International Journal of Innovative Research in Engineering
Volume 3, Issue 2 (March-April 2022), PP: 124-127.
www.theijire.com **ISSN No: 2582-8746**

Mental Health Issues Prediction during Covid Pandemic Using Machine Learning

Bhaskar Kumbhar[1], Khadija Morbiwala[2], Snehal Phapale[3], Radiowalla Ateef Imran[4]
[1,2,3,4]*Department of Information Technology, M.H. Saboo Siddik College of Engineering, Byculla, Mumbai – 400008, India.*

How to cite this paper:
Bhaskar Kumbhar[1], Khadija Morbiwala[2], Snehal Phapale[3], Radiowalla Ateef Imran[4], "Mental Health Issues Prediction during Covid Pandemic Using Machine Learning", IJIRE-V3I02-124-127.

Abstract: *It is been observed that during the COVID-19 pandemic, schools and colleges were forced to close, and many students have experienced severe or mild symptoms of mental health disorders. Mental illness is a disorder characterized by anxiety, depression, and stress. It is becoming more prevalent due to the increasing number of health problems. In this paper, we aim to conduct a timely assessment of the effects of the COVID-19 pandemic on the mental health of students. This research work proposes to get user inputs and display results about how good or badly has the pandemic affected the mental health of the person. The dataset used is related to social life and mental health of students, it will help raise awareness among them and provide greater attention to work related stress, depression and proper benefits can be provided to students suffering from a mental illness. The data used contains different labels such as age, Mode of online classes, time spend on fitness, self-study, sleep etc. Decision Tree classification algorithm was applied for classification of the data as it was found to be more accurate. The research presents an analysis of applying the aforementioned machine learning algorithm on the target groups and also suggests directions for future work.*

Keyword- *COVID-19, Decision Tree classification, Machine learning, Mental Health, Students.*

I. INTRODUCTION

Mental health refers to cognitive, behavioural, and emotional [1] well-being. It affects how we think, feel, and act. It also helps determine how we handle stress, relate to others, and make choices [2]. Mental health is important at every stage of life, from childhood and adolescence through adulthood [3]. According to the World Health Organization (WHO) 'Mental health is a state of well-being in which an individual realizes his or her own abilities, can cope with the normal stresses of life, can work productively, and is able to make a contribution [4] to his or her community.' Mental illness, also called mental health disorders, refers to a wide range of mental health conditions — disorders that affect your mood, thinking and [5] behaviour. Examples of mental illness include depression, anxiety disorders, schizophrenia, eating disorders and addictive behaviours [6]. These conditions deeply impact day-to-day living and may also affect the ability to relate to others.

A study conducted in Bangalore in which a total of 324 college students participated, of whom 180 (55.6%) were male and 144 (44.4%) were female. After assessment of the psychometric scales, it was found that of the 324 students, 223 (68.8%) had high fear of COVID-19, 93 (28.7%) had moderate to severe depression, and 167 (51.5%) had mild to severe anxiety. Among the identified risk factors, having a family member who was infected with COVID-19 was significantly associated with anxiety and depression, with P values of .02 and .001, respectively. In addition, the correlations of the Fear of COVID-19 Scale with the Generalized Anxiety Disorder-7 scale and the Patient Health Questionnaire-9 were found to be 0.492 and 0.474, respectively [7].

The symptoms of mental illness in young adults are Confused thinking, Prolonged depression (sadness or irritability), Feelings of extreme highs and lows, Excessive fears, worries and anxieties, social withdrawal, Dramatic changes in eating or sleeping habits, Strong feelings of anger, strange thoughts (delusions), Seeing or hearing things that [8] aren't there (hallucinations), Suicidal thoughts, etc. It's tough to diagnose a mental illness because a misdiagnosis might lead to catastrophic consequences. As a result, proper attention must be made to accurately identify and treat the mental health issue. For our project, we have collected data from online available dataset on Kaggle. The dataset mainly consists of data of students. Young adults will benefit most from it since it will raise awareness about study-related mental illness. We have applied machine learning algorithm to create a model. It has been implemented on a website for users to get information about their mental illness. The webpage shows a probability and recommendation to the user based on the inputs provided.

II. RELATED WORK

In [9] M. P. Dooshima et al. used demographic, biological, psychological, and environmental factors to predict outcomes. To validate the obtained parameters, various mental health experts were consulted. The goal of this study was to create a prediction model based on identified risk factors in order to classify the risk of mental illness in the study's participants.

In [10] U. S. Reddy et al. have compared the relationship between various parameters in the dataset and used various algorithms to find the most accurate one.

In [11] Anu Priya et al. have applied machine learning algorithms 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. 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, although Random Forest was identified as the best model.

In [12] Alonso et al. have carried out a review of the published works related to techniques and algorithms of Data Mining in Mental Health until March 2018. To carry out the review, the scientific databases were used: Google Scholar, IEEE Xplore, PubMed Science Direct, Scopus and Web of Science.

