

Designing Answer Sets For Thai Advisory Chatbot With Different Talking Styles For COVID-19 Pandemic



Miss Sakolwan Peetaneelavat

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Science in Computer Science and
Information Technology
Department of Mathematics and Computer Science
FACULTY OF SCIENCE
Chulalongkorn University
Academic Year 2021
Copyright of Chulalongkorn University



จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY



จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

การออกแบบชุดคำตอบสำหรับเชทบอทให้คำแนะนำภาษาไทยด้วยรูปแบบการสนทนาที่แตกต่าง
สำหรับการแพร่ระบาดของโรคติดเชื้อไวรัสโคโรนา 2019



วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชาวิทยาการคอมพิวเตอร์และเทคโนโลยีสารสนเทศ ภาควิชาคณิตศาสตร์และวิทยาการ

คอมพิวเตอร์

คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2564

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

Thesis Title	Designing Answer Sets For Thai Advisory Chatbot With Different Talking Styles For COVID-19 Pandemic
By	Miss Sakolwan Peetaneelavat
Field of Study	Computer Science and Information Technology
Thesis Advisor	Assistant Professor SOMJAI BOONSIRI, Ph.D.

Accepted by the FACULTY OF SCIENCE, Chulalongkorn University in Partial Fulfillment of the Requirement for the Master of Science

----- Dean of the FACULTY OF SCIENCE
(Professor POLKIT SANGVANICH, Ph.D.)

THESIS COMMITTEE

----- Chairman
(Assistant Professor DITTAYA WANVARIE, Ph.D.)

----- Thesis Advisor
(Assistant Professor SOMJAI BOONSIRI, Ph.D.)

----- External Examiner
(Associate Professor PERAPHON SOPHATSATHIT, Ph.D.)

ศกถาวรณ ปิตะนีละวัต : การออกแบบชุดคำตอบสำหรับแชทบอทให้คำแนะนำ
ภาษาไทยด้วยรูปแบบการสนทนาที่แตกต่างสำหรับการแพร่ระบาดของโรคติดเชื้อไวรัสโค
โรนา 2019. (Designing Answer Sets For Thai Advisory
Chatbot With Different Talking Styles For COVID-19
Pandemic) อ.ที่ปรึกษาหลัก : ผศ. ดร.สมใจ บุญศิริ

งานวิจัยนี้เน้นไปที่การออกแบบชุดคำตอบของแชทบอทให้คำแนะนำภาษาไทยด้วย
รูปแบบการสนทนาที่แตกต่างสำหรับการแพร่ระบาดของโรคติดเชื้อไวรัสโคโรนา 2019 เพื่อเพิ่ม
ประสิทธิภาพการใช้งานของแชทบอทให้ดีขึ้น โดยทำการศึกษาวิธีการปรับแต่งภาษาโต้ตอบของ
แชทบอทให้มีลักษณะคล้ายคลึงกับคนที่มีความรู้ในสาขาไอทีและ อินเทอร์เน็ต ทั้งนี้
เพื่อเป็นการประเมินตรวจสอบความถูกต้องในการออกแบบภาษาโต้ตอบ ผู้เขียนจึงได้สร้าง
โมเดลจำแนกความสนใจต่อสิ่งภายนอกจากข้อความขึ้นมาสองโมเดล โดยทั้งสองโมเดลใช้การ
วิเคราะห์การถดถอยโลจิสติกในการพัฒนาเหมือนกันแต่แตกต่างกันที่การเลือกคุณสมบัติ
สำคัญ โมเดลแรกเลือกใช้การสกัดใจความสำคัญข้อความเป็นคุณสมบัติสำคัญเพียงอย่างเดียว
ในขณะที่โมเดลที่สองเลือกใช้ทั้งการสกัดใจความสำคัญข้อความและสัดส่วนของคำที่มี
ความหมายสื่อถึงการแสดงออกต่อสิ่งภายนอกเป็นคุณสมบัติสำคัญ จากการเปรียบเทียบ
ประสิทธิภาพการทำงานพบว่าโมเดลที่เลือกใช้คุณสมบัติทั้งสองอย่างมีความแม่นยำมากกว่า
โมเดลที่ใช้การสกัดใจความสำคัญข้อความเพียงอย่างเดียว

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

สาขาวิชา	วิทยาการคอมพิวเตอร์และ	ลายมือชื่อนิติ
	เทคโนโลยีสารสนเทศ
ปีการศึกษา	2564	ลายมือชื่อ อ.ที่ปรึกษาหลัก
	

6378018323 : MAJOR COMPUTER SCIENCE AND
INFORMATION TECHNOLOGY

KEYWORD chatbot, COVID-19, Personality trait, Interactive language
RD: design, Natural language processing

Sakolwan Peetaneelavat : Designing Answer Sets For Thai
Advisory Chatbot With Different Talking Styles For COVID-19
Pandemic. Advisor: Asst. Prof. SOMJAI BOONSIRI, Ph.D.

This research focuses on designing a set of answers for Thai advisory chatbot using a different conversational approach from conventional chatbot for COVID-19 pandemic in order to improve user's chatbot experience. The research studies how to modify the underlying interactive chatbot language to suit those extrovert and introvert personalities. To evaluate the correctness of answer set design, the author builds two classification models based on Extroversion Text Classification Model (ETCM). Both models utilize the same Logistic Regression algorithm, but different feature selection techniques. The first model relies solely on Term Frequency-Inverse Document Frequency (TF-IDF) property. In contrast, the second model utilizes both TF-IDF and a unique word component score for each speech segment as the principal features. Performance comparison of the study shows that the model utilizing both features is more accurate than that uses only TF-IDF.



Field of Study:	Computer Science and Information Technology	Student's Signature
Academic Year:	2021	Advisor's Signature

ACKNOWLEDGEMENTS

I would like to express my heartfelt gratitude and appreciation to my thesis advisor, Assistant Professor Somjai Boonsiri, Ph.D., for her tremendous guidance and counsel. Her assistance has been invaluable in reading manuscripts and encouraging me to persevere during the research process.

Sincere thanks to the dissertation committee members, Assistant Professor Dittaya Wanvarie, Ph.D., and Associate Professor Peraphon Sophatsathit, Ph.D., for their incisive comments and thoughtful suggestions.

Special thanks to everyone in the public Facebook group Thai Natural Language Processing for their helpful assistance in resolving the research's difficulties.

Above all, I want to thank my family, as well as Miss Anunya Prasongkiat, Miss Jirattha Palasmith, Mr. Nachanok Potiwattutham, Mr. Peeraphol Youpaisarn, and everyone else I love and care about. Their unwavering encouragement and mental support helped me get through the stressful and challenging period of my thesis research.

Sakolwan Peetaneelavat

TABLE OF CONTENTS

	Page
.....	iii
ABSTRACT (THAI)	iii
.....	iv
ABSTRACT (ENGLISH).....	iv
ACKNOWLEDGEMENTS	v
TABLE OF CONTENTS.....	vi
LIST OF TABLES	viii
LIST OF FIGURES	ix
CHAPTER 1 INTRODUCTION	1
1.1 CONCEPTUAL BACKGROUND.....	1
1.2 RESEARCH PROBLEMS	2
1.3 OBJECTIVES	2
1.4 EXPECTED OUTCOMES	3
1.5 SCOPE OF WORK.....	3
1.6 REMAINING SECTION.....	3
CHAPTER 2 LITERATURE REVIEWS.....	4
2.1 NATURAL LANGUAGE PROCESSING (NLP) FOR THAI LANGUAGE....	4
2.2 FUNDAMENTAL PERSONALITY TRAITS.....	5
2.3 MACHINE LEARNING FOR PERSONALITIES CLASSIFICATION BASED ON TEXT	7
2.4 DESIGNING A CHATBOT FOR COUNTERING COVID-19 PANDEMIC...9	
CHAPTER 3 METHODOLOGY	11
3.1 THE PROPOSED WORKFLOW.....	11
3.2 THE QUESTIONNAIRE DESIGN.....	12
3.2.1 GENERAL CHARACTERISTICS OF THE RESPONDENTS.....	12

3.2.2 QUESTIONS ABOUT ADDRESSING CONCERNS AND OBTAINING INFORMATION ON COVID-19.....	13
3.3 EXTROVERSION TEXT CLASSIFICATION MODEL (ETCM) DEVELOPMENT.....	13
3.3.1 DATA ACQUISITION	14
3.3.2 DATA PREPROCESSING	14
3.3.3 FEATURE SELECTION	15
3.3.4 MACHINE LEARNING MODEL: LOGISTIC REGRESSION.....	17
3.3.5 MODEL EVALUATION.....	17
CHAPTER 4 RESULT AND DISCUSSION.....	19
4.1 THE PERFORMANCE OF EXTROVERSION TEXT CLASSIFICATION MODEL (ETCM)	19
4.1.1 PERFORMANCE OF ETCM USING ONLY TF-IDF	19
4.1.2 PERFORMANCE OF ETCM USING TF-IDF AND PART OF SPEECH AND UNIQUE WORDS PORTION CALCULATION.....	20
4.2 DISCUSSION ABOUT ETCM MODEL.....	20
4.3 THE DESIGNING OF ANSWER SETS	21
CHAPTER 5 CONCLUSION.....	26
5.1 SUMMARY.....	26
5.2 LIMITATION AND FUTURE WORK	27
REFERENCES	28
APPENDIX.....	30
VITA.....	39

LIST OF TABLES

	Page
Table 1. Set of questions about general characteristics of the respondents	12
Table 2. Set of questions about addressing concerns and obtaining information on COVID-19.....	13
Table 3. Selected parts of speech and unique words with examples words in English and Thai	16
Table 4. Confusion Matrix for ETCM evaluation	17
Table 5. Confusion Matrix of ETCM using TF-IDF as a feature and testing with 200 samples.....	19
Table 6. Confusion Matrix of ETCM using TF-IDF, Part of Speech, and Unique Words as Selected Features and Testing on 200 samples.....	20
Table 7. Performance comparison of two versions of ETCM	20
Table 8. List of intents included in proposed Thai advisory chatbot with different talking style for COVID-19 pandemic.....	23
Table 9. Different answer sets	25
Table 10. The first answer sets (65% of Extroversion and 35% of Introversion)	30
Table 11. The second answer sets (40% of Extroversion and 60% of Introversion)..	33
Table 12. The third answer sets (50% of Extroversion and 50% of Introversion)	36

LIST OF FIGURES

	Page
Figure 1. The Big Five Factors [5].....	5
Figure 2. The differences between extroverts, ambiverts, and introverts [8]	6
Figure 3. Personality trait classification processing [11].....	8
Figure 4. The workflow of “COVID-19 Preventable” chatbot [15]	9
Figure 5. The proposed workflow.....	11
Figure 6. The fundamental procedures of ETCM.....	14
Figure 7. ETCM models execution time comparison	21
Figure 8. The results of the vote on the COVID-19 topics.....	22
Figure 9. The results of the vote on unapproachable COVID-19 topics	23

CHAPTER 1 INTRODUCTION

This chapter describes the conceptual background, research problems, objectives, expected outcomes, the scope of the work, and the remaining section.

