Predicting and Detecting Depression through Social Media

Chao-Yuan Liu[1], Shih-Chun Lin[2], I-Ting Lin[2], and Heng-Chien Liou[2]

[1]*Department of Mechanical Engineering, National Taiwan University*
[2]*Department of Electrical Engineering, National Taiwan University*

July 1, 2020

Abstract

In this project, we explore the applicability to predict and detect depression using social media data. We review necessary background knowledge and previous works to deepen the discussion. We use two vectorization and embedding methods, TF-IDFT and BERT, for our main experiment. The result shows promising future for further investigation.

Contents

1 Introduction

The development of engineering and technology has shed light on many old problems of biomedical concerns. Microfluidics enable as to control cells and molecules; bioinformatics makes precision medicine toward cancer possible; imaging science and biomedical optics make breakthrough in our basic understanding of life. Yet some of the topics remain underdeveloped, without receiving great attention from biomedical engineering experts, such as topics about mental disorders.

On the other hand, the omnipresence of information technology produce a large amount of data about people's everyday lives. Such data inevitably contain information about people's feelings, thoughts, and behaviors. Such scenario renders us a possibility to utilize such data to understand people's mental health.

With the advance of machine learning method, which enable machine to extract valuable information from data, we can use the data produced by information technology to understand people's mental health. We choose one of the most important mental health problems, depression, as our target. We use the data collected from social media, which is also an inseparable part for many people's lives.

In Section 2, we discuss the background of depression, its clinical definition, etiology, and significance, from the views of psychology, psychopathology, and public health. In Section 3, we review several previous attempts to detect or predict depression using social media data. With the knowledge from these two section, we can further explicitly discuss the opportunities and challenges for such a task in Section 4, and formally formulate our problem in Section 5.

Following our conceptual discussion in the first half of this paper, we can dive deep into practical aspects of this work. In Section 6, we summarize important information of the data used. And in Section 7, we consider several methods for handling such a task. We evaluate our experiments in Section 8 and conclude in Section 9.

2 Background: Depression and its Significance

Generally, depression can refers to a group of affective disorder, including Major Depressive Disorder, Persistent Depressive Disorder, and Seasonal Affective Disorder [1], while sometimes depression refers to Major Depressive Disorder (MDD) exclusively; without explicit indication, we adopt the latter usage in this work.

2.1 Clinical Definition

According to *American Diagnostic and Statistical Manual of Mental Disorders*, 5th edition (DSM-5), MDD is characterized by one or more major depressive episodes (MDEs) [1, 2]. MDE is a period during which individual experience significant changes in various aspects of functioning, especially those in affect and cognition, with five or six symptoms below and at least one of them being depressed mood or anhedonia, for two weeks or longer [1, 2].

The symptoms considered are

- low or depressed mood,
- anhedonia,
- significant changes in appetite,
- disturbed sleeping patterns,
- psychomotor retardation or agitation,

- reduced energy levels,
- feelings of guilt and worthlessness,
- reduced concentration/decision making,
- and ideas or acts of self-harm or suicide [1, 2].

2.2 Etiology

In psychopathology, *equifinality* often applies: same product results from possibly different causes. There may be a number of reason to develop depression, and the episode may look similar [3]. Psychopathologists have identified several biological, psychological, and social factors that play significant roles in the development of depression. We outline some of them and corresponding theories in this subsection [1].

2.2.1 Biological Dimensions

Neurotransmitter Systems Research has identified low levels of serotonin in the causes of mood disorders [4, 5]. Serotonin regulates our emotional reactions [3]. Under low levels of serotonin, we are expected to be more impulsive and with larger mood swings. Yet the level of serotonin should be considered in the context of interactions with other neurotransmitters. Imbalanced level of dopamine is also identified as potential cause of depression, since chronic stress reduce dopamine levels and produce depressive-live behavior [5].

The Endocrine System Much attention has shifted to the endocrine system and the "stress hypothesis" [6, 7], which focuses on the activity in the hypothalamic-pituitary-adrenocortical (HPA) axis, which produces stress hormones. HPA axis coordinate the endocrine system [3], including cortisol systems. Coritsol is called a stress hormone, which is elevated during stressful events, and so is in depressed patients. This finding results into a biological test for depression, the dexamethasone suppression test (DST) [3]. Cortisol can be harmful to neurons, which suggests that elevated level of cortisol can lead to the suppression of neurogenesis in hippocampus [8]. And low hippocampal volume may precede or even contribute to the onset of depression [3].

