Detection of Text-Based Depression on Thai Social Media Using Deep Learning

Jaturawit Ardharn and Werayuth Charoenruengkit*

Department of Computer Science, Faculty of Science, Srinakharinwirot University, Bangkok, Thailand
*Corresponding author, Email: werayuth@g.swu.ac.th

Abstract

Social media has become an increasingly influential platform for emotional expression, particularly during the COVID-19 pandemic, which has restricted physical mobility and fueled a surge in internet usage. However, depression is a growing mental health concern, and social media can be a breeding ground for suicidal ideation. Furthermore, individuals experiencing depression frequently communicate their emotional states through social media posts. The potential to detect depression markers within these messages could offer valuable insights into their emotional well-being, paving the way for the provision of appropriate support. This study aimed to develop a method for depression detection in Thai social media texts using three neural network algorithms including BiLSTM, Thai-BERT, and WangchanBERTa with Natural Language Processing (NLP) techniques. The authors evaluated the performance of these algorithms using a dataset of 3,100 Thai text messages that had been scraped from Twitter and labeled by psychologists. WangchanBERTa with the Newmm tokenizer achieved the best performance, with an F1-score of 74.2%, Recall of 74.2%, Precision of 74.7%, and Accuracy of 74.2%. These results suggest that neural network algorithms can be used effectively to detect depression in Thai social media texts and that the choice of tokenizer can significantly impact performance. Furthermore, a critical analysis of the dataset and limitations that may impact model performance was incorporated.

Keywords: *Depression, Social Media, Natural Language Processing, Text Classification, Deep Learning*

1. Introduction

Social media has become an integral part of daily life, significantly influencing human interactions and emotional expression (The Electronic Transactions Development Agency, 2022). The COVID-19 pandemic has further amplified the importance of social media as a means of communication and connection, particularly in the context of restricted physical mobility. As internet usage surges, the potential for utilizing social media data to identify and address mental health concerns, such as depression, has gained significant attention.

Depression is a prevalent mental health issue, affecting over 280 million people worldwide (World Health Organization, 2021). In Thailand, approximately 20 percent of individuals with depression still lack access to necessary services, despite the implementation of policies for screening, treating, and monitoring depression (Department of Mental Health, 2021). The consequences of depression can be severe, including an increased risk of suicide, underscoring the need for effective detection and intervention strategies (Komsan Kiatrungrit, 2016).

Several studies have demonstrated the efficient use of Natural Language Processing (NLP) techniques for identifying and classifying messages related to depression and other mental health conditions on social media platforms (Hämäläinen, Patpong, Alnajjar, Partanen, & Rueter, 2021; Husseini Orabi, Buddhitha, Husseini Orabi, & Inkpen, 2018; Oh Jihoon, Yun Kyongsik, Maoz Uri, Kim Tae-Suk, & Jeong-

Ho, 2019; Q. Cong et al., 2018; S. Mahasiriakalayot, T. Senivongse, & Taephant, 2022; Zogan, Wang, Jameel, & Xu, 2020). However, most analyses have focused on binary classification, which involves determining whether or not a certain social media post indicates depression. Given the diverse manifestations and varying severity of depression symptoms, a more nuanced approach is necessary to accurately detect and assess the level of depression expressed in social media texts.

The Thai language presents unique challenges for NLP due to the absence of word boundaries and limited resources compared to languages like English (Hämäläinen et al., 2021). These challenges complicate the segmentation of text into meaningful units and affect tasks such as part-of-speech tagging and syntax analysis, making it crucial to develop tailored approaches for processing Thai text in the context of depression detection.

This study aims to address the following research objectives:

1) Develop text classification algorithms using different neural network architectures, namely bi-directional long-short term (BiLSTM) (S. Mahasiriakalayot et al., 2022), Thai-Bidirectional Encoder Representations from Transformer (Thai-BERT) (Hämäläinen et al., 2021), and WangchanBERTa (Lowphansirikul, Polpanumas, Jantrakulchai, & Nutanong, 2021), for detecting depression within Thai social media texts, taking into account the specific challenges posed by the Thai language.
2) Assess and compare the performance of these text classification algorithms in combination with different word tokenization techniques to identify the most effective approach for Thai text processing in the context of depression detection.

To achieve these objectives, this study utilizes a depression dataset from (S. Mahasiriakalayot et al., 2022), consisting of 3,100 Thai Twitter messages labeled by psychologists. By focusing on these state-of-the-art neural network algorithms, the aim is to develop an effective and robust system for detecting depression in Thai social media texts. Furthermore, error analysis was performed to investigate the limitations of the algorithms and the dataset based on the classification results, providing valuable insights for future research and improvements in this domain.

