ISSN: 1906-9553 (Print)
ISSN: 2651-1053 (Online)

JOURNAL OF INFORMATION SCIENCE AND TECHNOLOGY

JIST

Volume 13. No.1 January - June 2023

Research Papers

JOURNAL OF INFORMATION SCIENCE AND TECHNOLOGY (JIST)

ISSN 2651-1053 (Online), ISSN 1906-9553 (Print)

The Journal of Information Science and Technology (JIST) is an academic journal established by the collaboration of 22 faculties that conduct courses related to Information Technology, namely, Bangkok University, Dhurakij Pundit University, King Mongkut's Institute of Technology Ladkrabang, King Mongkut's University of Technology North Bangkok, King Mongkut's University of Technology Thonburi, Mae Fah Luang University, Mahanakorn University of Technology, Mahasarakham University, Mahidol University, Nakhon Phanom University, Panyapiwat Institute of Management, Prince of Songkla University, Rangsit University, Siam University, Silpakorn University, Sripatum University, Thai-Nichi Institute of Technology, Walailak University, Burapha University, Phayao University, Ubon Ratchathani Rajabhat University and Khon Kaen University. According to the agreement of deans of all faculties (Council of IT Deans of Thailand (CITT)).

The journal was established in 2010 and plans to publish 2 issues per year (JAN – JUNE and JULY – DECEMBER per year). The journal was established first print journal publication in 2010 (Vol 1. No.1) with ISSN 1906-9553 (Print) and plans to publish 2 issues per year on during January - June and July - December. Also the journal was established first online journal publication in 2010 (Vol 1. No.1) with ISSN 2651-1053 (Online) and plans to publish 2 issues per year on during January - June and July - December.

ADVISORY BOARD

JOURNAL OF INFORMATION SCIENCE AND TECHNOLOGY (JIST)

ISSN 2651-1053 (Online), ISSN 1906-9553 (Print)

CONTACT ADDRESS

Council of IT Deans of Thailand (CITT),
Faculty of Engineering and Technology,
Mahanakorn University of Technology
140 Moo 1, Cheum-Sampan Road, Nongchok,
Bangkok, Thailand 10530

CALL FOR PAPER
Journal of Information Science and Technology (JIST)

https://ph02.tci-thaijo.org/index.php/JIST

Submissions
Online Journal Submission
https://ph02.tci-thaijo.org/index.php/JIST/

Publication Periodicity
Published 2 times a year in June and December.
(JAN – JUN and JULY – DEC).
(***Indexed by TCI tier 2***)

JIST Office Address
Council of IT Deans of Thailand (CITT), Faculty of Engineering and Technology, Mahanakorn University of Technology 140 Moo 1, Cheum-Sampan Rd., Nongchok, Bangkok, Thailand 10530

Peer Review Process
All submitted manuscripts must be reviewed by at least three expert reviewers via the double-blinded review system.

The **Journal of Information Science and Technology (JIST)** aims to be the forum through which researchers, faculties, and experts of the computer and information technology and others technological related fields share and discuss their high quality research work as well as innovation. Original research articles, practical applications and innovations in the broad area of computer and information technology are suitable for publication in JIST.

- **Soft Computing:**
  Artificial Intelligence, Fuzzy and Neural Computing, Genetic Algorithms, Intelligent Agents, Neural Networks, Natural Language Processing, Robotics and Automation, Image Processing
- **Human-Computer Interaction:**
  Graphical User Interfaces, Multimedia, Interactive Systems, Computer-Supported Cooperative Work, Human Cognitive Skills, Visualization, and Computer Simulation
- **Information Assurance and Security:**
  Cryptography, Forensics and Incident Response, Biometrics, Security Policies and Procedures
- **Information Systems:**
  Databases, Information Retrieval, Transaction Processing, Distributed and Object Databases, Data Warehousing, Knowledge Management, Expert Systems, Multimedia Information Systems, Digital Libraries, Geographical Information Systems
- **Networking:**
  Advanced Computer Networks, Internet Technology and Applications, Distributed Systems, Wireless and Mobile Computing, Data Compression, Network Security, Enterprise Networking, Digital Communications
- **Programming:**
  Object-Oriented Programming, Event-Driven Programming, Functional Programming, Logic Programming, XML and other Extensible Languages, Parallel Programming
- **Platform Technologies:**
  Advanced Computer Architecture, Parallel Architectures, Cluster / Grid Computing, RFID, Embedded Systems
- **System Integration and Architecture:**
  Software Acquisition and Implementation, System Needs Assessment, Software Economics, Enterprise Systems, Computing Economics
- **Social and Professional Issues:**
  Professional Practice, Social Context of Computing, Computers Ethics, IT Law, Intellectual Property, Privacy and Civil Rights
- **Web Systems and Technologies**:
  Programming for the WWW, E-commerce, E-Learning, Content Management Systems, E-Government and E-Service
- **Multidisciplinary:**
  Bioinformatics, Biomedical Information Systems, Nano Technology
- **e-Business and m-Business:**
  e-Business Process Aspects, Intelligent e-Business System, m-Business, e-Supply Chain Management & e-Logistics, e-Customer Relationship Management, e-Negotiation, e-Payment, e-Business Design and Developments
- **Business and Information System:**
  Information management for business applications, Enterprise systems and architecture, Business systems infrastructure design for information integration, Service-oriented architecture, Service-component architecture
- **Social and Business Aspects of Convergence IT and Ubiquitous Computing:**
  Economics of emerging technologies, Home and community computing, Economic and social complexities in CIT and UC, New forms of media and communications
- **Internet of Things (IoT):**
  IoT system architecture, IoT enabling technologies, IoT communication and networking protocols such as network coding, and IoT services and applications and technology development in different standard development organizations (SDO).
- **Cloud Computing:**
  Auditing, monitoring, scheduling, resource registration/discovery, Automatic reconfiguration, self healing/monitoring, Service level agreements, integration, management.
- **Fintech and blockchain:**
  Development of Blockchain Security Enhancement Technologies, Platform Technology that Supports Safe Transactions, Financial and Economic Structures, Blockchain's Impacts on Workstyles, Fast Fintech for Smartphone Application Service Platform.

Manuscript Preparation Guide (Publication Charge: Free)

The template for writing the journal is available for downloading (jist_template_eng-2020.doc or jist_template_thai-2020.doc). Minimum length of the paper should be in between 6-16 pages of A4 size. Text should be typewritten or printed with double-spaced in 11-point of A4 white paper with margins of 1.5" for top, 1.25" for left, and 1" for bottom and right sides. All pages must be numbered sequentially. Here are some guidelines.

- Abstract with length 100-200 words in length.
- Introduction which discusses existing problem and related research
- Materials and methods used for experimentation, Results of the experiment, Discussion and conclusion, References Style is IEEE.

JIST Journal of Information Science and Technology
Volume 13, NO 1 | JAN – JUN 2023 | 1-9
ISSN: 2651-1053 (Online)
ISSN: 1906-9553 (Print)

แอนดรอยด์แอปพลิเคชันสำหรับจดจำและตรวจสอบใบหน้าโดยใช้โมบายเฟซเน็ต

Android Application for Face Recognition and Verification using MobileFaceNets

เชี่ยวชาญ ยางศิลา

Chiawchan Yangsila

สาขาวิชาคอมพิวเตอร์และเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยราชภัฏร้อยเอ็ด

Computer and Information Technology Program, Faculty of Information Technology, Roi Et Rajabhat University

Received: November 11, 2022; Revised: April 18, 2023; Accepted: June 08, 2023; Published: June 30, 2023

ABSTRACT – This paper presents the development of Android application for face recognition and verification using MobileFaceNets: a case study of the activity attendance check of staff at Roi Et Rajabhat University (RERU). The four objectives and their results in the study are as follows. 1) The developed Android application for face recognition and verification using MobileFaceNets was applicable to the activity attendance check of the RERU staff. 2) The Manhattan function had the fastest processing speed. 3) The researcher did a distance analysis for 100% accuracy and found that image A was identical to image B if the distance was <= 0.752 (Euclidean distance) or <= 5.796 (Manhattan distance) or >= 0.717 (Cosine Coefficient). 4) The efficiency analysis revealed that the application processing was 100% accurate.

KEYWORDS: Android Application, Face Recognition and Verification, MobileFaceNets, TensorFlow Lite, Convolutional Neural Network

บทคัดย่อ -- **บทความนี้นำเสนอการพัฒนาแอนดรอยด์แอปพลิเคชันสำหรับจดจำและตรวจสอบใบหน้าโดยใช้โมบายเฟซเน็ต กรณีศึกษา การเช็คชื่อเข้าร่วมกิจกรรมมหาวิทยาลัยราชภัฏร้อยเอ็ด ซึ่งมีวัตถุประสงค์ 4 ข้อ ขอสรุปผลการทดลองตามวัตถุประสงค์ดังนี้ 1) การพัฒนาแอนดรอยด์แอปพลิเคชันสำหรับจดจำและตรวจสอบใบหน้าโดยใช้โมบายเฟซเน็ต กรณีศึกษา การเช็คชื่อเข้าร่วมกิจกรรมมหาวิทยาลัยราชภัฏร้อยเอ็ดสามารถนำไปใช้งานได้จริง 2) การทดลอง พบว่า ฟังก์ชันแมนฮัตตัน มีความเร็วในการประมวลผลมากที่สุด 3) การวิเคราะห์ระยะห่างที่ให้คำตอบถูกต้อง 100% สรุปว่า ภาพ ก จะเหมือนกันกับภาพ ข ถ้าระยะห่าง <= 0.752 (ยูคลิเดียน) หรือ <= 5.796 (แมนฮัตตัน) หรือ >= 0.717 (สหสัมพันธ์โคไซน์) วิเคราะห์โดยใช้ผู้วิจัยเอง และ 4) ผลการหาประสิทธิภาพของแอปพลิเคชัน พบว่า แอปพลิเคชันประมวลผลถูกต้อง 100%**

คำสำคัญ: แอนดรอยด์แอปพลิเคชัน, จดจำและตรวจสอบใบหน้า, โมบายเฟซเน็ต, เทนเซอร์โฟลลิต, โครงข่ายประสาทเทียมแบบสังวัตนาการ

Corresponding Author: cheawchang2@gmail.com

1. บทนำ

โทรศัพท์มือถือแบบสมาร์ทโฟนได้มีบทบาทมากขึ้นในชีวิตประจำวันซึ่งโทรศัพท์มือถือนั้นมีประสิทธิภาพมากพอที่จะประมวลผลได้เทียบเท่ากับคอมพิวเตอร์ขนาดเล็ก ในขณะเดียวกันเทคโนโลยีการเรียนรู้ของเครื่อง (Machine Learning) ก็มีบทบาทสำคัญมากในการเข้ามาจำแนกชนิดหรือประเภทของวัตถุโดยการใช้เทคนิคโครงข่ายประสาทเทียมแบบสังวัตนาการ (Convolutional Neural Network: CNN) ในการจำแนกรูปภาพ [1,2,3,4] อีกทั้งโทรศัพท์มือถือยังมีกล้องถ่ายรูปในตัวและสามารถนำไปใช้งานร่วมกับเทคโนโลยีที่กล่าวไว้ข้างต้นอีกด้วย

มหาวิทยาลัยราชภัฏร้อยเอ็ดเป็นสถาบันอุดมศึกษาเพื่อการพัฒนาท้องถิ่นซึ่งมีพันธกิจหลายพันธกิจ พันธกิจที่สำคัญคือการผลิตบัณฑิตซึ่งการผลิตบัณฑิตนั้นนอกจากจะพัฒนานักศึกษาตามโครงสร้างหลักสูตรแล้วยังมีการเสริมประสบการณ์ให้กับนักศึกษาด้วยกิจกรรมต่าง ๆ ที่ดูแลโดยสำนักกิจการนักศึกษา ในการจัดกิจกรรมแต่ละครั้งจะต้องมีการเช็คชื่อนักศึกษาที่เข้าร่วมกิจกรรม กระบวนการเช็คชื่อจะกระทำเมื่อสิ้นสุดการจัดกิจกรรมโดยให้นักศึกษาเข้าแถวรอรับคูปองสำหรับใช้ในการลงทะเบียนเข้าร่วมกิจกรรม ในการจัดกิจกรรมแต่ละครั้งมีนักศึกษาเข้าร่วมจำนวนมาก เจ้าหน้าที่ต้องจัดเตรียมคูปองให้เพียงพอและภายหลังจากนักศึกษาได้รับคูปองแล้วนักศึกษาจะต้องเข้าไปกรอกข้อมูลในระบบสารสนเทศด้วยตนเอง ผู้วิจัยจึงมีแนวคิดในการนำโทรศัพท์มือถือที่ใช้ระบบปฏิบัติการแอนดรอยด์ซึ่งเชื่อว่านักศึกษาส่วนใหญ่ใช้กันอยู่มาใช้สำหรับการเช็คชื่อการเข้าร่วมกิจกรรม กระบวนการเช็คชื่อก็คือให้ศึกษาแต่ละคนใช้โทรศัพท์มือถือของตัวเองถ่ายภาพใบหน้าของตัวเองแล้วแอปพลิเคชันที่พัฒนาขึ้นจะประมวลผล โดยถ้าภาพที่ถ่ายตรงกับภาพที่เคยลงทะเบียนไว้ในโทรศัพท์มือถือ แอปพลิเคชันก็จะยินยอมให้สามารถบันทึกข้อมูลการเข้าร่วมกิจกรรมไปเก็บในฐานข้อมูลได้ทันที โดยแอปพลิเคชันดังกล่าวจะใช้บริการ GPS เพื่อให้ทราบตำแหน่งที่นักศึกษาอยู่เพื่อให้ทราบว่าขณะนั้นนักศึกษาอยู่ในบริเวณจัดกิจกรรมหรือไม่ ผู้วิจัยได้ศึกษาความเป็นไปได้ในการเลือกเครื่องมือในการพัฒนาแอปพลิเคชันจากงานวิจัยต่าง ๆ ดังนี้

1.1 การค้นหาใบหน้าในรูปภาพ (Face detection) Salinda Hettiarachchi [5] กล่าวว่า Google ML Kit เป็นเครื่องมือที่ดีที่สุด ใช้เวลาเฉลี่ย 68 มิลลิวินาที

1.2 การตรวจสอบใบหน้า (Face Verification)มีหลายงานวิจัยดังนี้ Salinda Hettiarachchi [5] กล่าวว่าโมเดล FaceNet [6] เป็นอัลกอริทึมที่ให้ค่าความถูกต้องมากที่สุดประมาณ 95% ในขณะเดียวกันโมเดล MobileFaceNet [7] เป็นอัลกอริทึมที่ประมวลผลเร็วที่สุดประมาณ 24 มิลลิวินาที Sheng Chen และคณะ [7] สรุปว่า โมเดล MobileFaceNet มีขนาด 4 MB ประมวลผลได้เร็วเหมาะสำหรับการใช้งานในโทรศัพท์มือถือ มีค่าความถูกต้อง 99.55% ในขณะที่โมเดล MobileNetV2 ให้ค่าความถูกต้อง 98.58% ความเร็ว 49 มิลลิวินาที ขนาดของโมเดลอยู่ระหว่าง 1.7 MB – 6.9 MB และโมเดล FaceNet ค่าความถูกต้อง 99.63% ขนาดของโมเดล 30 MB ซึ่งถือว่ามีขนาดใหญ่ไม่เหมาะสำหรับใช้งานในโทรศัพท์มือถือ Adrian Rosebrock [8] ได้แก้ปัญหาเดียวกันนี้โดยใช้ OpenCV's face_recognition library ในภาษาไพทอน ด้วย nn4.small2 pre-trained โมเดล (ขนาดมากกว่า 30 MB) รันบน MacBook Pro ซีพียู 8 คอร์ 3.70 GHz พบว่า ใช้เวลาประมวลผลประมาณ 58.9 มิลลิวินาทีต่อเฟรม ค่าความถูกต้อง 93% ซึ่งถ้านำเทคนิคนี้มาใช้ในโทรศัพท์มือถือซึ่งมีความสามารถในการประมวลผลน้อยกว่าคอมพิวเตอร์ก็คงจะทำให้ระยะเวลาในการประมวลผลมากขึ้น Esteban Uri [9] ได้ทดลองนำโมเดล FaceNet แปลงไปเป็นโมเดลที่สามารถรันใน TensorFlow Lite จากนั้นได้พัฒนาแอนดรอยด์แอปพลิเคชันแล้วเรียกใช้งานโมเดลดังกล่าว พบว่าโมเดลสามารถทำงานได้ดีใช้เวลาในการประมวลผล 3.5 วินาทีรันใน Google Pixel 3 ซึ่งถือว่าใช้เวลาค่อนข้างมาก

จากองค์ความรู้ที่กล่าวมาข้างต้นผู้วิจัยจึงเลือกใช้ Google ML Kit เป็นเครื่องมือสำหรับการค้นหาใบหน้าภายในภาพและใช้โมเดล MobileFaceNet ซึ่งมีความเร็วในการประมวลผลมากที่สุดและให้ค่าความถูกต้อง 99.55% เป็นโมเดลสำหรับการตรวจสอบใบหน้า

2. วัตถุประสงค์

2.1 เพื่อพัฒนาแอนดรอยด์แอปพลิเคชันสำหรับจดจำและตรวจสอบใบหน้าโดยใช้โมบายเฟซเน็ต กรณีศึกษาการเช็คชื่อเข้าร่วมกิจกรรมมหาวิทยาลัยราชภัฏร้อยเอ็ด

2.2 เพื่อเปรียบเทียบความเร็วของฟังก์ชันความเหมือน ได้แก่ ฟังก์ชันระยะห่างยูคลิเดียน (Euclidean Distance) ฟังก์ชันระยะห่างแมนฮัตตัน (Manhattan Distance) และฟังก์ชันสหสัมพันธ์โคไซน์ (Cosine Coefficient)

2.3 เพื่อหาค่า threshold ที่เหมาะสมของแต่ละฟังก์ชัน

2.4 เพื่อหาประสิทธิภาพของแอปพลิเคชันที่พัฒนาขึ้น

3. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

3.1 ชนาธิป หมั่นเพียรสุขและสุพจน์ เฮงพระพรหม [10] ได้ทำวิจัยพบว่า ฟังก์ชันแมนฮัตตันให้ประสิทธิภาพที่ดีสำหรับข้อมูลที่มีจำนวนคุณลักษณะน้อย ๆ แต่เมื่อจำนวนคุณลักษณะเพิ่มมากขึ้น ฟังก์ชันยูคลิเดียนจะให้ประสิทธิภาพที่ดีกว่า แต่โดยภาพรวมแล้วฟังก์ชันแมนฮัตตันให้ประสิทธิภาพที่ดี ในขณะที่ฟังก์ชันสหสัมพันธ์ ถ้าข้อมูลมีจำนวนคุณลักษณะน้อย ๆ ทั้งเพียร์สันและโคไซน์ให้ประสิทธิภาพที่ไม่ต่างกัน แต่เมื่อจำนวนคุณลักษณะเพิ่มมากขึ้น โคไซน์จะให้ประสิทธิภาพที่ดีกว่า

1) ฟังก์ชันยูคลิเดียน (Euclidean Distance) เป็นการวัดระยะห่างปกติระหว่างจุด 2 จุดในแนวเส้นตรง ซึ่งอาจวัดได้ด้วยไม้บรรทัด ที่ได้มาจากทฤษฎีพีทาโกรัส ระยะห่างยูคลิเดียนระหว่างจุด p และ จุด q แสดงด้วย d(p,q) คำนวณได้ดังสมการ (1)

$$d(p,q) = \sqrt{\sum_{i=1}^{n} (p_i - q_i)^2} \qquad (1)$$

โดยที่ p_i และ q_i คือ จุด 2 จุดที่ต้องการคำนวณระยะห่าง n คือ จำนวนมิติ (จุด) ของข้อมูล

ค่า d(p,q) น้อยแสดงว่า 2 จุด p และ q มีความใกล้เคียงกันมาก (หากมีค่าเป็นศูนย์ หมายถึง ทั้ง 2 จุด คือจุดจุดเดียวกัน) แต่หากมีค่ามาก แสดงว่า 2 จุดนี้ มีความห่างกันหรือแตกต่างกันมาก

2) ฟังก์ชันแมนฮัตตัน (Manhattan Distance) เป็นการวัดระยะทางระหว่างจุดสองจุดตามแกนวัดมุมขวา ซึ่งลอกเลียนมาจากตารางเค้าโครงของถนนในแมนฮัตตัน ซึ่งทำให้สามารถใช้เส้นทางที่สั้นที่สุดระหว่างจุดสองจุดในเมือง คำนวณได้ดังสมการ (2)

$$d = \sum_{i}^{n} \left| X_i - Y_i \right| \qquad (2)$$

โดยที่ X_i และ Y_i คือจุด 2 จุดที่ต้องการคำนวณระยะห่าง n คือ จำนวนมิติ (จุด) ของข้อมูล

3) ฟังก์ชันสหสัมพันธ์โคไซน์ (Cosine Coefficient) [11] หรือบางครั้งเรียกว่า ความคล้ายคลึงโคไซน์ (CosineSimilarity) เป็นการวัดความคล้ายคลึงระหว่าง 2 เวคเตอร์ โดยการวัดมุมโคไซน์ของเวคเตอร์ทั้งสอง ซึ่งคำนวณได้จากสมการ (3)

$$\cos(q,r) = \frac{\sum_v q(v)r(v)}{\sqrt{\sum_v q(v)^2}\sqrt{\sum_v r(v)^2}} \qquad (3)$$

โดยที่ q และ r คือ 2 เวคเตอร์ที่ต้องการนำมาเปรียบเทียบ ค่าสหสัมพันธ์โคไซน์จะมีค่าอยู่ระหว่าง -1 ถึง 1 โดยมีความหมายดังนี้

ถ้าค่าเข้าใกล้ 1 หมายถึง ทั้ง 2 เวคเตอร์มีความสัมพันธ์กันมากไปในทิศทางเดียวกัน

ถ้าค่าเข้าใกล้ -1 หมายถึง ทั้ง 2 เวคเตอร์มีความสัมพันธ์กันมากไปในทิศทางตรงข้ามกัน

ถ้าค่าเข้าใกล้ 0 หมายถึง ทั้ง 2 เวคเตอร์ไม่มีความสัมพันธ์กัน

3.2 Convolutional Neural Network คือ การเรียนรู้เชิงลึกมีพื้นฐานมาจาก Neural Network ซึ่งเป็นสาขาหนึ่งของเทคโนโลยี Artificial Intelligence ที่มีแนวคิดมานานตั้งแต่ปี 1943 คือ การพัฒนาเครือข่ายของอัลกอริทึมให้ทำงานแบบเดียวกับเครือข่ายระบบประสาทของมนุษย์

Convolutional Neural Network (CNN) [12] หรือโครงข่ายประสาทเทียมแบบสังวัตนาการจะจำลองการมองเห็นของมนุษย์ที่มองพื้นที่เป็นที่ย่อย ๆ และนำกลุ่มของพื้นที่ย่อย ๆ มาผสานกัน เพื่อดูว่าสิ่งที่เห็นอยู่เป็นอะไร โดยการใช้ Matrix ชนิดพิเศษที่เรียกว่า Convolution Matrix ซึ่งทำหน้าที่สกัดเอาส่วนต่าง ๆ ของภาพออกมา เช่น เส้นขอบของวัตถุต่าง ๆ เพื่อให้สามารถเรียนรู้ลักษณะของภาพได้อย่างมีประสิทธิภาพและแม่นยำ

3.3 TensorFlow [13] คือ ไลบรารีสำหรับสร้างแบบจำลองการเรียนรู้เชิงลึกซึ่งมีการนำไปประยุกต์เพื่อเพิ่มประสิทธิภาพกับผลิตภัณฑ์หลากหลายประเภท เช่น เครื่องมือค้นหา (Search Engine), การแปลภาษา (Translation), คำบรรยายภาพ (Image Captioning) และการแนะนำ (Recommendations)

TensorFlow ถูกพัฒนาโดย Google เพื่อให้นักวิจัยและนักพัฒนาทำงานกับแบบจำลอง AI ได้ เมื่อพัฒนาและปรับปรุงชักระยะหนึ่ง TensorFlow ก็ถูกปล่อยออกมาให้คนทั่วไปใช้งานได้โดยเปิดโอเพนซอร์สในปี 2015 และปล่อยตัวสมบูรณ์ออกมาในปี 2017 พร้อมลิขสิทธิ์แบบ Apache Open Source ให้คนทั่วไปสามารถใช้งาน ดัดแปลง และแจกจ่ายตัวที่ถูกดัดแปลงมาแล้วโดยที่ไม่มีค่าใช้จ่าย

TensorFlow เป็นที่นิยม เนื่องจากถูกสร้างมาเพื่อให้ทุกคนเข้าถึงได้ง่าย TensorFlow ได้รวมเอา API ที่แตกต่างกันเพื่อ

สร้างสถาปัตยกรรมแบบ Deep Learning อย่าง Convolutional Neural Network (CNN) และ Recurrent Neural Network (RNN) TensorFlow ยังมี Graph เป็นตัวคำนวณหลักเพื่อช่วยให้นักพัฒนาเห็นภาพโครงสร้าง Neural Network รวมทั้งสามารถทำงานร่วมกับ Tensorboad เครื่องมือที่ช่วยให้นักพัฒนาหาข้อผิดพลาดของโปรแกรมและสุดท้าย TensorFlow สามารถทำงานได้ทั้งบน CPU และ GPU

3.4 TensorFlow Lite [14] (TFLite) คือ Tools ที่ช่วยให้นักพัฒนาสามารถรันแบบจำลองของ TensorFlow ทำ Inference บนโทรศัพท์มือถือ Android หรือ iOS อุปกรณ์ Edge, IoT Device, Raspberry Pi, Jetson Nano, Arduino และ Microcontroller ได้เนื่องจากแบบจำลองของ TFLite มีความซับซ้อนน้อยกว่าแบบจำลองของ Tensorflow ปกติ ทำให้แบบจำลองของ TFLite มีขนาดเล็กลง ทำงานได้เร็วขึ้น แต่อย่างไรก็ตามต้องยอมรับกับความแม่นยำของแบบจำลองที่ลดลง

3.5 MobileFaceNets [7] เป็นผลงานที่ยอดเยี่ยมของนักวิจัยที่ Watchdata Inc. ในกรุงปักกิ่ง ประเทศจีน พวกเขานำเสนอโมเดล CNN ที่มีประสิทธิภาพสูงซึ่งออกแบบมาโดยเฉพาะสำหรับการตรวจสอบใบหน้าแบบเรียลไทม์ที่มีความแม่นยำสูงบนอุปกรณ์มือถือ โมเดลมีความเร็วที่น่าประทับใจด้วยความแม่นยำที่สูงมากด้วยขนาดของโมเดลเพียง 4.0 MB ความแม่นยำที่ได้รับนั้นใกล้เคียงกับโมเดลอื่น ๆ ที่ซับซ้อนกว่ามาก เช่น โมเดล FaceNet

จากรูปที่ 1 โครงสร้างของ MobileFaceNets จะใช้ภาพที่มีความละเอียด 112 × 112 พิกเซล เป็นอินพุตและเริ่มสกัดคุณสมบัติของใบหน้าจากภาพดังกล่าว ในขณะที่คุณสมบัติที่ได้รับเป็นเอาท์พุตมีขนาด 512 มิติ โมเดลมีทั้งหมด 20 เลเยอร์

Input	Operator	t	c	n	s
$112^2 \times 3$	conv3x3	-	64	1	2
$56^2 \times 64$	depthwise conv3x3	-	64	1	1
$56^2 \times 64$	bottleneck	2	64	5	2
$28^2 \times 64$	bottleneck	4	128	1	2
$14^2 \times 128$	bottleneck	2	128	6	1
$14^2 \times 128$	bottleneck	4	128	1	2
$7^2 \times 128$	bottleneck	2	128	2	1
$7^2 \times 128$	conv1x1	-	512	1	1
$7^2 \times 512$	linear GDConv7x7	-	512	1	1
$1^2 \times 512$	linear conv1x1	-	128	1	1

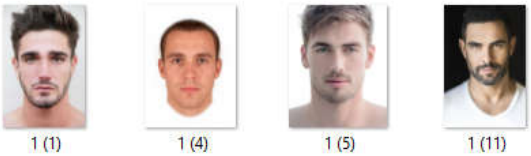
รูปที่ 1. แสดงโครงสร้างของ MobileFaceNets [7]

3.6 Google ML Kit [15] คือ SDK สำหรับ Machine Learning บน Android และ IOS ซึ่งประกอบไปด้วยสองส่วนหลัก ๆ คือ Base APIs และ Custom model ซึ่ง ML Kit นั้นสามารถใช้งานได้ทั้งแบบ Offline (On-device) และ Online (Google Cloud AI) ซึ่งแบบ Offline นั้นสามารถใช้ได้ฟรี แต่มีขีดจำกัดความสามารถอยู่บางส่วน Google ML Kit มาพร้อมกับ API พื้นฐานสำหรับงาน ML ทั่ว ๆไปหลายตัว แต่ในงานวิจัยนี้เลือกใช้เฉพาะการตรวจจับใบหน้า (Face detection) [16] ซึ่ง API ตรวจจับใบหน้าของ ML Kit จะช่วยให้เราสามารถตรวจจับใบหน้าในรูปภาพ ระบุลักษณะใบหน้าหลัก และทำความเข้าใจลักษณะของใบหน้าที่ตรวจพบได้

3.7 จีพีเอส (GPS) [17] ย่อมาจากคำว่า global positioning system หรือ ระบบตรวจสอบตำแหน่งของวัตถุ จีพีเอสจะบอกพิกัดตำแหน่งบนโลกโดยมีการติดต่อระหว่างอุปกรณ์รับสัญญาณจีพีเอสบนพื้นผิวโลกและดาวเทียมที่โคจรอยู่รอบโลก ซึ่งตำแหน่งดังกล่าวได้มาจากการคำนวณพิกัดของดาวเทียมระบุพิกัดที่ลอยอยู่ในอวกาศ จำนวน 24 ดวง ทั่วโลกการที่ระบบจีพีเอสจะทำงานได้นั้น ต้องประกอบไปด้วย 3 ส่วนหลัก ๆ คือ 1) สถานีฐาน 2) ดาวเทียมจีพีเอส และ 3) เครื่องรับสัญญาณจีพีเอส งานวิจัยนี้ใช้จีพีเอสเพื่อติดตามตำแหน่งของโทรศัพท์มือถือเพื่อให้ทราบว่าผู้ใช้งานอยู่พิกัดที่ต้องการหรือไม่

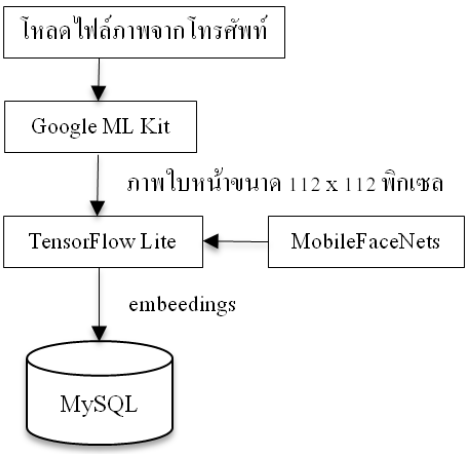
4. วิธีการวิจัย

4.1 ภาพใบหน้าที่นำมาใช้ในการทดลอง ได้นำมาจากเว็บไซต์ https://www.kaggle.com/datasets/ashwingupta3012/human-faces มีภาพใบหน้าทั้งหมด 7,219 ภาพ ซึ่งบางภาพซ้ำกัน ผู้วิจัยจึงคัดเลือกภาพที่ไม่ซ้ำกันและเป็นภาพหน้าตรง เป็นภาพผู้หญิง จำนวน 111 ภาพ และเป็นภาพผู้ชายจำนวน 399 ภาพ รวมทั้งสิ้น 510 ภาพ ทุกภาพจะถูกปรับให้ตั้งฉากเพื่อให้การประมวลผลไม่มี bias แล้วคัดลอกไปเก็บในโทรศัพท์



รูปที่ 2. ตัวอย่างภาพใบหน้าที่ใช้ทดลอง

4.2 พัฒนาแอนดรอยด์แอปพลิเคชัน โดยเรียกใช้งาน Google ML Kit, TensorFlow Lite และ MobileFaceNets เพื่อประมวลผลหาค่า embeedings (feature vector) ของแต่ละภาพจากข้อที่ 4.1 ซึ่งค่า embeedings ที่ได้จาก MobileFaceNets จะเป็นชนิดอาร์เรย์ 1 มิติ ขนาด 192 สมาชิก แต่ละสมาชิกเป็นชนิด float จากนั้นนำชื่อไฟล์ภาพและค่า embeedings ที่ได้ไปเก็บลงในฐานข้อมูล MySQL ขั้นตอนการทำงานดังรูปที่ 3



รูปที่ 3. แสดงขั้นตอนการประมวลผลหาค่า embeedings

รูปที่ 3 แสดงขั้นตอนการประมวลผลหาค่า embeedings จำนวน 1 ภาพ ซึ่งจะต้องประมวลผลจำนวน 510 รอบตามจำนวนภาพที่ใช้ทดลอง จะได้ฐานข้อมูลตามตัวอย่างในรูปที่ 4

filename	COL2	COL3	COL4	COL5	COL6
1 (1).jpg	-0.0182230980	0.0021961373	-0.0049122646	0.0061100530	-0.0229468230
1 (1005).jpg	0.0039141290	0.0003929171	-0.0130407360	0.0030874915	0.0128907060
1 (103).jpg	0.0010564298	-0.0080726845	-0.0118220460	-0.0021692850	-0.0020287314
1 (105).jpg	0.0013289162	0.0002461397	-0.0005984420	-0.0023782216	0.0012135789
1 (1053).jpg	-0.0121675220	-0.0011689715	-0.0002184557	-0.0007582008	0.0025738797
1 (1062).jpg	0.0043364903	-0.0035046497	-0.0052110700	-0.0083811890	0.0030967258

