

A Study and Prediction of Psychological Disorders Through Machine Learning

Sandeep Mishra¹, Garima Sharma², Chirag Rawat³, Khushi Mishra⁴,
Krishna Aggarwal⁵, Kaustubh Kumar Shukla^{6,*}

Abstract

Physical illness is very much visible but not psychological illness therefore, it requires more attention and care. Psychological disorders also known as psychiatric disorders refer to a wide range of conditions affecting a person's thought process, leading to significant changes in the behavior of an individual. The most prevalent psychological disorders include depression, anxiety disorders, and post-traumatic stress disorder (PTSD). Symptoms of psychological disorders vary greatly but include common symptoms like changes in mood and cognition. These disorders undermine an individual ability to function effectively in daily life, impacting their well-being. An individual's behavior and way of thinking are impacted by psychological disorders such as PTSD, anxiety, and depression, which affect their general well-being. They need the right medication, counseling, and assistance. Algorithms for machine learning are essential for both diagnosing and treating these conditions. Through the analysis of vast datasets, the study looks into two widely used supervised learning models, random forest, and decision trees, which can accurately detect and comprehend the origin and course of various conditions. By using these models in healthcare, professionals will be better able to identify these conditions and understand their underlying causes.

Keywords: Psychiatric disorders, random forest, decision trees, supervised, anxiety, depression, stress, treatment, prediction, algorithms

INTRODUCTION

Psychological or mental disorders are problems that affect a person's thoughts, feelings, and actions, resulting in discomfort and difficulties with day-to-day functioning. These diseases can vary in severity, ranging from more serious ailments, such as schizophrenia and bipolar disorders, to mood disorders, such as anxiety and sadness. These conditions can arise due to a combination of factors such as environmental stresses, life events, brain chemistry, and heredity. Therapy, medication, and lifestyle modifications are often used to control symptoms and enhance quality of life. Effective treatment of psychological problems and the advancement of general well-being depend heavily on early identification, professional intervention, and support from mental health providers [1].

*Author for Correspondence

Kaustubh Kumar Shukla
E-mail: dr.kkshuklaece@gmail.com

¹Assistant Professor, Department of Computer Science and Engineering, Dronacharya Group of Institutions, Greater Noida, Uttar Pradesh, India

²⁻⁵Student, Department of Computer Science and Engineering, Dronacharya Group of Institutions, Greater Noida, Uttar Pradesh, India

⁶Associate Professor, Department of Electronics and Communication Engineering, Dronacharya Group of Institutions, Greater Noida, Uttar Pradesh, India

Received Date: October 19, 2024

Accepted Date: October 22, 2024

Published Date: October 25, 2024

Citation: Sandeep Mishra, Garima Sharma, Chirag Rawat, Khushi Mishra, Krishna Aggarwal, Kaustubh Kumar Shukla. A Study and Prediction of Psychological Disorders Through Machine Learning. International Journal of Algorithms Design and Analysis Review. 2024; 2(2): 32–38p.

Machine Learning in the field of Artificial Intelligence has shown potential for predicting psychiatric diseases using various data sources [2]. Machine learning algorithms can detect trends and

forecast the probability of acquiring a psychiatric condition by analyzing massive datasets of clinical data including patient demographics, medical history, and symptoms [3]. These algorithms are trained to identify minor patterns and relationships that may not be noticeable to human doctors using historical data from patients with established diseases. Beyond clinical data, machine learning models may also make use of a wide range of data sources, including genetic markers, neuroimaging data, and social media activity. Combining these data sources can yield more thorough knowledge of a person's chance of developing psychiatric disorders.

One of the popular methods for using machine learning in prediction is supervised learning, in which the system is trained using labeled data, where each data point is connected to a particular diagnostic or result. By linking the input properties to the matching output labels, the system can anticipate new, unknown data. Supervised learning models include linear regression, decision trees, and k-nearest neighbors. The second alternative strategy is unsupervised learning, which includes linear discriminant analysis, in which the computer finds clusters and patterns in the data without the need for predetermined labels [4]. The third category is ensemble learning, which combines multiple individual models to improve overall prediction performance [5]. The ensemble learning model helps to achieve better accuracy, robustness, and generalization by integrating the prediction data of others. Bagging, stacking, random forest, boosting, etc. are the models that fall under the category of ensemble learning, as shown in Figure 1.