The findings of this study have the potential to clarify mental health interventions and policies, ultimately improving access to care and support for individuals with depression in Thailand. By leveraging the power of social media and advanced NLP techniques, this research contributes to the development of effective tools for early detection and timely intervention, addressing a crucial gap in the current mental health landscape.

2. Objectives

1) The principal aim of this paper is to develop a text classification algorithm tailored for the detection of depression within Thai social media texts.
2) To assess and compare the performance of text classification algorithms utilizing different word tokenization techniques.

3. Materials and Methods

In this section, the authors discuss the previous research related to Thai text classification problems. This paper constructs a text classification algorithm for depression detection from Thai language texts. An overview of the approach is shown in three steps in Figure 1. The first is data acquisition, wherein the dataset from (S. Mahasiriakalayot et al., 2022) was requested. The second step is data pre-processing. PyThaiNLP was utilized for data cleansing, removing emojis and URLs (these are not related to the analysis of Thai language texts), correcting misspellings to make the text easier to read and understand accurately, tokenizing words using three different word tokenizers, and stratifying the data into train and test sets. The final step was training the model. The authors employed the SMOTE technique before word embedding training three neural network algorithms and evaluating the models. The experiments were conducted by evaluating the data with three text classification algorithms. The following subsection describes the text classification algorithms and word tokenization techniques.

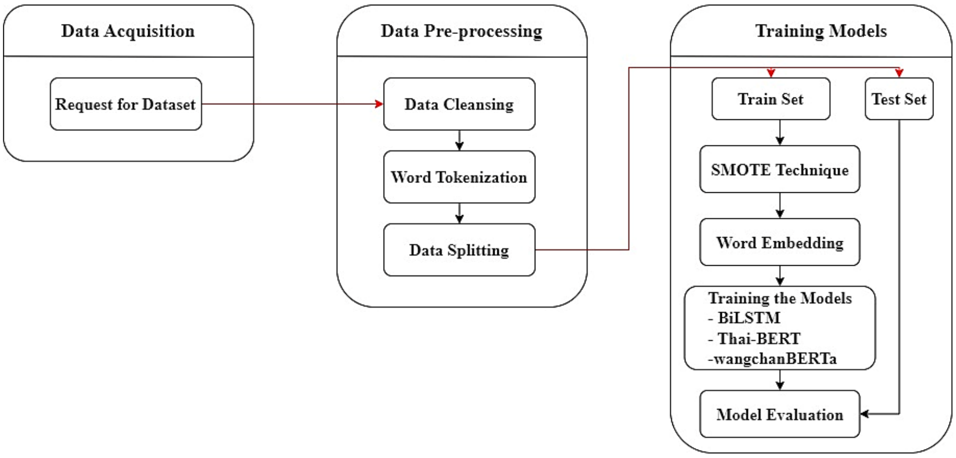

Figure 1 An overview of the approach, modified from (S. Mahasiriakalayot et al., 2022)

3.1 Neural Network Architectures for text classification

Text classification is a natural language processing (NLP) task that entails categorizing or labeling text documents into predefined categories or classes. Its objective is to automatically evaluate a given text document and classify it into one or more predetermined categories. This form of supervised learning entails training a model on a labeled dataset containing text examples and their corresponding categories.

Several neural network algorithms have been used for the depression detection task, e.g., long-shot term memory (LSTM), bi-directional long-shot term memory (BiLSTM) (Hämäläinen et al., 2021; Husseini Orabi et al., 2018), convolutional neural network (CNN) (Husseini Orabi et al., 2018) and bidirectional encoder representations form transformer (BERT) (Hämäläinen et al., 2021). Among these algorithms, LSTM, BiLSTM, and BERT are the most well-known and have been proven to achieve state-of-the-art results. In this paper, three text classification algorithms that have been developed are used and assessed in the following sections.

The first algorithm, BiLSTM, is a type of recurrent neural network (RNN) architecture utilized in natural language processing (NLP) and sequence modeling tasks. It extends the capabilities of the Long Short-Term Memory (LSTM) algorithm by addressing the vanishing gradient problem. Unlike LSTM, BiLSTM processes input sequences in both forward and backward directions. This is accomplished by employing two separate hidden layers: one that reads the input sequence from left to right (forward) and another that reads it from right to left (backward). The outputs from both directions are typically concatenated or combined in some manner. The bidirectional nature of the algorithm allows it to capture information from both past and future contexts, rendering it particularly effective for tasks where understanding the context in both directions is crucial. In NLP, for instance, comprehending the context of a word based on both preceding and succeeding words can significantly enhance the performance of the algorithms (Graves & Schmidhuber, 2005).

