Social Behavior Analysis and Thai Mental Health Questionnaire (TMHQ) Optimization for Depression Detection System

Konlakorn WONGAPTIKASEREE††††, Member, Panida YOMABOOT†††, Kantinee KATCHAPAKIRIN†††, and Yongyos KAEPITAKKUN†††††, Nonmembers

SUMMARY  Depression is a major mental health problem in Thailand. The depression rates have been rapidly increasing. Over 1.17 million Thai people suffer from this mental illness. It is important that a reliable depression screening tool is made available so that depression could be early detected. Given Facebook is the most popular social network platform in Thailand, it could be a large-scale resource to develop a depression detection tool. This research employs techniques to develop a depression detection algorithm for the Thai language on Facebook where people use it as a tool for sharing opinions, feelings, and life events. To establish the reliable result, Thai Mental Health Questionnaire (TMHQ), a standardized psychological inventory that measures major mental health problems including depression. Depression scale of the TMHQ comprises of 20 items, is used as the baseline for concluding the result. Furthermore, this study also aims to do factor analysis and reduce the number of depression items. Data was collected from over 600 Facebook users. Descriptive statistics, Exploratory Factor Analysis, and Internal consistency were conducted. Results provide the optimized version of the TMHQ-depression that contain 9 items. The 9 items are categorized into four factors which are suicidal ideation, sleep problems, anhedonic, and guilty feelings. Internal consistency analysis shows that this short version of the TMHQ-depression has good to excellent reliability (Cronbach’s alpha >.80). The findings suggest that this optimized TMHQ-depression questionnaire holds a good psychometric property and can be used for depression detection.

key words: depression detection, psychological inventory, social behavior analysis, TMHQ

1. Introduction

Depression is considered one of the major mental health problems. According to the World Health Organization, 4.4% of the world population or around 300 million people are affected by this mental disorder[1]. This statistic seems to be risen up especially in the lower socioeconomic country or developing country including Thailand. In Thailand, the Excellence Center for Depressive Disorder reported that among 53 million Thais who are aged over 15 years old, approximately 1.44 million of those are suffering from depression[2]. Depression is a result of an imbalance neurotransmitter as well as psychosocial stressor. This emotional illness can be defined as having an alteration in mood; being sad, lonely and apathy, for instance. It also includes having negative attitude toward oneself, changes or lost in vegetative signs such as loss of appetite, sleep or libido. Changes in activity level are also common in depression. When someone is depressed, it is likely that such person would lost interest in doing or participating in activity that once enjoyed. In addition, this emotional illness leads to a desire to escape, hide or disappear[3].

Depression not only affects an individual who suffers from this emotional illness but also is a burden to family and society socially and economically. It is estimated that depression and other psychological problems would cost the global economy up to $16 trillion between 2010 and 2030[4],[5]. Most importantly, depression is a major cause of suicide[1],[5]. Depression is preventable and treatable. It is important that depression is early detected because it would prevent this emotional illness from getting worse or last a long time. Early detection would also minimize disability and dysfunctional. It was found that depression detection program helped reduce depression prevalence and increase treatment response[6]. Early detection in depression is beneficial to various population such as teens, adult, elderly, pregnant and postpartum populations, and patients with other medical conditions such as stroke and coronary heart disease (CHD)[6],[7].

There are several depression screening tools available for mental and health care providers. Most of the tools are in form of self-rating questionnaire as it is convenience to use. Examples of well-known and widely used depression screening questionnaires are the Patient Health Questionnaire (PHQ)[8], the Hospital Anxiety and Depression Scales, the Beck Depression Inventory, the Geriatric Depression Scale (GDS), and the Edinburgh Postnatal Depression Scale (EPDS)[9]. It can be seen that the aforementioned screening tools are originally developed in English. However, in psychology, cultural and language background plays an important role in psychological assessment result. Thus, psychological assessment tool including depression screening that was developed in native language and designed for specific population might provide more accurate result.
Although there are several questionnaires available, evaluating depression using this method has its own limitation. Responses ascertain by this method might not represent the real symptoms and severity as it might affected from biases such as self-report bias which refers to a misunderstanding about what is being measured and social desirability bias [10] where the self-rater wants to ‘look good’ to others [11]. To detect depression passively from daily living activity could be a solution to somewhat resolve limitations from the discussed methods. It was hypothesized that the way people behave online might be a good source for self-reflection. During the past decade, literature in psychology and online behaviors has been rapidly increased. This includes the emotional aspect.