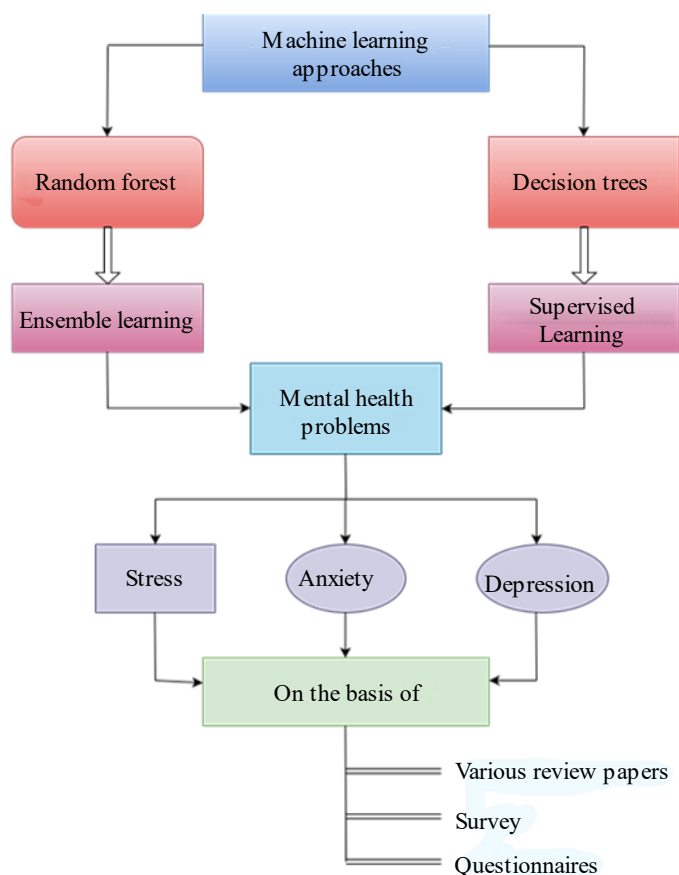


Figure 1. Machine learning approaches detecting mental health problems.

LITERATURE REVIEW

Numerous scholars have utilized random forest and decision tree models to predict a range of psychiatric conditions, such as anxiety, depression, bipolar disorder, schizophrenia, and post-traumatic stress disorder (PTSD) [6]. For instance, Leightley et al. (2019) [7] proposed a random forest model to showcase the likelihood that military soldiers who have experienced battle stress would acquire

PTSD. In addition, Ramos-Lima et al. [8] presented the application of machine learning techniques to predict PTSD survivors in a civil war in Uganda. They used a sample of 441 trauma-exposed subjects as the training dataset and 211 trauma-exposed subjects as the testing dataset [9]. Machine learning techniques, such as random forest with conditional inference, least absolute shrinkage, selection (LASSO), and logistic regression, are being applied to predict PTSD survivors. Based on the results obtained, using the random forest with conditional inference showed the highest accuracy of 77.25% compared with LASSO with an accuracy percentage of 74.88%, and logistic regression with an accuracy percentage of 75.36% [10]. Carozza et al. [11] collected speech samples from warzone-exposed veterans. Then, speech attributes that could help in predicting PTSD, such as slower monotonous speech and less change in tonality, were extracted from the clinical interviews. Random forest was used in the prediction, and the model achieved an accuracy of 89.1% with an AUC of 0.954. According to the study, the model identified people at risk for PTSD with high accuracy, demonstrating the usefulness of random forest for early identification [12]. Similarly, Lee et al. (2021) [13] used the decision tree model to forecast the intensity of depression symptoms in teenagers based on clinical, genetic, and demographic variables. The researchers showed that the decision tree model could successfully divide people into risk categories, offering insightful information for individualized treatment planning. Yuan H, Fan XS, et al. (2019) [9] developed a risk prediction model using a random forest approach from the Nationwide Inpatient Sample (NIS), which included 8 million hospital stays and 126 clinical and non-clinical data. The model was evaluated using SVM, random forest, Bagging, and Boosting algorithms after imbalanced data produced undesirable results. Gurjar et al. [14] collected data from 630 elderly individuals from a medical college and hospital in Kolkata, West Bengal, including 520 in special care. They applied various classification methods, including logistic, multiple layer perceptron, naive Bayes, random forest, random tree, J48, sequential random optimization, random subspace, and K-star. They found that random forest had the best accuracy rates of 91% and 89%. Hussein et al. (2012) [15] presented an efficient recommendation system for chronic diseases using data mining and decision tree approaches. The model, tested with 20 samples, found that RF outperformed other algorithms, indicating the effectiveness of data mining in improving the prediction accuracy.