An overview of the architecture of this study's BiLSTM is shown in Figure 2. After the word tokenization process, Thai2fit was utilized, which is a word-embedding model specifically designed for the Thai language. In the BiLSTM layer, which serves as the core of the BiLSTM model, two LSTMs are employed: one processes the sequence of words from left to right, and another processes the sequence from right to left. The text is encoded into numeric form based on the maximum length of the text in the dataset. Subsequently, in the dense layer, the output of the BiLSTM layer is transformed into a vector of dimension d. Finally, the output layer takes the output of the dense layer and produces the final output of the model.

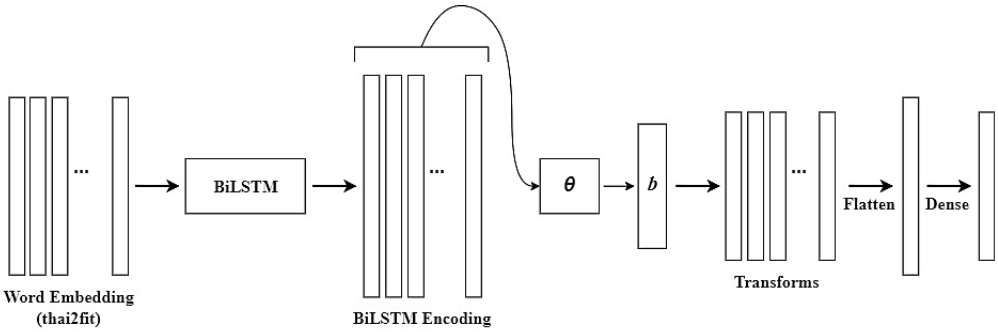

Figure 2 An overview of the architecture of the BiLSTM, modified from (Laosen, Laosen, & Ardharn, 2023)

The second algorithm, Thai-BERT, was developed by ThAIKeras using BERT, a transformer-based algorithm developed by Google that has demonstrated remarkable success in various natural language processing (NLP) tasks. BERT leverages the encoder component exclusively, responsible for transforming words within a sentence into vector representations. This enables the encoder to function as a Language Model (LM). BERT, therefore, expands existing encoders by incorporating an additional model. To facilitate its classification capabilities, BERT utilizes the generated vector representations from the encoder and employs calculations to arrive at an answer, echoing the output format of a traditional LM. It was trained on a dataset from the Thai language Wikipedia, which amounts to only a relatively small 515 megabytes. Consequently, the algorithm can only accommodate messages with a maximum length of 128 subwords, which impairs its performance with longer sequences (Hämäläinen et al., 2021). An overview of the architecture of the BERT family in this study is shown in Figure 3.

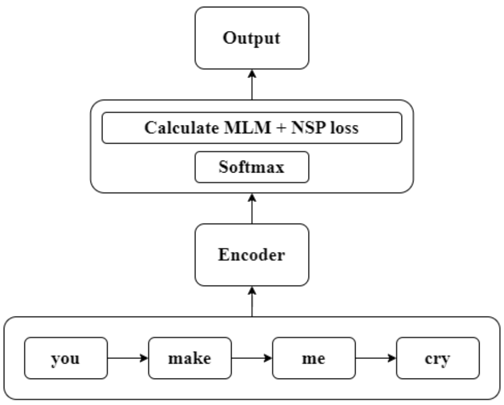

Figure 3 An overview of the architecture of the BERT family, modified from (Laosen et al., 2023)

The final algorithm, WangchanBERTa, is a Thai language pre-trained model. The architecture of this pretrained algorithm is based on RoBERTa, an improved version of BERT trained with more data and a different training strategy, developed by the Vidyasirimedhi Institute of Science and Technology (VISTEC). The language model has been trained on Thai language datasets obtained from various sources, including news, Wikipedia, social media messages, and information extracted from websites crawled on the internet. The total data size amounts to 78.5 GB. The performance of the language model was evaluated through fine-tuning and achieved the highest micro-averaged F1 score on 5 out of 6 datasets, which comprise a test dataset in the text classification and token classification problems, compared to the baseline model and the multilingual language model (Lowphansirikul et al., 2021).

