VER
Vascular & Endovascular Review

Comparative Analysis of Machine Learning Algorithms for Mental health issue classification

Swarnima Shrivastava[1], Varsha Thakur[2], Surendra Kumar Patel[3]

[1]Department of Information Technology Govt. N.P.G. College of Science Raipur, India
[2]Department of Computer Science Govt. N.P.G. College of Science Raipur, India
[3]Department of Information Technology Govt. N.P.G. College of Science Raipur, India

ABSTRACT

In today's fast paced life, stress is a very common problem but if this problem persists then it takes the form of a mental illness. The rising incidence of mental health conditions and the expanding amount of digital behavioral data have drawn a lot of attention to machine learning (ML)-based mental illness detection. In order to classify and predict mental health conditions like depression, anxiety, stress, personality disorder, bipolar disorder and suicidal this study compares a number of popular machine learning algorithms, including Random Forests, Naïve Bayes (NB), Multinomial NB, Decision Tree and Logistic Regression. Accuracy, precision, recall and F1-score are the metrics used to assess each algorithm using a structured dataset. The findings show that ensemble techniques like decision tree achieve highest accuracy which is 78%.

KEYWORDS: Mental health, depression, stress, machine learning.

How to Cite: Swarnima Shrivastava, Varsha Thakur, Surendra Kumar Patel, (2025) Comparative Analysis of Machine Learning Algorithms for Mental health issue classification, Vascular and Endovascular Review, Vol.8, No.12s, 205-213.

INTRODUCTION

Mental health issues are very common problem and the reason behind is depression. Depression affects the majority of people in one manner or another. As a result, people often ask questions about it, such as what depression is and how it's different from just feeling depressed. Millions of people worldwide suffer from depression, a prevalent mental health illness for which prompt care and support are essential. Conventional techniques for diagnosing depression mostly rely on subjective evaluations and self-reporting, which can be laborious and unreliable[1].

Individuals suffering from depression often display signs including mental illness, low mood, decreased interest and enjoyment, and in extreme situations, suicidal thoughts and actions [2]. Thus, identification and treatment of people having depression is required at earlier stage[3].

A behavioral disorder characterized by distress and impairment of physiological functions is referred to as mental health. It affects their attitude, ideas, and habits, as well as their general well-being. An estimated 971 million people worldwide suffer from mental illnesses, according to the study [4]. These mental illnesses can fall into a variety of categories, including bipolar disorder, schizophrenia, dementia, depression and anxiety. Most of these people are left untreated[5].

Artificial Intelligence (AI) is a superset of Machine Learning (ML) that shares many similarities with computational statistics, which is similarly concerned with computer-based prediction. The discipline receives methods, theory, and application domains from mathematical optimization, which has significant linkages to it. Although data mining and machine learning are sometimes confused [6], unsupervised learning is mainly concerned with investigative data analysis. It is used to discover significant anomalies by learning and establishing baseline behavioral profiles for different entities[7]. According to Arthur Samuel, the discipline's founder, without being explicitly programmed computers are able to learn and this field of study is known as machine learning. Regression and classification using known features previously discovered from the training data are the main goals of machine learning[8].

