Detection of Psychological Disorders using Twitter Data

Sahil Kadam
School of Informatics, Computing and Engineering
Indiana University, Bloomington, IN
saskadam@iu.edu

Uteerna Koul
School of Informatics, Computing and Engineering
Indiana University, Bloomington, IN
ukoul@iu.edu

Mohit Saraf
School of Informatics, Computing and Engineering
Indiana University, Bloomington, IN
sarafm@iu.edu

ABSTRACT
Analyzing human psych has intrigued human beings for a long time. Psychologists have been trying to master the skill of understanding and predicting human behavior. With the advent of computers and world wide web, the interactions among humans have changed drastically and this has also changed the way human behavior is pursued and analyzed. Nowadays social media plays an important role in the day to day life of people. From smallest to the biggest news, everything that happens in a person's life is being updated and shared with friends and family. Social media platforms can now be used for sharing news about an upcoming technology to publicize the new brands etc. Twitter is one of the most actively used social media platform. The presence of undetected psychological disorder in people can lead to serious health issues in the future. We are going to identify the psychological traits of a person based on their tweets. Our research focuses on using the Twitter data to identify the users who could be potentially suffering from a psychological disorder. We use a novel approach for data extraction, classification and focus on analyzing the tweets for identifying Depression, OCD, and Anxiety disorders. Our research can provide the medical professionals a head-start in identifying a potential patient and provide them with a pattern of the psychological disorder in social media users.

Keywords
Hashtags; OCD; Anxiety; Depression; Social media; Twitter; Psychological Disorders; Psychology; Mental Health.

1. INTRODUCTION
Social media has become a medium for communicating with others and for expressing self. Twitter is one such popular micro-blogging website. It allows users to exchange messages using a 140-word tweet. The tweets allow a user to keep track of the happenings around the world.

"A psychological disorder is a disorder of the mind involving thoughts, behaviors, and emotions that cause either self or others significant distress" [1]. Psychological disorders are a serious issue that needs to be dealt with. Approximately 1 in every 5 adults in US experiences mental illness in any given year [2]. According to the Centre for Disease Control and Prevention(CDC) estimated that by the year 2020, depression would be the second largest cause of disabilities in the world [3]. It is predicted that nearly 300 million people suffer from depression. Majority of people who suffer from these psychological disorders are left underdiagnosed. "Despite the devastating consequences, clinicians fail to diagnose up to 50% of depressed patients seen in primary care practices. If they are accurately diagnosed, only 22% receive adequate treatment,1 partly because practitioners fail to recognize depression's red flags. While diagnostic criteria appear to be relatively straightforward, diagnosis is often a multistage process." [4] There are no tests to diagnose the mental illness of a user. There is a need to diagnose these disorders and control them in the people. Conducting a survey of the social media profile of the user and looking up at the way they communicate, the psychologist can form the initial basis of whether to perform other tests on the user for psychological disorder detection or not. Social media sites can act as a platform to identify these disorders.

Consider the following tweet by a user:

I died years ago, but apparently, my body didn't get the message. #bipolar #anxiety #depression

The main aim of our research is to identify such tweets based on various hashtags and then classify the users in the 3 different disorder groups i.e. OCD, Depression, and Anxiety or classify them as normal.

In this study, we used some keywords such as "I", "me", "my", "#ocd", "ocd", "#depression", "depression", "#anxiety", "anxiety" which are related to the psychological disorders to extract the tweets. We identify the users who have been tweeting the most about these disorders and predict whether they are likely to suffer from the disorder currently or in future. The main motivation for this research was the inherent need to figure out the mental state of a person. We are trying to put together a basic model which can be used in the future to predict the disorders in a user and form as the basis of evaluation for the future diagnosis of the problem. The method that we have developed here focuses on the context of the tweet and the time frame during which the tweets were collected. After monitoring the tweets, we gathered that a couple of users were consistently tweeting about some disorders, and hence we kept track of those users for the past 6 months data and classified them based on the disorder they might have now or are likely to get in the future.

2. LITERATURE REVIEW
There has been a considerable amount of work done to identify and understand the connection of depression in individuals in researches conducted by various departments like psychiatry, sociolinguistics, etc. One such research is done by Ronald M. Epstein et al. wherein their main intention is to improve the recognition of psychological disorders and before and during the process of treatment [5]. This research idea prompted us to look for a way to detect if there are any psychological disorders in a user and if they exist then the data obtained can be used by the psychologists in their treatment.

