Weibo Depression Posts Detection by Natural Language Processing

Chengbo Li [1,*,†], Honghong Liu [2,†], Bochen Yin [3,†], Jiaying Yang [4,†]

[1] Department of Statistics, College of Liberal Arts & Sciences, University of Illinois at Urbana-Champaign, Illinois

[2] Department of Mathematics, South China Normal University, Guangzhou, Guangdong, China

[3] Department of Math, School of Physical Science, University of California Irvine, California

[4] Department of Jinan University-University of Birmingham Joint Institute, Jinan University, Guangdong, China

* Corresponding Author Email: chengbo4@illinois.edu

† These authors contributed equally

Abstract. The goal of this paper is to detect depression based on the posts on social media. The dataset combined both tweet dataset and Weibo dataset scraped from the social media. To classify emotions, researchers have been using traditional models such as Bagging, Support Vector Machines, Decision tree, Multinomial Naïve Bayesian and K-nearest neighbor. In this paper, K-nearest neighbor is chosen based on the better precision in result. The main challenge is Chinese context translation and complexity of the context of the post. Finally, an UI page is designed to complete the mission to input a Weibo ID and output of the depression classification. Our approach achieves relatively higher quality results compared to the previous models in literature, while combining depression detection with the real-time social media posts. With this system, we demonstrate the practicability of our project by predicting the depression situation of Internet users through the model, to bring help to the depressed people or people with potential depression tendencies in society.

Keywords: Depression Detection, Natural Language Process, Machine Learning.

1. Introduction

"Every individual has a fundamental right to health" stated by Dr. Tedros Adhanom Ghebreyesus, who is the director-general of World Health Organization (WHO) [1]. Along with the development of industries and technologies, some health issues, especially mental health issues, start to appear in society and further become a severe and hostile topic. Among all the mental health disorders which can be detected, depression is one of the most common ones- which 5% of adults suffer from- and a leading cause of disability globally. Depression is defined by the long presence of depressive mood and the loss of interest in daily activities every day for at least two weeks. It is a major contributor to the health diseases worldwide. According to WHO's 2013 to 2030 Mental Health Action Plan, WHO is making efforts to teach individuals and groups about how to prevent and mitigate depression [2].

On the other hand, the strive of Web 2.0 Era led people to new means to express their feelings through Social Media. In De Choudhury et al. 's research in May 2013, they emphasized that the use of language in people's social media posts has the ability to express and reflect their emotions. They also demonstrated the potential of utilizing social media to analyze the trend of depression in society when people are categorized as different groups [3]. Similar to Facebook and Twitter in the United States, Weibo- which is a social media platform widely used by Chinese Internet users- contains a large group of people from all walks of life; and each user's emotional state can be expressed by his posts on Weibo.

This research, therefore, is aimed to use Natural Language Processing and emotional analysis to detect Weibo users' emotional state by analyzing posts of individuals and to provide a reminder to them or the user. Through using this program, the users can be detected as depression and can be told that they need immediate mitigation.

2. Related work

In this section, related works that perform sentiment analysis with machine learning and depression detection on social media will be mentioned.

2.1. Research for depression detection

Depression can be detected in various aspects because depression can manifest itself in people not only in patient self-reports and clinical opinion but also in their daily behavior. For example, Alghowinem, S, et al.'s paper built a classification model (depressed vs. non-depressed) based on the eye movement [4]. With the popularity of social software in recent decades, social platforms have almost become the place where people most often express their emotions.

Depression detection based on users' behavior on social media gains more attention, mainly analyzing the emotion by text or image content of users' posts. De Choudhury, M. et al. analyze their mental condition through behavioral attributes relating to social activity, sentiment, textual content and linguistic styles, ego network, and antidepression, which then is leveraged to train a model to predict the potential of depression [5]. Information on all aspects of a person's life could be gathered from the posts and accounts on social media, which explains why social media has been emerging as significant sources of information. It is also an efficient data source for depression detection.

Language is a powerful approach to express emotion; specific utterances convey specific emotion. As a result, mining on textual content of users' posts has become one of the most popular topics for analyzing the emotions of users. De Choudhury et al. explore the unexplored capacity of social media to monitor and judge of the depressive disorder in users [5], and there are plenty of models of sentiment analysis based on the micro-blog social network, where post contents are chosen to be the optimal data set for depression analysis and detection [6].

