

# A MODEL TO DETECT SOCIAL NETWORK MENTAL DISORDERS USING AI TECHNIQUES.

<sup>1</sup>Dr.Tummapudi.Subha Mastan Rao,<sup>2</sup>Dr.Vijaya sri kompalli,<sup>3</sup>Udayasri kompalli

<sup>1</sup> Associate Professor, CMR Technical Campus, Hyderabad, Telangana

<sup>2</sup> Associate Professor, Koneru Lakshmaiah Educational Foundation, vaddeswaram, A.P

<sup>3</sup> Associate Professor, NRI institute of Technology ,Agirapalli,A.P

[mastan1061@gmail.com](mailto:mastan1061@gmail.com), [kompallivsri@gmail.com](mailto:kompallivsri@gmail.com), [mudayasri77@gmail.com](mailto:mudayasri77@gmail.com)

**Abstract:**

Nowadays the users of social network are increasing drastically worldwide. This platform become very useful for sharing information, discussing on various issues, Even majority of their active time they are spending on social medias like tweeter, face Book etc. Due to this physical human relations are damaging, and users are addicted to internet and frequent checking of tweeter, Facebook etc, Net compulsion. Recent surveys telling that there is a relation between mental health and social network behaviour. Still it is unclear how this mental illness and social networks are related. In this paper we are going propose a model to find out mental disorders using social network data analysis, in this work we have collected the data from twitter and manually labelled that data into two classes one is depressive and other is normal then then data pre-processing was performed then it is divided into training and testing sets, training data is used to build the model by making naive bayes classifier to learn from the data. Once model is build it was tested with the testing set and obtained results with high Accuracy around 92.3. So usually doctor need to find metal disorders they will fire some questions to the patient based on that doctor detects mental illness but in this model we can able to detect mental disorders without consulting patient based on their social network behaviour analysis.

**1. Introduction**

Social network is nothing but an internet based software application which allows family members, friends, colleagues, clients connect to one other to share information like photos, videos, messages, Documents, Location etc altogether it changes entirely the way we are communicating. In recent years the popularity of social media increases drastically throughout the world. Even it is reaching to common man , uneducated persons also can able to use these social media comfortably. Here information exchange is very fast with in no time any information can reach throughout the world. Various kinds of social media available in current market are shown in Fig.1.



Fig.1 Different Social media in Current Market

This platform is having various advantages like it allows people around the world to connect to one another in real time and can able to share information which intern helps to build individual social capital. Social networks are very much useful for brand expansion. Here they can meet massive customers and they can promote their products either by directly selling over site are sending links of their site where products are available. Social networks providing fun to their users. On an average daily every user spending 1.30 hours here users are enjoying by receiving comments, likes hearts from their friends. Students also using social media to discuss about their subject, Assignments there by increasing their marks percentage due to peer learning.

Some people are shy to talk directly with friends and family members but social media allowing them to meet many people online and shy can be reduced to maximum extent.

Social networks are helping senior citizens to meet their old friends, recent studies telling that social network groups are forming very fast among users whose age is greater than 65.

Politicians also using social networks very effectively for their election campaigns, opinion polling on some social issues, solving grievances in real time using social network.

Along with these many advantages social network also having many disadvantages like due to information overhead social network distracting many people even they are wasting 70 hrs of their valuable time per month by looking into the information available. Sleep cycles of many users also damaging because of this social media addiction. This happening because human brain having tendency to think about others and creates lot of disturbances in users mind, even the blue rays observed by eyes during the usage can damage sleep cycle.

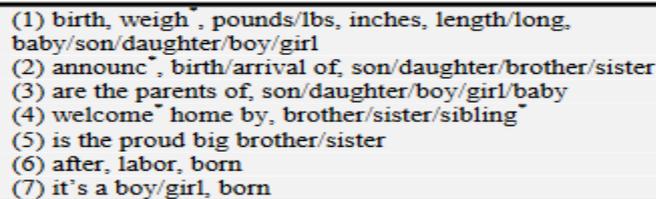
University of Pittsburgh revels in their survey that social network disorder causing depression and like symptoms such as isolating from the social life due to spending too much time, Every minute checking mobile for notifications from various sources like Facebook, tweeter etc, felling worst when there is no internet access, notifying the new location that you have visited for others to know, believing life on social media than reality, creating tweeter account for pets, every food taken posting in tweeter, felling upset when no responded on your post posted in Facebook, social annoy , waking up in the midnight and checking phone for Facebook updates, tweeter updates etc.

