Hybrid Machine Learning Approaches for Predicting and Diagnosing Major Depressive Disorder

Mr. N. Balakrishna[1], Dr. M. B. Mukesh Krishnan[2], Dr. D. Ganesh[3]
Research Scholar, Department of Networking and Communication,
SRM Institute of Science and Technology, Kattankulathur, Tamilnadu, India[1]
Professor, Department of Networking and Communications,
SRM Institute of Science and Technology, Kattankulathur, Tamilnadu, India[2]
Associate Professor of CSE, Mohan Babu University,
(Erstwhile Sree Vidyanikethan Engineering College (Autonomous), Tirupati, Andhra Pradesh, India[3]

Abstract*—Major Depressive Disorder (MDD) is common and debilitating, requiring accurate prediction and diagnosis. This study uses clinical, demographic, and EEG data to test hybrid machine learning methods for MDD prediction and diagnosis. EEG data reveals brain electrical activity and can identify MDD patterns and traits. The study aimed to enhance Major Depressive Disorder (MDD) prediction and diagnosis using hybrid machine learning methods, focusing on EEG data alongside clinical and demographic information. Employing various algorithms like CatBoost, Random Forest, XG Boost, XGB Random Forest, SVM with a linear kernel, and logistic regression with Elasticnet regularization, the study found that CatBoost achieved the highest accuracy of 93.1% in MDD prediction and diagnosis, surpassing other models. Additionally, the ensemble model combining XGBoost and Random Forest showed strong performance in ROC analysis, effectively discriminating between individuals with and without MDD. These findings underscore the potential of EEG data integration and hybrid machine learning techniques in accurately identifying and classifying MDD patients, paving the way for personalized interventions and targeted treatments in depressive disorders.

Keywords*—Major Depressive Disorder (MDD); hybrid machine learning; cat boost; random forest; XG boost; XGB random forest; SVM; logistic regression; EEG data

I. Introduction

Depression is a widespread mental health problem that has serious consequences for sufferers' lives and those around them. According to the World Health Organization's projection, depression is expected to become the leading mental health condition by 2030. The severity of depression can lead to tragic outcomes, including suicide [1]. However, the current diagnostic procedures lack reliable and clinically useful tools for characterizing depression accurately [2]. This limitation introduces biases and challenges in the diagnostic process. While the expertise and motivation of clinicians are crucial, factors such as patient education, cognitive ability, and honesty in symptom reporting also play vital roles [3]. Achieving an accurate diagnosis of depression's severity necessitates extensive background knowledge and comprehensive clinical training. So there is a need to develop machine learning algorithms to automatically predict the severity of depression using various computational techniques [4]. These systems aim to improve the diagnostic process and provide valuable insights for effective intervention and treatment strategies.

Millions of people all over the world suffer from Major Depressive Disorder (MDD), a serious mental illness with a significant negative impact [5]. Depression is characterized by an inability to lift one's mood, a loss of hope, and a lack of interest in once-pleasurable things. The key to successful treatment and management of major depressive disorder is an early and precise diagnosis of the condition. However, diagnosing MDD solely based on subjective assessments and clinical interviews can be challenging due to the complexity and heterogeneity of the disorder.

Recent advancements in the field of mental health research have seen the application of machine learning techniques for predicting and diagnosing psychiatric disorders, including Major Depressive Disorder (MDD). Machine learning algorithms can analyze large datasets, uncover hidden patterns, and make accurate predictions [5]. However, traditional approaches often rely solely on quantitative data or clinical assessments, which may overlook crucial aspects of MDD's multifaceted nature. To overcome these limitations, researchers have started exploring hybrid machine-learning approaches that integrate diverse data sources and combine multiple algorithms. By leveraging the complementary strengths of different techniques and data types, these hybrid models aim to improve the accuracy and robustness of MDD prediction and diagnosis. This integration of diverse data enables a more comprehensive understanding of MDD, leading to enhanced clinical decision-making and more effective treatment strategies.

Hybrid machine learning approaches encompass a wide range of techniques, including ensemble learning, feature selection, and dimensionality reduction. To improve classification accuracy, ensemble learning techniques blend the outputs of several different base models. Feature selection algorithms help identify the most relevant clinical variables, while dimensionality reduction techniques enable data visualization and interpretation.

The objective of this research is to investigate the application of hybrid machine learning approaches for predicting and diagnosing MDD. By leveraging the power of diverse data sources and integrating multiple machine-learning techniques, these approaches aim to enhance the accuracy, reliability, and clinical utility of MDD diagnosis. Additionally, the research will address challenges such as interpretability, data privacy, and generalization of findings to real-world clinical settings.

In this implementation work, we aim to detect Major Depressive Disorder (MDD) by conducting a comprehensive analysis of a dataset comprising clinical, demographic, and electroencephalogram (EEG) data. The first step involves preprocessing the collected data, which includes performing necessary cleaning, normalization, and handling of missing values or outliers to ensure data quality. Next, metrics for attribute selection were developed, including "Information Gain," "Gain Ratio," "Gini Decrease," and "$\chi 2$" are applied to identify the most relevant and informative features for MDD prediction and diagnosis specifically based on EEG data. Once the feature selection process is complete, we proceed to implement various machine learning models, including CatBoost, Random Forest, XGBoost, XGBRandom Forest, SVM with a linear kernel, and logistic regression with ElasticNet regularization. To optimize the models' performance, grid search is employed to fine-tune their hyperparameters. Finally, a vote classifier is constructed to combine the predictions from multiple models using ensemble learning techniques, thereby improving the overall accuracy in predicting and diagnosing MDD.

Ultimately, the successful implementation of hybrid machine learning approaches for MDD prediction and diagnosis holds great promise in improving patient outcomes. Early identification of individuals at risk for MDD, accurate diagnosis, and personalized treatment strategies can significantly contribute to reducing the burden of this debilitating disorder. This research presents a novel approach to the analysis of EEG data for Major Depressive Disorder (MDD) diagnosis by integrating a suite of attribute selection metrics, including Information Gain, $\chi 2$, Gain Ratio, and Gini Decrease, which enhances the model's predictive power through refined feature selection. The pioneering use of the NG Deluxe 3.0.5 system for artefact rejection significantly elevates the quality of EEG data preprocessing, a crucial factor for accurate analysis. Furthermore, the hybrid machine learning model, including CatBoost, excels in managing categorical data, providing robust predictions validated through k-fold cross-validation. Ensemble techniques, like XGBoost and Random Forest, further showcase the potential of this innovative approach in computational psychiatry, particularly in discriminating MDD with high accuracy. Finally, the utilization of ensemble models like XGBoost and Random Forest, complemented by ROC performance analysis, offers new insights into the capabilities of hybrid models in distinguishing MDD, setting a precedent for future research and clinical application

The rest of the paper is organized as the Basic preliminaries and associated work are discussed in Section II, a hybrid machine learning model is offered in Section III, results and discussion are described in Section IV, and the conclusion and future work are presented in Section V, which is followed by references.

II. Basic Preliminaries and Literature Works

Utilizing machine learning to predict and diagnose Major Depressive Disorder (MDD) is discussed in this section. MDD is a common mental illness with poor mood and other symptoms [33]. Traditional diagnostic methods have limits; thus machine learning is used. Machine learning algorithms without human code analyze and predict data automatically. Hybrid machine learning methods improve prediction accuracy and generalization by mixing various algorithms or data sources. Hybrid machine learning integrates many models and data to enhance MDD prediction and diagnosis.

A. Major Depressive Disorder (MDD)

Major depressive disorder (MDD) is characterized by long-lasting melancholy, a lack of interest in or enjoyment from formerly rewarding activities, and an overpowering sense of powerlessness [6]. It affects mood, energy, sleep, hunger, and quality of life, as well as mental and physical health. A person with this sort of depression may lose focus, feel guilty or worthless, and frequently consider suicide. Genetics, biology, environment, and psychology all contribute to this syndrome. Psychotherapy, medication, and behavioral modifications can treat and reverse MDD symptoms.

B. Types of Major Depressive Disorder

Major Depressive Disorder (MDD) encompasses a spectrum of manifestations with varying symptoms, durations, and features [6]. These include Melancholic Depression, characterized by severe depressive symptoms such as profound loss of pleasure or interest in activities, morning mood worsening, weight loss or reduced appetite, excessive guilt, and psychomotor disturbances. Psychotic Depression presents with coexisting MDD and psychosis, featuring hallucinations and delusions, often centered around depressive themes like remorse over a perceived disaster. Atypical Depression is marked by emotional reactivity, leading to transient feelings of pleasure in response to positive events, alongside symptoms like increased hunger, weight gain, hypersomnia, and social withdrawal due to rejection sensitivity. Seasonal Affective Disorder (SAD) is a type of MDD that recurs in autumn and winter, characterized by sadness, increased sleep, weight gain, and reduced interest in activities, with symptoms typically improving in spring and summer. Postpartum Depression, affecting women after childbirth, involves feelings of sadness, anxiety, and exhaustion, hindering the mother's ability to care for herself and her child, with onset ranging from days to a year post-delivery.

C. Scalp Electrode Positions for Accurate EEG Analysis in Major Depressive Disorder

The specific electrode placements for studying Major Depressive Disorder (MDD) using electroencephalography (EEG) can vary depending on the research protocol and the specific study design. However, there are commonly used electrode placements that are relevant to studying MDD. The International 10-20 system is a widely adopted standard for electrode placement in EEG studies. It involves placing electrodes on specific locations of the scalp based on a defined grid system [7]. While the 10-20 system does not directly target MDD, it provides a standardized approach for electrode placement in EEG research.

