Journal of Electronics, Electromedical Engineering, and Medical Informatics
Homepage: jeeemi.org; Vol. 7, No. 3, July 2025, pp: 581-596 e-ISSN: 2656-8632

RESEARCH ARTICLE OPEN ACCESS

Applied Machine Learning in EEG data Classification to Classify Major Depressive Disorder by Critical Channels

Sudhir Dhekane[1,2], **Anand Khandare**[1]

[1] Thakur College of Engineering and Technology, Mumbai, India
[2] D.J. Sanghvi College of Engineering, Mumbai, India

Corresponding author: Sudhir Dhekane (e-mail: sudhir.dhekane@tcetmumbai.in, sudhir.dhekane@djsce.ac.in),
Author(s) Email: Anand Khandare (e-mail: anand.khandare@thakureducation.org)

Abstract **The electroencephalogram (EEG) stands out as a promising non-invasive tool for assessing depression. However, the efficient selection of channels is crucial for pinpointing key channels that can differentiate between different stages of depression within the vast dataset. This study outcome a comprehensive strategy for optimizing EEG channels to classify Major Depressive Disorder (MDD) using machine learning (ML) and deep learning (DL) approaches, and monitor effect of central lobe channels. A thorough review underscores the vital significance of EEG channel selection in the analysis of mental disorders. Neglecting this optimization step could result in heightened computational expenses, squandered resources, and potentially inaccurate classification results. Our assessment encompassed a range of techniques, such as Asymmetric Variance Ratio (AVR), Amplitude Asymmetry Ratio (AAR), Entropy-based selection employing Probability Mass Function (PMF), and Recursive Feature Elimination (RFE) where, RFE exhibited superior performance, particularly in pinpointing the most pertinent EEG channels while including central lobe channels like Fz, Cz, and Pz. With this accuracy between 97 to 99% is recorded by Electroencephalography Neural Network (EEGNet). Our experimental findings indicate that, models using RFE achieved enhancement in accuracy to classifying depressive disorders across diverse classifiers: EEGNet (96%), Random Forest (95%), Long Short-Term Memory (LSTM: 97.4%), 1D-CNN with 95%, and Multi-Layer Perceptron (98%) irrespective of central lobe incorporation. A pivotal contribution of this research is the development of a robust Multilayer Perceptron (MLP) model trained on EEG data from 382 participants, achieved accuracy of 98.7%, with a perfect precision score of 1.00, F1-Score of 0.983, and a Recall-Score of 0.966, to make it an enhanced technique for depression classification. Significant channels identified include Fp1, Fp2, F7, F4, F8, T3, C3, Cz, T4, T5, and P3, offering critical insights about depression. Our findings shows that, optimized EEG channel selection via RFE enhances depression classification accuracy in the field of brain-computer interface.**

Keywords: Multi-Layer Perceptron; Recursive Feature Elimination; Channels; Depression

1. Introduction

Mental health encompasses emotional, social, and psychological well-being, affecting thoughts, feelings, and behaviors. Maintaining good mental health is crucial for functioning effectively across life stages, So, making early detection of mental disorders essential for balanced mental health [1]. Non-invasive technologies use external headgear to measure brain signals without the need for surgery, leading to increased popularity [2]. The widespread application of digital processing for electroencephalography (EEG) signals spans diverse fields, including mental task and emotion classification. Given the multitude of EEG channels available, the necessity for efficient channel selection algorithms has become evident, with varying relevance across applications. The primary objectives of the channel selection process encompass: (i) Diminishing computational complexity in EEG signal processing by pinpointing pertinent channels for extracting crucial features, (ii) Curbing overfitting arising from unnecessary channels to enhance overall performance, and (iii) Trimming setup time in specific applications. The crux of insightful information concerning the functional state of the human brain resides in five distinct brain waves, characterized by different frequency bands: delta

Manuscript received February 24, 2025; Accepted April 20, 2025; date of publication May 5, 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i3.719

(0–4 Hz), theta (3.5–7.5 Hz), alpha (7.5–13 Hz), beta (13–26 Hz), and gamma (26–70 Hz). Delta waves correlate with deep sleep, theta waves with the profound reflective state (body sleeping/mind wakeful), alpha waves with dreaming and relaxation, beta waves with heightened waking attention, and gamma waves with the brain's decision-making mode. Mental illnesses introduce unexpected disruptions in brain waves, necessitating substantial signal processing efforts for abnormal state diagnosis [3]. Typically, acquired EEG signals manifest as multi-channel data. When classifying these signals, two options emerge working on a subset of channels selected based on specific criteria or working on all channels [4]. Fig. 1 provides an overview of the general

Fig. 1. General process of EEG signal classification based on Optimized Vs All EEG channel

EEG signal classification process, emphasizing channel selection. Within this signal-processing context, the imperative to reduce channel numbers arises from the time-intensive setup process associated with a large channel count, causing inconvenience to subjects. EEG channel selection algorithms are needed with varying importance from one application to another[5]. Various techniques are contributed by researchers in optimizing the channels such as normalized mutual information (NMI) to optimally select EEG channels, achieving high accuracy in emotion detection while reducing channel count [6]. The correlation coefficient method can effectively determine the best channel combination, enhancing motor imagery decoding accuracy for both healthy individuals and ALS patients [7]. The suggested sparse common spatial pattern algorithm for EEG channel selection can be tailored to achieve optimal classification accuracy by filtering out noisy and irrelevant channels [8]. A method that optimizes EEG channel selection for brain-computer interfaces by using regularized common spatial patterns (CSP) and multi-band signal decomposition, achieving high accuracy with fewer channels [9]. Another study examines EEG channel selection methods for motor imagery in brain-computer interfaces, revealing that using a subset of channels (10-30% of total) can offer comparable performance to using all channels, highlighting efficiency in signal processing and system performance [10]. StEEGCS demonstrates superior performance compared to various cutting-edge EEG channel selection methods when tested on real-world EEG datasets [11]. The study introduces an EEG channel selection method utilizing a Gumbel-softmax concrete selector layer, which optimizes both channel selection and neural network parameters simultaneously. It shows better performance across two EEG tasks compared to traditional task-specific benchmarks. Contributed research by various authors is tremendous support to propose enhanced approach for EEG channel selection and to understand its effect on classification and analysis of time for EEG data [12]. In this paper author presents a comprehensive survey of various channel selection techniques for EEG signal processing in a wide range of applications, with the key purpose of reducing computational complexity, mitigating overfitting, and reducing setup time [5]. A latest study recommended an optimal channel selection method kernel-target alignment (KTA) to improve the viability and effectiveness of EEG-based assessments for depression [13]. One more study compared various single-channel EEG measures and found that a combination of linear and nonlinear measures could achieve up to 92% classification accuracy in discriminating between depressive and control subjects, demonstrating that single-channel EEG analysis can provide depression detection at the level of multichannel EEG analysis. EEG recordings from 13 depressive patients and 13 matched controls examination of 30-channel EEG using linear and nonlinear methods. Logistic regression analysis, paired with leave-one-out cross-validation, was used to determine classification accuracy for each individual EEG channel. Solo-channel electroencephalographic study can provide insight of depression at the level of multiple electroencephalographic channel analysis [14]. One of the study aimed to find a simple method for detecting depression using single-channel EEG signals, and found that a combination of linear (SASI) and nonlinear (DFA) analysis of a single EEG channel (Pz) can provide high accuracy (91.2%) in differentiating depressive and healthy individuals [15]. The proposed method uses EEG signals and a stochastic search algorithm to recognize major depressive disorder with high accuracy, providing a possible solution for an intelligent, computer-aided diagnostic tool to aid clinicians in early MDD diagnosis [16]. Traditional methods for identifying depression often rely on subjective scales, which can lack objectivity and precision. Addressing these limitations, recent advancements in EEG-based depression recognition have shown promise in providing more accurate and objective assessments [17]. The researchers investigated the use of EEG data and machine learning methods to detect mild depression, finding that the beta frequency band and left parietotemporal lobe region were most relevant, and a combination of the GSW feature selection method and KNN classifier performed best [18]. The research proposes an adaptive channel fusion method using

Manuscript received February 24, 2025; Accepted April 20, 2025; date of publication May 5, 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i3.719

improved focal loss functions for improved depression recognition from EEG signals [19]. Another study, presents a machine learning-based framework for detecting depression using EEG signals from a publicly available dataset, achieving a high classification accuracy of 96.36% using a BF-Tree classifier and a feature vector length of 12, outperforming existing state-of-the-art approaches [20]. The study used 3-channel EEG signals and linear/nonlinear features to classify depression patients and healthy controls, achieving 72.25% accuracy and suggesting the potential for early depression diagnosis. Collected EEG signals from 3 electrodes (Fp1, Fpz, Fp2)e extracted 3 linear features (Min-Max-Center Value)) and 3 nonlinear features (correlation-dimension, Renyi-entropy, C0-complexity) from the electroencephalograms [21]. Author used single-channel EEG signals and machine learning to discriminate between major depressive disorder patients and healthy controls with 97.28% accuracy [22]. One of the study used a combination of Independent Component Analysis (ICA) and sLORETA methods to analyze the resting EEG data [23]. The major finding is that, theta & alpha action in depressed subjects at parietal & occipital positions may redirect a diminished neural activation in these regions of brain[24]. The emotional and physical symptoms significantly hinder an individual's ability to perform at work or in social settings. Depression frequently presents as chronic sorrow, a lack of interest in once-enjoyable activities of interest, and more [25]. Traditional methods for diagnosis which rely on the