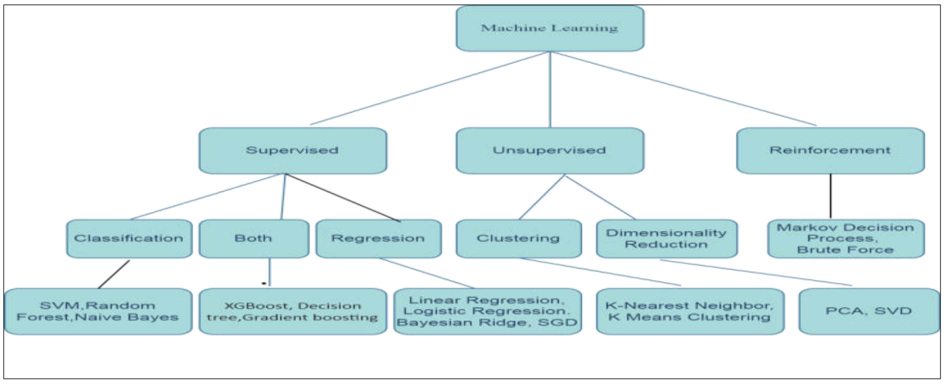

Fig. 1. Machine Learning Algorithms

LITERATURE REVIEW

Alalalmeh et al. [9] investigate the spread and impact of stress, depression and anxiety amongst students of university in UAE , using the Depression, Anxiety, and Stress Scale-21 Items (DASS-21). A bilingual DASS-21 survey administered through Google Forms is used in this investigation. The study examined 332 students, with most female participants. Descriptive statistics, chi-squared tests, Mann-Whitney tests, Kruskal-Wallis tests, and multinomial logistic regression were used in the analysis, which was carried out using SPSS version 29.0, to examine the connections between sociodemographic factors and mental health scores. This study examines the differences in depression, anxiety, and stress among UAE university students by gender, age, and academic year.

Ding et al. [10]proposed a Deep Integrated Support Vector Machine(DISVM) algorithm to detect depression of college student by using Sina Weibo social media data. This research uses a text mining approach for depression detection. This paper uses an integrated SVM for depression identification model. DISVM outperform other state of the art algorithms such as SVM, KNN,RBF-NN.

Zhou et al. [11]proposed a model named Joint Attention Multi-Scale Fusion Network(JAMFN). This model shows multi scale behavioral information of depression and it also guide about how multi modalities are combined together. This model uses BiLSTM and series of convolutional network on D-vlog dataset to perform experiment. In order to increase the model's learning efficiency, future research will concentrate on ways to reduce the impact of the mislabeling samples in the training samples.

Prama et al.[12]proposed a method which effectively categorize tweet into “non-depressive”, “mild”, “moderate”, “severe” depression using LSTM based deep learning model and measure its intensity. The dataset was constructed work together with a psychologist and categories of depression were created using the rules laid out in the DSM-5-TR. This model perform superior when compared with baseline models.

Tao et al [13] proposed a framework DepMSTAT, to detect depression. It is spatio-temporal converter framework which is capable, low-covariance and multimodal included framework that utilizes depression dataset (D-Vlog) which contains visual and audio data both. Four components comprise the framework: a depression classifier module, a Spatial-Temporal Attentional Transformer (STAT) module, a token generating module, and a data preprocessing module. Temporal and spatial information can be extracted and fused from speech and facial data using the suggested Spatio-Temporal Attentional Transformer (STAT) module. Every experiment and piece of code is built using the PyTorch framework.

A comprehensive review of machine and deep learning techniques for depression recognition is given by Tahir et al.[14] paper purpose is to indicate research gap in depression detection by applying machine and deep learning on social media data. A methodology of detecting depression, how social media comments like tweets are recognized as depressed or non depressed tweets by machine and deep learning algorithms? What are the scope and application area of depression detection? A generic architecture is proposed in this paper and datasets used for evaluation in this field are given thoroughly.

Kwok et al. [15] proposes a framework for depression detection utilizing electroencephalogram (EEG) data which was the first attempt to assess machine learning fairness. Three EEG datasets—Mumtaz, MODMA, and Rest—are experimented with various deep learning techniques that are CNN, GRU and LSTM networks. Several bias reduction techniques are used in this work before, during, and after processing and evaluate their performance. The experimental findings demonstrate that bias is present in current EEG datasets and depression diagnosis algorithms and that various bias mitigation strategies address unfairness at varying degrees across various fairness metrics.

Jin et al. [16] presented a study in which they tracked 302 hospitalized patients with depression over a six-month period using wearable technology. They gathered information of physical activity, heart rate, and sleep patterns, which were subsequently compared to the Hamilton Depression (HAMD) and Hamilton Anxiety (HAMA) scores. These vital signs differed significantly between instances of mild and severe depression, according to the data. Significant improvements in the monitoring and medication of depression are feasible with wearable technology. These tools can improve the understanding and organization of depression by combining ongoing physiological data with medical evaluations, which could revolutionize mental health care by making it more accurate, individualized, and proactive.

