iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** IV **Month of publication:** April 2023

DOI: https://doi.org/10.22214/ijraset.2023.50685

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue IV Apr 2023- Available at www.ijraset.com

Social Network Mental Disorders Detection via Online Social Media Mining

Mrs. V. Mythily[1], Hariharan S[2], Kavin U[3], Kowsik K[4], Sakthivel S[5]
[1]*Assistant Professor,* [2, 3, 4, 5]*UG Scholar, Department of Computer Science and Engineering, Nandha Engineering College (Autonomous), Erode, Tamilnadu, India*

Abstract: ***The rapid growth in popularity of the social networking results in problematicusage. An increasing number of social network mental disorders (SNMDs), like Information Overload, Cyber-Relationship Addiction, and Net Compulsion, are recently mentioned. Mental disorders symptoms are normally observed today passively, which are resulting in delayed clinical process of intervention. In this study, users alleged that mining online social behavior yields an opportunity for active identification of SNMDs in an early stage. It is challenging for finding SNMDs since mental status cannot be directly observed from social activity logs data. This approach, novel and innovative to practice the SNMD detection, is not relying on self-revealing of the mental factors using questionnaires in the Psychology. Depression detection taken from tweets records are processed out after proper preprocessing steps. Moreover, to address the above said problems, this project proposed a new depression detection classification model which use K-Nearest Neighbor and a Support Vector Machine based classification. Python 3.7 is used as the language for project development.***

Keywords: Social Media, Support Vector Machine, K-Nearest Neighbor, Mental Disorder.

I. INTRODUCTION

Everybody have experimented sadness at different times. Still, depression is very different and unavoidable in this scenario. Depression is a psychiatric disorder that needs to be coping with the medication.

A. Definition of Depression

According to the World in Data -Website, Depressive disorders occur with varying severity. The WHO's International Classification of Diseases (ICD-10) denotes this as set of disorders ranging as mild to moderate (to severe). The Institute for Health Metrics and Evaluation (IHME) adopt the definitions by disaggregating to i) mild, ii) persistent depression (dysthymia) and iii) major depressive disorder (severe). All forms of the depressive disorder experience some of the below symptoms:

1) Reduction of concentration and attention
2) Reduction of self-esteem and self-confidence
3) Ideas of guilt/unworthiness (even in a mild type of episode)
4) Bleak/pessimistic views of the future
5) Ideas or acts of self-harm/suicide
6) Sleep disturbance
7) Appetite diminish

Depression is one of the leading causes of disability worldwide. Almost 70% of people with mental disorders remain untreated in the developing countries with almost one million people taking their lives every year. Moreover, according to the World Health Organization (WHO), one in thirteen globally suffers from anxiety. The WHO reports that the anxiety disorders are the most common mental disorders around the world with specific phobia, major depressive disorders and social phobia are the most common anxiety disorders.

B. Depression on Social Media

Social media platforms become an integral part of the people's life. They reflect our personal life. People want to share joy, happiness and sadness on the social media. These platforms are used for the researchers for identifying the causes of depression. Twitter knows when you're suffered/ depressed and possibility of creating an artificial intelligence model which can scan the Twitter feeds and tell if you're at risk of depression or receive notices from the third parties, for instance, warning you that you want to seek help, just based on an automated scan of the tweets. The people who are suffering from the depression mostly wants to be isolated from others, which makes them hope for finding a solution in By sharing the feelings and illness believes that would be able to yield relief and able to solve their worries and anxieties. Instead, it is being found that this social communication and sharing increases the worries and anxieties. Amongst many other social network sites Facebook is one of the fastest growing.

Around the world, more than two billion users have an account in social network web sites. And over two seventy million Facebook members are from The India, which makes top country Facebook audience size. Maximum users are used as a floor in socializing and presenting which making Facebook as the most online access web site. Social network site (SNS) made bold and high attention in the last years since they were launched. Communication using social network sites becomes convenient for interacting and accessing with the aid of Internet. Maximum people are surfing to obtain information and depend on social network site to retrieve recent news and different opinion on various subject matters containing sharing, feelings, subjectivity, assessments, evaluation approaches, observations, influences, point of views and ideas using different kind of materials and marketing.