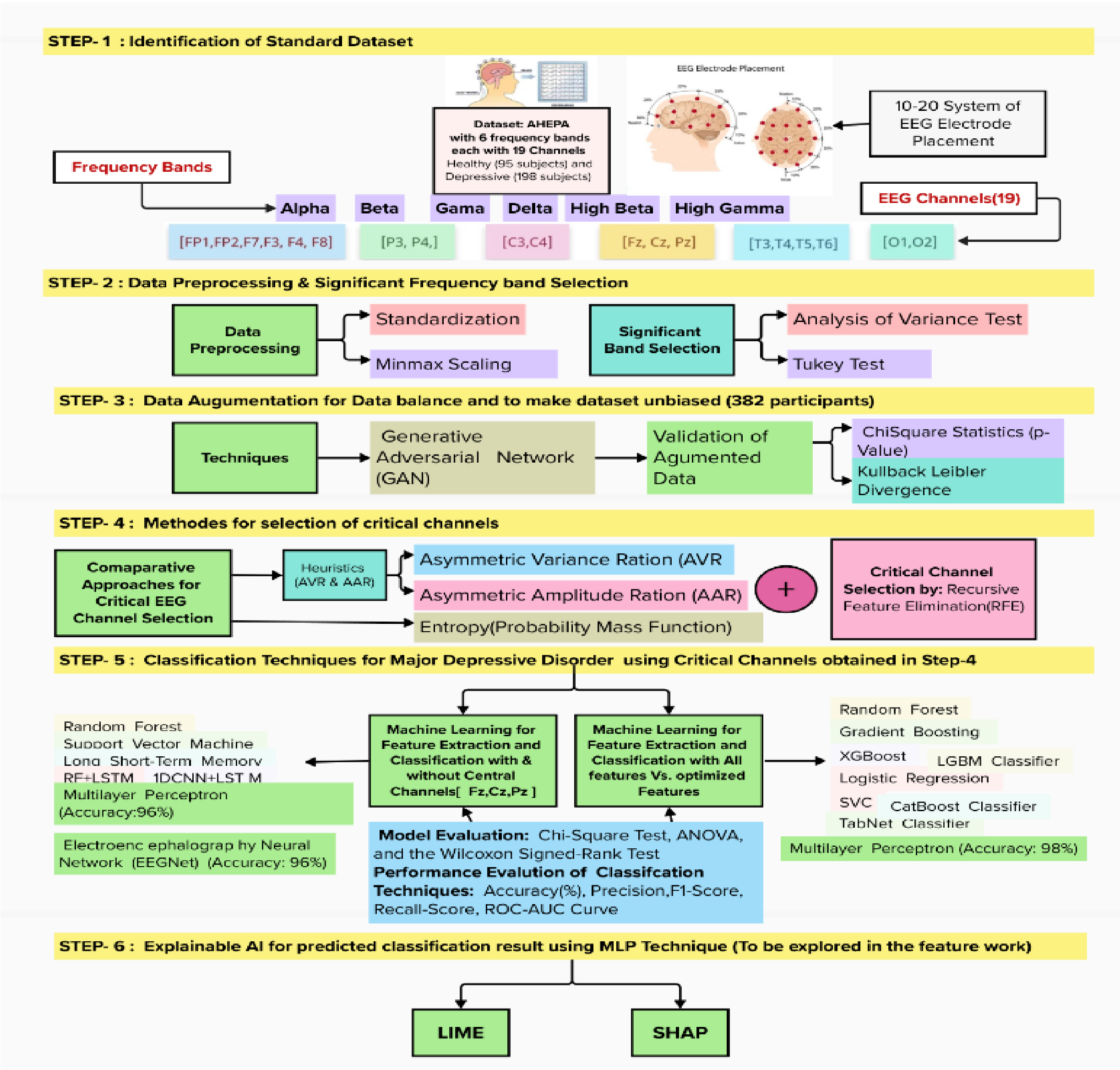


Fig. 2. Architecture of proposed methodology to classify Major Depressive Disorder (MDD)

Manuscript received February 24, 2025; Accepted April 20, 2025; date of publication May 5, 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i3.719

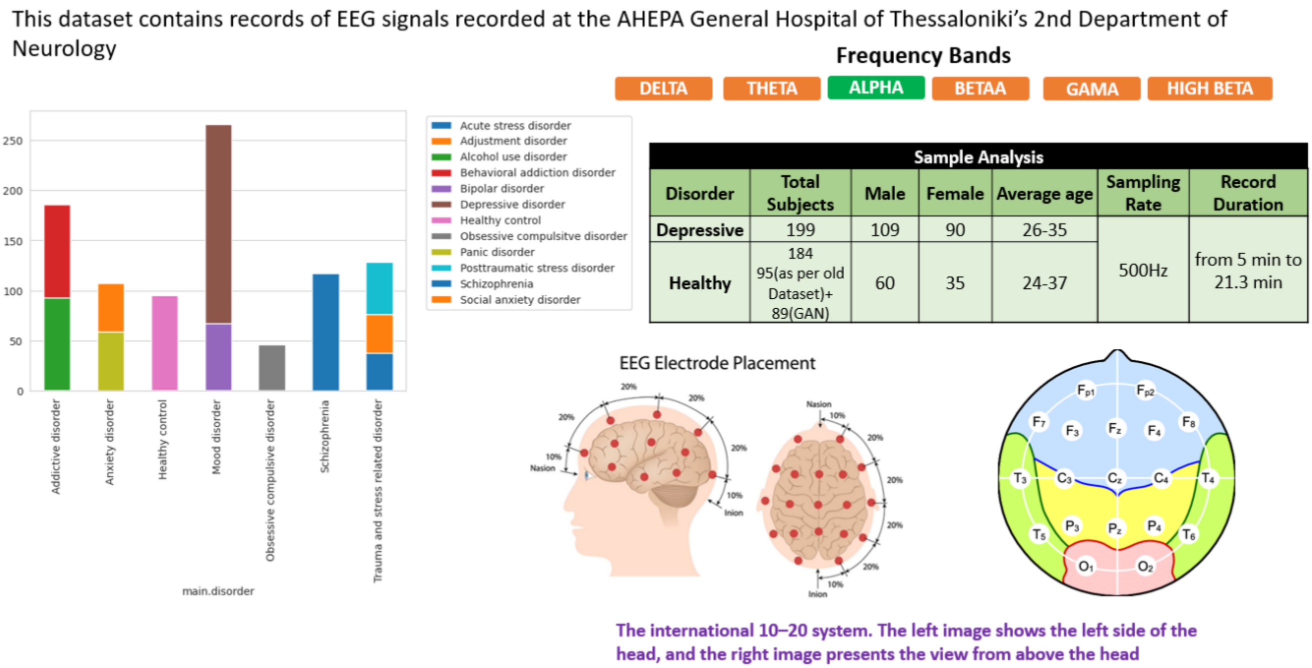


Fig. 3. **International 10-20 Electrode placement and Dataset Summary**

symptoms reported by people, can make accurate diagnosis challenging. It can result in false diagnosis or delayed treatment [26]. The timely detection of depression can greatly improve treatment plans and help prevent serious consequences like the alarmingly high risk of suicide among young adults [27].

2. Methodology

After the systematic review, it is been clearly noted that, without optimum channel selections w.r.t specific mental illness there is no sense in proceeding for classification of subjects undertaken for mental illness (MDD) analysis. We have proposed and implemented the "Hybrid technique Power Spe ctrum Analysis (AVR+AAR), Entropy Calculation using Probability mass Function (PMF) & deep learning approaches to MDD using EEG data by channel optimization using RFE.

The proposed methodology shown in Fig. 2 explore a structured approach to classify MDD based on EEG data. The process starts with collecting EEG recordings from a standardized dataset that is AHEPA dataset. To handle class imbalance, the data is subjected to augmentation using Generative Adversarial Networks (GANs) and validated with Chi-square tests and p-value analysis, ensuring a statistical validation of the augmented data. Data pre-processing techniques-standardization, minimax scaling, ANOVA, and Tukey tests-are used for normalizing the data. Heuristic measures, like Asymmetric Variance Ratio (AVR) and Amplitude Asymmetry Ratio (AAR), techniques are combined with channel selection techniques such as Recursive Feature Elimination (RFE) and entropy-based analysis for the selection of the most critical EEG channels. The features from important frequency bands, like Alpha, Beta, Gamma, Delta, High Beta, and High Gamma, are also extracted. A comparative heuristic approach is used to fine-tune channel selection by integrating AVR, AAR, and entropy measures. Further methodology focuses on, designing and training classifiers. The machine learning algorithms such as, Random Forest, Gradient Boosting, Support Vector Machines (SVM), Logistic Regression, AdaBoost, and Naïve Bayes implemented for feature extraction and classifications. In addition to that, deep learning models such as, Long Short-Term Memory (LSTM), Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), and EEGNet are implemented and evaluated. Then the models use different performance metrics such as accuracy, precision, recall, F1-score, and ROC-AUC curve that helps to ensure, the evaluation of the implemented models and reliable classification results. This integrated pipeline allows the development of an optimized approach for stages detection in MDD in the proposed feature work, through a combination of advanced data pre-processing and feature selection along with robust classification techniques.