1.1 CONCEPTUAL BACKGROUND

Since January 2020, Thailand has been undergoing the first wave of the COVID-19 pandemic catastrophe. However, it is presently into the fifth wave as of January 2022, with over 10,000 new cases each day and an epidemic in multiple provincial districts [1]. In addition, the newest specie of COVID-19, Omicron, is capable of spreading easier in the air, which increases the daily number of infections even in people who already have four doses of vaccines. Therefore, although Thailand's pandemic situation is relatively stable right now (May 2022), discretionary news monitoring is still required. However, according to the survey results, most people, particularly adolescents and young adults, prefer to follow COVID-19 news from unofficial sources over following sources directly from the Centers for Disease Control. If people follow news from unauthorized sources, they would likely quickly end up with incorrect information. However, filtering news is difficult in practice because people constantly face the issue of encountering contradictory information from various sources. By the way, it is difficult to determine which information is more reliable. Additionally, if they cannot find the necessary information immediately, this may cause them to experience anxiety and tension.

It would be excellent to get over this problem if there are some artificial intelligence (AI) machines. Thus, developing a chatbot that provides accurate information is a helpful approach to circumvent the difficulties and restrictions users face when searching for and receiving assistance regarding the COVID-19 outbreak. The chatbot also has more advantages in terms of performance compared with the human help call center which is able to provide assistance service 24/7, parallelly, and invariably. Although the chatbot is an automated machine, conversational satisfaction must be appropriately considered. Basically, a Question and Answer (Q&A) chatbot would respond to users as the retrieved information developer assigned, which may

use the blog language or too much formal word form; this method can decrease user experience since the users are overwhelmed with the bore as they need to read the long answers.

Similar to people, a chatbot's personality could be reflected in its responses' language selections and grammatical structures. According to psychological understanding, individuals typically get along well with similar personality traits or lifestyles. Previous research stated that this principle applies to human-machine interaction as well. Moreover, the experiment result on the chatbot and user engagement was reported that matching users with congruent chatbots, modified to be more personalized, improves the user experience and engagement [2].

It is common practice for developers to manually assign the response message when working on designing responding languages for the chatbot. Therefore, they may lack the expertise necessary to determine which choice of words and grammatical structures can identify the personalities they plan to manipulate for the chatbot.

Consequently, the author's effort contributed to developing the Extroversion Text Classification Model (ETCM), intended to assist in testing phrases of the designed responding languages for the chatbot. ETCM is trained with the Logistic Regression model, the Term Frequency-Inverse Document Frequency (TF-IDF) model, and a unique word component score for each speech segment.

1.2 RESEARCH PROBLEMS

1. The COVID-19 chatbot which are currently available do not, as of yet, meet the requirements for satisfactory user experiences.
2. There is only one personality that can be explicitly designed a chatbot.

1.3 OBJECTIVES

The research is centered on the designing answer sets for enhancing the user experience of the chatbot in the domain of COVID-19 by optimizing chatbot interactions for users with various talking styles by demonstrating the implementation of machine learning techniques along with the responding language design processes.

1.4 EXPECTED OUTCOMES

The outcome anticipated from this research is to design the customizing chatbot that has varied its personality through its multiple answer sets, each with its own characteristic speaking style.

1.5 SCOPE OF WORK

The following describes the areas that are the focus of this research:

1. The COVID-19 epidemic in Thailand will be the primary topic of discussion in both the domain of the proposed chatbot and the material that will be delivered.
2. The answer sets of proposed chatbot will exclusively converse with Thai speakers while interacting with people.

1.6 REMAINING SECTION

The remaining section of this work is organized as follows: Chapter 2 contains literature reviews. The methodology is outlined in Chapter 3. Chapter 4 displays the result and discussion. In Chapter 5, the conclusion of this research is presented.

CHAPTER 2 LITERATURE REVIEWS

In this chapter, the works and theories which relate to designing answer sets for Thai advisory chatbot with different talking style for COVID-19 pandemic are described as follows:

2.1 Natural Language Processing (NLP) for Thai language

2.2 Fundamental personality traits

2.3 Machine learning for personalities classification based on text

2.4 Designing a chatbot for countering COVID-19 pandemic

2.1 NATURAL LANGUAGE PROCESSING (NLP) FOR THAI LANGUAGE

In recent years, the surge in popularity of artificial intelligence (AI) devices has increased businesses and brands employing deep learning methodologies for product quality monitoring. Natural Language Processing, or NLP for short, is a subfield of artificial intelligence that enables computers to understand human languages. It may be taught to do a wide variety of text-related tasks, including text categorization and sentiment analysis.

In addition, NLP can function effectively with any language worldwide; for Thai, WANGCHANBERTa is an enhanced version of the BERT-based language transformer concept designed exclusively for Thai. It outperforms the majority of prior Thai NLP language models primarily in terms of word chunks and sentence boundaries. In addition, it may circumvent specific text patterns that are common among youngsters, such as numerous superfluous letters or emoji attachments.

WANGCHANBERTa also includes a variety of models that might be beneficial for text preprocessing, which is a necessary procedure in every NLP activity. Text preprocessing helps remove noise, which can affect the performance of categorizing prediction. For example, tokenization is the process involving chopping

up a passage of text or sentences into individual words. WANGCHANBERTa provides various versions of tokenizers carry out downstream activities utilizing Thai datasets which are well trained on the Thai Wikipedia Dump with using the analysis of PyThaiNLP, a popular python library for Thai NLP programming, in a variety of text structure's level, e.g., word-level, subword-level and syllable-level [3].

2.2 FUNDAMENTAL PERSONALITY TRAITS

People's distinctive habits of thinking, feeling, and behaving are reflected in the features that make up their personalities. Characteristics of a person's personality suggest regularity and dependability [4]. There are many types of personality trait dimension factors. The five factors, or Big Five factors, are the most widely used to describe a person's characteristics including agreeableness, conscientiousness, extroversion, neuroticism, and openness as shown in Figure 1.

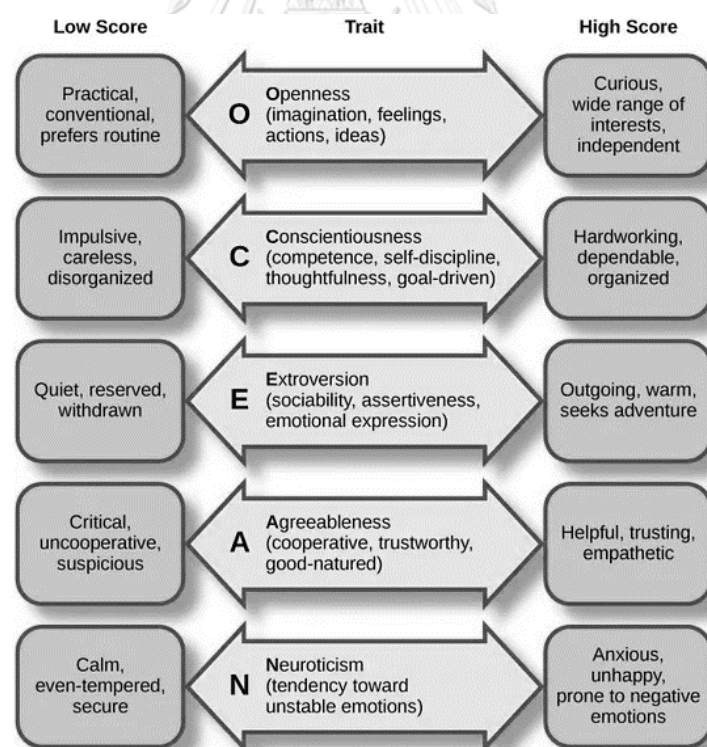


Figure 1. The Big Five Factors [5]

Personality traits get the more attention in nowadays. It could tell by an increasing number of individuals are taking the Myers–Briggs Type Indicator (MBTI) personality test to determine their personality traits or their Extroversion. Normally, there are typically two well-known personality types: extrovert and introvert. Each personality type determines a person's unique identity, which influences their physical and emotional conduct. It may also give a solution guideline for every scenario, such as socializing and tension. In addition, there is a unique personality type known as an "ambivert" who possesses a balance of both introversion and extroversion, with varying tendencies depending on the circumstances [6] [7]. Figure 2 depicts illustrative examples of each personality type's characteristic behaviors.

Extrovert	Ambivert	Introvert
Enjoys spending time with others	Enjoys others, but needs alone time	Enjoys spending time alone
Is outgoing	Can be outgoing in the right context	Is reserved
Prefers to work on teams	Can be on teams after working independently	Prefers to work independently
Likes to talk things out with others	Can process outloud or internally	Likes to process internally

Figure 2. The differences between extroverts, ambiverts, and introverts [8]

In addition, an individual's use of language might provide insight into their level of Extroversion. For example, the Linguistic Inquiry and Word Count (LIWC) research [9] suggests that extroverted people are more likely to have a more abstract linguistic style, as demonstrated by their use of a more significant number of state verbs and adjectives. They also like to use optimistic language and vocabulary while expressing themselves in communities and socialization. On the other hand, people who are more introverted are more inclined to focus on clarity and compactness in their language use. As a result, they favour using action verbs that allude to particular information and try to avoid using many needless verbiages.

Mark Trommer, a senior organizational consultant at the UA and an MBTI certified practitioner, mentioned that when people communicate with one another, it might be challenging if their styles are distinct from one another [9]. It could say that people are more likely to get along with others who have personalities similar to their own since they would have the same styles of thinking and making-decision pattern. Numerous studies demonstrate that this concept also applies to the interaction between humans and robots [10]. Personality may be put into any social bot by designing its interactions, such as its action motions and response language. Research reveals that the user experience is enhanced when a user is connected with a chatbot that shares similar personality traits, such as being an extrovert or an introvert [2].

2.3 MACHINE LEARNING FOR PERSONALITIES CLASSIFICATION BASED ON TEXT

Text classification is the one of important task in text mining which performs document categorization into one or multi-categories. Text mining is critical useful and be widely used in many business industries. For example, companies can perform text mining process with their customers' feedback comments to the products in order to measure their satisfaction. Practically, text classification tasks can be performed by various methodology, however, implementation of machine learning is the most popular techniques since there are many researchers did the studies on training machine learning models for text classification.