2.2. Research for emotion classification via text

Text-based emotion classification is a very important research field in natural language processing. Emotion classification is the automated process of identifying sentences in context and labeling them as happy, bored, neutral etc., based on the expression with words related to the specific emotions. Emolable is a Semi-Automatic Methodology for Emotion Annotation of Social Media Text, which aims to extract emotion from the information resources [7].

In order to predict depression by social media posts, it is important to train a model to classify the emotion of text and obtain binary classification with the depression and other emotions.

In the preprocessing of the raw dataset, the first step is to use the package named Term Frequency - Inverse Document Frequency (TF-IDF). It is usually applied to avert situations where the stop words (e.g., "the", "a") augment the noise in the dataset and deficiency of textual content [8]. The second one is Valence Aware Dictionary for Sentiment Reasoner (VADER), a package to attribute the polarity of language and to allocate them according to multiclass sentiment analysis and to achieve great accuracy in this research of classifying tweets' sentiment.

Recently, more and more papers on emotion classification have focused on social media research, like tweet sentiment analysis mentioned above. As an example, Pang, B. et al. mainly discusses to detect depression grounded in sentiment analysis of Micro-blog Social networks, combined with psychology, which aims to calculate the tendency of depression with ten depression features and the precision is around 80 percent in the end [6]. To handle emotion detection problems, a classification model will be the most suitable choice. Pang, B. et al. also focus on the sentiment classification based on the Twitter dataset, as well as the dataset used in this paper [9]. There are a variety of classification models such as Naïve Bayes, K-nearest Neighborhood, and Support Vector Machine. In Laeeq, F., et al. paper, F. Laeeq and N. M. Tabrez has done research on the sentiment classification from Twitter social media and have achieved the average probability of correct classification of K-NN, Naïve Bayes, and Decision Tree Classifiers are 77.50%, 80%, and 78% separately [10].

3. Methodology

3.1. Data collecting and labeling:

The beginning step of implement the sentiment analysis is training data set acquisition. The data set includes two parts: Tweet dataset and Weibo posts scrapped from Chinese social media.

The Tweet dataset was taken from Kaggle with 40000 rows of tweet contents and labeled with sentiments, including empty, angry, sadness, hate, boredom, neutral, enthusiasm, happiness, love, relief, fun, and surprise. These emotions, including neutral, surprise, love, fun, happiness, relief, and enthusiasm, are replaced by other emotions. Empty, sadness, worry, hate, boredom and anger are replaced by depression, which is helpful in identifying whether a post is depressed or not.

Figure 1 is the algorithm of composition of dataset.

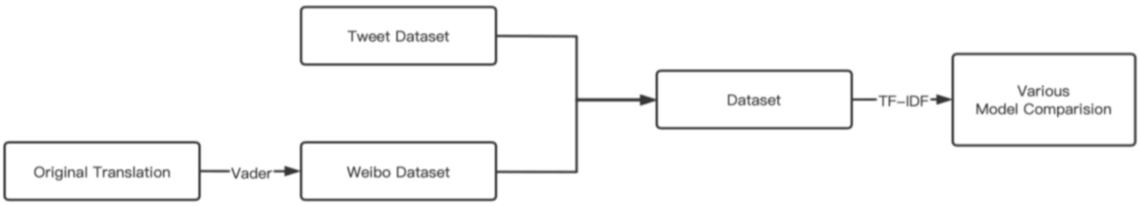

Figure 1. Algorithm

3.2. Data pre-processing

After collecting the data, the next stage is preprocessing, which is the stage of preparing the data for analysis. There are several steps in this stage, i.e., cleansing, translation and filtering.

3.2.1. Cleansing

In order to get data, which is easier to analyze, the raw dataset has to be cleaned up to reduce noise. The elements should be removed are the punctuation and the characters which are not needed. In this stage, the URLs beyond the scope of sentimental analysis of text, the repeated tags which would increase the weight of the emotion in the tag contents, the meaningless mentioned user's name and the reposted weibos which contain others' contents of reposts would be deleted. Table 1 clearly shows the comparison of data before and after cleaning.