In MDD studies, researchers often focus on specific regions of interest related to emotional processing and mood regulation. Common electrode placements for MDD studies may include specific regions. Fig. 1 will represent the Electrode Positions and specific regions for Accurate EEG Analysis in Major Depressive Disorder.

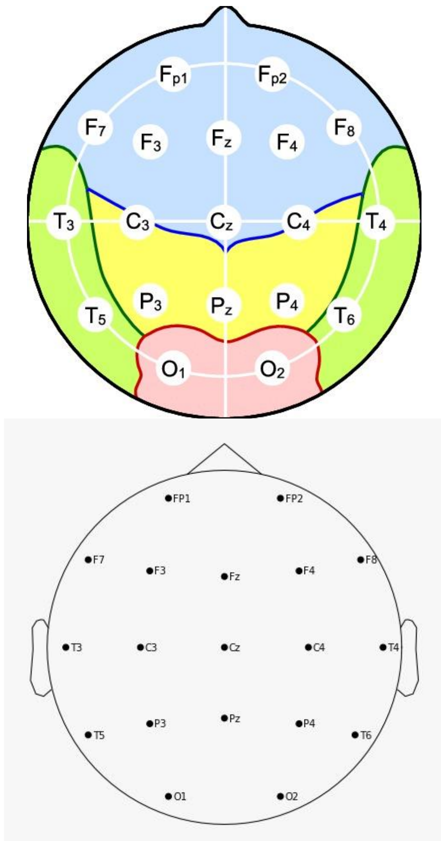

Fig. 1. Electrode positions and regions for accurate EEG analysis in major depressive disorder.

1) Frontal region: Electrodes placed over the prefrontal cortex (F3, F4, Fz) are of interest, as this area is associated with emotional regulation and cognitive processes relevant to MDD [7].

2) Temporal region: Electrodes placed over the temporal lobes (T3, T4, T5, T6) are important for capturing neural activity related to emotional processing and perception [7].

3) Central region: Electrodes placed over the central region (C3, C4, Cz) can provide insights into motor and cognitive functions that are implicated in MDD [7].

4) Parietal region: Electrodes placed over the parietal lobes (P3, P4, Pz) can capture neural activity related to attention and sensory processing, which may be relevant to MDD [7].

D. Literature Review on Major Depressive Disorder (MDD)

A literature analysis on hybrid machine learning approaches for prediction and diagnosis of Major Depressive Disorder (MDD) shows an increasing interest in using different algorithms and data sources to improve accuracy and resilience. Combining SVM, ANN, and other machine learning methods has been the focus of much research. decision trees and ensemble approaches to develop hybrid MDD prediction models. Hybrid models combine algorithm strengths to collect more features and increase performance. To improve model predictiveness, researchers have used clinical assessments, demographic data, genetic data, brain imaging data, and electronic health records. Hybrid models predict and diagnose MDD with higher sensitivity, specificity, and accuracy than traditional methods by merging multidimensional data and machine learning algorithms. However, more research is needed to standardize hybrid model creation and validation, address interpretability concerns, and ensure clinical application. Machine learning architectures, longitudinal data analysis, and real-time monitoring systems may promote hybrid machine learning for MDD prediction and diagnosis [34].

Many studies have examined utilizing ML to study mental illnesses. Depression, imaging, and ML approaches are covered in [5], a historical viewpoint. It also summarizes imaging and ML depression investigations. Linear, nonlinear, and relevance vector regression techniques are being studied. Survey examines one mental health aspect. This study did not compare algorithms or depression screening scales. An MHMS literature review examined machine learning (ML) and sensor data [8]. Several types of machine learning were applied to research depression, anxiety, bipolar disorder (BD), migraine, and stress [31]. These comprised supervised, unsupervised, semi-supervised, transfer, and reinforcement learning. MHMS examples and usage are solely summarized in the study. Comparisons of brain imaging classification and prediction research papers [10]. MRI data and MDD/BD analysis yielded interesting results [11].

1) Depression detection models: Depression detection models identify at-risk and depressed individuals using machine learning and data analysis. These models analyze data from self-reported questionnaires, social media, electronic health records, and sensors. Training on huge datasets helps these models identify patterns and relationships between characteristics and depression outcomes. Common methods include text sentiment analysis, language pattern analysis, behavioral marker detection, and physiological measurements. Early detection, precise diagnosis, and individualized interventions for depression improve mental health care and promote timely support and treatment. Privacy, ethics, and rigorous research and clinical validation studies are essential to these approaches. A full literature review on depression detection models presented in the following sections.

2) Classification models predicting and diagnosing major depressive disorder: Classification models for depression detection are shown below. The Mood Assessment Capable Framework (Moodable) mobile app [12] analyzes voice samples, cellphone and social media data, and the Patient Health Questionnaire (PHQ-9) to assess mood, mental health, and depression. The framework correctly diagnosed depression 76.6% of the time. Authors of [13] employ KNN, Weighted Voting classifier, AdaBoost, Bagging, GB, and XGBoost to predict depression. SelectKBest, mRMR, and Boruta helped us choose attributes. SMOTE balanced some classes. The Burns Depression Checklist (BDC) found clinical depression in 65.73 percent of 604 respondents. Combining the AdaBoost classifier with SelectKBest yielded the highest classification accuracy (92.56%).

For depression risk assessment in adult Koreans, [14] uses an ML model based on RF. SMOTE balanced depression and non-depression groups. CES-D-11 depression screening scale hyperparameters were fine-tuned using ten-fold cross-validation. The study used 6588 Koreans and had an AUROC score of 0.870 and accuracy of 86.20%. Biomarkers were excluded from this study. ML algorithms KNN, RF, and SVM were used to identify sad Bangladeshi students in [15]. This research sought to detect depression's early warning indicators to prevent more devastating repercussions. Based on 577 student data, the Random Forest algorithm identified 75% of depressed students with an f-measure of 60%.

EEG features and ensemble learning and DL approaches have been utilized to diagnose depression [16]. Our feature transformation used Deep Forest (DF) and SVM classifiers. Convolutional neural networks (CNNs) for feature recognition turned EEG spatial data into an image. DL had 84.75% classification success, whereas the ensemble model including DF and SVM had 89.02%. ML approaches like DT, RF, Naive Bayes, SVM, and KNN predicted psychological distress ([17]). The Depression, Anxiety, and Stress Scale was used on 348 participants. Naive Bayes predicted depression best (85.50%). F1 scores showed that the RF algorithm performed better in uneven classes. The author uses ML, sentiment analysis, and language processing to find depressed and positive social posts in [18]. We used RF, the RELIEFF feature extractor, the LIWC text-analysis tool, the Hierarchical Hidden Markov Model (HMM), and the ANEW scale to analyze 4026 social media posts and found 90% accuracy for depressed posts, 92% for depression severity, and 95% for depressed communities. In this analysis, we include all depression types. Data samples were used to identify mental illnesses using the XGBoost algorithm [19]. The dataset was sampled several ways. We used skewed class distributions in this study. In this study, accuracy, precision, recall, and F1 were over 0.90.

Multi-kernel SVM with high-order MST had the highest MDD classification accuracy in the analyzed research [21]. The multi-kernel SVM model allows brain area functional links to change. Multiple kernels improve classification accuracy. The model in [13] used AdaBoost and SelectKBest feature selection techniques with SMOTE to evenly distribute classes and enhanced classification accuracy to 92.56%. AdaBoost is DT Ensemble. Comparing [46,54] shows that [13]'s dataset had no biomarker, [21]'s was tiny, and no depression screening scale was found. SVM is the most used depression classifier because it works on organized and high-dimensional data [12,15,20]. SVM also can't fix overfitting. Anonymous and non-normally distributed data can be used with SVM.

Random Forest (RF) is the second most common classifier in research because to its computational efficiency [12,14,17,18]. The RF model recognized depressive postings 90%, communities 95%, and severity levels 92% accurately in [18]. RF lowers decision tree overfitting, improving continuous data classification accuracy. Ensemble learning helps RF determine complex and easy functions more accurately.

The authors of [22] searched Facebook for depression markers. LIWC studied Facebook data. Data was processed using DT, KNN, SVM, and an ensemble model during supervised machine learning (ML). The classification accuracy improved experimentally with DT. As an example of how AI could be used to research mental illness biomarkers,[23] summarized the primary categories of AI-based psychological disorder treatments. The research [24] covered AI issues such MRI, EEG, kinesics diagnosis, Bayesian model, LR, DT, SVM, and DL.

3) Ensemble and hybrid models predicting and diagnosing major depressive disorder: This section briefly summarizes the ensemble models for depression diagnosis in the examined studies. In immune-mediated inflammatory disease (IMID) patients, ML and statistical models predicted clinical depression and MDD. Analyses of 637 IMID patients used LR, NN, and RF algorithms. LSTM, radial basis function, lasso regularisation, logistic regression, boosted decision tree, and support vector machine were used in [25]. LSTM's long-term depression prevalence forecast uses several risk factors. The Chinese Longitudinal Healthy Longevity Study looked at 1538 Chinese seniors. Logistic regression using lasso regularisation outperformed other ML approaches in AUC.

