INTELLIGENT DEPRESSION DETECTION SYSTEM

B.Yamini [1], S.Aruljothi [2], S.M Madhumitha[3]

[1]*Assistant Professor, Computer Science Department, JEPPIAAR SRR Engineering College, Tamil Nadu*
[2,3]*UG Student, Computer Science Department, JEPPIAAR SRR Engineering College, Tamil Nadu*

---***---

Abstract-- Depression is becoming a serious leading mental health problem worldwide. It is a cause of psychological disorders and economic burden to a country. Even though it is a serious psychological problem, less than a half of those who have this emotional problem gained access to mental health service. This could be a result of many factors including having lack awareness about the disease. New solutions are needed to tackle this issue. We propose a system is to develop prediction models to classify users according to depression levels from Facebook data by employing Natural Language Processing (NLP) techniques where people use Facebook as a tool for sharing opinions, feelings and life events.

Keywords -- Natural Language Processing, Statistical analysis, Facebook site, Depression screening, User Generated Content (UGC).

1. Introduction

Social media can be exploited due to the large amount of information, which refers to user behavioral attributes. Getting use of that information to predict the social media users' mental health level can help psychiatrist, family or friends to get the right medical advice and therapy on time to the depressed user. WHO ranks the depression as one of the most devastating diseases in the world. In addition, about two thirds of depressed people do not seek appropriate treatments, which lead to major consequences. The medical science relies on asking the patients questions about their circumstances, which does not diagnose the depression in a precise way. The patient has to attend more than one session during a period of three weeks. The classification of a not depressed condition as a depressed is a False Positive problem. However, researchers found that the Electronic Health Record (EHR) systems are not optimally designed to handle integrating behavioral health and primary care. EHRs lack to support documenting and tracking data for behavioural health conditions such as depression. Most of the people use social media to express their feelings, emotions. Many researchers have been successfully proving that social media has been successfully used to maintain people's mental health. By mining the social media posts of users, we may get a complete image of the user natural behaviour, thinking style, interactions, guilt feeling, worthlessness, loneliness, and helplessness. Retrieving such behavioural attributes, show symptoms of depression on the social media users, which could be used to predict if the user is depressed or not. Psychiatrist, parents, and friends, could track the user depression by the proposed tool, which will save the time before the depressed user could get into major depression phase.

2. Related Work

Several approaches have been studied for collecting social media data with associated information about the users' mental health. Participants are either recruited to take a depression survey and share their Facebook or Twitter data, or data is collected from existing public online sources. These sources include searching public Tweets for keywords to identify (and obtain all Tweets from) users who have shared their mental health diagnosis, user language on mental illness related forums, or through collecting public Tweets that mention mental illness keywords for annotation. The approaches using public data have the advantage that much larger samples can, in principle, be collected faster and more cheaply than through the administration of surveys, though survey-based assessment generally provides a higher degree of validity. We first compare studies that attempt to distinguish mentally ill users from neurotypical controls.

A. Prediction Based on Survey Responses

Psychometric self-report surveys for mental illness have a high degree of validity and reliability. In psychological and epidemiological research, self- report surveys are second only to clinical interviews, which no social media study to date has used as an outcome measure. We discuss five studies that predict survey assessed depression status by collecting participants' responses to depression surveys in conjunction with their social media data. The most cited study used Twitter activity to examine network and language data preceding a recent episode of depression. The presence of depression was established through participants reporting the occurrence and recent date of a depressive episode, combined with scores on the Center forEpidemiologic Studies Depression Scale Revised (CES-D) and Beck's Depression Inventory (BDI). This study revealed several distinctions in posting activity by depressed users, including: diurnal cycles, more negative emotion, less social interaction, more self- focus, and mentioning depression-related terms throughout the year preceding depression onset. Reece et al. predicted user depression and posttraumatic stress-disorder (PTSD) status from text and Twitter meta-data that preceded a reported first episode with relatively high Areas under the Receiver Operating Characteristic (ROC) curve (AUCs) of .87 (depression) and .89 (PTSD). Data were aggregated to weeks, which somewhat outperformed aggregation to days, and modeled as longitudinal trajectories of activity patterns that differentiated healthy from mentally ill users. Tsugawa et al predicted depression from Twitter data in a Japanese sample, using the CES-D as their assessment criterion. Using tweets from the most recent 6–16 weeks preceding the administration of the CES-D was sufficient for recognizing depression; predictions derived from data across a longer period were less accurate. While most studies have used Twitter, used Facebook status updates for the prediction. Mothers self-reported a specific postpartum depression (PPD) episode and completed a screening survey. A model using demographics, Facebook activity, and content of posts before childbirth accounted for 35.5% of the variance in PPD status. Schwartz et al used questions from a personality survey to determine users' continuous depression scores across a larger sample of Facebook users (N = 28 749) than used in other studies (which typically range in the low hundreds). This study observed seasonal fluctuations of depression, finding that people were more depressed during winter months. This study also provided a shortlist of the words, phrases and topics (clusters of semantically coherent words) most associated with depression. Survey responses provide the most reliable ground- truth data for predictive models in this emerging literature.