รูปที่ 4. ตัวอย่างข้อมูลที่เก็บในฐานข้อมูล

รูปที่ 4 คอลัมน์ filename หมายถึงชื่อไฟล์ภาพ COL2 COL3 หมายถึง สมาชิกของ embeedings ตัวที่ 1 และ 2 เป็นต้น จำนวนแถวเท่ากับจำนวนภาพ คือ 510 แถว

4.3 นำข้อมูลจากฐานข้อมูลตามข้อที่ 4.2 มาคำนวณหาระยะห่างโดยใช้ฟังก์ชันตามข้อที่ 3.1 ได้แก่ ฟังก์ชันยูคลิเดียน ฟังก์ชันแมนฮัตตันและฟังก์ชันสหสัมพันธ์โคไซน์ ตัวอย่างการประมวลผล เช่น ภาพที่ 1 เปรียบเทียบกับทุกภาพ ภาพที่ 2 เปรียบเทียบกับทุกภาพ ไปเรื่อย ๆ จนครบ 510 ภาพ ฉะนั้นจำนวนครั้งในการประมวลผลต่อฟังก์ชัน เท่ากับ 510x510= 260,100 ครั้ง ทำการเก็บข้อมูลระยะเวลาที่ใช้ในการประมวลผลของแต่ละฟังก์ชันเพื่อสรุปผลการทดลอง ตัวอย่างข้อมูลที่ได้แสดงดังรูปที่ 5

filename	compare	distance	distance_man	distance_cosine
1 (1).jpg	1 (1005).jpg	1.379	11.147	0.049
1 (1).jpg	1 (103).jpg	1.063	8.274	0.435
1 (1).jpg	1 (105).jpg	1.327	10.406	0.120
1 (1).jpg	1 (1053).jpg	1.231	9.799	0.243

รูปที่ 5. ตัวอย่างผลการคำนวณระยะห่าง

รูปที่ 5 ข้อมูลแถวแรก หมายถึง ไฟล์ภาพ 1 (1).jpg กับ ไฟล์ 1 (1005).jpg ระยะห่างจากฟังก์ชันยูคลิเดียน เท่ากับ 1.379 ระยะห่างจากฟังก์ชันแมนฮัตตัน เท่ากับ 11.147 และระยะห่างจากฟังก์ชันสหสัมพันธ์โคไซน์ เท่ากับ 0.049 เป็นต้น

ข้อมูลตามรูปที่ 5 แต่ละฟังก์ชันจะมีจำนวน 260,100 แถว (510x510) นำข้อมูลดังกล่าวไปวิเคราะห์หาค่า threshold หรือระยะห่างที่ให้ผลลัพธ์ถูกต้อง 100% ของแต่ละฟังก์ชัน วิเคราะห์โดยนักวิจัยเอง ผลการวิเคราะห์ดังตารางที่ 1

ตารางที่ 1. แสดงผลการวิเคราะห์ค่า threshold ที่เหมาะสม

ฟังก์ชัน	threshold
ฟังก์ชันยูคลิเดียน	<= 0.752
ฟังก์ชันแมนฮัตตัน	<= 5.796
ฟังก์ชันสหสัมพันธ์โคไซน์	>= 0.717

จากข้อมูลในตารางที่ 1 สรุปว่า ภาพ ก จะเหมือนกันกับภาพ ข ถ้าระยะห่าง <= 0.752 (ฟังก์ชันยูคลิเดียน) หรือ <= 5.796 (ฟังก์ชันแมนฮัตตัน) หรือ >= 0.717 (ฟังก์ชันสหสัมพันธ์โคไซน์) ซึ่งค่า threshold ที่ได้มาผู้วิจัยก็จะนำไปใส่ในแอปพลิเคชันต่อไป

ตารางที่ 2. เปรียบความเร็วของแต่ละฟังก์ชัน

ฟังก์ชัน	ความเร็ว (วินาที)
ฟังก์ชันยูคลิเดียน	12.880
ฟังก์ชันแมนฮัตตัน	10.261
ฟังก์ชันสหสัมพันธ์โคไซน์	23.398

จากข้อมูลในตารางที่ 2 สรุปว่า ฟังก์ชันแมนฮัตตัน มีความเร็วมากที่สุด ฉะนั้นงานวิจัยนี้จะเลือกใช้ฟังก์ชันแมนฮัตตัน โดยกำหนดค่า threshold <= 5.796

4.4 พัฒนาแอนดรอยด์แอปพลิเคชัน โดยเรียกใช้งาน Google ML Kit, TensorFlow Lite และ MobileFaceNets เพื่อใช้สำหรับการเช็คชื่อเข้าร่วมกิจกรรม โดยแอปพลิเคชัน ดังกล่าวมีความสามารถดังนี้

1) การลงทะเบียนใบหน้า

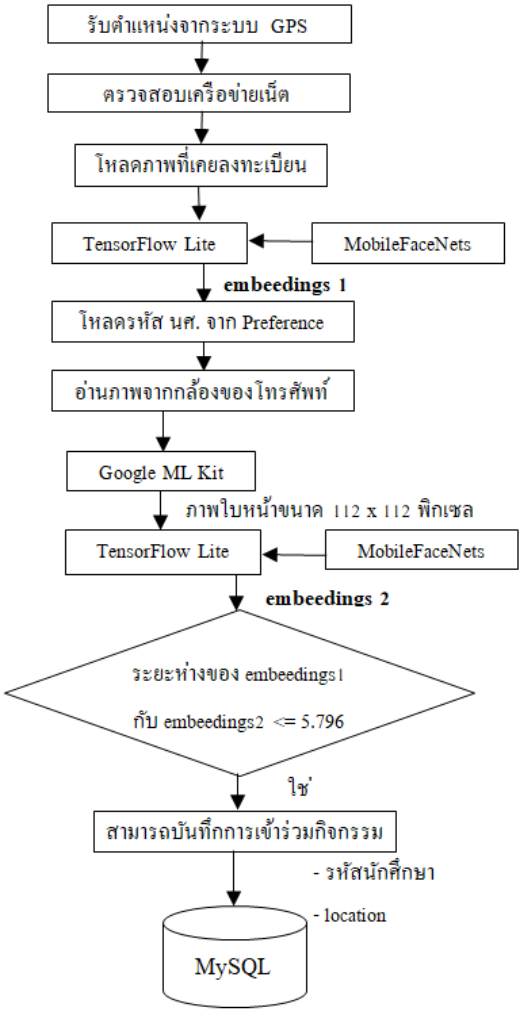
รูปที่ 6. หน้าจอหลักของแอปพลิเคชัน

จากรูปที่ 6 หมายเลข 1 ใช้สำหรับสลับกล้องหน้า-หลัง หมายเลข 2 สำหรับลงทะเบียนใบหน้าและหมายเลข 3 ใช้สำหรับบันทึกข้อมูลการเข้าร่วมกิจกรรม เมื่อผู้ใช้แตะหมายเลข 2 ผลลัพธ์ดังรูปที่ 7

รูปที่ 7. ผลลัพธ์เมื่อแตะหมายเลข 2

จากรูปที่ 7 ให้กรอกรหัสนักศึกษา ปรับใบหน้าให้ตรงแล้วแตะปุ่ม OK แอปพลิเคชันจะทำการบันทึกใบหน้าไปเป็นไฟล์ภาพขนาด 112 x 112 พิกเซล เก็บในโทรศัพท์มือถือพร้อมกับบันทึกรหัสนักศึกษาในตัวแปรชนิด Preference

2) การเช็คชื่อเข้าร่วมกิจกรรม ขั้นตอนการทำงานของโปรแกรมดังรูปที่ 8



รูปที่ 8. แสดงขั้นตอนการทำงานของแอปพลิเคชัน

4.5 นำแอปพลิเคชันไปทดสอบกับนักศึกษามหาวิทยาลัยราชภัฏร้อยเอ็ด จำนวน 30 คน วิธีการทดสอบก็คือ ให้นักศึกษาแต่ละคนใช้โทรศัพท์มือถือของตัวเองลงทะเบียนภาพใบหน้าของตัวเองจากนั้นนำโทรศัพท์มือถือไปถ่ายภาพใบหน้าของผู้ร่วมทดสอบคนอื่นแล้วสรุปผลว่าแอปพลิเคชันให้คำตอบอย่างไร รายชื่อนักศึกษาที่มาร่วมทดสอบแอปพลิเคชันมีดังตารางที่ 3 ซึ่งนักศึกษาทุกคนยินยอมให้เผยแพร่ในบทความวิจัย

ตารางที่ 3. รายชื่อนักศึกษาที่เข้าร่วมทดสอบแอปพลิเคชัน

ที่	ชื่อ-นามสกุล	สาขาวิชา
1	นายเทวา มาศวรรณา	คอมพิวเตอร์และเทคโนโลยีสารสนเทศ
2	นายชาญวิทย์ วรบุตร	คอมพิวเตอร์และเทคโนโลยีสารสนเทศ
3	นายไกรวิชญ์ แสงจันทร์ศรี	คอมพิวเตอร์และเทคโนโลยีสารสนเทศ
4	นายทศพร มาศวรรณา	คอมพิวเตอร์และเทคโนโลยีสารสนเทศ
5	น.ส.วัชราภรณ์ น้ำคำ	คอมพิวเตอร์และเทคโนโลยีสารสนเทศ
6	นายธเนศ เหง่าง่า	คอมพิวเตอร์และเทคโนโลยีสารสนเทศ
7	นายอรรถวัต พรหมคุณ	คอมพิวเตอร์และเทคโนโลยีสารสนเทศ
8	น.ส.ณัฐพร สีแสงหน่อม	คอมพิวเตอร์และเทคโนโลยีสารสนเทศ
9	น.ส.นิพาดา ธงสันเทียะ	คอมพิวเตอร์และเทคโนโลยีสารสนเทศ
10	น.ส.ปนัดดา ศรีบู้อ่ำ	คอมพิวเตอร์และเทคโนโลยีสารสนเทศ
11	นายสุวิโรจน์ ทวีศักดิ์	คอมพิวเตอร์และเทคโนโลยีสารสนเทศ
12	นายธีระศักดิ์ ยอดดี	คอมพิวเตอร์และเทคโนโลยีสารสนเทศ
13	น.ส.อนุชศรา พานาดี	การประถมศึกษา
14	น.ส.วิภาวี หล่าบุญทัน	การประถมศึกษา
15	น.ส.อมรรัตน์ จอดนอก	เคมี
16	นายธนชัย ไกยเกษ	วิทยาศาสตร์การกีฬา
17	น.ส.ละอองดาว เหลากลม	เคมี
18	นายเชษฐา ศรีขัดเค้า	วิทยาศาสตร์การกีฬา
19	นายวรภาส ละลอกน้ำ	คอมพิวเตอร์ธุรกิจ
20	น.ส.อุมาพร จันทร	การประถมศึกษา
21	นายพิชิตชัย ทับทิมไสย์	วิทยาการคอมพิวเตอร์
22	น.ส.ชลธิชา ไชยมาตย์	คอมพิวเตอร์และเทคโนโลยีสารสนเทศ
23	น.ส.ดุสิตา คาอามาตย์	คอมพิวเตอร์และเทคโนโลยีสารสนเทศ
24	นายอโรชา สมบัติศรี	วิทยาการคอมพิวเตอร์
25	นายกรกฤษ ศรีวะลา	วิทยาการคอมพิวเตอร์
26	นายสุพลชัย หงษ์ทอง	คอมพิวเตอร์และเทคโนโลยีสารสนเทศ
27	นายชยณัฐ หงส์ทอง	วิทยาศาสตร์การกีฬา
28	น.ส.ภัทราภรณ์ สุนสันเขตต์	ภาษาอังกฤษธุรกิจ
29	น.ส.รัตนาภรณ์ จำปาหอม	วิทยาการคอมพิวเตอร์
30	นายภัคพล ธรรมมนตรี	วิทยาการคอมพิวเตอร์

ผลการทดลองดังตารางที่ 4

ตารางที่ 4. ผลการทดสอบแอปพลิเคชัน

ที่	ชื่อ-นามสกุล	ถ่ายตนเอง	ถ่ายคนอื่น
1	นายเทวา มาศวรรณา	✓	✕
2	นายชาญวิทย์ วรบุตร	✓	✕
3	นายไกรวิชญ์ แสงจันทร์ศรี	✓	✕
4	นายทศพร มาศวรรณา	✓	✕
5	น.ส.วัชราภรณ์ น้ำคำ	✓	✕
6	นายธเนศ เหง่าง่า	✓	✕
7	นายอรรถวัต พรหมคุณ	✓	✕
8	น.ส.ณัฐพร สีแสงหน่อม	✓	✕
9	น.ส.นิพาดา ธงสันเทียะ	✓	✕
10	น.ส.ปนัดดา ศรีบู้อ่ำ	✓	✕

ที่	ชื่อ-นามสกุล	ถ่ายตนเอง	ถ่ายคนอื่น
11	นายสุวิโรจน์ ทวีศักดิ์	✓	×
12	นายธีระศักดิ์ ยอดดี	✓	×
13	น.ส.อนุชศรา พานาดี	✓	×
14	น.ส.วิภาวี หล่าบุญทัน	✓	×
15	น.ส.อมรรัตน์ จอดนอก	✓	×
16	นายธนชัย ไกยเกษ	✓	×
17	น.ส.ละอองดาว เหลากลม	✓	×
18	นายเชษฐา ศรีขัดเค้า	✓	×
19	นายวรภาส ละลอกน้ำ	✓	×
20	น.ส.อุมาพร จันทร	✓	×
21	นายพิชิตชัย ทับทิมไสย์	✓	×
22	น.ส.ชลธิชา ไชยมาตย์	✓	×
23	น.ส.ดุสิตา คาอามาตย์	✓	×
24	นายอโรชา สมบัติศรี	✓	×
25	นายกรกฤษ ศรีวะลา	✓	×
26	นายสุพลชัย หงษ์ทอง	✓	×
27	นายชยณัฐ หงส์ทอง	✓	×
28	น.ส.ภัทราภรณ์ สุนสันเขตต์	✓	×
29	น.ส.รัตนาภรณ์ จำปาหอม	✓	×
30	นายภัคพล ธรรมมนตรี	✓	×

หมายเหตุ ✓ หมายถึง แอปฯ สรุปใบหน้าที่ถ่ายตรงกับใบหน้าที่ลงทะเบียน × หมายถึง แอปฯ สรุปว่าไม่ตรงกัน

5. สรุปผลการทดลองและการอภิปราย

บทความนี้นำเสนอการพัฒนาแอนดรอยด์แอปพลิเคชันสำหรับจดจำและตรวจสอบใบหน้าโดยใช้โมบายเฟซเน็ต กรณีศึกษา การเช็คชื่อเข้าร่วมกิจกรรมมหาวิทยาลัยราชภัฏร้อยเอ็ด ซึ่งมีวัตถุประสงค์ 4 ข้อ ขอสรุปผลการทดลองตามวัตถุประสงค์ดังนี้

1) การพัฒนาแอนดรอยด์แอปพลิเคชันสำหรับจดจำและตรวจสอบใบหน้าโดยใช้โมบายเฟซเน็ต กรณีศึกษาการเช็คชื่อเข้าร่วมกิจกรรมมหาวิทยาลัยราชภัฏร้อยเอ็ด สามารถนำไปใช้งานได้จริง การใช้งานแอปพลิเคชัน ดังรูปที่ 6 และรูปที่ 7

2) จากข้อมูลตารางที่ 2 สรุปว่า ฟังก์ชันแมนฮัตตัน มีความเร็วมากที่สุด รองลงมาได้แก่ ฟังก์ชันยูคลิเดียนและฟังก์ชันสหสัมพันธ์โคไซน์ ตามลำดับ ฉะนั้นงานวิจัยนี้จึงเลือกใช้งานฟังก์ชันแมนฮัตตันเพื่อประสิทธิภาพที่ดีที่สุดของแอปพลิเคชัน

3) จากข้อมูลในฐานข้อมูลตามรูปที่ 5 นำไปวิเคราะห์เพื่อหาค่า threshold หรือระยะห่างที่ให้ผลลัพธ์ถูกต้อง 100% ของแต่ละฟังก์ชัน พบว่า ภาพ ก จะเหมือนกันกับภาพ ข ถ้าระยะห่าง <= 0.752 (ยูคลิเดียน) หรือ <= 5.796 (แมนฮัตตัน) หรือ >= 0.717 (สหสัมพันธ์โคไซน์) การวิเคราะห์จะวิเคราะห์โดยใช้ผู้วิจัยเอง

4) ผลการหาประสิทธิภาพของแอปพลิเคชัน ตามตารางที่ 4 พบว่า แอปพลิเคชัน ประมวลผลถูกต้อง 100%

6. ข้อเสนอแนะ

1) ในปัจจุบันระบบปฏิบัติการที่มีการใช้งานในโทรศัพท์มือถือนอกจากระบบปฏิบัติการแอนดรอยด์แล้ว ยังมีระบบปฏิบัติการ iOS ของค่าย apple อีกหนึ่งระบบปฏิบัติการที่มีผู้ใช้งานจำนวนหนึ่งในสัดส่วนที่น้อยกว่าระบบปฏิบัติการแอนดรอยด์ ฉะนั้นแนวทางในการพัฒนาแอปพลิเคชันที่มีวัตถุประสงค์เดียวกันบนระบบปฏิบัติการ iOS จึงเป็นแนวทางที่น่าสนใจอีกหนึ่งประเด็นงานวิจัย

เอกสารอ้างอิง

[1] R. L. Galvez, A. A. Bandala, E. P. Dadios, R. R. P. Vicerra and J. M. Z. Maningo, "Object detection using convolutional neural networks," TENCON 2018 – 2018 IEEE Region 10 Conference, pp. 2023-2027, 2018.

[2] ภูมิภักดิ์ พรมรังกา, สหรัฐ แหวนทอง และ ชูพันธุ์ รัตนโภคา, "โมบายแอปพลิเคชันสำหรับจำแนกสายพันธุ์นกด้วยวิธีการเรียนรู้เชิงลึก," Journal of Information Science and Technology (JIST) Volume 12, NO 1, pp. 37-46, JAN – JUN 2022.

[3] นัศพ์ชาณัณ ชินปัญช์ธนะ, "แบบจำลองการเรียนรู้ท่าทางมนุษย์จากการเคลื่อนไหวด้วยเทคนิคโครงข่ายประสาทเทียมแบบสังวัตนาการและแบบหน่วยความจำระยะสั้นแบบยาว," Journal of Information Science and Technology (JIST) Volume 12, NO 1, pp. 27-36, JAN – JUN 2022.

[4] วัชรชัย คงศิริวัฒนา, พลกฤต หวังศิริกำโชค และวายุภักดิ์ ชันติโก, "ระบบตรวจสอบการเข้าชั้นเรียนและประเมินความสนใจผ่านลักษณะอารมณ์ทางใบหน้าด้วยกล้องเว็บแคม," วารสารวิทยาศาสตร์ลาดกระบัง ปีที่ 30, ฉบับที่ 2, หน้า 42-57, กรกฎาคม-ธันวาคม 2564.

[5] Salinda Hettiarachchi, " Analysis of different face detection and recognition models for Android, " M.S. thesis, Mid Sweden University, Faculty of Science, Sweden, 2021.

[6] F. Schroff, D. Kalenichenko, and J. Philbin, "FaceNet: A unified embedding for face recognition and clustering," IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, pp. 815–823, 2015.

[7] Sheng Chen, Yang Liu, Xiang Gao and Zhen Han, "MobileFaceNets: Efficient CNNs for Accurate Real-Time Face Verification on Mobile Devices," Apr, 2018. [Online], Available: https://arxiv.org/ftp/arxiv/papers/1804/1804.07573.pdf. [Accessed Aug. 1,2022].

[8] A. Rosebrock, "OpenCV Face Recognition," Sep, 2018. [Online], Available: https://www.pyimagesearch.com/2018/09/24/opencv-face-recognition. [Accessed Aug.1,2022].

[9] E. Uri, "Converting Sandberg's FaceNet pre-trained model to TensorFlow Lite (using an "unorthodox way")," Jun 19, 2020. [Online], Available: https://medium.com/@ estebanuri/converting-sandbergs-facenet-pre-trained-model-to-tensorflow-lite-using-an-unorthodox-way- 7ee3a6ed02a3. [Accessed Aug. 1,2022].

[10] ชนาธิป หมั่นเพียรสุขและสุพจน์ เฮงพระพรหม, "ประสิทธิภาพของฟังก์ชันความเหมือนต่อขั้นตอนวิธีเพื่อนบ้านใกล้ที่สุดเคสำหรับการจำแนกประเภทข้อมูล," การประชุมวิชาการระดับชาติ ครั้งที่ 8 มหาวิทยาลัยราชภัฏนครปฐม, นครปฐม, หน้า 473-479, 2559.

[11] Lillian Lee, "Measures of distributional similarity," Proceedings of the 3 7 th annual meeting of the Association for Computational Linguistics on Computational Linguistics, pp. 25-32, 1999.

[12] S. Albawi, T. A. Mohammed and S. Al-Zawi, "Under standing of a convolutional neural network," In International Conference on Engineering and Technology (ICET), pp. 1-6, 2017.

[13] "TensorFlow," TensorFlow, 2015. [Online]. Available:https://www.tensorflow.org. [Accessed Aug. 1,2022].

[14] "TensorFlowLite," TensorFlow, 2018. [Online]. Available: https://www.tensorflow.org/lite/guide. [Accessed Aug. 1,2022].

[15] "ML Kit คืออะไร," www.mindphp.com, 2562. [Online]. Available: https://www.mindphp.com/คู่มือ/73-คืออะไร/7110-what-is-ml-kit.html. [Accessed Aug. 1,2022].

[16] "Face detection," developers.google.com, 2022. [Online]. Available: https://developers.google.com/ml-kit/vision/face-detection. [Accessed Aug. 1,2022].

[17] ปุญญรัตน์ ปุญญา, "การพัฒนา MSU–GPS สำหรับการเดินทางในมหาวิทยาลัยมหาสารคาม," ปริญญาวิทยาศาสตรมหาบัณฑิต สาขาวิชาสื่อนฤมิต,มหาวิทยาลัยมหาสารคาม, 2557.

JIST Journal of Information Science and Technology
Volume 13, NO 1 | JAN – JUN 2023 | 10-16
ISSN: 2651-1053 (Online)
ISSN: 1906-9553 (Print)

การพัฒนาระบบวิเคราะห์และนับจำนวนผลมะนาวด้วยเทคโนโลยีคอมพิวเตอร์วิทัศน์

Development of Counting and Analysis Lemons System with Computer Vision Technology

เอกรัตน์ สุขสุคนธ์

Aekkarat Suksukont

สาขาเทคโนโลยีคอมพิวเตอร์และนวัตกรรม คณะวิทยาศาสตร์และเทคโนโลยี

มหาวิทยาลัยเซาธ์อีสท์บางกอก

Program in Computer Technology and Innovation, Faculty of Science and Technology,

Southeast Bangkok University

Received: January 11, 2023; Revised: June 14, 2023; Accepted: June 24, 2023; Published: June 30, 2023

ABSTRACT – This article presents development of counting and analysis lemons system with computer vision technology for test the performance of analysis system for color and counting of lemons to facilitate farmers. Firstly, designing the system and training the system with images of soft lemons and mature lemons 574 images, which image processing techniques were applied to this process. Then test the performance of the pressing system with 400 lemons. Divided into 200 soft lemons and 200 mature lemons. From the experiment counting of soft lemons system analysis accuracy 92% there is an error of 8% and counting of mature lemons system analysis accuracy 98% there is an error of 2% respectively.

KEYWORDS: Image processing, Comparison, Fruit, Color model, Computer vision

บทคัดย่อ -- **บทความนี้นำเสนอการพัฒนาระบบวิเคราะห์และนับจำนวนผลมะนาวด้วยเทคโนโลยีคอมพิวเตอร์วิทัศน์ เพื่อนำไปใช้ทดสอบประสิทธิภาพการทำงานของระบบวิเคราะห์สีและนับผลของมะนาวเพื่อช่วยอำนวยความสะดวกให้แก่เกษตรกร เริ่มจากการออกแบบระบบและเทรนระบบด้วยภาพของผลมะนาวอ่อนและผลมะนาวแก่ จำนวน 574 ภาพ ซึ่งเทคนิคการประมวลผลภาพถูกนำมาใช้กระบวนการนี้ จากนั้นทดสอบประสิทธิภาพของระบบด้วยการนำผลมะนาว จำนวน 400 ลูก โดยแบ่งเป็นผลมะนาวอ่อน จำนวน 200 ลูก และผลมะนาวแก่ จำนวน 200 ลูก จากการทดลอง พบว่า ระบบสามารถวิเคราะห์และนับผลมะนาวอ่อนมีความแม่นยำ 92 เปอร์เซนต์ มีความผิดพลาด 8 เปอร์เซนต์ และสามารถวิเคราะห์และนับผลมะนาวแก่มีความแม่นยำ 98 เปอร์เซนต์มีความผิดพลาด 2 เปอร์เซนต์ ตามลำดับ**

คำสำคัญ: การประมวลผลภาพ, การเปรียบเทียบ, ผลไม้, แบบจำลองสี, คอมพิวเตอร์วิทัศน์

Corresponding Author: ekk_ele@hotmail.com

1. บทนำ

มะนาวเป็นพืชเศรษฐกิจที่สำคัญในประเทศไทยที่เกษตรกรในหลายพื้นที่นิยมปลูกและยึดเป็นอาชีพหลัก เพื่อจำหน่ายทั้งในประเทศและส่งออกจำหน่ายยังต่างประเทศ ประเทศไทยถือได้ว่ามีพื้นที่ที่เหมาะสมในการทำการเกษตรกรรม ซึ่งในปัจจุบันความเจริญก้าวหน้าของโลกยุคปัจจุบันได้มีการนำเอาเทคโนโลยีมาใช้ในชีวิตประจำวันในหลาย ๆ ด้าน เช่น ด้านโรงงานอุตสาหกรรม หน่วยงานราชการสำนักงาน บริษัทต่าง ๆ หรือแม้แต่ด้านการเกษตร ทั้งนี้ผู้วิจัยมีความสนใจที่จะสร้างระบบวิเคราะห์และนับจำนวนผลมะนาวอ่อนและผลมะนาวแก่เพื่อช่วยอำนวยความสะดวกแก่เกษตรกรในหลายพื้นที่ที่ยึดการปลูกผลมะนาวเป็นอาชีพในการสร้างรายได้และช่วยอำนวยความสะดวกในการวิเคราะห์และนับผลมะนาว

จากงานวิจัยของ [1] การพัฒนาเครื่องต้นแบบคัดแยกดอกดาวเรืองโดยการประมวลผลภาพ เป็นงานวิจัยที่ได้จากการพัฒนาต่อจากโปรแกรมวิเคราะห์ขนาดของดอกดาวเรืองด้วยวิธีการประมวลผลภาพ โดยการใช้การหาสัณฐาน (Shape) ของดอกดาวเรืองและวิธีการจัดกลุ่มด้วยการวัดระยะห่างของข้อมูล (Euclidean Distance) ซึ่งสามารวิเคราะห์ขนาดของดอกดาวเรืองได้ 4 ขนาด ระบบสายพานคัดแยกดอกดาวเรืองขึ้นโดยประกอบด้วยชุดสายพาน 2 ชุด ได้แก่ ชุดสายพานวางและบีบดอกดาวเรือง และชุดสายพานคัดแยกดอกดาวเรือง โดยควบคุมการทำงานของอุปกรณ์ด้วยคอมพิวเตอร์ขนาดเล็ก (Raspberry Pi 3 Model B) และรับส่งข้อมูลระหว่าโปรแกรมกับชุดสายพานผ่านอุปกรณ์สื่อสารไร้สายระยะใกล้หรือโมดูลบูลทูธ งานวิจัยของ [2] เครื่องคัดแยกขนาดมะนาว มีจุดมุ่งหมายเพื่อสร้างเครื่องคัดแยกขนาดมะนาว เพื่อใช้ในอุตสาหกรรมขนาดเล็กช่วยลดต้นทุนและประหยัดเวลาในการคัดแยกขนาดมะนาว จากการทดสอบ ผลปรากฏว่าเครื่องคัดแยกขนาดมะนาวสามารถทำงานได้อย่างต่อเนื่องสามารถคัดแยกมะนาวได้มีประสิทธิภาพ งานวิจัยของ [3] ระบบการวิเคราะห์การสุกของผลไม้ด้วยเทคโนโลยีการประมวลผลภาพ โดยมีวัตถุประสงค์เพื่อตรวจสอบการสุกของผลไม้ตามระยะเวลาจริงได้ และช่วยอำนวยความสะดวกให้แก่เกษตรกร เพื่อลดปัญหาการเน่าเสียของผลไม้ที่อาจจะเกิดขึ้น ในงานวิจัยนี้ใช้ซอฟต์แวร์ NI Vision Builder ทำงานร่วมกับกล้องวิดีโอ และใช้เทคโนโลยีการประมวลผลภาพ เข้ามาช่วยในการวิเคราะห์และประมวลผล จากการทดลองพบว่า ระบบการวิเคราะห์การสุกของผลไม้ด้วยเทคโนโลยีการประมวลผลภาพ สามารถนับจำนวนของผลไม้มีความถูกต้อง คิดเป็น 100 เปอร์เซนต์ และสามารถวิเคราะห์การสุกของผลไม้ คิดเป็น 95 เปอร์เซนต์ งานวิจัยของ [4] เครื่องคัดแยกสีอัตโนมัติบนระบบสายพานลำเลียงควบคุมด้วยอาดูยโน่ (Arduino UNO R3) เป็นงานวิจัยที่ได้ออกแบบตัวเครื่องที่สามารถปฏิบัติงานให้แยกชนิดวัตถุที่เป็นสีแดง สีเขียว สีน้ำเงินและสีอื่น ๆ โดยเคลื่อนย้ายวัตถุไปยังถาดที่กำหนดไว้อัตโนมัติอาศัยการควบคุมของอาดูยโน่ ผลการวิจัยพบว่า สามารถคัดแยกวัตถุสีแดง สีเขียว สีน้ำเงินและสีอื่นๆ โดยเวลาเฉลี่ยของวัตถุที่ทดสอบปรากฏว่าวัตถุสีแดง 9.96 วินาที วัตถุสีเขียว 13.47 วินาที วัตถุสีน้ำเงิน 16.58 วินาทีและวัตถุอื่น ๆ 9.04 วินาที และสามารถหมุนถาดเพื่อรับวัตถุในตำแหน่งที่ต้องการได้ถูกต้อง งานวิจัยของ [5] เครื่องคัดแยกน้ำหนักและสีมะม่วงอัตโนมัติด้วยไมโครคอนโทรลเลอร์ โดยเครื่องสามารถคัดแยกได้ 2 ส่วน คือ น้ำหนักและสีของผลมะม่วง โดยการใช้สายพานในการลำเลียงและตรวจจับน้ำหนักด้วยโหลดเซลล์และนำมาต่อกับสายพานจะลำเลียง เพื่อวัดค่าสีหาความสุกดิบของมะม่วง หลังจากนั้นจะส่งข้อมูลไปยังไมโครคอนโทรลเลอร์ (Mega 2560) ซึ่งทำหน้าที่ประมวลผลและสั่งเปิดช่องรับสำหรับคัดแยกมะม่วงตามที่กำหนดไว้ ผลการทดลองคัดแยกมะม่วงโดยวัดเป็นค่าความผิดพลาดโดยใช้ตัวอย่างมะม่วงพันธุ์เขียวเสวยจำนวน 100 ผล ผลการทดสอบคัดเกรดมะม่วงมีความผิดพลาดเท่ากับ 0.00 เปอร์เซนต์ และการคัดแยกความสุกดิบ มีความผิดพลาด เท่ากับ 0.00 เปอร์เซนต์ ส่วนมะม่วงพันธุ์น้ำดอกไม้สีทองจำนวน 100 ผล ผลการทดสอบคัดเกรดมะม่วงมีความผิดพลาด เท่ากับ 3 เปอร์เซนต์ และการคัดแยกความสุกดิบ มีความผิดพลาด เท่ากับ 3 เปอร์เซนต์ งานวิจัยของ [6], [7] นำเสนอเกี่ยวกับระบบคัดแยกสีของวัตถุต่างชนิดกัน โดยใช้เทคนิคการสกัดคุณลักษณะเด่น (Feature) ของแบบจำลองสีชนิดต่าง ๆ มาใช้เป็นส่วนสำคัญในการระบุชนิดของวัตถุนั้น ๆ งานวิจัยของ [8] และ [9] ในงานวิจัยเหล่านี้ได้นำหลักการประมวลผลภาพมาใช้ในการแยกแยะชนิดของผลไม้ชนิดต่าง ๆ เพื่อให้ระบบสามารถระบุชนิดของผลไม้ โดยที่สามารถแยกแยะได้ไม่ว่าจะเป็นผลไม้ชนิดเดียวกันหรือผลไม้ต่างชนิดกันก็ตาม

