



Mental Health Prediction using Machine Learning

Shivam Garg, Er. Nidhi Sengar, Prof.(Dr.) Amita Goel, Dr. Vasudha Bahl

Department of IT

Maharaja Agrasen Institute of Technology, Delhi, India

Abstract : Algorithm built using jupyter Notebook and Machine Learning that predict or manage mental health. The contents of this paper is an analysis of using machine learning models to predict Mental Health disorder in people using health care data. Worldwide, depression affects millions of individuals and is a crippling affliction. People may find it challenging to live regular, happy lives at best, and at worst, it may result in suicide. Primary care physicians frequently discover that they must treat mental health problems like depression despite having little or no formal training in how to do so.

Index Terms - Machine Learning, Mental Health Prediction.

I. INTRODUCTION

Mental health disorders have a significant impact on individuals and society as a whole. Worldwide, about 1 billion people [1] struggle with a mental illness. Healthcare systems are dealing with a rising demand for mental health services as well as a shortage of qualified employees as a result of the worldwide mental health emergency being significantly worsened by the Coronavirus Disease 2019 pandemic. Effective intervention and treatment planning depend on the early identification and precise prognosis of mental health disorders. In recent years, machine learning (ML) techniques have emerged as promising tools for predicting mental health outcomes using various data sources. This research paper provides a comprehensive review of the existing literature on the prediction of mental health using ML algorithms.

II. LITERATURE SURVEY

In their research work, Sumathi and Poorna used a variety of machine learning techniques to predict children's mental health issues [8]. Professionals are observing the causes, symptoms, and psychological tests of mental health issues. The data collection, which has 60 cases, was collected from a clinical psychologist. The classification procedure has chosen a number of traits and properties. To test the accuracy of several machine learning algorithms' predictions, this issue was given to them.

According to the data in Table 4, the average one-dependence estimator (AODE), a machine learning approach, has a 71% accuracy rate. MLP, on the other hand, displays the highest accuracy, at 78%[8]. The multiclass classifier is at 58% accuracy, followed by the logical analysis tree (LAT), which has a 70% accuracy rate. Radial basis function network (RBFN), another machine learning method, reports accuracy of 57%. In this experiment[7], K-star and functional tree (FT) both achieved an accuracy score of 42%.

Table 4
Accuracy of machine learning techniques in predicting the mental health problems among children.

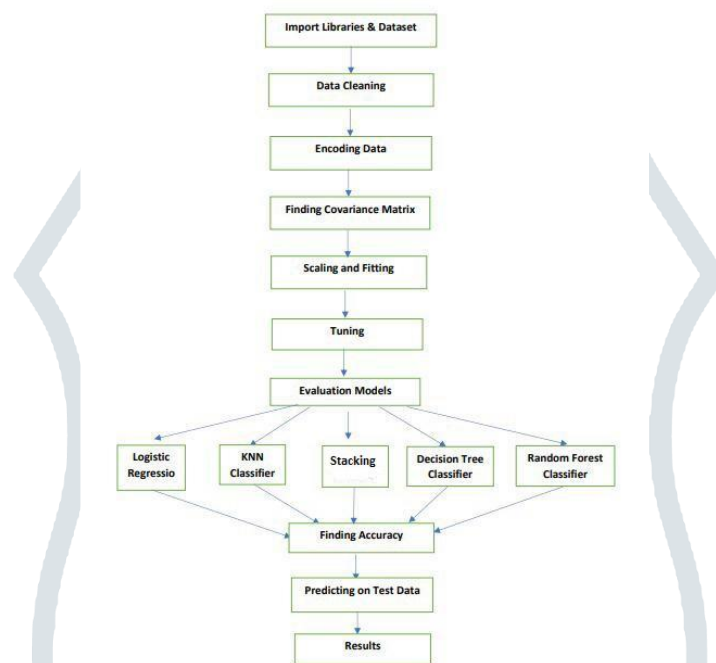
Machine learning	Accuracy (%)
Average one-dependence estimator (AODE)	71
Multilayer perceptron	78
Logical analysis tree (LAT)	70
Multiclass classifier	58
Radial basis function network (RBFN)	57
K-star	42
Functional tree (FT)	42

The classification of childhood-onset schizophrenia has been done, according to the work by Greenstein et al. [2]. The information consists of clinical data, brain magnetic resonance imaging data, and genetic data. The chance of a mental disease is determined by the authors using a random forest method. Because it has lower error rates than other techniques, random forest is chosen in this study. The classification results in an accuracy of 73.7%. [5]

In one of the studies carried out by Jo et al., they identified 48 people with schizophrenia and 24 healthy controls using network analysis and machine learning techniques [3]. The probabilistic brain tractography was used to reconstruct the network properties. Then, schizophrenic patients and health controls are labelled using machine learning. According to the findings, the multinomial naive Bayes model and the random forest model both had accuracy values of 66.9% and 68.6%, respectively. The accuracy score for XGBoost is 66.3%, and the accuracy for the support vector machine is 58.2%. [6] The majority of machine learning systems do well in predicting both healthy controls and those with schizophrenia.