Another research in this area is done by Munmum De Choudhary et al. wherein they used crowdsourcing to compile the set of Twitter users who have been reported to be suffering from clinical depression [6]. This paper depicts the way in which the user behaves pre-depression, during depression and in the post-depression phase of their life. Their findings can be used in the healthcare sector to help people deal with depression and enable them to be more proactive about their mental health. This paper takes those patients into consideration who are already suffering from depression whereas in our research we are predicting the onset of depression in a user.

Another research by Moreno MA et al. conducted a research on the college undergrad students to determine the depression symptoms[7]. In this research, the authors used negative binomial regression analyzer to figure out the relationship between depression disclosure and the demographics. They finally concluded that those students who receive reinforcement from their friends are more likely to come forward and discuss their issues than those who are not provided with any support. This idea formed as the basis for selecting the user to predict the psychological disorder that they might be in the initial stages of having.

A different research paper by Park G et al. focuses on compiling a written language formed from more than 60000 users and then predicting their personality using words [8]. This method can be used as a faster method for personality detection. This research paper also provided us with the idea to create a different corpus of words for checking the psychological disorders in the users as the words used in the tweets may or may not always be in one language. Gathering the words from multiple tweets related to the disorder can help us in easily detecting the disorder levels in the other users under consideration.

There is not a lot of research done for OCD. The paper by Rachelle Pavelko et al. discusses the trivialization and oversimplification of OCD as a mental disorder[9]. It also discusses the effect of the trivialization on users with severe OCD. The study measured user's reactions to different stimuli. This stimulus consisted of 11 features adapted from Angermeyer and Matschinger [10] and Wirth and Bodenhausen [11]. The paper concludes that people with OCD do not trivialize non-clinical tweets. We included the non-trivialization of OCD tweets in our model.

3. DATA

We build a dataset of tweets by using the Twitter Streaming API. The tweets obtained using the streaming API were repetitive. The tweets were extracted based on some hashtags which are considered essential for data fetching. The tweets were selected based on "I", "me", "my", "#depression", "depression", "#anxiety", "anxiety", "#ocd" and "ocd". We have a corpus of 40000 tweets wherein for each of these disorder categories we fetched around 10000 tweets. We also fetched tweets depending on the username of the person, for all those users whose tweets consistently included "depression", "ocd", or "anxiety". According to a research conducted by Stephanie Rude et al. the people who were formerly depressed and current depressed user tend to use the word "I" more than the user who has never been depressed [12].

The raw data was then preprocessed. The preprocessing involved removing the repetitive tweets, removing the "RT" keyword for the retweet and the username of the person whose tweet is retweeted. The data also contained some links to the twitter pages which acted as noise and were removed using the regular expressions. The hashtags "#ocd", "#depression" and "#anxiety" associated with those tweets were also removed from the training data.

We gathered the most commonly used words with respect to the disorder from the corpus of labeled tweets. We selected these words based on our understanding and intuition of the disorders ocd, depression and anxiety. We created a dictionary of words for each of these disorders.

4. FEATURES

Features can be defined as the important attributes that help in understanding the context of data under study. For twitter data, usually the most popular features used are based on the frequency of words being used. The most common method used for representing these features is Tf-Idf frequency measures. We are using this as features in our machine learning model using the TfidfVectorizer () method in Scikit-learn library. We preprocess the data by removing all English stop words, words occurring in more than 80% of the tweets and words occurring in less than three tweets by setting the max_df=0.8 and min_df=3 to take into consideration the most important features.

Apart from this, we consider the sentiment score of each tweet as it can act as an important indicator of the extent of the psychological disorder for a person. We use the Afinn wordlist for scoring the sentences for positivity, negativity or neutrality. We use Label Encoding to convert categorical Afinn score to numerical feature using LabelEncoder() method in Scikit-learn library.

<u>Afinn:</u>

"AFINN is a list of English words rated for valence with an integer between minus five (negative) and plus five (positive). The words have been manually labeled by Finn Årup Nielsen in 2009-2011". [13].

