DEPRESSION RECOGNIZER FOR TWITTER

A literature review report submitted in partial fulfilment of the requirements for the award of the degree of

Bachelor of Technology

in

Computer Science & Engineering

Submitted by

Adarsh Ajit
Alan Loovees
Akshay Susanth
Albert Francis Xavier

FOCUS ON EXCELLENCE

Federal Institute of Science And Technology (FISAT)®
Angamaly, Ernakulam

Affiliated to

APJ Abdul Kalam Technological University
CET Campus, Thiruvananthapuram

January 2021

FEDERAL INSTITUTE OF SCIENCE AND TECHNOLOGY (FISAT)

Mookkannoor(P.O), Angamaly-683577

FOCUS ON EXCELLENCE

CERTIFICATE

This is to certify that literature review report for the project entitled "**DEPRESSION RECOGNIZER FOR TWITTER**" is a bonafide report of the project presented during VIIth semester (CS451 - Seminar and Project Preliminary) by **Adarsh Ajit(FIT17CS005), Akshay Susanth(FIT17CS011), Alan Loovees(FIT17CS012), Albert Francis Xavier(FIT16CS012)**, in partial fulfilment of the requirements for the award of the degree of Bachelor of Technology (B.Tech) in Computer Science & Engineering during the academic year 2020-21.

		Dr. Prasad J C
Staff in Charge	Project Guide	Head of the Department

ABSTRACT

The rise of social media has essentially changed the way people interact with one another forever, and has become integral to our everyday lives. Social media use has seen a large increase over the past decade with the emergence of smartphones and progress in technology. For some people, social media is essential to even function on a daily basis. Online social networks have emerged as new platforms for the creation of a project that works to capture and analyze linguistic markers in social media posts such that it provides the building blocks to an algorithm that can effectively predict depression. Analyzing linguistic markers in social media posts allows for a low profile assessment which will complement traditional services and would leave a way earlier awareness of depressive signs how it can be hindered or stopped altogether.

Contribution by Author

I personally went through the proposed model and referred two papers which are - **A Comparative Study of Users' Microblogging Behavior on Sina Weibo and Twitter** and **Mining Twitter Data for Depression Detection**. After referring two reference papers I was able to come up with the advantages of proposed method over other methods mentioned in the reference paper.

For the proposed methodology, I contributed to the selection of Multinomial Naive Bayes as classifier. I researched about various similar papers and how they were implemented. After going through many journals and conferences I gained a solid understanding about the application and relevance of the proposed method.

Adarsh Ajit

Contribution by Author

In connection with the literature review, I did a deep research about the topic and domain, searched for relevant research papers and selected the best and relevant ones, they are- **Recognizing Depression from Twitter Activity** and **A Multiclass Depression Detection in Social Media Based on Sentiment Analysis**. Then studied these research papers, understood the aspects behind each paper and drew out results from each paper and compared them with others to derive the best methodologies for this project. For the proposed methodology, I contributed to the selection of activity features from twitter users and towards the multi-class classification of depression.

Akshay Susanth

Contribution by Author

A wide research and read through of papers and publications related to human emotions, depression and impact of social media was conducted. The documentation of the Twitter API was read as a preliminary familiarisation on the usage of the API to obtain the dataset for the project. Papers related to the use of Twitter data for emotion and sentiment recognition and analysis were studied, they are - **Emotion and Sentiment Analysis from Twitter text** and **Emotion Detection from Tweets using AIT-2018 Dataset**.

Alan Loovees

Contribution by Author

To study more about depression recognition from social networking sites, I did a research on two papers namely **Machine Learning-based Approach for Depression Detection in Twitter Using Content Activity Features** and **Identifying Depression on Twitter**. Both these papers gave me an insight and new ideas for our proposed project.
The paper **Machine Learning-based Approach for Depression Detection in Twitter Using Content Activity Features** gave me the approach to create a Document Term Matrix(DTM) in which features fetched from tweets are applied and merged. Ways to collect and select features are also obtained from this paper. I got the idea to use Bag of Words approach from the paper **Identifying Depression on Twitter**. This approach would be very helpful for our proposed system.

Albert Francis Xavier

ACKNOWLEDGMENT

The project **DEPRESSION RECOGNIZER FOR TWITTER** wouldn't have been possible without the kind support and help of many individuals. I use this opportunity to show gratitude towards each and every one who has made an impact on this project, directly or indirectly.

First of all, I thank God Almighty whose blessings were always with me and helped me to complete this project work successfully. I express my sincere gratitude to **Dr. GEORGE ISSAC**, Principal, FISAT, and to **Dr. PRASAD J C**, Head of Department, Computer Science and Engineering, FISAT, for providing the ambience, infrastructure and permissions to avail all required facilities, for undertaking this project in a systematic way.

I am extremely grateful to **Ms. SRUTHY SURESH**, our Project Guide, for her valuable suggestions, support, encouragement and constant guidance throughout all the phases of the project.

I am thankful to our Project Coordinators **Mr. Jyothish K John** and **Mr. Anuranj P** for their constant support in every phase of this project. I am lovingly and gratefully indebted to my Parents, Teachers, Siblings, Friends and my well-wishers who were there always for helping me in doing this project.

Adarsh Ajit
Alan Loovees
Akshay Susanth
Albert Francis Xavier

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Overview

Human beings are social creatures. In-order to thrive in life we want the companionship of others, the interactions and also the strength of our connections features a huge impact on our mental health and happiness. While working as a group can boost self-worth and can in turn reduce stress, anxiety and depression. Making new friends and connecting with our loved ones can provide comfort, joy, prevent loneliness and even add years to your life. On the other hand, one can pose a serious risk of being depressed if he or she is lacking strong social connections. Surrounding oneself with a small group of people who you can talk to whenever you are lonely is greater than having a huge friend circle.

In this tech driven world, many people consider social media platforms like Facebook, Twitter, Snapchat, YouTube, and Instagram to seek out and connect with one another. While each has its benefits, spending too much time engaging with social media can actually make you feel more lonely and isolated—and can propel the chances of mental health problems such as anxiety and depression.

Figure 1.1: Major Depressive Disorder (MDD)

It's important to recollect that social media can never be a replacement for real-world human connection. It requires in-person contact with others to trigger the hormones that alleviate stress and cause you to feel happier, healthier, and more positive. The most ironic fact for a technology that's designed to bring people closer together is that if a person is spending an excessive amount of time on social media the chances of him/her feeling sadness, dissatisfaction, frustration, or loneliness is high. All these parameters impact one's life, it's high time that we re-examine our online habits and find a healthier balance.

We propose a system to detect depression in twitter users. Twitter Developer API is used to retrieve information such as number of following, total number of posts and time of posts.The Data preprocessing includes cleaning the collected data, annotating the data, and text normalization. Data cleaning includes removing all user mentions, repeated words, punctuations, URLs, pictures and videos. Text normalization includes tokenization, stop-words removal, stemming, replacing emojis with their emotion keywords. A Document Term matrix is created for each account, TF-IDF is used to assign weights. Features applied on the DTM are then merged with account measures extracted from the social network and user activities. The results of the merge is given into a Multinomial Naive Bayes classifier to classify the users into three classes- Highly depressed, Depressed and Not Depressed.

1.2 Problem Statement

Health is a very important part of each and one of our well being. Early detection of diseases can help contain it, and provide appropriate treatment, especially mental health illnesses like depression. However, at the early stages of depression, 70% of the patients wouldn't consult doctors, which may take their condition to advanced stages. Social media platforms play an important role in our daily lives, especially for the 18 to 30 year old demographic. These platforms are used by the public to express their opinions on current affairs, products and brands. This gives an excellent opportunity to analyze and detect depression among users on one such platform, Twitter.

1.3 Objective

The project works to develop a social media-based depression recognition tool using existing research. The target is to detect and analyze sentiment and emotion, specifically depression expressed by people on Twitter.

Chapter 2

Literature Review

A Base paper was selected initially and a through study of the base paper was conducted and then literature survey was conducted to get more information on the domain and the topic. 8 literature papers related to depression recognition from social media were selected and a comparison study was conducted. The architecture and methodology of the project **DEPRESSION RECOGNIZER FOR TWITTER** is based on the comparison study conducted and the inferences made from it. A brief description of the papers are shown below:

2.1 Base Paper

2.1.1 A Depression Recognition Method for College Students Using Deep Integrated Support Vector Algorithm

The psychological impact of depression on people is very large. The concentration and learning ability of people with depression will decrease accordingly, and the efficiency of work will be greatly reduced, which will greatly affect their lives.

As a social networking tool in China, Weibo is one of the most popular personal and media publishing platforms in China. As Weibo is a platform for users to share their feelings, express their opinions, and interact with other users, it contains a large amount of user personal information and emotional dynamics. Obtaining and analyzing these Weibo contents can help in mining the personal emotions of the users.

Methodology: This study verifies the feasibility of identifying depression among college students using their social network data. They used user data from 1,000 college student Weibo users as the target sample. Feature extraction is done using deep neural networks and SVM classifier is used for dimensionality reduction.

Input data was constructed by counting the frequency of specific words in Weibo. 20 most frequently used words by both depressed and normal users were found. Integrated SVM was selected as the depression recognition model. The integration strategy used the AdaBoost algorithm.