B. Prediction Based on Self-declared Mental Health Status

A number of studies use publicly accessible data. 'Self-declared' mental illness diagnosis on Twitter (identified through statements such as 'I was diagnosed with depression today') is one such source of publicly-available data. We review seven studies of this kind. Helping to facilitate studies of this kind, a Computational Linguistics and Clinical Psychology (CLPsych) workshop was started in 2014 to foster cooperation between clinical psychologists and computer scientists. 'Shared tasks' were designed to explore and compare different solutions to the same prediction problem on the same data set. In the 2015 CLPsych workshop, participants were asked to predict if a user had PTSD or depression based on self- declared diagnoses on Twitter (PTSD n = 246, depression n = 327, with the same number of age- matched control and gender-matched control). Participating teams built language topic models (e.g. an anxiety topic contained the words: feel, worry, stress, study, time, hard), sought to identify words most associated with PTSD and depression status, considered sequences of characters as features, and applied a rule-based approach to build relative counts of N-grams present in PTSD and depression statuses of all users. The latter resulted in the highest prediction performance. All approaches found that it was harder to distinguish between PTSD and depression versus detecting the presence of either condition (compared to controls), suggesting overlap in the language associated with both conditions. On a shared dataset similar to the 2015 CLPsych workshop, the prediction of anxiety was improved by taking gender into account in addition to 10 comorbid conditions. Other studies have used psychological dictionaries (Linguistic Inquiry and Word Count; LIWC) to characterize differences between mental illness conditions, with some success. On the same dataset, Preotic-Pietro et al observed that estimating the age of users adequately identified users who had self-declared a PTSD diagnosis, and that the language predictive of depression and PTSD had large overlap with the language predictive of personality. This suggests that users with particular personality or demographic profiles chose to share their mental health diagnosis on Twitter, and thus that the results of these studies (mostly, prediction accuracies) may not generalize to other sources of autobiographical text.

C. Prediction Based on Forum membership

Online forums and discussion websites are a second source of publicly-available text related to mental health. They offer a space in which users can ask for advice, receive and provide emotional support, and generally discuss stigmatized mental health problems openly. We review three such studies here. The forum (reddit) posts were used to study the mental well-being of U.S. university students. A prediction model was trained on data gathered from reddit mental health support communities and applied to the posts collected from 109 university subreddits to estimate the level of distress at the universities. The proportion of mental health posts increased over the course of the academic year, particularly for universities with quarter-based, rather than semester-based, schedules. The language of 16 subreddits covering a range of mental health problems was characterized using LIWC and other markers of sentence complexity. De Choudhury et al examined posts of a group of reddit users who posted about mental health concerns and then shifted to discuss suicidal ideation in the future. Several features predicted this shift: heightened self- focus, poor linguistic style matching with the community, reduced social engagement, and expressions of hopelessness, anxiety, impulsiveness, and loneliness.

D. Prediction based on Annotated Posts

The publicly-available text involves manually examining and annotating Tweets that contain mental health keywords. Annotator's code social media posts according to pre-established (a priori, theory driven) or bottom-up (determined from the data) classifications; annotations can be predicted from the language of posts. 46 Big data in the behavioural sciences 1 that is, not using cross- validation. Most annotation studies on depression focus on identifying posts in which users are discussing their own experience with depression. Annotators are provided with guidelines on how to recognize a broad range of symptoms of depression that are derived from clinical assessment manuals such as the DSM-5, or a reduced set of symptoms, such as depressed mood, disturbed sleep and fatigue .Annotation has also been used to differentiate between mentions of mental illness for the purpose of stigmatization or insult as opposed to voicing support or sharing useful information with those suffering from a mental illness. In general, annotations of posts are a complementary (but labor-intensive) method that can reveal life circumstances associated with mental illness not captured by traditional depression diagnostic criteria.

