], "max_depth": [3,5], "learning_rate": [0.05, 0.1], "subsample": [0.8, 1.0], "colsample_bytree": [0.8, 1.0], reg_lambda: [1.0, 5.0]

3.3.4 แบบจำลองแบบผสม (Hybrid Model)

นำค่าความคลาดเคลื่อน (Residual) จากแบบจำลอง ARIMA เป็นข้อมูลนำเข้าให้แบบจำลองการเรียนรู้ของเครื่อง (Residual-based)

3.3.4.1 ARIMA- Random Forest ดำเนินการหาผลต่างข้อมูลเชิงเส้นที่ได้จาก ชุดข้อมูลARIMA training fitted กับ Order_Qtt ซึ่งได้ค่า Residual ดังสมการที่ 1

ResidualRF= Order_Qtt – ARIMA (1)

เมื่อได้ค่า Residual ใช้เป็น Target variable ในการวิเคราะห์ด้วย Random Forest จากนั้นนำผลการคาดการณ์ที่ได้จาก Residual Random Forest ($R_{RF}$) บวกกลับไปในข้อมูลเชิงเส้นที่ได้จาก ARIMA ดังสมการที่ 2

ARIMA-RF = ARIMA + $R_{RF}$ (2)

3.3.4.2 ARIMA-XGBoost ดำเนินการหาผลต่างข้อมูลเชิงเส้นที่ได้จากชุดข้อมูล ARIMA training fitted กับ Order_Qtt ซึ่งได้ค่า Residual ดังสมการที่ 3

ResidualXG= Order_Qtt – ARIMA (3)

เมื่อได้ค่า Residual ใช้เป็น Target variableในการวิเคราะห์ด้วย XGBoost จากนั้นนำผลการคาดการณ์ที่ได้จาก Residual XGBoost ($R_{XG}$) บวกกลับไปในข้อมูลเชิงเส้นที่ได้จาก ARIMA ดังสมการที่ 4

ARIMA-XGBoost = ARIMA + $R_{XG}$ (4)

**3.4 การประเมินผลแบบจำลอง**

งานวิจัยประเมินผลแบบจำลองด้วยค่ารากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ย (Root Mean Square Error) และค่าเปอร์เซ็นต์ความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (Mean Absolute Percentage Error ) และการทดสอบ Overfitting โดยพิจารณาจากอัตราส่วนของค่าความคลาดเคลื่อนระหว่างการประเมินระหว่างการฝึกและชุดข้อมูลทดสอบ

**3.5 การวิเคราะห์ความสำคัญของตัวแปร**

แบบจำลองที่มีประสิทธิภาพในการคาดการณ์สูงที่สุดจะนำไปวิเคราะห์ต่อด้วย Feature Importance Analysis

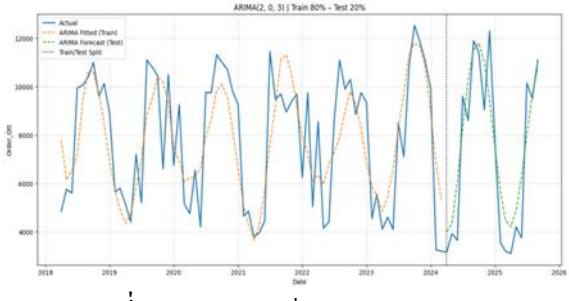
## 4. ผลการดำเนินงานวิจัย

**4.1 ผลการเตรียมข้อมูลและการวิเคราะห์เบื้องต้น**

ข้อมูลประวัติการสั่งซื้อยาตำราหลวงรายเดือน ตั้งแต่เดือนเมษายน พ.ศ. 2561 ถึงเดือนกันยายน พ.ศ. 2568 รวมทั้งสิ้น 90 ชุดข้อมูล มีค่าเฉลี่ยการสั่งซื้อต่อเดือนเท่ากับ 7,718.54 กล่อง ส่วนเบี่ยงเบนมาตรฐานมีค่า 2879.71 กล่องต่อเดือน จำนวนการสั่งซื้อสูงสุดเท่ากับ 12,530 กล่อง และจำนวนการสั่งซื้อต่ำสุดเท่ากับ 3,100 กล่อง

**4.2 ผลการทดสอบคุณสมบัติของข้อมูลสำหรับแบบจำลองเชิงสถิติ ARIMA**

การทดสอบคุณสมบัติความเป็น Stationary ของข้อมูลด้วยวิธี Augmented Dickey-Fuller (ADF) ให้ค่า p-value เท่ากับ 0.047 ซึ่งมีค่าน้อยกว่าระดับนัยสำคัญ 0.05 และค่า ADF Statistic เท่ากับ -2.885 แสดงให้เห็นว่าข้อมูลมีความเป็น Stationary และสามารถนำไปใช้ในการสร้างแบบจำลอง ARIMA ได้โดยไม่จำเป็นต้องทำผลต่างเพิ่มเติม (d = 0) และผลจากการเลือกพารามิเตอร์ ARIMA (p, d, q) จากการพิจารณาค่า ACF และ PACF จะได้รูปแบบ ARIMA (p, d, q) เป็น ARIMA (2, 0, 3) ดังแสดงในภาพที่ 1 และได้ผลประเมินในตารางที่ 1



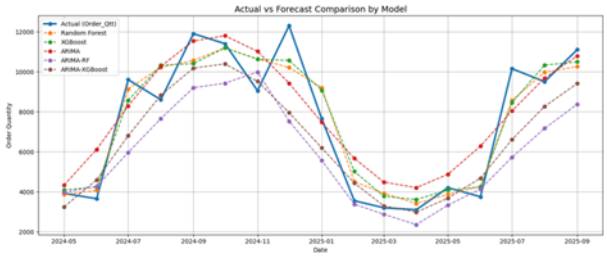

**ภาพที่ 1**: รูปแบบข้อมูลที่ได้จาก ARIMA (2, 0, 3)

**ตารางที่ 1**: ผลการประเมินแบบจำลอง ARIMA (2,0,3)

| แบบจำลอง | RMSE | MAPE |
|---|---|---|
| ARIMA (2,0,3) | 1540.10 | 24.22 |

**4.3 ผลการคาดการณ์จากแบบจำลองการเรียนรู้ของเครื่อง**

เมื่อพิจารณา RMSE และ MAPE ของผลการคาดการณ์ในทุกแบบจำลองพบว่าแบบจำลองแบบเดี่ยว Random Forest มีค่า MAPE และ RMSE ต่ำที่สุด แสดงดังภาพที่ 2 และตารางที่ 2




**ภาพที่ 2:** การเปรียบเทียบผลการคาดการณ์แบบจำลอง

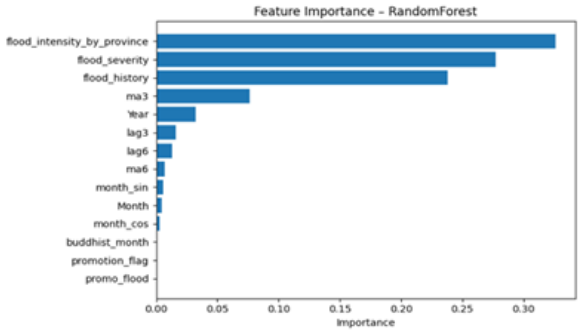
**ตารางที่ 2:** ผลการประเมินแบบจำลองการเรียนรู้ของเครื่อง

| แบบจำลอง | RMSE | MAPE |
|---|---|---|
| Random Forest | 1034.28 | 10.86 |
| XGBoost | 1344.77 | 11.36 |
| ARIMA-Random Forest | 1721.53 | 16.56 |
| ARIMA-XGBoost | 2210.503 | 20.04 |

ผลการทดสอบ Overfitting ของแบบจำลองการเรียนรู้ของเครื่องแบบเดี่ยว Random Forest พบว่าค่าความคลาดเคลื่อนอัตราส่วน Test/CV_Ratio มีค่า 0.9104

**4.4 ผลการวิเคราะห์ความสำคัญของตัวแปร**

ผลการวิเคราะห์ความสำคัญของตัวแปรของแบบจำลองการเรียนรู้ของเครื่อง Random Forest พบว่า ตัวแปรที่มีอิทธิพลต่อการคาดการณ์ปริมาณการสั่งซื้อยาตำราหลวงประกอบด้วยกลุ่มตัวแปรภายนอกที่มีลักษณะเปลี่ยนแปลงตามช่วงเวลา ได้แก่ flood_intensity_by_province, flood_severity และ flood_history นอกจากนี้ กลุ่มตัวแปรเชิงเวลายังมีนัยสำคัญต่อแบบจำลองอย่างชัดเจน ไม่เพียงแต่ค่าเฉลี่ยเคลื่อนที่ 3 เดือน (ma3) เท่านั้น แต่จากกราฟยังพบว่าตัวแปรเชิงเวลาอื่นๆ ที่สะท้อนรูปแบบในอดีต ได้แก่ ปี (Year), ค่าย้อนหลัง 3 เดือน (lag3), ค่าย้อนหลัง 6 เดือน (lag6) และค่าเฉลี่ยเคลื่อนที่ 6 เดือน (ma6) ล้วนมีบทบาทสำคัญต่อแบบจำลองเช่นกัน ดังแสดงในภาพที่ 3




**ภาพที่ 3:** ผลการวิเคราะห์ความสำคัญของตัวแปรของแบบจำลองการเรียนรู้ของเครื่อง Random Forest

## 5. สรุปและอภิปรายผล

จากผลการวิจัยแสดงให้เห็นว่าแบบจำลองการเรียนรู้ของเครื่องแบบเดี่ยว Random Forest มีประสิทธิภาพในการคาดการณ์ปริมาณการสั่งซื้อยาตำราหลวงได้ดีที่สุด มีค่า RMSE ต่ำที่สุดซึ่งสะท้อนถึงความสามารถในการลดความคลาดเคลื่อนขนาดใหญ่ในช่วงเวลาที่มีความผันผวนสูงได้ดีและเมื่อพิจารณาค่า MAPE ซึ่งเป็นตัวชี้วัดความคลาดเคลื่อนเชิงสัดส่วน พบว่า Random Forest ให้ค่าต่ำที่สุดแสดงให้เห็นถึงความเสถียรและความแม่นยำโดยรวมของการคาดการณ์ในช่วงเวลาส่วนใหญ่ ผลลัพธ์ดังกล่าวสามารถอธิบายได้จากบริบทของข้อมูลเป็นข้อมูลอนุกรมเวลาที่มีความผันผวนและได้รับอิทธิพลจากปัจจัยภายนอกที่ไม่เป็นเชิงเส้น Random Forest สามารถจับรูปแบบข้อมูลที่ซับซ้อนได้ดีกว่าแบบจำลองเชิงสถิติ ผลการวิจัยนี้สอดคล้องกับงานวิจัยของ Kane M. et al. [8] ที่พบว่า Random Forest มีประสิทธิภาพในการคาดการณ์อุบัติการณ์ของไข้หวัดนกสูงกว่าแบบจำลองสถิติ

ในขณะที่แบบจำลองแบบผสมอาศัยแนวคิดในการแยกโครงสร้างเชิงเส้นออกด้วยแบบจำลอง ARIMA และนำค่าความคลาดเคลื่อนที่เหลือ (Residual) มาวิเคราะห์ข้อมูลที่ไม่เป็นเชิงเส้นต่อด้วยแบบจำลองการเรียนรู้ของเครื่อง ซึ่งช่วยลดความคลาดเคลื่อนขนาดใหญ่ในบางช่วงเวลาสอดคล้องกับการศึกษาของ Teja & Rayalu [3] ที่คาดการณ์การเสียชีวิตของประชากรชาวอินเดียต่อโรคหัวใจพบว่า ช่วงอายุ 50-69 ปี ที่ให้ผลว่าแบบจำลองแบบผสมมีประสิทธิภาพในการคาดการณ์สูงกว่าแบบจำลองเชิงสถิติ

อย่างไรก็ตามการแยกขั้นตอนดังกล่าวอาจเพิ่มความซับซ้อนให้แบบจำลองแบบผสม (Hybrid Model) เมื่อใช้กับชุดข้อมูลที่มีขนาดจำกัด โดยขนาดของข้อมูลดังกล่าวอาจไม่เพียงพอที่จะเอื้อให้แบบจำลองแบบผสมเรียนรู้ความสัมพันธ์แบบแยกส่วนทั้งเชิงเส้นและไม่เป็นเชิงเส้นได้อย่างเต็มประสิทธิภาพส่งผลให้ความคลาดเคลื่อนเพิ่มขึ้นซึ่งสะท้อนออกมาในรูปของค่า MAPE และ RMSE ที่สูงกว่าแบบจำลองการเรียนรู้ของเครื่องแบบเดี่ยว

โดยแบบจำลองการเรียนรู้ของเครื่องแบบเดี่ยว Random Forest มีคุณสมบัติ Ensemble Learning ที่สามารถจัดการข้อมูลที่มีความผันผวนทั้งที่เป็นเชิงเส้นและไม่เป็นเชิงเส้นและได้รับอิทธิพลจากปัจจัยภายนอกได้อย่างมีประสิทธิภาพภายใต้บริบทข้อมูลที่มีขนาดจำกัด

เมื่อพิจารณาความสำคัญของตัวแปรของแบบจำลองการเรียนรู้ของเครื่องแบบเดี่ยวพบว่า XGBoost มีแนวโน้มให้น้ำหนักกับตัวแปรบางตัวมากเกินไป โดยเฉพาะตัวแปรด้านประวัติการเกิดอุทกภัย ส่งผลให้แบบจำลองมีความไวต่อรูปแบบเฉพาะของข้อมูลในช่วงฝึกและทำให้ประสิทธิภาพในการทำนายข้อมูลใหม่ลดลงในทางตรงกันข้าม Random Forest สามารถผสานข้อมูลจากตัวแปรหลายมิติได้อย่างสมดุลมากกว่าจึงชี้ให้เห็นว่าการผสานข้อมูลปัจจัยด้านอุทกภัยและตัวแปรเชิงเวลา ช่วยเพิ่มเสถียรภาพของคาดการณ์ปริมาณการสั่งซื้อยาตำราหลวงอย่างมีนัยสำคัญภายใต้บริบทของข้อมูลที่มีความผันผวน

จากการวิเคราะห์ความเสี่ยงของการเกิดภาวะการเรียนรู้ข้อมูลมากเกินไป (Overfitting) พบว่าแบบจำลอง Random Forest มีความสามารถในการนำรูปแบบที่เรียนรู้จากข้อมูลในอดีตไปประยุกต์ใช้กับข้อมูลใหม่ได้อย่างเหมาะสม โดยสะท้อนจากค่าความคลาดเคลื่อนบนชุดข้อมูลทดสอบที่ไม่สูงกว่าค่าที่ได้จากการประเมินระหว่างการฝึกแบบจำลองอย่างมีนัยสำคัญ ผลดังกล่าวแสดงให้เห็นว่าแบบจำลองไม่มีแนวโน้มเกิดปัญหา Overfitting และมีความสามารถในการ Generalize ภายใต้บริบทของข้อมูล ซึ่งสนับสนุนความเหมาะสมของการนำแบบจำลองไปประยุกต์ใช้ในสถานการณ์จริงด้านการวางแผนการจัดซื้อและการบริหารจัดการคลังยา

## เอกสารอ้างอิง


[1] Hfocus.(2017).*กำเนิดยาตำราหลวงของไทย.* https://www.hfocus.org/content/2017/03/13686

[2] Sen, N., Temur, L. O., & Atilla, D. C. (2024). *"Yellow fever vaccine demand forecasting with Arima, Sarima, linear regression, and xgboost."* IEEE Access, 12, 197557–197576.

[3] Teja, M. D., & Rayalu, G. M. (2025). *"Hybrid time series and machine learning models for forecasting cardiovascular mortality in India: An age specific analysis"*. BMC Public Health, 25(1).

[4] กองยา สำนักงานคณะกรรมการอาหารและยา.(2025). *ยาสามัญประจำบ้านแผนปัจจุบัน* พ.ศ.2568. https://drug.fda.moph.go.th/announcement-ministry/law15082568

[5] ฮิมวัง.(16 พ.ค. 2567).*ยาสามัญประจำบ้าน จากจุดเริ่มต้นยาตำราหลวง ในสมัยรัชกาลที่ 5* https://www.silpa-mag.com/history/article_88541

[6] Kontopoulou, V. I., Panagopoulos, A. D., Kakkos, I., & Matsopoulos, G. K. (2023). "*A review of Arima vs. Machine Learning Approaches for time series forecasting in Data Driven Networks"*. Future Internet, 15(8), 255.

[7] Sutthichaimethee, J. (n.d.). *Application of ARIMA Model for Research*. Suddhiparitad, 24(72), 101-120.

[8] Kane, M. J., Price, N., Scotch, M., & Rabinowitz, P. (2014). *"Comparison of Arima and Random Forest time series models for prediction of avian influenza H5N1 outbreaks*". BMC Bioinformatics, 15(1).

[9] Thejovathi, M., Rao, Dr. M., Priyadharsini, E. J., Siddi, S., Karthik, B., & Abbas, S. H. (2025). *Optimizing product demand forecasting with hybrid machine learning and time series models: A Comparative Analysis of XGBoost and Sarima*. SSRN Electronic Journal.

[10] Maleki, N., Lundström, O., Musaddiq, A., Jeansson, J., Olsson, T., & Ahlgren, F. (2024). *Future Energy Insights: Time-Series and Deep Learning Models for City Load Forecasting*.

[11] Fourkiotis, K. P., & Tsadiras, A. (2024). *"Applying machine learning and statistical forecasting methods for enhancing pharmaceutical sales predictions"*. Forecasting, 6(1), 170–186.

# การแนะนำรูปแบบการบันทึกความเห็นและคำสั่งคดีโดยเทคนิคการเรียนรู้ด้วยเครื่อง
# Machine Learning-based Recommendation for Legal Opinion and Order Drafting Formats

*กฤษณพงศ์ กิ้มซื่อ (Kritsanaphong Kimsue)[1] และณัฐวี อุตกฤษฎ์ (Nattavee Utakrit)[2]*
*ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล*
*มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ*
*s6707021856159@email.kmutnb.ac.th,nattavee.u@itd.kmutnb.ac.th*

**บทคัดย่อ**

*การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อพัฒนาระบบต้นแบบสำหรับแนะนำรูปแบบการบันทึกความเห็นและคำสั่งคดี โดยใช้กรณีศึกษาคดียาเสพติด ระบบถูกออกแบบให้ทำงานเป็นลำดับขั้น โดยบูรณาการ Generative NER, Hybrid RAG และการปรับแต่งโมเดลด้วย LoRA เข้าด้วยกัน การดำเนินงานประกอบด้วย (1) สร้างคลังข้อมูลกฎหมายโดยแปลงตัวบทประมวลกฎหมายยาเสพติด พ.ศ. 2564 เป็นเวกเตอร์ด้วยโมเดล BAAI/bge-m3 ขนาด 1,024 มิติ จัดเก็บใน ChromaDB แบบ Hierarchical Chunking พร้อมกำกับ Metadata ลำดับชั้นอำนาจทางกฎหมาย 4 ระดับ (2) สร้างชุดข้อมูลจำลองสำนวนคดียาเสพติด 3,056 รายการ ผ่านกระบวนการสังเคราะห์ด้วย Gemini API และตรวจสอบความถูกต้องทางกฎหมายด้วยระบบเชิงกฎ (Rule-based Engine) ได้ค่า Weighted F1-Score ด้านความสม่ำเสมอของป้ายกำกับที่ร้อยละ 84.4 และ (3) ปรับแต่งโมเดลภาษาไทย Typhoon-S 4B ด้วยเทคนิค LoRA SFT ร่วมกับ Chain-of-Thought Prompting สำหรับสกัดข้อเท็จจริงสำคัญ ค้นคืนบทบัญญัติกฎหมายที่เกี่ยวข้อง และสังเคราะห์ร่างความเห็นตามโครงสร้างแบบ อก.04 ผลการประเมินพบว่าระบบสกัดข้อมูลเอนทิตีทางกฎหมายได้ด้วยค่า NER F1-Score = 0.700 และค้นคืนมาตรากฎหมายได้ด้วยค่า Hit Rate@1 และ Hit Rate@3 = 0.667 ส่วนการปรับแต่งด้วย LoRA ช่วยยกระดับคุณภาพข้อความจาก SacreBLEU 8.90 เป็น 56.10 ซึ่งผ่านเกณฑ์ที่กำหนด ผลการวิจัยยืนยันว่าสถาปัตยกรรมแบบผสมผสานนสามารถลดปัญหาข้อมูลหลอน (Hallucination) ได้*

**คำสำคัญ:** ประมวลกฎหมายยาเสพติด การให้เหตุผลเชิงกฎหมาย โมเดลภาษาขนาดใหญ่ การสืบค้นข้อมูลแบบเสริม การรู้จำเอนทิตีชื่อเฉพาะ

**Abstract**

*The study aims to develop a prototype system for recommending legal opinion and case order drafting formats, using drug-related cases as a case study. The system integrates Generative NER, Hybrid RAG, and LoRA fine-tuning into a sequential pipeline. Implementation involves: (1) constructing a legal knowledge base by encoding the Narcotics Code B.E. 2564 into 1,024-dimensional vectors using BAAI/bge-m3, stored in ChromaDB with section-level hierarchical chunking and four-tier legal authority metadata; (2) synthesizing 3,056 drug case files via Gemini API with rule-based legal validation, achieving 84.4% Weighted F1-Score for label consistency; and (3) fine-tuning Typhoon-S 4B using LoRA SFT with Chain-of-Thought Prompting to extract key legal facts, retrieve relevant statutory provisions, and generate structured drafts conforming to the อก.04 template. Evaluation results show NER F1-Score of 0.700, Hit Rate@1 and Hit Rate@3 of 0.667, and SacreBLEU improvement from 8.90 to*

*56.10 after fine-tuning. Results confirm that this hybrid architecture effectively mitigates hallucination and enhances legal text generation quality.*

**Keywords:** Narcotics Code, Legal Reasoning, LLM, RAG, NER.

## 1. บทนำ

ในกระบวนการยุติธรรมทางอาญาที่เกี่ยวข้องกับคดียาเสพติดเอกสารความเห็นและคำสั่งคดีมีบทบาทสำคัญในการสรุปข้อเท็จจริง ประเมินข้อกฎหมาย และบันทึกเหตุผลประกอบการตัดสินใจ อย่างไรก็ตาม ในทางปฏิบัติเอกสารเหล่านี้มักถูกจัดทำในลักษณะข้อความอิสระ (Free-form Text) ซึ่งก่อให้เกิดความไม่สม่ำเสมอของถ้อยคำทางกฎหมายและยากต่อการประมวลผลซ้ำ ปัญหาดังกล่าวทวีความซับซ้อนยิ่งขึ้นในคดียาเสพติด เนื่องจากกฎกระทรวงและระเบียบที่เกี่ยวข้องมีการปรับปรุงอย่างต่อเนื่อง หากพึ่งพากระบวนการสืบค้นแบบดั้งเดิมอาจนำไปสู่การอ้างอิงบทกฎหมายที่ล้าสมัยได้

แม้ว่าโมเดลภาษาขนาดใหญ่ จะมีความสามารถในการประมวลผลข้อมูลไร้โครงสร้าง แต่การนำมาใช้โดยตรงกับงานกฎหมายยังมีข้อจำกัดสำคัญ ได้แก่ ปัญหาข้อมูลหลอน (Hallucination) และขอบเขตความรู้ที่หยุดอยู่เพียงช่วงเวลาที่ฝึกสอน เทคโนโลยี Retrieval-Augmented Generation (RAG) [1] จึงถูกนำมาใช้ร่วมกับ LLMs เพื่อสืบค้นข้อกฎหมายจริงและเป็นปัจจุบันมาเสริมการสร้างข้อความ ลดโอกาสเกิดความคลาดเคลื่อนจากข้อจำกัดของโมเดลเพียงอย่างเดียว

งานวิจัยนี้มุ่งพัฒนาระบบต้นแบบสำหรับแนะนำรูปแบบการบันทึกความเห็นและคำสั่งคดียาเสพติด โดยมีผลงานสนับสนุนหลัก (Contributions) 3 ด้าน ได้แก่ (1) การออกแบบ Hybrid Pipeline ที่บูรณาการ Generative NER, Hybrid RAG และ LoRA Fine-tuning ร่วมกับ Legal Syllogism Prompting สำหรับงานร่างเอกสารทางกฎหมายในบริบทคดียาเสพติดไทยโดยเฉพาะ (2) การสร้างชุดข้อมูลจำลองสำนวนคดียาเสพติด 3,056 รายการ ผ่าน Rule-based Pipeline ที่ตรวจสอบความถูกต้องเชิงกฎหมายโดยอัตโนมัติ ได้ค่า Weighted F1-Score ด้านความสม่ำเสมอของป้ายกำกับร้อยละ 84.4 และ (3) กลไก Legal Authority Re-ranking ใน RAG ที่กำหนดลำดับชั้นอำนาจทางกฎหมาย 4 ระดับ เพื่อบังคับให้ตัวบทกฎหมายหลักถูกนำเสนอก่อนเสมอ ส่งผลให้ค่า Hit Rate@1 และ Hit Rate@3 มีความเสถียรเท่ากันที่ 0.667 ระบบถูกประเมินโดยพิจารณาความถูกต้องของการค้นคืนข้อความกฎหมาย (Hit Rate@K) ความถูกต้องของการสกัดข้อมูลข้อเท็จจริง (NER F1-Score) และคุณภาพของข้อความที่สร้างขึ้น (SacreBLEU)

## 2.ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

### 2.1 การประมวลผลธรรมชาติทางกฎหมายและโมเดลภาษาขนาดใหญ่

การประมวลผลภาษาธรรมชาติทางกฎหมาย (Legal NLP) เป็นการประยุกต์ใช้เทคนิค NLP และการเรียนรู้ของเครื่องกับข้อมูลประเภทกฎหมาย ซึ่งมีโครงสร้างภาษาเฉพาะทางและความซับซ้อนทั้งเชิงความหมายและบริบทสูงกว่าข้อความทั่วไป โมเดลภาษาขนาดใหญ่ (LLMs) [2] ที่พัฒนาบนสถาปัตยกรรม Transformer [3] ได้เข้ามามีบทบาทสำคัญในงานด้านกฎหมาย เช่น การตอบคำถาม การสรุปสาระสำคัญ และการสกัดข้อมูลสำคัญ อย่างไรก็ตาม การนำ LLMs มาใช้โดยตรงยังพบปัญหาข้อมูลหลอนและขอบเขตความรู้ที่หยุดอยู่เพียงช่วงเวลาที่ฝึกสอน

### 2.2 เทคนิค Retrieval-Augmented Generation และการจัดเก็บเวกเตอร์

Retrieval-Augmented Generation เป็นแนวทางที่ผสานการค้นคืนข้อมูลจากแหล่งความรู้ภายนอกเข้ากับความสามารถในการสร้างข้อความของ LLMs ทำให้ผลลัพธ์มีความสอดคล้องกับข้อเท็จจริงมากขึ้นและลดปัญหา Hallucination งานวิจัยนี้แปลงเอกสารกฎหมายเป็นเวกเตอร์ด้วยโมเดล BAAI/bge-m3 ขนาด 1,024 มิติ ผ่านการทำ L2 Normalization ก่อนจัดเก็บใน ChromaDB แบบ Section-level Chunking และกำกับ Metadata ลำดับชั้นอำนาจทางกฎหมาย 4 ระดับ ได้แก่ (1) ประมวลกฎหมาย (2) กฎกระทรวง (3) ระเบียบและคำสั่งที่เกี่ยวข้อง และ (4) แนวคำพิพากษาศาลฎีกา เพื่อให้ระบบสามารถจัดลำดับ

จอวการนำเสนอเอกสารอ้างอิงได้อย่างถูกต้องตามหลักของกฎหมาย

**2.3 งานวิจัยที่เกี่ยวข้อง**

Hanwiboonwat [4] พัฒนาแชตบอตกฎหมายไทยโดยใช้ LLM ร่วมกับชุดข้อมูล QA ด้านกฎหมาย พบว่าแนวทาง LLM+QA+EX ให้ค่า BERTScore F1 = 0.9367 และทำคะแนนข้อสอบทนายความได้ 46/80 ข้อ ใกล้เคียง GPT-4o ที่ได้ 47 ข้อ

Promwang และ Boonma [5] พัฒนาระบบ RAG สำหรับกฎหมายไทยโดยเปรียบเทียบ Sparse (BM25), Dense (SentenceTransformer) และ Hybrid Retrieval ร่วมกับ BGE CrossEncoder พบว่า Hybrid ให้ผลดีที่สุด โดยมี Recall@1 = 73.3% และ ROUGE-L เฉลี่ย 0.47–0.61

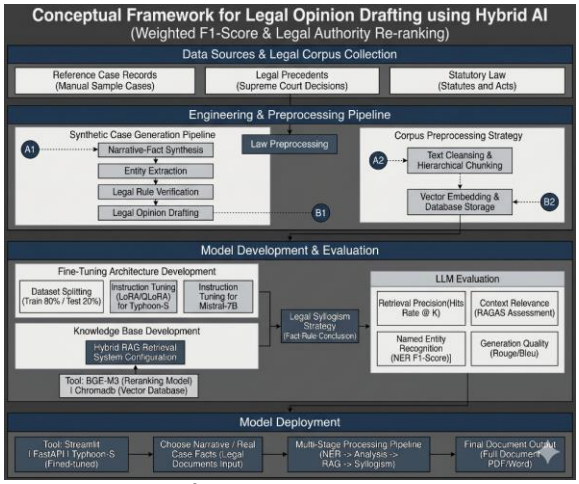
Akarajaradwong และคณะ [6] นำเสนอ NitiBench กรอบประเมินถาม–ตอบกฎหมายไทย ครอบคลุมกฎหมายการเงิน (CCL) และภาษี (Tax) ระบบ RAG บน NitiBench-CCL ให้ E2E F1 = 0.595 ขณะที่ชุด Tax ให้ E2E F1 = 0.457

Karamitsos และคณะ [7] นำเสนอ LegNER โมเดล NER เฉพาะโดเมนกฎหมาย ฝึกด้วยคดีศาล 1,542 คดี ครอบคลุม 6 ประเภทเอนทิตี ให้ F1-score = 99.4%

Akarajaradwong และคณะ [8] เสนอเทคนิค GRPO ร่วมกับ Semantic-similarity Reward โมเดล Typhoon2 บน NitiBench-CCL ให้ Citation F1 = 0.6828 และ Joint Score = 0.7773

## 3. วิธีการดำเนินการวิจัย

งานวิจัยนี้แบ่งกระบวนการดำเนินงานออกเป็น 4 ส่วนหลัก ได้แก่ การรวบรวมข้อมูลและสร้างคลังกฎหมาย การเตรียมและประมวลผลข้อมูล การพัฒนาและประเมินแบบจำลอง และการนำระบบไปใช้งานจริง ดังแสดงสถาปัตยกรรมภาพรวมในภาพที่ 1 : กรอบแนวคิดงานวิจัย แสดงสถาปัตยกรรมแบบ 4 ระยะ ได้แก่ (1) รวบรวมข้อมูลกฎหมายและสร้างสำนวนคดีจำลองผ่าน Gemini API (2) เตรียมข้อมูลด้วย Rule-based Validation และแปลงเป็น Vector Embedding ด้วย BGE-M3 จัดเก็บใน ChromaDB (3) พัฒนาและประเมินแบบจำลอง NER, Hybrid RAG และ LLM Fine-tuning ด้วย LoRA ร่วมกับ CoT Prompting และ (4) นำส่งผลลัพธ์ในรูปแบบ JSON ตามโครงสร้าง อก.4




**ภาพที่ 1:** กรอบแนวคิดงานวิจัย

**3.1 การรวบรวมและเตรียมข้อมูล**

การสร้างคลังข้อมูลกฎหมาย: ผู้วิจัยรวบรวมประมวลกฎหมายยาเสพติด พ.ศ. 2564 พระราชกฤษฎีกากฎกระทรวง และแนวคำพิพากษาศาลฎีกาที่เกี่ยวข้อง แล้วนำมาแบ่งย่อยตามโครงสร้างมาตรา (Section-level Chunking) จากนั้นแปลงเป็นเวกเตอร์ด้วยโมเดล BAAI/bge-m3 ขนาด 1,024 มิติ ผ่านการทำ L2 Normalization ก่อนจัดเก็บใน ChromaDB พร้อมกำกับ Metadata ลำดับชั้นอำนาจทางกฎหมาย 4 ระดับ รวมทั้งสิ้น 1,040 Chunks

การสร้างชุดข้อมูลจำลอง: เนื่องจากสำนวนคดีจริงมีข้อจำกัดด้านความลับทางราชการและการคุ้มครองข้อมูลส่วนบุคคล ผู้วิจัยจึงสังเคราะห์ชุดข้อมูลจำลอง 3,056 สำนวน ผ่าน Pipeline 5 ขั้นตอน ได้แก่ (1) สร้างพฤติการณ์คดีด้วย Gemini API (2) สกัดเอนทิตีและสร้างเอกสารประกอบสำนวน (3) ตรวจสอบความสมเหตุสมผลเชิงตรรกะด้วย Rule-based Engine ใน 4 มิติ (4) สืบค้นบริบทกฎหมายผ่าน RAG และ (5) สังเคราะห์ร่างความเห็นตาม อก.04 ด้วย Gemini 2.5 Pro ร่วมกับ Chain-of-Thought Prompting จากนั้นแบ่งชุดข้อมูลในอัตราส่วน 80:20 ได้

ชุดฝึกสอน 2,445 รายการ และชุดทดสอบ 611 รายการ ทั้งนี้เลือกใช้ Train/Test Split แทน K-Fold Cross Validation เนื่องจากการปรับแต่ง LLM แต่ละรอบต้องใช้ทรัพยากร GPU สูงมาก และงานสร้างข้อความไม่สามารถเปรียบเทียบผลระหว่าง Fold ได้โดยตรงเหมือนงานจำแนกประเภท [9]