Personality trait classification based on the text is an example of text classification research including NLP implementation. As depicted in Figure 3, it is possible to predict a person's personality attribute from their written content. Personal datasets and sample texts from any social network, such as Twitter or Facebook, must be collected first. Then, a series of preprocessing steps must be performed to prepare them as training data

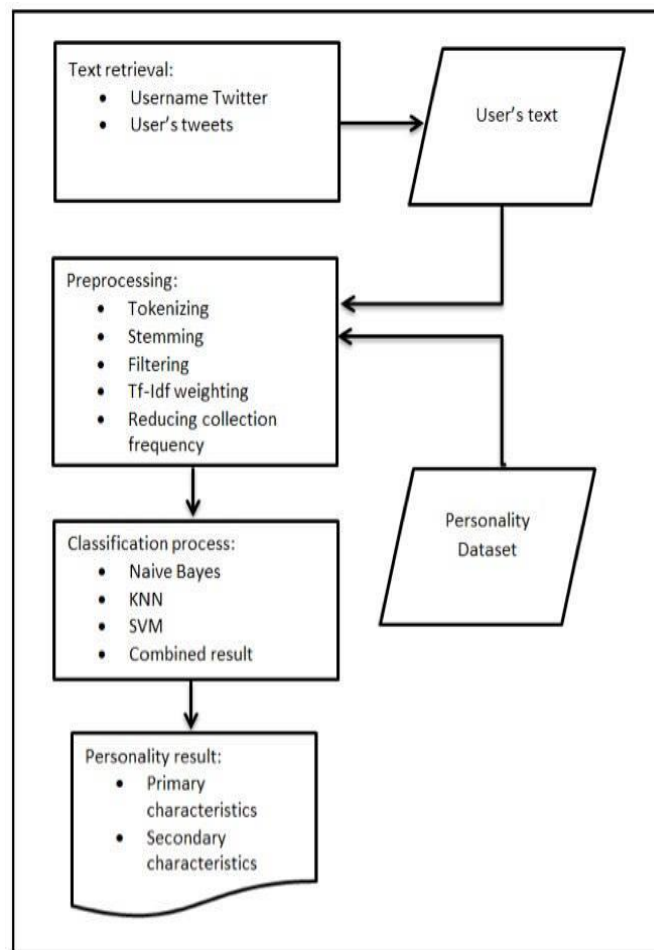


Figure 3. Personality trait classification processing [11]

In this study [11], they conducted an experiment utilizing three classification techniques: Naive Bayes (NB), K-Nearest-Neighbors (KNN), and Support Vector Machine (SVM). In addition, this study discovered that the new classification procedure that incorporated the three aforementioned methodologies provided the greatest performance and precision.

In [12], the authors proposed a Voting classifier that combined 2 traditional machine learning classification models, logistic regression and stochastic gradient descent, and the experiment results demonstrated that it provides superior performance compared to any traditional machine learning classification model in terms of accuracy.

In [13], the authors did the studies about training the text classification models for proactive personality prediction by applying five machine learning algorithm; NB, KNN, SVM, XGBoost and Logistic Regression (LR). According to model evaluation, it was found SVM and LR provides the excellent performance compared with the others.

2.4 DESIGNING A CHATBOT FOR COUNTERING COVID-19 PANDEMIC

A chatbot designed to fight the COVID-19 pandemic needs to be able to provide users with the following functionality in order to be effective: reliable information proclamation, risk assessment for COVID-19 infection by fastidious symptoms (such as cough, nasal congestion, difficulty breathing, fatigue, and fever), practical self-protect suggestion, and supporting for mental temperament [14].

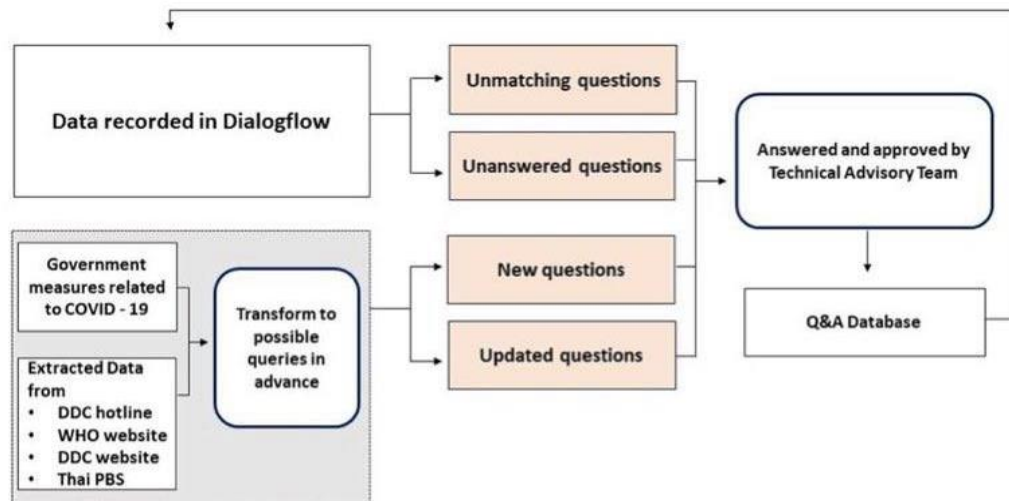


Figure 4. The workflow of "COVID-19 Preventable" chatbot [15]

In 2019, the personnel of the Department of Disease Control (DDC) has developed their own chatbot called "COVID-19 Preventable" [15] for answering inquiries about COVID-19 during the pandemic. The development process applied the Design Science Research Methodology (DSRM) and the workflow of this proposed

chatbot is described in Figure 4. "COVID-19 Preventable" is a retrieved-based chatbot for which the information data is taken directly from the government and official website; subsequently, the technical advisory team worked on query design and the questions were recorded in a database for use with Dialogflow, a chatbot creation software.



CHAPTER 3 METHODOLOGY

In this chapter, the methodology for designing answer sets for Thai advisory chatbot with different talking style for COVID-19 pandemic are described as follows:

3.1 The proposed workflow

3.2 The questionnaire design

3.3 Extroversion Text Classification Model (ETCM) development

3.1 THE PROPOSED WORKFLOW

Figure 5 illustrates the steps of proposed workflow for designing answer sets for Thai advisory chatbot with different talking style for COVID-19 pandemic development.

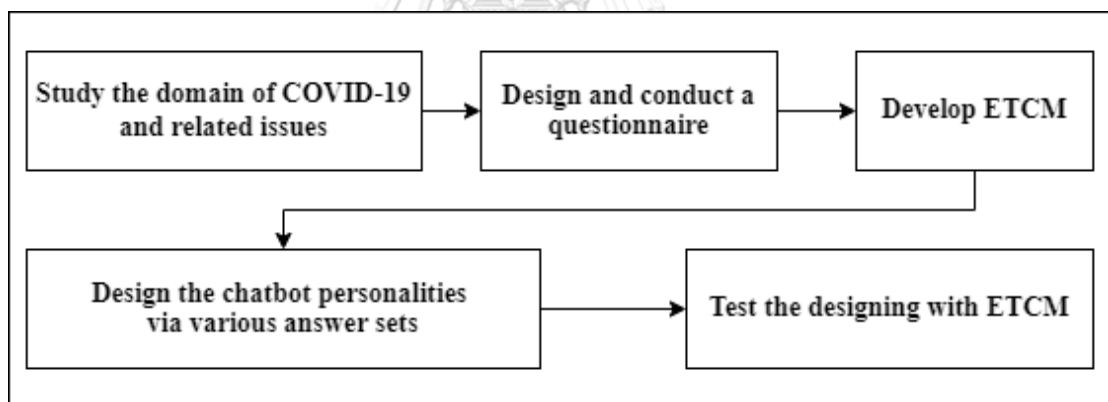


Figure 5. The proposed workflow

In order to scope the functional demand and non-functional requirement of the proposed chatbot, it is necessary first to understand the present COVID-19 pandemic condition in Thailand and the basic information about personality features. Hence, the initial duty is to determine the domain issues using research papers and websites. The second task is conducting a questionnaire about concerns and getting information regarding COVID-19, followed by an analysis.

The next phase is creating an Extroversion Text Classification Model, or ETCM, which would be the foundation for the chatbot's interactive language design.

The following phase is to design a prototype of the proposed chatbot that represents three essential chatbot framework properties: the bot's scope, its personality, and a prioritized list of must-have features. The next phase is conducting research on the appropriate theories for designing a chatbot so that it could have a personality.

As the final step, ETCM would analyze the design of interactive language in order to guarantee that the planned response text would be able to identify the individual's personality features correctly.

3.2 THE QUESTIONNAIRE DESIGN

The surveying of the opinions of the target user representatives is the purpose of the questionnaire that was created in response to concerns and information regarding COVID-19. This questionnaire is divided into 2 sections:

3.2.1 General characteristics of the respondents

3.2.2 Questions about addressing concerns and obtaining information on COVID-19

3.2.1 GENERAL CHARACTERISTICS OF THE RESPONDENTS

In this subsection, there are questions pertaining to the respondents' basic information. The sequence of questions that comprise this part are shown in Table 1.

Table 1. Set of questions about general characteristics of the respondents

Question Sequence No.	Question
1	Age range
2	Occupation
3	Personality trait (Extrovert, Ambivert, and Introvert)

3.2.2 QUESTIONS ABOUT ADDRESSING CONCERNS AND OBTAINING INFORMATION ON COVID-19

The purpose of this subsection is to inquire about the COVID-19-related search issues that respondents encounter. The sequence of questions comprising this section is displayed in Table 2.

Table 2. Set of questions about addressing concerns and obtaining information on COVID-19

Question Sequence No.	Question
1	The degree of concern about the spread of COVID-19
2	Top 3 concerns about the spread of COVID-19
3	The primary source for receiving information about COVID-19
4	The level of satisfaction with the ability to receive news from the chosen source in the current situation
5	Issues considered to be inaccessible or lacking information

3.3 EXTROVERSION TEXT CLASSIFICATION MODEL (ETCM) DEVELOPMENT

As shown in Figure 6, There are five fundamental procedures involved in the Extroversion Text Classification Model (ETCM) development.