3. Methodology

The Depression prediction model uses Natural Language Processing. The model consists of two datasets, and seven main operators. The first dataset is the training dataset which contains the manually trained depressed posts and not-depressed posts. The second dataset consists of the patient Facebook posts and it is changed for every individual to test the prediction of the model. The first operator is the Select Attributes which selects which attributes of the training dataset should be kept and which attributes should be removed. The second and the third operators are the Nominal to Text, this operator changes the type of selected nominal attributes to text, also it maps all values of the attributes to corresponding string values, it is used in the training dataset and the test set. The fourth and the fifth process are Process Process Documents and it is used in the training dataset and the test set which generates word vectors from string attributes and it consists of four operators. The four operators of Process Document operator are Tokenize, Filter Stop- words, Transform Cases, and Stem. The Tokenize operators splits the text of a document into a sequence of tokens. The filter Stop-words filters English stop words from a document by removing every token which equals a stop-word from the built-in stop-word list. The Transform Cases operator transforms all characters in a document to lower case. The Stem operator stems English words using the Porter stemming algorithm intending to reduce the length of the words until a minimum length is reached. The sixth operator is the Validation operator which contains which applies on the training dataset which consists of two main sections, training, and testing. The seventh and last operator is Apply model which connect the test dataset and training dataset to give us the final result of the prediction using one of the classifiers in the patients. The accuracy of the classification depends on the training set used to run the Natural Language Processing.

4. Proposed System

The approach being taken up in this paper is modular in its' organization. Individual component of the work flow has been segregated into stand-alone steps, to improve quality of implementation. The work flow starts with data collection step, which utilizes Facebook API for a generation of dataset. Natural language processing facilitates much better than average classification for sentimental analysis done by the human. Following the creation of datasets, the data pre processing module which systematically churns the data through tokenization, stemming and stop words removal. POS tagger then identifies essential pieces of the text to be utilized. After this the text classifier is trained on the processed text data from Facebook, in the training phase. In the testing phase, class prediction is made on the test dataset to identify potential Text demonstrating depression tendencies.

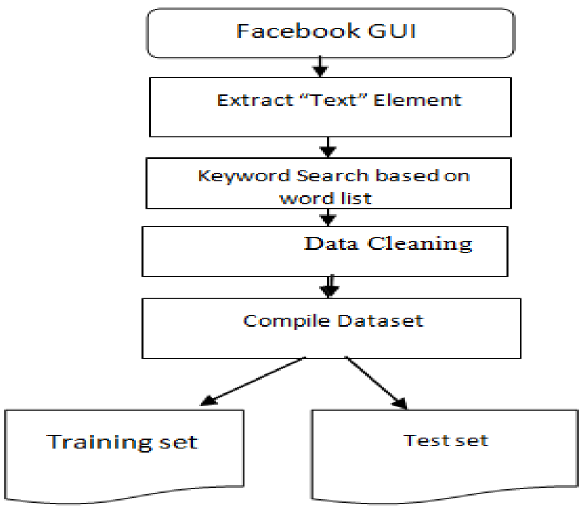

Figure 1 Data Collection

A) Data Collection

Sentiment analysis task starts with a collection of relevant data from various sources. In this Facebook is considered as the data source for analysis, in the form of User Generated Content. This portion covers tasks from streaming the data, to compile the training and test datasets.

1) Facebook API: By Facebook API, we collect the user generated content specifically images.
2) Keyword List: Using a pre-created wordlist for detecting trigger words symbolizing poor mental well- being, Text from all over the world are collected at random. These keywords specific texts are mixed with a general batch of non-weighted Texts.
3) Extracting Text from JSON: The collected text in the JSON objects are parsed to extract only the text field of the Text. Other meta-data related to any particular Text is removed.
4) Data Cleaning: To avoid errors in encoding textual data, the Text is purged for links (http) and non-ASCII characters like emoticons. Result is a clean dataset, rid of non-confirmative character types.
5) Generate CSV File for Train and Test Set: The cleaned text data from individual Text is added to the training and test dataset, in a vectorized format. Classification labels for the training and test datasets are manually added, to create a CSV file using comma as the delimiter.

