**Classifying Depression, Anxiety and Stress based on their Severity Using Machine Learning**

**ABSTRACT:**

Over the past few decades, many people are suffering from psychological health issues such as anxiety, depression, and stress. It is crucial to detect mental health condition on a timely basis and cure before it turns into a severe problem. In this research, we analyze the performance of Classification algorithms to predict severity levels of depression, anxiety, and stress. Severity levels can be measured by using the data based upon the Depression, Anxiety, and Stress Scale (DASS-42) questionnaire which consists of 42 questions and Ten Item Personality Inventory. The intention of this research is to use Machine Learning algorithms to classify mental health issues of people into categories based on their severity. Severity levels can be measured based on the data received by using DASS questionnaire and Ten Item Personality Inventory.

Algorithms used in this research are Decision tree, Naive Bayes, K-Nearest Neighbour (KNN). It is observed that, Decision Tree Algorithm achieved better performance in terms of accuracy compared to KNN and Navie Bayes.

**1.INTRODUCTION:**

Humans are, by nature, becoming ambitious nowadays and seek every possible opportunity to grow professionally. Depression, Anxiety and Stress have become so common that people now believe them to be part and parcel of professional life. The World Health Organization (WHO) has observed that depression is the most prevalent mental disorder affecting more than 300 million people and the severity of the issue has led many health researchers to focus their studies in this area. Differentiating depression, anxiety and stress from one another is problematic for machines; hence, an appropriate learning algorithm is required for an accurate diagnosis. According to WHO, a healthy person possesses a healthy brain along with physical wellness. [1] The standard diagnosis criterion Depression, Anxiety and Stress Scale (DASS-42), which has 42 questions, is used for screening the symptoms related to these mental illnesses [5-6]. The main symptoms of depression [3] from a clinical point of view are loss of memory; lack of concentration; an inability to make decisions; loss of interest in recreational activities; overeating and low appetite and weight loss; feelings of guilt, worthlessness, helplessness, restlessness and irritation; as well as suicidal thoughts. In case [2], these symptoms were found to have a significant effect on important areas of an individual’s life – such as in education, employment and social activities, and this provides a vital clue for forming a clinical diagnosis. The symptoms of Anxiety [3] are irritability, fatigue, insomnia, panic, increased heart rate, sweating, and difficulty concentrating. The symptoms of stress [4] are feeling upset or agitated, an inability to relax, low energy levels, chronic headaches. Thus, stress, anxiety and depression have many common symptoms including insomnia, chest pain, fatigue, increased heart rate and inability to concentrate, all of which makes classification challenging for machines.

**2. RELATED STUDIES:**

Many researchers have worked on predicting anxiety and depression with machine learning algorithms, such as Random Forest Tree (RFT), the Support Vector Machine (SVM) and the Convolution Neural Network (CNN) for the collection and subsequent classification of data from blog posts. For encoding the text, various techniques have been used, that is topic modelling, Bag-of-Words (BOW) and Term Frequency–Inverse Document Frequency (TF– IDF). Moreover, Python programming has been used for modelling experiments, with the best results among all the classifiers [2] being produced by the CNN, whose accuracy and recall scores were found to be 78% and 0.72, respectively. Different machine learning algorithms such as Logistic Regression, Catboost, Naïve Bayes, RFT and SVM were applied for classification in [7]. In this study, 470 seafarers were interviewed and information on the occupations, socio-demographics and health of the participants was collected via 16 characteristics including age, academic qualifications, monthly income, employment status, BMI, duration of service, family type, marital status, presence (if any) of hypertension, diabetes or ischemic heart disease, job profile, rank within the organization, types of vessels posted to and dummy variables for academic qualifications and marital status. As a result, the researchers found that Catboost produced the highest levels of accuracy and precision among all the classifiers – i.e. 82.6% and 84.1%, respectively.

