

Behavior Features for Automatic Detection of Depression from Facebook Users

Siranuch HEMTANON^{a,1}, Saifon AEKWARANGKOON^b
and Nichnan KITTPHATTANABAWON^a

^a*School of Informatics, Walailak University, Nakorn Si Thammarat, 80160 Thailand.*

^b*School of Nursing, Walailak University, Nakorn Si Thammarat, 80160 Thailand.*

Abstract. Major depressive disorder is one of common mental disorders globally. It is best to be early detected and cured. This work introduces a method to detect depressive disorder at risk via a behavior made on Facebook platform. A set of features related to Facebook main functions including amount of posting, sharing, commenting and replying is designed to represent users' activities in a numerical value form. The collected data with periodic and consecutive aspects are gathered without interpreting content. Thus, the data are easier to be collected with less privacy issue. To distinct between positive and negative depression-at risk, PHQ-9 questionnaire, a standard tool commonly used to screen depression patient in Thailand, was used. These features hence are used in supervised learning classification algorithm for detecting a risk of being depressive disorder. From the experiment of 160 Thai Facebook users, the statistical result indicated that depression-positive users tend to do consecutive actions and rarely reply to other comments. Moreover, they often have activities in late night. The classification experiment shows that the designed features based on users' activities from Facebook with deep learning algorithm yields about 87% in terms of F-measure. After analyzing the data, we thus split data regarding users' gender and removed obviously low active data, and the F-measure from classification went up to 91.4 which improves for 4 points.

Keywords. Depression, Detection, Behavior features, Social network activity, Facebook

1. Introduction

Depression (also called clinical depression or major depressive disorder) is a mood disorder causing a persistent negative thought, physical disturbance, behaviors change, and feeling of sadness and loss of interest [1]. It is one of the globally most common mental disorders [2, 3]. Depression can happen at any age. It causes severe symptoms affecting how ones feel, think, and handle daily activities including sleeping, eating, or working. To be diagnosed with clinical depression, the symptoms must be present for at least two weeks [4]. The study [5] shows that the earlier that treatment begins, the more effective it is. Thus, it is best for patients to be early diagnosed. Although it is impossible to physically examine all people at risk, surveillance screening test should be done.

¹ Corresponding Author: Nichnan KITTPHATTANABAWON, School of Informatics, Walailak University, Nakorn Si Thammarat, Thailand. 80160; E-mail: knichcha@wu.ac.th.

Since depression symptoms can alter patient's thinking process, their behaviors are changed in several aspects such as having less activities from losing interest to do thing and being uneasy easily from feeling hopeless. Though it is difficult to monitor every physical action for depression screening for every person, it is possible to find a sign in behavior on digital based media. The digital-based media refers to interact to ones' mobile phone, actions in an electronic game and behavior in social network applications. The signs of the disorder shown in social media can be easier to be recorded and analyzed.

In this work, we aim to analyze users' behavior via social network application to realize what the signs of depressive disorder. Since Facebook is a famous and widely used social network in Thailand as there are 26.9 million Thai users in 2019, we focus our feature design based on Facebook functions including posting, sharing, commenting and replying. The designed features are coped to be measurable as a numeric value without interpreting content; thus, it is easier to be gather, possibly automatically, with less privacy. We expect that the designed features regarding users' behavior on Facebook will help in detecting non-aware cases of depressive disorder and lead them for earlier treatments as well as prevention of a possible suicidal attempt.

2. Related Works

There are several researches on screening mental disorder using information technology. A summary of these works is provided in Table 1. From the summary, we found that most of the work handled a text related to mental disorder and suicidal risk. There are two main techniques including natural language processing to handle text and machine learning to create a decision-making model for determining the result. Their focused texts are from online sources such as twitter posts and news. Most of the works applies text-based classification techniques, such as Naïve-Bayes (NB) [6, 7, 8, 9], Decision-tree (DT) [7, 8, 9, 10], Support Vector Machine (SVM) [6, 8], Genetic algorithm (GA) [11] and Deep learning [12], to detect a post that is risky for expressing suicide attempt or shows a sign of mental disorder symptom. These works provide great help to reduce the risk of mental disorder; however, they require patients to make a post about what they have in mind. In a view of natural language processing (NLP), complexity of a target language also plays a crucial role in accuracy performance. Thus, the same technique applied to different language may result in noticeably different performance, and some languages including Thai still lack necessary resources and theoretical studies to improve its NLP performance.