### 3.2 การพัฒนาแบบจำลองและระบบต้นแบบ

ในขั้นตอนการพัฒนาแบบจำลอง ผู้วิจัยออกแบบระบบให้ผสานการทำงานระหว่างแบบจำลองภาษา Typhoon-S ฐานความรู้ทางกฎหมาย และกลไก RAG เพื่อใช้ในการสกัดข้อเท็จจริงสำคัญ ค้นคืนข้อกฎหมายที่เกี่ยวข้อง และจัดโครงสร้างผลลัพธ์ให้อยู่ในรูปแบบที่กำหนด ดังแสดงในตารางที่ 1

**ตารางที่ 1:** ค่าพารามิเตอร์ในการปรับแต่งโมเดล

| พารามิเตอร์ | ค่าที่กำหนด |
|---|---|
| LoRA Rank (r) | 64 |
| LoRA Alpha ($\alpha$) | 128 |
| Learning Rate | $4\times10^{-5}$ |
| Epochs | 3 |
| Max Sequence Length | 2,048 Tokens |
| Optimizer | AdamW (fused) |
| Scheduler | Cosine (warmup 5%) |
| Precision | BF16 |
| Effective Batch Size | 32 |

### 3.3 การประมวลผลการทำงานของระบบ

การประเมินผลการทำงานของระบบดำเนินการบนชุดทดสอบที่แยกออกจากชุดฝึกสอน เพื่อวิเคราะห์ความสามารถของแบบจำลองในด้านการค้นคืนข้อกฎหมาย การสกัดข้อมูลข้อเท็จจริงสำคัญ และการสร้างผลลัพธ์ให้อยู่ในรูปแบบที่กำหนด ตัวชี้วัดที่ใช้ในการประเมินประกอบด้วยค่า Hit Rate@K สำหรับประสิทธิภาพการค้นคืนข้อมูล ค่า F1-Score สำหรับการสกัดข้อมูลสำคัญ และค่า SacreBLEU [13] สำหรับคุณภาพของข้อความและความถูกต้องของรูปแบบผลลัพธ์ที่ระบบสร้างขึ้น

## 4 ผลวิจัยและอภิปรายผล

### 4.1 ผลการเตรียมข้อมูลและพัฒนาระบบต้นแบบ

ในการวิจัยครั้งนี้ ผู้วิจัยได้จัดเตรียมชุดข้อมูลจำลองสำนวนคดียาเสพติดจำนวนทั้งสิ้น 3,056 รายการ ดังแสดงในตารางที่ 2

**ารางที่ 2:** จำนวนข้อมูลที่ใช้ในการวิจัย

| รายการ | จำนวน | ร้อยละ |
|---|---|---|
| ชุดฝึกสอน (Train Set) | 2,445 | 80% |
| ชุดทดสอบ (Test Set) | 611 | 20% |
| รวมข้อมูลทั้งหมด | 3,056 | 100% |

เนื่องจากชุดข้อมูลที่ใช้เป็นข้อมูลจำลองที่สังเคราะห์ขึ้น ผู้วิจัยจึงดำเนินการตรวจสอบความถูกต้องเชิงกฎหมายด้วย Rule-based Engine โดยอัตโนมัติใน 4 มิติ ได้แก่ ความครบถ้วนของพยานหลักฐาน ความสอดคล้องระหว่างพฤติการณ์คดีกับฐานความผิด ความชอบด้วยกระบวนพิจารณาตามกฎหมาย และความถูกต้องของตรรกะในร่างความเห็น ผลการตรวจสอบพบว่าชุดข้อมูลมีค่า Weighted F1-Score ด้านความสม่ำเสมอของป้ายกำกับที่ร้อยละ 84.4 อย่างไรก็ตาม ข้อมูลจำลองมีโครงสร้างภาษาที่เป็นระเบียบกว่าสำนวนจริง ซึ่งอาจส่งผลต่อความแม่นยำเมื่อนำไปใช้กับสำนวนที่พนักงานสอบสวนเขียนด้วยภาษาพูด

### 4.2 ผลการประเมินประสิทธิภาพของระบบ

การประเมินประสิทธิภาพดำเนินการบนชุดทดสอบ 611 รายการ ผลการประเมินสรุปได้ดังตารางที่ 3

**ตารางที่ 3:** ผลการประเมินประสิทธิภาพของระบบ

| ตัวชี้วัด | ค่าที่ได้ |
|---|---|
| Hit Rate@1 | 0.667 |
| Hit Rate@3 | 0.667 |
| NER F1-Score | 0.700 |
| SacreBLEU (Base Model) | 8.90 |
| SacreBLEU (+ LoRA SFT) | 56.10 |

ด้านการค้นคืนข้อมูลกฎหมาย: ระบบมีค่า Hit Rate@1 และ Hit Rate@3 เท่ากันที่ 0.667 สะท้อนให้เห็นว่ากลไก Legal Authority Re-ranking ทำงานได้อย่างเสถียร มาตราที่ตรงตามเงื่อนไขทางคดีจะถูกบังคับจัดลำดับให้อยู่ในอันดับที่ 1 เสมอ

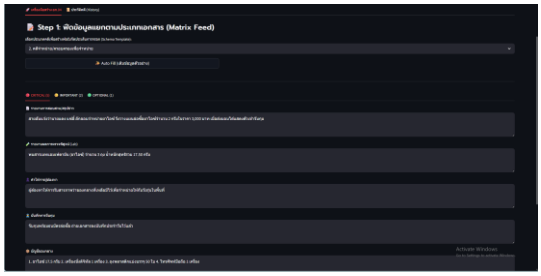
ด้านการสกัดข้อมูลข้อเท็จจริงสำคัญ: ค่า NER F1-Score = 0.700 แสดงถึงความสามารถในการระบุข้อมูลเฉพาะทาง เช่น ชื่อผู้ต้องหา ชนิดยาเสพติด และปริมาณสุทธิ ซึ่งเป็นผลจากการปรับแต่งด้วย LoRA บนชุดข้อมูลเฉพาะทาง

ด้านคุณภาพของผลลัพธ์: การปรับแต่งด้วย LoRA SFT ร่วมกับ CoT Prompting ช่วยยกระดับค่า SacreBLEU จาก 8.90 เป็น 56.10 สะท้อนให้เห็นว่าแบบจำลองสามารถเรียนรู้รูปแบบการวินิจฉัยและลำดับการให้เหตุผลทางกฎหมายได้อย่างมีประสิทธิภาพ

เพื่อให้เห็นระดับประสิทธิภาพของระบบในบริบท Thai Legal NLP ตารางที่ 4 แสดงการเปรียบเทียบเชิง Task กับงานวิจัยที่เกี่ยวข้อง ทั้งนี้แต่ละงานมีขอบเขต Task และชุดข้อมูลที่แตกต่างกัน จึงไม่สามารถเปรียบเทียบโดยตรงได้

### 4.3 ผลการนำระบบไปใช้งาน

ระบบถูกพัฒนาในรูปแบบเว็บแอปพลิเคชัน โดยผู้ใช้กรอกข้อมูลสำนวนคดีแยกตามประเภทเอกสาร ระบบจะประมวลผลผ่าน Pipeline ตามกรอบแนวคิดงานวิจัย แล้วสร้างร่างความเห็นและคำสั่งในรูปแบบ อ.ก.4 โดยอัตโนมัติ ดังแสดงในภาพที่ 2 และ 3 ผลลัพธ์ครอบคลุมทั้งการปรับบทความผิด การอ้างอิงมาตรากฎหมาย และคำสั่งริบของกลางได้ถูกต้องตามโครงสร้างที่กำหนด




**ภาพที่ 2:** หน้าจอการกรอกข้อมูลสำนวนคดีแยกตามประเภทเอกสาร



**ภาพที่ 3:** ผลลัพธ์ร่างความเห็นและคำสั่งคดีในรูปแบบ อ.ก.4 ที่ระบบสร้างขึ้นอัตโนมัติ

## 5. สรุปผล อภิปราย และข้อเสนอแนะ

### 5.1 สรุปผลการวิจัย

งานวิจัยนี้พัฒนาระบบต้นแบบสำหรับแนะนำรูปแบบการบันทึกความเห็นและคำสั่งคดียาเสพติด โดยเชื่อมต่อ NER, RAG และ LLM ที่ปรับแต่งด้วย LoRA เข้าเป็น Pipeline เดียว ผลการทดสอบบนชุดข้อมูล 611 รายการได้ NER F1-Score = 0.700, Hit Rate@1 และ @3 = 0.667 และ SacreBLEU เพิ่มจาก 8.90 เป็น 56.10 หลัง Fine-tuning

### 5.2 อภิปรายผลการวิจัย

การใช้ RAG ร่วมกับ Legal Authority Re-ranking ทำให้ระบบดึงตัวบทกฎหมายที่ถูกต้องมาประกอบการร่างได้สม่ำเสมอ ซึ่งลดปัญหา Hallucination ได้ดีกว่าการใช้ LLM อย่างเดียว ค่า Hit Rate@1 และ @3 ที่เท่ากันที่ 0.667 สะท้อนว่ากลไกนี้ทำงานได้เสถียร

ข้อจำกัดหลักคือชุดข้อมูลที่ใช้เป็นข้อมูลจำลอง ซึ่งมีภาษาเป็นระเบียบกว่าสำนวนที่พนักงานสอบสวนเขียนจริง และยังไม่ได้ผ่านการประเมินโดยผู้เชี่ยวชาญด้านกฎหมาย ผลที่ได้จึงยังเป็นการวัดในสภาพแวดล้อมที่ควบคุมได้เท่านั้น

### 5.3 ข้อเสนอแนะ

งานวิจัยในระยะต่อไปควร (1) ทดสอบบนสำนวนคดีจริงที่ผ่านการปกปิดข้อมูลส่วนบุคคล (2) ขยายไปยังฐานความผิดประเภทอื่น (3) ให้พนักงานอัยการหรือนิติกรประเมินผลลัพธ์จริง และ (4) เพิ่มกลไกป้อนกลับจากผู้ใช้งานเพื่อปรับปรุงต่อเนื่อง

## เอกสารอ้างอิง


[1] Gao, Y., Zhou, K., Zhang, W. X., and Wen, J. R. (2023). Retrieval-Augmented Generation for Large Language Models: A Survey. arXiv preprint arXiv:2312.10997.

[2] Zhao, W. X., Zhou, K., Li, J., Tang, T., Wang, X., Hou, Y., and Wen, J. R. (2023). A Survey of Large Language Models. arXiv preprint arXiv:2303.18223.

[3] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., and Polosukhin, I. (2017). Attention is All You Need. Advances in Neural Information Processing Systems, 30, 5998–6008.

[4] Hanwiboonwat, S. (2024). Enhancing Large Language Models for Thai Legal Chatbots. Master's thesis, Chulalongkorn University, Bangkok.

[5] Promwang, P., and Boonma, P. (2025). Development of a Retrieval-Augmented Generation System for Legal Data in Thai Language. Data Science and Engineering, 6(1), 225–243.

[6] Akarajaradwong, P., Pothavorn, P., Chaksangchaichot, C., Tasawong, P., Nopparatbundit, T., Pratai, K., and Nutanong, S. (2025). NitiBench: Benchmarking LLM Frameworks on Thai Legal Question Answering Capabilities. Proceedings of EMNLP 2025, 34304–34327.

[7] Karamitsos, I., Roufas, N., Al-Hussaeni, K., and Kanavos, A. (2025). LegNER: A Domain-Adapted Transformer for Legal Named Entity Recognition and Text Anonymization. Frontiers in Artificial Intelligence, 8, 1638971.

[8] Akarajaradwong, P., Chaksangchaichot, C., Pothavorn, P., Thamrongrattanarit-Rutherford, A., Chuangsuwanich, E., and Nutanong, S. (2025). Aligning LLMs for Thai Legal Question Answering with Efficient Semantic-Similarity Rewards. Proceedings of NLLP Workshop 2025.

[9] Piyawat, P., and others. (2026). Typhoon-S: Minimal Open Post-Training for Sovereign Large Language Models. arXiv preprint arXiv:2601.18129.

[10] Es, S., James, J., Espinosa-Anke, L., and Schockaert, S. (2024). RAGAS: Automated Evaluation of Retrieval Augmented Generation. Proceedings of EACL 2024. arXiv preprint arXiv:2309.15217.

[11] Wei, J., Wang, X., Schuurmans, D., Bosma, M., Ichter, B., Xia, F., and Zhou, D. (2022). Chain-of-Thought Prompting Elicits Reasoning in Large Language Models. Advances in Neural Information Processing Systems, 35, 24824–24837.

[12] Hu, E. J., Shen, Y., Wallis, P., Allen-Zhu, Z., Li, Y., Wang, S., and Chen, W. (2022). LoRA: Low-Rank Adaptation of Large Language Models. Proceedings of ICLR 2022.

[13] Post, M. (2018). A Call for Clarity in Reporting BLEU Scores. Proceedings of the Third Conference on Machine Translation (WMT), 186–191.

# การพัฒนาเครื่องมือประเมินระดับความมั่นคงปลอดภัยของข้อมูลส่วนบุคคลสำหรับองค์กรตามมาตรฐาน ISO/IEC 27701 และพระราชบัญญัติคุ้มครองข้อมูลส่วนบุคคล พ.ศ. 2562

# Development of an Assessment Tool for Personal Data Security Management in Organizations According to ISO/IEC 27701 Standard and Thailand's Personal Data Protection Act (PDPA)

*นิธิ สกุลพานิช (Nithi Sakunphanish)[1] และ ณัฐวี อุตกฤษฎ์ (Nattavee Utakrit)[2]*
*ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล*
*มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ*
*[1]s6707021856086@email.kmutnb.ac.th, [2]nattavee.u@itd.kmutnb.ac.th*

**บทคัดย่อ**

*งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาเครื่องมือประเมินระดับความมั่นคงปลอดภัยของข้อมูลส่วนบุคคลสำหรับองค์กร เพื่อให้สอดคล้องกับมาตรฐานสากล ISO/IEC 27701 และพระราชบัญญัติคุ้มครองข้อมูลส่วนบุคคล พ.ศ. 2562 (PDPA) โดยกระบวนการวิจัยประกอบด้วยการวิเคราะห์ข้อกำหนด การออกแบบชุดคำถามประเมินจำนวน 80 ข้อ และการพัฒนาประมวลผลอัตโนมัติผ่าน Google Apps Script เพื่อระบุช่องว่างความปลอดภัย (Gap Analysis) ผลการวิจัยพบว่าเครื่องมือมีความเที่ยงตรงเชิงเนื้อหา (IOC = 0.67–1.00) ผลการประเมินความเหมาะสมโดยผู้เชี่ยวชาญอยู่ในระดับดี ($\bar{x}$ = 4.43, S.D. = 0.66) โดยด้านความครอบคลุมได้รับการประเมินสูงสุดในระดับดีเยี่ยม ($\bar{x}$ = 4.87) และผลการทดลองใช้งานกับองค์กรนำร่อง 6 แห่ง พบว่าผู้ใช้งานมีความพึงพอใจในระดับมาก ($\bar{x}$ = 3.88) แสดงให้เห็นว่าเครื่องมือนี้สามารถช่วยองค์กรประเมินและยกระดับธรรมาภิบาลข้อมูลส่วนบุคคลได้อย่างเป็นรูปธรรม*

**คำสำคัญ:** เครื่องมือประเมิน การปฏิบัติตามมาตรฐาน ISO/IEC 27701 พระราชบัญญัติคุ้มครองข้อมูลส่วนบุคคล

**Abstract**

*This research aims to develop a tool for assessing personal data security levels within organizations, aligning with the ISO/IEC 27701 international standard and Thailand's Personal Data Protection Act (PDPA). The methodology includes analyzing regulatory requirements, designing an 80-item assessment questionnaire, and developing an automated system using Google Apps Script to perform gap analysis. The results indicate that the tool demonstrated high content validity (IOC = 0.67–1.00). Expert evaluation of the tool's appropriateness yielded a good overall rating ($\bar{x}$ = 4.43, S.D. = 0.66), with the coverage dimension receiving the highest score at an excellent level ($\bar{x}$ = 4.87). Pilot testing across six organizations showed a high level of user satisfaction ($\bar{x}$ = 3.88), demonstrating that the tool effectively supports organizations in assessing and enhancing their personal data governance in a practical and measurable manner.*

**Keywords:** Assessment Tool, Compliance, ISO/IEC 27701, Personal Data Protection Act (PDPA).

## 1. บทนำ

ปัจจุบันข้อมูลส่วนบุคคลกลายเป็นทรัพยากรที่มีค่ามหาศาลในการขับเคลื่อนองค์กร แต่ในขณะเดียวกันก็มีความเสี่ยงสูงที่จะเกิดการละเมิดข้อมูล ซึ่งส่งผลกระทบอย่างรุนแรงทั้งในด้านการเงินและชื่อเสียง จึงได้มีการประกาศใช้พระราชบัญญัติคุ้มครองข้อมูลส่วนบุคคล พ.ศ. 2562 (PDPA) ในประเทศไทย และมีการพัฒนามาตรฐานสากล ISO/IEC 27701 ขึ้น เพื่อเป็นแนวทางในการบริหารจัดการข้อมูลส่วนบุคคลให้มีความมั่นคงปลอดภัยและโปร่งใส อย่างไรก็ตาม การนำข้อบังคับทางกฎหมายและมาตรฐานสากลมาประยุกต์ใช้จริงในองค์กรนั้นมีความซับซ้อนสูง และต้องอาศัยการตีความที่ถูกต้องเพื่อให้เกิดผลในทางปฏิบัติ [1]

ด้วยเหตุนี้ องค์กรจำนวนมากจึงประสบปัญหาการขาดแคลนเครื่องมือที่ช่วยประเมินตนเองอย่างเป็นระบบ ทำให้ไม่ทราบถึงช่องว่าง หรือระดับความพร้อมที่แท้จริงในการคุ้มครองข้อมูลส่วนบุคคล ผู้วิจัยจึงได้พัฒนาเครื่องมือประเมินระดับความมั่นคงปลอดภัยของข้อมูลส่วนบุคคลขึ้น โดยบูรณาการข้อกำหนดจากทั้ง PDPA และ ISO/IEC 27701 เข้าด้วยกัน เพื่อให้องค์กรมีเครื่องมือที่ทันสมัย ใช้งานง่าย และสามารถประมวลผลความสอดคล้องได้โดยอัตโนมัติ อันจะช่วยให้องค์กรสามารถปรับปรุงกระบวนการทำงานได้อย่างตรงจุดและสร้างความเชื่อมั่นให้แก่เจ้าของข้อมูลได้อย่างยั่งยืน

## 2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

### 2.1 มาตรฐาน ISO/IEC 27701 Privacy Information Management System (PIMS)

การบริหารจัดการด้านการคุ้มครองข้อมูลส่วนบุคคล (Personal Data Protection Management) ในงานวิจัยนี้มีรากฐานทางทฤษฎีจากมาตรฐาน ISO/IEC 27701 [2] ซึ่งเป็นส่วนขยายของระบบบริหารจัดการความมั่นคงปลอดภัยสารสนเทศ (ISMS) โดยกำหนดแนวทางการจัดตั้งระบบบริหารจัดการข้อมูลส่วนบุคคล (Privacy Information Management System: PIMS) และแยกแยะหน้าที่ความรับผิดชอบระหว่างผู้ควบคุมข้อมูลส่วนบุคคล (PII Controller) และผู้ประมวลผลข้อมูลส่วนบุคคล (PII Processor) เพื่อให้องค์กรสามารถระบุและควบคุมความเสี่ยงด้านความเป็นส่วนตัวได้อย่างเป็นระบบ

### 2.2 พระราชบัญญัติคุ้มครองข้อมูลส่วนบุคคล พ.ศ. 2562 หรือ PDPA (Personal Data Protection Act)

พระราชบัญญัติคุ้มครองข้อมูลส่วนบุคคล พ.ศ. 2562 (PDPA) [3] กำหนดหน้าที่ตามกฎหมายที่องค์กรต้องปฏิบัติตามเพื่อคุ้มครองข้อมูล และสิทธิของเจ้าของข้อมูลส่วนบุคคล การปฏิบัติตามกฎหมาย PDPA ไม่ได้จำกัดเพียงการมีนโยบายความเป็นส่วนตัว (Privacy Policy) เท่านั้น แต่ต้องครอบคลุมมาตรการรักษาความมั่นคงปลอดภัยทั้งในเชิงบริหารจัดการ (Administrative Safeguards) เชิงเทคนิค (Technical Safeguards) และเชิงกายภาพ (Physical Safeguards) เพื่อป้องกันการละเมิดข้อมูลส่วนบุคคลในทุกมิติ [4]

### 2.3 เทคโนโลยีที่เกี่ยวข้อง

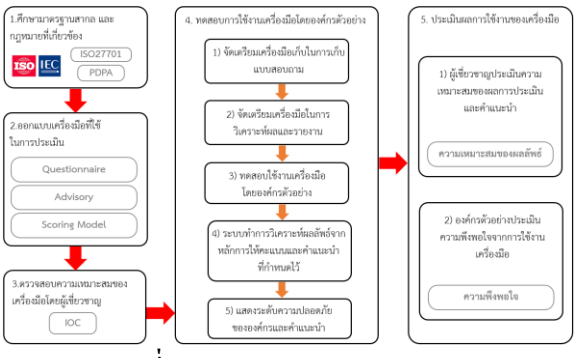
เครื่องมือที่พัฒนาขึ้นมุ่งเน้นความรวดเร็วและใช้งานง่าย โดยอาศัยชุดเครื่องมือของ Google LLC ตลอดกระบวนการทำงาน การตอบแบบสอบถามดำเนินการผ่าน Google Forms [5] ซึ่งเข้าถึงได้โดยไม่ต้องติดตั้งซอฟต์แวร์เพิ่มเติม คำตอบที่ได้รับจะถูกจัดเก็บใน Google Sheets [6] โดยอัตโนมัติ จากนั้น Google Apps Script [7][8] ที่พัฒนาด้วยภาษา JavaScript [9][10] จะประมวลผลคะแนนตาม Scoring Model คำนวณระดับความสอดคล้อง (Compliance %) รายหมวดหมู่และโดยรวม เชื่อมโยงผลคะแนนกับคำแนะนำการพัฒนาที่ตรงกับสถานะขององค์กร จัดทำรายงาน PDF และจัดส่งให้ผู้ประเมินผ่านอีเมลโดยอัตโนมัติ

## 3. วิธีการดำเนินการวิจัย

### 3.1 กรอบแนวทางดำเนินงานวิจัย

การวิจัยนี้ใช้วิธีดำเนินการวิจัยและพัฒนา (Research and Development) เพื่อสร้างเครื่องมือประเมินความมั่นคงปลอดภัยของข้อมูลส่วนบุคคล โดยแบ่งกระบวนการดำเนินงานออกเป็น 5 ขั้นตอนหลัก ได้แก่ การศึกษาข้อมูลที่เกี่ยวข้อง การออกแบบเครื่องมือประเมิน การตรวจสอบความเหมาะสมโดยผู้เชี่ยวชาญ การทดสอบ

การใช้งานโดยองค์กรตัวอย่าง และการประเมินผลการใช้งานของเครื่องมือ




**ภาพที่ 1:** กรอบแนวทางการดำเนินงานวิจัย

**3.2 การศึกษาข้อมูลที่เกี่ยวข้อง**

ผู้วิจัยดำเนินการสกัดข้อกำหนดและมาตรการควบคุมจากมาตรฐาน ISO/IEC 27701 (Annex A และ Annex B) และบทบัญญัติในพระราชบัญญัติคุ้มครองข้อมูลส่วนบุคคล พ.ศ. 2562 (PDPA) เพื่อนำมาจำแนกหมวดหมู่และกำหนดตัวชี้วัดความสอดคล้อง (Compliance Indicators) ที่จำเป็นสำหรับองค์กร

**3.3 การออกแบบเครื่องมือประเมิน**

เครื่องมือที่พัฒนาขึ้นประกอบด้วย 3 ส่วนหลัก ได้แก่ แบบสอบถามการประเมิน คำแนะนำการพัฒนา และแนวทางการจัดระดับผลการประเมิน โดยผู้วิจัยดำเนินการออกแบบชุดคำถามประเมินจำนวน 80 ข้อ อ้างอิงข้อกำหนดจากมาตรฐาน ISO/IEC 27701 เฉพาะ Clause 6, 7 และ 8 ซึ่งครอบคลุมมาตรการรักษาความมั่นคงปลอดภัยสารสนเทศ ข้อกำหนดสำหรับผู้ควบคุมข้อมูล (PII Controller) และข้อกำหนดสำหรับผู้ประมวลผลข้อมูล (PII Processor) ตามลำดับ ร่วมกับหมวดที่เกี่ยวข้องโดยตรงจากพระราชบัญญัติคุ้มครองข้อมูลส่วนบุคคล พ.ศ. 2562 โดยแบ่งออกเป็น 3 หมวดหมู่หลัก ได้แก่ คำถามสำหรับสอบถามข้อมูลทั่วไปขององค์กร คำถามสำหรับผู้ควบคุมข้อมูลส่วนบุคคล และคำถามสำหรับผู้ประมวลผลข้อมูลส่วนบุคคล รูปแบบคำถามเป็นปลายปิด (Yes/No และ Multiple Choice) พร้อมกำหนดระดับคะแนนแต่ละข้อเพื่อคำนวณระดับความสอดคล้องต่อมาตรฐานและกฎหมายพร้อมทั้งกำหนดโมเดลการให้คะแนนแบบการวิเคราะห์ช่องว่าง (Gap Analysis) และเกณฑ์การแบ่งระดับความสอดคล้อง (Compliance Level) เพื่อให้องค์กรสามารถระบุสถานะความปลอดภัยได้อย่างชัดเจน

ผู้วิจัยออกแบบคำแนะนำการพัฒนาจำนวน 65 คำแนะนำสำหรับแต่ละข้อคำถาม โดยออกแบบให้ระบบแสดงคำแนะนำโดยอัตโนมัติเมื่อองค์กรตอบในลักษณะที่บ่งชี้ว่ายังไม่ได้ปฏิบัติหรือปฏิบัติในระดับต่ำ

แนวทางการจัดกระดับผลการประเมิน ระบบแปลงคำตอบของแต่ละข้อเป็นคะแนนตามเกณฑ์ 3 ระดับ ได้แก่ Compliance (C) = 1 คะแนน เมื่อองค์กรดำเนินการครบถ้วน Partial-Compliance (PC) = 0.5 คะแนน เมื่อดำเนินการบางส่วน และ Non-Compliance = 0 คะแนน เมื่อยังไม่ดำเนินการ จากนั้นคำนวณค่า Compliance (%) โดยใช้สูตร (คะแนนที่ได้รับจริง / คะแนนเต็มสูงสุด) × 100 แยกรายหมวดหมู่และโดยรวม เพื่อสะท้อนระดับการคุ้มครองข้อมูลส่วนบุคคลขององค์กรในแต่ละมิติ

**3.4 การตรวจสอบความเหมาะสมของเครื่องมือโดยผู้เชี่ยวชาญ**

ผู้วิจัยดำเนินการตรวจสอบความตรงของเนื้อหา (Content Validity) ของแบบสอบถามด้วยค่าดัชนีความสอดคล้อง (Index of Item-Objective Congruence: IOC) โดยผู้เชี่ยวชาญ 3 ท่าน ซึ่งมีประสบการณ์ด้านการจัดการความปลอดภัยข้อมูลส่วนบุคคล มีความเชี่ยวชาญด้าน ISO/IEC 27701 กฎหมาย PDPA และเป็นผู้เชี่ยวชาญหรือที่ปรึกษาด้านความมั่นคงปลอดภัยสารสนเทศ หรือการคุ้มครองข้อมูล ผู้เชี่ยวชาญแต่ละท่านประเมินข้อคำถามโดยให้คะแนน +1 (ตรงกับวัตถุประสงค์) 0 (ไม่แน่ใจ) หรือ -1 (ไม่ตรงกับวัตถุประสงค์) จากนั้นคำนวณค่า IOC โดยใช้สูตร $IOC = \Sigma R / N$ โดยกำหนดเกณฑ์ผ่านที่ค่า $IOC \geq 0.50$ ข้อใดที่ได้ค่าต่ำกว่าเกณฑ์จะได้รับการปรับปรุงหรือตัดออกก่อนนำไปทดสอบจริง

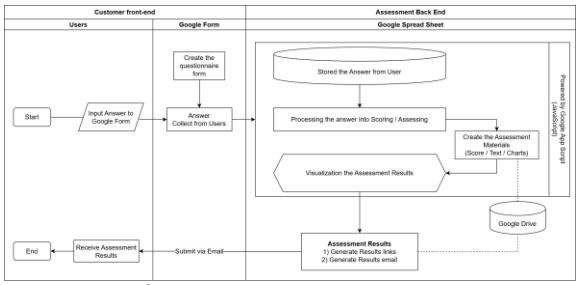
เพื่อยืนยันคุณภาพของคำแนะนำการพัฒนา ผู้วิจัยให้ผู้เชี่ยวชาญ 3 ท่านประเมินคำแนะนำทุกรายการด้วย Likert Scale 5 ระดับ ใน 2 ด้าน คือ ด้านความถูกต้อง (Accuracy) และด้านความสามารถนำไปปฏิบัติได้ (Practicability) โดยใช้เกณฑ์การแปลผลตั้งแต่ระดับ

"ควรปรับปรุง" (ค่าเฉลี่ย < 3.00) จนถึง "ดีเยี่ยม" (ค่าเฉลี่ย ≤ 4.50)

**3.5 การทดสอบการใช้งานโดยองค์กรตัวอย่าง**

ผู้วิจัยคัดเลือกองค์กรตัวอย่างจำนวน 6 องค์กร จากภาคเอกชนที่มีลักษณะธุรกิจแตกต่างกัน ครอบคลุม 5 ประเภทธุรกิจ ได้แก่ ธุรกิจการจัดงานอีเว้นท์ ธุรกิจจำหน่ายระบบและอุปกรณ์ไฟฟ้า ธุรกิจรับเหมาก่อสร้าง ธุรกิจให้คำปรึกษาด้านการก่อสร้างและตรวจสอบที่อยู่อาศัย และธุรกิจให้คำปรึกษาและตรวจสอบระบบความมั่นคงปลอดภัยสารสนเทศ การคัดเลือกองค์กรที่มีความหลากหลายดังกล่าวมีวัตถุประสงค์เพื่อทดสอบความสามารถของเครื่องมือในการรองรับองค์กรที่มีบริบทแตกต่างกัน ทั้งในด้านขนาด ลักษณะการประมวลผลข้อมูลส่วนบุคคล และระดับความเชี่ยวชาญด้านเทคโนโลยีสารสนเทศ ตั้งแต่องค์กรที่ไม่มีความเชี่ยวชาญด้านเทคโนโลยีเป็นพิเศษจนถึงองค์กรที่มีความเชี่ยวชาญด้านความมั่นคงปลอดภัยสารสนเทศโดยตรง

การทดสอบดำเนินการโดยตัวแทนขององค์กร เช่น ผู้บริหารขององค์กร ผ่านระบบ Google Forms โดยใช้ Google Sheets ทำหน้าที่จัดเก็บคำตอบ และ Google Apps Script ดำเนินการประมวลผลตาม Scoring Model คำนวณค่า Compliance (%) เชื่อมโยงผลคะแนนกับคำแนะนำ จัดทำรายงานผลการประเมินในรูปแบบ PDF และจัดส่งให้ตัวแทนขององค์กรผ่านอีเมลโดยอัตโนมัติ




**ภาพที่** 2: ไดอะแกรมเบื้องหลังระบบอัตโนมัติ

**3.6 การประเมินผลการใช้งานของเครื่องมือ**

ผู้วิจัยดำเนินการประเมินผลการใช้งานของเครื่องมือ แบ่งเป็น 2 ส่วน ได้แก่ การประเมินความเหมาะสมของเครื่องมือโดยผู้เชี่ยวชาญ ซึ่งประเมินใน 3 ด้าน ได้แก่ ความเหมาะสม (Appropriateness) ความชัดเจน (Clarity) และความครอบคลุม (Coverage) ด้วย Likert Scale 5 ระดับ โดยใช้เกณฑ์แปลผลตั้งแต่ "ควรปรับปรุง" (< 3.00) จนถึง "ดีเยี่ยม" (4.50–5.00) และการประเมินความพึงพอใจจากองค์กรตัวอย่าง โดยก่อนจัดส่งแบบสอบถามความพึงพอใจ ผู้วิจัยตรวจสอบความตรงของเนื้อหาด้วยค่า IOC จากผู้เชี่ยวชาญชุดเดิม จากนั้นจัดส่งแก่ตัวแทนองค์กรตัวอย่างองค์กรละ 1 คน วิเคราะห์ผลด้วยค่าเฉลี่ย (Mean) และส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation) โดยแปลผลตามเกณฑ์ของ Best W. John (1997)

## 4. ผลการดำเนินงาน

**4.1 ผลการประเมินคุณภาพของเครื่องมือโดยผู้เชี่ยวชาญ**

จากการนำชุดคำถามประเมินจำนวน 80 ข้อ และชุดคำแนะนำในการปรับปรุงจำนวน 65 รายการ เสนอต่อผู้เชี่ยวชาญเพื่อตรวจสอบความเที่ยงตรงเชิงเนื้อหา พบว่าข้อคำถามทั้งหมดมีค่าดัชนีความสอดคล้อง (IOC) อยู่ระหว่าง 0.67 ถึง 1.00 ซึ่งสูงกว่าเกณฑ์ที่กำหนด (0.50) โดยร้อยละ 86.25 ของข้อคำถามได้รับคะแนนความเที่ยงตรงสูงสุด

**ตารางที่ 1:** ผลการวิเคราะห์ความดัชนีความสอดคล้อง (IOC) ของข้อคำถามชุด ข้อมูลพื้นฐานขององค์กร

| IOC | จำนวนข้อคำถาม | รายการข้อคำถาม | ผลการพิจารณา |
|---|---|---|---|
| IOC = 1 | 5 ข้อ | O1, O2, O2.1, O3, O4 | ผ่านการประเมิน |
| **รวม** | **5 ข้อ** | | **ผ่านการประเมินทุกข้อ** |

