

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 7, Issue 5, May 2018

Predicting Depression Level Using Social Media Posts

Namrata Sonawane¹, Mayuri Padmane¹, Vishwaja Suralkar¹, Snehal Wable¹, Prakash Date²

B. E Students Department of Computer Engineering, Cummins College of Engineering, Karvenagar Pune, India¹

Professor, Department of Computer Engineering, Cummins College of Engineering, Karvenagar Pune, India²

ABSTRACT: Depression is a major issue which is increasing day by day. Worldwide mental illness is primary cause of depression. Millions of people each year suffered from depression and only fraction of them undergoes adequate treatment. It is serious medical illness which corresponds to user's ability to work, study, participation in social activities and having fun. Depression is serious challenge in personal and public health. One of the major solutions to this problem is detailed study of individual's behaviour attributes. These attributes are available on various social networking sites such as Facebook ,twitter,instagrametc.Social networking platform is best way to know person behaviour,thinking style,mood,egoistic networks ,opinions etc. The use of social networking sites is increasing especially by young generation.The people on social media express their feelings, daily activities, opinions about various topics etc.So,social networking sites are used as screening tool to predict depression levels. These social networking platforms givesperson'sexperiences,opinions,socialization,personality.People feelings could be positive, negative or neutral. In order to determine depression levels, person's negative response is important because it tells us about the negativism. Earlier method of diagnosis of patient is not so relevant but by using user generated content on social media such as Facebook definitely helps to predict mental health levels and depression of particular individual. Our project aim is to extract information from social media posts and by having clear understanding of person's behavioural attributes and attempted questionnaires, depression levels of user is predicted. By using machine learning algorithms such as Naive Bayes, depression level of user is classified into different levels and provides doctor's location near to user location.

KEYWORDS: Naive Bayes Classifier, Sentiment Analysis, SNS (Social Networking Sites), Text Mining, UGC (User generated Content)

I. INTRODUCTION

Web application is a part of this project. In this project the social media posts (Facebook) are considered. Social Network Sites (SNS) is a source of data and screening tool to classify the users according to user generated contents (UGC). By using machine learning algorithms such as Naive Bayes, depression level of user is classified into different levels and provides doctor's location near to user location.

As depression is very serious problem which is increasing day by day, many people are suffering from this problem. In India, out of total population 7.5% of it facing this problem. It seems to be major issue and that is the reason it motivate us to worked on it .Earlier diagnosis of depressed patient were done on basis of questionnaires and its behaviour reported by his relatives or friends. But the result was not so qualitative and accurate. In contrast with that, social media is powerful tool for predicting depression levels of an individual.

Nowadays, many people are using social media platforms such as Facebook, twitter. They share their thoughts, feelings, emotions, feelings of guilt, worthlessness, helplessness and egoistic nature of individual etc. Whatever they post is related to their daily activities & happenings. Social media helps to know about individual's thinking, mood, activities & socialization. So by analysing the social media data & applying some algorithms on it, we can able to determine the depression levels of that particular individual. So it will help to diagnose that person before he/she gets more affected to it. This motivates us to do the project, so that this will help not only to psychiatrist but also used by

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 7, Issue 5, May 2018

patients who want to do self-diagnosis.

II. OBJECTIVE AND SCOPE

1. User data is extracted from face book using Facebook developer site.
2. Our aim is to investigate how SNS user's posts can help to classify user according to mental health level.
3. GPS will be used for user's location.

III. PROBLEM STATEMENT

Web application is a part of this project. In this project the social media posts (Facebook) are considered. Social Network Sites (SNS) is a source of data and screening tool to classify the users according to user generated contents (UGC). By using machine learning algorithms such as Naive Bayes, depression level of user is classified into different levels and provides doctor's location near to user location.

IV. LITERATURE SURVEY

For predicting depression levels we need to have complete understanding of user behaviours through social media posts. Many researches have been done on this depression level problem in order to solve it. There is need to find a relationship between positive & negative posts along with symptoms and level of depression.

Twitter as a screening tool in [1] to predict the Major Depressive Disorder (MDD) level. It is unipolar disorder which is characterized by persistent feeling of sadness. It is implemented by using CES-D (Centre for Epidemiologic Studies Depression Scale) screening test in order to minimize the major depressive disorder. They collected twitter data using crowdsourcing technique. The results shows that depressed individuals have few social activities, negative feelings and intense expression of religious idea. They build a prototype which predicts depression of individuals using Support Vector Machine(SVM) classifier.

They used method of crowdsourcing to extract posts of twitter in [2]. They developed a probabilistic model which takes trained data as a input and analyze whether particular post indicates depression or not. This model is Support Vector Machine(SVM). This model considers various parameters like social activity, emotion ,egocentric social networks & language used in twitter posts. Depression was measured using the CES-D (Center for Epidemiologic Studies Depression Scale) screening test .This model provides an accuracy more than 70% with precision of 0.82 to determine whether post is depression indicative or not.

They proposed a statistical metric name as Social Media Depression Index (SMDI). It is used to predict depressive indicative postings on twitter & also helps to categorized depression levels. They justify the idea of using social networking sites (SNS) to detect depression in [3]. They collected Facebook posts to determine the depression levels of user.

They used SNS as a tool to create records for each user to use when diagnosing depression. Using this tool they got complete behaviours image of the user. They proposed a tool that predict and diagnose the depression of Facebook users regarding their activities in a limited period of time.