Based on dataset segregation factors, academics have demonstrated great accuracy in making predictions using SVM, neural networks, and logistic regression models [16]. An SVM model was presented by Shi X, Jiang D, et al. (2019) [17] to forecast the probability of schizophrenia based on medical history in the family. Liu et al. [18] developed a logistic regression model to predict the chance of relapses in patients receiving maintenance therapy for bipolar disease. By recognizing signs associated with recurrence, these models can aid medical professionals in monitoring and treating patients' symptoms, thereby proving their predictive power. Therefore, popular researchers in this field concluded that by leveraging the random forest model and the decision tree model along with other proposed supervised and unsupervised machine learning models, researchers can enhance the accuracy and generalizability of predictive models, making the prediction rate higher in the psychology field [19]. The use of these models has reformed the healthcare industry, and further research is needed to refine these proposed models by addressing their limitations and integrating them into clinical practice.

METHODOLOGY

Researchers employ various techniques to forecast psychiatric diseases [20]. These include preprocessing the data to ensure its quality and relevance; selecting features using methods such as recursive feature elimination; training the model via cross-validation; hyperparameter tuning through grid search or random search; evaluation of the model using metrics such as accuracy, precision, recall, F1 score, and ROC-AUC; and interpreting the results using methods such as feature importance plots or decision tree visualization. These techniques aid in the development of reliable and precise prediction models for diagnosing psychiatric diseases [21]. Researchers can learn more about the characteristics that are most crucial for predicting psychiatric diseases by merging these approaches [22–25]. With the

implementation of these approaches, the accuracy rate and precision percentage of prediction for these models were determined for different disorders.

To obtain a more comprehensive understanding of the current state of psychological disorders, we asked respondents of different age groups about their experiences with these conditions, the effects on their sleep patterns, whether they had previously been diagnosed with these conditions, and how stressful they felt overall on a scale of one to ten [25–28]. Approximately 130 persons submitted answer records to the poll, which included a maximum of 30 questions. The results section contains a predicted overview of all the data that were produced.

RESULTS

On analyzing various research papers, the accuracy and precision percentages of the random forest model and the decision trees model for psychiatric diseases depression, stress, and anxiety are depicted in Table 1 [22, 24, 26].

Table 1. Accuracy and precision percentages of the random forest model.

S.N.	Psychiatric Disorder	Model	Accuracy Rate	Precision Percentage
1.	Depression	Random forest	92.7	92.7
		Decision trees	79.5	79.6
2.	Stress	Random forest	88.8	88.8
		Decision trees	79.4	75.1
3.	Anxiety	Random forest	84.8	84.3
		Decision trees	73.5	73.3

Here, the random forest algorithm helped to analyze depression at 92.6%, anxiety at 88.7%, and stress at 84%.

Survey Results

Based on the questionnaire survey conducted, the ensuing deductions are concluded, as shown in Figure 2.

