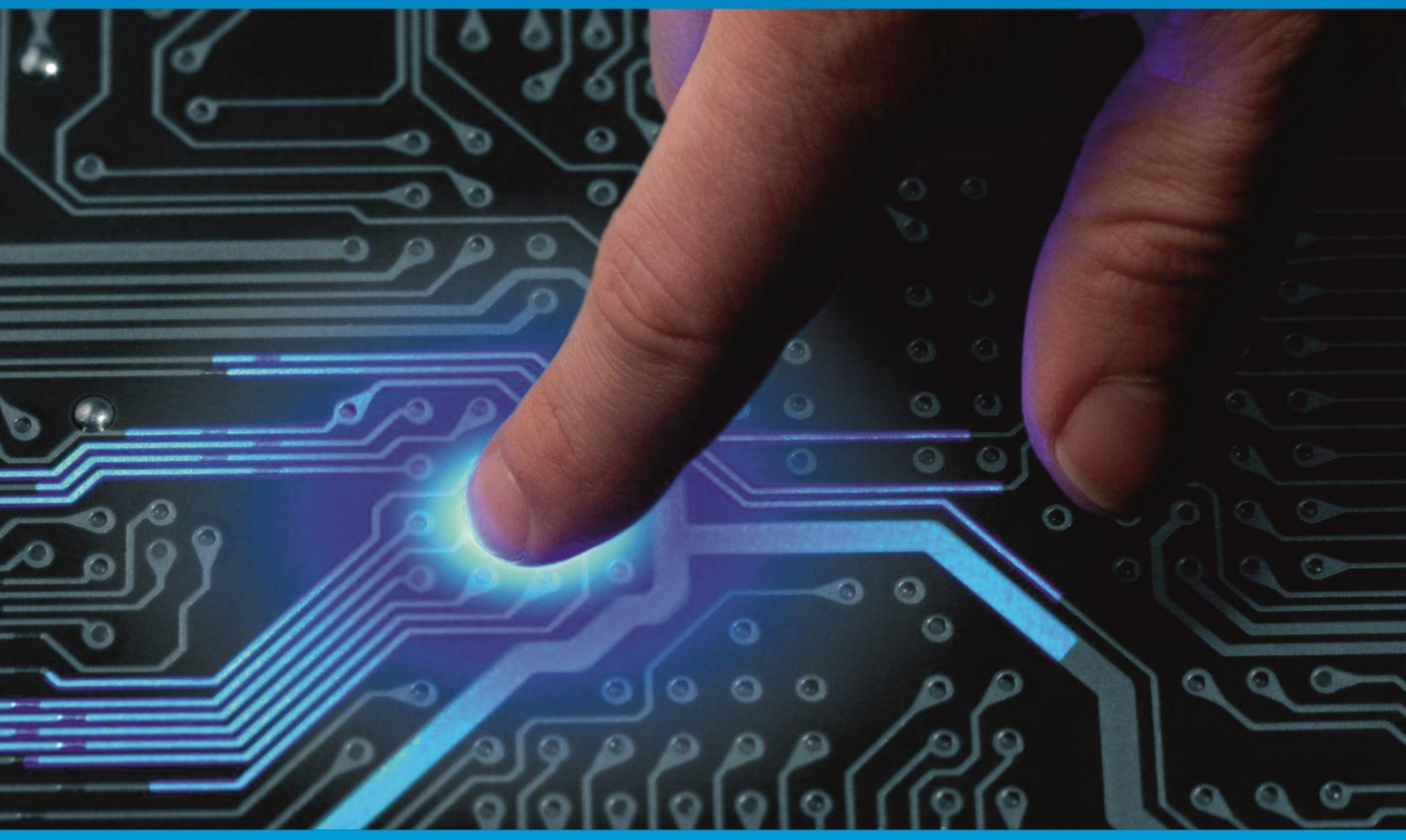
IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 9, Issue 5, May 2021

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 7.488

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 5, May 2021 ||

| DOI: 10.15680/IJIRCCE.2021.090524 |

Predicting Stress, Depression and Anxiety in modern life using machine learning

Nikita Khandagale, Avhale Adarsh, Kambale Mahesh, Dange Saurabh, Prof. A.N,Navathe

Department of Computer Engineering, Amrutvahini College of Engineering, Sangamner, Ahmednagar, Maharashtra, India

ABSTRACT: In the recent time people faces people faces so much psychological problem such stress, depression and anxiety are more common. In this paper, prediction of stress depression and anxiety using machine learning we use various machine learning algorithm to predict any person has how much amount of stress, depression and anxiety. For this project date is collected from various people such as employed, unemployed, man, women across various culture and religion for the collection of this data PHQ(public health questionnaire) were used Which consist of 21 question in which stress depression and anxiety consist of 7 question each. Which also known as DASS21 questionnaire. Stress Depression and anxiety were predicted at the 5 level of severity by various Machine learning algorithm. Further all Calculation and experimental result it is found that decision tree classifier and Random forest tree has more accuracy for the reason Decision Tree classifier were used for final implementation of project.