Munmund et al. [12] examined whether social media would be able to detect and diagnose the major depressive disorder or not. Data such as social engagement, emotion, language styles, ego network, and mentions of antidepressant medications were collected from Twitter users over one year. They found some significant online behaviors that could predict the onset of depression. They reported that decrease in social activity, increase negative emotion, highly clustered ego networks, heightened relational and medical concerns, and greater expression of religious involvement were predictors of the onset of depression. They also reported that the model developed from this study could predict depression before its onset with 70% accuracy and 0.74 precision. They suggested that this method could be a new approach to identify people who are at risk of depression.

Xinyu Wang et al. [13] detected social network users’ depression by applying data mining to psychology knowledge. They employed a sentiment analysis method. Vocabulary and man-made rules were utilized for calculating tendency. Their depression detection model included 10 selected features based on related literature. The 10 features could be categorized into three dimensions which are microblog, interactions and behaviors. Validated by Bayes, Trees and Rules, ROC Area and F-measure, they reported that the proposed model has 80% accuracy in detecting depression. They also reported that in cases of an incorrect prediction, it was caused by the insufficient data in microblogs.

Andrew G. Reece and Christopher M. Danforth [14] employed machine learning for depression screening on Instagram. Color analysis, metadata components, and algorithmic face detection were performed in order to develop a model. They found that their proposed model had 70% accuracy in predicting depression even before being diagnosed. They highlighted that this could be a new method for depression screening and detection.

In Thailand, the Thai Mental Health Questionnaire (TMHQ) is one of the widely-used mental health inventories [15]. This DSM-IV based mental health evaluation tool is a 70-question self-rating scale. It measures five aspects of mental and psychopathology including depression, anxiety, psychosomatic, psychotic and social function. Each set of question can be used separately [15]. TMHQ was originally developed in Thai. Its psychometric properties were yielded from Thai sample. This suggests that this mental health inventory is suitable for Thai population.

TMHQ has appropriate psychometric property. The reliability coefficients ranged from 0.82 to 0.94. Internal consistency of the depression scale is .82. Although hold appropriate psychometric property, the length of the questionnaire could be disadvantageous as it is time-consuming. Moreover, one of the symptoms of depression is lack of motivation and having low energy level. Long questionnaire might affect motivation in answering questions.

To address this shortcoming, this research aims to analyze the social behavior of the user on Facebook, which is the most popular social network platform in Thailand. Social behavior factors, including a number of posts, interaction with others, privacy settings, and day and time of posting are extracted for detecting depressive signs. TMHQ is used as the baseline of this research to assess the depression level. Thus, shorter version of TMHQ depression scale is needed. For this reason, research also aims to conduct depression detection system (DDS) for recognizing the depression by optimizing the TMHQ using the statistical technique. Factor analysis is a standard data reduction method widely accepted and used in psychological and social sciences research. This statistical technique extracts maximum common variance from all variables and puts them into a common factor. This statistical technique is used with continuous data such as score from psychological measurement and allows researcher to investigate latent variables which are the variables that cannot be directly observed [16].

Given the data in this research is yield from psychological assessment tools, the data type is continuous. Thus, Factor Analysis is an appropriate method to reduce the number of factors for further analysis. Over 600 data were collected from participants who are depressed individual and non-depressed individual. Exploratory factor analysis technique was conducted in order to find the significant value of each question, categorized and shorten the standardized questionnaire. Consequently, 20 questions in the TMHQ was optimized to 9 questions. Furthermore, we also investigated internal consistency to confirm that the analyzed results have acceptable psychometric property that the optimized version of TMHQ-Depression can be used as the screening tool for the depression detection system.

2. Methodology

To achieve the goals of this research, we divided research work into two parts: Social behavior analysis (SBA), and TMHQ optimization.