จากงานวิจัยที่กล่าวมาข้างต้น ผู้วิจัยจึงได้ออกแบบและพัฒนาระบบวิเคราะห์และนับจำนวนผลมะนาวด้วยเทคโนโลยีคอมพิวเตอร์วิทัศน์ เพื่อศึกษากระบวนการประมวลผลด้วยภาพและนำไปใช้ในระบบวิเคราะห์และนับจำนวนผลมะนาวอ่อน

ผลและมะนาวแก่ ทฤษฎีนี้จะทำให้ทราบค่าของสีผลมะนาวอ่อนและผลมะนาวแก่และช่วยในการวิเคราะห์ค่าสีร่วมกับการใช้ภาษา Java และนำเทคโนโลยีคอมพิวเตอร์วิทัศน์มาประยุกต์ใช้งานร่วมกันในการสร้างโปรแกรมวิเคราะห์และนับจำนวนผลของมะนาวอีกด้วย

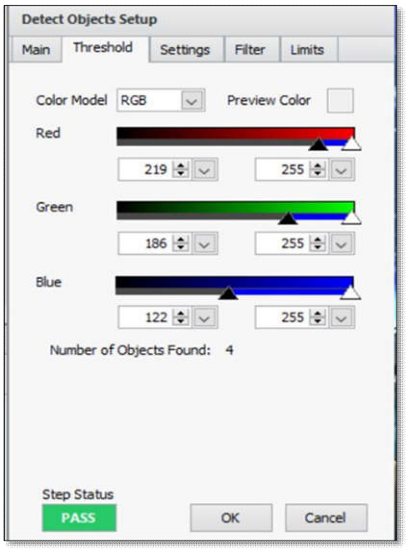
2. ทฤษฎี

2.1 การประมวลผลภาพ (Image Processing)

การประมวลผลภาพ คือ กระบวนการจัดการและวิเคราะห์รูปภาพให้เป็นข้อมูลในแบบดิจิทัล โดยใช้ระบบของคอมพิวเตอร์เพื่อให้ได้ข้อมูลที่เราต้องการทั้งในเชิงคุณภาพและปริมาณ (ขนาด รูปร่าง สี) หลังจากนั้นเราสามารถนำข้อมูลไปวิเคราะห์ และสร้างเป็นระบบต่าง ๆ เช่น ระบบดูแลการจราจรบนท้องถนน ระบบตรวจสอบคุณภาพของผลิตภัณฑ์ในกระบวนการผลิต ระบบเก็บข้อมูลรถที่เข้าและออกอาคาร ระบบตรวจจับใบหน้า เป็นต้น การประมวลผลภาพเกิดจากการกระทำการอย่างใดอย่างหนึ่งกับภาพต้นฉบับ (Input Image) เพื่อให้ได้ภาพผลลัพธ์ (Output Image) มีลักษณะของภาพเป็นไปตามที่ต้องการ ซึ่งการกระทำการกับภาพที่ใช้ในการประมวลผลภาพดิจิตอลมีอยู่มากมายหลายแบบ ความเข้าใจเกี่ยวกับคุณลักษณะและการแยกแยะประเภทของการกระทำการกับภาพ จะช่วยให้เราสามารถคาดคะเนภาพผลลัพธ์ที่จะได้จากการกระทำการแต่ละแบบ หรือประมาณความซับซ้อนของการกระทำการกับภาพที่จะนำไปใช้ได้การกระทำการกับภาพในการประมวลผลภาพดิจิตอล สามารถแบ่งออกได้ 2 ประเภทคือ การประมวลผลภาพแบบจุด (Point Image Processing) และการประมวลผลภาพแบบบริเวณ (Local Image Processing)

2.2 แบบจำลองสี (Color Model)

แบบจำลองสี เป็นสิ่งที่ใช้อ้างอิงถึงระดับสีต่างๆ สำหรับคอมพิวเตอร์แล้วเราจะไม่ใช้แบบจำลองที่เป็น Analytical Model เหมือนกับระบบที่ใช้ในทางวิทยาศาสตร์ ซึ่งใช้วิธีการวัดระดับที่อยู่ในรูปของพลังงานช่วงของสเปคตรัม (Spectrum) จากการทดลองที่เป็นการศึกษาแบบ Psychophysical ที่มีการรับรู้ของมนุษย์เข้ามาเกี่ยวข้องแบบจำลองสีมีหลายแบบด้วยกัน เช่น แบบจำลองสี RGB แบบจำลองสี HSV และแบบจำลองสี YCbCr เป็นต้น รูปที่ 1 เป็นการกำหนดค่าสี RGB เพื่อใช้ในการวิเคราะห์ผล



รูปที่ 1. ตัวอย่างแบบจำลองสี RGB

2.3 คอมพิวเตอร์วิทัศน์ (Computer Vision)

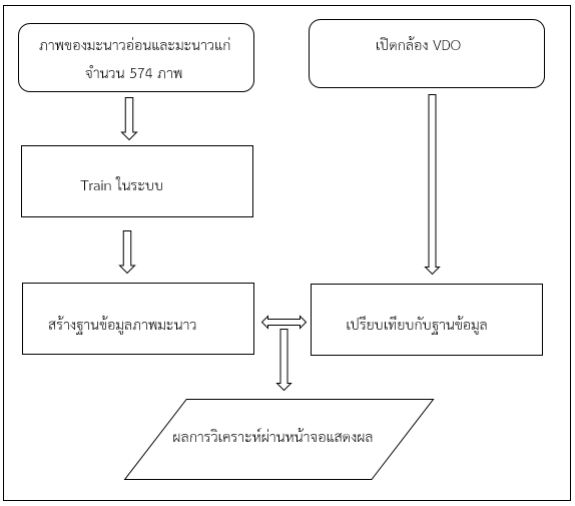
คอมพิวเตอร์วิทัศน์ได้รวมเอากล้องวิดีโอ, การประมวลผลที่ Edge หรือคลาวด์, ซอฟต์แวร์ และปัญญาประดิษฐ์ (AI) เข้าไว้ด้วยกัน เพื่อทำให้ระบบสามารถตรวจจับและระบุวัตถุได้ เทคโนโลยีมากมายที่ช่วยให้ AI รวมถึง CPU สำหรับการประมวลผลทั่วไป และคอมพิวเตอร์วิทัศน์และหน่วยประมวลผลวิสัยทัศน์ (VPU) ในการส่งมอบการเร่งความเร็ว ระบบคอมพิวเตอร์วิทัศน์มีประโยชน์ในสภาพแวดล้อมที่หลากหลาย สามารถจดจำวัตถุและผู้คนได้อย่างรวดเร็ว วิเคราะห์ข้อมูลประชากรศาสตร์ ตรวจสอบผลิตภัณฑ์ที่ผลิตและอื่น ๆ อีกมากมาย คอมพิวเตอร์วิทัศน์ใช้การเรียนรู้เชิงลึกเพื่อสร้างเครือข่ายประสาทที่นำทางระบบในการประมวลผลและการวิเคราะห์ภาพ เทคนิคโครงข่ายประสาทแบบคอนโวลูชัน (CNN) และมีชุดเครื่องมือ Open VINO ที่ช่วยให้มีการอนุมานการเรียนรู้เชิงลึกสำหรับการจัดประเภทรูปภาพ และการตรวจจับวัตถุ เมื่อได้รับการฝึกอบรมอย่างเต็มรูปแบบ โมเดลคอมพิวเตอร์วิทัศน์จะสามารถทำการรับรู้วัตถุ ตรวจจับและจดจำบุคคล หรือแม้กระทั่งติดตามการเคลื่อนไหว

3. การดำเนินการทดลอง

3.1 การออกแบบ

การพัฒนาระบบวิเคราะห์และนับจำนวนผลมะนาวด้วยเทคโนโลยีคอมพิวเตอร์วิทัศน์ ได้นำทฤษฎีการประมวลผลภาพหรือ Computer Vision เข้ามาใช้งานในระบบ ซึ่งการเทรนภาพ

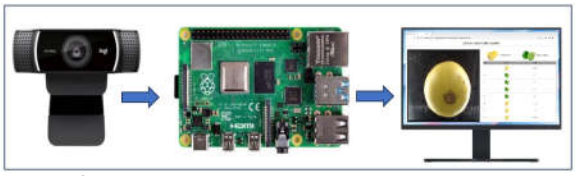
ของมะนาวอ่อนและมะนาวแก่ ในงานวิจัยนี้จะใช้ภาพของมะนาวแป้น ซึ่งในประเทศไทยนิยมปลูกมะนาวสายพันธุ์นี้กันอย่างแพร่หลาย โดยมีกรอบแนวคิด ดังรูปที่ 2



รูปที่ 2. แผนภาพกรอบแนวคิดของระบบ

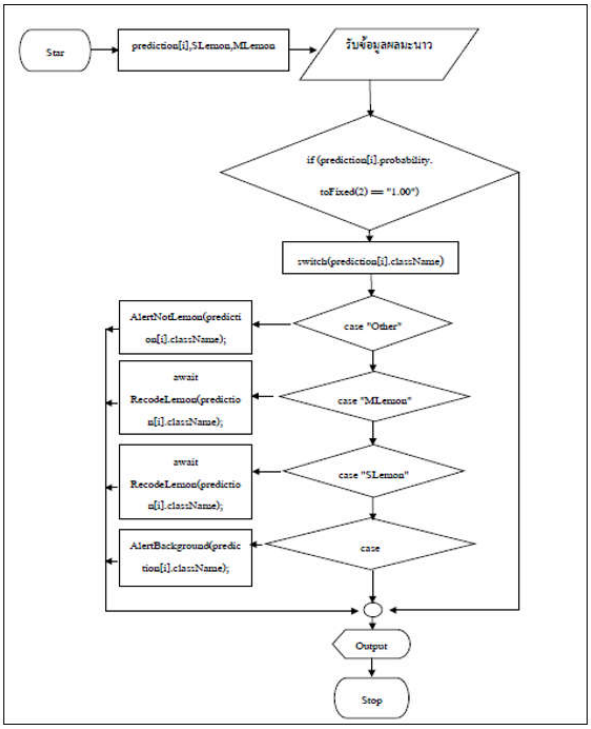
3.2 วิธีดำเนินการวิจัย

1) การทำงานของระบบวิเคราะห์และนับจำนวนผลมะนาวด้วยเทคโนโลยีคอมพิวเตอร์วิทัศน์ ต้องทำการกำหนดลักษณะการทำงานของระบบควบคุมการทำงานของอุปกรณ์อินพุต และเอาต์พุต กำหนดโครงสร้างของระบบวิเคราะห์และนับจำนวนผลมะนาวด้วยเทคโนโลยีคอมพิวเตอร์วิทัศน์ โดยทำการเชื่อมต่อกล้องวิดีโอ (Web Camera) เข้ากับคอมพิวเตอร์ขนาดเล็ก (Raspbery Pi 4) จากนั้นเชื่อมต่อกับจอสดงผล (Monitor) ดังรูปที่ 3



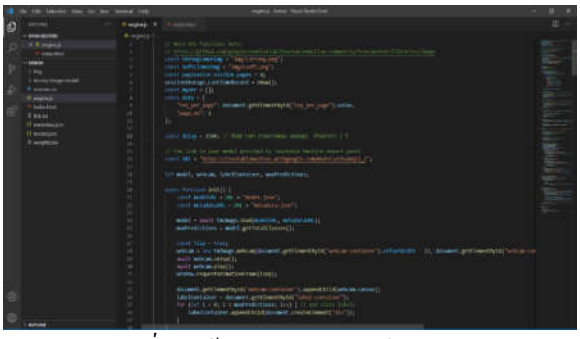
รูปที่ 3. การต่อระบบวิเคราะห์และนับจำนวนผลมะนาว

การกำหนดสถานะของผลลัพธ์แบ่งออกเป็น 4 กรณี คือ กรณีผลมะนาวอ่อน กรณีผลมะนาวแก่ กรณีจับภาพพื้นหลังและกรณีวัตถุอื่น โดยที่หากการรับภาพจากกล้องวิดีโอและเข้าสู่กระบวนการประมวลผล ค่าใดที่มีผลลัพธ์ใกล้เคียงกับภาพที่ทำการเทรนจะแสดงค่าเป็น 1 ซึ่งคือผลลัพธ์ที่แสดงว่าเป็นผลของมะนาวอ่อนหรือผลของมะนาวแก่ ดังรูปที่ 4



รูปที่ 4. การทำงานของระบบวิเคราะห์และนับจำนวนผลมะนาว

2) การออกแบบซอฟต์แวร์ โดยใช้ Visual Studio Code (VS Code) ในการเขียนภาษา Node.js หรือ JavaScript เพื่อสร้างระบบวิเคราะห์และนับจำนวนผลมะนาวด้วยเทคโนโลยีคอมพิวเตอร์วิทัศน์ ดังรูปที่ 5 โดยจะทำการวิเคราะห์จากสีและรูปทรงข้อมูลที่นำมาวิเคราะห์จะถูกเก็บอยู่ในคลาส (Class) ที่ได้สร้างไว้ใน Teachable Machine จากนั้นจะทำการเก็บข้อมูลที่วิเคราะห์ได้ไว้ในตารางเก็บค่า เพื่อทำการสรุปผลว่ามีผลมะนาวอ่อนและผลมะนาวแก่ทั้งหมดกี่ผลแล้วจึงแสดงผลออกทางหน้าต่างของโปรแกรม



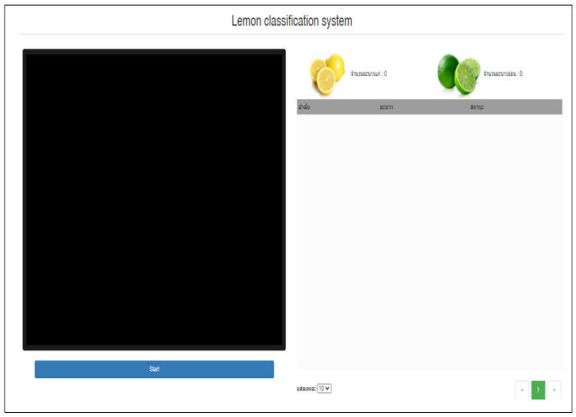
รูปที่ 5. หน้าต่างของการเขียนโปรแกรม

3) หลังจากนั้นนำภาพผลมะนาวอ่อนและผลมะนาวแก่ในระดับสีที่แตกต่างกัน จำนวน 574 ภาพ นำมาทำการเทรนใน

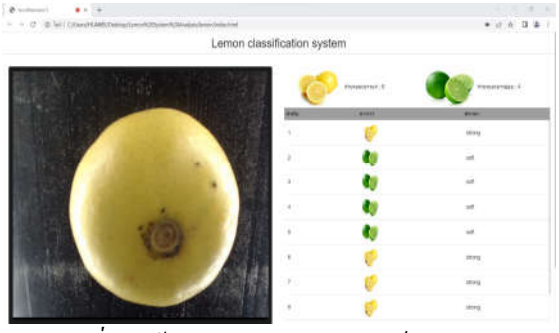
ระบบ เพื่อสร้างเป็นฐานข้อมูลของผลมะนาวในลักษณะการวางผลของมะนาวที่แตกต่างกัน โดยมีระยะการถ่ายภาพห่างจากกล้อง 30 เซนติเมตร ดังรูปที่ 6

รูปที่ 6. กระบวนการเทรนภาพผลมะนาวอ่อนและผลมะนาวแก่

4) การทดสอบประสิทธิภาพของระบบวิเคราะห์และนับจำนวนผลมะนาวด้วยเทคโนโลยีคอมพิวเตอร์วิทัศน์จะนำมาใช้งานร่วมกับกล้องวิดีโอ ซึ่งสามารถใช้งานแบบเวลาจริง (Real Time) โดยนำผลมะนาวจำนวนทั้งหมด 400 ลูก มาทดสอบผ่านกล้องวิดีโอ จากนั้นระบบจะทำการวิเคราะห์และนับผลมะนาวแสดงผลผ่านหน้าต่างของโปรแกรม ดังรูปที่ 7 ถึงรูปที่ 9



รูปที่ 7. หน้าต่างของโปรแกรมวิเคราะห์และนับผลมะนาว



รูปที่ 8. หน้าต่างแสดงผลการวิเคราะห์ผลมะนาว

รูปที่ 9. หน้าต่างแสดงผลการนับผลมะนาว

4. ผลการศึกษา

จากการทดลอง เมื่อทดสอบประสิทธิภาพของระบบวิเคราะห์และนับจำนวนผลมะนาวด้วยเทคโนโลยีคอมพิวเตอร์วิทัศน์ เมื่อนำผลมะนาวอ่อนและผลมะนาวแก่ รับภาพผลมะนาวผ่านกล้องวิดีโอและนำไปวิเคราะห์ผลผ่านคอมพิวเตอร์ขนาดเล็ก ซึ่งผลการวิเคราะห์และนับจำนวนผลมะนาวอ่อน ดังตารางที่ 1 และผลการวิเคราะห์และนับจำนวนผลมะนาวแก่ ดังตารางที่ 2 ตามลำดับ

ตารางที่ 1 ผลการวิเคราะห์และนับจำนวนผลมะนาวอ่อน

ครั้งที่	จำนวนมะนาวอ่อน	จำนวนที่วิเคราะห์ได้	จำนวนที่วิเคราะห์ไม่ได้	ความแม่นยำ (เปอร์เซนต์)
1	40	38	2	95.00
2	80	74	6	92.50
3	120	110	10	91.66
4	160	148	12	92.50
5	200	184	16	92.00

ตารางที่ 2 ผลการวิเคราะห์และนับจำนวนผลมะนาวแก่

ครั้งที่	จำนวนมะนาวแก่	จำนวนที่วิเคราะห์ได้	จำนวนที่วิเคราะห์ไม่ได้	ความแม่นยำ (เปอร์เซนต์)
1	40	40	0	100.00
2	80	80	0	100.00
3	120	118	2	98.33
4	160	158	2	98.75
5	200	196	4	98.00

จากตารางที่ 1 ผลการวิเคราะห์และนับจำนวนผลมะนาวอ่อน พบว่า ระบบสามารถนับจำนวนของผลมะนาวอ่อนโดยรวมเมื่อนำจำนวนของผลมะนาวจำนวน 200 ลูก มาวิเคราะห์ในระบบ ผลปรากฏว่าระบบสามารถวิเคราะห์ผลมะนาวอ่อนได้ 184 ลูก และไม่สามารถวิเคราะห์ผลมะนาวอ่อนได้ จำนวน 16 ลูก คิดเป็นความแม่นยำของระบบอยู่ที่ 92 เปอร์เซนต์ จากตารางที่ 2 ผลการวิเคราะห์และนับจำนวนผลมะนาวแก่ พบว่า ระบบสามารถนับจำนวนของผลมะนาวแก่โดยรวมเมื่อนำจำนวนของผลมะนาวจำนวน 200 ลูก มาวิเคราะห์ในระบบ ผลปรากฏว่าระบบสามารถวิเคราะห์ผลมะนาวแก่ได้ 196 ลูก และไม่สามารถวิเคราะห์ผลมะนาวแก่ได้จำนวน 4 ลูก คิดเป็นความแม่นยำของระบบอยู่ที่ 98 เปอร์เซนต์ ทั้งนี้การนับและวิเคราะห์ผลมะนาวที่ผิดพลาดเกิดจากแสงที่ตกกระทบลงบนผลของลูกมะนาวที่เกิดการซ้อนทับกัน มุมของผลมะนาวที่กล้องจับภาพได้มีความมืดเกินไป รวมทั้งการเคลื่อนที่ผ่านของผลมะนาวบนสายพานลำเลียงที่เร็วเกินไป ทำให้กล้องวิดีโอจับภาพของผลมะนาวที่ส่งภาพมายังระบบประมวลผลไม่ทันจึงทำให้เกิดความผิดพลาดในการวิเคราะห์และประมวลผล

การนำเสนอประสิทธิภาพของระบบวิเคราะห์และนับจำนวนผลมะนาวด้วยเทคโนโลยีคอมพิวเตอร์วิทัศน์ ได้สร้างแบบสอบถามโดยใช้หลักการของลิเคิร์ท สเกล จากการทำแบบสอบถามของผู้เข้าร่วมประเมิน คือ กลุ่มของประชากรที่เป็นเกษตรกรและผู้ที่ประกอบอาชีพค้าขาย จำนวน 10 ท่าน ผลของแบบประเมิน ดังตารางที่ 3

ตารางที่ 3 แบบสอบถามความพึงพอใจของผู้ทดลองใช้ระบบ

แบบสอบถามความพึงพอใจของผู้ใช้ระบบ	X	S.D.	แปลผล
1. ประสิทธิภาพของระบบ	4.70	0.54	พึงพอใจมากที่สุด
2. ความสะดวกของผู้ใช้งาน	4.30	0.30	พึงพอใจมาก
3. ความแม่นยำในการวิเคราะห์ผล	4.80	0.82	พึงพอใจมากที่สุด
4. การนำไปใช้ประโยชน์ในชีวิตประจำวัน	4.40	0.46	พึงพอใจมาก
5. ความเสถียรของซอฟต์แวร์	4.70	0.54	พึงพอใจมากที่สุด

5. สรุป

การพัฒนาระบบวิเคราะห์และนับจำนวนผลมะนาวด้วยเทคโนโลยีคอมพิวเตอร์วิทัศน์โดยที่ระบบสามารถวิเคราะห์และนับผลมะนาวอ่อนจากจำนวน 200 ลูก มีความแม่นยำถึง 92 เปอร์เซนต์ และสามารถวิเคราะห์และนับผลมะนาวแก่จากจำนวน 200 ลูก มีความแม่นยำถึง 98 เปอร์เซนต์ ทั้งนี้การนับและวิเคราะห์ผลมะนาวที่ผิดพลาดเกิดจากแสงที่ตกกระทบลงบนผลนะนาวที่เกิดการซ้อนทับกัน ลักษณะการเรียงกันของผลมะนาวที่กล้องจับภาพได้มีความมืดเกินไป รวมทั้งการเคลื่อนที่ผ่านกล้องของผลนะนาวบนสายพานลำเลียงที่เร็วเกินไป ทำให้กล้องวิดีโอจับภาพของผลมะนาวที่ส่งภาพมายังระบบที่รวดเร็วทำให้ระบบประมวลผลไม่สามารถวิเคราะห์ได้ จึงทำให้เกิดความผิดพลาดในการวิเคราะห์และประมวลผล ส่วนผลการประเมินความพึงพอใจในด้านต่างๆ อยู่ในระดับที่พึงใจมากถึงระดับที่พึงพอใจมากที่สุด ตามลำดับ

ในการพัฒนาระบบในขั้นต่อไป ควรแก้ไขปรับปรุงในส่วนของอัลกอลอริทึมที่ใช้ในการวิเคราะห์ผล โดยอาจใช้โมเดลสีในแบบต่าง ๆ หรือสร้างอัลกอริทึมการประมวลผลขั้นสูง เข้ามาช่วยในการประมวลผลเพื่อเพิ่มประสิทธิภาพของระบบ รวมทั้งการนำระบบสายพานลำเลียงเข้ามาช่วยในการคัดแยกชนิดของมะนาวอ่อนและมะนาวแก่ออกจากกันก็จะทำให้ระบบคัดแยกผลมะนาวมีประสิทธิภาพในการนำไปใช้งานได้มากยิ่งขึ้น

เอกสารอ้างอิง

[1] Tanut Bh. and Kamsuwan N., A Prototype Development of Marigold Classification by Image Processing, Journal Science and Technology Nakhon Sawan Rajabhat University. 2019; 11(13): 79-92.

[2] Kajornlap W. and Janpho J., Lemon Size Sorting Machine, Thesis in Production Techniques. 2015.

[3] Suksukont A., Santingamwong Ch., Khamkhaek W., Pansuwan S., Fruits Ripening Analysis System using Image Processing Technology, Journal of Information Science and Technology. 2022; 12(1): 61-66.

[4] Priwthaisong K., Automatic Color Sorting Machine on Conveyor Systems Controlled by Arduino, Association of Privatr Higher Education Institutions of Thailand. 5(1); 2016: 15-21.

[5] Pangerd S., Jeeranantasin N., Julakasemsak P., Cobal R., Duangmeun W., Automatic Mango Weight and Color Sorting Machine Equipped with Microcontrolle, Research Journal of Rajamangala University of Technology Krungthep. 2022; 16(2): 99-105.

[6] Jing L., Junzheng W. and Jiali M., Color moving object Detection Method Based on Automatic Color Clustering, Proceedings of the 33rd Chinese Control Conference. July 28-30, 2014: 7232-7235.

[7] Gokul P.R., Raj S. and Suriyamoorthi P., Estimation of Volume and Maturity of Sweet Lime Fruit using Image Processing Algorithm, International Conference on Communications and Signal Processing (ICCSP). April 2-4, 2015: 1227-1229.

[8] Pandey C., Sethy P.K., Biswas P., Behera S.K. and Khan M.R., Quality Evaluation of Pomegranate Fruit using Image Processing Techniques, International Conference on Communication and Signal Processing (ICCSP). July 28-30, 2020: 38-40.

[9] Sotirov S.I., Zhelyazkov S.P., Marudova M.G., Zsivanovits G. and Tokamkov D.M., Embedded System for Fruit Image Processing, International Scientific Conference Electronics. September 16-18, 2020: 1-3.

JIST Journal of Information Science and Technology
Volume 13, NO 1 | JAN – JUN 2023 | 17-25
ISSN: 2651-1053 (Online)
ISSN: 1906-9553 (Print)

Mining Users' Intentions from Thai Tweets Using BERT Models

Nattapong Sanchan
School of Information Technology and Innovation, Bangkok University,
Pathum Thani, Thailand.

Received: January 11, 2023; Revised: June 14, 2023; Accepted: June 24, 2023; Published: June 30, 2023

ABSTRACT – In this paper, we explore the mining of users' intentions in text. We viewed that being able to identify the intentions of users expressed in textual data provides us to specifically know aims and what users want to do. In the experiment, we collected tweets, constructed a Thai intention corpus, and performed a binary classification task on the corpus. We investigated the intent classification results derived through the application of 3 different Bidirectional Encoder Representations from Transformers (BERT), Word Embedding, and Bag of Words models. The results revealed that BERT Based EN-TH Cased model outperforms other models in both classification and processing time aspects. It achieves the F1 Score of 0.81 and performs the classification task faster than other BERT models up to 15%.

KEYWORDS: Intent Mining, Intention Mining, Intent Classification, Intent Detection, Text Mining, Natural Language Processing

1. Introduction

Due to the exceptional growth of data, data mining has been playing significant role in extracting meaningful knowledge or patterns in a large amount of data. The extracted knowledge can be beneficially used in several areas. For example, in business areas, companies can predict customers' behaviors and their purchasing trends so that the companies are able to make business decisions, plan business strategies, enhance companies' day-to-day operations, and finally generate cost-effective solutions to increase profits. For financial institutions, banks can analyze their historical data to predict whether to give loans to customers [1]. This provides the financial institutions to save costs and time in the loan approval process. In medical areas, data mining can be also used to find better treatments or cures for different diseases. Moreover, governments can also use data mining techniques to help understand the public's opinions regarding their launched policies. For these reasons, data mining is an important component that facilitates us to discover hidden knowledge in order to be beneficially applied to several applications.

Nowadays, another trend in utilizing data mining is the analysis of user-generated content on social media. Sentiment analysis and opinion mining in textual data have been widely explored to discover opinions expressed in the data. However, the study of sentiment analysis and opinion mining seems to be insufficient. Opinions expressed may not fully specify the purposes or directions that users want to achieve. Being able to identify the intentions of users expressed in the data provides us to specifically know the aims and what users want to do. This leads to the exploration of intent mining.[1]

Whereas intent mining has been largely explored in several languages, there is no resource officially available to conduct such research in Thai. In this paper, we explore the mining of intentions in Thai tweets by collecting Thai tweets and constructing a Thai intent corpus for our experiment. We view that intentions can be discovered in two forms: *explicit intention* and *implicit intention*. When a speaker has an explicit intention, for instance; *I am going shopping*, the listener clearly knows the aim of the speaker that he or she is going to do something. For the implicit intention, the intention is semantically hidden in the statement. For instance, the statement *Despite the fact that we are all unique, we should all be treated fairly and with equal rights*, implies that

[1] The term intent mining and intention mining are used interchangeably in some research work.

Corresponding Author: nattapong.sa@bu.ac.th

the author wants everyone to be treated equally. We collected tweets that express intentions and later annotated those tweets.

In the classification task, we utilized 3 different Bidirectional Encoder Representations from Transformers (BERT) models, that were pre-trained using different Thai datasets, to classify whether the given tweets are *intent* or *not intent*. We observed which model would yield the best result in the classification of intentions. Moreover, we also compared the classification results against word embedding and Bag of Words models. The final results revealed that BERT Based EN-TH Cased Model outperforms the classification of intentions in Thai tweets and performs less time than other BERT models.

2. Related Work

The analysis of intention in natural language has been studied since early 1970. In [2], the term intention was defined as a concept of goals and plans, and later this concept was used in the development of a question-answering system in [3]. Additionally, intention was also defined as the information related to goals, plans, and beliefs that a speaker wants to deliver to a computer, in the form of natural language. The goals can be exemplified as *finding out the location of a shopping mall*, *calling a taxi*, or *finding out how to cook Spaghetti Carbonara* with respect to the plans: *to use the Google Map*, *phone to a local taxi service*, and *search for a recipe on the Internet*. This is known as the concept of *explicit intent* tweets in this paper. The following sections discuss related work that has applied different approaches for detecting intentions in text.

2.1 Term-Based Approach

Tweets have been regarded as a valuable resource for mining user intentions [4–7]. Related work has studied intent mining from commercial tweets as aiming to find new business opportunities between buyers and sellers. [4] developed a system to classify whether tweets contain commercial intentions. The tweets are considered intent when they contain at least one verb and express intentions to perform a commercial activity in a recognizable way. By employing a certain number of relevant intent keywords, together with different classification algorithms, the researchers reported the best recall score of 77.4% and the precision of 57.1% using a Complement Naïve Bayes classifier and a linear logistic regression classifier respectively. However, the researchers did not report the generalization of the classifiers to data in other domains. They only focused on commercial tweets related to buying and selling intentions.

[5] worked on a multi-classification method to classify intent tweets into different categories. They employed intent indicators expressing the intent of the users, which is formed by, for instance; a subject I followed by a verb want to in the sentence *I want to buy an iPhone*. In addition, they also defined the intent keywords which can be a noun, a verb (or a set of verbs), or compound nouns. For instance, the keywords *buy an iPhone* in the previous example. These two definitions were used to identify intent tweets and generate a semi-supervised graph for multiclassification.

2.2 Word Embedding

Word Embedding is a technique that represents words in continuous-valued vectors to capture the internal semantic and syntactic information in text [8]. Examples of well-known word embedding models are GloVe [9], Word2Vec [10], Fasttext [11], and Universal Sentence Encoder (USE) [12]. In intention mining, [11] developed a system to detect intentions from users' utterances through the pre-processing, vectorization, and classification processes. Fastext was used to generate word embedding vectors which were derived from a convolutional network classifier. The results indicate that there is a significant improvement in the classification using Fasttext word embedding. Moreover, [13] enriched their word embedding model using WordNet, The Paraphrase Database (PPDB), and the Macmillan Dictionary to help capture semantic representations in intent text. The researchers revealed that the intent detection results derived from the enriched model outperform those of the original GloVe vectors.

2.3 BERT

BERT is a language presentation model that has been pre-trained using a vast amount of unlabeled text, including Wikipedia text and book corpora. BERT is bidirectionally trained, meaning that, during the training process, the model considers text sequences from either left to right and right to left [14]. Figure 1 illustrates a general architecture of BERT where text representation can be extracted. In the figure, a sentence is tokenized before special tokens *[CLS]* and *[SEP]* are added to indicate the beginning and end of a sentence. *[PAD]* is added to sentences that are shorter than the maximum ones. Next, in an encoding step, each token is mapped to a unique integer

representing the token in text. Finally, the vector representation of each term is derived.

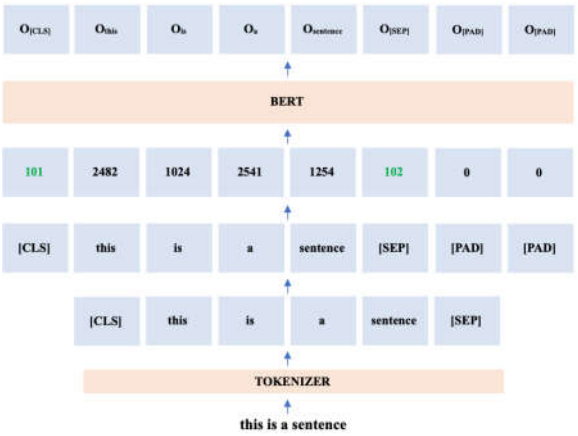


Figure 1. The representation of text extracted from BERT.

BERT has demonstrated its potential to provide state-of-the-art outcomes in a wide range of NLP tasks such as automatic summarization [15], sentiment analysis [16], machine translation [17], and intent classification [18]. However, focus on Thai intent classification. No research work officially reports Thai intent classification results. We explore the utilization of BERT, Word Embedding, and the Bag of Word models in the mining of intentions in Thai text. The following sections summarized key contributions made in this paper.

- Thai intention corpus. Currently, there is no Thai corpus available for conducting mining intention research, especially both explicit and implicit intentions. We, therefore, collected and annotated tweets related to gender non-discrimination. This will open research opportunities for NLP research communities to conduct research in Thai intent mining. We also provide an annotation guideline that could benefit the research communities for creating and expanding other genres of Thai intention datasets in future work.