Sleep and Circadian Rhythms Depressed people are found to enter rapid eye movement (REM) sleep more quickly, also with more intense REM activity [5]. Unusually short and long sleep duration were also associated with depression in adults [9].

Brain Structure and Function Depressed individuals exhibit greater right-sided anterior activation of their brains, especially in the prefrontal cortex [3, 10], with less left-sided activation and less alpha wave activity.

2.2.2 Psychological Dimensions

Life Events and Stress Stressful and traumatic life events, are one of the most striking contribution to psychological disorders, which is described in diathesis-stress model of psychopathology [3]. The contexts and meanings to individuals behind the events might be even more important [3, 11]. Humiliation, loss, and social rejection are the most potent stressful like events leading to depression [12].

[1] The organization of this subsection is greatly influenced by [3].

Learned Helplessness Psychologist Seligman suggested that people become anxious and depressed when they decide that they have no control over the stress in their lives [13], an important model called the learned helplessness theory of depression. Such an attributional style is internal, stable, and global [3].

Beck's cognitive triad In contrast to Seligman's model, Beck suggested that depression results from the tendency to interpret everyday events in negative ways [14, 15]. Beck catogorized two types of cognitive errors, arbitrary inference and overgeneralization. Beck suggested that depressed people make cognitive errors about themselves, their immediate world, and their future, the depressive cognitive triad [14].

2.2.3 Social and Cultural Dimensions

Marital Relations Depression is strongly influenced by interpersonal stress [16], and especially those in marital relations.

Sex and/or Gender Almost 70% of individual with depression are women [17]. The imbalance might result from the cultural roles taken by different sexes [3]. The culturally indeuced dependence and passivity might result into women's feeling of uncontrollability and helplessness. They also experience more discrimination, poverty, sexual harassment, and abuse. Women tend to place greater value on intimate relationships, which can be both protective and producing risk, depending on the strength of social networks. Disruptions in such relationships are damaging to women [18]. Women also tend to ruminate more [19].

Social Support People who lived alone has 80% higher risk of depression [20]. Having closed supportive friendship or not is a strong predictor of depression [21]. Studies also established the possibility of speed recovery supporting by social support [22], which later leads to the development of interpersonal psychotherapy.

2.2.4 An Integrative Theory

To summarized the discussion in this subsection, we can integrate all the factors outlined above into a single model, as shown in Figure 1.

2.3 Burden on Public Health

The lifetime risk of MDD is about 15%, which shows its relevance in the aspects of public health [23]. Moreover, depression is also greatly associated with disability, from fourth leading cause of disability of 2000 to the first health-related cause of disability of 2017 [23]. The mortality risk of depressed individuals are overall two-times greater than general population.

The cost from undiagnosed and untreated depression is even greater, with higher suicide rate and more usage of health care resource [23]. It is thus necessary to improve our methods to find potential patients and provide the assistance needed.

3 Related Works

With the advance of computation, automatic detection of depression has drawn lots of interest in recent years. Specifically, following the unceasing trend of using social media - such as Facebook, Twitter - in daily