2.1 Social Behavior Analysis (SBA)

Facebook users who were aged over 18 years old and willing to share their microblogs for depression research were recruited into the study through the internet. Participation was voluntary and the volunteers did not receive any
incentive for taking part in this research. The process to do the depression classifier model in social network is described in Fig. 1. The volunteers were asked to complete 20 items self-report depression screening derived from the TMHQ, only depression domain. The detail of this measure can be found in [15]. Total scores were classified into two groups; depress and non-depressed. This classification is used as depression criteria in this research.

After the completing the TMHQ, all microblogs which were created during a period of one month prior to the date completing the questionnaire were collected. The microblogs are represented in the Graph API JSON format [17]. Attributes Extractor Module parsed the microblogs to list of attributes, shown in Fig. 2.

With the aim of this paper, we aim to find the suitable technique to classify the depression from social network data. We investigated three types of algorithms, which can be categorized into three groups: linear, rule-based, and neural network techniques. Support Vector Machine (SVM), Random Forest, and deep learning technique are proposed in the depression classifier model. These three techniques are totally difference in term of utilization, and all of techniques have different strengths and weaknesses. The extracted attributes were fed as the input of the depression classifier model. The prediction label indicated the possible presence of depression of the user.

2.2 TMHQ Optimization

Depression detection system is a screening tool that used for recognizing depression, shown in Fig. 3. TMHQ-depression, a standardized mental health inventory, is used as the baseline of our purposed system because it is widely used for depression evaluation at the first step of depression’s treatment in a psychiatric setting in Thailand. 20 of 70 questions in TMHQ were focused to assess depression level, shown in Table 1.

TMHQ-depression data management is shown in Fig. 3. Data was collected from social media users. Participation was voluntary. Descriptive statistics were yielded to gain insight into the sample’s depression level. Factor analysis was conducted in order to categorized the TMHQ-depression items and reduce the number of the questionnaire. The full version of TMHQ can be found in [15].

In statistics in psychology, Factor Analysis is a statistical method that aims to explore and uncover the structure of the given data. It is commonly used for shortening psychological measures. This data reduction technique helps identify latent constructs and factor structure of the data. Principal Component Analysis (PCA) is a more specific data reduction technique that uses an orthogonal transformation to convert a set of data, depressive questions in this case, into a
The general distance and relatedness between principal components are displayed in form of Eigenvalue. Eigenvalues are the variances of the factors. The values are used to decide the numbers of factors that should be retained. Components with eigenvalue less than 1 will be excluded because they pay a little contribution to the explanation of variances in the variables and can be ignored as redundant with more important factors.

In factor analysis, factor rotation is a linear transformation process that makes the result more understandable and interpretable. This includes data rotating, flipping, and compressing. In this particular data set, Direct Oblimin orthogonal rotation method was employed. This rotation technique will be used when the factors are correlated. For example, psychological attributes. In this paper, it was hypothesized that the factors discovered from the analysis would be correlated as they are symptoms of depression. Thus, Direct Oblimin orthogonal rotation method was applied [19].

Internal consistency or reliability analysis is a statistic process that examines the reliability of a set of psychological questionnaire. It indicates the degree of relatedness of the item with the factor. It shows that the items in the same factor or construct measure the same thing. In other words, it tells how reliable the questionnaire is. Cronbach’s alpha (\( \alpha \)) is widely used to measure internal consistency. Rule of Thumb for Cronbach’s alpha is \( \alpha > .7 \) acceptable, \( \alpha > .8 \) good and \( \alpha > .9 \) excellent. The results from this study indicate that the optimized TMHQ-depression has good to excellent internal consistency and can be used as the screening tool for detecting the depression level.

In addition, mean and standard deviation for the total raw score of the optimized TMHQ were also ascertained in order to categorize level of depression. Consequently, this process provide 4 levels of depression ranging from non-depressed, mild-depressed, moderate depressed, and severely depressed.

