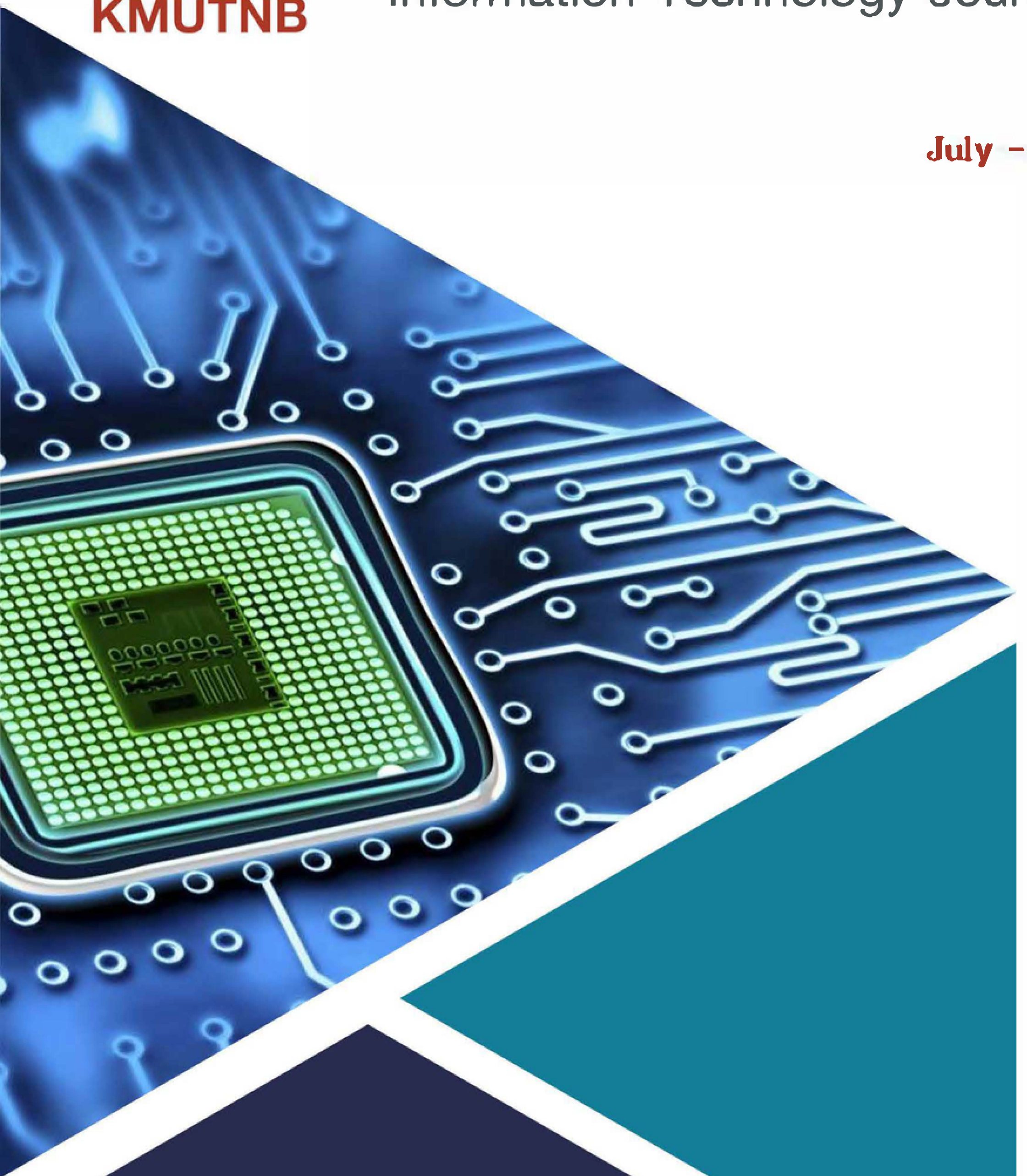
วารสารเทคโนโลยีสารสนเทศ มจพ.

Information Technology Journal KMUTNB

KMUTNB

ITJ

Vol. 20 No. 2
July - December 2024



ISSN 1685-8573

TCI Group 1

วารสารเทคโนโลยีสารสนเทศ มจพ.

Information Technology Journal KMUTNB

(IT Journal KMUTNB)

ปีที่ 20 ฉบับที่ 2 เดือนกรกฎาคม - ธันวาคม 2567

Vol. 20, No. 2, July – December 2024

ISSN 1685-8573

รศ.ดร.อรุณี อินทรไพโรจน์	มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี
รศ.ดร.สุพจน์ นิตย์สุวัฒน์	มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
รศ.ดร.สุมิตรา นวลมีศรี	มหาวิทยาลัยราชภัฏสวนสุนันทา
รศ.ดร.พณณา ตั้งวรรณวิทย์	มหาวิทยาลัยราชภัฏเพชรบูรณ์
ผศ.ดร.อุไรวรรณ อินทร์แหยม	มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี
ผศ.ดร.เมธิญาณินฆ์ คำขาว	มหาวิทยาลัยเทคโนโลยีราชมงคลพระนคร
อ.ดร.อลิสา คงทน	มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี
ดร.ชูชาติ หฤไชยศักดิ์	ศูนย์อิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ (NECTEC)

ฝ่ายประสานงานและฝ่ายจัดการ

ผศ.ดร.วัชรีวรรณ จิตต์สกุล	มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
นางสาวพรพิมล ฝ้ายเทศ	มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

กำหนดออก: ออกเผยแพร่ปีละ 2 ฉบับ

ฉบับที่ 1 เดือนมกราคม - มิถุนายน (บทความภาษาไทย)

ฉบับที่ 2 เดือนกรกฎาคม - ธันวาคม (บทความภาษาอังกฤษ)

กำหนดการรับและพิจารณาบทความ

รับพิจารณาบทความอย่างต่อเนื่อง

วัตถุประสงค์

1. เพื่อเผยแพร่ผลงานวิจัย/พัฒนาและผลการวิชาการด้านเทคโนโลยีสารสนเทศและวิทยาการคอมพิวเตอร์ ของคณาจารย์ เจ้าหน้าที่ และนักศึกษาของมหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ และสถาบันระดับอุดมศึกษาอื่น ๆ รวมทั้งนักวิจัย/พัฒนาจากหน่วยงานต่าง ๆ ทั้งภาครัฐและเอกชนภายในประเทศ
2. เพื่อเป็นศูนย์รวมแลกเปลี่ยนและกระจายองค์ความรู้ในนวัตกรรมใหม่ๆเกี่ยวกับเทคโนโลยีสารสนเทศ และวิทยาการคอมพิวเตอร์
3. เพื่อเป็นศูนย์รวมข่ายงานให้นักเทคโนโลยีสารสนเทศและวิทยาการคอมพิวเตอร์ได้เสนอแนวความคิดผลงานวิจัย/พัฒนาต่าง ๆ อันจะเป็นประโยชน์ต่อสถาบันและประเทศชาติ

เว็บไซต์

https://ph01.tci-thaijo.org/index.php/IT_Journal

อีเมล

itjournal@it.kmutnb.ac.th

เจ้าของ

คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล

มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

1518 ถนนประชาราษฎร์ 1 แขวงวงศ์สว่าง เขตบางซื่อ กรุงเทพฯ 10800

โทรศัพท์ 02-555-2726, 02-555-2704

บรรณาธิการแถลง

วารสารเทคโนโลยีสารสนเทศ มจพ. เป็นวารสารที่เผยแพร่ผลงานทางวิชาการ โดยรับบทความวิจัยและบทความวิชาการทั้งภาษาไทยและภาษาอังกฤษ เกี่ยวกับเทคโนโลยีสารสนเทศ ซึ่งบทความจะต้องได้รับการประเมินโดยผู้ทรงคุณวุฒิภายในหรือภายนอกมหาวิทยาลัย จำนวน 3 ท่าน โดยกระบวนการพิจารณานี้ผู้ทรงคุณวุฒิไม่ทราบข้อมูลของผู้ส่งบทความ (Double Blinded) และบทความจะต้องมีความคิดริเริ่มสร้างสรรค์ คุณค่าทางวิชาการ ความสมบูรณ์ของเนื้อหา และโครงสร้าง ภาษาที่ใช้ ความชัดเจนของสมมติฐาน/วัตถุประสงค์ ความชัดเจนของการนำเสนอและการจัดระเบียบบทความ ความถูกต้องทางวิชาการ การอภิปรายผลและการอ้างอิงที่ถูกต้องตามหลักวิชาการ ในนามของกองบรรณาธิการวารสารเทคโนโลยีสารสนเทศ มจพ. ขอขอบคุณทุกท่านที่ให้ความสนใจส่งบทความเพื่อพิจารณาในวารสารเทคโนโลยีสารสนเทศ มจพ. ปีที่ 20 ฉบับที่ 2 เดือนกรกฎาคม - ธันวาคม 2567 ในฉบับนี้กองบรรณาธิการได้คัดสรรบทความที่น่าสนใจมาเพื่อนำเสนอให้ผู้ที่สนใจเกี่ยวกับเทคโนโลยีสารสนเทศทุกท่าน

(ผู้ช่วยศาสตราจารย์ ดร.ศักดิ์ชาย ตั้งวรรณวิทย์)
บรรณาธิการวารสารเทคโนโลยีสารสนเทศ มจพ.

สารบัญ	หน้า

Applying Fuzzy Multi-Criteria Decision Making for the Selection of Competitive Factors in Robotics Based on System Development Life Cycle

Orasa Sirasakamol* Punnita Taecharoen* Narin Jiwitan** Kitti Puritat*** and Kannikar Intawong****

Received: August 27, 2023
Revised: January 22, 2024
Accepted: March 7, 2024

* Corresponding Author: Punnita Taecharoen, E-mail: punnita@rmutl.ac.th

DOI: 10.14416/j.it/2024.v2.001

Abstract

Robots are a technology that plays an important role in many industries today. Competition in the robot market is high, so choosing the right competitive factors for robots is important. This study uses a Fuzzy Analytic Hierarchy Process (F-AHP) method to select appropriate competitive factors for the robot. It considers competitive factors in robots and system development life cycles (SDLC). The research results found that important competitive factors for robots include: All 5 experts focus on (C1) interaction factors giving robots flexibility and adaptability to changing environmental conditions while (SC1–5) image processing gives robots the ability to sense and respond to the environment while competing effectively. Combining these two factors improves competitive performance and allows the robot to maximize its time in changing situations. Among the alternative scenarios compared with factors at levels 1 and 2, option A1 (Robotic 1) performed the best in terms of achieving optimal time, as indicated by its calculated weight (Ni) = 0.508 and All 5 students determined the weight values according to Table 1, focusing on factors (C2) Posture factors affect a robot's performance in competition. Maintaining the correct posture helps the robot be agile and efficient in carrying out its tasks and factors (SC2-6) in level 2 Climbing steep slopes gives robots the ability to burn off energy and maintain stability in difficult conditions. Efficient operation in steep conditions is an important competitive factor. Among the alternative scenarios compared with factors at level 1 and level 2, option A1 (Robotic 1) performed the best in terms of achieving optimal time, as indicated by its calculated weight (Ni) = 0.573. This study shows that the F-AHP is an effective tool for selecting appropriate competitive factors for robots.

Keywords: Robots, Competitive factors, AHP, Fuzzy, SDLC.

1. Introduction

In the realm of robotic competitions, the context extends beyond the physical arena where robots compete. It encompasses the broader objectives of fostering technological advancement, encouraging collaboration among participants and inspiring the next generation of engineers and innovators. These competitions provide a stage for participants to showcase their robots' agility adaptability, and efficiency in addressing specific tasks or challenges [1]. The educational landscape must recognize that robots are not merely mechanical entities but sophisticated systems equipped with artificial intelligence sensors and actuators. They are designed not only to perform tasks efficiently but also to adapt to dynamic and unforeseen challenges. Whether in the classroom research lab or industry setting, comprehending the context of robotics allows educators and students alike to harness the full potential of these technological marvels.

* Department Electrical, Faculty of Engineering, Rajamangala University of Technology Lanna.
** Department of Business Information Systems, Faculty of Business Administration and Liberal Arts, Rajamangala University of Technology Lanna.
*** Department of Library and Information Science, Faculty of Humanities, Chiang Mai University.
**** School of Public Health, Faculty of Public Health, Chiang Mai University.

This article aims to study the specific factors of robots used in competitions in the following aspects: Interaction Mobility Navigation Manipulation and Intelligence to determine the most significant factors for competition consideration, the selection process involves analyzing the quantities with the highest importance. The selection process of prototype factors uses the F-AHP technique to analyze the hierarchy of matrix equations.

2. Theories and Related Research

Decision-making involves the utilization of criteria or various tools to assist in the process of making choices in order to minimize the chances of errors or enhance the accuracy of decisions. Probability and conditional reasoning play a significant role in decision-making. This is because individuals' rationale in pursuing their objectives often necessitates selecting the best possible outcomes or rewards.

2.1 System Development Life Cycle (SDLC)

Software Development Life Cycle (SDLC) is a process that emphasizes software development through various stages. Generally, SDLC comprises the following steps:

1) Requirement Analysis: This step involves analyzing and gathering all the necessary requirements for the software including understanding the problem and user needs.

2) Design: Taking the requirements obtained from the previous step the team plans and designs the structure or architecture of the software.

3) Implementation: The team begins to create the software model based on the designed plan.

4) Testing: The developed model is tested to ensure it functions correctly and any errors are identified and fixed.

5) Deployment and Maintenance: The model that has been tested is deployed for real-world use and the team is responsible for maintaining and updating the software to keep it current and responsive to user needs

SDLC is a process that helps the team with a framework for developing a model and enables efficient tracking and control of the progress of development [2].

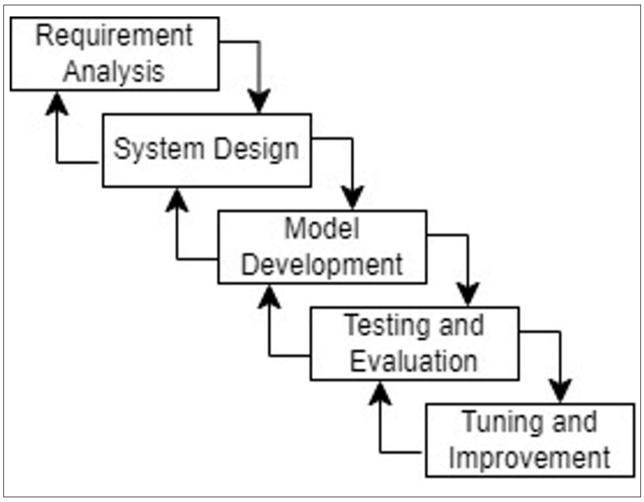


Figure 1. *System Development Life Cycle (SDLC).*

All of these process steps are methods that can be used to enhance decision-making efficiency regarding the selection of competition factors in robotics. This is an important aspect of research related to robotics.

2.2 Multi Criteria Decision Making (MCDM)

Multi Criteria Decision Making (MCDM) is a popular method used for decision-making analysis to select alternatives that best align with or are most suitable according to multiple criteria. The significant characteristic of MCDM is the involvement of stakeholders in the decision-making process and the weighting of all options for accurate results.

AHP (Analytic Hierarchy Process) and F-AHP (Fuzzy Analytic Hierarchy Process) and Fuzzy Logic are widely used in various sectors such as education industry and engineering to solve decision-making problems. Multi-criteria decision-making is an approach that aids in selecting the most appropriate criteria for predefined alternatives [3].

The traditional approach of MCDM is insufficient to handle linguistic uncertainties [4]. Therefore, it is recommended to use Fuzzy-based MCDM to deal with ambiguity in the decision-making process. Additionally, the Fuzzy approach leads to more plausible outcomes.

2.3 Fuzzy Analytic Hierarchy Process (F-AHP)

The Fuzzy Analytic Hierarchy Process (F-AHP) is a decision-making analysis method used for selecting or prioritizing alternatives in complex decision-making scenarios. It involves transforming real numbers into fuzzy numbers and structuring the decision-making process into a hierarchical framework using matrix equations [5]. The F-AHP process consists of five steps:

1) Convert real numbers into fuzzy numbers.

2) Populate the matrix with fuzzy numbers to pairwise compare alternatives.

3) Calculate the weights for each criterion using the theory of the Analytic Hierarchy Process.

4) Calculate the weights for each alternative by combining steps 1 to 3.

5) Convert or defuzzify the fuzzy numbers to real numbers and determine the values of the alternatives by multiplying criterion weights with alternative weights.

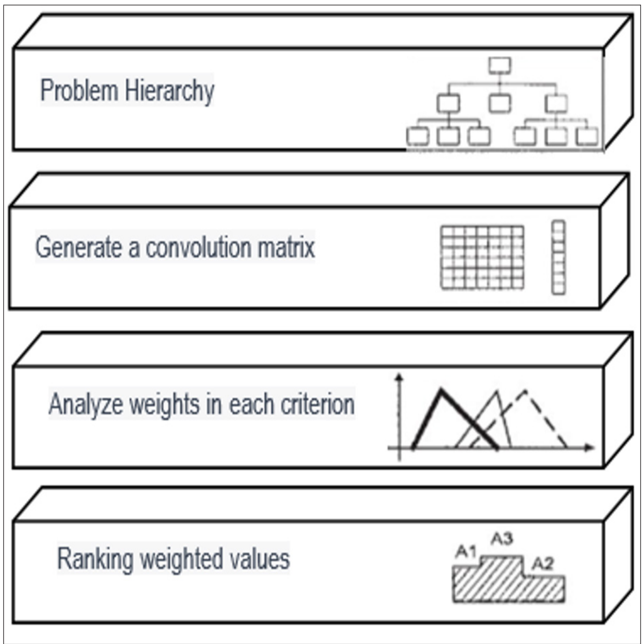


Figure 2. *Hierarchical Cluster Analysis Process.*

2.4 Fuzzy Set and Fuzzy Number

Fuzzy Set and Fuzzy Number: Zadeh (1965) [6] introduced the Fuzzy Set Theory to manage uncertainty arising from imprecision and vagueness. Mathematical operations and programming can be applied to the domain of Fuzzy Set which is a class of objects with a continuous degree of membership. Such sets are characterized by a membership function that assigns a membership level ranging from 0 to 1 to each object (Kahraman et al., 2019) [7]. The tilde symbol (~) is placed over a symbol if it denotes a Fuzzy set. Fuzzy numbers represented as triangular fuzzy membership functions are denoted as l, m, and u, signifying the minimum value, the value with the highest trend, and the maximum possible value respectively. This explains the concept of Fuzzy membership function of triangular fuzzy numbers as depicted in Figure 3.

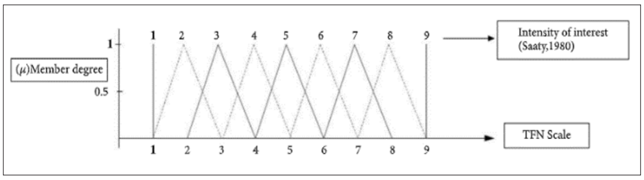


Figure 3. *Membership Function of Triangulars.*

$$u\tilde{A}(\mathrm{X}) = \tilde{M} = (l, m, u) \qquad (1)$$

E. Ertugrul Karsak [8] conducted a study on the problem of robot selection using a multi-criteria decision-making approach based on Choquet Integral. The main characteristic of Choquet Integral is its ability to aggregate interactions between various features for decision analysis. This problem is often overlooked in previous studies on robot selection. The proposed decision framework utilizes data extracted beforehand from experts in analyzing a reference set of robot alternatives using the same selection criteria as the current assessment. It employs Choquet Integral to determine the most suitable robot.

Athale and Chakraborty (2016) [9] conducted a study comparing the effectiveness of using Multi-Criteria Decision Making (MCDM) to rank industrial robot selection problems. They concluded that for the specified robot selection problem it is advisable to place greater importance on selecting appropriate criteria.

Changwon Kim and Yeesock Kim (2020) [10] introduced a multi-robot navigation strategy using a multi-objective decision-making algorithm namely Fuzzy Analytic Hierarchy

Process (F-AHP). The objective was to analyze and select the most suitable positions as sub-goals from various points on the mobile robot's detection boundary. This analysis was carried out considering the following three objectives: distance to the target destination collision safety and robot orientation adjustment to face the target. Alternative solutions were evaluated by determining the relative importance weights of the objectives as the F-AHP algorithm alone was insufficient for multi-robot navigation. Game theory with participation was incorporated to enhance the algorithm's efficiency. The proposed multi-robot navigation algorithm was tested on up to 12 mobile robots in various simulated scenarios with changing factors.

Bipradas Bairagi and Balaram Dey (2016) [11] utilized the Fuzzy Analytic Hierarchy Process (F-AHP) method to evaluate and select robots for automated foundry operations. In this approach AHP is integrated with fuzzy techniques to define the characteristics of the purchase order based on similarity. In each case (F-AHP) is used to assess the covering weights of the selection criteria under consideration to evaluate and select robots. The real-life problem of selecting robots for practical foundry operations was referred to in order to demonstrate and verify the suitability and potential of the utilized method. An analysis and comparison of the research results showed that the F-AHP-based approach serves as a valuable and efficient representation suitable for selecting the best robots in various environmental conditions.

3. Research Methodology

The study employed the F-AHP model to investigate the decision-making process for selecting factors related to competitive robots. The analytical hierarchy process depended on the system development process. The analysis of results was carried out based on the following principles of operation.

3.1 Research Framework

The research process follows a structured framework consisting of sequential steps.

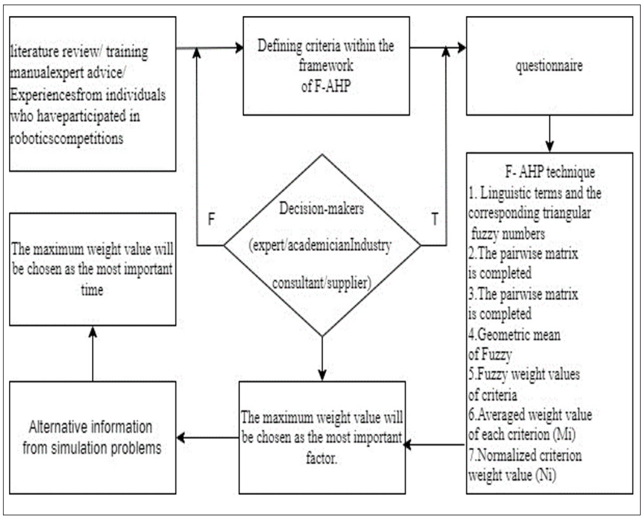


Figure 4. *Research Framework.*

3.1.1 Study of information related to robot factors to effectively utilize information for the evaluation and preparation of robotics competitions both domestically and internationally. The primary goal of these competitions is to provide the education sector with insights into the contextual characteristics and behavior of participating robots. Additionally, the competitions aim to foster the development and application of skills and knowledge to enhance the educational experience. A comprehensive consideration of all factors is essential to achieving these objectives.

1) Five experts were selected for this research based on the following criteria.

1.1 Possesses a deep understanding of robot design and operation in competition settings for a minimum of 5 years.

1.2 Demonstrates experience in related fields or competitions, including mechanical, automation, programming and robotics aspects. For instance, involvement in guiding students to compete at the World Skills level with a minimum of 10 instances.

1.3 Holds academic expertise in robotics and related technology with a background of 5 years or more in a professional setting.

1.4 Proficient in decision analysis using F-AHP demonstrating versatility with various decision-making theories such as AHP, F-AHP, ANP, PROMETHEE, TOPSIS. Capable of applying decision making theories in multiple formats.

2) Five students are interested in participating in the competition and studying in the Robotics and Automation program.

Table 1. *Data of Factors Related to Robotics Competition [12, 13].*

Criteria	Sub-Criterial
1) Interacting factors (C1)	Remembering faces (SC1-1) Posture (SC1-2) Robot vision (SC1-3) Image system (SC1-4) Process image (SC1-5) Voice recognition (SC1-6) Interacting with users (SC1-7) Speech recognition (SC1-8)
2) Posture factors (C2)	Getting up (SC2-1) Walking (SC2-2) Standing (SC2-3) Running (SC2-4) Jumping (SC2-5) Climbing steep slopes (SC2-6) Navigating obstacles (SC2-7) Using wings for flying (SC2-8) Using fins for swimming (SC2-9)
3) Accuracy factors (C3)	Characteristic (SC3-1) Range (SC3-2) Sensor (SC3-3) Disturbances (SC3-4)
4) Object movement factors (C4)	Picking up (SC4-1) Extending arms and legs (SC4-2) Bending arms and legs (SC4-3)
5) Intelligence (C5)	Brain simulation (SC5-1) Perception simulation (SC5-2) Logical method (SC5-3) Calculation method (SC5-4) Statistical method (SC5-5) integration method (SC5-6)
6) Drive system configuration factors (C6)	Wheel-based mobility (SC6-1) Belt-driven mobility (SC6-2) Leg-based mobility (SC6-3) Flight-based mobility (SC6-4) Aquatic mobility (SC6-5)

3.1.2 The most crucial aspect of robotics competitions is the ability to effectively address the challenges presented in each round. Inappropriate choices during competition can result in wasted time for each round. Each round of the robotics competition introduces different problem statements, requiring participants to be familiar with the identity and components of their robots. Therefore, experts or those with prior competition experience carefully consider the primary and secondary factors influencing the robot's performance in each round. They assess which factors contribute most to bringing the robot closer to the target quickly as dictated by the specific problem statement.

The top level 0, signifies the ultimate goal of this research which is to enable robots to reach the target level in the shortest possible time. While at level 1, the factor for selecting factors from [14] on page 43 necessitates participants to thoroughly study this information. It comprises 6 factors ranked in order of importance to be utilized in decision-making for robotics competition. These criteria are C1) Interactivity C2) Mobility C3) Accuracy C4) Object Handling C5) Intelligence and C6) Drive System Design (only one factor is exemplified for analysis in the paper). At level 2, enhances the primary factors and provides additional explanation and confidence-building aspects for decision-making. This is subdivided into 27 factors as depicted in Table 1.

Analysis at level 2 will not be presented on paper as the analytical approach is similar to that at level 1 and the paper may not provide sufficient space. Finally at level 3, alternative factors have simulated various competition scenarios [15] to decide on the optimal factors for consideration. This aids in achieving the fastest competition times in the next round.

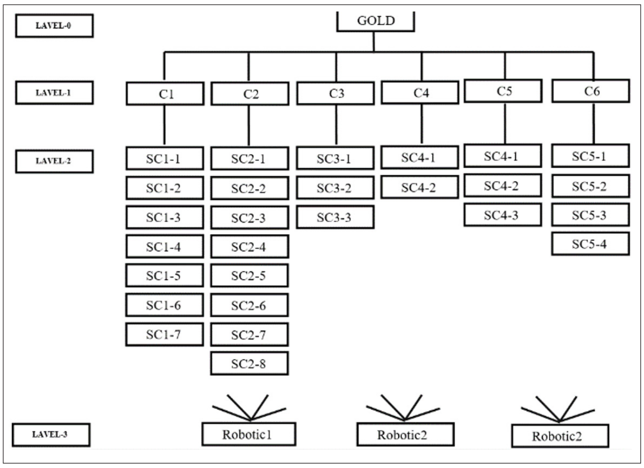


Figure 5. *Problem Decomposition into Layers of F-AHP Model for Factors of Competitive Robotics.*

In analyzing the researcher has analyzed the main factors and alternatives to show the comparisons in F-AHP analysis the researcher has shown the steps and methods of analysis only for the 5 experts. Therefore, all the data of every factor and expert is kept in Microsoft Excel.

3.2 Linguistic Variables and Corresponding Fuzzy Triangles.

After receiving all criteria and options the researcher therefore a questionnaire according to the criteria.

Step 1 involves decision-makers using linguistic terms to compare criteria or alternatives. This is done by converting the values from qualitative assessments into numerical form through linguistic variables. The average score of assessors is determined based on the principles of the triangular membership function which ranges from 1 to 9 [16], [17].

The real number values (l, m, u) constituting the triangular number are "l", the smallest probable value, "m", the most probable number, and "u", the largest probable value. The membership function of a triangular Fuzzy number is defined as follows.

