A REVIEW ON DEPRESSION DETECTION USING TWITTER DATASET

Charul, Mrs. Shruti Sharma, Dr. Parul Gupta
M.Tech. Scholar, Mentor, Co-Mentor
College: YMCA University of Science and Technology, Faridabad, Haryana (INDIA)

Abstract - Social media has recently emerged as a premier method to disseminate information online. Through these online networks, tens of millions of individuals communicate their thoughts, personal experiences, and social ideals. We therefore explore the potential of social media to predict, even prior to onset, Major Depressive Disorder (MDD) in online personas. We employ a crowdsourced method to compile a list of Twitter users who profess to being diagnosed with depression. Using up to a year of prior social media postings, we utilize a Bag of Words approach to quantify each tweet [1]. Lastly, we leverage several statistical classifiers to provide estimates to the risk of depression. Our work posits a new methodology for constructing our classifier by treating social as a text-classification problem, rather than a behavioral one on social media platforms. By using a corpus of 2.5M tweets, we achieved an 81% accuracy rate in classification, with a precision score of .86. We believe that this method may be helpful in developing tools that estimate the risk of an individual being depressed, can be employed by physicians, concerned individuals, and healthcare agencies to aid in diagnosis, even possibly enabling those suffering from depression to be more proactive about recovering from their mental health.

Index Terms—*Depression, Machine Learning, Social Media, Twitter*

I. INTRODUCTION

Depression is the mood disorder characterized by low mood, feeling sadness and lost of interests in things that a person loved to do in the past. According to the psyclogy when the person likes nothing to do and feel hopeless with no urge to live. In the depression the person forgets the difference between what is right and wrong for him and losses interest in friends and family also. Depressed person can have suicidal as well as criminal tendency that can be reason from having frustration , stress and lack of calmness for a long period of time. People with depression run away with deadlines and responsibilities and cut off with the people around them. They try to avoid things that takes lot of attention and effort. They can not pretend to have a simple nature in front of others because of their mood swings.

Social media is easiest computer technology to interact with people and share informations , ideas and interests to each other. It is a large platform to collect information related to human behaviours. Its is not only useful for people to connect each but also helpful for the psychiatrists to study about the human behaviours. There study helps to reach to the people having negativity ,stress, anxiety ,sleep disorder, body weight changes and suicidal thoughts. That helps to help these patients wether they are diagnosed or not. Social media is very popular in 20th century and almost everybody use it because of it gives valuable results about the mentally ill patients. In these days social media interest among people is increasing on large scale that is why it is very easy to collect the information of the people of their day to day activities, their emotions and feelings etc.

Our study builds upon prior mentioned work and contributes towards enhancing lexical methods for text classification. With our present work we: (1) further explore the capability for individual social media status updates to be utilized as a feature in determining or furthering a diagnosis of depression or not; (2) examine, compare, and analyze the effectiveness of several supervised statistical models to predict text classification; and (3) demonstrate that we may use these features to further the identification of depressive disorders in a cohort of individuals who may otherwise have slipped under the radar.

AIMS

This study aims to establish the feasibility of consistently detecting, identifying, and pursing the diagnosis of individuals Twitter posts, henceforth referred to as ‘tweets’, Using solely these tweets, we aim to design and implement an automated computational classifier which may be able to parallel the performance and precision of a concerned human individual. The feasibility of this automated predictions will be cross-validated and critiqued through standard Precision, Recall, and F1 scores, as well as Recipient Operating Classification curves.

II. LITERATURE SURVEY

The research that has been already done to predict the depression uses the common criteria of patients to check their depression using social media. Today there are different platform of virtual networking technology like twitter, facebook, instagram are the most famous social networking sites. All the study says that the prediction should be done on network that is popular such that maximum behovior attributes can be collected. Several studies have been done to collect and information from the

social media along with the information about metal health of patients as well.

Different papers tells the different classification of mentally ill patients. Survey justifies following information about the prediction of depression :