3. Experiment and Results

3.1 Social Behavior Analysis with AI

This study has been granted an ethical approval from the Center of Ethical Reinforcement for Human Research, Mahidol University with the COA No. MU-CIRB 2017/143.1809. Thirty-five Facebook users who aged over 18 years old and willing to donate their microblogs for this research were classified by TMHQ into 2 groups which were 22 depressed and 13 non-depressed. These TMHQ profiles were used as a gold label for this study. 1,105 posts were collected and attributes were extracted.

All attributes are listed in Table 2 [19]. Some attributes were transformed into numerical values by the values computing module such as N_allPost which refers to a total number of posts collected one month prior to the date the participant answered the TMHQ depression domain. Some language-related attributes, however, required more complicated processing such as N_Neg which refers to a total number of Negative posts. In order to process the value of the attribute value under the lack of NLP resource in the Thai language, we translated the post into English first. Then, system segmented the words, tagged part of speech and analyzed sentiments in order to determine the values of language-related attributes. We used Google Cloud Translation API [20] as a language translator and NLTK python library [21] as the NLP resources to manipulate the trans-
Three algorithms, SVM, Random Forest, and Deep Learning, are used to detect the depression based on the attributes in Table 2. Nonetheless the size of collected data was too small, we could not separate training and test sets as conventional validation without losing significant modelling. Eight-cross-validation was applied to estimate model prediction performance. The evaluation results are presented in Table 3.

Apart from the accuracy, we also found important factors that might be one of the depressive sign. For example, people who post a lot of micro-blog on Monday with neutral sentiment possibly suffer from depression. Moreover, people who post negative sentimental microblog without using emoticon and set privacy to be ’only me’ tended to suffer from depression. In contrast, for those who frequently forwarded others’ posts, shared their own memory, actively posted between 6 AM and 12 AM and usually tagged friends were more likely not suffer from depression. Figure 4 illustrated the important factors recognized by proposed methods.

### 3.2 TMHQ Optimization

According to the Fig. 3, 636 Facebook users who were Thai were recruited to take part in this study. Their participants were voluntary. Participants did not receive any incentives for taking part in the research project. Participants were asked to complete 20 items of the TMHQ-depression online. Their identities were strictly kept anonymous. Using the TMHQ original scoring and criteria of depression, more than half of the participants reported depression symptoms in moderate or high level. Table 4 illustrates the descriptive data of participants in this experiment.

From the collected data, only data from the participants who were depressed were used in this data analysis. Thus, the data from 454 participants were analyzed. Exploratory factor analysis (EFA), a standard statistical technique, is used to explore a number of factors that should be retained from all of the 20 TMHQ-depression items. Then, principal component analysis (PCA) is analyzed to reduce attribute space from a larger number of variables to a smaller number of factors.

Table 5 shows the results of EFA. The components, which have the eigenvalue above 1, were selected. Based on the results in Table 4, top 4 components, which are “Suicidal ideation”, “Anhedonic”, “Sleep problems”, and “Guilty feeling” were selected as the essential component for classifying the depression level. Then, with the answer of 20 questions, the result of each question was computed the factor loading, presented in Table 6.

From the analyzed results of factor loading, top 2 of each component are selected because it had a moderate to strong loading on their primary component. Nonetheless, question No. 1 (“I feel sad”) was retained in the final version of the optimized questionnaire. The question No. 1 has a direct meaning of depression. It is one of the questions that can confirm the depression level. Consequently, 9 questions were retained in optimized version of TMHQ, shown in Table 7.

Furthermore, in order to examine the psychometric property of the optimized version of the TMHQ, internal consistency was analyzed. Data from the whole sample (659 participants) was used in this process. Results show satisfactory internal consistency in all four aspects which can be
seen in Table 8. With the high consistency value, it is confirmed that DDS can be used to recognize the depression level. To complete the DDS, depression criteria is adopted as a rule to determine the depression level. Mean and SD values from the 659 collected data were used to calculate the criterion of depression level, displayed in Table 9.