Zhang et al. [3] proposed a model DepITCM which uses audio and visual dataset AVEC2017 and AVEC2019 for depression detection. This approach enables the capturing of global characteristics while emphasizing the finer-grained blending of geographical, temporal, and channel information. Furthermore, this research, which draws inspiration from multitask learning approaches, enhances overall performance by incorporating a secondary task (the regression task) into the primary task of depression classification. The ability of deep learning (DL) models trained on videos of clinical interviews done by a virtual agent to screen for depression has been the subject of recent study.

Niu et al.[17] Proposed a depression scale dictionary decomposition framework that contains a Bidirectional Dictionary Decomposition (BDD) module and a Bidirectional Multimodal Fusion (BMF) module. In order to promote the perception of depression cues, the BDD module may semantically break down audio and video sequences into information relevant to and unrelated to depression scores along token and channel dimensions using dictionaries created using the depression scale. Furthermore, the BMF module employs linear layers and graph convolution to accomplish cross-modal mixing, which is used to aggregate audio and video sequences for depression level prediction while taking into account the corresponding properties of tokens and channels. The validation on the DAICWOZ, AVEC 2013, and AVEC 2014 datasets shows how effective this model is.

METHODOLOGY

This paper has used five machine learning algorithms -Logistic Regression, Naïve Bayes, Multinomial Naive Bayes, Decision Tree and Random Forest to categorize mental health problems in anxiety, bipolar, depression, normal, personality disorder, stress and suicidal. Tensor-flow and keras framework is used in preprocessing and other crucial steps.

A. Importing Libraries

Importing important libraries such as NumPy, Pandas, matplotlib, Tensor Flow and Keras.

B. Loading dataset

Dataset used is “combined data.csv” from Kaggle. Dataset has 03 attributes and 53042 instances.

C. EDA (Exploratory Data Analysis)

A crucial phase in data analysis is exploratory data analysis (EDA), which uses statistical tools and visualizations to identify patterns, trends, and relationships.

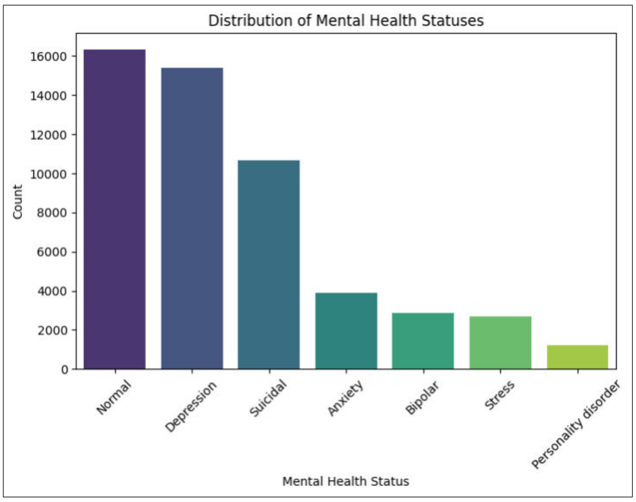

Fig. 2 Distribution of Mental health status

Data Preprocessing:

One crucial step in the data mining process is data preparation. This procedure prepares raw data for additional analysis by converting it into a comprehensible format. The goal is to raise the caliber of the data and make it suitable for the intended uses. Preprocessing data aids in converting it into a format that may be used. Data reduction, cleansing, integration and transformation are tasks that aid in data preprocessing. By managing missing values and reducing noise through the use of binning, regression, and clustering, data cleaning eliminates incomplete data. To create a single dataset, data integration brings together information from several sources[18].