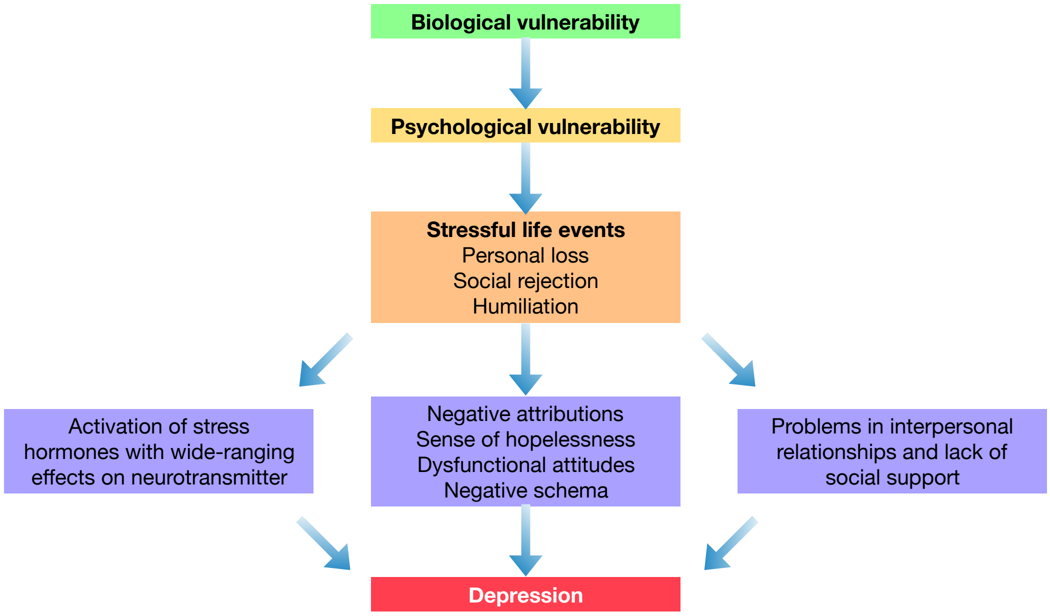

Figure 1: An integrative model of depression proposed by [3]. The image is drawn according to the FIGURE 7.7 in [3].

lives, many works have developed methods to detect depression and other mental health related concerns using the tremendous amount of data generated by social media.

De Choudhury, Gamon, Counts, and Horvitz's pioneering work marks the first notable attempt to use social media to detect and diagnose major depressive disorder in individuals [24]. They employed crowdsourcing to collect data from users who report being diagnosed with depression. The statistical classifier built can successfully estimate the risk of depression, *before* the reported onset. They further outlined the important behavioral cues, including the decrease in social activity, raised negative affect, highly clustered egonetworks, heightened relational and medical concerns, and greater expression of religious involvement [24].

Much more than what De Choudhury's team has done, the data collected by Shen et al. include more modalities [25]. They collected the text-based information from user profile, profile images, number of tweets, frequency of social interaction, posting time distribution, and other more specific feature extracted from tweets [25]. All these features are concatenated and learned via dictionary learning, in an attempt to get multimodal joint sparse representation of the extracted features [25].

The two works mentioned above used data with multi-modalities, with the reason to apply almost every facets of social media use. In contrast, the work by Husseini Orabi and colleagues and the work by Cornn tried to achieve great potential by exploiting simple data [26, 27].

Husseini Orabi, Buddhitha, Husseini Orabi, and Inkpen identified the lack of using deep learning methods in related research [26]. They attributed such phenomenon to the difficulties of obtaining sufficient amounts of annotated data. Therefore, they wished to find the most efficient ways for this task among several common architecture used in natural language processing [26]. Using the proposed optimization framework, they successfully estimate the risk of depression simply using tweets.

Cornn provided another work worth a mention, in which the scraped comments from Reddit are considered [27]. Cornn explored the applicability of various machine learning methods, including logistic regression, support vector machines, BERT-based models [28], and convolutional neural networks (CNN) with and without embeddings [27].

In addition to the works mentioned above, [29, 30] provide comprehensive reviews on related tasks.

4 Opportunities and Challenges

In addition to the related works mentioned before (Section 3), in this section, we conceptually discuss how is detecting and predicting depression through social media possible. We also outline several challenges faced in such tasks.

4.1 The Detectable-through-Social-Media Factors around Depression

In addition to the already-profound literature showing that such a task is empirically possible and promising, we can take a step backward and think about it in theoretical sense. As discussed in Section 2.2, there are lots factors might involved in the development of depression. A sensible question to ask is whether these factors are detectable or predictable through social media. And this is indeed the case as further shown by various works.

For example, in psychological dimensions, the trait *neuroticism*, which is correlated with almost all mental disorder, can be assessed through social media language [31]. Personal life events, especially those major, are also detectable using data from social media [32, 33].

4.2 The Idiosyncratic Nature of Social Media Data

Depending on the source, social media data gathered have different properties making standard methods ill-suited. Taking Twitter as example, a tweet is limited to length of 140 characters, and is often filled with noise, such as spelling errors, abbreviations, acronyms, initialisms, and special characters. These irregularities pose further challenges on the methods used.

5 Problem Formulation and Related Problems

In this section, we discuss how to formulate suitable problems. The formulation is neither too narrow nor too broad, in a sense that it can help up framework the solution practically and without losing flexibility in selecting appropriate methods.

Conceptually, we can at least formulate and distinguish the following two problems.

Problem 1. Depression Prediction
Using social media data to identify those individuals susceptible to depression, before the onset.

Problem 2. Depression Detection
Using social media data to identify those individuals with depression, after the onset.