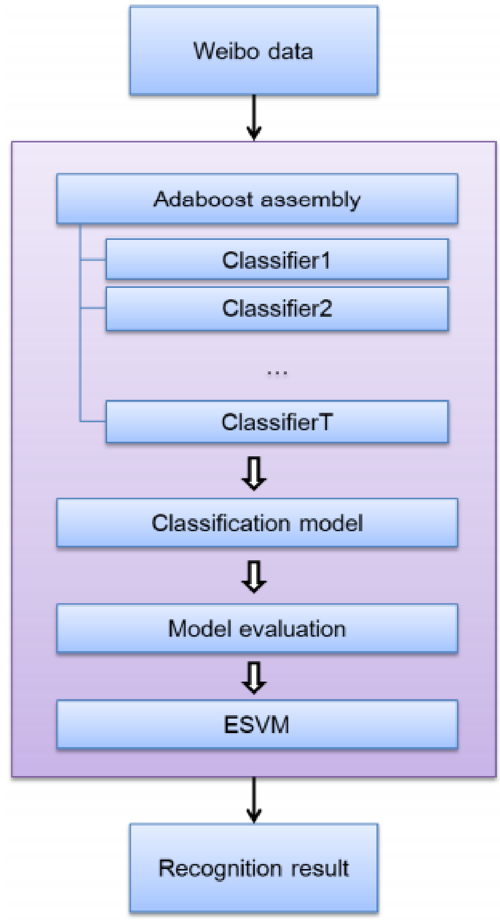


Figure 2.1: DISVM-based depression recognition framework.

Input	Givenatrainingsampleset $X = \{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}, y_i \in \{-1, 1\}, i \in (1, 2, \ldots, n)$, base classifier SVM, maximum number of iterations $T_{\max}$. Initialize the weight of X as $X_1(i) = 1/n$.
Step1	For $t = 1,2,\ldots,T_{\max}$
Step2	Learning under X_t, the t-th weak learner $L_t = L(X, X_t)$ is obtained from the $X_1(i)$ training set with weight distribution.
Step3	Calculate L_t error rate ζ_t by $\zeta_t = \sum_{i=1}^{X_t(i)} (L_t(x_i) \neq y_i)$
Step4	If $\zeta_t \geq 0.5$,break;
Step5	Calculate the weight w of L_t according to formula $w_t = \frac{\ln \frac{1-\zeta_t}{\zeta_t}}{2}$
Step6	Update the weight of each sample $X_{t+1}(i) = \frac{X_t(i)}{X_t} \bullet \begin{cases} \exp(-w_t), l_t(x_i) = y_i \\ \exp(w_t), l_t(x_i) \neq y_i \end{cases}$
Output	$L(X) = sign\left(\sum_{t=1}^{T_{amx}} w_t L_t(x)\right)$

Figure 2.2: The running steps of the DISVM algorithm.

Results: The DISVM algorithm proposed in this paper can give low-quality samples low weight, thereby improving the model recognition rate. The figure shows that when the time dimension of the crawl is 24 months, the recognition rate does not change much, and there is a downward trend in the recognition rate later. Therefore, an excessively long observation period is not conducive in identifying the depression tendency of Weibo users.[1]

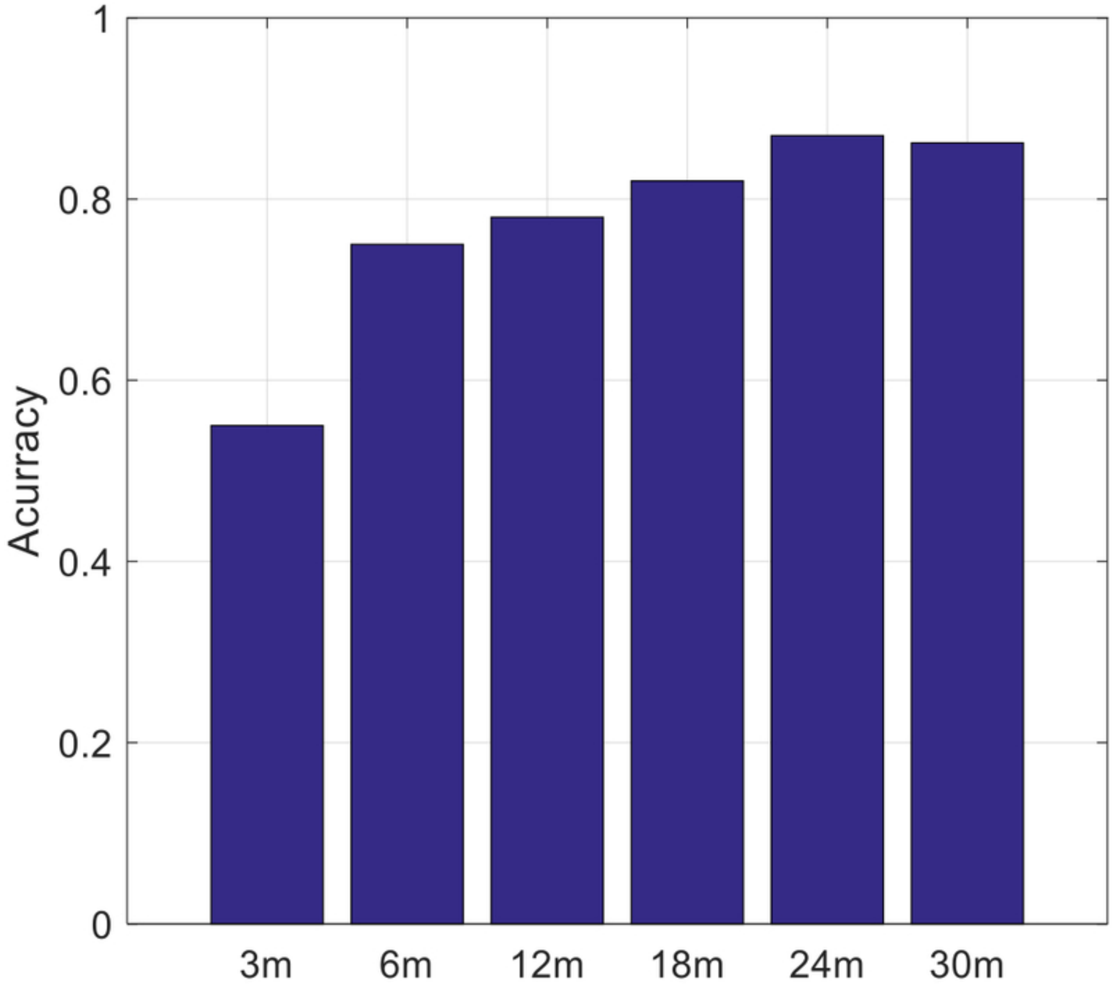


Figure 2.3: Comparison of recognition accuracy in different time dimensions

Algorithm	FP	FN	TP+TN	Precision
RBF-NN	73	19	471	0.8366
SVM	76	21	466	0.8277
KNN	81	23	459	0.8152
DISVM	56	11	496	0.8810

Table 2.1: Recognition accuracy of four comparison algorithms on the training set.

Algorithm	FP	FN	TP+TN	Precision
RBF-NN	17	6	107	0.8231
SVM	19	7	104	0.8000
KNN	20	7	103	0.7923
DISVM	13	5	112	0.8615

Table 2.2: Identification accuracy of the four comparison algorithms on the test set.

2.2 Literature Papers

2.2.1 A Comparative Study of Users' Microblogging Behavior on Sina Weibo and Twitter

Social media platforms like Twitter allow people to tweet, share and discuss short messages on the Internet. Today, Twitter users publish more than 200 million tweets each day. In China, a social media platform called Sina Weibo is leading the microblogging market since twitter is not available in that region. Both these platforms has a similar functionality. A few common features are : 140 character limit to these microposts, both these platforms has a follower-followee hierarchy.

Data Collection: Microposts were collected over a span of two months via the Sina Weibo Open API and Twitter Streaming API respectively. For the twitter data, it initially started with a seed set of 56 twitter users. This set was expanded gradually and now the set consist of more than 24 million tweets posted by over 1 million users. For the collection of data from Sina Weibo there was a limitation since the working of Sina Weibo API is different from that of the twitter API.Thus the most recent public microposts were collected. The Sina Weibo data now consist of more than 22 million posts from over 6 million users.

Methodology: To compare the user behaviour on both these platforms Sina Weibo and Twitter, each micropost is analyzed including the words and URLs so as to compare user behavior on these platforms to gain more insights about the user.

More than 40 million microblog posts were deeply analysed and microblogging behavior on Sina Weibo and Twitter were compared based on five aspects (i) access behavior, (ii) syntactic content analysis, (iii) semantic content analysis, (iv) sentiment analysis, (v) temporal behavior.

Results: The people who use Twitter tend to be more active during weekdays. Moreover, user interests change between weekends and weekdays. Intensive usage of hashtags and URLs which is characteristic for the Twitter users. They appear to be more desirous to let their posts appear within the discussions.The number of negative posts against organisations and individuals are found more in twitter.

On both platforms, the Web interface is the most well liked to access the microblogging services: 43.1% of the posts are published via the online on Sina Weibo and 38.5% on Twitter. Other popular clients on Sina Weibo are mainly designed for mobile devices like the iPhone (7.6%) and Nokia devices (9.4%).