3.2 Word Tokenization Techniques

In English, words in a sentence are separated by white spaces, making it easier to tokenize words. On the other hand, word tokenization in Thai is an important technique in natural language processing. Because the words in the sentence are stuck together, there are no white spaces between words. This makes it difficult to execute word tokenization. If the words are divided incorrectly, it will change the meaning or may have no meaning. For example, the word “ตากลม” if we tokenize “ตา|กลม” means “round eye” and “ตาก|ลม” means “feel the wind”. Therefore, there is a comparison of the use of three different word tokenization techniques in this paper.

The first technique, Newmm is a default word tokenization algorithm in PyThaiNLP (Thai Natural Language Processing in Python), which is dictionary-based maximum matching with Thai words and utilizes Thai character clustering (Phatthiyaphaibun et al., 2023).

The second technique, Attacut, is a word tokenization library in Python developed by the PyThaiNLP project for the Thai language. Attacut is designed to offer a fast and accurate method for segmenting Thai text into words. The name "Attacut" is derived from the term "Attacut-cut," where "cut" refers to the segmentation or tokenization of text. The library is built on top of deep learning techniques and employs a neural network architecture to predict word boundaries in Thai sentences (Chormai, Prasertsom, & Rutherford, 2019).

The last technique, Byte Pair Encoding (BPE), is a data compression technique adapted for subword tokenization in NLP. It is utilized to represent words as sequences of subword units to address vocabulary-related challenges and has been widely used in various natural language processing tasks, including machine translation, text summarization, and sentiment analysis. One notable application of BPE is in the training of language models, where it helps handle the open vocabulary problem by allowing the algorithm to generate or predict subword units rather than entire words (Sennrich, Haddow, & Birch, 2015).

3.3 Dataset Construction

3.3.1 Data Acquisition

The dataset used in this paper has been requested from a previous experiment conducted by the authors (S. Mahasiriakalayot et al., 2022), collected from Twitter in the Thai language during January – April 2021 by searching hashtag #ซึมเศร้า (#depressed), #โรคซึมเศร้า (#depressive disorder), #ภาวะซึมเศร้า (#depression) and #ฆ่าตัวตาย (#suicide). A total of 3,100 tweets were obtained with data privacy protection to ensure participant anonymity and comply with ethical data handling practices. Usernames associated with each tweet were pseudonymized before data dissemination to labelers.

3.3.2 Data labeling

The dataset was labeled by psychologists from the Faculty of Psychology at Chulalongkorn University. They categorized the data into five classes: Anhedonic, Suicidal Ideation, Guilty Feelings, Sleep Problems, and Other (i.e., not a sign of depression). The criteria for each class were derived from the Thai Mental Health Questionnaire (TMHQ) optimization (Wongaptikaseree, Yomaboot, Katchapakirin, & Kaewpitakkun, 2020). Table 1 shows the resulting dataset along with samples from each class. It should be noted that this is simply a translation of the original Thai tweets.

Table 1 Examples of tweets in the dataset of each class

Sign of Depression	Tweet messages (count)	Example (English Translation)
Anhedonic	1,274	The things around us are even more hurtful😭😔.
Suicidal Ideation	551	I took 50 pills to commit suicide but didn't die.
Guilty Feelings	185	I'm sorry that I couldn't be a good daughter, Mom😔.
Sleep Problems	40	I don't want to take medicine to help myself sleep well.
Other	1,050	This world is big. But loneliness is bigger than the world.

3.3.3 Data Pre-processing

The dataset was cleaned before using the training algorithms. This paper is interested only in Thai language texts, including the removal of emojis and URLs, as these elements are not integral to the linguistic features of the Thai language and could negatively impact the model's performance. Additionally, spelling errors were corrected to ensure accurate and efficient data processing and analysis. Each tweet was tokenized by using three tokenizers from section 3.2 (Newmm, Attacut, and BPE), and the tokens were compared with the vocabulary in the Thai2Fit word vector corpus from PyThaiNLP, containing 51,358 words embedded in 300 dimensions. The authors utilized a word cloud to visualize word frequency in each class, as depicted in Figure 4.

After pre-processing the data, the scikit-learn library was used to stratify the dataset into 90% training data and 10% test data. Since the dataset was class imbalanced and the number of data in some classes was small, counting text messages in Table 1, an enhancement of the Synthetic Minority Oversampling Technique (SMOTE) (Chawla, Bowyer, Hall, & Kegelmeyer, 2011) was used to increase the amount of data in small classes. The total number of training data was 5,735 embedded words.