B) *Data Pre-processing:*

The CSV file is read and several data pre-processing steps are performed on it. Natural language processing has been utilized for pre-processing methods applied on the extracted data:

1) Tokenization: Tokenization is a process of dividing a string into several meaningful substring, such as units of words, sentences, or themes. In this case, the first column of the CSV file containing the tweet is extracted and is converted into individual tokens.
2) Stemming: Stemming involves reducing the words to their root form. This would help us to group similar words together. For implementation, Porter Stemmer is used.
3) Stop Words Removal: The commonly used words, known as stop words need to remove since they are of no use in the training and could also lead to erratic results if not ignored. Nltk library has a set of stop words which can be used as a reference to remove stop words from the text.
4) POS Tagger: To improve the quality of the training data, the tokenized text is assigned the respective parts of speech by using POS Tagger. This would be used to extract only the adjectives, nouns and adverbs since other parts of the speech are not of much significance.

After all these pre-processing steps, a bag of words is formed. Bag of words calculates the number of occurrences of each word, which is then used as a feature to train a classifier.

C) Training

The classifier requires two parameters: training set and label. The training set in this case is the set of tweets which needs to be further processed in order to feed into a classifier. The set of tweets need to convert into vector format for further processing. The set of labels corresponding to each text is also fed into the classifier in the form a vector. Saving the Classifier and the Count Vectorized Object: Since training needs to be done once, the trained classifier object needs to be loaded into a pickle file. Same is applicable with the Count Victimizer object. Thus, both these objects are dumped into a pickle file for further use.

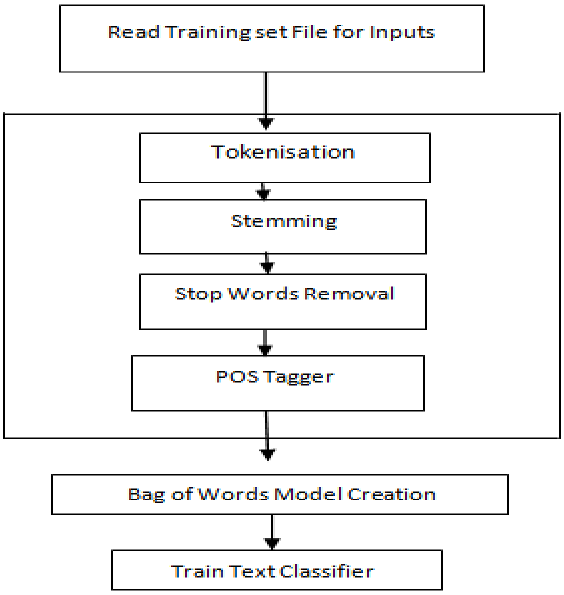

Fig. 2 Training Phase

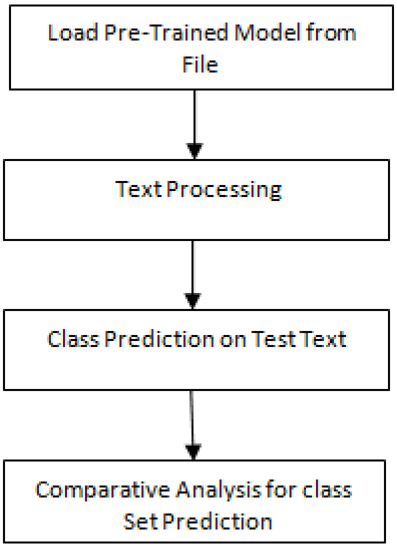

Figure 3 Testing Phase

D) Testing

1) *Loading:* Classification models are loaded from the pickle file, to be used for prediction on test dataset.
2) *Data Preprocessing*: The test dataset is preprocessed in a manner similar to the training data.
3) *Class Prediction on Test Text*: Each tweet is classified into a depressed or not depressed or neutral class.
4) *Comparison Based on the class prediction*: At last the compare the all class prediction set and finalize the result whether the person in depressed state or not. Further depression state is classified into three state (MINIMAL, MAXIMAL AND SEVERE) and also, we give a suggestion according to their person depression state and nearby doctors location.

