



# PREDICTING DEPRESSION LEVEL USING SOCIAL MEDIA POSTS



Hrishabh Patidar <sup>1</sup>✉ , Jayesh Umre <sup>2</sup>✉

<sup>1</sup> Student, Department of Computer Science and Engineering, Acropolis Technical Campus, Indore

<sup>2</sup> Professor, Department of Computer Science and Engineering, Acropolis Technical Campus, Indore



DOI: <https://doi.org/10.29121/granthaalayah.v8.i12.2020.1972>

**Article Type:** Research Article

**Article Citation:** Hrishabh Patidar, and Jayesh Umre. (2020). PREDICTING DEPRESSION LEVEL USING SOCIAL MEDIA POSTS. International Journal of Research - GRANTHAALAYAH, 8(12), 234-237. <https://doi.org/10.29121/granthaalayah.v8.i12.2020.1972>

**Received Date:** 11 October 2020

**Accepted Date:** 31 December 2020

**Keywords:**

Naive Bayes Algorithm  
Sentiment Analysis  
Text Mining

## ABSTRACT

Depression is a major concern snowballing day by day. There can be various causes of depression but mental illness is the main problem. A lot of people suffer from depression and a very few of them go through treatment. One out of six people between ages 10 to 19 years are suffering from depression. At its worst, depression can lead to suicide. Depression reduces user's ability to do work study or socialize. One solution to this problem is study of individual's behaviour through social media. We could know a person's opinion, thinking, mood etc. through his social media. These attributes of user can be collected from different social networking sites like Instagram, Facebook, and Twitter etc. Social networking sites can be used as an analysis tool to predict depression level. Our projects aim is to gather information of user from their social media posts and predict his depression level.

## 1. INTRODUCTION

Social media plays a vital part in this project, posts from different social networking sites are considered. Social networking site is the source of data and according to the data produced by the user they are classified into different groups. Depression is very serious problem around the world. According to World Health Organization (WHO) more than 300 million people of all ages suffer from depression. Depression and bad mental health have been ignored issue from ages. When it comes to countries, according to WHO India is the most depressed country in the world then comes China and USA. These are few of the most affected countries by anxiety and mental disorder.

Earlier diagnosis of depressed patients was done by questioning and his behaviour report by his friends and relatives. But the results were not accurate and qualitative. Whereas social media can be used to produce more qualitative and accurate results. Nowadays people use different social media such as Facebook, Instagram, and Twitter. They share their thoughts, emotion and inner feeling. We can know about the daily activity, mood and opinion of the individual. By analysing user's activity (data) and applying machine learning algorithms on it, we can predict the depression level of that individual.

## 2. PROBLEM STATEMENT

Using machine learning algorithm like Naive Bayes on user generated content, user depression level is classified into different levels.

## 3. BACKGROUND SURVEY

We should be able to understand the complete behaviour of user by his social media post to predict his depression level. A lot of research has been done on this problem. To solve this problem, we need to distinguish between positive and negative post and know the other symptoms of depression.

Twitter as a screening tool, for predicting Major Depressive Disorder Level (MDDL). It is a disorder that is found commonly in unhappy people. It is enforced by using CES-D (Centre for Epidemiologic Studies Depression Scale) screening test in order to diminish the major depressive disorder. Twitters user data is collected by them using the crowdsourcing technique. The results show that depressed people have very few social activities and the usually spend time by themselves. They also built a prototype model which estimates the depression level of an individual using Support Vector Machine (SVM) classifier. By crowdsourcing they extracted users twitter posts. They made a predictive model which takes training data as input and analyse the post whether it is depressing or no on various factors. The model is trained by Support Vector Machine (SVM). There is various parameter included in model like daily social activity, social networking & emotion of frequent twitter posts. Depression was measured using Centre for Epidemiologic Studies Depression Scale (CESDS) screening test. That model had an accuracy of around 70% with a precision of 0.82 to determine the depression of the post. When only status was used to predict depression, Naive Bayes gave much higher accuracy as compares to Support Vector Machine (SVM) and it also took emojis under consideration to predict user's emotion.

## 4. NAIVE BAYES CLASSIFIER

In machine learning, Naive Bayes is based on bayes theorem with independent assumption between features.it is not a single theorem but a family of theorem working on same principle.

The main assumption of naive bayes algorithm is each feature makes independent and equal contribution to the outcome. The probability of event occurring given the probability of another event already occurred is called bayes theorem; it can be mathematically represented as following equation