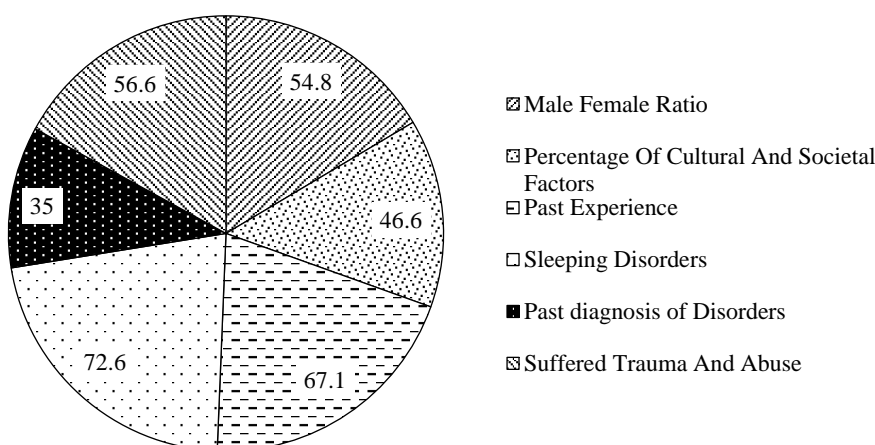


Figure 2. Mental health survey analysis.

The study's male-to-female participation ratio was approximately 54.8:45.2, with 57% of participants in the 19–20 age range. Interestingly, the findings revealed that almost 63% of people in this age group rated their stress levels as rating 7–8 out of 10. In addition, 67.1% of respondents reported having ever had abrupt mood swings, 72.6% reported having sleep issues, such as insomnia, and 59.2% reported

finding it difficult to maintain positive relationships with others that affect emotional and physical well-being [29]. Approximately 56.6% of the individuals experience trauma and/or abuse throughout their lifetime. Furthermore, 46.6% of respondents believed that cultural and societal influences had a significant impact on their mental health [30]. The fact that 35% of individuals have been diagnosed with these diseases is disheartening. The difficulties people face daily that hurt their physical and mental health were represented in this questionnaire survey, which was completed in the present study [31–33].

CONCLUSION

In summary, the prediction of psychiatric diseases can be greatly enhanced by the application of machine learning techniques, such as decision trees and random forests. These algorithms assist healthcare providers in making quicker and more accurate diagnoses by sifting through massive datasets and searching for patterns in the data. Compared to conventional diagnostic techniques, random forests, and decision trees provide several benefits. They are well suited for analyzing a wide variety of elements that lead to psychiatric problems because they can manage complicated and high-dimensional data. Moreover, by offering insights into the fundamental connections between many factors, these algorithms can aid doctors in comprehending the underlying causes of these illnesses. In addition, compared to conventional techniques that depend on subjective assessment, machine learning algorithms provide a more consistent and objective approach to diagnosis. These algorithms can offer a more complete picture of a person's risk of developing a psychiatric condition by combining a variety of data sources such as behavioral, environmental, and genetic components. Further research and development in this field may result in the development of more advanced and customized prediction models, which may enhance mental health outcomes and facilitate individualized treatment plans. The potential for improving diagnostic precision, treatment efficacy, and patient care in psychology and psychiatry through the use of machine learning algorithms in mental health practices is substantial.

REFERENCES

1. Vaishnavi K, Nikhitha Kamath U, Ashwath Rao B, Subba Reddy NV. Predicting mental health illness using machine learning algorithms. *J Phys Conf Ser. IOP Publishing*, 2161. DOI: 10.1088/1742-6596/2161/1/012021.
2. Pintelas EG, Kotsilieris T, Livieris IE, Pintelas P. A review of machine learning prediction methods for anxiety disorders. *Proceedings of the 8th International Conference on Software Development and Technologies for Enhancing Accessibility and Fighting Info-Exclusion*. New York (NY): Association for Computing Machinery; 2018. p. 8–15. doi:10.1145/3218585.3218587.
3. Priya A, Garg S, Tigga NP. Predicting anxiety, depression and stress in modern life using machine learning algorithms. *Procedia Comput Sci*. 2020;167:1258–67. DOI: 10.1016/j.procs.2020.03.442.
4. van der Meer D, Hoekstra PJ, Van Donkelaar M, Bralten J, Oosterlaan J, Heslenfeld D, et al. Predicting attention-deficit/hyperactivity disorder severity from psychosocial stress and stress-response genes: A random forest regression approach. *Transl Psychiatry*. 2017;7:e1145–. DOI: 10.1038/tp.2017.114. PubMed: 28585928.
5. Acharya N, Kar P, Ally M, Soar J. Predicting co-occurring mental health and substance use disorders in women: An automated machine learning approach. *Appl Sci*. 2024;14:1630. DOI: 10.3390/app14041630.
6. Luo J, Chen Y, Tao Y, Xu Y, Yu K, Liu R, et al. Major depressive disorder prediction based on sleep-wake disorders symptoms in US adolescents: A machine learning approach from national sleep research resource. *Psychol Res Behav Manag*. 2024;17:691–703. DOI: 10.2147/PRBM.S453046. PubMed: 38410378.
7. Leightley D, Williamson V, Darby J, Fear NT. Identifying probable post-traumatic stress disorder: Applying supervised machine learning to data from a UK military cohort. *J Ment Health*. 2019;28:34–41. DOI: 10.1080/09638237.2018.1521946. PubMed: 30445899.
8. Ramos-Lima LF, Waikamp V, Antonelli-Salgado T, Passos IC, Freitas LHM. The use of machine learning techniques in trauma-related disorders: A systematic review. *J Psychiatr Res*. 2020;121:159–72. DOI: 10.1016/j.jpsychires.2019.12.001. PubMed: 31830722.