**ตารางที่ 2:** ผลการวิเคราะห์ความดัชนีความสอดคล้อง (IOC) ของข้อคำถามชุด PII Controller

| IOC | จำนวนข้อคำถาม | รายการข้อคำถาม | ผลการพิจารณา |
|---|---|---|---|
| IOC = 1 | 36 ข้อ | C1, C1.2, C2, C3, C4, C5, C6, C6.2, C6.3, C6.4, C6.5, C6.6, C6.7, C6.8, C7, C8.1, C8.2, C8.4, C8.5, C9, C10, C10.1, C13, C14, C15, C16, C17, C18, C18.1, C19, C20, C21, C21.1, C22, C22.1, C22.2 | ผ่านการประเมิน |
| IOC = 0.67 | 8 ข้อ | C1.1, C6.1, C7.1, C8, C8.3, C11, C12, C21.2 | ผ่านการประเมิน |
| **รวม** | **44 ข้อ** | | **ผ่านการประเมินทุกข้อ** |

**ตารางที่ 3:** ผลการวิเคราะห์ความดัชนีความสอดคล้อง (IOC) ของข้อคำถามชุด PII Processor

| IOC | จำนวนข้อคำถาม | รายการข้อคำถาม | ผลการพิจารณา |
|---|---|---|---|
| IOC = 1 | 28 ข้อ | P1, P1.1, P1.2, P1.3, P1.4, P1.5, P1.6, P1.7, P3, P4, P5, P6, P6.1, P7, P8, P8.1, P9, P10, P11, P12, P13, P14, P15, P16, P17, P18, P18.2, P18.3, P18.4 | ผ่านการประเมิน |
| IOC = 0.67 | 2 ข้อ | P2, P18.1 | ผ่านการประเมิน |
| **รวม** | **31 ข้อ** | | **ผ่านการประเมินทุกข้อ** |

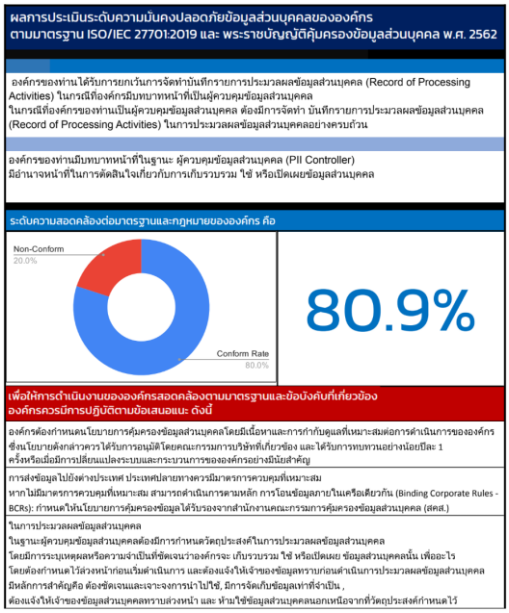
ผู้วิจัยให้ผู้เชี่ยวชาญ 3 ท่านประเมินคุณภาพคำแนะนำทั้ง 65 รายการใน 2 ด้าน ได้แก่ ด้านความถูกต้อง (Accuracy) และด้านความสามารถนำไปปฏิบัติได้ (Practicability) ผลการประเมินภาพรวมอยู่ในระดับดีเยี่ยม ($\bar{x}$ = 4.72, S.D. = 0.51) เมื่อพิจารณารายหมวด พบว่าหมวดข้อมูลทั่วไปขององค์กรได้รับการประเมินสูงสุด ($\bar{x}$ = 4.83) รองลงมาคือหมวด PII Processor ($\bar{x}$ = 4.73) และหมวด PII Controller ($\bar{x}$ = 4.69) ตามลำดับ โดยทุกหมวดอยู่ในระดับดีเยี่ยม ยกเว้นด้านความสามารถนำไปปฏิบัติได้ของหมวด PII Controller ($\bar{x}$ = 4.60) ที่อยู่ในระดับดี ซึ่งสะท้อนถึงความซับซ้อนในเชิงปฏิบัติของข้อกำหนดในหมวดดังกล่าว

**ตารางที่ 4:** สรุปผลการประเมินคุณภาพคำแนะนำ

| หมวดหมู่ | $\bar{x}$ ความถูกต้อง | $\bar{x}$ นำไปปฏิบัติได้ | $\bar{x}$ ภาพรวม | S.D. | ระดับ |
|---|---|---|---|---|---|
| ข้อมูลทั่วไป | 4.78 | 4.89 | 4.83 | 0.35 | ดีเยี่ยม |
| PII Controller | 4.77 | 4.60 | 4.69 | 0.55 | ดีเยี่ยม |
| PII Processor | 4.78 | 4.68 | 4.73 | 0.48 | ดีเยี่ยม |
| **ภาพรวมทั้งหมด** | **4.78** | **4.72** | **4.72** | **0.51** | **ดีเยี่ยม** |

### 4.2 ผลการทดสอบของระบบประเมินอัตโนมัติ

ระบบที่พัฒนาขึ้นด้วย Google Apps Script สามารถประมวลผลข้อมูลจากแบบประเมินและจัดทำรายงานผล (Gap Analysis Report) ได้อย่างถูกต้องแม่นยำ โดยสามารถระบุระดับความสอดคล้องแยกตามหมวดหมู่ของ ISO/IEC 27701 และ PDPA ช่วยให้องค์กรทราบถึงจุดอ่อนและแนวทางการปรับปรุงในรูปแบบรายงานสรุปได้ทันทีหลังเสร็จสิ้นการประเมิน




**ภาพที่ 3**: ตัวอย่างรายงานผลการประเมิน

จากการทดสอบกับองค์กรตัวอย่างทั้ง 6 แห่ง พบว่า PII Controller มีระดับความสอดคล้องระหว่าง 59.5% ถึง 80.9% และ PII Processor มีระดับความสอดคล้องระหว่าง 27.7% ถึง 72.2% องค์กรส่วนใหญ่มีความเข้าใจในบทบาทหน้าที่ตาม PDPA และมีการดำเนินการด้านการคุ้มครองข้อมูลส่วนบุคคลในระดับพื้นฐานบางส่วนแล้ว อย่างไรก็ตาม ช่องว่างที่พบร่วมกัน ได้แก่ การขาด Data Processing Agreement ที่ครบถ้วน RoPA ที่ยังไม่สมบูรณ์ และมาตรการรักษาความมั่นคงปลอดภัยตลอด Data Lifecycle ที่ยังไม่ครอบคลุม ซึ่งระบบสามารถระบุช่องว่างและจัดส่งคำแนะนำที่ตรงกับสถานะขององค์กรแต่ละแห่งได้โดยอัตโนมัติ

### 4.3 ผลการประเมินความเหมาะสมของผลลัพธ์

จากผลการประเมินความเหมาะสมของเครื่องมือโดยผู้เชี่ยวชาญ 3 ท่าน ใน 3 ด้าน ภาพรวมอยู่ในระดับดี ($\bar{x}$ = 4.43, S.D. = 0.66) ด้านที่ได้รับการประเมินสูงสุดคือด้านความครอบคลุม (Coverage) ระดับดีเยี่ยม สะท้อนว่าเครื่องมือครอบคลุมทั้ง ISO/IEC 27701 และ PDPA ได้ครบถ้วนในทุกมิติ ทั้ง Data Lifecycle, Data Governance และทั้งสองบทบาทขององค์กร รองลงมาคือด้านความชัดเจน (Clarity) และด้านความเหมาะสม (Appropriateness) ตามลำดับ โดยทั้งสองด้านอยู่ในระดับดี

**ตารางที่ 5:** สรุปผลการประเมินความเหมาะสมของผลลัพธ์โดยผู้เชี่ยวชาญ

| ข้อที่ | รายการประเมิน | $\bar{x}$ | S.D. | ระดับคุณภาพ |
|---|---|---|---|---|
| 1 | ด้านความเหมาะสม (Appropriateness) | 4.30 | 0.59 | ดี |
| 2 | ด้านความชัดเจน (Clarity) | 4.33 | 0.75 | ดี |
| 3 | ด้านความครอบคลุม (Coverage) | 4.87 | 0.34 | ดีเยี่ยม |
| **ค่าเฉลี่ยรวม** | | 4.43 | 0.66 | ดี |

### 4.4 ผลการวิเคราะห์ความพึงพอใจ

ผู้วิจัยดำเนินการวิเคราะห์ผลการประเมินความพึงพอใจโดยใช้สถิติเชิงพรรณนา (Descriptive Statistics) ตามที่กำหนดไว้ในขอบเขตงานวิจัย จากการนำเครื่องมือทดสอบใช้งานจริงกับองค์กรนำร่อง 6 แห่ง พบว่าผู้ใช้งานมีความพึงพอใจโดยรวมอยู่ในระดับ "มาก" ($\bar{x}$ = 3.88, S.D. = 0.72) ตามเกณฑ์ของ Best W. John (1997) และสูงกว่าเกณฑ์ที่กำหนดในสมมติฐาน ($\mu > 3.50$)

**ตารางที่ 6:** ผลการประเมินความพึงพอใจองค์กรตัวอย่าง

| ข้อที่ | รายการประเมิน | $\bar{x}$ | S.D | ระดับความพึงพอใจ |
|---|---|---|---|---|
| 1 | ระบบมีขั้นตอนการใช้งานที่เข้าใจง่ายและชัดเจน | 4.00 | 0.89 | มาก |
| 2 | การนำทางผ่านแบบสอบถาม (Navigation) สะดวกและลื่นไหล | 4.00 | 0.63 | มาก |
| 3 | ระบบสามารถใช้งานได้โดยไม่ต้องอาศัยความเชี่ยวชาญด้านเทคนิคสูง | 3.83 | 1.17 | มาก |
| 4 | คำถาม/ข้อความในแบบประเมินมีความชัดเจน ไม่คลุมเครือ | 3.67 | 1.03 | มาก |
| 5 | ระยะเวลาในการทำแบบประเมินภายในระบบมีความเหมาะสม ไม่ใช้เวลามากเกินไป | 3.67 | 0.52 | มาก |
| 6 | คะแนน/การจัดระดับ (Scoring Model) และคำแนะนำที่ได้รับสะท้อนสถานะจริงขององค์กรและนำไปปฏิบัติได้ | 3.67 | 0.52 | มาก |
| 7 | รายงานผล (Report) ชัดเจน เข้าใจง่าย และช่วยให้องค์กรเห็นช่องว่าง (Gap) ในการคุ้มครองข้อมูลส่วนบุคคลได้ | 3.83 | 0.41 | มาก |
| 8 | ท่านมีความเชื่อมั่นในความถูกต้องและความน่าเชื่อถือของผลการประเมินที่ได้รับ | 4.50 | 0.55 | มาก |
| 9 | รายงานผลมีประโยชน์สำหรับการวางแผนปรับปรุงการบริหารจัดการข้อมูลส่วนบุคคลขององค์กร | 3.83 | 0.75 | มาก |
| 10 | ความพึงพอใจของท่านต่อระบบประเมินในภาพรวม | 3.83 | 0.41 | มาก |
| **ค่าเฉลี่ยรวม** | | **3.88** | **0.71** | **มาก** |

## 5. สรุปผลและอภิปรายผล

งานวิจัยนี้พัฒนาเครื่องมือประเมินที่บูรณาการมาตรฐาน ISO/IEC 27701 และ PDPA เข้าด้วยกัน โดยผลการตรวจสอบจากผู้เชี่ยวชาญยืนยันว่าชุดคำถาม 80 ข้อมีความเที่ยงตรงเชิงเนื้อหา (IOC = 0.67–1.00) คำแนะนำ 65 รายการมีคุณภาพระดับดีเยี่ยม ($\bar{x}$ = 4.72) และผลการทดสอบกับองค์กรนำร่อง 6 แห่ง พบว่าผู้ใช้งานมีความพึงพอใจในระดับมาก ($\bar{x}$ = 3.88) สะท้อนว่าเครื่องมือที่พัฒนาขึ้นสามารถตอบโจทย์การใช้งานจริงได้

งานวิจัยนี้แสดงให้เห็นว่าการบูรณาการข้อกำหนดจากมาตรฐาน ISO/IEC 27701 และพระราชบัญญัติคุ้มครองข้อมูลส่วนบุคคล พ.ศ. 2562 เข้าด้วยกันในรูปแบบเครื่องมือประเมินที่ใช้งานได้จริงนั้นเป็นสิ่งที่เป็นไปได้และได้รับการยอมรับในระดับดี เครื่องมือที่พัฒนาขึ้นช่วยให้องค์กรสามารถประเมินความสอดคล้องได้ครบทั้งมิติกฎหมายและมาตรฐานสากลในกระบวนการเดียว พร้อมรับทราบช่องว่าง และแนวทางการพัฒนาผ่านรายงาน PDF อัตโนมัติได้อย่างรวดเร็ว

ข้อเสนอแนะในการปรับปรุงเพื่อการวิจัยครั้งต่อไปสำหรับการพัฒนาในอนาคต ควรทบทวนและรวมข้อคำถามที่ซ้ำซ้อนให้กระชับขึ้น นำระบบ Logic Branching มาใช้ และเพิ่ม Tooltips/Guidance Notes เพื่อลดความกำกวมในการตีความ นอกจากนี้ควรพัฒนาให้สามารถจำแนกคำแนะนำตามขนาดและความพร้อมขององค์กรและเพิ่มระบบตรวจสอบหลักฐาน เพื่อยกระดับความน่าเชื่อถือของผลการประเมินให้เทียบเท่ากับการตรวจประเมินโดยบุคคลภายนอก (Third-party Audit)

## เอกสารอ้างอิง


[1] DPO Centre. (2025). ISO 27701:2025 update: What's changed and why it matters. https://www.dpocentre.com/blog/iso-27701-2025-whats-changed/

[2] International Organization for Standardization & International Electrotechnical Commission. (2019). ISO/IEC 27701:2019 Security techniques — Extension to ISO/IEC 27001 and ISO/IEC 27002 for privacy information management — Requirements and guidelines. ISO.

[3] พระราชบัญญัติคุ้มครองข้อมูลส่วนบุคคล พ.ศ. 2562. (27 พฤษภาคม 2562). ราชกิจจานุเบกษา. เล่ม 136 ตอนที่ 69 ก หน้า 52-95.

[4] ประกาศคณะกรรมการคุ้มครองข้อมูลส่วนบุคคล เรื่อง มาตรการรักษาความมั่นคงปลอดภัยของผู้ควบคุมข้อมูลส่วนบุคคล พ.ศ. 2565. (20 มิถุนายน 2565). ราชกิจจานุเบกษา. เล่ม 139 ตอนพิเศษ 140 ง หน้า 28-31.

[5] Google Workspace Learning Center. (2026). Google Forms training and help. https://support.google.com/docs/topic/9054603

[6] Google Workspace Learning Center. (2026). Google Sheets training and help. https://support.google.com/docs/topic/9054603

[7] Google for Developers. (2026, February 3). Apps Script documentation. https://developers.google.com/apps-script

[8] Google for Developers. (2025, December 11). Extending Google Sheets. https://developers.google.com/apps-script/guides/sheets

[9] Mozilla Developer Network. (2026). JavaScript. https://developer.mozilla.org/en-US/docs/Web/JavaScript

[10] Haverbeke, M. (2024). Eloquent JavaScript: A modern introduction to programming (4th edition). https://eloquentjavascript.net/

# การพยากรณ์ผลสัมฤทธิ์ทางการเรียนในระดับประกาศนียบัตรวิชาชีพ ด้วยบุคลิกภาพ 5 องค์ประกอบ และเกรดวิชาหลักมัธยมศึกษาตอนต้น

# Predicting Academic Achievement of Vocational Certificate Using the Big Five Personality Factors and Core Subject Grades in Lower Secondary School

*อุไรวรรณ มีแสง (Uraiwan Meesang)[1] และ ธนพล เจนสุทธิเวชกุล (Tanapon Jensuttiwetchakul)[2]*
*สาขาวิชาระบบสารสนเทศเพื่อการจัดการ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล*
*มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ*
*[1]m.uraiwan.som@gmail.com, [2]tanapon.j@itd.kmutnb.ac.th*

## บทคัดย่อ

*การวิจัยนี้พัฒนาและเปรียบเทียบแบบจำลองการจำแนกข้อมูล 5 รูปแบบ ได้แก่ XGBoost, Random Forest, SVM, Naïve Bayes และ Neural Network สำหรับพยากรณ์ผลสัมฤทธิ์ทางการเรียน 5 ระดับ โดยเก็บข้อมูลจากนักเรียน ปวช. จำนวน 500 คน ใน 5 สาขาวิชา วิเคราะห์ด้วย RapidMiner แบ่งข้อมูล 70/30 ผลพบว่า XGBoost มีความแม่นยำสูงสุด 85.47% รองลงมาคือ Random Forest (85.05%) และ SVM (78.12%) ปัจจัยที่มีน้ำหนักสูงสุดคือ เกรดคณิตศาสตร์ (0.83) ภาษาอังกฤษ (0.11) และ Neuroticism (0.077) ซึ่งเป็นตัวแปรบุคลิกภาพที่มีนัยสำคัญเพียงมิติเดียว ผลการวิจัยสามารถนำไปพัฒนาเครื่องมือแนะแนวที่ช่วยให้นักเรียนเลือกสาขาได้สอดคล้องกับศักยภาพและบุคลิกภาพของตน*

**คำสำคัญ:** เทคนิคการจำแนก, บุคลิกภาพ 5 องค์ประกอบ, การพยากรณ์ผลสัมฤทธิ์ทางการเรียน, XGBoost, ความไม่มั่นคงทางอารมณ์

## Abstract

*This research develops and evaluates five classification models (XGBoost, Random Forest, SVM, Naïve Bayes, and Neural Network) to predict the academic achievement of vocational students across five levels. Based on a dataset of 500 students, the models were evaluated using a 70/30 data split. The results indicate that XGBoost achieved the highest predictive accuracy at 85.47%, outperforming Random Forest and SVM. The most critical predictors were Mathematics (0.83) and English (0.11) grades, alongside Neuroticism (0.077), which was the only significant personality trait. Ultimately, this study offers a foundation for designing predictive guidance systems that help students select academic paths corresponding to their abilities and personalities.*

**Keywords:** Data Classification, Big Five Personality Factors, Academic Achievement Prediction, XGBoost, Neuroticism.

## 1. บทนำ

ปัจจุบันผู้เรียนระดับประกาศนียบัตรวิชาชีพ(ปวช.) มีอัตราการออกกลางคันสูง สาเหตุสำคัญเกิดจากการตัดสินใจเลือกสาขาวิชาตามความนิยม และการชักจูงจากบุคคลรอบข้าง โดยขาดการพิจารณามิติด้านบุคลิกภาพและความถนัด ประกอบกับระบบการแนะแนวส่วนใหญ่มักประเมินเพียงผลการเรียนเดิมซึ่งไม่สะท้อนศักยภาพที่แท้จริงอย่างรอบด้าน

ด้วยเหตุนี้ ผู้วิจัยจึงมีแนวคิดพัฒนาแบบจำลองพยากรณ์ผลสัมฤทธิ์ทางการเรียนใน 5 สาขาวิชายอดนิยม โดยบูรณาการข้อมูลบุคลิกภาพ 5 องค์ประกอบร่วมกับผลการเรียนวิชา

พื้นฐาน โดยการพัฒนาและการเปรียบเทียบประสิทธิภาพของอัลกอริทึมที่มีความซับซ้อนกับอัลกอริทึมพื้นฐาน ได้แก่ XGBoost, SVM, Random Forest, Neural Network และ Naïve Bayes เพื่อศึกษาอิทธิพลของบุคลิกภาพที่มีต่อผลการเรียน โดยผลลัพธ์ที่ได้จะถูกนำไปต่อยอดเป็นระบบสนับสนุนการตัดสินใจเลือกสาขาวิชาเพื่อลดปัญหาการเรียนผิดสาขา ลดอัตราการออกกลางคัน และส่งเสริมให้ผู้เรียนประสบความสำเร็จในเส้นทางวิชาชีพได้อย่างยั่งยืน

## 2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ผู้วิจัยได้ศึกษาค้นคว้าเอกสาร และงานวิจัยที่เกี่ยวข้อง เพื่อนำทฤษฎี แนวคิดมาประยุกต์ใช้กับงานวิจัยดังนี้

**2.1 ทฤษฎีบุคลิกภาพ 5 องค์ประกอบ (Big Five Personality Factor)** พัฒนาโดย John, Naumann & Soto (2008) เป็นโดยใช้แบบวัดบุคลิกภาพ 44 ข้อ วัด 5 มิติ ได้แก่ การเปิดรับประสบการณ์ ด้านการมีจิตสำนึก ด้านการเปิดตัว ด้านความเป็นมิตร และด้านความไม่มั่นคงทางอารมณ์ ซึ่งได้รับการยอมรับอย่างกว้างขวางในการศึกษาความสัมพันธ์กับผลการเรียน [1] ทั้งนี้ พรสรัญ รุ่งเจริญกิจกุล. (2565) [2] ยังพบว่า การมีบุคลิกภาพที่ประสบความสำเร็จต้องพัฒนาจากคุณลักษณะภายใน เพื่อเป็นรากฐานที่มั่นคงในการพัฒนาทักษะด้านอื่นๆ ต่อไป

**2.2 เทคนิคการวิเคราะห์ข้อมูล** จากศึกษาและนำมาใช้ในงานวิจัยครั้งนี้ ได้แก่ Naïve Bayes ประมวลผลรวดเร็วแต่ต้องการข้อมูลปริมาณมาก Random Forest จัดการความซับซ้อนได้ดีแต่ใช้ทรัพยากรสูง Neural Network เหมาะกับความสัมพันธ์ที่ไม่เป็นเชิงเส้น XGBoost ประมวลผลแบบขนานได้รวดเร็วแต่มีพารามิเตอร์ต้องปรับแต่งจำนวนมาก และ SVM มีความแม่นยำสูงผ่านการค้นหาระนาบเกินแม้ข้อมูลน้อยแต่ไม่เหมาะกับชุดข้อมูลขนาดใหญ่

**2.3 การเพิ่มข้อมูลสังเคราะห์ (SMOTE)** ผู้วิจัยเลือกใช้เทคนิค SMOTE เพื่อแก้ปัญหาความไม่สมดุลของข้อมูล เนื่องจากสามารถรักษาข้อมูลสำคัญในชุดข้อมูลขนาดเล็กได้ดีกว่าการลดจำนวนข้อมูล (Liu, F., et al., 2024) [3]

**2.4 การประเมินประสิทธิภาพ** งานวิจัยนี้มีตัวแปรเป้าหมายแบบพหุคลาส การคำนวณจึงต้องพิจารณาทีละคลาสเทียบกับคลาสที่เหลือทั้งหมด โดยที่ K คือคลาสทั้งหมด และ i คือ คลาสที่กำลังพิจารณา มีสมการดังนี้

2.4.1 ค่าความแม่นยำ (Accuracy) (Alnoman A., 2024) [4]

$$\text{Accuracy} = \frac{\sum_{i=1}^{k} TP_i}{N}$$

2.4.2 ค่าความเที่ยงตรง (Precision) (Grandini et al., 2020) [5]

1.) หาความเที่ยงตรงแต่ละคลาส

$$\text{Precision} = \frac{TP_i}{TP_i+FP_i}$$

2.) หาค่าความเที่ยงตรงแบบมาโคร

$$\text{Precision}_{\text{Macro}} = \frac{1}{K}\sum_{i=1}^{K} Precision_i$$

2.4.3 ค่าระลึก (Recall) (Sarker, 2021) [6]

1.) หาค่าระลึกแต่ละคลาส

$$\text{Recall} = \frac{TP_i}{TP_i+FN_i}$$

2.) หาค่าระลึกแบบมาโคร

$$\text{Recall}_{\text{Macro}} = \frac{1}{K}\sum_{i=1}^{K} Recall_i$$

2.4.4 ค่าความสมดุลระหว่างความเที่ยงตรงและค่าระลึก (F1-Score) (Tharwat, 2021) [7]

1.) หาค่า F1-Score แต่ละคลาส

$$\text{F1}_{\text{i}} = 2 \times \frac{Precision_i \times Recall_i}{Precision_i+Recall_i}$$

2.) หาค่าความสมดุลแบบมาโคร

$$\text{F1}_{\text{Macro}} = \frac{1}{K}\sum_{i=1}^{K} F1_i$$

**2.5 การหาค่าสัมประสิทธิ์สหสัมพันธ์** เพื่อการวัดระดับความสัมพันธ์ระหว่างตัวแปร ว่าเป็นไปในทิศทางใดและมีความสัมพันธ์มากน้อยเพียงใด โดยพิจารณาตามเกณฑ์ของ John W. Best [8] ดังนี้

**ตารางที่ 1** การแปลความหมายค่าสัมประสิทธิ์สหสัมพันธ์

| ค่าสัมประสิทธิ์สหสัมพันธ์ | ระดับความสัมพันธ์ |
|---|---|
| 0.80-1.00 | สูงมาก |
| 0.60-0.79 | สูง |
| 0.40-0.59 | ปานกลาง |
| 0.20-0.39 | ต่ำ |
| 0 – 0.19 | ต่ำมากหรือไม่มี |

งานวิจัยด้านการพยากรณ์ผลสัมฤทธิ์ที่ผ่านมามักอาศัยเพียงผลการเรียนเดิม งานวิจัยนี้จึงริเริ่มบูรณาการมิติด้านบุคลิกภาพ 5 องค์ประกอบเข้ามาร่วมวิเคราะห์ เพื่อยกระดับความแม่นยำและทำให้การพยากรณ์มีความสมบูรณ์ยิ่งขึ้น

## 3. วิธีการดำเนินงานวิจัย

**3.1 การรวบรวมข้อมูล** ผู้วิจัยเก็บรวบรวมข้อมูลผลการเรียนวิชาพื้นฐาน ได้แก่ วิทยาศาสตร์ คณิตศาสตร์ และภาษาอังกฤษ ร่วมกับแบบประเมินบุคลิกภาพ 5 องค์ประกอบตามแนวคิด John, Naumann & Soto [1] ฉบับปรับปรุงปี พ.ศ. 2551 [1] ผ่าน Google Form จากนักเรียนระดับ ปวช. จำนวน 500 คน ใน 5 สาขาวิชา ได้แก่ ช่างยนต์ ช่างกลโรงงาน ช่างไฟฟ้า การบัญชี และเทคโนโลยีธุรกิจดิจิทัล โดยดำเนินการภายใต้การขอความยินยอมและปกปิดตัวตนตามหลักจริยธรรมการวิจัย

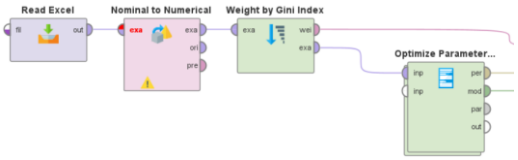
**3.2 การจัดเตรียมข้อมูล** ผู้วิจัยได้ตรวจสอบความสมบูรณ์ของชุดข้อมูล จัดรูปแบบ และกำหนดชนิดตัวแปรให้เหมาะสม พร้อมทั้งวิเคราะห์ค่าสัมประสิทธิ์สหสัมพันธ์ (Correlation Matrix) เพื่อคัดกรองตัวแปรที่มีความซ้ำซ้อนออก และดำเนินการแบ่งข้อมูลออกเป็นชุดฝึกสอนร้อยละ 70 และชุดทดสอบร้อยละ 30 และทำการปรับสมดุลข้อมูลด้วยเทคนิค SMOTE กับชุดฝึกสอนทั้งหมด 4 ครั้ง เพื่อเพิ่มจำนวนข้อมูลในกลุ่มผลการเรียนกลุ่ม A, C, D และ F ให้เท่ากับกลุ่ม B ที่มีขนาดใหญ่ที่สุด เนื่องจากกลุ่มผลการเรียน F ถูกแบ่งเข้ามาในชุดฝึกสอนเพียง 2 ตัวอย่าง ผู้วิจัยจึงได้กำหนดค่า K =1 เพื่อให้สอดคล้องกับข้อจำกัดทางคณิตศาสตร์ของอัลกอริทึมที่ค่า K ต้องน้อยกว่าจำนวนตัวอย่าง ดังตารางที่ 2

**ตารางที่ 2** การปรับสมดุลให้กับชุดข้อมูล

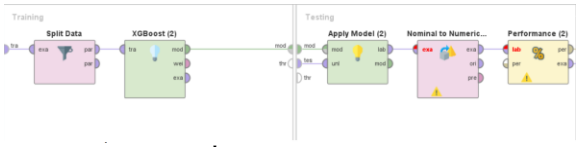
| ครั้งที่ | ค่า K | จำนวนคลาส | | | | | ผลรวม |
|---|---|---|---|---|---|---|---|
| | | A | B | C | D | F | |
| ก่อนปรับ | | 9 | 329 | 148 | 11 | 3 | 500 |
| 1 | 1 | 9 | 329 | 148 | 11 | 329 | 826 |
| 2 | 3 | 329 | 329 | 148 | 11 | 329 | 1146 |
| 3 | 5 | 329 | 329 | 148 | 329 | 329 | 1464 |
| 4 | 9 | 329 | 329 | 329 | 329 | 329 | 1645 |

อย่างไรก็ตาม ผู้วิจัยตระหนักดีว่าการสังเคราะห์ข้อมูลในกลุ่มที่มีตัวอย่างจำกัด อาจเพิ่มความเสี่ยงต่อปัญหาภาวะเรียนรู้เกิน จึงต้องอาศัยการประเมินประสิทธิภาพอย่างรัดกุมในขั้นตอนการสร้างแบบจำลองร่วมด้วย

**3.3 การสร้างแบบจำลอง** ผู้วิจัยนำชุดข้อมูลเข้าสู่โปรแกรม RapidMiner Studio โดยกำหนดให้ผลสัมฤทธิ์ทางการเรียนเป็นตัวแปรเป้าหมาย และแปลงชนิดข้อมูลให้เหมาะสมกับการประมวลผล จากนั้นได้พัฒนาแบบจำลอง 5 รูปแบบ ได้แก่ Naïve Bayes, Random Forest, Neural Network, XGBoost และ SVM




**ภาพที่ 1** ตัวอย่างการสร้างแบบจำลอง




**ภาพที่ 2** แบบจำลอง XGBoost

เพื่อให้แบบจำลองมีประสิทธิภาพสูงสุด ผู้วิจัยได้ประยุกต์ใช้เทคนิคการค้นหาแบบตาราง (Grid Search Optimization) เพื่อหาชุดค่าพารามิเตอร์ที่เหมาะสมที่สุดสำหรับแต่ละแบบจำลอง ได้ผลลัพธ์ดังตารางที่ 3

**ตารางที่ 3** ค่าพารามิเตอร์ที่ดีที่สุด แยกตามแบบจำลอง

| อัลกอริทึม | ค่าพารามิเตอร์ที่เหมาะสม |
|---|---|
| XGBoost | Max Depth = 100, Learning Rate = 0.3 |
| Random Forest | Max Depth = 39,<br>Number of Trees = 90 |
| SVM | Kernel = RBF |
| Neural Network | Learning Rate = 0.400, Cycles = 51,<br>Momentum = 0.100 |
| Naïve Bayes | Laplace Correction = False |

ถึงแม้ว่าค่าพารามิเตอร์ในตารางที่ 3 บางตัวแสดงค่าขอบเขตที่กว้างมาก เช่น แบบจำลอง XGBoost มีค่า Max Depth = 100 นั้น มีวัตถุประสงค์เพื่อให้แบบจำลองสามารถเรียนรู้ความสัมพันธ์ของชุดข้อมูลขนาดเล็กได้อย่างเต็มที่ โดยผู้วิจัยได้ควบคุมภาวะการเรียนรู้เกินและประเมินเสถียรภาพของแบบจำลองบนชุดฝึกสอนด้วยเทคนิค 10-fold Stratified Cross-Validation เพื่อรักษาสัดส่วนการกระจายตัวของคลาสในแต่ละกลุ่มย่อยให้คงที่ ส่งผลให้การประเมินเสถียรภาพมีความแม่นยำสูง ก่อนที่จะนำแบบจำลองที่ดีที่สุดไปประเมินประสิทธิภาพกับชุดทดสอบ

**3.4 การประเมินประสิทธิภาพ** ผู้วิจัยได้นำผลการพัฒนาแบบจำลองมาเปรียบเทียบประสิทธิภาพการทำงาน โดยใช้ค่าความแม่นยำ (Accuracy) ค่าความเที่ยงตรง (Precision) ค่าระลึก (Recall) และค่าเฉลี่ยฮาร์โมนิก (F1-Score) ทั้งนี้ยังเปรียบเทียบน้ำหนักความสำคัญ (Weight by Gini Index) อีกด้วย

**3.5 การนำไปใช้งาน** ผู้วิจัยได้ออกแบบหน้าต่างการใช้งานและสร้างแบบจำลองกำหนดค่าพารามิเตอร์ที่ดีที่สุดด้วยภาษาไพทอน และสร้างเป็นต้นแบบระบบพยากรณ์ผลสัมฤทธิ์ทางการเรียนระดับชั้นปวช.