Token Visualization :

Identified tokens in particular mental health problem such as depression

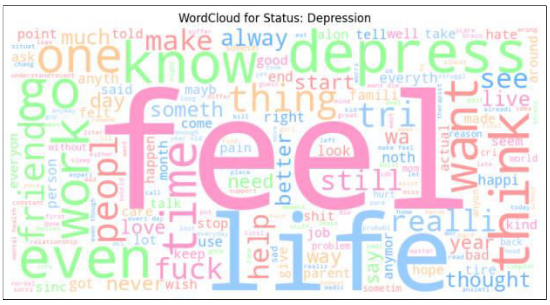

Fig. 3. Word Cloud for depression

Vectorization:

It is an important step in feature extraction like TF-IDF.

Modeling:

Trained Dataset is evaluated based on confusion matrix for each classifier. Parameters are precision, recall and F1-score.

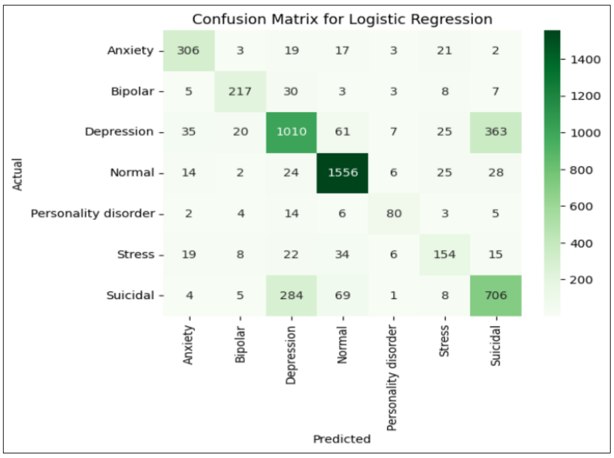

Fig. 4. Confusion matrix for Logistic Regression

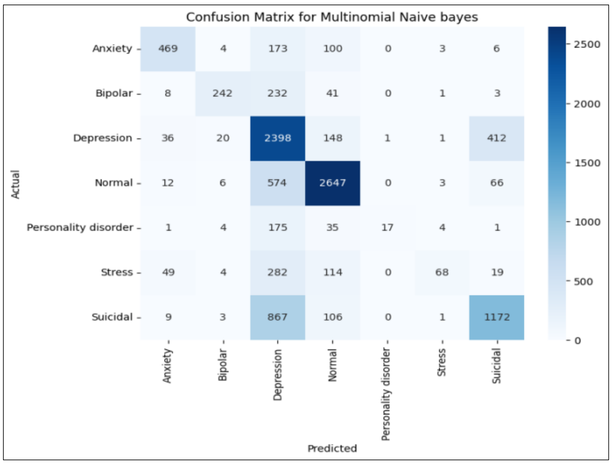

Fig. 5. Confusion matrix for Multinomial NB

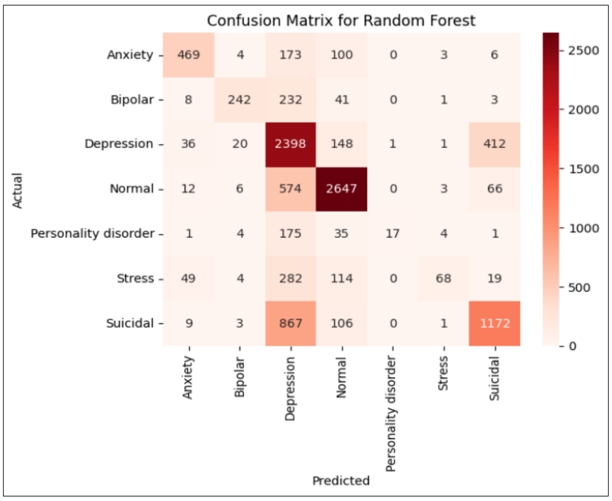

Fig. 6. Confusion matrix for Random Forest

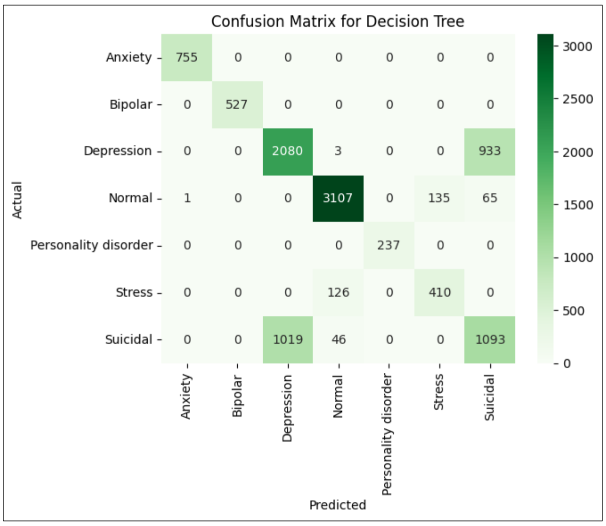

Fig. 7. Confusion matrix for Decision Tree

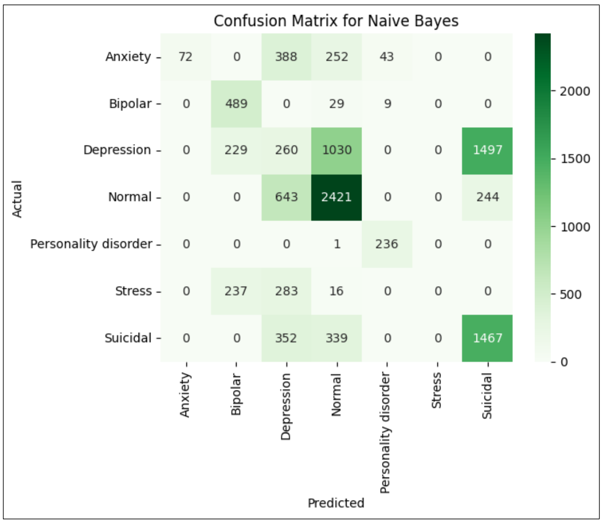

Fig. 8. Confusion matrix for Naïve Bayes

TABLE I. PRECISION, RECALL AND F1-SCORE FOR MACHINE LEARNING ALGORITHMS

Algorithm	Mental health status	Precision	Recall	F1-Score
Logistic Regression	Normal	0.89	0.94	0.92
	Depression	0.72	0.66	0.69
	Anxiety	0.79	0.82	0.81
	Suicidal	0.63	0.66	0.64