Table 1. Cleansing Example

Deleted elements	Before	After
URLs	I'm so happy. http://wx2.sinaimg.cn/large/008cVke6ly1h2sz5z5pyqj32bc3344qq.jpg	I'm so happy.
Repeated tags	#Depression# Sad. #Depression# I'm so nervous.	#Depression #Sad. I'm so nervous.
Mentioned user's name	@abcd I'm angry.	I'm angry.
Reposted weibos	I'm not happy.//@abcd: I'm so happy.	I'm not happy.

3.2.2. Translation

Although many resources connected to the English language can be called, such as lemmatization, stemming, identifying the language, and a range of dictionaries, there is a problem that needs to be solved in order to properly classify the dataset afterwards. The task of identifying the depression tendency of Chinese language sentences becomes extremely difficult due to the dearth of good resources for the Chinese language. Therefore, translating from Chinese to English is the first step in finding a solution. In order to achieve this, a more accurate emotion analysis has been obtained using the google-trans-new library. Google-trans-new is a free Python library that calls the Google Translate API which implements the detection and translation. There is an example of the data cleansing Weibo posts, input “#Depression# I love this world and hate this world.” The output is “I love this world and hate this world.”

3.2.3. Filtering

Since the ultimate goal is to detect the depression level of Weibo users, a large number of homogeneous datasets with high levels are needed. After the Chinese texts are translated into English, the English version of Weibo posts has been divided into two categories: depression and other emotions, which utilizes the python library–Vadersentiment. Valence Aware Dictionary and Sentiment Reasoner, a well-known rule-based library in the field of sentiment analysis, computes the text sentiment using a list of lexical characteristics annotated as positive or negative on the basis of semantic orientation. The Vader sentiment use sentiment lexicons that contain intensity measures for each word and put greater emphasis on guidelines that seize the spirit of regularly used social media language. In the process of constantly improving the performance of the model, the training depression dataset, which contains the negative emotions, is obtained through the VADER library. It leads to the outcome of lower accuracy at the beginning, and then the library has been changed to filter the weibo posts.

3.3. Feature Extraction

The dataset was initially preprocessed following the feature extraction, and the TF-IDF values were acquired. For text mining as well as information retrieval, the TF-IDF weighting approach is widely employed to rate the frequency of word occurrences in text documents. This statistical technique combines the ideas of document and word frequency. The former is a notion that is weighted by frequency a word emerges in a document, and the latter is the total number of documents in which a phrase appears.TF-IDF is utilized to evaluate the significance of a word to a document, which increases with frequency increases in the text. It varies inversely with the corpus's frequency of occurrence. The fundamental tenet of TF-IDF is that a word or phrase has strong classification capacity if it appears frequently in one sentence and infrequently in others. This allows the dataset to be more effectively divided into depression and other emotions.

3.4. Machine Learning Approach

In the machine learning approach, four classification operators, including Decision Tree (DT), Multinomial Naïve Bayesian, Bagging, K-Nearest Neighbor (KNN), have been used for comparison.

3.4.1. Decision Tree

The decision tree algorithm is an approach that close to the value of discrete functions, it is also a representative categorization approach. After the machining the dataset and the generation of comprehensible regulation and decision trees by induction algorithms, the decision trees are used to process latest data. In fact, a decision tree is the procedure of categorization of features by a set of laws.

3.4.2. Multinomial Naïve Bayesian

Based on Bayes’ theorem and the assumption of feature condition independence, Naïve Bayesian is a classification model which calculates the probability of classification through the features and selects the situations with the high probability for classification. Therefore, it is a machine learning classification operator on the basis of probability theory and it also belongs to supervised learning for the definite goal of classification.

3.4.3. Bagging

Bagging algorithms (also known as Bootstrap Aggregating algorithms) are common group learning algorithms in machine learning, which can be combined with other classification and regression methods to improve its accuracy and stability, which reduces the variance of results to avoid the occurrence of over-fitting. The underlying classifier's stability affects how well it performs. Bagging can assist in minimizing mistakes brought on by random oscillations in the training data when the basic classifier is unstable.