C. Social Network Sites

YouTube, Snapchat and Instagram all are behind Facebook in day to day accessing rate as well as popularity. J. Clement identified and found that mostly between 25 to 40 years of age are interacting with Facebook. It is shown that monthly activities rate of Facebook depicts how often people access Facebook. Due to numerous causes, this study has proposed scrutinization of the user's true manner and behavors to bring their differences/ level of depression from the SNS users. Results accumulated that 945 data of participants (which 780 Facebook Members, 165 Non-Facebook members) showed that Facebook Members got significantly higher on the narrow mindedness, selfishness, no-respect for elder people, the quality of being outgoing/ socially confident, stress depression than Non-Facebook member.

D. Depression

Depression is one of the major issues that rises day today. Worldwide mental health problem is primarily which comes from depression. Every year, millions of people newly getting depression problems, few of them are seriously taking quality treatment. Depression drives the user's inability in participation of their works enthusiastically and found less enjoyment in daily activities. A person who is having this mental illness is timid as well as shy to expose themselves as well. Depression is becoming a bing problem in both individual and community health. One of the finest of the problem's solution is to research the personal's actions more precisely in details. These actions information are to be retrieved from their daily activities through accessing of social networking. Social networking is the good platform to find the information of person towards their; feeling, behavior, emotional, point of views etc. the users emotional state can be a various state of level.

E. Social Media

Furthermore, the Facebook members seem to have higher numbers of the depression symptoms than the Non-Facebook Members. Amongst the Facebook members, majority are college students. The study found the users who were diagnosed with mental illness desire are isolating themselves from others to seek Social media to comfort their feelings by sharing their thought. There are more ways to detect sentiment in tweets. Detecting earlier depression may be a huge step for addressing the mental illness and offer support to people suffering from terrible mental illness. Creating a model for detecting depression in tweets In Machine Learning, there are more ways for sentiment analysis such: Bayesian classifiers, decision-based systems, sample-based methods, support vector machine and neural networks. After reading some papers about different Machine Learning and artificial intelligence techniques for detection of depression on Social Media, these authors tried to apply sentiment analysis through the powerful theorem from probability theory called Bayes Theorem. The model is written in python and it tells whether the given tweet is depressive or not. This project uses a) K-Nearest Neighbor and b) Support Vector Machine based classification. This project proposed a new depression detection model based on the collected tweets.

II. RELATED WORKS

A. Title: Psychological Stress Detection From Cross Media Microblog Data Using Deep Sparse Neural Network

In this paper [1] Huijie Lin has proposed in this paper Long-term stress lead to severe physical as well as mental problems. Traditional psychological stress detection relies usually on the active individual participation, which detect labour-consuming, time-costing and hysteretic. With fast development of social networks, people become more willing in sharing moods using microblog platforms. In this paper, authors proposed a novel automatic stress detection method in the cross-media microblog data. They constructed three-level framework to formulate problems. They obtained first a set of low-level features from the tweets. Then defined/ extracted middle-level representations based upon the psychological and art theories: linguistic attributes using the tweets' texts, visual attributes using tweets' images, and social attributes from tweets' comments, retweets as well as favourites. Then, a Deep Neural Network was designed to learn the stress categories incorporating cross-media attribute.

Experiment results showed that new method is effective and efficient in detecting the psychological stress from micro blog data. Psychological stress of the people is severely rising. Nowadays, most of the people are feeling increasingly stressed under rapid pace of the modern life. According to worldwide survey conducted by Regus Business Tracker in 2002, over half of the business population (53.9%) has experienced appreciable rise in stress over last two years. Though stress is the positive aspect in daily life, excessive stress amount is rather harmful to the physical/ mental health. Long-term stress leads to severe physical/mental problems, such as clinical depressions, insomnia and evensuicide. According to China's Centre for Disease Control and Prevention, suicide is the top reason of death among youth, and more stress is considered to be major suicide factor. All these reveal that rapid growth of stress becomes a great challenge to the health and quality of life. So, there is significant importance to detect stress before it turns in to a severe problem. While traditional psychological stress detection is based mainly in face to face interviews conducted by psychologists, it is usually labor-consuming, time-costing, and hysteretic. In this paper [1], authors presented the three-level framework to detect stress fromcross-media micro blog data. By combining a Deep Neural Network to incorporate various features from cross-media micro blog data, this framework is feasible and also efficient in stress detection. Using the framework, new method helps in automatic detection of psychological stress with social networks.In future work, they planned to investigate the social correlations in psychological stress to improve further the detection performance.