The classification of users posts into positive,negative or neutral is done by this framework based on sentimental analysis. Majority of the posts from both these microblogging platforms was neutral. About 83.4% and 82.4% of the Sina Weibo and Twitter posts were classified as neutral. Also users of Sina Weibo was observed to have a a rising interest in persons and organizations during the weekend, whereas the interests of Twitter users focus more on locations. [2]

Type of access	Fraction of posts in weibo	Fraction of posts in twitter
Web	54.9	66.2
Mobile	45.1	33.8
Microblogging activity	90.6	96.7
Activity on another platform	9.4	3.3

Table 2.3: Number of posts published via different categories of access clients

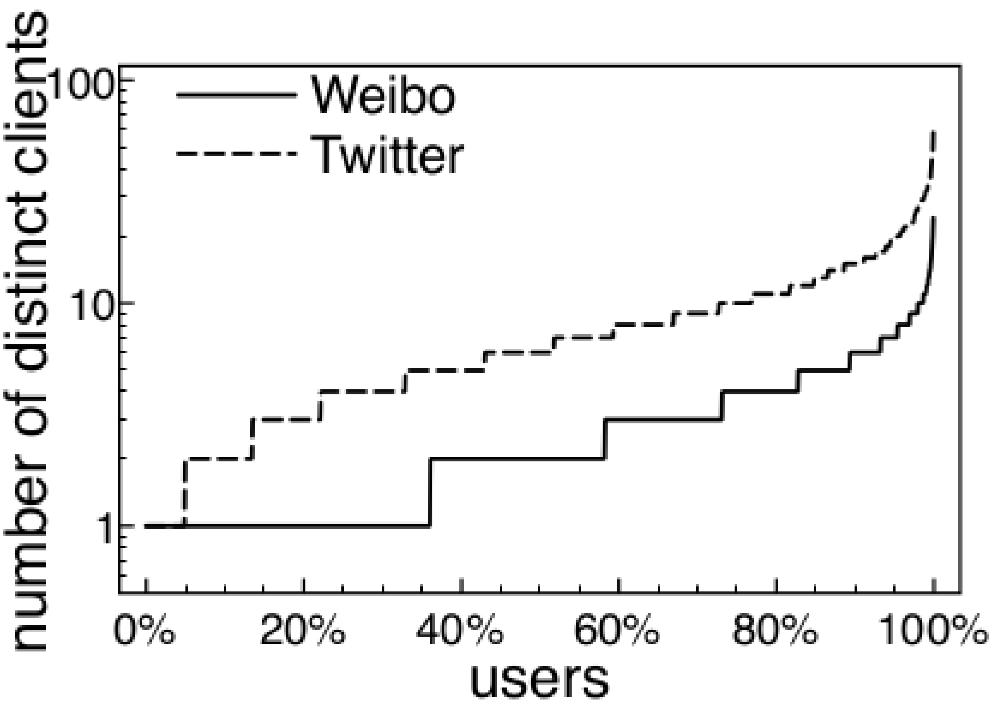


Figure 2.4: Number of distinct access clients for individual users

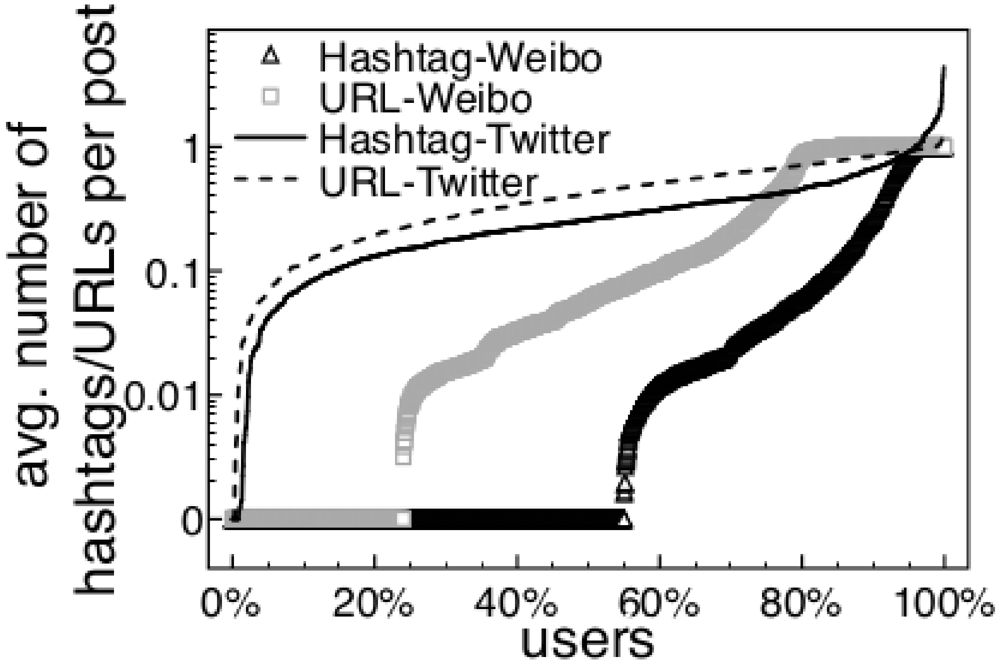


Figure 2.5: Comparison of writing style for individual users

2.2.2 Mining Twitter Data for Depression Detection

This paper presents a unique approach for the classification of health tweets for depression, anxiety etc from a collection of tweets which gives a better understanding of the health status in a living environment. Since depression can be cured if detected in its early stages this new platform provides interaction through patients, many of them participating in decision making for better results of health care treatment. After retrieval of these health tweets, some essential steps are performed on it.

Methodology: The dataset has been extracted from twitter streaming which consists of all tweets as per now, but we will classify tweets consisting of keywords such as Depression, Anxiety and Mental Illness, for achieving our goal. Pre-processing of retrieved data is the primordial step.

It takes one string as an input, which goes through all the steps of pre-processing such as removing slangs, emojis, hashtags etc. After that in feature extraction tweets are analyzed from various sentiment analyzers textBlob and SentiWordNet (in python script). TextBlob is used to extract the polarity and subjectivity of the sentence. SentiWordNet dictionary is also used to calculate the positivity and negativity of the sentence.

Multinomial Naive Bayes and Support Vector Regression methods are used as classifiers to distinguish the health tweets for Depression, Anxiety from the mixed tweets.

Results: Tweets are validated according to positive, neutral and negative score of sentence with the aim of determining the accuracy of these analyzers. Multinomial Naive Bayes (MNB) and Support Vector Regression (SVR) classifiers are

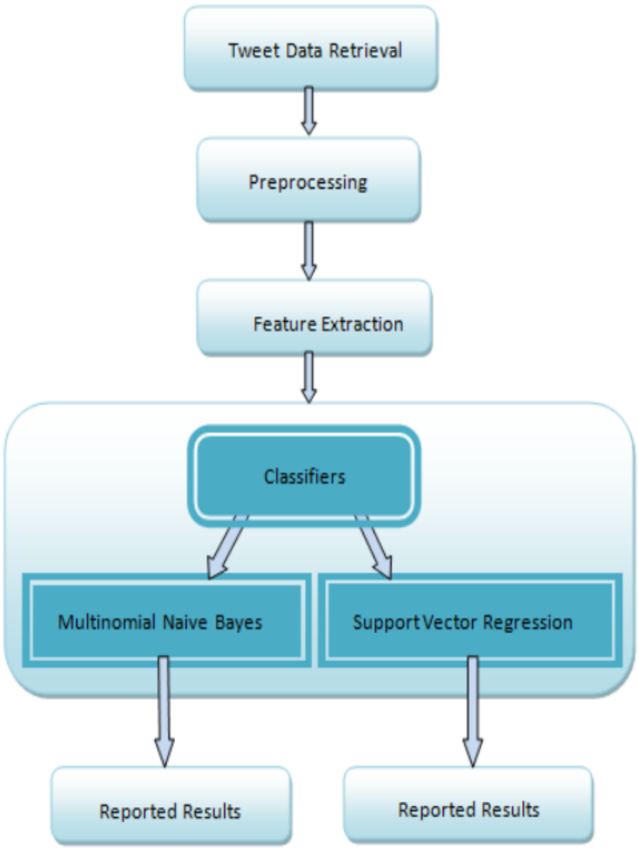


Figure 2.6: Framework of the system

implemented to give a comparison of their accuracies. [3]

2.2.3 Emotion and Sentiment Analysis from Twitter text

This work focused on detecting and analyzing sentiments and emotions expressed by people from tweets and use them for generating recommendations. The authors collected tweets and replies on few specific topics. The dataset contained text, user, emotion, sentiment information, etc. Sentiment and emotion were detected from tweets and their replies and the influence scores of users were measured based on various user-based and tweet-based parameters.

The authors used this information to generate recommendations for users based on their twitter activity. Some interesting features of this paper were that they, (i) included replies to tweets in the dataset and measurements, (ii) introduced agreement score, sentiment score and emotion score of replies in the calculation of the influence score for recommendations, (iii) generated general and personalized recommendation containing list of users who agreed on the same topic and expressed similar emotions and sentiments towards that particular topic.

Methodology: In this paper, they detected the emotions and sentiments from tweets and their replies and formed an emotion network based on texts posted by users. From the emotion network, influential people were detected for both positive and negative emotions. Finally, a trust network was computed based on

Figure 2.7: Sample tweets and replies

emotion similarities and influences by generating recommendations.

Results: The authors found out that the users who reply to tweets normally share similar emotion or sentiment and agree to the tweet content. When they found influential users of a network, the result was more specific to a certain issue or topic. The personalized recommender provided customized recommendations to suggest users from the network using information about other people who share similar sentiment or emotion on a topic as the user himself/herself.[4]

Algorithm	Full text	NAVA text
Naive Bayes (Sentiment)	66.86	61.15
SVM (Sentiment)	23.32	23.32
Random Forest (Sentiment)	55.23	52.01
Naive Bayes (Emotion)	47.34	43.24
SVM (Emotion)	14.48	14.48
Random Forest (Emotion)	35.66	37.26

Table 2.4: Sentiment and emotion classification accuracy comparison (in percentage %).