5) Result

To analyze the performance of our proposed model, we calculated precision and recall. The precision gives us the percent of the positive correct prediction of processed images and recall give us the percent of the false prediction of processed images. The performance was tested with up to 100 images. Table 1 shows computing metrics of SVM classifier. Table 2 shows the value for each True positive, True

Negative, False Positive, False Negative of SVM classifier in depression prediction.

TABLE 1

Computing Metrics

Metrics	Text in Images	Classifier Output
True Positive(TP)	Positive data	Positive Data
True Negative(TN)	Negative Data	Negative Data
False Positive(FP)	Negative Data	Positive Data
False Negative(FN)	Positive Data	Negative Data

TABLE 2

Comparative Analysis Of Related Work

Total Image	TP	TN	FP	FN	Precision	Recall
10	5	3	2	1	0.714	0.833
30	10	12	5	3	0.666	0.769
50	21	11	9	9	0.7	0.7
70	30	25	11	4	0.731	0.882
100	45	39	8	8	0.849	0.849

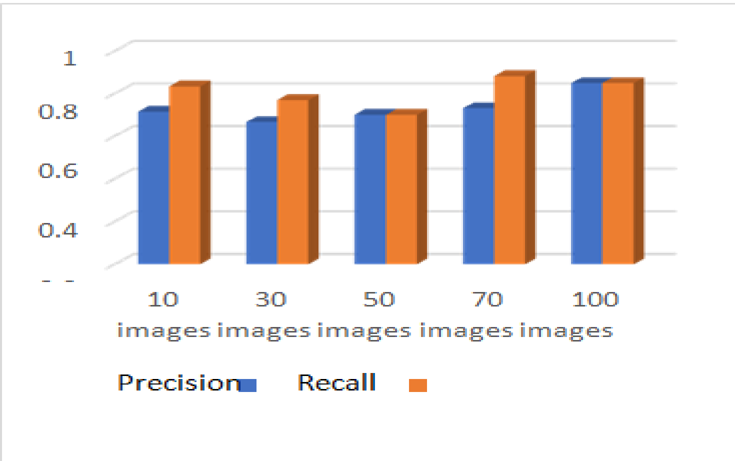

Fig 3. A Sample Bar Graph Using Precision and Recall Value.

Evaluating the system comparing with related work in Table 6 where we compare the performance metrics of the proposed system and the related works. Our proposed system got the best precision and least in recall.

5. Discussion

The greatest potential value of social media analysis may be the detection of otherwise undiagnosed cases. However, studies to date have not explicitly focused on successfully identifying people unaware of their mental health status. In screening for depression, multi-stage screening strategies have been recommended as a means to alleviate the relatively low sensitivity (around 50%) and high false positive rate associated with assessments by non-psychiatric physicians or short screening inventories. Social- media based screening may eventually provide an additional step in a mental health screening strategy. Studies are needed that integrate social media data collection with gold standard structured clinical interviews and other screening strategies in ecologically valid samples to test the incremental benefit of social media-based screening and distinguishing between mental health conditions.Self- reported surveys and clinical diagnoses provide snapshots in time. Online social media data may 'fill in the gaps' with ongoing in-the-moment measures of a broad range of people's thoughts and feelings. However, as depressed users may cease generating social media content, alternative uninterrupted data streams such as text messages and sensor data should also be tested for ongoing monitoring applications.

7. Conclusion:

Psychological depression is threatening people's health. It is non-trivial to detect depress level timely for proactive care. Therefore, we presented a framework for detecting users' psychological depression states from users' monthly social media data, leveraging face book post' content as well as users' social interactions. Employing real-world social media data as the basis, the project goal is to develop a web application which takes social media posts and predict output as various depression levels (minimal, maximal and severe).Using Natural Language Processing algorithm to increase the accuracy of system. And its deliver appropriate doctor's information depending upon location of user. According to his Facebook post system can find out user in depressed or not which is provided by the system.