Sau et al. (2017) manually collected data from the Medical College and Hospital of Kolkata, West Bengal on 630 elderly individuals, 520 of whom were in special care. After applying different classification methods Bayesian Network, logistic, multiple layer perceptron, Naïve Bayes, random forest, random tree, J48, sequential random optimization, random sub-space and K star they observed that random forest produced the best accuracy rate of 91% and 89% among the two data sets of 110 and 520 people, respectively. For feature selection and classification, WEKA tool were used in [1]. These days, social media is rapidly turning into a healthcare evaluation tool for predicting various types of illness. Saha et al. [8] selected topics and psycholinguistic attributes appearing in posts on the LiveJournal website. These were then inputted into a joint modelling framework, so as to categorize the mental problems occurring in online communities with an interest in depression. The proposed joint modelling framework outperformed the existing single task learning (STL) and multi task learning (MLT) baselines, and the study showed that discussions in online communities went beyond feelings of being depressed. Reece et al. [9] focused on the predictors of depression and Post Traumatic Stress Disorder (PTSD) among Twitter users. The Hidden Markov Model (HMM) was used to recognize increases in the probability of PTSD. Of the entire dataset, 31.4% and 24% were observed to be affected by depression and PTSD. Braithwaite et al. [10] collected tweets from 135 participants recruited from Amazon Mechanical Turk (M Turk) and applied decision tree classification to measure suicide risk. The accuracy level for the prediction of suicide rate was observed to be 92%.

Du et al. [11] extracted streaming data from Twitter and used psychiatric stressors to annotate tweets that had been deemed suicidal. The Convolution Neural Network (CNN) outperformed the Support Vector Machine (SVM) and extra trees (ET) etc. with a precision of 78% in recognizing tweets with suicidal tendencies.

The audio-text approach can also be used to model depression, where the researcher collects data from individuals with depression. The long short-term memory neural network model was used for detecting depression in [12], which observed that the context-free model produced the best results for audio (weighted, sequence and multi-model).

Depression was also predicted in [13] in the early stages through social media content. Data collection was carried out using CLEF eRisk. After evaluating five systems, it was discovered that a combination of machine learning and information retrieval gave the optimum result. In Hou et al., a big data approach was used to predict depression based on a person’s reading habits. The features of Chinese text were extracted in order to develop a book classifier and after applying five classifications, naïve Bayes was found to be the most appropriate [14].

Post-traumatic stress disorder has detected in [15] using supervised machine learning classifiers. Their study is on ex-serviceman UK militants, the parameters used in their study alcohol misuse, gender and deployment status. As results satisfactory sensitivity was obtained for multiple supervised Machine Learning classifiers, but the outcomes were not very sensitive to false negative diagnoses. Anxiety and mood disorder were detected in [16] by scanning patient facial emotions and applying cross validation and better precise results were found that is verified by different statistical measures. Imbalance classification was applied in [20] and ensemble machine learning methods were discussed in [21]. Different researchers have applied different machine leaning algorithms for the prediction of psychological disorders, and the performances of different algorithms have been found to vary, depending on the scenario; no fixed algorithm has been determined as most suitable in all cases. Thus, in the present study, all the machine learning algorithms were applied to identify the symptoms of anxiety, depression and stress.

**3. MATERIALS AND METHODOLOGY**

This research focused on classifying the anxiety, depression and stress levels using machine learning algorithms. Dataset was taken from Kaggle and was subsequently the tuples were classified using three machine learning algorithms – Decision Tree classifier, Naïve Bayes method and k – Nearest Neighbor classifier

* 1. **DATASET COLLECTION:**

Dataset collected from kaggle.com while the original source was Openpsychometrics.org. The data was hosted on OpenPsychometrics.org a nonprofit effort to educate the public about psychology and to collect data for psychological research

The authors of the dataset collected the data between 2017-2019 using the online version of the Depression Anxiety Stress Scales (DASS), see <http://www2.psy.unsw.edu.au/dass/>

Link to dataset on Kaggle

<https://www.kaggle.com/datasets/lucasgreenwell/depression-anxiety-stress-scales-responses>

**3.1.1 Attributes:**

The data for the study were collected through DASS-42, the Depression, Anxiety and Stress Scale questionnaire. DASS 42 comprises 42 questions, with 14 questions allocated to each of the scales of Stress, Anxiety and Depression. The possible answers for each question – which could be given in text or numeric form – are as follows:

0 = Did not apply to me at all

1 = Applied to me to some degree, or some of the time

2 = Applied to me to a considerable degree, or a good part of the time

3 = Applied to me very much, or most of the time

The questions asked in the DASS-42 questionnaire are described in table 1.