A. Dataset

A standard EEG dataset serves as the original source of the input. The dataset precise in Fig. 3. This dataset comprises EEG signal observations made in the second department of neurology at the AHEPA

Manuscript received February 24, 2025; Accepted April 20, 2025; date of publication May 5, 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i3.719

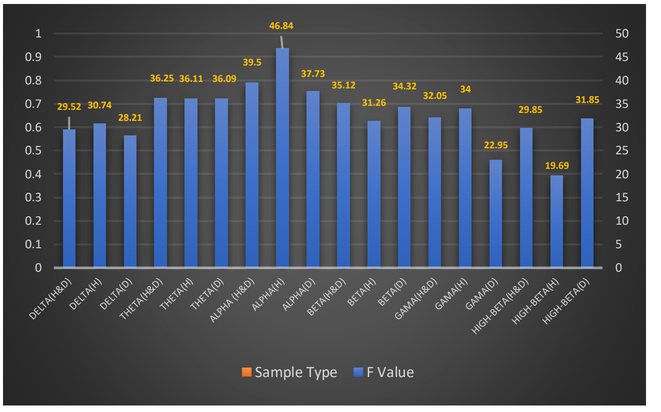
Journal of Electronics, Electromedical Engineering, and Medical Informatics
Homepage: jeeemi.org; Vol. 7, No. 3, July 2025, pp: 581-596 **e-ISSN: 2656-8632**

General Hospital in Thessaloniki [19]. This dataset consists of electroencephalograms of 293 patients divided into two groups: Healthy (95 subjects) and Depressive (198 subjects). For recording EEG, the dataset maker used a Nihon Kohen device (EEG-2100) with 19 electrodes situated on the head as per the 10–20 scheme: [Fp1-Fp2-F7-F3-Fz-F4-F8-T3-C3-Cz-C4-T4-T5-P3-Pz-P4-T6-O1-O2]. With sampling rate of 500 Hz for each signal. The period of EEG recordings ranged flanked by 5 minutes to 21 minutes. In the pre-processing, as an inactive EEG was noted with the eyes closed, the important insights like, artifact removal, normalization, standardization, and dimensionality reduction for classification are performed. Each step was carefully chosen to reduce noise, standardize features, and improve model performance. We used Z-score standardization and min-max normalization to handle EEG signal variability, ensuring better model stability. Additionally, Recursive Feature Elimination (RFE) and Principal Component Analysis (PCA) helped retain the most relevant features, reducing complexity while improving accuracy. The EEG electrode labels correspond to specific regions of the brain and follow a standardized naming convention. The letters denote different lobes: F stands for the frontal lobe, P for the parietal lobe, T for the temporal lobe, C for the central region (located between the frontal and parietal lobes), and O for the occipital lobe. The letter z is used to indicate the midline, which refers to the center of the head. Additionally, numbers are used to represent lateralization: odd numbers (such as F3 or C3) indicate electrodes placed on the left hemisphere, while even numbers (such as F4 or C4) refer to the right hemisphere. Labels with “z” (like Fz, Cz, or Pz) specify positions located directly on the midline of the scalp. Diagnosis is the target column with, ‘0’ indicating the patient is healthy and ‘1’ indicating the patient is depressed.

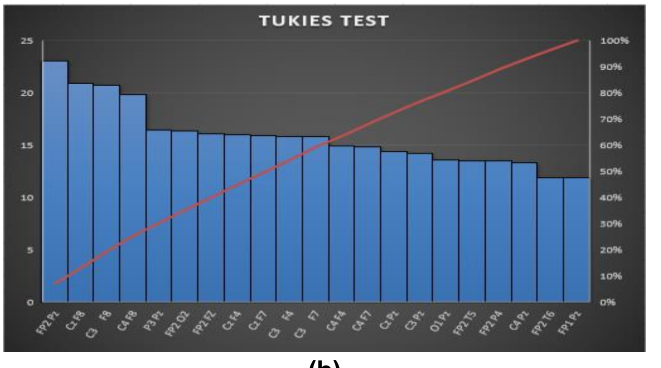
Generative Adversarial Networks (GANs) was used to increase the number of healthy patients from 95 to 184. This increased the dataset size to 382. The initial dataset was set with 19 channels records for each Frequency Band: Alpha, Beta, Delta, Gamma, High Beta and High Gamma. ANOVA and Tukey’s test was used to filter the Significant Frequency Band (Alpha) for further research.

B. Selection of Frequency Band

To select frequency band for the research ANOVA test is been carried out. further, Tukey’s test has been performed to confirm the significant band selection. Fig. 4(a), outcome the selection of Alpha band based on the F value calculated as maximum compared with all respective frequency bands (Alpha, Beta, Gama, Theta, High Beta etc.) Fig. 4(b) Tukey’s test confirms that alpha frequency band having all specified channel pairs are significant for further studies.



(a)



(b)

Fig. 4. Tests to select frequency band (a) ANOVA test, (b) Tukey’s test

4. Result

As we are working upon specific mental health disorder i.e. depression, our focus is to study part of the brain and effective activity analysis in the region of brain. Human emotional states are believed to exhibit distinct EEG signals, necessitating channel selection for efficient emotion classification and computational timesaving. A specific brain region is associated with emotions, rendering channels from unrelated areas irrelevant to emotion classification. Channel selection methodologies for emotion classification can be broadly classified into filtering and wrapper techniques. Rizon et al. [28] introduced an innovative asymmetric ratio (AR) channel selection method for recognizing human emotions from EEG signals. This method employs variance ratios between hemisphere channels as indicators for brain region assessment and emotion-associated channels. The electrical activity is accurately gauged by the spectral power ratios across hemisphere channels. The

Manuscript received February 24, 2025; Accepted April 20, 2025; date of publication May 5, 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i3.719

asymmetric variance ratio (AVR) is defined as Eq. (1) [28].

$$\mathrm{AVR} = (V(LHC) - V(RHC))/(V(LHC) + V(RHC)) \quad (1)$$

Where,V(LHC) represents the variance of the Left Hemisphere Channel, while V(RHC) denotes the variance of the Right Hemisphere Channel. The indices LHC and RHC range from 0 to N, where N is the number of electrodes that are evenly distributed across the left and right hemispheres of the brain. Additionally, the Amplitude Asymmetric Ratio (AAR) is a measure used to assess the asymmetry in brain activity between two events or regions. It is calculated using Eq. (2) [28].

$$\mathrm{AAR} = (P(i) - P(j))/(P(i) + P(j)) \quad (2)$$

Where, P(i) and P(j) are the probabilities associated with events i and j, respectively. This ratio helps in understanding the imbalance or dominance of brain activity across different hemispheres or during specific cognitive events.Here $P(i)$ is the spectral power of left hemisphere channel, $P(j)$ is the spectral power of right hemisphere channel, LHC = 0 to N, RHC = 0 to N, and N is the number of electrodes on left and right hemispheres.

A. Heuristics of AVR & AAR:

Asymmetric Variance Ratio (AVR): AVR is a metric used in EEG (electroencephalography) to quantify the asymmetry of brain activity between hemispheres. By calculating AVR for different EEG channels, researchers can identify channels that exhibit significant hemispheric differences during emotional experiences. Channels with higher AVR values indicate the stronger hemispheric asymmetry and such channels are selected for further analysis. Amplitude Asymmetric Ratio (AAR): Amplitude Asymmetry Ratio (AAR) is another metric used in EEG (electroencephalography) to quantify the asymmetry of brain activity between hemispheres. AAR for different EEG channels identifies channels where there are significant differences in amplitude between the left and right hemispheres. Channels exhibiting higher AAR values indicate stronger hemispheric asymmetry in amplitude and considered for further emotion analysis.

Above mentioned heuristics for AVR and AAR methods are need to take in consideration in case of critical channels selection in the case of depressive and healthy subjects. Here, we have proposed and implemented various technique such as Power Spectrum Analysis (AVR+AAR) ,Entropy Calculation using Probability mass Function (PMF) & Recursive Feature Elimination (RFE) for EEG Critical Channel selection, which has outcome to select and ranked EEG channels like most significant, more significant & average significant to be considered an optimal channels based on their high asymmetry ration between left and right hemispheres and the uncertainties measurement using high entropy measured channels. The pair or ranked channels in each rank may result in numerous accuracy in the classification of the subject being depressive or normal and these research further may progress in identifying major depressive stages hence, it may become one of the novel approach can be added in the field of brain computer interfaces. The Algorithm 1 shows steps to find AAR as per the following.