9. Yuan H, Fan XS, Jin Y, He JX, Gui Y, Song LY, et al. Development of heart failure risk prediction models based on a multi-marker approach using random forest algorithms. *Chin Med J (Engl)*. 2019;132:819–26. DOI: 10.1097/CM9.000000000000149. PubMed: 30829708.
10. Goutam D, Rani V, Saini HS. Mental health illness prediction with hybrid machine learning approach. *Int Res J Mod Eng Technol Sci*. 2024;6:3893–8.
11. Carozza L. Post-traumatic stress disorder and cognitive communication effects. In: Hyter YD, editor. *Language research in post-traumatic stress*. London: Routledge; 2024. pp. 227–54. DOI: 10.4324/9781003225270.
12. Jayasri Devi T, Gopi A. An efficient novel approach for early detection of mental health disorders through distributed machine learning paradigms from public societal communication. *Int J Intell Syst Appl Eng*. 2023;12:767–78.
13. Lee Y, Kim H, Lee Y, Jeong H. Comparison of the prediction model of adolescents' suicide attempt using logistic regression and decision tree: Secondary data analysis of the 2019 youth health risk behavior web-based survey. *J Korean Acad Nurs*. 2021;51:40–53. DOI: 10.4040/jkan.20207. PubMed: 33706330.
14. Gurjar S, Patil C, Suryawanshi R, Adadande M, Khore A, Tarapore N. Mental health prediction using machine learning. *Int Res J Eng Technol*. 2022;9(12):1117–1121.
15. Hussein AS, Omar WM, Li X, Ati M. Efficient chronic disease diagnosis prediction and recommendation system. 2012 IEEE-EMBS Conference on Biomedical Engineering and Sciences, Langkawi, Malaysia, 2012, pp. 209–14. DOI: 10.1109/IECBES.2012.6498117.
16. CACHEDA F, FERNANDEZ D, NOVOA FJ, CARNEIRO V. Early detection of depression: Social network analysis and random forest techniques. *J Med Internet Res*. 2019;21:e12554. DOI: 10.2196/12554. PubMed: 31199323.
17. Shi X, Jiang D, Huang Y, Wang X, Chen Q, Yan J, Tang B. Family history information extraction via deep joint learning. *BMC Med Inform Decis Mak*. 2019;19(Suppl 10):277. DOI: 10.1186/s12911-019-0995-5, PubMed: 31881967.
18. Liu B, Zhang Y, Fang H, Liu J, Liu T, Li L. Efficacy and safety of long-term antidepressant treatment for bipolar disorders – A meta-analysis of randomized controlled trials. *J Affect Disord*. 2017;223:41–48. DOI: 10.1016/j.jad.2017.07.023, PubMed: 28715727.
19. Airlangga G. Evaluating machine learning models for mental health diagnostics: A comparative analysis and visual insights. *KLIK: Kajian Ilmiah Inform Komputer*. 2024;4:2058–68.
20. Sau A, Phadikar S, Bhakta I. Prediction of spirometry parameters of adult Indian population using machine learning technology. *Multimed Tools Appl*. 2024;83:78023–78057. DOI: 10.1007/s11042-024-18470-y.
21. Jafari M, Sadeghi D, Shoeibi A, Alinejad-Rokny H, Beheshti A, García DL, et al. Empowering precision medicine: AI-driven schizophrenia diagnosis via EEG signals: A comprehensive review from 2002–2023. *Appl Intell*. 2024;54:35–79. DOI: 10.1007/s10489-023-05155-6.
22. Chaudhary A, Agarwal A, Ansari A, Arya D. Psychological disorder analysis using machine learning. *J Pharm Negat Results*. 2022;3326–3333.
23. Harris MN. Violent victimization of youth with mental disorders: Does lifestyles/routine activities or control perspectives mediate the relationship between mental illness and victimization? *Crime Delinq*. 2024. DOI: 10.1177/00111287241231747.
24. Gentili E, Franchini G, Zese R, Alberti M, Ferrara M, Domenicano I, et al. Machine learning from real data: A mental health registry case study. *Comput Methods Programs Biomed Update*. 2024;5:100132. DOI: 10.1016/j.cmpbup.2023.100132.
25. Vichare S, Pirjade ST, Parte S, Salunkhe D, Jaunjale S. Mental health prediction and support application. *Grenze Int J Eng Technol*. 2024;10: 2536–2543.
26. Kasaudhan H, Shukla KK, Kushwaha R, Sharma K, Gupta U, Sharma A. Early detection and analysis of lung cancer using artificial intelligence. *IEEE Int Conf Comput Power Commun Technol*. 2024;5:1470–1474. DOI: 10.1109/IC2PCT60090.2024.10486335.
27. Pandey S, Arora SV, Shukla KK, Priyanka P, Nivetha A, Chaudhary A. Optimizing 3D printing to increase the gain and bandwidth of microstrip antennas: A research on implementation and design.