## 4. ผลการวิจัย

### 4.1 การเปรียบเทียบปัจจัยที่ผลต่อการพยากรณ์

**ตารางที่ 4** ค่าสัมประสิทธิ์สหสัมพันธ์

| ตัวแปร | Subject | Math | Science | Eng | A | B | C | N | O |
|---|---|---|---|---|---|---|---|---|---|
| **Subject** | 1 | **.209** | **.318** | .140 | -.076 | **-.253** | -.053 | **.332** | .022 |
| **Math** | **.209** | 1 | **.665** | **.218** | **.256** | .050 | -.196 | **.365** | **.305** |
| **Science** | **.318** | **.665** | 1 | **.379** | .135 | .032 | -.150 | **.303** | **.301** |
| **Eng** | .140 | **.218** | **.379** | 1 | .006 | .133 | .119 | -.177 | .150 |
| **A** | -.076 | **.256** | .135 | .006 | 1 | **.480** | .074 | -.108 | **.423** |
| **C** | **-.253** | .050 | .032 | .133 | **.480** | 1 | **.208** | **-.420** | **.371** |
| **E** | -.053 | -.196 | -.150 | .119 | .074 | **.208** | 1 | **-.271** | **.242** |
| **N** | **.332** | **.365** | **.303** | -.177 | -.108 | **-.420** | **-.271** | 1 | .033 |
| **O** | .022 | **.305** | **.301** | .150 | **.423** | **.371** | **.242** | .033 | 1 |

จากตารางที่ 4 แสดงความสัมพันธ์ของตัวแปร ตามเกณฑ์ของ John W. Best ดังนี้

ความสัมพันธ์ระดับสูง พบว่า เกรดวิชาคณิตศาสตร์และวิทยาศาสตร์ (r = 0.665) ซึ่งสะท้อนให้เห็นถึงความสอดคล้องของทักษะและสมรรถนะการเรียนรู้ในกลุ่มวิชาการคำนวณและเหตุผล

ความสัมพันธ์ระดับปานกลาง พบในกลุ่มตัวแปรบุคลิกภาพ ได้แก่ ด้านความเป็นมิตร (A) กับด้านการมีจิตสำนึก (C) (r = 0.480) และด้านความเป็นมิตร (A) กับการเปิดรับประสบการณ์ (O) (r = 0.423)

ความสัมพันธ์ระดับต่ำและเชิงลบที่น่าสนใจ ได้แก่ ด้านการมีจิตสำนึก (C) มีความสัมพันธ์เชิงลบกับความไม่มั่นคงทางอารมณ์ (N) สูงที่สุด (r = -0.420) สะท้อนว่าผู้ที่มีความวิตกกังวลสูงมักมีปัญหาในการควบคุมและจัดการตนเอง ทั้งนี้ยังไม่พบตัวแปรอิสระคู่ใดที่มีค่าความสัมพันธ์สูงเกินกว่าเกณฑ์ 0.80 จึงสามารถสรุปได้ว่าชุดข้อมูลที่ใช้ในการศึกษานี้ ไม่ปรากฏปัญหาภาวะร่วมเส้นตรงพหุคูณและมีความเหมาะสมในการนำไปสร้างแบบจำลองต่อไป

**ตารางที่ 5** น้ำหนักความสำคัญของปัจจัย จาก XGBoost

| ปัจจัย | น้ำหนัก |
|---|---|
| เกรดวิชาคณิตศาสตร์ | 0.830 |
| เกรดวิชาภาษาอังกฤษ | 0.111 |
| เกรดวิชาวิทยาศาสตร์ | 0.095 |
| ด้านความไม่มั่นคงทางอารมณ์ (N) | 0.077 |
| สาขาวิชาที่เลือกเรียน | 0.062 |
| ด้านการเปิดตัว (E) | 0.038 |
| ด้านการเปิดรับประสบการณ์ (O) | 0.017 |
| ด้านการมีจิตสำนึก (C) | 0.016 |
| ด้านความเป็นมิตร (A) | 0.014 |

จากตารางที่ 5 พบว่าเกรดวิชาคณิตศาสตร์มีค่า Gini Index สูงสุดที่ 0.830 แสดงว่าเป็นตัวพยากรณ์ที่สำคัญที่สุด รองลงมาคือเกรดภาษาอังกฤษ (0.111) วิทยาศาสตร์ (0.095) และ Neuroticism (0.077) ในบรรดาตัวแปรด้านบุคลิกภาพ สะท้อนให้เห็นว่าความวิตกกังวลและความไม่มั่นคงทางอารมณ์มีผลต่อผลสัมฤทธิ์ในการเรียนสายอาชีพที่ต้องปฏิบัติงานจริงอย่างมีนัยสำคัญ

### 4.2 การเปรียบเทียบประสิทธิภาพอัลกอริทึม

ผลการเปรียบเทียบประสิทธิภาพทั้ง 5 แบบจำลองด้วย Accuracy, Precision, Recall และ F1-Score บนชุดข้อมูลทดสอบ 30% หลังจากปรับสมดุลด้วย SMOTE แล้ว

**ตารางที่ 6** ผลการเปรียบเทียบประสิทธิภาพอัลกอริทึม

| อัลกอริทึม | Accuracy | Precision | Recall | F1-Score |
|---|---|---|---|---|
| XGBoost | 85.47 | 85.51 | 85.46 | 85.48 |
| Random Forest | 85.05 | 85.05 | 85.06 | 85.05 |
| SVM | 78.12 | 78.80 | 78.12 | 78.46 |
| Naïve Bayes | 69.30 | 67.63 | 69.30 | 68.45 |
| Neural Network | 65.96 | 62.15 | 65.97 | 63.99 |

จากตารางที่ 6 พบว่า XGBoost ให้ประสิทธิภาพสูงสุด โดยมี Accuracy 85.47% ซึ่งเกินเกณฑ์ที่กำหนด (>75%) รองลงมาเป็น Random Forest (85.05%) และ SVM ( 78.12%) ตามลำดับ ขณะที่ Naïve Bayes และ Neural Networkให้ความแม่นยำต่ำกว่าเกณฑ์ที่ 69.30% และ 65.96% ตามลำดับ

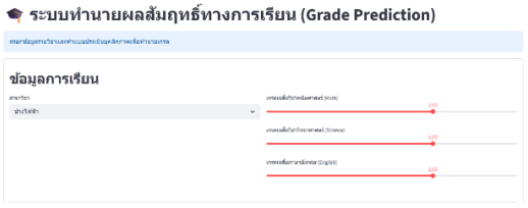
**ตารางที่ 7** การวิเคราะห์ความคลาดเคลื่อนของ XGBoost

| | True A | True B | True C | True D | True F | Precision |
|---|---|---|---|---|---|---|
| **Pred. A** | 305 | 21 | 7 | 0 | 0 | 91.59 |
| **Pred. B** | 12 | 231 | 62 | 8 | 1 | 73.57 |
| **Pred. C** | 6 | 57 | 230 | 16 | 0 | 74.43 |
| **Pred. D** | 6 | 16 | 27 | 305 | 0 | 86.16 |
| **Pred. F** | 0 | 4 | 3 | 0 | 328 | 97.91 |
| **Recall** | 97.71 | 70.21 | 69.91 | 92.71 | 99.70 | |

จากตารางที่ 7 เมื่อพิจารณาประสิทธิภาพรายคลาส พบว่าแบบจำลองสามารถทำนายคลาส F และ A ได้แม่นยำที่สุด เนื่องจากลักษณะทางพฤติกรรมและผลการเรียนของกลุ่มนี้มีความชัดเจนและแตกต่างจากกลุ่มอื่นอย่างมีนัยสำคัญ อย่างไรก็ตามคลาส B และ C ถือเป็นกลุ่มที่พยากรณ์ได้ยากที่สุด ซึ่งสะท้อนว่าผู้เรียนในสองกลุ่มนี้มีลักษณะบุคลิกภาพและผลการเรียนที่ใกล้เคียงกันจนแบบจำลองแยกแยะได้ยาก

### 4.4 ต้นแบบระบบทำนายผลสัมฤทธิ์ทางการเรียน

ผู้วิจัยได้นำแบบจำลอง XGBoost ที่ผ่านการฝึกฝนไปพัฒนาเป็นต้นแบบระบบทำนายผลสัมฤทธิ์ทางการเรียน โดยเมื่อผู้ใช้งานกรอกข้อมูลผลการเรียนและทำแบบประเมินบุคลิกภาพดังภาพที่ 3 และกดปุ่มพยากรณ์ระบบจะประมวลผลและแสดงผลสัมฤทธิ์ทางการเรียน ดังภาพที่ 4




**ภาพที่ 3** หน้าจอระบบพยากรณ์ผลสัมฤทธิ์




**ภาพที่ 4** ตัวอย่างผลการพยากรณ์ 5 ระดับ

## 5. สรุปผล อภิปรายผล และข้อเสนอแนะ

### 5.1 สรุปผลการวิจัย

จากการวิเคราะห์พบว่า XGBoostมีประสิทธิภาพสูงสุด ด้วยค่าความแม่นยำร้อยละ 85.47 โดยมีปัจจัยที่มีอิทธิพลต่อการพยากรณ์สูงสุดคือ เกรดวิชาคณิตศาสตร์ (0.83) และในมิติบุคลิกภาพพบเพียง ความไม่มั่นคงทางอารมณ์ (0.077) เท่านั้นที่มีนัยสำคัญต่อการทำนายผลสัมฤทธิ์ทางการเรียน

### 5.2 อภิปรายผลการวิจัย

5.2.1 ประสิทธิภาพอัลกอริทึมกลุ่ม แสดงให้เห็นว่าแบบจำลอง XGBoostและ Random Forest มีประสิทธิภาพในการพยากรณ์สูงกว่าแบบจำลองเดี่ยว เนื่องจากกระบวนการรวมผลลัพธ์ช่วยลดความแปรปรวนและป้องกันภาวะเรียนรู้เกิน (Overfitting) ได้ดีกว่า สอดคล้องกับงานวิจัยของ พิริยพงศ์ หริ่งรอด (2566) [9] และ Soheila Farokhi et al.(2566) [10] ทั้งนี้ XGBoost ให้ความแม่นยำสูงกว่า Random Forest เล็กน้อย เนื่องจากใช้กลไก Boosting ที่เรียนรู้และแก้ไขข้อผิดพลาดของแบบจำลองก่อนหน้าอย่างเป็นลำดับขั้น ผสานกับกลไก Regularization ที่จัดการกับชุดข้อมูลขนาดเล็กได้ละเอียดกว่าเทคนิค Bagging ของ Random Forest ในขณะที่ Neural Network และ Naïve bayes มีประสิทธิภาพต่ำกว่าสมมติฐาน เนื่องจากมีข้อจำกัดด้านปริมาณข้อมูลที่มีเพียง 500 ชุด ซึ่งไม่เพียงพอต่อพยากรณ์

5.2.2 ประสิทธิภาพของการจำแนกผลสัมฤทธิ์ทางการเรียนนั้นพบว่า อัลกอริทึมสามารถทำนายกลุ่ม “ไม่ผ่านเกณฑ์” ได้แม่นยำที่สุด เนื่องจากลักษณะข้อมูลของกลุ่มนี้มีความแตกต่างจากกลุ่มอื่นอย่างชัดเจน ทำให้แบบจำลองสามารถตัดสินใจได้ง่าย ในขณะที่กลุ่ม ดี และ พอใจ มีความแม่นยำลดลง เป็นผลมาจากชุดข้อมูลมีความใกล้เคียงกันทำให้แบบจำลองยากที่จะตัดสินใจ สอดคล้องกับงานวิจัยของ Sarker (2021) [6] ที่ได้กล่าวว่ามีข้อจำกัดในการจำแนกหากเป็นแบบพหุคลาสและมีความคลุมเครือระหว่างคลาส

5.2.3 ปัจจัยที่มีอิทธิพลต่อการพยากรณ์ พบว่า เกรดคณิตศาสตร์สอดคล้องกับ ทฤษฎีความสามารถทางปัญญา (Cognitive Ability Theory) ที่มองว่าทักษะตรรกะเป็นสมรรถนะพื้นฐานของสายอาชีพ [11] สำหรับความไม่มั่นคงทางอารมณ์ที่โดดเด่นเพียงมิติเดียว เนื่องจากการเรียนอาชีวศึกษาเน้นการปฏิบัติภายใต้ความกดดันสูง ผู้เรียนที่ควบคุมอารมณ์ได้ยากจึงได้รับผลกระทบต่อประสิทธิภาพการเรียนโดยตรงซึ่งสอดคล้องกับแนวคิดของ John, Naumann และ Soto (2008) [1] ที่ระบุว่าความไม่มั่นคงทางอารมณ์เป็นปัจจัยสำคัญในการปรับตัวเข้ากับสภาพแวดล้อม

5.2.4 ความสัมพันธ์ระหว่างตัวแปร พบว่า ความสัมพันธ์ระดับสูงระหว่างวิชาคณิตศาสตร์และวิทยาศาสตร์สะท้อนถึงความสำคัญในงานอุตสาหกรรม นอกจากนี้ความสัมพันธ์เชิงลบระหว่างความไม่มั่นคงทางอารมณ์กับผลสัมฤทธิ์ทางการเรียน ยืนยันว่าความวิตกกังวลเป็นอุปสรรคต่อการควบคุมตนเองและการพัฒนาตนตามทฤษฎีบุคลิกภาพ 5 องค์ประกอบ

ผลการทดสอบสมมติฐานพบว่า XGBoost, Random Forest และ SVM มีความแม่นยำผ่านเกณฑ์ร้อยละ 75 ในขณะที่ Naïve Bayes และ Neural Network ไม่ผ่านเกณฑ์เนื่องจากข้อจำกัดด้านปริมาณข้อมูล จุดเด่นของงานวิจัยนี้เมื่อเทียบกับอดีต คือการบูรณาการมิติด้านบุคลิกภาพร่วมกับผลการเรียนเดิม ซึ่งช่วยดักจับความเสี่ยงของผู้เรียนสายอาชีพที่ต้องเผชิญความกดดันได้ดีขึ้น ส่งผลให้ XGBoost มีความแม่นยำสูงถึงร้อยละ 85.47 ช่วยยกระดับระบบสารสนเทศเพื่อการแนะแนวให้รอบด้านและสะท้อนศักยภาพผู้เรียนได้อย่างแท้จริง

**5.3 ข้อจำกัดและข้อเสนอแนะ**

5.3.1 การวิจัยใช้ข้อมูลจากกลุ่มตัวอย่างเพียง 500 ชุด จากสถานศึกษาแห่งเดียว โดยไม่มีการทดสอบความแม่นยำกับข้อมูลสถานศึกษาอื่น

5.3.2 การเก็บข้อมูลบุคลิกภาพใช้มาตรวัดแบบรายงานตนเอง ซึ่งอาจมีความคลาดเคลื่อนจากความเป็นจริง

5.3.3 การใช้เทคนิค SMOTE สังเคราะห์ข้อมูลในกลุ่มที่มีตัวอย่างน้อยมาก อาจเพิ่มความเสี่ยงต่อปัญหาภาวะเรียนรู้เกิน

จากผลการศึกษาผู้วิจัยมีข้อเสนอแนะว่า สถานศึกษาควรใช้ XGBoost หรือ Random Forest พัฒนาระบบแนะแนว และเฝ้าระวังผู้เรียนที่เกรดคณิตศาสตร์ต่ำหรือมีความไม่มั่นคงทางอารมณ์สูง การวิจัยครั้งหน้าควรขยายกลุ่มตัวอย่างไปยังสถาบันอื่น เพิ่มตัวแปรสภาพแวดล้อม และประยุกต์ใช้อัลกอริทึมที่ทันสมัยยิ่งขึ้นเพื่อยกระดับประสิทธิภาพ

## เอกสารอ้างอิง


[1] John, O. P., Naumann, L. P., & Soto, C. J. (2008). Paradigm shift to the integrative Big Five trait taxonomy: History, measurement, and conceptual issues. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), Handbook of personality: Theory and research (3rd ed., pp. 114–158)

[2] พรสรัญ รุ่งเจริญกิจกุล. (2565). บุคลิกภาพที่ประสบความสำเร็จ พัฒนาได้อย่างไร. การเงินธนาคาร. https://moneyandbanking.co.th/2022/15543/

[3] Liu, F., et al. (2024). Addressing Class Imbalance of Health Data: A Systematic Literature Review on Modified Synthetic Minority Oversampling Technique (SMOTE) Strategies. Journal on Innovation and Visualization, 2(3), 2283.

[4] Alnoman, A. (2024). Foundational Subject Grades as Predictors of Cumulative GPA in Vocational Education. International Journal of Educational Technology, 12(3), 88–102.

[5] Grandini, M., Bagli, E., & Visani, G. (2020). Metrics for multi-class classification: An overview. arXiv. https://arxiv.org/abs/2008.05756

[6] Sarker, I. H. (2021). Machine learning: Algorithms, real-world applications and research directions. SN Computer Science, 2(3), Article 160.

[7] Tharwat, A. (2021). Classification assessment metric roles in machine learning applications. AI Communications, 34(1), 27-55.

[8] Best, J. W., & Kahn, J. V. (2006). Research in education (10th ed.). Pearson Education.

[9] พิริยพงศ์ หริ่งรอด. (2566). การพยากรณ์ผลสัมฤทธิ์ทางการเรียนด้วย Random Forest และ XGBoost. วารสารวิทยาการข้อมูลและระบบสารสนเทศ, 10(2), 45–58

[10] Farokhi, S., Yaramala, A., Huang, J., Khan, M. F. A., Qi, X., & Karimi, H. (2023). Enhancing the performance of automated grade prediction in MOOC using graph representation learning. 2023 IEEE 10th International Conference on Data Science and Advanced Analytics (DSAA), 1-10. https://doi.org /10.1109/dsaa60987.2023.10302642

[11] Roth, B., Becker, N., Romeyke, S., Schäfer, S., Domnick, F., & Spinath, F. M. (2015). Intelligence and school grades: A meta-analysis. Intelligence, 53, 118-137.

# การวิเคราะห์เชิงเปรียบเทียบเชิงลึกด้านความเหลื่อมล้ำในการออมของครัวเรือนไทย กรณีศึกษาตามภาคและกลุ่มรายได้โดยใช้ข้อมูล NSO นำเสนอผลด้วย Power BI

# In-Depth Comparative Analysis of Household Savings Inequality in Thailand by Region and Income Group: A Power BI-Based Study Using NSO Data

*อภิญญา ทองคำ (Apinya Thongkum)[1] และณัฐวี อุตกฤษฎ์ (Nattavee Utakrit)[2]*
*ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล*
*มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ*
*[1]S6707021857136@kmutnb.ac.th, [2]nattavee.u@itd.kmutnb.ac.th*

## บทคัดย่อ

*ปัจจุบันครัวเรือนไทยเผชิญความเปราะบางทางการเงินจากภาระหนี้สินที่สูงและระดับการออมที่ต่ำ ท่ามกลางความผันผวนทางเศรษฐกิจและโครงสร้างสังคมสูงวัย การศึกษานี้มีวัตถุประสงค์เพื่อวิเคราะห์ระดับและรูปแบบความเหลื่อมล้ำด้านการออมของครัวเรือนไทยตามภูมิภาคและกลุ่มรายได้ โดยใช้ข้อมูลทุติยภูมิจากการสำรวจภาวะเศรษฐกิจและสังคมของครัวเรือน (SES) ของสำนักงานสถิติแห่งชาติ ระหว่างปี พ.ศ. 2556–2566ซึ่งครอบคลุมมิติด้านรายได้ ค่าใช้จ่าย การออม และการเข้าถึงบริการทางการเงิน วิธีดำเนินการวิจัยเริ่มจากการทำความสะอาดข้อมูล (Data Cleaning) และคำนวณตัวชี้วัดสำคัญ ได้แก่ รายได้สุทธิ ค่าใช้จ่าย เงินออมเฉลี่ย และอัตราการออม พร้อมทั้งวัดความเหลื่อมล้ำด้วยดัชนีจีนี (Gini Coefficient) และนำเสนอผลลัพธ์ผ่านระบบสารสนเทศทางธุรกิจ (Business Intelligence) ด้วยโปรแกรม Power BI ผลการวิจัยพบว่า ความเหลื่อมล้ำด้านการออมของครัวเรือนไทยอยู่ในระดับปานกลาง โดยมีค่าดัชนีจีนีเฉลี่ยประมาณ 0.24 และมีความแตกต่างอย่างมีนัยสำคัญตามภูมิภาคและกลุ่มรายได้ โดยครัวเรือนรายได้สูงมีศักยภาพในการออมสูงกว่ากลุ่มรายได้ต่ำ การประยุกต์ใช้ Power BI ในการสร้างแดชบอร์ดช่วยให้เห็นรูปแบบความเหลื่อมล้ำเชิงพื้นที่ได้อย่างชัดเจน และช่วยในการตัดสินใจเชิงนโยบายเพื่อลดความเหลื่อมล้ำทางการเงินได้อย่างเป็นรูปธรรม*

**คำสำคัญ:** ความเหลื่อมล้ำด้านการออม, ครัวเรือนไทย, ดัชนีจีนี, Power BI, สำนักงานสถิติแห่งชาติ

## Abstract

*Thai households are currently facing financial vulnerability due to high debt burdens and low savings levels amid economic fluctuations and an aging society. This study aims to analyze the level and patterns of savings inequality among Thai households across regions and income groups using secondary data from the Household Socio-Economic Survey (SES) conducted by the National Statistical Office during 2013–2023, covering dimensions of income, expenditure, savings, and access to financial services. The research methodology includes data cleaning and the calculation of key indicators such as net income, expenditure, average savings, and savings rate, with inequality measured using the Gini Coefficient. The results reveal that savings inequality among Thai households is at a moderate level (average Gini coefficient of approximately 0.24), with significant variations across regions and income groups. Higher-*

*income households demonstrate greater saving capacity than lower-income groups. The application of Power BI dashboards enables clear visualization of spatial patterns of inequality and supports evidence-based policy decision-making aimed at reducing financial disparities.*

**Keywords:** Savings Inequality, Thai Households, Gini Coefficient, Power BI, NSOconsidering publishing in the NCCIT proceeding. The paper describes the format, the sizes, and font types used in each section.

## 1. บทนำ

เศรษฐกิจไทยในช่วงทศวรรษที่ผ่านมาเผชิญกับความผันผวนจากปัจจัยภายนอกและโครงสร้างประชากรที่เข้าสู่สังคมสูงวัย ส่งผลกระทบโดยตรงต่อพฤติกรรมการออมของครัวเรือน ข้อมูลจากสำนักงานสถิติแห่งชาติปี 2566 ระบุว่าครัวเรือนไทยมีเงินออมเฉลี่ยเพียงร้อยละ 10 ของรายได้ และมีสัดส่วนหนี้สินต่อรายได้สูงถึงร้อยละ 90 สะท้อนถึงความเปราะบางทางการเงินและความเหลื่อมล้ำในการเข้าถึงทรัพยากรระหว่างกลุ่มรายได้และภูมิภาคที่ยังคงสูงอยู่ (สำนักงานสถิติแห่งชาติ [NSO], 2566)

แม้จะมีการศึกษาด้านรายได้และการบริโภค แต่การวิเคราะห์เชิงลึกด้าน "การออม" ที่เชื่อมโยงมิติเชิงพื้นที่และเวลาด้วยเครื่องมือวิเคราะห์สมัยใหม่ยังมีจำกัด ผู้วิจัยจึงประยุกต์ใช้ Power BI เพื่อสร้าง Dashboard ในการวิเคราะห์และนำเสนอข้อมูลให้เห็นภาพรวมความเหลื่อมล้ำที่ชัดเจน เพื่อสนับสนุนการตัดสินใจเชิงนโยบายอย่างมีประสิทธิภาพ

### 1.1 วัตถุประสงค์การวิจัย

1.1 เพื่อวิเคราะห์เปรียบเทียบระดับความเหลื่อมล้ำในการออมของครัวเรือนไทยตามภูมิภาคและกลุ่มรายได้ โดยใช้ข้อมูลจากสำนักงานสถิติแห่งชาติ (NSO)

1.2 เพื่อพัฒนา Dashboard ด้วย Power BI ในการสื่อสารข้อมูลเชิงลึกเกี่ยวกับความเหลื่อมล้ำด้านการออม

### 1.2 สมมุติฐานการวิจัย

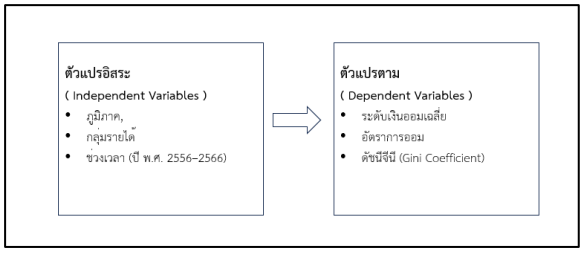
1.3.1 สมมติฐานที่ 1 (H1) ครัวเรือนในภูมิภาคที่แตกต่างกัน (กรุงเทพฯ ภาคกลาง ภาคเหนือ ภาคตะวันออกเฉียงเหนือ และภาคใต้) มีระดับเงินออมเฉลี่ยและอัตราการออมที่แตกต่างกันอย่างมีนัยสำคัญ

1.3.2 สมมติฐานที่ 2 (H2) กลุ่มรายได้ของครัวเรือนมีความสัมพันธ์เชิงบวกกับความสามารถในการออม โดยครัวเรือนกลุ่มรายได้สูงจะมีอัตราการออมที่สูงกว่าครัวเรือนกลุ่มรายได้ต่ำอย่างชัดเจน

1.3.3 สมมติฐานที่ 3 (H3) ความเหลื่อมล้ำด้านการออมของครัวเรือนไทยมีแนวโน้มเพิ่มสูงขึ้นในช่วงที่เกิดภาวะวิกฤตเศรษฐกิจ เมื่อเทียบกับช่วงภาวะเศรษฐกิจปกติ

1.3.4 สมมติฐานที่ 4 (H4) การนำเสนอข้อมูลผ่าน Dashboard ด้วย Power BI ช่วยให้ผู้ใช้งานสามารถระบุรูปแบบความเหลื่อมล้ำและเข้าใจข้อมูลเชิงลึกได้ดีกว่าการนำเสนอข้อมูลในรูปแบบตารางปกติ

### 1.3 กรอบแนวคิดการวิจัย




ภาพ 1 กรอบแนวคิดการวิจัยแสดงความสัมพันธ์ระหว่างตัวแปรอิสระ ได้แก่ ภูมิภาค กลุ่มรายได้ และช่วงเวลา ซึ่งส่งผลต่อตัวแปรตาม ได้แก่ เงินออมเฉลี่ย อัตราการออม และดัชนีจีนี เพื่อใช้ในการวิเคราะห์ความเหลื่อมล้ำด้านการออมของครัวเรือนไทย และนำเสนอผลผ่าน Power BI Dashboard

### 1.4 ขอบเขตการวิจัย

1.4.1 ขอบเขตด้านข้อมูล การศึกษานี้ใช้ข้อมูลทุติยภูมิจากการสำรวจภาวะเศรษฐกิจและสังคมของครัวเรือน (SES) ของสำนักงานสถิติแห่งชาติ ครอบคลุมปี พ.ศ. 2556–2566 โดยมีหน่วยวิเคราะห์เป็น “ครัวเรือน (Household)” มุ่งวิเคราะห์ความเหลื่อมล้ำในมิติรายได้ ค่าใช้จ่าย การออม และการเข้าถึงบริการทางการเงิน โดยใช้ตัวชี้วัดหลัก ได้แก่ เงินออมเฉลี่ย อัตราการออม และ

ดัชนีจีนี (Gini Coefficient) พร้อมจำแนกข้อมูลตามจังหวัด ภูมิภาค และกลุ่มรายได้ เพื่อวิเคราะห์เชิงพื้นที่และเวลา

1.4.2 ขอบเขตด้านการพัฒนาระบบ ผลการวิเคราะห์นำเสนอผ่านแดชบอร์ด Power BI ในรูปแบบ Interactive ที่รองรับการ Drill-down และการแสดงผลด้วยแผนที่ กราฟเส้น และกราฟแท่ง นอกจากนี้ มีการประเมินประสิทธิภาพของระบบโดยเจ้าหน้าที่ธนาคาร 30 คน และผู้ทรงคุณวุฒิ ครอบคลุมด้านความชัดเจน การใช้งาน และการนำไปใช้ประโยชน์

**1.5 ประโยชน์ที่คาดว่าจะได้รับ**

1.5.1 ได้ข้อมูลเชิงลึกเกี่ยวกับระดับความเหลื่อมล้ำในการออมของครัวเรือนไทยจำแนกตามภูมิภาคและกลุ่มรายได้ ซึ่งสามารถนำไปใช้เป็นฐานข้อมูลในการกำหนดนโยบายลดความเหลื่อมล้ำทางเศรษฐกิจได้อย่างตรงจุด

1.5.2 ได้รูปแบบการนำเสนอข้อมูลผ่าน Dashboard บน Power BI ซึ่งสามารถสื่อสารข้อมูลเชิงซ้อนให้เข้าใจง่าย ช่วยสนับสนุนการตัดสินใจของผู้บริหาร หน่วยงานรัฐ หรือผู้กำหนดนโยบายในเชิงกลยุทธ์ได้อย่างมีประสิทธิภาพ

## 2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

**2.1 แนวคิดเกี่ยวกับรายได้ครัวเรือน**

2.1.1 ความหมายของรายได้ (Income)

รายได้คือผลตอบแทนทางเศรษฐกิจทั้งในรูปเงินสดและสินทรัพย์ ซึ่งเป็นตัวกำหนดระดับการบริโภค ความมั่นคง และความสามารถในการรับมือกับความเสี่ยงของครัวเรือน (Case et al., 2020; Mankiw, 2019) โดยตามทฤษฎีของ Keynes (1936) ระบุว่าปริมาณการออมจะแปรผันตรงตามระดับรายได้ที่ได้รับในปัจจุบัน ตามสมการพื้นฐานคือ การออมเท่ากับรายได้หักด้วยการบริโภค ($S = Y - C$)

2.1.1 ความหมายของรายได้ (Income)

รายได้ของครัวเรือนประกอบด้วยรายได้จากการทำงาน ทรัพย์สิน และการโอน ซึ่งเป็นตัวบ่งชี้ศักยภาพในการบริโภคและการออมตามมาตรฐานระบบบัญชีประชาชาติ (SNA 2008) โดยโครงสร้างรายได้ที่สมบูรณ์เป็นพื้นฐานสำคัญในการวิเคราะห์พฤติกรรมทางเศรษฐกิจและความเหลื่อมล้ำของครัวเรือนอย่างเป็นระบบตามมาตรฐานสากล

**2.2 แนวคิดเกี่ยวกับค่าใช้จ่ายของครัวเรือน**

ค่าใช้จ่ายของครัวเรือนประกอบด้วยรายจ่ายเพื่อการบริโภคสินค้าบริการ และรายจ่ายที่ไม่ใช่เพื่อการบริโภค เช่น การชำระหนี้และภาษี ตามมาตรฐาน SNA 2008 ของสำนักงานสถิติแห่งชาติ โดยผลรวมของค่าใช้จ่ายทั้งสองประเภทใช้เป็นดัชนีสะท้อนพฤติกรรมการจัดการรายจ่ายและความสามารถทางการเงินของครัวเรือนในภาพรวมได้

**2.3 แนวคิดเกี่ยวกับการออม**

เงินออมคือรายได้ส่วนที่เหลือจากการบริโภคเพื่อเก็บไว้ใช้ในอนาคตและรองรับเหตุการณ์ไม่คาดคิด ซึ่งสะท้อนถึงความมั่นคงและวินัยทางการเงินของครัวเรือน โดยอัตราการออม (Saving Rate) เป็นตัวชี้วัดสำคัญที่แสดงถึงสัดส่วนรายได้ที่ถูกเก็บสะสมเพื่อสร้างเสถียรภาพทางเศรษฐกิจในระยะยาว

**2.4 ทฤษฎีที่เกี่ยวข้องกับการออม**

ทฤษฎีสมมติฐานรายได้สมบูรณ์ (Keynes): การออมแปรผันตามรายได้ปัจจุบัน หากรายได้สูงขึ้น สัดส่วนการออมจะเพิ่มขึ้นในอัตราที่มากกว่ารายได้

ทฤษฎีวัฏจักรชีวิต (Modigliani): บุคคลออมเงินเพื่อรักษาระดับการบริโภคให้คงที่ตลอดช่วงชีวิต โดยออมมากในวัยทำงานเพื่อไว้ใช้ในวัยเกษียณ

ทฤษฎีรายได้ถาวร (Friedman): พฤติกรรมการออมขึ้นอยู่กับรายได้ระยะยาวที่คาดว่าจะได้รับ (Permanent Income) มากกว่ารายได้ที่ผันผวนชั่วคราว

ทฤษฎีการออมเพื่อเหตุไม่คาดคิด (Carroll): การออมทำหน้าที่เป็นกลไกป้องกันความเสี่ยงจากความไม่แน่นอนในอนาคต เช่น สุขภาพหรือรายได้ที่ผันผวน

ทฤษฎีเศรษฐศาสตร์พฤติกรรม (Thaler): อคติทางจิตวิทยา เช่น การรักความสบายในปัจจุบัน (Present Bias) มีผลต่อการออม ซึ่งแก้ไขได้ด้วยการออกแบบทางเลือกเชิงนโยบาย

แนวคิดการเข้าถึงและความรู้ทางการเงิน (OECD): ความรู้ทางการเงินและการเข้าถึงบริการทางการเงินที่สะดวก ช่วยลดความเหลื่อมล้ำและเพิ่มประสิทธิภาพการออม

ทฤษฎีทุนมนุษย์และตลาดแรงงาน (Becker & Mincer): การศึกษาและทักษะช่วยเพิ่มรายได้ในระยะยาว ซึ่งส่งผลบวกโดยตรงต่อความสามารถในการออมอย่างมั่นคง

**2.5 ทฤษฎีวัฏจักรชีวิตในบริบทสังคมสูงวัย**

ทฤษฎีวัฏจักรชีวิต (Life-Cycle Hypothesis) ของ Modigliani อธิบายว่าบุคคลมีแนวโน้มออมเงินในช่วงวัยทำงาน และใช้เงินออมในช่วงวัยเกษียณ ดังนั้น โครงสร้างประชากรที่เข้าสู่สังคมสูงวัยจึงส่งผลโดยตรงต่อพฤติกรรมการออมในระดับครัวเรือน ในบริบทของประเทศไทยที่กำลังก้าวเข้าสู่สังคมสูงวัย สัดส่วนประชากรวัยทำงานที่ลดลงและผู้สูงอายุที่เพิ่มขึ้น อาจส่งผลให้ระดับการออมโดยรวมลดลง และเพิ่มความเหลื่อมล้ำระหว่างกลุ่มครัวเรือนที่มีศักยภาพในการออมแตกต่างกัน

**2.6 ความเหลื่อมล้ำเชิงพื้นที่ (Spatial Inequality)**

ความเหลื่อมล้ำเชิงพื้นที่ (Spatial Inequality) เป็นแนวคิดที่อธิบายความแตกต่างของการกระจายรายได้และความมั่งคั่งระหว่างพื้นที่ทางภูมิศาสตร์ ซึ่งเกิดจากความไม่เท่าเทียมในการเข้าถึงทรัพยากร โครงสร้างพื้นฐาน และโอกาสทางเศรษฐกิจ ตามแนวคิดภูมิศาสตร์เศรษฐกิจ (Economic Geography) โดยเฉพาะทฤษฎี Center–Periphery ระบุว่า พื้นที่ศูนย์กลาง (Center) เช่น เขตเมืองหรือกรุงเทพมหานคร มักมีความได้เปรียบด้านโครงสร้างเศรษฐกิจและการจ้างงานมากกว่าพื้นที่รอบนอก (Periphery) ซึ่งส่งผลให้เกิดความแตกต่างด้านรายได้และศักยภาพในการออมของครัวเรือน แนวคิดดังกล่าวสามารถนำมาอธิบายความเหลื่อมล้ำด้านการออมของครัวเรือนไทยที่มีความแตกต่างกันตามภูมิภาคได้อย่างชัดเจน

