Interdisciplinary Journal of Agriculture and Environmental Sciences
ISSN: 2837-9977 | Impact Factor: 6.52
Volume. 10, Number 1; January-March, 2023;
Published By: Scientific and Academic Development Institute (SADI)
8933 Willis Ave Los Angeles, California
https://sadipub.com/Journals/index.php/ijaes| editorial@sadipub.com

PREDICTING MENTAL HEALTH CONDITIONS THROUGH SOCIAL MEDIA ACTIVITY: DATA MINING TECHNIQUES

Olawale Joseph
Department of Computer Sciences, University of Lagos, Akoka, Lagos.
Department of Information and Communication Technology, Osun State University, Osogbo

Abstract: The use of social media platforms to gather real-time information has proven to be an effective tool for identifying various mental health problems early on. This study aimed to classify mental health conditions using data mining techniques by analyzing tweets related to mental health disorders retrieved with various keywords and pre-processed with sentiment analysis and natural language processing techniques. Using different data mining algorithms, including decision tree and random forest classifiers, the developed models accurately predicted mental health conditions such as depression, anxiety disorders, schizophrenia, drug abuse, and seasonal emotional disorders. By tracking social media activities in real-time, the developed models could help monitor mental health and recognize mental health problems early on. The potential benefits of data mining strategies in healthcare include personalized treatment options, evidence-based decision making, and new discoveries. Overall, the study highlights the potential of data mining techniques in identifying mental health conditions early and improving mental health care.

Keywords: Data Mining, Mental Health, social media, Twitter, Classification Models, Sentiment Analysis, Natural Language Processing, Decision Tree, Random Forest.

1.0 Introduction

The World Health Organization (WHO) defines mental health as a state of well-being in which the person knows how to deal effectively, fruitfully, and productively with the regular stresses of work and life, and is capable of making a difference to his or her immediate environment. Mental illness encompasses both physical, abnormal behavior, thinking, and feeling mental and psychological conditions [1]. A high level of impairment, for example, an affective disorder that ends in variations in anxiety disorders, dejection, or despair, and depression can be used to measure mental health.

Globally, 25% of the world's population suffers both in developing countries and in developed countries from mental health issues. Based on research carried out by the Global Mental Health Initiative Grand Challenges, probably the biggest challenge to worldwide mental health care is, actually the absence of an evidence-based set of preliminary preventive *Received Dec 24,* approaches and strategies [2], to be able to bridge this particular gap, the techniques of data mining are being employed in the mental health domain.

The data mining process can be considered as a group of tasks that is based on the automated process of examination, and the search for practical information and knowledge buried within the massive and voluminous quantity of data (big data) to draw out useful information for the objective of producing models useful in decision making and latest discoveries. The foremost objectives of data mining are descriptions and predictions [3], [4]. With data mining techniques, it provides ways of acquiring unbiased unseen patterns from evidence-based information, particularly within the public healthcare sector [5]. Data mining comes with great potential to allow healthcare systems to gain useful information from data and use such information more effectively and efficiently, therefore, cutting back on the probable expenses associated with decision making. Data mining strategies and methods are invaluable in the healthcare domain, the techniques of data mining have the power to positively change the way patients are treated, and as well assist us to advance knowledge much more rapidly [6], [7]. Patients can get highly personalized treatment options, therapists are going to receive assistance in making evidence-based choices, as well as scientists, will have the ability to explore new information and understanding that reveals the real reasons for Mental Health problems while creating much more effective treatment approaches [8].

The data to be mined is in the volume of petabytes and terabytes, eighty percent of which is unstructured, therefore it's tough to handle them, and process with database managing tools along with other traditional methods. The cost for Mental Health treatment globally is over two trillion dollars [9]. By improving, enhancing, filtering, and refining the quality of the treatment, we will greatly lower costs, by applying data-mining methods and tools in mental health detection and treatment, this particular cost can be greatly reduced [10], [11].Psychologists or experts in public health will benefit from using this model to filter out relevant data, and then to study their symptoms and progression patterns before this turns into a serious problem [12]. Most of the studies that have been conducted in the area of mental health uses stored data from mental health patients, the data is small with little information to mine from it. Using twitter comments is another approach to gather enough relevant data from people showing signs of mental health illness and to monitor such persons before it escalates.

Therefore, there is a need to develop models for mental health classification, and for the associated risks of mental illness based on information concerning similar risk factors, hence this particular study.

2.0 Related Works

Several kinds of researches have been conducted in the area of mental health prediction from the survey data, data repositories and databases, social media, and other online platforms where people share their views, opinions, and thoughts. Here are the reviews of some of their findings.

Reference [13] predicted the generalized anxiety disorders happening among women. The researchers used the Random Forest technique and found that results were over 90% accurate. This research concluded that the Random Forest approach is useful for the prediction of General Anxiety Disorders. The consistency was seen for the evaluation of specificity as Random Forest predicted accurately about the people who did not have General Anxiety Disorders.