4. Conclusion

In this research, depression detection system was introduced for recognizing the depression level. We proposed two methods to detect depression. Social behavior analysis was proposed to analyze factors in social network. The experiments results show that the use of behavioural information on Facebook, both in forms of messages and activities, could predict depression. However, the sample of this research is relatively small because Facebook has limited their permission to collect personal information and the process of gaining approval has become more complicated. Thus, the results getting from this study might not cover all relevant factors. Moreover, as the language-related features had to be translated from Thai to English for analyzing the process, there might be some errors due to this process because some important sentiment polar words might have been eliminated during the translation process.

TMHQ optimization was proposed as the second choice for those who do not want to share personal data. The main task of this method was to analyze the key components of the standard screening tool, which is TMHQ questionnaire. 659 participants participated in the experiment. 4 significant components were ascertained by a standard psychological data analysis process; Factor Analysis, EFA technique and Eigenvalue. Meanwhile, PCA technique investigated the factor loading of each question in order to decide 2 significant questions in each component. To confirm the reliability of optimized TMHQ, internal consistency was inspected. Results indicated the high consistency value of each component. Based on the results, we can summarize that the optimized TMHQ can be used as the screening tool in DDS for detecting the depression level.

Generally, in the machine learning field, Minimum redundancy-maximum-relevance (mRMR) is frequently used for data reduction. The method in machine learning requires a large value of data whereas method in psychology require a much smaller sample size and use statistic significant and power to determine the quality of data and results. This process was done in a short period of time given small sample size required. Off note, although this paper might not be a novel experiment in term of method used but it is the first experiment in Thai culture and language.

This research sought to gain insight and develop a tool
to understand psychological phenomena; depression by developing the so-called depression detection system (DDS). It was important that the experiment was done in a way that psychological procedure was remained and psychological standard was met. In the researchers’ point of view the most important point of this experiment is that it brought together experts from two different disciplines, psychology and intelligent information to achieve the experiment’s goal while psychological protocol has strictly remained. By this mean, the results are validated and are beneficial to psychologist in their clinical work, depression in particular.

For the improvement of this research, the results of this research can be used to design a new conversation flow for recognizing the depression level because some wording of current version of the questionnaire are not natural language. In some cases, participants attempt to understand the meaning of the question. Furthermore, we intend to collect more data to get more relevant and valid features. Manual annotating all complex attributes using crowdsourcing and deeper dimensions should also be analyzed in order to be able to create a better depression detection algorithm.

Acknowledgments

This paper is a part of the larger research which would have been impossible without the support of the Council of University Presidents of Thailand Funding in the project of Healthcare Innovation Hub: for Community and Aging Society. We would also like to express our thanks to the experts from two different disciplines, psychology and in-
telligent information. We would also like to express our thanks to the ethic-al committees of Mahidol University for granting an ethical approval (MU-CIRB 2017/142.1108) as well as the partici-pants who took part in this research. We would also like to sincerely thank our participants for their time taking part in this research.