**2.7 ตัวชี้วัดความเหลื่อมล้ำ**

ดัชนีจีนี (Gini Coefficient) ซึ่งเป็นเครื่องมือที่ใช้วัดระดับความไม่เท่าเทียมของการกระจายรายได้หรือการออมในประชากร โดยค่าดัชนีมีช่วงตั้งแต่ 0 ถึง 1 ซึ่งค่าที่ใกล้ 0 แสดงถึงความเท่าเทียมในการกระจายทรัพยากร ในขณะที่ค่าที่ใกล้ 1 แสดงถึงความเหลื่อมล้ำที่สูงขึ้น ดัชนีดังกล่าวสามารถประยุกต์ใช้ในการวิเคราะห์การกระจายตัวของเงินออมหรืออัตราการออมของครัวเรือนได้ เพื่อสะท้อนภาพรวมของความเหลื่อมล้ำด้านการออมในระดับประชากร

$$G = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n}|x_i - x_j|}{2n^2\mu}$$

โดยที่

$x_i$ = ค่าเงินออมของครัวเรือนที่ i

$n$ = จำนวนครัวเรือนทั้งหมด

$u$ = ค่าเฉลี่ยของเงินออม

**2.8 แนวคิดด้าน Data Analytics / BI / Dashboard**

2.8.1 Power Business Intelligence (Power BI) ซึ่งเป็นเครื่องมือวิเคราะห์ข้อมูลที่พัฒนาโดย Microsoft สำหรับการรวบรวม แปลง และวิเคราะห์ข้อมูลจากแหล่งข้อมูลหลากหลาย เช่น ฐานข้อมูล SQL ไฟล์ Excel และบริการบนคลาวด์ Power BI ช่วยให้ผู้ใช้สามารถสร้างรายงานเชิงโต้ตอบและแดชบอร์ดที่ช่วยนำเสนอข้อมูลเชิงลึก (Insights) ได้อย่างมีประสิทธิภาพ เครื่องมือนี้ยังมีความสามารถในการสร้างโมเดลข้อมูล การใช้ภาษา DAX สำหรับการคำนวณ และการใช้ Power Query ในกระบวนการแปลงข้อมูล ซึ่งช่วยเพิ่มความแม่นยำและความยืดหยุ่นในการวิเคราะห์ข้อมูล

2.8.2 การสร้างจินตทัศน์ข้อมูล (Data Visualization) ซึ่งเป็นกระบวนการแปลงข้อมูลดิบให้อยู่ในรูปแบบภาพ เช่น กราฟ แผนภูมิ หรือแผนภาพ เพื่อช่วยให้ผู้ใช้สามารถเข้าใจข้อมูลได้ง่ายขึ้น การนำเสนอข้อมูลด้วยภาพช่วยให้สามารถมองเห็นแนวโน้ม ความสัมพันธ์ของตัวแปร และความผิดปกติของข้อมูลได้อย่างชัดเจน จึงเป็นเครื่องมือสำคัญในการสื่อสารข้อมูลและสนับสนุนการตัดสินใจในองค์กร

2.8.3 Dashboard Capability หรือความสามารถของแดชบอร์ดในการแสดงผลข้อมูลสำคัญในรูปแบบที่เข้าใจง่าย โดยแดชบอร์ดที่มีประสิทธิภาพควรประกอบด้วยการจัดวางข้อมูลที่เหมาะสม การออกแบบกราฟที่สื่อความหมายชัดเจน การกำหนดตัวชี้วัดสำคัญ (Key Performance Indicators: KPI) รวมถึงการรองรับการวิเคราะห์เชิงลึกผ่านการ Drill-down และการโต้ตอบกับ

ข้อมูล (Interactive Features) ซึ่งช่วยให้ผู้ใช้งานสามารถติดตามผลการดำเนินงาน วิเคราะห์แนวโน้ม และตัดสินใจได้อย่างรวดเร็วและมีประสิทธิภาพ

**2.9 การเปรียบเทียบงานวิจัย**

จากการทบทวนงานวิจัยที่เกี่ยวข้อง พบว่า งานของวิลาวัณย์ (2563) มุ่งเน้นการวิเคราะห์อัตราการออมของครัวเรือนในระดับภูมิภาค โดยพบว่าปัจจัยด้านรายได้มีความสัมพันธ์เชิงบวกกับการออม ขณะที่งานของชนกพร (2565) เน้นการเข้าถึงบริการทางการเงินของกลุ่มรายได้น้อย ซึ่งมีผลต่อศักยภาพในการออม

อย่างไรก็ตาม งานวิจัยที่ผ่านมาโดยส่วนใหญ่ยังขาดการวิเคราะห์เชิงบูรณาการที่เชื่อมโยงมิติด้านรายได้ ค่าใช้จ่าย การออม และความเหลื่อมล้ำเชิงพื้นที่เข้าด้วยกัน รวมถึงยังขาดการใช้เครื่องมือวิเคราะห์ข้อมูลเชิงภาพ (Data Visualization) เพื่อสนับสนุนการตีความข้อมูลในเชิงลึก

ดังนั้น การศึกษานี้จึงมุ่งเติมเต็มช่องว่างดังกล่าว โดยการวิเคราะห์ความเหลื่อมล้ำด้านการออมในหลายมิติ พร้อมทั้งประยุกต์ใช้ Power BI Dashboard เพื่อเพิ่มประสิทธิภาพในการวิเคราะห์และการสื่อสารผลการวิจัยเชิงนโยบาย

## 3. วิธีการดำเนินการวิจัย

**3.1 กำหนดประชากรและกลุ่มตัวอย่าง**

3.1.1 ประชากรและกลุ่มตัวอย่าง

3.1.1.1 ประชากร: ครัวเรือนไทยทั่วประเทศที่ปรากฏในชุดข้อมูล SES ตั้งแต่ปี พ.ศ. 2556 – 2566

3.1.1.2 กลุ่มตัวอย่าง: ครัวเรือนที่มีข้อมูลสมบูรณ์เกี่ยวกับรายได้ รายจ่าย เงินออม และที่อยู่ จำแนกตามภูมิภาค (ภาคเหนือ กลาง อีสาน ใต้) และกลุ่มรายได้ แยกจังหวัด

3.1.1.3 กลุ่มตัวอย่างสำหรับการประเมินระบบ: ประกอบด้วยเจ้าหน้าที่ธนาคารจำนวน 30 คน และผู้ทรงคุณวุฒิ ซึ่งใช้ในการประเมินประสิทธิภาพของแดชบอร์ดด้านความชัดเจน การใช้งาน และการนำไปใช้ประโยชน์

3.1.2 แหล่งข้อมูล

3.1.2.1 ใช้ข้อมูลชุดสำรวจภาวะเศรษฐกิจและสังคมของครัวเรือน (Household Socio-Economic Survey: SES) ที่ดำเนินการโดยสำนักงานสถิติแห่งชาติ (NSO)

**3.2 เครื่องมือที่ใช้ในการวิจัย**

3.2.1 โปรแกรม Microsoft Excel สำหรับเตรียมและทำความสะอาดข้อมูล

3.2.2 โปรแกรม Power BI สำหรับการวิเคราะห์ข้อมูลและสร้างรายงาน Dashboard

3.2.3 โปรแกรมที่ใช้พัฒนารายงานในรูปแบบเว็บแอปพลิเคชัน : Microsoft Power Bi Could Service

**3.3 การเตรียมข้อมูล**

ข้อมูลที่ใช้ในการศึกษานี้เป็นข้อมูลที่ผ่านการประมวลผลและสรุป (Aggregated Data) จากการสำรวจภาวะเศรษฐกิจและสังคมของครัวเรือน (SES) ของสำนักงานสถิติแห่งชาติ ครอบคลุมช่วงปี พ.ศ. 2556–2566 โดยชุดข้อมูลที่นำมาวิเคราะห์ประกอบด้วย 83 รายการ และจำนวนตัวแปร (Attributes) ประกอบด้วยตัวแปรหลัก เช่น รายได้รวม รายจ่ายรวม เงินออม ภูมิภาค จังหวัด และกลุ่มรายได้

3.3.1 ตรวจสอบโครงสร้างข้อมูล

3.3.1.1 ตรวจสอบข้อมูล รายได้ รายจ่าย ข้อมูลพื้นที่

3.3.1.2 ตรวจสอบหน่วยข้อมูล

3.3.2 การรวมข้อมูล

3.3.2.1 รวมข้อมูลหลายปี

3.3.2.2 ตรวจสอบความสอดคล้องของข้อมูล ชื่อคอลัมน์, ชนิดข้อมูล, รูปแบบข้อมูล

3.3.3 การแปลงข้อมูล

3.3.3.1 คำนวณค่าเงินออม = รายได้ – ค่าใช้จ่าย

3.3.3.2 คำนวณอัตราการออม (%)

3.3.3.3 คำนวณค่า Gini เพื่อดูความเหลื่อมล้ำการออม

3.3.4 การจัดการ Missing Value และ Outliers

3.3.4.1 Missing Value ข้อมูลที่ขาดค่าหลัก เช่น รายได้หรือค่าใช้จ่าย จะถูกจัดการโดยวิธี Listwise Deletion เนื่องจากตัวแปรดังกล่าวมีความสำคัญต่อการคำนวณ

ตัวชี้วัดหลัก หากข้อมูลขาดหายเกินกว่าร้อยละ 10 ของตัวแปรสำคัญ จะถูกตัดออกจากการวิเคราะห์

3.3.4.2 Outliers ใช้เกณฑ์ Interquartile Range (IQR) ในการตรวจสอบค่าผิดปกติ โดยค่าที่ต่ำกว่า Q1 − 1.5IQR หรือสูงกว่า Q3 + 1.5IQR จะถูกพิจารณาเป็น Outliers และถูกตัดออกเพื่อป้องกันการบิดเบือนผลการวิเคราะห์

3.3.5 การปรับรูปแบบข้อมูล

3.3.5.1 แปลงหน่วยให้เป็นมาตรฐานเดียวกัน (บาท/เดือน)

3.3.5.2 จัดรูปแบบข้อมูลให้เหมาะสมก่อนนำเข้า Power BI

**4. การวิเคราะห์ข้อมูล**

4.1 ดาวน์โหลดข้อมูลชุด SES จากเว็บไซต์ของสำนักงานสถิติแห่งชาติ

4.2 ตรวจสอบความครบถ้วนและความสอดคล้องของข้อมูล เช่น รายได้ รายจ่าย การออม หนี้สิน

4.3 คัดกรองเฉพาะกลุ่มตัวอย่างที่มีข้อมูลครบถ้วน โดยแยกตาม ภูมิภาค (ภาคเหนือ ภาคตะวันออกเฉียงเหนือ ภาคกลาง ภาคใต้ กรุงเทพฯ) กลุ่มรายได้ (เช่น รายได้น้อย ปานกลาง สูง)

4.4 ทำความสะอาดข้อมูล (Data Cleaning) เช่น แก้ไขค่าผิดปกติ (Outliers) ลบข้อมูลที่ขาดหาย (Missing Values) แปลงหน่วยหรือประเภทข้อมูลให้สอดคล้องกัน

4.5 คำนวณตัวชี้วัดด้านการออม

4.5.1 จำนวนเงินออมเฉลี่ย (Savings)

4.5.2 อัตราการออม = (เงินออม / รายได้) x 100

4.5.3 Gini index เฉลี่ย

4.6 นำข้อมูลจาก Excel เข้าโปรแกรม Power BI

4.7 แสดงผลข้อมูลด้วย Power BI เพื่อแสดงผลในมิติต่างๆ ได้แก่

4.7.1 กราฟแท่งเปรียบเทียบระดับการออมระหว่างภูมิภาค

4.7.2 กราฟเส้นแสดงแนวโน้มการออมในแต่ละปี

4.7.3 กราฟวงกลมแสดงสัดส่วนกลุ่มรายได้

4.7.4 แผนที่ประเทศไทยแสดงข้อมูลการออมจำแนกตามจังหวัด

4.7.5 ตลอดจนเส้นแนวโน้มแสดงการเปลี่ยนแปลงตลอดช่วงปี พ.ศ. 2556–2566

4.7.6 ข้อมูลแดชบอร์ดแสดงผ่านหน้าเว็บเพจ

4.7.7 วัดประสิทธิภาพของระบบและความพึงพอใจของผู้ใช้

4.7.8 สรุปผลและตีความข้อมูลในบริบทของความเหลื่อมล้ำทางเศรษฐกิจ

**5. การวิเคราะห์ข้อมูลและสถิติที่ใช้ในการวิจัย**

ในการวิจัยเรื่อง การพัฒนาแดชบอร์ดเพื่อวิเคราะห์เชิงเปรียบเทียบความเหลื่อมล้ำในการออมของครัวเรือนไทยตามภาคและกลุ่มรายได้ โดยใช้ข้อมูลจากสำนักงานสถิติแห่งชาติ (NSO) และนำเสนอผลผ่าน Power BI ผู้วิจัยได้ใช้โปรแกรมสำเร็จรูปทางสถิติ SPSS ในการวิเคราะห์ข้อมูล โดยแบ่งการวิเคราะห์ออกเป็น 2 ส่วนหลัก ได้แก่

การวิเคราะห์แบบประเมินประสิทธิภาพและความพึงพอใจของผู้ใช้งานผู้วิจัยใช้การวิเคราะห์ค่าทางสถิติ ได้แก่ ค่าเฉลี่ย (Mean) และ ส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation: SD) เพื่อประเมินประสิทธิภาพของระบบแดชบอร์ดและระดับความพึงพอใจของผู้ใช้งานทั่วไปและผู้เชี่ยวชาญ โดยผลการวิเคราะห์ถูกนำเสนอในรูปแบบตารางควบคู่กับการอธิบายเชิงพรรณนา เพื่อให้เห็นระดับความคิดเห็นของผู้ตอบแบบสอบถามต่อระบบแดชบอร์ดที่พัฒนาขึ้น

การวิเคราะห์ข้อมูลทั่วไปของผู้ตอบแบบสอบถาม

ผู้วิจัยใช้ สถิติเชิงพรรณนา (Descriptive Statistics) ในการวิเคราะห์ข้อมูลพื้นฐานของกลุ่มตัวอย่าง โดยนำเสนอผลในรูปแบบของ ค่าความถี่ (Frequency) ร้อยละ (Percentage) ค่าเฉลี่ย (Mean) และส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation) เพื่ออธิบายลักษณะทั่วไปของผู้ตอบแบบสอบถามและแสดงภาพรวมของกลุ่มตัวอย่างที่ใช้ในการประเมินระบบแดชบอร์ด

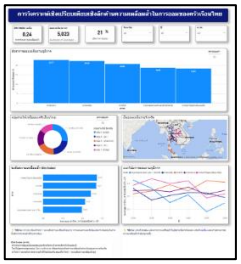
โดยการใช้สถิติดังกล่าวช่วยให้สามารถสรุปผลการประเมินประสิทธิภาพของระบบแดชบอร์ดและระดับความพึงพอใจของผู้ใช้งานได้อย่างเป็นระบบและชัดเจน

**4. ผลการวิเคราะห์ข้อมูล**

ก่อนดำเนินการวิเคราะห์ข้อมูล ผู้วิจัยได้ทำการตรวจสอบและจัดการข้อมูลเพื่อให้ผลการวิเคราะห์มีความถูกต้องและน่าเชื่อถือ โดยข้อมูลที่ขาดหาย (Missing Values) ในตัวแปรสำคัญ เช่น รายได้และค่าใช้จ่าย ถูกจัดการด้วยวิธี Listwise Deletion ซึ่งเป็นการตัดข้อมูลที่ไม่สมบูรณ์ออกจากการวิเคราะห์ เนื่องจากตัวแปรดังกล่าวมีผลโดยตรงต่อการคำนวณตัวชี้วัดหลัก สำหรับค่าผิดปกติ (Outliers) ผู้วิจัยใช้เกณฑ์ Interquartile Range (IQR) ในการตรวจสอบ โดยค่าที่อยู่นอกช่วง Q1 − 1.5IQR และ Q3 + 1.5IQR จะถูกพิจารณาเป็นค่าผิดปกติ และถูกตัดออกจากชุดข้อมูล เพื่อป้องกันการบิดเบือนผลการวิเคราะห์

การดำเนินการดังกล่าวช่วยให้ข้อมูลที่นำมาวิเคราะห์มีความเหมาะสม และไม่ส่งผลกระทบต่อความถูกต้องของผลลัพธ์ในภาพรวม

**4.1 ผลการวิเคราะห์ภาพรวมความเหลื่อมล้ำด้านการออมของครัวเรือนไทย**

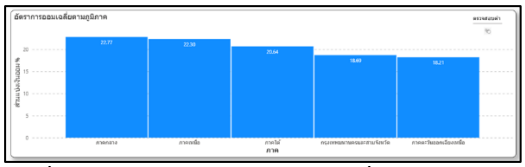


ภาพที่ 4.1 ผลการวิเคราะห์ภาพรวมความเหลื่อมล้ำด้านการออมของครัวเรือนไทย

จากการวิเคราะห์ข้อมูลการออมของครัวเรือนไทยโดยใช้ Power BI Dashboard (ภาพที่ 4.1) พบว่า ในภาพรวมครัวเรือนไทยมี ค่า Gini Index ด้านการออมเฉลี่ยเท่ากับ 0.24 ซึ่งสะท้อนถึงระดับความเหลื่อมล้ำด้านการออมในระดับปานกลาง โดยค่า Gini ที่มีค่าเข้าใกล้ 0 แสดงถึงการกระจายการออมที่ค่อนข้างเท่าเทียม ขณะที่ค่าเข้าใกล้ 1 แสดงถึงความเหลื่อมล้ำสูง

นอกจากนี้ ครัวเรือนไทยมี เงินออมเฉลี่ยประมาณ 5,023 บาทต่อเดือน และมี อัตราการออมเฉลี่ยคิดเป็นร้อยละ 21 ของรายได้ ซึ่งสะท้อนให้เห็นว่าครัวเรือนส่วนใหญ่ยังมีศักยภาพในการออมในระดับจำกัด เมื่อพิจารณาร่วมกับภาระค่าใช้จ่ายและโครงสร้างรายได้ของแต่ละกลุ่มครัวเรือน

**4.2 ผลการวิเคราะห์การเปรียบเทียบอัตราการออมเฉลี่ยตามภูมิภาค**

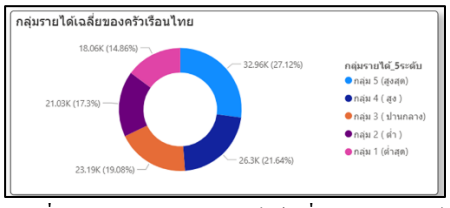


ภาพที่ 4.2 การเปรียบเทียบอัตราการออมเฉลี่ยตามภูมิภาค

จากกราฟแท่งแสดงอัตราการออมเฉลี่ยตามภูมิภาค พบว่า ภาคกลางมีอัตราการออมเฉลี่ยสูงที่สุด (ประมาณร้อยละ 22.77) รองลงมาคือ ภาคเหนือ (ร้อยละ 22.30) และภาคใต้ (ร้อยละ 20.84) ขณะที่ กรุงเทพมหานครและ 3 จังหวัด และ ภาคตะวันออกเฉียงเหนือ มีอัตราการออมเฉลี่ยต่ำกว่าภูมิภาคอื่น

ผลการวิเคราะห์นี้สะท้อนให้เห็นว่า ครัวเรือนในพื้นที่ที่มีโอกาสทางเศรษฐกิจและการจ้างงานสูง มีแนวโน้มสามารถจัดสรรรายได้เพื่อการออมได้มากกว่า ในขณะที่พื้นที่ที่มีรายได้เฉลี่ยต่ำหรือพึ่งพาภาคเกษตรกรรมเป็นหลัก ยังคงเผชิญข้อจำกัดในการออมอย่างมีนัยสำคัญ

**4.3 ผลการวิเคราะห์การกระจายกลุ่มรายได้เฉลี่ยของครัวเรือนไทย**



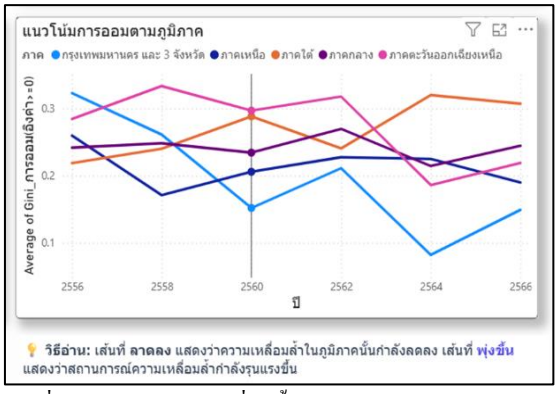

ภาพที่ 4.3 การกระจายกลุ่มรายได้เฉลี่ยของครัวเรือนไทย

จากกราฟเปรียบเทียบค่า Gini Index ด้านการออมจำแนกตามภูมิภาค พบว่า โครงสร้างรายได้ของครัวเรือนไทยกระจายอยู่ในทุกระดับ โดยกลุ่มรายได้ระดับสูง (ระดับ 4–5) มีสัดส่วนรวมกันมากกว่าครึ่งหนึ่งของครัวเรือนทั้งหมด ขณะที่กลุ่มรายได้ระดับต่ำ (ระดับ 1–2)

มีสัดส่วนรองลงมาแต่ยังเป็นประชากรกลุ่มใหญ่ของประเทศ

ความแตกต่างด้านรายได้ดังกล่าวสะท้อนศักยภาพในการออมที่ไม่เท่ากัน โดยครัวเรือนรายได้สูงสามารถออมได้มากกว่าครัวเรือนรายได้ต่ำ ซึ่งเป็นปัจจัยสำคัญที่ก่อให้เกิดความเหลื่อมล้ำด้านการออมของครัวเรือนไทย

**4.5 ผลการวิเคราะห์แนวโน้มความเหลื่อมล้ำด้านการออมตามช่วงเวลา**




ภาพที่ 4.5 แนวโน้มความเหลื่อมล้ำด้านการออมตามช่วงเวลา

จากกราฟแนวโน้มค่า Gini Index ด้านการออมของครัวเรือนไทย พบว่าค่า Gini Index ด้านการออมของครัวเรือนไทยในช่วงปี พ.ศ. 2556–2566 อยู่ในช่วงประมาณ 0.08–0.33 แสดงให้เห็นว่าความเหลื่อมล้ำด้านการออมแตกต่างกันตามภูมิภาคและมีความผันผวนตามช่วงเวลา

โดย ภาคตะวันออกเฉียงเหนือและภาคใต้มีความเหลื่อมล้ำสูงและผันผวนมากที่สุด ขณะที่ ภาคกลางมีการกระจายการออมค่อนข้างเสถียร เมื่อเทียบกับภูมิภาคอื่น.

ภาษาอังกฤษและตัวเลขใช้ Times New Roman ขนาด 11 จุด ตัวหนาและจัดแบบชิดซ้าย ส่วนชื่อหัวข้อภาษาไทยใช้ตัวอักษร Angsana New ขนาด 14 จุด ตัวหนา เว้นหนึ่งบรรทัดก่อนขึ้นหัวข้อลำดับที่ 2 ไม่ต้องเว้นบรรทัดหลังหัวข้อลำดับที่ 2 ให้ขึ้นเนื้อหาได้เลย (ให้ดูตัวอย่างจากการเขียนเอกสารนี้ประกอบ)

**4.6 วิเคราะห์ข้อมูลการประเมินผลคุณภาพแดชบอร์ด**

4.6.1 แบบแสดงความคิดเห็นของผู้ใช้งานทั่วไปที่มีต่อการวิเคราะห์เชิงเปรียบเทียบเชิงลึกด้านความเหลื่อมล้ำในการออมของครัวเรือนไทย กรณีศึกษาตามภาคและกลุ่มรายได้ โดยใช้ข้อมูล **NSO** และนำเสนอผลด้วย **Power BI**

**ตารางที่ 4-1** ผลการวิเคราะห์ข้อมูลทั่วไปของเจ้าหน้าที่ธนาคาร

| ด้านที่ 1 ข้อมูลทั่วไปของเจ้าหน้าที่ธนาคารแบบสอบถาม | จำนวน | ร้อยละ % |
|---|---|---|
| **1. เพศ** | | |
| ชาย | 15 | 50.00 |
| หญิง | 15 | 50.00 |
| **รวม** | 30 | 100.00 |
| **ด้านที่ 1 ข้อมูลทั่วไปของเจ้าหน้าที่ธนาคารแบบสอบถาม** | **จำนวน** | **ร้อยละ %** |
| **2. อายุ** | | |
| 18–24 ปี | 2 | 6.70 |
| 25–34 ปี | 10 | 33.30 |
| 35–44 ปี | 12 | 40.00 |
| 45–54 ปี | 5 | 16.70 |
| 55 ปีขึ้นไป | 1 | 3.30 |
| **รวม** | 30 | 100.0 |
| **3. ประสบการณ์ใช้งาน Dashboard** | | |
| มี | 18 | 60.00 |
| ไม่มี | 12 | 40.00 |
| **รวม** | 30 | 100.00 |

จากตารางที่ 4-1 ผลการวิเคราะห์ข้อมูลทั่วไปของเจ้าหน้าที่ธนาคาร จำนวน 30 คน มีสัดส่วนเพศชายและเพศหญิงเท่ากัน คือเพศละ ร้อยละ 50ส่วนใหญ่อยู่ในช่วงอายุ 35–44 ปี มากที่สุด รองลงมาคือช่วงอายุ 25–34 ปี และ 45–54 ปี ตามลำดับ นอกจากนี้ ผู้ตอบแบบสอบถามส่วนใหญ่ มีประสบการณ์ในการใช้งาน **Dashboard** ร้อยละ 60 แสดงว่ากลุ่มตัวอย่างมีความคุ้นเคยกับระบบการแสดงผล

**ตารางที่ 4-2** ผลการวิเคราะห์ข้อมูลด้านความชัดเจนและการใช้งานง่าย

| ด้านที่ 2 ความชัดเจนและการใช้งานง่าย | $\overline{X}$ | S.D. | ระดับการประเมิน |
|---|---|---|---|
| 1. Dashboard แสดงข้อมูลได้ชัดเจนและเข้าใจง่าย | 4.20 | 0.81 | ดีมาก |
| 2. กราฟและแผนภาพช่วยให้เข้าใจข้อมูลได้รวดเร็ว | 4.60 | 0.67 | ดีมาก |
| 3. การใช้สีและรูปแบบการแสดงผลมีความเหมาะสม | 4.50 | 0.63 | ดีมาก |
| 4. Dashboard ใช้งานได้ง่าย ไม่ซับซ้อน | 4.23 | 0.82 | ดีมาก |
| 5. การจัดวางองค์ประกอบบนหน้าจอมีความเหมาะสม | 4.33 | 0.61 | ดีมาก |
| รวม | 4.37 | 0.71 | ดีมาก |

จากตารางที่ 4-2 ผลการวิเคราะห์ข้อมูลด้าน ความชัดเจนและการใช้งานง่าย พบว่าความชัดเจนและการใช้งานง่ายของ Dashboard โดยรวมอยู่ในระดับ ดีมาก ($\bar{X}$= 4.37, **S.D.** = 0.71) แสดงให้เห็นว่าผู้ใช้งานสามารถเข้าใจและใช้งานระบบได้อย่างมีประสิทธิภาพ ข้อที่มีค่าเฉลี่ยสูงสุดคือ กราฟและแผนภาพช่วยให้เข้าใจข้อมูลได้รวดเร็ว รองลงมาคือ การใช้สีและรูปแบบการแสดงผลมีความเหมาะสม และ การจัดวางองค์ประกอบบนหน้าจอเหมาะสม ซึ่งทั้งหมดอยู่ในระดับดีมาก โดยสรุป ผู้ตอบแบบสอบถามมีความพึงพอใจต่อความชัดเจนและการใช้งานของ Dashboard ในระดับดีมาก โดยเฉพาะการใช้กราฟและแผนภาพที่ช่วยสื่อสารข้อมูลได้อย่างรวดเร็วและเข้าใจง่าย

**ตารางที่ 4-3** ผลการวิเคราะห์ข้อมูลด้านการเปรียบเทียบและการวิเคราะห์ข้อมูล

| ด้านที่ 3 การเปรียบเทียบและการวิเคราะห์ข้อมูล | $\bar{X}$ | S.D. | ระดับการประเมิน |
|---|---|---|---|
| 1. การเลือกดูข้อมูลตามภาคและกลุ่มรายได้ทำได้สะดวก | 4.03 | 0.89 | ดี |
| 2. การโต้ตอบกับ Dashboard (เช่น การกรองข้อมูล) ทำได้อย่างราบรื่น | 4.10 | 0.88 | ดี |
| 3. Dashboard ช่วยให้เห็นความแตกต่างด้านการออมระหว่างภูมิภาคได้ชัดเจน | 4.20 | 0.76 | ดี |
| 4. Dashboard ช่วยให้เปรียบเทียบข้อมูลระหว่างกลุ่มรายได้ได้ง่าย | 4.27 | 0.64 | ดีมาก |
| 5. Dashboard สนับสนุนการวิเคราะห์เชิงเปรียบเทียบได้อย่างมีประสิทธิภาพ | 4.20 | 0.71 | ดี |
| รวม | 4.16 | 0.78 | ดี |

จากตารางที่ 4-3 ผลการวิเคราะห์ข้อมูลด้าน การเปรียบเทียบและการวิเคราะห์ข้อมูล พบว่าการเปรียบเทียบและการวิเคราะห์ข้อมูล โดยรวมอยู่ในระดับ ดี ($\bar{X}$= 4.16, **S.D.** = 0.78) แสดงว่า Dashboard สามารถสนับสนุนการวิเคราะห์ข้อมูลได้อย่างเหมาะสม ข้อที่มีค่าเฉลี่ยสูงสุดคือ Dashboard ช่วยให้เปรียบเทียบข้อมูลระหว่างกลุ่มรายได้ได้ง่าย รองลงมาคือการช่วยให้เห็นความแตกต่างด้านการออมระหว่างภูมิภาคและการสนับสนุนการวิเคราะห์เชิงเปรียบเทียบ โดยสรุป Dashboard มีจุดเด่นด้าน การเปรียบเทียบข้อมูลระหว่างกลุ่มรายได้และภูมิภาคได้อย่างชัดเจน ช่วยให้ผู้ใช้งานวิเคราะห์ข้อมูลได้สะดวกมากขึ้น

**ตารางที่ 4-4** ผลการวิเคราะห์ข้อมูลด้านการนำไปใช้ประโยชน์และความพึงพอใจโดยรวม

| ด้านที่ 4 การนำไปใช้ประโยชน์และความพึงพอใจโดยรวม | $\bar{X}$ | S.D. | ระดับการประเมิน |
|---|---|---|---|
| 1 Dashboard ช่วยให้เข้าใจความเหลื่อมล้ำในการออมได้ดีขึ้น | 4.00 | 0.83 | ดี |
| 2. Dashboard สามารถนำไปใช้ประกอบการตัดสินใจได้ | 4.30 | 0.75 | ดีมาก |
| 3. Dashboard มีประโยชน์ต่อการวิเคราะห์เชิงวิชาการหรือเชิง | 4.20 | 0.92 | ดี |
| 4. โดยรวมแล้ว Dashboard มีประโยชน์ต่อผู้ใช้งาน | 4.47 | 0.51 | ดีมาก |
| 5. ผู้ใช้งานมีความพึงพอใจต่อ Power BI Dashboard โดยรวม | 4.30 | 0.75 | ดีมาก |
| รวม | 4.25 | 0.75 | ดีมาก |

จากตารางที่ **4-4** ผลการวิเคราะห์ข้อมูลด้าน การนำไปใช้ประโยชน์และความพึงพอใจโดยรวม พบว่าการนำไปใช้ประโยชน์และความพึงพอใจโดยรวม อยู่ในระดับดีมาก ($\bar{X}$= 4.25, **S.D.** = 0.75) แสดงว่า Dashboard มีประโยชน์ต่อการใช้งานและได้รับความพึงพอใจจากผู้ใช้ ข้อที่มีค่าเฉลี่ยสูงสุดคือ Dashboard มีประโยชน์ต่อผู้ใช้งานโดยรวม รองลงมาคือ การนำไปใช้ประกอบการตัดสินใจ และ ความพึงพอใจต่อ Power BI Dashboard โดยรวม โดยสรุป ผู้ตอบแบบสอบถามเห็นว่า Dashboard ช่วยสนับสนุนการทำงานและการตัดสินใจได้อย่างมีประสิทธิภาพ และมีความพึงพอใจต่อการใช้งานในระดับดีมาก

4.6.2 แบบแสดงความคิดเห็นของผู้ทรงคุณวุฒิที่มีต่อการวิเคราะห์เชิงเปรียบเทียบเชิงลึกด้านความเหลื่อมล้ำในการออมของครัวเรือนไทย กรณีศึกษาตามภาคและกลุ่มรายได้

ตารางที่ 4-5 ความสอดคล้องของเนื้อหาและตัวชี้วัดกับวัตถุประสงค์การวิจัย

| ด้านที่ 1 ความสอดคล้องของเนื้อหาและตัวชี้วัดกับวัตถุประสงค์การวิจัย | $\bar{X}$ | S.D. | ระดับการประเมิน |
|---|---|---|---|
| 1ตัวชี้วัดการออมสะท้อนระดับการออมของ | 4.0 | 1.00 | ดี |
| 2.ตัวชี้วัดอธิบายความแตกต่างด้านการออมระหว่างครัวเรือนได้ | 4.00 | 0.00 | ดี |
| 3. ตัวชี้วัดสะท้อนความเหลื่อมล้ำในการออมได้ชัดเจน | 4.33 | 0.57 | ดี |
| 4. ตัวชี้วัดสอดคล้องกับแนวคิดเศรษฐศาสตร์ครัวเรือน | 4.33 | 0.57 | ดี |
| 5. ตัวชี้วัดสอดคล้องกับกรอบแนวคิดการ | 3.6 | 1.15 | ดี |
| รวม | 4.0 | 0.50 | ดี |