Figure 4 Word cloud of tweets in the dataset of each class

4. Results and Discussion

The experiment comprises word tokenization and text classification processes. Both steps were conducted using Python 3, PyThaiNLP, scikit-learn machine learning libraries (Pedregosa et al., 2011), and Transformers version 4.6.0 (Wolf et al., 2019). The evaluation of text classification results was compared among three classification algorithms: BiLSTM, Thai-BERT, and WangchanBERTa. In the BiLSTM algorithm, the authors followed related work (S. Mahasiriakalayot et al., 2022), which used BiLSTM with the ReLU activation function as a baseline. The default parameters provided from the sklearn's library and HuggingFace's transformers (Wolf et al., 2019) were used for the classification. During the model training phase, the authors experimented with the following hyperparameters: learning rate (1e-4 and 1e-5), number of epochs (50 and 100), and dropout rate (0 to 1). Based on the experimental results, a learning rate of 1e-5, a dropout rate of 0.2, and 100 epochs for the training process were selected. However, the authors opted for a shorter training duration of 50 epochs for the Thai-BERT and WangchanBERTa models.

4.1 Experiment 1: The results of comparing text classification performance

This experiment aimed to compare the performance of three classification algorithms: BiLSTM, Thai-BERT, and WangchanBERTa with default word tokenization. The performance of each multiclass classification was then calculated using the weighted-averaging method, as reported in Table 2. WangchanBERTa exhibited the best performance, achieving the highest F1-score of 73.2%, Recall of 73.2%, Precision of 73.3%, and Accuracy of 73.2%. The other algorithms showed similar metric values.

Table 2 Classification performance for each classification algorithm

Classification Techniques with default tokenizer	F1-score	Recall	Precision	Accuracy
BiLSTM+ReLU (Attacut)	0.683	0.684	0.685	0.684
Thai-BERT (Newmm)	0.686	0.687	0.686	0.687
WangchanBERTa (BPE)	**0.732**	**0.732**	**0.733**	**0.732**

4.2 Experiment 2: Comparison of text classification performance in various word tokenizer

This experiment aimed to compare the performance of three classification algorithms: BiLSTM, Thai-BERT, and WangchanBERTa, with various word tokenization methods: Newmm, Attacut, and BPE. Since Thai-BERT and WangchanBERTa outperformed BiLSTM in Experiment 1 with the ReLU activation function, these two outperforming classification algorithms, Thai-BERT and WangchanBERTa, were used to experiment with various tokenizers. The performance of each multiclass classification and word tokenizer was then calculated using the weighted-averaging method, as reported in Table 3. WangchanBERTa with the Newmm tokenizer demonstrated the best performance, achieving the highest F1-score of 74.2%, Recall of 74.2%, Precision of 74.7%, and Accuracy of 74.2%. When considering the results with other tokenizers, it was found that WangchanBERTa outperformed the Thai-BERT algorithm.

Table 3 Classification performance for each classification algorithm

Classification Techniques	F1-score	Recall	Precision	Accuracy
Thai-BERT (Newmm)	0.686	0.687	0.686	0.687
Thai-BERT (Attacut)	0.665	0.665	0.666	0.665
Thai-BERT (BPE)	0.663	0.665	0.665	0.665
WangchanBERTa (Newmm)	**0.742**	**0.742**	**0.747**	**0.742**
WangchanBERTa (Attacut)	0.722	0.723	0.727	0.723
WangchanBERTa (BPE)	0.732	0.732	0.733	0.732

Since the dataset in this study involves an imbalanced classification problem, two diagnostic tools were utilized: the confusion matrix and ROC curves. With ROC curves, the closer the curve is to the upper-left corner, the better the model's predictions. The area under the ROC curve (AUC) provides a quantitative measure of this performance. A straight-line ROC curve indicated poor prediction. Figures 5 and 6 depict the confusion matrices and ROC curves for each classification algorithm using the Newmm tokenizer. Additionally, ROC curves demonstrated that WangchanBERTa had a larger AUC (Area under the Curve) than Thai-BERT, indicating superior performance.