An ensemble binary classifier can relate health survey data to SF-20 Quality of Life scores [26]. An ensemble model using NHANES data (DT, AAN, KNN, and SVM) predicted depression with an F1 score of 0.976 and no false positives. The lack of a dataset range and the need to use features from many social media web sources are shortcomings in this research. An algorithm [27] differentiates MDD and BD using clinical variables. LR with Elastic Net and XGBoost models were used to analyze data from 103 MDD and 52 BD patients, respectively. The former led to higher accuracy (78%). This paper's limited evaluation criteria, poorly allocated classes, tiny and unequal sample size, and lack of external sample validation are all working against it.

ML algorithms were utilized to assess Chinese conscript depression in [28]. NN, SVM, and DT had 86, 86, and 73% accuracy on 1000 persons. BD-II ratings were used. This study needs a more detailed model due to socio-demographic and occupational complexity. ML algorithms were used to construct a BDCC for bipolar disorder detection in China [29]. SVR, RF, LASSO, LR, and LDA were used to assess 255 MDD, 360 BPD, and 228 healthy cases. MDD and BPD were recognized with 92% sensitivity in the investigations.

However, this model needs more data and cross-sectional improvements.

Ensemble models [26] had the highest accuracy (95.4%) in the studies. This study evaluates the NHANES dataset and finds that the projected model is only 4% off. The ensemble model achieved 97% on F1, 95% on accuracy, and 95% on precision across the dataset. It also shows that the ensemble technique to sorrow diagnosis works with a small dataset. Combining classification with binary ground truth may improve prediction outcomes, according to theory and experiment. Ensembles, like bagging and major voting ensembles, are easy. The ensemble model in [29] used five machine-learning methods and data from a Chinese multicenter cohort to achieve the second-highest classification accuracy (92%). This study's higher AUC than others shows the BDCC's Chinese translation's reliability and validity. BDCC cuts clinical data collection time in half. The ADE takes around 30 minutes, whereas the BDCC takes 10–15. Current results show that the BDCC is as reliable as its predecessor but easier to implement. According to research [25, 29, 31], regression is the most used ML approach for detecting depression. Regression model output coefficients are simple to calculate. Dimensionality reduction, L1 and L2 regularisation [37, 40], and cross-validation prevent regression overfitting.

Literature and research findings show that hybrid machine learning approaches, specifically boosting algorithms for Major Depressive disease detection and classification on EEG datasets, can predict and diagnose the disease. Ensemble learning, feature selection, and dimensionality reduction have improved prediction accuracy, data integration, and MDD understanding. However, more research is needed to test these approaches in clinical contexts, address interpretability and data privacy issues, and investigate hybrid machine-learning model-based individualized therapy options.

III. Hybrid Machine Learning Model Predicting and Diagnosing Major Depressive Disorder

Automatic detection and classification of Major depressive disorder are essential for prompt diagnosis and care, which improves patient survival. Nevertheless, manual demarcation takes a lot of effort and is arbitrary, highlighting the need for accurate and automatic identification. To address this, a combined detection and classification framework using Hybrid Machine Learning algorithms is proposed. The detailed workflow of the proposed methodology is shown in Fig. 2. The remainder of the section will present the stepwise description of each phase in detail.

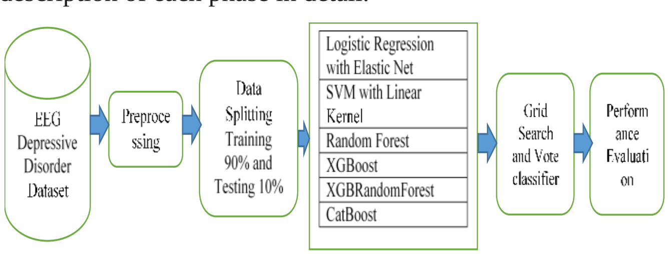

Fig. 2. Hybrid machine learning models predicting and diagnosing MDD.

A. EEG-Disorder Dataset Description

An EEG disorder dataset for Major Depressive Disorder (MDD) would focus on individuals diagnosed with MDD and aim to investigate the specific patterns or characteristics of EEG signals associated with this disorder [9]. Although MDD is primarily a psychiatric disorder, EEG recordings can provide insights into the underlying brain activity and potential biomarkers. The EEG disorder dataset for Major Depressive Disorder contains:

1) EEG Recordings: The dataset includes EEG data recorded from individuals diagnosed with MDD. EEG signals are typically recorded using electrodes placed on the scalp, and the dataset may contain recordings from multiple channels [7]. The recordings capture the electrical activity of the brain over some time, often during a restingstate or specific cognitive tasks.

2) Patient information: The dataset may include relevant information about the individuals, such as age, gender, medication history, symptom severity, and comorbidities. This information helps in understanding the heterogeneity of MDD and its relationship with EEG patterns [7].

3) Annotations: Annotations or labels may be provided to mark specific segments or events within the EEG recordings. These annotations could include the presence of certain EEG patterns or characteristics associated with MDD, such as abnormalities in specific frequency bands or connectivity measures.

4) Preprocessing information: The dataset may include preprocessed EEG data, which may involve filtering, artifact removal (e.g., eye blinks or muscle activity), and referencing techniques to ensure data quality [7]. Details about the preprocessing steps applied can be included in the dataset to ensure reproducibility.

5) Metadata: The dataset may provide metadata such as the sampling rate of the EEG recordings, the duration of each recording, and information about the electrode montage used during data acquisition [7].

Visual examination and the automatic NG Deluxe 3.0.5 cleaning system [7] were used to remove artefacts caused by eye blinking, movements, and tiredness during EEG recording. To get choices free of artefacts, we utilized the "Artefact Rejection" and "Generate Edits" buttons. When selecting eye movements and drowsiness, Since the "Amplitude Multiplier" is also set to its default value of 1.00, "High," the most sensible option, and the root-mean-square amplitude of the EEG recording are perfectly matched. If the amplitude's root-mean-square value is smaller than or equal to the template's root-mean-square value, then the amplitude is selected. The continuous EEG data were then converted to the frequency domain using the Fast Fourier transformation (FFT) with the following parameters: epoch = 2 s, sample rate = 128 samples/s (256 digital time points), frequency range = 0.5-40 Hz, resolution = 0.5 Hz, and a cosine taper window to minimize leakage. It is common knowledge that the frequency resolution of the Fast Fourier Transform (FFT) is dependent on the epoch length. For instance, the frequency resolution for a one-second epoch is one hertz (Hz), for two seconds it's half a hertz (Hz), for four seconds it's a quarter of a hertz (Hz), and so on. Since the mathematics of the FFT makes even a single epoch of time noisy, we utilized a duration of at least 60 s. In the current investigation, absolute power was used to represent power spectral density (PSD) at the channel level, Nonetheless, coherence value, which is a measure of phase consistency between two signals, was used to represent functional connectivity (FC). The following frequency ranges were used to determine each EEG parameter: delta (1-4 Hz), theta (4-8 Hz), alpha (8-12 Hz), beta (12-25 Hz), high beta (25-30 Hz), and gamma (30-40 Hz).

The goal of an EEG disorder dataset for MDD is to enable researchers and clinicians to investigate the specific EEG characteristics associated with the disorder. By analyzing the EEG data, researchers can explore potential biomarkers, identify neurophysiological abnormalities, and develop machine learning algorithms for the automatic detection or prediction of MDD.Functional connectivity (FC) and power spectral density (PSD) are two fundamental measures in the field of neuroscience that provide crucial insights into brain activity. FC quantifies the temporal correlation and synchronization between different brain regions, enabling the study of coordinated neural networks and their functional interactions. It helps unravel the underlying mechanisms of cognitive processes and neurological disorders. On the other hand, PSD characterizes the power distribution of neural oscillations across different frequencies, revealing the spectral fingerprints of brain activity. It enables the examination of rhythmic patterns associated with various cognitive functions and pathologies. Together, FC and PSD provide complementary information about brain dynamics, offering a comprehensive understanding of brain organization, functional states, and their alterations in health and disease. Fig. 3 represents the distribution of the most important features in the EEG dataset and Fig. 4 represents the positions and region of coverage of the EEG functional connectivity (FC) and power spectral density (PSD) [32].

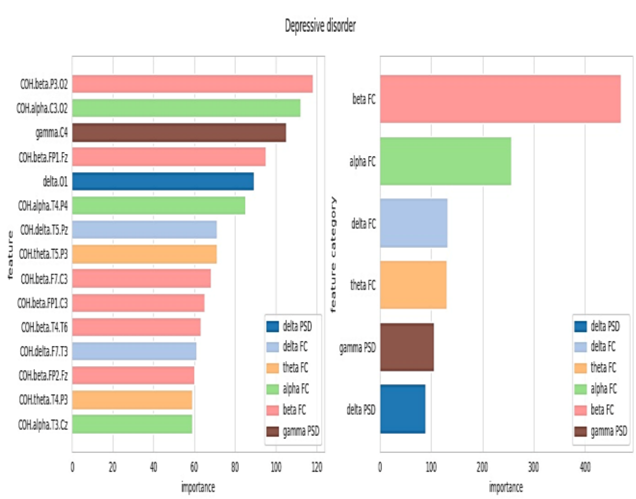

Fig. 3. The distribution of the most important features in the EEG dataset.