## 2. Literature Survey

Choudhury et al [1] in this paper he proposed a model which helps to find out depression in a person by performing analysis on web based social networking posts, behaviour, login log off frequency, comments on other posts, various issues rising by persons and checking for various kinds of anti-depression medicines based on these data analysis his model can able to identify weather person having depression or not. To achieve this they have used support vector machine classifier and PCA for feature selection.

Choudhury et al.[2] contributed good effort to identify postpartum depression from newly born mothers. In this work they are collected data from twitter specifically

from newly born mothers. To identify newly born mothers they have adopted patterns



(1) birth, weigh , pounds/lbs, inches, length/long, baby/son/daughter/boy/girl  
(2) announc , birth/arrival of, son/daughter/brother/sister  
(3) are the parents of, son/daughter/boy/girl/baby  
(4) welcome home by, brother/sister/sibling  
(5) is the proud big brother/sister  
(6) after, labor, born  
(7) it's a boy/girl, born

Once tweets are identified they are labelled as various classes saviour , mild ,Low then data trained to SVM,K-NN classifiers and build the model and results are evaluated.

O' Dea et al.[3] in this paper author build a model to detect suicidal tendency of depressive users by analysing their tweets. In this they have taken 2.5k tweets and used decision tree classifier to build the model. Once model was build it was tested with testing set and evaluated its performance around various parameters like precision, recall, F-Score etc with this model patients can alert in advance to take proactive decisions to come out from life risk events.

Aldarwish et al.[4] was a good researcher who performed analysis on Social network sites for understanding depression levels by using user generated content at different social networks once data is collected from different sources that data is trained to SVM classifier once model is build it is used to perform multi class classification based on the results from content analysis.

Reference	Data Source	Algorithm used	Quality Metrics	Performance
[5]	Twitter	PCA,SVM	Accuracy	72%
[6]	Facebook	Logistic Regression	Pseudo-R2	.36
[7]	Twitter	SVM	Accuracy	.69
[8]	Twitter	Log Linear Classification	Precision	.76
[9]	Twitter	Nueral Networks	AUC	.77
[10]	Twitter(Self Declared)	Naïve Bayes	AUC	.70

Nguyen et al.[11] proposed a model to differentiate depressive users and normal users of the twitter by using statistical procedures ,temperament and understand their behaviour by analysing their social behaviour through various linguistic approaches using LIWS features . In this paper they are also used 132 labelled mood tags .

Park et al. [12][13]in this paper these people interviewed two group of people depressed and non-depressed then they also performed content analysis on the tweets of interviewed users and their friends. To analyse them they used LIWC where emotion of each word and sentiment is identified based on that count estimated depression levels.

Bachrach et al.[14] proposed a model by analysing face book data by considering users and their friends profile, relationship between friends, posts they have posted responses they have given to the posts posted by other users, number of times they have attended to the gatherings, number of times their photos are tagged based on this data they have identified weather user is suffering with depression or not for that they have used SVM machine learning model to perform classification between normal user and depressive user.

Ortigosa et al. [15] they proposed a model to find out sentiment of the user posts posted in face book into three classes like positive, negative, neutral they can able to detect emotional changes using this analysis. To implement this they have used a hybrid approach such as combination of linguistic and machine learning and got an accuracy 83.25.

3. Proposed model:

In this paper we have proposed a model which can able to use data collected from twitter through its API and those tweets are manually labelled as Depression and normal tweets. Credentials used to collect data through twitter API are as follows

Consumer Key (API Key) - 1kfxTGceYiH1CLHcO4wzTkEBz

Consumer Secret (API Secret)- CZ0NFyrEORTgVsYleV3loBC51LAatDO0IEt6JNdt9FxEiMpaOe

Access Token: 724523442758885380-GMYKT7DUEVqOGWfFym9SI1KyDjJg1gO

Access Token Secret: DSUn4Mjr80IWeQpnHmkYElgoDiiTeERzYNnHLGrTVPdbu

The format of the data appears as follows shown in Table.2.