	Bipolar	0.84	0.79	00.82
	Stress	0.63	0.60	00.61
	Personality Disorder	0.75	0.70	00.73
Multinomial NB	Normal	0.83	0.80	00.81
	Depression	0.51	0.80	00.62
	Anxiety	0.80	0.62	00.70
	Suicidal	0.70	0.54	00.61
	Bipolar	0.86	0.46	00.60
	Stress	0.84	0.13	00.22
	Personality Disorder	0.94	0.07	00.13
	Normal	0.83	0.80	00.81

Random Forest	Depression	0.51	0.80	00.62
	Anxiety	0.80	0.62	00.70
	Suicidal	0.70	0.54	00.61
	Bipolar	0.86	0.46	00.60
	Stress	0.84	0.13	00.22
	Personality Disorder	0.94	0.07	00.13
Decision Tree	Normal	0.95	0.94	00.94
	Depression	0.67	0.69	00.68
	Anxiety	1.00	1.00	10.00
	Suicidal	0.52	0.51	00.51
	Bipolar	1.00	1.00	10.00
	Stress	0.75	0.76	00.76
	Personality Disorder	1.00	1.00	10.00
Naïve Bayes	Normal	0.59	0.73	00.65
	Depression	0.13	0.09	00.11
	Anxiety	1.00	0.10	00.17
	Suicidal	0.46	0.68	00.55
	Bipolar	0.51	0.93	00.66
	Stress	0.00	0.00	00.00
	Personality Disorder	0.82	1.00	00.90

RESULT ANALYSES

After analyzing all evaluation parameters it is concluded that Decision Tree exhibit maximum accuracy of 78% among all algorithms.

