Mental Health (Potential SPROJ Topic)

**Research Question**

**Can topic modelling and sentiment analysis be used to effectively differentiate tweets/their users with a very common mental health disorders - Depression**

Depression is a mental illness that is not taken very seriously especially in developing countries such as Pakistan. The average Pakistani citizen never gets exposure to mental illness education at school or at home. Getting help for mental illness is considered Taboo in Pakistan and the people who are diagnosed with depression do not take medication to overcome it. They just continue with their lives without the proper treatment and potentially keep getting worse. With Social media platforms becoming an integral part of a person’s life, it is safe to assume that they reflect the user’s personal life. These platforms can be used by researchers to identify the cause of depression and detect it. Detecting depression earlier can be a huge step to address the mental issue and offer support to the people suffering from this terrible illness in Pakistan. The social platform Twitter was chosen because of two reasons. One being that by using a Twitter developer account, tweets can easily be scrapped and used in the model and two, twitter is a platform where people mainly speak out about what they are feeling or thinking about, something that Facebook and Instagram do not completely achieve. Therefore this paper will be used to examine whether topic modeling and sentiment analysis of Tweets can be used to differentiate people who are clinically depressed versus people who are not

**Literature Review**

**Article 1:**

In this paper Horvitz and a team of researchers helped develop a model that can scan tweets and predict depression in Twitter users, with an accuracy they claim to be 70%. When the model scans your tweets, it misses some signals and doesn’t diagnose many people about 30% who really will get depression. And the system has a “false positive” issue, Horvitz said, causing it to incorrectly predict that healthy Twitter users will get depression in about 10% of cases.

The group of researchers however noticed that while there were some tweets which had clear signs of mental distress, others were not that clear. Words like anxiety, severe, suicidal and attacks were commonly used by depressive people however words like amazing, love, fun, favorite, movie could also be indications of depression. The volume of tweets mattered too, as did the percentage of exchanges—users who are depressed begin to tweet less, and tweet less at other people, indicating a possible loss of social connectedness, said Horvitz.

A study by University of California San Diego builds upon this research in which Michael Conway is creating models that will track depression in communities.

<https://time.com/1915/how-twitter-knows-when-youre-depressed/>

The Microsoft team found 476 Twitter users, 171 of whom were seriously depressed. They went back into users’ Twitter histories as far as a year in advance of their depression diagnosis, examining their tweets for language, level of engagement, mentions of certain medications, and other factors, using computer models to sift through a total of 2.2 million tweets. By comparing depressed Twitter users’ feeds with the non-depressed user sample class, they came up with a method for predicting depression diagnoses before they happened. When they tested the model on a different set of Twitter users, it showed 70% accuracy in predicting depression before its onset.

Microsoft’s researchers looked at factors like the number of tweets users made per day, what time of day users tweeted, how often users interacted with each other, and what kind of language tweeters were using. For example, seemingly depressed tweeters were more likely to post messages late at night (between 9pm and 6am) compared with healthy tweeters, who were most active during the day and after work hours.

**Article 2:**

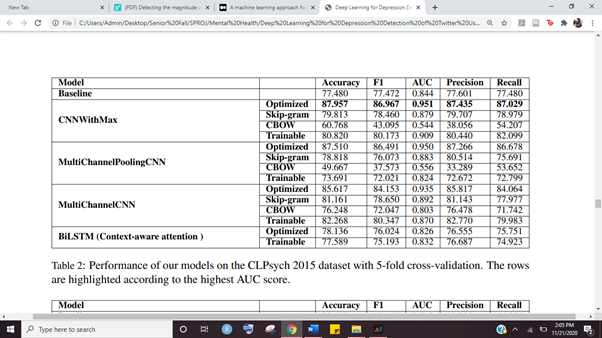
<https://medium.com/datadriveninvestor/a-machine-learning-approach-for-detection-of-depression-and-mental-illness-in-twitter-3f3a32a4df60>

Sentiment analysis through a powerful theorem from probability theory called Baye’s Theorem. The model will be written in python and it will tell whether a given tweet is depressive or not. Dataset used: https://www.kaggle.com/kazanova/sentiment140

The code: https://github.com/viritaromero/Detecting-Depression-in-Tweets

**Article 3:**

<https://www.aclweb.org/anthology/W18-0609/>



**Article 4:**

<https://www.researchgate.net/publication/336887898_Detecting_the_magnitude_of_depression_in_Twitter_users_using_sentiment_analysis/link/5db8e860a6fdcc2128eba2b7/download>

Python has been used to parse json data extracted from twitter, merge selected data and provide a formatted data as a csv file. R has been used for base emotion analysis, sentiment calculation and depression score calculation. The data has been visualized using ggplot in R which provided a better understanding of the final data