Li et al. [14] researched electroencephalogram (EEG). In this study, the investigators explored the influential and effective frequencies and brain regions that have a greater impact on moderate depressions. The authors used Bayes Net, Support Vector Machine(SVM), KNearest Neighbor(KNN), Random Forest(RF), Logistic Regression(LR), Best First Search, Linear Forward Regression, Greedy Stepwise (GSW), and Rank Search techniques. These all were applied based on the Correlation Feature Selection (CFS) to guess the future

predictions. This research concluded that the GSW based upon the CFS and KNN provided optimal performances. On the other hand, the beta frequencies proved to be more efficient to detect mild depressions. The alpha and theta frequency bands also provided accuracies of 92% and AUC above 0.950 for beta frequency bands of Emo_block and Neu_block. In the same way, [15]worked on the development of a predictive model for the prediction of depression among senior citizens of India using machine learning classifiers.Data was collected at Bagbazar, Kolkata, service region of Bagbazar Urban Health and Training Centre (UHTC). The data was collected from 1st April 2016 through 30th April 2016 through sixty elderly people applying the Geriatric Depression Scale (GDS) to gather training data. Five Classification algorithms had been compared about four metrics of Accuracy, ROC area, Precision as well as Root Mean Square Error (RMSE). The Machine Learning (LR),
Multilayer Perceptron (MLP), Support Vector Machines (SVM) as well as Decision Trees (DT). The outcomes exhibited SVM had the highest accuracy while the Naïve Bayes (NB) had Receiver Operating Characteristics (ROC) and lower RMSE.

In their research, [16] researched people living alone to monitor their daily living activities. The proposed design was discreet and simple and it used a detection system using passive infrared sensors. The technique used was Neural Networks, Decision Tree (C4.5 DT), Bayesian Networks, and Support Vector Machine (SVM). According to the results of this research, Neural networks beat the other algorithms. After this C4.5 DT worked well and is effective to detect normal conditions and mild depressions with up to 96% accuracy.

[1] developed a model for the prediction of risks of mental health illness in Nigeria using
Naïve Bayes' and the Decision Trees' Classifiers**.** Data were collected from thirty individuals with a nearly equal division of no, minimal, high and moderate risk of mental illness instances. The results showed that there were three types of mental illness-related risk factors: physical, behavioral, and psychological. The findings also revealed that the model for Decision Trees Classifiers probably found the most relevant factors, such as the lack of close relationships, in mental illness.

In recent years, there have been researches on the sentimental analysis of the data obtained from social network websites. [17] executed an analysis of Twitter data for knowing the sentiments of people using Twitter. They used a python programmed model by using a natural language processing technique. The textual analysis of the Twitter data proved to provide sentimental analysis of the Twitter users, they classified the users in negative and positive classes. Natural language processing methods were useful in word processing and the grades for the behavior of emotions analyzes were classified as positive and negative. In the same way, [18] used the technique of short comments analysis for predicting mental illness. The classes were defined to execute the analysis. A class of known mentally ill persons and a class of healthy people were used for the analysis of their postings on social network websites. More than 90% accuracy was seen to conduct the analysis. Two classes of 12,000 comments were used. 12,000 for negative and 12,000 for the positive class to detect the mental illness. According to this research, the statements posted everyday can help in the early detection of mental illness. This was the application of sentimental analysis to predict mental health diseases in social network website users.

However, most existing works focused on few stored data collected from patients, some use Twitter data using binary classification while this work is based on scrapping Twitter data in real-time which further classified text into 3 categories of positive comments; which shows signs of mental health problem, negative comment; which shows sign of no mental health issue and neutral comment; which means the tweet does not contain using enough information to be classified as either a positive or negative tweet using multi-class text classification.

3.0 Methodology 3.1 Architecture

Figure 1 below shows an overview of the system architecture and different stages in achieving the objectives of this work.