3.3.1 Data Acquisition

3.3.2 Data Preprocessing

3.3.3 Feature Selection

3.3.4 Machine Learning model: Logistic Regression

3.3.5 Model Evaluation

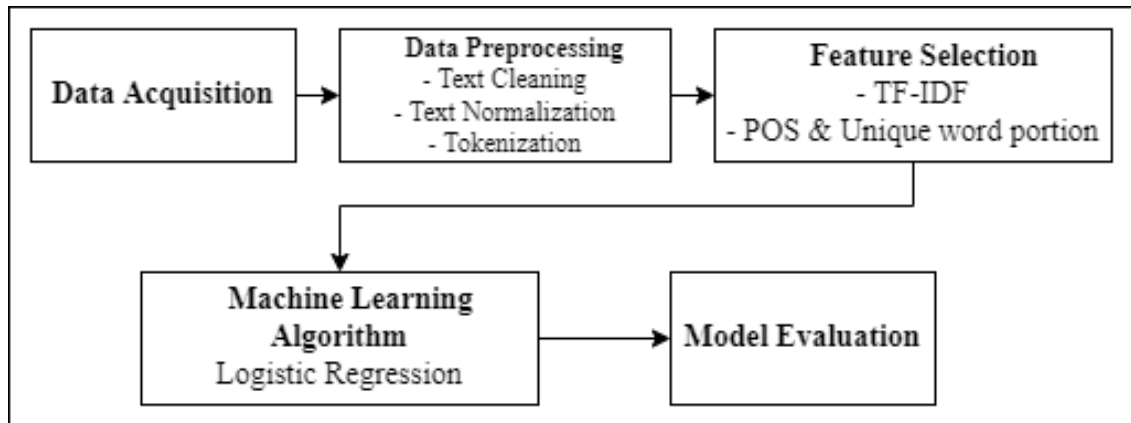


Figure 6. The fundamental procedures of ETCM

3.3.1 DATA ACQUISITION

Column “content,” and column “personal_type” exist in 600 records of Train data, 200 recordings of Validation data, and 200 records of Testing data, respectively. The first column, “content” consists of postings and comments from the public MBTI Facebook group. The second column, “personal_type,” indicates the personality type of each content poster. There is two valid values for “personal_type”; “E” and “I”, with E representing Extrovert and I representing Introvert.

3.3.2 DATA PREPROCESSING

Following the completion of data collection, every piece of information included in the column labelled “content” is preprocessed utilizing text cleaning and normalization, most likely leading to noise reduction procedures. Text cleaning helps remove unnecessary symbols like the asterisk (*) and question mark (?), along with hashtags, punctuation, and separators (\n or \t). Text normalization is putting words back in their typical sequence and spelling after they have been misspelt or rearranged.

After that, the contents are word segments being tokenized at the word level using WANGCHANBERTa's provided models, and the results are saved in the "content token" variable.

Stop words are frequently used terms in a language that are often removed from the original document because their presence does not contribute anything to the meaning of the text. The removal of stop words is another essential step in the preprocessing stage that was skipped over in this investigation if the sentimental meaning of the text is the vital factor of the analysis [16]. However, in this research, the emotional meaning of the text is not the primary purpose. The collection of stop words is one of the essential word values for the classification of extroversion in text. Hence, the author decided to skip the process of stop words removal and kept them in the content data since the proposed ETCM was intended to perform the Extroversion from the text by considering every word in the sentence.

3.3.3 FEATURE SELECTION

The author of this work made the decision to develop the suggested ETCM into two different variants. The initial iteration of the model has the stand-alone component be the Term Frequency-Inverse Document Frequency (TF-IDF) algorithm.

Term Frequency (TF) is a statistic that measures the frequency with which specific terms and phrases appear in a given document. The $TF(t)$ formula divides the total number of words in the text by the number of occurrences of the term t . In the interim, the Inverse Document Frequency (IDF) approach may be used to determine a phrase's importance. $IDF(t)$ equals the logarithm of the total number of documents divided by the number of documents containing the word t . The TF-IDF score of each word is then calculated by multiplying its TF values by its IDF value [16].

The second version of ETCM is that adding another selected feature which is Part of Speech and Unique words portion score. By dividing the total number of words in the text by the total number of words analyzed, the proportion of occurrences of each word that meet the requirements for each kind of word may be

determined. The sum of these proportions determines the score for each section of speech and unique word as shown as below formula.

$$\sum \frac{\text{Number of marked word as each word type in each content}}{\text{Number of words in content}}$$

As can be seen in Table 3, there are a total of ten different types of parts of speech and unique words chosen based on researches on extroversion and verbal behavior [17].

Table 3. Selected parts of speech and unique words with examples words in English and Thai

Word Types	Examples in English	Examples in Thai
Adjective	Beautiful, Delicious	สวย, อร่อย
Positive word	Excellent, Fun	ดีเยี่ยม, สนุก
Negative word	Terrible, Boring	แย่, น่าเบื่อ
Social word	Friend, Community	เพื่อน, ชุมชน
Exclusive word	But, Except	แต่, ยกเว้น
Certainty word	Certainly, Most	แน่นอน, ส่วนใหญ่
Tentative word	Perhaps, May	บางที, อาจจะ
Action verb	Swim, Hit	ว่ายน้ำ, ตี
State verb	Love, Hate	ชอบ,เกลียด
Quantity word	Few, Many	น้อย, มาก

3.3.4 MACHINE LEARNING MODEL: LOGISTIC REGRESSION

When it comes to challenges involving text categorization, one of the most prominent machine learning algorithms is called logistic regression. This technique is particularly useful for binary classifications. It calculates an estimate of the probability that is associated with a discrete result given a variable that is included in the input [18, 19].

3.3.5 MODEL EVALUATION

Table 4 is described the confusion matrix, True positives, also known as TPs, are denoted by the top-left rectangle in the image. These are examples in which the model accurately predicted that extroverts make up the content owners. In a similarly, the bottom right one displays true negatives (TNs), which indicates that the model accurately predicts that persons who own content are introverted. On the other hand, the columns at the top right and bottom left reflect false positives (FPs) and false negatives, respectively (FNs). Both of these subsections detail the length of time during which the model made inaccurate predictions about the Extroversion of testing contents.

Table 4. Confusion Matrix for ETCM evaluation

	Actual: E	Actual: I
Predicted: E	True Positives (TPs)	False Positives (FPs)
Predicted: I	False Negatives (FNs)	True Negative (TNs)

For the purpose of evaluating and contrasting the efficacy of various proposed models, the confusion matrix is utilized in the computation of accuracy, precision, recall, and F1-score. The objective of this comparison is to identify the models that

provide the highest levels of performance. The following is an illustration of how the accuracy, precision, recall, and F1-score calculations are applied:

$$\textit{Accuracy} = \frac{\textit{TPs} + \textit{TNs}}{\textit{P} + \textit{N}}$$

$$\textit{Precision} = \frac{\textit{TPs}}{\textit{TPs} + \textit{FPs}}$$

$$\textit{Recall} = \frac{\textit{TPs}}{\textit{TPs} + \textit{FNs}}$$

$$\textit{F1 - Score} = \frac{2 * \textit{Precision} * \textit{Recall}}{\textit{Precision} + \textit{Recall}}$$

CHAPTER 4 RESULT AND DISCUSSION

In this chapter, the result and discussion of designing answer sets for Thai advisory chatbot with different talking style for COVID-19 pandemic are described. The issues covered in this chapter are divided into the following categories:

4.1 The performance of Extroversion Text Classification Model (ETCM)

4.2 Discussion about ETCM model

4.3 The designing of answer sets

4.1 THE PERFORMANCE OF EXTROVERSION TEXT CLASSIFICATION MODEL (ETCM)

As indicated before, the author of this research developed two distinct versions of the Extroversion Text Classification Model (ETCM), each of which applied a distinct method for selecting features. In the first model, the TF-IDF is the only differentiating characteristic, but the second model also considers parts of speech and unique words.

4.1.1 PERFORMANCE OF ETCM USING ONLY TF-IDF

ETCM's efficiency while employing TF-IDF alone is represented as the confusion matrix in Table 5.

Table 5. Confusion Matrix of ETCM using TF-IDF as a feature and testing with 200 samples

	Actual: E	Actual: I
Predicted: E	52	45
Predicted: I	37	66

4.1.2 PERFORMANCE OF ETCM USING TF-IDF AND PART OF SPEECH AND UNIQUE WORDS PORTION CALCULATION

The Confusion Matrix of ETCM was developed using TF-IDF, Part of Speech, and Unique Words as the Selected Features in Table 6.

Table 6. Confusion Matrix of ETCM using TF-IDF, Part of Speech, and Unique Words as Selected Features and Testing on 200 samples

	Actual: E	Actual: I
Predicted: E	56	45
Predicted: I	33	66

4.2 DISCUSSION ABOUT ETCM MODEL

According to Table 7, the second version of ETCM, which incorporated a segment of unique words and parts of speech as an extra feature, got a slightly higher confusion matrix score.

Table 7. Performance comparison of two versions of ETCM

	ETCM using only TF-IDF	ETCM using TF-IDF and Part of Speech and Unique words portion calculation
Accuracy (Model Score)	0.59	0.61
Precision	0.58	0.63
Recall	0.54	0.55
F1-Score	0.56	0.59

It was revealed that taking into consideration word choice in the text might improve the model's capacity to predict extroversion from the text in terms of

performance metrics. ETCM using TF-IDF and Part of Speech and Unique words portion calculation (POS) also consumes less execution time and CPU execution time (shown in Figure 7). Execution time is the total number of seconds required to run the code, meanwhile, CPU Execution time) is the time between the beginning and finish of a program's execution. This time does not include the time spent waiting for I/O or executing other applications.

The authors had to manually assemble a substantial amount of sample data and a unique vocabulary of parts of speech in order to perform this research. Thus, the amount of data accessible for analysis was limited. The result, however, reveals that it is still operational and may improve the performance of ETCM via a greater degree of accuracy. Therefore, if we continue building this dictionary so that the model may learn more languages and test with more sample data, ETCM would certainly deliver a much higher degree of performance than the proposed models in this research.

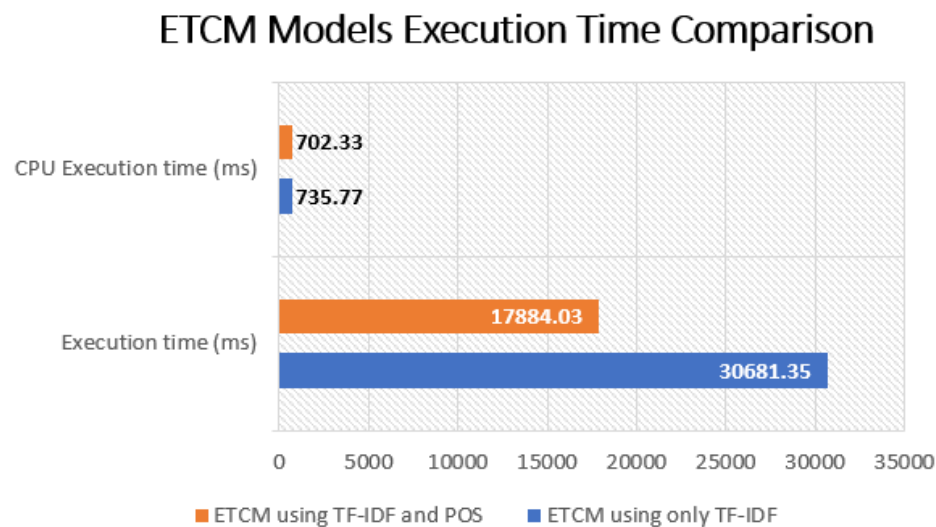


Figure 7. ETCM models execution time comparison

4.3 THE DESIGNING OF ANSWER SETS

Accord to the conducted questionnaire analysis about concerned and unapproachable topic about COVID-19, Figure 8 presents the results of the vote on the concerned COVID-19 issue.