In addition, the mental disorder may affect the patients to be unstable and result in rarely making a post, especially depressive disorder which makes one losing interest in doing anything [5]. Thus, solely focusing on detecting the mental risks from the text may be insufficient to detect all the at-risk people. Detecting people with depression from online sources may require additional information such as how they interact with other people in their social network.

Table 1. Related work on detection of mental disorder and suicide risk

Research Title	Techniques		Dataset
	NLP	ML	
An Automatic Screening for Major Depressive Disorder from Social Media in Thailand (2019) [6]	Word vector	NB, SVM	Facebook posts

Research Title	Techniques		Dataset
	NLP	ML	
Detection of Suicide-Related Posts in Twitter Data Streams (2018) [7]	Lexicon-based	NB, DT	Twitter posts
Google and suicides: what can we learn about the use of internet to prevent suicides? (2018) [8]	-	Regression analysis	Suicide data mortality from all American state
Detecting suicidality on Twitter (2015) [10]	-	SVM, Logistic Regression	Twitter posts
Analyzing the connectivity and communication of suicidal users on twitter (2016) [12]	-	TF-IDF, Unigram	Twitter posts
Natural Language Processing of Social Media as Screening for Suicide Risk (2018) [13]	Word vector	Deep Learning	Social media data
Mining Twitter for Suicide Prevention (2014) [14]	grammatical structures	NB, DT, K-NN	Twitter posts
Multi-class machine classification of suicide-related Communication on Twitter (2017) [15]	-	NB, DT	Twitter posts
The association of suicide-related Twitter use with suicidal behavior: A cross-sectional study of young internet users in Japan (2015) [16]	-	Association	Twitter posts

3. Methodology

This work focuses on Depression at risk screening via users' behavior in Face-book. We first gather users' behavior on Facebook for both positive and negative in depression into two groups. Statistics of their behavior are analyzed and calculated following our designed features. Then, we use the features to learn a classification model to detect depression-positive users. An overview of the process is illustrated in Figure 1.

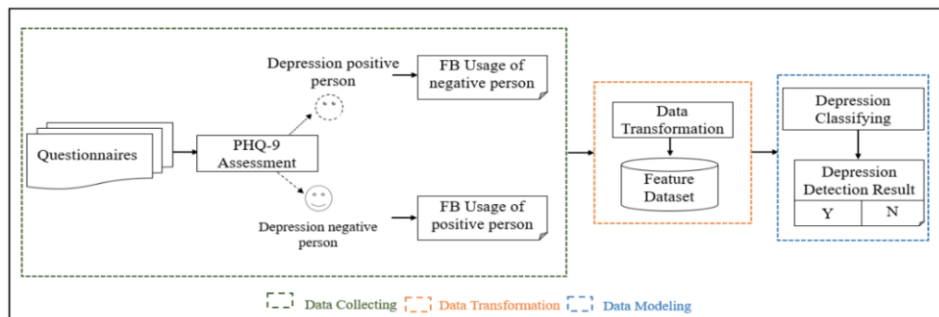


Figure 1. An overview of the proposed method.

3.1. Data Collection

To gather volunteers and differentiation between positive and negative in risk of depression, we design a questionnaire for asking for anonymous personal information and a test for depression. The chosen depression test in this work is PHQ-9 questionnaire [4] which is widely used and accepted for detecting signs of depression in Thailand. The details of questionnaire and other required information are provided in Table 2.

Table 2. The acquired information from participants

Question type	Question	Note
Personal Information	<ul style="list-style-type: none"> - Gender (Male, Female, Sexual Oriental, Alternative gender) - Age (year-olds) - Occupation 	
Depression testing	<ul style="list-style-type: none"> - PHQ-9 (0-3 scale answer where 0 means not at all, and 3 refers to often): <ol style="list-style-type: none"> 1. Little interest or pleasure in doing things 2. Feeling down, depressed, or hopeless 3. Trouble falling or staying asleep, or sleeping too much 4. Feeling tired or having little energy 5. Poor appetite or overeating 6. Feeling bad about yourself or that you are a failure or have let yourself or your family down 7. Trouble concentrating on things, such as reading the newspaper or watching television 8. Moving or speaking so slowly that other people could have noticed - Or the opposite being so fidgety or restless that you have been moving around a lot more than usual 9. Thoughts that you would be better off dead, or of hurting yourself 	These questions [5] are to capture symptoms of depression that affect ones' mind emotional, cognition, physical, and behaviors.
Facebook Account	<ul style="list-style-type: none"> - Consensus for monitoring their account for their public activities including post, share and reply 	Monitoring as seeing their actions as their friend

With PHQ-9 questionnaire, depression result is defined into five classes as minimal depression, mild depression, moderate depression, moderately severe depression and severe depression. In this work, volunteers are split into two groups for positive depression (from mild depression, moderate depression and severe depression) and negative depression (no depression) based on their depression result. We ask them to allow us to monitor their Facebook accounts and collect data according to our designed features. The data collection is conducted for at least 14 days according to depression guideline [5].