In [13] Li, Xiaowei et al. have focused on automated method that predicts clinical outcomes in depression is essential for increasing the accuracy of depression recognition and treatments. This paper aims at better recognizing depression using the transformation of EEG features and machine learning methods.

***Table No. 1**: Summary of Related Work*

SN	Paper	Algorithms Used	Advantages and Disadvantages
1.	A Predictive Model for the Risk of Mental Illness in Nigeria Using Data Mining	Decision Tree (DT) and Naïve Bayes	Advantages: Two different supervised machine learning algorithms were used to formulate the predictive model for the risk of mental illness, namely: Naïve Bayes' and Decision Trees Classifiers. Disadvantages: They were used to train the development of the prediction model using the dataset containing ONLY 30 patients' risk factor records. This research does not focus on the work predominantly.
2.	Behavioral Modeling for Mental Health using Machine Learning Algorithms	Random forest, KNN, Naive bayes, Logistic regression and Decision tree	Advantages: This paper provides an intuitive understanding of the mental health analysis amongst different target groups. Here all types of algorithms like Random forest, KNN, Naive bayes, Logistic regression and Decision tree were used. Disadvantages: This paper focuses mainly on depression, stress and anxiety and only some aspects of work-related stress are considered.
3.	Machine Learning Techniques for Stress Prediction in Working Employees	Performance Metrics, Logistic Regression KNN Classifier Decision Trees Random Forest Classifier	Advantages: In this paper, it was found that people working in a tech company were slightly more at the risk of developing stress even if their role was not tech-based. These insights can be effectively used by corporates to frame better HR policies for their employees. Disadvantages: In this paper, accuracy of algorithm is only found out.
4.	Predicting Anxiety, Depression and Stress in Modern Life using Machine Learning Algorithms	Decision Tree (DT), Random Forest Tree (RFT), Naïve Bayes, Support Vector Machine (SVM) and K- Nearest Neighbor (KNN).	Advantages: In this paper, machine learning algorithms were applied to determine five different severity levels of anxiety, depression and stress. Disadvantage: It was found that classes were imbalanced.

Various researchers have used different machine learning algorithms to predict mental illnesses, and the performance of different algorithms has been found to vary depending on the scenario; no [14] single algorithm has been determined to be most suitable in all cases. As a result, in the current study, Decision Tree Algorithm is used to identify symptoms of anxiety, depression, and stress.

Decision Tree Algorithm:

Decision Trees are a type of Supervised Machine Learning where the data is continuously split according to a certain parameter. The tree can be explained by two entities, namely decision nodes and leaves. The leaves are the decisions or the final outcomes. And the decision nodes are where the data is split.

III. IMPLEMENTED SYSTEM

Dataset:

In this study, a cross-sectional survey is conducted with a sample size of 1182 students of different age groups from different educational institutions in Delhi National Capital Region (NCR).

Implementation:

Based on the above dataset, we have created a system with the primary aim of developing a web application where users can answer basic questions related to their social life. First, data analysis was carried out. Second, we visualised the data, and finally, we pre-processed it and label encoded it for better prediction results. We implemented decision tree and random forest algorithm to find out the more accurate algorithm. Decision tree algorithm gave us more accurate results; hence it was applied in our classification module. After finding out various insights that the data revealed, we created a model based [15] on the above-mentioned algorithm and used it on our web application designed.

In order to achieve our goal, we created a website where a user can simply fill out a form with questions based on the dataset gathered. The user will answer the questionnaire and the system will predict a result of his/her mental condition on the website. The prediction will be carried out with the help of the machine learning [16] model created.

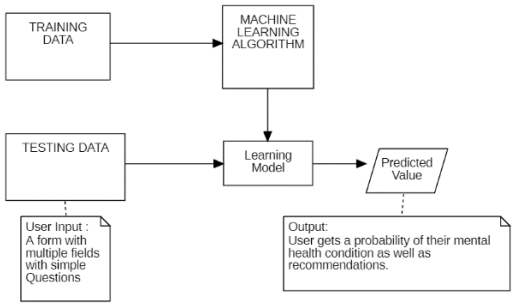


Fig. 1: Proposed system architecture

Since this project uses student's basic lifestyle questions during pandemic, it is very easy to answer and it will also create awareness among other students and encourage them to compare their result scores. This will benefit the students and parents with early diagnosis of a mental problem too.

IV. RESULT AND ANALYSIS

Fig 2, Fig 3 and Fig 4 are the outputs based on different inputs that are accepted from the user through a form in website. There is a different recommendation for different values of the probability of mental illness.