The fact that vaccinations are in first place with 173 points reveals that the vast majority of respondents have a reasonable amount of anxiety over them. The second item of concern, under "measures and risk areas" has a total of 134 points. The next topic, which is COVID-19's various new species, came in third place with 122 points. For how to care for and protect themselves during the pandemic, which is far behind in fourth place with 87 points and in last place with 78 points respectively, are testing and treatment during the pandemic.

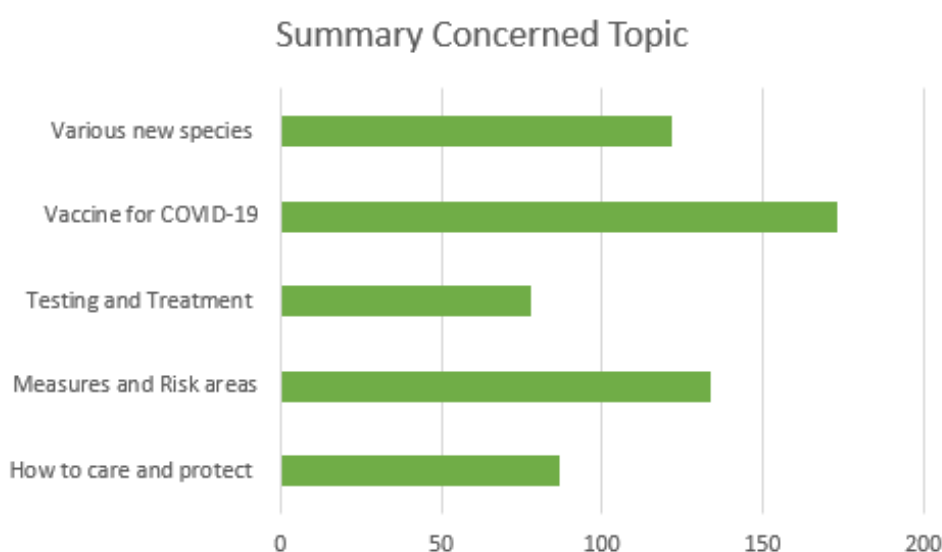


Figure 8. The results of the vote on the COVID-19 topics

The results of the vote on unapproachable COVID-19 subjects are displayed in Figure 9. With 162 points, it is clear that the majority of respondents believe the subject of vaccinations for COVID-19 to be the most insurmountable obstacle. This is evidenced by the fact that it is in the lead. Measures and risk areas are in second position with 103 points, which is a very distant second. The next subject, which is all about the many new species of COVID-19, has scored 88 points and has taken the neighboring third position. Testing and treatment, as well as how people may care for themselves and protect themselves, are tied for fourth place with 79 points each and are a long way behind the issue that came in last place, which had 35 points. Because the results of the two polls were so strikingly comparable, we can deduce that individuals are generally troubled by the fact that they are unable to obtain

information that is either straightforward or simple to grasp on the subject at hand and the advisory chatbot should be able to provide these five issues comprehensively.

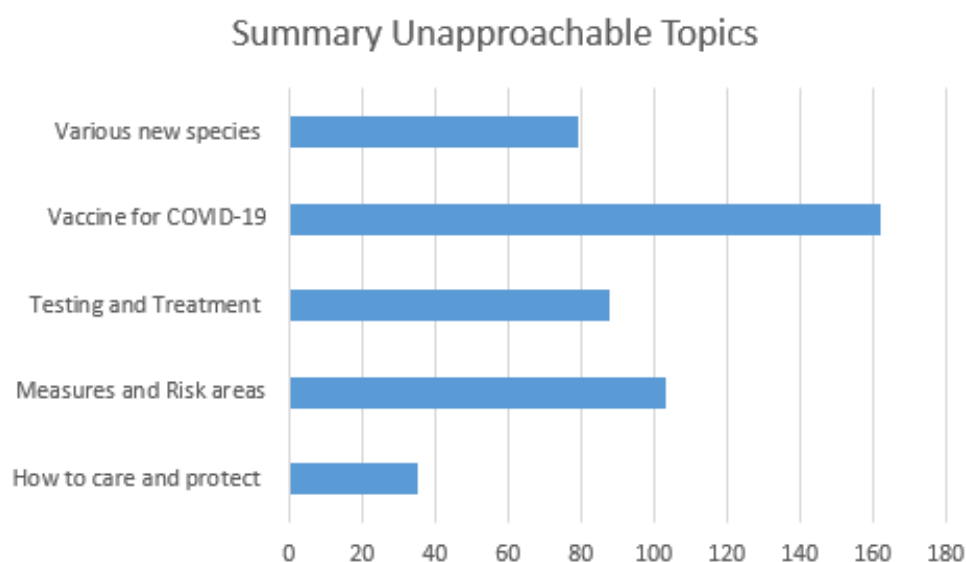


Figure 9. The results of the vote on unapproachable COVID-19 topics

As the result, the function of the Thai advisory chatbot is designed as shown in Table 8 according to the issue under consideration in order to cover all of COVID-19 issues which are focused by the target user representatives.

Table 8. List of intents included in proposed Thai advisory chatbot with different talking style for COVID-19 pandemic

Group	Sub-group	Intent Name	Keyword	Description
Update COVID-19 Situation	Daily Cases	daily	รายวัน/วันนี้	Show the statistic of COVID-19 daily cases
	Official Announcement	announce	มาตรการ/สั่ง/สั่งการ/บังคับ	Update for DDC Official Announcement
Information of COVID-19	General Characteristic	covid_char	โควิด/คือ/เป็นไง	Show a picture about general characteristic
	Spreading	spread_main	แพร่เชื้อ/ติด/ติดต่อ	Provide information about COVID-19 spreading
		spread_per	คน/ผู้ติดเชื้อ/กลุ่มเสี่ยง	Provide information about COVID-19 spreading by people

		spread_obj	สิ่งของ/จับ	Provide information about COVID-19 spreading by object
		spread_envi	อากาศ/ละออง	Provide information about COVID-19 spreading by environment
	Various species	species	สายพันธุ์	Show a picture that describes the differences of each species of covid.
Check-up risk and symptoms	Self-Assessment	self_assessment	ประเมิน/เสี่ยง	Provide information about self-assessment
	COVID-19 Testing	covid_test	ตรวจ/หาเชื้อ	Provide information about COVID-19 testing
		atk	atk/ATK/เอทีเค	Provide information about ATK
		pcr	pcr/PCR/พีซีอาร์	Provide information about PCR
Treatment Advices	Medicine	med	ยา/สมุนไพร	Provide information about medicine
	Home Isolation	home_iso	HI/home/iso/กักตัว	Provide information about home isolation
	Hospital	hospital	โรงพยาบาล/โรงพยาบาล	Provide information about hospital
	Hospitel	hospitel	โรงแรม/hospitel	Provide information about hospitel
Self-Pervent Advices	Mask & Alcohol	mask_al	หน้ากาก/แอลกอฮอล์	Provide advice about Mask & Alcohol
		mask	แมสก์/มาสก์/หน้ากาก	Provide advice about Mask
		alcohol	ล้างมือ/เจล	Provide advice about Mask Alcohol
	Transportation	transport	ขนส่ง/รถ/รถไฟฟ้า/bts/mrt	Provide advice avoid infection while using transportation
	Routines	routine	ป้องกันตัว	Provide advice about daily avoid infection
	Vaccines	vac	วัคซีน/เข็ม/ฉีด	Provide information about vaccine
		vac_kind	ยี่ห้อ/ตัวไหน	Provide information about kinds of vaccine

		vac_eff	อาการ/ข้างเคียง/ หลังฉีด	Provide information about effects of vaccine
	Place	place	สถานที่/ห้าง/ ชุมชน	Provide advice about avoid infection in public place

As part of the process of designing an interactive language for the proposed chatbot, one of the objectives was to adjust the personalities of this chatbot such that they were akin to those of extroverts and introverts via the use of several sets of text answer styles. For extrovert personalities, the bulk of answer text would consist of adjectives, positive phrases, terminology associated with social events or communities, many numbers in a single sentence, words expressing confidence, and state verbs. In the meantime, the collection of inversion response text consists mostly of terminology such as action verbs, negative words, number, quantification, exclusive words and tentative words.

Table 9 demonstrates the response text sets that can be used to imbue the chatbot with a variety of personalities. The prediction performed by ETCM accounts for a certain amount of Extroversion and introversion in an individual.

Table 9. Different answer sets

Answer sets	1	2	3
Extroversion (%)	65	40	50
Introversion (%)	35	60	50
Personality result	Extrovert	Introvert	Ambivert

CHAPTER 5 CONCLUSION

This chapter provides a summary of the research, limitations, and suggestions for further research.

5.1 SUMMARY

In May 2022, the COVID-19 pandemic situation in Thailand seems to have settled down for the time being but its most up-to-date variant, which is referred to as Omicron, can disperse through the air with greater ease. This has led to an increase in the daily number of infections, even among those who had received four doses of the vaccine at an earlier point in time. So, preventing and monitoring the news is still requested. Therefore, avoiding infection and keeping up with current circumstances is still recommended. The purpose of this research was to design answer sets for a Thai advisory chatbot with different talking styles and to improve the user experience by assigning personalities to the chatbot through the use of various responding languages.

ETCM is a machine learning system built by the author to aid in the design of the chatbot's answer sets. The following are the five essential phases of the suggested procedures. The approach began with the collection of information from public Facebook groups. Data The subsequent processes are preprocessing and selection of the distinguishing feature. Using the logistic regression model and the confusion matrix, we evaluate the model's performance. There are two versions of ETCM trained in this work. The difference between the two versions is applying different feature selections. The first version is simpler by using TF-IDF as the only feature. The second version has more complex features but it was found to outperform the first version on various criteria, such as accuracy, precision, recall, F1 score, and time complexity.

5.2 LIMITATION AND FUTURE WORK

For the authors to finish their research, they had to manually assemble a substantial number of the sample data as well as an original lexicon of parts of speech. As a result, there are certain limitations to the amount of data that is accessible for analysis. The result reveals that it is still functional and suggests that having a greater degree of accuracy may improve the performance of ETCM. As a consequence, ETCM will provide a higher level of performance if we proceed with the construction of this dictionary so that the model may learn more languages with the help of a wide variety of test data.