Both these two problems are important. For those individual that are vulnerable to, but not yet experience with, depression, we can provide essential supports to prevent them from such a suffering. An ounce of prevention is worth a pound of cure. On the other hand, for those individuals with depression but unaware of it, providing cares and treatments may improve the conditions.

However, in practice, we might not be able to distinguish two tasks easily, since such a distinction relies on what are included in data: we need at least the information about time of onset to distinguish the two tasks. Since it is not viable in our work, we abuse the notations above, and formulate the problem of concern as below.

Problem 3. Depression Prediction/ Detection
Using social media data to identify those individuals who are either susceptible to depression or already with depression.

In the remaining parts of this work, we may use the terms 'prediction' and 'detection' interchangeably, all referring to the problem defined in Problem 3.

After formulating our problem, we can also discuss how such a problem is different to related tasks, at least conceptually.

5.1 The Difference with Sentiment Analysis

One of the well-known problems related is the sentiment analysis [34]. The two problems may seem similar at first glance, however, there are distinctive differences between. First, the levels of the objects of concern are different: in depression prediction/ detection, we consider an individual; while in sentiment analysis, we often consider a single sentence/ image/ or document. Second, the constructs we want to measure are also different. Depression is supposedly a coherent configuration of individual's thoughts, feeling, and behaviors. Instead, in sentiment analysis, we only address the emotion encoded in behaviors, sentences in most cases.

Here is a hypothetical example to illustrate the difference. Suppose a depressed individual post a tweet saying "I just brought a bag of coal this morning". Such a sentence is emotion-free, therefore sentiment analysis doesn't show any anomaly; however, a good depression detector should identified the suicidal clue behind such a sentence, which is also one of the most distinctive indicators of depression.

5.2 Fine-Grained Classification of Depression

A natural extension of the problem we address is to classify different kinds of depressions (in broadest sense), including persistent depressive disorder, seasonal affective disorder, depressive psychosis, and so on. We believe such a tasks is also possible, potentially with the techniques for fine-grained classification from deep learning literature [35]. However, to enable such an advanced tasks, a much more sophisticated source of data is needed, which must at least include some label for different kinds of depressions, which is not the case for the dataset considered in this work.

6 Data

6.1 Source

Part of the publicly available data collected from Twitter by Shen et al. are used in the development of this project. [25].

6.2 Structure & Classification

The data consists of three datasets, the Depression dataset, the Non-depression dataset, and the Depression-candidate dataset. Each user information in the datasets consists of an anchor tweet, the account information obtained at the time the anchor tweet was tweeted and a timeline folder containing all tweets tweeted one month before the anchor tweet.

6.2.1 Depression dataset (D1)

D1 is constructed based on tweets between 2009 and 2016. Users were labeled as depressed if their anchor tweets satisfied the strict pattern "(I'm/ I was/ I am/ I've been) diagnosed with depression".

6.2.2 Non-depression dataset (D2)

D2 is constructed based on tweets in December 2016. Users were labeled as non-depressed if they had never posted any tweet containing the character string "depress".

6.2.3 Depression-candidate dataset (D3)

The unlabeled depression-candidate dataset D3 is constructed based on the tweets on December 2016. Users were obtained if their anchor tweets loosely contained the character string "depress". This dataset is not used in this project.

6.3 Usage

Data of 6201 users from the D1 and D2 datasets containing relatively complete information is used in this project. Approximately 80 percent of the data is used in the training of the model, while the rest is used to test the trained model.

6.3.1 Model training

15 to 20 tweets in the timeline folder from each users are used to train the model. Users from the D1 datasets account for approximately 40 percent of all users in this process.

6.3.2 Testing

15 to 29 tweets in the timeline folder from each users are used to test the trained model.

7 Methods

7.1 Vectorization & Embedding

In this project, we have adopted two methods to transform a given tweet to numerical vectors.

7.1.1 Term Frequency-Inverse Document Frequency (TF-IDF) transformation

TF-IDF transformation is a common method to vectorize a text. The underlying concept of TF-IDF transformation is to measure the importance of each term in a sentence. In order to do so, we first introduce two quantities:

- Term frequency (TF): The number of times a term occurs in a sentence, which is normalized to the length of sentence.
- Inverse document frequency (IDF): Measurement of how often the term is used in other sentences. IDF for those terms frequently used in other sentences is small, indicating that those terms may not be informational in this sentence (e.g. This, are, a, the).

