Neuroscience Letters 809 (2023) 137313

Contents lists available at ScienceDirect

Neuroscience Letters

journal homepage: www.elsevier.com/locate/neulet

Machine Learning and Electroencephalogram Signal based Diagnosis of Dipression

Adil O. Khadidos [a], Khaled H. Alyoubi [a,c], Shalini Mahato [b], Alaa O. Khadidos [d], Sachi Nandan Mohanty [e,*]

[a] Department of Information Technology, Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah, Saudi Arabia
[b] Department of Computer Science and Engineering, Indian Institute of Information Technology (IIIT), Ranchi, Jharkhand, India
[c] Department of Computer Science, College of Computers and Information Systems, Umm Al-Qura University, Makkah, Saudi Arabia
[d] Department of Information Systems, Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah, Saudi Arabia
[e] Department of Computer Science & Engineering, Vardhaman College of Engineering(Autonomous), Hyderabad, India

ARTICLE INFO

Keywords:
Depression
Band power
Detrended Fluctuation Analysis (DFA)
Temporal region

ABSTRACT

Depression is a psychological condition which hampers day to day activity (Thinking, Feeling or Action). The early detection of this illness will help to save many lives because it is now recognized as a global problem which could even lead to suicide. Electroencephalogram (EEG) signals can be used to diagnose depression using machine learning techniques. The dataset studied is public dataset which consists of 30 healthy people and 34 depression patients. The methods used for detection of depression are Decision Tree, Random Forest, Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), Bidirectional Long-Short Term Memory (Bi-LSTM), Gradient Boosting, Extreme Gradient Boosting (XGBoost) along with band power. Among Deep Learning techniques, CNN model got the highest accuracy with 98.13%, specificity of 99%, and sensitivity of 97% using band power features.

1. Introduction

In terms of the number of years people live with a disability, depression or Major Depressive Disorder (MDD) has grown significantly as global cause of disability. In past ten years, the percentage of persons affected by MDD has increased to 20%, with a male to female ratio of 1:1.5. The World Health Organization estimates that by 2030, MDD will overtake other diseases as the leading cause of disease burden, affecting more than 300 million people worldwide [1,2] Fig. 1; Block representation of the detection of depressed individual and healthy individuals.

A person with depression displays a lack of interest in routine pleasure and activities. This is accompanied by additional depressive symptoms, including restlessness and agitation, exhaustion and lack of energy, guilt or worthlessness, suicidal thoughts, etc Fig. 2.

Depression impairs cognitive ability, which is expected to be reflected in the bioelectrical activity in the brain. Under these conditions, EEG brain screening can assist in comprehending how the brain operates as well as any disparities in brain activity. EEG is affordable and widely accessible. Since the EEG uses the same millisecond time scale as neural activity, it provides better temporal resolution. The EEG signals are extremely intricate, non-stationary, and non-linear, making manual interpretation very challenging. Therefore, computer-aided signal processing is required to automatically distinguish between depression patients and healthy people Fig. 3..

A lot of research has been done to increase the depression prediction as many other mental health disease using various EEG features and classifiers [31–33]. To detect stress, numerous researchers have been using machine learning and deep learning models [3]. There have been analyses in which the features extracted inherently by deep learning conform to the suitable neurocognitive background information [4]. Perceptual domain was linked to decreased alpha activity and coherence in the posterior areas in the neurological disorders [5]. Three common

Abbreviations: EEG, Electroencephalogram; CNN, Convolutional Neural Network; RNN, Recurrent Neural Network; LSTM, Long short Term Memory; DFA, Detrended Fluctuation Analysis; GRU, Gated Recurrent Unit; Bi-LSTM, Bidirectional Long-Short Term Memory; XGBoost, Extreme Gradient Boosting; MDD, Major Depressive Disorder; ICA, Independent Component Analysis.

* Corresponding author.
E-mail addresses: akhadidos@kau.edu.sa (A.O. Khadidos), kalyoubi@kau.edu.sa (K.H. Alyoubi), shalinimahato@iiitranchi.ac.in (S. Mahato), aokhadidos@kau.edu.sa (A.O. Khadidos), sachinandan09@gmail.com (S. Nandan Mohanty).

https://doi.org/10.1016/j.neulet.2023.137313
Received 21 March 2023; Received in revised form 12 May 2023; Accepted 22 May 2023
Available online 29 May 2023
0304-3940/