3.2. Data Transformation

Since each social network platform provides different functions, collectible features regarding users' behavior should be specified differently. In this work, we focus on Facebook's users; thus, the designed features are according to Facebook functions specifically. The features are about what and how a user uses and interacts to other accounts, especially posting, commenting, sharing and replying. The amount of these actions is counted and stored. Moreover, the timing of these actions is also collected to

calculate for a gap between each action. We though do not collect the content of these actions in this phase since they are private information. There are three aspects as Facebook functions, time and consecutive action. The details of the collected features are given in Table 3.

Table 3. The Features based on Facebook usage behavior for detecting depression at risk

Aspect	Features	Details
Facebook functions	Posting (posts made by an account owner)	- Average number of daily posts - Standard deviation of daily posts
	Sharing (posts made by sharing other account's post)	- Average number of daily shares - Standard deviation number of daily shares
	Commenting (comment as replying in another person's post and share)	- Average number of daily comments - Standard deviation number of daily comments
	Replying (comment as replying in another person's comments in own post and share)	- Average percentage of daily replies - Standard deviation number of daily replies
Time	Weekday - Weekend	- Number of actions done in weekday (Mon-Fri) - Number of actions done in weekend (Sat-Sun)
	Period	- Number of actions done in morning (06:01-12:00) - Number of actions done in afternoon (12:01-18:00) - Number of actions done in evening (18:01-24:00) - Number of actions done in late night (00:01-06:00)
Consecution	Consequence	- Number of actions done consecutively (activities within 5-minute gap)

Since the main functions of Facebook are posting, commenting, replying to a comment and sharing a post, our features are designed regarding the functions. The activity for these functions includes posting a text, an image and a video or replying with a text and a sticker/emoticon. Differently, we do not count the number of replying since there is a vary number of obtained comments for each person. Thus, we choose the percentage of replying from a total number of obtained comments from other users. This work excludes the chatting function provided with Facebook messenger because we are allowed to monitor only the activities that users set as public or friend only. The designed features in different aspects are to combine such as an amount of posting (function) done in weekday (period) and in consequence (consecution).

Data of the features are collected as counting number representing their frequency; hence, it is less complex and explicitly comparable. This work collects only actions via Facebook of Thai users; thus, we ignore the content of the posts and the replies. Although the posts and replies may contain the keywords obvious for showing the sign of depression, we decide to focus solely on interaction between users in this work since it breaches less privacy and is simpler in terms of collection. Moreover, analyzing Thai text requires a lot more techniques of natural language processing in which is complex and remains challenge.

We can use the features for analyzing the behavior of users who are positive in being depression. Moreover, these features can be used to automatically identify users who may be at risk of being depression using an automatic classification method.

3.3. Data Modeling

Once data are transformed, they are used to generate a model for an automated classification as a depression detection model. To generate the model, this work applies supervised learning which requires labeled dataset for training. The training dataset includes input data and their response value. Supervised Learning is to map an input to a particular class from a label of an output, and the mapping becomes a model for classification. Supervised Learning algorithms thus help to make predictions for new unseen data by referring to the generated model.

In this work, a set of features for behavior usage of Facebook users is designed for supervised learning input while designate classes of the output are positive-depressive and negative-depressive state of a user. The supervised learning hence generates a model as a depress detection model to map the significant features that have an impact towards the depressive status classes. There are several available supervised learning algorithms such Naïve-Bayes (NB) [6, 7, 8, 9] Support Vector Machine (SVM) [6, 8] and Deep learning [12]. They all have their own advantages and limitations from a different training approach. Without adjusting their default algorithm, we will apply the feature set and compare a performance of the generated models.

4. Experiments

We designed an experiment to explore the potentials of the designed features. We provided results that we extracted from the dataset using the features. Moreover, we used them in an automated classification with supervised machine learning techniques to see how they perform in detection of depression.

4.1. Data and Settings

This study has been granted with a Certificate of Ethical Approval from the Human Research Ethics Committee of Walailak University with the Project No. WU-EC-IN-0-187-62. Participants were volunteers who willingly applied to our experiment. They were first required to take a PHQ-9 test to determine their depression state. After selection, we had 160 Thai participants who commonly use Facebook and allowed us to monitor their account for 14 days. All participants are over 18-year-olds. The questionnaire results showed that there are 120 participants who are positive in depression-at risk and the rest are negative.