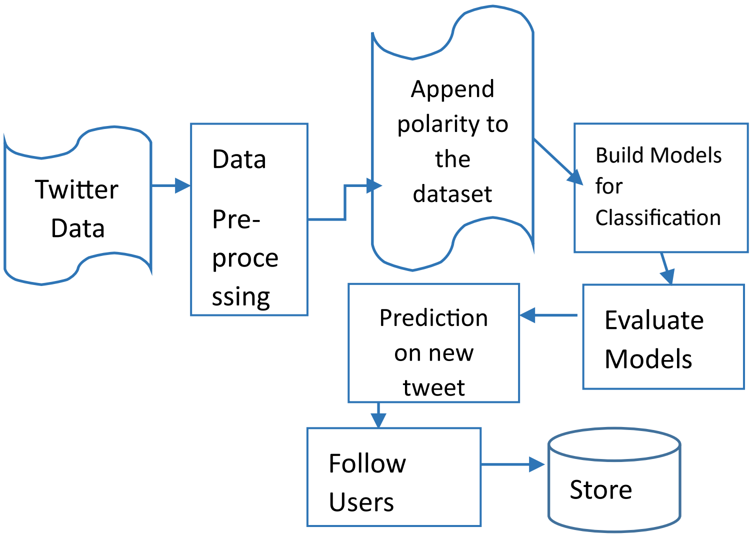


Figure 1: Basic Architecture of the System

3.2 Data Mining Classification Techniques

In a classification task, the label or the target variable in a labeling function retains the categorical (i.e. the target variable is discrete). The target variable of the categorical values is normally set. The mapping method then tries to determine one of the groups in which the goal variable is dependent on input variables, a classification algorithm will attempt to determine what the mental health condition of the person is, for example, given a set of input variables [19], [20], [21]. In this case, the mental health state is a label and has three possible values positive, negative, or neutral meanings. The problem can be classified as a conditional labeling function if the symbol only includes two values: positive or negative, yes or no, or 1 or 0. The tag can also include multiple class attributes that could be referred to as a multiclass classification problem, an algorithm, for example, will try to determine, with a collection of input characteristics, if the shape itself is square or circle, or triangle.In this work, four classification algorithms were used namely; Decision Trees, Random Forest, Naïve Bayes, and K-Nearest Neighbor.

3.3 Evaluation Metrics

Execution Time: This is the total time spent by the processor to execute a code. We run the code five times to get the average execution time.

Accuracy: For the test results, the percentage of correct prediction is the accuracy. This is determined by dividing the sum of accurate predictions by the sum of total predictions. This can be easily determined as stated below.

$$accuracy = \frac{correct^{prediction}}{all\_prediction} \quad \ldots\ldots\ldots\ldots\ldots\ldots (1)$$

Precision: Precision is the fraction of relevant examples (true positives) in all the predicted examples which belong in a certain class.

$$Precision = \frac{true-positive}{(true\_positve+false\_positive)} \quad \ldots\ldots.(2)$$

Recall: For samples of the same class, recall can be defined as the fraction of these samples that were predicted correctly to belong to this class.

$$Recall = \frac{true^{positive}}{(true\_positive + false\_negative} \ldots\ldots\ldots\ldots (3)$$

F-measure:

(2 *x Precision x Recall*) / (*Precision* + *Recall*) …. (4) to provide additional analysis for each model.

Confusion Matrix: The measurement criteria for classification systems are greatly varied. For the question area, metrics used for a research review must be sufficient. The output of a classification system, based on test data for which positive (i.e. true) values are known is represented via a confusion matrix. For classifying positive, neutral, and negative cases, a confusion matrix is used. **4.0 Implementation and Result**

4.1 Dataset Collection

One of the fastest-growing micro-blogging sites on the webspace where millions of people express their views and sometimes deep feelings publicly is Twitter. People send their views inform of tweets, which are short text messages, with a maximum of two hundred and eighty (280) characters in length. There are two ways, through Twitter Application Programming Interface (API) by which data can be extracted from Twitter. This data extraction can either be through Streaming API or Representational State Transfer (REST) API. Over fourteen thousand tweets (14000) were gathered in real-time through streaming API for analysis in this work based on mental health-related hashtags. To start the streaming process for data gathering and extraction, "Initial Query" is issued using Twitter API. The process of data extraction from Twitter using streaming API is initiated using definite keywords and access keys. Corpus is different tweets stored in JavaScript Object Notation (JSON) format [22]. More than 10 different keywords and phrases thatare related to this study were used to extract the data. Keywords such as: 'anxiety', 'depression', 'mental health', 'suicide', 'schizophrenia', 'bipolar', 'stress', 'sad', 'antidepressants', 'pills for depression, etc.

4.2 Data Pre-Processing and Transformation

The JSON format was available for extracted tweets. A few JSON-format items were filtered, like Tweet Text, Date& Time, and Place and Username. Hashtags (#) and @ characters have been removed, followed by RT@. Special characters have also been excluded from regular expressions. The HTTP:// and web address below have also been omitted from the file. Then all the tweets have been turned into letters. Blank spaces have replaced the following symbols/text:

i. Clean certain text that starts from http[^\\s]+ ii. Clean hashtag #, @, RT@ iii.Clean Unicode characters /([\ud800-\udbff][\udc00-\udfff])|./g iv. Change upper case to lowercase v. Get rid of hyperlinks

The transformation of data is the mechanism by which the data is transformed into the system's correct format. Data processing typically involves transforming source data into the desired format suitable enough to run it through different algorithms. In this scenario, the data obtained will be after the data cleaning stage / pre-processing steps, and this will be the data to be sent to the data mining algorithm.