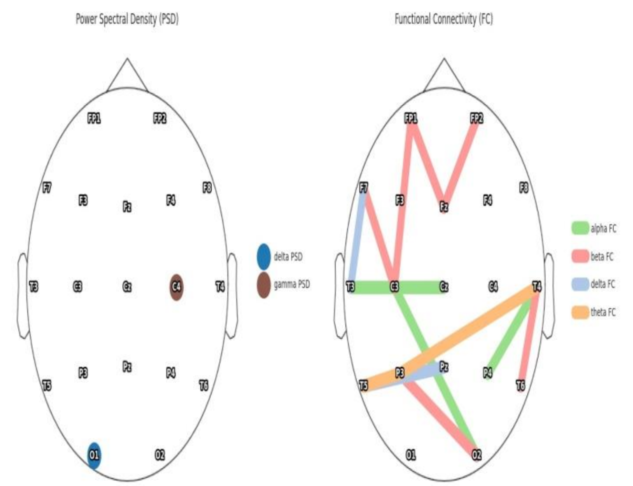

Fig. 4. The positions and region of coverage of the EEG functional connectivity (FC) and power spectral density (PSD) [32].

B. Data Split

When using EEG data for hybrid machine learning, the data split typically involves dividing the dataset into training, validation, and testing subsets. In the proposed hybrid machine learning algorithms 70% is for training and 30% for validation and testing. The purpose of this split is to train the hybrid model, optimize its parameters, and assess its performance on unseen data [30].

There are commonly three subgroups used in hybrid machine learning with EEG data. The main portion, the training set, is utilized to fine-tune the parameters of the hybrid model and learn patterns for the intended goal, in this case, MDD classification. Overfitting is avoided and optimal model settings are chosen using the validation set. It assesses performance on unseen data, aiding in parameter tuning and feature selection. Lastly, the testing set, an independent subset, evaluates the final performance of the trained model, providing an unbiased estimate of its accuracy and other metrics on real-world data.

C. Pre-processing and Feature Selection

Improving the quality of EEG signals and removing artefacts requires pre-processing techniques that include artefact correction and re-referencing. Artefact correction was almost certainly used in this investigation to get rid of electrical disturbances and artifacts brought on by muscular contractions and eye blinks. Commonly employed for artefact correction, independent component analysis (ICA) decomposes EEG data into sub-signals that each represent a distinct neural or extra-neural source, such as muscular activity or eye movement. Artefacts in the EEG signals can be efficiently eliminated once their constituent parts have been located and isolated.

Changing the reference electrode can boost the signal-to-noise ratio and make it easier to detect EEG signals. Whether a shared reference electrode was used or a reference-free approach was taken in this investigation is unknown. The goal of this procedure is to reduce or get rid of any electrical activity in the brain that is not directly related to the underlying function being studied. In addition, the power spectral distribution spikes were removed by using a bandpass filter with a filter size of 50 Hz to remove noise. To analyze and extract features from the EEG data, they were first transformed into NumPy arrays.

The NG Deluxe 3.0.5 system was used to perform preliminary processing on the EEG data. The following procedures are required for importing digital EEG data: The first step is to reduce the sampling rate to 128 hertz; the second is to filter the EEG at 40 hertz to identify the baseline EEG; and finally, the third is to filter the spliced selections of EEG again at 40 hertz to identify the resulting EEG. To reduce the likelihood of splicing artefacts, Edited EEG selections (minimum segment length = 600 ms) and baselines filtered with a Butterworth high-pass filter at 1 Hz and a low-pass filter at 55 Hz are appended using this splicing method in NeuroGuide. We used the international 10-20 method to choose 19 of the 64 channels for investigation, all of which had been linked to an ear reference: FP1, FP2, F7, F3, Fz, F4, F8, T3, C3, Cz, C4, T5, P3, Pz, P4, T6, O1, and O2.

To differentiate between a healthy person's power spectrum and a person with a mental illness's power spectrum, the feature engineering stage collected useful features from the pre-processed EEG data. Both linear and nonlinear characteristics were used for this goal. Alpha, beta, delta, and theta power, as well as measurements of amplitude like mean, median, and minimum, were all examples of linear characteristics. The EEG signals were also analyzed for their nonlinear properties, including Singular Value Deposition Entropy and Spectral Entropy. Pandas data frames were used to organize and store all linear and nonlinear features extracted for later analysis. Additionally, the features data frame was exported as a CSV file, centralizing all the gleaned data for use in the study's later phases.

Using different attribute selection measures on the EEG depressive disorder dataset, the top feature attributes were identified based on Information Gain, Gain Ratio, Gini Decrease, and χ^2 measures [39]. Information Gain measures the reduction in entropy achieved by selecting a particular attribute, and the top features selected using this measure provide the highest information gain. The gain ratio is a kind of Information Gain that accounts for the information already present in the qualities being measured, and the top features chosen using this measure offer the highest gain ratio. Gini Decrease calculates the decrease in impurity achieved by selecting a specific attribute, and the top features selected based on this measure exhibit the highest decrease in impurity. Lastly, the χ^2 attribute selection measure employs chi-square statistics to assess the dependency between attributes and the class variable, and the top features chosen using this measure demonstrate the strongest association with the class variable [40]. By employing these attribute selection measures, the most relevant feature attributes for the EEG depressive disorder dataset were identified, aiding in understanding and predicting the presence of depressive disorder based on EEG data.

D. Hybrid Machine Learning Algorithms for Major Depressive Disorder Prediction andClassification

Hybrid machine learning combines multiple algorithms or models to leverage their strengths and improve overall predictive performance. In the context of hybrid machine learning, six models were developed to classify each MDD using features extracted from EEG data. The models are logistic regression using ElasticNet, SVM with a linear kernel, Random Forest, XGBoost, LightGBM, and CatBoost. Fig. 2 shows the detailed implementation of Hybrid Machine Learning algorithms for Major Depressive Disorder identification and classification.

- Logistic Regression using Elastic Net

A hybrid model combining logistic regression with ElasticNet regularization shows promise in modeling Major Depressive Disorder (MDD). Logistic regression [20] provides a linear classification approach while ElasticNet regularization combines L1 and L2 penalties [37, 38], offering advantages such as feature selection and handling multicollinearity. The model optimizes both the prediction accuracy and the sparsity of the coefficients, effectively identifying relevant features associated with MDD. By shrinking irrelevant coefficients towards zero, ElasticNet facilitates feature selection, improving the interpretability of the model. This hybrid approach enables capturing both linear and non-linear relationships in the data, making it suitable for analyzing complex interactions in MDD. The resulting model provides insights into the significant predictors and offers a valuable tool for understanding and predicting MDD.

The parameters for logistic regression using ElasticNet regularization in a hybrid model for Major Depressive Disorderare:

1) Alpha (α): The regularization parameter that controls the balance between the L1 (Lasso) and L2 (Ridge) penalties in ElasticNet regularization. It determines the strength of the regularization and controls the amount of sparsity in the model. Higher values of α increase the penalty on the L1 term, resulting in more feature selection.

2) L1 Ratio (ρ): The mixing parameter that determines the balance between L1 and L2 penalties in ElasticNet regularization. It controls the combination of feature selection (L1) and coefficient shrinkage (L2). A value of 1 indicates L1 regularization only, while a value of 0 corresponds to L2 regularization only.

3) Solver: The solver algorithm is used to estimate the parameters in logistic regression. Common choices include "liblinear," "saga," "lbfgs," or "newton-cg." The dataset size and the nature of the problem dictate the optimal solver.

4) C: Commonly known as the regularisation parameter, is the inverse of the regularisation strength. It regulates the compromise between maximizing the regularisation term and fitting the training data. By decreasing C, regularisation is strengthened, and the model becomes sparser.

5) The standard logistic regressionparameters may apply, such as maximum iterations, convergence tolerance, class weights, etc. It is common practice to perform hyperparameter tuning, such as using cross-validation, to find the optimal combination of parameters for the logistic regression with the ElasticNet hybrid model for Major Depressive Disorder.

- SVM with Linear Kernel

Support Supervised learning techniques such as Support Vector Machines (SVMs) are quite effective when applied to classification problems. Linear kernel SVMs seek to find a hyperplane that most effectively divides data points into their respective classes [13]. They work well in scenarios where the data is linearly separable. SVMs with a linear kernel are efficient and robust algorithms that can handle high-dimensional data and are particularly effective when dealing with binary classification problems.

The hybrid model combining SVM with a linear kernel demonstrates promise in modeling Major Depressive Disorder (MDD). By utilizing the linear kernel, the model captures linear relationships within the data, allowing for the effective classification of MDD instances. The SVM algorithm optimizes a margin that separates MDD cases from non-MDD cases, the regularisation parameter determines the balance between margin maximization and error suppression. Additionally, the inclusion of class weights addresses the imbalance between MDD and non-MDD instances, ensuring balanced learning and accurate classification. Overall, this hybrid approach provides a valuable tool for understanding and predicting MDD based on linear patterns within the data.