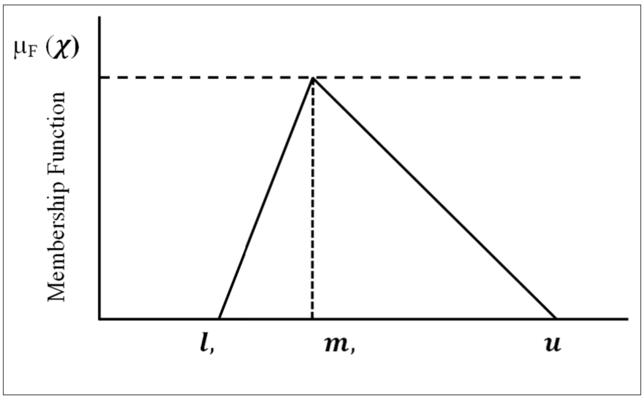


Figure 6. *Membership function of a triangular number* $\mu\tilde{A}(\chi) = \widetilde{M} = (\boldsymbol{l}, \boldsymbol{m}, \boldsymbol{u})$

Set up the Triangular Fuzzy Numbers (TFN's). Each expert makes a pair-wise comparison of the decision criteria and gives them relative scores. In this method the Fuzzy conversion scale is as in Table 2 this scale has been employed in Fuzzy prioritization approach [18].

Table 2. *Linguistic terminology 1-9 and Corresponding Numerical Scale through Fuzzy Triangles.*

F-AHP Scales	Linguistic Trust	Triangular Fuzzy Nun's (l, m, u) Scales	Reciprocal TEN's (1/u, 1/m, 1/l) Scales
$\tilde{1}$	Equal Important (Eq. Imp.)	(1,1,1)	(1,1,1)
$\tilde{3}$	Weak Important (W. Imp.)	(2,3,4)	(1/4,1/3,1/2)
$\tilde{5}$	Fairly Important (F. Imp.)	(4,5,6)	(1/6,1/5,1/4)
$\tilde{7}$	Strong Important (S. Imp.)	(6,7,8)	(1/8,1/7,1/6)
$\tilde{9}$	Absolute Important (A. Imp.)	(9,9,9)	(1/9,1/9,1/9)
$\tilde{2}$ $\tilde{4}$ $\tilde{6}$ $\tilde{8}$	The intermittent values between two adjacent scales	(1,2,3) (3,4,5) (5,6,7) (7,8,9)	(1/3,1/2,1/1) (1/5,1/4,1/3) (1/7,1/6,1/5) (1/9,1/8,1/7)

3.2.1 Linguistic Terms in Data Creation and Analysis to create pairwise comparison matrices for the factors as outlined in Table 2 can be identified as follows:

1) If experts assign equal importance to criterion C1 as C1, itself the value becomes "Eq. Imp" The resulting pairwise comparison is 1, 1, 1.

2) If experts assign equal importance to criterion C1 as C2, the value becomes "W. Imp" The resulting pairwise comparison is 2, 3, 4.

3) If experts assign less importance to criterion C1 compared to C3, the value becomes 1/W. Imp The resulting pairwise comparison is 1/4, 1/3, 1/2.

4) If experts assign less importance to criterion C1 compared to C4, the value becomes 1/W. Imp The resulting pairwise comparison is 1/4, 1/3, 1/2.

5) If experts assign less importance to criterion C1 compared to C5, the value becomes "W. Imp" The resulting pairwise comparison is 1/4, 1/3, 1/2.

6) If criterion C1 has the same importance as C6 criterion, the value becomes " A. Imp" " The resulting pairwise comparison is 1/9, 1/9, 1/9.

Table 3. *Pairwise Comparison Matrix using Buckley's (1985) method with Geometric Mean used for weight calculation.*

Matrix analogy display						
CRI	**C1**	**C2**	**C3**	**C4**	**C5**	**C6**
C1	Eq. Imp	1/F.Imp	1/F.Imp	1/F.Imp	1/ W.Imp	Eq. Imp
C2	F. Imp	Eq. Imp	1/F.Imp	1/ W.Imp	Eq. Imp	1/ W.Imp
C3	F. Imp	F. Imp	Eq. Imp	W. Imp	1/ W.Imp	W. Imp
C4	F. Imp	W. Imp	1/ W.Imp	Eq. Imp	1/ W.Imp	W. Imp
C5	W. Imp	Eq. Imp	1/ W.Imp	W. Imp	Eq. Imp	1/ W.Imp
C6	Eq. Imp	W.Imp	1/ W.Imp	1/ W.Imp	1/ W.Imp	Eq. Imp

$$\tilde{A}(l,m,u)^{-1}\left(\frac{1}{u},\frac{1}{m},\frac{1}{l}\right) \qquad (2)$$

This process of generating pairwise comparison matrices is applied to all criteria and levels. The results from the comparison process are then used for further analysis.

$$\tilde{a}_{ij} = \frac{\sum_{k=1}^{k} \tilde{d}_{ij}^{k}}{k} \qquad (3)$$

If decision makers are more than one each decision maker's preference will be averaged and the average will be calculated according to equation 3.

Table 4. *For consistent linguistic translation using a pairwise comparison matrix.*

Pairwise Comparison Matrix						
CRI	**C1**	**C2**	**C3**	**C4**	**C5**	**C6**
C1	(1,1,1)	(1/6,1/5, 1/4)	(1/6,1/5, 1/4)	(1/6,1/5, 1/4)	(1/4,1/3, 1/2)	(1,1,1)
C2	(4,5,6)	(1,1,1)	(1/6,1/5, 1/4)	(1/4,1/3, 1/2)	(1,1,1)	(1/4,1/3, 1/2)
C3	(4,5,6)	(4,5,6)	(1,1,1)	(2,3,4)	(1/4,1/3, 1/2)	(2,3,4)
C4	(4,5,6)	(2,3,4)	(1/4,1/3, 1/2)	(1,1,1)	(1/4,1/3, 1/2)	(2,3,4)
C5	(2,3,4)	(1,1,1)	(1/4,1/3, 1/2)	(2,3,4)	(1,1,1)	(1/4,1/3, 1/2)
C6	(1,1,1)	(2,3,4)	(1/4,1/3, 1/2)	(1/4,1/3, 1/2)	(1/4,1/3, 1/2)	(1,1,1)

3.2.2 Step 2: Due to the averaging configuration the pairwise comparison support matrix will be updated based on the transformations indicated in Table 4.

3.2.3 Step 3 involves substituting the values of the transformed pairwise comparison triangle analyses which have been converted into an absolute form from the matrix table. The geometric mean value of the comprehensive set of pairwise comparison values for the "interrelation aspect" criterion ($\tilde{R}_i$)will be calculated using the following formula.

$$\tilde{R}_i = \left(\prod_{j=1}^{n} \tilde{d}_{ij}\right)^{\frac{1}{n}} \qquad (4)$$

3.2.4 Step 4 involves calculating the covering cluster weights for each criterion which can be obtained from Equation 5. Compute the vector sum of each entity then find the exponentiation (-1) of the vector sum representing the numbers within the covering cluster triangle.

Table 5. *Geometric Mean of Fuzzy Logic.*

CRI	$\tilde{R}_i$		
C1	0.324	0.372	0.445
C2	0.589	0.693	0.849
C3	1.587	2.054	2.570
C4	1.000	1.308	1.698
C5	0.794	1.000	1.260
C6	0.561	0.693	0.891

Total = =(D62*G62*J62*M62*P62*S62) ^ (1/6)

Reverse P (-1) =D85^ (-1)

$$Increasing(INCR) = REVERSEP(-1) \quad (5)$$

To find the fuzzy weight of the criteria $\tilde{W}_i$ multiply each $\tilde{R}_i$ with this reverse vector.

$$\tilde{W}_i = \tilde{r}_i \otimes (\tilde{r}_i \oplus \tilde{r}_2 \oplus .... \oplus \tilde{r}_n)^{-1} \quad (6)$$

$$\tilde{W}_i = [(0.756 \times 0.094);(0.855 \times 0.117);(1.017 \times 0.150)]$$

$$\tilde{W}_i = [0.071;0.097;0.138]$$

Therefore, the relative fuzzy weights of each criterion are presented in Table 6.

Table 6. *Relative Fuzzy Weights of Each Criterion.*

CRI	$\tilde{W}_i$		
C1	0.042	0.061	0.092
C2	0.077	0.113	0.175
C3	0.206	0.335	0.529
C4	0.130	0.213	0.350
C5	0.103	0.163	0.260
C6	0.073	0.113	0.184

3.2.5 Step 6: The values obtained from the relative fuzzy weights of each criterion are still ambiguous triangular numbers. Therefore, the ambiguity must be resolved by the centripetal method using the equation introduced by Chou and Chang [19] using equation 7.

$$\tilde{M}_i = \frac{lw_i + mw_i + uw_i}{3} \quad (7)$$

$\tilde{M}_i$ is an unambiguous number but needs to be normalized by following Equation 8.

$$N_i = \frac{M_i}{\sum_{i=1}^{n} M_i} \quad (8)$$

Table 7. *Average Weight Criteria ($\tilde{M}_i$) and Normal Weight Criteria ($\tilde{N}_i$).*

CRI	$\tilde{M}_i$	$\tilde{N}_i$
C1	0.065	0.060
C2	0.121	0.113
C3	0.357	0.333
C4	0.231	0.163
C5	0.175	0.163
C6	0.123	0.115
Total	1.081	

3.3 Alternative Criteria

This article explores three possible scenarios, with Robot 1 [20] representing a contender designed for sheep herding competitions. In rural areas sheep farming is a prevalent practice serving both agricultural and tourism purposes. One crucial aspect that sheep farm owners must address is safeguarding their flocks from stray dogs that may infiltrate the farm and pose a threat to the sheep. Such incidents can result in substantial losses. The upcoming competition challenges participants to employ creativity and innovation in devising the robot's movement to efficiently manage the situation. Assuming the robot's role is to herd the sheep it must guide the sheep (represented by white ping pong balls) into the sheepfold while driving away the simulated wild dogs (represented by orange ping pong balls) from the sheepfold

Robot 2 [20] represents a participant in the basketball robot game. Basketball is a widely played international sport where two teams compete to score points by shooting the ball into the opposing team's basket. The team with the highest score at the end of the game is declared the winner. In the case of the basketball robot competition, it simulates the human basketball game. Each team has one robot with exceptional abilities in capturing and shooting the basketball. The competition is divided into two halves, each lasting 3 minutes. The team with the highest score at the end of the game emerges as the winner of the competition.

Robot 3 [20] serves as a representative in the search and rescue competition. The robot is tasked with assisting disaster-stricken individuals at designated locations by deploying life-sustaining bags. The transfer of life-sustaining bags requires the robot to load them once and place them in predetermined positions. The robot must navigate within the white area or the safe zone while avoiding entry into hazardous zones or areas marked with black lines. Upon successful completion of the mission the robot is required to return to the starting point. To assess various aspects of the competing robot it is crucial to prepare scenarios and challenges. These challenges serve as models that help evaluate the capabilities of the robot [21].

3.3.1 Linguistic Terms in Data Creation and Analysis

Table 8. *Transforming the Languages Referenced by Pairwise Comparison Matrices into Numerical Terms.*

Alts	A1	A2	A3
A1	(1,1,1)	(2,3,4)	(2,3,4)
A2	(4,5,6)	(1,1,1)	(1,1,1)
A3	(9,9,9)	(1,1,1)	(1,1,1)

3.3.2 The geometric means of the fuzzy comparison ($\tilde{W}_i$) and the alternative relative fuzzy weights for each criterion ($\tilde{R}_i$) are mentioned in Table 9.

Table 9. *Aalternative relative fuzzy weights ($\tilde{R}_i$).*

Alts	$\tilde{R}_i$			$\tilde{W}_i$		
A1	1.587	1.710	1.817	0.460	0.520	0.580
A2	0.550	0.585	0.630	0.160	0.178	0.201
A3	1.000	1.000	1.000	0.290	0.304	0.319
total	3.138	3.295	3.447			
Reverst	0.319	0.304	0.290			
Increase	0.290	0.304	0.319			

Table 10. *Average Weight Criteria ($\tilde{M}_i$) and Normal Weight Criteria ($\tilde{N}_i$).*

Alts	$\tilde{M}_i$	$\tilde{N}_i$
A1	0.520	0.518
A2	0.179	0.179
A3	0.304	0.303

3.3.3 Weight of Alternative Criteria

Table 11. *Displaying the Weight of Each Choice for Each Criterion.*

Criteria	Weight	A1	A2	A3
C1	0.06	0.597	0.202	0.059
C2	0.113	0.307	0.255	0.147
C3	0.333	0.113	0.327	0.235
C4	0.215	0.597	0.202	0.147
C5	0.163	0.425	0.425	0.296
C6	0.115	0.518	0.179	0.101
Total		0.373	0.300	0.313

4. Evaluation Results

Following the conclusion of the Level 3 decision analysis the criteria for analysis were applied to both experts and all students. This article provides an example from Expert 5 only. This selective focus is due to the extensive nature of explaining the entire process to everyone in the document. The final results for all participants will be consolidated and presented in the summary in 4.1 and 4.2

4.1 All 5 experts focus on (C1) interaction factors giving robots flexibility and adaptability to changing environmental conditions while (SC1–5) image processing gives robots the ability to sense and respond to the environment while competing effectively. Combining these two factors improves competitive performance and allows the robot to maximize its time in changing situations [21]. Among the alternative scenarios compared with factors at level 1 and 2, option A1 (Robotic 1) performed the best in terms of achieving optimal time, as indicated by its calculated weight (Ni) = 0.508, as shown in Table 12.

***Table 12.** Summarizing evaluations by decision-makers based on experts 5.*

Scores of Alternatives to related Criteria					
	Weights (Ni)	**Level2**	**A1**	**A2**	**A3**
C1	0.336	SC1-1 (0.116)	0.014	0.014	0.007
		SC1-2 (0.125)	0.012	0.007	0.009
		SC1-3 (0 .111)	0.01	0.005	0.007
		SC1-4 (0 .119)	0.011	0.004	0.005
		SC1-5 (0 .136)	0.017	0.100	0.015
		SC1-6 (0 .117)	0.011	0.011	0.011
		SC1-7 (0 .121)	0.011	0.007	0.009
		SC1-8 (0 .125)	0.012	0.007	0.003
C2	0.298	SC2-1 (0.132)	0.007	0.007	0.007
		SC2-2 (0.114)	0.006	0.004	0.005
		SC2-3 (0.092)	0.005	0.004	0.002
		SC2-4 (0.100)	0.006	0.004	0.005
		SC2-5 (0.099)	0.006	0.003	0.003
		SC2-6 (0.118)	0.007	0.002	0.002
		SC2-7 (0.104)	0.006	0.003	0.004
		SC2-8 (0.128)	0.007	0.002	0.002
		SC2-9 (0.112)	0.006	0.00.	0.002
C3	0.275	SC3-1 (0.119)	0.014	0.014	0.014
		SC3-2 (0.111)	0.013	0.008	0.01
		SC3-3 (0.136)	0.015	0.015	0.015
		SC3-4 (0.242)	0.013	0.01	0.006
C4	0.245	SC4-1 (0.133)	0.022	0.018	0.028
		SC4-2 (0.128)	0.019	0.012	0.03
		SC4-3 (0.109)	0.017	0.021	0.026
C5	0.296	SC5-1 (0.136)	0.015	0.015	0.03
		SC5-2 (0.128)	0.012	0.016	0.02
		SC5-3 (0.124)	0.012	0.018	0.026
C6	0.221	SC6-1 (0.123)	0.008	0.011	0.013
		SC6-2 (0.135)	0.007	0.008	0.011
		SC6-3 (0.133)	0.012	0.012	0.012
		SC6-4 (0.104)	0.008	0.01	0.012
		SC6-5 (0.105)	0.012	0.012	0.012

Level1							
Alts	**C1**	**C2**	**C3**	**C4**	**C5**	**C6**	**Total**
A1	0302	0.307	0.212	0.322	0.452	0.200	0.508
A2	0.065	0.032	0.047	0.058	0.039	0.047	0.507
A3	0.059	0.032	0.045	0.051	0.049	0.053	0.485

4.2 All 5 students determined the weight values according to Table 1, focusing on factors (C2) Posture factors affect a robot's performance in competition. Maintaining the correct posture helps the robot be agile and efficient in carrying out its tasks and factors (SC2-6) in level 2 Climbing steep slopes gives robots the ability to burn off energy and maintain stability in difficult conditions. Efficient operation in steep conditions is an important competitive factor. Among the alternative scenarios compared with factors at level 1 and level 2, option A1 (Robotic 1) performed the best in terms of achieving optimal time, as indicated by its calculated weight (Ni) = 0.573, as shown in Table 13.

***Table 13.** Summarizing evaluations by decision-makers based on students 5.*

Scores of Alternatives to related Criteria					
Level1	**Weights (Ni)**	**Level2**	**A1**	**A2**	**A3**
C1	0.299	SC1-1 (0. 133)	0.014	0.014	0.007
		SC1-2 (0. 105)	0.012	0.007	0.009
		SC1-3 (0. 125)	0.01	0.005	0.007
		SC1-4 (0. 133)	0.011	0.004	0.005
		SC1-5 (0. 127)	0.017	0.100	0.015
		SC1-6 (0. 119)	0.011	0.011	0.011
		SC1-7 (0. 129)	0.011	0.007	0.009
		SC1-8 (0. 128)	0.012	0.007	0.003
C2	0.335	SC2-1 (0.131)	0.007	0.007	0.007
		SC2-2 (0.113)	0.006	0.004	0.005
		SC2-3 (0.084)	0.005	0.004	0.002
		SC2-4 (0.091)	0.006	0.004	0.005
		SC2-5 (0.111)	0.006	0.003	0.003
		SC2-6 (0.145)	0.027	0.002	0.002
		SC2-7 (0.129)	0.006	0.003	0.004
		SC2-8 (0.106)	0.007	0.002	0.002
		SC2-9 (0.127)	0.006	0.003	0.002

Scores of Alternatives to related Criteria					
Level1	Weights (Ni)	Level2	A1	A2	A3
C3	0.235	SC3-1 (0.120)	0.014	0.014	0.014
		SC3-2 (0.129)	0.013	0.008	0.01
		SC3-3 (0.137)	0.015	0.015	0.015
		SC3-4 (0.134)	0.013	0.01	0.006
C4	0.289	SC4-1 (0.138)	0.03	0.024	0.019
		SC4-2 (0.133)	0.028	0.018	0.011
		SC4-3 (0.109)	0.026	0.017	0.021
C5	0.222	SC5-1 (0.118)	0.026	0.013	0.013
		SC5-2 (0.142)	0.017	0.01	0.013
		SC5-3 (0.144)	0.022	0.011	0.015
C6	0.256	SC6-1 (0.131)	0.011	0.007	0.009
		SC6-2 (0.142)	0.008	0.005	0.006
		SC6-3 (0.144)	0.009	0.009	0.009
		SC6-4 (0.103)	0.01	0.006	0.008
		SC6-5 (0.130)	0.011	0.011	0.011

Level1							
Alts	C1	C2	C3	C4	C5	C6	Total
A1	0.365	0.322	0.289	0.330	0.254	0.392	0.573
A2	0.210	0.257	0.330	0.354	0.336	0.245	0.466
A3	0.202	0.400	0.4323	0.220	0.256	0.303	0.468

5. Summary

This paper is to study the prototype factors of robots suitable for competition and use the Multi Criteria Decision Making (MCDM) method with the Fuzzy Analytic Hierarchic Process (F-AHP) model to select and evaluate the main 6 factors, 35 sub-factors, and 3 alternatives It also forms the basis for decision-making in robotics competitions and serves as a repository of information for individuals in academia and related commerce.

The researcher has studied related literature to collect basic information about robots and factors in robot competition and System Development Life Cycle (SDLC). From the literature study Collect factors related to robots used in competition by considering factors that affect the success of robots in competing in the market. and from general educational institutions participating in the competition. Then determine the weight of the factors based on the importance of each factor to the success of the robot.

The method used for the comparison was a face-to-face questionnaire using the question entry method. For example, which factor is more important than which factor by how much according to the level of consideration? Compare from table. Both sides enter scores to determine weights by developing a decision model.

Next is to analyze the results of the decision model to select appropriate competitive factors for the robot in the competition. Data analysis will get factors at level 1 and level 2. Then the weight of the data obtained will be analyzed against level 3, which is comparing factors with alternative problems to use in making the best time.

In addition, the researcher uses additional research methods such as collecting field data from real competitions to support the results of this research. Analysis and evaluation show that experts identify the interaction aspect as the best weight for the competitive robot factor while beginners identify the nature of the gesture. Both sides estimated the weight of Robotic 1 in the same way.

It also emphasizes the importance of accepting robots as complex systems equipped with artificial intelligence and the need to understand the context in which robots compete. Education and industry decision making in robotics involves using criteria and tools to reduce errors and increase decision accuracy, with probability and conditional reasoning playing a key role.

6. Discussion

6.1 Differences in academic experiences and knowledge can be assumed to lead to varying perspectives.

1. It is conceivable that experts may emphasize factors related to efficiency and effectiveness in competition, while students may focus on aspects influencing the design or development process.

2. Experts may possess a deep understanding of the rules and requirements of the competition whereas students may have a more limited comprehension.

6.2 The analysis with a shared viewpoint on the proposed solution for addressing the time management issues of the robot's factors at level 3 indicates a consensus. Both sides of the assessment agree that Robotic 1 is ranked as the top choice based on the highest weighted criteria. What the analysis with a shared viewpoint means is that both parties agree on the prioritization of Robotic 1 as the preferred solution.

6.3 The Fuzzy Analytic Hierarchy Process (F-AHP) method is used for ranking when comparing both sides. Serves to distribute weight, resulting in higher values during analysis. Therefore, using vague numerical values in the analysis will reduce the risk of decision making.

6.4 Therefore F-AHP employs the fuzzy set coverage theory for pairwise comparisons instead of crisp numerical values. This grants F-AHP the capability to make decisions under coverage considering the uncertainties of various factors. This process mimics human-like thinking enhancing decision-making efficiency.

7. References

[1] C. Marknual, *Research Intelligence, Robot industry and the development of the Thai manufacturing sector robotic-series1*. Available Online at https://www.krungsri.com/th/research/research-intelligence/ri-robotic-series1-landscape, accessed on 30 June 2023.

[2] D. Gupta, A. Ahlawat, and K. Sagar. "Usability Prediction & Ranking of SDLC Models Using Fuzzy Hierarchical Usability Model." *De Gruyter,* Vol. 7, No. 1, pp. 161-168, June, 2017.

[3] H. R. Weistroffer and Y. Li. "Multiple criteria decision analysis software." *Multiple Criteria Decision Analysis,* pp. 1301-1341, 2016.

[4] T. Tsoutsos, M. Drandaki, N. Frantzeskaki, E. Iosifidis and I. Kiosses. "Sustainable energy planning by using multi-criteria analysis application in the island of Crete." *Energy Policy, Elsevier,* Vol. 37, No. 5, pp. 1587-1600, May, 2009.

[5] G. Kabir and M.A.A. Hasin. "Multi-criteria inventory classification through integration of fuzzy analytic hierarchy process and artificial neural network." *International Journal of Industrial and Systems Engineering,* Vol. 14, No. 1, pp. 74-103, January, 2013.

[6] L.A. Zadeh. "Fuzzy Sets." *Information and Control,* Vol. 8, No. 3, pp. 199-249, 1965.

[7] C. Kahraman. *Knowledge-Based Systems.* Available Online at https://www.journals.elsevier.com/knowledge-based-systems/editorial-board/professor-cengiz-kahraman, accessed on 30 June 2023.

[8] E.E. Karsak. "Robot selection using an integrated approach based on quality function deployment and fuzzy regression." *International Journal of Production Research,* Vol. 46, No. 3, pp. 723-738, February, 2008.

[9] Ch. Sudip and P.S. Aithal. "A Smart IDE for Robotics Research." *International Journal of Management, Technology, and Social Sciences (IJMTS),* Vol. 7, No. 1, pp. 513- 519, June, 2022.

[10] C. Kim, Y. Kim, and H. Yi. "Fuzzy Analytic Hierarchy Process-Based Mobile Robot Path Planning." *Multidisciplinary Digital Publishing Institute,* Vol. 9, No. 2, February, 2020.

[11] B. Bairagi and B. Dey. "Selection of robotic systems in fuzzy multi criteria decision-making environment." *International Journal of Computational Systems Engineering,* Vol. 2, No. 1, pp. 32-42, 2016.

[12] A. Narkglom, *Robot industry in Thailand.* Available Online at chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://waa.inter.nstda.or.th/prs/pub/Robot-Whitepaper-Cover.pdf, accessed on 25 July 2023.

[13] P. Luenam. "Prioritized factors using fuzzy analytic hierarchy process: understanding concepts and its application." *Modern Management Journal,* Vol. 11, No. 1, pp. 1-12, January-June, 2015.

[14] P. Meesad, *LACC System and Artificial Neural Networks.* Available Online at http://202.44.34.134/teacher/FileDL/phayung36255362552.pdf, accessed on 25 October 2009.

[15] Development Bureau OBEC Building, *Simulated problems used in robot competitions.* Available Online at https://www.sillapa.net/rule61/robot-68.pdf, accessed on 30 June 2023.

[16] CODECRUCKS, *Linguistic variables and hedges in fuzzy logic.* Available Online at https://codecrucks.com/linguistic-variables-and-hedges-in-fuzzy-logic/, accessed on 30 July 2023.

[17] R.L.N. Murty, S.G. Kondamudi, M.V. Suryanarayana, and P. Giribabu. "Application of Buckley's fuzzy AHP to identify the most important factor affecting the unorganized micro-entrepreneurs borrowing decision." *International Journal of Management (IJM),* Vol. 11, No. 6, pp. 665-674, June, 2020.

[18] Y. Fu, M. Li, H. Luo, and G.Q. Huang. "Industrial robot selection using stochastic multicriteria acceptability analysis for group decision making." *Robotics and Autonomous Systems,* Vol. 122, December, 2019.

[19] S.Y. Chou and Y.H. Chang. "A decision support system for supplier selection based on a strategy-aligned fuzzy SMART approach." *Expert Systems with Applications,* Vol. 34, No. 4, pp. 2241-2253, May, 2008.

[20] T.T. Mac, C. Copot, D.T. Tran, and R.D. Keyser. "Heuristic approaches in robot path planning: A survey." *Robotics and Autonomous Systems,* Vol. 86, pp. 13-28, December, 2016.

[21] A.A. Ali, A.T. Rashid, M. Frasca, and L. Fortuna. "An algorithm for multi-robot collision-free navigation based on shortest distance." *Robotics and Autonomous Systems,* Vol. 75, pp. 119-128, 2016.

Modern Marketing Platform for Surin Farmer Group

Jeeranun Tasuntia*, Anyawee Chaiwachirakumpol*, Wiphasith Hiranrat*,
and Khanistha Silachot**

Received: September 22, 2023
Revised: January 8, 2024
Accepted: March 8, 2024

* Corresponding Author: Anyawee Chaiwachirakumpol, E-mail: anyawee_chai@hotmail.com

DOI: 10.14416/j.it/2024.v2.002

Abstract

The objective of this research was to create a modern marketing platform for the Surin Farmer Group and evaluate user satisfaction with the platform. The platform development process involved integrating the Line application (Line), which is highly popular among farmers, to provide convenience and user-friendly functionality. The Software Development Life Cycle (SDLC) process was employed to assist in system development, and the tools used for system development included Visual Studio Code, PHP as the programming language, JavaScript Library (jQuery), and MySQL database management system. Additionally, the LINE Frontend Framework or LIFF, a technology used for connecting Line applications, was incorporated. The principles of a LINE business account (Line OA) were applied to enhance the development into an online marketing platform. The study included a sample of individuals interested in buying and selling products online in Surin province. The sample groups consisted of 20 farmers from Surin province and 50 general users. The research findings indicated that the developed platform is effective and responsive to user needs. The satisfaction evaluation results revealed that the sellers (farmers) had a satisfaction rating of 4.32, S.D. = 0.45, indicating significantly high satisfaction levels. On the other hand, the buyers (general users) had a satisfaction rating of 4.40, S.D. = 0.39, also indicating significantly high satisfaction levels.

Keywords: Modern Marketing, Online Marketing, Line OA.