Table 1. Questions related to depression, anxiety and stress.

|  |  |  |
| --- | --- | --- |
| Depression | Anxiety | Stress |
| I couldn’t seem to experience any positive feeling at all.  I just couldn't seem to get going.  I felt that I had nothing to look forward to.  I felt sad and depressed.  I felt that I had lost interest in just about everything.  I felt I wasn't worth much as a person.  I felt that life wasn't worthwhile.  I couldn't seem to get any enjoyment out of the things I did.  I felt down-hearted and blue.  I was unable to become enthusiastic about anything  I felt I was pretty worthless  I could see nothing in the future to be hopeful about  I felt that life was meaningless  I found it difficult to work up the initiative to do things | I was aware of dryness of my mouth  I experienced breathing difficulty  I had a feeling of shakiness (eg, legs going to give way)  I found myself in situations that made me so anxious I was most relieved when they ended  I had a feeling of faintness  I perspired noticeably (eg, hands sweaty) in the absence of high temperatures or physical exertion  I felt scared without any good reason  I had difficulty in swallowing  I was aware of the action of my heart in the absence of physical exertion  I felt I was close to panic  I feared that I would be "thrown" by some trivial but unfamiliar task  I felt terrified  I was worried about situations in which I might panic and make a fool of myself  I experienced trembling | I found myself getting upset by quite trivial things  I tended to over-react to situations  I found it difficult to relax  I found myself getting upset rather easily  I felt that I was using a lot of nervous energy  I found myself getting impatient when I was delayed in any way  I felt that I was rather touchy  I found it hard to wind down  I found that I was very irritable  I found it hard to calm down after something upset me  I found it difficult to tolerate interruptions to what I was doing  I was in a state of nervous tension  I was intolerant of anything that kept me from getting on with what I was doing  I found myself getting agitated |

Along with this questions, ten personality types are also taken into consideration because the goal here is to see how different personality types are related to individual emotional states.

The Ten Item Personality Inventory was administered [10]

TIPI1 Extraverted, enthusiastic.

TIPI2 Critical, quarrelsome.

TIPI3 Dependable, self-disciplined.

TIPI4 Anxious, easily upset.

TIPI5 Open to new experiences, complex.

TIPI6 Reserved, quiet.

TIPI7 Sympathetic, warm.

TIPI8 Disorganized, careless.

TIPI9 Calm, emotionally stable.

TIPI10 Conventional, uncreative.

The TIPI attributes have values 1, 2, .., 7 based on how much the person relates to the personality type

1 = Disagree strongly

2 = Disagree moderately

3 = Disagree a little

4 = Neither agree nor disagree

5 = Agree a little

6 = Agree moderately

7 = Agree strongly

Count of Attributes in the dataset is 58

42 questions - (Q1A, Q2A,…………..Q42A)

10 personality types - (TIPI1,TIPI2,……….TIPI10)

Education - 1 = Less than high school, 2 = High school, 3 = University degree, 4 = Graduate degree

Urban - What type of area the person grew up?

1 = Rural , 2 = Suburban, 3 = Urban

Gender - 1 = Male, 2 = Female, 3 = Other

Age - how old the person is?

Married - 1 = Never married, 2 = Currently married, 3 = Previously married

Familysize - size of the person’s family

Count of Records in the dataset are: 39775

* 1. **DATA PREPROCESSING:**

Technique which is used to transform the raw data in a useful and efficient format.

**3.2.1 Data Cleaning:**

The data can have many irrelevant attributes and attributes with missing values. Data cleaning involves handling attributes with missing values, noisy and redundant data. In this research the attributes which are not necessary are removed and there were no attributes with missing values after the above step.

**3.2.2Data Reduction:**

Data reduction is a process that decreases the size of the original data and reflects it into a much smaller size.