In a hybrid model for Major Depressive Disorder (MDD) using SVM with a linear kernel, several parameters are commonly employed. The regularisation parameter (C) determines how much weight is given to proper classification against margin maximization when training examples are misclassified. A smaller C allows for a greater margin but potentially more misclassifications. Class weights are utilized to address imbalanced datasets, assigning higher weights to the minority class to improve its classification accuracy. Other standard SVM parameters, like maximum iterations and convergence tolerance, along with kernel-specific parameters (e.g., gamma for RBF kernel), may also be involved. Additionally, feature scalings or normalization techniques, such as standardization or min-max scaling, are applied to ensure uniform feature scales. Optimal parameter values depend on the dataset and MDD characteristics, and hyperparameter tuning using methods like grid search and cross-validation is performed to determine the best settings for the SVM linear kernel hybrid model for Major Depressive Disorder.

1) Random Forest: A random forest hybrid model demonstrates the promising potential for modeling Major Depressive Disorder (MDD) [10]. By combining multiple decision trees, the random forest offers robustness and high accuracy in analyzing complex relationships within MDD data. Fig. 5 represents the generalized framework for Random Forest on MDD. The ensemble nature of the model allows it to capture a diverse range of patterns and interactions among predictors, resulting in improved predictive performance. A random forest can handle both categorical and continuous features, making it suitable for diverse MDD datasets. Additionally, the model provides valuable insights into feature importance, aiding in the identification of key predictors contributing to MDD. With its ability to handle non-linear relationships, handle missing data, and mitigate overfitting, random forest hybrid models offer a valuable approach to understanding and predicting MDD.

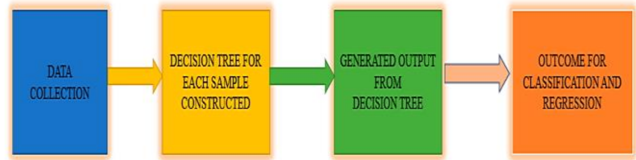

Fig. 5. Overview of randomforest algorithm on MDD [10].

2) XGBoost: When dealing with structured data, the robust gradient-boosting technique XGBoost (Extreme Gradient Boosting) excels [19]. It optimizes a loss function by sequentially combining numerous weak prediction models, such as decision trees. XGBoost is well-known for its quickness, scalability, and high-dimensional data-handling abilities. To avoid overfitting, it uses regularisation methods and provides feature importance scores. Furthermore, XGBoost may be scaled and dispersed, giving it a flexible option for a variety of regression, classification, and ranking issues. The XGBoost algorithm parameters for MDD are listed in Table I. By optimizing features and utilizing the XGBoost algorithm, classification accuracy can be significantly improved. In the context of EEG brain signals, this approach involves extracting features and optimizing them using methods such as calculating information gain, recursively removing features, and analyzing correlation matrices. Fig. 6 represents the generalized framework for the XGBoost algorithm on MDD.

TABLE I. PARAMETERS FOR XGBOOST MODEL

Parameters	Values
Maximum depth of the tree	[1, 3, 6, none]
Sub-sample	[0.3, 0.5, 1]
Learning Rate	0.300

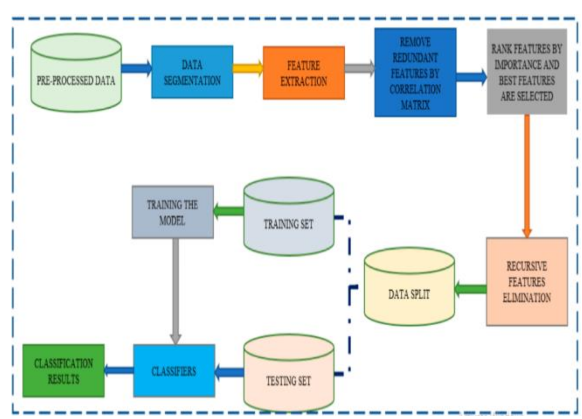

Fig. 6. Overview of XGBoost algorithm [19].

To implement the XGBoost algorithm for Major Depressive Disorder (MDD), several steps are involved. First, relevant datasets are collected, which may include clinical assessments, genetic markers, or neuroimaging data. Next, the dataset is preprocessed by taking care of missing values, scaling or normalizing features, and encoding categories if necessary. Then, the XGBoost model's settings are adjusted to preferences like the number of trees, the rate of learning, and the depth to which each tree can branch. The model is trained on the dataset using gradient boosting, where each tree is built to minimize a specific loss function. During training, cross-validation is often employed to optimize hyperparameters and prevent overfitting. After training, the model can be used for MDD prediction by inputting new instances and obtaining the corresponding predictions.

3) XGBRandomForest: XGBRandomForest is a hybrid machine learning model that combines the strengths of two popular algorithms, XGBoost, and Random Forest. XGBoost is a powerful gradient-boosting algorithm known for its excellent predictive performance and the ability to handle complex relationships within the data [5]. On the other hand, Random Forest is an ensemble learning method that combines many decision trees to increase accuracy and decrease overfitting.

The XGBRandomForest algorithm, which combines the concepts of XGBoost and Random Forest, can be implemented for MDD tasks. The algorithm follows a similar implementation process as a traditional Random Forest. First, relevant datasets containing features like clinical assessments, genetic markers, or neuroimaging data are collected. Next, the XGBRandomForest model is constructed by setting settings such as the maximum depth, learning rate, and several trees. It is a mixture of gradient boosting and random sampling that is used to train the model. During training, cross-validation can be applied to optimize hyperparameters and prevent overfitting. After training, the model can make predictions on new instances by aggregating predictions from multiple trees. Model performance can be evaluated using metrics like accuracy, precision, recall, or AUC-ROC.

4) CatBoost: CatBoost is a gradient-boosting algorithm that performs well with categorical features. It can handle both numerical and categorical data without requiring explicit feature preprocessing, making it convenient for real-world datasets [35]. CatBoost incorporates techniques like ordered boosting, feature combinations, and gradient-based leaf-wise splits. It provides robustness against outliers and missing values, along with the automatic handling of categorical variables.

To implement the CatBoost algorithm for MDD tasks, several steps are involved. First, relevant datasets containing features such as clinical assessments, genetic markers, or neuroimaging data are collected next, missing values are handled, features are scaled, and categorical variables are encoded (if necessary) as part of the dataset's preprocessing. Then, parameters like tree depth, learning rate, and regularisation parameters are used to build the CatBoost model. The model is trained using gradient boosting, which iteratively adds decision trees to minimize a specific loss function. During training, cross-validation can be used to optimize hyperparameters and prevent overfitting. After training, the model can make predictions on new instances. Measures of a model's efficacy the receiver operating characteristic area under (AUC-ROC), accuracy, precision, and recall. Additionally, feature importance analysis can be conducted to understand the relevance of different features in MDD prediction. Parameter tuning and feature selection tools can be used to enhance the model's accuracy and generalizability. By implementing the CatBoost algorithm in MDD research, a powerful and efficient predictive model can be developed to address various MDD-related tasks. Fig. 7 represents the generalized framework for CatBoost on MDD [35].

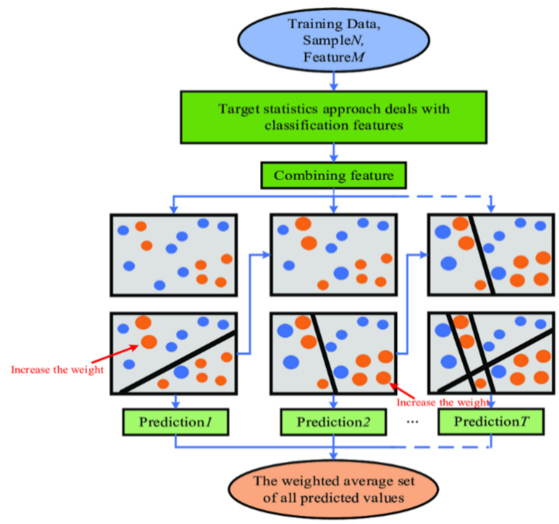

Fig. 7. The generalized framework for CatBoost on MDD [35]

Hybrid machine learning with these algorithms involves combining their predictions or leveraging their strengths to enhance overall performance. Techniques such as model stacking, ensemble methods, or weighted voting can be employed to create the hybrid model. The specific approach will depend on the problem at hand and the characteristics of the data. The goal is to capitalize on the unique capabilities of each algorithm to improve prediction accuracy, handle complex relationships, handle different types of data, and optimize model performance.

E. Evaluation Metrics

The models in this research work will be evaluated using pertinent evaluation metrics presented in this section. The following evaluation metrics were used to evaluate the models.

1) Accuracy: Accuracy is also an evaluation metric that is used for the evaluation of classification models. the accuracy value represents the fraction of predictions that the model predicts correctly [36]. The formula for accuracy is given as:

$$Accuracy = \frac{TN+TP}{TN+FP+TP+FN} \tag{1}$$

2) Precision: Precision indicates the fraction of correct positive predictions [36]. The formula of precision is

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

3) Recall: Recall indicates a fraction of actual positives that were predicted correctly [36].