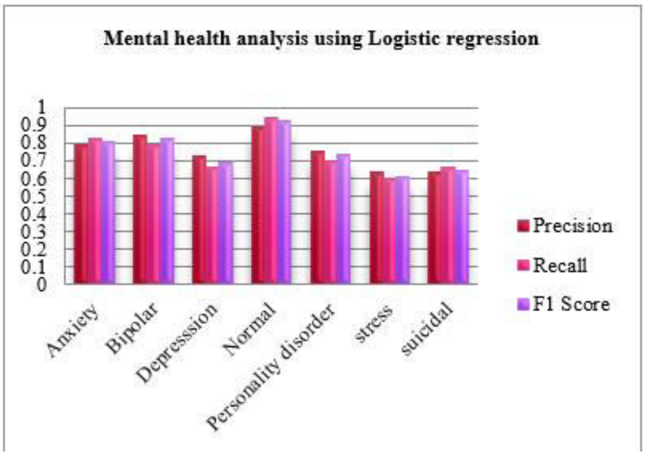

Fig. 9 Mental health analysis for each status using Logistic Regression

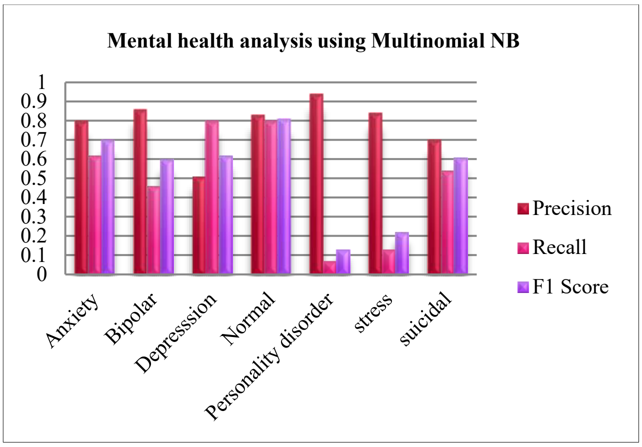

Fig. 10 Mental health analysis for each status using Multinomial NB

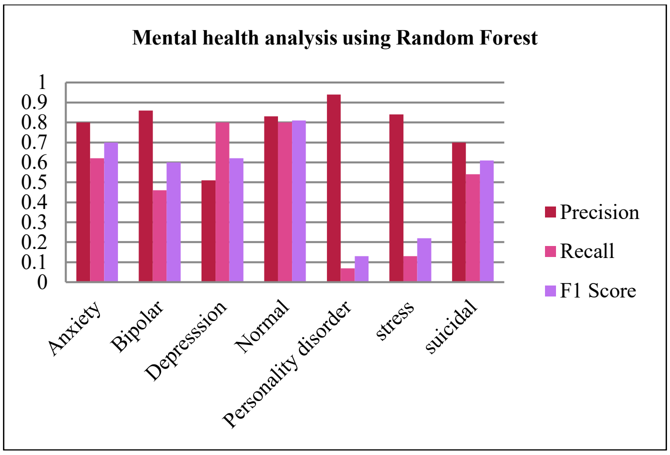

Fig. 11. Mental health analysis for each status using Random Forest

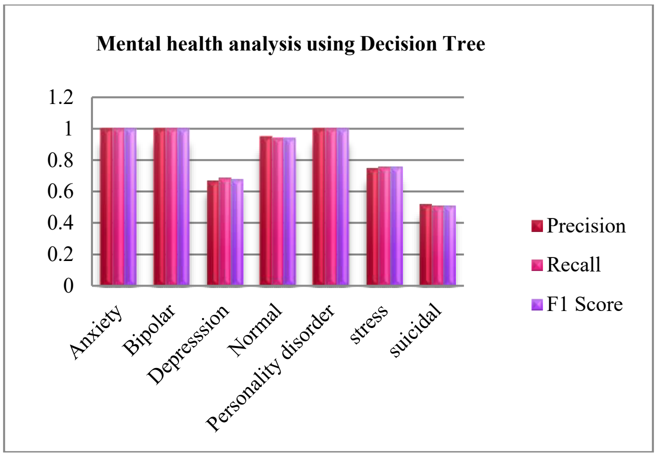

Fig. 12 Mental health analysis for each status using Decision Tree

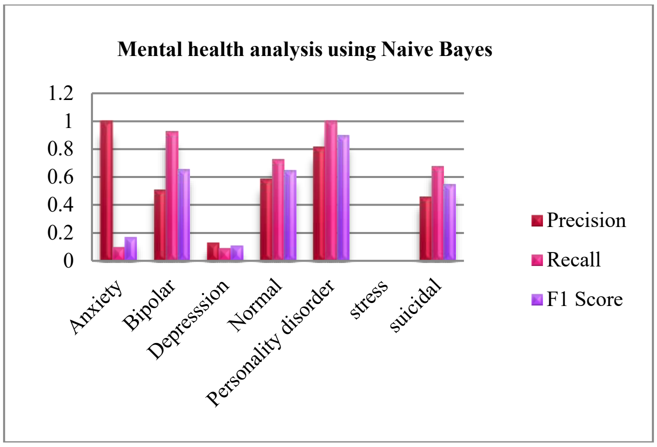

Fig. 13 Mental health analysis for each status using Naive Bayes

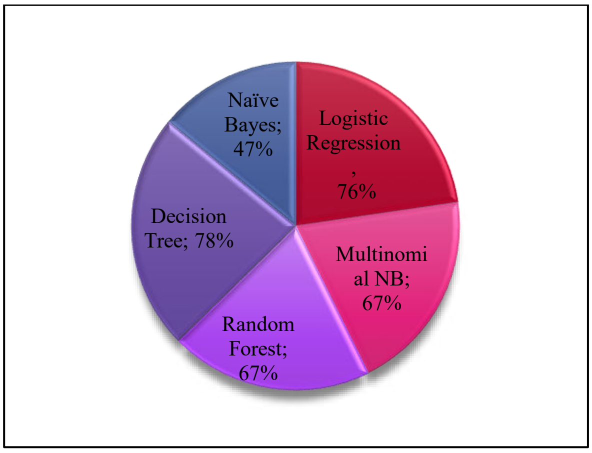

Fig. 14 Accuracy of machine learning algorithms

CONCLUSION

In order to assess how well different machine learning algorithms detect mental diseases such as depression, anxiety, bipolar disorder, personality disorder, stress, and suicidal thoughts, this study has conducted a thorough comparison of these algorithms. We have empirically demonstrated that no single algorithm consistently performs better than others across all assessment metrics using a textual dataset. More interpretability and quicker training timeframes were offered by conventional models such as Random Forests and Logistic Regression, which are essential in actual clinical situations. The study also emphasizes how crucial feature selection, data quality, and ethical issues are when using AI for mental health diagnosis. The best algorithm to use ultimately relies on the particular application context, including whether computing efficiency, explainability, or