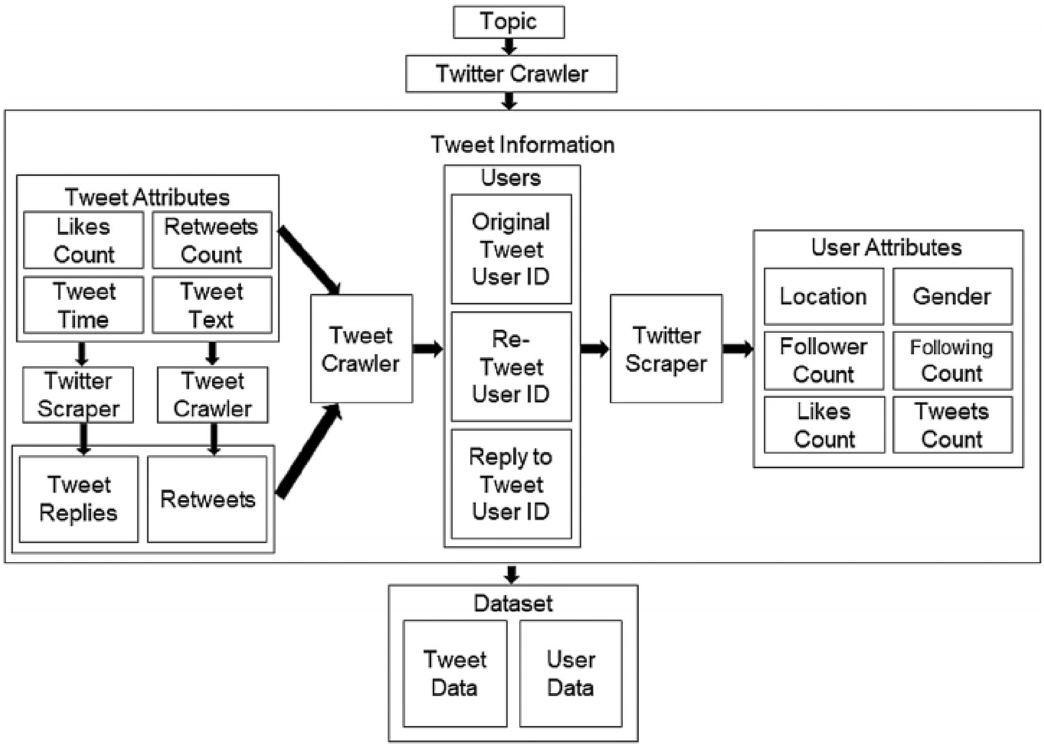


Figure 2.8: Dataset creation workflow.

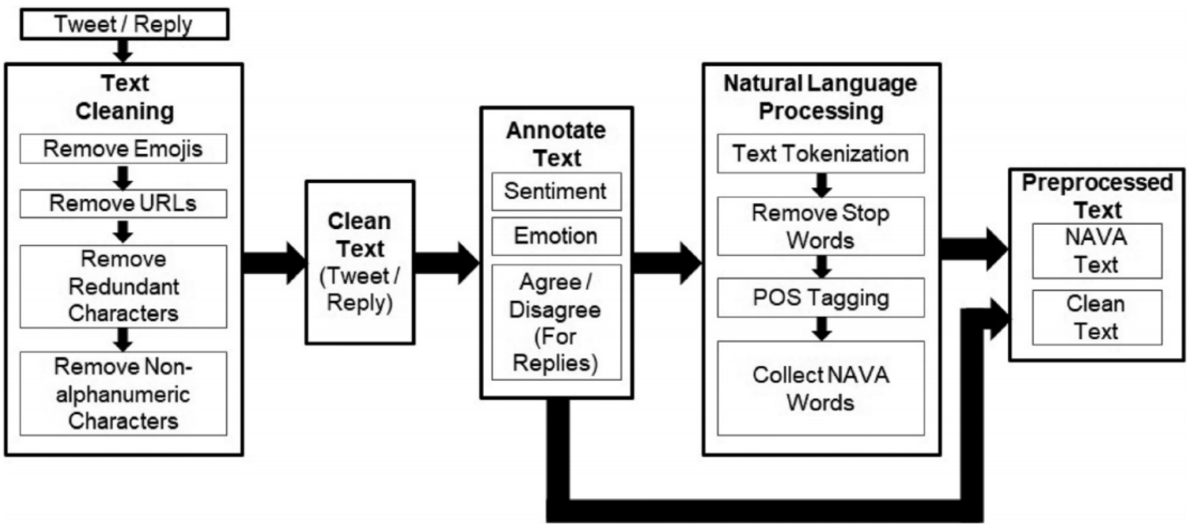


Figure 2.9: Preprocessing workflow.

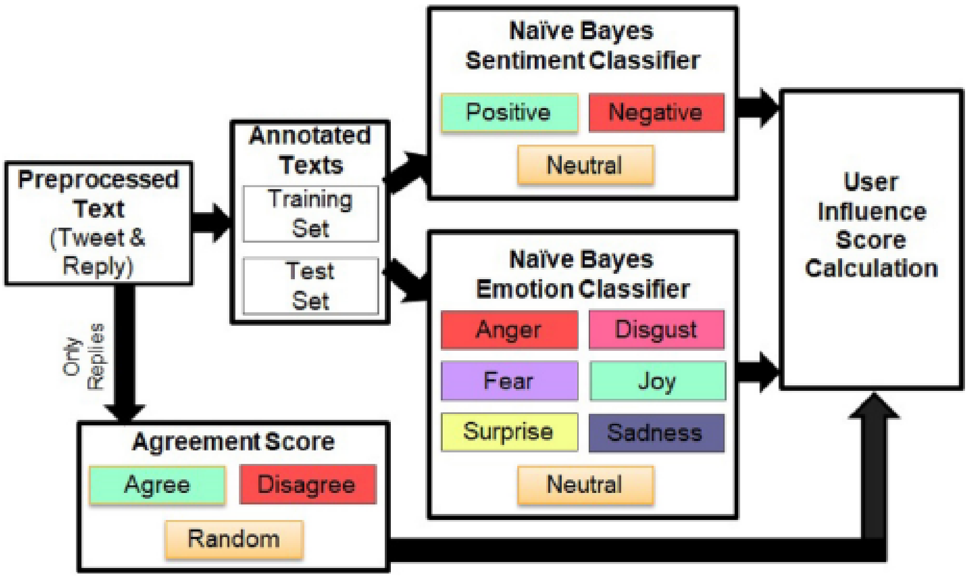


Figure 2.10: Classification influence score calculation workflow.

2.2.4 Emotion Detection from Tweets using AIT-2018 Dataset

In the world of today's social networks, users share their opinions and emotions in their way through different medium like Twitter, Instagram, Facebook, and many more. Millions of users, in their everyday lives, express their views and opinions and also their emotions on or about a particular thing through social networks. This gave the researchers an excellent opportunity to analyze the emotions of social networking users' activities. These large numbers of data, generated by social networks contain feelings, opinions and emotions of people from day to day.

Different emotional analytical research on the social platform has been underway over the years. It is a challenge to analyze the correct emotion from social data, as different people have different thoughts. This makes it clear about the need to work on these problems and it offers many possibilities for future research into the hidden identification of emotions of users in general or emotions of users on a specific topic etc.

Methodology: Lexical affinity method was used combined with learning-based methods to automatically classify multi-class emotions from the dataset. WordNet-Affect and EmoSenticNet emotion lexicon was used to extract the emotion containing words as features separately. Each tweet is represented into a vector of features for the training of a classifier from labeled data. They employed two well-known techniques to create the feature vector from the 2 features set. These are term frequency (TF) and term frequency and inverse document frequency (TFIDF). The authors worked using the AIT-2018 dataset.

Results: They have taken 1000 tweets from each class from the dataset, 4000 samples in total. They applied the training dataset into three base Machine Learning algorithms: Naive Bayes, Decision Tree and Support Vector Machine (SVM). A precision of 90% was found on anger class using the random forest as the main classifier and logistic regression to find the features needed. [5]

Comment	Tweet
Real Tweet	@SatisfyingTaste @TheAnimalVines I used to make the peanut butter energy balls all the time. My famjam loved them! Btw my cats keep loving them as well :D :D #recipes #yummy 100 âðŸ https://bit.ly/2orTQLP
Preprocessed Tweet	i used to make the peanut butter energy balls all the time my famjam loved them by the way my cats keep loving them as well recipes yummy joy joy

Figure 2.11: Preprocessing sample.

