International Journal of Scientific Research in Computer Science, Engineering and Information Technology

ISSN : 2456-3307

Available Online at : www.ijsrcseit.com

doi : https://doi.org/10.32628/CSEIT23903114

Using Machine Learning to Predict Mental Illness

Rhitik Nimbalkar[1], Rakesh Oza[1], Akash Pabale[1], Khushi Pandey[1], Afsha Akkalkot[2]

[1]TE Students, Department of Computer Engineering, Zeal College of Engineering and Research, Pune, Maharashtra, India

[2]Assistant Professor, Department of Computer Engineering, Zeal College of Engineering and Research, Pune, Maharashtra, India

ARTICLEINFO

Article History:

Accepted: 20 June 2023

Published: 05 July 2023

Publication Issue

Volume 9, Issue 4

July-August-2023

Page Number

01-05

ABSTRACT

Depression and anxiety are omnipresent but largely unnoticeable to us. The statistics indicate that most families on the planet are undoubtedly impacted by them, although many people suffer in silence. Now, this is somewhat due to the difficulty in discussing these diseases, but it's also due to the difficulty in obtaining care. In order to forecast mental health issues, this survey presents a current thorough assessment of machine learning algorithms. We will also go over the difficulties, restrictions, and future directions for using machine learning in the field of mental health.

Keywords : Survey Paper, Machine Learning Algorithms

I. INTRODUCTION

How effectively a person's ailments are treated depends on how well they are mentally. In order to spot any health-related irregularities, it is essential to track the mental health traits of different populations. Contrary to popular belief, stress and sadness can affect people of all ages and socioeconomic statuses. In order to avoid catastrophic illness, it is crucial to periodically evaluate the mental health of different categories. Soon, healthcare practitioners will be required to consider a patient's mental health profile in order to provide better care and hasten recovery.

Chronic illnesses, bipolar disorder, and schizophrenia are some of the most severe mental health conditions that manifest gradually over time and have observable early warning symptoms. These conditions could be avoided or managed more effectively. Early detection of abnormal mental states will allow for the provision of additional care and therapeutic interventions.

They have an impact on men and women of all ages, including small children and the elderly. Anxiety and depression disorders have a significant negative influence on health and wellbeing. They can lead to a variety of somatic symptoms, including gastritis, acid reflux, palpitations, insomnia or hypersomnia, tremors, significant weight loss or gain, as well as a variety of psychosocial symptoms, including low mood, social withdrawal, decreased workplace productivity,

Copyright © 2023 The Author(s): This is an open-access article distributed under the terms of the Creative Commons Attribution **4.0 International License (CC BY-NC 4.0)** which permits unrestricted use, distribution, and reproduction in any medium for non-commercial use provided the original author and source are credited.

suicidal ideation or attempt, and difficulty concentrating.

Smartphones, social media, neuroimaging, and wearable technology have made it possible for medical professionals and mental health researchers to obtain a ton of data quickly. The ability to analyse these data with machine learning has grown. Advanced probabilistic and statistical methods are used in machine learning to build computers that can independently learn from data. This makes it possible to more accurately forecast outcomes from data sources and to more easily and correctly identify data trends. Machine learning has aided fields including natural language processing, speech recognition, computer vision, and artificial intelligence by enabling researchers and developers to extract vital data from datasets, provide individualized experiences, and create intelligent systems.

Social media has given young people a forum to discuss current events and share their personal struggles. Using sentiment analysis on data that has been timely gathered from social networking sites (in this case, Twitter), it is feasible to predict depression, anorexia, and other analogous mental diseases in young people. Data mining, sentiment analysis, or emotion AI is the process of extrapolating illogical information, conclusions, and characteristics from text. It is covered under the NLP (Natural Language Processing) discipline. This area of computer science study is developing quite quickly.

II. METHODS AND MATERIAL

Stacking, logistic regression, K-nearest neighbour classifier, decision tree classifier, and random forest classifier are a few of the machine learning techniques that are used to evaluate the models after that.

Methods	Accuracy (%)
Logistic Regression	79.63
KNeighbours Classifier	80.42
Decision Tree classifier	80.69
Random Forests	81.22
Stacking	81.75

III. RESULTS AND DISCUSSION

It is necessary to indent each paragraph. All paragraphs must be justified, which calls for both left and right alignment.

A. Logistic Regression

Logistic regression is a well-known machine learning technique that belongs to the supervised learning approach. We can use this technique to forecast a certain dependent variable by employing a collection of unbiased variables. Logistic regression is used to forecast the output of a specific structured variable. As a result, the result must be a discrete or categorical value. It may be 0 or 1, Yes or No, true or false, and so on, but rather than offering precise values between 0 and 1, it provides probabilistic values that fall inside that range.