Table.2.The format of the source data

Target	Ids	Date	Flag	User	text
This field consists of polarity includes {positive, Negative, neutral }	The identity number of the tweet	The date when that tweet was posted	It contains query, if no query then it returns NO_Query	It contains the user details who posted that tweet	It contains the text message posted

A snap shot of twitter data set will be as follows in fig 2.

```

RT @TeluguVijayFans: Sarkar Very good hold in Telugu states after initial 2 big days. 4 Days AP/TG Gross 11.25 CR approx & Share 5.90 CR...
RT @GouriGKishan: The Film Which Is Needed For The State ! #Sarkar https://t.co/su6BQjk36r
RT @Irshad5676: #Sarkar beats #Kaala Tamil Nadu gross in 3 days...
#Thalapathy rage 🙌🙌
RT @manickamtagore: Watched #Sarkar with my kids ! Younger son loved it! More politics but it had kids magic which is Mr Vijay USP ! Good s...
RT @bharathi1: Adei kirukka. Idhu Kalaignar TV illada. Appa kashta pattu instalment katti vangi vechadhu da. #Sarkar https://t.co/detvha8TBE
RT @MalayalamReview: Wom effected very badly for #Sarkar - today also below average performance In Kerala Boxoffice

Crucial Weekend*****
RT @rhevanth95: Undisputed King of BOX OFFICE !
Records created within 4 days of release and has surpassed "almost" EVERY movie that rele...
RT @cinema_babu: #Sarkar TN 4 days Gross : 75 Cr+. All Set to breach 100 Cr mark by the end of Sunday . Will be First film in the state to...
RT @GBK_Official: Spi Cinemas All Shows Sold Out From Online Booking Only #SARKAR
Day 5 On A Double Rage 🙌
#SarkarPuratchiArambam https://...
RT @Ajaychairman: #Sarkar TN Gross
    
```

Fig2. A snap shot of tweeter data set

Once data is collected it is undergoing various phases as shown in Fig3.first manually each and every tweet is labelled as depressive and normal then they are sent for data cleaning in that process special symbols, numbers expect a..z,A...Z remaining all the data removed from the source.

Then second step tokenization means dividing the data into various meaning full words but all meaning full words are not useful for decision making only nouns and verbs are useful in most of the cases so in third step stop words are removed such as {is, and, the etc..} by using predefined stop word corpus available in python natural language tool kit NLTK. Then stemming is used to bring every word to its root like {actions -> Action, Enjoying-> Enjoy} finally text is converted into bag of words matrix. This data then given to Naïve bayes classifier around 800 tweets split into training and testing model was build by learning from training set and it was tested using testing and giving 92.3% accuracy.

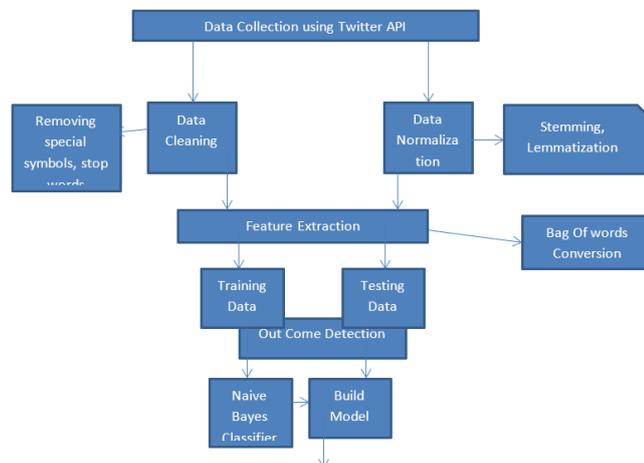


Fig .3. Proposed Architecture Tweet Classification Algorithm:

- 1) List of tweets collected from Twitter API=[T1,T2,...Tn]
- 2) Remove Special Symbol

Tweet= re.Sub[a-z,A-Z]

- 3) Tokenization
- 4) Stem the tweets brining every word to its parent word
- 5) Remove Stop words by Comparing with NLTK stop words
- 6) Convert Bag of Words

Build Naïve bayes Classifier

$$7) \quad p\left(\frac{A}{B}\right) = \frac{P\left(\frac{B}{A}\right)p(A)}{P(B)}$$

Where P(A)- P(A/B) is the posterior probability ,P(A) is the prior probability(B|A) likelihood probability, P(B) prior probability of predictor.