**Principle Component Analysis (PCA):**

Data reduction technique which groups the important variables into a component taking the maximum information present within the data and discards the other, not important variables. For each issue Depression, Anxiety and Stress, PCA is applied to convert 14 attributes which contain scores for 14 DASS questions relating to that issue into single attribute and set of 10 attributes which contain relevance scores for 10 personality types are made into single attribute. When classifying the overall mental state, the 42 attributes containing scores for 42 DASS questions are converted into one attribute remaining attributes are reduced as above mentioned.

**3.3 CLASSIFICATION:**

**3.3.1 Classes**

Count of Class labels in the dataset : 4

Mental\_state – the overall mental state of the person

Stress\_state – severity of stress the person is facing

Depression\_state – severity of depression state the person is facing

Anxiety\_state – severity of anxiety state the person is facing

Each class label has 5 different values : 0. Normal

1. Mild
2. Moderate
3. Severe
4. Extremely Severe

A tuple is classified to any one of the above mentioned severity levels for each Depression, Anxiety, Stress and Overall mental state.

The dataset was divided into the ratio 60:40, representing the training and test sets, respectively. The working principles of each machine learning algorithm are described below:

**3.3.2 Decision Tree:**

A decision tree is a structure that includes a root node, branches, and leaf nodes.Each internal node denotes a test on an attribute, each branch denotes the outcome of a test, and each leaf node holds a class label. The topmost node in the tree is the root node.

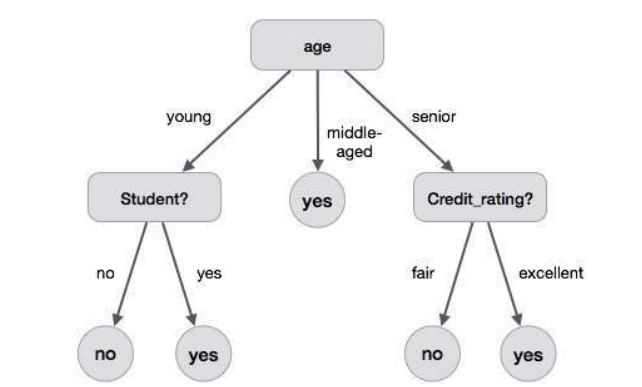
The following decision tree is for the concept buy\_computer that indicates whether a customer at a company is likely to buy a computer or not. Each internal node represents a test on an attribute. Each leaf node represents a class.

Fig.1. Decision tree

**3.3.3 K- Nearest Neighbour (KNN):**

KNN is one of the most straightforward algorithms adopted in machine learning for classification and regression problems. Based on closest measures, KNN takes information and classifies recent information points. The information is then allotted to the class with the foremost closest neighbour.

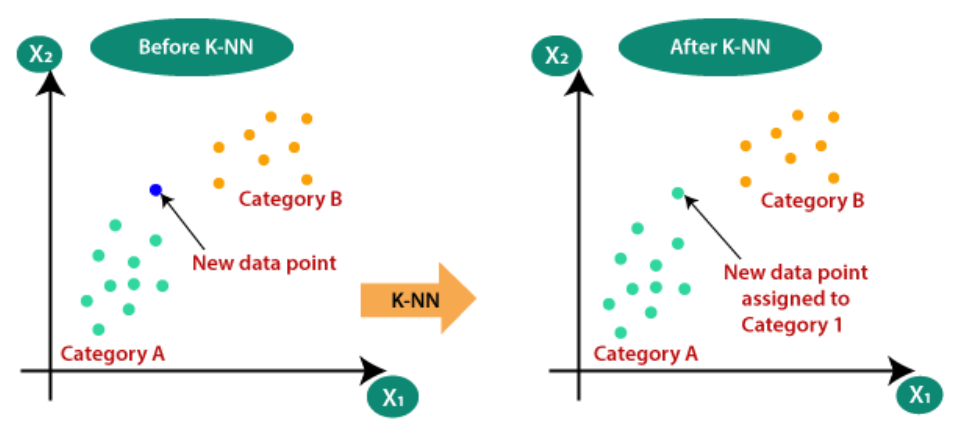
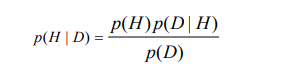