They present the sentiment analysis for language learning using Naive Bayes Classifier in [4]. Proposed system was used for Facebook social media platform. They mainly focus on user's Facebook status updates. It does not include any face book posts like photo stories, application stories etc. They used Naive Bayes algorithm for classifying the status into positive or negative status. This system also extract emoji's which helps to understand person's emotions. Accuracy of system was so high due to used of Naive Bayes algorithm.

They extracted twitter data by using unigram, bigram feature extractors in [5]. Extracted data from twitter is analysed by various algorithms such as Naive Bayes, support vector machine, Maximum entropy etc. Data that is extracted is consist of textual data, emotions etc. On these data, reprocessing and then analysis is carried out. They classify it using Naive Bayes algorithm which classifies it into various depression levels.

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 7, Issue 5, May 2018

V.EXISTING SYSTEM APPROACH

Existing works demonstrated that leverage social media for healthcare, and in particular stress detection, is feasible. There are some limitations exist in facebook content based stress detection. Users do not always express their stressful states directly in facebook post. . Although no stress is revealed from the post itself, from the follow-up interactive comments made by the user and her friends, we can find that the user is actually stressed from work. Thus, simply relying on a user’s Facebook post content for stress detection is insufficient. Users with high psychological stress may exhibit low activeness on social networks. Stress detection performance is low.

VI.PROPOSED SYSTEM APPROACH

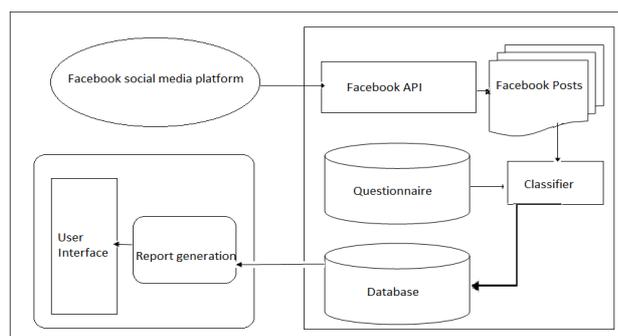


Fig.1 Block Diagram of Proposed System

System architecture is used to design and develop an web application, which provide an easy and convenient way to get information about depression levels of user by using machine learning algorithms and according to the location of user, the information about doctor is provided. The extraction class will performed the extraction of textual data from facebook through facebook graph API.Preprocessing class is used to preprocess the extracted data.Data must be clear,right and it is preprocessed for taking care of missing or repetitive attributes.The data ought to be complete and reliable data to deliver the best result from the data mining methodology. Preprocessing of data takes place by using techniques such as tokenization, lower case conversion, word stemming and words removal.Term frequency (tf) has been computed to measure term occurrence. In proposed system user is on Facebook, according to his Facebook post system can find out user in stressed or not as well as different quaternaries which is provided by the system. If user’s are not on Facebook they can attempt only quaternaries which is provided by the system according to that we can find out user’s in stressed or not. system.

VII. SYSTEM ANALYSIS

In our experimental setup, we are identified that in our system number stressed user and number of non-stressed.

Sr.No	No. Stressed User	No. Non-Stressed User
1	30	50

Table1: Stressed and Non-Stressed User

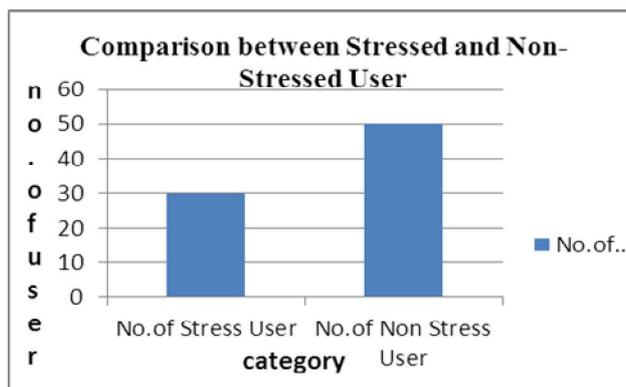
From above data, we can see the stressed and non-stressed user in the graph; we see 30 users are in stressed and 50 users in the non-stressed user.

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 7, Issue 5, May 2018



Graph 1: Stressed and Non-Stressed User

VII. CONCLUSION

Psychological stress is threatening people's health. It is non-trivial to detect stress timely for proactive care. Therefore we presented a framework for detecting users' psychological stress 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 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 his Facebook post system can find out user in stressed or not as well as different quaternaries which is provided by the system.

ACKNOWLEDGMENT

This work is supported in a depression level system using social media post of any state in india. Authors are thankful to Faculty of Engineering and Technology (FET), Savitribai Phule Pune University, Pune for providing the facility to carry out the research work.

REFERENCES

- [1] Munmun De Choudhury, Michael Gamon, Scott Counts, Eric Horvitz Predicting Depression via Social Media 2013
- [2] Christos Troussas, Maria Virvou, Kurt Junshean Espinosa, Kevin Llaguno, Jaime Caro Sentiment analysis of Facebook statuses using Naive Bayes classifier for language learning. 2013
- [3] Alec Go, Richa Bhayani, Lei Huang Twitter sentiment classification using distant supervision 2009
- [4] Altug Akay, Andrei Dragomir and Bjorn-Erik Erlandsson Mining Social Media Big Data for Health 2015
- [5] Megan A Moreno , Henry Young, Tara Becker Feeling bad on Facebook: Depression disclosures by college students on a social networking site 2011
- [6] Munmun De Choudhury , Scott Counts, Eric Horvitz Social Media as a Measurement Tool of Depression in Populations 2013