- Investigation on Different BERT models. Currently, there are three BERT models that were pre-trained with the Thai dataset. One is Google's BERT-Based Multilingual Uncased model and the other two are those from the NLP research communities. At present, there is no observation on which model would provide the best intent mining result in Thai tweets. We, therefore, observe the classification results using the three models.
- Although BERT provides state-of-the-art results, one major drawback is the size of the pre-trained model which leads to the difficulty of the development time and the deployment of the system. [19] proposed a compact version of BERT which is claimed to reduce the processing speed and maintain the accuracy of results. Currently, there is no report to support this claim in Thai intent mining. We aim to observe this claim by comparing the processing time performed through each BERT model.

3. Methodology

In this paper, we view intention mining as a binary classification task. We compared the experimental results among 3 different pre-trained BERT models, word embedding, and Bag of Words approaches through 4 different classification algorithms. Figure 2 illustrates the experimental pipelines which will be discussed in the following sections.

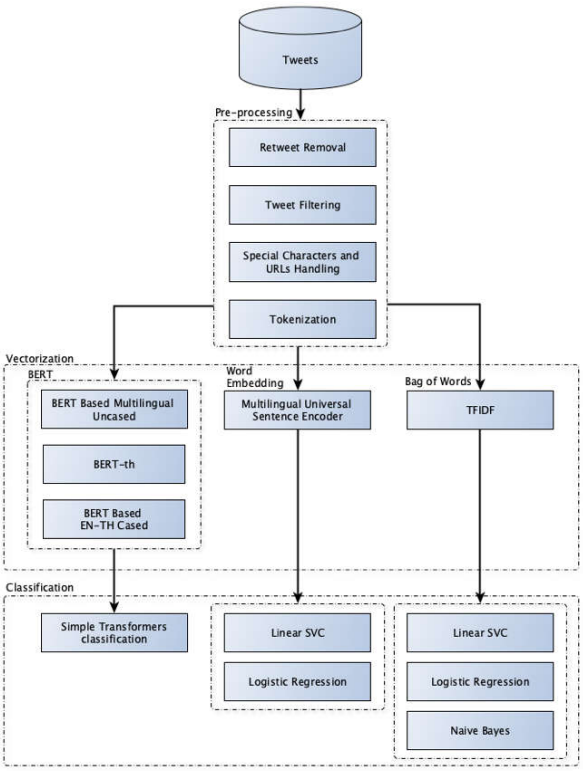


Figure 2. The experimental pipeline of the intent mining shapes 8 pipelines that generate vector representations through BERT, Word Embedding, and Bag of Words approaches.

3.1 Tweet Collection and Annotation

In order to conduct the experiment, we collected Thai tweets via the Twitter hashtag #MarriageEquality which discussed gender non-discrimination, same-sex marriage, and the legislation of same-sex marriage in Thailand, posted before February 2022. Originally, we collected 252,999 tweets. However, as the majority of tweets are redundant tweets and retweets, we used preprocessing techniques to clean the tweets and we obtained distinct 4,310 tweets in total, consisting of 382 intent tweets and 3928 non-intent tweets.

To annotate the data, an annotation guideline was given to two annotators who are fluent in English and aged above 18 years old. We partially reported the inter-annotator agreement with Cohen's Kappa [20] which accounts for 0.53. According to Landis and Koch, this reached the agreement level of moderate agreement, indicating that the annotation is quality [21]. The following section elaborates on how we annotated the collected tweets into one of the two classes: *intent* and *not intent*. Note that, in this paper, we translated Thai tweets and described the exampled tweets in English.

1) Intent

We employed a previous work [4] to identify intent tweets based on the three conditions: 1) tweets have at least one verb that describes an action of the user; 2) elaborate the speaker's attention to perform an activity that relates or would relate to oneself (3) in a recognizable way. When the tweets are considered intent, they can be either *explicit intent* or *implicit intent*:

1.1) Explicit Intent: As its name implies, we refer to explicit intent as the tweets that explicitly express the intention of the authors. Usually, these tweets simply contain the intent keywords such as *want*, *confirm*, *expect*, *call for*, *agree*, *want to see*, *support*, etc. The following tweets exemplify the explicit intent.

- Example 1: "I *want* all marriages to be equal because genders are not divided."
- Example 2: "I still *confirmed* the original statement that what we asked for was the fundamental rights that everyone should have."

1.2) Implicit Intent: This category is opposite to explicit intent in which the tweets do not directly express the intent of the authors. In other words, intent can be hidden and semantically expressed such as in the forms of interrogative (Example 3) and imperative sentences (Example 4).

- Example 3: "How many sets of morals or committees must be used in order to be able to recognize that marriage equality is a fundamental right, not a privilege."
- Example 4: "Become a part of one million names to protect all of the loves and for gender diversity."

2) Not Intent

Tweets are considered as *not intent* when they do not indicate any intentions in the context. Usually, they only describe options related to people, situations, events, or activities. The following examples exemplify the *not intent* tweets*:*

- *Example 5: "The Equal Marriage Act hanged for another 60 days."* This is an example of a news headline that does not indicate an intention.
- *Example 6: "The Chiang Mai team is accepting volunteers to organize a great marriage equality campaign event on February 14th*." In this example, no intention is indicated in this tweet. The main idea of this tweet is the description of what is happening.
- *Example 7: "If claiming marriage equality is against religious principles, the law should be repealed because sin should be judged by morality. Law and religion should be distinguished."* This tweet describes the opinions of the speaker, not the intention.

3.2 Preprocessing

Preprocessing techniques are the processes for cleaning textual data. We applied them for cleaning tweets as follows.

1) Retweet Removal

Retweeting is the action of reposting a user's tweet. It generally happens when a user likes and wants to share the tweet in his or her timeline. Retweets are usually indicated by having RT in the tweets. As retweets are redundant, we remove them from the dataset.

2) Tweet Filtering

We found the other kind of redundant tweets in the dataset that should be also removed. Such tweets mostly express the same content, are copied from other users' accounts, and are posted several times by the same users. To filter out these redundant tweets, a similarity metric, the Levenshtein distance, was used. We compared tweets from one to the others and computed the similarity scores among them. We observed that the tweets having similarity scores greater than or equal to 0.7 are usually redundant tweets.

3) Special Characters and URLs Handling

Generally, special characters, URLs, and emotions appearing in tweets do not usually contribute to the analysis of intention. We, therefore, removed URLs, emoticons, and characters @#/(){}? from the dataset

4) Tokenization

Tokenization is a process of breaking down a phrase, sentence, paragraph, or an entire textual document

into smaller units i.e., phrases or words. These units are called tokens. In the detection of intentions, we applied a Stacked Ensemble Filter and Refine for Word Segmentation (SEFR CUT) [22] to tokenize Thai tweets into words. After the preprocessing was completed, the tweets were fetched and transformed into vector representations in the next step.

3.3 Vectorization

To investigate which models would provide the best classification results, we generated vector representations for Thai tweets using BERT, Word Embedding, and Bag of Words models.

1) BERT

In this paper, we investigated the classification results derived through different BERT models. Currently, there are three BERT models that support the Thai language, including BERT Based Multilingual Uncased, BERT-th, and BERT Based EN-TH Cased.

- BERT Based Multilingual Uncase: a BERT that was pre-trained with Wikipedia text over 102 languages using Masked Language Modeling (MLM). With MLM, 15% of the tokens are masked and were trained to predict the mask tokens. This allows the model to learn the sentence representation bidirectionally [14].

- BERT-th: a BERT model that was pre-trained using a 2.39 gigabyte-sized Thai Wikipedia dataset. To use the model, we employed the Hugging Face transformer to generate BERT-th vector representations from Thai tweets [23-24].

- BERT Based EN-TH Cased: a compact version of multilingual BERT. [19] described that due to the large size of the original BERT, it is difficult to meet the production requirements when the system is implemented and deployed. Therefore, they proposed a smaller version of the multilingual BERT, implemented by decreasing the vocabulary size which leads to the reduction of the total number of parameters. For this reason, the model can process faster while retaining state-of-the-art results. We additionally observed the classification efficiency and the processing time performed through each model. Later, we compared the processing time of the classification task performed through the BERT models.

2) Word Embedding

To investigate the performance of the classifier using the word embedding approach, we employed the Multilingual Universal Sentence Encoder from the pre-trained Tensor Flow saved model [12]. This model is an extension of Universal Sentence Encoder which was trained and optimized with a great number of sentences, phrases, and paragraphs. We employed the model to generate word embedding vectors. For instance, for each input tweet, the model encodes it and generates a 512-dimensional vector. We later used them as the input to the classification algorithms.

3) Bag of Word

Bag of Words is a representation of an unordered set of words that describes the occurrence of words within a document. In this paper, we represented tweets in the Bag of Word model and extracted Term Frequency-Inverse Document Frequency (TF-IDF). Equation 1 - 2 illustrates the equations of TF-IDF where tf_i indicates the number of times a word i occurs. IDF measures the commonality of a word i that occurs in the entire document set. N represents the number of documents and n_i indicates the number of documents containing the word w_i. The IDF value closer to 0 indicates the more common the word is. The high value of the multiplication of TF and IDF, the rarity of the word occurrence. The final output derived from the model is the TF-IDF vectors which will be fed into the classification algorithms.

$$idf(w_i) = log\frac{N}{n_i} \quad (1)$$

$$tf_i \cdot idf(w_i) \quad (2)$$

3.4 Classification

We employed four classification algorithms to classify intent tweets. The first algorithm is Simple Transformers Classification. Simple Transformers is a transformer-based library that can perform several machine-learning tasks, including text classification. After deriving vector representations through BERT models, we employed a text classification module in the library to classify intentions. Moreover, we also observed the results derived from Linear Support Vector Classification (Linear SVC) and Logistic Regression since recent work has reported that Linear SVC yields significant results in text classification tasks [25–29]. The final algorithm is a probabilistic algorithm based on the Bayes Theorem, Naïve Bayes. We investigated this algorithm as several systems use it as a baseline.

Table 1. The classification results of intent tweets are derived from each pipeline.

Models	Processing Time (Mins)	Precision	Recall	F1 Score
BERT				
BERT Based Multilingual Uncase	23.69	0.59	0.73	0.65
BERT-th	24.91	0.74	0.83	0.78
BERT Based EN-TH Cased	21.01	0.79	0.83	**0.81**
Word Embedding: Universal Sentence Encoder (USE)				
Linear SVC	0.66	0.69	0.83	0.75
Logistic Regression	0.56	0.88	0.69	0.77
Bag of Words				
Linear SVC	0.10	0.48	0.70	0.76
Logistic Regression	0.09	0.87	0.60	0.71
Naïve Bayes	0.07	0.46	0.50	0.48

4. Experiment and Evaluation

In the evaluation of the classifiers, we utilized the 5-fold Cross Validation with a stratified option. In this setting, the examples in both classes are equally split in each fold. We followed the assignment of K = 5 from Kasthuriarachchy, Chetty, Karmakar, and Walls which explored the intent classification within a limited dataset [18]. We reported the classification results using three evaluation metrics, including Precision, Recall, and F1 Score as shown in Equation 3 to Equation 5. Note that TP, FP, and FN indicate True Positive, False Positive, and False Negative examples respectively. We calculated averaged precision, recall, and F1 Score from the aggregated confusion matrix derived in each fold. Precision is a ratio of *intent* examples that are correctly predicted by the classifier. Recall refers to from all *intent* examples how many *intent* examples are correctly captured. The final metric, the F1 score, is the harmonic mean of precision and recall.

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

$$Recall(R) = \frac{TP}{TP + FN} \quad (4)$$

$$F1\ Score = \frac{2PR}{P + R} \quad (5)$$

Table 1 presents the results of intent classification derived in each pipeline. In this experiment, we compared the classification results derived from three main vectorization approaches, including BERT, Word Embedding, and Bag of Words. By applying BERT, we explored different BERT models that were pre-trained using Thai datasets. Overall, the experimental results indicated that BERT Based EN-TH Cased model outperforms other models in detecting intentions in Thai tweets. The model achieved the averaged precision, recall, and F1 Score of 0.79, 0.83, and 0.81 respectively. Additionally, the intent classification using Word Embedding and Bag of Words yield nearly the same classification results, except those processed through Naïve Bayes.

Moreover, we also explored the processing time in the classification task performed by each model. It is a fact that the smaller size of the model commonly performs faster. However, emphasize large models like BERT. The processing time is longer due to the complexity of the model. Thus, Abdaoui et al. proposed the BERT Based EN-TH Cased which has a smaller number of language parameters so that it requires less memory and therefore loads faster [19]. In this paper, we compared the classification time among the three BERT models. We proved that the BERT Based EN-TH Cased model processes less time than other BERT models while still maintaining state-of-the-art results. It was able to classify the intent tweets faster than BERT-th and BERT Based Multilingual Uncased up to 11.32% and 15.66% respectively.

Aside from assessing the performance of classifiers with the averaged precision, recall, and F1 Score, we also conducted an error analysis to investigate possible factors that would lead to the miss-classification results. The examination of these factors could enhance the intent classification in future work. The tweets shown in Table 2 exemplify the miss-classification examples derived through the application of the BERT Based EN-TH Case model, which provides the best intent classification results.

Table 2. Miss-Classification Examples

Tweet Numbers	Tweets
1	"Regardless of gender, they should be able to love each other freely. We shouldn't block them just by gender."
2	"We are all human. We have equal rights. The era in that a male marries a female has ended. Male-male and female-female couples should have equal rights."
3	"We did not agree that you (the assemblymen) pass it (the Act) to the Cabinet, but the Pheu Thai Party confirmed in its resolution that they agreed with the Equal Marriage Act."

After investigating the classification results, we found that the classifier is substandard in understanding hidden semantics in the context. To illustrate, from the examples in Table 2, Tweet 1 illustrates an implicit intent tweet in which the intention is indirectly stated in the context. The sentence fragments "*should be able to love each other*" and "*shouldn't block them just by gender*" imply that the author intends to support gender non-discrimination. Likewise, the indication of should have equal rights in the second tweet also infers the author's intention. By understanding the semantics of the tweets, the classier would be able to interpret the overall meanings of the tweets. Moreover, the missing world knowledge identification is problematic in the detection of intention in our dataset. As shown in Tweet 3, the author expresses name entities i.e. the "*Cabinet*" and "*Phue Thai*" Party which are a part of the word knowledge in the gender equality domain. Additionally, the terms "*you*" and "*it"* refer to the assemblymen and the Equal Marriage Act led to the misclassification results.

5. Discussion and Conclusion

While intent mining has been extensively studied in several languages, Thai has very limited resources for this genre of research. In this paper, we collected, annotated, and constructed a Thai intention dataset that is related to gender non-discrimination. For future research, we view that the dataset derived from this study could benefit the research communities for researching mining intentions in Thai text. Additionally, our annotation guideline could provide the researchers for creating and expanding other genres of intention datasets in future work.

Moreover, in the classification task, we explored 8 pipelines that generate vector representations through BERT, Word Embedding, and Bag of Words. The main algorithms used are Simple Transformers Classification, Linear SVC, logistic regression, and Naïve Bayes. The results revealed that BERT Based EN-TH Cased model outperforms other models in both classification and processing time aspects.

In our error analysis, we found that lacking hidden semantic and word knowledge identification approaches impedes the success of the intent classification task. In future work, further investigation could be performed on adding a semantic layer to BERT models in order to observe whether semantics would increase the classification performance. In addition, the investigation of world knowledge as new features could enhance the classification task. For instance, [30] used domain-specific and common-sense knowledge to build features. However, this may require time and effort as constructing domain-specific knowledge needs human experts. Furthermore, the work could be extended by performing an automatic summarisation to generate the Thai intent summaries regarding gender equality. The ability to generate summaries of the intentions would provide users to digest a large number of intentions expressed in tweets conveniently.

References

[1] Gupta, V. Pant, S. Kumar, & P. K. Bansal, "Bank Loan Prediction System using Machine Learning," 2020 9th International Conference System Modeling and Advancement in Research Trends (SMART*)*, 4, 12, 2020, pp. 423 – 426.

[2] Schank RC, Abelson RP. Scripts, plans, goals, and understanding. Hillsdale, New Jersey: Lawrence Erlbaum Associates; 1997.

[3] Mckevitt P, "Analysing coherence of intention in natural language dialogue," Ph.D. thesis. Exeter: University of Exeter; 1991.

[4] Hollerit, B., Kröll, M., & Strohmaier, M, "Towards linking buyers and sellers: detecting commercial intent on Twitter," In Proceedings of the 22nd International Conference on World Wide Web. 2023, pp. 629-632.

[5] Wang J, Cong G, Zhao XW, Li X, "Mining User Intents in Twitter: A Semi-Supervised Approach to Inferring Intent Categories for Tweets, " The Twenty-Ninth AAAI Conference on Artificial Intelligence. '05, 2015, pp.339-345.

[6] Zhao XW, Guo Y, He Y, Jiang H, Wu Y., & Li, X, "We Know What You Want to Buy: A Demographic-based System for Product Recommendation on Microblogs," The 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 08, 2014 pp.1935-1944.

[7] Purohit H, Dong G, Shalin V, Thirunarayan K, Sheth A, "Intent Classification of Short-Text on Social Media," IEEE International Conference on

Smart City/Socialcom/ Sustaincom (Smartcity). 12, 2015, pp. 222-228.

[8] Li, Y., Yang, T, "Word Embedding for Understanding Natural Language: A Survey,"In: Srinivasan, S. (eds) Guide to Big Data Applications. Studies in Big Data, Vol 26. Springer, Cham.

[9] Pennington J, Socher R, Manning CD, "GloVe: Global Vectors for Word Representation," In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)* 10, 2014, pp. 1532-1543.

[10] Mikolov T, Sutskever I, Chen K, Corrado GS, Dean J, " Distributed representations of words and phrases and their compositionality". Advances in neural information processing systems. The 26*th* International Conference on Neural Information Processing Systems, 05, 12, 2013, pp. 3111-3119.

[11] Bojanowski P, Grave E, Joulin A, Mikolov, "Enriching word vectors with subword information," Transactions of the association for computational linguistics. 2017, pp. 135-146.

[12] Balodis K, Deksne D, "Intent detection system based on word embeddings," International Conference on Artificial Intelligence: Methodology, Systems, and Applications. 12, 08, 2018, pp. 25-35.

[13] Chidambaram, M., Yang, Y., Cer, D., Yuan, S., Sung, Y. H., Strope, B., & Kurzweil, R, "Learning Cross-Lingual Sentence Representations via a Multi-task Dual-Encoder Model," In Proceedings of the 4th Workshop on Representation Learning for NLP (RepL4NLP-2019), '02, 08, 2019, pp. 250 – 259.

[14] Kim JK, Tur G, Celikyilmaz A, Cao B, Wang Y Y, "Intent detection using semantically enriched word embeddings," The 2016 IEEE spoken language technology workshop (SLT). 13-16, 12, 2016, pp. 414-419.

[15] Devlin, J., Chang, M. W., Lee, K., & Toutanova, K, " BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Vol 1, 2019, pp. 4171 – 4186.

[16] Grail Q, Perez J, Gaussier E, "Globalizing BERT-based transformer architectures for long document summarization." The 16*nd*Conference of the European Chapter of the Association for Computational Linguistics. 19-23, 04, 2021, pp. 1792-1810.

[17] Hoang M, Bihorac OA, Rouces J, "Aspect-based sentiment analysis using BERT," The 22*nd* Nordic Conference on Computational Linguistics.

[18] Imamura K, Sumita E. "Recycling a pretrained BERT encoder for neural machine translation," The 3*rd* Workshop on Neural Generation and Translation. '04, 11, 2019, pp. 23-31.

[19] Kasthuriarachchy, B., Chetty, M., Karmakar, G. & Walls, D. Pre-trained Language Models with Limited Data for Intent Classification. In 2020 International Joint Conference on Neural Networks (IJCNN), Glasgow, United Kingdom, 19 – 24 July 2022, pp. 1 - 9.

[20] Abdaoui A., Pradel C., & Sigel, G, "Load What You Need: Smaller Versions of Multilingual BERT," In Proceedings of SustaiNLP: Workshop on Simple and Efficient Natural Language Processing, 20, 11, 2020, pp. 119 – 123.

[21] Cohen, J. (1960). A coefficient of agreement for nominal scales. Educational and psychological measurement, 20(1), pp. 37 – 46.

[22] Landis, J.R. & Koch, G.G. (1997). The Measurement of Observer Agreement for Categorical Data. Biometrics, 33, pp.159.

[23] Limkonchotiwat, P., Phatthiyaphaibun, W., Sarwar, R., Chuangsuwanich, E., & Nutanong, S, "Domain Adaptation of Thai Word Segmentation Models using Stacked Ensemble.," In Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP), 16 – 18, 11, 2020, pp. 3841 – 3847.

[24] Debut, L., Sanh, V., Chaumond, J., Delangue, C., Moi, A., Cistac, P., Rault, T., Louf, R., Funtowicz, M., Davison, J., Shleifer, S., von Platen, P., Ma, C., Jernite, Y., Plu, J., Xu, C., Le Scao, T., Gugger, S., Rush, A, "Transformers: State-of-the-Art Natural Language Processing," In Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations, 16 – 20, 11, 2020, pp. 38 – 45.

[25] T. Team, "Thaikeras: BERT Pre-training in Thai Language.".[Online]. Available: https://huggingface.co/Geotrend/bert-base-en-th-cased. [Accessed: May 1, 2020].

[26] Kaewnoo, P., & Senivongse, T, "Identification of Software Problem Report Types Using Multiclass Classification," Proceedings of the 2019 3rd International Conference on Software and e-Business. 9 – 11, 12, 2019, pp. 104 –109.

[27] Abdullah Y. Muaad, G. Hemantha Kumar, J. Hanumanthappa, J.V. Bibal Benifa, M. Naveen Mourya, Channabasava Chola, M. Pramodha, R. Bhairava, "An effective approach for Arabic document classification using machine learning," Global Transitions Proceedings, 3(1)(2022), pp. 267 – 271.

[28] Sikarwar, S.S., & Tiwari, D.N, "Analysis of the Sentiments of Amazon Reviews Dataset By Using Linear SVC and Voting Classifier,"

International Journal of Scientific & Technology Research, 2020, pp. 461 – 465.

[29] Gulati, K., Kumar, S. S., Boddu, R. S. K., Sarvakar, K., Sharma, D. K., & Nomani, M. Z. M, "Comparative analysis of machine learning-based classification models using sentiment classification of tweets related to COVID-19 pandemic" Materials Today: Proceedings, 51, pp. 38 – 41.

[30] Sanchan N., Bontcheva K. & Aker, A. "An Adoption of a Contradiction Detection Task to Assist the Summarization of Online Debates," 5th International Conference on Information Technology (InCIT), 21 – 22, 10, 2020, pp. 185 – 190.

[31] Gabrilovich, E., & Markovitch, S, "Feature Generation for Text Categorization Using World Knowledge," IJCAI International Joint Conference on Artificial Intelligence, 30, 07, 2005, pp. 1048 – 1053.

JIST Journal of Information Science and Technology
Volume 13, NO 1 | JAN – JUN 2023 | 26-37
ISSN: 2651-1053 (Online)
ISSN: 1906-9553 (Print)

การใช้เกมสอนทักษะการตัดสินใจตามสถานการณ์เพื่อป้องโรคไข้ดินและไข้ฉี่หนู (รุ่นเบื้องต้น)

Gamification for situated decision making for disease control – melioidosis and leptospirosis – a preliminary version

นฤมล วรรณประทีป[1], ปิยะรัตน์ ศิลปศุภกรวงศ์ [1]และ ชนัดดา ตั้งวงศ์จุลเนียม[2]*

Narumol Vannaprathip[1], Piyarat Silapasuphakornwong[1] and Chanatda Tungwongjulaniam[2]*

[1]*คณะเทคโนโลยีสารสนเทศและนวัตกรรม และศูนย์นวัตกรรมพิเศษเฉพาะทาง มหาวิทยาลัยกรุงเทพ*

[2]*สำนักโรคติดต่อทั่วไป กรมควบคุมโรค กระทรวงสาธารณสุข*

Received: May 27, 2023; Revised: June 17, 2023; Accepted: June 24, 2023; Published: June 30, 2023

ABSTRACT – The primary objective of proactive healthcare is to make the community environment free from disease and people have good health. One mechanism is to increase the number of community volunteers with basic knowledge and efficient skills in public health, in order to conducting the communication and promoting the correct well-being of people in the community. However, traditional on-site training is resource-intensive and limited by the COVID-19 pandemic. To accomplish the goal, gamification is proposed as a means to educate volunteers about endemic diseases and develop their decision-making skills for disease prevention. Starting from the game of Leptospirosis and Melioidosis, the experimental trials showed that individuals without prior knowledge could achieve an knowledge as 83.8%, and the prototype game received a high user satisfaction rating, scoring 4 from 5. In the future, the game can be used as a screening tool for assessing volunteer competency, ensuring sufficient volunteers for proactive disease surveillance in critical areas.

KEYWORDS: Decision-making, Melioidosis, leptospirosis, Simulation, Gamification

บทคัดย่อ -- **เป้าหมายหลักของการดำเนินการด้านสุขภาพในเชิงรุกคือการทำให้สิ่งแวดล้อมในชุมชนปลอดจากโรค และประชาชนมีสุขภาพอนามัยที่ดี ซึ่งกลไกหนึ่งคือการเพิ่มจำนวนอาสาสมัครในชุมชน (อสม.) ที่มีความรู้พื้นฐานและทักษะในการดูแลสุขภาพสาธารณสุขที่มีประสิทธิภาพ เพื่อดำเนินการสื่อสารและผลักดันการมีสุขภาวะที่ถูกต้องของคนในชุมชน เดิมทีการอบรมโดยผู้เชี่ยวชาญสิ้นเปลืองทรัพยากรต่างๆ ทั้งงบประมาณ แรงงาน และเวลา แต่เนื่องด้วยการระบาดของ COVID-19 การอบรมแบบเดิมไม่สามารถดำเนินการได้ ดังนั้น บทความนี้จึงนำเสนอการนำเกมสื่อดิจิทัลมาใช้ในการฝึก อสม. เพื่อให้มีความรู้เกี่ยวกับโรคที่ระบาดในพื้นที่และพัฒนาทักษะการตัดสินใจในการป้องกันโรคผ่านกระบวนการทางเกม (gamification) โดยเริ่มต้นจากเกมโรคไข้ฉี่หนูและโรคไข้ดิน การทดลองพบว่าผู้ไม่มีความรู้พื้นฐานสามารถพัฒนาความรู้ได้ถึงระดับ 83.8% และผู้เล่นเรียนรู้และประทับใจเกมต้นแบบโดยให้คะแนนเฉลี่ย 4 จาก 5 ในอนาคต เกมสามารถใช้เป็นเครื่องมือคัดกรองความรู้ของอาสาสมัครเพื่อการประเมินทักษะเบื้องต้นและเพื่อให้มีอาสาสมัครเพียงพอสำหรับการเฝ้าระวังโรคเชิงรุกในพื้นที่ที่สำคัญ**

คำสำคัญ: การประมวลผลภาพ, การเปรียบเทียบ, ผลไม้, แบบจำลองสี, คอมพิวเตอร์วิทัศน์

**Corresponding Author: narumol.v@bu.ac.th*

1. บทนำ

แนวโน้มของการเกิดโรคติดต่ออุบัติใหม่ อุบัติซ้ำ มีเพิ่มขึ้น และมากกว่าร้อยละ 70 เป็นโรคติดต่อจากสัตว์สู่คน สร้างความสูญเสียทั้งทางร่างกายและจิตใจของผู้ป่วยและคนรอบข้าง รวมทั้งมีผลกระทบต่อสภาพเศรษฐกิจและสังคมของประเทศ ด้วยสภาพการณ์ที่เป็นอยู่ปัจจุบัน มนุษย์ใกล้ชิดกับสัตว์มากขึ้น ทั้งการเลี้ยงสัตว์เพื่อเป็นเพื่อนหรือการเลี้ยงเพื่อค้าขายสร้างรายได้ ประกอบกับปัจจัยทางด้านการบุกรุกผืนป่า การทำลายทรัพยากรธรรมชาติ ที่อาจทำให้เชื้อและพาหะนำโรคสามารถนำโรคมาสู่คนได้มากขึ้น สำหรับประเทศไทยโรคติดต่อระหว่างสัตว์และคนที่ยังเป็นปัญหาทางสาธารณสุขที่สำคัญ ได้แก่ โรคพิษสุนัขบ้า, โรคเลปโตสไปโรสิส(Leptospirosis) หรือ โรคไข้ฉี่หนู, โรคเมลิออยโดสิส (Melioidosis) หรือโรคไข้ดิน, และโรคไข้หูดับ เป็นต้น หากเราไม่มีการวางแผนและดำเนินงานป้องกันควบคุมโรคก็อาจก่อความเสียหายทั้งทางด้านสุขภาพของประชาชน เกิดการเจ็บป่วยเสียชีวิต หรือเกิดการแพร่ระบาดของเชื้อจนควบคุมไม่อยู่ก็เป็นไปได้

โดยโรคไข้ดิน หรือโรคเมลิออยด์ [1] และโรคไข้ฉี่หนู หรือ โรคเลปโตสไปโรสิส (Leptospirosis) [2] ทั้งสองโรค เป็นโรคที่เกิดจากการติดเชื้อแบคทีเรียทั้งสองชนิดได้แก่ *Burkholderia pseudomallei* และ leptospira ตามลำดับ แหล่งของเชื้อมาจากดินและน้ำ ซึ่งผู้ป่วยอาจได้รับเชื้อจากการสัมผัสโดยตรง หรือ อาจมีสัตว์เป็นพาหะได้ด้วย ทั้งนี้ โรคทั้งสองเป็นภัยคุกคามแก่ประชาชนอย่างยิ่งในประเทศเขตร้อนชื้น จากข้อมูลการเฝ้าระวัง ข้อมูล ณ วันที่ 1 มกราคม - 11 พฤศจิกายน 2565 สถานการณ์โรคไข้ดิน พบผู้ป่วย 2,876 ราย จาก 67 จังหวัด คิดเป็นอัตราป่วย 4.35 ต่อแสนประชากร พบเสียชีวิต 51 ราย คิดเป็นอัตราตาย 0.08 ต่อแสนประชากร สถานการณ์โรคไข้ฉี่หนู พบผู้ป่วย 2,886 ราย จาก 66 จังหวัด คิดเป็นอัตราป่วย 4.36 ต่อแสนประชากรเสียชีวิต 34 ราย คิดเป็นอัตราป่วยตาย ร้อยละ 1.2 ซึ่งเพิ่มสูงขึ้น [3]

ในประเทศไทยจะพบโรคทั้งสองนี้มากในภาคอีสาน พบเชื้อได้ในดินและแหล่งน้ำทั่วไป การติดต่อเกิดได้ 3 ช่องทาง (รูปที่ 1) โดยโรคไข้ดินติดจาก 1) การดื่มน้ำที่ไม่ต้มสุกที่มีเชื้อปนเปื้อนเข้าไป หรือการรับประทานเนื้อสัตว์หรือเครื่องในสัตว์ที่มีฝีหนองโดยไม่ปรุงสุก 2) การหายใจเอาฝุ่นดินเข้าสู่ปอดหรือทางการสำลักน้ำ 3) การย่ำน้ำขณะมีแผล ส่วนโรคไข้ฉี่หนูติดทางผิวหนัง เกิดจากการสัมผัสน้ำหรือแม้กระทั่งสัตว์พาหะนำโรคซึ่งมีสารคัดหลั่งที่มีเชื้อปนมาด้วย หรือเกิดจากการแช่น้ำนาน ๆ เชื้อจะเข้าผ่านทางบาดแผล, รอยถลอก, หรือรอยขีดข่วน หรือ เมื่อฝนตกหนักเชื้อสามารถไชเข้าทางผิวหนังที่อ่อนนุ่มอย่างเยื่อบุตา จมูก และปาก หรือจากการลงแช่น้ำ ลุยน้ำ ย่ำดินโคลน โดยไม่สวมอุปกรณ์ป้องกัน เป็นต้น โดยกลุ่มเสี่ยงของโรคได้แก่ ผู้ที่ต้องประกอบอาชีพสัมผัสกับดินและน้ำโดยตรง เช่น เกษตรกร ชาวนา ชาวสวน ชาวประมง กลุ่มเสี่ยงพิเศษของโรคไข้ดิน รวมถึงผู้มีโรคประจำตัวได้แก่ผู้ป่วยโรคเบาหวาน โรคไตวายเรื้อรัง โรคธาลัสซีเมียชนิดรุนแรง และ โรคมะเร็ง โดยตัวความรุนแรงของโรคไข้ดินเกิดจาก หลังจากติดเชื้อประมาณ 1-2 สัปดาห์ ผู้ป่วยที่เริ่มมีอาการ จะเป็นอาการทั่วไปที่ไม่จำเพาะ เช่น มีไข้ เหนื่อย น้ำหนักลด ปวดศีรษะ ปวดเมื่อยตัว ปวดกล้ามเนื้อโดยเฉพาะที่น่องหรือโคนขา ซึ่งอาการระยะแรกนี้จะมีความคล้ายกับโรคติดเชื้ออื่น ๆอย่างมาก จึงทำให้มีการสับสนในการวินิจฉัยกับโรคอื่นๆ เช่น โรคปอดอักเสบ โรควัณโรค อีกทั้งระยะฟักตัวของโรคนี้มีช่วงกว้างอาจใช้เวลาตั้งแต่ 1 วัน ไปจนถึงหลายปี อาการโรคไข้ฉี่หนูอาจมีตาแดง ตัวเหลืองตาเหลือง ปัสสาวะออกน้อย ไอเป็นเลือด จนเสียชีวิตในที่สุด เมื่อแพทย์สืบประวัติเสี่ยงจากคนไข้ คนไข้จำไม่ได้ว่าเคยมีพฤติกรรมเสี่ยงมาก่อน เพราะเวลาที่ได้รับเชื้อจริงๆนั้นเนิ่นนานเกินไป หรือจุดเสี่ยงมากกว่า 1 แห่งทำให้ผู้ป่วยไม่สามารถจำได้ ซึ่งเป็นความยากในการวินิจฉัยโรครวมไปถึงการตระหนักรู้ถึงภาวะของโรคของผู้ป่วยเองในการที่จะรับการรักษาได้อย่างทันท่วงที และในที่สุดนำไปสู่การเสียชีวิต นอกจากนี้ หากผู้ป่วยซื้อยามารับประทานเองหรือเข้ารับการรักษาล่าช้า ก็อาจเกิดภาวะแทรกซ้อนและเสียชีวิตได้