Algorithm 1: (Amplitude Asymmetry Ratio (AAR))

Input: LC: left_channels, RC: right_channels
Output: LC,RC, p-value
Abbreviations: H_AAR: healthy_aar D_AAR: depressive_aar, L: left, R: Right, HL: = Healthy Left Channel, HR: = Healthy Right Channel, DL: = Depressive Left Channel, DR: = Depressive Right Channel.

Procedure:
LC = ['Fp1', 'F7', 'F3', 'T3', 'C3', 'T5', 'P3', 'O1']
RC = ['Fp2', 'F8', 'F4', 'T4', 'C4', 'T6', 'P4', 'O2']
H_AAR ← []
D_AAR ← []
FOR (L,R) ← (LC, RC):

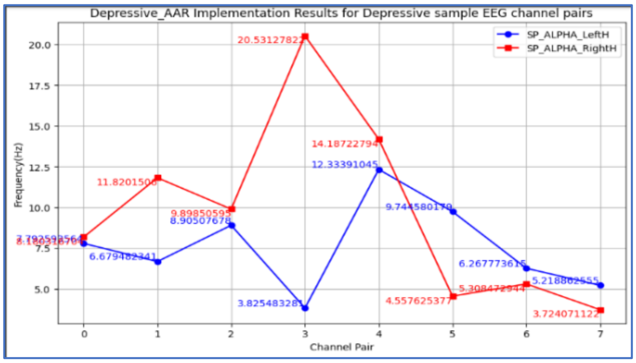
$$AAR_{healthy} = \sum_{HL,HR=1}^{n} \left(\frac{HL - HR}{HL + HR + \epsilon} \right)$$

FOR (L,R) ← (LC, RC):

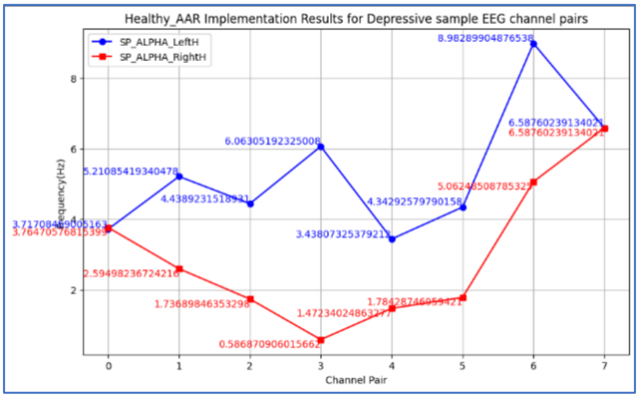
$$AAR_{depressive} = \sum_{DL,DR=1}^{n} \left(\frac{DL - DR}{DL + DR + \epsilon} \right)$$

p_values ←[]
FOR i ← LC:
T-Test ($AAR_{healthy}$, $AAR_{depressive}$)
append (p-value , p_values)
results←(LC,RC, p-value)
Sort (results)
PRINT (results).

Manuscript received February 24, 2025; Accepted April 20, 2025; date of publication May 5, 2025
Digital Object Identifier (DOI): https://doi.org/10.35882/jeeemi.v7i3.719



(a)



(b)

Fig. 5. AAR Score (a) Healthy & (b) Depressive subjects

Fig. 5(a) shows visualization of channels and their scores obtained for AAR (Depressive) and Fig. 5(b) scores AAR for healthy subjects. Table 1, showing the experimentation results of power spectrum analysis on Electrode pair considered for both the subjects normal as well as depressive. The results reveals notable differences in the selection of EEG channel pairs based on AVR and AAR scoring methods, particularly in the most significant category. When using the AVR score, the most significant pairs are located in the temporal (T5-T6) and occipital (O1-O2) regions, indicating that these areas may be more responsive or informative under this metric. In contrast, the AAR score emphasizes frontal (F7-F8) and temporal (T3-T4) channels as most significant, suggesting that it captures neural activity patterns from slightly different brain regions. In the more and average significant categories, there is some overlap but also considerable variation. For instance, channels like F3-F4 appear consistently across both scoring methods, though often at a lower level of significance. This suggests that while certain regions are generally informative (like the frontal lobe), the scoring method strongly influences which areas are prioritized. Overall, these differences highlight that AVR and AAR metrics may reflect distinct neural dynamics, which could impact how EEG data is interpreted or used in applications such as cognitive state monitoring or brain-computer interfaces.

B. Entropy Calculation using Probability mass Function (PMF)

Choosing based on entropy entails assessing the unpredictability of an EEG channel, considering the EEG signal as a random variable. The entropy for channel c is calculated by Eq. (3) [28],

$$H(c) = -\sum_{i=1}^{n} (p(x_i)\, log_2\, p(x_i\ )) \tag{3}$$

where p(xi) represents the probability mass function of the channel over n trials. The n channels with the topmost entropy are elected as input to further

Table 1. AVR and AAR Method Implementation Results

Channel Selection Methods →			AVR	AAR			
Channel Pair No	**Electrode Pair**	**Array pair in the implementation**	**AVR Score in(Hz)**	**Depressive (FFT)**	**Depressive (Spectral Power)(Hz)**	**Spectral Power in the Alpha Band (Hz) (Left Hemisphere)**	**Spectral Power in the Alpha Band (Right Hemisphere)**
1	FP1 & FP2	0,1	1.19	-0.012	-0.024	7.792593564	8.180316704
2	F7 & F8	2,6	1.38	-0.141	-0.277	6.679482341	11.8201506
3	F3 & F4	3,5	1.22	-0.026	-0.052	8.90507678	9.89850595
4	T3 & T4	7,11	0.89	-0.396	-0.685	3.825483281	20.53127822
5	C3 & C4	8,10	1.2	-0.034	-0.069	12.33391045	14.18722794
6	T5 & T6	12,16	1.96	0.187	0.362	9.744580179	4.557625377
7	P3 & P4	13,15	1.42	0.041	0.082	6.267773615	5.308472944
8	O1 & O2	17,18	1.54	0.084	0.167	5.218862555	3.724071122

Manuscript received February 24, 2025; Accepted April 20, 2025; date of publication May 5, 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i3.719

procedure of classification. Table 2 represents entropy calculated for each channel.

C. Major Depressive Disorder Classification based on optimized EEG channels:

EEG channel optimization using Recursive Feature Elimination (RFE) is a systematic approach to identifying the most relevant EEG channels for classifying depressive and normal subjects. This process reduces the dimensionality of the feature space by eliminating irrelevant or redundant channels while retaining those that, contribute most significantly to classification accuracy, thereby improving model performance and computational efficiency.

Table 2. EEG Channels Entropy score experiment results

Channel Name	Entropy of Channel	Channel Name	Entropy of Channel
FP1	2.938665479694	**C4**	2.983472691221
FP2	2.939698577005	**T4**	3.058298546160
Fz	3.044672090277	**T5**	3.010856415941
F5	2.957788026382	**P3**	2.579697725962
F2	3.045236095710	**Pz**	2.752029285544
F4	2.980541971599	**P4**	2.764383032989
F8	2.993311601390	**T6**	2.786933766242
T3	3.074479057257	**O1**	3.055173204750
C3	2.903351653257	**O2**	2.8846970207772658
Cz	2.8959230865298675		

The RFE process begins by training a machine learning classifier on the complete set of EEG channels. In this study, we employ the RandomForestClassifier, which is well suited for EEG data due to its ability to rank features based on importance. The significance of individual channel is quantified built on its influence to reducing the Gini impurity in decision trees within the Random Forest model. For a given feature j, the importance score I_j is computed as per Eq. (4) [29],

$$I_j = \frac{1}{T}\sum_{t=1}^{T} \Delta\mathrm{Gini}(\mathrm{j},\mathrm{t}) \quad (4)$$

Where ΔGini (j, t) denotes the decrease in Gini impurity caused by feature j in tree t, and T represents the total number of trees in the RandomForest ensemble. RFE iteratively eliminates the least important features, as determined by the classifier, while retraining the model on the remaining features at each step. This repetitive process remains until a pre-defined number of features is reached or till model performance reaches a plateau.

Specifically, if F= {f1, f2… fn } denotes the set of EEG channels, at each iteration k, the feature set is reduced as per Eq. (5) [29],

$$F_{k+1} = F_k / \mathrm{F}_{\min} \quad (5)$$

Where F_{min} represents the channel with the lowest importance score in the current set F_k. The model is then retrained with the reduced set of features, and its performance is evaluated. Once the optimal number of EEG channels is identified, the classifier is retrained on this reduced feature set. Performance is evaluated using standard metrics such as accuracy, precision, recall, and F1-score. By limiting the feature space to the most relevant channels, the model is expected to achieve comparable or improved performance with reduced computational complexity. The result of comparative analysis of all above implemented channel optimization methods shown in following Table 3. Channel Optimization using Recursive Feature Elimination is as per steps defined in Algorithm 2. Recursive Feature elimination methods is been tested separately on all classifiers because of different channel set obtained by this method.