Fig.2: Output with Probability 0.31

Fig. 3: Output with Probability 0.44

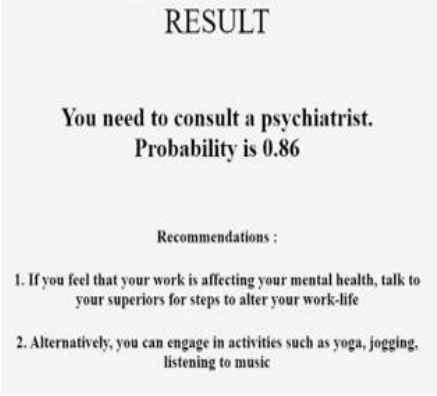
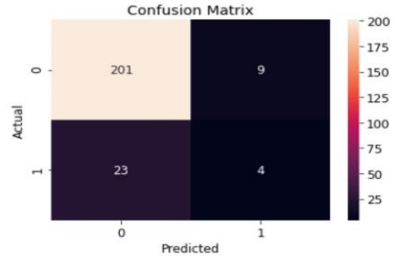


Fig. 4: Output with Probability 0.86




```
Classification Accuracy: 0.8649789029535865
Classification Error: 0.13502109704641352
False Positive Rate: 0.04285714285714286
Precision: 0.3076923076923077
AUC Score: 0.5526455026455026
Cross-validated AUC: 0.5975892977122317
```

Fig. 5: Confusion Matrix

We have designed the system in a way that a probability less than 0.30 indicates that the user does not suffer from a mental condition. Probability between 0.3 and 0.63 indicates that the user may face a mental illness in the future and a probability greater than 0.63 shows that he/she suffers from mental illness.

Fig 5 shows the confusion matrix denoting the accuracy of the algorithm. It shows that 9 instances have been correctly classified as Positive while 201 have been correctly classified as negative.

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

Using the values in fig 2 we get the accuracy of the Decision Tree algorithm as 0.86 i.e., 86%.

V. CONCLUSION

Student mental health in higher education has been an increasing concern. The COVID-19 pandemic situation has brought this vulnerable population into renewed focus. Multiple stressors were identified that contributed to the increased levels of stress, anxiety, and depressive thoughts among students [17]. This included fear and worry about their own health and of their loved ones. This study aims to conduct a timely assessment of the effects of the COVID-19 pandemic on the mental health of college students [18]. To do the prediction, we have trained the data first. The decision tree algorithm was then used to train a model. A prediction accuracy of 86% indicates that the use of this Machine Learning method for stress and mental health condition prediction yields significant results. When a user answers the questionnaire on our website, he or she obtains a probability of their mental health condition as well as recommendations. People who are hesitant to approach humans for mental health diagnosis can also use our model.

VI. FUTURE SCOPE

In the future, we may be able to develop a system that predicts a particular mental illness that an individual is suffering from. Also, because the number of responses in our case is limited, a much more specific and larger dataset can be used as a training model [19]. We can also customise the survey taken in order to obtain responses in [20] the appropriate format and to increase the number of attributes based on relevance [21]. Moreover, different machine learning techniques like KNN Classifier, Naive Bayes classifier, Logistic Regression can be used to analyse the model's efficiency and increase accuracy of the model.

References

[1] Mhambe Priscilla Dooshima, Egejuru Ngozi Chidozie, Balogun Jeremiah Ademola, Olusanya Olayinka Sekoni, Idowu Peter Adebayo. "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. doi: 10.11648/j.iji.20180601.12

[2] U. S. Reddy, A. V. Thota and A. Dharun, "Machine Learning Techniques for Stress Prediction in Working Employees," 2018 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC), Madurai, India, 2018, pp. 1-4

[3] Anu Priya, Shruti Garg, Neha Prerna Tigga, "Predicting Anxiety, Depression and Stress in Modern Life using Machine Learning Algorithms" 2019 International Conference on Computational Intelligence and Data Science (ICCIDS), Mesra, Ranchi-835215, India.

[4] Alonso, Susel Góngora; de la Torre-Díez, Isabel; Hamrioui, Sofiane; López-Coronado, Miguel; Barreno, Diego Calvo; Nozaleda, Lola Morón; Franco, Manuel (2018). Data Mining Algorithms and Techniques in Mental Health: A Systematic Review. Journal of Medical Systems, 42(9), 161– . doi:10.1007/s10916-018-1018-2

[5] Li, Xiaowei; Zhang, Xin; Zhu, Jing; Mao, Wandeng; Sun, Shuting; Wang, Zihan; Xia, Chen; Hu, Bin (2019). Depression recognition using machine learning methods with different feature generation strategies. Artificial Intelligence in Medicine, (), S0933365719300296– . doi:10.1016/j.artmed.2019.07.004

[6] https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7473764/

[7] https://www.kaggle.com/kunal28chaturvedi/covid19-and-its-impact-on-students

[8] https://www.kaggle.com/rohan2002/mentalhealth-treatment prediction?select=tuned_mental_health_random.h5

[9] D.Filip & C. Jesus. (2015). A Neural Network Based Model for Predicting Psychological Conditions International Conference on Brain Informatics and Health 252-261.

[10] https://pubmed.ncbi.nlm.nih.gov/32805704//