III. PROPOSED METHODOLOGY PROJECT IMPLEMENTATION

First process will be the import of all the libraries like numpy, pandas, matplotlib, seaborn, sklearn. Followed by the import of data and providing the right location of the excel sheet. The data that is taken will not always be pure functional which is cleaned of the inconsistency, noises, Missing data in the column. Columns can also be normalized further. For making the data understandable by the models it is encoded.



Encoding data is a crucial step in preparing categorical or textual data for machine learning algorithms. Machine learning models typically require numerical input, so encoding categorical variables allows the algorithms to process and understand these features. After encoding of data comes forming of the correlation matrix.

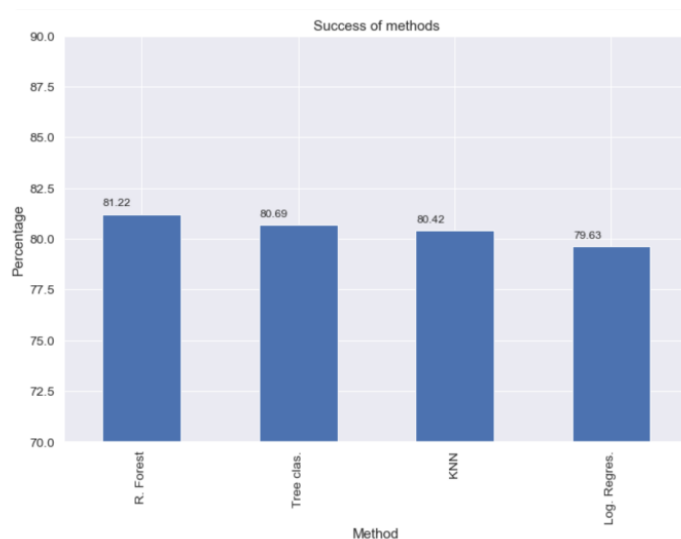
A correlation matrix is a statistical tool used in machine learning and data analysis to quantify the relationship between variables. It provides a matrix of correlation coefficients that measure the strength and direction of the linear relationship between pairs of variables in a dataset. The correlation coefficient values range between -1 and 1, where -1 indicates a strong negative correlation, 0 indicates no correlation, and 1 indicates a strong positive correlation.

Scaling and fitting is done. This project does Features Scaling, because it is extremely different from the other ones. Then comes

Tuning, it refers to the process of optimizing and preparing the dataset to improve the performance of machine learning models. Data tuning involves various techniques and steps to enhance the quality, balance, and generalizability of the data.

In the later stage comes the Implementation of the Evaluation models, it is a crucial step to assess their performance and determine their suitability for the task at hand. There are several evaluation metrics and techniques available to measure the effectiveness of a model. This project has implementation of logistic regression, KNN classifier, Decision tree, Random forest Classifier.

IV. RESULTS



V. CONCLUSION

There are many factors that can affect our mental health such as Age, Regional information, Gender, Education level, Anxiety level and Depression level, Anxiety frequency, Income, BMI, Insurance coverage. Disorder can be caused by yourself, others and even the environment.

Finding techniques to help manage health is key. The model should be designed to predict mental health. It's far from ideal, but I hope to see more of it someday.

VI. REFERENCES

1. D. Greenstein, J. D. Malley, B. Weisinger, L. Clasen, and N. Frontiers in Psychiatry, vol. 3, Gogtay, "Using multivariate machine learning methods and structural MRI to classify childhood onset schizophrenia and healthy controls," 2012.
2. Y. T. Jo, S. W. Joo, S. H. Shon, H. Kim, Y. Kim, and J. "Diagnosing schizophrenia with network analysis and a machine learning method," The International Journal of Methods in Psychiatric Research, vol. 29, no. 1, 2020.
3. Paton, F. et al. A quick review of the research supporting the various models of care is needed to improve outcomes for those experiencing mental health crises. *Health Technol. Assess.* 20, 1–162 (2016).
4. Hyland, S. et al. A Systematic analysis for the Global Burden of Disease Study 2017. *Lancet* 392, 1789 -1858 (2018).

5. Early prediction of circulatory failure in the intensive care unit using machine learning. *Nat. Med.* 26, 364–373 (2020).
6. Simon, G. et al. Predicting suicide attempts and suicide deaths following outpatient visits using electronic health records. *Am. J. Psychiatry* 175, 951–960 (2018).
7. Poulin, C. et al. Analysing the content of clinical notes to predict the likelihood of suicide.
8. M. Sumathi and B. Poorna, "Utilising machine learning to predict children's mental health issues," *Advanced Computer Science and Applications International Journal*, vol. 7, no. 1, 2016.