4.3 Classification of Tweets

Pre-Processed tweets are assigned sentiment with a different polarity which could either be positive, negative, or neutral. Polarity score is calculated by the summation of each word of the text present in the dictionary Three different classes of tweets have been listed. In this study, classification is an important step. The classes are as follows:

i. Positive: Probably with symptoms and signs of any mental disorder we are looking at.
ii. Negative: Mental health possibly not affected.

iii. Neutral: Those who do not think about themselves and who generally speak about mental disorders to raise awareness or chat to relatives/friends or the classifier cannot rightly classify.

4.4 Models Implementation and Results

The dataset was divided into a train (75%) and a test set (25%) to train the algorithms. Models were later evaluated by infusing the test dataset into the models to test their performances. The experimental outcomes for the four classifiers are as presented below.

4.4.1 Naïve Bayes Classifier

Table 1: Prediction report for Naïve Bayes Classifier

Label	precision	recall	f1-score
Negative (-1)	0.88	0.86	0.87
Neutral (0)	0.86	0.89	0.87
Positive (1)	0.88	0.84	0.86
weighted average	0.87	0.87	0.87

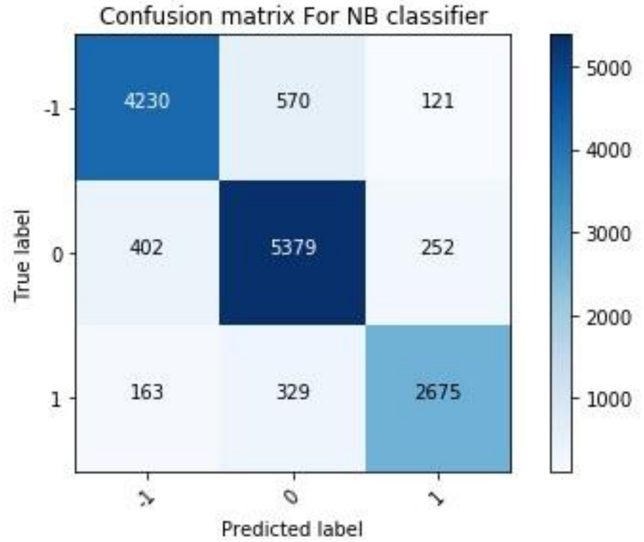


Figure 2: Confusion Matrix for Naïve Bayes Classifier

Negative (-1) True Positive: 4230

Neutral (0) True Positive: 5379

Positive (1) True Positive: 2675

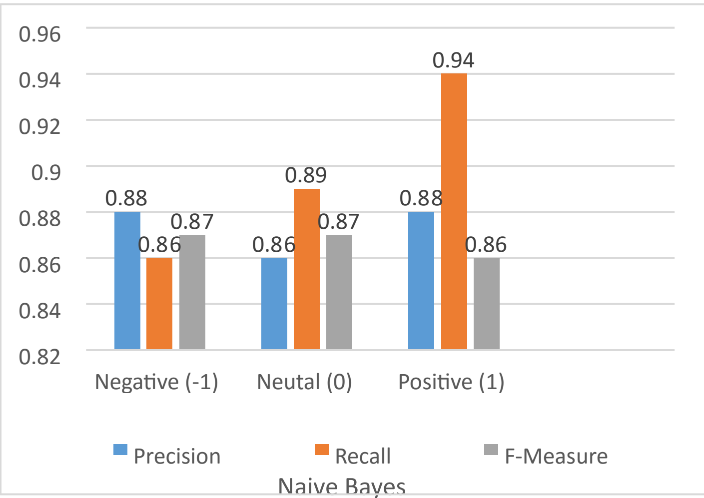


Figure 3: Bar Chart Visualization of Naïve Bayes Model for Precision, Recall, and F1 Score

4.4.2 Decision Tree Classifier

Table 2: Prediction report for Decision Tree Classifier

Label	precision	recall	f1-score
Negative (-1)	0.96	0.95	0.95
Neutral (0)	0.94	0.98	0.96
Positive (1)	1.00	0.92	0.96

weighted average	0.96	0.96	0.96

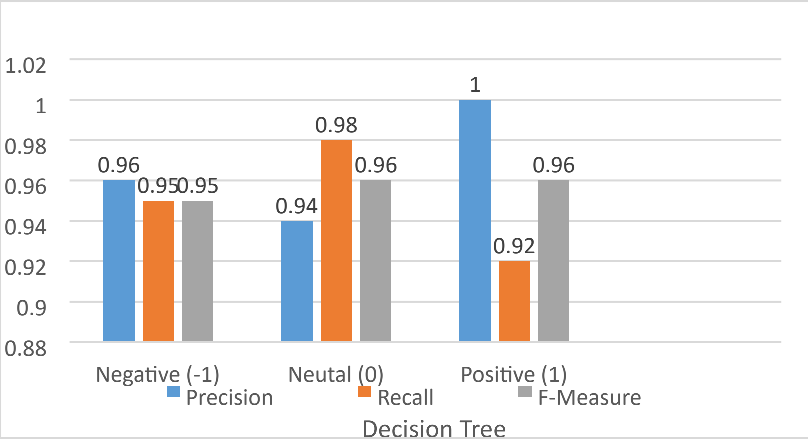


Figure 4: Bar Chart Visualization of Decision Tree Model for Precision, Recall, and F1 Score