จากตารางที่ 4-5 ผลการวิเคราะห์ความสอดคล้องของเนื้อหาและตัวชี้วัดกับวัตถุประสงค์การวิจัยพบว่าผู้ทรงคุณวุฒิมีความคิดเห็นว่าความสอดคล้องของเนื้อหาและตัวชี้วัดอยู่ในระดับ ดี ($\bar{X}$ = 4.06, **S.D.** = 0.50) สะท้อนว่าตัวชี้วัดสามารถอธิบายความเหลื่อมล้ำด้านการออมได้สอดคล้องกับแนวคิดทางเศรษฐศาสตร์ครัวเรือน โดยสรุปเนื้อหาและตัวชี้วัดที่ใช้ในการวิจัยมีความเหมาะสมและสอดคล้องกับวัตถุประสงค์ของการศึกษา สามารถนำไปใช้ในการวิเคราะห์ข้อมูลได้อย่างมีประสิทธิภาพ

**ตารางที่ 4-6** ความเหมาะสมของข้อมูลและวิธีการวิเคราะห์

| ด้านที่ 2 ความเหมาะสมของข้อมูลและวิธีการวิเคราะห์ | $\bar{X}$ | S.D. | ระดับการประเมิน |
|---|---|---|---|
| 1 ข้อมูลจากสำนักงานสถิติแห่งชาติมีความเหมาะสมกับการวิจัย | 4.67 | 0.57 | ดีมาก |
| 2.สูตรการคำนวณเงินออมมีความถูกต้องตามหลักวิชาการ | 4.00 | 0.00 | ดี |
| 3. ตัวชี้การแบ่งกลุ่มรายได้สามารถสะท้อนความแตกต่างด้านการออมได้ | 4.00 | 1.00 | ดี |
| 4. การจำแนกข้อมูลตามภูมิภาคเหมาะสมต่อการวิเคราะห์เชิงเปรียบเทียบ | 4.00 | 1.73 | ดี |
| 5. กระบวนการเตรียมข้อมูลไม่ส่งผลกระทบต่อความถูกต้องของผลการวิเคราะห์ | 3.33 | 0.57 | ปานกลาง |
| รวม | 4.00 | 0.40 | ดี |

จากตารางที่ **4-6** ผลการวิเคราะห์ข้อมูลความเหมาะสมของข้อมูลและวิธีการวิเคราะห์ พบว่าผลการวิเคราะห์ความเหมาะสมของข้อมูลและวิธีการวิเคราะห์โดยรวมอยู่ในระดับ ดี ($\bar{X}$ = 4.00, **S.D.** = 0.40) แสดงว่าข้อมูลและแนวทางการวิเคราะห์มีความเหมาะสมต่อการวิจัย ข้อที่มีค่าเฉลี่ยสูงสุดคือ ข้อมูลจากสำนักงานสถิติแห่งชาติมีความเหมาะสมกับการวิจัย อยู่ในระดับ ดีมาก สะท้อนถึงความน่าเชื่อถือของแหล่งข้อมูลและความเหมาะสมในการนำมาใช้ศึกษา ขณะที่กระบวนการเตรียมข้อมูล มีค่าเฉลี่ยต่ำสุดอยู่ในระดับ ปานกลาง ซึ่งอาจควรมีการตรวจสอบหรือปรับปรุงขั้นตอนการเตรียมข้อมูลให้รัดกุมมากยิ่งขึ้น

**ตารางที่ 4-7 ความเหมาะสมของการนำเสนอผลด้วย** Dashboard (Power BI)

| ด้านที่ 3 ความเหมาะสมของการนำเสนอผลด้วย Dashboard (Power BI) | $\bar{X}$ | S.D. | ระดับการประเมิน |
|---|---|---|---|
| 1. Dashboard สามารถสื่อสารผลการวิเคราะห์ความเหลื่อมล้ำได้อย่างชัดเจน | 4.33 | 0.57 | ดี |
| 2.รูปแบบกราฟและแผนภาพช่วยให้เข้าใจข้อมูลได้ง่าย | 3.67 | 0.57 | ดี |
| 3. Dashboard เอื้อต่อการเปรียบเทียบข้อมูลระหว่างภาคและกลุ่มรายได้ | 3.67 | 1.15 | ดี |
| 4. การนำเสนอผลด้วย Power BI มีความเหมาะสมต่อการสื่อสารข้อมูลเชิง | 3.67 | 1.15 | ดี |
| 5. Dashboard สามารถใช้สนับสนุนการตัดสินใจหรือข้อเสนอเชิงนโยบาย | 3.33 | 1.52 | ปานกลาง |
| รวม | 3.73 | 0.94 | ดี |

จากตารางที่ 4-7 ผลการวิเคราะห์ข้อมูลความเหมาะสมของการนำเสนอผลด้วย Dashboard (Power BI) พบว่าผลการวิเคราะห์ความเหมาะสมของการนำเสนอผลด้วย Dashboard (Power BI) โดยรวมอยู่ในระดับ ดี ($\bar{X}$ = 3.73, S.D. = 0.945) แสดงว่า Dashboard สามารถสื่อสารผลการวิเคราะห์ได้อย่างเหมาะสม ข้อที่มีค่าเฉลี่ยสูงสุดคือ Dashboard สามารถสื่อสารผลการวิเคราะห์ความเหลื่อมล้ำได้อย่างชัดเจน ขณะที่การใช้กราฟและการเปรียบเทียบข้อมูลระหว่างภาคและกลุ่มรายได้อยู่ในระดับดีเช่นกัน อย่างไรก็ตาม การใช้ Dashboard เพื่อ สนับสนุนการตัดสินใจเชิงนโยบาย มีค่าเฉลี่ยต่ำสุดอยู่ในระดับปานกลาง จึงอาจควรมีการปรับปรุงการออกแบบเพื่อเพิ่มประสิทธิภาพในการใช้งานเชิงบริหารมากยิ่งขึ้น

**4.7 การประยุกต์ใช้** Dashboard **เพื่อการวิเคราะห์เชิงพยากรณ์**

นอกจากการวิเคราะห์ข้อมูลเชิงย้อนหลัง (Descriptive Analysis) การศึกษานี้ได้พิจารณาการพัฒนา Dashboard ให้รองรับการวิเคราะห์เชิงพยากรณ์ (Predictive Analysis) และการวิเคราะห์ฉากทัศน์ (Scenario Analysis) เพื่อเพิ่มประสิทธิภาพในการใช้งานเชิงนโยบาย โดยสามารถประยุกต์ใช้เทคนิคการพยากรณ์แนวโน้ม เช่น Time Series Forecasting เพื่อคาดการณ์การเปลี่ยนแปลงของอัตราการออมในอนาคต รวมถึงการจำลองสถานการณ์ (What-if Analysis) เช่น การเปลี่ยนแปลงรายได้หรือค่าใช้จ่ายของ

ครัวเรือน เพื่อประเมินผลกระทบต่อระดับความเหลื่อมล้ำ แนวทางดังกล่าวช่วยให้ผู้บริหารสามารถใช้ Dashboard เป็นเครื่องมือสนับสนุนการตัดสินใจเชิงนโยบายได้อย่างมีประสิทธิภาพมากยิ่งขึ้น แทนการพิจารณาข้อมูลย้อนหลังเพียงอย่างเดียว

## 5 สรุปผลการวิจัย อภิปรายผล และข้อเสนอแนะ

### 5.1 สรุปผลการวิจัย

**ภาพรวมความเหลื่อมล้ำด้านการออม** ความเหลื่อมล้ำด้านการออมของครัวเรือนไทยอยู่ในระดับปานกลาง (Gini Index 0.24) มีเงินออมเฉลี่ย 5,023 บาทต่อเดือน หรือร้อยละ 21 ของรายได้ สะท้อนถึงการกระจายตัวที่ยังไม่เท่าเทียมและเป็นประเด็นสำคัญที่ต้องได้รับความใส่ใจเชิงนโยบาย

**ความแตกต่างเชิงพื้นที่และภูมิภาค** อัตราการออมมีความแตกต่างตามภูมิภาคชัดเจน โดยภาคกลางมีอัตราการออมสูงสุด (ร้อยละ 22.77) ขณะที่กรุงเทพฯ และภาคตะวันออกเฉียงเหนือมีอัตราต่ำกว่า สะท้อนว่าพื้นที่ที่มีโอกาสทางเศรษฐกิจสูงสามารถจัดสรรรายได้เพื่อการออมได้ดีกว่าพื้นที่เกษตรกรรม

จากผลการวิเคราะห์พบว่า ภาคตะวันออกเฉียงเหนือและภาคใต้มีความผันผวนของความเหลื่อมล้ำด้านการออมสูงกว่าภูมิภาคอื่น ซึ่งสามารถอธิบายได้จากโครงสร้างเศรษฐกิจของพื้นที่ที่พึ่งพาภาคเกษตรกรรมและทรัพยากรธรรมชาติเป็นหลัก

โดยรายได้ของครัวเรือนในภูมิภาคดังกล่าวมีความอ่อนไหวต่อปัจจัยภายนอก เช่น ความผันผวนของราคาสินค้าเกษตร สภาพภูมิอากาศ และเหตุการณ์วิกฤต เช่น การแพร่ระบาดของโรค COVID-19 ซึ่งส่งผลกระทบต่อรายได้และความสามารถในการออมของครัวเรือนอย่างมีนัยสำคัญ

ในขณะที่ภาคกลางซึ่งมีโครงสร้างเศรษฐกิจที่หลากหลายและมีโอกาสในการจ้างงานสูงกว่า มีความสามารถในการกระจายรายได้และการออมได้อย่างสม่ำเสมอมากกว่า ส่งผลให้ระดับความเหลื่อมล้ำมีความเสถียรเมื่อเปรียบเทียบกับภูมิภาคอื่น

ดังนั้น ข้อเสนอแนะเชิงนโยบายควรมุ่งเน้นการลดความเปราะบางทางเศรษฐกิจในพื้นที่ที่พึ่งพาภาคเกษตรกรรม เช่น การส่งเสริมรายได้เสริม การพัฒนาทักษะแรงงาน และการเพิ่มการเข้าถึงบริการทางการเงิน เพื่อช่วยลดความผันผวนของรายได้และเพิ่มศักยภาพในการออมของครัวเรือนในระยะยาว

**โครงสร้างรายได้และศักยภาพการออม** ครัวเรือนกลุ่มรายได้สูงมีศักยภาพการออมสูงกว่ากลุ่มรายได้ต่ำอย่างชัดเจน โดยกลุ่มรายได้ระดับสูง (4 และ 5) มีสัดส่วนรวมกันมากกว่าครึ่งหนึ่งของกลุ่มตัวอย่าง ทำให้การสะสมความมั่งคั่งกระจุกตัวและกลายเป็นฐานของความเหลื่อมล้ำในระยะยาว

**ความผันผวนของความเหลื่อมล้ำตามช่วงเวลา** ค่า Gini Index ด้านการออมมีความผันผวนตามสถานการณ์เศรษฐกิจในช่วงปี 2556–2566 โดยภาคตะวันออกเฉียงเหนือและภาคใต้มีความผันผวนสูงและรุนแรงขึ้นในช่วงหลัง ในขณะที่ภาคกลางมีการกระจายการออมที่มีเสถียรภาพมากที่สุด

**ประสิทธิภาพของ Power BI Dashboard** ผลการประเมินจากผู้ใช้งานพบว่าแดชบอร์ดมีความชัดเจนและใช้งานง่ายในระดับดีมาก (ค่าเฉลี่ย 4.56) โดยเฉพาะการใช้สีและรูปแบบที่เหมาะสม ช่วยให้การโต้ตอบและเปรียบเทียบข้อมูลระหว่างกลุ่มรายได้ทำได้สะดวกและมีประสิทธิภาพสูง

**การนำไปใช้ประโยชน์และความพึงพอใจ** ผู้ใช้งานมีความพึงพอใจโดยรวมในระดับดีมาก เนื่องจากแดชบอร์ดมีประโยชน์ต่อการวิเคราะห์เชิงนโยบายและช่วยสนับสนุนการตัดสินใจได้จริง พร้อมทั้งช่วยให้เข้าใจภาพรวมของความเหลื่อมล้ำด้านการออมได้อย่างชัดเจนเป็นรูปธรรม

**ความสอดคล้องของเนื้อหาและตัวชี้วัด** ผู้ทรงคุณวุฒิประเมินว่าเนื้อหาและตัวชี้วัดมีความสอดคล้องกับวัตถุประสงค์การวิจัยในระดับดี (ค่าเฉลี่ย 4.06) โดยตัวชี้วัดดัชนีจีนีสามารถสะท้อนปรากฏการณ์ความเหลื่อมล้ำได้เหมาะสมตามกรอบแนวคิดเศรษฐศาสตร์ครัวเรือน

**ความเหมาะสมของข้อมูลและวิธีการวิเคราะห์** ข้อมูลจากสำนักงานสถิติแห่งชาติ (NSO) และวิธีการวิเคราะห์มีความเหมาะสมในระดับดี แหล่งข้อมูลมีความน่าเชื่อถือสูง และสูตรการคำนวณถูกต้องตามหลักวิชาการ แม้ควรเพิ่มความรัดกุมในขั้นตอนการเตรียมข้อมูลให้มากยิ่งขึ้น

**คุณภาพการนำเสนอผลด้วยแดชบอร์ด** การนำเสนอผ่าน Power BI มีความเหมาะสมในระดับดี (ค่าเฉลี่ย 3.73) สามารถสื่อสารผลการวิเคราะห์เชิงเปรียบเทียบได้อย่างชัดเจน แต่อาจต้องปรับปรุงการออกแบบเพิ่มเติมเพื่อให้ตอบสนองต่อการใช้งานเชิงบริหารและนโยบายให้ดียิ่งขึ้น

**5.2 อภิปรายผลการวิจัย**

ผลการวิจัยสอดคล้องกับทฤษฎีรายได้และวัฏจักรชีวิตที่ชี้ว่าครัวเรือนรายได้สูงมีศักยภาพการออมมากกว่า โดยปัจจัยด้านภูมิภาคและความรู้ทางการเงินเป็นตัวแปรสำคัญที่เอื้อต่อการสะสมความมั่งคั่งในระยะยาว ทั้งนี้การนำเสนอผ่าน Power BI Dashboard ช่วยยกระดับการสื่อสารข้อมูลเชิงซ้อนให้ชัดเจนและเข้าใจง่ายกว่าสถิติรูปแบบเดิม ซึ่งสอดคล้องกับแนวคิดด้าน Data Visualization ที่สนับสนุนการตัดสินใจเชิงนโยบายได้อย่างมีประสิทธิภาพ

**5.3 ข้อเสนอแนะในการวิจัยครั้งต่อไป**

5.3.1 ข้อเสนอแนะเชิงนโยบาย

จากผลการวิจัย สามารถเสนอแนวทางเชิงนโยบายได้ดังนี้

5.3.1.1 ภาครัฐควรส่งเสริมมาตรการเพิ่มรายได้และลดภาระค่าใช้จ่ายของครัวเรือนรายได้น้อย เพื่อเพิ่มศักยภาพในการออมอย่างยั่งยืน

5.3.1.2 ควรพัฒนาการเข้าถึงบริการทางการเงินและความรู้ทางการเงินในภูมิภาคที่มีความเหลื่อมล้ำสูง เช่น ภาคตะวันออกเฉียงเหนือและภาคใต้

5.3.1.3 สนับสนุนการใช้เครื่องมือวิเคราะห์ข้อมูลและ Dashboard ในการติดตามสถานการณ์ความเหลื่อมล้ำด้านการออม เพื่อช่วยให้ผู้กำหนดนโยบายสามารถตัดสินใจบนพื้นฐานของข้อมูลเชิงประจักษ์

5.3.1.1 ส่งเสริมการออกแบบนโยบายการออมที่คำนึงถึงความแตกต่างเชิงพื้นที่และกลุ่มรายได้ มากกว่าการใช้มาตรการแบบเดียวกันทั้งประเทศ

5.3.2 ข้อเสนอแนะสำหรับการวิจัยในอนาคต

5.3.2.1 ควรขยายการวิเคราะห์ไปสู่ระดับจังหวัดหรือระดับครัวเรือนรายบุคคล เพื่อสะท้อนความเหลื่อมล้ำเชิงพื้นที่ได้ละเอียดมากขึ้น

5.3.2.2 อาจนำตัวแปรอื่น เช่น หนี้สิน อายุ ระดับการศึกษา หรืออาชีพ มาวิเคราะห์ร่วมกับการออม

5.3.2.3 ควรเปรียบเทียบผลการวิเคราะห์กับประเทศอื่นในภูมิภาค เพื่อศึกษาความเหลื่อมล้ำด้านการออมในบริบทสากล

5.3.2.4 พัฒนาระบบ Dashboard ให้สามารถอัปเดตข้อมูลแบบเรียลไทม์ เพื่อใช้เป็นเครื่องมือเฝ้าระวังทางเศรษฐกิจในระยะยาว

## เอกสารอ้างอิง


[1] กชพร จำรัสรักษ์. (2564). พฤติกรรมการออมของบุคลากรสำนักงานป้องกันและปราบปรามการฟอกเงิน.การค้นคว้าอิสระ มหาวิทยาลัยรามคำแหง.

[2] ชนกพร ชลาลัย. (2565). การเข้าถึงบริการทางการเงินของผู้มีรายได้น้อย.การค้นคว้าอิสระ จุฬาลงกรณ์มหาวิทยาลัย.

[3] ปาจรีย์ ขาวเกื้อ. (2564). การออกแบบผลิตภัณฑ์เงินฝากแบบลุ้นรางวัลเพื่อส่งเสริมการออม [วิทยานิพนธ์ปริญญามหาบัณฑิต, มหาวิทยาลัยธรรมศาสตร์]. TU Digital Collections

[4] วราภรณ์ ทองสุข. (2563). พฤติกรรมการใช้จ่ายของครัวเรือนไทย. วารสารสุทธิปริทัศน์, 34(109), 115-128.

[5] วิลาวัณย์ ศรีสมบัติ. (2563). การวิเคราะห์อัตราการออมของครัวเรือนในแต่ละภูมิภาค. วารสารเศรษฐศาสตร์และนโยบายสาธารณะ, 11(21), 45-62.

[6] อริสิริพิมพ์ บุญเพิ่มพูน. (2564). พฤติกรรมการออมเพื่อวัยเกษียณของประชากรกลุ่ม Generation Y ในเขตกรุงเทพมหานคร [วิทยานิพนธ์ปริญญามหาบัณฑิต, มหาวิทยาลัยธรรมศาสตร์]. TU Digital Collections.

[7] สุนิษา วัฒนะ. (2564). การใช้แอปพลิเคชันด้านการเงินกับพฤติกรรมการออมของประชากรวัย ทำงานในเขตกรุงเทพมหานคร. วารสารวิชาการนวัตกรรมและการจัดการ, 6(2), 45-58.

[8] ณัฐวัฒน์ หาญกิจ. (2564). การวิเคราะห์ข้อมูลรายได้และรายจ่ายของครัวเรือนไทยด้วย โปรแกรม Power BI: กรณีศึกษาข้อมูลจากสำนักงานสถิติแห่งชาติ. วารสารเทคโนโลยีสารสนเทศและการจัดการ, 8(1), 12-25

[9] รัชนี กัลยาเลิศ และ นันทวรรณ บุญช่วย. (2563). ปัจจัยที่ส่งผลต่อการวางแผนการเงินส่วน บุคคลของลูกค้าธนาคารธนชาต จำกัด (มหาชน) ในเขตพื้นที่กรุงเทพมหานครและปริมณฑล. วารสารบริหารธุรกิจและสังคมศาสตร์มหาวิทยาลัยรามคำแหง, 3(2), 85-99.

[10] สิทธา อุปนิกขิต. (2564). การแสวงหาข่าวสารและการใช้ประโยชน์ในการออมเงินของประชาชนในจังหวัดนนทบุรี. วารสารการสื่อสารมวลชน, 9(1), 110-125. (เผยแพร่ผ่านระบบ ThaiJO).

[11] Alabi, A., et al. (2024). Predictive analytics for household savings trends: A machine learning approach to financial forecasting. International Journal of Financial Studies, 12(1), 45-60.

[12] Lopez, R., et al. (2021). Income-based saving behavior study: Global cross-sectional evidence from the Findex database. Journal of Financial Counseling and Planning, 32(2), 210-225.

[13] Khatuwal, S., & Puri, R. (2022). Visualizing gender gaps in saving patterns: Data-driven insights for financial inclusion. [Research Paper]. Global Economic Review.

[14] Dicarlo, M., & Niño, J. (2023). Financial literacy and BI dashboard interventions: Empowering households through interactive data tools. Technology in Society, 74, 102-118.

[15] Hernandez, L., et al. (2023). Regional disparities in financial inclusion: Evidence from emerging economies. World Development Perspectives, 29, Article 100482.

[16] Kim, H., & Lee, J. (2021). Using Power BI in municipal budget transparency: A case study of open government data initiatives. Government Information Quarterly, 38(4), 101-115.

[17] Konkid, P., & Junlabuddee, S. (2025). Power BI-based monitoring of financial planning behavior: A case study of working-age population in Thailand. Journal of Applied Information Technology, 11(1), 88-103

[18] Nguyen, T., & Tran, M. (2022). Household budget visualization using BI tools: Improving financial awareness in urban areas. Journal of Digital Economy and Sustainable Development, 4(2), 120-135.

[19] Sharma, R., & Joshi, P. (2023). BI for visualizing public financial aid: Enhancing transparency in social welfare distribution. International Journal of Public Administration and Data Science, 9(3), 56-72.

[20] Wongsrikaew, K., & Sunalai, P. (2023). Adoption of BI dashboards for community banking: A study on user engagement and financial literacy. Asian Journal of Information Systems, 15(1), 210-226.

# การพยากรณ์ยอดขายสินค้ากลุ่มผลิตภัณฑ์ดูแลผ้า ผ่านช่องทางพาณิชย์อิเล็กทรอนิกส์
# Forecasting sales of fabric care products through the E-Commerce channel

*ณัฐพร ก้องสกุลไกร (Nattaporn Kongsakunkrai)[1] และ ศักดิ์ชาย ตั้งวรรณวิทย์ (Sakchai Tangwannawit)[2]*
*ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล*
*มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ*
*[1]s6707021857171@email.kmutnb.ac.th, [2] sakchai.t@itd.kmutnb.ac.th*

**บทคัดย่อ**

*งานวิจัยนี้มีวัตถุประสงค์เพื่อศึกษาประสิทธิภาพของแบบจำลองการพยากรณ์ยอดขายสินค้าอุปโภคบริโภคบนแพลตฟอร์ม E-Commerce โดยใช้ข้อมูลยอดขายรายวันในระดับราย SKU และเปรียบเทียบแบบจำลอง Prophet, ARIMA และ SARIMA ภายใต้บริบทของข้อมูลที่มีความผันผวนสูงและได้รับอิทธิพลจากกิจกรรมส่งเสริมการขาย การศึกษาใช้ตัวชี้วัดความแม่นยำ ได้แก่ MAPE และ RMSE ในการประเมินผล ผลการศึกษาพบว่า Prophet ให้ประสิทธิภาพการพยากรณ์ดีที่สุดในภาพรวม โดยให้ค่าความคลาดเคลื่อนต่ำที่สุดใน 5 จาก 6 SKU ในขณะที่ ARIMA ให้ผลดีกว่าในบาง SKU ที่มีรูปแบบยอดขายค่อนข้างเสถียร และ SARIMA ไม่แสดงความได้เปรียบอย่างชัดเจน นอกจากนี้ การกำหนด Promotion Flags และการปรับพารามิเตอร์ของ Prophet ได้แก่ seasonality prior scale และ changepoint prior scale ช่วยลดค่า MAPE และ RMSE ได้อย่างสม่ำเสมอเมื่อเทียบกับการใช้ค่าเริ่มต้นของโมเดล โดยเฉพาะการตั้งค่า seasonality prior scale ในระดับต่ำและการเพิ่มความยืดหยุ่นของ changepoints ซึ่งช่วยให้โมเดลสามารถอธิบายการเปลี่ยนแปลงของแนวโน้มยอดขายที่เกิดจากกิจกรรมทางการตลาดได้ดียิ่งขึ้น สรุปได้ว่าการพยากรณ์ยอดขายบน E-Commerce ควรใช้แนวทางการเลือกแบบจำลองที่เหมาะสมในระดับราย SKU ร่วมกับการออกแบบตัวแปรอธิบายและการปรับแต่งพารามิเตอร์อย่างเป็นระบบ ซึ่งจะช่วยเพิ่มความแม่นยำของการพยากรณ์และสนับสนุนการตัดสินใจด้านการวางแผนความต้องการสินค้าและการบริหารสินค้าคงคลังได้อย่างมีประสิทธิภาพ*

**คำสำคัญ:** การพยากรณ์ยอดขาย พาณิชย์อิเล็กทรอนิกส์ Prophet ARIMA SARIMA โปรโมชั่น

**Abstract**

*This study aims to evaluate the performance of sales forecasting models for consumer goods in an E-Commerce context using daily sales data at the SKU level. The models compared in this study include Prophet, ARIMA, and SARIMA, under conditions characterized by high demand volatility and strong promotional effects. Forecasting accuracy was assessed using MAPE and RMSE. The results indicate that Prophet outperforms the benchmark models in overall performance, achieving the lowest forecasting errors in five out of six SKUs. ARIMA demonstrates superior performance in specific SKUs with relatively stable demand patterns, while SARIMA does not exhibit a clear advantage in the dataset examined. Furthermore, the inclusion of promotion flags and systematic tuning of Prophet hyperparameters—particularly the seasonality prior scale and changepoint prior scale—significantly improves forecasting accuracy compared to the default model configuration. Lower seasonality prior values combined with more flexible changepoint settings enable the model to better capture structural changes in sales trends driven by promotional activities. In conclusion, the findings suggest that sales forecasting in E-Commerce environments benefits from a SKU-level model selection approach combined with effective feature engineering and parameter tuning. Prophet,*

*when appropriately configured, serves as a robust primary forecasting model for consumer goods, supporting more accurate demand planning and inventory management decisions in practice.*

**Keywords:** Sales Forecasting, E-Commerce, Prophet, ARIMA, SARIMA, Promotional Effects.

## 1. บทนำ

การวางแผนและควบคุมการผลิตมีบทบาทสำคัญต่อความสามารถในการแข่งขันขององค์กร เนื่องจากช่วยให้การใช้ทรัพยากรเป็นไปอย่างมีประสิทธิภาพ ลดความล่าช้าในกระบวนการผลิต ควบคุมต้นทุนการผลิตและต้นทุนสินค้าคงคลัง รวมถึงสนับสนุนการตัดสินใจของฝ่ายบริหารให้สามารถตอบสนองต่อความต้องการของลูกค้าได้อย่างทันท่วงที [1] ขณะเดียวกันการเติบโตของการค้าผ่านช่องทางพาณิชย์อิเล็กทรอนิกส์ (E-Commerce) ทั้งในรูปแบบ Marketplace และ Social Commerce ทำให้ความผันผวนของยอดขายเพิ่มสูงขึ้น ส่งผลให้การพยากรณ์ยอดขายมีความสำคัญมากยิ่งขึ้น

ปัจจุบันมีแบบจำลองการพยากรณ์หลากหลายรูปแบบ เช่น ARIMA ซึ่งอาศัยแนวคิดของตัวแบบสโตแคสติกที่มีโครงสร้างสหสัมพันธ์ในตัวเอง SARIMA ซึ่งพัฒนาจาก ARIMA เพื่อรองรับข้อมูลที่มีฤดูกาล และ Prophet ซึ่งเป็นเครื่องมือพยากรณ์อนุกรมเวลาที่พัฒนาโดย Facebook ที่เหมาะสมกับข้อมูลที่มีแนวโน้ม ฤดูกาล และเหตุการณ์พิเศษประกอบกัน [1]

ดังนั้น งานวิจัยนี้จึงมีวัตถุประสงค์เพื่อศึกษาและเปรียบเทียบประสิทธิภาพของแบบจำลองการพยากรณ์ ได้แก่ ARIMA, SARIMA และ Prophet สำหรับการพยากรณ์ยอดขายสินค้าผ่านช่องทาง E-Commerce โดยมุ่งหวังนำแบบจำลองที่มีความแม่นยำสูงสุดไปใช้เป็นเครื่องมือสนับสนุนการวางแผนการจัดซื้อวัตถุดิบและการผลิตสินค้า ช่วยเพิ่มประสิทธิภาพในการบริหารต้นทุนและเป็นแนวทางให้ผู้ประกอบการสามารถนำไปประยุกต์ใช้กับข้อมูลอื่น ๆ ได้ต่อไป

## 2. เอกสารและงานวิจัยที่เกี่ยวข้อง

### 2.1 แนวคิดและทฤษฎีที่เกี่ยวข้อง

การพยากรณ์เป็นการคาดการณ์เหตุการณ์หรือผลลัพธ์ที่อาจเกิดขึ้นในอนาคต ซึ่งไม่สามารถควบคุมได้โดยตรง เป็นข้อมูลประกอบการตัดสินใจเพื่อช่วยให้ผู้บริหารสามารถจัดการอนาคตได้อย่างเหมาะสม และนำไปใช้ในการวางแผนเพื่อให้องค์กรบรรลุเป้าหมายที่ตั้งไว้ [2]

อนุกรมเวลา หมายถึง การเก็บรวบรวมข้อมูลของตัวแปรหนึ่งตามลำดับเวลา หลังจากข้อมูลของตัวแปรที่สนใจถูกเก็บรวบรวมมาแล้ว ค่าสถิติเบื้องต้น ได้แก่ ค่าเฉลี่ย (Mean) ความแปรปรวน (Variance) และความแปรปรวนร่วมระหว่างช่วงเวลา (Autocovariance) จะต้องสามารถคำนวณได้เสมอ อนุกรมเวลาของตัวแปรหนึ่ง จะประกอบไปด้วย 4 ส่วน คือ แนวโน้ม วัฏจักร ความผันแปรจากฤดูกาล และความผันผวนจากเหตุการณ์ไม่ปกติ โดยสามารถประเมินความแม่นยำในการพยากรณ์ได้หลายวิธี เช่น Mean Absolute Percentage Error (MAPE) หรือ Root Mean Square Error (RMSE) เป็นต้น

Prophet เป็นเครื่องมือสำหรับการพยากรณ์ข้อมูลแบบอนุกรมเวลา สามารถใช้งานได้ทั้งบน R และ Python ซึ่งถูกพัฒนาโดยกลุ่มนักวิทยาศาสตร์ข้อมูลภายในองค์กร Facebook (ที่รู้จักในปัจจุบันว่า Meta) ได้นำมาใช้ในงานพยากรณ์ภายใน Facebook ในหลายโครงการ และได้ทำการปล่อยให้ใช้งานกันแบบสาธารณะในปี 2017 อัลกอริทึมของ Prophet จะทำงานได้ดีกับข้อมูลที่มีหลายฤดูกาล มีปริมาณข้อมูลมาก และอธิบายคุณลักษณะของฤดูกาลของข้อมูลได้ [3] ซึ่งคุณลักษณะเด่น ๆ ของ Prophet ได้แก่ มีความยืดหยุ่น สามารถจำแนกแนวโน้มและฤดูกาล วิเคราะห์ผลกระทบจากวันหยุด และการแสดงผลการวิเคราะห์ข้อมูลในรูปแบบกราฟต่างๆ

### 2.2 งานวิจัยที่เกี่ยวข้อง

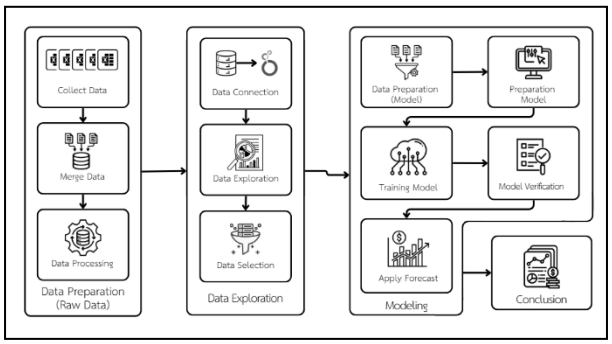
งานวิจัยที่เกี่ยวข้องแสดงให้เห็นถึงการนำแบบจำลองการพยากรณ์อนุกรมเวลาหลากหลายรูปแบบ เช่น ARIMA, SARIMA, Prophet และ Deep Learning ไปประยุกต์ใช้ในหลายบริบท ทั้งด้านการเงิน พลังงาน สุขภาพ และธุรกิจ E-Commerce โดยผลการศึกษาส่วน

ใหญ่พบว่า Prophet มีประสิทธิภาพที่ดีในการพยากรณ์ข้อมูลที่มีแนวโน้มและฤดูกาลชัดเจน โดยเฉพาะข้อมูลรายวันและรายสัปดาห์ [4] [5] และในหลายกรณี Prophet ให้ผลลัพธ์เหนือกว่า SARIMA และ ARIMA ในด้านความแม่นยำ เช่น การพยากรณ์ยอดขายสินค้า [6] การพยากรณ์ความต้องการพลังงาน [7] และการเข้าชมเว็บไซต์ E-Commerce [8]

อย่างไรก็ตาม งานวิจัยบางส่วนชี้ให้เห็นข้อจำกัดของ Prophet ในการรับมือกับเหตุการณ์ไม่ปกติหรือความผันผวนรุนแรง เช่น วิกฤต COVID-19 หรือความขัดแย้งทางภูมิรัฐศาสตร์ [9] รวมถึงในบางกรณี ARIMA หรือ SARIMA ให้ผลลัพธ์ที่ดีกว่าสำหรับข้อมูลรายเดือนหรือบริบทเฉพาะทาง [4] [10] นอกจากนี้ งานวิจัยจำนวนมากเสนอว่าแบบจำลองแบบผสมหรือ Hybrid Model เช่น SARIMA-Prophet, LSTM-Prophet หรือ NeuralProphet สามารถลดค่าความคลาดเคลื่อนและจัดการทั้งแนวโน้มระยะยาวและความผันผวนระยะสั้นได้ดีกว่าแบบจำลองเดี่ยว [11] [12] [13]