Algorithm 2: EEG Channel Optimization using Recursive Feature Elimination

Input: D: Dataset
Output: S_ch , Acc
Abbreviations: X: EEG Features, Y: Labels (0: Normal, 1: Depressive) , Xtr, Xte: Training and Test Features, Ytr, Yte: Training and Test Labels, C: RandomForestClassifier, R: Recursive Feature Elimination (RFE), k: Number of Selected Channels, S_ch: Selected Channels, Acc: Accuracy, Ypred: Predicted Labels

Procedure:
1. Recursive Feature Elimination (Dataset)
   1.1 D←Datset
   1.2 X,Y←D
   1.3 Xtr, Xte, Ytr, Yte ← train_test_split(X,Y, test_size=0.3)
   1.4 C← RandomForestClassifier
2. R← RFE(C,n_features_to_select=k)R.fit(Xtr, Ytr)
3. Sch←{indices wher, R. ranking=1)}
   3.1 C.fit (Xtr [: , S_ch], Ytr)
   3.2 Ypred ←C. predict(Xte [: , S_ch])
   3.3Acc←accuracy_score(Yte, Ypred)
4. return S_ch , Acc

Manuscript received February 24, 2025; Accepted April 20, 2025; date of publication May 5, 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i3.719

Journal of Electronics, Electromedical Engineering, and Medical Informatics
Homepage: jeeemi.org; Vol. 7, No. 3, July 2025, pp: 581-596 **e-ISSN: 2656-8632**

Table 3. Optimized Channels/Channels Set to classify MDDS

Methods/Channel Significance	High Significant	Above Average Significant	Less Significant	No Significant
AVR	T5 & T6, O1 & O2	P3 & P4, F7 & F8	F3 & F4, FP1 & FP2	T3 & T4
AAR	T3,T4, F7, F8	T5-T6, C3-C4, O1-O2	F1,F2,F3, F4	P3, P4 & Pz
Entropy	T3,T4, F7, F8 & Fz	T5,T6,C3,C4 O1 & O2	F1,F2,F3, F4	P3, P4 & Pz
RFE	F7,F4,T5,P3,Pz	Not Applicable		

Table 4. Classification Accuracy to test MDDS over Normal Subject

Sr. No.	Classifier	Accuracy with All Channels (Epoch)	Accuracy with Optimized Channels Methods (AVR, AAR & Entropy) (without Central Lobe Channels [Fz,Cz,Pz])	Accuracy with Optimized Channels Methods (AVR, AAR & Entropy) (with Central Lobe Channels [Fz,Cz,Pz])	Accuracy with Optimized Channels Methods (RFE) (without Central Lobe Channels [Fz,Cz,Pz])	Accuracy with Optimized Channels Methods (RFE) (without Central Lobe Channels [Fz,Cz,Pz])
1	Random Forest	92%	86%	88	95%	95%
2	Support Vector Machine	80%	74%	76%	77%	77%
3	Long Short-Term Memory	94.81% ((200-250))	90.91%	89.61%	97.40%	93.51%
4	RF+LSTM	95% (200-250)	86%	88%	94%	87%
5	1D-CNN+LSTM	97% (50)	92%	92%	95%	95%
6	Multilayer Perceptron	98% (100)	98%	98%	98%	98%
7	Electroencephalography Neural Network (EEGNet)	97 to 99% (50)	91%	92%,	96%	96%

As per the analysis in the Table 4, the machine learning and deep learning approach such as Random Forest is concluding classification accuracy 95% with or without considering central lobe channels, LSTM having classification accuracy recorded as 97% using optimized method RFE & without consideration of central lobe channels, 1D-CNN reported classification accuracy 95% with or without consideration of central lobe channels, LSTM & Multilayer Perceptron & EEGNet are best classifiers over the other implemented various machine learning approaches concluding model accuracy is between 95% to 99% while consideration of optimized channels by RFE optimization techniques over the other optimized techniques without ambiguity in consideration of central lobe channels. Hence LSTM, MLP & EEGNet outperforming with or without considering central lobe channels. The other way is explored as follow.

D. Classification result on Entire Dataset vs. Optimized feature (EEG Channels) set w.r.t different classifiers

1. Feature Selection

Feature choice is a machine learning process that involves selecting the most related features from a dataset to use in a model. In this study, feature selection was done using Recursive Feature Elimination. Top-ranked feature subsets determined via Recursive Feature Elimination (RFE) have been employed for assessing the performance of numerous machine-learning models. Based on the model's empirical feedback, RFE, a wrapper procedure, incrementally minimizes less essential features, remaining only the most significant ones. To identify interactions between traits and the target variable, the dataset first served to train the Random Forest Classifier. Then, attributes were classified by slowly eliminating the least noteworthy ones through Recursive Feature Elimination (RFE). RFE iteratively clipped the feature set until solely the top 15 features were retained, using the Random Forest Classifier as the foundational model. The importance of these traits was classified, with lower numbers implying more significance. Considering most models supplied good accuracy, feature counts from ten to fifteen were implemented for comparing the models. The same top-

Manuscript received February 24, 2025; Accepted April 20, 2025; date of publication May 5, 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i3.719

Journal of Electronics, Electromedical Engineering, and Medical Informatics
Homepage: jeeemi.org; Vol. 7, No. 3, July 2025, pp: 581-596 e-ISSN: 2656-8632

Table 5. Summary of Hyperparameter Settings and Tuning Strategies

Classifier	Key Hyperparameters	Tuning Strategy	Optimization Method	Reproducibility Measures
Random Forest (RF)	n_estimators=100, max_depth=10, min_samples_split=2, min_samples_leaf=1, max_features=sqrt, criterion=gini	GridSearch CV	Bootstrap sampling	Fixed random seed (42)
Gradient Boosting	n_estimators=200, learning_rate=0.1, max_depth=6, min_samples_leaf=1, subsample=0.8	GridSearch CV	Deviation loss function	Scikit-learn library
XGBoost	n_estimators=200, learning_rate=0.1, max_depth=6, min_child_weight=1, gamma=0.0, subsample=0.8, colsample_bytree=0.8	Randomized Search CV	Multi-class objective function	XGBoost library
Logistic Regression	penalty=L2 (Ridge), C=1.0, solver=lbfgs, max_iter=1000	GridSearch CV	Regularization for generalization	Scikit-learn library
SVC (Support Vector Classifier)	kernel=RBF, C=1.0, gamma=scale, tol=1e-3, max_iter=1000	GridSearch CV	Margin maximization	Scikit-learn library
MLP Classifier	hidden_layer_sizes=(64), activation=ReLU, solver=Adam, learning_rate=0.001, batch_size=32, epochs=50, loss=categorical cross-entropy	Randomized Search CV	Deep learning optimization	TensorFlow/Keras
TabNet Classifier	decision_steps=3, batch_size=64, virtual_batch_size=32, momentum=0.98, gamma=1.5	GridSearch CV	Feature selection entropy	PyTorch-based TabNet
CatBoost Classifier	iterations=1000, learning_rate=0.1, depth=6, l2_leaf_reg=3.0, border_count=32, leaf_estimation_method=Newton	Randomized Search CV	Gradient boosting optimization	CatBoost library
LightGBM Classifier	num_leaves=31, learning_rate=0.05, max_depth=-1, boosting=gbdt, min_data_in_leaf=20, feature_fraction=0.8	GridSearch CV	Leaf-wise tree growth	LightGBM library

ranked features were applied uniformly across all models without recalculating rankings for each, ensuring a consistent basis for comparison.

2. Model Selection

The features in the dataset were scaled using Min-Max Scaler. The data was also standardized. Nine distinct machine and deep learning models were implemented to fit the data as per listed and the accuracies and performance achieved by these models after training on the entire dataset are summarized in Table 6. To achieve the best classification performance, we tuned hyper parameters using Grid Search Cross-Validation (GridSearchCV) and Randomized Search Cross-Validation with a 5-fold cross-validation method. Table 5 on above page is summarizing the selected models

Manuscript received February 24, 2025; Accepted April 20, 2025; date of publication May 5, 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i3.719

Table 6. Performance of Different Algorithms on Entire Dataset

Sr No.	Model Name	Accuracy (in %)	Precision	F1-Score	Recall-Score
1)	Random Forest	88.31	0.8000	0.8615	0.9333
2)	Gradient Boosting	92.21	0.8529	0.9063	0.9667
3)	XGBoost	93.51	0.9032	0.9180	0.9333
4)	Logistic Regression	92.21	0.8333	0.9091	1.0000
5)	SVC	72.73	0.5918	0.7342	0.9667
6)	MLP Classifier	98.70	1.0000	0.9831	0.9667
7)	TabNet Classifier	50.65	0.4259	0.5476	0.7667
8)	CatBoost Classifier	92.21	0.8750	0.9032	0.9333
9)	LGBM Classifier	94.81	0.9063	0.9355	0.9667

and their hyperparameters settings and tuning Strategies.