diagnostic precision are the top priorities. To ensure responsible and efficient deployment, future research should concentrate on integrating multimodal data sources, reducing model bias, and testing these systems in clinical settings.

REFERENCES

1.S. Kanoujia and P. Karuppanan, "Depression Detection in Speech Using ML and DL Algorithm," in *2024 IEEE International Conference on Interdisciplinary Approaches in Technology and Management for Social Innovation (IATMSI)*, Gwalior, India: IEEE, Mar. 2024, pp. 1–5. doi: 10.1109/IATMSI60426.2024.10503510.
2.D. M. Maurer, "Screening for depression," *Am. Fam. Physician*, vol. 85, no. 2, pp. 139–144, Jan. 2012.
3.L. Zhang *et al.*, "DepITCM: an audio-visual method for detecting depression," *Front. Psychiatry*, vol. 15, Jan. 2025, doi: 10.3389/fpsyt.2024.1466507.
4."Global, regional, and national incidence, prevalence, and years lived with disability for 328 diseases and injuries for 195 countries, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016 - PubMed." Accessed: Apr. 29, 2025. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/28919117/
5.M. Ahmad Wani, M. A. ELAffendi, K. A. Shakil, A. Shariq Imran, and A. A. Abd El-Latif, "Depression Screening in Humans With AI and Deep Learning Techniques," *IEEE Trans. Comput. Soc. Syst.*, vol. 10, no. 4, pp. 2074–2089, Aug. 2023, doi: 10.1109/TCSS.2022.3200213.
6.P. Louridas and C. Ebert, "Machine Learning," *IEEE Softw.*, vol. 33, no. 5, pp. 110–115, Sep. 2016, doi: 10.1109/MS.2016.114.