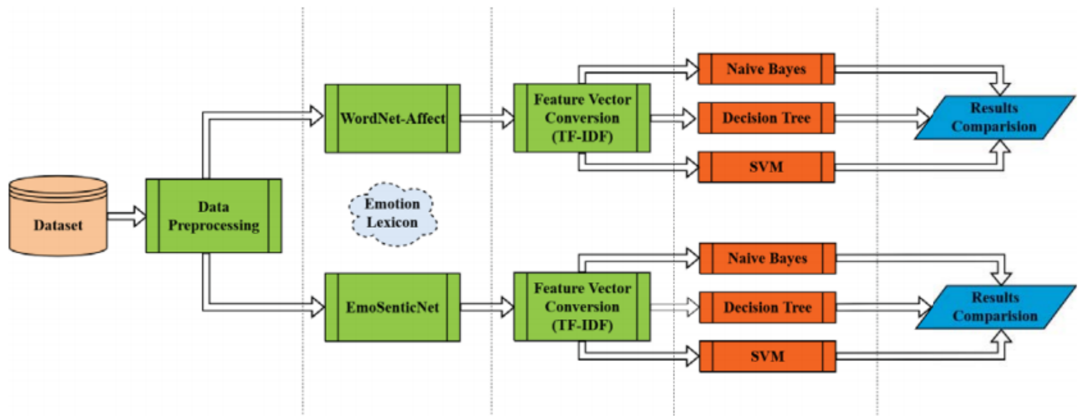


Figure 2.12: Methodology of Emotion Detection from Tweets using AIT-2018 Dataset

Emotions	Feature selection process	Naive Bayes	Decision Tree	SVM
Joy	WordNet	47%	52%	49%
	EmoSenticNet	80.32%	80.32%	81.96%
Fear	WordNet	62%	56%	48%
	EmoSenticNet	73.33%	80%	73.33%
Anger	WordNet	37%	33%	50%
	EmoSenticNet	75%	85.71%	89.28%
Sadness	WordNet	43%	43%	43%
	EmoSenticNet	97.05%	78.43%	96.07%

Table 2.5: Precision of SVM, Naive Bayes, Decision Tree using TF-IDF

Emotions	Feature selection process	Naive Bayes	Decision Tree	SVM
Joy	WordNet	85%	83%	83%
	EmoSenticNet	94.23%	80.32%	92.59%
Fear	WordNet	20%	39%	41%
	EmoSenticNet	91.66%	70.58%	91.66%
Anger	WordNet	30%	33%	30%
	EmoSenticNet	84%	70.58%	89.26%
Sadness	WordNet	39%	31%	31%
	EmoSenticNet	82.5%	86.95%	85.21%

Table 2.6: Recall of SVM, Naive Bayes, Decision Tree using TF-IDF

Feature selection process	Naive Bayes	Decision Tree	SVM
WordNet	49.23%	45.64%	47.17%
EmoSenticNet	86.42%	80.09%	88.23%

Table 2.7: Accuracy of different classifiers

2.2.5 Machine Learning-based Approach for Depression Detection in Twitter Using Content and Activity Features

Methodology: This study analyzes activity features from Twitter tweets to detect users with depression. Here the system uses depression detection using activity and content features (DDACF) classification model. In this method, all tweets for depressed and non-depressed accounts, as well as information of user account and activities such as number of followers, number of following, total number of posts, number of mentions and number of retweets are retrieved. And then all tweets of an account are assembled in one document.

Text preprocessing is applied to all documents. And a corpus is created and tweets in each document are tokenized. In the next phase, normalization is applied, where all characters are turned to lower case and punctuations, retweets, mentions, links, unrecognized emojis and symbols are removed. Normalization includes removing stop words, such as first-person pronouns like "I," "me" etc. Stemming is also applied and then our required document term matrix (DTM) is created for each account. The matrix that was created indicates the frequency of words in each tweet. Each row in DTM indicates a document of tweets and each column indicates all words used in all accounts. TF-IDF is used to measure the words' weight.

Features applied on the Document Term Matrix(DTM) are then merged with account measures extracted from the social network and user activities. The resultant merge are then treated as independent variables in a classification algorithm

Figure 2.13: User activity features extracted from user account

to predict the dependent variable of an outcome of interest. Ultimately, we decide upon the DT, a linear and radial kernel support vector classifier, and an NB algorithm. DDACF classification model is described in Figure 2.14

Twitter platform has massive amount of information about the user, various features can be extracted from the activity histories and tweets of Twitter users. Features are extracted from the text after text preprocessing, when the text is in the desired format, as shown in Figure 2.15.

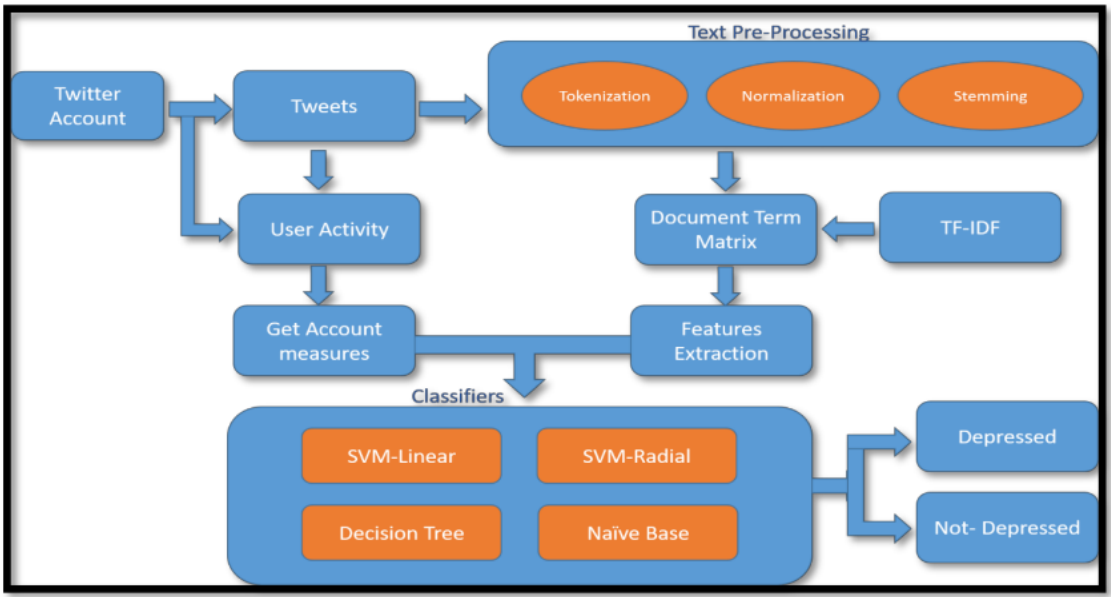


Figure 2.14: DDACF classification models

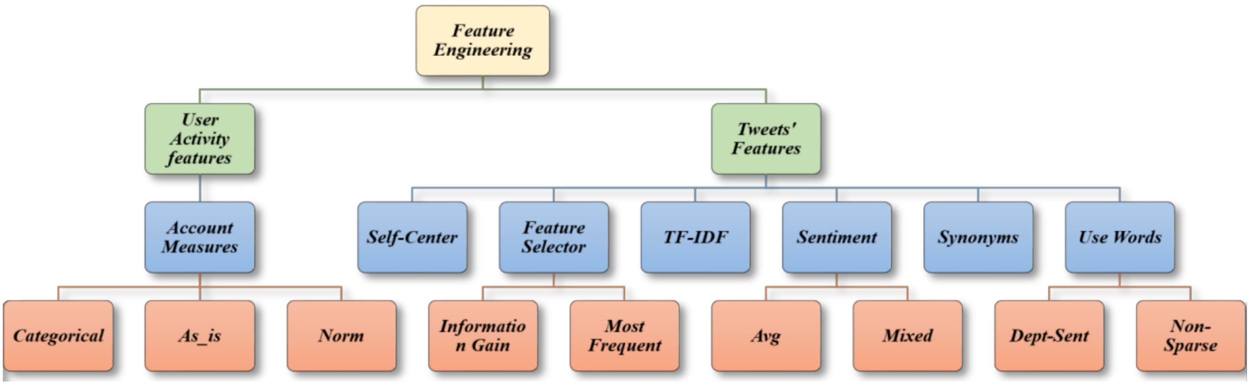


Figure 2.15: Visualization of features used in the study

Results: This paper defines a binary classification problem as identifying whether a person is depressed, based on his tweets and Twitter profile activity. SVM model has achieved optimal accuracy metric combinations. It gives an accuracy of 82.5%. The results obtained from various classifiers like Decision Tree, Naive Bayes etc is summarized in Table 2.8

The main contribution of this study lies in exploiting a rich, diverse, and discriminating feature set that contains both tweet text and behavioral trends of different users.

Classifier	Accuracy	Precison	Recall
DT	77.5	0.65	0.590909
NB	80	0.653846	0.809524
SVM	82.5	0.73913	0.85

Table 2.8: Evaluation measures of different classifiers

2.2.6 Recognizing Depression from Twitter Activity

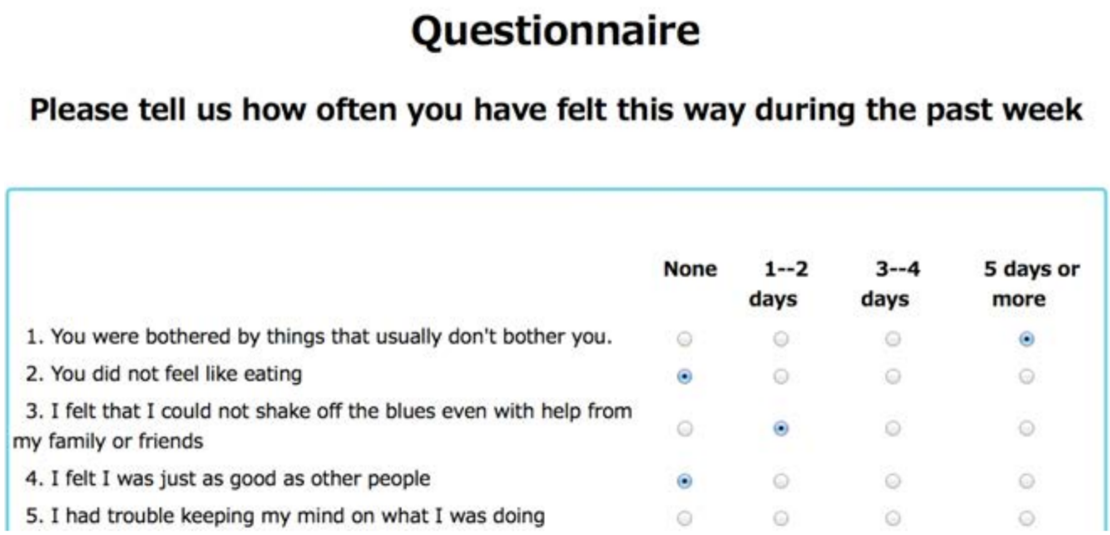
Methodology: This study investigated the degree to which the activity history of Twitter users is useful in recognizing depression. The results from a web based questionnaire is used for measuring the degree of depression in twitter users and collected their activity history. The activity history was collected using Twitter API. The statistics on the activity history of the participants are shown in the Table 2.9.