1. Introduction

Surin Province is well-known for its large-scale agricultural activities, specifically for rice cultivation [1]. The province is committed to using organic farming practices and, in addition to rice, grows a wide variety of crops. The community produces a range of agricultural products, including goods derived from agricultural processing. These community products comprise processed agricultural goods and items made from leftover agricultural materials. For example, sweet potato is a popular alternative crop for Surin farmers, and during certain seasons, there is an oversupply that leads to low prices [2]. As a result, Surin has started to process sweet potato into health-oriented food products to manage surplus production efficiently. After processing, the surplus is often used as animal feed. Moreover, the silk industry in Surin is thriving. Silk cocoons, silk fabrics, and silk products are well-known and have a diverse market. Traditional items like "Ton-Morn" (a type of silk cocoon), handicrafts, jewelry, and various other goods made from silk are all part of the product range. The silk industry is a significant aspect of Surin's economy, and its products are sold in various forms to meet customer demand. The primary occupation of Surin's farmers is agriculture, primarily crop cultivation. However, during idle periods in farming, many households engage in supplementary activities such as silk weaving [3]. There is a wide range of agricultural products and outputs. Sometimes, oversupply issues arise, and having only one sales channel is insufficient for effective product distribution. Enhancing sales skills and expanding marketing channels could alleviate this challenge, providing farmers in Surin with increased income opportunities.

The Modern Marketing Platform for Surin Farmer Group

* Major of Computer Technology, Faculty of Agriculture and Technology, Rajamangala University of Technology Isan, Surin University.
** Department of Business Data Management, Faculty of Business Administration and Management, Ubon Ratchathani Rajabaht University.

focuses on online marketing to adapt to the current global changes. The development process employs information systems using the Systems Development Life Cycle (SDLC) [4]. The strategy involves expanding sales channels widely, by incorporating the LINE Official Account business model. This model uses the LINE Front-end Framework with Web View and Rich Menu to display menus on the LINE application. The application acts as an online store that helps businesses reduce their costs and gain access to new customers. The LINE application offers an easy-to-use communication channel that is popular for detailed discussions beyond image-based communication. This makes it convenient for customers to browse and purchase products, thus promoting the marketing and sales of agricultural products from Surin Province.

2. Research Objectives

2.1. To develop a modern marketing platform.

2.2. To determine results of user satisfaction evaluations using modern marketing platforms.

3. Theoretical Background and Related Researches

To create modern marketing platforms for Surin farmer groups, we conducted thorough studies of various theories and documents.

3.1 Modern Marketing Management (MMM)

This pertains to business management, which encompasses utilizing modern technology to engage with customers through various means, including marketing and operations. The objective is to analyze the genuine needs of customers, gathering information to enhance products and services. It considers several factors, including marketing management, operations, and technology. [5]

3.2 Concepts about the marketing mix

The marketing concept centers around satisfying the needs and desires of customers while also achieving the organization's goals. This approach distinguishes itself from the production and sales concepts. The production concept prioritizes efficiency in the production and distribution of products. To effectively sell products or services, it's crucial to prioritize the needs of the customer. This can be achieved through various marketing strategies, such as high-powered personal selling and aggressive advertising. For small businesses, it's paramount to apply marketing concepts that emphasize meeting the needs of customers. Here are some ways to do this [6].

1) Prior to successfully managing a small business, owners must possess a thorough understanding of its operations.

2) Emphasize meeting customer needs by analyzing their preferences to provide pertinent products.

3) Who constitutes your customer base? These individuals possess distinct requirements. What is their expenditure capacity for goods and services?

4) Grasp the characteristics of buyers that impact their purchasing choices.

5) Recognizing your competitors and comprehending their strategies is imperative. A thriving business must deliver superior products and services.

3.3 System development guidelines

In the development of modern marketing platforms, the System Development Life Cycle (SDLC) is used as a development guideline. The SDLC consists of a total of seven steps [7].

1) Problem Definition: This process involves identifying the causes of issues in current operations, exploring the possibility of developing a new system, and determining the needs and requirements of system analysts and users. This information can be gathered through interviews and by collecting data from various operations to summarize the requirements.

2) Analysis: Analysis involves examining a system by gathering information and creating a logical model. The system analyst designs the system based on user requirements, determining the desired behavior, format of display, and data storage. The analysis is presented through a data flow diagram and a data dictionary. Let's review this process again to ensure a thorough understanding.

3) Design: The process of system design involves creating

a new system that meets the needs of its users. Systems analysts are responsible for designing the system's functionality, as well as determining the appropriate hardware and information technology to be used. They also design the program interface that connects users with the system and set up the network characteristics for connecting computers. Additionally, they establish security measures and estimate the various expenses that will be incurred.

4) Development: This phase involves a collaborative effort between programmers and system analysts to create the system using the components gathered from steps 2 (Analysis) and 3 (Design). Documentation must also be prepared, and users must be trained simultaneously.

5) Testing: System testing is a crucial process to ensure that the system is fully functional before it is put into practical use. The developer first tests the data by creating simulated data. They then check the system's operation to identify any errors. If any errors are found, the system goes back to the development stage for further improvements.

6) Implementation: Once the system has been tested and its functionality and user-friendliness confirmed, proceed with the system installation for practical use.

7) Maintenance: Maintenance refers to the process of enhancing and modifying a system once it has been installed and used. This may be necessary if newly created programs encounter issues at some point or if they have been used for an extended period. Additionally, changes in the organization's structure or expansion may require modifications to the program to align with the organization's structure. Maintenance can be divided into four types: corrective maintenance, adaptive maintenance, perfective maintenance, and preventive maintenance, each serving a specific purpose.

3.4 Computer language used for system development.

Creating a modern marketing platforms, the tools and computer languages utilized for development were as follows:

1) PHP is a programming language that allows for embedded scripting. This means that PHP commands can be added to web pages alongside HTML tags and saved as files with the extensions .php, .php3, or .php4. The syntax used in PHP combines elements from C, Perl, and Java. PHP commands are typically incorporated into HTML documents. When using PHP language, a client opens a web browser to request a PHP file. The browser then asks the server for the PHP file, which is searched for and run through the PHP engine. The engine processes the data in the database and sends the results back to the client. These results are converted into HTML and sent back to the web browser for display [8].

2) HTML is a file format saved in ASCII code and created using text editing programs like Notepad or word processing software. HTML files are made up of tags, which are HTML commands enclosed within < and > symbols. These tags come in two types: container tags and empty tags. A container tag has an opening tag and a closing tag. The closing tag will have a "/" preceding the tag, like <h1>...</h1>. Meanwhile, empty tags only have an opening tag, such as <HR>, which can be written in uppercase or lowercase letters without affecting the web browser's display. This information is consistent across various sources [9].

3) CSS is a style sheet language that uses a specific syntax, like HTML and XHTML. It is used to format and decorate HTML and add spacing to make web pages more flexible and visually appealing. It allows for easy and quick editing of the appearance of information in the presentation. CSS is standardized by W3C (World Wide Web Consortium). To set various properties of HTML or XHTML documents, we can assign values to HTML elements like <body>, <p>, and <h1>. This helps to standardize our website [10].

4) JavaScript is a programming language that is distinct from HTML and CSS, which are display languages or markup languages. JavaScript is a processing language that can perform calculations, use variables, and execute arithmetic operations such as addition, subtraction, multiplication, and division. These capabilities are not present in HTML, so JavaScript is required. JavaScript is a small program that can be added to HTML code. It is important to note that JavaScript is a case-sensitive language, so capitalization matters.

For instance, 'Myworld' and 'MyWorld' are different names because the 'W' is in uppercase and lowercase, respectively. As a result, it is essential to be careful when naming variables [11].

3.5 MySQL database management system (MySQL)

MySQL is a widely used database often paired with PHP. It is favored by many for its medium size and high level of user satisfaction. Additionally, MySQL is available for free. However, interacting with the database through the command line can be challenging. To address this issue, tools were developed to assist PHP programmers in managing MySQL databases more efficiently. One such tool is a web application called phpMyAdmin, written in PHP, which can handle multiple tasks with MySQL. It interacts with users graphically, making it effortless to use. This tool saves time by reducing errors and eliminates the need for typing commands [12].

3.6 Line Developer management and connection tools

The LINE Frontend Framework, also known as LIFF, is a platform developed by LINE Corporation that helps developers create web applications that can work within the LINE application and chat. It uses modern web technologies like HTML, CSS, and JavaScript. The main function of LINE LIFF is to expose applications within specific areas of chats or the LINE application. Users can access these applications through buttons or links configured in the chat, without the need to install a new application. This makes it easier for users to connect with LINE applications and enables developers to develop and update their applications without requiring users to download and install new versions. LINE LIFF allows developers to create applications with various functionalities like displaying news, products, schedules, personal information, or connecting to other services like location sharing, e-commerce, hotel reservations, and various application forms. [13].

3.7 Related research

The development of an integrated marketing communication strategy for holistic promotion of agricultural tourism in Nan Province is to propose a comprehensive strategy that builds upon the existing marketing communication activities. The research findings suggest that to successfully promote tourism in Nan Province, the strategy should be all-encompassing. The proposed strategies include developing a centralized website that serves as a comprehensive information hub for complete tourist destinations, utilizing online social media platforms such as Facebook, Instagram, Twitter and LINE official account, organizing special events like product exhibitions or travel activity showcases, public relations through television programs, promoting sales through travel agencies like Agoda and Traveloka, creating landmarks for check-ins, encouraging tourists to capture photos, and YouTube advertising. All these activities will be incorporated under a unified strategy and plan to transition marketing communication into a new format. This approach aims to meet the real needs of consumers and expand the consumer base for tourism [14].

A new e-commerce platform has been developed for online product deposit through a collaborative website with portable devices. It consists of mobile applications for general users and a website for system administrators. The results of user satisfaction across all groups indicate that the system has been highly effective, with an average rating of 4.30 (x). This suggests that the presented system can be used to conduct online business, creating job opportunities and generating income. This is particularly beneficial for those interested in buying and selling or seeking additional income in the agricultural sector, providing them with the maximum possible outcome [15].

The purpose of developing a website for the Office of Behavioral Control in Surin Province through a systematic process is two-fold. First, it aims to improve the website's functionality and secondly, to evaluate user satisfaction with its usage. Based on the research findings, system administrators have given an average satisfaction rating of 4.98, which falls within the "very good" range. The standard deviation is 0.02. General users have also been given an average satisfaction rating of 4.86, which is also within the "very good" range. However, the standard deviation is 0.35 for the general users. Both the system administrators and general users found the website to be effective and easy to use within the defined scope of work. [16]

4. Research Methodology

4.1 Problem Definition

A researcher visited Surin Province to gather data from a group of farmers. Through their study of agriculture, product sales, marketing basics, and social media usage, they discovered that the main products of this group are jasmine rice, silk, and various silk-based processed products. Sales primarily occur through closed market systems or via middlemen. The most used social media platforms are Line and Facebook.

4.2 System Analysis

After collecting information during the field visit, the researcher analyzed the data and proposed a new marketing channel for farmers in Surin Province. The idea involves utilizing the popular social media application Line to sell products. This modern marketing platform could greatly benefit farmer groups in Surin Province. The platform aims to offer a convenient way for farmers to manage their stores and for buyers to easily find and purchase products from these groups. It involves two main players: sellers (farmers) and buyers.

Sellers:

1. Login: The seller can log in to their account to access their store.

2. Manage Store: The seller can add, delete, and modify existing product listings in their store.

3. View Order List: When customers place orders, the seller can view the list of orders.

4. Confirm Payment and Update Delivery Status: The seller can confirm payment and update the delivery status when products are shipped to customers.

5. View Summary Reports: The seller can view summary reports of all orders within their store.

Buyers:

1. Search for Stores or Products: Easily search for stores or products that you are interested in.

2. Log into Your Account: After finding what you want, log into your account.

3. Place Your Order: Once logged in, place your order.

4. Make Your Payment: After placing your order, you will receive an order summary and will need to make the payment to complete the transaction.

5. Check Your Delivery Status: When a tracking number becomes available, check the delivery status of your order.

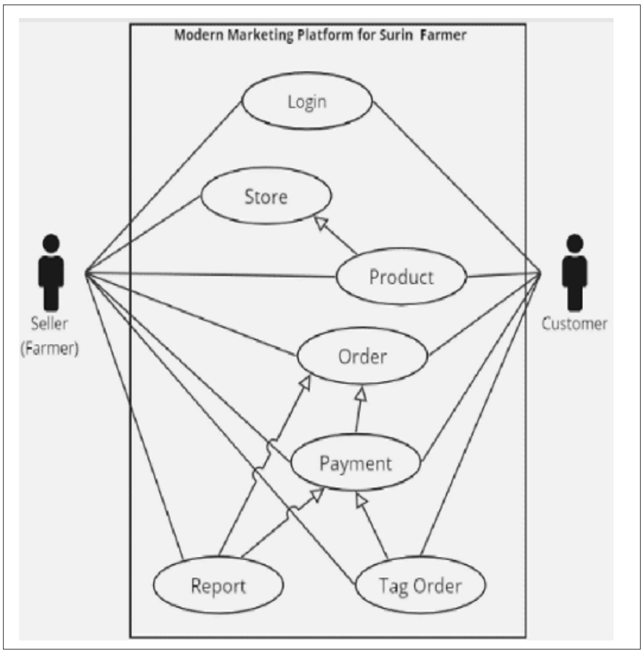


Figure 1. *Use Case Platform Overview.*

4.3 System Design

After obtaining details from data analysis, the researcher proceeded with designing a platform using a system design program. This platform can be accessed through the Line system and can connect to the website for managing store and product information, as well as ordering information. It also facilitates the issuing of various reports. Buyers can conveniently use the widely used Line system for their transactions.

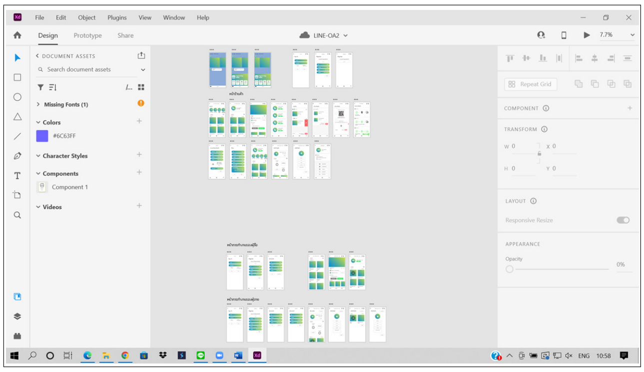


Figure 2. *User Interface Design.*

4.4 Development

1). To create a functional website, it's necessary to develop various components. One of these is the seller (farmer) section, which allows for the management of store and product information, as well as order processing. Additionally, various reports can be generated from this section. To create these components, Visual Studio Code software will be used, along with the Bootstrap framework for the User Interface section. On the server side, scripting languages such as PHP and JavaScript will be utilized, while HTML and CSS will be used on the user side. This system will be designed to run on the Windows operating system.

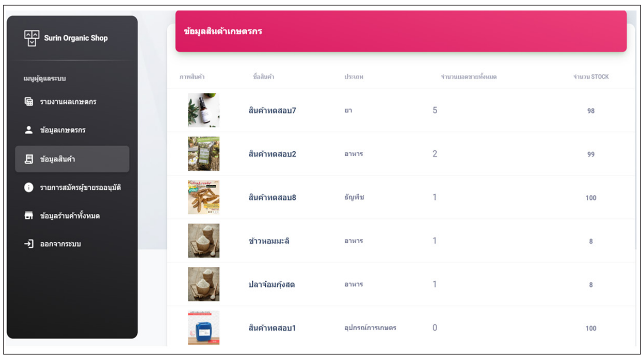


Figure 3. *Product Management.*

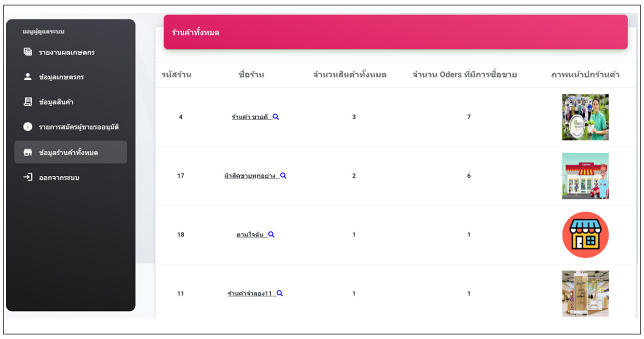


Figure 4. *Store Management.*

2). The Line application uses the LINE Frontend Framework (LIFF) to create Line OA components. This technology is used to develop the main functionality for both sellers (farmers) and buyers accessing the application. Using LIFF, users can search for products, find stores, place orders, and track all orders within the LINE application. To enable the Line Chatbot to work seamlessly with web applications, a necessary process involves connecting a Web Application and the Line Chatbot Platform.

Figure 5. *Part of use via Line.*

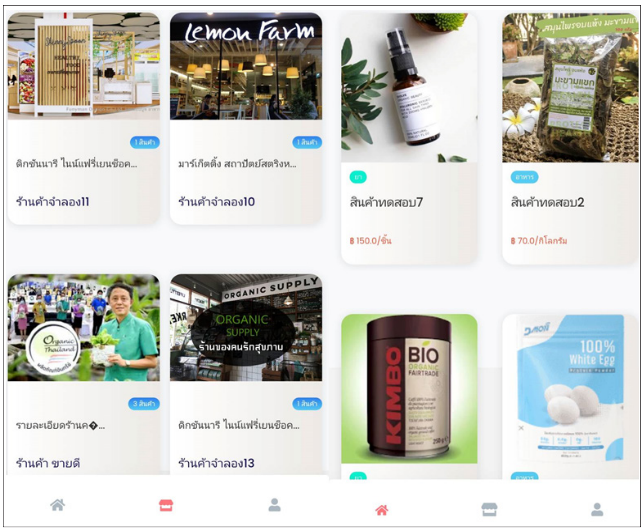


Figure 6. *Store and Product via Line.*

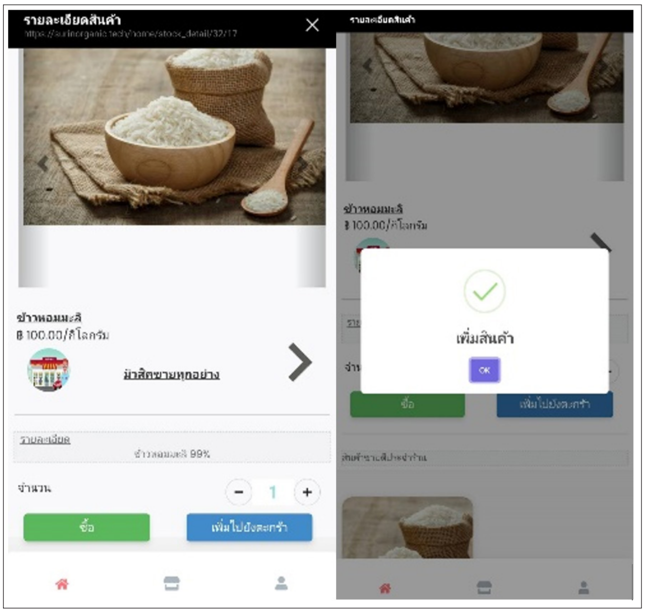


Figure 7. *Add product in cart.*

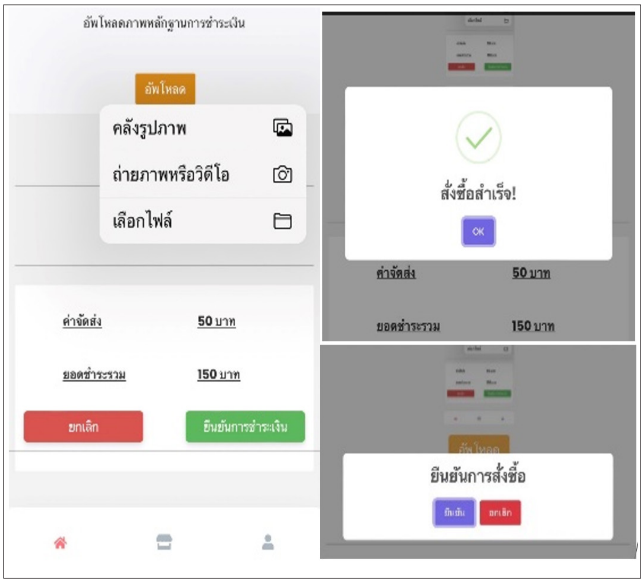


Figure 8. *Confirm order and payment.*

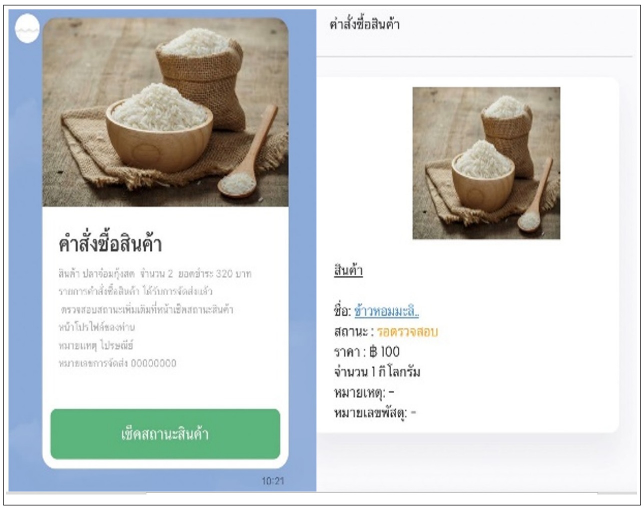


Figure 9. *Check product status.*

4.5 Testing

During system testing, the test is divided into two parts: 1) sellers (farmers) and 2) buyers. Both groups utilize the Black Box Testing method. The results indicate that the system operates within its specified limits and can be effectively used.

4.6 Implementation

When creating contemporary marketing platforms for farmer groups in Surin Province, we conducted a satisfaction assessment. We analyzed this data using the Likert Technique, employing descriptive statistics to determine the median value of the data through the mean (Mean) and measuring the distribution of the data through the standard deviation value (Standard Deviation) [17]. Here are the details of our analysis:

Score 5 indicates a very high level of satisfaction.

Score 4 indicates a high level of satisfaction.

Score 3 indicates a moderate level of satisfaction.

Score 2 indicates a fair level of satisfaction.

Score 1 indicates improved satisfaction results.

System users are rated using Likert's method, which consists of 5 levels of scores [17].

4.21 - 5.00 indicates a very good level.

3.41 - 4.20 indicates a good level.

2.61 - 3.40 indicates a medium level.

1.81 - 2.60 indicates a fair level.

1.00 - 1.80 indicates an improvement level.

To interpret the questionnaire results, calculate the average score for each class and apply basic statistics by determining the range (highest value minus lowest value). To find the width of the class rate, use the following formula.

Class interval = upper-class limit - lower-class limit. Use the collected data to analyze the statistical results by computing the mean and standard deviation (SD) [17].

Mean Formula

$$(\bar{x}) = \frac{\sum x}{n}$$

$(\bar{x})$ = Mean

$\sum x$ = Sum of all data points

n = Number of data points

Standard Deviation (SD) formula

$$SD = \sqrt{\frac{n\sum x^2-(\sum x)^2}{n(n-1)}}$$

S.D. = Standard Deviation

n = Number of data point

x = Each of the values of the data

$\sum x$ = Sum of all data points

5. Data Collection

During the data collection phase, a sample group was chosen from the population. The sample group consisted of people who were interested in buying and selling products from Surin Province through online channels. The sampling method used was random sampling, where individuals were selected without any specific criteria, but they had to be part of the overall population interested in the study. The group of selected samples included 20 farmers from Surin Province who were engaged in selling various agricultural products and 50 general consumers interested in purchasing products from the farmers. These consumers were part of the group that used the new-age marketing platform. Afterwards, an evaluation was conducted using a questionnaire to assess the usability and effectiveness of the platform among the group of farmers from Surin Province. The research team then analyzed and interpreted the evaluation results based on the statistics.

6. Performance results

6.1 System development results

A Modern Marketing Platform was created for the Surin Farmers Group using Visual Studio Code software. The website's User Interface section was developed with the Bootstrap framework, while PHP was used as a scripting language on the server side and JavaScript on the user side. HTML and CSS were utilized for the site's operation, which runs on the Windows operating system. LINE Frontend Framework or LIFF was employed to integrate Line OA, the technology used by the Line application. The platform involves two main players: sellers (farmers) and buyers.

Figure 10. *Login via Line.*

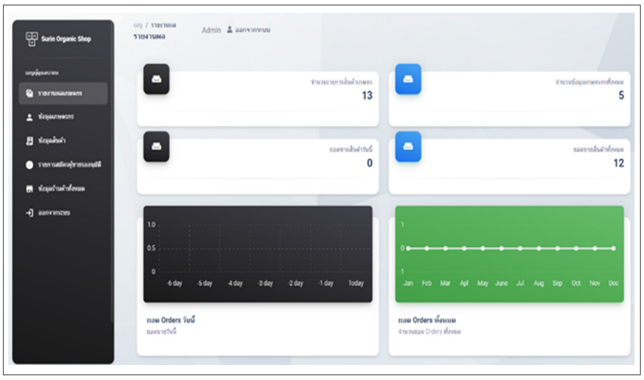


Figure 11. *Information Management.*

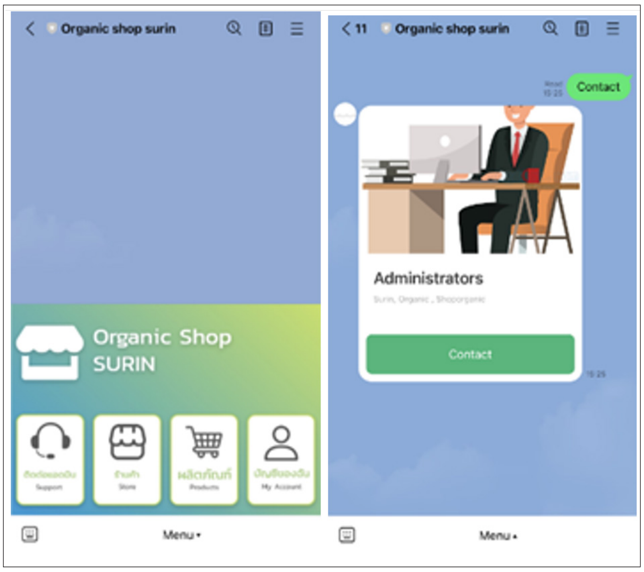


Figure 12. *Bayer use via Line.*

6.2 Satisfaction Assessment Results

To assess the level of satisfaction among the Surin Farmers Group who use Modern Marketing Platforms, separate the evaluation results into two parts: the results of the satisfaction evaluation from the sellers (farmers), as displayed in Table 1, and the results of the buyer satisfaction evaluation, as displayed in Table 2.

Table 1 shows that 20 sellers (farmers) are highly satisfied with the modern marketing platform, with an average rating of 4.32 and a standard deviation of 0.45.

Table 2 shows that 50 buyers are highly satisfied with the modern marketing platform, with an average rating of 4.40 and a standard deviation of 0.39.

Table 1. *Satisfaction Assessment Results of Seller (farmers) (n=20).*

Questions	$\bar{x}$	S.D.
System Capabilities	4.40	0.09
1. Exactness	4.39	0.61
2. Handily	4.28	0.67
3. Suitability	4.44	0.62
4. Ability content	4.50	0.51
Working According to Function	4.25	0.22
5. Easy to use commands	4.33	0.69
6. Errors protection	4.28	0.67
7. Working speed	3.94	1.06
8. Satisfaction	4.44	0.78
System Design	4.36	0.14
9. Suitability on display	4.39	0.50
10. Menu position	4.22	0.55
11. Color of letters and background	4.56	0.62
12. Explanatory texts	4.39	0.70
13. order of data on display	4.22	0.73
Security System	4.19	0.04
14. Checking login errors	4.22	0.81
15. Conditions for checking users	4.17	0.62
Average	**4.32**	**0.45**

Table 2. *Satisfaction Assessment Results of Buyer (n=50).*

Questions	$\bar{x}$	S.D.
System Capabilities	4.43	0.11
1. Exactness	4.52	0.58
2. Handily	4.33	0.62
3. Suitability	4.33	0.73
4. Ability content	4.52	0.58
Working According to Function	4.38	0.14
5. Easy to use commands	4.56	0.58
6. Errors protection	4.22	0.85
7. Working speed	4.41	0.69
8. Satisfaction	4.33	0.73
System Design	4.39	0.08
9. Suitability on display	4.44	0.64
10. Menu position	4.41	0.64
11. Color of letters and background	4.33	0.68
12. Explanatory texts	4.48	0.64
13. order of data on display	4.30	0.67
Security System	4.41	0.05
14. Checking login errors	4.44	0.51
15. Conditions for checking users	4.37	0.56
Average	**4.40**	**0.39**

7. Discussion and Conclusion

Based on the research findings, the following summary and discussion can be presented.