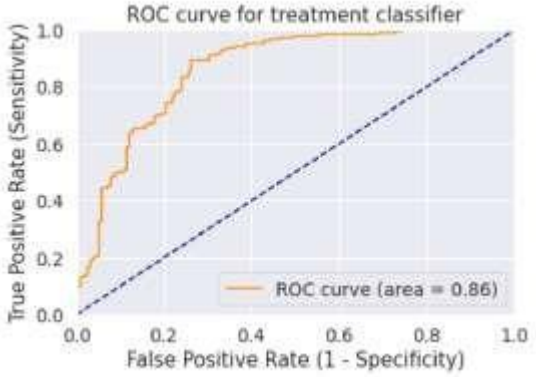

Figure 1 : ROC of Logistic Regression

B. K nearest neighbour classifier

A fundamental machine learning algorithm that uses the Supervised Learning method is called the K-Nearest Neighbour. Existing cases and fresh case/data will be comparable in the K-NN approach. KNN is a

non-parametric algorithm that does not assume anything about the distribution or the emphasised data. Additionally, it functions with several classes.

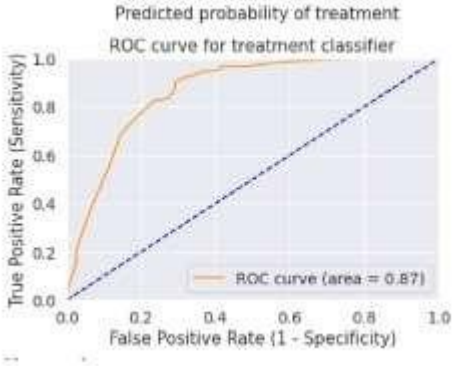

Figure 2 : ROC of KNN C. Decision tree classifier

The most popular supervised machine learning method used in data mining is the decision tree. Using a decision tree is a visual representation of a statistical likelihood or the order of events, actions, or consequences.

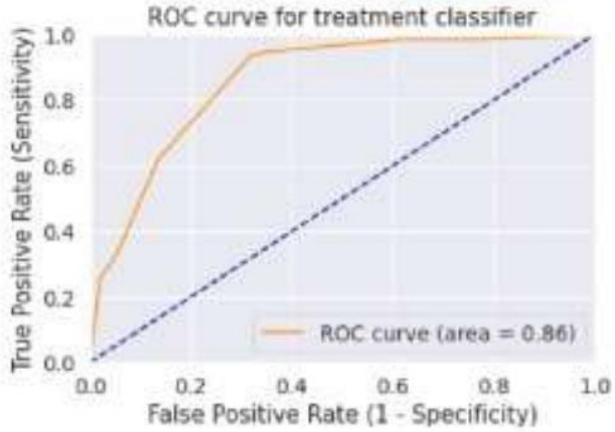

Figure 3: ROC of Decision Tree

D. Random Forest Classifier

Using a method called random forest, which is based on supervised machine learning, classification and regression issues can be solved. However, classification is a common usage for it. It is known as a random forest because it aggregates numerous decision trees to form a "forest" and feeds random features from the input dataset to them.

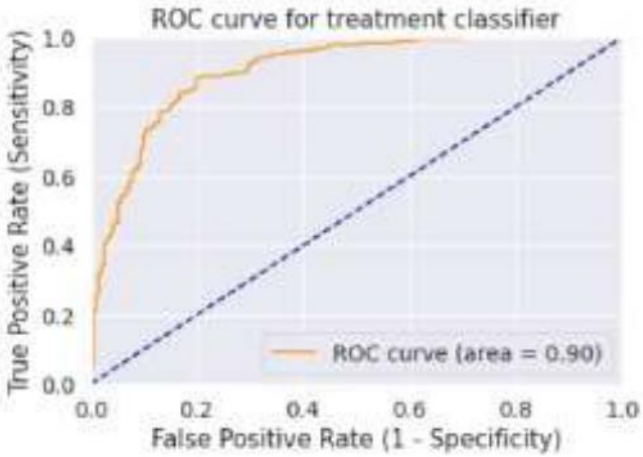

Figure 4 : ROC of Random Forest Classifier

E. Stacking

Stacking Generalisation, also referred to as "Stacking," is a machine learning ensemble method. Similar to bagging and boosting, it entails integrating predictions from several machine learning models on the same dataset.