The TF-IDF score for a given term in a sentence is $TF \times IDF$. Finally, the procedure of TF-IDF transformation is listed as followed:

- Collect a set of tweets as a corpus (from training set).
- Each tweet is vectorized by calculating TF-IDF score of each term in the tweet and this score is the value of the corresponding dimension in the vector.
- Sum over all the vectors of transformed tweets to obtain the feature of this user.

7.1.2 Bidirectional Encoder Representations from Transformers (BERT)

BERT [28] is a powerful, deep leaning based language model proposed by Google in 2018. BERT is pretrained by two tasks in an unsupervised manner:

- Mask language model (Cloze test)
- Next sentence prediction

According to the above-mentioned tasks, BERT has a strong comprehension of text. Therefore, we want to take the advantages of BERT to vectorize a given tweet, obtaining a semantic and numerical representation of the tweet. To deploy BERT model in our task, we have fine-tuned BERT in five epochs and treated the first output of the BERT model as the feature.

7.2 Classifier

In our experiment, we use four different algorithms to classify the input feature, that is, predict whether the user is a potential patient of depression. In this section, we will give a brief introduction to these four algorithms.

7.2.1 K-Nearset Neighbor (KNN)

In this algorithm, the prediction of a embedded feature is decided by its neighbors in the high-dimensional space. If most of the K neighbors adjacent to the tested feature are positive then the algorithm will declare this user as positive, and vice versa.

7.2.2 Support Vector Machine (SVM)

SVM is designed to decide a linear hyperplane that can split most of the positive and negative features. The prediction of the tested feature is based on which region does it fall into.

7.2.3 Multi-Layer Perceptron (MLP)

MLP is a deep learning based classifier, there are multiple hidden layers between input and output neurons, each neuron accounts for some specific information in the feature. The weight of these neurons are updated by gradient descent. After several epochs of training, MLP is able to do the binary classification of tested feature.

7.2.4 Gradient Boosting Decision Tree (GBDT)

GBDT is one of a decision tree algorithm. The advantage of decision tree algorithm is that it is fast and explainable. After training, decision tree will develop lots of nodes to separate different features based on some learned conditions. We can inspect these explainable conditions and speculate what information is important to classify depression.

8 Experiments & Results

Without explicitly stated, the data used in this section are from the dataset collected by Shen et al. [25]. In the dataset, there are a few users posting less than five tweets before their anchor tweets. These data are removed during preprocessing, leaving 2303 users with depression and 3926 users without depression.

8.1 Settings

As metioned before, both TF-IDF and BERT are used to vectorize the tweets, which result two sets of features, 11333 and 768 features, respectively. Next, 5-Fold Cross Validation is applied for validating all models. Before training, the dataset is split into 5 folds equally. For each round, 4-fold data are used to train a model and the other one is for testing. Through this method, we can obtain five different outcomes for every model.

8.2 Results

In this work, we design two experiments. In Subsection 8.2.1, the experiment is to select a more suitable classifiers for this task. In Subsection 8.2.2, the experiment is to reveal the information in the tweets via analyzing the feature importance.

8.2.1 Comparing Classifiers

In this subsection, we use all 12101 features to train four different models and evaluate their performance and time consumption. The comparison is shown in Table 1.

To quantize the outcomes with unbalanced input, precision, recall and f1-score are applied. Precision is the ratio of correctly predicted positive observations to the total predicted positive observations. Recall is the ratio of correctly predicted positive observations to the all positive observations in actual class. F1 Score is the weighted average of Precision and Recall.

As shown in Table 1, the f1-scores of four classifiers are nearly the same while the time consumptions are notably different. The time comsumption of Lightgbm are five to ten times less than other algorithms. Note that features will increase as long as the number of tweets increases. As the result, Lightgbm is more suitable for this task since its time consumption is the least among the four algorithms. In Subsection 8.2.2, Lightgbm are applied for the rest experiments.

Classifier	Precision	Recall	F1 score	Time cost
KNN	0.978	0.960	0.968	$>$ 10 min
Linear SVM	0.971	0.963	0.967	5 $\sim$ 10 min
MLP	0.982	0.947	0.964	5 $\sim$ 10 min
Lightgbm	0.979	0.961	0.970	$<$ 1 min

Table 1: Classifier Comparison

8.2.2 Comparing Feature Extraction Methods

In this subsection, we will focus on the features with high correlation to depression or non-depression. Table 2 shows the evaluation of three feature sets.