7.1 The study focused on creating Modern Marketing Platforms for the Surin Farmers Group. The system development cycle theory was employed as the guiding principle for the development process. The results demonstrated the platform's effectiveness and its potential for extensive use. These findings align with the research conducted by Thinnapat et al. [16], who applied a similar approach in developing the website for the Surin Probation Office. The use of system development tools proved valuable in creating a functional and practical system that met the users' needs. Similarly, a study by Jeeranun et al. [4] found that the Scheduling System with the Drag and Drop Technique at Sai Kaew Witthaya School in Surin Province effectively operated within the defined scope and objectives. This success is attributed to the clear steps involved in the system development cycle. Furthermore, the Line application has introduced trading channels that provide convenience and promote sales for both farmers and buyers. In a study by Chutiphon et al. [15], they developed a new E-commerce platform for product preordering, which included a mobile app for general users and a website for system administrators. The research indicated that this platform has the potential to create job opportunities and generate income for those interested in pursuing business ventures.

7.2 The evaluation of user satisfaction with the Modern Marketing Platforms for the Surin Farmers Group has yielded positive results. Both the sellers (farmers) and the buyers (general users) expressed a high level of satisfaction with the system. Overall, both sample groups reported very high satisfaction levels, indicating that the system effectively meets their needs. The operations are based on the principles of the system development cycle theory. The process commences with field visits to study and collect initial data from groups of farmers. This aids in developing a system that caters to the needs of users and supports flexibility based on their regular usage. Similarly, as per the research conducted by Adisak et al. [14], guidelines for integrated marketing communication to promote Argo-tourism in Nan province require a successful integrated marketing communication approach. This approach should build upon existing marketing communication activities and lead to new formats that meet the actual needs of consumers.

The Modern Marketing Platform for Surin Farmer Group can function effectively within its defined scope and meet user requirements. However, it is important to promote additional aspects among the farming community to improve the efficiency of this marketing approach. One key aspect is to enhance knowledge related to marketing and public relations. This can be achieved by leveraging other social media platforms to establish networks with consumers, retailers, and supporters. Additionally, using data to make decisions on crop selection, overall management, and production planning could be beneficial in increasing sales in the future.

8. References

[1] W. Nantakaew, *Rice growing area of Surin province.* Available Online at https://thainews.prd.go.th/th/news/detail/TCATG200723195102937, accessed on 28 June 2021.

[2] R. Ruttanaprasert. R. Puttha, Y. Butdaphong, and R. Sennoi. "Influence of chicken manure on growth and root yield of different sweet potato varieties." *Agriculture and Technology Journal,* Vol. 1, No. 3, pp. 42-51, September-December, 2020.

[3] M. Phayakprakhon. "Weaver of folk art with social change at Surin Province." *The Journal of Social Communication Innovation,* Vol. 3, No. 2, pp. 16-24, July-December, 2015.

[4] J. Tasuntia, B Nabumrung, A. Chaiwachirakumpol, C. Kunok, and C. Sanmee. "Scheduling System with Drag and Drop Technique Sai Kaew Witthaya School Surin Province." *Agriculture and Technology Journal,* Vol. 1, No. 2, pp. 70-85, May-August, 2020.

[5] S. Khumhaeng, S. Na Pathum, and P. Laksitamas. "The Modern Marketing Based of New Startup Thailand 4.0." *Journal of Humanities and Social Sciences Thonburi University,* Vol. 12, No. 29, pp. 45-55, September-December, 2018.

[6] W. Thosuwanjinda. *Organizational Secrets: Modern Organizational Behavior.* Bangkok: Thumniti, 1992.

[7] Computer Science Textbook Production, Department. *System Analysis and Design.* Bangkok: SE-Education Public Company Limited, 2008.

[8] C. Kaewpet, *Introduction to Web Programming.* Available Online at http://ba2it.rmutsb.ac.th/ Suphan, accessed on 28 July 2021.

[9] P. Vijitboonyaruk. "HTML: a web language." *Executive Journal,* Vol. 31, No. 3, pp. 199-202, 2011.

[10] W. Tumtong. *Basics of information development on the Internet with HTML5 and CSS3.* Bangkok, Phappim, 2014.

[11] S. Sangakong and P. Naksong. *Develop your website*

and add features with JavaScript. Bangkok, Chulalongkorn University Press, 2019.

[12] B.Pasilatesang. *Web development guide with PHP 5 and MySQL 5.* Bangkok, SE-Education Public Company Limited, 2007.

[13] Twinsysnergy. *LIFF (LINE Front-end Framework) with ReactJs and Ngrok.* Available Online at https://twinsynergy.co.th, accessed on 17 February 2022.

[14] A. Champathong, N. Hiranwijitporn, K. Tikamat, and S. Kaewphrae. "Guidelines for integrated marketing communication to promote Agro-tourism in Nan province." *Western University Research Journal of Humanities and Social Science,* Vol. 6, No. 1, January-April, 2020.

[15] C. Huanthanom, N. Saladkaew, and S. Chunwiphat. "A Development of New E-commerce Platform for Product Preordering." *The Journal of Industrial Technology,* Vol. 16, No. 1, January-April, 2020.

[16] T. Opor, U. Thaworn, J. Tasuntia, A. Khunlerd, and A. Chaiwachirakumpol. "The Development of Website Surin Probation Office." *Proceedings of the 13th Rajamangala Surin National Conference,* pp. B444-B457, 2022.

[17] L. Saisot and A. Saiyot. *Educational Research Techniques.* Bangkok, Suweeriyasan, 1995.

Thinking Skills Level Classification of Scientific Questions Using Bidirectional LSTM

Nanthawan Yaemsawat* and Nuanwan Soonthornphisaj*

Received: December 14, 2023
Revised: July 3, 2024
Accepted: July 19, 2024

* Corresponding Author: Nanthawan Yaemsawat, E-mail: nanthawan.ya@ku.th

DOI: 10.14416/j.it/2024.v2.003

Abstract

Science education with a suitable learning activity can help students enhance their thinking skills. Examination is one of the assessment tools to evaluate the student learning outcome in the domain of thinking skills. The Revised Bloom's Taxonomy, a well-known theory used to describe cognitive domains, divides thinking skills into two categories: basic and advanced thinking skills. Classifying questions according to their level of thinking abilities is an important task for teachers to design effective assessment tools. The objective of this study is to propose a model for classifying Thai language questions in science subjects. Initially, we used three algorithms: Bidirectional LSTM (BiLSTM), Naive Bayes (NB), and Support Vector Machine (SVM) for selecting Thai word tokenization algorithms. Then, we compare the model's performance using different feature sets. The combination of the question, training choice, and length of choice features with BiLSTM obtained an accuracy of 70%. Moreover, we employed part-of-speech (POS) tagging for feature selection. According to the findings, using nouns, verbs, adjectives, and adverbs enhances accuracy by 80.24%. This study shows the ability to use a model to categorize science questions to assist teachers in choosing questions that are appropriate to encourage higher-order thinking skills in students.

Keywords: Text Classification, Deep Learning, Learning Taxonomy, Science Education.

1. Introduction

Science education is aimed at developing important scientific skills in students, such as problem-solving, critical thinking, creative thinking, and decision-making. According to "The Future of Jobs Report 2023" from the World Economic Forum, analytical and creative thinking are the most significant skills for workers [1]. As a result, students must develop these skills, as they are necessary for the 21st century. The revised Bloom's taxonomy classifies learning outcomes into six levels of cognitive ability. These include remembering, understanding, applying, analyzing, evaluating, and creating [2], with each step representing a higher degree of cognitive growth. The questions at the top three levels are regarded as higher-order thinking abilities, whereas the questions at the bottom three levels are considered lower-order thinking skills [3]. Therefore, student assessments are essential to assess students' knowledge and skills. The teacher needs to design questions that are relevant to their learning outcome.

At present, many students in Thailand appear to lack 21st-century skills, especially higher-order thinking skills such as critical thinking [4]. As a result, teachers must create innovative methods of instruction for their students. Measurement from questions is a way of determining how learning activities help students improve their thinking skills. Therefore, teachers must create or select suitable questions from the question bank. The main issue is that teachers need to understand which questions require higher-order thinking abilities and which require lower-order thinking skills to assess and design effective tools. In addition, it was discovered that there are numerous institutions in Thailand, particularly in the countryside. Due to government budget constraints and limited teacher employment framework, teachers,

** Department of Computer Science, Faculty of Science, Kasetsart University Bangkok.*

such as those with a degree in social sciences, must teach subjects in which they are not proficient, such as science. As a result, the design or selection of questions from the bank may be inappropriate. As a result of this research, we hope to help teachers understand the level of thinking of each question designed or chosen from the question bank so that they can choose questions that are appropriate for the instructional level objectives and encourage higher-order thinking skills.

Science questions can be classified automatically using machine learning (ML) techniques. According to previous studies, [5] implemented a K-Nearest Neighbors (KNN) and a Support Vector Machine (SVM) to classify questions into six levels based on the revised Bloom's taxonomy. A dataset containing 1,000 questions from an operating system course. They compared the performance of the two classifiers using the same dataset. The findings indicated that SVM outperformed KNN, with the highest F1-score of 92.3%. In addition, In the past few years, the classification of questions has become increasingly dependent on deep learning (DL). For example, [6] sought to improve the automatic revised Bloom's taxonomy CLOs and exam questions classification model by combining the proposed LSTM model with contextual domain embeddings. The dataset includes fields in computer science, electrical engineering, and business administration domains. They used different pre-trained embeddings to learn efficient word representations for their datasets and found that pre-trained embeddings, "namely, "Wiki Word Vectors", provided the highest accuracy (74% for the CLO dataset and 87% for exam questions). Difference classifiers from ML and DL were used for comparison, but the LSTM classifier performed the best.

Previous research has applied ML or DL to classify questions in English. Nevertheless, research that uses this method to classify science questions written in Thai is still scarce. To solve the problem, we propose a model for classifying science questions in Thai based on learning level, which is divided into 2 levels: basic thinking skills (lower-order thinking skills) and advanced thinking skills (higher-order thinking skills), using BiLSTM, NB, and SVM classifiers to compare performance. In this study, we describe approaches to effectively classifying science questions using ML and DL algorithms to develop an appropriate model that can assist teachers in creating suitable tools to measure and evaluate students, as well as improve students' higher-order thinking skills. As a result, utilizing this proposed model to classify science questions will significantly benefit Thai educators.

The remainder of this paper is as follows: Previous studies that classified questions using deep learning and machine learning are discussed in Section 2. Section 3 illustrates preprocessing and the processes of ML and DL to learn from the dataset, whereas Section 4 discusses the results of the experiment. Finally, Section 5 concludes this research paper.

2. Related Work

2.1 Bloom's Taxonomy

Benjamin Bloom was an educational psychologist who formulated a learning taxonomy that depicts the hierarchical behavioral characteristics of learners. He stated that every individual has three separate learning domains: cognitive, affective, and psychomotor [7].

In 2001, a team of psychologists under the direction of Anderson and Krathwohl—researchers, measurement, and evaluation experts—published A Taxonomy for Teaching, Learning, and Assessment [2]. Bloom's taxonomy is improved within the Cognitive Domain to facilitate the creation of learning objectives. After revision, this taxonomy gained widespread recognition in the academic community. Figure 1 shows six cognitive behavior levels the instructor desires students to possess. In addition, cognitive domain behavior has been divided into two levels based on the level of thinking skills: basic and advanced thinking skills. Note that the first three levels of thinking behavior: remembering, understanding, and applying are considered basic thinking skills. Analyzing, evaluating, and creating are considered advance thinking skills [3].

2.2 ML in Question Classification

Many studies use ML to classify computer programming

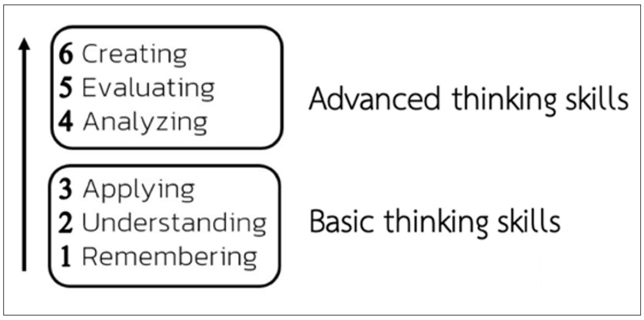


Figure 1. *The revised Bloom's Taxonomy.*

questions based on revised Bloom's taxonomy. For example, [8] aimed to classify programming questions utilizing KNN and NB classifiers as well as feature selection based on Chi-Square, Mutual Information, and Odds Ratio. They found that feature selection plays a key role in classification performance. The best F1 score result obtained from the KNN classifier is 90% using the Mutual Information feature selection, with an F1 measurement of 89%. [9] utilizes the revised Bloom's taxonomy to categorize online content for programming languages into three classes based on difficulty: beginner, intermediate, and advanced level, allowing learners to select content that best meets their requirements. The performance of the bag of words and POS tagging was compared using a random forest for classification. The results demonstrated that the use of the revised Bloom's taxonomy verbs and synonyms enhanced performance. The maximum accuracy obtained is 98%.

Some studies used ML to classify questions from different domains. [10] used ML to classify questions in a variety of areas, such as chemistry, biology, social marketing, math, law, and others, using the SVM, KNN, and Naïve Bayes (NB) classifiers. The dataset was collected from a variety of sources, including 600 questions. They observed that E-TFIDF, an upgraded version of the traditional TF-IDF feature, improves classifying questions better than traditional TF-IDF, while the SVM classifier excelled by up to 86%. Two years later, they compared feature extraction, including TF-IDF, TFPOS-IDF, and Word2Vec-TFPOSIDF. The question set was obtained from Yahya et al. (2012), and a few additional questions were gathered manually. The results revealed that Word2Vec-TFPOSIDF provided the best performance in this study, followed by TF-IDF and TFPOS-IDF [11].

Question classification based on revised Bloom's taxonomy has also been applied to languages other than English. [12] compared the performance of classification questions with two sets of questions translated from English to Chinese using Random Forests (RF), Logistic Regression, and XGBoost classifiers. POS tagging is used to generate keywords for training. The experimental results demonstrated that RF and selection keywords improve accuracy. The highest score was 86%. [13] categorized primary and high school questions. The dataset used would consist of mathematics and Indonesian language questions totaling 670. They trained both text and non-text questions to compare the results and use a pre-train model named IndoBert, which was learned from Indonesian text. The classifiers used were SVM and NB. The evaluation results show that the SVM classifier provided the highest value for mathematics, which is 82%, whereas the NB classifier provided the highest value for Indonesian language subjects at 63%. Both results are trained from text and questions.

Previous research has shown that ML can classify questions based on revised Bloom's taxonomy. In addition, feature selection plays an important role in improving classification performance. There is also research that uses questions in English and other languages from different domains to show that we can use ML to categorize questions in a variety of languages and subjects.

2.3 Deep Learning in Questions Classification

Word embedding is an essential approach for converting text to vectors before training deep learning algorithms. Many studies in the past have utilized various models to construct word embedding. [14] classified questions in English that were translated from Turkish using the Word2vec technique with a CBOW and skip-gram models. The dataset collects questions from numerous categories, such as animal, creative, food, and many more. They analyze models using CNN, LSTM, and a combination of CNN-LSTM and CNN-SVM classifiers to compare outcomes. The accuracy of the test data is quite high, with CNN and skip-gram, a type of Word2vec model, achieving 94%. [15] sought to identify an appropriate word embedding technique for five datasets to be pre-trained for contextual and non-contextual data. For non-contextual word embedding, Fast Text has an accuracy of 0.82, whereas for contextual RoBERTa, it has an accuracy of 0.85. Up to six techniques were compared. Once the appropriate word embedding technique was applied to the CNN classifier, the accuracy was found to be 86%, which was higher than previous studies in which the word embedding was in the dataset.

Some studies employ feature extraction for DL classification [16] employed the TF-TDF approach to extract features. They categorized 141 questions based on Bloom's taxonomy from a variety of domains and compared two classifiers, the Artificial Neural Network dataset (ANN) and SVM. The experimental findings showed that using TF-TDF to extract features enhanced classification performance, with the ANN classifier having the best performance of 85.2%. [17] applied the ETFPOS-IDF approach for examination question categorization, which is an improved version of the TFPOS-IDF proposed by [11] Three datasets were collected from a variety of fields, including computing, social science, business, and others. Three classifiers were used: SVM, ANN, and Random Forest. The results demonstrate that ETFPOS-IDF outperforms all previous studies' schemes in test question categorization, achieving 0.749 in accuracy and 0.746 in F1 score.

In addition, BERT (Bidirectional Encoder Representations from Transformers), a transformer-based AI deep learning technique, is also utilized in text categorization. [18] used BERT to classify computer education questions from the Canterbury Question Bank using Bloom's taxonomy. In the experiment, it was found that the dataset containing certain classes had an imbalance issue. As a consequence, the Application, Synthesis, and Evaluation classes must be eliminated to resolve class imbalance concerns. In Experiment 3, reducing the number of classes to be classified to the remaining three based on Bloom's taxonomy level produced the highest accuracy at 82.61%.

DL algorithms such as LSTM, CNN, and RNN are increasingly being utilized to categorize questions. Word2vec, ETFPOS-IDF, and other approaches have been developed and deployed. Despite the quantity of such research, the vast majority of data sets are in English or other non-Thai languages. To gain a suitable model, we must also investigate the work in Thai.

2.4 Thai Text Classification

ML and DL have been applied to classify Thai text in the education field based on revised Bloom's taxonomy, both in classifying questions and content written in blogs. [19] used multiple classifiers, such as NB, decision tree, SVM, and multilayer perceptron, to classify the revised Bloom's taxonomy-based questions in Thai. Datasets were collected from of several websites in Bloom's cognitive domain literature. They focus on feature selection. Cleaning data, word segmentation, part-of-speech tagging, and feature selection are all applied to each question. This experiment shows that verbs, adverb, adjectives, conjunctions, and question tags should be selected as features in Thai's exam classification. The highest accuracy of 71.2% with the SVM classifier. [20] desired to classify blogs according to information and communication technology (ICT) using the revised Bloom's taxonomy. Three classes are used to classify simple, moderate, and difficult blog content. Multiple classifiers, DL and ML, are utilized. The training data consists only of textual content. Utilize a dictionary and name entity recognition (NER) to tokenize words. The training text frequently contained English computer terminology. So, they preprocess in both languages. Deep neural networks (DNN) were found to have the highest classification performance of all classifiers in this experiment, with an accuracy of 87%.

In the educational domain, there are few Thai text classifications. We attempted to investigate more text processing in other domains, such as sentiment analysis. [21] classified Thai children's tales consisting of 1,115 sentences divided into three sentiments—negative, neutral, and positive—using combining BiLSTM and CNN classifiers with several combinations of the features, including POS tagging, semantic, and word embedding. The experiment result shows that the combination of POS tagging, word embedding, and semantic features provided the highest classification accuracy at 78.89%. [22] proposed a social media sentiment analysis model.

They collected a dataset from the Wongnai food platform website, which had 2000 positive and negative restaurant reviews. They employ the longest word pattern method for word segmentation and the Word2Vec method for word embedding. Gated Recurrent Unit (GRU), CNN, and LSTM classifiers were used to compare results. The results demonstrated that LSTM was more effective at classification than CNN and GRU, with an accuracy of 84%.

In recent years, there has been an increase in the amount of research that uses BiLSTM to categorize Thai text. [23] employed a BiLSTM classifier to analyze the text categorization of comments on Thai tourism-related YouTube channels. BiLSTM can narrow down the 43 remaining tourist categories to 33. The top four categories for places of interest are temples, food, coffee, and green tea, in that order. Performance of the proposed text classification at 85.78%. [24] used a BiLSTM classifier to detect clickbait. The Thai clickbait corpus consists of 30000 headlines collected from trendy websites shared on online social media and non-clickbait from newspapers, community blogs, and online magazines. DL models are used to compare outcomes. However, BiLSTM with word-level embedding performs the highest, with a 98% accuracy rate.

We found that the research on question classification in Thai using the revised Bloom's taxonomy is very limited. Therefore, we wish to develop a question classification model to be useful to Thai educators.

3. Proposed Method

This section describes the approach to preprocessing data, training data to build models, and evaluating experimental results. Figure 2 provides an overview of the modeling for the classification of Thai scientific questions based on learning level.

3.1 Dataset

This study, dataset contains the Thai questions in a science course. At the level of secondary education, grades 7-12, 1246 questions from the database of the Office of the Basic Education Commission (OBEC), Ministry of Education, have already been labeled for all six cognitive levels by educators. The majority of the questions are multiple-choice. The questions were separated into two categories. The first three levels of the revised Bloom's Taxonomy assessed basic thinking skills (lower-order thinking skills), while the top three levels assessed advanced thinking skills (higher-order thinking skills). Questions are distributed by their level labels. Table 1 demonstrates that the questions have been divided into two classes, so there is no class imbalance issue.

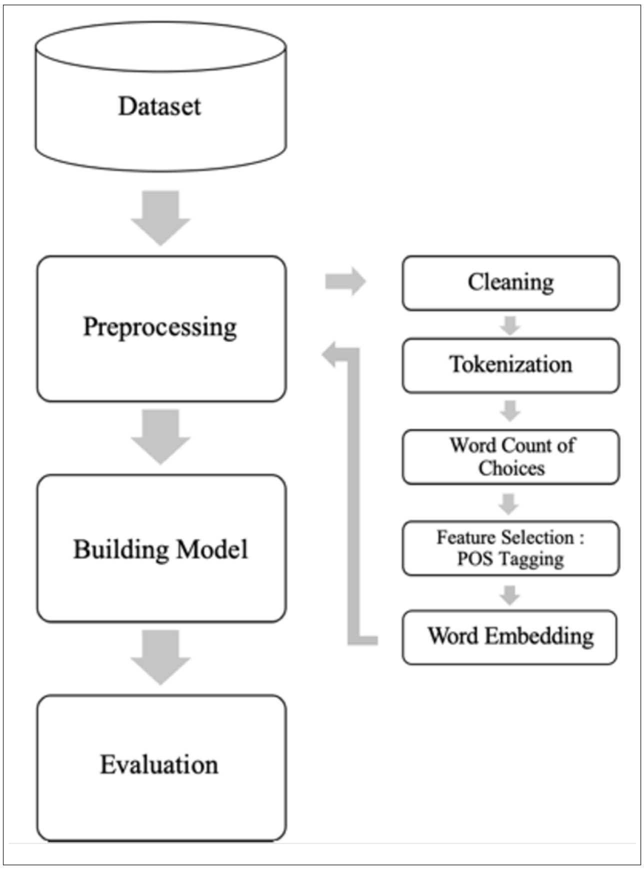


Figure 2. Overview of the proposed scientific question classification model.

3.2 Preprocessing

3.2.1 Cleaning and Tokenization

Based on the Natural Language Processing (NLP) tasks, we must preprocess the text data. Then, each question was cleaned by removing English letters and punctuation. We use the Newmm algorithm to tokenize words. The reasons are explained in the Experiment 1 section. We also compare the word sizes with and without removing stop words. In Table 2, it can be seen that if the stop word isn't removed, the average number of words per question is doubled, with 8 words per question for removing stop words and 14-15 words per question for non-removing stop words. However, based on the results of experiment 1, we have decided not to remove the stop words in this study because their performance is similar.

Table 1. *Class distribution of the dataset.*

Class	Cognitive Level	Question	Total
Basic Thinking Skills	Remembering	28	666
	Understanding	539	
	Applying	99	
Advanced Thinking Skills	Analyzing	573	580
	Evaluating	3	
	Creating	4	
Total			1,246

Table 2. *The size of words in the dataset.*

Class	Question	Word Size			
		Stop word removal	Avg.	Without stop word removal	Avg.
Basic Thinking Skills	666	5,405	8.07	9,666	14.51
Advanced Thinking Skills	580	4,990	8.60	9,100	15.69
Total	1,246	10,395	8.32	18,766	15.06

3.2.2 Word Embeddings

Word embedding, which represents a "word" as a "number", is one method to create machine-understandable features from words. The format of these numbers is a vector [25]. There are numerous word embedding models, including Word2vec, GloVe, FastText, and more. Each model employs a unique algorithm for vector generation.

Word embedding is an approach with which we represent documents and words before training. This study will utilize Thai2Vec from PythaiNLP, which has a function to generate vectors of words trained using the Word2vec family of techniques. The vocabulary was used to generate a 51,556-word vector from 300 dimensions [26], which will convert questions and multiple-choice answers into vectors for further training. In addition, to prevent the model from memorizing training data for overfitting, we set the vector length to 180 and scaled the dimension to the magnitude of the vector using the Principal Component Analysis (PCA) method [27].

3.3 Question Classification Algorithms

There are three main question classification approaches, including rule-based, machine-learning-based, and hybrid-based approaches [28]. We focus on machine learning approaches, including traditional methods and deep learning methods. In this paper, we experiment by using two standard algorithms of traditional methods, including Naïve Bayes and Support Vector Machines, and one approach of deep learning model, including Bidirectional Long Short-Term Memory.

3.3.1 Bidirectional Long-Short-Term Memory

BiLSTM is bidirectional long-short-term memory. The LSTM is a sort of recurrent neural network in which the previous step's RNN output is supplied as input to the next phase. It was invented by Hochreiter & Schmidhuber [29]. It solves the issue of long-term dependency on RNN, which cannot anticipate words stored in long-term memory. The LSTM can hold data for a long time and may be used for time series data processing, prediction, and classification. It has a chain structure made up of a neural network and multiple memory units known as cells. BiLSTM is a bidirectional LSTM, which means that the signal is propagated backward and forward. Figure 3 describes the BiLSTM layer's architecture, where X_i is the input token, Y_i is the output token, and A and A' are LSTM nodes. The combination of A and A' LSTM nodes is the final output of Y_i.

The BiLSTM architecture consists of two unidirectional LSTMs that process the sequence in both forward and backward directions. This architecture may be viewed as having two independent LSTM networks, one receiving the token sequence as it is and the other receiving it in reverse order. Both of these LSTM networks provide a probability vector as output, and the result is the sum of these probabilities [30]. It can be written as equation 1.

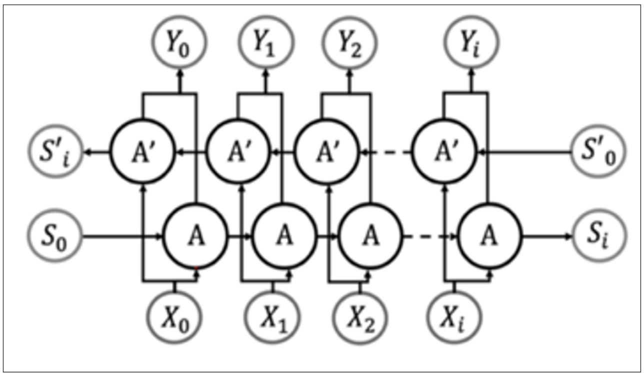


Figure 3. *BiLSTM layer architecture.*

$$p_t = p^f_t + p^b_t \quad (1)$$

where p_t is the final probability vector of the network, p^f_t is the probability vector from the forward LSTM network, and p^b_t is the probability vector from the backward LSTM network.

3.3.2 Naïve Bayes

Naïve Bayes (NB) is a probabilistic classifier. This approach applied Bayes' theorem, which was developed by Thomas Bayes, an English mathematician [26]. This algorithm makes new assumptions that differ from Bayes' theorem. As a result, the word naive is utilized.

Bayes' theorem finds the probability of event A occurring when event B has already occurred. which can be found in the equation 2.

$$P(A|B) = \frac{P(A)P(B|A)}{P(B)} \quad (2)$$

where P*(A)* is the probability of event A, and $P(B|A)$ is the probability of event B occurring when event A is known to have already occurred.

NB identifies a class by assuming that each feature's occurrence is independent, resulting in equation 3.

$$P(class|a_1, a_2, \ldots, a_n) = \frac{P(class)P(a_1, a_2, \ldots, a_n|class)}{P(a_1, a_2, \ldots, a_n)} \quad (3)$$

where a_i is the value of any feature that appears in the sample to be categorized, and P(a_1, a_2, ..., a_n) is the probability of occurrence of a feature with value a_1 followed by a feature with value a_2 until the occurrence of features with value a_n.