- 2024 Third International Conference on Smart Technologies and Systems for Next Generation Computing (ICSTSN), Villupuram, India, 2024. p. 1–5. DOI: 10.1109/ICSTSN61422.2024.10671029.
28. Shukla KK, Muthumanickam T. Micro cantilever arrays optimization and analysis for healthcare applications. 2022 International Conference on Applied Artificial Intelligence and Computing (ICAAIC), Salem, India, 2022. p. 1681–1689. DOI: 10.1109/ICAAIC53929.2022.9792771.
 29. Peng B, Wang R, Zuo W, Liu H, Deng C, Jing X, et al. Distinct correlation network of clinical characteristics in suicide attempters having adolescent major depressive disorder with non-suicidal self-injury. *Transl Psychiatry*. 2024;14:134. DOI: 10.1038/s41398-024-02843-w, PubMed: 38443348.
 30. Cohen TR, Fronk GE, Kiehl KA, Curtin JJ, Koenigs M. Clarifying the relationship between mental illness and recidivism using machine learning: A retrospective study. *PLOS One*. 2024;19:e0297448. DOI: 10.1371/journal.pone.0297448. PubMed: 38394314.
 31. Katiyar K, Fatma H, Singh S. Predicting anxiety, depression, and stress in women using machine learning algorithms. In: Hemanth D, Gupta M, editors. *Combating Women's Health Issues with Machine Learning: Challenges and Solutions*. Boca Raton (FL): CRC Press; 2023. pp. 22–40. DOI: 10.1201/9781003378556.
 32. Song Y, Tian Y, Fan C, Zheng Q, Huang L, Zhou Z. Employing decision trees to predict cyberbullying victimization among Chinese adolescents and identify subgroups and their shared characteristics. *Curr Psychol*. 2024;43:19518-19531. DOI: 10.1007/s12144-024-05696-0.
 33. Menzies RE, Richmond B, Sharpe L, Skeggs A, Liu J, Coutts-Bain D. The ‘revolving door’ of mental illness: A meta-analysis and systematic review of current versus lifetime rates of psychological disorders. *Br J Clin Psychol*. 2024;63:178-196. DOI: 10.1111/bjc.12453, PubMed: 38197576.