4.2. Outstanding Features for Detecting Depression

In this section, we provided results that we extracted from the dataset using the designed features. The results of extracted data of Facebook usage regarding depression are given in Figure 2-4.

From the data, the highly different statistics are from the number of late-nights actions, number of afternoon actions and standard deviation (SD) of daily posts, respectively. In details from function-based data, SD values of daily posts and shares were noticeably different while average (AVG) values of posts and shares were slightly different. This finding is understandable since depression-negative users tentatively had

consistent numbers of daily posts and shares while depression-positive users may have unstable mood that affects their inconsistent behavior. The depression-positive users had more likely post because it is a process of releasing a state of negative feelings or suffering. In the contrary, AVG values of comments and reply percentage are more usable than their SD values. This is because depression-positive users scarcely conducted a reply or comment, but depression-negative users are more likely to interact to other users. From time-based features, we can see the time when the two groups are active. The depression-positive users had more activity numbers during late night since it is a time when depression is more likely to occur, including loneliness and insomnia. As a result, people with depression spend more time online. Last, consecutive actions from depression positive users were also significantly higher than those of depression-negative users.

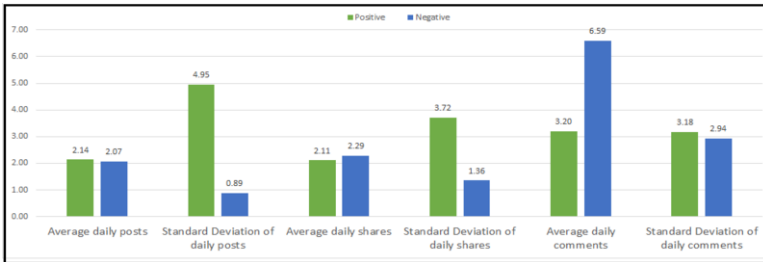


Figure 2. Data results of Facebook usage regarding depression according to the features of Facebook functions including posts, shares and comments.

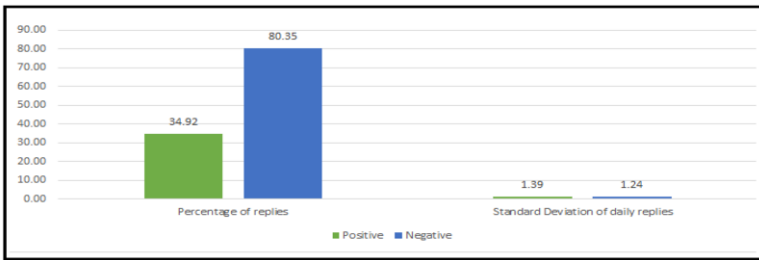


Figure 3. Data results of Facebook usage regarding depression of reply function

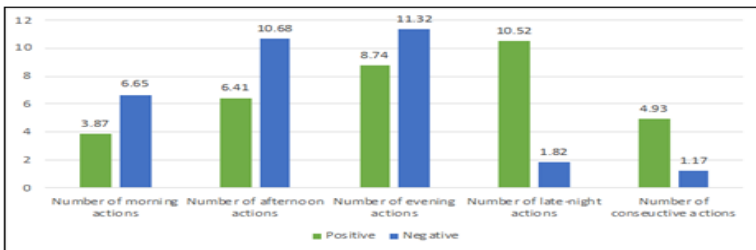


Figure 4. Data results of Facebook usage regarding depression according to the designed time and consecutiveness features

4.3. Performance of Classification Models

With the designed features given in Table 3, the data are in a form of numerical value and can be easily used in supervised machine learning. Four classification algorithms as SVM, NB, DT and deep learning with default setting are used to train the classifying model. With small set of data, 5-fold cross validation was applied. The evaluation results are shown in Table 4.

Table 4. Classification results from designed features

Algorithm	Precision	Recall	F-measure
SVM	81.4	88.1	84.6
NB	80.8	86.8	83.7
DT	77.9	80.3	79.1
Deep learning	84.4	89.7	87.0

From the results, deep learning performed the best in all measurements among the chosen algorithms. The results though were still low; thus, we decided to split training data based on participant gender and removed the data with less than 5 activities in total. There were 102 and 33 remained participants for female and male, respectively. Then, we experimented with the gender-based model and gained evaluation results as given in Table 5.