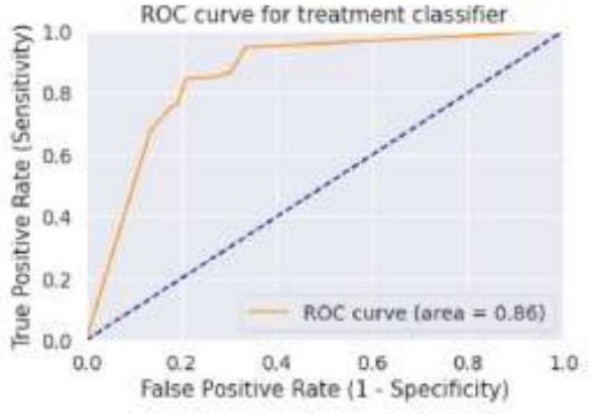

Figure 5 : ROC of Stacking

IV. RESULTS

This study discovered five machine learning techniques, including logistic regression, random forest, decision tree stacking, and k-nearest neighbour classifier. We also assessed their ability to recognise mental health issues. The classifiers were first run with all 27 of the text document extractions attributes, and then with 8 additional attributes that were selected using the feature selection methodology. The accuracy

of a given test set is measured by the proportion of occurrences from the test set that can be correctly identified using the classifier. Figure 5 demonstrates that stacking's accuracy is higher than that of other classifiers. The ability of a classifier to accurately classify the test data set will determine how accurate the classifier is. Utilising the area under the receiver operating curve, we calculated that. A perfect test will have an area of 1 in the ROC area, whereas a useless test will have an area of 0.5. The graph of five classifiers on ROC Area values is shown in Figure 1. Because the ROC area of all the classifiers utilised is between 0.8 and 0.9, we found that the classifiers were superior to other classifiers at predicting the state of mental health.

V. CONCLUSION

Before choosing the machine learning technique that best fits the target domain, it is vital to compare the various ones because there are many of them available. Today, there are numerous specialised systems in the medical profession that can accurately and promptly forecast disease, enabling effective and timely treatment. The dataset on diverse mental health diseases was classified using five distinct machine learning approaches in the proposed work. Before choosing the machine learning technique that best fits the target domain, it is vital to compare the various ones because there are many of them available. Today, there are numerous specialised systems in the medical profession that can accurately and promptly forecast disease, enabling effective and timely treatment. The dataset on diverse mental health diseases was classified using five distinct machine learning approaches in the proposed work.

VI. REFERENCES

[1]. T Nagar. “Prediction of mental health problems among children using machine learning techniques”

[2]. Konda Vaishnavi, Unikhitha, B Ashwath Rao and N V Subba Reddy. “Predicting Mental Health Illness using Machine Learning Algorithms”

[3]. Jetli Chung and Jason Teo. “Mental Health Prediction Using Machine Learning: Taxonomy, Applications, and Challenges”

[4]. M Srividya, S Mohanavalli, and N Bhalaji. “Behavioral modeling for mental health using machine learning algorithms.”

[5]. M. Hamilton, “Development of a rating scale for primary depressive illness,” British Journal of Social and Clinical Psychology, vol. 6, no. 4, pp. 278–296, 1967.

[6]. Theodoros Iliou, Georgia Konstantopoulou, Mandani Ntekouli, Christina Lymperopoulou, Konstantinos Assimakopoulos, Dimitrios Galiatsatos, and George Anastassopoulos. Iliou machine learning preprocessing method for depression type prediction. Evolving Systems, 10(1):29–39, 2019.

[7]. A. E. Tate, R. C. McCabe, H. Larsson, S. Lundstr¨om, P. Lichtenstein, and R. Kuja-Halkola, “Predicting mental health problems in adolescence using machine learning techniques,” PLoS One, vol. 15, no. 4, Article ID e0230389, 2020.

[8]. G. Miner, L. Miner, M. Goldstein et al., Practical Predictive Analytics and Decisioning Systems for Medicine: Informatics Accuracy and Cost - Effectiveness for Healthcare Administration and Delivery Including Medical Research, Academic Press, Inc., Cambridge, MA, USA, 1st edition, 2016.

[9]. M.-J. Wu, I. C. Passos, I. E. Bauer et al., “Individualized identification of euthymic bipolar disorder using the cambridge neuropsychological test automated battery (CANTAB) and machine learning,” Journal ofAffective Disorders, vol. 192, pp. 219–225, 2016.

[10].Stevie Chancellor, Eric PS Baumer, and Munmun De Choudhury. Who is the” human” in human-centered machine learning: The case of predicting

mental health from social media. Proceedings of the ACM on Human-Computer Interaction, 3(CSCW):1–32, 2019.

Cite this article as :

Rhitik Nimbalkar, Rakesh Oza, Akash Pabale, Khushi Pandey, Afsha Akkalkot, "Using Machine Learning to Predict Mental Illness", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN : 2456-3307, Volume 9, Issue 4, pp.01-05, July-August-2023. Available at doi :
https://doi.org/10.32628/CSEIT23903114
Journal URL : https://ijsrcseit.com/CSEIT23903114