REFERENCES

1. Organization, W.H. *Thailand: WHO Coronavirus Disease (COVID-19) Dashboard*. 2020 [cited 2021 20 Sep.]; Available from: <https://covid19.who.int/region/searo/country/th>.
2. Johnson, M.S.a.L.W., *Making conversations with chatbots more personalized*. *Computers in Human Behavior*, 2021. **117**: p. 106627.
3. Lalita Lowphansirikul, C.P., Nawat Jantrakulchai and Sarana Nutanong, *WangchanBERTa: Pretraining transformer-based Thai Language Models*. *CoRR*, 2021. **abs/2101.09635**.
4. Edward Diener, R.E.L., and Jorden A. Cummings. *16.1 Personality Traits*. *INTRODUCTION TO PSYCHOLOGY* [cited 2022 4 May]; Available from: https://openpress.usask.ca/introductiontopsychology/chapter/personality-traits/?fbclid=IwAR00d3WL4AjLUxCU7r5bIc_wfU8Gd5TafJg4Erz5ylt6DZ5tGhEsxzglQ-s.
5. Pasztor, J.a.B., Gerda, *Attitudes of Generation Z Towards Instagram & Facebook – A Comparative Study*. 2021.
6. Muller, R. *Both Introverts and Extroverts Feel Drained From Socializing — Here's How to Deal*. 2019 [cited 2021 20 Sep.]; Available from: <https://thriveglobal.com/stories/introvert-extrovert-ambivert-social-anxiety-stress/>.
7. Megan, S. *7 Signs That You May Be An Ambivert & How To Thrive*. *Personal Growth 2020* [cited 2021 20 Sep.]; Available from: <https://www.mindbodygreen.com/articles/ambivert-meaning-and-signs>.
8. Edwards, V.V. *What is an Ambivert? Take the Quiz to See if You're an Introvert, Extrovert or Ambivert*. [cited 2021 17 Sep.]; Available from: <https://www.scienceofpeople.com/ambivert/>.
9. Trommer, M. *Introverts and Extroverts: Can We Get Along?* 2015 [cited 2022 8 July]; Available from: <https://news.arizona.edu/blog/introverts-and-extroverts-can-we-get-along>.
10. Speranza, S.a.R., Carmine Tommaso and Bruno, Barbara and Sgorbissa, Antonio, *A Model for the Representation of the Extraversion-Introversion Personality Traits in the Communication Style of a Social Robot*, in *2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. 2020. p. 75-81.
11. Pratama, B.Y.a.S., Riyanarto, *Personality classification based on Twitter text using Naive Bayes, KNN and SVM*, in *2015 International Conference on Data and Software Engineering (ICoDSE)*. 2015. p. 170-174.
12. Yousaf, A.a.U., Muhammad and Sadiq, Saima and Ullah, Saleem and Mirjalili, Seyedali and Rupapara, Vaibhav and Nappi, Michele, *Emotion Recognition by Textual Tweets Classification Using Voting Classifier (LR-SGD)*. *IEEE Access*, 2021. **9**: p. 6286-6295.
13. Wang, P.a.Y., Meng and Zhan, Xiangping and Tian, Mei and Si, Yingdong and Sun, Yu and Jiao, Longzhen and Wu, Xiaojie, *Predicting Self-Reported Proactive Personality Classification With Weibo Text and Short Answer Text*. *IEEE Access*, 2021. **PP**: p. 1-1.
14. Miner, A.S., L. Laranjo, and A.B. Kocaballi, *Chatbots in the fight against the*

- COVID-19 pandemic*. NPJ Digit Med, 2020. **3**: p. 65.
15. Chonnatee Rodsawang, P.T., Theeraporn Intawong, Apisit Sonong, Yosita Thitiwatthana and Suthat Chottanapund, *Designing a Competent Chatbot to Counter the COVID-19 Pandemic and Empower Risk Communication in an Emergency Response System*. *Outbreak, Surveillance, Investigation & Response (OSIR) Journal*, 2020. **13**(2).
 16. Faishal, D.E.C.a.A.F., *Classification of Big Five Personality Behavior Tendencies Based On Study Field with Twitter Analysis Using Support Vector Machine*, in *7th International Conference on Information Technology, Computer, and Electrical Engineering (ICITACEE)*. 2020. p. 140-145.
 17. Vermeulen, C.J.B.a.M.T.a.I.E., *The Language of Extroversion: Extraverted People Talk More Abstractly, Introverts Are More Concrete*. *Journal of Language and Social Psychology*, 2013. **32**(2): p. 191-201.
 18. Ravinder Ahuja, A.C., Shruti Kohli, Shaurya Gupta, and Pratyush Ahuja, *The Impact of Features Extraction on the Sentiment Analysis*. *Procedia Computer Science*, 2019. **152**: p. 341-348.
 19. Thomas W. Edgar, D.O.M., *In Research Methods for Cyber Security*. 2017.



APPENDIX

Table 10. The first answer sets (65% of Extroversion and 35% of Introversion)

Group	Sub-group	Intent Name	Response Text
Update COVID-19 Situation	Daily Cases	daily	ได้เลย!! นี่คือยอดจำนวนผู้ติดเชื้อโรคระบาดโควิด 19 ของวันนี้ที่คุณต้องการ
	Official Announcement	announce	เพื่อนๆสามารถอ่านรายละเอียดเกี่ยวกับข้อมาตรการบังคับใหม่ล่าสุดจากกรมควบคุมโรคได้ที่ลิงค์ข้างล่างนี้ค่ะ https://ddc.moph.go.th/viralpneumonia/measure.php
Information of COVID-19	General Characteristic	covid_char	Non-Text Response
	Spreading	spread_main	สงสัยเกี่ยวกับการแพร่ระบาดแบบไหนคะ? (s1-s3) S1: การแพร่ระบาดจากคนสู่คน S2: การแพร่ระบาดผ่านวัตถุ S3: การแพร่ระบาดผ่านสภาพอากาศ
		spread_per	หากมีการใกล้ชิด หรือสัมผัสกับผู้ติดเชื้อ หรือผู้ที่มีความเสี่ยง ให้เฝ้าดูอาการประมาณ 3 วัน จากนั้นอาจจะลองตรวจหาเชื้อด้วยตัวเองค่ะ
		spread_obj	แน่นอนว่าเชื้อโควิด-19 สามารถแพร่กระจายผ่านสิ่งของได้และอยู่ได้นานหลายวันหากไม่มีการทำความสะอาด เชื้อสามารถแปดเปื้อนผ่านมือแล้วจับโทรศัพท์มือถือ ใช้มือสัมผัสสิ่งของแล้วนำมาแตะบริเวณหน้า และหยิบอาหารเข้าปาก เป็นการติดเชื้อได้เช่นเดียวกับการสัมผัสเชื้อจากผู้ป่วยโดยตรง จึงควรต้องระมัดระวัง ทำความสะอาดพื้นผิวที่อาจแปดเปื้อนบ่อยๆ และล้างมือบ่อยๆ
		spread_envi	การแพร่เชื้อส่วนใหญ่ยังเป็นยังเป็นการอยู่ใกล้ชิด ไอจามรดกัน, มือไปสัมผัสพื้นผิวที่ปนเปื้อนเชื้อแล้วนำมาถูร่างกายผ่านการขยี้ตา และจุมูกและปากที่เป็นเนื้อเยื่ออ่อน (Mucosa) และเชื้อไวรัสโควิดสายพันธุ์โอมิครอนสามารถแพร่กระจายผ่านละอองฝอยที่ฟุ้งแบบเครื่องพ่นยาขยายหลอดลม
	Various species	species	Non-Text Response
Check-up risk and symptoms	Self-Assessment	self_assessment	ลิงค์นี้คือเว็บไซต์สำหรับการประเมินความเสี่ยงของตนเองเบื้องต้นค่ะ https://covid19.rajavithi.go.th/test/th_index.php
	COVID-19 Testing	covid_test	สงสัยเกี่ยวกับการตรวจหาเชื้อแบบไหนคะ? (t1-t2) t1: Antigen Test Kit (ATK) t2: RT-PCR (Polymerase chain reaction)
		atk	Non-Text Response

		pcr	Non-Text Response
Treatment Advices	Medicine	med	ยาสำหรับรักษาโรคโควิด-19 ได้แก่ Favipiravir, Remdisivir, Molnupivir และ Paxlovid นอกจากนี้จะมียาบรรเทาอาการอื่นๆ เพิ่มเติมแล้วแต่อาการของผู้ป่วยคนนั้น ส่วนยาสมุนไพรอย่างฟ้าทะลายโจรสามารถทานได้เป็นทางเลือก แม้จะยังไม่ถูกใช้เป็นตัวยาเพื่อรักษาโรคโควิด-19 โดยตรง
	Home Isolation	home_iso	เมื่อทราบว่าตนเองติดเชื้อและมีอาการไม่หนักมากถึงไม่มีอาการเลย แนะนำให้กักตัวแบบ home isolation ดังรูป
	Hospital	hospital	สำหรับผู้ป่วยติดเชื้อระดับสีเขียวและไม่สะดวกทำ home isolation อย่างแน่นอนหรือคิดว่าตัวเองมีอาการมากกว่าระดับสีเขียว ควรเข้ารับการรักษาที่โรงพยาบาลค่ะ ซึ่งแต่ละโรงพยาบาลจะมีการคิดค่าบริการแตกต่างกันไปแล้วแต่สถานที่นะคะ
	Hospital	hospital	สำหรับผู้ที่เข้ารับการรักษาด่วนใน Hospital แพทย์ต้องทำการประเมินอย่างละเอียดก่อน ผู้ที่เข้าเกณฑ์มีลักษณะดังนี้ <ul style="list-style-type: none"> - ผู้ป่วยยืนยันโควิด 19 ที่ไม่มีอาการ - หลังนอนโรงพยาบาล 4-7 วัน เมื่อไม่มีภาวะแทรกซ้อนให้พักต่อที่ Hospital จนครบ 10 วัน (และครบ 14 วันกรณีสงสัยเชื้อกลายพันธุ์) ผู้ป่วยโควิด 19 ที่มีอาการ แต่ไม่มีภาวะเสี่ยง หรือภาวะร่วม - หลังนอนโรงพยาบาล 4-7 วัน เมื่ออาการดีขึ้น ให้พักต่อที่ Hospital จนครบ 10.วัน (และครบ 14 วันกรณีสงสัยเชื้อกลายพันธุ์) ผู้ป่วยโควิด 19 ที่มีอายุน้อยกว่า 50 ปีที่ไม่มีอาการ หรือไม่มีภาวะเสี่ยง หรือภาวะอื่นร่วม - เข้าพัก รักษาสังเกตอาการที่ Hospital จนครบ 10.วัน (และครบ 14 วันกรณีสงสัยเชื้อกลายพันธุ์) โดยผู้ป่วยที่เข้ารับการรักษาด่วนที่ Hospital ควรต้องได้รับการตรวจ X-Ray ทุกราย หากปกติ ควรได้รับการรักษาด่วนที่โรงพยาบาล - แนะนำสิ่งจำเป็นในการเตรียมตัวก่อนเข้ารับการรักษาที่ Hospital -> บัตรประจำตัวประชาชนม, อุปกรณ์และเอกสารด้านสุขภาพ, เสื้อผ้าสำหรับ 10 หรือ 14 วัน (แล้วแต่บุคคล), ของใช้ส่วนตัวที่จำเป็น และอุปกรณ์เพื่อผ่อนคลาย เช่น หนังสือ คอมพิวเตอร์ เป็นต้น
Self-Pervent Advices	Mask & Alcohol	mask_al	กรุณาเลือกหัวข้อย่อๆดังต่อไปนี้ค่ะ (m1-m2) m1: แนะนำการใส่ Mask ที่เหมาะสม m2: แนะนำแอลกอฮอล์ล้างมือ