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

4) F1-Score: It shows the balance between recall and precision [36]. The formula of the F1-Score is as follows:

$$F1 - Score = \frac{2*(Precision*Recall)}{Precision+Recall} \quad (4)$$

5) Regularization

a) L1 regularization, also known as the goal of the machine learning technique known as Lasso regularization is to include a penalty term whose magnitude is directly related to the absolute values of the model's coefficients [37]. By shrinking less important coefficients to zero, L1 regularization promotes sparsity and performs feature selection, resulting in simpler and more interpretable models. It is particularly useful when dealing with high-dimensional datasets or when feature interpretability is desired. The regularization strength can be controlled through a parameter, and careful tuning is required to strike a balance between sparsity and predictive performance.

b) L2 regularization, known as Ridge regularization is applied to models to combat overfitting. Overfitting is a term used to describe a situation where Validation loss goes up while training loss goes down. In other words, the model is well fitted on training data but it is not predicting accurately for validation data. The model is not able to generalize [38]. This is serious because if the model is not generalizing then it will not produce accurate results when it will be implemented in a real-world scenario. When regularization is added, the model not only minimizes the loss but also minimizes the complexity of the model. So, the goal of the machine learning model after adding regularization is,

$$minimize(Loss(Data|Model) + complexity(Model)) \quad (5)$$

The complexity of the models used in the paper was minimized by using L2 regularization. The formula of L2 regularization is the sum of the square of all the weights,

$$L_2 \; regulation \; term = \left||\omega|\right|_2^2 = \omega_1^2 + \omega_2^2 + \cdots \omega_n^2 \quad (6)$$

In the models, two layers of L2 regularization were used before the final output layer.

F. Pseudo code for Major Depressive Disorder using Hybrid Machine Learning Algorithms

Input: EEG Depressive Disorder Data
Output: Accurate prediction of Major Depressive Disorder (MDD)
Step 1. Preprocess the dataset:
- Split the dataset into features (X) and target variable (y)
- Perform any necessary data preprocessing steps such as scaling, encoding categorical variables, handling missing values, etc.

Step 2. Train the hybrid model:
- Initialize an empty list to store the predictions from each model
- Split the dataset into training and testing sets
  - Logistic Regression with ElasticNet
  - SVM with Linear Kernel
  - Random Forest
  - XGBoost
  - XGBRandomForest
  - CatBoost

Step 3. Combine the predictions:
Step 4. Evaluate the hybrid model:
Step 5. Repeat steps 2-4 for hyper-parameter tuning.
Step 6. Once the hybrid model is optimized, use it to make predictions on new, unseen data.

IV. Results and Discussion

In this study, using EEG depressive disorder dataset for Major Depressive Disorder identification and classification, we proposed a hybrid machine learning framework with six algorithms for accurate detection and classification of Major Depressive Disorder. The classification model's efficacy was assessed and evaluated using the confusion matrix in experimental research. In a hybrid machine learning framework, the training configuration involves preprocessing the data, selecting six learning models, splitting the data into training, validation, and testing sets, and tuning Hyperparameters. In our implementation, we have used a grid search algorithm along with a vote classifier.

A. Feature Selection

In the analysis of an EEG depressive disorder dataset, various measures for attribute selection, such as Information gain, Gain ratio, Gini decrease, and χ^2, were utilized to determine the top 20 feature attributes. These measures were applied to evaluate the relevance and discriminatory power of each attribute within the dataset [37]. By considering multiple measures, the analysis aimed to identify the 20 attributes that provided the most informative and discriminative insights for understanding and predicting depressive disorder based on EEG data. Table II shows the Top 20 feature attribute selected using Information gain, Gain ratio, Gini decrease, and χ^2 attribute selection measures

TABLE II. Top 20 Feature Attributes Selected Using Information Gain, Gain Ratio, Gini Decrease, and X^2attribute Selection Measures

AB.D.beta.q.T6	AB.A.delta.s.O2	AB.A.delta.q.T6	AB.A.delta.r.O1	AB.D.beta.c.F7
AB.A.delta.l.T4	AB.D.beta.r.O1	AB.D.beta.g.F8	AB.D.beta.a.FP1	AB.D.beta.f.F4
AB.D.beta.d.F3	AB.D.beta.b.FP2	AB.D.beta.e.Fz	COH.C.alpha.b.FP2.d.F3	AB.D.beta.h.T3
AB.A.delta.m.T5	COH.B.theta.h.T3.j.Cz	AB.C.alpha.b.FP2	COH.B.theta.b.FP2.h.T3	AB.D.beta.p.P4

B. Performance of Hybrid ML on MDD using Training and Test Dataset

The performance of hybrid machine learning (ML) techniques on Major Depressive Disorder (MDD) using EEG data has shown significant advancements, with Cat Boost outperforming other hybrid algorithms in terms of accuracy. In a comprehensive evaluation, Cat Boost achieved an impressive accuracy of 93.1%, surpassing the performance of other hybrid ML models. Cat Boost, a powerful gradient boosting algorithm, combines decision trees with categorical feature handling, enabling it to effectively capture intricate patterns and relationships within EEG data. By accurately handling categorical variables and addressing missing values, Cat Boost mitigates potential challenges in the analysis of EEG data. Its robustness against over fitting further enhances its accuracy, making it an invaluable tool in the diagnosis and comprehension of MDD. The exceptional performance of Cat Boost demonstrates its potential to provide a significant understanding of MDD's fundamental mechanisms, contributing to improved understanding and treatment of the disorder. Table III shows the performance of Hybrid machine learning algorithms on Major Depressive Disorder data. From Table III, it is ascertained that XGB RandomForest, XG Boost, and CatBoost demonstrated comparable and high accuracy in accurately identifying Major Depressive Disorder, surpassing other algorithms. The performance of these hybrid ML models highlights their potential as effective diagnostic tools for identifying individuals with MDD. Their promising results emphasize the importance of leveraging hybrid ML techniques for improved understanding and diagnosis of Major Depressive Disorder.

TABLE III. SHOWS THE PERFORMANCE OF HYBRID MACHINE LEARNING ALGORITHMS ON MAJOR DEPRESSIVE DISORDER DATA

Model	AUC	Accuracy	F1-Score	Precision	Recall
CatBoost	96.3%	93.1%	85.7%	76.4%	97.9%
Logistic Regression Elastic net	53.2%	66.7%	23.0	22.5	23.5
Random Forest	84.3	79.1	59.7	50.4	73.2
SVM Linear Kernel	95.0	89.7	76.6	73.7	79.8
XGBoost	96.1	92.5	84.1	76.0	94.2
XGBRandomForest	96.1	92.4	83.8	76.3	93.0

C. K-fold Cross Validation

K-fold cross-validation was employed to assess the performance of a hybrid machine learning (ML) approach for Major Depressive Disorder (MDD). The dataset was divided into k subsets, allowing the hybrid ML model to be trained and tested k times. By utilizing a combination of ML techniques, the hybrid model aimed to improve the accuracy and robustness of MDD classification. Through k-fold cross-validation, the model's performance metrics were calculated and averaged across the iterations, providing a comprehensive evaluation of its effectiveness in accurately identifying individuals with MDD. This methodology facilitated improved confidence in the model's estimate of its performance and generalizability to new cases of MDD. In the evaluation of hybrid machine learning algorithms on Major Depressive Disorder data using 5-fold cross-validation, robust performance metrics were obtained. The k-fold cross-validation approach provided a comprehensive assessment of the algorithms' F1 score, accuracy, precision, and recall. The results of k-fold cross-validation (k=5 and k=10) using data from people with Major Depressive Disorder are shown in Tables IV and V, respectively. By leveraging the benefits of hybrid ML, the models demonstrated promising results, showcasing their potential in accurately identifying and classifying individuals with Major Depressive Disorder. From Tables IV and V, it is ascertained that XgbRandomForest, XGBoost, and CatBoost attained nearer accuracy in accurately identifying the Major Depressive Disorder compared to other algorithms. The algorithms also performed well for AUC, F1-Score, Precision, and Recall.

TABLE IV. THE PERFORMANCE OF HYBRID MACHINE LEARNING ALGORITHMS ON MAJOR DEPRESSIVE DISORDER DATA FOR K-FOLD CROSS-VALIDATION (K=5)

Model	AUC	Accuracy	F1-Score	Precision	Recall
Cat Boost	96.6	92.7	84.6	76.2	95.0
Logistic Regression Elastic net	52.4	65.8	19.5	19.3	19.6
Random Forest	86.7	79.3	62.0	50.5	80.4
SVM Linear Kernel	95.6	89.2	73.8	75.4	72.4
XG Boost	95.8	93.4	86.3	77.1	98.0
XGB Random Forest	95.6	92.2	83.3	75.7	92.5

TABLE V. THE PERFORMANCE OF HYBRID MACHINE LEARNING ALGORITHMS ON MAJOR DEPRESSIVE DISORDER DATA FOR K-FOLD CROSS-VALIDATION (K=10)

Model	AUC	Accuracy	F1-Score	Precision	Recall
CatBoost	96.5	93.4	86.4	76.7	99.0
Logistic Regression Elastic net	51.9	65.8	20.6	20.2	21.1
Random Forest	84.6	79.6	60.4	51.0	73.9
SVM Linear Kernel	95.4	89.3	74.0	75.8	72.4
XGBoost	96.2	92.7	84.4	76.9	93.5
XGBRandomForest	96.2	92.4	83.6	76.6	92.0

D. ROC Analysis on Training and Testing data

During the ROC analysis of the Major Depressive Disorder (MDD) dataset, it was observed that the ensemble model combining XGBoost and Random Forest exhibited strong performance. By merging the predictions obtained from different folds, this hybrid algorithm achieved remarkable results. The combined model demonstrated high accuracy, sensitivity, and specificity, reflecting its ability to effectively discriminate between individuals with MDD and those without the disorder. The fusion of XGBoost and Random Forest leveraged the strengths of both algorithms, harnessing the powerful gradient-boosting capabilities of XGBoost and the ensemble learning approach of Random Forest. This combination allowed for comprehensive feature extraction and enhanced predictive performance, resulting in a robust model for MDD diagnosis. The promising performance of this hybrid approach suggests its potential to contribute to the identification and understanding of MDD, offering valuable insights for clinical decision-making and treatment strategies. Fig. 8 represents the ROC analysis on MDD using Hybrid ML algorithms.