We have considered most recent 3000 tweets for each user over a period of 6 months. We assigned weights to each tweet based on the dictionary of most commonly used words for each disorder.

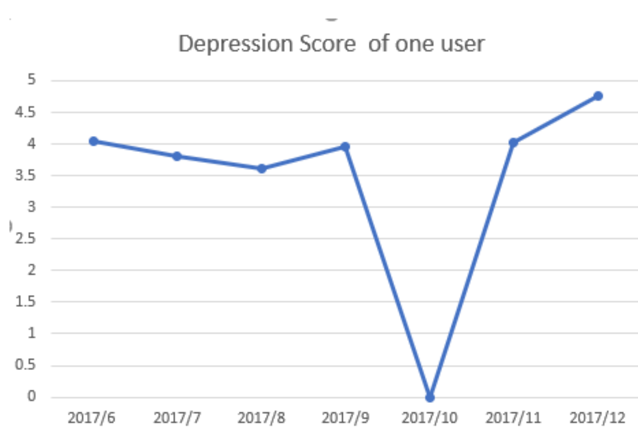

Figure 1. Depression Score of a user

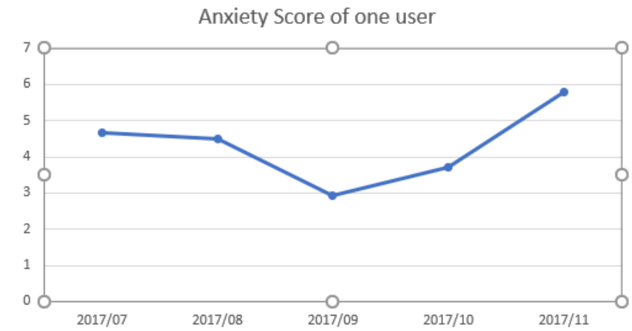

Figure 2. Anxiety Score of a user

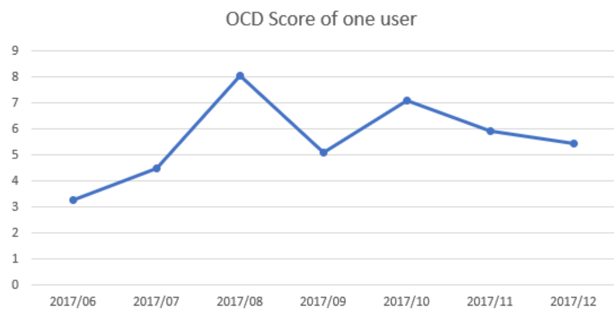

Figure 3. OCD Score of a user

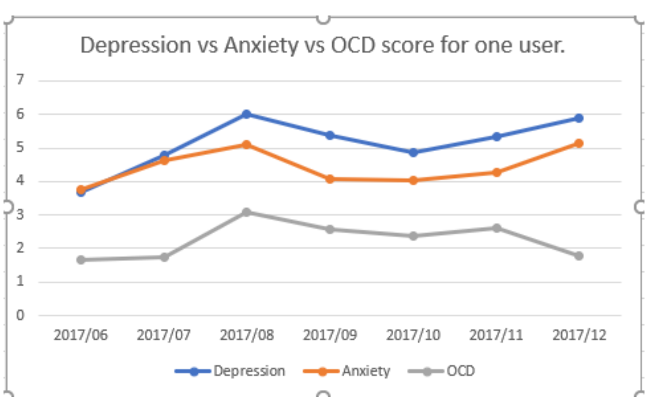

Figure 4. Depression vs Anxiety vs OCD Score of a user

5. MACHINE LEARNING METHOD AND ANALYSIS

We will be testing different classifiers like Support Vector Machines, Logistic Regression and Random Forest Classifier on our data. We are also analyzing twitter data of certain users with 3000 tweets for each user who consistently tweeted about the disorder categories and studied their tweet trends over a period of 6 months related to their disorder using the assigned weights for each tweet based on the dictionary of words related to the disorder.

5.1 Logistic Regression Classifier:

In statistics, Logistic regression can be defined as the regression model where we have categorical data as the dependent variable. In our analysis, we are using multinomial logistic regression as we have multiple categories of output such as depression, ocd, anxiety, and general. One benefit of using Logistic Regression is that it does not assume statistical independence of features like naïve Bayes classification model.