Fig.2 .K-Nearest Neighbour

**3.3.4 Naïve Bayes:**

 This classifier superintends machine learning algorithms using Bayes theorem and works on the premise that features are analytically independent. This theorem depends on naïve assumption, in which input factors are independent of each other [24-25]. The formula for naive bayes is given as follows:

p(H│D) = this is posterior

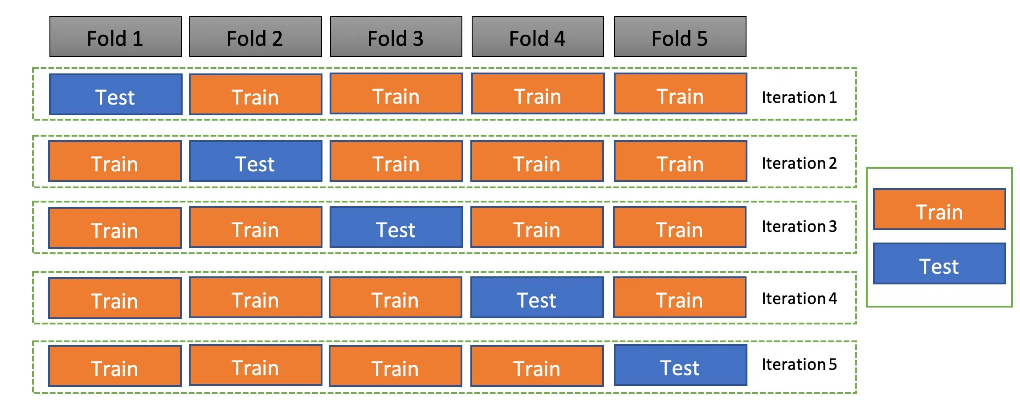
p(H) = this is the prior i.e. what you believed before you saw the evidence

p(D|H) = this the likelihood of seeing that evidence if your hypothesis is correct

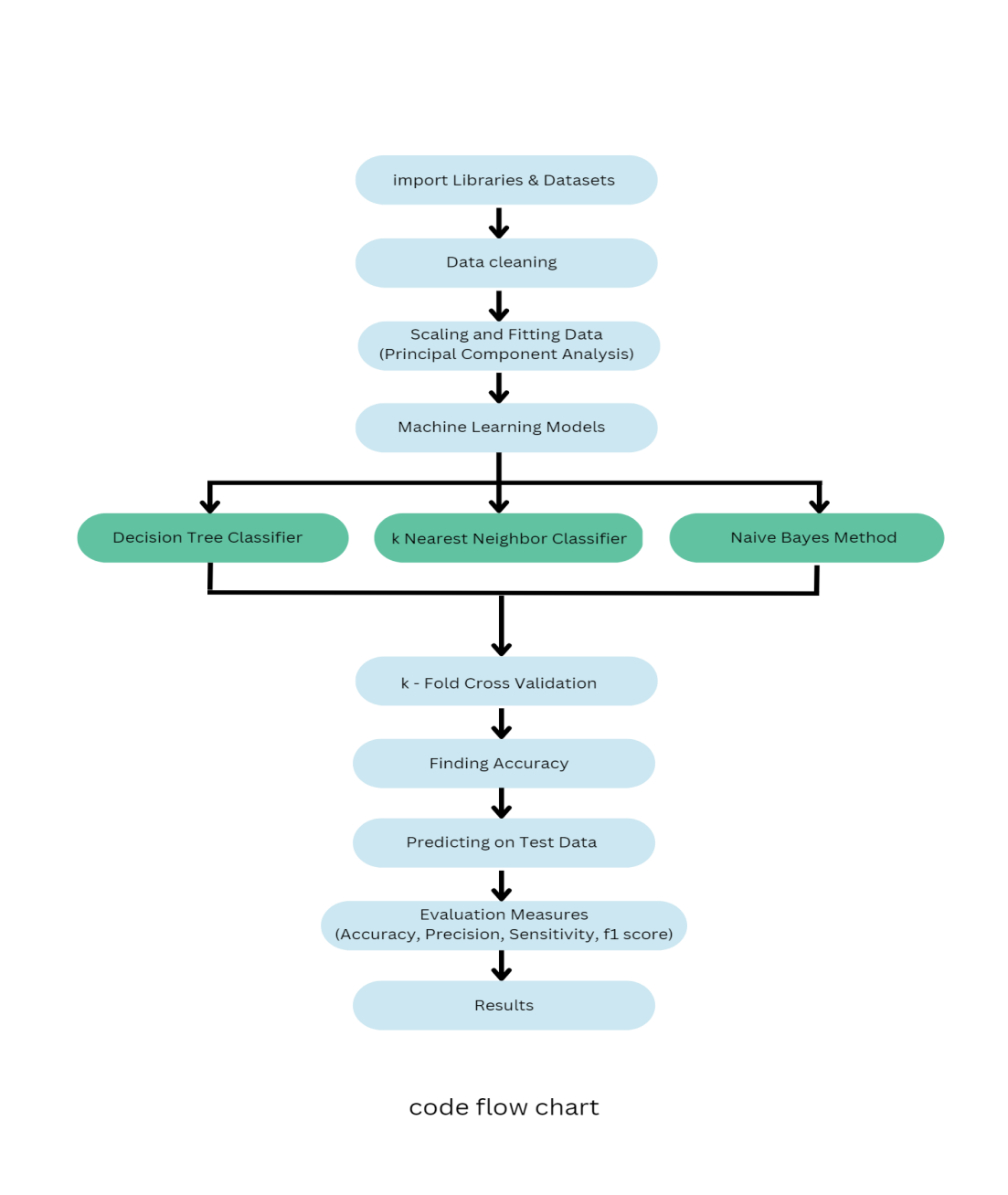
p(D) = this is the normalizing of that evidence under any circumstances.