To optimize the performance of classifiers used in our study, we conducted hyperparameter tuning by exploring a predefined search range for key parameters in Support Vector Machine (SVM), Random Forest (RF), and Multi-Layer Perceptron (MLP) models. For SVM, we tuned the regularization parameter C, which controls the trade-off between maximizing the margin and minimizing classification errors, with values [0.1, 1, 10]. Additionally, we explored different Gamma settings, including 'scale' and 'auto', which influence the kernel function's impact on the decision boundary. For Random Forest, we varied the number of estimators (n_estimators) between 50, 100, and 200, as a higher number of trees can improve accuracy but increases computational cost. We also tuned the max_depth parameter, selecting values [5, 10, 20], to control model complexity and prevent overfitting. For MLP (Multi-Layer Perceptron), we adjusted the learning rate (0.001, 0.01) to balance convergence speed and model stability, and explored different numbers of hidden neurons (32, 64, 128) to optimize network capacity for EEG feature representation. By systematically tuning these parameters, we ensured that each classifier was optimized for robust and reliable MDD classification.
We used Grid Search Cross-Validation (GridSearchCV) and Randomized Search Cross-Validation to find the best hyperparameters. 5-fold cross-validation helped ensure reliable model evaluation. To improve reproducibility, we set a fixed random seed (42) and listed the software libraries used. This will make it easy for reviewers to understand the hyperparameter tuning process.

Feature Selection is applied, to identify the channels in the human brain, which can help detect depression. These classifiers underwent training on the dataset with reduced features with MLP demonstrating the highest accuracy amongst them with a commendable accuracy of 98.70%. It also displayed strong performance metrics with a Precision Score of 1.00, F1-Score of 0.983051 and Recall Score of 0.966667. Table 8 on next page summarizes the accuracies of the algorithms with different feature counts which clearly shows that, the MLP Classifier is the best model among all the algorithms, achieving a consistent accuracy of 98.70%. These findings demonstrate that, the identified channels can be effectively used for depression detection. Given the serious nature of depression, early detection through these channels can play a crucial role in timely intervention and treatment. By utilizing these methods, we can contribute to improving mental health outcomes and potentially mitigate the impact of depression through early-stage identification and support. The 11 most important channels are identified and mentioned in Table 7.

As discussed earlier in Fig. 2 which shows a flowchart that summarizes the complete process in order to provide an in-depth understanding of the steps

Table 7. Accuracies of Different Algorithms after Feature Selection

Sr. No	1	2	3	4	5	6	7	8	9	10	11
EEG Channel	Fp 1	Fp 2	F 7	F 4	F 8	T 3	C 3	C z	T 4	T5	P3

that are required. It makes the flow of the process easier to comprehend by providing a graphical representation of how it works from data collection to the final classification. The MLP model achieved high classification performance with only one misclassification, correctly identifying 47 non-depressed and 29 depressed subjects. This indicates strong model accuracy and reliability in distinguishing between the two classes as per confusion matrix. Achieving a commendable accuracy of 98.70%, the model also achieved strong performance metrics with a precision of 1.00, recall score of 0.966667 and F1-Score of 0.983051. Confusion matrix is a crucial method to check how classification models work.

The predictions are classified into four groups - True Positives (TP), False Positives (FP), True Negatives (TN) and False Negatives (FN). This division is crucial in the calculation of accuracy, precision, F1-score and recall score which give us better insights into the model's performance. The examination of the confusion matrix gives us a measure of the model's overall accuracy in addition to the specific regions where errors may arise resulting in greater emphasis on improvements in efficiency.

Manuscript received February 24, 2025; Accepted April 20, 2025; date of publication May 5, 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i3.719

Table 8. Accuracies of Different Algorithms after Feature Selection

Sr. No.	Model Name	Accuracy % (10 features)	Accuracy % (11 features)	Accuracy % (12 features)	Accuracy % (13 features)	Accuracy % (14 features)	Accuracy % (15 features)
1)	Random Forest	87.01	90.91	90.91	92.21	89.61	89.61
2)	Gradient Boosting	90.91	92.21	92.21	92.21	90.91	93.51
3)	XGBoost	92.21	96.10	94.81	94.81	93.51	93.51
4)	Logistic Regression	89.61	90.91	89.61	89.61	89.61	90.91
5)	SVC	76.62	74.03	74.03	74.03	74.03	72.73
6)	MLP Classifier	96.10	98.70	98.70	98.70	98.70	98.70
7)	TabNet Classifier	68.83	61.04	72.73	61.04	41.56	38.96
8)	CatBoost Classifier	88.31	94.81	92.21	94.81	93.51	93.51
9)	LGBM Classifier	92.21	92.21	94.81	97.40	96.10	93.51

The ROC-AUC curve (Receiver Operating Characteristic - Area Under the Curve) is important to evaluate the performance of binary classifiers. It plots the True Positive Rate (TPR) against the False Positive Rate (FPR). This visualizes the relationship between recall (TPR) and fall-out (FPR) across different thresholds. The AUC (Area Under the Curve) measures the ability of the model to classify between classes. A higher AUC score indicates better model performance. This curve is particularly useful for comparing models and assessing their effectiveness. Fig. 6 shows the ROC-AUC curve of the MLP model. The high AUC score of 0.9943 of the MLP model demonstrates its powerful classifying power. The effectiveness of the chosen channels in detecting depression is indicated by the strong performance of the model. Therefore, these channels can help in the effective and quick identification of depression. This facilitates quick action and therapy.

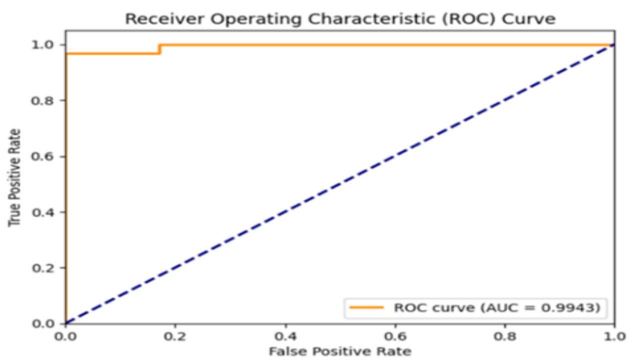


Fig. 6. ROC-AUC Curve of MLP Model

The AUC score is calculated by as per the Eq. (6).

$$\int_0^1 TPR(x)dFPR(x) \quad (6)$$

5. Discussion

To ensure our findings are reliable and meaningful, we used strong statistical validation techniques, including the Chi-Square Test, ANOVA, and the Wilcoxon Signed-Rank Test. These tests helped us to confirm that, the differences in model performance were not random but statistically significant. The Chi-Square Test checked whether classification accuracy was independent of different feature selection methods and classifiers. ANOVA with post-hoc Tukey's HSD showed that certain models and feature selection techniques had a significant impact on performance. The Wilcoxon Signed-Rank Test further confirmed that optimized feature selection methods (AVR, AAR, and RFE) led to noticeable improvements in classification accuracy. These statistical analyses provide strong proof that, combining feature selection techniques with advanced machine learning models significantly enhances EEG classification performance.

To comprehensively evaluate the performance of classifiers used in Major Depressive Disorder (MDD) classification, we employed multiple evaluation metrics, including Accuracy, Precision, Recall, F1-Score, and the ROC-AUC Curve. Accuracy (%) provides an overall measure of correctness by assessing the proportion of correctly classified instances. However, in the presence of class imbalance, accuracy alone may not provide an accurate representation of model performance. To address this limitation, Precision (Positive Predictive Value) was included to measure the proportion of correctly identified MDD cases among all predicted positive cases, ensuring that false positives are minimized. Recall (Sensitivity) was used to evaluate the model's ability to correctly detect MDD cases, minimizing false negatives, which is crucial in medical

Manuscript received February 24, 2025; Accepted April 20, 2025; date of publication May 5, 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i3.719

diagnosis to avoid undetected cases. Since both precision and recall are essential in clinical applications, we incorporated the F1-Score, which provides a balanced measure by computing their harmonic mean, making it suitable for datasets with imbalanced class distributions. Additionally, the ROC-AUC Curve (Receiver Operating Characteristic - Area under Curve) was utilized to assess the model's overall discriminatory power, providing a threshold-independent evaluation of classification performance. A higher AUC indicates better differentiation between MDD and non-MDD cases, making it a robust metric for clinical applications. By integrating these complementary evaluation metrics, we ensured a rigorous assessment of classifier performance, enabling a more reliable and clinically interpretable approach to EEG-based MDD classification. To make sure robustness of model, different validation methods, including k-fold cross-validation, leave-one-out cross-validation (LOO-CV), stratified k-fold cross-validation, and hold-out validation. K-fold cross-validation (k=5) helped tune the model while balancing bias and variance. Leave-one-out cross-validation provided an accurate measure of generalization, especially for small datasets, but required high computational power. Stratified k-fold cross-validation ensured class balance in each fold, improving reliability and Hold-out validation (80-20 split) tested how well the model performed on unseen data. By combining these methods, we reduced overfitting risks, improved reproducibility, and ensured our models were reliable for EEG classification.