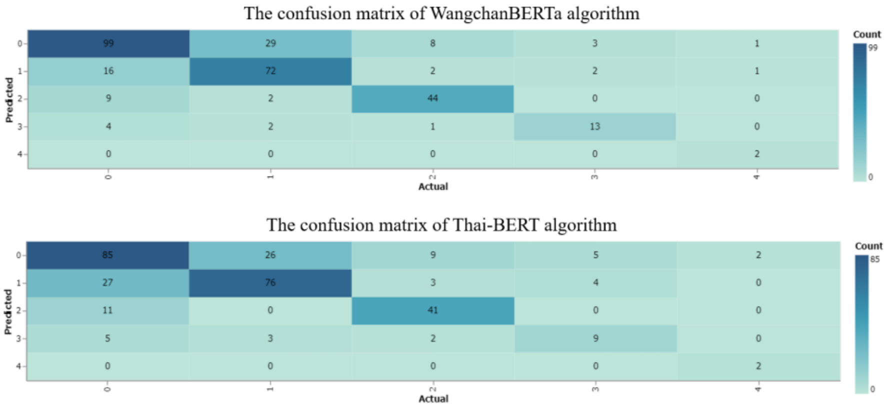

Figure 5 Confusion matrix of WangchanBERTa (up) and Thai-BERT (down) algorithms with Newmm tokenizer

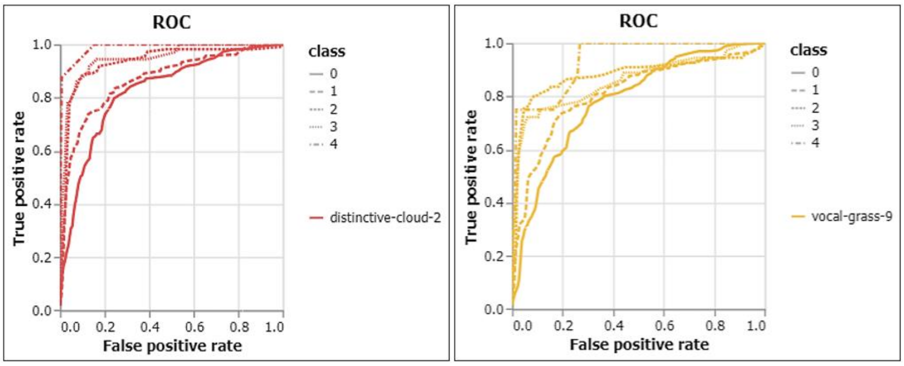

Figure 6 ROC curves of WangchanBERTa (red) and Thai-BERT (yellow) algorithms with Newmm tokenizer

The classes were numerically assigned from 0 to 4 as follows: 0 for Anhedonic, 1 for Other, 2 for Suicidal Ideation, 3 for Guilty Feelings, and 4 for Sleep Problems. Analysis of the confusion matrix revealed higher rates of misclassification for classes 0 (Anhedonic) and 1 (Other) compared to the remaining classes. Delving into the class-wise performance analysis, Tables 4 and 5 show the performance of algorithms by class was then calculated using the weighted-averaging method. WangchanBERTa still outperformed ThaiBERT for all signs of depression.

Table 4 Classification performance of WangchanBERTa algorithm by class

Sign of Depression	F1-score	Recall	Precision	Accuracy
Anhedonic	0.739	0.773	0.707	0.774
Other	0.727	0.686	0.774	0.826
Suicidal Ideation	0.800	0.800	0.800	0.929

Guilty Feelings	0.684	0.722	0.650	0.961
Sleep Problems	0.667	0.500	1.000	0.994

Table 5 Classification performance of ThaiBERT algorithm by class

Sign of Depression	F1-score	Recall	Precision	Accuracy
Anhedonic	0.667	0.664	0.669	0.726
Other	0.707	0.724	0.691	0.797
Suicidal Ideation	0.766	0.746	0.786	0.919
Guilty Feelings	0.487	0.500	0.474	0.939
Sleep Problems	0.667	0.500	1.000	0.994

4.3 Discussion

The experiments in the current paper built upon previous experiments by the authors, which used BiLSTM with the ReLU activation function as a baseline (S. Mahasiriakalayot et al., 2022). After constructing BiLSTM, Thai-BERT, and WangchanBERTa algorithms, it was found that the results did not outperform the previous experiment, even though the same dataset and configuration were used. Particularly, the LSTM model could not be reproduced as in the previous experiment. Therefore, a more in-depth examination of the dataset was performed. For future improvement, error analysis was executed to gain a better understanding of the strengths and weaknesses of both the models and the dataset. Below are some interesting topics that emerged from the analysis.