**3.3.5** **k-Fold Cross Validation:**

Evaluating a Machine Learning model can be quite tricky. Usually, we split the data set into training and testing sets and use the training set to train the model and testing set to test the model. We then evaluate the model performance based on an error metric to determine the accuracy of the model. This method however, is not very reliable as the accuracy obtained for one test set can be very different to the accuracy obtained for a different test set. K-fold Cross Validation(CV) provides a solution to this problem by dividing the data into folds and ensuring that each fold is used as a testing set at some point. K-fold CV has been used to find the correct accuracy given by each algortihm.

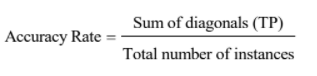


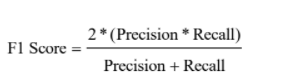
**3.3.6 Code Flowchart:**

****

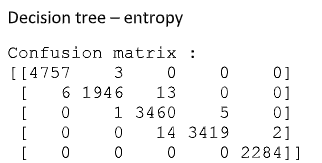
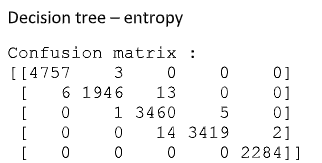
**4. RESULTS AND DISCUSSION**

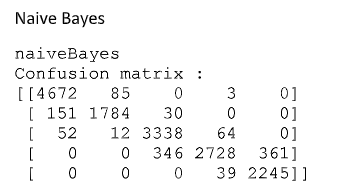
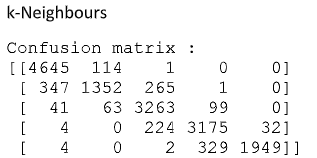
The application of all three models – i.e. Decision Tree (DT), Naïve Bayes (NB), and K-Nearest Neighbour (KNN) – to all three classes of Stress, Anxiety, and Depression, resulted in the confusion matrices decpicted below. The rows in the matrices show the actual classes whilst the columns show the predicted classes. The equations to calculate accuracy, precision, recall and f1 score are mentioned below:

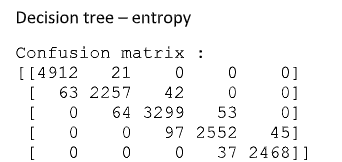
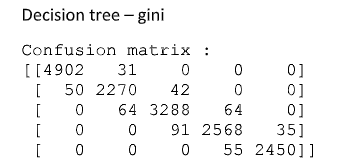
p