KEYWORDS: PHQ,DASS21,Decision tree Classifier, Random Forest Classifier, psychology.

I .INTRODUCTION

Humans are, by nature, becoming ambitious nowadays and seek every possible opportunity to grow professionally. Anxiety, depression, stress frustration and dissatisfaction have become so commonplace 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 worldwide, and the severity of the issue has led many health researchers to focus their studies in this area. Differentiating anxiety, depression 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. The standard diagnosis criterion for depression is the Patient Health Questionnaire (PHQ), whilst the Depression, Anxiety and Stress Scale (DASS 21), which has 21 questions, is used for screening the symptoms related to these mental illnesses. The main symptoms of depression from a clinical point of view are loss of memory; lack of concentration; an inability to make decisions; loss of interest in recreational activities and hobbies including sex; overeating and weight gain; 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 GAD (Generalised Anxiety Disorder) are irritability, nervousness, fatigue, insomnia, gastro intestinal problems, panic, and a sense of impending danger, increased heart rate, sweating, rapid breathing and difficulty concentrating. The symptoms of stress are feeling upset or agitated, an inability to relax, low energy levels, chronic headaches, frequent overreaction and persistent colds or infections. 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. This paper is structured as follows: Section 2 explores related studies on anxiety, depression and stress along with the methods and techniques that were adopted. Section 3 describes the materials and methodology used in the research herein, whilst Section 4 shows the results that were gained after applying the classification algorithms. Finally, section 5 is the conclusion, which summarises the study in its entirety

II.RELATED WORK

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

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 5, May 2021 ||

| DOI: 10.15680/IJIRCCE.2021.090524 |

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 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 organisation, 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. These days, social media is rapidly turning into a healthcare evaluation tool for predicting various types of illness. Saha et a selected topics and psycholinguistic attributes appearing in posts on the LiveJournal website. These were then inputted into a joint modelling framework, so as to categorise 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 recognise 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. collected tweets from 135 participants recruited from Amazon Mechanical Turk (MTurk) and applied decision tree classification to measure suicide risk. The accuracy level for the prediction of suicide rate was observed to be 92%. 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 recognising 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 , which observed that the context-free model produced the best results for audio (weighted, sequence and multi-model). Depression was also predicted 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. Post-traumatic stress disorder has detected in 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 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 and ensemble machine

III. PROPOSED ALGORITHM

A. *Decision Tree Classifier*

- Decision tree is predictive machine learning algorithm
- Decision tree classifier creates the model by designing the decision tree.
- In this paper Decision tree takes total score and type of test as input variable and final result of severity of disease as target variable and according to that create model and predict result.

B. *Description of the Proposed Algorithm:*

Aim of the proposed algorithm is to predict the result by considering the given dataset and finally gives the appropriate output.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 5, May 2021 ||

| DOI: 10.15680/IJIRCCE.2021.090524 |

Step 1: :Collection of data for designing the dataset:-

Initially sample test of stress, depression and anxiety were conducted on Google form and manually result were calculated almost 100 people data were collected according to that dataset is created. Target variable has five level of severity such as Normal, Mild, Moderate, Severe and Extremely Severe.

Step 2: Model creation and prediction of result :

According dataset created it is found that Decision tree classifier and random forest classifier has more accuracy than any other algorithm hence model were created using Decision tree classifier.
And finally test conducted from UI application Data is feed to the Algorithm and predicts the result.

3.2 ALGORITHM FLOW

Step 1: Create dataset
Step 2: Building an Machine learning Model.
Step 3: Final selection of Algorithm whose accuracy is more
Step 4: Conduct a test using a UI application
Step 5: Data collected feeds to the Model
Step 6: Finally Predicts the Result.

IV.MATERIALS AND METHODOLOGY

This research focused on detecting anxiety, depression and stress using the Depression, Anxiety and Stress Scale questionnaire (DASS 21). Data were collected from a total of 100 participants via Google forms and subsequently classified using five machine learning algorithms – namely Decision Tree, Random Forest Tree, Naïve Bayes, Support Vector Machine and KNN.

