SOCIAL MENTAL DISORDER DETECTION USING SOCIAL MEDIA MINING

[1]Ajay Jadhav, [2]Karthik Morab, [3]Harshavardhana I, [4]Niranjana H.B

[1,2,3 & 4]Students of Computer Science and Engineering, R V College of Engineering

Abstract : As society becomes more technically advanced, we are more into the digital world. This has also brought numerous mental health issues for the individuals world-wide. Social networking platforms like Facebook and Twitter allows users to share their opinions and thoughts almost every day. This offer users a level of anonymity and provides users to be more uninhibited in their expressions. The objective of this paper is to propose depression detection and analysis of users based on Twitter data using sentiment analysis. Our study to date shows that such dataset that is readily available was too simplistic and most of the labelled depressive data had a word depression, which proves to be major obstacle to the development of this project, so we also aim to create our own a dataset that is designed precisely for depression identification based on tweets. We have evaluated the efficiency of this model using different approaches and the test accuracy was around 83% with test loss of 0.409 and also show that this proposed technique can significantly improve accuracy and classification error rate.

***Key Words* - Depression, Social Network Sites (SNS), Data Analysis, Natural Language Processing (NLP), Neural Network (CNN).**

I. INTRODUCTION

The improvement and development of internet and digital-communication technology, Specifically the emerging social media, has revived the way people communicate and interact with each other electronically. Now-a-days many people use social media to express emotions, feelings and thoughts on platforms like twitter, Facebook and Instagram. These applications not only allow to share multi-media messages but also allows users to express their feelings, sentiment and emotions related to any issue, area or subject. These platforms allow user a level of anonymity which provides users to be more uninhibited in their expressions. On one side it allows user on social media talk openly without any restrictions and discuss any topic or subject online. Whereas on the other side this generates an opportunity for social mental disorder detection based on the tweets made by the user. From medical perspective it creates opportunity to identify potential user with depression so that doctors can provide them appropriate support and help them to recover from the mental disorder.

One of the major social mental disorder is the depression which lead a person to major illness or even suicide. Around 4.4% of world's population is estimated to suffer from depression. Suicide is one of the major reasons for individuals with age between 15-29.

we aim to achieve detection of online social behavior and actively identify mental issues at an early stage and try to minimize the suicidal impacts of a social media users by providing them with necessary recovery technique. Facebook and Twitter are used by highest number of users in the world. Until 2019, there were 323 million active users monthly on Twitter whereas Facebook had more than 2.13 billion users. On these platforms user mostly post about their lifestyle, thoughts, sadness and happiness. By scraping data from the user's profile, we can get a overview of a user's personality. From this data we can extract data such as thinking style, sleeping hours, day-to-day transactions, loneliness, sadness, can be extracted. This data can help us to identify whether a particular individual has depression or not. In the paper Nadeem et al. [19] classified user activity into four different classifier and the result was that among many, the naïve Bayes model gave good accuracy as compared to others. Some authors such as Tsugawa et al. [9] also used this approach, but however it was clear that a detailed evaluation was required to evaluate the degree of disorder like depression through user activity history on social networking platforms.

Figure 1.1: Example of depressive tweet

In the proposed model, at first, we collect live tweets from particular Twitter ID. Collected information are then processed and analyzed. We apply various tasks using Natural Language Processing. The model of machine learning is trained and tested. We then classify the result into three different categories (Negative, neutral or positive). However, the initial model that was developed was able to attain good accuracy during validation, this was due to data was too simplistic and most of the labelled depressive entries had a word depression in it. Our research in this area of study showed such dataset was not available, this proved to be a big obstacle in the construction of this project. So, we decided to create our own dataset of twitter. We applied various techniques to build twitter dataset. Finally, we train and test this dataset.

II. LITERATURE SURVEY

Many researchers have utilized the advantage of social media sites to extract mental health status of users. So, there are numerous literature that analyzes and predict the structure of depression. Orabi et al. [1] mentioned that social networking sites basically show daily activities of a particular user on many levels. This paper showed how deep neural network can be used in this case. The idea was to detect depression using natural language processing: Recurrent Neural Network (RNN) and Convolution Neural Network (CNN), given a limited number of random data.

Choudhury et al. [2] considers that the social mental disorder such as depression is a real test of a person's well-being and general well-being. This paper tried to investigate whether online social media can be used to analyze and detect any sign related to depression with the help of a web-based social networking sites which evaluate behavioral credentials that reflect social engagement, dialect, emotion and semantic pattern and signs of anti-depressant. The Zhang et al. [3] have considered that when people are at serious risk of suicide, and if they are not identified by means of online network such as online microblogging sites then it is possible to consider the powerful intervention program to protect their sufferings.