Feature	Dimension	Precision	Recall	F1 score
TF-IDF	11333	0.779	0.607	0.682
BERT	768	0.979	0.962	0.971
TF-IDF+BERT	12101	0.979	0.961	0.970

Table 2: Feature Comparison

The BERT feature set and the TF-IDF+BERT feature set obtain almost the same performance, which implies that the 768 features of BERT completely cover the 11333 features of TF-IDF. To further analyze, all the tree-based algorithm, including Lightgbm, can compute the information gain for every decision and thus provide the importance of features. Figure 2 shows the 20 most important features in Bert feature set and TF-IDF+BERT feature set evaluated by Lightgbm. The highly overlapping of the most important features verifies again how crucial the BERT feature set is.

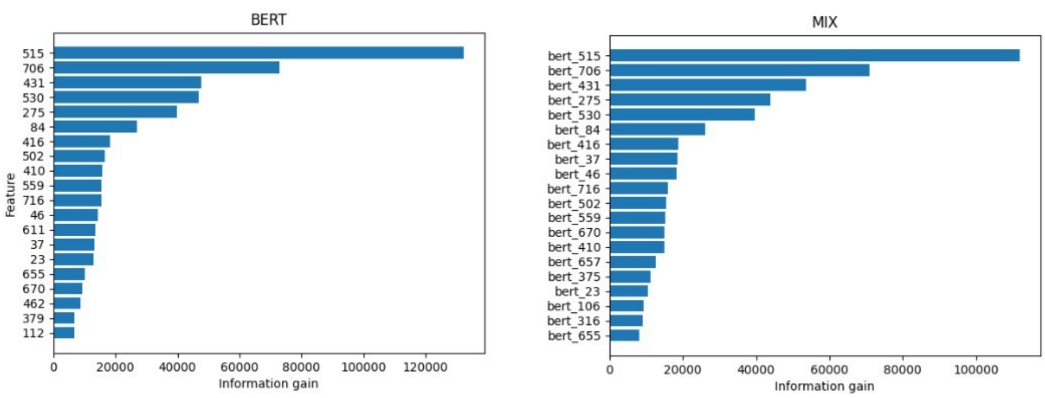

Figure 2: Feature Importance of BERT

With high performance, unfortunately, the features extracted by BERT is not easy to understand since BERT encodes the input tweets to a low dimensional output features. To reveal more information about the features, we turn to observe the feature importance of the TF-IDF feature set though its outcome performance is not good enough. Figure 3 shows the 20 most important features in the TF-IDF feature set.

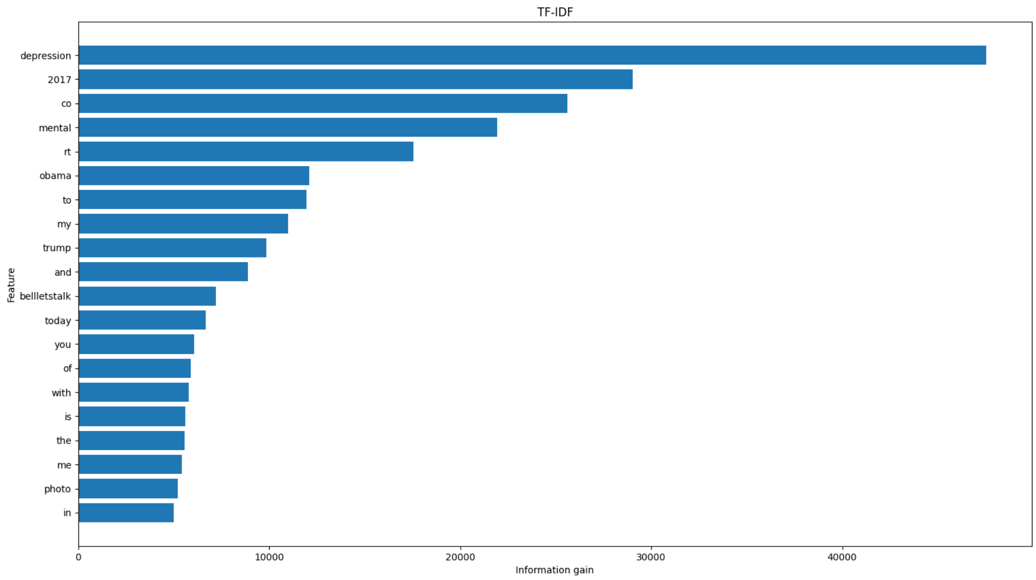

Figure 3: Feature Importance of TF-IDF

Some terms, e.g. depression, mental, and bellletstalk (Bell Let's Talk is a campaign to raise awareness and combat stigma surrounding mental illness), are comprehensible that depression sufferers may use more frequently. Besides, depression patients tend to use more first-person pronouns such as "me", "my" and "I" [36], which can be observed in our experiment as well.