	Mean	Median	Standard Deviation
No. of tweets	2749.1	3191	850.2
No. of tweets per day	18	7.8	27.5
Days spanned by tweets	569.2	348.2	502

Table 2.9: Basic statistics about activities of participants

The CES-D questionnaire was used to measure the depression, a sample of the questionnaire is shown in the Figure 2.16. The scores of each participants along with the time taken to answer was also recorded and those completed in very small time were excluded. The features extracted for predicting depression are shown in the Table 2.10.

SVM classifiers are used for estimating the level of active depression in users. The classification accuracy was evaluated by 10-fold-cross-validation. Several classifiers were constructed and the classification accuracy of each was determined. Precision, Recall, Accuracy and F-measure were the evaluation parameters used.



Questionnaire

Please tell us how often you have felt this way during the past week

	None	1--2 days	3--4 days	5 days or more
1. You were bothered by things that usually don't bother you.	○	○	○	◉
2. You did not feel like eating	◉	○	○	○
3. I felt that I could not shake off the blues even with help from my family or friends	○	◉	○	○
4. I felt I was just as good as other people	◉	○	○	○
5. I had trouble keeping my mind on what I was doing	○	○	○	○

Figure 2.16: A screenshot of the questionnaire

No.	Name	Description
F1	Bag of words	Frequencies of words used in the tweet
F2	Topic	Ratio of tweet topics found by LDA
F3	Positive	Ratio of positive-affect words contained in the tweet
F4	Negative	Ratio of negative-affect words contained in the tweet
F5	Hour	Hourly posting frequency
F6	Tweet frequency	Tweets per day
F7	No. of words	Average number of words per tweet
F8	RT	Overall retweet rate
F9	Mention	Overall mention rate
F10	URL	Ratio of tweets containing a URL
F11	followee	Number of users following
F12	follower	Number of users followed

Table 2.10: Features used for predicting depression

Results:

A significant difference between the depressed and non-depressed group was found for the relative ratio of negative words, posting frequency, retweet rate and ratio of tweets containing a URL. A significant difference can also be seen between groups in the number of followees and followers, suggesting it as a useful feature. No big differences were seen in mention frequencies suggesting that the mention frequency is not a robust feature to detect depression. [6]

The Figure 2.18 shows that the distribution of the word frequencies are biased.

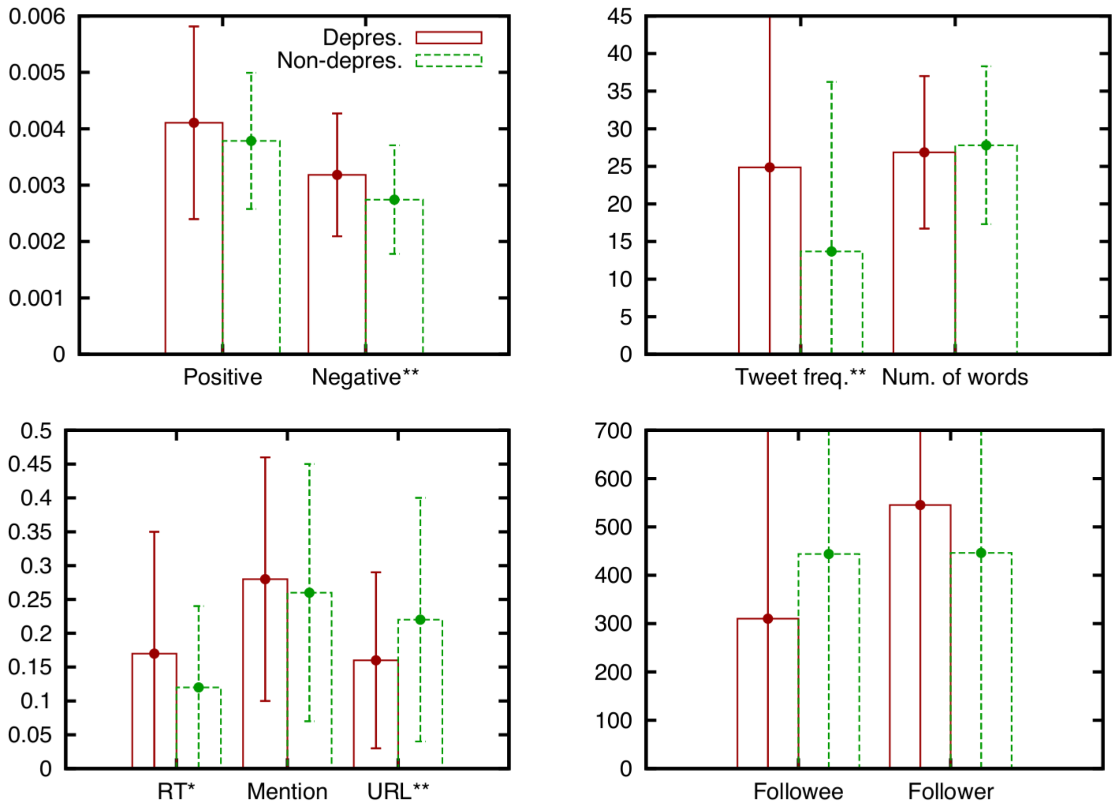


Figure 2.17: Comparison of features between participants with and without depression

Features	Precision	Recall	F-measure	Accuracy
bag of words (20,000 words)	0.04	0.005	0.024	61%
bag of words (2,000 words)	0.58	0.13	0.21	62%
5 topics	0.50	0.13	0.21	61%
10 topics	0.55	0.35	0.43	64%
20 topics	0.54	0.33	0.41	63%
F3 + F4 + F6 + F8 + F10 + F11 + F12	0.57	0.33	0.41	64%
F2 + F3 + F4 + F6 + F8 + F10 + F11 + F12	0.61	0.37	0.46	66%

Table 2.11: Classification accuracy of constructed models

The figure shows that the frequency of posting was highest at 11 p.m and lowest at 4 a.m . But no big differences were found in the posting frequency of two classes.