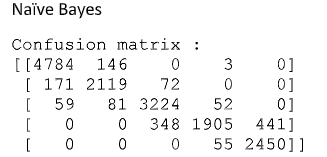
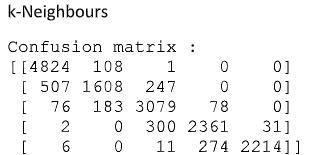


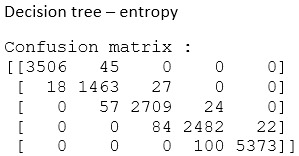
r

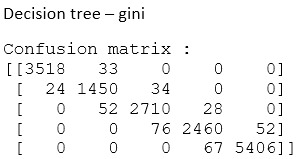
**Stress :**

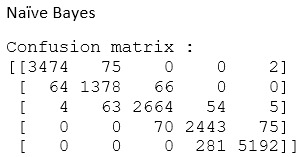
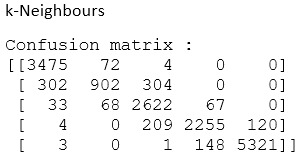


Anxiety :

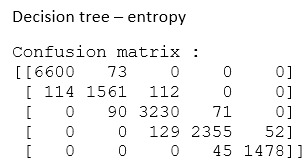


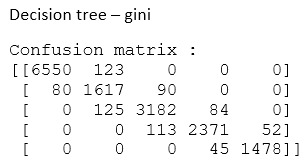
Depression :

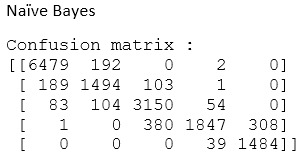


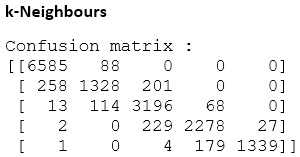


|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Classifier** | **Issue** | **Accuracy** | **Precision** | **Recall** | **F1 score** |
| Decision Tree (Entropy) | Depression  Anxiety  Stress  Mental state | 0.9763  0.9734  0.9972  0.9568 | 0.9766  0.9734  0.9972  0.9567 | 0.9763  0.9734  0.9972  0.9568 | 0.9764  0.9734  0.9972  0.9567 |
| Decision Tree (Gini index) | Depression  Anxiety  Stress  Mental state | 0.9769  0.9728  0.9971  0.9552 | 0.9770  0.9728  0.9971  0.9557 | 0.9769  0.9728  0.9971  0.9557 | 0.9770  0.9728  0.9971  0.9554 |
| Naïve Bayes | Depression  Anxiety  Stress  Mental state | 0.9522  0.9102  0.9281  0.9084 | 0.9536  0.9134  0.9311  0.9118 | 0.9522  0.9102  0.9281  0.9084 | 0.9526  0.9079  0.9270  0.9089 |
| k-Nearest Neighbour | Depression  Anxiety  Stress  Mental state | 0.9160  0.8853  0.9040  0.9255 | 0.9165  0.8862  0.9050  0.9254 | 0.9160  0.8853  0.9040  0.9255 | 0.9131  0.8832  0.9020  0.9244 |

**Mental Health State:**







**5. CONCLUSION :**

In this project, machine learning algorithms were applied to classify anxiety, depression and stress into five severity levels. Data was based upon a standard questionnaire measuring the common symptoms of anxiety, depression and stress (DASS-42). Three classification technique were applied – Decision Tree, Naïve Bayes and K - Nearest Neighbour. Decision Tree classifier has produced the highest accuracy among all algorithms for all issues. Accuracy was same when two different attribute selection measures ‘Entropy’ and ‘Gini index’ were used. By the overall results from algorithms, Decision Tree classifier was found to be the best algorithm to classify mental health issues into 5 different severity levels.

**6. REFERENCES** **:**

[1] Sau, A., Bhakta, I. (2017) “ Predicting anxiety and depression in elderly patients using machine

learning technology.” Healthcare Technology Letters 4 (6): 238-43.