B. Title: Daily Stress Recognition From Mobile Phone Data, Weather Conditions And Individual Traits

In this paper [2] Andrey bogomolov have proposed researches which proved that stress reduces quality of life and causes many diseases. For this reason, researchers are devising stress detection systems with physiological parameters. Still, these systems require the obtrusive sensors are carried continuously by the users. In this paper, they proposed alternative approach which provide evidence that daily stress are reliably recognized based upon behavioral metrics, derived from the user's mobile phone usage activity and then with additional indicators, such as weather conditions (data pertaining to transitory properties of environment) and then personality traits (data concerning permanent dispositions among individuals). Their multi-factorial statistical-models, which areperson independent, obtains accuracy score of 72.3% for a two-class daily stress recognition problems. The model is efficient enough for applying most of the multimedia applications because of the highly minimized low dimensional feature space (32d). For sake of transitory individual property, mobile usage patterns have important advantage among alternative methods: they are not only minimal unobtrusive but also increase limited privacy problems when compared to, e.g., voice analysis/the exploitation of the data from physiological sensors. Moreover and most importantly, automatic stress detection with use of mobile data takes advantage of the extensive usage diffusion of these devices, and it is applied in real world scenarios and exploited for variety of applications that delivered using same device. For example, applications used to inform design of clinical decision support systems or self monitoring applications of stress levels in work settings and in other routine daily life.

III. METHODOLOGY

In this project, users argue that mining online social behavior provides an opportunity to actively identify SNMDs at an early stage. It is challenging to detect SNMDs because the mental status cannot be directly observed from online social activity logs. TF-IDF based classification is being carried out here with the collected tweet data set records. Accuracy score is found out after classification for given training and testing data records. This approach, new and innovative to the practice of SNMD detection, does not rely on self-revealing of those mental factors via questionnaires in Psychology. Depression detection taken from tweets records are carried out after proper preprocessing steps.

1) Processing time will be more if the dataset contains more records.
2) Removal of Unicode characters are not carried out.
3) TF-IDF values are prepared for sentences with Unicode values also.
4) Accuracy score is not much higher for given dataset records.

All the existing system approaches are carried out in proposed system. Removal of Unicode characters are carried out. To address the issue, this project proposed a depression detection classification model using K-Nearest Neighbor and a Support Vector Machine based classification. Accuracy score and confusion matrix is prepared for both KNN and SVM based classification and found to be more that existing system. The following modules are present in the project.

- Dataset collection
- Find depression based on TF-IDF
- SVM classification
- KNN classification
- SVM/KNN classification

A. Dataset Collection

In this module, the tweet dataset from twitter which contains attributes (message, label (1/0) are taken. Null value records are eliminated during preprocessing work. Unicode characters are removed from the paragraphs. The dataset download link (https://github.com/viritaromero/Detecting-Depression-in-Tweets).

B. FIND Depression Based on TF-IDF

In this module, sentiment tweet CSV (comma separated values) records with message column and label column are taken. The text contents are split into words and term frequency – inverse document frequency is calculated. Then based on the words (mostly occurred) in both training, the classification is being done on testing records. Then the model is trained with training data and then predicted with test data. Of which, most of the records are classified as depression or not with accuracy score.

C. SVM Classification

In this module, 80% of the data in given data set is taken as training data and 20% of the data is taken as test data. The text (categorical) columns are converted into numerical values. Then the model is trained with training data and then predicted with test data. Of which, most of the records are classified as depression or not with accuracy score and confusion matrix.

D. KNN Classification

In this module, 80% of the data in given data set is taken as training data and 20% of the data is taken as test data. The text (categorical) columns are converted into numerical values. Then the model is trained with training data and then predicted with test data. Of which, most of the records are classified as depression or not with accuracy score and confusion matrix.

IV. EXPERIMENT RESULTS AND FINDINGS

The proposed system has following advantages.

5) Processing time will be less even if the dataset contains more records.
6) Removal of Unicode characters are carried out.
7) KNN classification is carried out to improve accuracy.
8) SVM will perform better in accuracy terms.