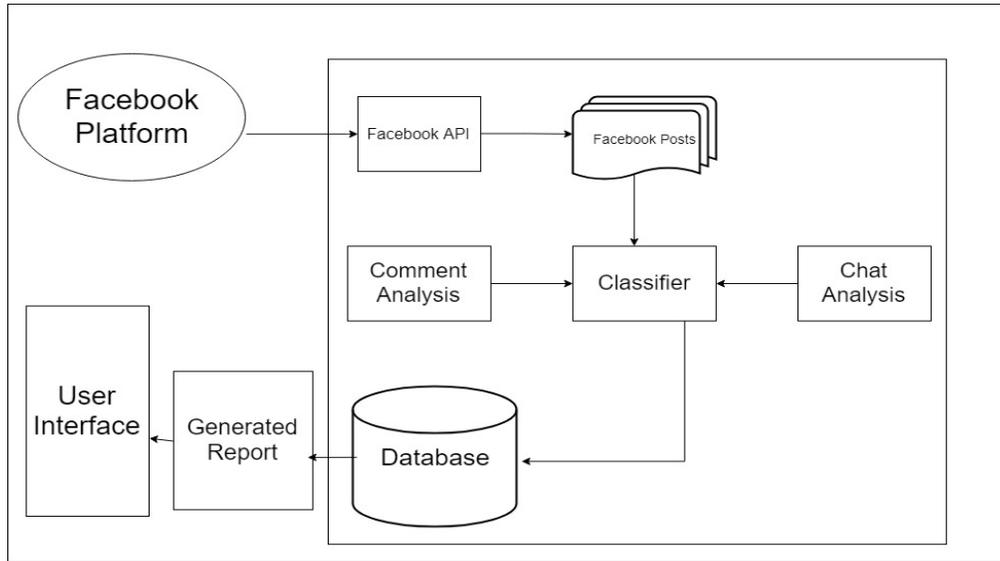
$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Where A and B are occurred and occurring events respectively

## 5. EXISTING SYSTEM

The already existing system demonstrate the social media for healthcare and specially stress detection, Facebooks content-based stress detection have some limitation. Stress is not revealed by the user directly from the post, follow-up like comment from user and his friends can also be used. So, completely relying on the users Facebooks posts is inadequate. Users with high depression level may not be highly active on social network; therefore, their stress is low.

## 6. PROPOSED SYSTEM



**Figure 1:** Proposed system block diagram

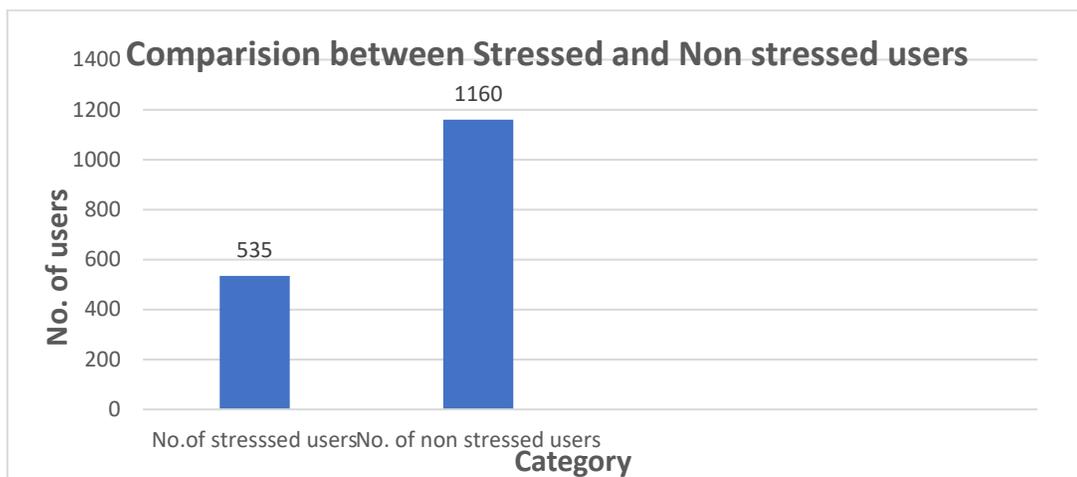
We should design a system which provides much easy flowing way to determine depression level of user using Naïve Bayes algorithm. The extraction of textual data is done by the extraction class from Facebook with the help of Facebook graph API. After extraction, the data is pre-processed. The missing or repetitive attributes are taken care in pre-processing. Techniques like tokenization, lower case conversion, and word stemming and words removal are used for Pre-processing of data. In proposed system according to users Facebook post model can find out whether he/she in depressed or not. But only analysing post won't give accurate result so we also analyse the comments by user and his friends and his chats are also analysed as user will definitely share his depression with his friend. On basis of these analyses the users can be classified as stressed and non-stressed.

## 7. SYSTEM ANALYSIS

In our experiment we identified No. of Stressed and Non-Stressed users.

Serial no.	No. of stressed users	No. of non-stressed users
1	535	1160

From above table we can plot a graph between Stressed and Non-Stressed users. We see 535 users are in stressed and 1160 users in the non-stressed user.



## **8. CONCLUSION**

Mental stress is damaging people's health and making them prone to more diseases so, it is important to detect stress level. For detecting stress level of people, we created a framework that analyses his social media (Facebook) status/posts and his other social activities. We used real world data to train our model. We used Naïve Bayes Algorithm as it gave the highest accuracy on the data. The final goal of the project was to develop a web application where depression level of comment is predicted.

## **SOURCES OF FUNDING**

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

## **CONFLICT OF INTEREST**

The author have declared that no competing interests exist.

## **ACKNOWLEDGMENT**

None.

## **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] Namrata Sonawane, Mayuri Padmane, Vishwja Suralkar, Snehal Wable, Prakash Date via Predicting Depression Level Using Social Media Posts.