- 8) Predict Using Testing set on model

**4. Results &Discussion:**

Once model was build it was tested with testing set and produced results with very good Accuracy around 92.3% when compared with existing machine learning techniques this model producing good results. Due to adoption of advanced pre-processing and transformation we can able to produce good results with naïve bayes and those comparative analysis with existing algorithms are shown below table.4

$$\text{Accuracy} = \frac{TN+TP}{TN+TP+FP+FN}$$

TP: True positive, TN: True Negative, FP: False Positive, FN: False Negative

S.No	Model	Accuracy
1	Proposed Naïve Bayes	92.3
2	Random Forest	48.52
3	SVM	50
4	KNN	81.46

For better visualization the above results are shown in graph

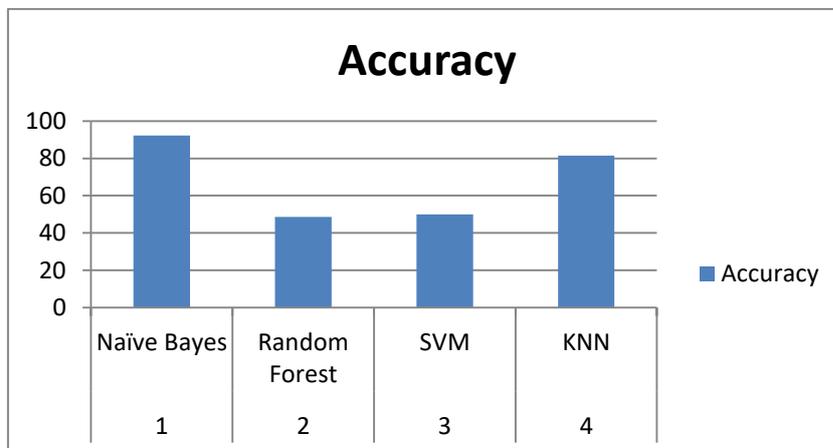


Fig .4 Comparative analysis among various machine learning algorithms

## 5. Conclusion:

In this paper we have conducted Deeper literature survey and implemented a optimized naïve bayes model for analysing tweets to detect weather a specific user is suffering with depression or not we have used tweeter API to collect tweets and are manually labelled as two classes depression and normal using python we have implemented the model and able to achieve good accuracy around 92.3% with testing set and effectively able to detect mental disorders through social media.

## 6. References:

- [1] De Choudhury, Munmun, et al. "Predicting depression via social media." *Seventh international AAAI conference on weblogs and social media*. 2013.
- [2] De Choudhury, Munmun, Scott Counts, and Eric Horvitz. "Predicting postpartum changes in emotion and behavior via social media." *Proceedings of the SIGCHI conference on human factors in computing systems*. ACM, 2013.
- [3]. O'Dea, Bridianne, et al. "Detecting suicidality on Twitter." *Internet Interventions* 2.2 (2015): 183-188.
- [4] Aldarwish, Maryam Mohammed, and Hafiz Farooq Ahmad. "Predicting depression levels using social media posts." *2017 IEEE 13th international Symposium on Autonomous decentralized system (ISADS)*. IEEE, 2017.
- [5] De Choudhury, Munmun, et al. "Predicting depression via social media." *Seventh international AAAI conference on weblogs and social media*. 2013.
- [6] Eric, Munmun De Choudhury Scott Counts, and Horvitz Aaron Hoff. "Characterizing and Predicting Postpartum Depression from Shared Facebook Data." (2014).
- [7] Tsugawa, Sho, et al. "Recognizing depression from twitter activity." *Proceedings of the 33rd annual ACM conference on human factors in computing systems*. ACM, 2015.
- [8] Coppersmith, Glen, Mark Dredze, and Craig Harman. "Quantifying mental health signals in Twitter." *Proceedings of the workshop on computational linguistics and clinical psychology: From linguistic signal to clinical reality*. 2014.
- [9] Benton, Adrian, Margaret Mitchell, and Dirk Hovy. "Multi-task learning for mental health using social media text." *arXiv preprint arXiv:1712.03538* (2017).
- [10] Nadeem, Moin. "Identifying depression on Twitter." *arXiv preprint arXiv:1607.07384* (2016).
- [11] Nguyen, Thin, et al. "Affective and content analysis of online depression communities." *IEEE Transactions on Affective Computing* 5.3 (2014): 217-226.
- [12] Park, Minsu, David W. McDonald, and Meeyoung Cha. "Perception differences between the depressed and non-depressed users in twitter." *Seventh International AAAI Conference on Weblogs and Social Media*. 2013.
- [13] Wee, Jieun, et al. "The influence of depression and personality on social networking." *Computers in Human Behavior* 74 (2017): 45-52.
- [14] Bachrach, Yoram, et al. "Personality and patterns of Facebook usage." *Proceedings of the 4th annual ACM web science conference*. ACM, 2012.
- [15] Ortigosa, Alvaro, José M. Martín, and Rosa M. Carro. "Sentiment analysis in Facebook and its application to e-learning." *Computers in human behavior* 31 (2014): 527-541