We performed a 5-fold cross-validation on our data using LogisticRegression() method from Scikit-learn library with multiclass parameter as Multinomial and solver as ‘lbfgs’ to handle multinomial loss in our model. We achieved an accuracy of 83.40% for this. Also, when we tested the model on 66-33% train test split we achieved an accuracy of 84.51%. Figure 5. represents the confusion matrix for the Logistic Regression.

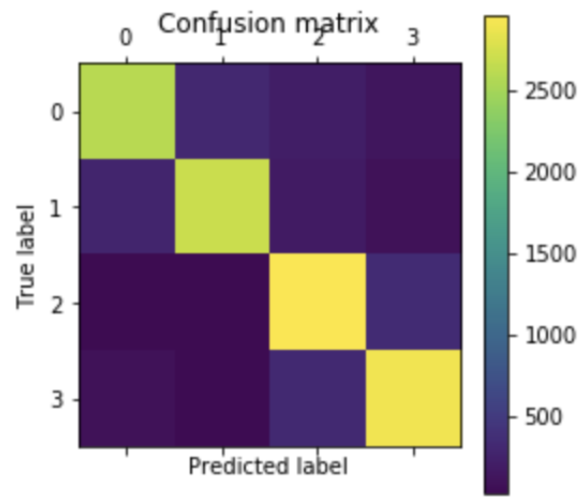

Figure 5. Confusion Matrix for Logistic Regression Classifier

5.2 Support Vector Machine(SVM) Classifier:

Support Vector Machines are sophisticated machine learning models used for classification and regression analysis of data. In our analysis, we are using Support Vector Machine model with a linear kernel. We are using SVC () method from Scikit-learn library to implement our model. We used penalty parameter as 0.9 by checking for different values of C as this was giving the best result for our model. We performed 5-fold cross-validation on our data and achieved an accuracy of 84.10% for our model. Also, when we tested the model on 66-33% train test split we achieved an accuracy of 84.63%. Figure 6. represents the confusion matrix for the support vector machine classifier.

5.3 Random Forest Classifier:

“Random forests or random decision forests are an ensemble learning method for classification, regression, and other tasks, that operate by constructing a multitude of decision trees at training time and outputting the class that is the mode of the classes (classification) or mean prediction (regression) of the individual trees. Random decision forests correct for decision trees' habit of overfitting to their training set.”[14]

We used RandomForestClassifier() method from Scikit-learn library to implement our model and selected 150 as the no of estimators or trees for our model after testing for different values of no of estimators.

We performed 5-fold cross-validation on our data and achieved an accuracy of 84.15% for our model. Also, when we tested the model on 66-33% train test split we achieved an accuracy of 86.08%. Figure 7. presents the confusion matrix for the Random Forest.