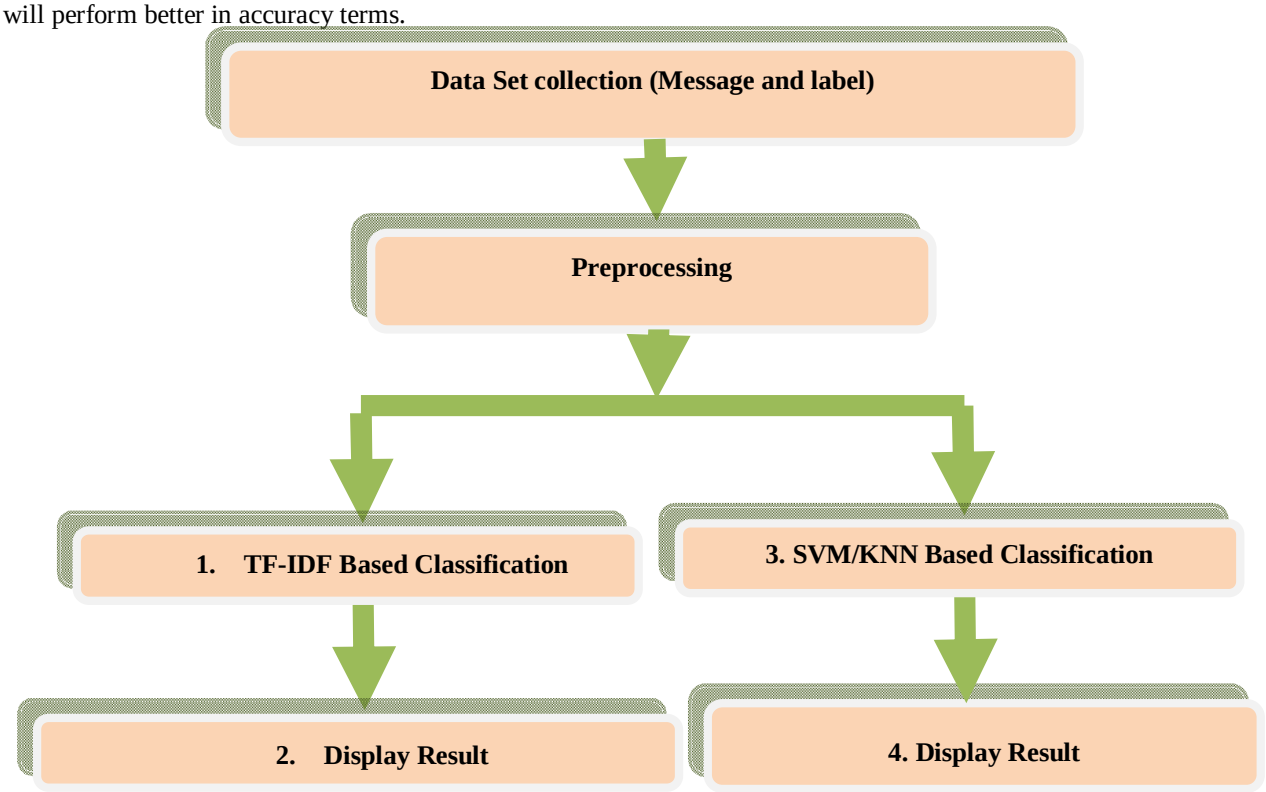


FIG 4.1 architecture Diagram

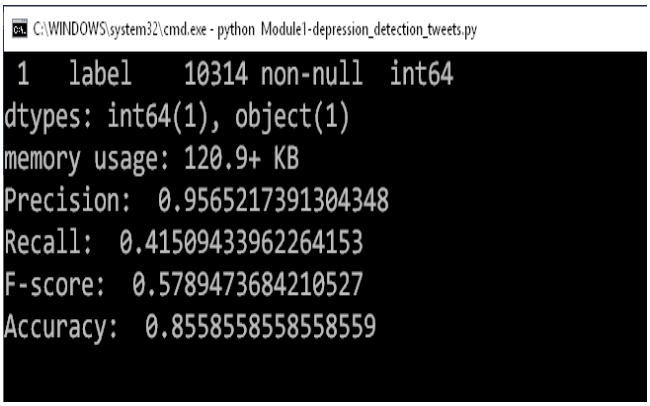


FIG 4.2 TF-IDF ACCURACY SCORE (1)

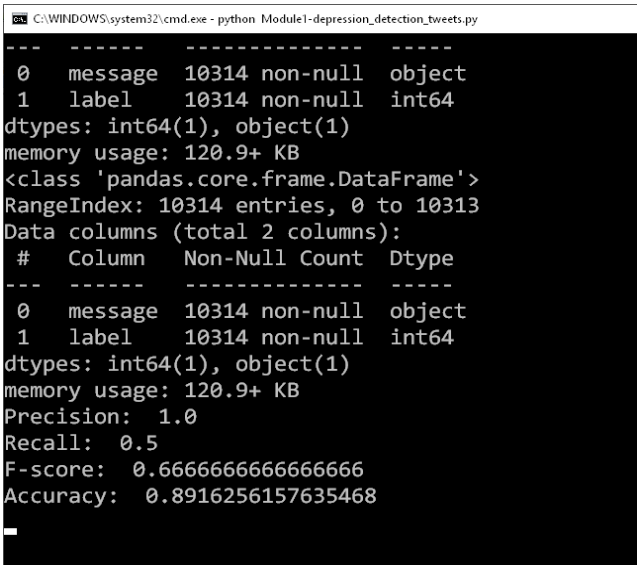


FIG 4.3 ACCURACY SCORE (2)