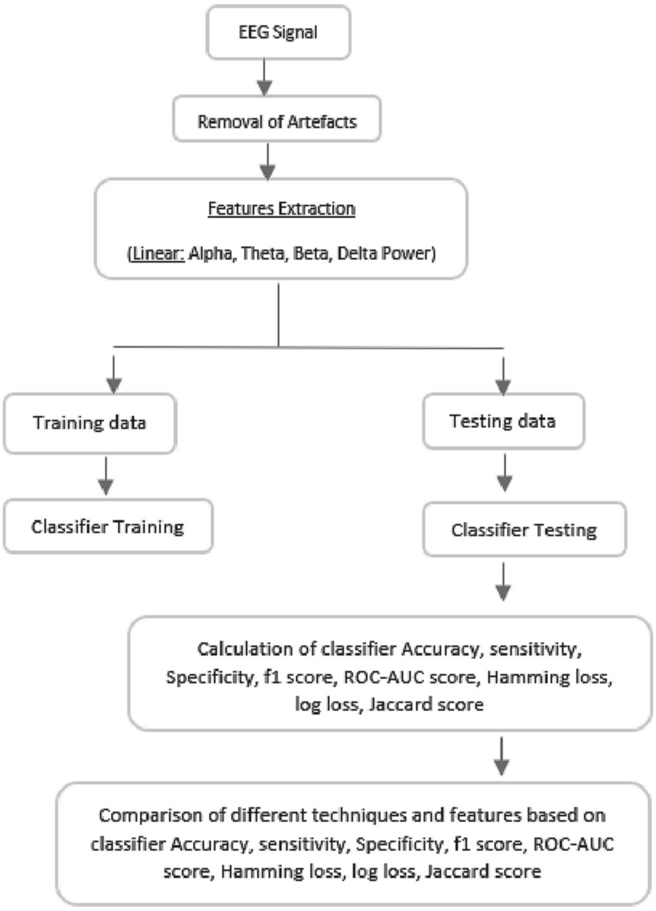

Fig. 1. Block representation of the detection of depressed individual and healthy individuals

bands—theta (4–8 Hz), alpha (8–13 Hz), and beta (13–30 Hz)—have always found to have significant information between depressed individuals and healthy controls [6]. Methods for grouping extracted features are suggested, which aids in shrinking the space while maintaining feature quality [7]. EEG Signals features are extracted and correlation matrix along with data benefit assessment and recursive feature reduction process is used for tuning them to fed into XGBoost [8]. Depressed individuals have elevated resting state EEG functional connectivity when compared to controls, with statistically significant differences observed in male subjects but not in female subjects [9] Fig. 4.

In children and adolescents, RF outscored all other ML algorithms with precision of 99% and accuracy of 95% in predicting depressed classes [10]. The mean accuracy achieved when using raw EEG-data, conducting frequency domain pre-processing to split the data into its distinct frequency domains, and creating EEG data was 88.14% [11]. Using the Logistic Regression Model, the linear mechanisms spectral asymmetry index, alpha power variability, and relative gamma power yielded an accuracy of 81% [12]. CNN and GRU can achieve an accuracy of 89.63% and 88.56% respectively on a public EEG depression dataset [13]. CNN and RNN achieved an average accuracy of 74.12% as they were able to retrieve task related features, mine inter channel correlations and incorporate that relevant information from frames [14] Fig. 5..

2. Material and Methods

Publicly available EEG signal sample data was utilized [15]. In accordance to the experiment structure granted by human ethics council of Hospital University Sains Malaysia (HUSM), Kelantan, Malaysia, a sample of 34 MDD patients were chosen for the study. It consists of 17 females and 17 males with a mean age of 40.3 ± 12.9 years. The permission forms for research participants have been signed once they had received information regarding the experiment's design. The MDD patients satisfied the Diagnostic and Statistical Manual-IV, an internationally accepted set of criteria for diagnosing depression (DSM-IV) [16]. A second set of 30 healthy age-matched controls was gathered. They served as the control group. It consists of 21 men and 9 females with the mean age of 38.3 ± 15.6 years. The people who were in good health were checked for mental issues. Fig. 1 shown the overall block diagram of study for diagnosis of depression.

2.1. Data acquisition

To collect EEG data, a cap with 19 electro-gel sensors was employed. Longer recordings and better patient care were possible because the electro-gel sensors' reduced need for corrections compared to hydro-sensors. The international 10–20 system was used in this work to put the EEG sensors on the scalp [17]. Nineteen electrodes were placed in scalp using the international 10-20 system as shown inFig. 2. The EEG data was transformed to digital form at 256Hz with band pass filtration from 0.1 to 70Hz. The eliminate power line noise, a 50 Hz notch filter was applied. Artifacts like muscle movement, eyeblinks were removed applying Independent Component Analysis (ICA).