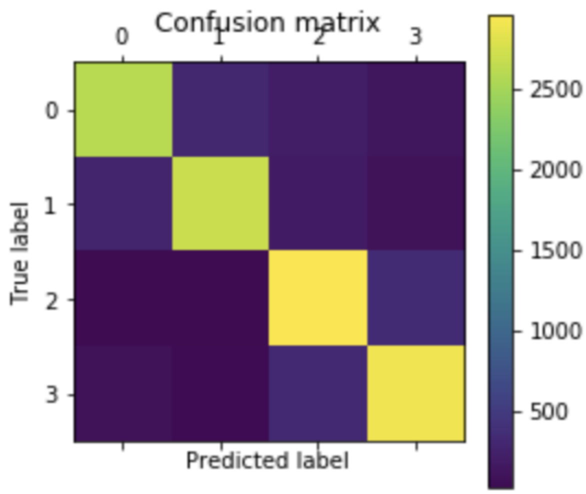

Figure 6. Confusion Matrix for Support Vector Machine Classifier

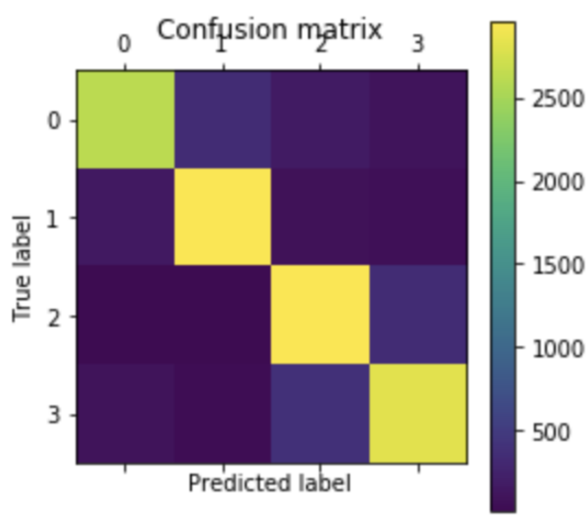

Figure 7. Confusion Matrix for Random Forest Classifier

We are showing the trends for some of the users over here which can help in assessing the extent of distress or disorder the person was going through during this period.

6. RESULTS

From our analysis, we found out that Random Forest was giving the best accuracy amongst the three Machine learning models we used and least accuracy for Logistic Regression. Thus, we can use random forests as the primary model for our analysis.

The following is the classification report for each of the classifier on the given data with precision, recall f1-score and support metrics for each of the categories.

On detailed analysis of this report, we found out that the maximum number of correct classifications were for ocd and anxiety category whereas misclassification was high for depression category.

Also, since we are getting a very high recall value and f-1 score along with precision thus we can say that our results are statistically significant.

	precision	recall	f1-score	support
anxiety	0.87	0.80	0.83	3250
ocd	0.88	0.83	0.85	3258
depression	0.80	0.89	0.84	3335
general	0.84	0.86	0.85	3357
avg / total	0.85	0.85	0.85	13200

Figure 8. Classification Report for Logistic Regression Classifier

	precision	recall	f1-score	support
anxiety	0.86	0.80	0.83	3250
ocd	0.88	0.83	0.85	3258
depression	0.80	0.90	0.85	3335
general	0.85	0.86	0.85	3357
avg / total	0.85	0.85	0.85	13200

Figure 9. Classification Report for Support Vector Machine

	precision	recall	f1-score	support
anxiety	0.91	0.80	0.85	3250
ocd	0.88	0.91	0.89	3258
depression	0.82	0.89	0.85	3335
general	0.84	0.84	0.84	3357
avg / total	0.86	0.86	0.86	13200

Figure 10. Classification report for Random Forest Classifier

We also find out a high degree of correlation between anxiety and depression and we think that can be the reason for the higher number of misclassifications for depression and higher accuracy for anxiety.

	Depression Score	OCD Score	Anxiety Score
Depression Score	1.000000	0.283206	0.418928
OCD Score	0.283206	1.000000	0.230400
Anxiety Score	0.418928	0.230400	1.000000

Figure 11. Correlation Matrix

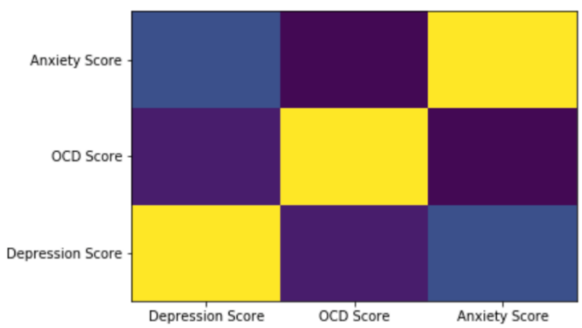

Figure 12. Correlation Heat Map

7. ACKNOWLEDGEMENT

The authors would like to thank Prof. Vincent Malic for providing constant feedback and guidance on the project. We would also like to thank Ms. Ashley Dainas for her constant feedback and support.

8. TASK DISTRIBUTION

Step 1: Data Collection:-
The data collection work was equally distributed between all the team members. Following is the details of data collection process.
1) Sahil Kadam - Depression
2) Uteerna Koul - Anxiety
3) Mohit Saraf - OCD, General

Step 2: Data Cleaning:-
The person that had collected the data was also responsible for cleaning the data.
1) Sahil Kadam - Depression
2) Uteerna Koul - Anxiety
3) Mohit Saraf - OCD, General

Step 3: Features Extraction:-

The features extraction process was completed by Sahil Kadam.

Step 4: Creating Machine Learning Model:-

This task was equally divided into the team members. Each team member created unique machine learning model.

1) Sahil Kadam - SVM.

2) Uteerna Koul - Logistic Regression.