7.M. I. Jordan and T. M. Mitchell, "Machine learning: Trends, perspectives, and prospects," *Science*, vol. 349, no. 6245, pp. 255–260, Jul. 2015, doi: 10.1126/science.aaa8415.
8.Y. Xin *et al.*, "Machine Learning and Deep Learning Methods for Cybersecurity," *IEEE Access*, vol. 6, pp. 35365–35381, 2018, doi: 10.1109/ACCESS.2018.2836950.
9.S. O. Alalalmeh *et al.*, "Assessing mental health among students in the UAE: A cross-sectional study utilizing the DASS-21 scale," *Saudi Pharm. J.*, vol. 32, no. 4, p. 101987, Apr. 2024, doi: 10.1016/j.jsps.2024.101987.
10. Y. Ding, X. Chen, Q. Fu, and S. Zhong, "A Depression Recognition Method for College Students Using Deep Integrated Support Vector Algorithm," *IEEE Access*, vol. 8, pp. 75616–75629, 2020, doi: 10.1109/ACCESS.2020.2987523.
11. L. Zhou, Z. Liu, Z. Shangguan, X. Yuan, Y. Li, and B. Hu, "JAMFN: Joint Attention Multi-Scale Fusion Network for Depression Detection," in *INTERSPEECH 2023*, ISCA, Aug. 2023, pp. 3417–3421. doi: 10.21437/Interspeech.2023-183.
12. T. T. Prama, Md. S. Islam, M. M. Anwar, and I. Jahan, "AI-Enabled Deep Depression Detection and Evaluation Informed by DSM-5-TR," *IEEE Trans. Comput. Soc. Syst.*, vol. 11, no. 5, pp. 6453–6465, Oct. 2024, doi: 10.1109/TCSS.2024.3382139.
13. Y. Tao, M. Yang, H. Li, Y. Wu, and B. Hu, "DepMSTAT: Multimodal Spatio-Temporal Attentional Transformer for Depression Detection," *IEEE Trans. Knowl. Data Eng.*, vol. 36, no. 7, pp. 2956–2966, Jul. 2024, doi: 10.1109/TKDE.2024.3350071.
14. W. B. Tahir, S. Khalid, S. Almutairi, M. Abohashrh, S. A. Memon, and J. Khan, "Depression Detection in Social Media: A Comprehensive Review of Machine Learning and Deep Learning Techniques," *IEEE Access*, vol. 13, pp. 12789–12818, 2025, doi: 10.1109/ACCESS.2025.3530862.
15. A. M. H. Kwok, J. Cheong, S. Kalkan, and H. Gunes, "Machine Learning Fairness for Depression Detection using EEG Data," Jan. 30, 2025, *arXiv*: arXiv:2501.18192. doi: 10.48550/arXiv.2501.18192.
16. Y. Jin and Y. Huang, "Long-Term Vital Sign Tracking Study of Depression Patients Based on Wearable Devices," *Int. J. Crowd Sci.*, vol. 9, no. 1, pp. 56–63, Jan. 2025, doi: 10.26599/IJCS.2024.9100044.
17. M. Niu, X. Wang, J. Gong, B. Liu, J. Tao, and B. W. Schuller, "Depression Scale Dictionary Decomposition Framework for Multimodal Automatic Depression Level Prediction," *IEEE Trans. Circuits Syst. Video Technol.*, pp. 1–1, 2025, doi: 10.1109/TCSVT.2025.3533480.
18. "Data Preprocessing in Data Mining." Accessed: Jun. 17, 2024. [Online]. Available: https://www.tutorialspoint.com/data-preprocessing-in-data-mining