3.3.3 Support Vector Machines

Another classic technique used in this study is Support Vector Machines. SVM was first heard in COLT-92 in 1992, when Boser, Guyon, and Vapnik introduced it [31]. This method of supervised learning is utilized for classification and regression. It is a classifier that aims to divide the instants into the training data set by finding linear equations that separate the data in each class to achieve high-performance values (optimal hyperplane). From Figure 4, a hyperplane must have the longest distance and width determined from the hyperplane to a class instance (maximum margin).

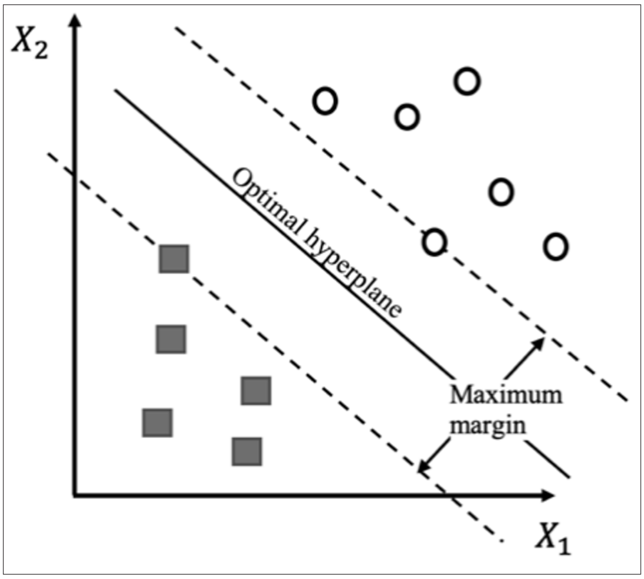


Figure 4. *SVM algorithm for learning on training data with two features.*

SVM attempts to find the support vector at the closest data points from each of the classes [31], which leads to finding the optimal hyperplane that can then be defined in Equation 4.

$$W.X + b = 0 \quad (4)$$

where w represents the weight vector, x is the input feature vector, and b represents the bias. For all components of the training set, w and bwould meet inequalities 5 and 6, respectively:

$$w.x_i + b \geq 1\ ;\ if\ Y_{i=}\ +1 \tag{5}$$

$$w.x_i + b \leq 1\ ;\ if\ Y_{i=}\ -1. \tag{6}$$

where Y_i ϵ{-1,1} In the case in Figture. 5, the support vectors are H_1 and H_2 When the support vector is identified, Equation 7 can be used to get the value that maximizes the margin.

$$Margin = \frac{1}{\|w\|^2} \tag{7}$$

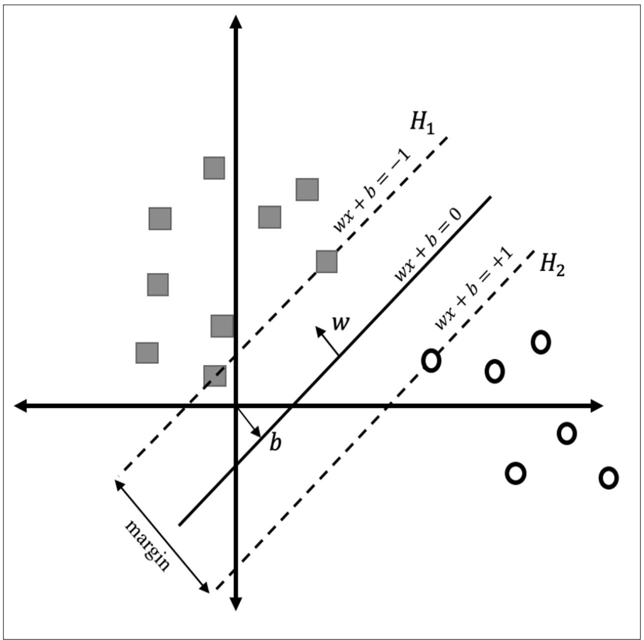


Figure 5. *Support vector.*

3.4 Model Evaluation

The model evaluation utilizes the K-Fold method (K = 10), which is more standardized than the Train-Test split because the model is constructed and evaluated 10 times, mitigating the issue of overfitting and producing accurate results. Precision, Recall, F1-Score, and Accuracy are the most commonly used evaluation metrics for text classification models to assess classification performance. True positives (TP), false positives (FP), true negatives (TN), and false negatives (FN) are used to calculate these metrics. This combination of four numerals generates the confusion matrix depicted in Figure 6.

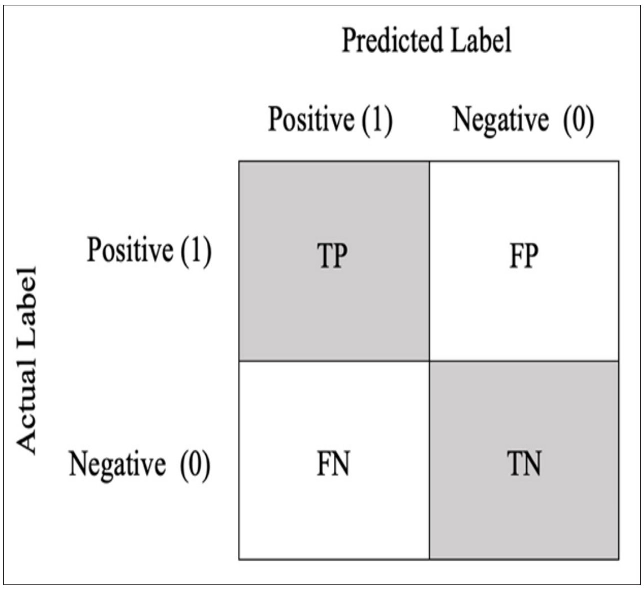


Figure 6. *Confusion matrix.*

Precision is the proportion of correctly predicted values for a particular class relative to the total number of predicted values for that class.

$$Precision = \frac{TP}{TP+FP} \tag{8}$$

Recall is the ratio of all predicted values for a particular class to the actual values for that class.

$$Recall = \frac{TP}{TP+FN} \tag{9}$$

F1-Score achieves a balance between precision and recall. The ratio of precision to recall is optimal.

$$F1 - Score = \frac{2xPrecision\ x\ Recall}{Precision+Recall} \tag{10}$$

Accuracy is a ratio of instances appropriately classified in relation to all values.

$$Accuracy = \frac{(TP+TN)}{TP+FP+TN+FN} \tag{11}$$

3.5 Experimental Results

3.5.1 Experiment 1

We consider Thai word tokenization algorithms, Newmm [32], Longest Matching (LM) [33], and Deepcut [34], to look at their potential because Thai has no word spaces. The result of word splitting is significant because it affects the meaning of words. As a result, we must ensure that the appropriate algorithm is selected. Furthermore, we compare performance with and without stop words to determine whether stop words are required.

We employ both machine learning and deep learning algorithms (NB, SVM, and BiLSTM) to train data. Table 3 displays the results of the three algorithms. The highest accuracy with the BiLSTM algorithm is 55% for Newmm and Deepcut. However, we chose the Newmm approach for data splitting because it splits words faster than Deepcut. When comparing the results of removing or not removing stop words, it appears that not removing stop words has a bit higher results in BiLSTM and NB, thus we decided not to delete stop words in this study.

Table 3. *Performance of the NB, SVM, and BiLSTM models using different preprocessing techniques.*

Classifier	Types of dataset	Accuracy		
		Newmm	LM	Deepcut
BiLSTM	Stop word removal	0.53	0.53	0.54
	Without stop word removal	**0.55**	0.53	0.55
NB	Stop word removal	0.51	0.51	0.49
	Without stop word removal	0.52	0.51	0.51
SVM	Stop word removal	0.55	0.55	0.54
	Without stop word removal	0.53	0.54	0.54

3.5.2 Experiment 2

The objective of this experiment

1. To identify a feature set that can be used to train a model. We consider all three parts: 1. contexts in the question 2. multiple choices 3. length of choices.

2. To compare the performance of the model with different feature sets to improve the model's classification accuracy.

Part 1. Contexts in the question

In previous experiments, only questions were used for training. A question also contains other contexts, such as situations that are textual in addition to the question, images, and tables; however, we didn't train images and tables because we only want to learn text. So, context refers to situations only. We counted the frequency of contexts in each class, which is shown in Table 4, and found that context is not useful for categorizing classes because both classes have almost equal contexts. In the basic thinking skills class question, 64.40% of the contexts were found, which is similar to 57.71% in the advanced thinking skills class. Therefore, contexts will not be trained.

Table 4. *Percentage of contexts appearing in each class's questions.*

Class	Contexts	
	Yes	No
Basic Thinking Skills	64.40%	35.60%
Advanced Thinking Skills	57.71%	42.29%

Part 2. Multiple choices

Each question has multiple choices that may help in question classification because the words in the choices help make the question more complete and clearer about what learning level is needed. Table 5 shows that the question is classified as advanced thinking skills, but it only asks yes or no. However, the objective of the question is not only wants a yes or no answer but also the correct reasons, which are shown in the multiple choice. That is why the question is in the advanced thinking skills class. As a result, we will select multiple choices as one of the features to evaluate the model's performance.

Table 5. *Example of a question that needs Advanced Thinking Skills.*

Context	ภารโรงโรงเรียนแห่งหนึ่ง ตัดกิ่งไม้เพื่อให้ต้นไม้แตกกิ่ง โดยเมื่อตัดกิ่งไม้ที่มีใบแล้ว นำกิ่งไม้วางไว้ในที่ร่มตอนกลางวัน ซึ่งกิ่งไม้นั้นยังสดอยู่ [A school janitor prunes the tree to make the tree branch. After cutting he places them in the shade during the day when the branch is still fresh.]
Question	จากข้อมูล ยังคงมีการสังเคราะห์แสงในใบไม้ หรือไม่ [From the given scenario, does the photosynthesis continue in the leaves?]
Multiple-choice	1. มี เพราะไมโทคอนเดรียเกิดปฏิกิริยาภายในเซลล์ [1. Yes, because there is reaction in mitochondria] 2. ไม่มี เพราะไซโทพลาสซึมไม่เกิดปฏิกิริยาภายในเซลล์ [2. No, because no reaction found in cytoplasm] 3. มี เพราะคลอโรพลาสต์ยังทำงานจึงเกิดกิจกรรมภายในเซลล์ [3. Yes, because chloroplasts are still active. Therefore, there is activity inside the cell.] 4. ไม่มี เพราะแวคิวโอลไม่ทำงานจึงไม่เกิดกิจกรรมภายในเซลล์ [4. No, because vacuoles do not work. So, no activity inside the cell]

Part 3. Length of choices

We also observed that the length of multiple choice can help in classifying classes. Table 6 compares the average word counts of choices and questions in the two classes. In terms of question length, the mean word count for both classes is similar. But in terms of the average word count of choices, the advanced thinking skills class has a higher average word count than the basic thinking skills class, at a difference of 16.83 words. Therefore, we will also train the word count of choices as another feature.

Table 6. *Average word count of choices and questions.*

Class	Average word count	
	Choice	Question
Basic Thinking Skills	20.74	14.23
Advanced Thinking Skills	37.57	15.47

Due to this experiment, we need to use the choices feature, but some of the questions in the dataset contain no choices. Consequently, it is necessary to eliminate those questions. Thus, the dataset was reduced by approximately 20%, leaving a total of 992 questions for training in Table 7.

Then, questions and multiple choices were preprocessed. After tokenizing the words in the choices, we receive the length of the choices in each question and then normalize those values, which is another feature that will be utilized for training.

Table 7. *The class distribution of the dataset after eliminating questions that have no choices.*

Class	Question
Basic Thinking Skills	486
Advanced Thinking Skills	506
Total	992

We use three classifiers for training: BiLSTM, NB, and SVM. We built three models to compare results using different feature sets. Model 1 employs questions with choices; Model 2 employs questions and the length of choices; and Model 3 combines questions, choices, and the length of the choices. The results indicated that all three models improved by approximately 20% from experiment 1. The BiLSTM classifier performed better than the other two classifiers, particularly in model 3. As shown in Table 8, the highest accuracy in this experiment was 0.70. This shows that multiple choices and the length of choices contribute to the model's ability to classify more accurately.

Table 8. *Performance of the model using different feature sets.*

Feature Set	Classifer	Accuracy	Precision	Recall	F1
Questions + Choices	BiLSTM	0.68	0.68	0.68	0.68
	NB	0.65	0.70	0.78	0.66
	SVM	0.59	0.75	0.57	0.49
Questions + Length of choices	BiLSTM	0.67	0.68	0.68	0.67
	NB	0.54	0.67	0.17	0.26
	SVM	0.65	0.77	0.52	0.53
Questions + Choices + Length of Choices	**BiLSTM**	**0.70**	**0.70**	**0.70**	**0.70**
	NB	0.61	0.76	0.41	0.39
	SVM	0.59	0.74	0.56	0.51

3.5.3 Experiment 3

We investigate the impact of POS tagging on feature selection to improve the model's performance in categorizing questions. The addition of multiple choice as a feature increased word size by 39% to 30,769 words. As a result, we wish to eliminate terms that do not affect question classification and select group words that help the classifier categorize questions more accurately.

POS tagging assigns a label for each word with related grammatical elements such as nouns, verbs, adjectives, or adverbs. Previous studies used POS tagging for selecting features to train the model. They found that it helps to improve accuracy [12], [19].

We choose the ORCHID POS tags set , in which the Thai text corpus is over 400,000 words and all word types are labeled, for tagging each word in the dataset. Then, the words in each tag are counted and ranked. The results show that the most POS tagging is NCMN, according to VACT and VSTA, as shown in Table 9.

We create five groups for the POS tag combinations to select features for model training. For example, Group 1 consists of words tagged POS: verb, noun, adposition, and adjective; Group 2 consists of words tagged POS: verb, noun, adposition, adjective, and determiner. Each group chooses a tag from the ORCHID POS tags, as shown in Table 10. The BiLSTM classifier is used for training with different POS tag groups to compare efficiency.

[1] https://pythainlp.github.io/docs/2.3/api/tag.html

Table 9. *Top 10 Occurrences of POS tagging.*

Rank	Part-of-Speech tag	Abbreviation	Word (%)
1	Common noun	NCMN	62.84
2	Active verb	VACT	11.86
3	Stative verb	VSTA	3.93
4	Determiner, cardinal number expression	DCNM	2.59
5	Attributive verb	VATT	2.52
6	Cardinal number	NCNM	2.44
7	Adverb with normal form	ADVN	1.89
8	Unit classifier	CNIT	1.89
9	Preposition	RPRE	1.84
10	Subordinating conjunction	JSBR	1.41

Table 9. *Top 10 Occurrences of POS tagging.*

Part of the Speech tag	ORCHID POS tags	Group 1	2	3	4	5
Verb	VACT, VSTA	✓	✓	✓	✓	✓
Noun	NCMN, CMTR, CNIT	✓	✓		✓	✓
Adposition	RPRE	✓	✓			✓
Adjective	VATT	✓	✓	✓	✓	✓
Determiner	DDAC, DDAN, DIBQ		✓			
Adverb	ADVN			✓	✓	✓
Subordinating Conjunction	JSBR			✓		

The results of POS tag combinations show that group 4, which uses nouns, verbs, adjectives, and adverbs, had the highest performance. As illustrated in Figure 7, the highest was 80.24%. Type 1 has an accuracy of 79.84%, indicating that prepositions are not required as features. For groups 2 and 5, even though more POS tags were chosen, the classification remained lower than group 4. Group 3 demonstrates the importance of nouns as a learning feature because no nouns were chosen, resulting in a lower categorization efficiency than other groups.

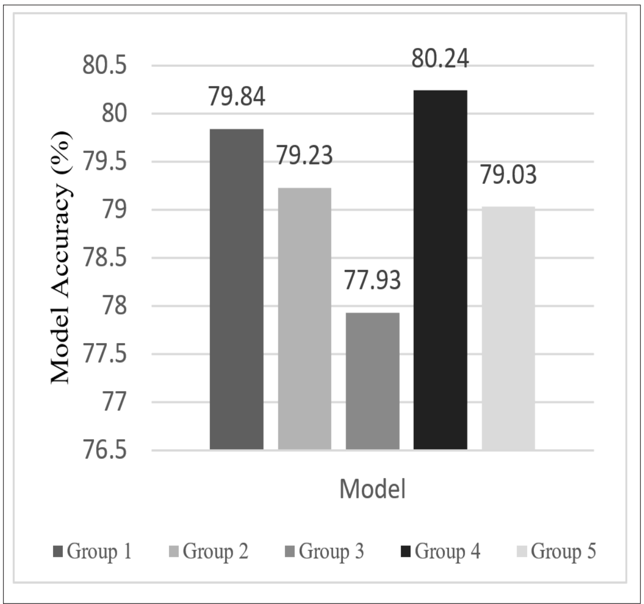


Figure 7. *Performance of feature selection from POS tag combinations in each group.*

As a result, the POS tags chosen as features will be listed in Table 11. The vocabulary has been reduced by 17% from the total number of terms, leaving 3,293 words out of 3,972.

Table 11. *The best group of POS tags as a feature for model training.*

Type	Part of the Speech tag
Nouns	NCMN
Verbs	VACT, VSTA
Adjectives	VATT
Adverbs	ADVN

Table 12 shows the experimental result when compared to the NB and SVM classifiers. The classification performance of the NB and SVM classifiers was similar to that of experiment 2, however, the accuracy of the BiLSTM classifier increased to 80%. This demonstrates that feature selection from POS tagging had an influence on the BiLSTM classifier and improved the model's performance.

Table 12. *The accuracy of classifiers following feature selection with POS tagging includes nouns, verbs, adjectives, and adverbs.*

Classifier	Accuracy	Precision	Recall	F1
Proposed BiLSTM model	**0.80**	**0.81**	**0.80**	**0.80**
NB	0.65	0.77	0.50	0.58
SVM	0.65	0.70	0.65	0.65

3.6 Experimental setting

Experiment 3 provides a suitable classification model for the questions using the BiLSTM classifier. Figture 8 depicts the model's architecture. We experiment several times to select a value for the hyperparameter until a suitable value is found, as shown in Table 13.

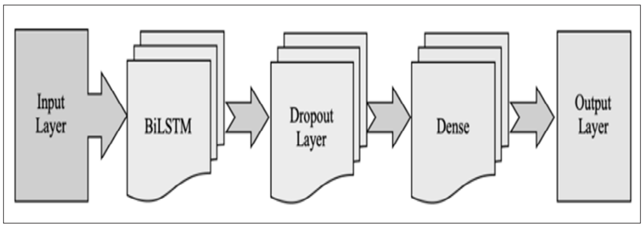


Figure 8. *Architecture of the proposed BiLSTM model.*

Table 13. *The hyperparameters tuned in this research.*

Hyperparameter	Optimal value
Embedding	180
The first Bidirectional-LSTM	10
The second Bidirectional-LSTM	10
epoch	50
Batch size	5
Optimizer	Adam
Activation in a dense layer	Softmax
Rate of the dropout layer	0.5
Learning rate	0.001
Dense output	2

4. Results and Discussion

In this paper, we aim to present a model for efficiently classifying Thai science questions. The results of the three experiments are shown in tables 3, 8, and 12, respectively. We improved the accuracy of the model by adding feature word selection to experiment 3, thereby optimizing the model. Using POS tagging to select features, Word2vec to convert words into vectors, and BiLSTM to classify resulted in the highest accuracy of 0.80, while precision, recall, and F1-scores were 0.81, 0.80, and 0.80, respectively, as shown in Table 12. Table 14 displays the values separated by class.

Table 14. *Class-by-class analysis of the proposed BiLSTM model.*

Class	Precision	Recall	F1-score
Advanced Thinking Skills	0.84	0.75	0.79
Basic Thinking Skills	0.77	0.85	0.81

The confusion matrix of the model presented in this study is depicted in Figture 9. To evaluate the K-fold assessment method, the model selects questions at random so that the class classification accuracy in each fold can be determined and then averaged across all folds (10 folds). The dataset contains 992 questions; there are two folds with 100 questions and eight folds with 99 questions for the test set. The model's classification findings revealed that in the basic thinking skills class, the model correctly classified 85%, while in the advanced thinking skills class, the model correctly classified 75% of the questions from all folds on average.

We test the model with an unseen test set of 36 questions: 27 in the Advanced Thinking Skills class and 9 in the Basic Thinking Skills class. These questions were provided by OBEC. However, the model has never been evaluated with this dataset before. When the model was employed to test question classification, Table 15 and the confusion matrix in Fig. 10 reveal that 30 questions were properly identified, with an accuracy of 0.83. The model correctly classified 24 questions from the Advanced Thinking Skills classification, resulting in 89%, and 6 questions from the Basic Thinking Skills classification, representing 67%.

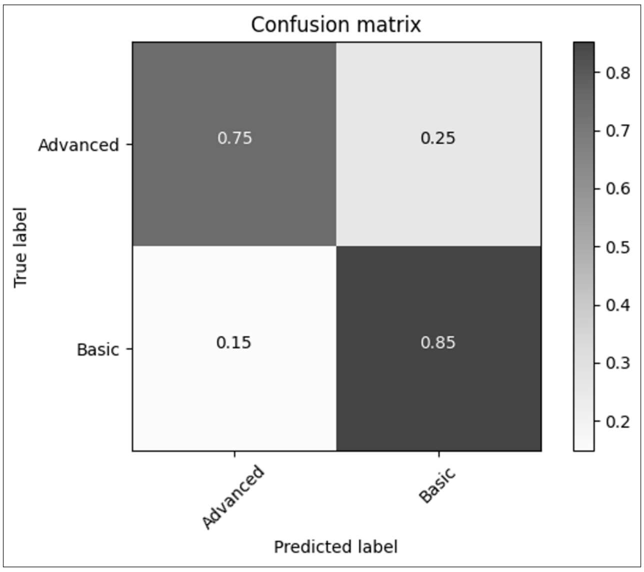


Figure 9. *Confusion matrix of the proposed model for data testing in the K-fold method.*

Table 15. *The proposed model's performance for unseen data testing.*

Class	Accuracy	Precision	Recall	F1-score
Advanced Thinking Skills	0.83	0.89	0.89	0.89
Basic Thinking Skills		0.67	0.67	0.67

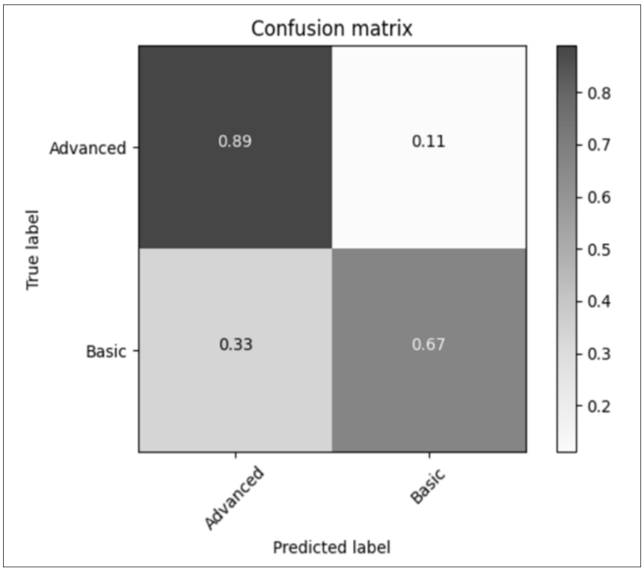


Figure 10. *Confusion matrix of the proposed model for unseen data testing.*

5. Conclusion

This research was conducted to develop an efficient model for classifying secondary school science questions into basic and advanced thinking skills based on revised Bloom's taxonomy. We achieved an accuracy of 80% for question categorization by improving the feature set and applying techniques through 3 experiments. The performance of the model was evaluated by an accuracy metric using 10-fold cross-validation. The results demonstrated that multiple choices influence the model's training. When the choices and the length of words in the choices were trained alongside the questions, the model learned more effectively, and the selection of words in the choice also affected the model's learning. In this study, the POS tagging method was used to tag POS terms such as nouns, verbs, adjectives, and adverbs, resulting in improved model performance. Previous research also used POS tagging for feature selection [9], [12], [19] showing that feature selection can increase classification performance [8], [9]. In addition, the BiLSTM algorithm was found to have better classification efficiency compared to Naïve Bayes and SVMs, resulting in an optimal model using BiLSTM to classify questions. This is consistent with previous work that used BiLSTM to classify Thai text [24], [27] showing the BiLSTM algorithm's high efficiency in classifying Thai text.

In the future, it may be possible to increase the size of the larger dataset so that the model can be classified as closely to reality and standardized as possible, as well as incorporate additional deep learning techniques for additional experiments, such as combining CNN with BiLSTM to produce a more

6. Rererences

[1] World Economic Forum, *Future of Jobs Report 2023.* Available Online at https://www.weforum.org/reports/the-future-of-jobs-report-2023, accessed on 30 September 2023.

[2] L.O. Wilson. "Anderson and Krathwohl–Bloom's taxonomy revised." *Understanding the New Version of Bloom's Taxonomy*, 2016.

[3] J. Irvine. "A Comparison of Revised Bloom and Marzano's New Taxonomy of Learning." *Research in Higher Education Journal,* Vol. 33, 2017.

[4] K. Changwong, A. Sukkamart, and B. Sisan. "Critical thinking skill development: Analysis of a new learning management model for Thai high schools." *Journal of International Studies,* Vol. 11, No. 2, pp. 37-48, 2018.

[5] S. K. Patil and M. M. Shreyas. "A Comparative Study of Question Bank Classification based on Revised Bloom's Taxonomy using SVM and K-NN." *2017 2nd International Conference On Emerging Computation and Information Technologies (ICECIT),* Tumakuru, India, pp. 1-7, 2017.

[6] S. Shaikh, S. M. Daudpotta, and A. S. Imran. "Bloom's Learning Outcomes' Automatic Classification Using LSTM and Pretrained Word Embeddings." *IEEE Access,* Vol. 9, pp. 117887-117909, 2021.

[7] M. Forehand. "Bloom's taxonomy." *Emerging perspectives on learning, teaching, and technology,* Vol. 41, No. 4, pp. 47-56, 2010.

[8] N. Ghalib and D. S. Hammad. "Classifying Exam Questions Based on Bloom's Taxonomy Using Machine Learning Approach." *Technologies for the Development of Information Systems (TRIS-2019),* pp. 260-269, 2020.

[9] J. Chandra and B. Thomas. "The Effect of Bloom's Taxonomy on Random Forest Classifier for cognitive level identification of E-content." *2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE),* pp. 1-6, 2020.

[10] M. Mohammed and N. Omar. "Question Classification Based on Bloom's Taxonomy Using Enhanced TF-IDF." *International Journal on Advanced Science, Engineering and Information Technology,* Vol. 8, No. 4-2, pp. 1679-1685, 2018.

[11] M. Mohammed and N. Omar. "Question classification based on Bloom's taxonomy cognitive domain using modified TF-IDF and word2vec." *PLOS ONE,* Vol. 15, No. 2, 2020.

[12] J. Huang, Z. Zhang, J. Qiu, L. Peng, D. Liu, P. Han, and K. Luo. "Automatic Classroom Question Classification Based on Bloom's Taxonomy." *Proceedings of the 13th International Conference on Education Technology and Computers,* pp. 33-39, 2022.

[13] Hasmawati, A. Romadhony, and R. Abdurohman. "Primary and High School Question Classification based on Bloom's Taxonomy." *2022 10th International Conference on Information and Communication Technology (ICoICT),* pp. 234-239, 2022.

[14] S. Yilmaz and S. Toklu. "A deep learning analysis on question classification task using Word2vec representations." *Neural Computing and Applications,* Vol. 32, No. 7, pp. 2909-2928, 2020.

[15] M. O. Gani, R. K. Ayyasamy, A. Sangodiah, and Y. T. Fui. "Bloom's Taxonomy-based exam question classification: The outcome of CNN and optimal pre-trained word embedding technique." *Education and Information Technologies,* Vol. 28, pp. 15893-15914, 2023.

[16] M. Ifham, K. Banujan, B. T. G. S. Kumara, and P. M. A. K. Wijeratne. "Automatic Classification of Questions based on Bloom's Taxonomy using Artificial Neural Network." *2022 International Conference on Decision Aid Sciences and Applications (DASA),* pp. 311-315, 2022.

[17] M. O. Gani, R. K. Ayyasamy, S. M. Alhashmi, A. Sangodiah, and Y. T.Fui. "ETFPOS-IDF: A Novel Term Weighting Scheme for Examination Question Classification Based on Bloom's Taxonomy." *IEEE Access,* Vol. 10, pp. 132777-132785, 2022.

[18] J. Zhang, C. Wong, N. Giacaman, and A. L. Reilly. "Automated Classification of Computing Education Questions using Bloom's Taxonomy." *Proceedings of the 23rd Australasian Computing Education Conference,* pp. 58-65, 2021.