4.3.1 Size of the dataset

From the experimental results, the authors found that the dataset is both small and imbalanced. There are 5 classes, with some classes being 20 times smaller than others. The dataset revolves around signs of depression, which leads to certain sentences having similar meanings or common words that can cause misclassification. Frequent errors were observed in predictions between the 'Anhedonic' and 'Other' classes. For example (English Translation), the tweet 'No one understands us' (Anhedonic class) shares many common words with the tweet 'Even though no one in the world understands us, just us understanding ourselves is enough' (Other class), leading to misclassification.

4.3.2 Confusion over labeling and vocabularies

Analysis of the word cloud revealed that words related to depression appeared in every class. Additionally, each class possessed its own distinct vocabulary, which may not be prominently featured in the word cloud. For instance, in the 'suicidal ideation' class, vocabulary related to suicide or death may be present but with smaller font sizes, indicating lower frequency. When such vocabularies occur in the same text, it can lead to confusion for the models. For instance, consider the tweet text (English Translation) "suffering", labeled as the "Other" class by the labeler but belonging to the "Anhedonic" class as per the ground truth. This confusion results in the model struggling to predict the appropriate class for this text.

5. Conclusion

In this study, the authors developed and compared text classification algorithms using different neural network architectures (BiLSTM, Thai-BERT, and WangchanBERTa) and word tokenization techniques (Newmm, Attacut, and BPE) for detecting depression in Thai social media texts. The findings of this study demonstrate the superiority of the WangchanBERTa algorithm in combination with the Newmm tokenizer, which achieved the highest performance metrics (F1-score: 74.2%, Recall: 74.2%, Precision: 74.7%, and Accuracy: 74.2%) among the evaluated models. The analysis of confusion matrices and ROC curves further confirmed the effectiveness of WangchanBERTa, highlighting its ability to capture the nuances and complexities of the Thai language in the context of depression detection.

Despite the aforementioned, this study also revealed several challenges and limitations that warrant further investigation. The discrepancy between the current results and the previous experiment (S. Mahasiriakalayot et al., 2022), despite using the same dataset and configuration, emphasizes the need for rigorous testing and validation of proposed algorithms across different studies and datasets to ensure their reliability and generalizability. Moreover, the error analysis uncovered issues related to the size and imbalance of the dataset, as well as confusion over labeling and vocabulary, which can adversely affect the performance of depression detection algorithms.

To address these challenges, future research should focus on expanding the dataset to include a more diverse and balanced range of depression-related expressions, employing data augmentation techniques, and developing more sophisticated algorithms that can effectively handle imbalanced datasets as well as capture the subtle differences between similar expressions. Furthermore, establishing clearer and more consistent labeling guidelines, in collaboration with mental health professionals, can help reduce the ambiguity and confusion in the dataset and improve the accuracy and reliability of the depression detection algorithms.

The implications of the findings in this study extend beyond the realm of academic research and have the potential to inform real-world applications in mental health support and intervention. By demonstrating the feasibility and effectiveness of automated depression detection in Thai social media texts, this study lays the groundwork for the development of early warning systems and targeted support mechanisms. Mental health professionals and organizations can leverage these tools to identify individuals at risk of depression and provide timely interventions, ultimately improving mental health outcomes in Thailand.

In conclusion, the present study contributes to the advancement of mental health research and NLP in non-English languages, highlighting the importance of language-specific approaches and appropriate word tokenization techniques in depression detection tasks. While the findings of this study demonstrate the potential of the WangchanBERTa algorithm and the Newmm tokenizer, the challenges and limitations identified in this study underscore the need for continued research, interdisciplinary collaboration, and the adoption of rigorous testing and validation practices. By addressing these issues and further refining the proposed approaches, it is possible to work towards the development of accurate, reliable, and generalizable depression detection models for the Thai language, ultimately contributing to the improvement of mental health outcomes in Thailand and beyond.

6. Acknowledgements

The authors would like to thank Suwaroj Mahasiriakalayot, Twittie Senivongse, and Nattasuda Taephant for contributing to the dataset in this paper. Conference participation and presentation support for this research were provided by the Graduate School of Srinakharinwirot University, Thailand, for the fiscal year 2024.