ซึ่งจากการสำรวจข้อมูลเรื่องความรู้เกี่ยวกับโรคทั้งสองในหมู่ประชาชนในพื้นที่เสี่ยง ยังพบว่า ประชาชนมีความรู้จักโรคไข้ดินน้อยมาก ส่วนโรคไข้ฉี่หนูมีความรู้แต่ไม่ตระหนัก ทั้งนี้ผู้เสียชีวิตจากไข้ฉี่หนูส่วนใหญ่เกิดจากการย่ำดินที่ชื้นแฉะโดยไม่สวมรองเท้าป้องกัน ลุยน้ำท่วมขัง ทำนา/หาปลา ลงแช่น้ำเป็นเวลานาน รวมถึงเมื่อมีอาการป่วยก็วางใจ และไม่รีบไปพบแพทย์ในทันที [4] อนึ่งโรคทั้งสองเนื่องจากมีสาเหตุต้นกำเนิดโรคจากปัจจัยเสี่ยงคล้ายกัน

รูปที่ 1. ปัจจัยและพฤติกรรมเสี่ยงการติดเชื้อโรคไข้ฉี่หนูและโรคไข้ดิน

ดังนั้น การป้องกันในเชิงรุก คือการลงพื้นที่และทำการฝึกอบรม อาสาสมัครหมู่บ้าน (อสม.) เรื่องความรู้ด้านโรคและการป้องกันโรค เพื่อให้ อสม. เป็นตัวแทนเผยแพร่ความรู้ในการป้องกันโรคไปยังผู้อยู่อาศัยในชุมชนต้นทาง รวมถึงเฝ้าระวังการเกิดโรค โดยการเยี่ยมเยือนผู้อยู่อาศัยในชุมชน จึงเป็นสิ่งจำเป็นในการให้ความรู้ทั้งสองโรคไปควบคู่กัน เพื่อสามารถป้องกันได้ในคราวเดียวกันจากการปฏิบัติตัวที่ถูกต้องเพียงครั้งเดียว

ประเทศไทยมีจุดแข็งในการสาธารณสุขอย่างหนึ่ง คือ การสื่อสารให้ความรู้แก่ประชาชนผ่าน อสม. ทำให้ประชาชนในแต่ละพื้นที่สามารถเข้าถึงความรู้และรู้จักการป้องกันตัวจากโรคระบาดและโรคต่างๆได้ [5] โดยผลสำเร็จที่เห็นได้ชัดของการดำเนินงานของ อสม. คือในช่วง COVID-19 ที่ระบาดในช่วงปี 2563-2565 ที่ระบบ อสม. เป็นกลไกสำคัญในการสร้างความเข้มแข็งอย่างมากในการดูแลสุขภาพสมาชิกในหมู่บ้าน ให้ข่าวสารและคอยส่งข้อมูลระหว่างสาธารณสุขและประชาชนในพื้นที่ได้ทันท่วงทีเป็นอย่างดี [6]

ดังนั้นการเพิ่มศักยภาพ อสม. ในการให้ความรู้และปรับเปลี่ยนพฤติกรรมคนในชุมชนจึงเป็นเรื่องสำคัญอย่างมากสำหรับการสาธารณสุขเชิงรุก เพราะไม่ใช่แค่การคัดแยกผู้ป่วยหรือ รักษาตามอาการโรคเมื่อประชาชนได้รับเชื้อแล้ว แต่เป็นการป้องกันไม่ให้เกิดการติดเชื้อขึ้นเลย หรือ กำจัดที่ต้นตอแห่งเชื้อให้หมดไปจากชุมชนด้วยซ้ำ ซึ่งเป็นการป้องกันการติดเชื้อและระบาดของโรคอย่างยั่งยืนและถาวรมากกว่า ซึ่งสำหรับการสร้างความรอบรู้ด้านสุขภาพสนับสนุนการป้องกัน การควบคุมโรคไข้ดินและโรคไข้ฉี่หนูมุ่งเป้าเบื้องต้นเพื่อลดจำนวนผู้ป่วยและลดจำนวนผู้เสียชีวิต เป็นไปตามนโยบายและทิศทางการดำเนินงานกระทรวงสาธารณสุขที่เน้นสุขภาพคนไทย เพื่อสุขภาพประเทศไทย ในการเพิ่มประสิทธิภาพการสื่อสารยกระดับการสร้างความรอบรู้ด้านสุขภาพในทุกมิติ ทำให้ประชาชนสามารถเข้าถึงข้อมูลข่าวสารได้อย่างถูกต้อง เป็นปัจจุบัน สะดวก รวดเร็ว เพื่อพัฒนาศักยภาพคนทุกช่วงวัยให้สามารถดูแลสุขภาพกาย ใจ ของตนเอง ครอบครัวและชุมชนให้แข็งแรง ซึ่งตอบเป้าหมายการพัฒนาที่ยั่งยืน (Sustainable Development Goals–SDGs) ในเป้าหมายการมีสุขภาพและความเป็นอยู่ที่ดี (Good Health And Well-Being) สอดคล้องกับการดำเนินงานบรรลุตามแผนยุทธศาสตร์ชาติระยะ 20 ปี ด้านสาธารณสุขการพัฒนาและเสริมสร้างศักยภาพทรัพยากรมนุษย์ แผนแม่บทการเสริมสร้างให้คนไทยมีสุขภาวะที่ดี แผนย่อยที่ 1 การสร้างความรอบรู้ด้านสุขภาวะและการป้องกันและควบคุมปัจจัยเสี่ยงที่คุกคามสุขภาวะ [4]

อย่างไรก็ตาม การฝึกอบรม อสม.ในพื้นที่ใช้ทรัพยากรค่อนข้างมากทั้งคน เวลา และอุปกรณ์ นอกจากนี้ การระบาดของโรค COVID-19 ทำให้มีข้อจำกัดมากขึ้นในการเข้าพื้นที่ ดังนั้นจึงจำเป็นต้องมีการใช้เทคโนโลยีเข้ามาช่วยในการฝึกอบรม แต่การใช้เทคโนโลยีทางไกลมาช่วยฝึกอบรมอาจจะสร้างความน่าเบื่อให้กับ อสม. ได้ ดังนั้น ผู้วิจัยจึงเสนอแนวคิดการใช้เกม (gamification) เป็นสื่อในการอบรม อสม. แทนการอบรมโดยผู้เชี่ยวชาญลงพื้นที่จริง เพื่อลดการเดินทางไปยังพื้นที่จริงลงและลดการใช้ทรัพยากรในการจัดการอบรมอย่างมหาศาล โดยที่ยังสามารถพัฒนาศักยภาพ อสม. ได้อย่างมีประสิทธิผลในระยะเวลาอันสั้นได้อีกด้วย ทั้งนี้ เนื่องจาก กรมควบคุมโรค มีเป้าหมายในการอบรมให้ความรู้และพัฒนาศักยภาพ อสม. ที่มีคุณภาพให้กับชุมชนเป็นสำคัญ ดังนั้นประสิทธิภาพของวิธีทดแทนการอบรมแบบลงพื้นที่ปกติด้วยเกมนั้น จึงจำเป็นต้องมีความชัดเจนว่าสามารถให้ความรู้แก่ผู้เล่นได้จริงด้วย ดังนั้นในบทความนี้ เราจึงทำการวิจัยเรื่องการใช้เกมมาเป็นสื่อในการให้ความรู้เรื่องการป้องกันโรคเชิงรุกเป็นหลัก บนสมมติฐานที่ว่า การใช้กระบวนการทาง gamification สามารถจัดการให้ความรู้แก่ผู้เล่นได้ดีกว่าหรือเทียบเท่าการอบรมโดยผู้เชี่ยวชาญแบบลงพื้นที่จริงได้ โดยเราได้ทำการทดลองเบื้องต้นเพื่อยืนยันสมมติฐานก่อนที่จะนำแผนนี้ไปใช้ในการพัฒนาเกมจริงอย่างเต็มรูปแบบต่อไป กลไกหนึ่งที่ทำให้

ความรู้เข้าถึงประชาชนอย่างแท้จริง ทำหน้าที่เป็นสื่อสู่ชุมชนเป็นตัวกลาง

ทั้งนี้ ในบทความนี้ มุ่งเป้าไปที่การป้องกันโรคเชิงรุกของ 2 โรคก่อนคือ โรคไข้ดินและไข้ฉี่หนู เพื่อให้เกิดการตระหนักรู้ถึงความเสี่ยงในการเกิดโรคไข้ดินและไข้ฉี่หนูรวมถึงเข้าใจสาเหตุของโรค และการสร้างทักษะการตัดสินใจเพื่อป้องกันโรค นอกจากนี้ ตัวเกมยังออกแบบมาเพื่อเป็นมาตรฐานการคัดกรองการประเมิน อสม. เบื้องต้นในการป้องกันโรค (รูปที่ 1) ใช้เพื่อเพิ่มจำนวน อสม. ที่มีความรู้ในการป้องกันโรคไข้ดินและโรคไข้ฉี่หนูให้เพียงพอกับการเฝ้าระวังเชิงรุกของโรคในแต่ละพื้นที่เป็นสำคัญ

โดยในบทความนี้ ได้นำเสนอแนวคิดการใช้เกมเพื่อเป็นสื่อในการอบรมให้ความรู้แก่ อสม. เพื่อการป้องกันโรคในเชิงรุก โดยใช้วิธีของ gamification เข้ามาตอบปัญหาในเชิงการเรียนรู้ เพื่อให้ อสม. ได้ความรู้อย่างแท้จริง โดยเราทำการทดลองเพื่อจำลองสถานการณ์ของ อสม. ที่ต้องเดินทางไปเยี่ยมเพื่อนบ้าน เพื่อฝึกการตัดสินใจในสถานการณ์จริง และทำแบบสอบถามเพื่อประเมินผลรับรองว่า แนวความคิดเราสามารถนำไปใช้ได้จริงในเบื้องต้นด้วยนั่นเอง

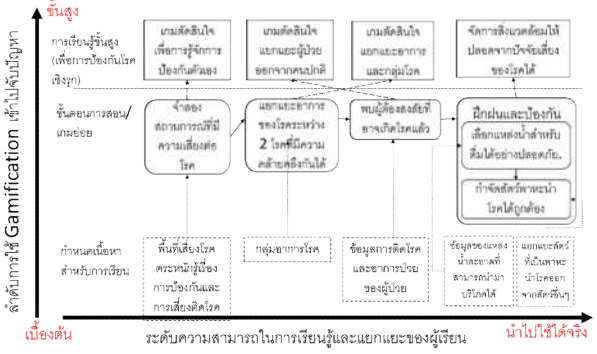
2. แนวคิดในการใช้ Gamification ในการให้ความรู้อบรมเรื่องการป้องกันโรคเชิงรุก

Gamification คือ การนำส่วนประกอบของเกมมาสร้างเกมเพื่อวัตถุประสงค์ใด ๆ ที่นอกเหนือจากการให้ความสนุกสนาน [7] ทั้งนี้ ในเชิงวิชาการ (Academic) วัตถุประสงค์จะเป็นไปเพื่อการใช้เกมเพื่อพัฒนาการเรียนรู้ของผู้เรียนโดยตรงเป็นหลัก ดังนั้น การใช้ gamification เพื่อเป็นสื่อในการช่วยฝึกอบรมนั้นนอกจากจะให้ความรู้ทางด้านการป้องกันโรคแล้ว ยังสร้างความสนุกสนาน ความผ่อนคลาย และไม่ทำให้ผู้เรียนรู้สึกว่าถูกบังคับจนเกินไป ตัวเกมช่วยอำนวยความสะดวกให้กับเจ้าหน้าที่ในการอบรม อสม. ใหม่ ๆ ในหลาย ๆ พื้นที่ในเวลาเดียวกัน

รูปที่ 2 แผนภาพแสดงแนวคิดขั้นตอนวิธีในการนำ Gamification เข้ามาจับปัญหาการเผยแพร่ความรู้ให้ผู้เล่น โดยแกน X เป็นลำดับการใช้ gamification เข้าไปจับกับปัญหา จากเบื้องต้นไปจนถึงขั้นสูง ซึ่ง เบื้องต้นได้แก่ 1) การกำหนดเนื้อหาสำหรับการเรียนรู้ 2) ขั้นตอนการสอน (โดยใช้รูปแบบเกมย่อย) 3) การเรียนรู้ขั้นสูง เพื่อจุดประสงค์การป้องกันโรคเชิงรุกได้ และแกน Y เป็นระดับความสามารถในการเรียนรู้และแยกแยะได้ของผู้เรียน จากในขั้นบทเรียนแบบง่ายเริ่มต้นที่ผู้เรียนแค่รับฟังเนื้อหา ไปจนถึง การที่ผู้เรียนสามารถนำไปปฏิบัติได้จริงและสามารถนำไปแก้ปัญหาในสถานการณ์และพื้นที่จริงได้อย่างคลอบคลุม ซึ่งสุดท้ายแล้ว เป้าหมายของการป้องกันโรคแบบเชิงรุก คือการที่ อสม. สามารถจัดการสิ่งแวดล้อมให้ปลอดจากปัจจัยเสี่ยงต่างๆของโรคได้อย่างสมบูรณ์ โดยทั้งนี้ จะดำเนินการไปถึงเป้าหมายที่ตั้งไว้นี้ได้ จะต้องผ่านการอบรมและเรียนรู้ฝึกฝนทักษะในวิถีทางที่ถูกต้องมาชักระยะหนึ่งด้วย ซึ่งเป็นกระบวนการที่ยากลำบาก ใช้เวลา และทรัพยากรมากในการอบรมแบบปกติโดยผู้เชี่ยวชาญ แต่ข้อดีคือ เนื่องจากเป็นการอบรมแบบรายบุคคล จึงได้บุคคลากรที่มีประสิทธิภาพสูงมากเช่นกัน

อนึ่ง การเรียนรู้โดยใช้ gamification สามารถออกแบบเป็นลักษณะของ ฐาน (station) ต่างๆ เพื่อให้ผู้เรียนตระหนักรู้และฝึกฝนทักษะที่จำเป็นไปได้ทีละอย่างตามเป้าหมายได้ เมื่อจบชั่วโมงการเรียน จะมีการประเมินความสามารถของผู้เรียนจากผู้เชี่ยวชาญ ได้ทั้งจาก การตอบคำถาม การทำข้อสอบ หรือให้ลองปฏิบัติให้ดู เป็นต้น ซึ่งใช้เวลาในการประเมินพอสมควร

เนื่องด้วย เราจะใช้เทคโนโลยีทางเกมดิจิทัล เพื่อการสื่อสารบทเรียนแทนการอบรมแบบใช้เกมจริงด้วยผู้เชี่ยวชาญ ดังนั้น เราจึงต้องออกแบบเลือกเกมในรูปแบบที่สามารถตอบวัตถุประสงค์การเรียนรู้ข้างต้นได้อย่างครบถ้วนเช่นกันโดยการออกแบบเกมสื่อดิจิทัล เป็นไปดังรูปที่ 3



รูปที่ 2. แผนภาพแสดงแนวคิดการนำ Gamification เข้ามาจับปัญหาการเผยแพร่ความรู้ให้ผู้เล่น

รูปที่ 3. วิธีทาง Gamification ที่ใช้ในเกมเพื่อมุ่งเป้าการป้องกันโรคเชิงรุก:

จากรูปที่ 3 เป็นการแสดงการนำวิธีการทาง Gamification มาประยุกต์ใช้ในเกมเพื่อมุ่งเป้าการป้องกันโรคเชิงรุก นั่นคือการป้องกันโรคและการทำลายแหล่งต้นกำเนิดของโรคเป็นสำคัญ โดยเกมที่จะนำมาใช้ เป็นเกมเพื่อการตอบจุดประสงค์ในการเรียนรู้ของผู้เรียนเป็นสำคัญ ซึ่งจะเป็นไปตามกระบวนการเรียนรู้ ดังนี้:

1) เกมเล่าเนื้อหาเรื่องราวเชิงโต้ตอบ (Interactive Narrative)-- เพื่อให้ผู้เรียนได้เรียนรู้เนื้อหาที่สำคัญสำหรับตัวโรค อาการของโรค แหล่งเพาะและติดเชื้อ รวมไปถึงการป้องกันโรคได้ในเบื้องต้น

2) เกมตอบคำถาม (Questions and Answers : Q&A) -- เพื่อให้ผู้เรียนได้ทบทวนสิ่งที่เรียนไปในบทแรก และคิดวิเคราะห์ข้อมูลเพื่อการจดจำบทเรียนเพื่อพร้อมสำหรับการนำไปประยุกต์ใช้ในบทถัดไป อนึ่ง อาจใช้ขั้นตอนนี้ เพื่อทำการทดสอบความรู้ของผู้เรียนด้วย เป็นในแนวของการสอบข้อเขียนเพื่อให้ได้ความรู้มาตรฐานที่บุคคลากรควรมี เป็นต้น

3) เกมฝึกตัดสินใจ (Decision Making) -- เพื่อให้ผู้เรียนสามารถแยกแยะพาหะ อาการของโรค รวมไปถึงจำแนกปัจจัยเสี่ยง ที่มีผลต่อการติดเชื้อโรคได้เป็นอย่างดี เพื่อในสถานการณ์จริงสามารถตระหนักรู้และนำไปสู่การจัดสิ่งแวดล้อมเพื่อการป้องกันและกำจัดปัจจัยเสี่ยงออกจากพื้นที่ได้อย่างสมบูรณ์

4) เกมสร้างสถานการณ์จำลอง (Simulation) - เพื่อให้ผู้เรียนสามารถใช้ความรู้ที่เรียนมาในการฝึกทักษะการปฏิบัติจริง การลองตัดสินใจในสถานการณ์ต่างๆ เพื่อฝึกซ้อมให้เกิดความชำนาญก่อนไปเจอสถานการณ์จริงในชุมชน เพื่อป้องกันสถานการณ์เลวร้ายที่เกิดจากการตัดสินใจผิดพลาดในสถานการณ์ของจริง (มาแฟ้ในเกมดีกว่าเกิดเหตุร้ายจริงๆ)

3. การทบทวนวรรณกรรม การใช้ Gamification ในสื่อการอบรมให้ความรู้เรื่องการป้องกันโรคเชิงรุก

มีการใช้ gamification ในการให้ความรู้ป้องกันโรคมาก่อนในหลายประเทศอย่างแพร่หลายแล้ว โดยมีจุดเด่นและจุดด้อยที่แตกต่างกันออกไป ดังแสดงในตารางที่ 1

จากตารางเราจะได้เห็นว่า แต่ละเกมมีข้อดีข้อเสียแตกต่างกันไป และทำให้เห็นถึงประเด็นที่ต้องแก้ไขในการออกแบบเกมที่เพื่อป้องกันโรค โดยที่

1) ควรจะเป็นเกมที่เล่นคนเดียว เพื่อให้มีบทบาทที่ชัดเจนในการตัดสินใจด้านต่าง ๆ

2) เกมควรจะสื่อความหมายให้ชัดเจนในตัวเอง ผู้เล่นไม่ต้องเสียเวลาเรียนรู้และรับรู้วัตถุประสงค์ของเกมได้อย่างรวดเร็วและง่ายดาย นอกจากนี้ การให้โอกาสผู้เล่นได้มีปฏิสัมพันธ์กับเกมน่าจะเป็นการเพิ่มการดึงดูดให้ผู้เล่นสนใจในการเล่นเกมอย่างต่อเนื่อง

3) เกมควรจะ 3.1) สอนให้ผู้เล่นเข้าใจว่าที่มาที่ไปของการติดเชื้อเป็นอย่างไร การกระทำอะไรที่ส่งผลต่อการติดเชื้อ 3.2) สอนให้ผู้เล่นรู้จักการปฏิบัติตัวเพื่อป้องกันโรค 3.3) สอนให้ผู้เล่นเข้าใจบทบาทส่วนตัวเพื่อผลต่อส่วนรวมด้วย

ความต้องการเหล่านี้นำไปสู่การออกแบบเกมเพื่อป้องกันโรคไข้ฉี่หนูและโรคไข้ดินแบบควบคู่กันไป ซึ่งสมมติฐานคือน่าจะมีประสิทธิภาพเทียบเท่าหรือดีกว่า การอบรมโดยใช้รูปแบบปกติแบบผู้เชี่ยวชาญลงพื้นที่ให้ความรู้

4. แนวคิดการออกแบบเกมเพื่อการป้องกันโรคคู่ควบของโรคไข้ดินและโรคไข้ฉี่หนู (proposed method)

จากรูปที่ 4 เราออกแบบเป็น 5 สถานการณ์ (scenario) คือ พื้นที่ที่แตกต่างกันได้แก่ นา สวน ไร่ และพื้นที่น้ำขัง และเป็นการเยี่ยมเยียนบ้าน ของ อสม. ในพื้นที่เพื่อผลักดันให้คนที่เป็นคนได้รับเชื้อรีบไปพบแพทย์ได้อย่างเร็วที่สุดเพื่อป้องกันการสูญเสียชีวิตจากการนิ่งนอนใจ เกมที่ใช้ จะแบ่งเป็น 6 เกมพื่อจุดประสงค์การเรียนรู้ที่ต่างกัน 4 ข้อ คือ

ตารางที่ 1. ตารางเปรียบเทียบจุดเด่นจุดด้อยของเกมเพื่อการให้ความรู้การป้องกันโรค.

ชื่อเกม	ประเทศ	โรคที่พุ่งเป้า	ลักษณะเกม	จุดเด่น	จุดด้อย
MEliOAPP [8]	มาเลเซีย	Melioidosis	Mobile application ที่ให้ความรู้เรื่องโรคไข้ดิน และมี Mini quiz game และ level game	เกมมีโหมดความรู้เรื่องโรคไข้ดินครบรอบด้าน	ภารกิจใน Level game ไม่มีความชัดเจนถึงการสอนวิธีในการป้องกันโรคไข้ดิน
ไข้ฉี่หนู [9]	มาเลเซีย	Leptospirosis	เกมอยู่ในรูปแบบฐานปฏิบัติการ 10 ด่าน ผู้เล่นต้องจับกลุ่มและเวียนเล่นไปตามแต่ละด่าน โดยต้องใช้ทักษะการแก้ปัญหาในการผ่านแต่ละด่านให้ได้	ทั้งสิบด่านให้ความรู้เรื่องไข้ฉี่หนูในแง่มุมต่าง ๆ อย่างละเอียด	ในแต่ละด่าน ผู้เล่นจะต้องร่วมกันใช้ความคิดสร้างสรรค์และความรวดเร็วในการแก้ไขปัญหาอาจทำให้ผู้เล่นไม่ได้มุ่งเน้นถึงการเรียนรู้เรื่องการป้องกันโรคไข้ฉี่หนูได้อย่างเต็มที่
Stay at home [10]	สวีเดน	COVID-19	Serious game ประกอบด้วยเกมย่อยหลาย ๆ ด่าน ที่มุ่งเน้นการแนะนำแนวทางการปฏิบัติตัวเพื่อป้องกันการติดเชื้อ	เกมช่วยเปลี่ยนทัศนคติและพฤติกรรมใหม่ให้ผู้เล่นสู่การใช้ชีวิตแบบ New-Normal	วิธีเล่นเข้าใจยากและมีการโต้ตอบน้อยไม่ดึงดูดผู้เล่นให้อยู่ในเกมและทำให้ผู้เล่นไม่ต่อเนื่องกับเกม
Plague Inc [11]	สหราชอาณาจักร	ไม่ระบุ	Simulation game ที่สอนให้ผู้เล่นเข้าใจการแพร่ระบาดของเชื้อโรคที่แพร่กระจายไปทั่วโลก (Balance strategy planning game)	ผู้เล่นเข้าใจการกระจายตัวของเชื้อโรคที่แพร่ระบาดได้ในภาพกว้าง	ไม่มีการสอนเรื่องการปฏิบัติตัวต่อการป้องกันโรคแบบเฉพาะเจาะจง เน้นความสนุกสนานของเนื้อเรื่องในเกมมากกว่า

ชื่อเกม	ประเทศ	โรคที่พุ่งเป้า	ลักษณะเกม	จุดเด่น	จุดด้อย
The Great Flu [12]	เนเธอแลนด์	ไข้หวัดใหญ่	Simulation game ที่ผู้เล่นต้องตัดสินใจทำภารกิจ เพื่อควบคุมและต่อสู้กับไวรัสกับไข้หวัดใหญ่	ตัวเกมพยายามสร้างความตระหนักรู้ถึงการแพร่กระจายของโรค รวมถึงผลกระทบที่จะเกิดขึ้นหลังจากที่เกิดโรคระบาดไปแล้ว	ผู้เล่นไม่สามารถเข้าใจผลกระทบ หรือการปฏิบัติตัวจริงว่าไวรัสแพร่กระจายได้ อย่างไร หรือการกระทำผู้เล่นส่งผลอย่างไรต่อการแพร่กระจายของเชื้อโรค
Pandemic [13]	อเมริกา	ไม่ระบุ	เป็น Board game เกมที่เล่นเป็นกลุ่ม โดยผู้เล่นแต่ละคนต้องรับบทบาทต่างกัน และร่วมกันทำสิ่งใด สิ่งหนึ่งเพื่อชะลอการแพร่กระจายของโรคเพื่อ ปกป้องมนุษยชาติจากไวรัสสี่ชนิดที่กำลังระบาดไปทั่วโลก	เกมทำให้ผู้เล่นเกิดความรู้สึกเสมือนจริง เมื่อ ต้องตัดสินใจเรื่องที่สำคัญต่อการแพร่ระบาดของโรค	บทบาทผู้เล่นที่ไม่สมดุลกัน (เช่น บทบาทที่ไม่สำคัญในเกม) อาจทำให้ ผู้เล่นที่มุ่งเน้นการมีส่วนร่วมส่วนบุคคลมากกว่าการมีส่วนร่วมกับสังคมรู้สึกเบื่อในการเล่นเกมได้

1) เกมเพื่อการตัดสินใจ เพื่อสอนวิธีการป้องกันโรคตามสถานการณ์ -- เราใช้โมเดลเกมแบบ Decision Making ควบคู่ไปกับการเล่น auto side scrolling คือ เป็นการให้ตัวละครเดินไปได้เองเรื่อยๆแบบอัตโนมัติ (คล้ายคุกกี้รัน หรือ มาริโอ้) และเจอจุดตัดสินใจเป็นระยะๆ โดยจะทำให้มีตัวเลือกและคะแนนที่ต่างกัน ระหว่างตัวเลือกถูกต้องและตัวเลือกที่ไม่ควรเลือก พร้อมขึ้นคำเฉลย และบอกวิธีการปฏิบัติตัวที่ถูกต้องควบคู่กันไปด้วย เพื่อเป็นการสอนให้ความรู้แก่ผู้เล่นในทันที

2) เกมสอนการแยกแยะสาเหตุและอาการ เพื่อให้ อสม. สามารถแยกแยะ 2 โรคนี้ออกจากกันได้อย่างถูกต้องชัดเจน เพื่อเป็นประโยชน์สำหรับการคัดกรองผู้ป่วยก่อนส่งโรงพยาบาล

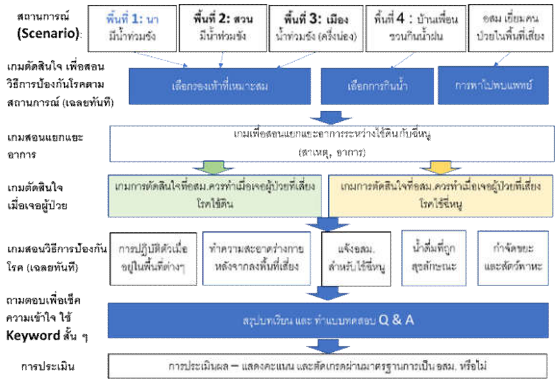
3) เกมตัดสินใจเมื่อเจอผู้ป่วย แบ่งเป็น 2 โหมดการเล่น คือ โหมดโรคไข้ดิน และโหมดโรคไข้ฉี่หนู เพื่อสร้างทักษะ และความตระหนักรู้และการปฏิบัติเพื่อผลักดันผู้ป่วยให้รีบไปพบแพทย์อย่างเร็วที่สุด

4) เกมสอนวิธีการป้องกันโรค สอนการจัดสิ่งแวดล้อมให้ปลอดปัจจัยเสี่ยงต่างๆ เป็นการป้องกันแบบเชิงรุกอย่างแท้จริง

5) เกมถามตอบ เพื่อเช็คความเข้าใจของเนื้อหาที่ อสม. ควรจะทราบ ทั้งนี้ สามารถใช้เกมตรงนี้ในการประเมินผล และมุ่งใช้เกมเพื่อเป็นเกณฑ์ตัดสิน สร้างมาตรฐานให้ อสม. อย่างมีคุณภาพ

6) การประเมินผลเพื่อคัดกรองผู้เล่นสุดท้ายว่ามีคุณสมบัติในมาตรฐานการเป็น อสม. ได้หรือไม่

โดยทั้งกระบวนการเกมทั้งหมด 6 เกมนี้ จะต้องเล่นเสร็จสิ้นภายใน 15-30 นาที (เฉลี่ยแต่ละเกมใช้เวลาประมาณ 2-5 นาที เพื่อให้ไม่เยิ่นเย้อ และผู้เล่นเกิดความเบื่อได้) [14]



รูปที่ 4. แผนภาพแสดงกระบวนการขั้นตอนในเกมที่เราออกแบบ (proposed method) เพื่อการป้องกันเชิงรุกของโรคไข้ดินและโรคไข้ขี้หนู

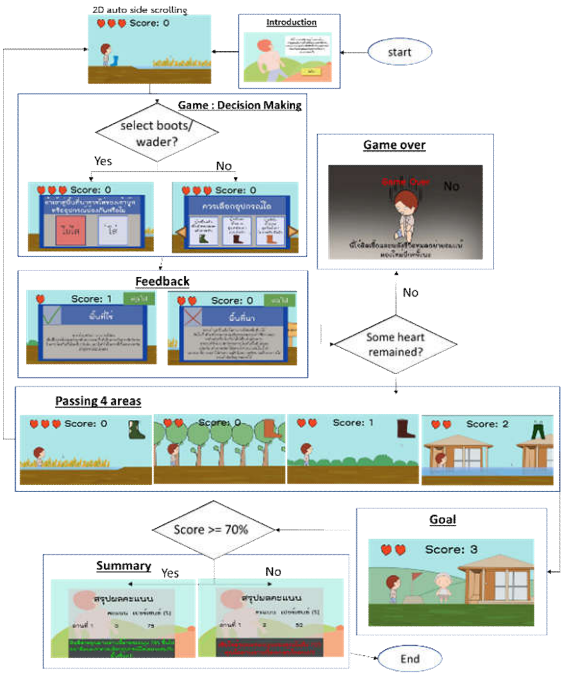
5. การทดลองสร้างเกมเพื่อการประเมินผลจากผู้ใช้จริงในเบื้องต้น

ทีมผู้วิจัยได้ทดลองสร้างเกมในส่วนที่ 1 (จากรูปที่ 4) เพื่อการทดสอบสมมติฐานของปัญหาที่ว่า การใช้เกมดิจิทัลเป็นสื่อสามารถให้ความรู้ได้เทียบเท่าหรือดีกว่าการอบรมโดยผู้เชี่ยวชาญลงพื้นที่โดยตรง เพื่อเป็นการยืนยันการใช้ได้ของเกมสื่อดิจิทัลนี้ว่าสามารถทดแทนวิธีการเดิมได้จริง โดยกระบวนการศึกษาวิจัยมีขั้นตอน ดังนี้

5.1 การสร้างเกม

เกมเป็นการจำลองสถานการณ์ที่ อสม. ต้องเดินทางไปเยี่ยมเพื่อนบ้านโดยต้องผ่านพื้นที่เสี่ยงต่างๆ 4 พื้นที่ได้แก่ นา สวน ไร่ และพื้นที่น้ำขัง และ อสม. ต้องตัดสินใจเลือกอุปกรณ์เพื่อป้องกันความเสี่ยงในการเกิดโรคไข้ดินและโรคไข้ขี้หนู เช่น การเลือกรองเท้าแบบต่างๆ แนวเกม คือ web-based adventure, 2D auto side scrolling เป็นแบบผู้เล่นคนเดียว (Single player) โดยเกมได้พัฒนามาด้วย Unity C# ในเวอร์ชั่นเบื้องต้นนี้ยังไม่มีการเก็บข้อมูลผู้ใช้ในเกม