3.4.4. K-Nearest Neighbor

K-Nearest Neighbor is a basic machine learning method whose idea is that to enter the test data in the condition of the known training data and labels and to focus on the comparison of the characteristics of the test and training data to find the most similar former K data in the training data. Then the corresponding category is the most classification in K data.

3.5. User Interface Design

User interface (UI) allows the users to interact with a forum using type-in textbox, buttons, checkboxes, and other elements to achieve their goals or to find and obtain the necessary information. The pursuit of UI design for this project was detecting a single user's depression level provided with his unique user ID- a series of numbers to help identify the user from other users. By achieving this mission by UI page, Weibo users' level of depression would be presented through a pie chart, and the page would provide one line of simple explanation on whether the user with this particular user ID is prone to depression or not.

IPyWidgets was used to achieve the inputting action of user ID and the initiation of processing Weibo user posts in the construction of UI. IPyWidgets is a Python library chosen to use for applying HTML interactive widgets, and it is designed particularly for JupyterNotebook [11]. Furthermore, two label widgets were created using it in this project: one is the text box in which user can type in uid (a series of unique number that every single Weibo user has), and the other is a button by clicking which people can trigger the presentation of a detection pie chart and one concise sentence describing the depression level of the user with the uid input.

4. Results

At the end of the project, the final result is an application based on the ipyWidgets expansion package and the UI for ordinary Weibo users to analyze the depression percentage of personal accounts based on any Weibo accounts.

4.1. Translation Accuracy

An important part of our project is Chinese to English translation. Based on the mainstream Translation websites in the market, we choose Google Translation and Youdao Translation. First of all, even though Google Translation is one of the most widely used Translation websites in the world, its main Translation database does not focus on Chinese-English Translation. Gabriel Fairman gives his opinion about Google Translate, that when using Google Translate the accuracy percentage can be close to 90%, but the problem is that 10% can make the content incomprehensible easily [12]. In contrast, Youdao, which focuses on Chinese-to-foreign language translation, has received more favorable comments. Youdao claims in their product," The device is also equipped with OCR technology with up to 99% of accuracy scanning rate, Neural Machine Translation, and Text-to-Speech technology" [13]. Finally, we chose to use Youdao to realize the translation process.

4.2. Model Evaluation

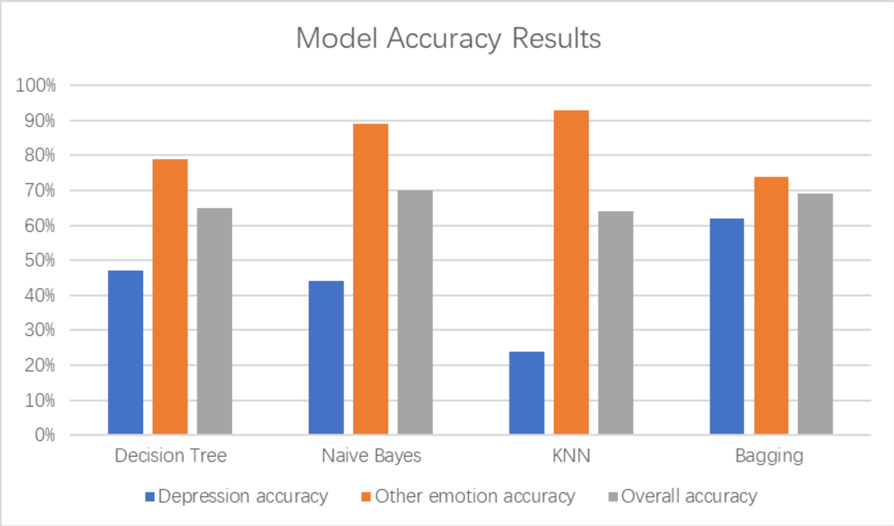

Figure 2. Model Accuracy Comparison

As shown in Figure 2, a total of four different models were applied and tested. They are Decision Tree, Naive Bayes, KNN, and Bagging. At the beginning of the project, the Decision Tree module was adopted first and always. For the initial Decision Tree module, the accuracy reached 65 percent. Only 47 percent accuracy was reached on depression analysis, while the accuracy of other emotions reached 79 percent. While the rest of the emotions were almost 80 percent accurate, depression-based emotions were only less than 50 percent accurate. It's not the best choice. Three different modules were tried based on the representation of the Decision Tree module. Naive Bayes had an overall accuracy of 70 percent, 44 percent based on depression, and 89 percent based on other emotions. KNN had an overall accuracy of 64 percent based on depression, 24 percent based on depression, and 93 percent accuracy based on other emotions. The Bagging model achieved an overall accuracy of 69 percent, 62 percent based on depression, and 74 percent based on other emotions. Based on the performance of the above four models, the most balanced Bagging model was selected as the final model.