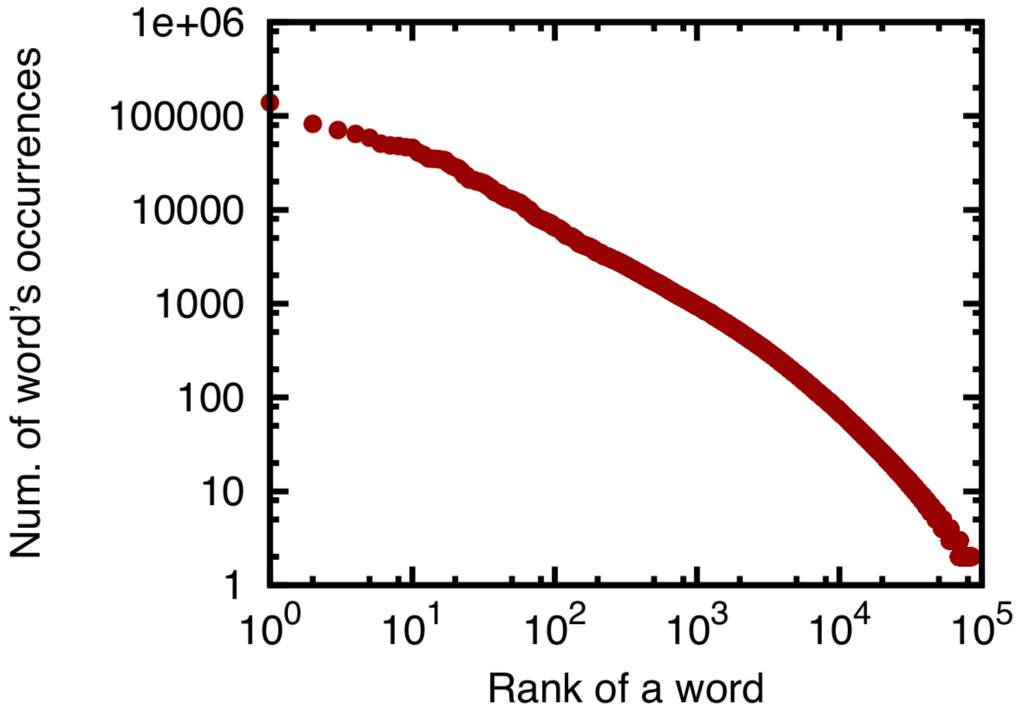


Figure 2.18: Distribution of word frequencies

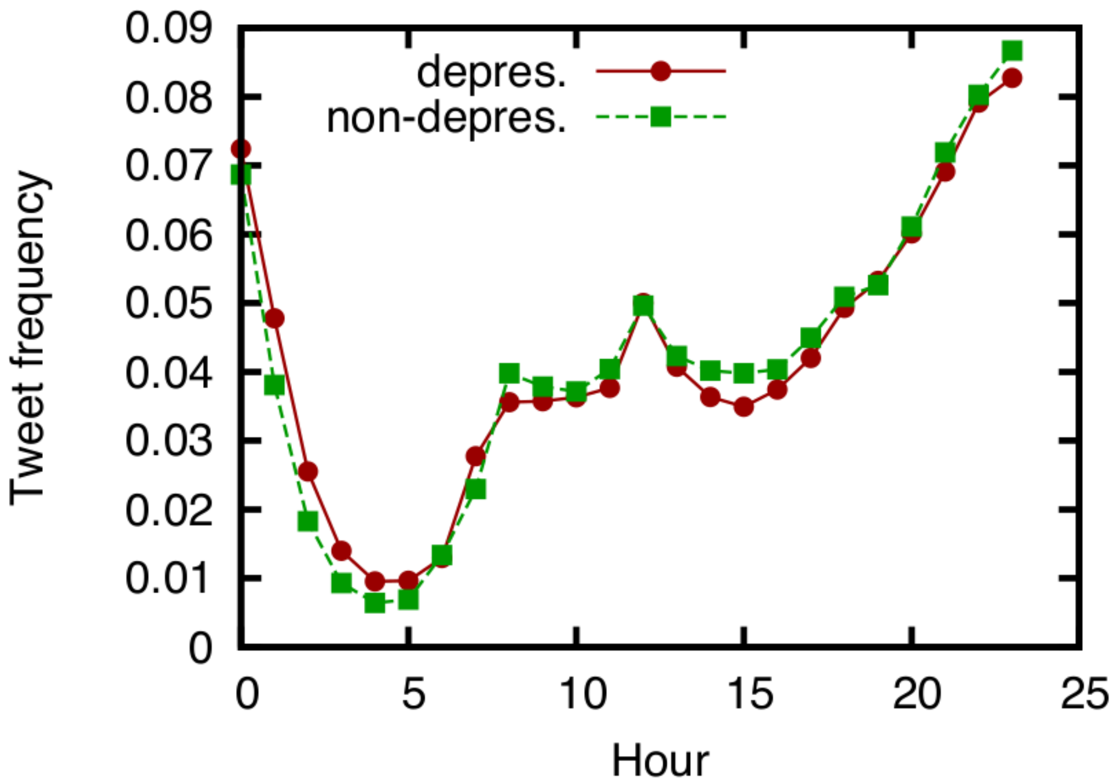


Figure 2.19: Hourly posting frequency: Comparison between participants with and without depression

Models using tweet topics as a feature achieved a high accuracy. This means it may be possible to improve estimation accuracy by reducing the dimensions to the broad feature of topics, rather than using the narrow feature of word frequencies.

The results showed that the activity history of users can be used to recognize the presence of depression with an accuracy of approximately 69%.

2.2.7 A Multiclass Depression Detection in Social Media Based on Sentiment Analysis

Methodology: This study classifies the diagnosed depressive users from Twitter into three classes High, Medium, Low depress stage. Recent tweets of 179 individuals, who have reported depression were extracted using Twitter API. Then they converted raw tweets into useful text by cleaning the data and then normalized the text. After which they assigned weights to the tokens based on their relative effectiveness using TF-IDF function. It gives a complete count of term occurrences. Thus 100 most frequently used words were collected from each user. Then by using a method called LIWC (Linguistic inquiry and word count), the words were classified into 14 psychological attributes such as social, family, friends etc. Then weights were assigned to each word based on the scale of happiness ranging from unhappy to happy i.e. for unhappy words 1 will be assigned and for happy words 9 will be assigned as weights. Thus a total weight will be computed, according to that weight classification is done by machine learning classifiers as High, Medium and Low.

Algorithm: Multi-class depression detection

Input: sw = string words, iw = input words, sd = string document, ww = word weight, and A = matrix.

1. For I ← 0 to n
2. do A[0,i] ← sw_i
3. do A[1,i] ← 0 // initialize all with zeros
4. For i←0 to n
5. input iw_i
6. If(iw_i ==x i)
7. Then A[1,i] ← 1
8. H←0, M←0, L←0
9. For j←0 to n
10. If ww[j] >= 1 and ww[j] <= 3.9
11. Then H ← H+1
12. Else if ww[j] = 4 to 6.9
13. Then M ← M+1
14. Else L ← L+1

15. If H¿M and H¿L
16. Then MaxVal ← H
17. Else if M¿H and M¿L
18. Then MaxVal ← M
19. Else MaxVal ← L

Output: Depression class of the tweet in the form of H, M and L

Results:

Model	Class	AUC	Precision	Recall	F-Score
1DCNN	H vs. M L	0.91	0.92	0.86	0.89
1DCNN	M vs. H L	0.83	0.85	0.54	0.66
1DCNN	L vs. H M	0.86	0.93	0.78	0.85
NN	H vs. M L	0.89	0.78	0.78	0.78
NN	M vs. H L	0.89	0.83	0.83	0.83
NN	L vs. H M	0.88	0.83	0.83	0.83
SVM	H vs. M L	0.86	0.77	0.93	0.84
SVM	M vs. H L	0.91	0.90	0.81	0.85
SVM	L vs. H M	0.86	0.93	0.78	0.85
RF	H vs. M L	0.80	0.90	0.60	0.72
RF	M vs. H L	0.83	0.85	0.54	0.66
RF	L vs. H M	0.83	0.84	0.84	0.84

Table 2.12: Performance evaluation of different models

The One-vs-All technique to differentiate the different level of depressed users were useful. The classification of word selections in the classes of High, Medium, and Low depression are the major findings. The top 100 words used by depressive users were utilized to build a classifier that has classified users with an accuracy of 91% . The suggested classification approach can be used to detect similar patterns on Twitter for timely handling of severe consequences. [8]

2.2.8 Identifying Depression on Twitter

Methodology: This study uses a crowdsourced method to compile a list of Twitter users who are diagnosed with depression. This study uses a dataset from CLPsych. The data was collected from several Twitter users who stated their random thoughts. This dataset was developed from the data of users with public Twitter accounts who posted a status update in which they speak of their state of mind.

For each user, about 3000 of their most recent public tweets are taken as part of the dataset, and each user is isolated from the others. It is to be noted that this 3000 tweet limit derives from Twitter's archival polices, and that most tweets concentrated long after a two month time span may possibly lower the effectiveness of a classifier.

In the last phase, each user and the users they interacted with had been removed from the dataset to ensure privacy. In addition to that, participants were required to sign a privacy agreement, institute security measures on the data, and obtain the approval of an ethics review board in order to secure the dataset. Data had been distributed in compliance with Twitter company policy and terms of service.

The system here compares the behaviors of the depressed user class that of the standard and utilize the Bag of Words (BOW) approach to tokenize the models. Then they use the signals derived from the BOW approach to develop, contrast several classifiers and yield a statistical analysis to evaluate the results of each one.

Results: Here authors aim to establish a method by which recognition of depression through analysis of large-scale records of user's linguistic history in social media. The system yields promising results with an 86% classification accuracy.

The ROC curve for various training models are shown in Figure 2.14

A multinomial Naive Bayes algorithm yielded a precision score of 0.82. Other methods used are Decision Trees, Linear Support Vector machine, Logistic Regression, and Naive Bayes algorithm. The results obtained are summarized in the Table 2.13

Classification Algorithm	Accuracy	Precision	Recall
Decision Trees	0.67	0.67	0.68
Linear Support Vector Classifier	0.82	0.83	0.83
Naïve Bayes w/ 2-grams	0.82	0.82	0.82
Logistic Regression	0.82	0.86	0.82
Naïve Bayes w / 1-gram	0.86	0.81	0.82
Ridge Classifier	0.79	0.81	0.79

Table 2.13: Evaluation measures of different classifiers

To quantitatively measure the performance of a classifier's ROC curve, the Area under Curve (AUC) metric is calculated and analyzed. It is found that the

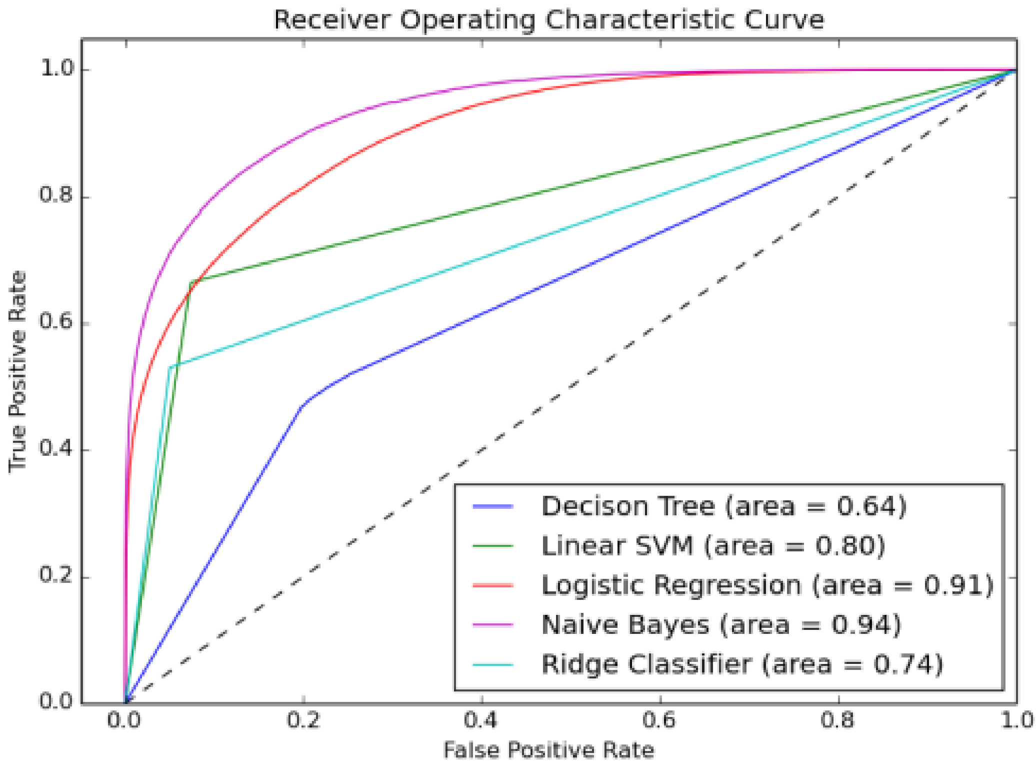


Figure 2.20: Receiver Operating Characteristic (ROC) curves for various training models

Naive Bayes approach scored the highest out of all classifiers with a ROC AUC score of 0.94 (Fig. 2.14). Logistic Regression scored with an AUC score of 0.91, trailing behind Linear SVM with 0.80, Ridge Classifiers with 0.74 and a Decision Tree giving a score of 0.64.