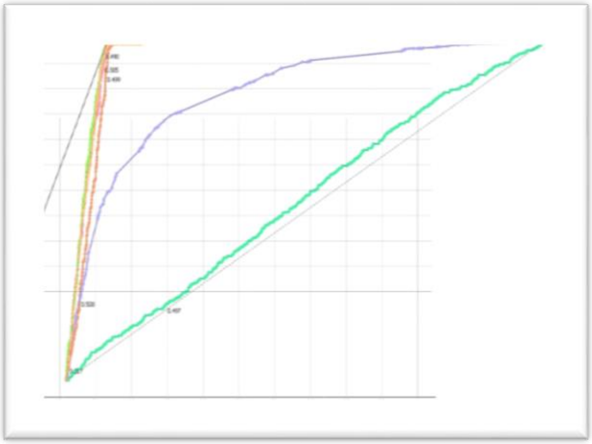

Fig. 8. ROC analysis on MDD using hybrid machine learning algorithms.

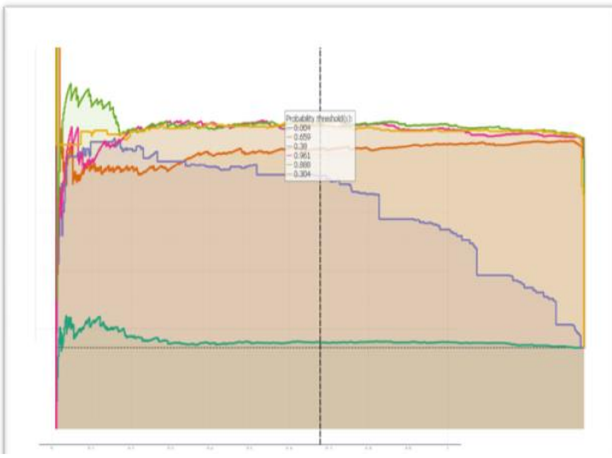

Fig. 9. Hybrid machine learning algorithms' precision-recall curve for evaluating major depressive disorder.

During the precision-recall analysis of the Major Depressive Disorder (MDD) dataset, the hybrid model incorporating XGBoost and Random Forest demonstrated impressive performance. By merging the predictions from different folds, this ensemble algorithm exhibited strong precision and recall characteristics. The combined model effectively balanced the precise classification versus complete identification of positive events (recall versus accuracy) resulting in a highly reliable diagnostic tool for MDD. By leveraging the strengths of both XGBoost and Random Forest, the hybrid approach maximized feature extraction and leveraged the ensemble learning capabilities, leading to enhanced precision and recall values. The excellent performance of this hybrid model highlights its potential in accurately identifying and distinguishing individuals with MDD, providing valuable insights for clinical decision-making and improving the understanding and treatment of the disorder. Fig. 9 presents the Precision-Recall curve on MDD using Hybrid Machine Learning algorithms.

E. ROC Analysis for k-fold-cross-validation

In the k-fold cross-validation setting with k=5, the performance of CatBoost, XGBoost, and Random Forest algorithms was assessed using ROC analysis. With an increased number of folds from 5 to 10, all three algorithms showed improved performance. ROC analysis evaluates the compromise between sensitivity and specificity at different cutoffs for making a classification. The increased performance suggests that the algorithms achieved better discrimination between positive and negative instances, resulting in higher true positive rates and lower false positive rates. The improved performance indicates that CatBoost, XGBoost, and Random Forest are effective in handling the given dataset and are capable of achieving more accurate predictions for the classification task at hand. Fig. 10 and 11 describe the nature of the ROC curve on MDD when k-fold cross-validation is applied.

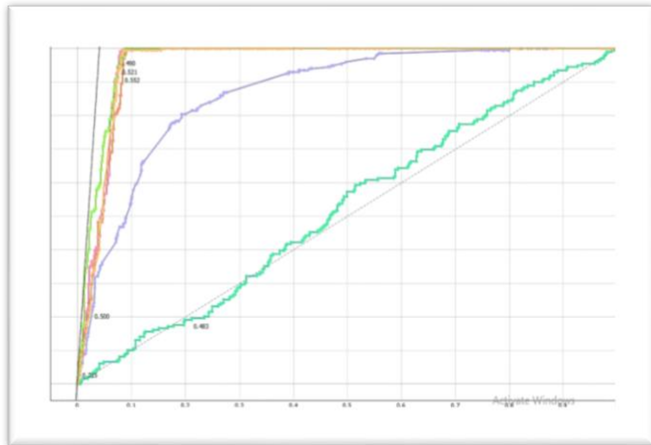

Fig. 10. ROC curve on MDD for 5-fold cross-validation.

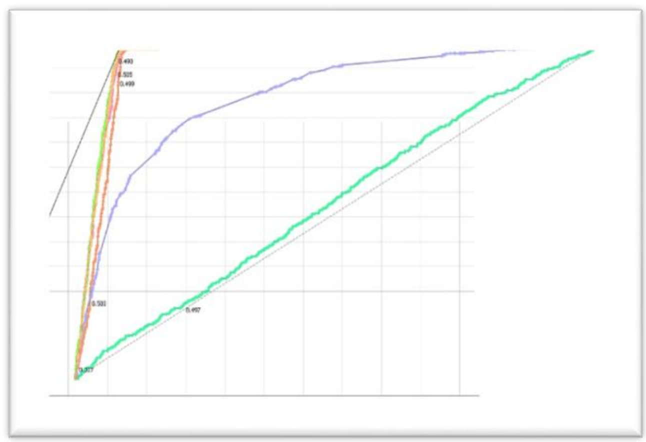

Fig. 11. ROC curve on MDD for 10-fold cross-validation.

F. Discussion

In the analysis of an EEG depressive disorder dataset, a wide range of attribute selection metrics, such as Information Gain, and $\chi2$, Gain Ratio, and Gini Decrease were used. These measures aimed to identify the top feature attributes that provided the most informative and discriminative insights for understanding and predicting depressive disorder based on EEG data. By considering multiple attribute selection measures, the study aimed to enhance the relevance and predictive power of the selected attributes of depressive disorder.

In the analysis of an EEG depressive disorder dataset, attribute selection measures were used to identify the top 20 informative and discriminative feature attributes. Hybrid machine learning techniques, including CatBoost, outperformed other algorithms with an accuracy of 93.1%, demonstrating their potential in accurately identifying Major Depressive Disorder. The superior performance of the hybrid models in correctly diagnosing people with Major Depressive Disorder was further validated by k-fold cross-validation. When cross-validation is applied there is an improvement in the reliability of the CatBoost model for making predictions on Major Depressive Disorder. The ensemble model combining XGBoost and Random Forest showed strong performance in ROC analysis, showcasing its ability to discriminate between individuals with MDD and those without. This hybrid approach holds promise for improving the understanding and diagnosis of Major Depressive Disorder.

The research study's advantages lie in its effective MDD diagnosis using EEG data and machine learning. Advanced artefact rejection techniques ensure high-quality EEG data, enhancing prediction accuracy and real-world clinical applicability. The hybrid machine learning approach combines models for improved accuracy and comprehensive MDD classification. CatBoost's handling of categorical variables simplifies preprocessing, crucial for clinical efficiency. Multi-metric feature selection enhances sensitivity and specificity in MDD detection. K-fold cross-validation boosts model reliability for confident clinical decision-making. The ensemble model excels in discriminating MDD cases, potentially enabling earlier and more precise diagnoses for timely interventions.

V. Conclusion

In conclusion, our study has yielded significant insights into the prediction and diagnosis of Major Depressive Disorder (MDD) using hybrid machine learning methods applied to EEG data along with clinical and demographic information. The standout performer among the algorithms tested was CatBoost, achieving an impressive accuracy rate of 93.1% in MDD prediction and diagnosis. This result notably surpassed the performance of other algorithms evaluated, highlighting the superiority of CatBoost in this context. Additionally, our ensemble model combining XGBoost and Random Forest demonstrated strong performance in ROC analysis, further supporting the effectiveness of hybrid machine learning approaches. The incorporation of attribute selection metrics such as Information Gain, $\chi 2$, Gain Ratio, and Gini Decrease also played a crucial role in identifying the most informative features for MDD prediction based on EEG data. Overall, our findings underscore the potential of hybrid machine learning techniques, particularly CatBoost, in improving MDD prediction and diagnosis accuracy, thereby facilitating the development of personalized interventions and targeted treatments for individuals with depressive disorders. These results not only contribute to advancing mental health diagnostics but also hold implications for enhancing patient outcomes and quality of care in the clinical setting.

VI. Future Work

The future research could focus on exploring additional attribute selection measures and fine-tuning the hybrid models to further enhance their accuracy and generalize their findings to larger and more diverse datasets. Additionally, integrating other modalities of data, such as genetic and environmental factors, could provide us with a deeper insight into the workings of the system of Major Depressive Disorder and lead to personalized interventions and targeted treatment strategies.