References

[1] WHO, “WHO | Depression,” 2017.
[2] Thailand Excellence Center for Depressive Disorder, “Accessibility to Mental Health Care in Depression Report,” 2018.
[3] A.T. Beck and B.A. Alford, Depression: Causes and Treatment, University of Pennsylvania Press, 2014.
[4] K. Kelland, "Mental health crisis could cost the world $16 trillion by 2030" [Reuters]. [Online]. Available: https://www.reuters.com/article/us-health-mental-global/mental-health-crisis-could-cost-the-world-16-trillion-by-2030-idUSKCN1MJ2QN. [Accessed: 03-Mar-2019].
[5] M. Marcus, T. Yasamy, M. van Ommeren, D. Chisholm, and S. Saxena, "Depression: A Global Public Health Concern," WHO Department of Mental Health and Substance Abuse, pp.6–8, 2012.
[6] J. Garland, “Early detection of depression in young and elderly people | British Columbia Medical Journal,” BCMJ, vol.44, no.9, 2002. [Online]. Available: https://www.bcmj.org/articles/early-detection-depression-young-and-elderly-people. [Accessed: 03-Mar-2019].
[7] J.W. Feightner and G. Worrall, “Early detection of depression by primary care physicians,” CMAJ, vol.142, no.11, pp.1215–20, 1990.
[8] K. Kroenke, R.L. Spitzer, and J.B.W. Williams, “The PHQ-9: Validity of a brief depression severity measure.” [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/pmid/11556941/.
[9] M.P. Pignone, B.N. Gaynes, J.L. Rushston, C.M. Burchell, C.T. Orleans, C.D. Mulrow, and K.N. Lohr, “Screening for depression in adults: A summary of the evidence for the U.S. Preventive Services Task Force,” Annals of Internal Medicine, 2002.
[10] H.J. Arnold and D.C. Feldman, “Social Desirability Response Bias in Self-Report Choice Situations,” Acad. Manag. J., vol.24, no.2, pp.377–385, 1981.
[11] R. Rosenman, V. Tennekoon, and L.G. Hill, “Measuring bias in self-reported data,” Int. J. Behav. Health Res., vol.2, no.4, p.320, 2011.
[12] M. De Choudhury and M. Gamon, “Predicting depression via social media,” Proc. Seventh Int. AAAI Conf. Weblogs Soc. Media, vol.2, pp.128–137, 2013.
[13] X. Wang, C. Zhang, Y. Ji, L. Sun, L. Wu, and Z. Bao, “A Depression Detection Model Based on Sentiment Analysis in Micro-blog Social Network,” PAKDD Work., pp.201–213, 2013.
[14] A.G. Reech and C.M. Danforth, “Instagram photos reveal predictive markers of depression,” EPJ Data Sci., vol.6, no.1, pp.1–34, 2017.
[15] S. Phatharayuttawat, T. Ngamthipwattana, and K. Sukhatungka, “The norm profile for the Thai mental health questionnaire.”
[16] B.G. Tabachnick and L.S. Fidell, Using Multivariate Statistics, vol.28, no.8, 2013.
[17] Facebook, “The Graph API Reference.”
[18] A. Aron, E.J. Coups, and E.N. Aron, Statistics for Psychology, Pearson, Boston., pp.1–819, 2006.
[19] K. Katchapakirin, W. Wongpatikaseree, P. Yomaboot, and Y. Kaewpitakkun, “Facebook Social Media for Depression Detection in the Thai community,” In 15th International Joint Conference on Computer Science and Software Engineering (JCSE8 2018), Thailand, 2018.
[20] Google Cloud, “Google Cloud Translation API Documentation.” [Online]. Available: https://cloud.google.com/translate/docs/.
[21] S. Bird, E. Klein, and E. Loper, Natural Language Processing with Python, O’Reilly Media, 2009.

Konlakorn Wongpatikaseree received his B.Sc. from Thammasart University in 2008. During 2008–2010, he received a scholarship to pursue a Master degree from the Thailand Advanced Institute of Science and Technology and Tokyo Institute of Technology. He received his Ph.D. in information science from Japan Advance Institute of Science and Technology. Now he is a lecturer at Department of Computer Engineering, Faculty of Engineering, Mahidol University, Thailand. His research interests are in intelligent systems, embedded systems, smart homes, activity recognition, and healthcare systems. He is a member of IEICE.

Pania Yomaboot is a licensed clinical psychologist at the Department of Psychiatry, Faculty of Medicine Siriraj Hospital, Mahidol University. Her research interests are in clinical psychology, depression, personality, psychometric property, music psychological and the application of AI in psychology. She has a strong background in psychology as she holds B.Sc. in Psychology with a second class honour, M.Sc. in Clinical Psychology and she is working toward the final step of her PhD in psychology at Goldsmiths, University of London.
Kantinee Katchapakirin received her B.Eng. from Ubon Ratchathani University in 1999 and M.Eng. from Asian Institute of Technology, Thailand in 2003. Now she is a first-year Ph.D. student in the Computer Science and Information Technology Department of the Asian Institute of Technology. Her current research project focuses on Computational Thinking Development for Kids. Her research interests are AI in Education, Semantic Web Technologies, Ontology, Knowledge Management, Natural Language Processing.

Yongyos Kaewpitakkun received Ph.D. degree from Japan Advanced Institute of Science and Technology, Japan in 2016. His research interests lie in the field of natural language processing and artificial intelligent.