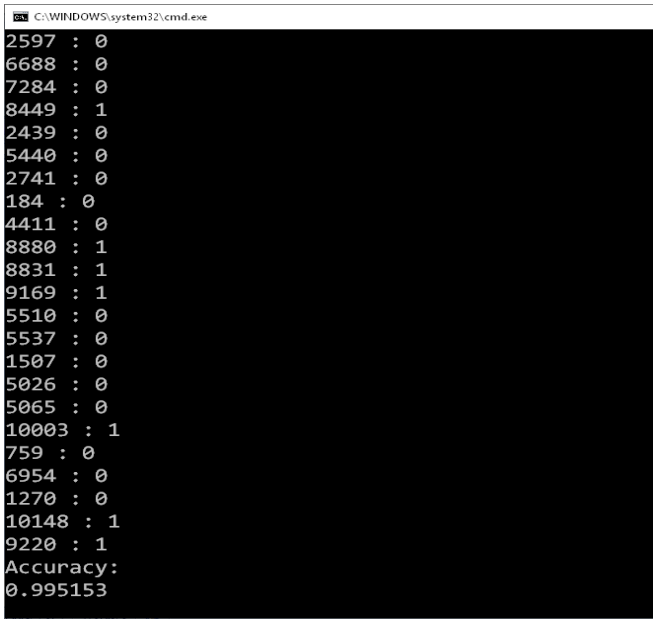


FIG 4.4 SVM ACCURACY SCORE

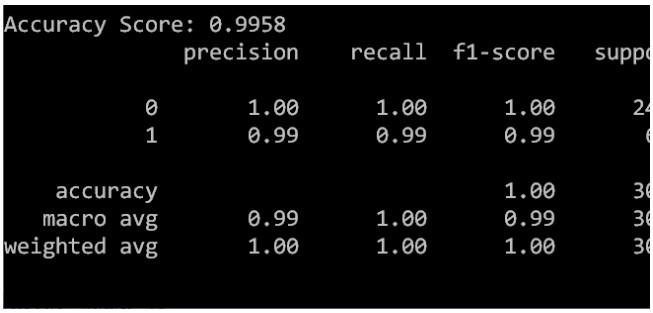


FIG 4.5 KNN ACCURACY SCORE

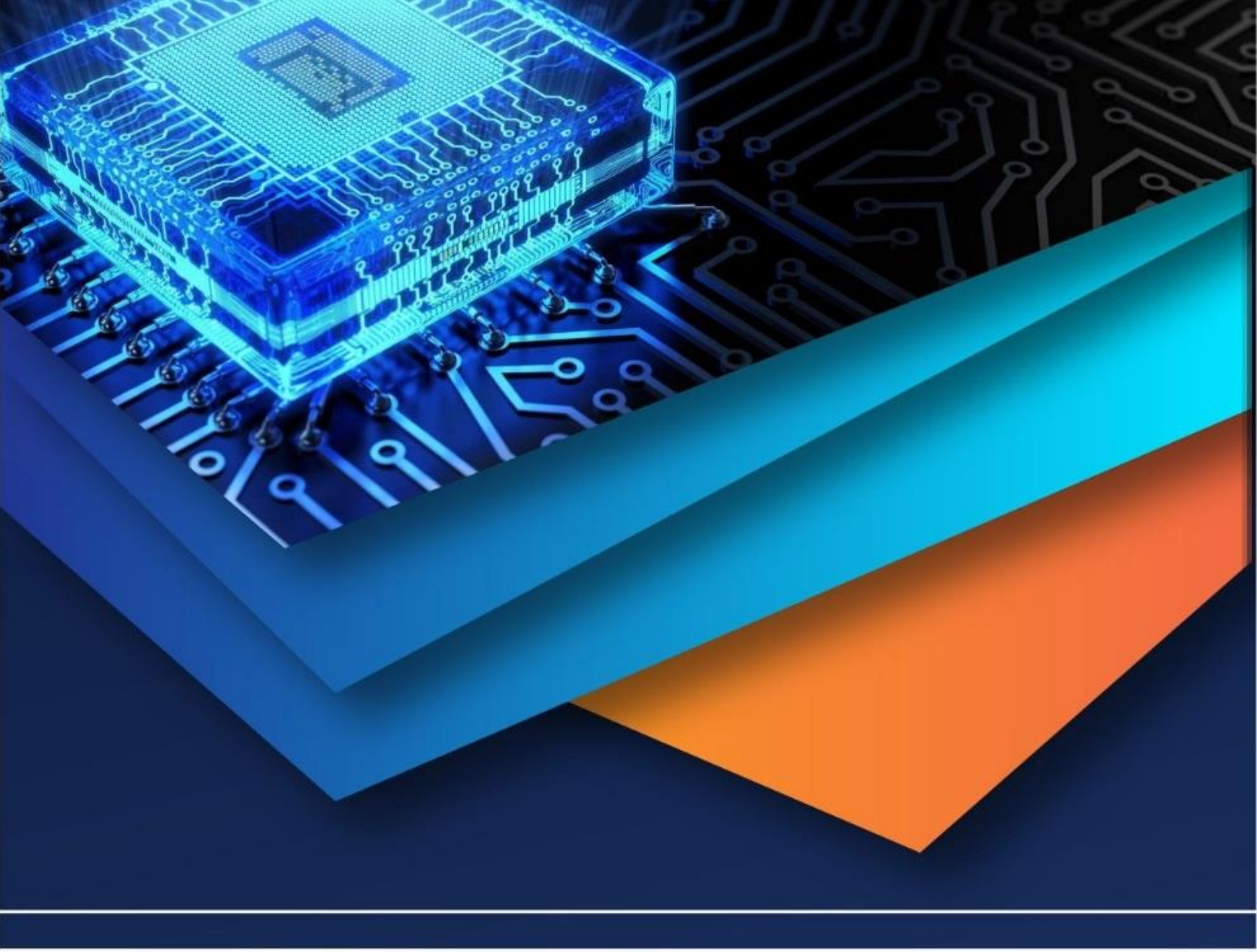
V. CONCLUSION

In this project term frequency inverse document frequency based classification for depression detection is carried out in the existing system. Removal of Unicode characters are carried out in the proposed system for new classification methods. Accuracy score as well as confusion matrix along with depression category to test records are found out and displayed. To address classification issue, a) K-Nearest Neighbor and b) Support Vector Machine based classification is used here. This project proposed the depression detection model based on tweets collected with TF-IDF, SVM and KNN methods. KNN and SVM classification is used in proposed system for yielding better classification results. The project is designed using Python Language 3.7.

REFERENCES

[1] T. Huijie Lin, "Psychological stress detection from cross-media microblog data using deep sparse neural network, conference: 2014 ieee international conference on multimedia and expo (icme).
[2] Andrey Bogomolov , "Daily Stress Recognition from Mobile Phone Data, Weather Conditions and Individual Traits" October 2014. http://dx.doi.org/10.1145/2647868.2654933.