9 Conclusion & Outlook

This work aims to predict and detect depression users in Twitter. We compare different models and select a suitable algorithm for this task. Next, we analyze the contribution of all features to reveal more information from tweets. As the results show, our experiments validate the depression prediction and detection is possible, Finally, we expect the techniques can be combined with other technology, such as Woebot [37], a automated conversational agent to deliver cognitive behavior therapy [38, 39], to immediately provide mental health assistance to those vulnerable to depression.

Appendix: Division of Labours

The division of labours roughly goes as below, without explicit clear cut.

- Chao-Yuan Liu: coordination, literature review and survey, paper writing (Section 6)
- Shih-Chun Lin: data handling, experiment, paper writing (Section 8 and 9)
- I-Ting Lin: data handling, feature extraction, experiment, paper writing (Section 7)

- Heng-Chien Liou: general affairs, literature review and survey, paper writing (Abstract, Section 1, 2, 3, 4, and 5)

References

[1] American Psychiatric Association et al. *Diagnostic and statistical manual of mental disorders (DSM-5®)*. American Psychiatric Pub, 2013.

[2] Edward S Friedman and Ian M Anderson. Classification, causes, and epidemiology. In *Handbook of Depression*, pages 1–12. Springer, 2014.

[3] David H Barlow, Vincent Mark Durand, and Stefan G. Hofmann. *Abnormal psychology: An integrated approach*. Cengage Learning, 8 edition, 2018.

[4] Michael E Thase, Ripu Jindal, and Robert H Howland. *Biological aspects of depression.* The Guilford Press, 2002.

[5] Michael E Thase. *Neurobiological aspects of depression.* The Guilford Press, 2009.

[6] Shaojie Bai, J Zico Kolter, and Vladlen Koltun. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling. *arXiv preprint arXiv:1803.01271*, 2018.

[7] Charles B Nemeroff. Psychoneuroimmunoendocrinology: The biological basis of mind-body physiology and pathophysiology. *Depress Anxiety*, 30(4):285–7, 2013.

[8] Erica R Glasper, Timothy J Schoenfeld, and Elizabeth Gould. Adult neurogenesis: optimizing hippocampal function to suit the environment. *Behavioural brain research*, 227(2):380–383, 2012.

[9] Long Zhai, Hua Zhang, and Dongfeng Zhang. Sleep duration and depression among adults: A meta-analysis of prospective studies. *Depression and anxiety*, 32(9):664–670, 2015.

[10] Wendy Heller, Jack B Nitschke, and Dana L Lindsay. Neuropsychological correlates of arousal in self-reported emotion. *Cognition & Emotion*, 11(4):383–402, 1997.

[11] George W Brown. *Life events and measurement.* Guilford Press, 1989.

[12] Scott M Monroe, George M Slavich, Katholiki Georgiades, et al. The social environment and life stress in depression. *Handbook of depression*, 2(1):340–60, 2009.

[13] Martin EP Seligman. Helplessness. *On depression, development and death*, 1975.

[14] Aaron T Beck. *Depression: Clinical, experimental, and theoretical aspects.* Hoeber Medical Division, Harper & Row, 1967.

[15] Aaron T Beck. *Cognitive therapy and the emotional disorders.* Penguin, 1979.

[16] Erin S Sheets and W Edward Craighead. Comparing chronic interpersonal and noninterpersonal stress domains as predictors of depression recurrence in emerging adults. *Behaviour research and therapy*, 63:36–42, 2014.

[17] Ronald C Kessler and Evelyn J Bromet. The epidemiology of depression across cultures. *Annual review of public health*, 34:119–138, 2013.

[18] Kenneth S Kendler and Charles O Gardner. Sex differences in the pathways to major depression: a study of opposite-sex twin pairs. *American Journal of Psychiatry*, 171(4):426–435, 2014.

[19] Susan Nolen-Hoeksema. *Sex differences in depression.* Stanford University Press, 1990.