7. References

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2011). SMOTE: Synthetic Minority Over-sampling Technique. arXiv:1106.1813. doi:10.48550/arXiv.1106.1813

Chormai, P., Prasertsom, P., & Rutherford, A. (2019). AttaCut: A Fast and Accurate Neural Thai Word Segmenter. arXiv:1911.07056. doi:10.48550/arXiv.1911.07056

Department of Mental Health, M. o. P. H. (2021). Report on Access to Services for Patients with Depression in Fiscal Year 2021. Retrieved from https://thaidepression.com/www/report/main_report/

Graves, A., & Schmidhuber, J. (2005). Framewise phoneme classification with bidirectional LSTM and other neural network architectures. Neural Netw, 18(5-6), 602-610. doi:10.1016/j.neunet.2005.06.042

Hämäläinen, M., Patpong, P., Alnajjar, K., Partanen, N., & Rueter, J. (2021). Detecting Depression in Thai Blog Posts: a Dataset and a Baseline. CoRR, abs/2111.04574. Retrieved from https://arxiv.org/abs/2111.04574

Husseini Orabi, A., Buddhitha, P., Husseini Orabi, M., & Inkpen, D. (2018). Deep Learning for Depression Detection of Twitter Users.
Komsan Kiatrungrit. (2016, 2016). Social Media Use and Suicide. Retrieved from https://www.rama.mahidol.ac.th/ramachannel/en/article/%e0%b8%9b%e0%b8%b1%e0%b8%8d%e0%b8%ab%e0%b8%b2%e0%b8%81%e0%b8%b2%e0%b8%a3%e0%b9%83%e0%b8%8a%e0%b9%89%e0%b8%aa%e0%b8%b7%e0%b9%88%e0%b8%ad%e0%b8%aa%e0%b8%b1%e0%b8%87%e0%b8%84%e0%b8%a1%e0%b8%ad%e0%b8%ad/
Laosen, N., Laosen, K., & Ardharn, J. (2023, 9-12 May 2023). Intent Classification from Online Forums for Phuket Medical Tourism. Paper presented at the 2023 20th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON).
Lowphansirikul, L., Polpanumas, C., Jantrakulchai, N., & Nutanong, S. (2021). WangchanBERTa: Pretraining transformer-based Thai Language Models. CoRR, abs/2101.09635. Retrieved from https://arxiv.org/abs/2101.09635
Oh Jihoon, Yun Kyongsik, Maoz Uri, Kim Tae-Suk, & Jeong-Ho, C. (2019). Identifying depression in the National Health and Nutrition Examination Survey data using a deep learning algorithm. Journal of Affective Disorders, 257, 623-631. doi:https://doi.org/10.1016/j.jad.2019.06.034
Phatthiyaphaibun, W., Chaovavanich, K., Polpanumas, C., Suriyawongkul, A., Lowphansirikul, L., Chormai, P., . . . Udomcharoenchaikit, C. (2023, December). PyThaiNLP: Thai Natural Language Processing in Python, Singapore.
Q. Cong, Z. Feng, F. Li, Y. Xiang, G. Rao, & Tao, C. (2018, 3-6 Dec. 2018). X-A-BiLSTM: a Deep Learning Approach for Depression Detection in Imbalanced Data. Paper presented at the 2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM).
S. Mahasiriakalayot, T. Senivongse, & Taephant, N. (2022, 22-25 June 2022). Predicting Signs of Depression from Twitter Messages. Paper presented at the 2022 19th International Joint Conference on Computer Science and Software Engineering (JCSSE).
Sennrich, R., Haddow, B., & Birch, A. (2015). Neural Machine Translation of Rare Words with Subword Units. arXiv:1508.07909. doi:10.48550/arXiv.1508.07909
The Electronic Transactions Development Agency, M. o. D. E., and Society (2022). Thailand Internet User Behavior 2022. Retrieved from https://www.etda.or.th/th/Useful-Resource/publications/iub2022.aspx
Wolf, T., Debut, L., Sanh, V., Chaumond, J., Delangue, C., Moi, A., . . . Rush, A. M. (2019). HuggingFace's Transformers: State-of-the-art Natural Language Processing. arXiv:1910.03771. doi:10.48550/arXiv.1910.03771
Wongaptikaseree, K., Yomaboot, P., Katchapakirin, K., & Kaewpitakkun, Y. (2020). Social Behavior Analysis and Thai Mental Health Questionnaire (TMHQ) Optimization for Depression Detection System. IEICE Transactions on Information and Systems, 103, 771-778. doi:10.1587/transinf.2019IIP0003
World Health Organization. (2021). Depressive disorder (depression). Retrieved from https://www.who.int/news-room/fact-sheets/detail/depression#:~:text=An%20estimated%203.8%25%20of%20the,among%20women%20than%20among%20men.
Zogan, H., Wang, X., Jameel, S., & Xu, G. (2020). Depression Detection with Multi-Modalities Using a Hybrid Deep Learning Model on Social Media. CoRR, abs/2007.02847. Retrieved from https://arxiv.org/abs/2007.02847