REFERENCES

[1] S. R. Chowdhury, M. Imran, M. R. Asghar, S. Amer- Yahia, and C. Castillo, "Tweet4act: Using incident-specific profiles for classifying crisis-related messages," in 10th Proc. Int. Conf. Inform. Syst. Crisis Response Manag., 2013.

[2] T. Davidson, D. Warmsley , M.W.Macy, and I.Weber,"Automated hate speech detection and the problem of offensive language," in Proc.11thInt.Conf.WebSocialMedia,2017,pp.512–515 T. Hofmann, "Probabilistic latentsemantic indexing ," inProc.An nu.Int.ACMSIGIRConf.Res.Develop.Inf.Retrieval,1999,pp. 50–57.

[3] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent dirichlet allocation," J. Mach. Learn., vol. 3, pp. 993–1022, 2003.

[4] Y. Wang, E. Agichtein, and M. Benzi, "TM-LDA: Efficient online modeling of latent topic transitions in social media," in Proc. ACM SIGKDDInt.Conf.Knowl.Discovery DataMining, 2012, pp.123–131.

[5] S. Sidana, S. Mishra, S. Amer-Yahia, M. Clausel, and M. Amini, "Health monitoring on social media over time," in Proc. 39th Int. ACM SIGIR Conf. Res. Develop. Inf. Retrieval, 2016, pp. 849–852.

[6] K. W. Prier, M. S. Smith, C. Giraud-Carrier, and C. L. Hanson,"Identifying health-related topics on twitter," in Social Computing, Behavioral-Cultural Modeling and Prediction. Berlin, Germany: Springer, 2011, pp. 18–25.

[7] C. Cortes and V. Vapnik, "Support-vector networks," Mach.Learn., vol. 20, no. 3, pp. 273–297, 1995. [Online]. Available: http://dx.doi.org/10.1007/BF00994018

[8] M. De Choudhury, "Anorexia on tumblr: A characterization study," in Proc. 5th Int. Conf. Digital Health, 2015, pp. 43–50.

[9] M. De Choudhury, A. Monroy-Hern andez, and G. Mark, ""narco" Emotions: Affect and desensitization in social media during the mexican drug war," in Proc. SIGCHI Conf. Human Factors Comput.Syst., 2014, pp. 3563–3572.

[10] U. Pavalanathan and M. De Choudhury, "Identity management and mental health discourse in the social media," in Proc. 24th Int.Conf. World Wide Web, 2015, pp. 315–321.

[11] F. Bouillot, P. Poncelet, M. Roche, D. Ienco, E. Bigdeli, and S.Matwin, "French presidential elections: What are the most effi-cient measures for tweets?" in Proc. 1st Edition Workshop Politics Elections Data, 2012, pp. 23–30.

[12] L. Hemphill and A. J. Roback, "Tweet acts: How constituents lobby congress via twitter," in Proc. ACM Conf.Comput. Supported Cooperative Work Social Comput., 2014, pp. 1200–1210.

[13] A. Ceron, L. Curini, and S. M. Iacus, "Using sentiment analysis to monitor electoral campaigns: Method matters- evidence from the United States and italy," Social Sci. Comput. Rev., vol. 33, no. 1 pp. 3–20, 2015.

[14] P. Barbera, "Birds of the same feather tweet together: Bayesian ideal point estimation using twitter Data," Political Anal., vol. 23, no. 1, pp. 76–91, 2015.

[15] L. Jiang, M. Yu, M. Zhou, X. Liu, and T. Zhao, "Target- dependent twitter sentiment classification," in Proc. Annu. Meeting Assoc.Comput. Linguistics: Human Language Technol., 2011, pp. 151–160.

[16] B. J. Jansen, M. Zhang, K. Sobel, and A. Chowdury, "Twitter power: Tweets as an electronic word of mouth," J. Amer. Soc. Inf. Sci.Technol., vol. 60, no. 11, pp. 2169–2188, 2009.

[17] S. Wang, M. J. Paul, and M. Dredze, "Exploring health topics in Chinese social media: An analysis of sina weibo," in Proc. AAAI Workshop World Wide Web Public Health Intell., 2014.

[18] A. Culotta, "Estimating county health statistics with twitter," in Proc. SIGCHI Conf. Human Factors Comput. Syst., 2014, pp. 1335–1344.