[20] Laura Pulkki-Råback, Mika Kivimäki, Kirsi Ahola, Kaisla Joutsenniemi, Marko Elovainio, Helena Rossi, Sampsa Puttonen, Seppo Koskinen, Erkki Isometsä, Jouko Lönnqvist, et al. Living alone and antidepressant medication use: a prospective study in a working-age population. *BMC Public Health*, 12(1):236, 2012.

[21] George W Brown and Tirril Harris. Social origins of depression: a reply. *Psychological medicine*, 8(4):577–588, 1978.

[22] Gabor I Keitner, Christine E Ryan, Ivan W Miller, Robert Kohn, Duane S Bishop, and Nathan B Epstein. Role of the family in recovery and major depression. *The American Journal of Psychiatry*, 1995.

[23] Raymond W Lam. *Depression.* Oxford University Press, 2018.

[24] Munmun De Choudhury, Michael Gamon, Scott Counts, and Eric Horvitz. Predicting depression via social media. In *Seventh international AAAI conference on weblogs and social media*, 2013.

[25] Guangyao Shen, Jia Jia, Liqiang Nie, Fuli Feng, Cunjun Zhang, Tianrui Hu, Tat-Seng Chua, and Wenwu Zhu. Depression detection via harvesting social media: A multimodal dictionary learning solution. In *IJCAI*, pages 3838–3844, 2017.

[26] Ahmed Husseini Orabi, Prasadith Buddhitha, Mahmoud Husseini Orabi, and Diana Inkpen. Deep learning for depression detection of twitter users. In *Proceedings of the Fifth Workshop on Computational Linguistics and Clinical Psychology: From Keyboard to Clinic*, pages 88–97, New Orleans, LA, June 2018. Association for Computational Linguistics.

[27] Kali Cornn. Identifying depression on social media. https://web.stanford.edu/class/archive/cs/cs224n/cs224n.1194/reports/custom/15712307.pdf, 2019.

[28] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*, 2018.

[29] Sharath Chandra Guntuku, David B Yaden, Margaret L Kern, Lyle H Ungar, and Johannes C Eichstaedt. Detecting depression and mental illness on social media: an integrative review. *Current Opinion in Behavioral Sciences*, 18:43–49, 2017.

[30] Stevie Chancellor and Munmun De Choudhury. Methods in predictive techniques for mental health status on social media: a critical review. *npj Digital Medicine*, 3(1):1–11, 2020.

[31] Gregory Park, H Andrew Schwartz, Johannes C Eichstaedt, Margaret L Kern, Michal Kosinski, David J Stillwell, Lyle H Ungar, and Martin EP Seligman. Automatic personality assessment through social media language. *Journal of personality and social psychology*, 108(6):934, 2015.

[32] Jiwei Li, Alan Ritter, Claire Cardie, and Eduard Hovy. Major life event extraction from twitter based on congratulations/condolences speech acts. In *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, pages 1997–2007, 2014.

[33] Smitashree Choudhury and Harith Alani. Personal life event detection from social media. In *Workshop Proceedings.* CEUR, 2014.

[34] Rushlene Kaur Bakshi, Navneet Kaur, Ravneet Kaur, and Gurpreet Kaur. Opinion mining and sentiment analysis. In *2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom)*, pages 452–455. IEEE, 2016.

[35] Tianjun Xiao, Yichong Xu, Kuiyuan Yang, Jiaxing Zhang, Yuxin Peng, and Zheng Zhang. The application of two-level attention models in deep convolutional neural network for fine-grained image classification. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 842–850, 2015.

[36] Mohammed Al-Mosaiwi and Tom Johnstone. In an absolute state: Elevated use of absolutist words is a marker specific to anxiety, depression, and suicidal ideation. *Clinical Psychological Science*, 6(4):529–542, 2018.

[37] Kathleen Kara Fitzpatrick, Alison Darcy, and Molly Vierhile. Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (woebot): a randomized controlled trial. *JMIR mental health*, 4(2):e19, 2017.

[38] Barbara Olasov Rothbaum, Elizabeth A Meadows, Patricia Resick, and David W Foy. *Cognitive-behavioral therapy*. Guilford Press, 2000.

[39] Jesse H Wright, Gregory K Brown, Michael E Thase, and Monica Ramirez Basco. *Learning cognitive-behavior therapy: An illustrated guide*. American Psychiatric Pub, 2017.