2.2. Feature: Band power

Band power were extracted from all 19 channels: delta (0.5–4 Hz), theta (4–8 Hz), alpha (8–13 Hz), and beta (13–30 Hz) [18]. The Welch periodogram was used to calculate the EEG signal's power spectrum. The signals were divided into small segments in this method, and each segment had a 50% overlap. The signal power of each band was determined by taking the average of all modified periodograms. Alpha waves travel at a slower and are larger than beta waves. They are related with a calm state of mind and suggest that the brain is preparing to react if necessary. On eyes closing or pleasent imagination, alpha brainwaves increase. Whenever we close our eyes or imagine something pleasant, our alpha brainwaves increase. Theta brainwaves are related with mental inefficiency because they depict a day dreamy, spacey state of mind. Theta brain wave activity is a relatively calm condition at very low levels, signifying the twilight zone between awake and sleeping. Beta brainwaves are quick, smaller brainwaves linked to cerebral and intellectual activity, as well as outwardly focused concentration. This is essentially a condition of vigilance. The slowest and biggest amplitude brain waves we experience when sleeping are delta brainwaves. In general, dominant brainwave states are associated to different degrees of awareness.

2.3. Classifier

Eight different classifiers have been used in this study as shown in Fig. 3 and a comparative analysis has been done among all of them.

2.3.1. Random Forest

Random Forest is a flexible, supervised, and straightforward machine learning method that gives excellent results in most situations even without hyper-parameter tuning [19]. Because of its simplicity, it is one of the most often used algorithms, and it may be used to tackle regression and classification issues. It averages the results of numerous decision trees applied to different subsets of the dataset. Python's GridSearchCV mechanism was used to fine-tune the optimal random forest parameters. After using the model on all 4 bandpower obtained from 19 channels, accuracy of 96.70% was achieved.

2.3.2. Convolutional Neural Network

The CNN used in this study consists of the three basic layers namely, convolutional layer, a pooling layer, and a fully-connected layer as

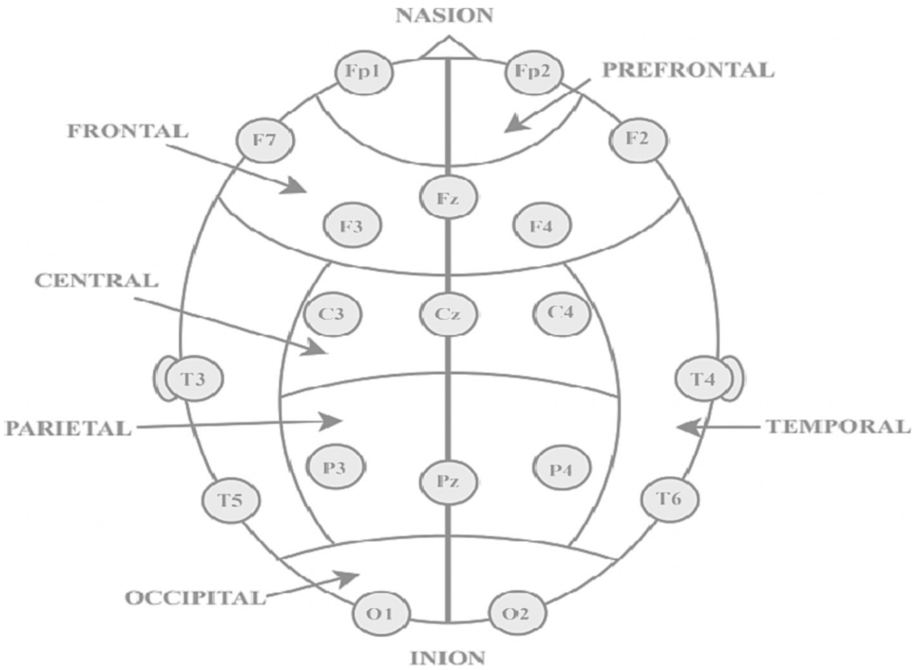

Fig. 2. Fig. 2: Standard 10–20 electrode system [34]