เกมเริ่มต้นด้วยการเล่าเรื่องผ่าน cut scene ว่าตัวละคร อสม. ชื่อพี่โจ้ต้องเดินทางไปเยี่ยมเพื่อนบ้าน โดยจะต้องเดินทางผ่านพื้นเสี่ยงทั้ง ซึ่งเป็นพื้นที่เสี่ยงต่องการติดเชื้อ โดยตอนต้นของเกม ผู้เล่นจะมีค่าชีวิต (จำนวนหัวใจ) เท่ากับ 3 ผู้เล่นจะทำการเดินทางไปเรื่อย ๆ และก่อนเข้าพื้นที่ ผู้เล่นจะต้องตัดสินใจว่าจะ



รูปที่ 5. แผนภาพแสดง storyboard ของเกมในด่านที่ 1 (เกมเพื่อการตัดสินใจ) เรื่องการสอนวิธีการป้องกันโรคตามสถานการณ์ (เลือกรองเท้าที่ถูกต้องตามแหล่งที่จะทำงาน)

ใส่อุปกรณ์ป้องกันก่อนเข้าพื้นที่หรือไม่ ถ้าเลือกที่จะใช้อุปกรณ์ผู้เล่นจะต้องทำการเลือกอุปกรณ์ป้องกันประเภทต่าง ๆ ได้แก่ 1) รองเท้าบู๊ตพื้นแข็งหนาแข็งกระชับ 2) รองเท้าบู๊ตนินจาพื้นบางสูงเหนือเข่าแนบผิว 3) รองเท้าบู๊ตส้มสูงถึงต้นขาไม่กระชับผิว 4) รองเท้าบูทสีฟ้าพื้นบางสูงแค่เข่าและไม่แนบกับผิว และ 5) ชุดเอี๊ยมลุยน้ำ หลังจากนั้นตัวละครจะเดินผ่านพื้นที่เสี่ยงโดยหน้าจอจะแสดงอุปกรณ์ป้องกันที่ผู้เล่นเลือกไว้ เมื่อผ่านพื้นที่เสี่ยง ผู้เล่นจะได้รับข้อความแจ้งว่าอุปกรณ์ที่เลือกนั้นเหมาะสมกับพื้นที่ที่เข้าไปหรือไม่ รูปที่ 4 แจ้งผู้เล่นเมื่อผู้เล่นเลือกอุปกรณ์ไม่เหมาะสมกับพื้นที่ พร้อมแจ้งผลกระทบ หากผู้เล่นเลือกที่จะไม่ใช้อุปกรณ์ป้องกันหรือผู้เล่นเลือกอุปกรณ์ป้องกันไม่ถูกต้อง ค่าชีวิตในเกมจะลดลง หากค่าชีวิตลดลงจนเหลือ 0 จะเกิด game over และผู้เล่นจะต้องเริ่มเล่นเกมใหม่ หากผู้เล่นเลือกอุปกรณ์ป้องกันได้ถูกต้อง ผู้เล่นจะได้คะแนนและสามารถเดินทางต่อเพื่อผ่านพื้นที่เสี่ยงอื่น ๆ ต่อไป เมื่อผู้เล่นเดินทางผ่านพื้นที่เสี่ยงทั้ง 4 พื้นที่แล้ว เกมจะแสดงหน้าสรุปคะแนนที่ได้ในการเล่นครั้งนั้น ๆ หากได้คะแนนเกิน 70% ผู้เล่นสามารถผ่านด่านได้ โดยเกณฑ์ที่ใช้ในวัดการผ่านการอบรมเป็นไปตามเกณฑ์ การพิจารณาการแปลผลคะแนนความรู้ [15]

5.2 รูปแบบการประเมินผล (ทดสอบว่าแนวคิดของเราใช้การได้)

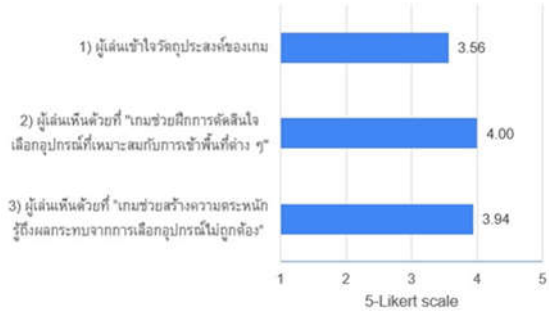
เราสร้างแบบสอบถามความพึงพอใจของผู้เล่นเพื่อทดสอบทิศทางความเห็นและทัศนะคติต่อเกม (Feedback) และรวมถึงประสิทธิภาพการให้ความรู้ของเกม โดยเราสุ่มกลุ่มตัวอย่างที่เป็นผู้ที่ไม่เคยมีความรู้เรื่องโรคไข้ดินหรือโรคไข้ฉี่หนูมาก่อนเลย โดยเลือกอายุที่มีความใกล้เคียงกับ อสม. ในพื้นที่จริง คือ ช่วง 35-65 ปี เป็นกลุ่มตัวอย่างในการทดสอบเล่นเกม

5.3 ผลการทดลอง

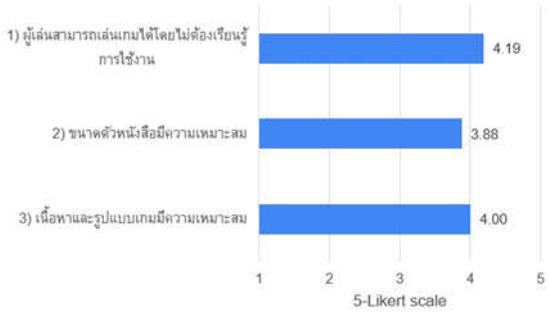
เราทำการศึกษาเบื้องต้นเพื่อค้นหาว่าเกมนำเสนอเนื้อหาการสอนและความพึงพอใจของผู้เล่นอย่างไร ในการศึกษาครั้งนี้มีผู้เข้าร่วมการทดลองทั้งหมด 16 คน: ชาย 6 คนและหญิง 10 คน อายุโดยเฉลี่ยของผู้เข้าร่วมการทดลองคือ 33 ปีซึ่งเป็นอายุใกล้เคียงกับ อสม. ผู้เข้าร่วมการทดลองแต่ละคนจะถูกขอให้เล่นเกมและตอบแบบสอบถามเพื่อให้คะแนนเกมด้านเนื้อหาและความพึงพอใจต่อเกมโดยใช้ 5-Likert scale โดยที่ 1 คือคะแนนเห็นด้วยน้อยสุด และ 5 คือคะแนนเห็นด้วยมากสุด โดยอาสาสมัครใช้เวลาในการเล่นเกมไม่เกิน 2 นาที และทำแบบสอบถามไม่เกิน 2 นาที รวมแล้ว 4 นาที

ในแบบสอบถามมีคำถามแบ่งเป็นสองส่วน: ความรู้และความเข้าใจในเกมและความพึงพอใจของผู้ใช้ที่มีต่อเกม ในส่วนความรู้และความเข้าใจในเกม ผู้เข้าร่วมการทดลองให้คะแนนในด้าน 1) ความเข้าใจในวัตถุประสงค์ของเกม 2) ระดับการยอมรับในเรื่อง "เกมช่วยฝึกการตัดสินใจเลือกอุปกรณ์ที่เหมาะสมกับการเข้าพื้นที่ต่าง ๆ" และ 3) ระดับการยอมรับในเรื่อง “เกมช่วยสร้างความตระหนักรู้ถึงผลกระทบจากการเลือกอุปกรณ์ไม่ถูกต้อง” โดยมีค่าเฉลี่ยอยู่ที่ 3.56 , 4.00 และ 3.94 และส่วนเบี่ยงเบนมาตรฐานอยู่ที่ 0.81, 0.89, 0.77 ตามลำดับ (รูปที่ 6) ซึ่งจะเห็นได้ว่าผู้เล่นโดยเฉลี่ยเข้าใจวัตถุประสงค์ของเกม และเกมได้แสดงให้เห็นชัดเจนเรื่องการสอนการตัดสินใจและการเลือกเครื่องมือให้ถูกต้องได้เป็นอย่างดี ตอบโจทย์เรื่องการเข้าถึงและเข้าใจเนื้อหาที่กำหนดไว้ได้อย่างชัดเจน

ในส่วนความพึงพอใจของผู้ใช้ที่มีต่อเกม ผู้เข้าร่วมการทดลองให้คะแนนในด้าน 1) ผู้เล่นสามารถเล่นเกมได้โดยไม่ต้องเรียนรู้การใช้งาน 2) ขนาดตัวหนังสือมีความเหมาะสม 3) เนื้อหาและรูปแบบเกมมีความเหมาะสม โดยมีค่าเฉลี่ยอยู่ที่ 4.19, 3.88 และ 4.00 และส่วนเบี่ยงเบนมาตรฐานอยู่ที่ 0.83, 0.62, 0.63 ตามลำดับ (รูปที่ 7) แสดงการเปรียบเทียบความพึงพอใจของผู้ใช้ที่มีต่อเกม จะเห็นว่าผู้เล่นโดยเฉลี่ยเห็นด้วยโดยมีค่าใกล้เคียง 4 คือพึงพอใจในระดับดี ตอบโจทย์เรื่องการเรียนรู้ที่ง่ายโดยผู้เล่นไม่จำเป็นต้องมีพื้นฐานการเล่นเกมมาก่อน ขนาดตัวหนังสือ เนื้อหา และรูปแบบเกมมีความเหมาะสมในการเข้าถึงของคนที่อายุเทียบเท่า อสม.



รูปที่ 6. ความพึงพอใจในด้านความรู้และความเข้าใจในเกม ของกลุ่มตัวอย่าง



รูปที่ 7. ผลสำรวจความพึงพอใจโดยเฉลี่ยของกลุ่มตัวอย่างเกี่ยวกับสิ่งต่างๆของเกม

6. การอภิปรายผลจากผู้เชี่ยวชาญชำนาญการจากกรมควบคุมโรค (ผู้ใช้งานจริง)

ในการจัดการทางสาธารณสุขเชิงรุก เป้าหมายของแต่ละพื้นที่ในท้องถิ่น จำเป็นจะต้องมี อสม. 1 คนเพื่อรับผิดชอบดูแล 10 หลังคาเรือน ดังนั้น ในระดับหมู่บ้าน จึงมีความจำเป็นที่ต้องอบรม อสม. จากบุคคลท้องถิ่นในพื้นที่ ซึ่งเป็นหนึ่งในความยากลำบากและท้าทายของทีมผู้เชี่ยวชาญที่จะต้องลงพื้นที่ไปเพื่อให้ความรู้ให้ได้อย่างทั่วถึง และ สร้าง อสม. ให้เพียงพอกับความต้องการให้ได้ตามเป้าหมาย ในการลงพื้นที่แต่ละครั้งจะต้องมีการอบรมในรูปแบบของการบรรยายให้ความรู้ และการอบรมเชิงปฏิบัติการเป็นแต่ละฐาน เพื่อฝึกการทดลอง

ปฏิบัติงานจริงเพื่อให้เกิดความชำนาญ ทั้งนี้ การคิดตัวอย่างการปฏิบัติในแต่ละฐาน อาจต้องมีความสอดคล้องกับธรรมเนียมประเพณี วัฒนธรรมท้องถิ่นด้วยเป็นสำคัญ เพื่อการเรียนรู้ได้ชัดเจนและใช้งานได้จริงในพื้นที่อีกด้วย อนึ่ง จากสถิติจำนวน อสม. ที่ลงทะเบียนทั้งประเทศในปี 2566 มีจำนวน 254,743 คน [16] ในการจัดฝึกอบรมให้แก่อาสาสมัครสาธารณสุขหรือแกนนำชุมชนประเทศต้องใช้ทรัพยากรการอบรม 170 บาท/หัว/คน ซึ่งเป็นงบประมาณที่ไม่รวมค่าวิทยากรหรือค่าใช้จ่ายอื่น ซึ่งเมื่อคิดค่าใช้จ่ายในการฝึกอบรมเพื่อให้อาสาสมัครสาธารณสุขหรือแกนนำชุมชนมีความรู้และทักษะในการเฝ้าระวังป้องกันในชุมชนต้องใช้งบประมาณประมาณเป็นจำนวนถึง 43,306,310.- บาท/ปี ดังนั้นการจัดทำเครื่องมือเพื่อพัฒนาอาสาสมัครสาธารณสุขหรือแกนนำด้วยนวัตกรรมเกมสร้างความรู้ จึงสร้างทักษะจะช่วยลดค่าใช้จ่ายในการฝึกอบรมของภาครัฐได้เป็นอย่างมาก รวมถึงประหยัดทรัพยากรคนในการฝึกอบรมแและมีประสิทธิผลในการสร้าง อสม. ที่มีความรู้ความเชี่ยวชาญในการป้องกันโรคเชิงรุกได้ในระดับดีอีกด้วย (ตารางที่ 2)

ทั้งนี้ จุดแข็งของการอบรมโดยมีผู้เชี่ยวชาญลงพื้นที่โดยตรงคือ การได้บุคคลากรที่มีคุณภาพของความรู้และมีความพร้อมในการปฏิบัติการสูงมาก แต่ข้อจำกัดของการอบรมโดยผู้เชี่ยวชาญโดยตรง คือ 1) จำนวน อสม. ที่อบรมได้แต่ละครั้งมีจำนวนน้อยมาก 2) แม้ว่าจะสามารถวัดประสิทธิภาพของแผนปฏิบัติการสอนในฐานเพื่อการสอนทักษะการตัดสินใจเพื่อป้องกันโรคจากแบบประเมินหลังการอบรมความรู้ประชาชนแต่ปัญหาคือต้องจัดสรรงบประมาณเพื่อจัดการอบรมทบทวนความรู้หรืออบรม อสม. คนใหม่อยู่เสมอ

ซึ่งหากเป็นการใช้สื่อเกมในการอบรมแทน พบว่า จุดแข็งคือ 1) สามารถใช้อบรมคนจำนวนมากได้อย่างไม่จำกัดเวลาสถานที่ ทำให้ประหยัดงบประมาณเป็นจำนวนมาก เรียกได้ว่าลงทุนน้อยแต่ได้ผลเป็นการสร้างบุคคลากรได้จำนวนมากและรวดเร็วกว่ามาก 2)สามารถวัดประสิทธิภาพการเรียนรู้ของบุคคล จากการทำโจทย์ในเกมได้และยังสามารถเก็บเป็นสถิติในฐานข้อมูลได้อีกด้วย และ 3) ผู้เชี่ยวชาญหรือผู้ฝึกสอนสามารถติดตามพฤติกรรมการเรียนรู้ของผู้เรียนได้เพื่อใช้ในการพัฒนาปรับปรุงระบบในครั้งต่อไปได้อีกด้วย แต่จุดอ่อนของการใช้เกมเป็นสื่อแทนการสอนจริงคือ 1) ไม่สามารถปรับตามวัฒนธรรมพื้นถิ่นตามที่ อสม. นั้นถนัดหรือมีประสบการณ์มาได้ สามารถแค่เพียงสอนเนื้อหาได้แบบกลางๆเป็นแค่มาตรฐานเบื้องต้นเท่านั้น 2) ในการสร้างบุคคลากรมานั้น อาจเพียงแต่มีความรู้เท่านั้น ยังไม่สามารถฝึกทักษะในการเข้าปฏิบัติงานจริงได้ ซึ่งในอนาคต จะต้องมีการพัฒนาเกมที่ส่งเสริมการฝึกทักษะต่างๆไปในแต่ละทิศทางมากขึ้นด้วย

ทั้งนี้ อีกปัจจัยหนึ่งที่สำคัญคือ อสม. ส่วนใหญ่ มีอายุ 35-65 ปี ซึ่งส่วนใหญ่มักมีปัญหาเรื่องการอ่าน จึงจำเป็นต้องมีเสียงพากย์ประกอบคำบรรยายในเกมเพิ่มเติม เพื่ออำนวยความสะดวกให้กับผู้เล่นที่ไม่ถนัดในการอ่านเนื้อหาในเกม นอกจากนี้ ควรมีการปรับปรุงเนื้อหาให้มีความตรงประเด็นมากขึ้น และควรดำเนินการพัฒนาเกมในส่วนอื่น ๆ ที่ได้วางแผนไว้ เช่นเกมสอนการแยกแยะอาการไข้ดินและไข้ฉี่หนู รวมถึงเพิ่มสถานการณ์ใหม่ๆ ในเกมเพื่อฝึกการตัดสินใจ เช่นการตัดสินใจเมื่อพบเพื่อบ้านที่ติดเชื้อ การกำจัดสัตว์ที่เป็นพาหะ การเลือกบริโภคน้ำดื่มที่สะอาดเพื่อป้องกันโรค เป็นต้น และควรดำเนินการศึกษาทดลองประสิทธิภาพของเกมในการอบรม อสม. จริงในพื้นที่

7. บทสรุป

ผู้วิจัยได้เสนอแนวคิดการใช้เกมสื่อดิจิทัลเพื่อการฝึกอบรม อสม. ในพื้นที่ ทดแทนการอบรมจากผู้เชี่ยวชาญลงพื้นที่โดยตรง โดยเราได้ทำการทดลองสร้างเกมในด่านแรกขึ้นเพื่อทดสอบสมมติฐานว่า เกมสื่อดิจิทัลจะสามารถใช้อบรมให้ความรู้แก่ อสม. ได้ประสิทธิภาพดีหรือไม่ ซึ่งพบว่า ค่าเฉลี่ยของกลุ่มตัวอย่างตอบแบบสอบถามว่าอยู่ในระดับดี นั่นคือ สามารถใช้ทดแทนได้เป็นอย่างดี ทั้งนี้ เกมด่านแรกเป็นเกมเพื่อสอนการตัดสินใจในการเลือกอุปกรณ์ป้องกันที่เหมาะสมเพื่อป้องกันการติดเชื้อโรคไข้ดินและโรคไข้ฉี่หนู ซึ่งเป็นแนวคิดที่มีประสิทธิภาพสูงในการอบรมและเพิ่มจำนวน อสม. ในพื้นที่ได้อย่างมีประสิทธิภาพ โดยประหยัดทรัพยากรเพื่อฝึกอบรมได้จริงอย่างมหาศาล ทั้งนี้ ในอนาคตทางผู้วิจัยตั้งเป้าร่วมกับกรมควบคุมโรค ที่จะพัฒนาต่อยอดแนวคิดนี้เป็นเกมที่สมบูรณ์เพื่อเป็นเกมที่สามารถประเมินศักยภาพของ อสม. ได้อีกทางหนึ่ง เพื่อใช้เป็นเกณฑ์มาตรฐานเบื้องต้นในการวัดผลการเป็น อสม. ในพื้นที่ได้จริงต่อไป

กิตติกรรมประกาศ

ผู้วิจัยและคณะขอขอบคุณสพ.ญ. วิมวิการ์ ศักดิ์ชัยนานนท์ และสพ.ญ. ชฎาภรณ์ เพียรเจริญ กองโรคติดต่อทั่วไป กรมควบคุมโรค กระทรวงสาธารณสุข ที่ให้การสนับสนุนข้อมูลโรคไข้ฉี่หนูและโรคไข้ดินและร่วมตรวจสอบการออกแบบเกมเนื้อหาโรคไข้ดินและไข้ฉี่หนู และกลุ่มนักพัฒนาเกมรุ่นใหม่ได้แก่ นายกษิดิ์เดช สุทธิพัฒน์อนันต์ นายนพรุจ อินทร์สวรรค์ และนายอิทธิพล มั่นเกษวิทย์ ที่ให้การสนับสนุนการร่วมพัฒนาเกม

เอกสารอ้างอิง

[1] กรมควบคุมโรค, “ข้อมูลโรคเมลลิออยด์.” https://ddc.moph.go.th/disease_detail.php?d=99 (accessed Aug. 18, 2023).

[2] กรมควบคุมโรค, “เลปโตสไปโรซิส (Leptospirosis, Weil Disease).” https://ddc.moph.go.th/disease_detail.php?d=16 (accessed Jun. 17, 2023).

[3] Department of Disease Control (T H), “National Disease Surveillance (Report 506).” https://ddc.moph.go.th/viralpneumonia/index.php (accessed Apr. 28, 2023).

[4] กรมควบคุมโรค, แผนงานด้านการป้องกันควบคุมโรคและภัยสุขภาพ ระยะ 5 ปี (พ.ศ.2566-2570). กรุงเทพมหานคร: สำนักพิมพ์อักษรกราฟฟิคแอนด์ดีไซน์ จำกัด, 2565.

[5] ธ. วิเชียรประภา, พ. หอมสินธุ์, and ร. ศรีสุริยเวศน์, “ปัจจัยที่มีผลต่อพฤติกรรมของอาสาสมัครสาธารณสุขประจำหมู่บ้าน จังหวัดจันทบุรี,” *วารสารสาธารณสุข มหาวิทยาลัยบูรพา*, vol. 7, no. 2, pp. 53–68, 2555, Accessed: Jun. 17, 2023. [Online]. Available: https://www.tci-thaijo.org/index.php/phjbuu/article/download/45594/37734

[6] สำนักงานสาธารณสุขอำเภอสามชุก อำเภอสามชุก จังหวัดสุพรรณบุรี, *คู่มือ อาสาสมัครสาธารณสุข (อสม.)*. Accessed: Jun. 16, 2566. [Online]. Available: https://www.govesite.com/samc00749/document.php?fid=20190322124407IpkHQtn

[7] S. Deterding, R. Khaled, L. E. Nacke, and D. Dixon, “Gamification: Toward a Definition,” in *CHI 2011 gamification workshop proceedings*, ACM Press, 2011.

[8] H. Arshad, G. M. Zen, and R. Z. Abidin, “Integrating Interactive Multimedia Elements to Increase Melioidosis Awareness,” *International Journal of Advanced Trends in Computer Science and Engineering*, vol. 9, no. 3, pp. 4037–4042, Jun. 2020, doi: 10.30534/ijatcse/2020/229932020.

[9] N. N. Azhari *et al.*, “Gamification, a successful method to foster leptospirosis knowledge among university students: A pilot study,” *Int J Environ Res Public Health*, vol. 16, no. 12, Jun. 2019, doi: 10.3390/ijerph16122108.

[10] M. López Hernández, “Healthcare gamification Serious game about COVID-19; Stay at home,” 2020. Accessed: Jun. 16, 2023. [Online]. Available: https://www.diva-portal.org/smash/get/diva2:1481046/FULLTEXT01.pdf

[11] Plague Inc, “‘Plague Inc’ changing sides with anti-pandemic update,” *The Star*, 2020. https://www.thestar.com.my/tech/tech-news/2020/03/26/plague-inc-changing-sides-with-anti-pandemic-update (accessed Jun. 16, 2023).

[12] D. Macsai, “The Great Flu Video Game Turns Pandemic Into Pastime - Fast Company,” 2009. https://www.fastcompany.com/1332286/great-flu-video-game-turns-pandemic-pastime (accessed Jun. 16, 2023).

[13] Pandemic boardgame, “Pandemic is one of the best board games ever made. It could be fun to play right now!,” Mar. 12, 2020. https://www.vox.com/culture/2020/3/12/21175738/pandemic-board-game-coronavirus (accessed Jun. 16, 2023).

[14] S. Al-Rayes *et al.*, “Gaming elements, applications, and challenges of gamification in healthcare,” *Informatics in*

Medicine Unlocked, vol. 31. Elsevier Ltd, Jan. 01, 2022. doi: 10.1016/j.imu.2022.100974.

[15] เปรมจิตร แก้วมูล, "ความรู้และการปฏิบัติในการป้องกันการติดเชื้อสเตร็พโตค็อกคัสซูอิสของประชาชน." https://dric.nrct.go.th/Search/SearchDetail/217015 (accessed Oct. 15, 2562).

[16] กรมสนับสนุนบริการสุขภาพ, "อสม. อาสาสมัครสาธารณสุข." http://xn--y3cri.com/defaults/registered (accessed Jun. 16, 2023).

JIST Journal of Information Science and Technology
Volume 13, NO 1 | JAN – JUN 2023 | 38-44
ISSN: 2651-1053 (Online)
ISSN: 1906-9553 (Print)

An Investigation of Machine Learning Techniques for Loan Default Payments Prediction

*Jesada Kajornrit, Wilawan Inchamnam and Waraporn Jirapanthong**
Collage of Creative Design and Entertainment Technology
Dhurakij Pundit University, Bangkok, Thailand

Received: June 10, 2023; Revised: June 11, 2023; Accepted: June 24, 2023; Published: June 30, 2023

ABSTRACT – In banking business, loan default payments of individual customers are counted as risks that result in the loss of the business. Thus, some assessment mechanisms are needed to assess the risks of individual customers who apply for personal loan products. This paper presents an investigation of machine learning techniques to predict loan default payments based on individual customers information backgrounds. The paper emphasis on the ensemble techniques that mostly used in banking business. Besides the ensemble prediction models, the principal component analysis is also used for further investigation. The experimental results showed that all prediction models provided acceptable prediction of non-defaulting payment class, but provided unacceptable prediction of default payment class. That is because the imbalance nature of the data and the features used are not specific enough for the prediction models to classify the minor class from the major class. This paper acts as an initial study of the credit default payment analysis.

KEYWORDS: Loan Default Payments, Imbalance Data, Machine Learning, Ensemble Techniques, Dimensionality Reduction

1. Introduction

In banking business, individual customer loan evaluation is an important process to reduce the impact of default payments and to mitigate risks into an acceptable level. Presently, the process of loan evaluation is not only limit to human expert decision, but also additional modern analytics techniques. Loan default payments prediction model makes use of the current and historical information related to the customers to make a prediction about the customer ability to pay back on time [1]. Accurate prediction model is able to enhances the decision making of human experts with higher confidence. Thus, developing accurate loan default payments prediction system is an important task for the bank profitability and sustainability.

Presently, machine learning techniques have been taking over several businesses and, of course, banking business is no exception. By using machine learning prediction models, banks are able to predict the probability of loan default payments in advance, and thus helping them in mitigating risks. It is indisputable that the success of machine learning model depends on the high quality of training data. Unfortunately, most loan default payments data (or other related credit default payments data) are usually imbalance [2]. Number of default payments data are less than non-default payments data significantly. Furthermore, features of the dataset are varied among organizations. One way to assess the feasibility of developing a prediction system is to test the limitation of prediction techniques toward the dataset. This paper reports our feasibility study.

The paper is organized as follows; Sections 2 presents some related literatures and machine learning techniques. Section 3 describes the statistical views of the loan default payments dataset. Section 4 present

*Corresponding Author: waraporn.jir@dpu.ac.th

experimental results and some discussions. Section 5 is our conclusion.

2. Related Backgrounds

2.1 Machine Learning Techniques

Decision Trees (DT) are a non-parametric supervised learning method used for classification (and regression) tasks. A DT model is generated from labeled dataset by means of learning algorithms. The popular algorithms include ID3, C4.5, and CART. Some advantages of DT are that it is simple to understand, interpret, and visualize. The DT requires little data preparation [3].

Random Forest (RF) is a meta estimator that fits a number of decision tree classifiers on various sub-samples of the dataset and uses averaging to improve the predictive accuracy and control overfitting. This is called ensemble method. The goal of ensemble methods is to combine the predictions of several base estimators built with a given learning algorithm in order to improve generalizability and robustness over a single estimator [4].

Extremely Randomized Trees (ERT) is one that randomness goes one step further in the way splits are computed. As in Random Forests, a random subset of candidate features is used, but instead of looking for the most discriminative thresholds, thresholds are drawn at random for each candidate feature and the best of these randomly-generated thresholds is picked as the splitting rule. This usually allows to reduce the variance of the model a bit more, at the expense of a slightly greater increase in bias [5].

Both RF and ERT fall into an averaging method, that is, the driving principle is to build several estimators independently and then to average their predictions. On average, the combined estimators are usually better than any of the single base estimator because its variance is reduced.

The core principle of AdaBoost Tree (ADT) is to fit a sequence of weak learners on repeatedly modified versions of the data. The predictions from all of them are then combined through a weighted majority vote to produce the final prediction. Each subsequent weak learner is forced to concentrate on the examples that are missed by the previous ones in the sequence [6].

Gradient Boosted Decision Trees (GBT) is a generalization of boosting to arbitrary differentiable loss functions. GBT is an accurate and effective off-the-shelf procedure that can be used for both regression and classification problems in a variety of areas including web search ranking and ecology [7].

Both ADT and GBT fall into a boosting method, base estimators are built sequentially and one tries to reduce the bias of the combined estimators. The motivation is to combine several weak models to produce a powerful ensemble.

Principle Component Analysis (PCA) is linear dimensionality reduction using Singular Value Decomposition (SVD) of the data to project it to a lower dimensional space. The input data is centered but not scaled for each feature before applying the SVD. PCA technique falls into unsupervised machine learning type. PCA is practically used to visualize high-dimensional data. Also, it is used as a pre-processing step to reduce the dimensions of input vector before feeding to the prediction model [8].

2.2 Literatures Reviews

Many literatures related to loan default payments prediction prefer using interpretable machine learning models liked Decision Trees and its ensemble techniques. That is because such models provide good prediction results, relatively fast training time, require little data preprocessing, and provide human-understandable prediction mechanism. Some literatures are presented as follows.

Soni and Shankar [9] adopted Random Forest classification to predict bank loan defaulting. They claimed that the ensemble technique outperforms single model such as logistic regression, k-nearest neighbours, support vector machine, and decision tree classification.

Shaheen and ElFakharany [11] showed that Random Forest and Gradient Boosting Tree outperform single technique in prediction accuracy when apply these models to predict load default dataset

Fan [10] compared LightGBM to Random Forest algorithms to predict personal loan defaulting. He claimed that LightGBM showed the better prediction;

Lai [12] confirmed the performance of AdaBoost that is show better performance than those of XGBoost, Random Forest and K-Nearest Neighbors, and Neural Network in order to predict loan defaulting from real-world dataset of a prestigious international bank.

Barua et al. [13] explored the use of CatBoost algorithm for loan default prediction. CatBoost is a fast-learning algorithm which is capable of handling categorical data type. They compared to the Random Forest and Gradient Boosting Tree. They claimed that CatBoost achieved the highest accuracy amongst all other algorithms.

Al-qerem et al. [14] presented some different classification methods including Naïve Bayes, Decision Tree, and Random Forest for loan defaulting prediction. Furthermore, comprehensive different pre-processing techniques are being applied on the dataset, and three different feature extractions algorithms are used to enhance the accuracy and performance.

Patel et al. [15] used Logistic Regression, Gradient boosting, CatBoost Classifier, and Random Forest to forecast loan default. They claimed that Gradient boosting and CatBoost Classifier provide the equivalent accuracy and slightly better than Random Forest. While Logistic Regression provided unacceptable result.

Up to this point, however, accuracy of prediction is subject to dataset characteristics, especially, features of the dataset. Furthermore, such problem becomes difficult when it exhibits a profile of imbalanced data, because classifier may misclassify the rare samples from the minority class. To find out an appropriate solution of a specific dataset, some preliminary study is necessary.

3. Loan Default Payments Data

The dataset in this paper is bank loan default payments of individual customers. As the original dataset are confidential, the dataset is anonymized. The features of the dataset are personal information of the customers. The original dataset consists of sixteen features as shown in the Table 1. However, for practical reasons, some features such as LOAN_ DATE or ZIP are removed since they are not appropriate to use. The records that contained *NaN* values are removed. Table 2 shows some statistic of numeric features and Table 3 shows some statistics and numeric code of category features.

The dataset contains 42767 records of customers details which are labeled as default payment (1) or non-default payment (0). The dataset is divided into cross- validation data and final-validation data. Number of cross-validation data is 32075 records and number of final-validation data is 10692 records. The imbalance ratio of cross-validation data is 1:4.5 and 1:4.8 for final-validation data.

4. Experiment and Discussion

The entire loan defaulting payment dataset are divided into cross-validation dataset and final-validation dataset. The former dataset is used to determine the optimal hyper parameters of the models, and is also used to train the models for final validation. The later dataset is used to validate performance of the trained models. The proportion of the data division is about 75 to 25 percent.
The machine learning techniques in this experiment include Decision Tree Classifier (DT), Random Forest (RF), Extremely Randomized Tree (ERT), AdaBoost Tree (ABT), Gradient Boosting Tree (GBT). All algorithms are developed based-on Scikit-Learn library (*https://scikit-learn.org*). The validation measures include accuracy, precision, recall and F1-scores. The number of folds in cross-validation process is set to 5.

The results from the cross-validation process pointed out that “gini” is the good selection criterion for all selected models. The DT model selected 30 for the “max_depth” parameters. The RF model selected 20 for the “n_estimator” parameter and do not set limitation of the “max_depth” parameter. The ERT model was set the same as the RF model, except select 15 for the “n_estimator” parameter. The ABT model set “n_estimator” to 20, and the rest parameters are the same as the RF model. Finally, the GBT model sets “n_estimator” to 50. All optimized models are tested with final-validation dataset and the results are shown in the Table 4.

According to the Table 4, overall accuracy measures of the models are fall between 70 to 80 percent. The model performance can be ranked as RF > ERT > ABT > GBT > DT respectively. This results obviously show that ensemble techniques yield better accuracy than single model, in turn, Bagging technique yields better accuracy than Boosting techniques. However, the dataset is likely to imbalance, accuracy measure may not give the substantial performance of the prediction

In term of class_0 prediction (non-defaulting payment), overall result is acceptable. The precision measures are about 83 percent and the recall measures range between 79 to 95 percent. The model performance based on this measure are the same as one describing by the accuracy measures. The Bagging ensemble technique show the best premising prediction and all ensemble techniques provide better prediction than single model outstandingly.