Our research outcomes are compared with number of authors w.r.t to the methods, techniques, algorithms they used for classification of the depression with different channel selection approaches such as, techniques contributed by researchers in optimizing the channels such as normalized mutual information (NMI) to optimally select EEG channels, achieving high accuracy in emotion detection while reducing channel count [6] . The correlation coefficient method can effectively determine the best channel combination, enhancing accuracy [7]. Sparse common spatial pattern algorithm for EEG channel selection can be tailored to achieve optimal classification accuracy by filtering out noisy and irrelevant channels [8]. Another study examines EEG channel selection methods that using a subset of channels (10-30% of total) can offer comparable performance to using all channels, highlighting efficiency [10]. One more study compared various single-channel EEG measures and found that, a combination of linear and nonlinear measures could achieve up to 92% classification accuracy in discriminating between depressive and control subjects [14]. Another simple method for detecting depression using single-channel EEG signals, and found that a combination of linear (SASI) and nonlinear (DFA) analysis of a single EEG channel (Pz) can provide high accuracy (91.2%) in differentiating depressive and healthy individuals [15]. To detect mild depression, finding that the beta frequency band and left parietotemporal lobe region were most relevant, and a combination of the GSW feature selection method and KNN classifier performed best [18]. Machine learning-based framework for detecting depression using EEG signals from a publicly available dataset, achieving a high classification accuracy of 96.36% using a BF-Tree classifier and a feature vector length of 12, outperforming existing state-of-the-art approaches [20]. The study used 3-channel EEG signals and linear/nonlinear features to classify depression patients and healthy controls, achieving 72.25% accuracy and suggesting the potential for early depression diagnosis. Collected EEG signals from 3 electrodes (Fp1, Fpz, Fp2)e extracted 3 linear features (Min-Max-Center Value)) and 3 nonlinear features (correlation-dimension, Renyi-entropy, C0-complexity) from the electroencephalograms [21]. So, for detailed comparative study the introduction part provide the background of research on classification study of the depression. Our study has some limitations such as, Limited sample size can impact the generalizability of the study's findings, The process of selecting participants for the study may introduce selection bias, BCI technology is constantly evolving, and the validity and reliability of specific BCI measures for depression detection may still be an ongoing area of research, Interpreting the complex patterns of brain activity obtained through BCI can be challenging. So, despite these limitation, our research provides important insights into the viability of employing EEG-based Brain-Computer Interface (BCI) technologies in detecting depression. The research indicates that, even with a small sample size, machine learning models such as MLPs are capable of differentiating depressive patterns of brain activity. The research emphasizes the importance of using larger and more heterogeneous datasets to enhance generalizability and minimize selection bias. Moreover, it emphasizes the need to establish standardized procedures and interpretable models to advance the dependability of BCI measures in mental health assessment. As BCI technologies advance, our study sets the stage for future growth in non-invasive, neurotechnology-based mental health testing.

6. Conclusion

After the systematic review, it is been clearly observed that, channel selection is significant for the mental disorder analysis. It may affect to the unnecessary computational time, resources as well as false prediction or classification of undertaken mental health disorder. So, from the experimental results (AVR, AAR, Entropy using Probability mass Function (PMF) & RFE), RFE demonstrated and evaluated over the other

Manuscript received February 24, 2025; Accepted April 20, 2025; date of publication May 5, 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i3.719

optimized techniques without ambiguity in consideration of central lobe channels [Fz,Cz,Pz]) which significantly conclude in enhanced performance of mental depressive disorder classification by undertaking optimized EEG channels. The effectiveness of theses elected channels can be checked with different machine learning and deep learning models such as, Random Forest is concluding classification accuracy 95% with or without considering central lobe channels, LSTM having classification accuracy recorded as 97% using optimized method RFE & without consideration of central lobe channels, 1D-CNN reported classification accuracy 95% with or without consideration of central lobe channels, LSTM & Multilayer Perceptron & EEGNet are outperformed classifiers over the other implemented various machine learning approaches concluding model accuracy is between 95% to 99% with channel optimization technique RFE over the other optimized techniques without ambiguity in consideration of central lobe channels.

The MLP model proposed by this study has been trained on EEG signal data from 382 patients to diagnose their depression. Through careful data analysis and training, the following EEG channels were identified as primary features for depression detection: [Fp1, Fp2, F7, F4, F8, T3, C3, Cz, T4, T5, and P3]. The high accuracy of the model of 98.70% displays its ability in differentiating between depressed and non-depressed subjects. The ideal precision score of the model of 1.00 along with Recall-Score of 0.966 and F1-Score of 0.983 demonstrate its powerful and reliable performance. This helps us to conclude that, the proposed model may be used for the early detection of depression demonstrating significant potential for clinical applications. This research contributes in the field of brain computer interface to treat depression classification of the subject. To discuss about future work, though even we worked with small and dataset and enhanced accuracy by using stratified k-fold cross-validation to maintain class balance and by data augmentation, using GAN as per required class data to make it nearly balanced. In the further study, we have target, to scale dataset up to the mark by ensuring diversity in demographics information present in the current dataset. This we must have to explore with the objective of identification of MDD stages (Mild, Moderate, Server). In addition to that, to find efficiency in computational time of training period over entire dataset vs. training period with critical channels which will be ensuring the computational optimization in classification. Finally, explainable AI models such as LIME, SHAP, etc. will support our classification of MDD stages.

Author contributions

All authors contributed equally to the conception and design of the study. All authors did preparation of materials, data acquisition and analysis. All authors of this manuscript wrote and then revised on previous versions of the manuscript. All the authors read and approved the final manuscript.

Conflicts of interest

The authors declare no conflicts of interest.

Acknowledgment

The authors would like to express sincere gratitude to the Department of Computer Engineering, Thakur College of Engineering and Technology, for the invaluable support and resources provided throughout this research. The facilities, academic environment, and encouragement from faculty members have significantly contributed to the completion of this work. This study would not have been possible without the institution's commitment to advancing research and innovation in the field of brain computer interface.

References

[1] M. Ahmad, N. Wahid, R. A. Hamid, S. Sadiq, A. Mehmood, and G. S. Choi, “Decision Level Fusion Using Hybrid Classifier for Mental Disease Classification,” Computers, Materials and Continua, vol. 72, no. 3, pp. 5041–5058, 2022, doi: 10.32604/cmc.2022.026077.

[2] M. F. Mridha, S. C. Das, M. M. Kabir, A. A. Lima, M. R. Islam, and Y. Watanobe, “Brain-computer interface: advancement and challenges,” Sep. 01, 2021, MDPI. doi: 10.3390/s21175746.

[3] B. J. Baars and N. M. Gage, Cognition, Brain, and Consciousness. Elsevier, 2010. doi: 10.1016/C2009-0-01556-6.

[4] D. Garrett, D. A. Peterson, C. W. Anderson, and M. H. Thaut, “Comparison of linear, nonlinear, and feature selection methods for EEG signal classification,” IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 11, no. 2, pp. 141–144, Jun. 2003, doi: 10.1109/TNSRE.2003.814441.

[5] T. Alotaiby, F. E. A. El-Samie, S. A. Alshebeili, and I. Ahmad, “A review of channel selection algorithms for EEG signal processing,” EURASIP J Adv Signal Process, vol. 2015, no. 1, Dec. 2015, doi: 10.1186/s13634-015-0251-9.

[6] Z. M. Wang, S. Y. Hu, and H. Song, “Channel Selection Method for EEG Emotion Recognition Using Normalized Mutual Information,” IEEE Access, vol. 7, pp. 143303–143311, 2019, doi: 10.1109/ACCESS.2019.2944273.

[7] Tao Yang, Kai Keng Ang, and Kok Soon Phua, 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society : Learning from the Past, Looking to the Future :

Manuscript received February 24, 2025; Accepted April 20, 2025; date of publication May 5, 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i3.719

July 17-21, 2018, Hawaii Convention Center, Honolulu, Hawaii. IEEE, 2018.

[8] M. Arvaneh, C. Guan, K. K. Ang, and C. Quek, "Optimizing the channel selection and classification accuracy in EEG-based BCI," IEEE Trans Biomed Eng, vol. 58, no. 6, pp. 1865–1873, Jun. 2011, doi: 10.1109/TBME.2011.2131142.

[9] M. Arvaneh, C. Guan, K. K. Ang, and C. Quek, "OPTIMIZING EEG CHANNEL SELECTION BY REGULARIZED SPATIAL FILTERING AND MULTI BAND SIGNAL DECOMPOSITION."

[10] Abdullah, I. Faye, and M. R. Islam, "EEG Channel Selection Techniques in Motor Imagery Applications: A Review and New Perspectives," Dec. 01, 2022, MDPI.