[3] G. Bauer and P. Lukowicz. Can smartphones detect stress-related changes in the behaviour of individuals? In Pervasive Computing and Communications Workshops (PERCOM Workshops), 2012 IEEE. International Conference on, pages 423–426. IEEE, 2012.
[4] A. Bogomolov, B. Lepri, and F. Pianesi. Happiness recognition from mobile phone data. Proceedings of the 2013 International Conference on Social Computing (SocialCom 2013), pages 790–795, 2013.
[5] N. Bolger and E. A. Schilling. Personality and the problems of everyday life: The role of neuroticism in exposure and reactivity to daily stressors. Journal of Personality, 59(3):355–386, 1991.
[6] G. E. P. Box and D. R. Cox. An Analysis of Transformations. Journal of the Royal Statistical Society. Series B (Methodological), 26(2):211–252, 1964.
[7] L. Breiman. Random forests. Mach. Learn., 45(1):5–32, Oct. 2001.
[8] G. Chittaranjan, J. Blom, and D. Gatica-Perez. Mining large-scale smartphone data for personality studies. Personal Ubiquitous Comput., 17(3):433–450, Mar. 2013.
[9] J. Cohen. A Coefficient of Agreement for Nominal Scales. Educational and Psychological Measurement, 20(1):37–46, Apr. 1960.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)