Table 1. The dataset features

No	Features	Type	Use	Value Range
1	AGE	Integer	Y	[18, 71]
2	COMPANY_TYPE	Category	Y	5 Unique Values
3	CUSTID	Identifier	Y	
4	CUSTOMER_DOB	Category	N	-
5	EDUCATION	Category	Y	4 Unique Values
6	LOAN_AMOUNT	Integer	Y	[19415, 100513]
7	LOAN_DATE	Date	N	-
8	MARITAL_STATUS	Category	Y	4 Unique Values
9	NO_OF_DEPENDENT	Integer	Y	[0, 44]
10	SEX	Category	Y	2 Unique Values
11	STATE_NAME	Category	Y	21 Unique Values
12	TOTAL_MONTHLY_INCOME	Integer	Y	[0, 1500000]
13	YEARS_OF_EXPERIENCE	Integer	Y	[0, 70]
14	YRS_IN_PRESENT_JOB	Integer	Y	[0, 60]
15	ZIP	Category	N	-
16	LABEL	Category	Y	2 Unique Values

Table 2. The dataset statistics of numeric features

Statistics	Age	Loan Amount	Number of Dependent	TotalMonthly Income	Year of Experience	Year in Present Job
mean	33.371	49700.842	1.048	16784.980	6.450	6.209
std	9.589	6462.786	1.417	14917.690	6.614	6.156
min	18.000	19415.000	0.000	0.000	0.000	0.000
max	71.000	100513.000	44.000	1500000.000	70.000	60.000
skewness	0.674	-0.132	2.648	55.410	2.189	2.115
kurtosis	-0.311	0.482	36.034	5056.051	6.074	4.932

Table 3 The dataset statistic of category features

Code	Name	Counts
COMPANY_TYPE		
0	Government	6883
1	Individual	7047
2	Private limited company	18817
3	Public limited company	1856
4	Others	8162
EDUCATION		
0	High school	15610
1	Graduate	18787
2	Post-graduate	970
3	Others	7400

Code	Name	Counts
MARITAL_STATUS		
0	married	31365
1	single	11292
2	widowed	76
3	divorced	34
SEX		
0	male	4240
1	female	38527

Table 4. Experimental Results without PCA transformation

Models	Class 0			Class 1			Accuracy
	Precision	Recall	F1	Precision	Recall	F1	
DT	0.833	0.799	0.816	0.194	0.233	0.212	0.701
RF	0.832	0.959	0.891	0.268	0.072	0.113	0.806
ERT	0.832	0.924	0.876	0.225	0.106	0.144	0.783
ABT	0.834	0.862	0.848	0.208	0.174	0.189	0.744
GBT	0.834	0.820	0.827	0.202	0.218	0.210	0.716

Table 5. Experimental Results with PCA transformation

Models	Class 0			Class 1			Accuracy
	Precision	Recall	F1	Precision	Recall	F1	
DT	0.831	0.818	0.824	0.188	0.203	0.195	0.712
RF	0.830	0.943	0.883	0.212	0.074	0.109	0.793
ERT	0.831	0.919	0.873	0.210	0.103	0.138	0.779
ABT	0.829	0.871	0.850	0.182	0.137	0.156	0.745
GBT	0.832	0.832	0.832	0.193	0.193	0.193	0.722

However, in term of class_1 prediction (defaulting payment), overall result is unacceptable. All classifiers provide very low performance in both precision and recall measures. Precision measures range between 19 to 26 percent. and the recall measures fall under 23 percent. Considering the recall measures, the model performance is contrast to class_0 prediction results. The higher value comes from single classifier (DT) and the lower values come from ensemble technique, especially, Bagging algorithms.

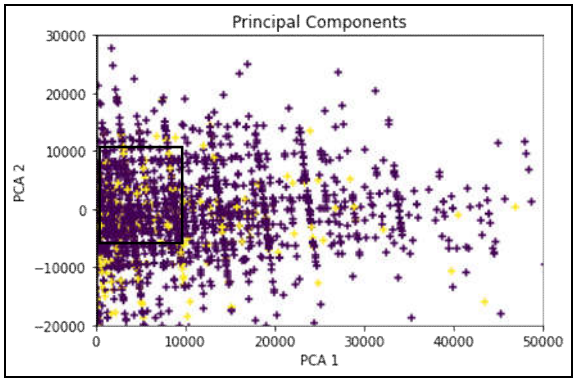
Some further analysis is conducted by unsupervised technique because of the unacceptable prediction of the class_1. This experiment used PCA technique to map high-dimensional default payments data into two dimensions in order to observe homogeneity of the data. The Figure 1 (a) shows the scatter plot between the first and second principal components of the data. The Figure 1 (b) shows the expansion view of the principal components in the rectangle area.

According to the Figure 1, it is possibly that the data points of the both classes are too close. Such data are not in clustered regions. This heterogeneous data is very difficult for classifier to create boundary decision between them. Further experiment is conducted a bit more on the hypothesis that “is it possible to use PCA as preprocessing step before prediction by machine lineaging”. After observed the variance of principle components (Figure 2), the data are transformed by PCA into three-dimensional data

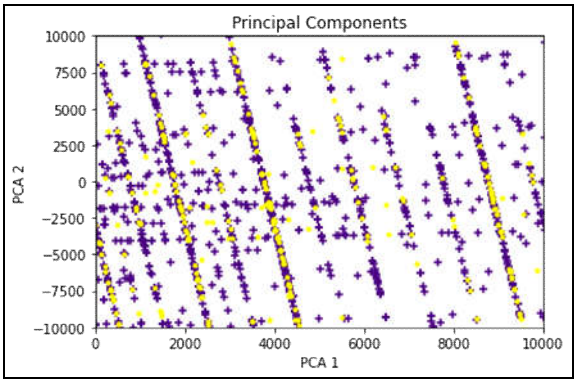
(dimensional reduction). The experimental result show in Table 5. Unfortunately, it seems that PCA techniques cannot improve the performance of class_1 prediction.

The experimental results aforementioned could be summarized as:

- The ensemble models provide better overall accuracy and prediction of the major class than the single model.



(a)



(b)

Figure 1. The first two PCA components of default payment data

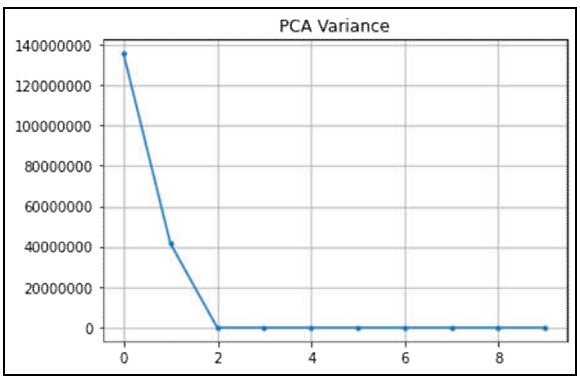


Figure 2. The variants of the PCA transformation.

- In turn, the Bagging techniques provide better overall accuracy and prediction of major class than Boosting techniques
- The ensemble techniques tend to increase the recall measure of the major class, especially, the Bagging technique, but do not improve the precision measure of the major class.
- The ensemble techniques tend to increase precision measure and decrease recall measure of the minor class, especially, Bagging technique.
- PCA transformation (dimensional reduction) may not be an appropriate for preprocessing step as the technique cannot separate the major class form the minor class.

So, the next step to achieve this challenge are as follows.

- Deep down analysis of dataset is need so as to find more relationships in the data. Some data preprocessing may be applied to transform values or to filter out unnecessary information.
- The number of features of the data may not be adequate to classify both classes. More features must be investigated form the data warehouse.
- As the data tend to be imbalance, some techniques such as SMOTE or ADASYN must be investigated (*https://imbalanced-learn.org*).

5. Conclusion

This paper presents an investigation of machine learning technique to predict customer loan default payments. Machine learning techniques include Decision tree, Random Forest, Extremely Randomized Tree, Adaboost Tree, and Gradient Boosted Tree. The experimental results show that prediction of non-default payment class is high accurate. But prediction of default payment class needs to be improved as the poor accuracy. Furthermore, Principal Component Analysis is used to visualize the homogeneity of the data. It showed that minor class data are mix up into major class region. Besides, the dataset is likely to imbalance. This resulted in poor performance of minor class prediction, especially ensemble techniques. Thus, the future study should focus on how to handle the imbalance problem and how to gather and extract more necessary features form the bank's data warehouse**.**

References

[1] A. K. I. Hassan and a. Abraham, "Modeling Consumer Loan Default Prediction Using Ensemble Neural Networks," In Proc. of Int. Conf. on Computing, Electrical and Electronic Engineering (ICCEEE), Khartoum, Sudan, 2013, pp. 719-724.

[2] T. Alam, K. Shaukat, I. A. Hameed, S. Luo, M. U. Sarwar, S. Shabbir, J. Li, and M. Khushi, "An Investigation of Credit Card Default Prediction in the Imbalanced Datasets," IEEE Access, Vol 8. October 2020, pp. 201173-201198.

[3] M. Dumont, R. Maree, L. Wehenkel, and P. Geurts, "Fast multi-class image annotation with random subwindows and multiple output randomized trees,", in Conf. Computer Vision Theory and Applications, Lisboa, Portugal, February 2009.

[4] L. Breiman, "Random Forests, "Machine Language, Vol. 45, No. 1, October 2001, pp 5–32.

[5] P. Geurts, D. Ernst., and L. Wehenkel, "Extremely randomized trees", Machine Learning, Vol. 63, No. 1, 2006, pp. 3-42.

[6] Y. Freund, and R. Schapire, "A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting," Journal of Computer and System Sciences, Vol. 55, No. 1, August 1997, pp. 119-139.

[7] J. H. Friedman. "Greedy function approximation: A gradient boosting machine," Ann. Statist, Vol. 29 No. 5, October 2001, pp. 1189 - 1232.

[8] O. Sornil and J. Kajornrit, "Improving performance of neural network models for intrusion detection using singular value decomposition", Recent Advances in Intelligent Systems and Signal Processing, WSES Press, 2003, pp. 338-342.

[9] A. Soni and K. C. P. Shankar, "Bank Loan Default Prediction Using Ensemble Machine Learning Algorithm" In Proc. of 2nd Int. Conf. on Interdisciplinary Cyber Physical Systems (ICPS), Chennai, India, 2022, pp. 170-175.

[10] S. Fan, "Design and implementation of a personal loan default prediction platform based on LightGBM model," In Proc. on 3rd Int. Conf. International Conference on Power, Electronics and Computer Applications, Shenyang (ICPECA), China, 2023, pp. 1232-1236.

[11] S. K. Shaheen and E. ElFakharany, "Predictive analytics for loan default in banking sector using machine learning techniques," In Prod. Of 28th Int. Conf. on International Conference on Computer Theory and Applications (ICCTA), Alexandria, Egypt, 2018, pp. 66-71.

[12] L. Lai, "Loan Default Prediction with Machine Learning Techniques," In Prod. of Int. Conf. International Conference on Computer Communication and Network Security (CCNS), Xi'an, China, 2020, pp. 5-9.

[13] S. Barua, D. Gavandi, P. Sangle, L. Shinde, and J. Ramteke, "Swindle: Predicting the Probability of Loan Defaults using CatBoost Algorithm," In Proc. of 5th Int. Conf. on Computing Methodologies and Communication (ICCMC), Erode, India, 2021, pp. 1710-1715.

[14] A. Al-qerem, G. Al-Naymat, and M. Alhasan, "Loan Default Prediction Model Improvement through Comprehensive Preprocessing and Features Selection," In Proc. of Int. Arab Conference on Information Technology (ACIT), Al Ain, United Arab Emirates, 2019, pp. 235-240.

[15] B. Patel, H. Patil, J. Hembram, and S. Jaswal, "Loan Default Forecasting using Data Mining," In Proc. of Int. Conf. for Emerging Technology (INCET), Belgaum, India, 2020, pp. 1-4.

JIST Journal of Information Science and Technology
Volume 13, NO 1 | JAN – JUN 2023 | 45-53
ISSN: 2651-1053 (Online)
ISSN: 1906-9553 (Print)

การประเมินความสามารถในการคำนวณท่วงท่าการเคลื่อนไหวของมนุษย์ โดยใช้ MediaPipe เพื่อการฟื้นฟูสมรรถภาพจากโรคหลอดเลือดสมองที่บ้าน

Enhancing Stroke Rehabilitation at Home: Evaluating MediaPipe for Human Pose Estimation

นรภัทร ลาภชูรัต[1], วรวัฒน์ เชิญสวัสดิ์[2] และ กิ่งกาญจน์ สุขคณาภิบาล[2]*

Norapat Labchurat[1], Worawat Choensawat[2] and Kingkarn Sookhanaphibarn[2]*

[1]คณะเทคโนโลยีสารสนเทศและนวัตกรรม สาขาวิชาเทคโนโลยีสารสนเทศและวิทยาการข้อมูล

School of Information Technology and Innovation Department of Information Technology and Data Science

[2]ศูนย์นวัตกรรมเฉพาะทาง มหาวิทยาลัยกรุงเทพ

Center of Specialty Innovation (CoSI), Bangkok University

Received: May 31, 2023; Revised: June 4, 2023; Accepted: June 25, 2023; Published: June 30, 2023

ABSTRACT – This research experiment evaluates the accuracy of MediaPipe, a computer vision technology, in estimating human pose compared to motion capture technology. The experiment utilizes poses derived from the Fugl-Meyer Assessment for upper extremity (FMA-UE). Webcams are installed at three positions: front 90º, left 40º, and right 140º of the participant, while participants wear motion capture suits. The experiment begins with simultaneous movements performed on both sides of the body, with data recorded from both systems simultaneously. We calculate angles of body joints and compare the accuracy percentage. The results indicate that MediaPipe's accuracy in estimating pose based on the FMA-UE was not significantly high. Due to limitations in estimating poses in the vertical plane from z-axis and the range resulted in wider angles than reality. However, the usage of normalization for 2D angle calculation presents the better accuracy. Therefore, MediaPipe can still be used for developing interactive applications and assist the doctor evaluate patient, but further research and development are needed for evaluating patients during FMA-UE exercises.

KEYWORDS: Computer vision, MediaPipe, Human pose estimation, Rehabilitation, Fugl-Meyer Assessment, Home-based rehabilitation, Upper limb rehabilitation

บทคัดย่อ -- งานวิจัยนี้เน้นการประเมินความแม่นยำของเทคโนโลยีคอมพิวเตอร์วิทัศน์ MediaPipe ในการประมาณค่าท่วงท่าของมนุษย์โดยเปรียบเทียบกับเทคโนโลยีโมชันแคปเจอร์ โดยจะใช้ท่วงท่าอ้างอิงมาจาก Fugl-Meyer Assessment for upper extremity (FMA-UE) ติดตั้งกล้องเว็บแคมสามจุดที่ด้านหน้า 90º ด้านซ้าย 40º และด้านขวา 140º ของผู้เข้าร่วม และผู้เข้าร่วมจะสวมใส่ชุดโมชันแคปเจอร์ตลอดการทดลอง เริ่มต้นการทดลองผู้เข้าร่วมจะทำท่วงท่าโดยการเคลื่อนไหวส่วนของร่างกายทั้งสองด้านพร้อมกัน และข้อมูลจากทั้งสองระบบถูกบันทึกพร้อมกัน ผู้วิจัยจะคำนวณองศาของข้อต่อในร่างกายและเปรียบเทียบร้อยละความแม่นยำ ผลการวิจัยพบว่าความแม่นยำของ MediaPipe ในการประมาณค่าท่วงท่า

**Corresponding Author: worawat.c@bu.ac.th*

อ้างอิงจาก FMA-UE ไม่สูงมาก เนื่องจากมีข้อจำกัดในการประมาณค่าท่วงท่าในระนาบดิ่งของแกนแนวลึกและมีช่วงขององศาที่ได้กว้างเกินความเป็นจริง อย่างไรก็ตามการทำให้ข้อมูลขององศาจากการประมาณค่าท่วงท่าเป็นมาตรฐานเดียวกันกับ MoCap ในการคำนวณแบบสองมิติแสดงให้เห็นว่าการคำนวณมีความแม่นยำมากขึ้น ดังนั้น MediaPipe ยังคงสามารถใช้ในการพัฒนาแอปพลิเคชันเชิงตอบโต้และช่วยเหลือแพทย์ประเมินผู้ป่วยเบื้องต้นได้ แต่จำเป็นจะต้องวิจัยและพัฒนาต่อยอดจาก MediaPipe เพื่อใช้ในการประเมินผู้ป่วยขณะทำท่วงท่าจาก FMA-UE

คำสำคัญ: คอมพิวเตอร์วิทัศน์, MediaPipe, การประมาณค่าท่วงท่าของมนุษย์, การฟื้นฟูสมรรถภาพ, แบบประเมิน Fugl-Meyer, การฟื้นฟูสมรรถภาพที่บ้าน, การฟื้นฟูสมรรถภาพรยางค์บน

1. บทนำ

จากผลการสำรวจในปี 2562 โรคหลอดเลือดสมองเป็นสาเหตุสำคัญอันดับต้น ๆ ของโลกและของไทยที่ทำให้เกิดการเสียชีวิตและความพิการ [1] ผู้ป่วยที่รอดชีวิตจากโรคหลอดเลือดสมองส่วนใหญ่จะเกิดอาการอ่อนแรงครึ่งซีก หรือเกิดการอ่อนแรงของกล้ามเนื้อในด้านหนึ่งของร่างกายซึ่งทำให้ไม่สามารถช่วยเหลือตัวเองได้ จึงต้องได้รับการทำกายภาพบำบัด การออกกำลังกาย และการฟื้นฟูสมรรถภาพเพื่อให้สามารถกลับมาใช้ชีวิตประจำวันได้อย่างปกติที่สุด [2]

การประเมินและการติดตามผลการฟื้นฟูของผู้ป่วยแบบเดิม ผู้ป่วยจะต้องพบแพทย์ในโรงพยาบาลหรือคลินิก ฉะนั้นผู้ป่วยจำเป็นต้องเดินทางไปยังสถานที่ห่างไกลเพื่อขอรับการฟื้นฟู [3] จึงเป็นเหตุให้ผู้ป่วยขาดการนัดพบแพทย์และไม่สามารถได้รับการฟื้นฟูสมรรถภาพอย่างสม่ำเสมอ อีกทั้งสถานการณ์ระบาดของโรคโควิด 19 ที่กำลังเกิดขึ้นกับโลกในปัจจุบัน ทำให้มีการเข้ารับการฟื้นฟูสมรรถภาพที่โรงพยาบาลหรือคลินิกลดน้อยลงและทำได้อย่างยากลำบาก [4] เนื่องจากความจำเป็นในการลดความเสี่ยงของการติดเชื้อไวรัสผ่านการสัมผัส หรือการลดความเสี่ยงระหว่างการเดินทาง

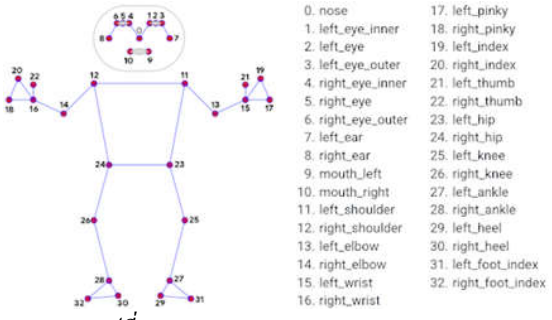
นอกจากนี้ การฟื้นฟูด้วยกระบวนการเดิม ๆ เป็นประจำอาจทำให้เกิดความเบื่อหน่ายแก่ผู้ป่วยได้ และการออกกำลังกายด้วยเทคโนโลยีความเป็นจริงเสมือน เทคโนโลยีจับการเคลื่อนไหวโมชันแคปเจอร์หรือการใช้หุ่นยนต์ฟื้นฟู ถึงแม้จะมีความแม่นยำสูง สามารถช่วยฟื้นฟู ประเมินสมรรถภาพ และลดความเบื่อหน่ายแก่ผู้ป่วยได้ แต่อุปกรณ์เหล่านี้ต้องติดตั้งโปรแกรมล่วงหน้าที่บ้านผู้ป่วย และมีต้นทุนที่สูง [3] ในบทความนี้จึงเสนอการใช้อุปกรณ์เว็บแคมจับภาพทั่วไปเพื่อเป็นอีกทางเลือกหนึ่งสำหรับการฟื้นฟู

ทางผู้วิจัยจึงเล็งเห็นความสำคัญในการทำให้ผู้ป่วยสามารถเข้าถึงการฟื้นฟูสมรรถนะได้อย่างง่ายที่บ้าน ส่งเสริมแรงจูงใจ และมีต้นทุนต่ำบนเว็บเบราว์เซอร์ด้วยเทคโนโลยีประมาณค่าท่วงท่าการเคลื่อนไหวของมนุษย์ (Pose estimation) ผ่านอุปกรณ์ที่สามารถส่งภาพไปในหน่วยประมวลผลกลาง (CPU) หรือหน่วยประมวลผลกราฟิกส์ (GPU) เช่น กล้องเว็บแคม กล้องไอแพดหรือกล้องที่สามารถจับค่าสี RGB เพื่อเปลี่ยนท่วงท่าออกมาเป็นอินพุตหรือคำสั่งสำหรับเกมออกกำลังกาย

อย่างไรก็ตาม คอมพิวเตอร์วิทัศน์ได้ใช้การประมาณค่าของท่วงท่าโดยการเรียนรู้ของเครื่องหรือ Machine Learning (ML) ผ่านภาพจากกล้องเว็บแคมซึ่งสามารถส่งข้อมูลได้เพียงรูปภาพเท่านั้น ดังนั้นมุมกล้องที่แตกต่างกันอาจจะทำให้ได้ความแม่นยำที่แตกต่างกัน และการออกแบบเพื่อใช้ท่วงท่าที่มาจาก FMA-UE จำเป็นจะต้องพิสูจน์ความแม่นยำในการประมาณค่าท่วงท่าที่มีการเคลื่อนไหวไปในทิศทางแนวแกนลึก (Z-Axis) ด้วยการคำนวณองศาแบบทั้งสองและสามมิติ และค้นหาวิธีการใช้งานคอมพิวเตอร์วิทัศน์ที่เหมาะสมก่อนจะนำมาใช้ในการพัฒนาเกมออกกำลังกาย หรือใช้เพื่อประเมินผู้ป่วยโรคหลอดเลือดสมองขณะออกกำลังกายได้จากที่บ้าน

การวิจัยนี้จึงใช้เครื่องมือคอมพิวเตอร์วิทัศน์ (Computer vision) ที่มีชื่อว่า MediaPipe ประมาณค่าท่วงท่าของมนุษย์ในขณะที่กำลังออกกำลังกาย และออกแบบท่วงท่าเพื่อการออกกำลังกายจาก FMA-UE ผู้วิจัยได้ทำการทดลองเปรียบเทียบความแม่นยำทั้งแบบสองมิติและสามมิติ ด้วยระบบจับการเคลื่อนไหวโมชันแคปเจอร์ เพื่อตรวจสอบความเป็นไปได้ใน

การสร้างแอปพลิเคชันให้กับผู้ป่วยโรคหลอดเลือดสมองออกกำลังกายเพื่อฟื้นฟูสมรรถนะได้จากที่บ้าน และตรวจสอบร้อยละความแม่นยำในการประเมินสมรรถภาพ เพื่อใช้เป็นเครื่องมือในการช่วยแพทย์สามารถเข้าถึง และให้การรักษาแก่ผู้ป่วยได้จากระยะไกลหรือที่เรียกว่า Telerehabilitation



รูปที่ 1. Blaze pose model (33 Landmarks)

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

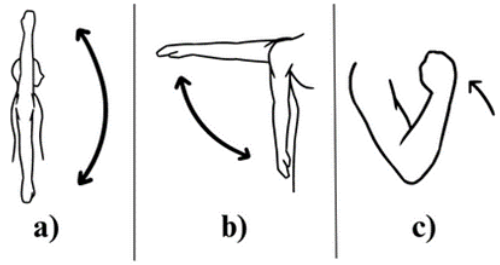
เทคโนโลยีคอมพิวเตอร์วิทัศน์หรือ Computer vision เป็นการใช้คอมพิวเตอร์เพื่อประมวลผลภาพโดยใช้ปัญญาประดิษฐ์ (Artificial Intelligence) หรือการเรียนรู้ของเครื่องเพื่อให้คอมพิวเตอร์สามารถเข้าใจและวิเคราะห์ภาพเช่นเดียวกับมนุษย์ได้ผ่านอุปกรณ์ที่เป็นกล้อง RGB [5] ซึ่งปัจจุบันมีโปรแกรมและเครื่องมือมากมายที่สามารถนำไปสร้างแอปพลิเคชันเพื่อการประมวลผลภาพได้ เช่น OpenCV, MediaPipe, Openpose, TensorFlow.js และอื่น ๆ [6] [7] [8] โดยในการวิจัยนี้เราได้เลือกใช้ MediaPipe ซึ่งได้พิสูจน์แล้วใน [9] ว่ามีความแม่นยำสูงที่สุดจากเครื่องมือที่ยกตัวอย่างมา

ในวรรณกรรมที่ผ่านมา ได้แสดงให้เห็นถึงการใช้เทคโลยีคอมพิวเตอร์วิทัศน์ และเทคนิคการประมาณค่าท่วงท่าของมนุษย์ในการพัฒนาแอปพลิเคชันเพื่อเสริมสร้างแรงจูงใจและการฟื้นฟูสมรรถภาพแก่ผู้ป่วยที่มีความพิการหรือเกมออกกำลังกายส่งเสริมสุขภาพ [10] และการตรวจสอบความถูกต้องในขณะที่ท่วงท่าจากการเล่นโยคะ [9] นอกจากการประมาณค่าท่วงท่ายังมีการใช้เทคนิคที่เกี่ยวข้องกับคอมพิวเตอร์วิทัศน์อย่างอื่นอีกเช่นการอ่านค่าสีจากพิกเซล (Pixel) ของรูปภาพในขณะที่กำลังสวมใส่อุปกรณ์เสริมอย่างเช่นถุงมือสี [11] หรือการถือสิ่งของเพื่อการใช้เทคนิคตรวจจับวัตถุ (Object detection) [12] แทนที่การตรวจจับพิกัดของร่างกายมนุษย์โดยตรง เทคนิคเหล่านี้เพิ่มความแม่นยำในการโต้ตอบ และผลลัพธ์ของการกระทำที่จะเกิดขึ้นภายในแอปพลิเคชัน

2.1 MediaPipe และการประมาณค่าท่วงท่า

MediaPipe เป็นเครื่องมือที่ผู้วิจัยเลือกใช้ในการวิจัยนี้ และเป็นเครื่องมือที่ถูกพัฒนาโดย Google ซึ่งมีเป้าหมายเพื่อให้นักพัฒนาสามารถสร้างและพัฒนาแอปพลิเคชันคอมพิวเตอร์วิทัศน์ได้ง่ายขึ้น อย่างเช่นเรื่องของการประมาณค่าท่วงท่าของมนุษย์ (Pose estimation) [9] ซึ่งเป็นกระบวนการหาตำแหน่งและทิศทางของส่วนต่างๆ ของร่างกายในภาพหรือวิดีโอ

การประมาณค่าท่วงท่าเป็นกระบวนการที่ใช้ในการตรวจจับการเคลื่อนไหวและระบุตำแหน่งส่วนต่าง ๆ ของร่างกาย โดยส่วนใหญ่ใช้เทคนิคการเรียนรู้ของเพื่อสร้างแบบจำลองที่สามารถระบุโครงสร้างกระดูกของมนุษย์ ซึ่ง MediaPipe มีการใช้การประมวลผลร่วมกันของ CPU และ GPU ผ่านเครื่องมือ OpenGL จึงทำให้มีการประมวลผลที่รวดเร็ว [13] และผลลัพธ์ของการประมาณค่าท่วงท่าจะได้เป็นพิกัด Landmarks จำนวน 33 จุด [14] ดังรูปที่ 1 นอกจากนี้ยังมีโมเดลอื่น ๆ เช่น การตรวจจับใบหน้า มือ และสิ่งของ เป็นต้น



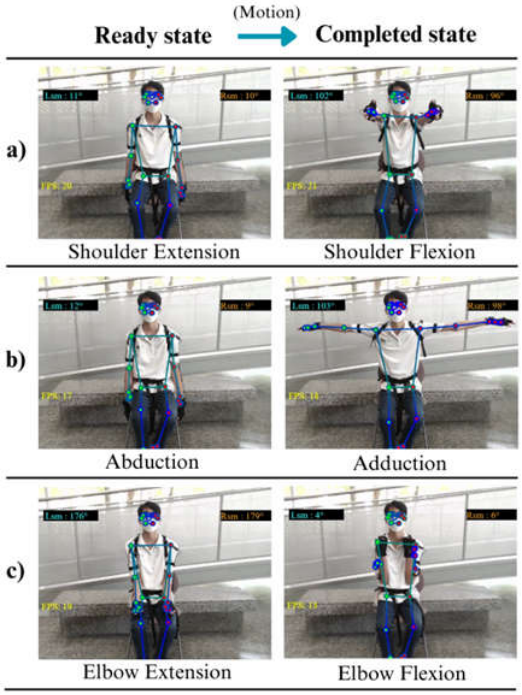
รูปที่ 2. ท่วงท่าที่ใช้ในการวิจัยจาก FMA-UE

2.2 Motion Capture (MoCap)

การวิจัยนี้ผู้วิจัยใช้ระบบจับการเคลื่อนไหวโมชันแคปเจอร์ประเภท Inertial Measurement Unit (IMU) [15] ซึ่งใช้เซนเซอร์ในการจับและบันทึกการเคลื่อนไหวของร่างกาย ซึ่งแต่ละเซนเซอร์ IMU ประกอบด้วยเซนเซอร์จับความเร่ง (accelerometer) และเซนเซอร์วัดความเร็วมุม (gyroscope) เพื่อวัดและบันทึกการเคลื่อนไหวของส่วนต่าง ๆ ของร่างกาย [16] และสามารถนำผลลัพธ์จากเซนเซอร์มาสร้างโมเดลที่แสดงการเคลื่อนไหวของร่างกายในรูปแบบสามมิติ ผ่านโปรแกรม

สำหรับโมชันแคปเจอร์ ด้วยเฟรมเรทมาตรฐานที่ 120 Fps (Frame per second)

ปัจจุบันมีการใช้งาน MoCap เพื่อเปรียบเทียบคุณภาพของเทคโนโลยีที่มีการประมาณค่าท่วงท่าของมนุษย์ [17] และใช้เพื่อการฟื้นฟูสมรรถภาพและประเมินผู้ป่วยโรคหลอดเลือดสมองเพื่อหาระดับของความพิการ หรือประเมินความบกพร่องของร่างกายในการเคลื่อนไหวซึ่งสามารถประเมินได้อย่างละเอียด [18] แต่ข้อเสียคือราคาที่ค่อนข้างจะสูง ส่วนในการวิจัยนี้ผู้วิจัยได้ใช้พิกัดข้อต่อของ MoCap คำนวณมุมและเป็นข้อมูลที่จะทำนาย (Predict) เพื่อหาความแม่นยำการประมาณค่าท่วงท่าของ MediaPipe



รูปที่ 3. ท่วงท่าที่ใช้ในการวิจัยจาก FMA-UE เมื่อทำการประมาณค่าท่วงท่าใน MediaPipe

3. วิธีดำเนินการวิจัย

ในส่วนนี้ผู้วิจัยจะอธิบายขั้นตอนในการดำเนินการวิจัย วิธีเก็บข้อมูลจากผู้เข้าร่วมการทดลอง และขั้นตอนการดำเนินการกับข้อมูลที่ได้มาเพื่อเตรียมนำไปหาผลลัพธ์ของการทดลองต่อไป

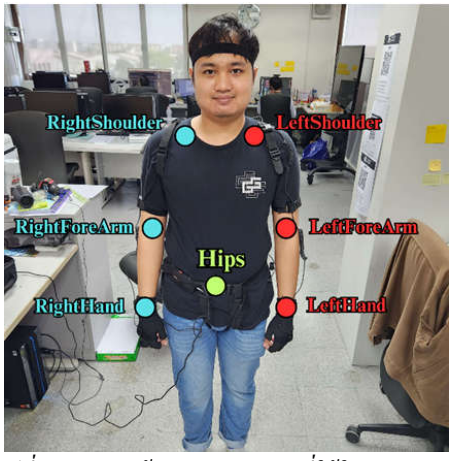
3.1 การออกแบบท่วงท่า

การออกแบบท่วงท่าที่ใช้ในการดำเนินการวิจัยได้อ้างอิงการเคลื่อนไหวจาก Fugl-Meyer Assessment-Upper Extremity หรือ FMA-UE ซึ่งเป็นเครื่องมือทางการแพทย์ที่ใช้ในการประเมินความสามารถของผู้ป่วยอัมพาตครึ่งซีกส่วนรยางค์บน [19] ทางผู้วิจัยเลือกใช้ท่วงท่าเหล่านี้เพื่อหาความเป็นไปได้ในการประเมินสมรรถนะผู้ป่วยระหว่างการฟื้นฟูสมรรถภาพด้วยเกมออกกำลังกายและการกระทำท่วงท่าที่มาจาก FMA-UE

ท่วงท่าที่ใช้ในการวิจัยนี้มีทั้งหมดสามท่วงท่าดังรูปที่ 2 และการเคลื่อนไหวดังรูปที่ 3 ได้แก่ a) การงอหรือเคลื่อนไหวไหล่ขึ้น (Shoulder Flexion), b) การเคลื่อนไหวของไหล่ไปในทิศทางด้านข้างหรือออกจากลำตัว (Abduction) และ c) การงอข้อศอก (Elbow Flexion) โดยภาพทางด้านซ้ายจะเป็นท่าเตรียมพร้อม และภาพด้านขวาจะเป็นท่าที่ต้องทำให้สำเร็จ