[11] C. Dai, D. Pi, and S. I. Becker, "Shapelet-Transformed Multi-channel EEG Channel Selection," ACM Trans Intell Syst Technol, vol. 11, no. 5, Sep. 2020, doi: 10.1145/3397850.

[12] T. Strypsteen and A. Bertrand, "End-to-end learnable EEG channel selection for deep neural networks with Gumbel-softmax," J Neural Eng, vol. 18, no. 4, p. 0460a9, Aug. 2021, doi: 10.1088/1741-2552/ac115d.

[13] J. Shen et al., "An Optimal Channel Selection for EEG-Based Depression Detection via Kernel-Target Alignment," IEEE J Biomed Health Inform, vol. 25, no. 7, pp. 2545–2556, Jul. 2021, doi: 10.1109/JBHI.2020.3045718.

[14] M. Bachmann et al., "Methods for classifying depression in single channel EEG using linear and nonlinear signal analysis," Comput Methods Programs Biomed, vol. 155, pp. 11–17, Mar. 2018, doi: 10.1016/j.cmpb.2017.11.023.

[15] M. Bachmann, J. Lass, and H. Hinrikus, "Single channel EEG analysis for detection of depression," Biomed Signal Process Control, vol. 31, pp. 391–397, Jan. 2017, doi: 10.1016/j.bspc.2016.09.010.

[16] H. Chang et al., "EEG-based major depressive disorder recognition by selecting discriminative features via stochastic search," J Neural Eng, vol. 20, no. 2, p. 026021, Apr. 2023, doi: 10.1088/1741-2552/acbe20.

[17] Z. Wang, C. Hu, W. Liu, X. Zhou, and X. Zhao, "EEG-based high-performance depression state recognition," Front Neurosci, vol. 17, 2023, doi: 10.3389/fnins.2023.1301214.

[18] X. Li, B. Hu, S. Sun, and H. Cai, "EEG-based mild depressive detection using feature selection methods and classifiers," Comput Methods Programs Biomed, vol. 136, pp. 151–161, Nov. 2016, doi: 10.1016/j.cmpb.2016.08.010.

[19] J. Shen et al., "Depression Recognition From EEG Signals Using an Adaptive Channel Fusion Method via Improved Focal Loss," IEEE J Biomed Health Inform, vol. 27, no. 7, pp. 3234–3245, Jul. 2023, doi: 10.1109/JBHI.2023.3265805.

[20] S. Khan et al., "A machine learning based depression screening framework using temporal domain features of the electroencephalography signals," PLoS One, vol. 19, no. 3 MARCH, Mar. 2024, doi: 10.1371/journal.pone.0299127.

[21] C. Te Wu et al., "Resting-state EEG signal for major depressive disorder detection: A systematic validation on a large and diverse dataset," Biosensors (Basel), vol. 11, no. 12, Dec. 2021, doi: 10.3390/bios11120499.

[22] B. Seifeddine and N. Dib, "A Single-Channel EEG-Based Discrimination of Major Depressive Disorder with Wavelets Features and Machine Learning," in 6th International Conference on Networking and Advanced Systems, ICNAS 2023, Institute of Electrical and Electronics Engineers Inc., 2023. doi: 10.1109/ICNAS59892.2023.10330501.

[23] V. A. Grin-Yatsenko, I. Baas, V. A. Ponomarev, and J. D. Kropotov, "Independent component approach to the analysis of EEG recordings at early stages of depressive disorders," Clinical Neurophysiology, vol. 121, no. 3, pp. 281–289, Mar. 2010, doi: 10.1016/j.clinph.2009.11.015.

[24] A. Velichko, M. Belyaev, Y. Izotov, M. Murugappan, and H. Heidari, "Neural Network Entropy (NNetEn): Entropy-Based EEG Signal and Chaotic Time Series Classification, Python Package for NNetEn Calculation," Algorithms, vol. 16, no. 5, May 2023, doi: 10.3390/a16050255.

[25] Z. Li, M. Ruan, J. Chen, and Y. Fang, "Major Depressive Disorder: Advances in Neuroscience Research and Translational Applications," Jun. 01, 2021, Springer. doi: 10.1007/s12264-021-00638-3.

[26] K. Kroenke, R. L. Spitzer, and J. B. W. Williams, "The PHQ-9 Validity of a Brief Depression Severity Measure."

[27] P. Riera-Serra et al., "Clinical predictors of suicidal ideation, suicide attempts and suicide death in depressive disorder: a systematic review and meta-analysis," Eur Arch Psychiatry Clin Neurosci, Oct. 2023, doi: 10.1007/s00406-023-01716-5.

[28] M. Murugappan, S. Yaacob, M. Rizon, and S. Yaacob, "Asymmetric ratio and FCM based salient channel selection for human emotion detection using EEG," 2008. [Online]. Available: https://www.researchgate.net/publication/228574665

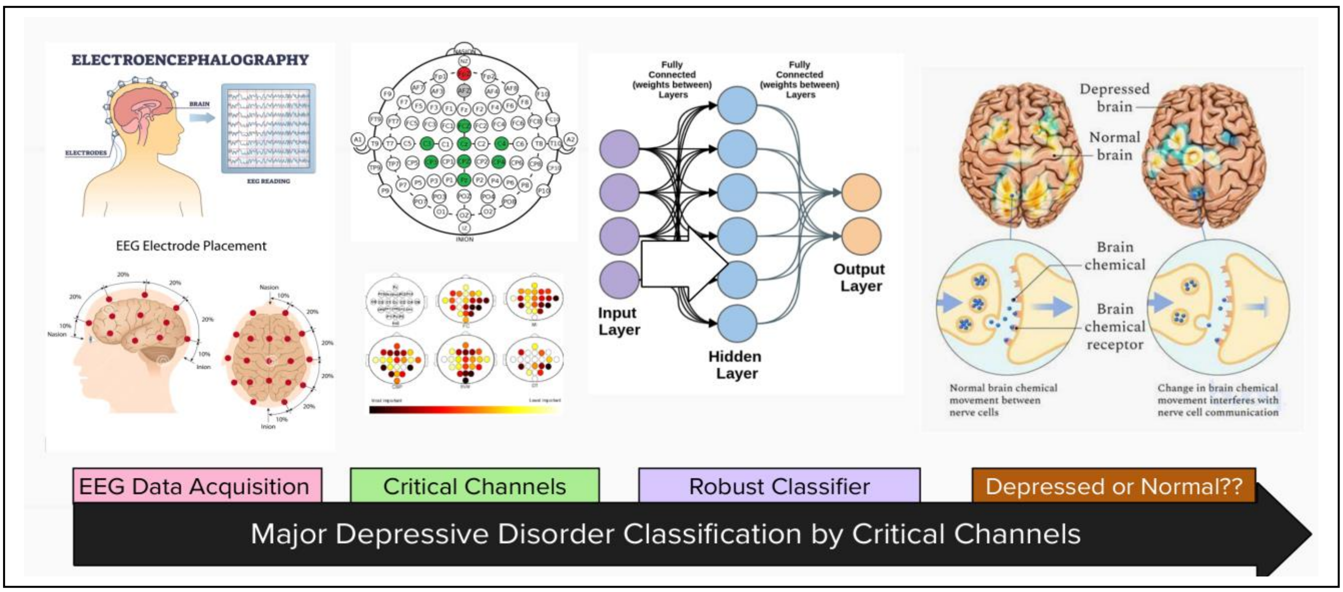
[29] Guyon, I., Weston, J., Barnhill, S. et al. Gene Selection for Cancer Classification using Support Vector Machines. Machine Learning 46, 389–422 (2002).

Manuscript received February 24, 2025; Accepted April 20, 2025; date of publication May 5, 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i3.719

AUTHOR BIOGRAPHY

SUDHIR DHEKANE received B.E in Computer Science and Engineering from Shivaji University Kolhapur Maharashtra, M.E. degree in Computer Engineering from University of Mumbai in India. He is currently working as an Assistant Professor in Artificial Intelligence (AI) & Data Science at Dwarkadas J. Snghvi College of Engineering affiliated to the University of Mumbai. His current research interest includes Machine Learning and Deep Learning, Artificial Intelligence, algorithms, Programming, Web Development etc. and Brain Computer Interface. He has published number of papers in the international journal and Conferences as a part of professional development. Further he has reviewed journals and articles for various conference.

DR. ANAND KHANDARE has received M.E. degree in Computer Engineering and Ph.D in CSE. He is currently working as Professor & Associate Dean (Planning & Operations-Digital Resources) at Thakur College of Engineering & Technology, Mumbai. His current research interest includes Machine Learning and Deep Learning, Database management, natural language Processing and Data Warehouse and Brain Computer Interface. He has received Got Silver category award for Infosys Campus Connect Program. He has been felicitated by Bhaktivedanta Hospital for successfully launching Application. He has earned funding for many projects and successfully completed the same. He has published number of papers in the international journal and Conferences as a part of professional development. Further he has reviewed journals and articles for various conference.



Manuscript received February 24, 2025; Accepted April 20, 2025; date of publication May 5, 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i3.719