โดยสรุป วรรณกรรมที่ผ่านมาแสดงให้เห็นว่า Prophet เป็นเครื่องมือพยากรณ์ที่มีความยืดหยุ่น ใช้งานง่าย และให้ความแม่นยำสูงในหลายบริบท อย่างไรก็ตาม ประสิทธิภาพของแบบจำลองยังขึ้นอยู่กับลักษณะข้อมูล ความถี่ของข้อมูล และการมีเหตุการณ์ผิดปกติ ซึ่งสนับสนุนแนวคิดของงานวิจัยนี้ในการเปรียบเทียบ Prophet กับ ARIMA และ SARIMA เพื่อค้นหาแบบจำลองที่เหมาะสมที่สุดสำหรับการพยากรณ์ยอดขายในบริบทของ E-Commerce

## 3. วิธีดำเนินการวิจัย




**ภาพที่ 1:** Conceptual Framework

งานวิจัยนี้ใช้ข้อมูลยอดขายสินค้าผ่านช่องทาง E-Commerce จากแพลตฟอร์ม Shopee ซึ่งอยู่ในรูปแบบไฟล์ Microsoft Excel โดยข้อมูลถูกจัดเก็บแยกเป็นรายเดือน ผู้วิจัยได้ทำการรวมไฟล์ทั้งหมดและเตรียมข้อมูลดิบด้วยโปรแกรม Microsoft Excel ก่อนนำเข้าสู่กระบวนการทำความสะอาดข้อมูล (Data Cleansing) ซึ่งประกอบด้วยการตรวจสอบและแก้ไขรหัสเฉพาะที่ใช้ระบุสินค้าแต่ละรายการ (Stock Keeping Unit: SKU) ที่ไม่ถูกต้อง และการแปลงข้อมูลสินค้าชุด ให้อยู่ในรูปของสินค้าเดี่ยว เพื่อให้ข้อมูลครอบคลุมทุก SKU อย่างถูกต้องและสอดคล้องกัน จากนั้นจัดรูปแบบข้อมูลให้อยู่ในลักษณะที่พร้อมต่อการวิเคราะห์และสร้างแบบจำลอง

ในขั้นตอนการสำรวจข้อมูล ผู้วิจัยได้โหลดข้อมูลไปยัง Google Drive และแปลงเป็น Google Sheet เพื่อเชื่อมต่อกับ Looker Studio สำหรับการแสดงผลและสำรวจแนวโน้มของข้อมูลยอดขายในมิติต่าง ๆ ขั้นตอนนี้ช่วยให้สามารถตรวจสอบความสมบูรณ์ของข้อมูล วิเคราะห์พฤติกรรมยอดขายเบื้องต้น และคัดเลือกข้อมูลที่เหมาะสมสำหรับการสร้างแบบจำลองการพยากรณ์

สำหรับการเตรียมข้อมูลเพื่อสร้างแบบจำลอง ข้อมูลถูกนำเข้าไปยัง Google Colab ด้วยภาษา Python โดยกำหนดตัวแปรหลักที่ใช้ในการวิเคราะห์ ได้แก่ วันที่ขาย รหัสสินค้า และจำนวนขาย พร้อมทั้งจัดการข้อมูลที่ขาดหาย กำหนดรูปแบบข้อมูล และรวมข้อมูลตามรหัสสินค้าและวันที่เดียวกัน จากนั้นคัดเลือกเฉพาะ SKU ที่มีช่วงข้อมูลต่อเนื่องไม่น้อยกว่า 36 เดือน เพื่อให้เหมาะสมกับการสร้างแบบจำลองอนุกรมเวลา

การสร้างแบบจำลองการพยากรณ์ใช้แบบจำลอง Prophet, ARIMA และ SARIMA โดยในกรณีของ Prophet มีการกำหนดตัวแปร Promotion Flags จำนวน 2 ชุด เพื่อสะท้อนรูปแบบโปรโมชั่นที่แตกต่างกัน พร้อมทั้งกำหนดค่าพารามิเตอร์ที่เกี่ยวข้อง ขณะที่แบบจำลอง ARIMA และ SARIMA ใช้ค่าพารามิเตอร์ที่กำหนดไว้ล่วงหน้า การประเมินประสิทธิภาพของแบบจำลองใช้ตัวชี้วัดความแม่นยำ ด้วยค่า MAPE และ RMSE

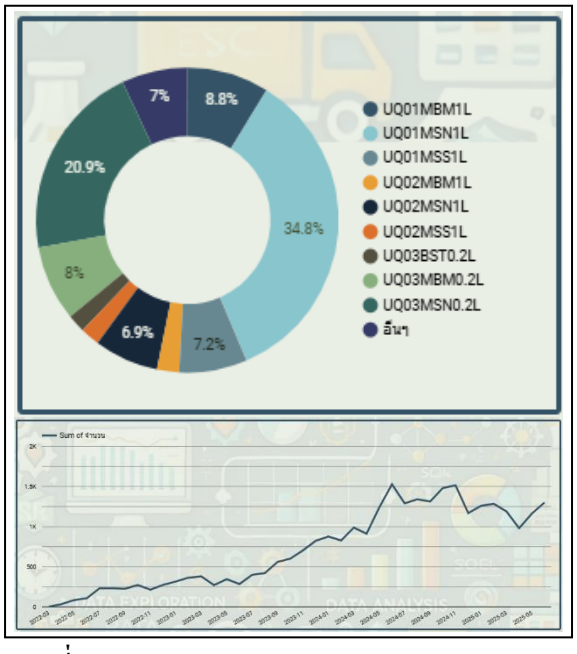
## 4. ผลการวิจัย

### 4.1 การสำรวจ และเตรียมข้อมูล

จากการตรวจสอบไฟล์ข้อมูลต้นฉบับ พบว่าข้อมูลยอดขายประกอบด้วยตัวแปรจำนวนทั้งสิ้น 37 คอลัมน์ ซึ่งครอบคลุมรายละเอียดด้านการสั่งซื้อ ราคา ค่าธรรมเนียม โปรโมชั่น และข้อมูลผู้ซื้อ อย่างไรก็ตาม ในการพยากรณ์ยอดขาย ผู้วิจัยได้คัดเลือกเฉพาะตัวแปรที่มีความจำเป็นต่อการสร้างแบบจำลองอนุกรมเวลา ได้แก่ วันที่ทำการสั่งซื้อ เลขที่อ้างอิงสินค้า (SKU) และจำนวนขาย

การสำรวจข้อมูลเชิงโครงสร้างพบมีการกำหนด SKU ในรูปแบบชื่อสินค้าและมีการเปลี่ยนแปลงไปมา รวมถึงการกำหนด SKU บางรายการอยู่ในรูปของสินค้าแบบชุด ซึ่งประกอบด้วยสินค้าเดี่ยวหลายรายการ ผู้วิจัยจึงทำการปรับปรุง SKU ให้อยู่ในรูปแบบที่สามารถเป็นตัวแทนของสินค้าได้

เมื่อพิจารณาระยะเวลาที่มียอดขายของสินค้าทั้งหมด พบว่ามีเพียง 6 SKU ที่มีข้อมูลต่อเนื่องมากกว่า 36 เดือน ซึ่งเป็นช่วงเวลาที่เหมาะสมต่อการสร้างแบบจำลองพยากรณ์ยอดขาย ได้แก่ UQ01MSM1L, UQ01MSN1L, UQ01MSS1L, UQ03MBM0.3L, UQ03MSN0.3L และ UQ03MSS0.3L จึงทำการคัดเลือกเฉพาะสินค้าเหล่านี้มาใช้ในการศึกษาต่อไป




**ภาพที่ 2:** ผลการวิเคราะห์ยอดขายตามรายการสินค้า และช่วงเวลา

การวิเคราะห์สัดส่วนยอดขายพบว่า สินค้า SKU UQ01MSN1L มีสัดส่วนการขายสูงที่สุดเมื่อเทียบกับสินค้าอื่น และภาพรวมของยอดขายทั้งหมดมีปริมาณสูงสุดในช่วงปี ค.ศ. 2024 ด้วยเหตุนี้ ผู้วิจัยจึงเลือกสินค้า UQ01MSN1L และข้อมูลในช่วงปี ค.ศ. 2024 เป็นกรณีศึกษาในการสำรวจรูปแบบโปรโมชั่นที่มีผลต่อยอดขายรายวัน ร่วมกับรูปแบบการจัดโปรโมชั่นของแพลตฟอร์ม Shopee

จากผลการสำรวจดังกล่าว ผู้วิจัยได้กำหนดตัวแปร Promotion Flags จำนวน 2 รูปแบบ สำหรับการพยากรณ์ด้วย Prophet โดยครอบคลุมโปรโมชั่นช่วงกลางเดือน โปรโมชั่นวันเงินเดือนออก โปรโมชั่น Double Date และโปรโมชั่นเฉพาะของช่วงเดือนธันวาคม พร้อมทั้งทำการปรับค่าพารามิเตอร์ของโมเดลในหลายระดับ ได้แก่ จำนวน Changepoints และค่า Seasonality prior เพื่อศึกษาผลกระทบของการตั้งค่าพารามิเตอร์ต่อความแม่นยำในการพยากรณ์ยอดขายรายวันของสินค้าอุปโภคบริโภค

**ตารางที่ 1:** การกำหนด Promotion Flags

| Promotions Flags | Set 1 | Set 2 |
|---|---|---|
| Double Date | ≤ 3 | x.x (ยกเว้น 12.12) |
| Mid Month | 13 – 17 | 15 - 17 |
| Pay Day | ≥ 25 | 25 – 27 |
| Dec Mega Sale | - | 12/12 – 15/12 |

### 4.2 การสร้างโมเดล และพยากรณ์ยอดขาย

ผลการศึกษาการสร้างแบบจำลองพยากรณ์ยอดขายโดยเปรียบเทียบแบบจำลอง Prophet, ARIMA และ SARIMA สำหรับสินค้า 6 SKU โดยมีระยะเวลาการพยากรณ์ 60 วัน พบว่า แบบจำลอง Prophet ให้ประสิทธิภาพโดยรวมดีที่สุด โดยสามารถให้ค่าความคลาดเคลื่อนต่ำที่สุดใน 5 จาก 6 SKU เมื่อพิจารณาจากตัวชี้วัด MAPE และ RMSE ขณะที่ ARIMA และ SARIMA ให้ผลลัพธ์ที่แตกต่างกันไปตามลักษณะของข้อมูลแต่ละรายการสินค้า แสดงให้เห็นว่าความเหมาะสมของแบบจำลองขึ้นอยู่กับรูปแบบอนุกรมเวลาของแต่ละ SKU

ในภาพรวม Prophet สามารถอธิบายโครงสร้างของข้อมูลยอดขายที่มีแนวโน้มและความผันผวนได้ดีกว่าแบบจำลองดั้งเดิม โดยเฉพาะสินค้าที่มีรูปแบบยอดขายไม่

สม่ำเสมอและได้รับอิทธิพลจากโปรโมชั่นอย่างชัดเจน ทั้งนี้ Prophet ให้ค่า RMSE ต่ำที่สุดในทุก SKU สะท้อนถึงความสามารถในการลดความคลาดเคลื่อนเชิงปริมาณ ซึ่งมีความสำคัญต่อการวางแผนการผลิตและการจัดซื้อในเชิงปฏิบัติ

อย่างไรก็ตาม พบว่า ARIMA และ SARIMA ให้ผลลัพธ์ที่ดีกว่าในบางกรณี ได้แก่ SKU UQ03MSS0.2L ซึ่ง ARIMA ให้ค่า MAPE ต่ำที่สุด และ SKU UQ01MBM1L ซึ่ง SARIMA ให้ค่า MAPE ต่ำที่สุด แสดงให้เห็นว่าแบบจำลองเชิงสถิติแบบดั้งเดิมยังคงมีความสามารถในการพยากรณ์ในบางรูปแบบของข้อมูล โดยเฉพาะกรณีที่มีโครงสร้างฤดูกาลชัดเจนหรือมีรูปแบบคงที่

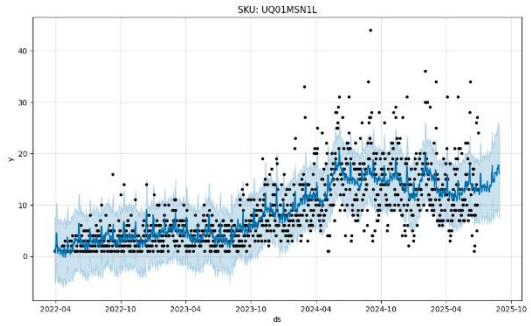
**ตารางที่ 2:** เปรียบเทียบประสิทธิภาพการพยากรณ์ของแต่ละโมเดล

| SKU | Prophet | | ARIMA | | SARIMA | |
|---|---|---|---|---|---|---|
| | MAPE | RMSE | MAPE | RMSE | MAPE | RMSE |
| UQ01MBM1L | 68.99 | 1.94 | 73.65 | 2.90 | 65.96 | 3.00 |
| UQ01MSN1L | 66.88 | 4.67 | 86.59 | 7.13 | 90.20 | 7.93 |
| UQ01MSS1L | 64.34 | 1.69 | 92.68 | 2.20 | 87.05 | 2.29 |
| UQ03MBM0.2L | 68.17 | 1.89 | 107.04 | 2.40 | 86.36 | 2.58 |
| UQ03MSN0.2L | 66.28 | 3.21 | 75.74 | 4.96 | 68.46 | 5.68 |
| UQ03MSS0.2L | 65.07 | 1.75 | 47.35 | 2.23 | 57.87 | 2.16 |

สำหรับการทดลองปรับค่าพารามิเตอร์ และ Promotion Flags ของ Prophet พบว่าการเพิ่มความยืดหยุ่นของ Trend (Changepoint) และการควบคุมความเรียบของฤดูกาล (Seasonality Prior) ช่วยเพิ่มความแม่นยำของการพยากรณ์อย่างสม่ำเสมอในทุก SKU เมื่อเทียบกับค่าเริ่มต้นของโมเดล โดยค่าพารามิเตอร์ Changepoint = 30 และ Seasonality prior = 0.05 ให้ผลลัพธ์ดีที่สุดหรือใกล้เคียงดีที่สุดในเกือบทุกกรณี

เมื่อเปรียบเทียบรูปแบบ Promotion Flags พบว่า Set 1 ให้ผลการพยากรณ์ที่แม่นยำกว่า Set 2 เล็กน้อย โดยเฉพาะใน SKU ที่มีความผันผวนสูง สะท้อนว่าการกำหนดตัวแปรโปรโมชั่นที่สอดคล้องกับพฤติกรรมการซื้อจริงของลูกค้า สามารถช่วยเพิ่มความสามารถของ Prophet ในการอธิบายความเปลี่ยนแปลงของยอดขายได้อย่างมีนัยสำคัญ

โดยสรุป แบบจำลอง Prophet ที่มีการปรับพารามิเตอร์และกำหนด Promotion Flags อย่างเหมาะสม เป็นแบบจำลองที่มีศักยภาพสูงในการพยากรณ์ยอดขายสินค้าผ่านช่องทาง E-Commerce เนื่องจากสามารถลดค่าความคลาดเคลื่อนทั้งเชิงสัดส่วนและเชิงปริมาณได้อย่างมีประสิทธิภาพ และสามารถนำไปประยุกต์ใช้ในการวางแผนการผลิตและบริหารสินค้าคงคลังในเชิงธุรกิจได้




**ภาพที่ 3:** ตัวอย่างกราฟการพยากรณ์จาก Prophet

## 5. สรุป

งานวิจัยนี้มีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพของแบบจำลองการพยากรณ์ยอดขายสินค้าอุปโภคบริโภคบนแพลตฟอร์ม E-Commerce ได้แก่ ARIMA, SARIMA และ Prophet โดยใช้ข้อมูลยอดขายรายวันในระดับราย SKU ผลการศึกษาพบว่า Prophet ให้ประสิทธิภาพดีที่สุดในภาพรวม โดยให้ค่าความคลาดเคลื่อนต่ำที่สุดใน 5 จาก 6 SKU เมื่อประเมินด้วย MAPE และ RMSE ขณะที่ ARIMA และ SARIMA ให้ผลลัพธ์ที่เหมาะสมเฉพาะบาง SKU ที่มีรูปแบบอุปสงค์ค่อนข้างเสถียร นอกจากนี้ การปรับพารามิเตอร์ของ Prophet และการกำหนด Promotion Flags ที่สะท้อนพฤติกรรมการซื้อจริง สามารถช่วยลดค่าความคลาดเคลื่อนของการพยากรณ์ได้อย่างมีนัยสำคัญ สะท้อนให้เห็นว่าการเลือกและปรับแต่งแบบจำลองในระดับราย SKU เป็นแนวทางที่เหมาะสมสำหรับการใช้งานจริงในธุรกิจ E-Commerce ซึ่งสอดคล้องกับผลการศึกษาที่พบว่า Prophet มีประสิทธิภาพเหนือกว่า SARIMA ในการพยากรณ์ยอดขายสินค้า [6] หรือ การพยากรณ์ราคา Bitcoin ที่ Prophet ก็ให้ผลลัพธ์ที่ดีกว่า ARIMA ในข้อมูลรายวันและรายสัปดาห์ [4]

ผลการทดลองชี้ให้เห็นว่า Prophet มีความเหมาะสมกับข้อมูลยอดขายบน E-Commerce ที่มีความผันผวนสูง

และได้รับอิทธิพลจากกิจกรรมส่งเสริมการขาย เนื่องจากโครงสร้างของโมเดลสามารถแยกองค์ประกอบของแนวโน้ม ฤดูกาล และผลกระทบจากเหตุการณ์พิเศษได้อย่างชัดเจน ในขณะที่ ARIMA และ SARIMA มีข้อจำกัดจากสมมติฐานเชิงเส้นและความนิ่งของข้อมูล อย่างไรก็ตาม แม้ในภาพรวม Prophet จะให้ผลลัพธ์ที่ดีที่สุด แต่ ARIMA และ SARIMA ก็ยังคงมีศักยภาพในบาง SKU ที่มีรูปแบบยอดขายค่อนข้างเสถียร ผลดังกล่าวสนับสนุนแนวคิดว่าการใช้โมเดลเดียวกับทุกสินค้าอาจนำไปสู่ความคลาดเคลื่อนที่ไม่จำเป็น และการเลือกแบบจำลองที่เหมาะสมราย SKU เป็นแนวทางที่เหมาะสมกว่าในบริบทของการพยากรณ์เชิงธุรกิจ ผลการศึกษายังยืนยันว่าการใช้แบบจำลองเดียวกับทุก SKU อาจไม่เหมาะสม และการเลือกแบบจำลองให้สอดคล้องกับลักษณะข้อมูลรายสินค้า รวมถึงการออกแบบตัวแปรและการปรับพารามิเตอร์ของโมเดล มีบทบาทสำคัญต่อความแม่นยำของการพยากรณ์มากกว่าการพึ่งพาโครงสร้างโมเดลเพียงอย่างเดียว

งานวิจัยในอนาคตควรมุ่งเน้นการบูรณาการตัวแปรภายนอกเพิ่มเติม เช่น ราคา ส่วนลด กิจกรรมทางการตลาด หรือปัจจัยเศรษฐกิจ เพื่อเพิ่มความสามารถในการอธิบายความผันผวนของยอดขาย ควบคู่กับการพัฒนาแบบจำลองเชิงผสมหรือแบบจำลองเชิงเรียนรู้ขั้นสูง เพื่อรองรับความสัมพันธ์ที่ไม่เป็นเชิงเส้น นอกจากนี้ ควรมีการประเมินแบบจำลองในหลายระดับช่วงเวลาและขยายขอบเขตการศึกษาไปยังสินค้าและแพลตฟอร์ม E-Commerce อื่น ๆ เพื่อเพิ่มความสามารถในการอ้างอิงทั่วไปและประโยชน์เชิงปฏิบัติของผลการวิจัยในบริบททางธุรกิจที่หลากหลายมากยิ่งขึ้น

## เอกสารอ้างอิง


[1] พิภพ ลลิตาภรณ์. (2553). *ระบบการวางแผนและควบคุมการผลิต*. สำนักพิมพ์ ส.ส.ภ.

[2] อัจฉรา จันทร์ฉาย. (2557). *เทคนิคการพยากรณ์เพื่อการจัดการ*. จุฬาลงกรณ์มหาวิทยาลัย.

[3] Rafferty, G. (2025). *Forecasting Time Series Data with Facebook Prophet*. Packt.

[4] Angelo, M. D., Fadhiilrahman, I., & Purnama, Y. (2023). Comparative Analysis of ARIMA and Prophet Algorithms in Bitcoin Price Forecasting. *Procedia Computer Science*, 227, 490–499.

[5] Chaturvedi, S., Rajasekar, E., Natarajan, S., & McCullen, N. (2022). A comparative assessment of SARIMA, LSTM RNN and Fb Prophet models to forecast total and peak monthly energy demand for India. *Energy Policy*, 168(113097).

[6] Negre, P., Alonso, R. S., Prieto, J., Garcia, O., & de-la-Fuente-Valentin, L. (2024). Prediction of footwear demand using Prophet and SARIMA. *Expert Systems with Applications*, 255(124512).

[7] Ghosh, I., & Jana, R. K. (2024). Clean energy stock price forecasting and response to macroeconomic variables: A novel framework using Facebook's Prophet, NeuralProphet and explainable AI. *Technological Forecasting and Social Change*, 200(123148).

[8] สิทธิ โง้ววัฒนา และ เอกสิทธิ์ พัชรวงศ์ศักดา. (2566). การพยากรณ์ปริมาณการเข้าชมเว็บไซต์สำหรับธุรกิจอีคอมเมิร์ซโดยใช้เทคนิคการเรียนรู้ของเครื่องและสร้างระบบแจ้งเตือนโดยอัตโนมัติ. *วารสารบัณฑิตศึกษา มหาวิทยาลัยธุรกิจบัณฑิตย์*, 3.

[9] Cheng, J., Tiwari, S., Khaled, D., Mahendru, M., & Shahzad, U. (2024). Forecasting Bitcoin prices using artificial intelligence: Combination of ML, SARIMA, and Facebook Prophet models. *Technological Forecasting and Social Change*, 198(122938).

[10] Fuller, B., Ssekitoleko, R., Kyozira, C., Colston, J. M., Makumbi, I., Bakainaga, A., Moore, C. C., & Kiirya, H. I. (2025). Comparison of ARIMA and FB-Prophet time series models for the prediction of national and regional malaria incidence in Uganda. *Malaria Journal*, 24(1), 289.

[11] Kenyi, M. G. S., & Yamamoto, K. (2024). A hybrid SARIMA Prophet model for predicting historical streamflow time series of the Sobat River in South Sudan. *Discover Applied Sciences*, 6(457).

[12] Albahli, S. (2025). LSTM vs. Prophet: Achieving Superior Accuracy in Dynamic Electricity Demand Forecasting. *Energies*, 18, 278.

[13] Long, B., Tan, F., & Newman, M. (2023). Forecasting the Monkeypox Outbreak Using ARIMA, Prophet, NeuralProphet, and LSTM Models in the United States. *Forecasting*, 5, 127–137.

# ปัจจัยที่มีอิทธิพลต่อการใช้งานปัญญาประดิษฐ์เชิงสร้างสรรค์อย่างต่อเนื่อง
# Factors Influencing Continued Use of Generative AI

*เมตตา นักธรรม (Metta Naktam)[1] และธนพล เจนสุทธิเวชกุล (Tanapon Jensuttiwetchakul)[2]*
*ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล*
*มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ*
*[1]s6707021856035@email.kmutnb.ac.th*, [2] tanapon.j@itd.kmutnb.ac.th

## บทคัดย่อ

*การวิจัยนี้มีวัตถุประสงค์เพื่อศึกษาปัจจัยที่มีอิทธิพลต่อการใช้งานปัญญาประดิษฐ์เชิงสร้างสรรค์อย่างต่อเนื่อง กรณีศึกษา ChatGPT โดยบูรณาการแบบจำลองการยืนยันความคาดหวัง ทฤษฎีความเป็นส่วนตัวและแนวคิดด้านลักษณะของ AI รวบรวมข้อมูลจากกลุ่มตัวอย่างผ่านแบบสอบถามออนไลน์ และวิเคราะห์ข้อมูลจากกลุ่มตัวอย่างจำนวน 312 คน ด้วยการวิเคราะห์องค์ประกอบเชิงสำรวจ (EFA) และทดสอบสมมติฐานการวิเคราะห์การถดถอยเชิงพหุคูณ (MRA) จากผลการศึกษาพบว่า ตัวแปรที่มีอิทธิพลทางตรงต่อความตั้งใจใช้งานปัญญาประดิษฐ์เชิงสร้างสรรค์อย่างต่อเนื่อง กรณีศึกษา ChatGPT ได้แก่ ความพึงพอใจ ($\beta = 0.489$,, p-value = 0.001) การรับรู้ถึงประโยชน์ ($\beta = 0.383$, p-value = 0.001) โดยข้อกังวลเรื่องความเป็นส่วนตัวไม่มีอิทธิพลต่อความตั้งใจใช้งานปัญญาประดิษฐ์เชิงสร้างสรรค์อย่างต่อเนื่อง นอกจากนี้คุณภาพของข้อมูล และความสามารถในการเรียนรู้ มีอิทธิพลทางอ้อมต่อความตั้งใจในการใช้งานผ่านการรับรู้ถึงประโยชน์และความพึงพอใจ โดยค่า Adjusted $R^2 = 0.679$ อธิบายความแปรปรวนของความตั้งใจในการใช้งาน ChatGPT ได้ร้อยละ 67.9*

**คำสำคัญ:** ปัญญาประดิษฐ์เชิงสร้างสรรค์ แชทจีพีที การยืนยันความคาดหวัง ข้อกังวลความเป็นส่วนตัว คุณภาพของข้อมูล ความสามารถในการเรียนรู้

## Abstract

*This research aims to examine the factors influencing users' continuance intention to use generative AI, focusing on ChatGPT. The research integrates the Expectation-Confirmation Model (ECM), Privacy Calculus Theory, and AI characteristics. Data was collected from 312 respondents through an online questionnaire and analyzed using Exploratory Factor Analysis (EFA) and Multiple Regression Analysis (MRA). The results show that satisfaction ($\beta = 0.489$, p-value = 0.001) and perceived usefulness ($\beta = 0.383$, p-value = 0.001) have a direct effect on continuance intention to use generative AI (ChatGPT). In contrast, privacy concerns do not have a significant effect on continuance intention. Additionally, information quality and learning capability have indirect effects on continuance intention through perceived usefulness and satisfaction. The model explains 67.9% of the variance in continuance intention (Adjusted $R^2 = 0.679$).*

**Keywords:** Generative Artificial Intelligence, ChatGPT, Expectation Confirmation, Privacy Concern, Information Quality, Learnability.

## 1. บทนำ

ในปัจจุบันปัญญาประดิษฐ์เชิงสร้างสรรค์ (Generative Artificial Intelligence หรือ Generative AI) เข้ามามีบทบาทสำคัญและเป็นเครื่องมือช่วยในการทำงานผ่านการสร้างเนื้อหาและวิเคราะห์ข้อมูลที่ซับซ้อน โดยเฉพาะ ChatGPT ได้กลายเป็นเครื่องมือสำคัญที่เปลี่ยนรูปแบบการทำงานทั่วโลก จากงานวิจัยของ Backlinko [1] รายงานว่า ChatGPT มีผู้ใช้งานราย เพิ่มขึ้นทั่วโลกแตะ 900 ล้านคนในปี 2025 อย่างไรก็ตามแม้จะมีการเปิดรับใช้งานอย่างแพร่หลายในเบื้องต้น แต่ความสำคัญของงานวิจัยยุคนี้คือจะทำอย่างไรให้คนใช้งานต่อไปได้อย่างยั่งยืน เพราะการใช้ต่อเนื่องไม่ได้ขึ้นอยู่กับกระแสในช่วงแรก แต่เกิดจากผลลัพธ์ที่ตอบโจทย์ความคาดหวังจริงหลังการใช้งาน

งานวิจัยนี้จึงเลือกใช้แบบจำลอง ECM เป็นแกนหลักในการวิเคราะห์ โดยดึงปัจจัยเรื่องคุณภาพข้อมูลและความง่ายในการ

เรียนรู้มาผสานกับข้อกังวลเรื่องความเป็นส่วนตัว เพื่อตรวจสอบว่าปัจจัยเหล่านี้ส่งผลต่อความพึงพอใจและความตั้งใจใช้งานต่อเนื่องในระยะยาวอย่างไร ผลการศึกษาที่ได้จะช่วยให้ผู้พัฒนาเทคโนโลยีและองค์กรเข้าใจพฤติกรรมผู้ใช้ในมิติต่างๆ ได้ชัดเจนขึ้น เพื่อนำไปปรับปรุงระบบให้สอดคล้องกับความต้องการจริงและสร้างความพึงพอใจให้ผู้ใช้งาน

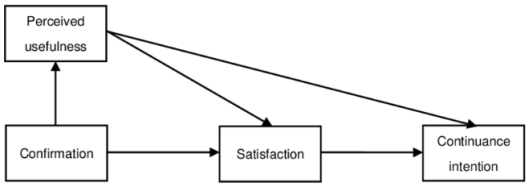
## 2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

### 2.1 แนวคิดเกี่ยวกับเทคโนโลยีปัญญาประดิษฐ์เชิงสร้างสรรค์

เทคโนโลยีปัญญาประดิษฐ์เชิงสร้างสรรค์ (Generative AI) เป็นสาขาหนึ่งของปัญญาประดิษฐ์ (AI) ที่กำลังได้รับความสนใจอย่างมากในปัจจุบัน เนื่องจากความสามารถในการสร้างเนื้อหาใหม่ที่มีคุณภาพสูงและมีความหลากหลาย สามารถประยุกต์ใช้งานได้อย่างมีประสิทธิภาพ เมื่อพิจารณาถึงการใช้งานปัญญาประดิษฐ์เชิงสร้างสรรค์อย่างต่อเนื่อง ในการศึกษาครั้งนี้ ผู้วิจัยได้มุ่งเน้นศึกษาปัจจัยที่ส่งผลต่อการตัดสินใจใช้งานเทคโนโลยีดังกล่าวอย่างต่อเนื่อง โดยได้บูรณาการแบบจำลองการยืนยันความคาดหวัง (ECM) ร่วมกับแนวคิดด้านลักษณะเฉพาะของเทคโนโลยี AI ตามกรอบงานวิจัยของ Bubas et al. [2] ซึ่งประกอบด้วยปัจจัยสำคัญ 2 ประการ ได้แก่ คุณภาพของข้อมูล (Information Quality) และความสามารถในการเรียนรู้ (Learnability) เพื่อให้ครอบคลุมมิติด้านประสิทธิภาพและการยอมรับเทคโนโลยีของผู้ใช้งาน

### 2.2 แบบจำลองการยืนยันความคาดหวัง

แบบจำลองการยืนยันความคาดหวัง (Expectation-Confirmation Model : ECM) จากงานวิจัยของ Bhattacherjee [3] กรอบแนวคิดที่อธิบายพฤติกรรมการใช้ระบบสารสนเทศอย่างต่อเนื่อง โดยดัดแปลงมาจากทฤษฎีการยืนยันความคาดหวัง (Expectation Confirmation Theory) ในศาสตร์พฤติกรรมผู้บริโภค ซึ่งชี้ให้เห็นว่าการศึกษาเกี่ยวกับระบบสารสนเทศส่วนใหญ่มุ่งเน้นไปที่การยอมรับเริ่มแรก แต่ความสำเร็จระยะยาวของระบบสารสนเทศกลับขึ้นอยู่กับการใช้อย่างต่อเนื่องมากกว่า โดยมีองค์ประกอบหลัก 4 ประการ ได้แก่ การยืนยันความคาดหวัง การรับรู้ประโยชน์ ความพึงพอใจ และความตั้งใจใช้งานต่อเนื่อง




**ภาพที่ 1** แบบจำลองการยืนยันความคาดหวัง (ECM)

### 2.3 ทฤษฎีความเป็นส่วนตัว

ทฤษฎีความเป็นส่วนตัว (Privacy Calculus Theory) เป็นแนวคิดที่ถูกพัฒนาขึ้นจาก Laufer and Wolfe [4] ในปี 1977 โดยมองว่าความเป็นส่วนตัวเปรียบได้กับสินค้าทางเศรษฐกิจที่อาจนำมาแลกเปลี่ยนกับสินค้าหรือผลประโยชน์อื่นได้ จากงานวิจัยของ Dinev e al. [5] พบว่าผู้ใช้ประเมินผลประโยชน์เทียบกับความเสี่ยงก่อนตัดสินใจใช้งาน หากผลประโยชน์สูงกว่าความเสี่ยง ความกังวลอาจไม่ส่งผลเชิงลบต่อการใช้งาน ดังนั้นในงานวิจัยนี้จึงได้นำของข้อกังวลเรื่องความเป็นส่วนตัว (Privacy Concerns) มาประยุกต์ใช้ในการพิจารณาถึงปัจจัยที่มีอิทธิพลการใช้งานปัญญาประดิษฐ์เชิงสร้างสรรค์อย่างต่อเนื่อง

### 2.4 การวิเคราะห์องค์ประกอบเชิงสำรวจ

การวิเคราะห์องค์ประกอบเชิงสำรวจ (Exploratory Factor Analysis) เป็นเทคนิคทางสถิติที่ใช้ในการศึกษาโครงสร้างความสัมพันธ์ของตัวแปรหลายๆตัว โดยมีจุดประสงค์ เพื่อจัดกลุ่มตัวแปรที่มีลักษณะคล้ายคลึงกันไว้ในองค์ประกอบเดียวกัน วิเคราะห์ KMO และ Bartlett's test เพื่อประเมินความสอดคล้องกลมกลืนของข้อมูลประกอบกับค่าน้ำหนักองค์ประกอบ Factor Loading ของแต่ละตัวแปร ซึ่งควรมีค่า 0.3 ขึ้นไป และการ Rotation เพื่อจัดกลุ่มตัวแปรที่มีความสัมพันธ์กันให้อยู่ในองค์ประกอบเดียวกัน