3.2 การคำนวณองศาของตำแหน่งในร่างกาย

การเลือกท่วงท่าในการวิจัยนั้นสำคัญต่อตำแหน่งของร่างกายที่ใช้ในการคำนวณผลลัพธ์องศาข้อต่อของ MoCap หรือผลลัพธ์องศาพิกัด Landmark ของ MediaPipe ทั้งนี้เพื่อให้มั่นใจว่าองศาที่ได้จากการเคลื่อนไหวนั้นมาจากตำแหน่งของร่างกายจุดเดียวกัน หรือใกล้เคียงกันมากที่สุด การกำหนดตำแหน่งของร่างกายสามจุดเพื่อเอาไว้ใช้คำนวณองศาจึงจำเป็นอย่างมากที่จะทำให้ผลลัพธ์การทดลองออกมาถูกต้องมากที่สุด



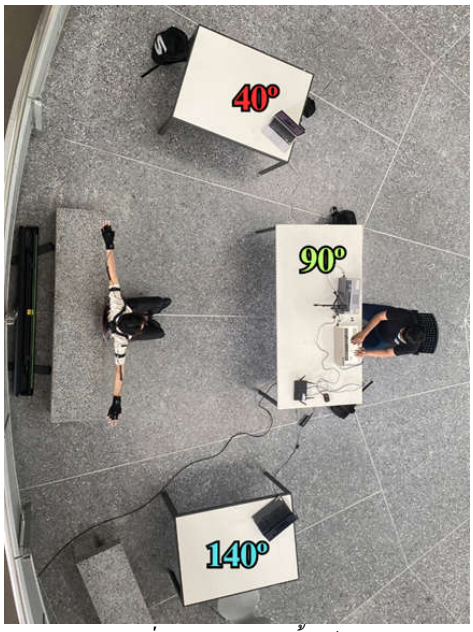
รูปที่ 4. ตำแหน่งข้อต่อของ MoCap ที่ใช้ในการทดลอง

3.2.1 การเลือกข้อต่อของ MoCap

ชื่อข้อต่อ MoCap ที่ผู้วิจัยเลือกใช้แสดงดังรูปที่ 4 เพื่อคำนวณหาองศาของไหล่ในท่วงท่าอไหล่ขึ้น และกางไหล่ออกจากลำตัวในแต่ละด้าน (แทนคำว่า [Side] ด้วยคำว่า Left ด้านซ้ายหรือ Right ด้านขวา) ได้แก่ Hips, [Side]Shoulder [Side]ForeArm ส่วนชื่อข้อต่อที่ใช้เพื่อคำนวณองศาของข้อศอกแต่ละด้านคือ [Side]Shoulder, [Side]ForeArm และ [Side]Hand ตามลำดับ

3.2.2 การเลือก Landmarks ของ MediaPipe

หมายเลข Landmark ของ MediaPipe ดังรูปที่ 1 เพื่อใช้พิกัดมาคำนวณหาองศาของท่วงท่างอไหล่ขึ้นและกางไหล่ออกจากลำตัว เฉพาะในส่วนด้านขวาของร่างกายได้แก่ [24, 12, 14] ส่วนด้านซ้ายของร่างกายได้แก่ [23, 11, 13] ส่วนหมายเลขที่เอามาคำนวณองศาของข้อศอกเฉพาะด้านขวาของร่างกายได้แก่ [12, 14, 16] ส่วนด้านซ้ายของร่างกายได้แก่ [11, 13, 15]



รูปที่ 5. องศาการติดตั้งกล้อง

3.3 ขั้นตอนและวิธีการทดลอง

การติดตั้งกล้องเพื่อส่งภาพให้ MediaPipe ผู้วิจัยได้แบ่งตำแหน่งให้คอมพิวเตอร์แต่ละเครื่องทำมุมกับผู้เข้าร่วมการทดลองดังรูปที่ 5 ได้แก่ มุมที่ 90° เน้นจับทางด้านหน้า เนื่องจากเป็นมุมที่สามารถติดตั้งกล้องเว็บแคมบนมอนิเตอร์ได้ และเป็นมุมปกติที่ผู้คนจะติดตั้งกล้องกัน, มุมที่ 40° และมุมที่ 140° เป็นมุมที่อยู่ด้านซ้ายและขวาตามลำดับ ซึ่งเป็นมุมเฉียงที่สามารถมองเห็นแขนทั้งสองด้านของร่างกายได้อย่างชัดเจน อีกทั้งภาพจากกล้องเว็บแคมเป็นแบบสองมิติ และทำการประมาณค่าท่วงท่าด้วย MediaPipe เป็นพิกัดสามมิติ ผู้วิจัยจึงตั้งสมมติฐานว่ามุมกล้องส่งผลต่อการประมาณค่าท่วงท่าในแต่ละมิติหรือไม่ การแบ่งมุมกล้องดังกล่าวจะสามารถแสดงความแม่นยำ ที่แตกต่างกันของการประมาณค่าในแต่ละมิติของแต่ละมุมกล้องได้

นอกจากนี้ผู้เข้าร่วมการทดลองจะต้องสวมใส่ชุดเซนเซอร์ IMU ของ Motion capture ตลอดการทดลอง และเมื่อการทดลองเริ่มขึ้น ผู้วิจัยจะทำการบันทึกการเคลื่อนไหวของ MoCap และบันทึกการประมาณค่าท่วงท่าของ MediaPipe ทั้งหมดพร้อมเพรียงกันแบบเรียลไทม์

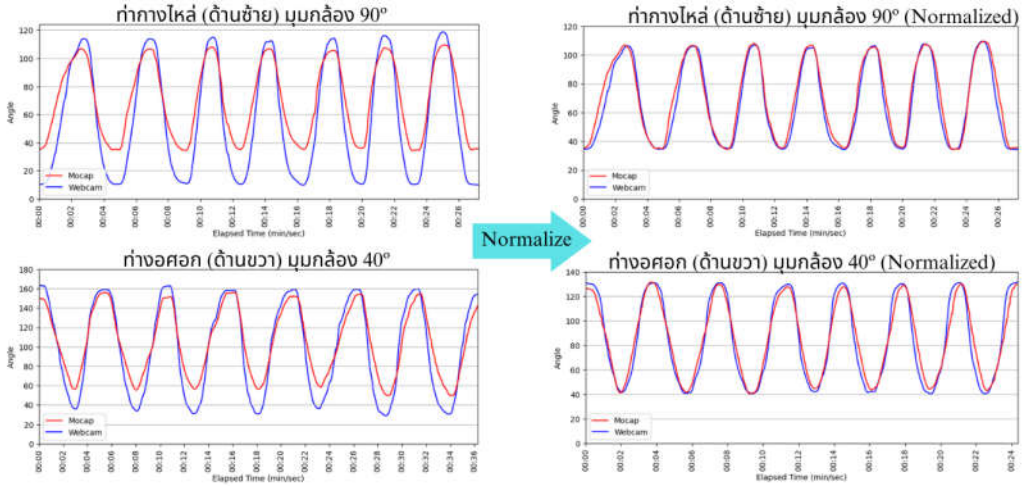
การบันทึกข้อมูลจะแบ่งออกไปในการทดลองแต่ละท่วงท่าดังรูปที่ 3 เริ่มต้นผู้เข้าร่วมการทดลองจะอยู่ในท่าเตรียมพร้อมเพื่อเข้าสู่สถานะ Ready state และทำการเคลื่อนไหวรยางค์ส่วนบนทั้งสองด้านของร่างกายพร้อมกัน (Synchronous) เพื่อทำช่วงของการเคลื่อนไหว (Range of motion) ของท่วงท่าให้สำเร็จ และเข้าสู่สถานะต่อไปคือ Completed state หลังจากนั้นจะทำการเคลื่อนไหวกลับมายังท่าเตรียมพร้อมอีกครั้ง และยิ่งกระทำท่วงท่าเดิมวนซ้ำมากก็ยิ่งทำให้มีข้อมูลเพื่อนำมาเปรียบเทียบมากขึ้นและได้ผลลัพธ์ที่ชัดเจนมากขึ้น ซึ่งในการวิจัยนี้จึงจะทำการวนซ้ำทั้งหมด 30 วินาที ซึ่งจะได้ประมาณ 7 รอบ หลังจากทำครบก็จะสิ้นสุดการทดลองของท่วงท่านั้น ๆ และทำท่วงท่าที่เหลือต่อจนครบทั้งหมดสามท่วงท่า จึงจะเปลี่ยนเป็นผู้เข้าร่วมการทดลองคนถัดไป

3.4 การประมวลผลการเคลื่อนไหว

พิกัด Landmark ที่ได้จาก MediaPipe จะถูกคำนวณองศาออกมาดังข้อ 3.2.2 และผลลัพธ์ที่ได้อาจจะมีความแปรปรวนเนื่องจากสีเสื้อผ้าของผู้เข้าร่วมการทดลอง สีของพื้นหลัง และสิ่งของที่อยู่ในพื้นหลัง อาจจะกระทบกับการประมาณค่าท่วงท่าของเทคโนโลยีคอมพิวเตอร์วิทัศน์ ผู้วิจัยจึงต้องทำการลดสัญญาณข้อมูลหรือ Smooth ข้อมูลขององศาด้วยหลักการ Moving Average (MV) [20] ตามขนาดหน้าต่าง

ขนาดหน้าต่าง (Window size) หรือช่วงที่เกลี่ย ถ้า FPS สูงสามารถกำหนดให้มีขนาดใหญ่ได้ แต่กลับกันถ้า FPS ต่ำ ช่วงที่เกลี่ยก็ควรมีขนาดเล็ก เพราะฉะนั้นการกำหนดขนาดมากเกินไปจะส่งผลต่อการ delay ของคำสั่งหรืออินพุตในระบบ จากการทดลองในการวิจัยนี้ ซึ่งเฟรมเรทของ MediaPipe เฉลี่ยใน

การวิจัยอยู่ที่ 25 FPS และจากการทดลองปรับขนาดต่าง ๆ ผู้วิจัยได้สังเกตว่าขนาดที่ดีที่สุดของการวิจัยนี้คือ 5 หรือ 1 ใน 5 ของเฟรมเรททั้งหมด และ Window size นี้ได้ให้รูปแบบของกราฟที่ไม่แตกต่างกับข้อมูลดิบมากนัก และมีผลลัพธ์ใกล้เคียงกับ MoCap มากที่สุด



รูปที่ 6. ตัวอย่างของข้อมูลจากการคำนวณองศาแบบสองมิติเมื่อนำมาทำการ Normalize

เฟรมเรทการบันทึกข้อมูลต่อวินาทีของ MoCap อยู่ที่120 fps เพราะฉะนั้นการบันทึกข้อมูลเฟรมเรทของระบบทั้งสองจะไม่เท่ากัน ทางผู้วิจัยจึงต้องจัดเรียงให้ข้อมูลของทั้งสองระบบมีจำนวนเฟรมที่เท่ากัน โดยวิธีทำคือจับเฟรมที่มีระยะเวลา (Elapsed time) ใกล้เคียงมากที่สุดเมื่อเทียบกับ MediaPipe มาสร้างเป็นชุดข้อมูลใหม่สำหรับ MoCap เพื่อที่จะสามารถเปรียบเทียบความแตกต่าง และใช้เครื่องมือสถิติคำนวณผลลัพธ์ออกมาได้เฟรมต่อเฟรม

4. การทดลองและผลการทดลอง

ในการทดลองครั้งนี้มีผู้เข้าร่วมทั้งหมด 5 คน ประกอบด้วยเพศชาย 3 คนและหญิง 2 คน อายุของผู้เข้าร่วมอยู่ในช่วง 20-23 ปี

4.1 ชุดข้อมูล

ผู้วิจัยแบ่งชุดผลการทดลองออกเป็นหัวข้อตามตัวแปรที่ใช้ในการสรุปผลได้แก่ ท่วงท่า มิติในการคำนวณ ด้านของร่างกาย และ มุมกล้อง ฉะนั้นการจับคู่ของข้อมูลทั้งสองระบบเพื่อใช้ในการหาค่าผลลัพธ์สถิติจะกระทำโดยตัวแปรที่กำหนดไว้เช่นเดียวกัน การแบ่งชุดข้อมูลลักษณะนี้ทำให้สามารถจำลองสถานการณ์ต่าง ๆ เพื่อให้ได้ผลการทดลองที่มีความละเอียด

4.2 การทดลอง

การหาความแม่นยำของการประมาณค่าท่วงท่าโดย MediaPipe ผู้วิจัยได้ใช้ผลลัพธ์องศาของทั้งสองระบบตามที่ได้อธิบายไปในข้อ 3.2 และใช้สูตรคำนวณหาค่าความผิดพลาดเฉลี่ยสัมบูรณ์หรือ Mean Absolute Error (MAE) ดังสมการที่ (1) คำนวณช่วงของการเคลื่อนไหวโดยใช้ค่าองศาที่ได้จาก MoCap ดังสมการที่ (2) และหาร้อยละความแม่นยำ (Accuracy) ในแต่ละชุดการทดลองดังสมการที่ (3)

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|Me_i - Mo_i| \quad (1)$$

$$ROM = Mo_{max} - Mo_{min} \quad (2)$$

$$Accurency = 1 - (MAE/ROM) \quad (3)$$

โดยกำหนดให้ *Mo* และ *Me* เป็นองศาที่ได้มาจาก MoCap และ MediaPipe ตามลำดับ i เป็นดัชนีของข้อมูลทั้งสอง n เป็นจำนวนเฟรมที่ใช้ในการเปรียบเทียบ และ *ROM* เป็น Range of motion หรือช่วงของการเคลื่อนไหว

นอกจากการวัดความแม่นยำขององศาปกติที่ได้จากการประมาณค่าท่วงท่า ผู้วิจัยยังได้นำข้อมูลทั้งหมดมาทำ Normalization หรือทำให้ข้อมูลขององศาจากการประมาณค่า

ท่วงท่าเป็นมาตรฐานเดียวกันกับ MoCap ดังตารางที่ 1 และรูปที่ 6 ตัวอย่างของข้อมูลเมื่อนำองศามาทำการ Normalize เพื่อตรวจสอบความแม่นยำเพิ่มเติมว่าหลังจากทำให้องศาเป็นมาตรฐานเดียวกัน ข้อมูลจะมีความแม่นยำมากขึ้นหรือไม่

4.3 ผลการทดลอง

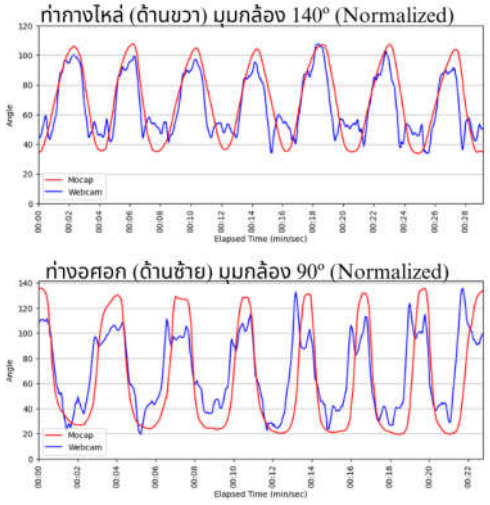
4.3.1 ผลลัพธ์จากการคำนวณสามมิติ

จากผลการทดลองทำให้ผู้วิจัยพบว่าการคำนวณท่วงท่าพิกัดสามมิติของ MediaPipe มีความแปรปรวนเป็นอย่างมาก

ตารางที่1. แสดงสถิติต่าง ๆ โดยเฉลี่ยของท่วงท่าที่ใช้ในการทดลองจากผู้เข้าร่วมการทดลองทั้งหมดห้าคน โดยแบ่งเป็นด้านซ้ายและด้านขวา แบ่งตามการคำนวณองศาแบบปกติและองศาแบบ Normalize และแบ่งตามมุมกล้องทั้งสาม

สถิติ	ท่วงท่า	ด้านของร่างกายในแต่ละมุมของเว็บแคม องศาปกติและหลังจาก Normalize											
		ซ้าย			ขวา			ซ้าย (Normalized)			ขวา (Normalized)		
		40º	90º	140º	40º	90º	140º	40º	90º	140º	40º	90º	140º
Mean	งอไหล่ขึ้น	31	32	22	23	31	28	8	10	4	4	11	6
Absolute	กางไหล่	26	18	30	30	18	29	8	4	9	8	5	11
Error	งอข้อศอก	15	46	21	20	47	13	6	16	7	8	16	7
Accuracy	งอไหล่ขึ้น	42	40	59	60	46	53	86	82	92	93	81	89
Percentage	กางไหล่	60	73	54	56	73	58	88	93	87	89	93	84
	งอข้อศอก	84	51	77	81	54	87	94	83	92	92	84	93

แม้จะSmooth และ Normalize ข้อมูลแล้วดังตัวอย่างกราฟที่แสดงในรูปที่ 7 เนื่องจากการใช้ ML ประมาณค่าท่วงท่าในแนวแกนลึก และเป็นขีดจำกัดของอุปกรณ์เว็บแคมที่จับได้เพียงภาพสองมิติ จึงทำให้การประมาณค่าท่วงท่าในระนาบตัดแนวดิ่ง (Sagittal Plane) ของคอมพิวเตอร์วิทัศน์จึงไม่มีประสิทธิภาพเท่ากับอุปกรณ์หรือฮาร์ดแวร์ที่มีการติดตั้งเซนเซอร์ตรวจจับความลึก



รูปที่ 7. ตัวอย่างของข้อมูลองศาสามมิติ

แต่ละมุมกล้อง ถ้าหากวิเคราะห์ด้วยสีพื้นหลังจากตารางดังกล่าวว่ามุมกล้องใดสามารถประมาณค่าท่วงท่าในแต่ละด้านได้ร้อยละความแม่นยำมากที่สุด ก็จะพบได้ว่าจากตารางที่ 1 จะมี Pattern หรือรูปแบบชัดเจนว่าผลลัพธ์ความแม่นยำที่ได้จากการคำนวณองศาสองมิติและหลังจาก Normalization ผลลัพธ์ในแต่ละมุมกล้องมีความสม่ำเสมออย่างมีนัยสำคัญ ซึ่งผู้วิจัยจะอธิบายต่อในย่อหน้าถัดจากนี้

องศาของท่วงท่ากางไหล่จะได้ความแม่นยำมากที่สุดเมื่อมุมกล้องหันเข้าหาผู้เข้าร่วมการทดลอง 90º เนื่องจากมุมและการเคลื่อนไหวของท่วงท่าขนานกับการมองเห็นของกล้องพอดี กลับกันท่วงท่าที่ใช้การยืดส่วนของร่างกายออกไปด้านหน้าอย่างท่วงท่างอข้อศอกและงอไหล่ขึ้น กลับมีความแม่นยำต่อการคำนวณองศาในมุมด้านข้างหรือมุมเฉียงมากกว่ามุมกล้องที่อยู่ด้านหน้า เพราะส่วนของร่างกายได้ยืดออกไปทางด้านหน้าและขนานกับการจับภาพของมุมกล้องด้านข้าง อย่างไรก็ตามท่วงท่างอไหล่ขึ้นกลับได้ความแม่นยำของด้านตรงข้ามมากว่าด้านของร่างกายด้านเดียวกับกล้อง นั่นเป็นเพราะผลลัพธ์ของการคำนวณองศาสองมิติในส่วนนั้นจะได้แคบกว่าความเป็นจริง จึงทำให้ความแม่นยำลดลง

นอกจากนี้การทำ Normalization กับข้อมูลองศาสองมิติได้ให้ผลลัพธ์การประมาณค่าท่วงท่าที่แม่นยำสูงขึ้นของทุกท่วงท่าจนความแม่นยำเกินร้อยละ 90 ดังตารางที่ 1 อีกทั้งลำดับความแม่นยำยังคงเป็นลำดับเดียวกับข้อมูลองศาปกติจึงแสดงให้เห็นถึงความสัมพันธ์ของข้อมูลที่ไม่เปลี่ยนแปลง

5. อภิปรายและสรุปผลการวิจัย

งานวิจัยนี้เน้นการทดลองความแม่นยำของเทคโนโลยีประมาณค่าท่วงท่าเมื่อเปรียบเทียบด้วยเทคโนโลยีจับการเคลื่อนไหวของร่างกายด้วยท่วงท่าที่ออกแบบมาจาก FMA-UE เพื่อตรวจสอบความเป็นไปได้ในการใช้เทคโนโลยีประมาณค่าท่วงท่าและ MediaPipe สร้างสื่อเชิงตอบโต้ แอปพลิเคชันหรือเกมออกกำลังกายเพื่อใช้ในการฟื้นฟูสมรรถภาพ จูงใจให้ผู้ป่วยออกกำลังกายมากขึ้น การประเมินสมรรถนะ และเปิดโอกาสให้ผู้ป่วยโรคหลอดเลือดสมองได้เข้ารับการรักษาและฟื้นฟูสมรรถภาพทางไกลจากที่บ้าน

จากผลการทดลองสรุปได้ว่าเทคโนโลยีการประมาณค่าท่วงท่าโดยโปรแกรม MediaPipe การประมาณค่าและคำนวณองศาโดยใช้พิกัดแนวลึกหรือ Z-Axis มีความแปรปรวนมากเกินไป ไม่เหมาะกับการนำมาใช้สำหรับการประเมินผู้ป่วย และไม่เหมาะกับการใช้เป็นอินพุตเพื่อออกคำสั่งภายในเกมออกกำลังกาย เพราะหากส่วนของร่างกายใด ๆ โดนบดบังแม้เพียงเล็กน้อยก็อาจจะทำให้การประมาณค่าในแนวแกนลึกผิดพลาดได้ ส่วนการคำนวณแบบสองมิติ ร่างกายที่ขนานกับกล้องจะได้ผลลัพธ์ที่แม่นยำที่สุด แต่โดยรวมแล้วยังคงมีร้อยละความแม่นยำที่ต่ำ เนื่องจากช่วงของการเคลื่อนไหวที่จับได้กว้างหรือแคบเกินจริง อย่างไรก็ตาม การใช้ Normalization เพื่อให้ข้อมูลองศาเป็นมาตรฐานเดียวกันกับ MoCap พบว่ามีความแม่นยำที่เพิ่มสูงขึ้นเป็นอย่างมาก เพราะ Normalization ช่วยปรับมาตรฐานช่วงของอศาให้เป็นมาตรฐานที่ถูกต้อง และทำให้สามารถใช้ MediaPipe เพื่อประเมินหรือช่วยแพทย์ประเมินผู้ป่วยในขณะทำกายภาพบำบัดด้วยท่วงท่าที่ออกแบบมาจาก FMA-UE

การคำนวณองศาแบบสองมิติผนวกกับการ Normalization จึงเหมาะสมเป็นอย่างมากในการพัฒนาแอปพลิเคชันหรือเกมออกกำลังกาย เพื่อการฟื้นฟูสมรรถภาพจากที่บ้านหรือการเข้ารับการรักษาทางไกล และลดจำนวนผู้ป่วยที่ขาดการเข้ารับการรักษา แต่วิธีการประเมินผู้ป่วยเบื้องต้นในระหว่างใช้งานเกมออกกำลังกายผ่านเว็บไซต์ จำเป็นจะต้องให้ข้อมูลระหว่างการทำกายภาพบำบัดในรูปแบบของวิดีโอ และแสดงกราฟองศาแต่ละส่วนของร่างกายเป็นข้อมูลประกอบ เพื่อให้แพทย์หรือผู้ดูแลทำการวินิจฉัยได้ด้วยตนเองต่อไป

6. ข้อเสนอแนะ

การทดลองของงานวิจัยนี้คือการหาร้อยละความแม่นยำของ MediaPipe เมื่อเปรียบเทียบกับเทคโนโลยีตรวจจับการเคลื่อนไหวของร่างกาย การทดลองปัจจุบันผู้วิจัยใช้ Moving Average ซึ่งอาจจะมีประสิทธิภาพไม่พอ ถ้าหากต้องการเพิ่มแนวโน้มที่จะนำเทคโนโลยีดังกล่าวเพื่อการประเมินผู้ป่วยที่กระทำท่วงท่าที่มีแนวแกนลึก อาจจะต้องมีการเพิ่มเทคนิคการคาดคะเนหรือการกรองข้อมูล (Filter) เพื่อเพิ่มความแม่นยำของการประมาณค่าท่วงท่า และการออกแบบท่วงท่าที่ใช้ในการประเมินหรือการออกแบบการประเมินผู้ป่วยอาจจะต้องเป็นท่วงท่าที่เน้นการเคลื่อนไหวระนาบในแนวตั้งและไม่มีการเคลื่อนไหวในแนวลึกเกี่ยวข้อง หรือใช้เทคนิคอื่นอย่างการตรวจจับสิ่งของร่วมเพื่อเพิ่มความแม่นยำ

เอกสารอ้างอิง

[1] สมศักดิ์ เทียมเก่า, “อุบัติการณ์โรคหลอดเลือดสมองประเทศไทย,” *วารสารประสาทวิทยาแห่งประเทศไทย,* vol. 37, p. 54–60, 2021.

[2] กองโรคไม่ติดต่อหรือสำนักสื่อสารความเสี่ยงฯ กรมควบคุมโรค, “วันอัมพาตโลก 2565 เน้นสร้างความตระหนักรู้เกี่ยวกับสัญญาณเตือนโรคหลอดเลือดสมองให้กับประชาชน,” [ออนไลน์], 2565. เข้าถึงได้: https://pr.moph.go.th/?url=pr/detail/all/02/180623. [เข้าถึงเมื่อ 27 เมษายน 2566].

[3] ตรีนุช อมรภิญโญเกียรติ, “บทบาทของการใช้หุ่นยนต์ฟื้นฟูผู้ป่วยโรคหลอดเลือดสมองในโรงพยาบาล ตากสิน,” *วารสารสาธารณสุขและวิทยาศาสตร์สุขภาพ*, vol. 5, p. 160–171, 2022.

[4] ปิยธิดา ชูรักษ์, “กรอบแนวคิดการกายภาพบำบัดผู้ป่วยโรคหลอดเลือดสมองในชุมชน,” *Thaksin University Online Journal*, vol. 1, p. 65, 2022.

[5] J. Adolf *et al*., "Automatic telerehabilitation system in a home environment using computer vision.," in *pHealth*, pp. 142–148, 2020.

[6] G. Bradski and A. Kaehler, *Learning OpenCV: Computer vision with the OpenCV library*. "O'Reilly Media, Inc.", 2008.

[7] N. Nakano *et al*., "Evaluation of 3 D markerless motion capture accuracy using OpenPose with multiple video cameras," *Frontiers in sports and active living*, vol. 2, art. 50, 2020.

[8] M. Nour, M. Gardoni, J. Renaud and S. Gauthier, "Real-time detection and motivation of eating activity in elderly people with dementia using pose estimation with TensorFlow and OpenCV," *Adv. Soc. Sci. Res. J*, vol. 8, p. 28–34, 2021.

[9] D. M. Kishore, S. Bindu and N. K. Manjunath, "Estimation of yoga postures using machine learning techniques," *International Journal of Yoga*, vol. 15, no. 2, p. 137 - 143, 2022.

[10] C. El-Habr, X. Garcia, P. Paliyawan and R. Thawonmas, "Runner: A 2 D platform game for physical health promotion," *SoftwareX*, vol. 10, p. 100329, 2019.

[11] J. W. Burke *et al*., "Vision based games for upper-limb stroke rehabilitation," in *2 0 0 8 International Machine Vision and Image Processing Conference*, pp. 159–164, IEEE, 2008.

[12] L. E. Sucar, R. Luis, R. Leder, J. Hernández and I. Sánchez, "Gesture therapy: A vision-based system for upper extremity stroke rehabilitation," *in 2 0 1 0 Annual International Conference of the IEEE Engineering in Medicine and Biology*, 2010.

[13] C. Lugaresi *et al*., "Mediapipe: A framework for building perception pipelines," *arXiv preprint arXiv:1906.08172*, 2019.

[14] A. Latreche *et al*., "Reliability and validity analysis of MediaPipe-based measurement system for some human rehabilitation motions," *Measurement*, vol. 2 1 4 , p. 112826, 2023.

[15] R. Sers *et al*., "Validity of the Perception Neuron inertial motion capture system for upper body motion analysis," *Measurement*, vol. 149, p. 107024, 2020.

[16] P. Palani, S. Panigrahi, S. A. Jammi and A. Thondiyath, "Real-time Joint Angle Estimation using Mediapipe Framework and Inertial Sensors," in *2 0 2 2 IEEE 2 2 nd International Conference on Bioinformatics and Bioengineering (BIBE)*, 2022.

[17] A. Fern'ndez-Baena, A. Susin and X. Lligadas, "Biomechanical validation of upper-body and lower-body joint movements of kinect motion capture data for rehabilitation treatments," in *2 0 1 2 fourth international conference on intelligent networking and collaborative systems*, 2012.

[18] M. Yahya *et al*., "Motion capture sensing techniques used in human upper limb motion: A review," *Sensor Review*, vol. 39, p. 504–511, 2019.

[19] A. R. Fugl-Meyer *et al*., "A method for evaluation of physical performance," *Scand. J. Rehabil. Med*, vol. 7, p. 13–31, 1975.

[20] I. Hen *et al*., "The dynamics of spatial behavior: how can robust smoothing techniques help?," *Journal of neuroscience methods*, vol. 133, p. 161–172, 2004.

CONTRIBUTORS' FORM

(to be modified as applicable and one signed copy attached with the manuscript)

Manuscript Title:

_____

Author(s):

_____

I/we certify that I/we have participated sufficiently in the intellectual content, conception and design of this work or the analysis and interpretation of the data (when applicable), as well as the writing of the manuscript, to take public responsibility for it and have agreed to have my/our name listed as a contributor. I/we believe the manuscript represents valid work. Neither this manuscript nor one with substantially similar content under my/our authorship has been published or is being considered for publication elsewhere, except as described in the covering letter. I/we certify that all the data collected during the study is presented in this manuscript and no data from the study has been or will be published separately. I/we attest that, if requested by the editors, I/we will provide the data/information or will cooperate fully in obtaining and providing the data/information on which the manuscript is based, for examination by the editors or their assignees. Financial interests, direct or indirect, that exist or may be perceived to exist for individual contributors in connection with the content of this paper have been disclosed in the cover letter. Sources of outside support of the project are named in the cover letter.

I/We hereby transfer(s), assign(s), or otherwise convey(s) all copyright ownership, including any and all rights incidental thereto, exclusively to the Journal, in the event that such work is published by the Journal. The Journal shall own the work, including 1) copyright; 2) the right to grant permission to republish the article in whole or in part, with or without fee; 3) the right to produce preprints or reprints and translate into languages other than English for sale or free distribution; and 4) the right to republish the work in a collection of articles in any other mechanical or electronic format.

We give the rights to the corresponding author to make necessary changes as per the request of the journal, do the rest of the correspondence on our behalf and he/she will act as the guarantor for the manuscript on our behalf.

All persons who have made substantial contributions to the work reported in the manuscript, but who are not contributors, are named in the Acknowledgment and have given me/us their written permission to be named. If I/we do not include an Acknowledgment that means I/we have not received substantial contributions from non-contributors and no contributor has been omitted.

	Name	Signature	Date signed	
1	_____	_____	_____	
2	_____	_____	_____	
3	_____	_____	_____	
4	_____	_____	_____	(up to 4 contributors for case report/images/review)
5	_____	_____	_____	
6	_____	_____	_____	(up to 6 contributors for original studies)

INFORMATION FOR AUTHORS

The Journal of Information Science and Technology (JIST) aims to be the forum through which researchers, faculties, graduate students and experts of the computer and information technology and other technological related fields share and discuss their high quality research work as well as innovation. Original research articles, practical applications and innovations in the broad area of computer and information technology are suitable for publication in JIST. Periodicity (Publication) is 2 issues per year (JAN - JUN and JULY - DECEMBER).

Accepted papers will be Double-blind peer reviewed by minimum 3 reviewers with "Accept" result.

1. Soft Computing
2. Human-Computer Interaction
3 Information Assurance and Security
4. Information Systems
5. Networking
6. Programming
7. Platform Technologies
8. System Integration and Architecture
9. Social and Professional Issues
10. Web Systems and Technologies
11. Multidisciplinary
12. e-Business and m-Business
13. Business and Information System
14. Social and Business Aspects of Convergence IT and Ubiquitous Computing
15 Internet of Things (IoT)
16 Cloud Computing
17 Fintech and blockchain

Manuscript Preparation Guide

The template for writing the journal is available for downloading (https://ph02.tci-thaijo.org/index.php/JIST/). Length of the paper should be in between 8-12 pages of A4 size. Text should be typewritten or printed with double-spaced in 11-point of A4 white paper with margins of 1.5" for top, 1.25" for left, and 1" for bottom and right sides. All pages must be numbered sequentially. Here are some guidelines.

- Abstract which length 100-200 words.
- Introduction which talks about the problem and related research.
- Materials and methods use for experimentation.
- Results of the experiment.
- Discussion and Conclusion, References

Submissions (Publication Charge: Free) (Indexed by TCI tier 2)

Please submit paper in MS Word via https://ph02.tci-thaijo.org/index.php/JIST/
(Online Journal Submission)

Contact Address

The Association of Council of IT Deans (CITT)
140 Moo 1, Cheumsampan Road, Nongchok
Bangkok, Thailand 10530
Tel: 02-988-3655 ext 4115 Fax: 02-988-4027
E-mail: jist@citt.or.th

สคทส. CITT
สมาคมสภาคณบดีคณะเทคโนโลยีสารสนเทศ
The Association of Council of Information Technology Deans of Thailand