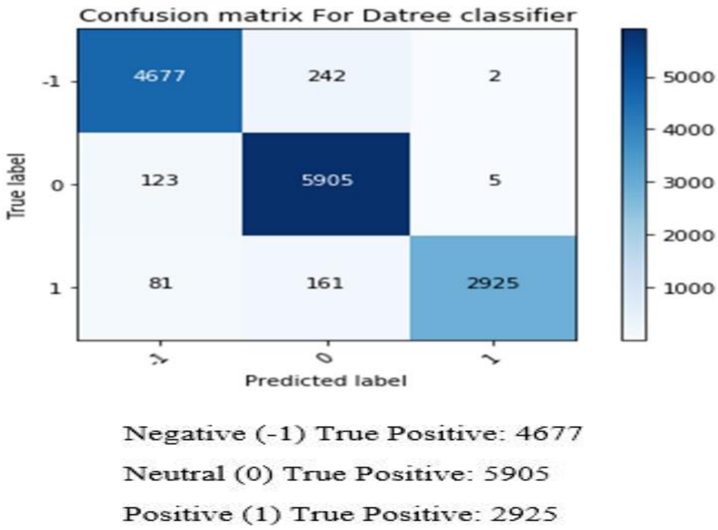


Figure 5: Confusion Matrix for Decision Tree Classifier

4.4.2 K-Nearest Neighbor (KNN) Classifier

Table 3: Prediction report for K-Nearest Neighbor (KNN)Classifier

Label	precision	recall	f1-score
Negative (-1)	0.74	0.98	0.84
Neutral (0)	0.82	0.80	0.81
Positive (1)	0.97	0.54	0.70
weighted average	0.83	0.80	0.80

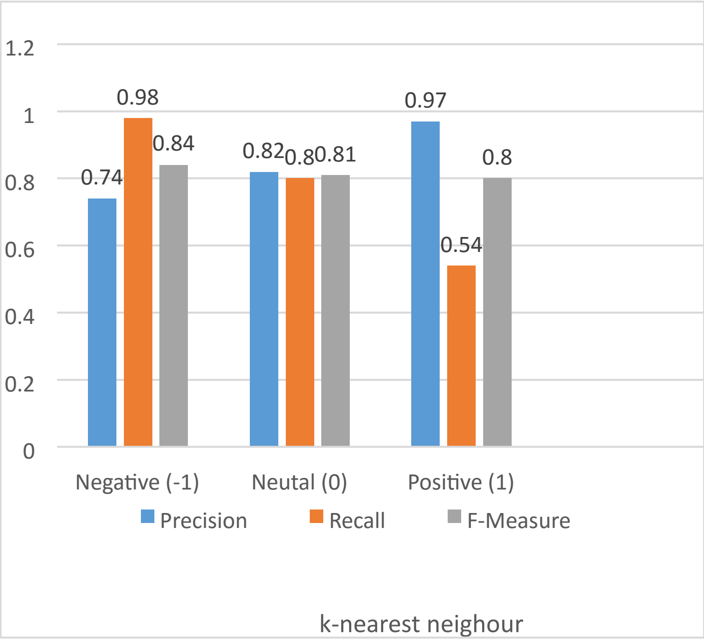


Figure 6: Bar Chart Visualization of K-Nearest Neighbor Model for Precision, Recall, and F1 Score

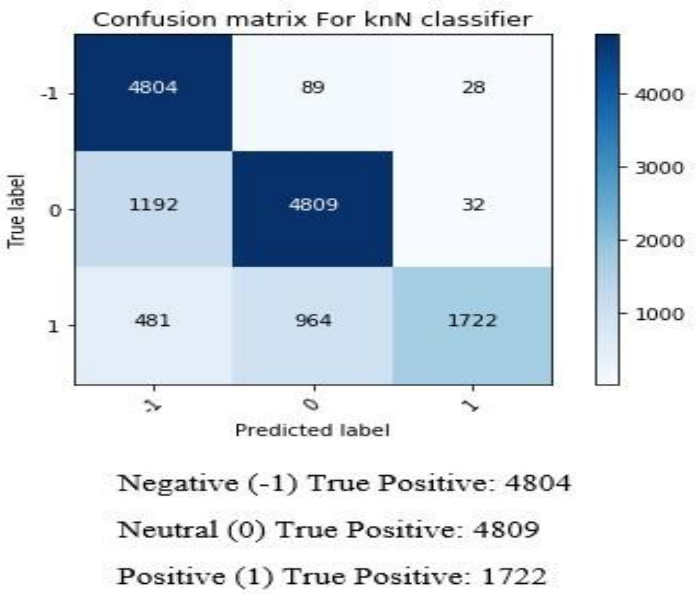


Figure 7: Confusion Matrix for KNN Classifier

4.4.2 Random Forest Classifier

Table 4: Prediction report for Random Forest Classifier

Label	precision	recall	f1-score
Negative (-1)	0.96	0.95	0.95
Neutral (0)	0.93	0.98	0.96
Positive (1)	0.99	0.93	0.96