		mask	Non-Text Response
		alcohol	ผลิตภัณฑ์แอลกอฮอล์ทำความสะอาดมือ จะต้องใช้ปริมาณแอลกอฮอล์มากกว่า 70% โดยปริมาตร จึงจะสามารถฆ่าเชื้อโรคได้
	Transportation	transport	เมื่อมีความจำเป็นต้องใช้บริการขนส่งสาธารณะร่วมกับผู้อื่น ควรระมัดระวังการสัมผัสกับสิ่งต่างๆภายในรถ หมั่นล้างมือ งดคุยกับเพื่อนหรือคนอื่นๆที่ร่วมเดินทางไปกับเรา และสวมหน้ากากอนามัยให้ถูกต้องตลอดเวลา
	Routines	routine	เพื่อนๆสามารถทำตามรูปด้านล่างนี้ได้เลย รับรองปลอดภัยัวร์
	Vaccines	vac	สงสัยเกี่ยวกับข้อมูลเกี่ยวกับวัคซีนด้านไหนคะ? (v1-v2) v1: วัคซีนยี่ห้อต่างๆ v2: อาการข้างเคียงหลังฉีด
		vac_kind	Non-Text Response
		vac_eff	การมีอาการมีไข้ ปวดศีรษะ ปวดแขนบริเวณที่ฉีด เป็นอาการข้างเคียงที่สามารถเกิดขึ้นได้หลังจากการฉีดวัคซีน และสามารถลองปรึกษาแพทย์ได้ในกรณีที่มีอาการมากกว่าที่ขยตัวอย่างมาคะ
	Place	place	หากพื้นที่ที่ต้องไปเป็นที่สาธารณะ คนแออัด แนะนำให้สวมหน้ากากอนามัยตลอดเวลา และล้างมือบ่อยๆคะ จากนั้นให้ลองประเมินอาการและความเสี่ยงเมื่อเวลาผ่าน 2-3 วันคะ

Table 11. The second answer sets (40% of Extroversion and 60% of Introversion)

Group	Sub-group	Intent Name	Response Text
Update COVID-19 Situation	Daily Cases	daily	นี่คือยอดจำนวนผู้ติดเชื้อโรคระบาดโควิด 19 ของวันนี้ค่ะ
	Official Announcement	announce	สามารถติดตามอำนาจมาตรการจากกรมควบคุมโรคได้ที่ลิงค์นี้ คลิก! https://ddc.moph.go.th/viralpneumonia/measure.php
Information of COVID-19	General Characteristic	covid_char	Non-Text Response
	Spreading	spread_main	กรุณาเลือกหัวข้อย่อยดังต่อไปนี้ค่ะ (s1-s3) s1: การแพร่ระบาดจากคนสู่คน s2: การแพร่ระบาดผ่านวัตถุ s3: การแพร่ระบาดผ่านสภาพอากาศ
		spread_per	หากทราบว่า มีเพื่อนหรือคนใกล้ชิดติดเชื้อหรือกลายเป็นกลุ่มคนที่มีความเสี่ยง ให้เฝ้าดูอาการ 3 วัน และลองตรวจหาเชื้อด้วยตัวเองค่ะ
		spread_obj	เชื้อ โควิด-19 สามารถแพร่กระจายผ่านสิ่งของได้และอยู่ได้นานหลายวันหากไม่มีการทำความสะอาด บางทีอาจมีการแปดเปื้อนผ่านมือที่จับสัมผัสสิ่งของแล้วนำมาแตะบริเวณหน้า หรือหยิบอาหารเข้าปาก ก็สามารถติดเชื้อได้ เช่นเดียวกับการสัมผัสเชื้อจากผู้ป่วยโดยตรง จึงควรทำความสะอาดพื้นผิวและล้างมือบ่อยๆ
		spread_envi	การแพร่เชื้อติดต่อหลักยังเป็นการอยู่ใกล้ชิด ไอจามรดกัน หรือมือไปสัมผัสพื้นผิวที่ปนเปื้อนเชื้อแล้วนำมาสู่วางกายผ่านการขี้น้ำ และจมูกหรือปากที่เป็นเนื้อเยื่ออ่อน (Mucosa) ยังไม่มีหลักฐานว่าสามารถแพร่เชื้อผ่านอากาศได้โดยตรง หากแต่ว่าเป็นสายพันธุ์โอมิครอน สามารถแพร่กระจายเชื้อผ่านละอองฝอยที่พุ่งแบบเครื่องพ่นยาขายหลอดลม
	Various species	species	Non-Text Response
Check-up risk and symptoms	Self-Assessment	self_assessment	สามารถลองประเมินความเสี่ยงการติดเชื้อโรคโควิด-19ได้ที่ลิงค์นี้ คลิก! https://covid19.rajavithi.go.th/test/th_index.php
	COVID-19 Testing	covid_test	กรุณาเลือกหัวข้อย่อยดังต่อไปนี้ค่ะ (t1-t2) t1: Antigen Test Kit (ATK) t2: RT-PCR (Polymerase chain reaction)
		atk	Non-Text Response
		pcr	Non-Text Response
Treatment	Medicine	med	ยาสำหรับรักษาโรคโควิด-19 มีหลักๆประมาณ 4 ชนิด ดังนี้

Advices			1.Favipiravir 2.Remdisivir 3.Molnupivir 4.Paxlovid นอกจากนี้อาจเพิ่มยาบรรเทาอาการอื่นๆแล้วแต่อาการ ส่วนยาสมุนไพรอย่างฟ้าทะลายโจร ไม่ใช่ยารักษาโดยตรง แต่สามารถทานได้เป็นทางเลือก
	Home Isolation	home_iso	หากติดเชื้อแต่อาการไม่หนักมาก หรือไม่มีอาการเลย แนะนำให้กักตัวแบบ home isolation ดังรูป
	Hospital	hospital	สำหรับผู้ป่วยติดเชื้อระดับสีเขียวแต่ไม่สะดวกทำ home isolation หรือมีอาการรุนแรง ไม่ต่ำกว่าระดับสีเขียว แนะนำให้เข้ารับการรักษาที่โรงพยาบาลค่ะ ซึ่งแต่ละโรงพยาบาลจะมีการคิดค่าบริการแตกต่างกันไปแล้วแต่สถานที่นะคะ
	Hospitel	hospitel	สำหรับผู้ที่เข้ารับการรักษาด้าน Hospitel แพทย์ต้องทำการประเมินอย่างละเอียดก่อนส่งตัวเข้ามาพัก โดยผู้ที่เข้าเกณฑ์ได้แก่ <ul style="list-style-type: none"> - ผู้ป่วยยืนยันโควิด 19 ที่ไม่มีอาการ - หลังนอนโรงพยาบาล 4-7 วัน เมื่อไม่มีภาวะแทรกซ้อนให้พักต่อที่ Hospitel จนครบ 10.วัน (และครบ 14 วันกรณีสงสัยเชื้อกลายพันธุ์) ผู้ป่วยโควิด 19 ที่มีอาการ แต่ไม่มีภาวะเสี่ยง หรือภาวะร่วม - หลังนอนโรงพยาบาล 4-7 วัน เมื่ออาการดีขึ้น ให้พักต่อที่ Hospitel จนครบ 10.วัน (และครบ 14 วันกรณีสงสัยเชื้อกลายพันธุ์) ผู้ป่วยโควิด 19 ที่มีอายุน้อยกว่า 50 ปีที่ไม่มีอาการ หรือไม่มีภาวะเสี่ยง หรือภาวะอื่นร่วม - เข้าพัก รักษาสังเกตอาการที่ Hospitel จนครบ 10.วัน (และครบ 14 วันกรณีสงสัยเชื้อกลายพันธุ์) โดยผู้ป่วยที่เข้ารับการรักษาด้าน Hospitel ควรต้องได้รับการตรวจ X-Ray ทุกราย หากผิดปกติ ควรได้รับการรักษาด้านที่โรงพยาบาล - สิ่งที่ต้องเตรียมก่อนเข้ารับการรักษา -> 1.บัตรประจำตัวประชาชน 2.อุปกรณ์และเอกสารด้านสุขภาพต่าง ๆ 3.เสื้อผ้าสำหรับ 10 หรือ 14 วัน (แล้วแต่บุคคล) 4.ของใช้ส่วนตัวที่จำเป็น 5.อุปกรณ์เพื่อผ่อนคลาย เช่น หนังสือ คอมพิวเตอร์ เป็นต้น
Self-Pervent Advices	Mask & Alcohol	mask_al	สงสัยเกี่ยวกับข้อมูลเกี่ยวกับด้านไหนคะ? (m1-m2) m1: แนะนำการใส่ Mask ที่เหมาะสม m2: แนะนำแอลกอฮอล์ล้างมือ
		mask	Non-Text Response
		alcohol	ผลิตภัณฑ์แอลกอฮอล์ทำความสะอาดมือ จะต้องมึปริมาณแอลกอฮอล์ไม่ต่ำกว่า 70% โดยปริมาตร เมื่อละลายกับน้ำจะสามารถแพร่ผ่านเชื้อหุ้มเซลล์เชื้อโรคได้ดี ทำให้โปรตีนเสียสภาพ เชื้อหุ้มเซลล์แตก และเข้าไปรบกวนระบบของเชื้อ

			โรค ส่งผลให้เชื้อโรคตาย
Transportation	transport		หากต้องใช้บริการขนส่งมวลชนต่างๆ สิ่งที่ต้องระวังคือการสัมผัสกับสิ่งต่างที่อาจมีเชื้อโรคปนเปื้อน ดังนั้นควรล้างมือทุกครั้งหลังใช้บริการ หลีกเลี่ยงการพูดคุยบนรถ และสวมหน้ากากอนามัยตลอดการเดินทาง
Routines	routine		นี่คือข้อเสนอแนะในการใช้ชีวิตประจำวันเพื่อหลีกเลี่ยงการติดเชื้อโรคโควิด-19
Vaccines	vac		กรุณาเลือกหัวข้อย่อยดังต่อไปนี้ค่ะ (v1-v2) v1: วัคซีนยี่ห้อต่างๆ v2: อาการข้างเคียงหลังฉีด
	vac_kind		Non-Text Response
	vac_eff		หากมีอาการมีไข้ ปวดศีรษะ ปวดแขนบริเวณที่ฉีด นั่นคืออาการข้างเคียงที่อาจเกิดขึ้นได้หลังจากการฉีดวัคซีน ไม่อันตรายเท่าไร แต่ถ้ามีอาการมากกว่านี้ หากกังวลสามารถลองปรึกษาแพทย์ได้ค่ะ
Place	place		ถ้าออกไปเที่ยวสนุกในที่ ๆ คนเยอะๆ ต้องคอยระวังให้มากๆ ค่ะ แนะนำให้สวมหน้ากากตลอดเวลา และล้างมือบ่อยๆ ค่ะ จากนั้นก็ให้คอยสังเกตอาการตัวเอง และเพื่อนๆ หลังจากผ่านไป 2-3 วันค่ะ