References

[1] Mathers, C.D.; Loncar, D. Projections of global mortality and burden of disease from 2002 to 2030. PLoS Med. 2006, 3, e442.

[2] World Mental Health Day: An Opportunity to Kick-Start a Massive Scale-Up in Investment in Mental Health. Available online: https://www.who.int/news/item/27-08-2020-world-mental-health-day-an-opportunity-to-kick-start-a-massive-scale-up

[3] World Health Organization. Depression and Other Common Mental Disorders: Global Health Estimates. online: https://apps.who.int/iris/bitstream/handle/10665/254610/WHO-MSD-MER-2017.2-eng.pdf.

[4] Turukmane, A. V. ., Tangudu, N. ., Sreedhar, B. ., Ganesh, D. ., Reddy, P. S. S. ., & Batta, U. . (2023). An Effective Routing Algorithm for Load balancing in Unstructured Peer-to-Peer Networks. *International Journal of Intelligent Systems and Applications in Engineering*, *12*(7s), 87–97.

[5] Patel, M.J.; Khalaf, A.; Aizenstein, H.J. Studying depression using imaging and machine learning methods. NeuroImageClin. 2016, 10, 115–123.

[6] Strunk, D.R.; Pfeifer, B.J.; Ezawa, I.D. Depression. In Handbook of Cognitive Behavioral Therapy: Applications; Wenzel, A., Ed.; American Psychological Association: Washington, DC, USA, 2021; pp. 3–31.

[7] Park SM, Jeong B, Oh DY, Choi CH, Jung HY, Lee JY, Lee D, Choi JS. Identification of Major Psychiatric Disorders From Resting-State Electroencephalography Using a Machine Learning Approach. Front Psychiatry. 2021 Aug 18;12:707581. doi: 10.3389/fpsyt.2021.707581. PMID: 34483999; PMCID: PMC8416434.

[8] Garcia-Ceja, E.; Riegler, M.; Nordgreen, T.; Jakobsen, P.; Oedegaard, K.J.; Tørresen, J. Mental health monitoring with multimodal sensing and machine learning: A survey. Pervasive Mob. Comput. 2018, 51, 1–26.

[9] https://www.kaggle.com/datasets/shashwatwork/eeg-psychiatric-disorders-dataset

[10] Kumar, T. P., & Kumar, M. S. (2021). Optimised Levenshtein centroid cross-layer defence for multi-hop cognitive radio networks. *IET Communications*, *15*(2), 245-256.

[11] Cho, G.; Yim, J.; Choi, Y.; Ko, J.; Lee, S.-H. Review of Machine Learning Algorithms for Diagnosing Mental Illness. Psychiatry Investig. 2019, 16, 262–269.

[12] Miyajima, A.; Tanaka, M.; Itoh, T. Stem/progenitor cells in liver development, homeostasis, regeneration, and reprogramming. Cell Stem Cell 2014, 14, 561–574.

[13] Zulfiker, M.S.; Kabir, N.; Biswas, A.A.; Nazneen, T.; Uddin, M.S. An in-depth analysis of machine learning approaches to predict depression. Curr. Res. Behav. Sci. 2021, 2, 100044.

[14] Na, K.-S.; Cho, S.-E.; Geem, Z.W.; Kim, Y.-K. Predicting future onset of depression among community dwelling adults in the Republic of Korea using a machine learning algorithm. Neurosci. Lett. 2020, 721, 134804.

[15] Choudhury, A.A.; Khan, R.H.; Nahim, N.Z.; Tulon, S.R.; Islam, S.; Chakrabarty, A. Predicting Depression in Bangladeshi Undergraduates using Machine Learning. In Proceedings of the 2019 IEEE Region 10 Symposium (TENSYMP), Kolkata, India, 7–9 June 2019; pp. 789–794.

[16] Singh, N.; Gunjan, V.K.; Roy, S.; Rahebi, J.; Farzamnia, A.; Saad, I. Multi-Controller Model for Improving the Performance of IoT Networks. *Energies* **2022**, *15*, 8738.

[17] Priya, A.; Garg, S.; Tigga, N.P. Predicting Anxiety, Depression, and Stress in Modern Life using Machine Learning Algorithms. Procedia Comput. Sci. 2020, 167, 1258–1267.

[18] Fatima, I.; Mukhtar, H.; Ahmad, H.F.; Rajpoot, K. Analysis of user-generated content from online social communities to characterise and predict depression degree. J. Inf. Sci. 2018, 44, 683–695.

[19] K. D. Lakshmi and L. Chakradhar, "Implementation of Latest Deep Learning Techniques for Brain Tumor Identification from MRI Images," *2023 8th International Conference on Communication and Electronics Systems (ICCES)*, Coimbatore, India, 2023, pp. 1166-1171, doi: 10.1109/ICCES57224.2023.10192620.

[20] Hilbert, K.; Lueken, U.; Muehlhan, M.; Beesdo-Baum, K. Separating generalized anxiety disorder from major depression using clinical, hormonal, and structural MRI data: A multimodal machine learning study. Brain Behav. 2017, 7, e00633.

[21] Guo, H.; Qin, M.; Chen, J.; Xu, Y.; Xiang, J. Machine-Learning Classifier for Patients with Major Depressive Disorder: Multifeature Approach Based on a High-Order Minimum Spanning Tree Functional Brain Network. Comput. Math. Methods Med. 2017, 2017, 4820935.

[22] Mahdy, N.; Magdi, D.A.; Dahroug, A.; Rizka, M.A. Comparative Study: Different Techniques to Detect Depression Using Social Media. In Internet of Things-Applications and Future; Springer: Singapore, 2020; pp. 441–452.

[23] Liu, G.-D.; Li, Y.-C.; Zhang, W.; Zhang, L. A Brief Review of Artificial Intelligence Applications and Algorithms for Psychiatric Disorders. Engineering 2020, 6, 462–467.

[24] He, L.; Niu, M.; Tiwari, P.; Marttinen, P.; Su, R.; Jiang, J.; Guo, C.; Wang, H.; Ding, S.; Wang, Z.; et al. Deep learning for depression recognition with audiovisual cues: A review. Inf. Fusion 2021, 80, 56–86.

[25] Su, D.; Zhang, X.; He, K.; Chen, Y. Use of machine learning approach to predict depression in the elderly in China: A longitudinal study. J. Affect. Disord. 2021, 282, 289–298.

[26] Tao, X.; Chi, O.; Delaney, P.J.; Li, L.; Huang, J. Detecting depression using an ensemble classifier based on Quality of Life scales. Brain Inform. 2021, 8, 2.

[27] Karoly, P.; Ruehlman, L.S. Psychological “resilience” and its correlates in chronic pain: Findings from a national community sample. Pain 2006, 123, 90–97.

[28] Zhao, M.; Feng, Z. Machine Learning Methods to Evaluate the Depression Status of Chinese Recruits: A Diagnostic Study. Neuropsychiatr. Dis. Treat. 2020, 16, 2743–2752.

[29] Ma, Y.; Ji, J.; Huang, Y.; Gao, H.; Li, Z.; Dong, W.; Zhou, S.; Zhu, Y.; Dang, W.; Zhou, T.; et al. Implementing machine learning in bipolar diagnosis in China. Transl. Psychiatry 2019, 9, 305.

[30] Aldabbas, H.; Albashish, D.; Khatatneh, K.; Amin, R. An Architecture of IoT-Aware Healthcare Smart System by Leveraging Machine Learning. Int. Arab J. Inf. Technol. 2022, 19, 160–172.

[31] Tennenhouse, L.G.; Marrie, R.A.; Bernstein, C.N.; Lix, L.M. Machine-learning models for depression and anxiety in individuals with immune-mediated inflammatory disease. J. Psychosom. Res. 2020, 134, 110126.

[32] Li, X.; La, R.; Wang, Y.; Hu, B.; Zhang, X. A Deep Learning Approach for Mild Depression Recognition Based on Functional Connectivity Using Electroencephalography. Front. Neurosci. 2020, 14, 192.

[33] D. Ganesh et al., "Implementation of Novel Machine Learning Methods for Analysis and Detection of Fake Reviews in Social Media," 2023 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS), Erode, India, 2023, pp. 243-250, doi: 10.1109/ICSCDS56580.2023.10104856.

[34] Alloghani, M.A.; Al-Jumeily, D.; Mustafina, J.; Hussain, A.; Aljaaf, A.J. A systematic review on supervised and unsupervised machine learning algorithms for data science. In Supervised and Unsupervised Learning for Data Science; Berry, M., Mohamed, A., Yap, B., Eds.; Springer: Cham, Switzerland, 2020

[35] CatBoost: unbiased boosting with categorical features support. Accessed on September 2021. Available at: https://catboost.ai/

[36] David M.W. Powers. “Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation”. In: (Oct. 2020).

[37] L1 Regularization: Tibshirani, R. (1996). Regression Shrinkage and Selection via the Lasso. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 58(1), 267-288

[38] L2 Regularization: Tikhonov, A. N. (1943). On the Stability of Inverse Problems. DokladyAkademiiNauk SSSR, 39(5-6), 195-198

[39] Bugata, P.; Drotar, P. On some aspects of minimum redundancy maximum relevance feature selection. Sci. China Inf. Sci. 2020, 63, 1–15.

[40] Kursa, M.B.; Jankowski, A.; Rudnicki, W.R. Boruta—A system for feature selection. Fundam. Inform. 2010, 101, 271–285.