4.3. Evaluation of composition datasets

At the beginning of the experiment, we only used the dataset "Emotion Detection from Text Predict Emotion from textual data: Multi-class text classification" by Pashupati Gupta, which contains 39,827 Twitter posts sentiment analysis datasets to build the initial models [14]. Later, the data of negative sentiment posts of Weibo users judged by the Vadersentiment model were added to fill the database. Using the decision tree model, we experimented with the effect of inputting Weibo data into the model. The final training results did not change much from the depression training set. Overall accuracy was all stable at 70 percent. However, that is not the only way to judge the model's effectiveness. Since the original data set is selected from the user posts on Twitter, the final goal of the project is to analyze the posts of Weibo users. There are certain differences between the two, both in language and culture. Based on this problem, we experimented to test the effect of actual analysis. We fed the model posts that had been artificially determined to be all depressed and had the machine analyze them to determine true accuracy. As an example of visualization, we fed the model a data set of 47 posts that people judged to be depressed. The most accurate result should show that all 47 posts are depressing.

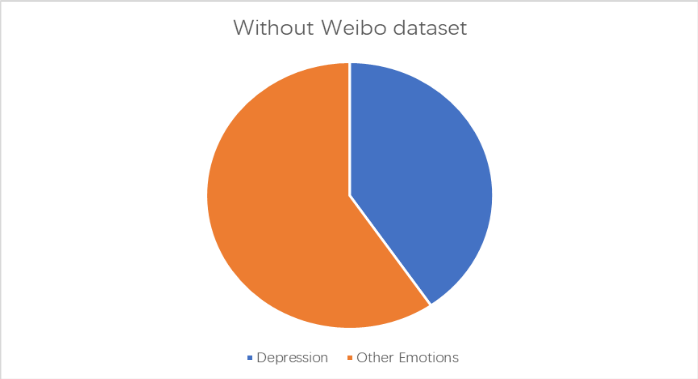

Figure 3. Experiment Result Without Weibo Dataset

For the model without Weibo data filling, 19 were judged as Depression, and 28 were judged as Other Emotions as shown in Figure 3.

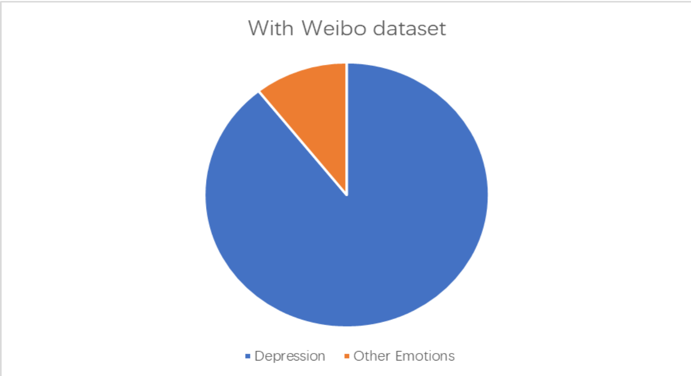

Figure 4. Experiment Result with Weibo Datase

For the model with Weibo data filling, 42 were judged as Depression, and 5 were judged as Other Emotions as shown in Figure 4.

We then carried out 10 experiments, all of which input all the artificially judged data as depression and let the two models make judgments. It was concluded that the correct rate of the model without Weibo data is 55%, and the correct rate of the model with Weibo data is 85%. It is concluded that filling the homologous Weibo data set can significantly improve the accuracy of the model prediction.