weighted average	0.99	0.96	0.96

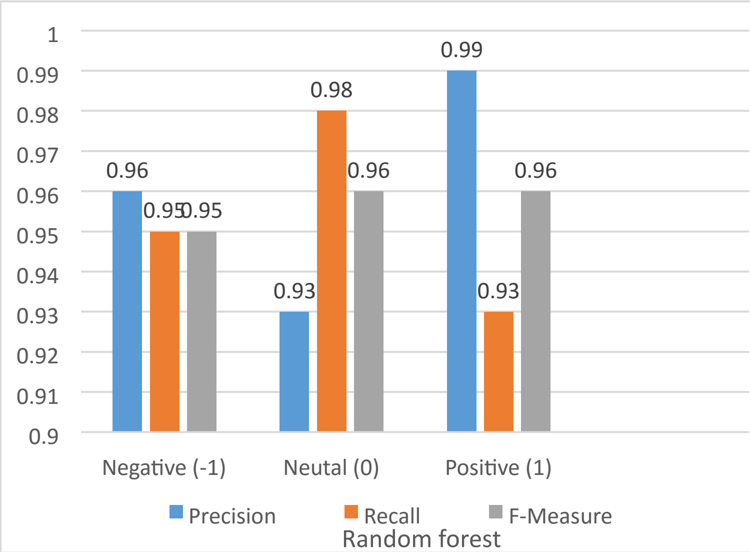
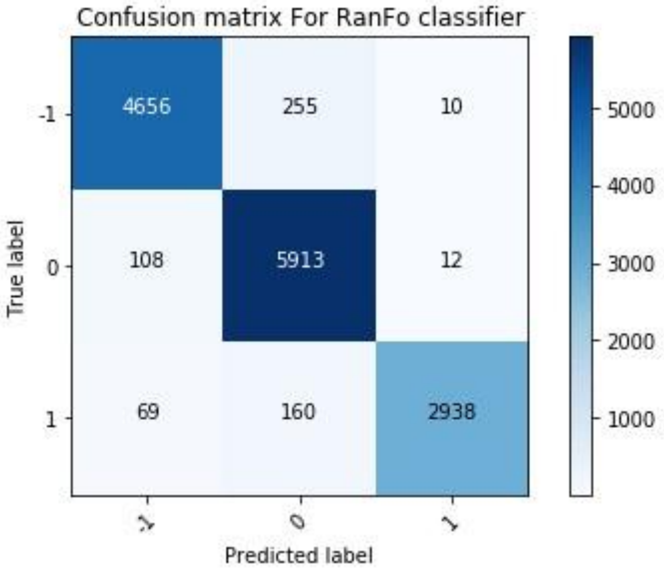


Figure 8: Bar Chart Visualization of Random Forest Model for Precision, Recall, and F1 Score



Negative (-1) True Positive: 4656

Neutral (0) True Positive: 5913

Positive (1) True Positive: 2938

Figure 9: Confusion matrix for RanFo

4.5 Algorithms Comparison and Discussion

The primary concern in extracting Twitter messages that contain very large data and timerelated information is the accuracy and reliability in finding target tweets. One of its goals was to correctly identify tweets in large data sets with natural language processing (NLP) which correspond to causal facts. An addiction parser was used to enhance the detailed extraction of information, and the related NLP techniques. The number of causal relations derived is minimal. But assessment indicated a high degree of accuracy. The use of the lexiconsyntactic relations from the dependence parser results in great precision, which is an important factor in the extraction of data from a wide range.

Also, the metrics for evaluation are precision, recall, and F1-score. Precision gives the percentage any of the classifiers correctly predict, while recall gives the percentage of correct detection of a particular class of tweet whether neutral, positive or negative. The F1-score value, the harmonic mean for precision, and how well the algorithms were able to recall different classes.

Considering the aforementionedmetrics of accuracy, precision, and recall, random forests did very well, but it takes so much time to build the trees in the forest. Random forest with an accuracy of (96.76%) takes a significant average amount of time to complete its execution (547.41s) when compared to other algorithms. Decision Tree follows closely by the accuracy of 96.54% with a significant reduction in execution time when compared to the random forest (99.96s). Naïve Bayes also have good accuracy (91.32%) with a very small execution time (3.30s), it has the best execution time among the algorithms used. K-Nearest Neighbor with an accuracy of 81.84% is the least performer of the algorithms but has the second best execution time of 55.63s.