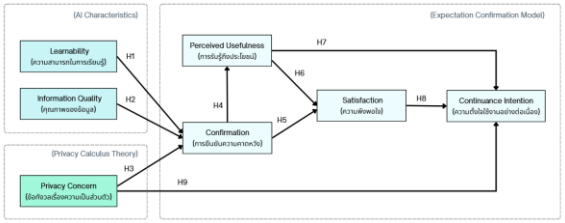
### 2.5 การวิเคราะห์การถดถอยพหุคูณ

การวิเคราะห์การถดถอยเชิงพหุคูณ (Multiple Regression Analysis) เป็นเทคนิคที่ใช้ในการศึกษาความสัมพันธ์ระหว่างตัวแปรตาม (Dependent Variable) และตัวแปรอิสระหลายตัว (Independent Variables) โดยใช้สมการเชิงเส้น เพื่ออธิบายผลกระทบของตัวแปรอิสระต่อการเปลี่ยนแปลงของตัวแปรตามได้

## 3. วิธีการดำเนินการวิจัย

### 3.1 กำหนดกรอบแนวคิด

งานวิจัยนี้ ผู้วิจัยพัฒนากรอบแนวคิดโดยบูรณาการแบบจำลองการยืนยันความคาดหวัง (ECM) ร่วมกับลักษณะเฉพาะของ Generative AI (ด้านคุณภาพข้อมูล และด้านความสามารถในการเรียนรู้) และปัจจัยด้านข้อกังวลเรื่องความเป็นส่วนตัว โดยแบบจำลอง ECM มีความเหมาะสมต่อการศึกษาการใช้งาน Generative AI เนื่องจากเทคโนโลยีนี้เป็นนวัตกรรมที่เน้นผลลัพธ์การใช้งาน ดังนั้นการใช้งานอย่างต่อเนื่องจึงไม่ได้ขึ้นอยู่กับความสะดวกเพียงอย่างเดียว แต่ขึ้นอยู่กับว่าผลลัพธ์ที่ Generative AI สร้างขึ้นนั้นสอดคล้องกับความคาดหวังของผู้ใช้หรือไม่ (Confirmation) ซึ่งส่งผลโดยตรงต่อการรับรู้ถึงประโยชน์ (Perceived Usefulness) ความพึงพอใจ (Satisfaction) และความตั้งใจใช้งานจริงอย่างต่อเนื่อง (Continuance Intention) ในระยะยาว ดังภาพที่ 2



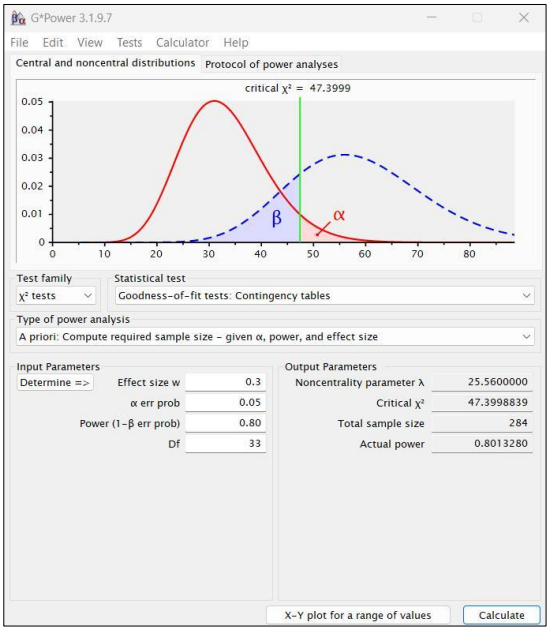

**ภาพที่ 2** กรอบแนวคิดการวิจัย

### 3.2 ประชากรและกลุ่มตัวอย่าง

3.2.1 ประชากรเป้าหมายในการวิจัยคือ บุคคลทั่วไปที่ทำงานอยู่ในพื้นที่จังหวัดกรุงเทพมหานคร ที่มีอายุตั้งแต่ 20 ปีขึ้นไป โดยการเลือกกลุ่มตัวอย่างในเขตกรุงเทพมหานครเพื่อควบคุมตัวแปรภายนอกด้านโครงสร้างพื้นฐานดิจิทัลและลดความคลาดเคลื่อนจากความเหลื่อมล้ำในการเข้าถึงเทคโนโลยี เพื่อให้มั่นใจว่าข้อมูลที่ได้รับสะท้อนจากกลุ่มผู้ใช้งานจริงที่มีประสบการณ์การใช้งาน

3.2.2 กลุ่มตัวอย่างคัดเลือกโดยใช้วิธีการสุ่มแบบสมัครใจ จากพนักงานที่มีประสบการณ์ใช้งาน ChatGPT เพื่อช่วยในการทำงาน โดยวิเคราะห์ข้อมูลในระดับบุคคล

3.2.3 การกำหนดขนาดกลุ่มตัวอย่างใช้โปรแกรม G*Power โดยกำหนดค่า effect size = 0.3 เป็นค่าขนาดกลางในการคำนวณสูตรไคสแควร์ของ Cohen ค่า Degree Freedom เป็น 33 ได้มาจากการคำนวณข้อคำถามจากแบบสอบถาม ค่าระดับนัยสำคัญที่ 0.05 เป็นค่าความคลาดเคลื่อนที่สมมติฐานจะไม่เป็นจริง โดยได้กลุ่มตัวอย่างจากการคำนวณจำนวน 284 คน




**ภาพที่ 3** การคำนวณขนาดของกลุ่มตัวอย่าง

### 3.3 เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการเก็บรวบรวมข้อมูลคือแบบสอบถามออนไลน์ผ่าน Microsoft Forms ซึ่งประกอบด้วยคำถามจำนวน 33 ข้อ โดยรูปแบบคำถามของตัวแปรดัดแปลงมาจากงานวิจัยดังนี้ คุณภาพข้อมูล ความสามารถในการเรียนรู้ และการรับรู้ถึงประโยชน์ จากงานวิจัยของ Bubas et al [2] ข้อกังวลเรื่องความเป็นส่วนตัว จากงานวิจัยของ Lim et al. [6] ความพึงพอใจ จากงานวิจัยของ Cheng and Jiang [7] การยืนยันความคาดหวังและความตั้งใจใช้งานต่อเนื่อง จากงานวิจัยของ Baig and Yadegaridehkordi [8] โดยสร้างเป็นคำถามประเภทสเกลเชิงเส้น (Likert Scale) 7 ระดับ เพื่อวัดระดับความคิดเห็นของผู้ตอบแบบสอบถาม เนื่องจากมีความเหมาะสมสำหรับการวัดทัศนคติที่มีความซับซ้อน โดยแบ่งแบบสอบถามออกเป็น 4 ส่วนได้แก่ ส่วนที่1 ด้านการยินยอมให้ข้อมูลกับผู้วิจัยตามแบบสอบถาม ส่วนที่2 คัดกรองผู้ตอบแบบสอบถามตามเกณฑ์ประสบการณ์ใช้งาน ChatGPT ส่วนที่ 3 ข้อมูลทั่วไปของผู้ทำแบบสอบถาม และส่วนที่4 เกี่ยวกับปัจจัยที่มีอิทธิพลต่อการใช้งานปัญญาประดิษฐ์เชิงสร้างสรรค์อย่างต่อเนื่อง กรณีศึกษา ChatGPT

## 4. ผลการดำเนินงาน

### 4.1 ผลการวิเคราะห์ข้อมูลทั่วไปของกลุ่มตัวอย่าง

ผลการสำรวจข้อมูล กลุ่มตัวอย่างจำนวน 312 คน ที่ทำงานอยู่ในจังหวัดกรุงเทพมหานคร พบว่า ส่วนใหญ่เป็นเพศหญิง กลุ่มอายุมากที่สุดอยู่ระหว่าง 35-39 ปี กลุ่มระดับการศึกษามากที่สุดอยู่ที่ระดับปริญญาตรีหรือเทียบเท่า และส่วนใหญ่ทำงานในกลุ่มธุรกิจอาหารและเครื่องดื่ม ซึ่งสะท้อนถึงกลุ่มผู้ใช้งานที่มีการใช้ Generative AI อย่าง ChatGPT เพื่อช่วยในการทำงาน ดังตารางที่ 1

**ตารางที่ 1** การวิเคราะห์ข้อมูลทั่วไปของกลุ่มตัวอย่าง

| ข้อมูลทั่วไป | จำนวน (คน) | ร้อยละ (%) |
|---|---|---|
| **เพศ** | | |
| ชาย | 111 | 35.58 |
| หญิง | 196 | 62.82 |
| ไม่ระบุเพศ | 5 | 1.60 |
| **รวม** | 312 | 100.00 |
| **อายุ** | | |
| 20 - 24 ปี | 4 | 1.28 |
| 25 - 29 ปี | 32 | 10.26 |
| 30 - 34 ปี | 54 | 17.31 |
| 35 - 39 ปี | 131 | 41.99 |
| 40 - 44 ปี | 44 | 14.10 |
| 45 - 49 ปี | 22 | 7.05 |
| 50 – 54 ปี | 9 | 2.88 |
| 55 ปีขึ้นไป | 16 | 5.13 |
| **รวม** | 312 | 100.00 |
| **ระดับการศึกษา** | | |
| ต่ำกว่าปริญญาตรี | 25 | 8.01 |
| ปริญญาตรีหรือเทียบเท่า | 180 | 57.69 |
| ปริญญาโทหรือสูงกว่า | 107 | 34.29 |
| **รวม** | 312 | 100.00 |
| **ประเภทธุรกิจที่ทำงาน** | | |
| การแพทย์ | 38 | 12.18 |
| ขนส่งและโลจิสติกส์ | 14 | 4.49 |
| เทคโนโลยีสารสนเทศ | 48 | 15.38 |
| ธนาคาร | 17 | 5.45 |
| ธุรกิจการเกษตร | 31 | 9.94 |
| อาหารและเครื่องดื่ม | 83 | 26.60 |
| ภาครัฐ | 16 | 5.13 |
| ธุรกิจอื่นๆ (16 ประเภท) | 65 | 20.83 |
| **รวม** | 312 | 100.00 |

### 4.2 ผลการวิเคราะห์องค์ประกอบเชิงสำรวจ (Exploratory Factor Analysis)

การวิเคราะห์องค์ประกอบเชิงสำรวจ เพื่อจัดกลุ่มตัวแปรที่มีความสัมพันธ์กันให้อยู่ในองค์ประกอบเดียวกัน จากการวิเคราะห์พบว่าค่า KMO เท่ากับ 0.904 ซึ่งมีค่ามากกว่า 0.50 แสดงถึงตัวแปรมีความสัมพันธ์กันในขนาดที่เหมาะสมในการนำไปวิเคราะห์เชิงสำรวจได้ และผลการทดสอบ Bartlett's Test of Sphericity พบว่ามีนัยสำคัญทางสถิติที่ระดับ 0.05 แสดงให้เห็นว่าตัวแปรมีความสัมพันธ์กันเพียงพอที่จะนำมาวิเคราะห์องค์ประกอบได้ ดังตารางที่ 2

**ตารางที่ 2** ค่า KMO และ Bartlett’s Test of Sphericity

| **KMO and Bartlett's Test** | | |
|---|---|---|
| Kaiser-Meyer-Olkin Measure of Sampling Adequacy. | | 0.904 |
| Bartlett's Test of Sphericity | Approx. Chi-Square | 7077 |
| | df | 253 |
| | Sig. | <.001 |

ผลการวิเคราะห์องค์ประกอบโดยใช้วิธี Maximum Likelihood และปรับหมุนแกนด้วยวิธี Oblimin (Oblique Rotation) และตัดข้อคำถามที่มีน้ำหนักองค์ประกอบ (Factor Loading) น้อยกว่า 0.3 และข้อคำถามที่สัมพันธ์กับองค์ประกอบอื่นออกแล้ว พบว่าสามารถจัดกลุ่มตัวแปรได้เป็น 7 องค์ประกอบ โดยทุกตัวแปรมีค่าน้ำหนักองค์ประกอบ (Factor Loading) มากกว่า 0.4 ซึ่งถือว่ามีความสัมพันธ์ที่ชัดเจนกับองค์ประกอบที่ถูกจัดกลุ่มการกระจายตัวของตัวแปรในแต่ละองค์ประกอบเป็นดังนี้ โดยองค์ประกอบที่ 2 องค์ประกอบที่ 3 และองค์ประกอบที่ 7 มีจำนวนตัวแปรมากจำนวน 4 ตัวแปร องค์ประกอบที่ 1 องค์ประกอบที่ 5 และองค์ประกอบที่ 6 มีตัวแปรจำนวน 3 ตัวแปร องค์ประกอบที่ 4 มีตัวแปรจำนวน 2 ตัวแปร ดังตารางที่ 3

**ตารางที่ 3** Rotated Factor Matrix

| **Rotated Factor Matrix** | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Factor | | | | | | | Uniqueness |
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
| CI2 | 0.991 | | | | | | | 0.097 |
| CI1 | 0.786 | | | | | | | 0.177 |
| CI4 | 0.755 | | | | | | | 0.248 |
| PC3 | | 0.955 | | | | | | 0.087 |
| PC2 | | 0.955 | | | | | | 0.082 |
| PC4 | | 0.922 | | | | | | 0.143 |
| PC1 | | 0.846 | | | | | | 0.247 |
| IQ5 | | | 0.854 | | | | | 0.203 |
| IQ6 | | | 0.801 | | | | | 0.351 |

**ตารางที่ 3** (ต่อ)

| Rotated Factor Matrix | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Factor | | | | | | | Uniqueness |
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
| IQ4 | | | 0.769 | | | | | 0.322 |
| IQ1 | | | 0.704 | | | | | 0.423 |
| SAT | | | | 0.670 | | | | 0.095 |
| SAT | | | | 0.581 | | | | 0.168 |
| LE5 | | | | | 0.830 | | | 0.254 |
| LE4 | | | | | 0.718 | | | 0.320 |
| LE1 | | | | | 0.401 | | | 0.522 |
| CON2 | | | | | | 0.627 | | 0.217 |
| CON3 | | | | | | 0.495 | | 0.249 |
| CON4 | | | | | | 0.411 | | 0.185 |
| PU4 | | | | | | | 0.535 | 0.321 |
| PU5 | | | | | | | 0.505 | 0.128 |
| PU3 | | | | | | | 0.456 | 0.204 |
| PU2 | | | | | | | 0.417 | 0.212 |

**4.3 ผลการทดสอบความเชื่อมั่นของแบบสอบถาม (Reliability Analysis)**

การวิเคราะห์ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน และค่าความเชื่อมั่น (Cronbach's Alpha) ของแต่ละตัวแปร ผลการวิเคราะห์ พบว่า ค่าความเชื่อมั่นของตัวแปรทั้งหมดอยู่ในช่วงระหว่าง 0.756 – 0.956 ซึ่งสูงกว่าเกณฑ์มาตรฐานที่กำหนดไว้ที่ 0.70 แสดงให้เห็นว่าแบบสอบถามมีความน่าเชื่อถืออยู่ในระดับดีถึงดีมาก และสามารถนำไปใช้ในการวิเคราะห์ข้อมูลในขั้นตอนถัดไปได้อย่างเหมาะสม โดยตัวแปรต่างๆ มีค่าเฉลี่ย (Mean = 5.49–6.00) สะท้อนว่าผู้ตอบแบบสอบถามมีความคิดเห็นในระดับเห็นด้วยต่อปัจจัยที่เกี่ยวข้องกับการใช้งานปัญญาประดิษฐ์เชิงสร้างสรรค์อย่างต่อเนื่อง ขณะที่ค่าส่วนเบี่ยงเบนมาตรฐาน (SD = 0.605–1.17) แสดงว่าความคิดเห็นของผู้ตอบแบบสอบถามมีการกระจายตัวในระดับปานกลาง โดยพบว่ามีแค่เพียงตัวแปร Privacy Concern ที่มีค่า SD เกิน 1 เนื่องจากพฤติกรรมของผู้ใช้งานและประสบการณ์ที่แตกต่างกันส่งผลให้มีการกระจายตัวในระดับสูง ดังตารางที่ 4

**ตารางที่ 4** ผลการทดสอบความค่าเชื่อมั่นของแบบสอบถาม

| ตัวแปร | Mean | S.D. | Cronbach's α |
|---|---|---|---|
| Continuance Intention (CI) | 5.830 | 0.924 | 0.927 |
| Privacy Concern (PC) | 5.540 | 1.170 | 0.956 |
| Information Quality (IQ) | 5.490 | 0.761 | 0.878 |
| Satisfaction (SAT) | 5.900 | 0.774 | 0.920 |
| Learnability (LE) | 6.000 | 0.605 | 0.756 |
| Confirmation (CON) | 5.760 | 0.759 | 0.906 |
| Perceived Usefulness (PU) | 5.920 | 0.747 | 0.911 |
| Continuance Intention (CI) | 5.830 | 0.924 | 0.927 |

**4.4 การวิเคราะห์สมการถดถอยพหุคูณ (Multiple Regression Analysis)**

จากกรอบแนวคิดการวิจัยในการศึกษา ผู้วิจัยได้ดำเนินการทดสอบสมมติฐานโดยใช้การวิเคราะห์สมการถดถอยพหุคูณ แบ่งออกเป็น 4 แบบจำลอง ซึ่งผลการวิเคราะห์พบว่าค่า VIF ของทุกองค์ประกอบอยู่ระหว่าง 1.01–2.79 ซึ่งต่ำกว่าเกณฑ์ที่กำหนดทั้งหมดและค่า Tolerance มีค่าเข้าใกล้ 1 แสดงว่าตัวแปรอิสระทุกตัวไม่มีระดับความสัมพันธ์ซึ่งกันและกัน สามารถทำการทดสอบสมมติฐานงานวิจัยได้ สรุปผลดังนี้

แบบจำลองที่ 1 วิเคราะห์ผลตัวแปรอิสระ 3 ตัวแปร ได้แก่ ความสามารถในการเรียนรู้ คุณภาพของข้อมูล และข้อกังวลเรื่องความเป็นส่วนตัว สามารถอธิบายความแปรปรวนของการยืนยันความคาดหวังได้ร้อยละ 49.1 ดังตารางที่ 5

**ตารางที่ 5** Model Summary และ Coefficients - CON

| Model | R | | $R^2$ | | Adjusted $R^2$ | Collinearity Statistics | |
|---|---|---|---|---|---|---|---|
| 1 | 0.701 | | 0.491 | | 0.486 | | |
| Model Coefficients – Confirmation (CON) | | | | | Stand. Estimate | VIF | Tolerance |
| Predictor | Estimate | SE | t | p | | | |
| Intercept | 0.2052 | 0.3633 | 0.565 | 0.573 | | | |
| LE | 0.3431 | 0.0568 | 6.042 | <.001 | 0.271 | 1.22 | 0.819 |
| IQ | 0.5439 | 0.0456 | 11.936 | <.001 | 0.539 | 1.23 | 0.810 |
| PC | 0.0917 | 0.0269 | 3.402 | <.001 | 0.139 | 1.01 | 0.986 |

แบบจำลองที่ 2 วิเคราะห์ผลตัวแปรอิสระเพียง 1 ตัวแปรคือ การยืนยันความคาดหวังมีอิทธิพล สามารถอธิบายความแปรปรวนของการรับรู้ถึงประโยชน์ ได้ร้อยละ 59.2 ดังตารางที่ 6

**ตารางที่ 6** Model Summary และ Coefficients - PU

| Model | R | | $R^2$ | | Adjusted $R^2$ | Collinearity Statistics | |
|---|---|---|---|---|---|---|---|
| 1 | 0.770 | | 0.593 | | 0.592 | | |
| Model Coefficients - Perceived Usefulness (PU) | | | | | Stand. Estimate | VIF | Tolerance |
| Predictor | Estimate | SE | t | p | | | |
| Intercept | 1.628 | 0.2035 | 8.00 | <.001 | | | |
| CON | 0.746 | 0.0351 | 21.26 | <.001 | 0.770 | 1.00 | 1.00 |

แบบจำลองที่ 3 วิเคราะห์ผลตัวแปรอิสระ 2 ตัวแปร ได้แก่ ได้แก่ การยืนยันความคาดหวังและการรับรู้ถึงประโยชน์ โดยสามารถอธิบายความแปรปรวนของความพึงพอใจได้ร้อยละ 70 ดังตารางที่ 7

**ตารางที่ 7** Model Summary และ Coefficients - SAT

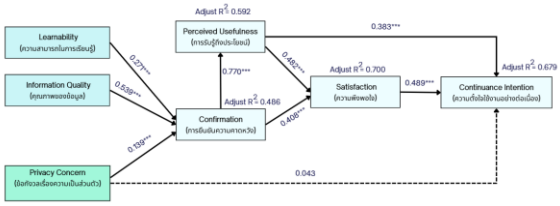
| Model | R | | $R^2$ | | Adjusted $R^2$ | Collinearity Statistics | |
|---|---|---|---|---|---|---|---|
| 1 | 0.838 | | 0.702 | | 0.700 | | |
| Model Coefficients – Satisfaction (SAT) | | | | | Stand. Estimate | VIF | Tolerance |
| Predictor | Estimate | SE | t | p | | | |
| Intercept | 0.591 | 0.1984 | 2.98 | 0.003 | | | |
| CON | 0.410 | 0.0488 | 8.39 | <.001 | 0.408 | 2.46 | 0.407 |
| PU | 0.500 | 0.0504 | 9.91 | <.001 | 0.482 | 2.46 | 0.407 |

แบบจำลองที่ 4 วิเคราะห์ผลตัวแปรอิสระ 3 ตัวแปร ได้แก่ การรับรู้ถึงประโยชน์ ความพึงพอใจ และข้อกังวลเรื่องความเป็นส่วนตัว สามารถอธิบายความแปรปรวนของความตั้งใจในการใช้งานต่อเนื่องได้ร้อยละ 67.9 ดังตารางที่ 8

**ตารางที่ 8** Model Summary และ Coefficients - CI

| Model | R | | $R^2$ | | Adjusted $R^2$ | Collinearity Statistics | |
|---|---|---|---|---|---|---|---|
| 1 | 0.826 | | 0.682 | | 0.679 | | |
| Model Coefficients - Continuance Intention (CI) | | | | | Stand. Estimate | VIF | Tolerance |
| Predictor | Estimate | SE | t | p | | | |
| Intercept | -0.6483 | 0.2896 | -2.24 | 0.026 | | | |
| PU | 0.4763 | 0.0666 | 7.15 | <.001 | 0.383 | 2.77 | 0.361 |
| SAT | 0.5874 | 0.0645 | 9.11 | <.001 | 0.489 | 2.79 | 0.359 |
| PC | 0.0340 | 0.0257 | 1.32 | 0.187 | 0.043 | 1.02 | 0.982 |

จากการทดสอบสมมติฐานทั้งหมด 7 ตัวแปร ได้โมเดลอธิบายการใช้งานปัญญาประดิษฐ์เชิงสร้างสรรค์อย่างต่อเนื่อง กรณีศึกษา ChatGPT ดังภาพที่ 4




**ภาพที่ 4** โมเดลอธิบายความตั้งใจใช้งานปัญญาประดิษฐ์เชิงสร้างสรรค์อย่างต่อเนื่อง

## 5. สรุป

จากวัตถุประสงค์ของงานวิจัยและโมเดลการศึกษาปัจจัยที่มีอิทธิพลต่อการใช้งานปัญญาประดิษฐ์เชิงสร้างสรรค์อย่างต่อเนื่อง กรณีศึกษา ChatGPT ผลการศึกษาพบว่า ปัจจัยด้านความพึงพอใจ ($\beta = 0.489, p < .001$) มีอิทธิพลสูงที่สุด รองลงมาคือการรับรู้ถึงประโยชน์ ($\beta = 0.383, p < .001$) ซึ่งสะท้อนว่าผู้ใช้ให้ความสำคัญกับประสิทธิภาพที่ได้รับจริงจากการใช้งานมากกว่ากระแส สอดคล้องกับแนวคิด ECM ที่ใช้การยืนยันผลลัพธ์หลังการใช้งานเป็นตัวตัดสินการใช้ต่อ ขณะที่ข้อกังวลเรื่องความเป็นส่วนตัวไม่มีอิทธิพลโดยตรงอย่างมีนัยสำคัญ ($\beta = 0.043, p > .05$) สะท้อนถึงปรากฏการณ์ Privacy Paradox ที่ผู้ใช้ยอมแลกความเสี่ยงกับประโยชน์ที่ได้รับ นอกจากนี้คุณภาพของข้อมูล และความสามารถในการเรียนรู้มีอิทธิพลทางอ้อมต่อความตั้งใจในการใช้งานผ่านการรับรู้ถึงประโยชน์และความพึงพอใจ โดยอธิบายความแปรปรวนของความตั้งใจในการใช้งาน ChatGPT ได้ร้อยละ 67.9

งานวิจัยนี้พิสูจน์ให้เห็นถึงปัจจัยและแบบจำลองการใช้งานปัญญาประดิษฐ์เชิงสร้างสรรค์อย่างต่อเนื่อง กรณีศึกษา ChatGPT ตามกรอบแนวคิด ECM ได้อย่างชัดเจน โดยผลการศึกษาช่วยยืนยันว่าการสร้างความพึงพอใจและประโยชน์จากการใช้จริงคือหัวใจสำคัญ ซึ่งเป็นแนวทางที่มีประสิทธิภาพในการพัฒนาและส่งเสริมการใช้งานปัญญาประดิษฐ์เชิงสร้างสรรค์อย่างยั่งยืนในอนาคต

## เอกสารอ้างอิง

[1] Backlinko Team. (2025, December 18). ChatGPT statistics 2026: How many people use ChatGPT?. Backlinko.https://backlinko.com/chatgpt-stats

[2] Goran Bubaš., et al (2023). “Development of an Assessment Scale for Measurement of Usability and User Experience Characteristics of Bing Chat Conversational AI”. https://doi.org/10.3390/fi16010004

[3] Anol Bhattacherjee. (2001). "Understanding Information Systems Continuance: An Expectation-Confirmation Model.” MIS Quarterly. Vol. 25, No. 3: 351–370. doi: 10.2307/3250921.

[4] Laufer, R. S. & Wolfe, M. (1977). Privacy as a Concept and a Social Issue:A Multidimensional Developmental Theory. Journal of Social Issues. 33(3), 22-42.

[5] Dinev, T., & Hart, P. (2006). An extended privacy calculus model for e-commerce transactions. Information Systems Research, 17(1), 61–80. https://doi.org/10.1287/isre.1060.0080

[6] Joon Soo Lim., et al. (2025). “The Role of User Empowerment, AI Hallucination, and Privacy Concerns in Continued Use and Premium Subscription Intentions: An Extended Technology Acceptance Model for Generative AI”https://doi.org/10.1080/08838151.2025.2487679

[7] Yang Cheng & Hua Jiang (2020). “How Do AI-driven Chatbots Impact User Experience? Examining Gratifications, Perceived Privacy Risk, Satisfaction, Loyalty, and Continued Use”. https://doi.org/10.1080/08838151.2020.1834296

[8] Maria Ijaz Baig and Elaheh Yadegaridehkordi (2025) “Factors Influencing academic staff satisfaction and continuous usage of generative artificial intelligence (GenAI) in higher education”. https://doi.org/10.1186/s41239-025-00506-4

# AUTHOR INDEX


| | Pages |
|---|---|
| *Adisak Yongpanya* | 321 |
| *Adisorn Na Ubon* | 106, 254, 260 |
| *Aimphagar Luangphusab* | 242 |
| *Amnach Khawne* | 199 |
| *Anan Malichai* | 334 |
| *Anawin Uthaisang* | 175 |
| *Anothai Arunthong* | 382 |
| *Anthika Wattanakulwiwat* | 30 |
| *Apichart Kanseenuan* | 328 |
| *Apinya Thongkum* | 532 |
| *Apirat Jirangkapanich* | 480 |
| *Bavornthat Dangtang* | 341 |
| *Benchapol Thongchua* | 248 |
| *Benjawan Surapol* | 277 |
| *Bunyawat Pattanasakkomjon* | 260 |
| *Burasakorn Yoosooka* | 341 |
| *Chaiyut Mattrai* | 230 |
| *Chakchai So-In* | 13, 137, 158, 175, 185, 213 |
| *Chalermkiet Tongwilai* | 100 |
| *Chanakit Mitsongkore* | 296 |
| *Chatphol Nawjampa* | 158 |
| *Chaveevan Pechsiri* | 106, 254, 260 |
| *Chayanon Premvijitr* | 242 |
| *Chonnika Phuthong* | 486 |
| *Ekasis Girachaipitack* | 254 |
| *Jittakron Saetung* | 394 |
| *Kamonlak Pinyakong* | 302 |
| *Kanoknipha Panset* | 213 |
| *Keerati Jantorn* | 455 |
| *Kiattisak Luithong* | 7 |
| *Komjak Kamsiang* | 131 |
| *Korawit Sutthi-at* | 158 |
| *Krairawee Nosri* | 205 |
| *Kritsanaphong Kimsue* | 513 |
| *Laksamee Noneluecha* | 236 |
| *Mahasak Ketcham* | 141, 500 |
| *Maleerat Maliyaem* | 283 |

# AUTHOR INDEX


| | Pages |
|---|---|
| *Manu Boonpleum* | 483 |
| *Mayurakran Chinnawong* | 55 |
| *Metta Naktam* | 551 |
| *Montean Rattanasiriwongwut* | 81, 100, 334, 375, 388, 394, 444 |
| *Monthicha JImthaisong* | 500 |
| *Nalinpat Bhumpenpein* | 7, 68, 112, 125, 131, 230, 449, 480, 507 |
| *Nansaphat Termsub* | 175 |
| *Napatwan Chaiban* | 213 |
| *Naravit Conjaiboon* | 137 |
| *Nataporn Koopipat* | 271 |
| *Natcharee Ariyanithitawesuk* | 188 |
| *Nattakan Tamcharoen* | 375 |
| *Nattapon Boondang* | 199 |
| *Nattaporn Kongsakunkrai* | 545 |
| *Nattavee Utakrit* | 55, 408, 426, 467, 493, 513, 519, 532 |
| *Natthanun Wicheanrat* | 1 |
| *Natthida Looksua* | 36 |
| *Nawaporn Wisitpongphan* | 49, 170, 248, 315 |
| *Nippich Yimporntana* | 151 |
| *Nithi Sakunphanish* | 519 |
| *Nitipat Sritharachiyanont* | 341 |
| *Nopsapoj Srisuwan* | 362 |
| *Pacharin Sirirak* | 368 |
| *Panatsaya Thongdee* | 42 |
| *Papatsara Siharach* | 449 |
| *Phakamon Malee* | 355 |
| *Phattapol Yamwankaew* | 170 |
| *Phattranit Vasuangsunit* | 81 |
| *Phayung Meesad* | 308, 420, 473 |
| *Phumsit Nanram* | 341 |
| *Phutthicha Rattana* | 185 |
| *Pipat Thongjapo* | 289 |
| *Pongpeera Chaipongpatsorn* | 467 |
| *Pongsri Prasomboon* | 94 |
| *Poramets Phetman* | 75 |
| *Prawit Ekweerapatarasakun* | 388 |

# AUTHOR INDEX


| | Pages |
|---|---|
| *Pudsadee Boonrawd* | 349, 461 |
| *Rachata Nanta* | 185 |
| *Rajabhat Raksujarit* | 401 |
| *Ratchada Pleanthong* | 106 |
| *Ratchanat Huansin* | 112 |
| *Rewadee Aminram* | 125 |
| *Run Kanchano* | 493 |
| *Sabsin Dilokwechpanich* | 507 |
| *Sakchai Tangwannawit* | 88, 94, 362, 368, 401, 545 |
| *Silaporn Chuntha* | 13 |
| *Sirikarn Netpookgana* | 137 |
| *Sitawee Sirithitiwithaya* | 461 |
| *Sojirat Songpradit* | 414 |
| *Suda Phuchamnong* | 432 |
| *Sunantha Sodsee* | 1, 30, 36, 151 |
| *Supakit Nakpomchin* | 321 |
| *Supakorn Srithongjun* | 408 |
| *Supakpong Jinarat* | 205, 223, 236, 271, 289, 302 |
| *Suphakit Awiphan* | 328 |
| *Suphitcha Chanrueang* | 321 |
| *Sureewan Jangjit* | 296 |
| *Sutat Gammanee* | 296 |
| *Sutthikarn Julsuwan* | 68 |
| *Suvil Chomchaiya* | 341 |
| *Tadchanon Polngam* | 349 |
| *Tanapon Jensuttiwetchakul* | 42, 75, 118, 355, 382, 525, 551 |
| *Tanyavat Kaewpradit* | 265 |
| *Teeradaj Kanching* | 308 |
| *Teeratuch Kongyuenyongwanit* | 473 |
| *Tewarach Rakngam* | 277 |
| *Thanagrit Wutthiamornthada* | 49 |
| *Thanatmeth Wuttisathittham* | 88 |
| *Theerachai Kasemphongjarat* | 283 |
| *Thitiworada Thabtimsai* | 13 |
| *Uraiwan Inyaem* | 242, 277, 341 |
| *Uraiwan Meesang* | 525 |
| *Wachirasak Pokakorn* | 61 |

# AUTHOR INDEX


| | Pages |
|---|---|
| *Wannachai Chouegthong* | 321 |
| *Waraporn Viyanon* | 265 |
| *Warong Naivinit* | 296 |
| *Watchareewan Jitsakul* | 432, 438, 455, 486 |
| *Weena Janratchakool* | 242, 277 |
| *Wilaiwan Phanthura* | 315 |
| *Wilasinee Chutiwat* | 426 |
| *Wittayadet Jaikwang* | 420 |
| *Witthawat Chankimha* | 321 |
| *Worawat Soiphet* | 223 |
| *Worawoot Samanpong* | 444 |




# NCCIT2027

## The 23rd National Conference on Computing and Information Technology

---

**Organized by**

Faculty of Information Technology and Digital Innovation

King Mongkut's University of Technology North Bangkok

---

**WWW.NCCIT.NET**

# Conference Organizers

www.nccit.net