Table 5. Classification results from designed features splitting model based on participant's genders

Algorithm	Precision	Recall	F-measure
SVM	86.5	90.3	88.4
NB	84.6	89.9	87.2
DT	80.6	83.7	82.1
Deep learning	90.2	92.6	91.4

The results of classification were improved with splitting model based on participant's genders. This shows that gender of participants is an important factor for differentiating behavior of Facebook users.

5. Conclusion and Future Works

This paper proposes a method to detect depression at risk person via a behavior of Facebook users. A set of features related to Facebook functions interact able to other users is designed to represent users' activities in Facebook platform in a numerical value form. The designed features include posting, sharing, commenting and replying amount with periodic focus without interpreting its content. The limitation of the acquired data from Facebook user account is the number of activities as frequency of actions in a specific periodic time and consequential manner. Therefore, the data are easier to be collected with less privacy issue. To distinct between positive and negative depression-at risk, PHQ-9 questionnaire, commonly used to screen depression patient in Thailand, was used. The collected data according to the designed features with a depression result are used to classify a likelihood of being depress disorder. These features hence are used in supervised learning classification algorithm for detecting a risk of being depress disorder.

From the experiment of 160 Thai Facebook users, we found that the designed features based on users' activities from Facebook with deep learning algorithm yields about 87% in terms of F-measure. After analyzing the data, we thus split data regarding users' gender and removed obviously low active data, and the F-measure from classification went up to 91.4 which improves for 4 points. From the data, the highly significant features in differentiation between depression-positive and negative users are from the number of late-nights actions, percentage of replies, number of afternoon actions, standard deviation (SD) of daily posts and consecutive actions, respectively. In the future, we plan to include text analysis in detection of depression including sentimental analysis of a post, vulgarity degree of a language and keyword detection that depress person often used.

References

- [1] DSM-5 American Psychiatric Association. Diagnostic and statistical manual of mental disorders. Arlington: American Psychiatric Publishing, 2013.
- [2] 6 statistics about "Depression" that you may never know. Cited: 2017, 16 May. Available from: <https://www.sanook.com/health/7137/>.
- [3] Mike J. 1985. Classification algorithms. Wiley-Interscience.
- [4] Depression and suicide assessment form (Including 2q 9q and 8q assessment). DEPARTMENT OF MENTAL HEALTH Ministry of Public Health. (in Thai) Retrieved November 18, 2018, from: <https://www.dmh.go.th/test/download/view.asp?id=22>.
- [5] Halfin A. Depression: the benefits of early and appropriate treatment. *Amer. J. Manag. Care*, 2007, 13(4 Suppl), S92-7.
- [6] Hemtanon S, Kittiphattanabawon N. An automatic screening for major depressive disorder from social media in Thailand. The 10th National & International Conference "Global Goals, Local Actions: Looking Back and Moving Forward" Suan Sunandha Rajabhat University, Bangkok, Thailand, 2019 1(10), 103-113.
- [7] Vioulès MJ, Moulahi B, Azé J, Bringay S. Detection of suicide-related posts in Twitter data streams. *IBM Journal of Research and Development*, 2018, 62(1), 7-1.
- [8] Chandler V. Google and suicides: what can we learn about the use of internet to prevent suicides? *Public health*, 2018, 154, 144-150.
- [9] Desmet B, Hoste V. Online suicide prevention through optimised text classification. *Information Sciences*, 2018 439, 61-78.
- [10] O'dea B, Wan S, Batterham PJ, Calear AL, Paris C, Christensen H. Detecting suicidality on Twitter. *Internet Interventions*, 2015 2(2), 183-188.
- [11] Wu C, Landgrebe DA, Swain PH. The decision tree approach to classification, 1975.
- [12] Colombo GB, Burnap P, Hodorog A, Scourfield J. Analysing the connectivity and communication of suicidal users on twitter. *Computer communications*, 2016, 73, 291-300.
- [13] Coppersmith G, Leary R, Crutchley P, Fine A. Natural language processing of social media as screening for suicide risk. *Biomedical informatics insights*, 2018, 10, 1178222618792860.
- [14] Abboute A, Boudjeriou Y, Entringer G, Azé J, Bringay S, Poncelet P. Mining twitter for suicide prevention. In *International Conference on Applications of Natural Language to Data Bases/Information Systems*, Springer, Cham, 2014, June, pp. 250-253.
- [15] Burnap P, Colombo G, Amery R, Hodorog A, Scourfield J. Multi-class machine classification of suicide-related communication on Twitter. *Online Social Networks and Media*, 2017, 2, 32-44.
- [16] Sueki H. The association of suicide-related Twitter use with suicidal behavior: a cross-sectional study of young internet users in Japan. *J. Affective Disorders*, 2015, 170, 155-160.