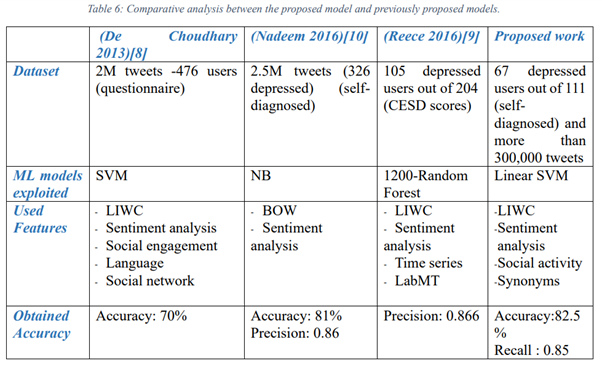
“Does there exist a method to evaluate the magnitude of depression in a person based on the emotional integrity of their tweets”.

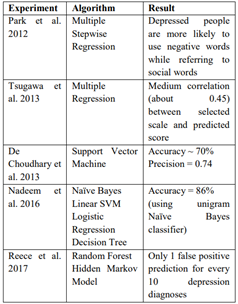
**Article 5:**

Social Media Analytics, Depression Detection, Machine Learning (ML), Support Vector Machine (SVM), Naive Bayes, Decision Tree, Feature Selection

<https://arxiv.org/ftp/arxiv/papers/2003/2003.04763.pdf>





**Article 6:** 

This basically explains the different techniques that have been used to detect depression using tweets.

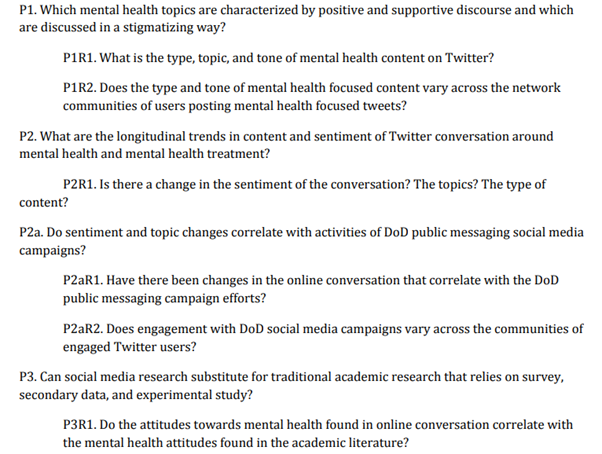
**Article 7:**

<https://www.aclweb.org/anthology/W14-3207/>

• We demonstrate the effectiveness of our automatically derived data by showing that statistical classifiers can differentiate users with four different mental health disorders: depression, bipolar, post traumatic stress disorder and seasonal affective disorder.

• We conduct a LIWC analysis of each disorder to measure deviations in each illness group from a control group, replicating previous findings for depression and providing new findings for bipolar, PTSD and SAD.

• We conduct an open-vocabulary analysis that captures language use relevant to mental health beyond what is captured with LIWC.



One area of public health where this kind of research could come in handy is in measuring public reactions to events. Tracking public Twitter feeds after profound or traumatic events could help scientists understand how we’re affected by the news. “We really didn’t used to have many tools available traditionally for that kind of fine-grained analysis,” says said Horvitz. “Now there’s a new direction for doing the science.”

School shootings, church/mosque attacks, protests, bombings

Throughout the literature, it could be identified that the most widely adopted feature engineering method is to extract lexical features using the Linguistic inquiry word count (LIWC) lexicon, which contains more than 32 categories of psychological constructs (Pennebaker et al., 2007). The lexicons have been used as one of the key feature extraction mechanisms in identifying insomnia (Jamison-Powell et al., 2012), distress (Lehrman et al., 2012), postpartum depression (De Choudhury et al., 2013), depression (Schwartz et al., 2014) and post-traumatic stress disorder (PTSD) (Coppersmith et al., 2014a). For each of these mental disorders to be identified, researchers had to extract features that overlap with each other, and are unique to a particular disorder.

character n-grams can be identified as the key feature extraction mechanism in detecting mental illnesses such as attention deficit hyperactivity disorder (ADHD), generalized anxiety disorder, and eight other mental illnesses

**Topic DISCARDED:**

* **Tagging the dataset would be extremely difficult**
* **Figuring out if a person has clinical depression is a data point that we would not have**
* **Training the model using these tweets will be an extremely biased method.**
* **Understanding depression in a Pakistani context is something that needs further research before Twitter data can be used to predict depression.**
* **Subjective thing and it would be weird to have diagnosis based on one tweet or a few tweets.**