Chapter 3

Design

The proposed methodology is a combination of the best methods specified in the literature papers. The architecture of the system and a brief description of the proposed methodology is given below.

3.1 Proposal

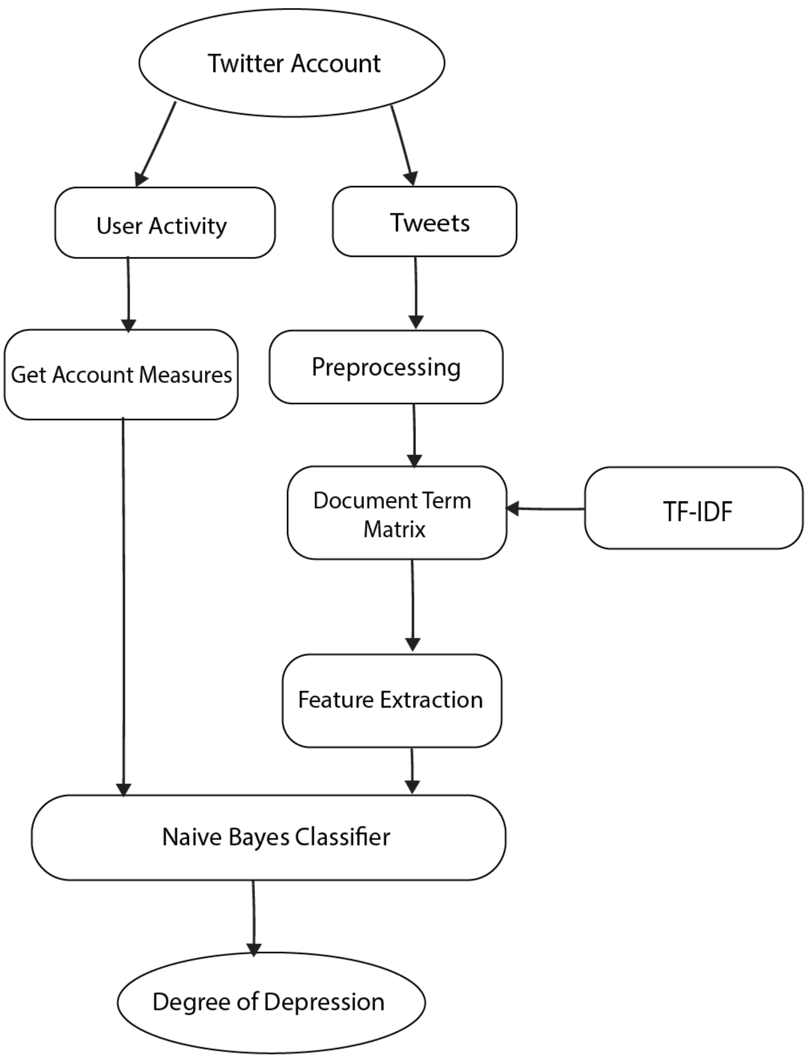


Figure 3.1: System Architecture

Algorithm:

1. Extract tweets and user activity from Twitter using Twitter API.
2. Extract account measures from the user activity.
3. Cleaning of the tweet data.
   - 3.1 Removal of user mentions
   - 3.2 Removal of repeated words
   - 3.3 Removal of punctuations
   - 3.4 Removal of URLs
   - 3.5 Removal of media
4. Preprocessing of the tweets.
   - 4.1 Case folding
   - 4.2 Tokenization
   - 4.3 Removal of stopwords
   - 4.4 Replacing emojis with keywords
   - 4.5 Stemming
5. Creation of Document Term Matrix
6. Features extracted from tweets and account measures are given as input to the classifier.
7. Multi-Class classification using Naive Bayes Classifier

Output: Degree of Depression

All tweets for depressed and non-depressed accounts, also information of user account and activities like number of followers, number of following, total number of posts and time of posts will be retrieved using the Twitter Developer API. The Data preprocessing includes cleaning the collected data, annotating the data, and text normalization. Text normalization includes Case folding, tokenization, stopword removal, stemming, replacing emojis with their emotion keywords. Case folding means converting all words in the dataset to lowercase. Tokenization means giving a token to each words which we want to treat as a group. These tokens can later be counted and can be used a feature. The stop words are words that do not contain any meaning for examples word coming under the category of pronouns, prepositions, conjunctions etc. So removing these reduces the size of the dataset and improves the efficiency while training. Stemming is the process of extracting the base form of a word. Data cleaning includes removing all user mentions, repeated words, punctuations, URLs, pictures and videos.

A Document Term matrix is created for each account, TF-IDF is used to assign weights. Features applied on the DTM are then merged with account measures extracted from the social network and user activities. The results of the merge is given into a Multinomial Naive Bayes classifier to classify the users into three classes- Highly depressed, Depressed and Not Depressed.

Chapter 4

Work Plan

4.1 Timeline and Goals

- **March 2021:** Extract tweets and user activity from Twitter.
- **April 2021:** Pre-processing and cleaning of the data collected.
- **May 2021:** Creation of the dataset.
- **May 2021:** Training of the dataset using the classifier.
- **June 2021:** Performance evaluation of the classifier.

4.2 Resources

- Twitter Developer API
- Natural Language Toolkit module in Python.
- Scikit-learn package in python
- Sklearn package

Chapter 5

Conclusion

In an era of increasing competition and hectic work load, the number of people with depression is increasing day by day. The authors of our IEEE base paper uses a Chinese social networking site called Sina Weibo to detect depression on Weibo users. By analyzing and studying various research papers, we saw the amazing potential of using Twitter as a tool for measuring and predicting major depressive disorder in individuals. Emotion detection is one of the challenging problems in machine learning. Detecting emotion from text is a challenging work and most of the research works have some kind of limitations most importantly, language ambiguity, multiple emotion bearing text, text which does not contain any emotion words etc. Yet we plan to try several approaches to detect emotion from twitter.

Our target is to analyze Twitter tweets from a comparatively newer perspective. We are planning to observe users behavior based on their text (tweets, replies) along with numerical scores like number of tweets,followers, etc. We plan to combine those values with the sentiment and emotion of users on certain topics to measure the effect.

After retrieving tweets, some essential steps are to be performed. Pre-processing is performed on the tweet strings, so that influence of the redundant words and punctuation's can be avoided. It is considered to be primordial task, as it improves the performance of the work. After that, feature extraction such as stemming, Parts Of Speech vector creation and sentiment extraction are performed to calculate sentiment positive and negative scores.

We also plan to use a Bag of Words approach towards quantifying this dataset, and to create a dimensional feature space as our input vector. Most researches focused on the text of the tweets alone whereas we plan to consider text from tweets, replies and emoticons to find the state of mind of users. We think that a multinomial approach to the Naive Bayes' algorithm would deliver optimal accuracy metric combinations.

Bibliography

[1] Y. Ding, X. Chen, Q. Fu and S. Zhong, "A Depression Recognition Method for College Students Using Deep Integrated Support Vector Algorithm," in IEEE Access, vol. 8, pp. 75616-75629, 2020, doi: 10.1109/ACCESS.2020.2987523.

[2] Gao, Qi Abel, Fabian Houben, Geert-Jan Yu, Yong. "A Comparative Study of Users' Microblogging Behavior on Sina Weibo and Twitter". 7379. (2012).

[3] P. Arora and P. Arora, "Mining Twitter Data for Depression Detection," 2019 International Conference on Signal Processing and Communication (ICSC), NOIDA, India, 2019, pp. 186-189.

[4] Sailunaz, Kashfia Alhajj, Reda. "Emotion and Sentiment Analysis from Twitter Text". Journal of Computational Science. 36. 10.1016/j.jocs.2019.05.009. (2019).

[5] F. M. Shah, A. S. Reyadh, A. I. Shaafi, S. Ahmed and F. T. Sithil, "Emotion Detection from Tweets using AIT-2018 Dataset," 2019 5th International Conference on Advances in Electrical Engineering (ICAEE), Dhaka, Bangladesh, 2019, pp. 575-580.

[6] Alsagri, Hatoon Ykhlef, Mourad. "Machine Learning-based Approach for Depression Detection in Twitter Using Content and Activity Features". (2020).

[7] Tsugawa, S., Kikuchi, Y., Kishino, F., Nakajima, K., Itoh, Y., and Ohsaki, H. "Recognizing depression from twitter activity" 3187-3196 (ACM, 2015).

[8] Mustafa, Raza Ashraf, Noman Ahmed, Fahad Ferzund, Javed Shahzad, Basit Gelbukh, Alexander. (2020). "A Multiclass Depression Detection in Social Media Based on Sentiment Analysis".

[9] Nadeem, M. "Identifying depression on Twitter". Preprint at arXiv:1607.07384 (2016).