Table 12. The third answer sets (50% of Extroversion and 50% of Introversion)

Group	Sub-group	Intent Name	Response Text
Update COVID-19 Situation	Daily Cases	daily	ได้เลย!! นี่คือยอดจำนวนผู้ติดเชื้อโรคระบาดโควิด 19 ของวันนี้ที่คุณต้องการใช้มั๊ยคะ?
	Official Announcement	announce	เพื่อนๆสามารถอ่านรายละเอียดเกี่ยวกับข้อมาตรการบังคับใหม่ล่าสุดจากกรมควบคุมโรคได้ที่ลิงค์ข้างล่างนี้ค่ะ https://ddc.moph.go.th/viralpneumonia/measure.php
Information of COVID-19	General Characteristic	covid_char	Non-Text Response
	Spreading	spread_main	สงสัยเกี่ยวกับการแพร่ระบาดแบบไหนคะ? (s1-s3) s1: การแพร่ระบาดจากคนสู่คน s2: การแพร่ระบาดผ่านวัตถุ s3: การแพร่ระบาดผ่านสภาพอากาศ
		spread_per	หากทราบว่ามีเพื่อนหรือคนใกล้ชิดติดเชื้อหรือกลายเป็นกลุ่มคนที่มีความเสี่ยงให้เฝ้าดูอาการ 3 วัน และลองตรวจหาเชื้อด้วยตัวเองค่ะ
		spread_obj	แน่นอนว่าเชื้อโควิด-19 สามารถแพร่กระจายผ่านสิ่งของได้และอยู่ได้นานหลายวันหากไม่มีการทำความสะอาด เชื้อสามารถแปดเปื้อนผ่านมือแล้วจับโทรศัพท์มือถือ ใช้มือสัมผัสสิ่งของแล้วนำมาแตะบริเวณหน้า และหยิบอาหารเข้าปาก เป็นการติดเชื้อได้เช่นเดียวกับการสัมผัสเชื้อจากผู้ป่วยโดยตรง จึงควรต้องระมัดระวัง ทำความสะอาดพื้นผิวที่อาจแปดเปื้อนบ่อยๆ และล้างมือบ่อยๆ
		spread_envi	การแพร่เชื้อส่วนใหญ่ยังเป็นยังเป็นการอยู่ใกล้ชิด ไอจามรดกัน, มือไปสัมผัสพื้นผิวที่ปนเปื้อนเชื้อแล้วนำมาสู่ร่างกายผ่านการขยี้ตา และจมูกและปากที่เป็นเนื้อเยื่ออ่อน (Mucosa) และเชื้อไวรัสโควิดสายพันธุ์โอมิครอนสามารถแพร่กระจายผ่านละอองฝอยที่ฟุ้งแบบเครื่องพ่นยาขยายหลอดลม
	Various species	species	Non-Text Response
Check-up risk and symptoms	Self-Assessment	self_assessment	ลิงค์นี้คือเว็บไซต์สำหรับการประเมินความเสี่ยงของตนเองเบื้องต้นค่ะ https://covid19.rajavithi.go.th/test/th_index.php
	COVID-19 Testing	covid_test	สงสัยเกี่ยวกับการตรวจหาเชื้อแบบไหนคะ? (t1-t2) t1: Antigen Test Kit (ATK) t2: RT-PCR (Polymerase chain reaction)
		atk	Non-Text Response
		pcr	Non-Text Response

Treatment Advices	Medicine	med	ยาสำหรับรักษาโรคโควิด-19 ได้แก่ Favipiravir, Remdisivir, Molnupivir และ Paxlovid นอกจากนี้จะมียาบรรเทาอาการอื่นๆ เพิ่มเติมแล้วแต่อาการของผู้ป่วยคนนั้น ส่วนยาสมุนไพรอย่างฟ้าทะลายโจรสามารถทานได้เป็นทางเลือก แม้จะยังไม่ถูกใช้เป็นตัวยาเพื่อรักษาโรคโควิด-19 โดยตรง
	Home Isolation	home_iso	เมื่อทราบว่าตนเองติดเชื้อและมีอาการไม่หนักมากถึงไม่มีอาการเลย แนะนำให้กักตัวแบบ home isolation ดังรูป
	Hospital	hospital	สำหรับผู้ป่วยติดเชื้อระดับสีเขียวและไม่สะดวกทำ home isolation อย่างแน่นอนหรือคิดว่าตัวเองมีอาการมากกว่าระดับสีเขียว ควรเข้ารับการรักษาที่โรงพยาบาลค่ะ ซึ่งแต่ละโรงพยาบาลจะมีการคิดค่าบริการแตกต่างกันไปแล้วแต่สถานที่นะคะ
	Hospital	hospital	<p>สำหรับผู้ที่เข้ารับการรักษาดำเนิน Hospitel แพทย์ต้องทำการประเมินอย่างละเอียดก่อนส่งตัวเข้ามาพัก โดยผู้ที่เข้าเกณฑ์ได้แก่</p> <ul style="list-style-type: none"> - ผู้ป่วยยืนยันโควิด 19 ที่ไม่มีอาการ - หลังนอนโรงพยาบาล 4-7 วัน เมื่อไม่มีภาวะแทรกซ้อนให้พักต่อที่ Hospitel จนครบ 10.วัน (และครบ 14 วันกรณีสงสัยเชื้อกลายพันธุ์) ผู้ป่วยโควิด 19 ที่มีอาการ แต่ไม่มีภาวะเสี่ยง หรือภาวะร่วม - หลังนอนโรงพยาบาล 4-7 วัน เมื่ออาการดีขึ้น ให้พักต่อที่ Hospitel จนครบ 10.วัน (และครบ 14 วันกรณีสงสัยเชื้อกลายพันธุ์) ผู้ป่วยโควิด 19 ที่มีอายุน้อยกว่า 50 ปีที่ไม่มีอาการ หรือไม่มีภาวะเสี่ยง หรือภาวะอื่นร่วม - เข้าพัก รักษาสังเกตอาการที่ Hospitel จนครบ 10.วัน (และครบ 14 วันกรณีสงสัยเชื้อกลายพันธุ์) โดยผู้ป่วยที่เข้ารับการรักษาดำเนิน Hospitel ควรต้องได้รับการตรวจ X-Ray ทุกราย หากผิดปกติ ควรได้รับการรักษาดำเนินที่โรงพยาบาล <p>สิ่งที่ต้องเตรียมก่อนเข้ารับการรักษา -> 1.บัตรประจำตัวประชาชน 2.อุปกรณ์และเอกสารด้านสุขภาพต่าง ๆ 3.เสื้อผ้าสำหรับ 10 หรือ 14 วัน (แล้วแต่บุคคล) 4.ของใช้ส่วนตัวที่จำเป็น 5.อุปกรณ์เพื่อผ่อนคลาย เช่น หนังสือ คอมพิวเตอร์ เป็นต้น</p>
Self-Pervent Advices	Mask & Alcohol	mask_al	สงสัยเกี่ยวกับข้อมูลเกี่ยวกับด้านไหนคะ? (m1-m2) m1: แนะนำการใส่ Mask ที่เหมาะสม m2: แนะนำแอลกอฮอล์ล้างมือ
		mask	Non-Text Response

		alcohol	ผลิตภัณฑ์แอลกอฮอล์ทำความสะอาดมือ จะต้องมีความเข้มข้นมากกว่า 70% โดยปริมาตร จึงจะสามารถฆ่าเชื้อโรคได้
	Transportation	transport	เมื่อมีความจำเป็นต้องใช้บริการขนส่งสาธารณะร่วมกับผู้อื่น ควรระมัดระวังการสัมผัสกับสิ่งต่างๆภายในรถ หมั่นล้างมือ งดคุยกับเพื่อนหรือคนอื่นๆที่ร่วมเดินทางไปกับเรา และสวมหน้ากากอนามัยให้ถูกต้องตลอดเวลา
	Routines	routine	เพื่อนๆสามารถทำตามรูปด้านล่างนี้ได้เลย รับรองปลอดภัยชัวร์
	Vaccines	vac	สงสัยเกี่ยวกับข้อมูลเกี่ยวกับวัคซีนด้านไหนคะ? (v1-v2) v1: วัคซีนยี่ห้อต่างๆ v2: อาการข้างเคียงหลังฉีด
		vac_kind	Non-Text Response
		vac_eff	การมีอาการมีไข้ ปวดศีรษะ ปวดแขนบริเวณที่ฉีด เป็นอาการข้างเคียงที่สามารถเกิดขึ้นได้หลังจากการฉีดวัคซีน และสามารถลองปรึกษาแพทย์ได้ ในกรณีที่มีอาการมากกว่าที่ยกตัวอย่างมาค่ะ
	Place	place	ถ้าออกไปเที่ยวสนุกในที่ๆคนเยอะๆ ต้องคอยระวังให้มากๆนะ แนะนำให้สวมหน้ากากตลอดเวลา และล้างมือบ่อยๆนะ จากนั้นก็ให้คอยสังเกตอาการตัวเอง และเพื่อนๆหลังจากผ่านไป 2-3 วันค่ะ

VITA

NAME Sakolwan Peetaneelavat

DATE OF BIRTH 25 March 1998

PLACE OF BIRTH Bangkok, Thailand.

INSTITUTIONS ATTENDED

2002-2006	Sesawech School
2006-2016	Khema Siri Memorial School
2016-2020	Degree of Bachelor of Science in Computer Science, Department of Mathematics and Computer Science, Faculty of Science, Chulalongkorn University
2020-present	Degree of Master of Science in Computer Science and Information Technology, Department of Mathematics and Computer Science, Faculty of Science, Chulalongkorn University

HOME ADDRESS 59/101 Soi Ratchaphruek 9, Bang Chueak Nang Sub-district, Taling Chan District, Bangkok, 10170, Thailand.

AWARD RECEIVED The Hitachi Trophy 2020 Competition. Issued by Hitachi Asia (Thailand) Co., Ltd (Jul 2020)

- Honorable mention
- Excellent Pitching
- Popular Vote