To differentiate user-generated content (UGC) and SNS, many researchers have used different algorithm techniques. Ahmed and AL Darwish. [4] used the Naïve Classifier and support vector networks. Developed a web-based software that categorize different users of SNS into four different categories of depression. Nguyen et al. [5] made use of machine leaning and mathematical strategies to differentiate message from online between depression and category that psychological processes, personality and substance subjects eliminated from the user content formed by individual from these categories.

Another study of Hassan et al. [6] used machine learning approaches to investigate emotions to make comparisons of two or three subdivision: SVM, Maximum Entropy (ME) and Naïve Bayes (NB). They found that SVM showed higher results than NB and ME. Park et al. [6] investigated practices targeted on social networking sites which is web-based, to analyze whether a person is depressed or non-depressed. Semi-organized and informal gathering were conducted with 15 different and active Twitter users, in which 50% of them were depressed and the other 50% were non-depressed.

In spite of the fact that most of the paper is to examine the emotional process, the temporal, and to the linguistic style for detection of depression, we found the following deficiencies in the literature cited above. There are very less papers that have used sentiment analyzer of tweeter to detect social mental disorder. Also, most of the studies use readily available twitter dataset through online. However, these datasets are very simple and most of the depressive data have a word depression in it. So, we aim to develop a dataset that is specially created for depression detection of people using their twitter data.

III. METHODOLOGY

3.1. Module - 1

Initially We built the depression detector using sentient analyzer. The methodology used is discussed below. In this module we develop a sentiment analysis of twitter data that is nothing but classifying tweets of different twitter users into negative, neutral or positive based on sentiment of the tweet.

3.1.1. Input (Entering Keyword)

In this step we take raw data of a particular user of Twitter with the help of a python library called "Tweepy". It is a package for scaping twitter data using API.

3.1.2. Tweets Retrieval:

The tweets are retrieved from Twitter with the help of Twitter API. The Tweets are retrieved either based on the screen name of the user, userid, trending hashtags etc.

3.1.3. Data Pre-processing:

In this step the following processes are carried:

i. **Tokenization**: In this step the sequence of text is fragmented into words or some meaningful elements which is called as "Tokens".

ii. **Normalization**: in this step, if there are any abbreviations such as TTYL, whose full form is be 'talk to you later' within a text is replaced with actual meaning. Then words are transformed into lower case. If there are any reaped characters, then such characters are replaced by single character.

3.1.4. Data cleaning:

i. Removing @mentions.
ii. Removing '#' hash tag.
iii. Removing RT(Re-Tweets).
iv. Removing hyperlink.

3.1.5. Classification Algorithm:

Sentiment classification: In this step, sentiments of the tweets are detected first based on subjectivity and polarity of the tweets retrieved. The polarity and the subjectivity values for each tweet is determined and classification of tweets is carried out based on the polarity score. when the score is negative then the tweet is categorized as negative, when the score is equal to zero the tweet is categorized as neutral and the tweets with score greater than one are categorized as positive tweets. As per the understanding from the previous papers the we hope negative sentiments would further be classified as depressed, anxiety and other mental issues. After performing all the analysis, we represent the result with help of visual graphs such as pie chart or bar chart.

Even though we got sentiments of the user via tweet, this model is only limited for analyzing individual users. As our agenda is to detect it over an entire twitter network to detect if there is sign of any of the mental issues so we have proposed module 2.

3.2. Module -2

In this model we will create our own dataset using various techniques with proper ratio of depressed and non-depressed dataset and then train and test our model.

Step 1:

In this step a script is constructed that use following technique to identify tweets that are depressive:

i. With the help of 'Twint' API, we acquire tweets based on twitter-hashtags such as #hopelessness, #depression, #depressed, #loneliness etc.
ii. Removal of duplicate entries based on the ID of twitter accounts.
iii. Remove entries which contain any medical messages or educational messages with hashtags such as #Depressivehealth, #mental-health, #happiness, #happy, #joyless etc.
iv. Remove entries which contain following characteristics which might probably be promotional or advertising informative messages
    a) Tweets which contain @mentions
    b) Tweets more than three or four hashtags
    c) Tweets with URLs
v. Delete tweets with less than 26 characters or 6 words
vi. Then finally eliminate all the twitter hashtags from text extracted from twitter. This is very important step because if there is any depressive word in the text or tweet then the machine learning model can easily cheat with help of this word, so we want to train the model which mainly focusses on the content of tweet rather than the existence of such hashtags with word 'depression'.
vii. lastly, the results are stored into a CSV files and they are again manually reviewed.