Table 5: Comparing the Average Performance of the Algorithms.

Algorithms	Accuracy (%)	Execution Time (s)	Precision	Rec-all	F1-score
NB	91.32	3.30	0.87	0.87	0.87
DT	96.54	99.96	0.96	0.96	0.96
K-NN	81.84	55.63	0.83	0.80	0.80
RF	96.76	547.41	0.99	0.96	0.96

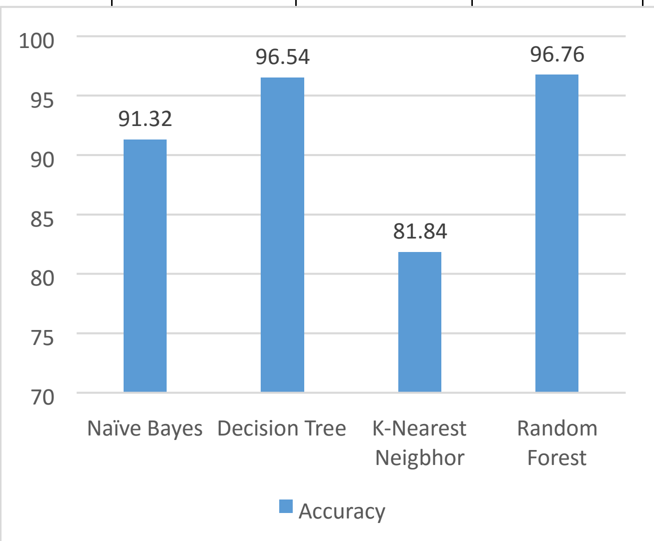


Figure 10: Bar chart Showing accuracy of each algorithm

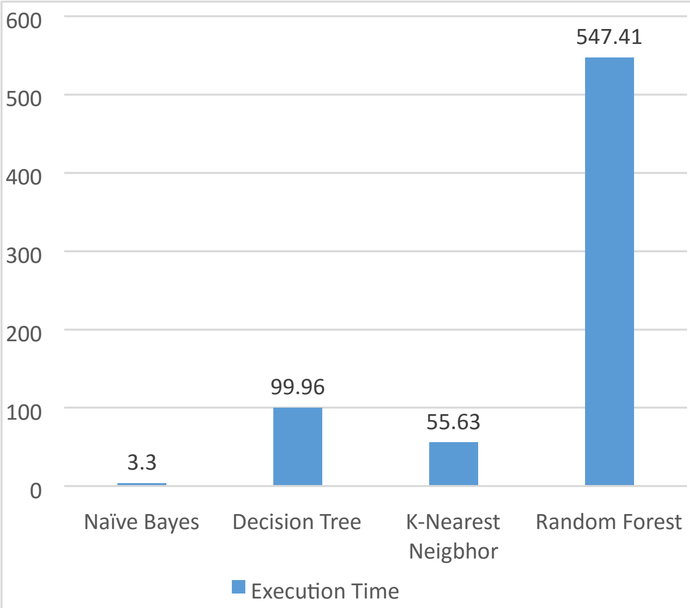


Figure 11: Bar chart Showing execution time for each algorithm

5.0 Conclusion

A lot of data and information is revealed through different social media users' accounts. Tweets from different Twitter users were scrapped, pre-processed, and this data was used to train the algorithms for the models.After the development of models for mental health risk classification. Decision Tree algorithm with an accuracy over 96% and Random Forest with

96% accuracy outperforms other algorithms because they aggregate multiple steps and decisions to give the best prediction. Naïve Bayes follows with an accuracy of 91% accuracy, and K-Nearest Nearest; 81% accuracy because it only captures data samples near to itself The goal was to provide a simple, reliable, and scalable platform for user recognition, the study of the language and feeling, trends, patterns and, sentiment of their writings by analyzing their tweets and classifying them as positive, negative, or neutral users, as determined by the system. These models could also be used by following Twitter users to monitor their activity and use of language as a powerful tool in the treatment of mental health. It has been found that people dealing with depression and other mental disorders most often use self-referenced terms more often and their vocabulary is more negative.

The results show that social media and real-life correlate. The model can also be integrated into the Health-Related Information Management System, which tracks and maintains clinical information that can be fed to the classification prediction model for mental health illness,enhances clinical decisions concerning risk for mental illness, and analyses clinical information concerning the risk of mental illness in a remote location in real-time. The system is important for the assessment, early detection, and tracking of patients, mental health workers, and policymakers.

References

Mhambe P.D., Egejuru N.C., Balogun J.A., Olusanya O.S., andIdowu P.A., (2018). A Predictive Model for the Risk of Mental Illness in Nigeria Using Data Mining. *International Journal of Immunology.* Vol. 6, No. 1, 2018, pp. 5-16.