[19] K. Anekboon. "Feature Selection for Bloom's Question Classification in Thai Language." *Proceedings of the 2018 Computing Conference,* Vol. 1, pp. 152-162, 2019.

[20] C. Chootong and J. Charoensuk. "Cognitive level classification on information communication technology skills for blog." *International Journal of Electrical and Computer Engineering,* Vol. 12, No. 6, pp. 6387-6396, 2022.

[21] T. S. N. Ayutthaya and K. Pasupa. "Thai Sentiment Analysis via Bidirectional LSTM-CNN Model with Embedding Vectors and Sentic Features." *2018 International Joint Symposium on Artificial Intelligence and Natural Language Processing (iSAI-NLP),* pp. 1-6, 2018.

[22] C. Jitboonyapinit. "Development of Sentiment Analysis Model Based on Thai Social Media Using Deep Learning Techniques." *Huachiew Chalermprakiet Science and Technology Journal,* Vol. 2, 2022.

[23] S. Khruahong, O. Surinta, and S. C. Lam. "Sentiment Analysis of Local Tourism in Thailand from YouTube Comments Using BiLSTM." *Mult-disciplinary Trends in Artificial Intelligence,* pp. 169-177, 2022.

[24] P. Klairith and S. Tanachutiwat. "Thai Clickbait Detection Algorithms Using Natural Language Processing with Machine Learning Techniques." *2018 International Conference on Engineering, Applied Sciences, and Technology (ICEAST),* pp. 1-4, 2018.

[25] S. S. Birunda and R. K. Devi. "A review on word embedding techniques for text classification." *Innovative Data Communication Technologies and Application: Proceedings of ICIDCA,* pp. 267-281, 2021.

[26] W. Phatthiyaphaibun, K. Chaovavanich, C. Polpanumas, A. Suriyawongkul, L. Lowphansirikul, and P. Chormai. *PyThaiNLP: Thai Natural Language Processing in Python,* Available Online at https://pythainlp.github.io/docs/2.3/api/tag.html, accessed on 29 September 2022.

[27] T. Kurita. "Principal component analysis (PCA)." *Computer Vision: A Reference Guide,* pp. 1-4, 2019.

[28] S. Jayalakshmi and A. Sheshasaayee. "Question Classification: A Review of State-of-the-Art Algorithms and Approaches." *Indian Journal of Science and Technology,* Vol. 8, 2015.

[29] A. Chugh, *Deep Learning | Introduction to Long Short Term Memory.* Available Online at https://www.geeksforgeeks.org/deep-learning-introduction-to-long-short-term-memory, accessed on 10 August 2022.

[30] A. Taparia, *Bidirectional LSTM in NLP.* Available Online at https://www.geeksforgeeks.org/bidirectional-lstm-in-nlp/, accessed on 5 August 2022.

[31] V. Jakkula. "Tutorial on support vector machine (svm)." *School of EECS, Washington State University*, Vol. 37, No. 2.5, pp. 3, 2006.

[32] K. Jearanaitanakij, N. Kueakool, P. Limwanichsin, T. Kullawan, and C. Yongpiyakul. "LCS-based Thai Trending Keyword Extraction from Online News." *Naresuan University Engineering Journal,* Vol. 17, No. 2, pp. 54-61, 2022.

[33] P. Prakrankamanant. *Data augmentation for Thai natural language processing using different tokenization,* M.S. Thesis, Chulalongkorn University, Bangkok, Thailand, 2021.

[34] M. Jaiwai, K. Shiangjen, S. Rawangyot, S. Dangmanee, T. Kunsuree, and A. Sa-nguanthong. "Automatized educational chatbot using deep neural network." *2021 Joint International Conference on Digital Arts, Media and Technology with ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunication Engineering,* pp. 5-8, 2021.

Green Space Network in Nakhonratchasima City Municipality, Thailand

Yaowaret Jantakat* and Kittikhun Srithumma**

Received: October 18, 2023
Revised: August 12, 2024
Accepted: August 16, 2024

* Corresponding Author: Yaowaret Jantakat, E-mail: yaowaret.ja@rmuti.ac.th

DOI: 10.14416/j.it/2024.v2.004

Abstract

Urban green space still is one critical role for developing sustainability of cities that well-planed and well-designed green networks can create attractive settings for daily life, distinctive local identities for places and can help guide future settlement growth. Therefore, this study aims to plan green space network (GSN) or urban greenways' network in Nakhonratchasima City Municipality (NCM), Thailand. This GSN in NCM used GIS techniques such as the least-cost path analysis in ArcGIS, it was analyzed by the selected output of road nodes' and green space nodes' location. As results, the obtained GSN in NCM was found that Mittraphap Road (Highway no.2), cuts through the middle of NCM area as main potential green space location to connect other minor potential green space locations.

Consequently, this obtained GSN can help not only in local decision-making for planning and designing green city of NCM, but also in Thai city municipal areas to use its future approach or case study for their green areas.

Keywords: Green space network, Urban landscape planning, Geographical information system.

1. Introduction

According to 1 of 17 Sustainable Development Goals (SDGs) has set sustainable cities and communities (Goal no.11), focuses on implementing inclusive, resilient and sustainable urban development policies and practices that prioritize access to basic services, affordable housing, efficient transportation and green spaces for all [1]. This SDG goal has been set targets no. 11.7 is relevant to provide access to safe and inclusive green and public spaces, by 2030, provide universal access to safe, inclusive, and accessible, green and public spaces, in particular for women and children, older persons and persons with disabilities [2]. A part of this target stresses the "creating of green public spaces" which cannot be achieved without significantly transforming the way we build and manage our urban spaces [3]. Moreover, this target is proposed by two main indicators [4]: 1) ratio of land consumption rate to population growth rate, at comparable scale (to be developed) and 2) Area of public and green space as a proportion of total city space. Therefore, green space network is a strategic network connecting various habitats and species, urban and rural green spaces to each other and the communities around them that it offers a wide range of social, health economic and environmental benefits [5], for example, planning and design of urban green networks in Stockholm [6], metropolitan Melbourne's open space network [7], the Baltimore green network [8], the connectivity of urban public green space in an empirical study of Wuhan [9], the green network of Madrid community [10], construction of urban green space network in Kashgar city of China [11], and Significance of urban green space network Formysuru city, India [12].

From the reasons and importance mentioned above, this study aims to apply Geographical Information System (GIS) for analyzing green space network from finding the most possible and suitable spaces including ground checking. This study focused on area of Nakhonratchasima City

* Department of Information and Communication Technology, Faculty of Sciences and Liberal Arts, Rajamangala University of Technology Isan.
** Office of Nakhonratchasima City Municipality, Nakhon Ratchasima 30000, Thailand.

Municipality (NCM) in Thailand because report of NCM land use between 2007-2019 [13], [14] were found that there was the least ratio of green areas about 0.16% and ratio of the most urban and built-up areas (about 83.64 - 87.67%). Although there was more NCM green spaces in year 2020 [15], was about 3.30 km2 or 8.74% of NCM area. However, there is still very small areas in NCM. This affect to ratio of using green areas (per person) in NCM because NCM population has increased every year [16]. Moreover, NCM has been facing challenges in terms of biodiversity loss, green space loss and conflicts between green and blue and grey spaces (access to green spaces, such as forests, parks, grasslands, and recreational areas, and blue spaces, such as rivers, oceans, seas, and lakes, improves health and wellbeing [17]). Therefore, this study requires framework of Green Space Network (GSN) in NCM to see how to connect green spaces in a fragmented city as a functional network. Consequently, this obtained results of NCM-GSN framework will help for planning green areas' increasing, restoring and accessing where will be able to provide ecosystem services as functional green infrastructure further.

2. Theoretical background and related academic papers

2.1 Definition and Importance of Green Space Network

This study focuses on green space in urban area and uses definition of De Hass [18]: urban land, partly or completely covered with grass, trees, shrubs, or other vegetation including parks, community gardens and cemeteries, but also rooftop gardens and vertical gardens, meadows and woods, blue-green zone (it is urban water such as ditches, canals, inland waterways and rivers and riverbanks). Moreover, urban green spaces provide many functions in urban context that benefits people's quality of life so there is a wide consensus about the importance and value of urban green spaces in cities towards planning and constructing sustainable or eco-cities of 21st century [19]. One critical strategy ever used for planning green city as Green Space Network (GSN) that would help to reduce green areas' landscape fragmentation and increase the shape complexity of green space patches and connectivity [20]. Generally, GSN is a strategic network connection between artificial and natural areas that is always implemented in green city development to increase, restore, and link green ways for accessibility of convenient all people [19]. Therefore, the study of urban GSN relates to the recreational and the cultural essence of an urban fabric which helps in maintaining and restoring the urban ecology [12]. A fair distribution of the urban green spaces all over the city caters accessibility for each and every inhabitant of the city [19]. Having sufficient public space allows cities and regions to function efficiently and equitably that provides the rights of way required for streets and infrastructure (and their connectivity) as well as the green space necessary for recreation and the provision of ecosystem services [4]. Furthermore, the diagram below illustrates the idea of maximising the benefits an individual space can deliver and the value that can be added by connecting them into a green network [21] as Figure 1.

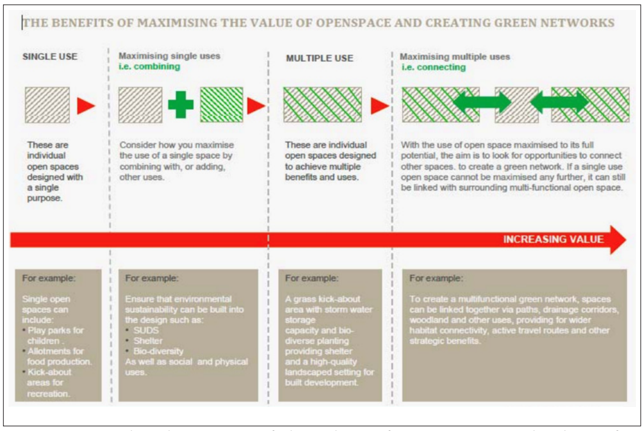


Figure 1. *The diagram of the idea of maximising the benefits an individual space [21].*

2.2 Methods for Green Space Network

There are a variety of approaches and analysis of urban GSN from many related academic papers that are more specifically, the following are referred to be some examples:

- Xiu, Ignatieva and Bosch [6] used the concept of GSN that was expanded into a concrete analytical framework for studying green and blue linkages, as well as social and ecological connections and integrations. This study built a set of criteria to select node patches: 1) the land patches which can be potential habitats in biotope map for crested tit and common toad, and commonly used habitats in sociotype map for humans; 2) large areas of habitats suffer more acutely from landscape fragmentation, so we selected larger habitats as high priority of connection. The minimum area of patches was 1 ha; 3) the patches' central point should be located inside of area because irregular polygons of habitats may center outside the graphs.

- Yashaswini and Shankar [12] analyzed Urban GSN of Mysuru city in India was evaluated with a methodological framework to understand the scenario of GSN in densely built urban area. The urban green space per capita and accessibility of all the present green spaces are studies to know its functionality. As it is well known fact that the quality of life, human well-being are the important factors influenced by urban green and open spaces accessibility and also helps in improving the strategic lifestyle in the urban areas. Eventually, contributions from UGS to the urban areas are very broad and multidimensional like, social, cultural, recreational, economic, environmental and aesthetical.

- Mougiiakou and Photis [22] proposed methodological framework includes accessibility of residents to large green spaces, in relation to population density of blocks. It is important to identify the blocks that meet the fewer criteria, according to Accessible Natural Greenspace Standards (ANGSt), or those that are most populated and have less access to quality green, with an easy and quick way. At the same time, the output of initial evaluation can directly be used in the next step, as a criterion for the formation of the cost raster. In this way, blocks' degraded parts can be improved from the green corridors. The formation of a cost raster, which will include the whole number of the criteria (ecological, environmental, urban, and bioclimatic), will finally lead to the creation of connection paths; not only to the formation of green corridors but also green spaces (existing and new) and of the - necessary - stepping stones and to the interconnected smaller patches.

- Comber, Brunsdon and Green [23] used a network analysis in GIS to access for different religious and ethnic groups, was compared with benchmark standards that form part of the UK government guidance on greenspace provision. Whilst the specific results are locally important (Indian, Hindu and Sikh groups were found to have limited access to greenspace in the city), the study shows how a GIS-based network analysis in conjunction with statistical analysis of socio-economic data can be used to analyze the equity of access to community goods and services.

- Saura and Pascual-Hortal [24], He and his team [25] and Scotland's Nature Agency [26] studied a habitat connectivity index for measuring functional connectivity. The metric used to calculate it is the Equivalent Connected Area (Probability of Connectivity or PC). The index is a value that allows us to see change over time in each region. Its calculation has considered the habitat area, the size and number of patches of habitat, how the habitat patches are arranged in the area under consideration, and the effect the different elements of the landscape has on species movements between habitat patches. Habitat connectivity has been shown as a percentage of the total habitat area as this allows a rapid visual comparison of the degree of connectivity in relation to the amount of habitat present.

3. Methodology

3.1 Study area

Nakhonratchasima City Municipality (NCM) is selected as case study for this study (Figure 2) where covers 37.78 sq.km. The study area is in the administrative bound of Ni-Muang sub-district, Muang district, Nakhon Ratchasima province in Thailand. These areas establish in northern east region of Thailand and located in Southeast Asia of Equator from 14°56′-15°00′N to 102°01′-102°08′E.

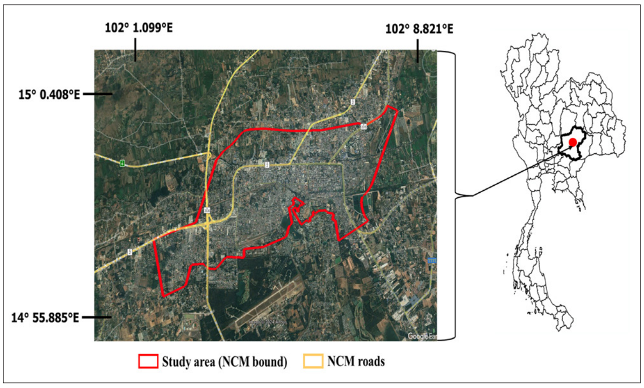


Figure 2. *Nakhonratchasima City Municipality (NCM).*

3.2 Materials and Methods

This study collected available GIS data and sources for analyzing green space network in NCM as Table 1.

Table 1. *Data collection and sources.*

No.	GIS data collection	Sources
1	Land use data in year 2021	Land Development Department (LDD)
2	Road data	NCM office
3	Green space data	
4	Others supplementary data e.g., building data	

Remark: - All data collection were checked and updated by true ground checking and data of Google Earth program in year 2023.

Later, this study set the process for planning GSN based on connection of suitable spaces (road and green space locations). In this analysis, we used GIS techniques in ArcGIS as Figure 3 and more details as follows:

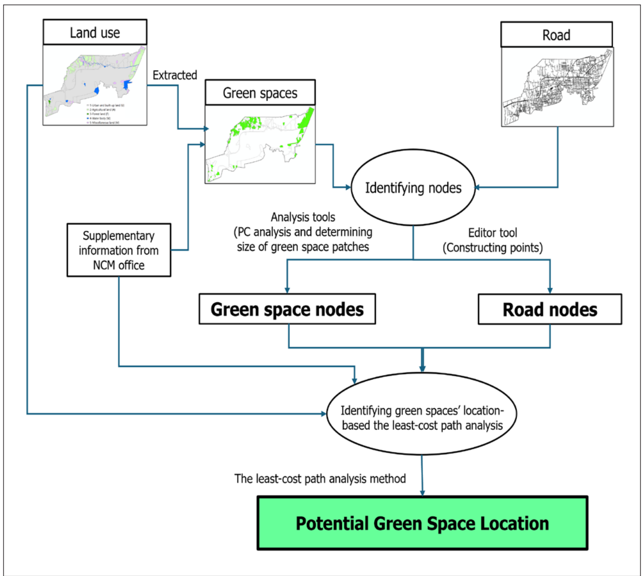


Figure 3. *The process used to plan green space network in NCM.*

3.2.1 Creating Nodes for Roads and Green Spaces

1. Road nodes

Road nodes were created by using Editor menu (Constructor Points) to generate points along a road line in every 100 m. And then a destinating point were extracted to use in GSN analysis because this study focused on road accessibility and easement of traveling based on Mougiiakou and Photis [22] and perspective of ecosystem service in NCM context. This suitable road locations were used for finding the most possibility for GSN in NCM to integrate with green space location in the next topic.

2. Green space nodes

Green space nodes were created by using probability of connectivity (PC). Generally, PC is based on a probabilistic connections model, in which the dispersal probability p_{ij} characterizes the feasibility of a step between patches i and j, where a step is defined as a direct movement of a disperser between two habitat patches without passing by any other intermediate habitat patches. This is PC equation as below:

$$PC = \frac{\sum_{i=1}^{n}\sum_{i=1}^{n} a_i.a_j.p_{ij}^{*}}{A_L^2} \qquad (1)$$

Where n is the total number of habitat nodes in the landscape, a_i and a_j are the attributes of nodes i and j, A_L is the maximum landscape attribute, and $p*_{ij}$ is the maximum product probability of all paths between patches i and j. PC ranges from 0 to 1 and increase with improved connectivity.

The coefficient of variation (CV) is used to determine the threshold where the difference in the most stable importance index levels is most pronounce and to determine the final study threshold. Its formula is [27]:

$$CV = \frac{S}{\bar{dx}} \qquad (2)$$

Where S is the standard deviation of the importance value of the patch connectivity index, and $\bar{dx}$ is the mean of the importance value of the connectivity index. A more considerable CV value indicates a better dispersion of the importance value of the green space patches under this threshold and a more noticeable difference in rank, making it easier to select green space patch nodes.

3.2.2 Identifying Potential Green Space Location

In this paper, the least-cost path analysis method was used by calculating the cumulative cost from source to destination nodes of green spaces and roads based on the connectivity or patch area throughout the landscape. This analysis can be used to identify habitat connections that will preserve or improve connectivity [28] - [30] and then are determined by the minimum cost path between ecological nodes in a network of green spaces and identify potential green spaces [31].

Moreover, this study modified an important level of land use type from Li et al. [11] for supporting identification of potential green space location as shown in Table 2. This table presents the range of the important values of land use type that this study gave the highest weight on road factors because we focus on convenient access to travel by using vehicles and walking. Greenland, construction and others were set as the ordered important level of green space network in NCM, respectively.

Table 2. *Land use type classification and important weights for green space network in NCM.*

Factors	Description	Range of Impedance Values
Road	Available NCM-road	4
Greenland	Forest, agriculture, gardens, lawn, and others green spaces (little vegetation is present, though some areas have some planted vegetation (primarily shrubs and grasses)	3
Urban and built-up land	Residences, public facilities, municipal utilities, warehouses, industries.	2
Other	Lands are used for water, agriculture, abandon etc.	1

Remark: Modified from Li et al. [11]

4. Results and Discussions

4.1 Distribution of The Existing Green Spaces

The random sampling-based true ground checking on 19-25 October 2013 for NCM land use as Figure 4A. The result of kappa coefficient showed 89.90%. And then this NCM land use was particularly extracted as GIS layer of green spaces including forest, agriculture and gardens as Figure 4B). The existing NCM land use included forest land (0.06 km2 or 0.16%), agriculture land (3.24 km2 or 8.58%), urban and built-up land (33.12 sq.km. or 87.67%), water body (1.04 sq.km. or 2.75%) and miscellaneous land (0.32 sq.km. or 0.85%). This NCM land use output showed that NCM green spaces (e.g., forest and agriculture in NCM) is still very small size about 3.30 km^2 (3,300,000 m^2) or 8.74% of NCM area. When it was compared to the minimum standard set by World Health Organization (WHO) of 9 m^2 green open space per city dweller [32], NCM green space (28.46 m^2 per city dweller) is still under such minimum WHO standard. However, we shouldn't neglect green spaces that they should preserve and find ways to increase green space.

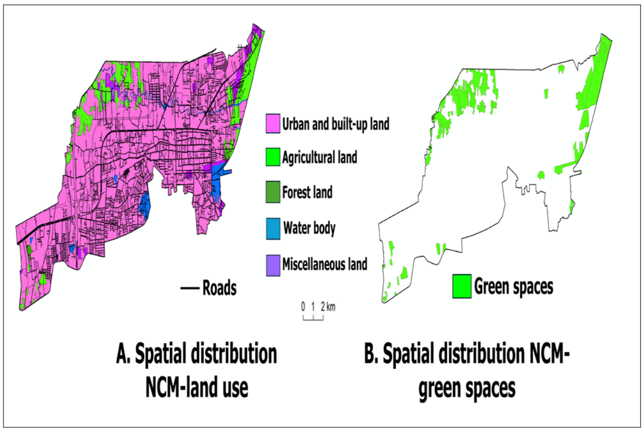


Figure 4. *Spatial distribution NCM-land use and green spaces.*

4.2 Selection of Road Locations

In NCM area, there are 31 roads, include 3 highways and 28 urban roads. This study created new points along NCM roads using GIS techniques. There was the generated total 4, 835 points that were divided into 3,902 starting points and 933 destinating points along NCM roads (Figure 5A). However, this study found that there are 132 of 933 destinating points (14.15%) where were selected for analyzing potential GSN in NCM (as Figure 5B). These 132-selected destinating point based on criteria: road accessibility, easement of traveling and ecosystem services. Similarly, research of Marshall et al. [33] considered criteria in such factors to really benefit for city people. Moreover, we integrated perspective of ecosystem service in NCM context to select road nodes that related research of

Stanford et al. [34] gave importance to key social ecological characteristics.

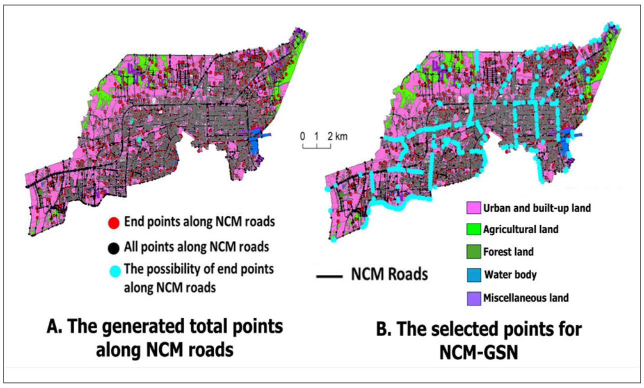


Figure 5. End points and total points along NCM roads.

4.3 Selection of Green Space Locations

Based on Figure 4B, there were 40 green space patches with an area range from minimum (700 m^2 or 0.01 km^2) to maximum (770,000 m^2 or 0.77 km^2). Therefore, the number of components (or connected region) was 40 when the threshold value was 0.01 km^2 and was 1 when threshold value was 0.77 km^2. Therefore, the threshold for connectivity was set at 0.01-0.77 km^2, with 0.01 gaps. And then Spearman' rank correlation coefficient was used for analyzing changes of the connectivity threshold. In this analysis, the probability of connectivity was determined the most stable when the threshold changes. After identifying probability of connectivity as a criterion for selecting nodes in a greenfield patch. We calculated its coefficient of variation as shown in Figure 6, when the threshold value is 0.01 km., the coefficient of variation (CV) of the significant value of probability of connection was the greatest (2.75), and the rank difference between patches in the research region was the greatest and most appropriate for selecting green space patch nodes as Figure 6. Consequently, there were 19 of 40 green patch nodes (47.50%) where were selected for analyzing potential GSN in NCM (as Figure 7). This obtained output is related to output of Li et al. [11] and Yashaswini and Shankar [12], have CV of the significant value of probability of connection in the same direction although distance thresholds are different each other.

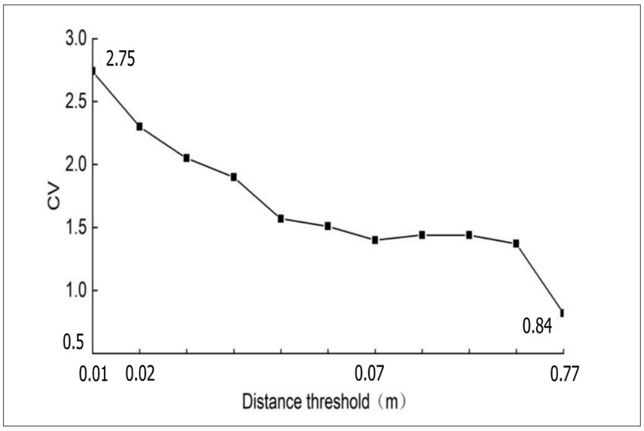


Figure 6. Coefficient of variation under different distance thresholds.

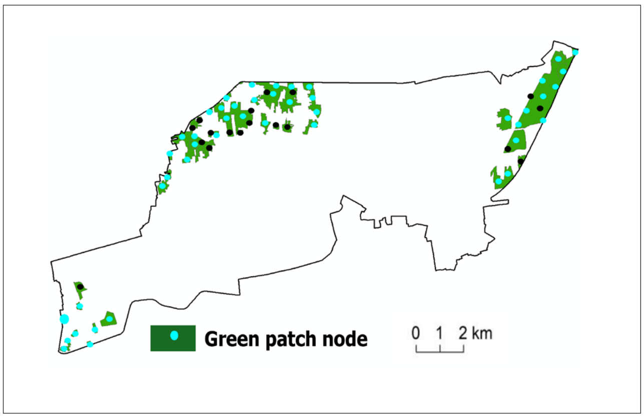


Figure 7. Spatial distribution of the selected green patch nodes in NCM.

4.4 Potential Green Space Network (GSN)

The potential GSN was selected by analyzing the output of road (Figure 5B) and green space locations (Figure. 7) to create GSN in NCM-based the least-cost path analysis. These GSN outputs were found that Mittraphap Road (Highway no.2), cuts through the middle of NCM area as main potential green space location to connect other minor potential green space locations as Figure. 8. This obtained GSN in NCM is likely common network typology of least cost user (see more details from Hellmund [35]). Based on the result of GSN in NCM above, it relates to the output of Li et al. [11], Yashaswini and Shankar [12] and Marshall et al. [33], most obtained GSN is always mainly proposed from middle point to spread around in city areas. Moreover, it is likely GSN output of Kong et al. [29] in project of scenarios for urban green space networks based on the gravity model and graph theory.

Figure 8. *GSN construction in NCM.*

5. Conclusions

This existing NCM land use output showed that NCM green spaces (e.g., forest and agriculture in NCM) are still very small size about 3.30 km^2 (3,300,000 m^2) or 8.74% of NCM area. When it was compared to the minimum standard set by World Health Organization (WHO) of 9 m^2 green open space per city dweller.

For potential GSN-based the least-cost path analysis, it was analyzed by the selected output of road nodes' and green space nodes' location. The output of the obtained GSN in NCM-based was found that Mittraphap Road (Highway no.2), cuts through the middle of NCM area as main potential green space location to connect other minor potential green space locations.

Consequently, this obtained GSN can help not only in local decision-making for planning and designing green city of NCM, but also in Thai city municipal areas to use its future approach or case study for their green areas.

6. Acknowledgement

I appreciate to thank for LDD and office of NCM that gave data for this study and suggestions. Moreover, this paper is a part of the research project, named Comparison of tree biodiversity in urban street using application of geo-informatics technology between Nakhon Ratchasima Municipality (Nakhon Ratchasima, Thailand and Honolulu (Hawaii, USA)' funded by Thailand Science Research and Innovation (TSRI), supported fund under contract No. FRB640019 / NMA / 13 in year 2021. And thanks for Department of Information and Communication Technology (ICT), Faculty of Sciences and Liberal Arts, RMUTI that has supported time and facilities.

7. References

[1] United Nations (UN), *The Sustainable Development Goals Report 2023: Special Edition.* Available Online at https://unstats.un.org/sdgs/report/2023/, accessed on 9 October 2023.

[2] United Nations Human Settlements Programme, *Urban Sustainable Development Goals (SDGs).* Available Online at https://data.unhabitat.org/pages/sdgs, accessed on 9 October 2023.

[3] United Nations (UN), *Sustainable Development Goals.* Available Online at https://www.undp.org/content/undp/en/home/sustainable-development-goals.html, accessed 9 October 2023.

[4] Sustainable Development Solutions Network (SDSN), *Indicators and monitoring framework.* Available Online at https://indicators.report/indicators/i-70/, accessed 9 October 2023.