4.1 Participants:

This study was conducted on a total of 100 participants aged between 20 and 60 years, both males and females, employed and unemployed and with a wide range of responsibilities from household chores to professional duties who were asked to complete a questionnaire

4.2 Questionnaires:

The data for the study were collected through DASS-21, the Depression, Anxiety and Stress Scale questionnaire. DASS 21 comprises 21 questions, with 7 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 applied to me
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 time.
3 applied to me very much or most of the time

Following the data collection, the participants' responses were encoded using numeric values of 0 to 3, and the scores were then calculated by adding the values associated with each question and the below formula:

score Sum of rating points of each class = *2

Once the final scores had been calculated, these were labelled according to severity – i.e. Normal, Mild, Moderate, Severe and extremely severe as Below Table

	Anxiety	Depression	Stress
Normal	0-7	0-9	0-14
Mild	8-9	10-13	15-18
Moderate	10-14	14-20	19-25
Severe	15-19	21-27	26-33
Extremely Severe	20+	28+	29+

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 5, May 2021 ||

| DOI: 10.15680/IJIRCCE.2021.090524 |

SYSTEM DESIGN:

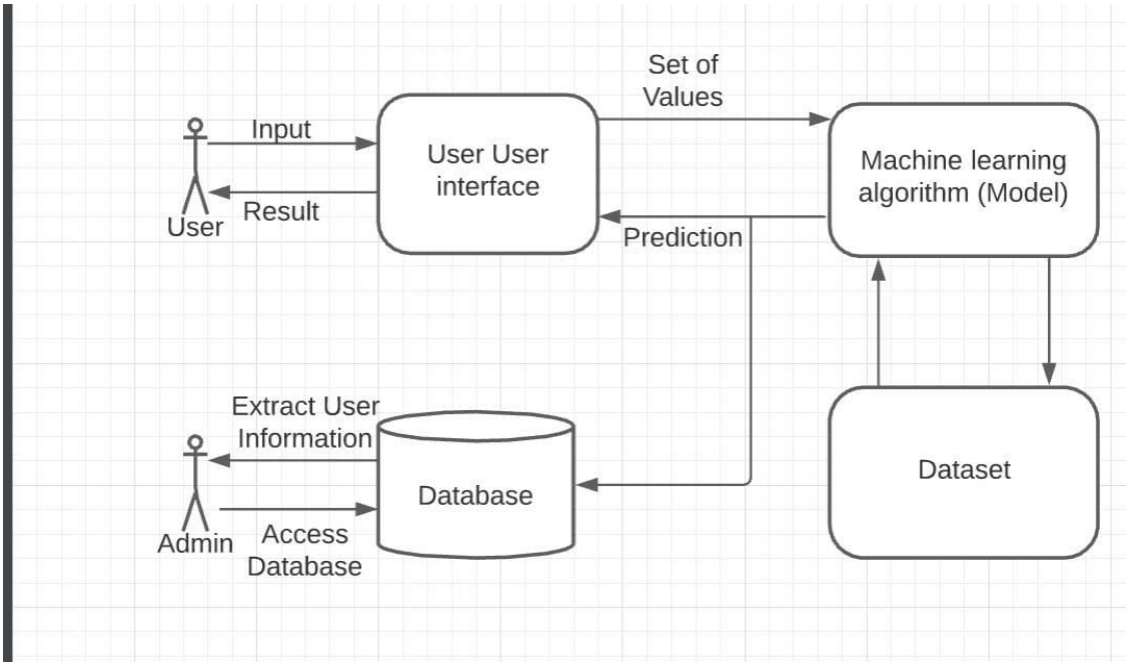


Figure 1 System architecture

V.EXPERIMENTAL RESULTS

1.Registration for test

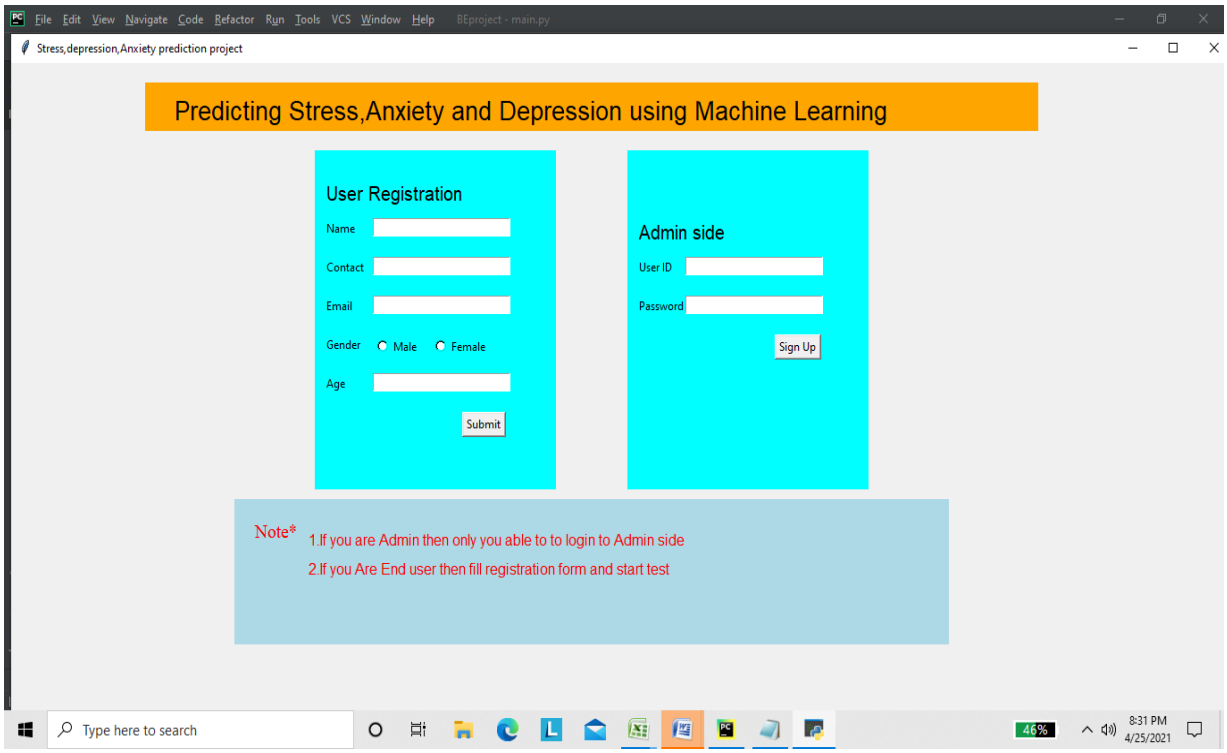


Figure 2 Login Page

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 5, May 2021 ||

| DOI: 10.15680/IJIRCCE.2021.090524 |

2.Test Type

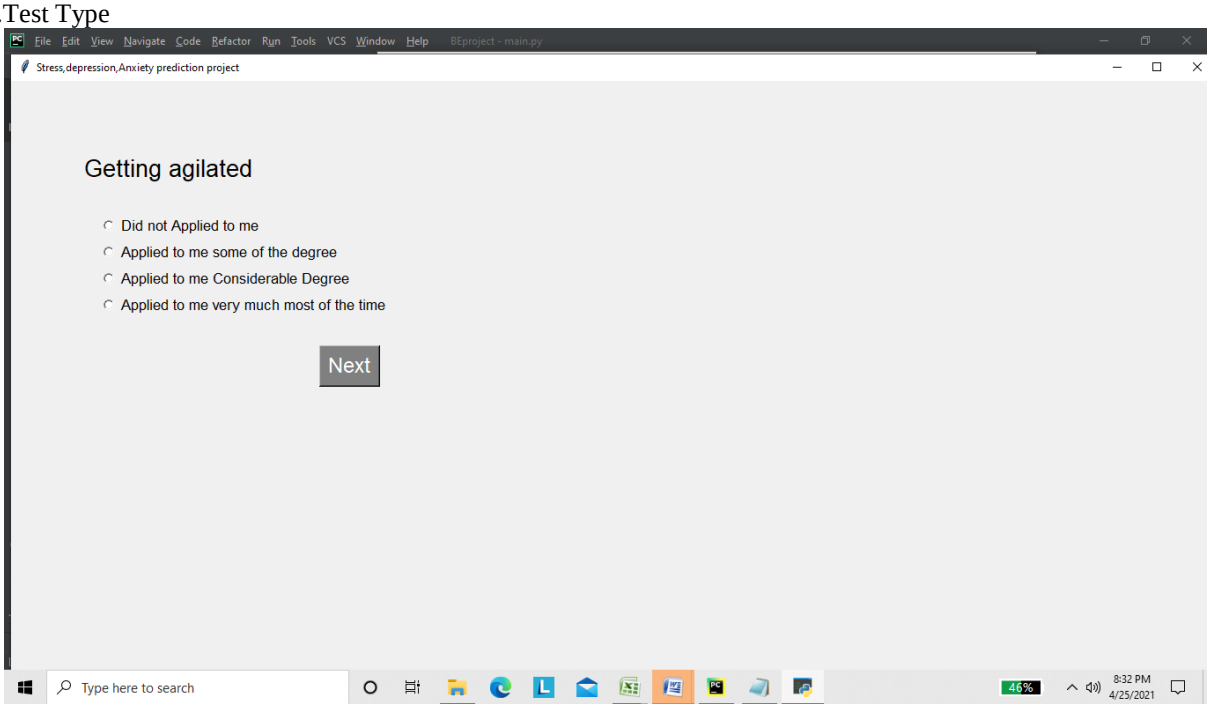


Figure 3 Questions format

3.Final Result

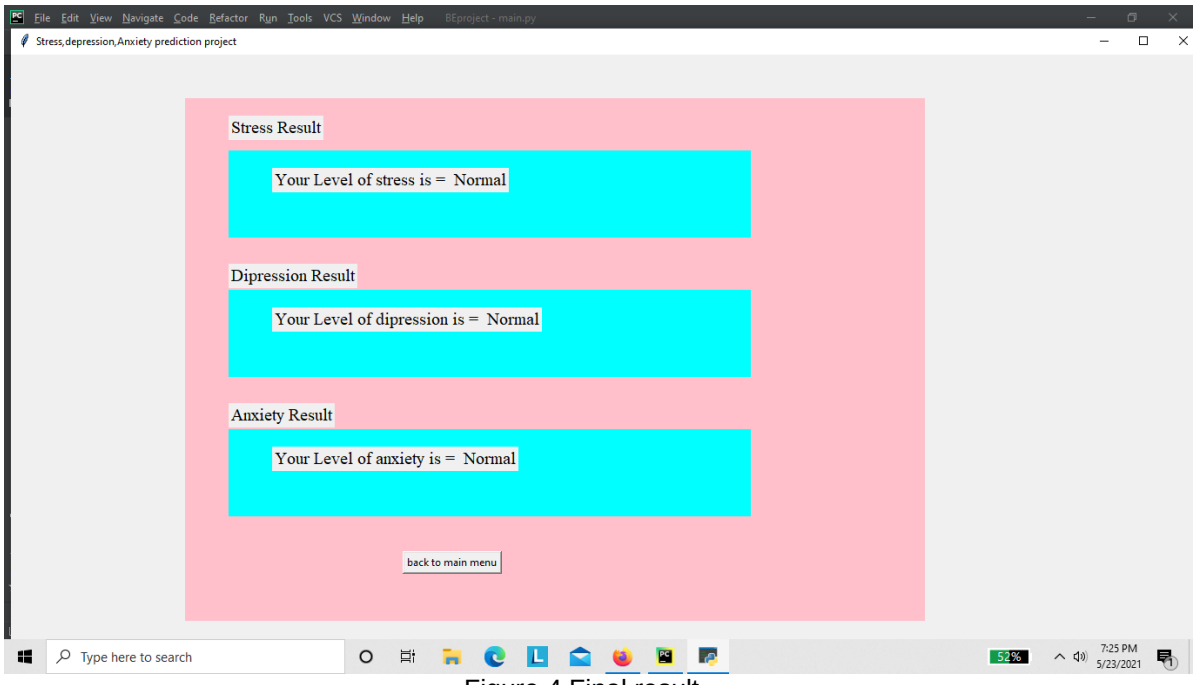


Figure 4 Final result

VI. CONCLUSION AND FUTURE WORK

In this paper, machine learning algorithms were applied to determine five different severity levels of anxiety, depression and stress. Data were collected using a standard questionnaire measuring the common symptoms of anxiety, depression and stress (DASS-21). Subsequently, five different classification techniques were applied – Decision Tree (DT), Random Forest Tree (RFT), Naïve Bayes, Support Vector Machine (SVM) and K- Nearest Neighbour (KNN). The accuracy of naïve Bayes was found to be the highest,

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 5, May 2021 ||

| DOI: 10.15680/IJIRCCE.2021.090524 |

although Random Forest was identified as the best model. Due to the fact that this problem produced imbalanced classes, the best-model selection was made on the basis of the f1 score, which is used for cases of imbalanced partitioning. The important variables were found to be 'scared_without_any_good_reason', 'Life_was_meaningless' and 'Difficult_to_relax' for the scales of Anxiety, Depression and Stress, respectively. As such, these variables were considered to be most important in detecting psychological disorder.

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.

INNO SPACE

SJIF Scientific Journal Impact Factor

Impact Factor:
7.488

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details