3) Mohit Saraf - Random Forest.

Step 5: Analysis of the result:-

This task was completed by the team together.

Step 6: Literature Review:-

This task was evenly divided among the team members. The team sat together and discusses the papers.

Step 7: Technical Paper Draft:-

This task was carried out by Uteerna Koul.

Step 8: Technical Paper Proof Reading:-

This task was jointly undertaken by Sahil Kadam and Mohit Saraf.

9. CONCLUSION

We conclude that using tweet sentiment along with tweet text can be an effective way to detect psychological disorder for any user.

Machine learning algorithms like SVM, Logistic Regression and Random Forest classifier can achieve higher accuracy for classifying Psychological disorders using this method, with random forest being the best.

We also concluded that there is high correlation between anxiety and depression between the tweets of a user for these categories and thus both can coexist within the same user.

10. FUTURE SCOPE

We can further extend this study by collecting more data on the given disorder categories. Demographic data analysis can be performed along with the given features to get a better sense of the geographical distribution of the users with the disorder category. We can do more linguistic analysis of the text and perform user surveys along with this to get better perspective of the user.A disorder related corpora can be built and used for further research.

11. REFERENCES

[1] Psychological Disorder. Alleydog.com. Retrieved from: https://www.alleydog.com/glossary/definition-cit.php?term= Psychological+Disorder

[2] Any Mental Illness (AMI) Among Adults. Retrieved from https://www.nimh.nih.gov/health/statistics/ prevalence/any-mental-illness-ami-among-us-adults.shtml

[3] Murray CJL, Lopez AD. The Global Burden of Disease: A Comprehensive Assessment of Mortality and Disability from Diseases, Injuries and Risk Factors in 1990 and Projected to 2020. Geneva, Switzerland; World Health Organization. 1996

[4] Guido R. Zanni. 2011. Depression: Underdiagnosed and Undertreated. Retrieved from http://www.pharmacytimes.com/publications/issue/2011/january2011/counseling-0111

[5] Epstein RM, Duberstein PR, Feldman MD, Rochlen AB, Bell RA, Kravitz RL, et al (2010) “I didn’t know what was wrong:” How people with undiagnosed depression recognize, name and explain their distress. J Gen Intern Med 25(9): 954–961. doi.org/10.1007/s1160601013670

[6] Munmun De Choudhury, Michael Gamon, Scott Counts, and Eric Horvitz. 2013. Predicting depression via social media. In Proceedings of the International AAAI Conference on Weblogs and Social Media (ICWSM)

[7] Moreno, M.; Jelenchick, L.; Egan, K.; Cox, E. et al. 2011. Feeling bad on Facebook: depression disclosures by college students on a social networking site. Depression and Anxiety 28(6):447–455

[8] Park G, Schwartz HA, Eichstaedt JC, Kern ML, Kosinski M, Stillwell, DJ et al. (2015) Automatic personality assessment through social media language. J Pers Soc Psychol 108 (6): 934–952. http://doi.org/10.1037/pspp0000020

[9] Rachelle Pavelko. 2016. Tweeting and Trivializing: How the Trivialization of Obsessive–Compulsive Disorder via Social Media Impacts User Perceptions, Emotions, and Behaviors. magination, Cognition and Personality: Consciousness in Theory, Research, and Clinical Practice 2016, Vol. 36(1) 41–63

[10] Angermeyer, M. C., & Matschinger, H. (1996). The effect of personal experience with mental illness on the attitude towards individuals suffering from mental disorders. Social Psychiatry and Psychiatric Epidemiology

[11] Wirth, J. H., & Bodenhausen, G. V. (2009). The role of gender in mental-illness stigma: A national experiment. Psychological Science, 20 (2), 169–173

[12] Rude, S., Gortner, E.-M., & Pennebaker, J. 2004. Language use of depressed and depression-vulnerable college students. Cognition and Emotion, 18(8), 1121–1133. doi: https://doi.org/10.1080/10810730.2017.1367338

[13] Finn Årup Nielsen. 2011. Afinn. Retrieved from http://www2.imm.dtu.dk/pubdb/views/publication_details.php?id=6010

[14] Random Forest. 2017. WikipediA: The Free Encyclopedia. Retrieved from https://en.wikipedia.org/wiki/Random_forest