4.4. Future application

The result is a model that can be fed into any Chinese text to analyze the likelihood of depression, so it can be used by many Internet companies. First of all, many Users of Weibo express their pessimistic remarks on Weibo under the pressure of life, and many of them even commit suicide. If this project or similar projects are adopted by Weibo, it can help patients with depression to a certain extent or locate some potential users with depression tendencies. Secondly, WeChat, a chat software that has developed rapidly in China in recent years, can also be used in this project. Emotional analysis and reminders of the messages sent by all users in real-time can help users with depression tendencies to reduce the distress caused by depression to a certain extent.

5. Conclusion

With the development of technology, the young generation's life and social pressure gradually increased. As more and more people suffer from depression, many choose to post on Weibo or other social media to explain their problems and vent their emotions. Many people are already depressed and don't even know it. Many people suffer from depression but can only express their true thoughts on the Internet. To prevent the number of people suffering from depression, and to help those already suffering from depression, we have proposed a system that enables ordinary users to analyze the depression degree of users' posts through the model by only inputting the specific Weibo accounts. Our method only needs to input a Weibo account, which can be easily obtained by any user. We show the different model effects we applied, as well as the final output. Using machine learning and NLP algorithms, our model was able to predict depression in individual posts with nearly 70 percent accuracy. We demonstrated the practicability of our project by predicting the depression situation of Internet users through the model, to bring help to depressed people or people with potential depression tendencies in society.

References

[1] World Health Organization. (2022, May 23). Director-general's report to member states at the 75th World Health Assembly – 23 may 2022. World Health Organization. https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-address-at-the-75th-world-health-assembly---23-may-2022

[2] World Health Organization. (2021, September 13). Depression. World Health Organization. https://www.who.int/news-room/fact-sheets/detail/depression

[3] De Choudhury, M., Counts, S., & Horvitz, E. (2013, May). Social media as a measurement tool of depression in populations. In Proceedings of the 5th annual ACM web science conference (pp. 47-56).

[4] Alghowinem, S., Goecke, R., Wagner, M., Parker, G., & Breakspear, M. (2013, September). Eye movement analysis for depression detection. In 2013 IEEE International Conference on Image Processing (pp. 4220-4224). IEEE.

[5] De Choudhury, M., Gamon, M., Counts, S., & Horvitz, E. (2013, June). Predicting depression via social media. In Seventh international AAAI conference on weblogs and social media.

[6] Wang, X., Zhang, C., Ji, Y., Sun, L., Wu, L., & Bao, Z. (2013, April). A depression detection model based on sentiment analysis in micro-blog social network. In Pacific-Asia Conference on Knowledge Discovery and Data Mining (pp. 201-213). Springer, Berlin, Heidelberg.

[7] Canales, L., Daelemans, W., Boldrini, E., & Martínez-Barco, P. (2019). EmoLabel: semi-automatic methodology for emotion annotation of social media text. IEEE Transactions on Affective Computing.

[8] Patel, A., & Meehan, K. (2021, June). Fake News Detection on Reddit Utilising CountVectorizer and Term Frequency-Inverse Document Frequency with Logistic Regression, MultinominalNB and Support Vector Machine. In 2021 32nd Irish Signals and Systems Conference (ISSC) (pp. 1-6). IEEE.

[9] Pang, B., Lee, L., & Vaithyanathan, S. (2002). Thumbs up? Sentiment classification using machine learning techniques. arXiv preprint cs/0205070.

[10] Laeeq, F., Nafis, M., & Beg, M. R. (2017). Sentimental Classification of Social Media using Data Mining. International Journal of Advanced Research in Computer Science, 8(5).

[11] Jupyter Notebook - IPyWidgets. Tutorials Point. (n.d.). https://www.tutorialspoint.com/jupyter/jupyter_notebook_ipywidgets.html

[12] Fairman, G. (2022, January 31). How accurate is Google Translate in 2022? https://www.bureauworks.com/blog/how-accurate-is-google-translate-in-2022/

[13] Youdao. 2022. A magic scanning pen that is revolutionizing language learning: https://www.prnewswire.com/news-releases/a-magic-scanning-pen-that-is-revolutionizing-language-learning-youdao-dictionary-pen-3-global-version-301527914.html#:~:text=The%20device%20is%20also%20equipped, Text%2Dto%2DSpeech%20technology

[14] Gupta, P. 2021. Emotion detection from text. Kaggle. https://www.kaggle.com/datasets/pashupatigupta/emotion-detection-from-text