Step 2:

In this step we manually check each and every csv files generated by past script, which basically contain tweets that are depressive which are filtered using hashtag 'depressive'. If there is any text which is not depressive, we label it as non-depressive. The csv has a target attribute which is set to 1 initially, and we set non-depressive data to have label with 0 to represent non-depressive tweet. One important step here is to remove tweets that are not English.

Step 3:

In this step to represent non-depressive texts, additional tweets from various sources are collected. It covers feelings or emotions such as surprise, love, joy, happy, along with neutral tweets.
Various non-depressive data is analyzed and collected based on certain criteria of intensity values of the text. Following non-depressive text are considered

i. Anger with intensity value less than 3.5.
ii. Fear with intensity value less than 0.45.
iii. Joy with intensity value greater than 0.55.
iv. Sadness with intensity value less than 0.4.

Sadness with intensity value greater than 0.75 is added into depressive data.

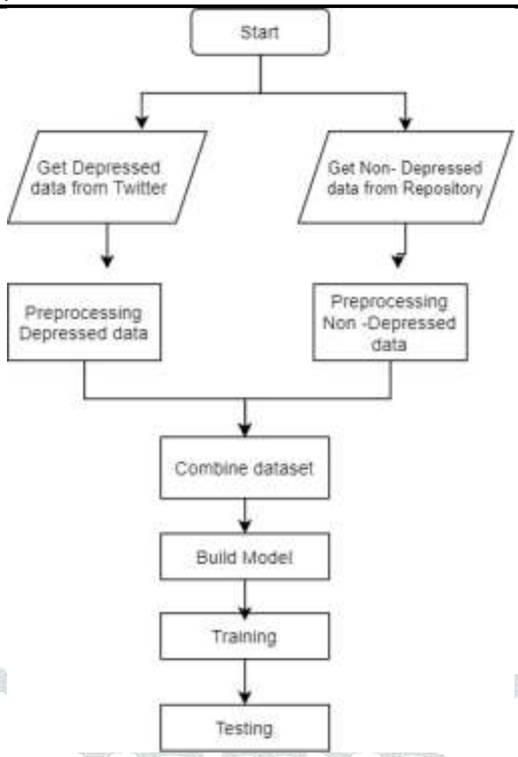


Figure 3.1: Flow chart for Model 2.

Step 4:
In this phase, we combine dataset generated by step 2 and step 3. Result of this step is a final dataset in CSV format. This CSV file contain appropriately 50% of depressive tweet and 50% of non-depressive tweets. This final dataset becomes a good resource to train the model with good quality of the depressive and non-depressive dataset.

Step 5:
In this step, the final dataset is preprocessed with custom techniques to work with this unique twitter data and then model is trained and tested.
Following tasks are performed in final step

i. Tokenization
ii. load pretrained word vectors and built vocabulary
iii. Loading the data in batches
iv. Finally Model and training using Concat Pooling model.
v. Testing the Model.

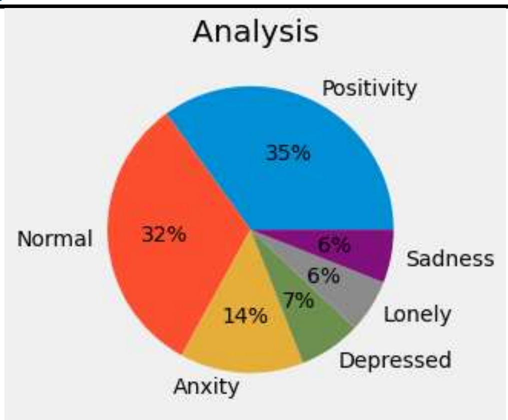
IV. FINDINGS & RESULTS

The sentiment analysis Model is tested on 80 Twitter user's account. Few significant users are given below.

Table 4.1: Analysis of Twitter User

User	No of post	Positive	Negative	Neutral	Result
Ravikumar	26	20%	73%	7%	Depressed
Nitzjotwani	88	28%	15%	57%	Neutral
iamVkohli	203	61%	6%	32%	Positive
GemmaXX03	81	35%	48%	17%	Negative
gellyman226	103	20%	29%	51%	Normal
kariss_cox	153	24%	49%	27%	Depressed

The above table contains users which different disorder classified based on positive, negative and neutral data. The visual representation of one the Twitter users is given below.



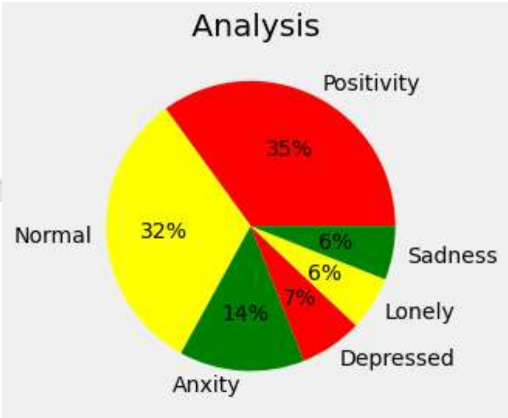


Figure 4.2: Pie chart for analysis of various Mental disorders.