S.No	Title	Author(s)	Journal/ Organisation	Publication Year	Problems Identified
1.	Predicting Depression Levels Using Social Media Posts	Maryam Mohammed Aldarwish	IEEE 13th International Symposium On Autonomous Decentralised Systems	2017	In this paper according to World Health Organization (WHO) , around 350 million people are affected by depression . The medical science relies on asking the patients questions about their situations, which does not diagnose the depression in a precise way . The patient has to attend more than one session during a period of two weeks. The classification of a not depressed condition as a depressed is a False Positive problem. However researcher 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 behavioral health conditions such as depression. The reasons behind getting lower accuracy and recall are finding depressed individuals which are active in Twitter or Facebook in the same time, is challenging task .Also all the depressed patients from Saudi Arabia where all their tweets and SNS posts in Arabic using slang Arabic words which give different meanings.
2.	Predicting Depression Level Using Social Media Posts	Namrata Sonawane, Mayuri Padmane, Vishwja Suralkar, Snehal Wable, Prakash Date	International Journal of Innovative Research in Science, Engineering and Technology	May 5, 2108	In this paper psychological concept has been taken It is non-trivial to detect stress timely for proactive care. They presented a framework for detecting users psychological stress states from users monthly social media data, leveraging posts content as well as users social interactions. Employing the real world to social media data as the basis, the goal was to develop a web application which takes social media posts and questionnaire test as a input and predict output as various depression levels. Using Naive Bayes classifier algorithm to increase accuracy of system. To deliver appropriate doctors information depending upon location of user. According to user posts system can find out user in stressed or not as well as different quaternaries which is provided by the system.
3.	Predicting Depression From Language-Based Emotion Dynamics: Longitudin-al Analysis of Facebook and Twitter Status Updates	Elizabeth M Seabrook.	Journal Of Medical Internet Research	May 20 2018	In this paper social media is used in different ways by different people, but for many individuals, status updates provide snapshots of their lived experience. Studies to date have primarily considered how emotional expression over time is explored as another window of insight into the psychological health of social media users. Depression, including major depressive disorder (MDD) and dysphoria, are characterized by persistent low mood (including sadness or emptiness) or anhedonia (inability to experience pleasure from activities that are usually enjoyable). At a broad level, the frequent expression of negative affect within social media status updates has been associated with higher levels of depression symptoms. Frequently expressing positive affect, on the other hand, tends to be associated with lower levels of depression and greater levels of well-being . specific topics, keywords, and linguistic features (especially negative emotions) are able to identify depression-indicative posts with high accuracy, many of these features may also be present in posts that are non indicative of depression. For example, Mowery et al found considerable signal discrepancies over 70% of tweets identified in their sample containing words related to depression were not actually indicative of depression. The study suggests that instability in the negative emotion expressed on Facebook provides insight into the presence of depression symptoms for social media users, and greater variability of negative emotion expression on Twitter may be protective for mental health. The findings also point to possible differences across the online culture created by a particular social media platform, such that different platforms may provide different insights into mental health.

4.	Detecting depression and mental illness on social media: an integrative review	Sharath Chandra Guntuku, David B Yaden, Margaret L Kern, Lyle H Ungar	Science Direct	July 05 2017	This paper says that the widespread use of social media may provide opportunities to help reduce undiagnosed mental illness. A growing number of studies examine mental health within social media contexts, linking social media use and behavioral patterns with stress, anxiety, depression, suicidality, and other mental illnesses. The greatest number of studies of this kind focus on depression. Depression continues to be under-diagnosed, with roughly half the cases detected by primary care physicians and only 13–49% receiving minimally adequate treatment. Automated analysis of social media potentially provides methods for early detection. If an automated process could detect elevated depression scores in a user, that individual could be targeted for a more thorough assessment, and provided further resources, support, and treatment. Studies to date have either examined how the use of social media sites correlates with mental illness in users or attempted to detect mental illness through analysis of the content created by users. For mental health perspective, clear guidelines on mandated reporting are needed. There are open questions around the impact of misclassifications, and how derived mental health indicators can be responsible.

III. CONCLUSION

The aim of our survey was to study the work on the prediction of depression that has been done in various ways. Using questioning techniques , assessment of data that has been collected by user generated content ,checking depression by their levels and providing sessions for the depressed or people who found to be affected by depression.

IV. REFERENCES

[1]. US Burden of Disease Collaborators, "The state of US health, 1990-2010: burden of diseases, injuries, and risk factors.," JAMA, vol. 310, no. 6, pp. 591–608, 2013.

[2]. US Census Bureau, "Census 2010," US Census Bureau, 2010. [Online]. Available: http://quickfacts.census.gov/qfd/states/13/13135.html.

[3]. L. A. Pratt and D. J. Brody, "Depression in the U.S. Household Population, 2009-2012.," NCHS Data Brief, no. 172, pp. 1–8, 2014.

[4]. L. Andrade, J. J. Caraveo-Anduaga,P. Berglund, R. V. Bijl, R. De Graaf, W. Vollebergh, E. Dragomirecka, R. Kohn, M. Keller, R. C. Kessler, N. Kawakami, C. Kili??, D. Offord, T. B. Ustun, and H. U. Wittchen, "The epidemiology of major depressive episodes: Results from the International Consortium of Psychiatric Epidemiology (ICPE) Surveys," Int. J. Methods Psychiatr. Res., vol. 12, no. 1, pp. 3–21, 2003.

[5]. M. Hewitt and J. H. Rowland, "Mental health service use among adult cancer survivors: Analyses of the National Health Interview Survey," J. Clin. Oncol., vol. 20, no. 23, pp. 4581–4590, 2002.

[6]. D. Michelson, C. Stratakis, L. Hill, J. Reynolds, E. Galliven, G. Chrousos, and P. Gold, "Bone Mineral Density in Women with Depression," N. Engl. J. Med., vol. 335, no. 16, pp. 1176–1181, 1996.

[7]. S. E. Chance, S. L. Reviere, J. H. Rogers, M. E. James, S. Jessee, L. Rojas, C. A. Hatcher, and N. J. Kaslow, "An empirical study of the psychodynamics of suicide: A preliminary report," Depression, vol. 4, no. 2, pp. 89–91, 1996.

[8]. R. S. Levine, I. Goldzweig, B. Kilbourne, and P. Juarez, "Firearms, youth homicide, and public health.," J. Health Care Poor Underserved, vol. 23, no. 1, pp. 7–19, 2012.

[9]. R. C. W. Hall, D. E. Platt, and R. C. W. Hall, "Suicide risk assessment: A review of risk factors for suicide in 100 patients who made severe suicide attempts: Evaluation of suicide risk in a time of managed care," Psychosom. J. Consult. Liaison Psychiatry, vol. 40, no. 1, pp. 18–27, 1999.