Vijiyarani, S. and Sudha, S., (2013). Disease Prediction in Data Mining Technique – A Survey. *International Journal of Computer Applications & Information Technology*, vol. II, no. I, pp. 17–21.

Alonso, S.G., Torre-Díez, I., Hamrioui, S., López-Coronado, M., Barreno, D.B., Nozaleda, L. M., and Franco, M., (2018). Data Mining Algorithms and Techniques in Mental Health: A Systematic Review. I*n Journal of Medical Systems*42: 16

Dipnall, J. F., Pasco, J. A., Berk, M., Williams, L. J., Dodd, S., Jacka, F. N., and Meyer, D. (2016). Fusing data mining, machine learning, and traditional statistics to detect biomarkers associated with depression. *PloS one*, *11*(2).

Granholm, E., Ben-Zeev, D., Link, P., Bradshaw, K., and Holden, J. L., (2012). Mobile assessment and treatment for schizophrenia (MATS): A pilot trial of an interactive textmessaging intervention for medication adherence, socialization, and auditory hallucinations." *Schizophrenia Bull*., vol. 38, no. 3, pp. 414–425.

Chandra. E. B, and Bindu. A G., (2019). Novel approach for Psychiatric Patient Detection and Prediction using Data Mining Techniques.*International Journal of Engineering Research & Technology (IJERT).*Volume 7, Issue 05 Special Issue

World Health Organization (WHO). Trastornosmentales. 2019; Available from:http://www.who.int/mediacentre/factsheets/fs396/es/ (last accessed June 2020) [8] Mathew, J., Mekkayil, L., Ramasangu, H., Karthikeyan, B. R., and Manjunath, A. G., (2016). Robust algorithm for early detection of Alzheimer's disease using multiple feature extractions. *IEEE Annu. India Conf.* 2016:1–6.

Ertek, G., Tokdil, B., and Günaydın, İ., (2014). Risk factors and identifiers for Alzheimer's disease. A data mining analysis. *Ind. Conf. Data Min.*;1–11

Fernández-Llatas, C., García-Gomez, J. M., Vicente, J., Naranjo, J. C., Robles, M., and Benedí, J. M., (2011). Behavior patterns detection for persuasive design in nursing homes to help dementia patients. *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS*.;6413–6417. [11] Jung, Y., and Yoon, Y. I., (2017). Multi-level assessment model for wellness service based on human mental stress level. *Multimed. Tools Appl.* 76(9):11305–11317.

Tovar, D., Cornejo, E., Xanthopoulos, P., Guarracino, M. R., and Pardalos, P. M., (2012). Data mining in psychiatric research. *Psychiatr. Disord*. 829:593–603.

Husain, W., Xin, L. K., Rashid, N. A., and Jothi, N., (2016) 'Predicting Generalized Anxiety Disorder Among Women Using Random Forest Approach', in *IEEE 3rd International Conference On Computer and Information Sciences (ICCOINS)*, pp. 37–42.

Li, X., Hu, B., Sun, S., and Cai, H.(2016), EEG-based mild depressive detection using feature selection methods and classifiers. *Comput. Methods Programs Biomed*. 136:151–161. [15] Bhakta, I and Sau,

A. (2016). Prediction of Depression among Senior Citizens using Machine Learning Classifiers. *International Journal of Computer Applications* 144 (7): 11–16.

Kim, J. Y., Liu, N., Tan, H. X., and Chu, C. H., (2017) Unobtrusive monitoring to detect depression for elderly with chronic illnesses. *IEEE Sens. J.* 17(17):5694–5704, 2017.

Bhavsar, H., & Manglani, R. (2019). Sentiment Analysis of Twitter Data using Python. *International Research Journal of Engineering and Technology (IRJET)* Mar 2019eISSN, 2395-0056.

Baba, T., Baba, K., and Ikeda, D. (2019). Detecting mental health illness using short comments. In *International Conference on Advanced Information Networking and Applications* (pp. 265-271). Springer, Cham.

Witten H.I., and Frank E. (2005), Data Mining: Practical Machine Learning Tools and Techniques, Second edition, Morgan Kaufmann Publishers.

Cairney, J., Veldhuizen, S., Vigod, S., Streiner, D. L., Wade, T. J., and Kurdyak, P., (2018) Exploring the social determinants of mental health service use using intersectionality theory and CART analysis. *J. Epidemiol. Commun. Health.* 68(2):145–150.

Matthews, M., Murnane, E. L., Cosley, D., Chang, P., Guha, S., Frank, E., and Gay, G., (2016). Self-monitoring practices, attitudes, and needs of individuals with bipolar disorder: implications for the design of technologies to manage mental health. *Journal of the American Medical Informatics Association*, *23*(3), 477-484.

Twitter, (2020). Developers Documentation. Retrieved Jan. 10, 2020, from https://dev.twitter.com/overview/documentation