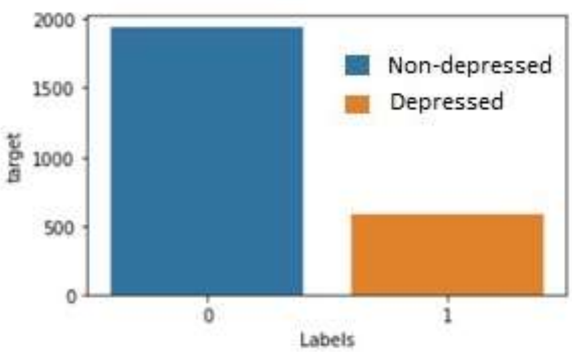


Figure 4: Distribution of new dataset.

The final result of module 2 was good, with around 83% test accuracy. After training the model and testing the Test Loss of model was 0.409 and Test accuracy was around 83%.

V. CONCLUSION

In this paper, it has been shown that sadness, unhappiness, anxiety or depression can cause a person to have serious mental disorder. Even on the path of suicide. It is also shown that how we can utilize social networking sites to identify depression of the users. This is because of social media platforms like as Facebook and Twitter allow user to express their opinions, feelings, emotions and also day to day activities that reflect user's personality traits and characteristic of behavior. In this project, we used Twitter as an online tool to demonstrate how we can use this tool to evaluate and identify social mental disorder like depression. In this model, we take any twitter user account ID and then perform analysis based on the posts, captions and comments and using this data identify the levels of susceptibility to disorder. This model of machine learning classifies this data into positive, negative and neutral based on which we classified users into different categories of mental disorders. But this sentiment analyzer was not predicting accurately for some tweets because the dataset was very simple and almost all the entries related to depression had a word depression in it. Because of this reason, we created a new dataset by our own using various techniques. After training the machine learning model on this dataset, the result evaluated was around 85% accuracy, which can be considered good in this case and would be helpful to detect depression on the social networking platforms. This project can be implemented in different social networking platforms like Facebook or Twitter, which will help us to identify users with depression at earliest stage.
The limitation of this project is that the model identifies only one single language English, so in future work we can extend our model to different popular languages.

REFERENCES

[1] A. Husseini Orabi, P. Buddhitha, M. Husseini Orabi, and D. 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, 2018.

[2] M. De Choudhury, M. Gamon, S. Counts, and E. Horvitz, "Predicting Depression via Social Media," In: ICWSM, vol. 13. 2013. p. 1–10.

[3] L. Zhang, X. Huang, T. Liu, A. Li, Z. Chen, and T. Zhu, "Using Linguistic Features to Estimate Suicide Probability of Chinese Microblog Users," Human Centered Computing Lecture Notes in Computer Science, pp. 549–559, 2015..

[4] M. M. Aldarwish and H. F. Ahmad, "Predicting Depression Levels Using Social Media Posts," 2017 IEEE 13th International Symposium on Autonomous Decentralized System (ISADS), 2017.

[5] Nguyen T, et al. Affective and content analysis of online depression com-munities. IEEE Trans Affect Comput. 2014;5(3):217–26.

[6] A. U. Hassan, "Sentiment Analysis of Social Networking Sites (SNS) Data Using Machine Learning Approach for the Measurement of Depression," 2017 International Conference on Information and Communication Technology Convergence (ICTC), 2017, doi:10.1109/ictc.2017.8190959.

[7] Park M, McDonald DW, Cha M. Perception differences between the depressed and non-depressed users in Twitter. In: ICWSM, vol. 9. 2013. p. 217–226.

[8] Albert Biffet and Eibe Frank. Sentiment Knowledge Discovery in Twitter Streaming Data. Discovery Science, Lecture Notes in Computer Science, 2010, Volume 6332/2010, 1-15, DOI: 10.1007/978-3-642-16184-1_1.

[9] Sho Tsugawa, Yusuke Kikuchi, Fumio Kishino, Kosuke Nakajima, Yuichi Itoh, and Hiroyuki Ohsaki. 2015. Recognizing depression from twitter activity. In HFCI. 3187–3196.

[10] Tuka Al Hanai, Mohammad M Ghassemi, and James R Glass. 2018. Detecting Depression with Audio/Text Sequence Modeling of Interviews.. In Interspeech. 1716–1720.