[2] Tyshchenko, Y. (2018) “Depression and anxiety detection from blog posts data.”Nature Precis. Sci.,

Inst. Comput. Sci., Univ. Tartu, Tartu, Estonia.

[3] https://adaa.org/understanding-anxiety/depression/symptoms

[4] https://www.webmd.com/balance/stress-management/stress-symptoms-effects\_of-stress-on-the-body#1

[5] Oei, T. P., Sawang, S., Goh, Y. W., Mukhtar, F. (2013) “Using the depression anxiety stress scale 21

(DASS-21) across cultures.” International Journal of Psychology 48 (6): 1018-1029.

[6] Kroenke, K., Spitzer, R. L.,Williams, J. B. (2001) “The PHQ‐9: validity of a brief depression severity

measure.” Journal of general internalmedicine 16 (9): 606-613.

[7] Sau, A., Bhakta, I. (2018) “Screening of anxiety and depression among the seafarers using machine

learning technology.” Informatics in Medicine Unlocked :100149

[8] Saha, B., Nguyen, T., Phung, D., Venkatesh, S. (2016) "A framework for classifying online mental health-related communities with an interest in depression."IEEE journal of biomedical and health informatics 20 (4): 1008-1015.

[9] Reece, A. G., Reagan, A. J., Lix, K. L. M., Dodds, P. S., Danforth, C. M., Langer, E. J.(2016) "Forecasting the Onset and Course of Mental Illness with Twitter Data."Scientific reports 7 (1): 13006.

[10]Braithwaite, S. R., Giraud-Carrier, C., West, J., Barnes, M. D., Hanson, C.L. (2016) "Validating machine learning algorithms for Twitter data against established measures of suicidality."JMIR mental health 3 (2): e21.

[11]Du, J., Zhang, Y., Luo, J., Jia, Y., Wei, Q., Tao, C., Xu, H. (2018) "Extracting psychiatric stressors for

suicide from social media using deep learning."BMC medical informatics and decision making 18 (2): 43.

[12]Al Hanai, T., Ghassemi, M. M., Glass, J.R. (2018) "Detecting Depression with Audio/Text Sequence Modeling of Interviews." InInterspeech : 1716-1720.

[13]Ramiandrisoa, F., Mothe, J., Benamara, F., Moriceau, V. (2018) "IRIT at e-Risk 2018."E-Risk workshop: 367-377.

[14] Hou, Y., Xu, J., Huang, Y., Ma, X. (2016) "A big data application to predict depression in the university based on the reading habits."3rd IEEE International Conference on Systems and Informatics (ICSAI): 1085-1089.

[15] Leightley, D., Williamson, V., Darby, J., Fear, N.T. (2019) " Identifying probable post-traumatic stress disorder: applying supervised machine learning to data from a UK military cohort." Journal of Mental Health. 28 (1): 34-41.

[16]Young, C., Harati, S., Ball, T., Williams, L. (2019) "Using Machine Learning to Characterize Circuit-Based Subtypes in Mood and Anxiety Disorders."Biological Psychiatry 85 (10): S310.

[17] Li, L., Zhang, X. (2010) "Study of data mining algorithm based on decision tree." In2010 International Conference On Computer Design and Applications IEEE 1: V1-155.

[18] Paul, A., Mukherjee, D. P., Das, P., Gangopadhyay, A., Chintha, A. R., Kundu, S. (2018)"Improved random forest for classification."IEEE Transactions on Image Processing 27 (8): 4012-4024.

[19] Rodriguez-Galiano, V. F., Ghimire, B., Rogan, J., Chica-Olmo, M., Rigol-Sanchez, J. P. (2012) “An assessment of the effectiveness of a random forest classifier for land-cover classification.”ISPRS Journal of Photogrammetry and Remote Sensing 67: 93-104.

[20] Liu, X. Y., Wu, J., Zhou, Z. H. (2009)"Exploratory undersampling for class-imbalance learning."IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics) 39 (2): 539-50.

[21]Dietterich, T. G. (2000) "Ensemble methods in machine learning."InInternational workshop on multiple classifier systemsSpringer, Berlin, Heidelberg: 1-15.