[5] ABERDEEN City Council, *Green Space Network.* Available Online at https://www.aberdeencity.gov.uk/services/environment/open-spaces/green-space-network, accessed on 9 October 2023.

[6] N. Xiu, M. Ignatieva, and C. K. Bosch, *Planning and Design of Urban Green Networks in Stockholm.* Available Online at https://core.ac.uk/download/pdf/220132708.pdf, accessed on 10 October 2023.

[7] Victorian Planning Authority, *Metropolitan Open Space Network.* Available Online at https://vpa.vic.gov.au/wp-content/uploads/2018/02/Open-Space-Network-Provision-and-Distribution-Reduced-Size.pdf, accessed on 10 October 2023.

[8] City of Baltimore Department of Planning, *The Baltimore Green Network.* Available Online at https://planning.baltimorecity.gov/sites/default/files/GreenNetworkPlan_FullDocument_Reduced_2018_0926%20FN%20and%20KK%20edits%20(1).pdf, accessed on 10 October 2023.

[9] Y. Dong, H. Liu, and T. Zheng. "Does the Connectivity of Urban Public Green Space Promote Its Use? An Empirical Study of Wuhan." *International Journal*

of Environmental Research and Public Health, Vol. 17, No. 1, 2020. doi:10.3390/ijerph17010297.

[10] A. V. Vazquez, E. F. Pablos, R. V. L. Diez, and O. L. Zaldivar. "Green space networks as natural infrastructures in PERI-URBAN areas." *Urban Ecosystem,* Vol. 24, pp. 187-204, 2021. doi:10.1007/s11252-020-01019-w

[11] X. Li, G. Xia, T. Lin, Z. Xu, and Y. Wang. "Construction of urban green space network in Kashgar city of China." *Land,* Vol. 11, No. 10, 2022. Doi:10.3390/land11101826

[12] S. Yashaswini and B. Shankar. "Significance of urban green space setwork for Mysuru city." *Journal of Positive School Psychology,* Vol. 6, No. 2, pp. 462-472, 2022.

[13] Y. Jantakat and P. Juntakut. "Application of geographic information system for monitoring land use changes in Nakhonratchasima City Municipality." *Proceedings of The 5th Regional Conference on Graduate Research,* pp. 217-224, 2020.

[14] Y. Jantakat and P. Juntakut. "Spatio-temporal comparison for studying land use change and diversity in Nakhonratchasima City Municipality, Muang District and Nakhonratchasima Province." *Thai Science and Technology Journal,* Vol. 29, No. 4, pp. 728-743, July-August, 2020.

[15] Land Development Department (LDD), *Land use year 2020.* Available Online at https://dinonline.ldd.go.th/, accessed on 15 October 2023.

[16] Office of Nakonratchasima City Municipality, *Local Development Plan (2023-2027) Nakhon Ratchasima Municipality.* Available Online at https://www.koratcity.go.th/plan2566-2570, accessed on 11 October 2023.

[17] K. L. Ebi and K. Bowen. "Green and blue spaces: crucial for healthy, sustainable urban futures." *The LANCET,* Vol. 401, No. 10376, pp. 529-530, 2023. doi:10.1016/S0140-6736(23)00096-X

[18] W. D. Hass, J. Hassink, and M. Stuiver. "The role of urban green space in promoting inclusion: Experiences from the Netherlands." *Frontiers in Environmental Science,* Vol. 9, July, 2021. doi:10.3389/fenvs.2021.618198

[19] B. C. Bilgili and E. Gokyer, *Urban green space system planning.* Available Online at https://www.intechopen.com/Chapters/37556, accessed on 10 October 2023.

[20] H. Li, W. Chen, and W. He. "Planning of Green Space Ecological Network in Urban Areas: An Example of Nanchang, China." *International Journal Environmental Research and Public Health,* Vol. 12, No. 10, pp. 12889-12904, 2015. doi:10.3390/ijerph121012889

[21] NatureScot, *Green Networks in Development Planning.* Available Online at https://www.nature.scot/sites/default/files/2019-10/SNH%20Information%20note%20-%20Green%20Networks%20in%20Development%20Planning%20-%20March%202012.pdf, accessed on 15 October 2023.

[22] E. Mougiiakou and Y. Photis. "Urban green space network evaluation and planning: Optimizing accessibility based on connectivity and raster GIS analysis." *European Journal of Geography,* Vol. 5, No. 4, pp. 19-46, 2014.

[23] A. Comber, C. Brunsdon, and E. Green. "Using a GIS-based network analysis to determine urban greenspace accessibility for different ethnic and religious groups." *Landscape and Urban Planning,* Vol. 86, No. 1, pp. 103-114, 2008. doi:10.1016/j.landurbplan.2008.01.002

[24] S. Saura and L. P. Hortal. "A new habitat availability index to integrate connectivity in landscape conservation planning: Comparison with existing indices and application to a case study." *Landscape and Urban Planning,* Vol. 83, No. 2, pp. 91-103, 2007. doi:10.1016/j.landurbplan.2007.03.005

[25] M. He, Y. Wu, X. Liu, B. Wu, and H. Fu. "Constructing a multi-functional small urban green space network for green space equity in urban built-up areas: A case study of Harbin, China." *Heliyon,* Vol. 9, No. 11, e21671, November, 2023. doi.org/10.1016/j.heliyon.2023.e21671

[26] Scotland's Nature Agency, *Habitat Connectivity Index.*

Available Online at https://www.nature.scot/professional-advice/land-and-sea-management/managing-land/habitat-networks/habitat-connectivity-index, accessed on 20 October 2023.

[27] J. H. Xu. *Mathematical Methods in Modern Geography.* 2nd Edition, High Education Press, Beijing, 2006.

[28] F. Adriaensen, J.P. Chardon, G. D. Blust, E. Swinnen, S. Villalba, H. Gulinck, and E. Matthysen. "The application of 'least-cost'modelling as a functional landscape model." *Landscape and Urban Planning,* Vol. 64, No. 4, pp. 233-247, August, 2003. https://doi.org/10.1016/S0169-2046(02)00242-6

[29] F. Kong, H. Yin, N. Nakagoshi, and Y. Zong. "Urban green space network development for biodiversity conservation: Identification based on graph theory and gravity modeling." *Landscape and Urban Planning,* Vol. 95, No. 1-2, pp. 16–27, 2010. https://doi.org/10.1016/j.landurbplan.2009.11.001

[30] Z. Wu and H. Wang. "Establishment and optimization of green ecological networks in Yangzhou City." *Chinese Journal of Ecology,* Vol. 34, No. 7, pp. 1976-1985, 2015.

[31] B.J. Goodwin and L. Fahrig. "Effect of landscape structure on the movement behaviour of a specialized goldenrod beetle, Trirhabda borealis." *Canadian Journal of Zoology,* Vol. 80, No. 1, pp. 24-35, January, 2002.

[32] V. S. Singh, D. N. Pandey, and P. Chaudhry. *Urban Forests and Open Green Spaces: Lessons for Jaipur, Rajasthan (India).* Available Online at https://urbanforestrysouth.org/resources/library/citations/urban-forests-and-open-green-spaces-lessons-for-jaipur-rajasthan-india, accessed on 20 October 2023.

[33] A.J. Marshall, M. J. Grose, and N.S.G. Williams. "From little things: More than a third of public green space is road verge." *Urban Forestry & Urban Greening,* Vol. 44, August, 2019. https://doi.org/10.1016/j.ufug.2019.126423

[34] H.R. Stanford, J. Hurley, G.E. Gerrard, and H. Kirk. "Finding the forgotten spaces: Using a social-ecological framework to map informal green space in Melbourne, Australia." *Land Use Policy,* Vol. 141, June, 2024. https://doi.org/10.1016/j.landusepol.2024.107114

[35] P. Hellmund. *Quabbin to Wachusett Wildlife Corridor Study.* Harvard Graduate School of Design, Cambridge, MA., 1989.

Invention The Program of Hands Motion Detector for Translating Disabilities People's Sign Language

Santi Pattanavichai* and Nopparuj Phongtulee*

Received: July 3, 2024
Revised: August 22, 2024
Accepted: August 27, 2024

* Corresponding Author: Santi Pattanavichai, E-mail: pattanavichai@gmail.com

DOI: 10.14416/j.it/2024.v2.005

Abstract

The number of people with disabilities is currently increasing, and in Thai society, the proportion of individuals with disabilities is notably high, approximately 3% of the population (Krungthep Turakij, 2022). Providing support and facilitating accessibility for people with disabilities is of great importance. As part of these efforts, we developed a motion detection program aimed at translating sign language for individuals with disabilities. This program utilizes object detection techniques, such as MediaPipe, in combination with machine learning. It accesses the camera to capture sign language gestures and compares them with models trained on a sign language gesture dataset. The program then displays the corresponding meaning on the screen, with the goal of aiding communication for people with disabilities. The results indicate that the model performs well when tested with a benchmark dataset of 75,000 sign language gestures. However, challenges arise when gestures are unclear, incorrect, not included in the dataset, or closely resemble other gestures, which may lead to misclassification. Additionally, limitations due to insufficiently powerful training equipment have caused a delay in processing and displaying gesture meanings, with a lag time of approximately 5-10 seconds. Despite these challenges, the model achieves an accuracy of 76.40%, which is considered satisfactory. The program is also capable of translating the detected gestures into the Thai language.

Keywords: Hands Motion Detector, Sign Language, Machine Learning, Media Pipe, Datasets, Pillow library, Disabilities People.

* *Faculty of Science and Technology, Rajamangala University of Technology Thanyaburi, RMUTT*

1. Introduction

In today's Thai society has the number of people with disabilities is increasing. For example, in Thai society, the ratio of people with disabilities, it is approximately 3 percent [1] from the all population ratio. This makes facilitating people with disabilities is an important matter and should be developed for communication. Because of the disabled person, it deserves quality and convenience in living and no different from normal people Both in terms of living and understanding of communication.

Disabled people who are hearing impaired and the speech impaired considered a type of person Disabled people. They still have communication barriers because of the disabled person unable to communicate and can understand speech like normal people in society who due to hearing or speaking problems. There has been an invention Sign Language for the disabled or Sign Language for use in communication among the disabled people Sign language can be seen on television programs. or news reporting that must be communicated to people with disabilities to understand. But for normal people in society May not have knowledge understanding of sign language communication for people with disabilities people who is considered a barrier to communication between people with disabilities and other people in society.

Sign language for the disable, it is considered nonverbal. Which is the use of moving the body with the fingers, hands and arms are symbols to convey meaning into various meanings that need to be communicated instead of spoken words. It may have meaning. or must use spelling methods to convey meaning. [2], [3].

The organizing team therefore conducted a study and research on sign language for the disabled using techniques for Object Detection such as Media Pipe combined with machine learning or Machine Learning and developed into a system to translate sign language for the disabled people. To reduce communication barriers between people with disabilities and other people in society.

This paper is organized as follows. Section I introduces the topic and elaborates of a study and research on sign language for the disabled people. Section II describes previous studies related to sign language and machine learning. Section III discusses the proposed the program for using techniques for Object Detection such as MediaPipe combined with machine. Section IV describes the performance of the program for using techniques for Object Detection such as MediaPipe combined with machine. Section and this program can translate to Thai language. Finally, Section V, we discuss the conclusions and proposed future work.

2. Related Work

2.1 Sign language

The system that the group developed each one focuses on recognizing a single sign language gesture parameter [4].

These parameters are the position of the hands. Shape, movement and orientation and non-manual gestures in the form of facial expressions [5], [6].

It's part of a full animation show of animated human characters. Sign language, so humans viewing the animation may be able to identify something that is signed based on the words spoken or the movement of the surrounding arms. In this case, it is the difference between the quality of the hand shape perceptions in the first evaluation study appear less dramatic. When hand shapes are part of the full animation of sign language [7], [8].

2.2 Machine Learning

Data is collected in the form of images to be used for machine learning. This program used object detection technology such as Media Pipe together with machine learning in order to remember and consider the posture of movement. Using OpenCV image processing technology. In order to do Machine learning using KERAS and receiving image values from the camera in order to process and display them in the form character. Popular Machine Learning Algorithms, it is based on biological neurons, which are the foundation of artificial neural networks [9], [10].

2.2.1 Media Pipe It is object detection technology. This program used the translate sign language python language with Media Pipe which it is owned by Google and it is an open source AI platform that can be used as a pipeline for detecting and recognizing complex faces, hands, and gestures. Using acceleration in identification and processing So it came out to be an accurate and fast solution. Now Google AI has taken it to the next level. and is ready to introduce everyone to "Media Pipe Holistic," a solution that will allow devices that can detect multiple parts of the body at the same time to actually be developed [10] - [12].

MediaPipe Holistic uses a trade-off between the three sensing points, and its efficiency depends on the speed and quality of the data exchange. Combining the three sensing points results in a single, cohesive topology that captures 540 + motion keypoints (33 gesture points, 21 points on each hand, and 468 points on the face) at an unprecedented level, and can be processed in near real-time for mobile display.

But that's just the tip of the iceberg. MediaPipe also makes use of OpenCV, a powerful open-source library for computer vision. OpenCV has lots of tools and algorithms for working with images and videos. By using OpenCV, MediaPipe can easily add features like video capture, processing, and rendering to its pipelines. MediaPipe also teams up with TensorFlow, Google's machine learning tool, to make adding pre-trained or custom models easy. This makes tasks like recognizing faces or understanding speech easy. MediaPipe can also support popular languages like C++, Java, and Python, so it's simple to add to your projects.

2.2.2 OpenCV is the image processing technology. It is a Library of Programming functions, mainly aimed at real-time computer vision. It was originally developed by Intel, but later became supported by Willow Garage followed by Itseez (later acquired by Intel). OpenCV is a cross-platform library. (Cross-Platform) and free to use under the Open-Source BSD License [13] - [15]. OpenCV is a library of programming functions used primarily for images processing.

Opencv is a huge open-source library for computer vision, machine learning, and image processing. Now, it plays a major role in real-time operation which is very important in today's systems. By using it, one can process images and videos to identify objects, faces, or even the handwriting of a human.

When it is integrated with various libraries, such as NumPy, python is capable of processing the Opencv array structure for analysis. To Identify an image pattern and its various features we use vector space and perform mathematical operations on these features.

2.2.3 KERAS is the tool for Machine Learning or Deep Learning, we need to choose. Library (Library) to suit our work. In order to make the performance of our model good. KERAS is an open-source neural network written in Python. KERAS is mainly used for doing Deep Learning where the Input (data import) is an image. Because it is ready-made, but it will be a little customizable and it is a finished product. Highly adaptable, KERAS gives every programmer freedom through the integration of low-level machine learning languages such as Tensor Flow or Theano, which means that anything written in the source language may be executed by KERAS [16], [17].

Keras is an open-source deep-learning framework that gained attention due to its user-friendly interface. Keras offers ease of use, ùexibility, and the ability to run seamlessly on top of TensorFlow. In this article, we are going to provide a comprehensive overview of Keras.

Keras is a high-level, user-friendly API used for building and training neural networks. It is designed to be user-friendly, modular, and easy to extend. Keras allows you to build, train, and deploy deep learning models with minimal code. It provides a high-level API that is intuitive and easy to use, making it ideal for beginners and experts alike.

To develop a model that can help with prediction, we can use KERAS or TensorFlow, which are deep-learning libraries. KERAS is a high-level deep learning API developed by Google for implementing neural networks. It is written in Python and is used to simplify the implementation of neural networks, and supports multiple neural network computation backbends, which use a high level of abstraction in the Python frontend. This makes KERAS slower than other deep learning frameworks.

2.2.4 Pillow Library It is the library for translating to Thai Language. The developer encountered a problem with the library that was selected for use in the Image Processing section, OpenCV, which was unable to display the language in the system like UTF-8, causing various display sections to be unable to display Thai language. The developer therefore chose to use the Pillow Library, which is used in the Image Processing section, to work with OpenCV, which is to use OpenCV to access working with the camera and use Pillow to do all the text display work. Therefore, we allowing the system to display Thai language [18], [19].

2.3 Comparisons with Previous Research

In this study, we compare our work with the paper titled Development of Thai Sign Language Detection and Conversion System into Thai with Deep Learning. That paper utilized the MediaPipe framework and Bidirectional Long Short-Term Memory (BiLSTM) to compare the performance of LSTM and BiLSTM models. While it demonstrated that BiLSTM provided better results than LSTM, it did not present an implementation in a programmatic format as our study does.

Our research integrates a broader set of tools, including MediaPipe, OpenCV, LSTM, KERAS, and the Pillow library, to develop a program specifically designed for individuals with disabilities [20]. The primary objective of our research is to create a practical tool that can be used by people with disabilities in real-world scenarios. The combination of these tools enabled us to successfully develop a program that meets this objective.

3. The Working Structure of Program Hands Motion Detector for Translating Disabilities People's Sign Language

3.1 Flow Chart shows the flow of the program

From the flow chart, it shows the sequence of the program's operations when starting to use. The program will receive values from the device's camera through OpenCV and then use the Media Pipe to capture points as defined, including right and left fingers, right and left hands, shoulders, and faces, by capturing the data in Array format. After that, it will use the AI that has been performed. Practice it. to compare with the values currently captured.

After that, if the detected gesture Is it closest to which word? The word will be displayed on the screen and the whole system will be repeated again. If there is still movement. If the user is not moving anymore The program will stop working. and if the user presses to shut down the system The system will shut down immediately shown in Figure 1.

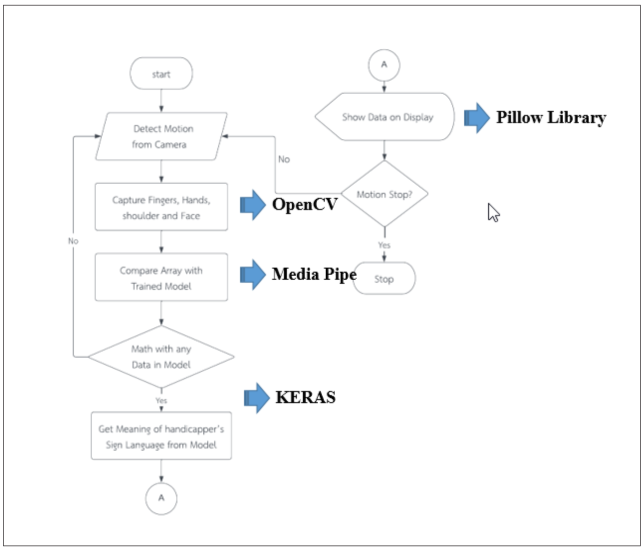


Figure 1. Figure for Flow Chart shows the flow of the program.

From the flow chart, it shows how Machine Learning works. It uses the dataset that the developer has collected to train the AI. The model being built uses LSTM layers to handle sequence data [21], [22]. It is characterized as a "time series" where the model's input is a sequence of data with 30 time steps (number of frames) and 1662 features (number of landmarks at various points to be used in learning) in each time step. The resulting Dense transforms the nature of the data to produce output of the correct class. The activation function used is ReLU (Rectified Linear Unit) [23].

After that, the data set will be divided into 2 sets, the ratio is 80:20, with 80 sets used for learning and 20 used for testing. Then use the Adam algorithm to divide and improve the weights in the train, then train a total of 1,000 times, then export to get the Action.h5 file to use in the next program shown in Figure 2.

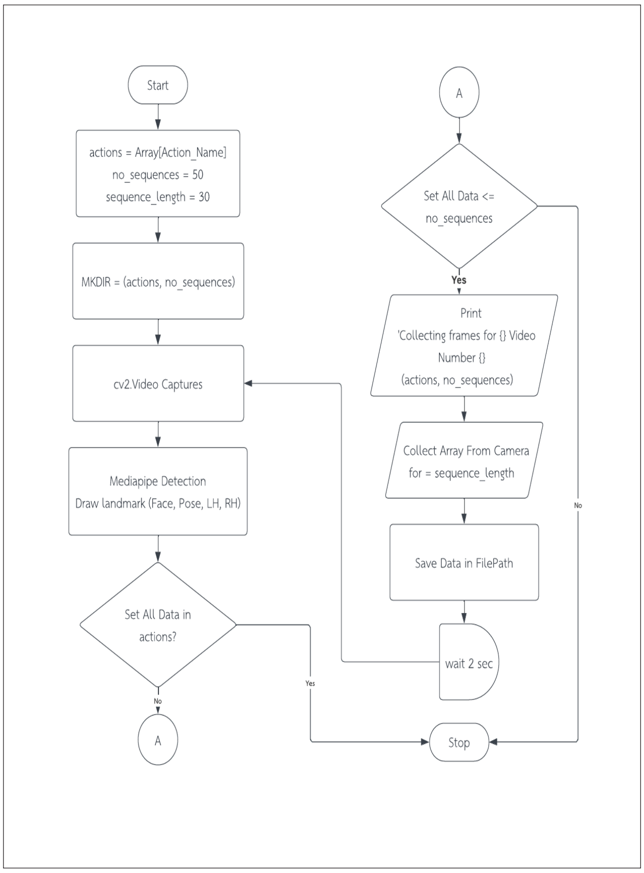


Figure 2. Flow Chart shows the machine learning sequence based on KERAS.

Based on the flow diagram, the developer designed the Dataset storage system, starting with creating variables to store gesture meanings in Array format and creating variables to define the number of cycles. (no_sequences) to a value of 50, meaning that 50 cycles will be stored, 30 frames per cycle (sequence_length). After that, create an actions folder to store the Dataset separated by poses and the number of cycles. The machine learning based the approach typically trains a sentiment classifier using features by KERAS and LSTM [24].

Then use OpenCV to activate the camera and use Media Pipe to detect points on the body, edit the face, shoulder, left hand, and right hand, then create 1,662 points and create an array in the npy extension format.

Then begin collecting the Dataset, sorting them according to the folder of actions, repeating 50 rounds, 30 frames each, and then saving the Array file (.npy) in the order of the frames shown in Figure 3.

From the Context Diagram, it shows the basic data flow that occurs within the system. There is a lot of information from people with disabilities sign language gestures. Then the system will move to search from the dataset then make a comparison. Before using the information shown results through information to users and to communicate with people with disabilities shown in Figure 4.

From the Data Flow Diagram, it shows the flow of data. and various subsystems that occur within the program shown in Figure 5.

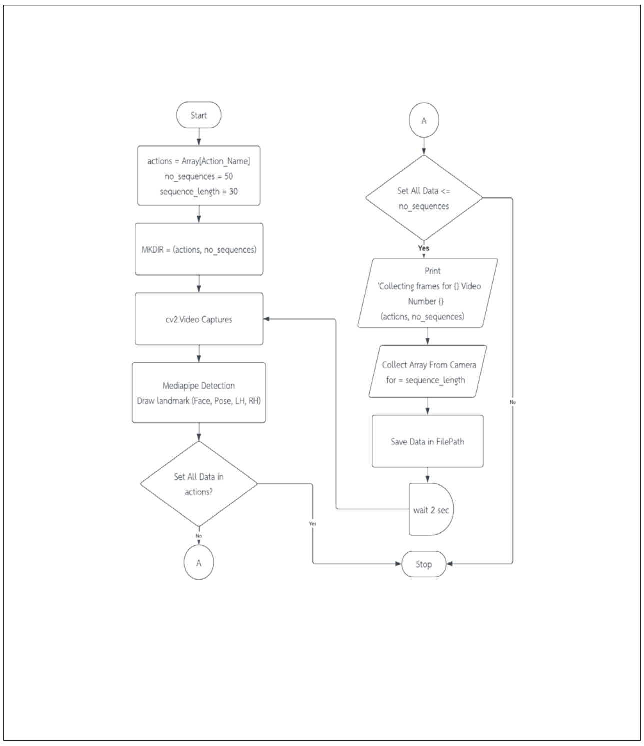


Figure 3. *Flow Chart Dataset Storage System.*

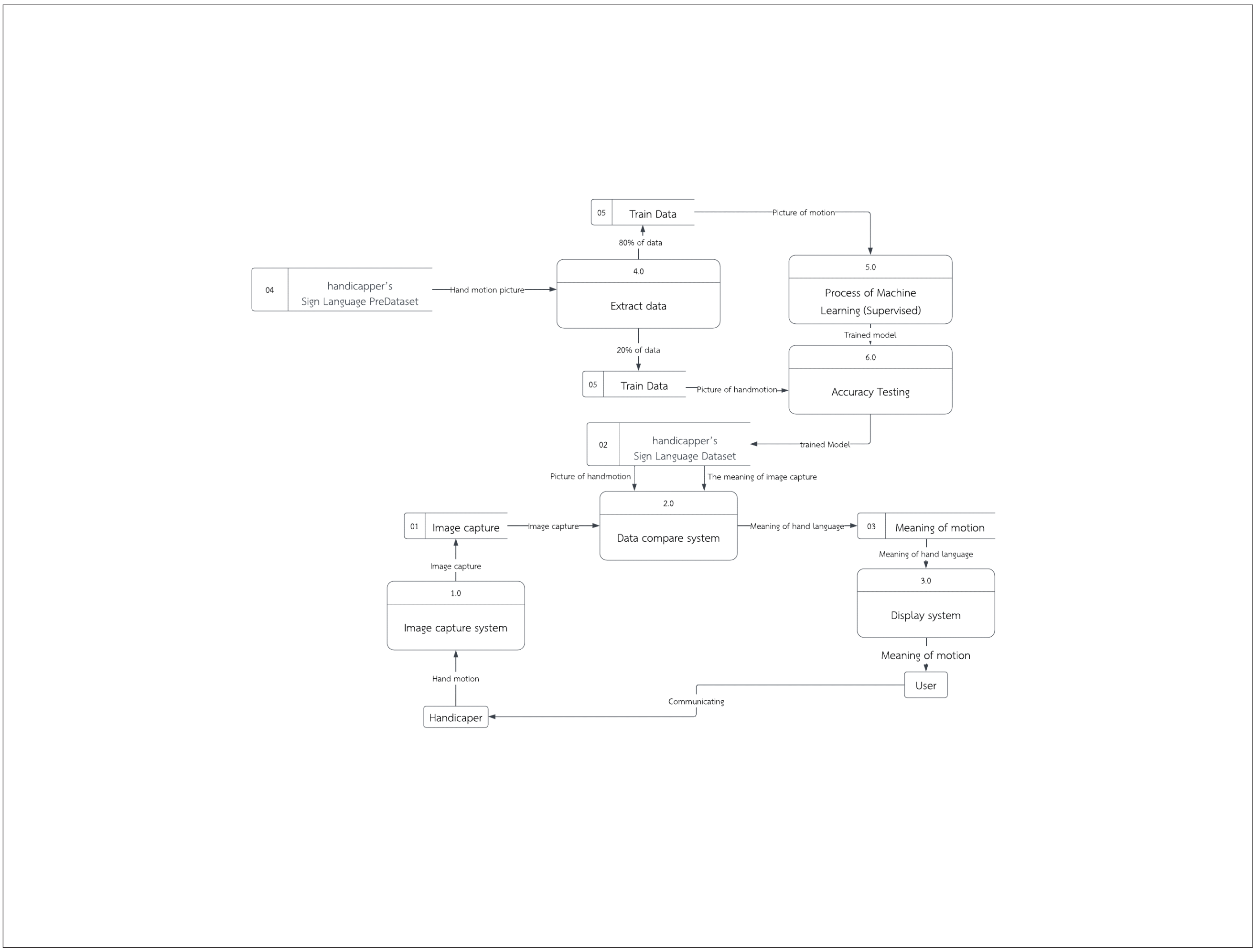


Figure 4. *Data Flow Diagram shows the flow of data in the system.*

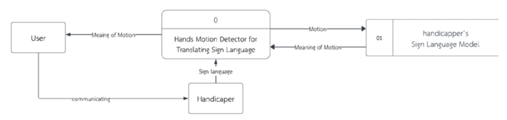


Figure 5. *Context Diagram.*

3.2 Implementation of Datasets

Begin developing the system by collecting a dataset to track hand and body gestures according to the available dataset and using machine learning technology. This program used Datasets Benchmark from MNIST on website https://www.kaggle.com/datasets/datamunge/sign-language-mnist [25]. By giving examples to teach, remember and be able to separate different gestures from each other. In this step, the developer has stored a dataset of 50 sign Thai language words shown in Table 1 that can be used in everyday life, including:

Table 1. *Words stored as a Datasets.*

สวัสดี	ขอบคุณ	ขอโทษ	ใช่	ไม่
สบายดี	ไม่เป็นไร	ชื่อ	นามสกุล	ภาษามือ
ผู้พิการทางการได้ยิน	ผู้มีการได้ยินปกติ	ไม่สบาย	น้ำมูก	หิวน้ำ
เจ็บคอ	อุณหภูมิ	ฉันพบกับคุณ	คุณพลกัน	ไว้พบกันใหม่
โชคดี	วัน	วันจันทร์	วันอังคาร	วันพุธ
วันพฤหัสบดี	วันศุกร์	วันเสาร์	วันอาทิตย์	เดือน
วันนี้	พรุ่งนี้	เมื่อวาน	สัปดาห์	ปี
เวลา	คน	ผู้ใหญ่	ผู้ชาย	ผู้หญิง
เด็ก	ฉัน	คุณ	พวกเรา	อะไร
ทำไม	ที่ไหน	เมื่อไหร่	อย่างไร	หิว

3.3 The source code of the Jupyter Notebook Program

This program is implemented in Python Language on the Jupyter Notebook Program for the functionality of OpenCV, to train dataset on machine learning and the results display section in this program shown in Figure 6, 7, 8.

```
font_path = "Kanit-Medium.ttf"
font_size = 20
font = ImageFont.truetype(font_path, font_size)
mp_holistic = mp.solutions.holistic
holistic = mp_holistic.Holistic(min_detection_confidence=0.5,
min_tracking_confidence=0.5)
round_interval_seconds = 2
for action in actions:
draw_styled_landmarks(image, results)
for sequence in range(no_sequences):
pil_image = Image.new('RGB', (640, 480), color=(255, 255, 255))
draw = ImageDraw.Draw(pil_image)
draw.text((120, 200), 'เริ่มการบันทึกสำหรับ {} วิดีโอที่ {}'.format(action,sequence),
fill=(0, 255, 0), font=font)
initial_message_image = cv2.cvtColor(np.array(pil_image),cv2.COLOR_RGB2BGR)
cv2.imshow('Processed Image', initial_message_image)
cv2.waitKey(2000)
for frame_num in range(sequence_length):
ret, frame = cap.read()
image, results = mediapipe_detection(frame, holistic)
draw_styled_landmarks(image, results)
pil_image = Image.fromarray(cv2.cvtColor(image,cv2.COLOR_BGR2RGB))
draw = ImageDraw.Draw(pil_image)
draw.text((15, 12), 'กำลังบันทึกเฟรมสำหรับ {} วิดีโอที่'.format(action, sequence),
fill=(0, 0, 255), font=font)
keypoints = extract_keypoints(results)
npy_path = os.path.join(DATA_PATH, action, str(sequence),str(frame_num))
np.save(npy_path, keypoints)
result_image = cv2.cvtColor(np.array(pil_image),cv2.COLOR_RGB2BGR)
cv2.imshow('Processed Image', result_image)
if cv2.waitKey(10) & 0xFF == ord('q'):
break
time.sleep(round_interval_seconds)
cap.release()
cv2.destroyAllWindows()
```

Figure 6. *Part of the Dataset collection function.*

```
from keras.layers import Flatten, Dropout
model = Sequential()
model.add(LSTM(64, return_sequences=True, activation='relu', input_shape=(30, 1662)))
model.add(LSTM(128, return_sequences=True, activation='relu'))
model.add(LSTM(64, return_sequences=False, activation='relu'))
model.add(Flatten())
model.add(Dropout(0.25))
model.add(Dense(128, activation='relu'))
model.add(Dense(64, activation='relu'))
model.add(Dense(actions.shape[0], activation='softmax'))
model.compile(optimizer='Adam', loss='categorical_crossentropy', metrics=['categorical_accuracy'])
from keras.callbacks import EarlyStopping, ModelCheckpoint
es = EarlyStopping(monitor = 'val_acc', min_delta = 0.01, patience = 4,verbose = 1)
mc = ModelCheckpoint("realth.keras", monitor = "val_acc", verbose = 1,save_best_only = True)
cb = [es, mc]
model.fit(X_train, y_train, epochs=100, callbacks=[cb], validation_split =0.3)
model.summary()
```

Figure 7. *Train Model section.*

```
sequence = []
sentence = []
predictions = []
thai_font = ImageFont.truetype(thai_font_path, font_size)
with mp_holistic.Holistic(min_detection_confidence=0.5,
min_tracking_confidence=0.5) as holistic:
while cap.isOpened():
image, results = mediapipe_detection(frame, holistic)
print(results)
draw_styled_landmarks(image, results)
keypoints = extract_keypoints(results)
sequence.append(keypoints)
sequence = sequence[-30:]
if len(sequence) == 30:
res = model.predict(np.expand_dims(sequence, axis=0))[0]
print(actions[np.argmax(res)])
predictions.append(np.argmax(res))
if np.unique(predictions[-10:])[0] == np.argmax(res):
if res[np.argmax(res)] > threshold:
if len(sentence) > 0:
if actions[np.argmax(res)] != sentence[-1]:
sentence.append(actions[np.argmax(res)])
else:
sentence.append(actions[np.argmax(res)])
if len(sentence) > 5:
sentence = sentence[-5:]
pillow_image = Image.fromarray(cv2.cvtColor(image,
cv2.COLOR_BGR2RGB))
draw = ImageDraw.Draw(pillow_image)
draw.rectangle([(0, 0), (640, 60)], fill=(245, 117, 16))
draw.text((3, 10), ' '.join(sentence), font=thai_font, fill=(255,
255, 255))# Convert Pillow image back to OpenCV format for display
image = cv2.cvtColor(np.array(pillow_image), cv2.COLOR_RGB2BGR)
cv2.imshow('Sign Language Translator', image)
key = cv2.waitKey(1)
if key & 0xFF == ord('q'):
break
cap.release()
cv2.destroyAllWindows(
```

Figure 8. *Results display section.*

4. The Performance of Program Hands Motion Detector for Translating Disabilities People's Sign Language

4.1 Sign language gesture detection

From the development of a sign language gesture detection system used OpenCV to enable access to the device's camera. Then created a program window to work with the Media pipe to detect points. Various features that have been specified then create colored dots for observation, it is called "Landmark" which the developer has used to detect, including Face or face, Pose or posture around the shoulder, Left Hand or left hand and Right Hand or right hand by setting the confidence in a program window value to 0.5, resulting shown in Figure 9.

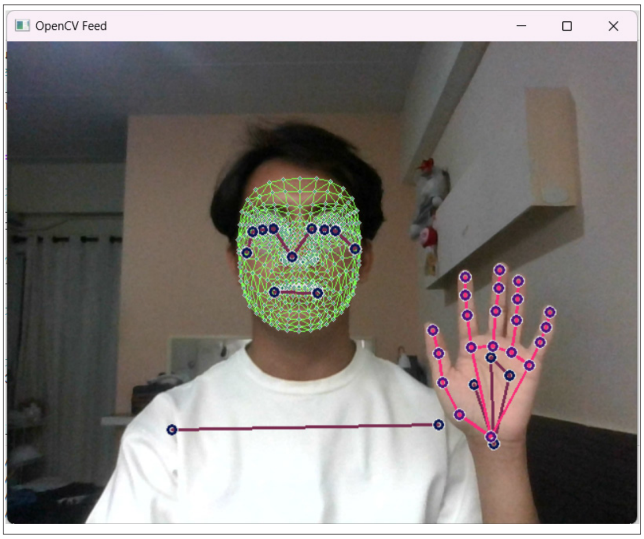


Figure 9. *Train Model section.*

The program made it possible to send images captured from the Video Capture program to check the arrangement of important points again. Through the operation of the Jupyter Notebook program as shown in Figure 10.

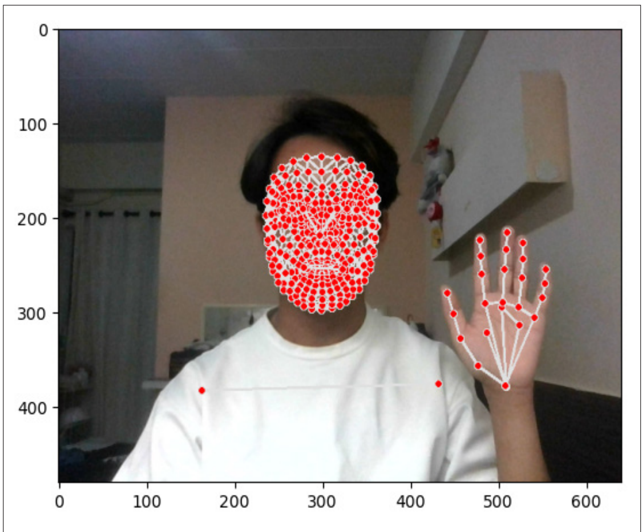


Figure 10. *Image obtained from the gesture detection program.*

After that, a function will be created to extract keypoints from the results obtained from the gesture tracking and face capture systems Media pipe. This function worked by receiving all the results from the Media pipe system and then selecting the keypoints of interest in the form of an array that stores the X, Y, and Z values of various Keypoints and saves them to a file in Array format with the extension .npy. If the Keypoints were not detected from the capture, they will substitute the value 0 into the Array to prevent A NULL value occurred shown in Figure 11.

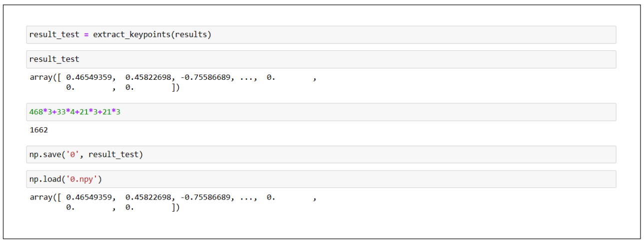


Figure 11. *Example of displaying values in an array stored from a keypoint.*

4.2 Datasets Benchmark collection system

The Datasets Benchmark collection system started by creating a folder to store the array obtained from collecting the Dataset through Python commands. It starts by creating an array according to the name of the pose you want to store. and create a folder to store the number of rounds and the frame that is stored in the system. You will get a folder as shown in Figure 12.

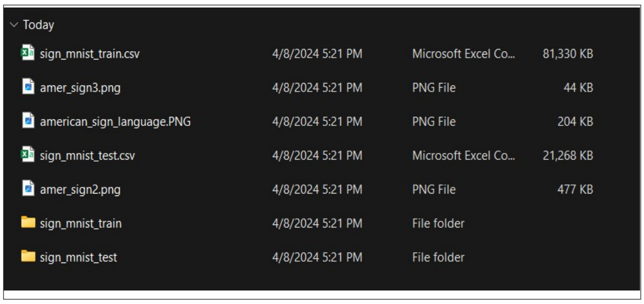


Figure 12. *Folder for storing datasets benchmark.*

The developer chosen to create the folder via command. Because it will allow the system to iterate over the collection of the Dataset according to the array and the values in the variables of the pose name, number of cycles, and number of frames. This making it is possible to collect Datasets continuously, easily and with greater accuracy. rather than creating folders manually.

4.3 Results of training for AI

During the training, a log file will be created to track the results of the training via Localhost:6006 of the Tensor Board. It will be created according to the path of the developed program shown in Figure 13.

The results of the trial found that the model that had been trained had an accuracy of 76.40%, resulting in a model file for use in further development, file name Model.h5.

Figure 13. *Logs recorded during training.*

After that, the collected Dataset, which had 75,000 Array formats (50 words, 50 times each word, 30 frames at a time) will be used in the Long-Short Term Model format, RNN format. To make it possible to recognize patterns for a long time, it is effective for prediction problems. It is sequential because previous data can be collected and used in processing and then trained with KERAS using the Adam algorithm.

In the MNIST dataset benchmark, it is suggested that a training set of approximately 60,000 datasets samples is sufficient. However, in our experiments using a training set of 50 words, we obtained 75,000 datasets samples, achieving an accuracy of 76.40%, which is considered adequate. When the training set size was increased to 100 words, resulting in 150,000 datasets samples, the accuracy improved by an additional 5-10%. Despite this improvement, the increased computational requirements pose a challenge for implementation on smartphone devices, given the current performance limitations.

This will result in a trained model. To be used in programs that detect in real time shown in Figure 14.

Testing: Test the system. By testing sign language translation to be accurate. At not less than 75 percent of all poses, by checking the training results each time through TensorFlow that was accessed through Localhost:6006 which will be displayed through the form of a line graph. This is the accuracy value obtained from each training session shown in Figure 15.

From the work of the model that has been tested, the developer has introduced the model to use with the system in order to show the performance results, it was found that the model could work correctly and accurately, but there was a slowness in processing. This is caused by the limitations of the equipment used to train the model, including its ability to display results. This causes a delay of approximately 5-10 seconds after performing the corresponding gesture. The display format is as shown in Figure 16, 17, 18.

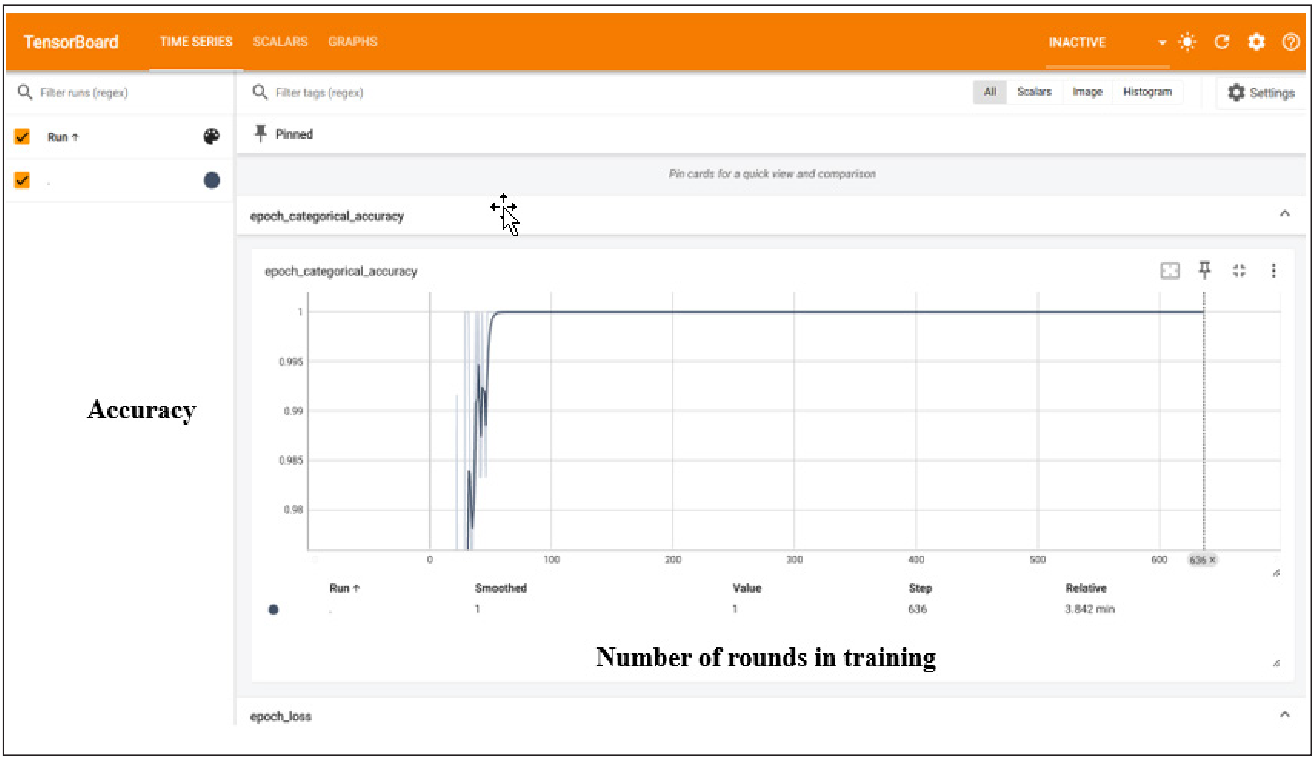


Figure 14. *Line graph showing the accuracy of training.*

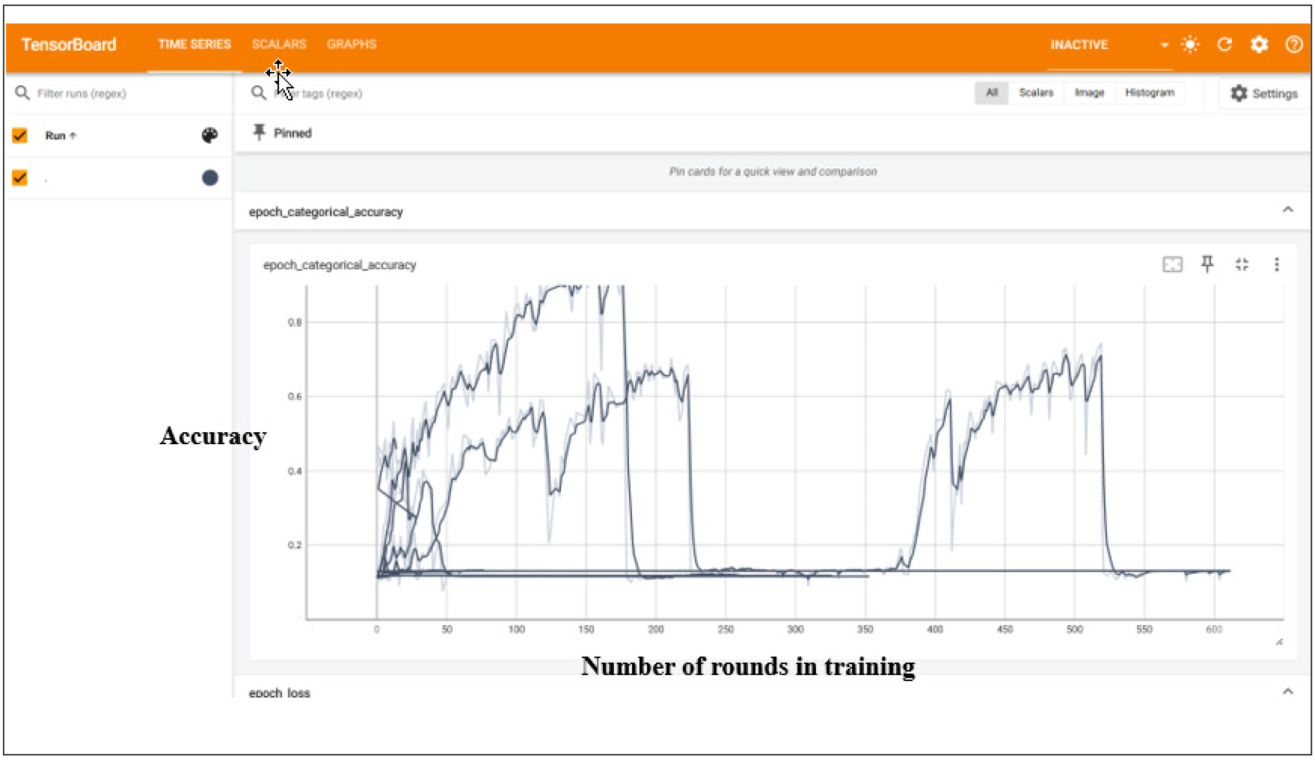


Figure 15. *Line graph showing the accuracy of training.*

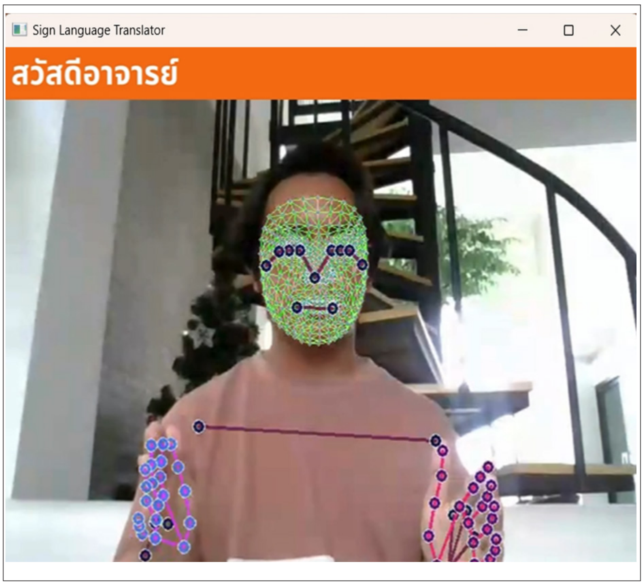


Figure 16. *Image of the display system hello teacher.*

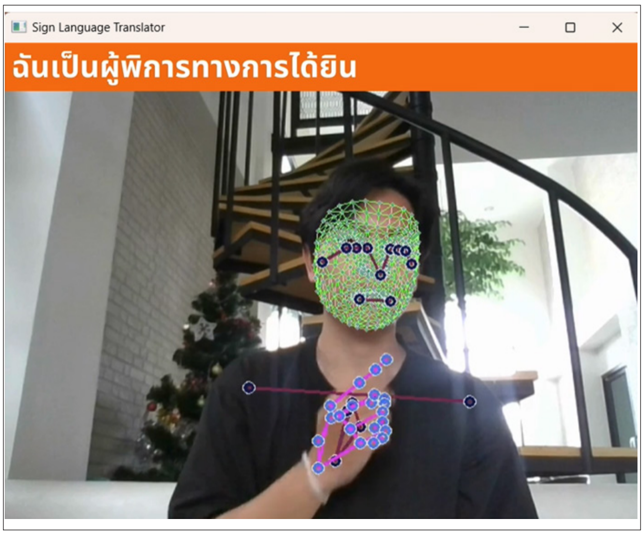


Figure 17. *Image of the display system I am hearing impaired.*

Figure 18. *Example of a dataset collection image (hello, thank you, sorry, and how are you).*

5. Conclusions

From the results of using the Model together with the display section, it was found that the Model can work well. Within the training dataset gesture experiment but problems may be encountered if gestures are unclear, gestures are wrong, gestures that are not in the dataset, and gestures that are similar. It may cause the system to classify the wrong gestures. Including limitations from insufficiently effective training depend on equipment causing the system to be delayed in comparison and display the meaning of the gesture for approximately 5-10 seconds, depending on the difficulty and the similarity of gestures.

As a result of training the model, the accuracy of the work was 76.40%, which is considered to be in the good range. and higher than the target value that has been set according to the scope of the project.

The operation of the display is clear. There is a display from the camera showing various points on the body that are important points of the model and can correctly display the meaning of the gesture as the model sends the values. It can be accessed. and can stop using the system via directly closing the program.

The developer encountered a problem with the library that was selected for use in the Image Processing section, OpenCV, which was unable to display the language in the system like UTF-8, causing various display sections to be unable to display Thai language. The developer therefore chose to use the Pillow Library, which is used in the Image Processing section, to work with OpenCV, which is to use OpenCV to access working with the camera and use Pillow to do all the text display work. Therefore, allowing the system to display Thai language.

The future work, we can use this program to apply on smartphones but it must depend on the efficiency of smartphones. We plan to expand the vocabulary from 50 words to 100 words and run the enhanced model on our server and through an API accessible on smartphones. This expansion aims to improve the model's accuracy. By using a larger dataset and more efficient training equipment, we anticipate achieving higher accuracy and faster performance in the program.

Acknowledgements

Compliance with Ethical Standards Funding This research was financially supported by Rajamangala University of Technology Thanyaburi, RMUTT. Authors' Contributions

The author confirms sole responsibility for the following: study conception and design, data collection, analysis and interpretation of results, and manuscript preparation.

6. References

[1] Bangkokbiznew, *Thailand.* Available Online at https://www.bangkokbiznews.com/corporate-moves/lifestyle/judprakai/1041542, accessed on 27 April 2024.

[2] S. Naqvi, J. Ohene-Djan, and R. Spiegel. "Testing the effectiveness of digital representations of sign language content." *the Instructional Technology and Education of the Deaf Symposium Rochester,* New York, 2005.

[3] J. Whitehill. *Automatic Real-time Facial Expression Recognition for Signed Language Translation.* Master's thesis. University of the Western Cape. Computer Science, 2006.

[4] V. M. Segers. "The efficacy of the eigenvector approach to south african sign language identification." *In South African Telecommunication Networks and Applications Conference 2009*, pp. 363-366, 2009.

[5] P. Li. *Hand shape estimation for South African Sign Language.* Master's thesis University of the Western Cape Computer Science, 2010.

[6] C. Rajah. *Chereme-based recognition of isolated, dynamic gestures from South African Sign Language with Hidden Markov Models.* Master's thesis University of the Western Cape Computer Science, 2006.

[7] C.-S. Lee, Z. Bien, G. T. Park, W. Jang, J. S. Kim, and S. K. Kim. "Real-Time recognition system of Korean sign language based on elementary components." *In Proceedings of the 6th IEEE International Conference on Fuzzy Systems 1997,* pp. 1463-1468, 1997.

[8] Y. H. Lee and C. Y. Tssai. "Taiwan sign language (TSL) recognition based on 3D data and neural networks." *Expert Systems with Applications,* Vol. 36, No. 2, pp. 1123-1128, 2007.

[9] N. D. Sandhya and K. R. Charanjeet. "A review on Machine Learning Techniques." *International Journal on Recent and Innovation Trends in Computing and Communication (IJRITCC),* Vol. 4, No. 3, pp. 395-399, 2019.

[10] 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, pp. 137-143, May-August, 2022.

[11] Y. Quiñonez, C. Lizarraga, and R. Aguayo. "Machine Learning solutions with MediaPipe." *2022 11th International Conference On Software Process Improvement (CIMPS),* pp. 212-215, 2022.

[12] MediaPipe Pose. Available online: https://google.github.io/mediapipe/solutions/pose.html, accessed on 28 December 2023.

[13] 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." *Advances in Social Sciences Research Journal,* Vol. 8, No. 3, pp. 28-34, 2021.

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

[15] OpenCV Open Source Computer Vision library. Available Online at http://opencv.willowgarage.com/wiki/, accessed on 27 April 2024.

[16] M. Ashraf, S. M. Ahmad, N. A. Ganai, R.A. Shah, M. Zaman, S. A. Khan, and A. Shah. "Prediction of cardiovascular disease through cutting-edge deep learning

technologies: an empirical study based on TENSOR FLOW, PYTORCH and KERAS." *In International Conference on Innovative Computing and Communications. Proceedings of ICICC Springer Singapore 2020,* Vol. 1, pp. 239-255, 2020.

[17] S. W. B. Tan, P. K. Naraharisetti, S. K. Chin, and L. Y. Lee. "Simple Visual-Aided Automated Titration Using the Python Programming Language." *Journal of Chemical Education,* Vol. 97, pp. 850-854, 2020.

[18] D. Ulus, V. Kraevo, and A. Sheffer. "D-Charts: Quasi-Developable Mesh Segmentation, Computer Graphics." *Eurographics 2005,* Vol. 24, No. 3, pp. 981-990, 2005.

[19] A. Sheffer, B. Levy, M. Mogilnitsk, and A. Bogomyakov. "ABF++: Fast and Robust Angle Based Flattening." *ACM Transactions on Graphics,* Vol. 24, No. 2, pp. 311-330, 2005.

[20] C. Damrongekarun, L. Pisitpipattana, S. Waijanya, and N. Promrit. "Development of Thai Sign Language Detection and Conversion System into Thai with Deep Learning." *KKU Science Journal,* Vol. 51, No. 3, pp. 216-225, 2023.

[21] X. Shi, Z. Chen, H. Wang, and D. Y. Yeung. "Convolutional LSTM network: a machine learning approach for precipitation nowcasting." *Advances in Neural Information Processing Systems 28 (NIPS 2015),* 2015.

[22] S. Kim, S. Hong, M. Joh, and S. Song. "Deep Rain: ConvLSTM network for precipitation prediction using multichannel radar data." *7th International Workshop On Climate Informatics,* pp. 20-22, September, 2017.

[23] N. K. Mai, D. Yang, I. Shin, H. Kim, and M. Hwang. "Dprelu: Dynamic parametric rectified linear unit." *The 9th International Conference on Smart Media and Applications,* pp. 121-125, 2020.

[24] A. Kumar and A. Abraham. "Opinion Mining to Assist User Acceptance Testing for Open-Beta Versions." *Journal of Information Assurance and Security,* Vol. 12, No. 4, pp. 146-153, 2017.

[25] Tecperson, *Sign Language MNIST.* Available Online at https://www.kaggle.com/datasets/datamunge/sign-language-mnist, accessed on 27 April 2024.

Information Technology Journal

KMUTNB

คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล
มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
1518 ถนนประชาราษฎร์ 1 แขวงวงศ์สว่าง เขตบางซื่อ กรุงเทพฯ 10800
Tel: 02-555-2726
Website: http://www2.it.kmutnb.ac.th/journal