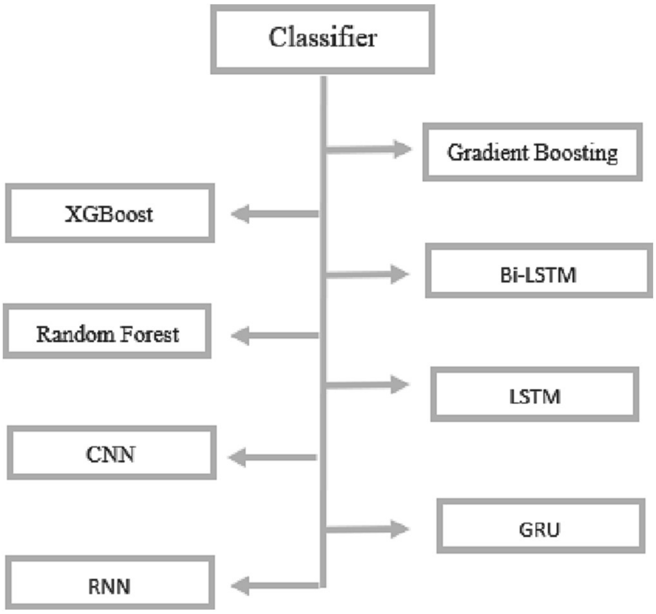

Fig. 3. Eight different classifiers used in the study

mentioned in Fig. 4. Convolutional layer utilises the filter to extract essential features from the incoming data for processing [20]. Filter is a weighted vector that is changed during training and is used to convolve the input. The parameters are determined through trial and error while maximizing the model's accuracy [21]. The convolutional layer's goal is to extract significant properties from the incoming EEG signals so that the algorithm may be trained. The rectified linear unit is the activation function after the convolution technique. In this study, a max-pooling technique is used. The outputs of the last fully connected layer were sent into a final layer with sigmoid function for determining the class likelihood as normal or depression. An accuracy of 98.13% after adopting CNN on all characteristics, outperforming all other model.

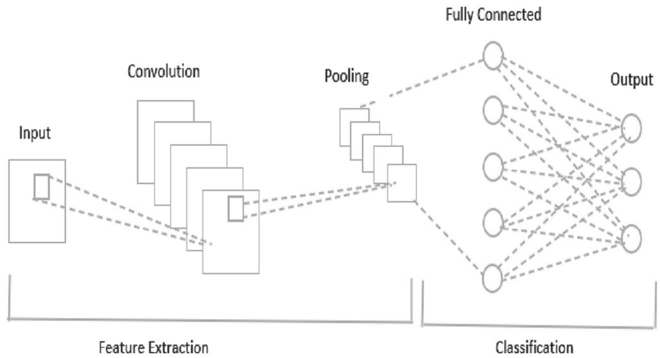

Fig. 4. Basic CNN Architecture

2.3.3. *Long Short-Term Memory*

Because of the gating mechanism used in LSTM cells, the LSTM model performs better with time series data because it only preserves meaningful information and discards irrelevant information [22].

Memory cells are unique hidden units in the LSTM design that are utilized to recall the past data for a longer period. An LSTM design includes forget, learn, remember, and utilise gates, which evaluate if an information is valuable enough to keep. The LSTM unit employs four distinct functions, including sigmoid, hyperbolic tangent, multiplication, and sum, to facilitate updating the weight matrix during back propagation process [23]. The LSTM model implemented in this study has three LSTM layers that give an accuracy of 96.7% on 4 bandpower obtained from 19 channels.

2.3.4. *Bidirectional Long-Short Term Memory*

Bi-LSTMs demonstrates to be particularly beneficial when context of the information is needed. For tasks like sentiment classification, it is very helpful. Information moves from the rear to the front in a unidirectional LSTM [24]. Bi-LSTMs are better at comprehending context. The bidirectional-LSTM network, which is composed of two distinct hidden blocks, not only relays information from front to rear, but also from back to front, and the output value is determined jointly. As a result, when working with EEG data with pre-and-post relationships,

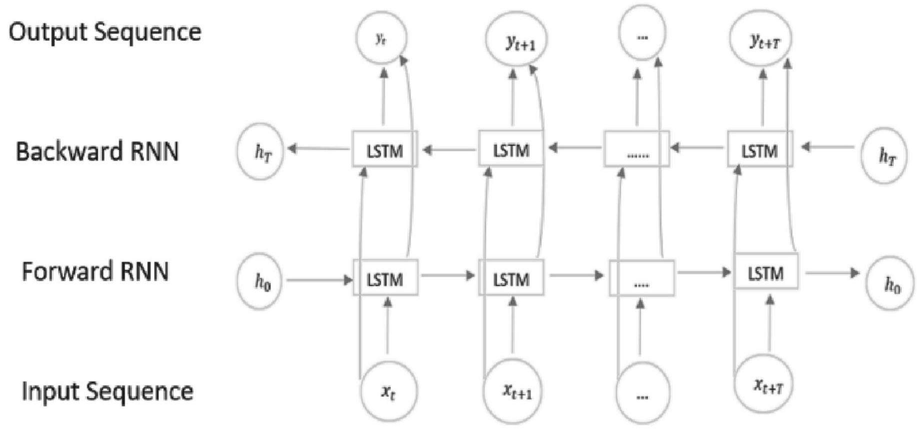

Fig. 5. Bi-LSTM Architecture

this deep design can obtain exceptional results [25]. In this study, there is a bidirectional LSTM layer giving the accuracy of 96.93% on 4 bandpower obtained from 19 channels.

2.3.5. Extreme Gradient Boosting

XGBoost is a modified Gradient Boosting Decision Tree technique that can successfully create boosted trees and execute in parallel as shown in Fig. 6. In XGBoost, there are two types of boosted trees: regression trees and classification trees [26]. The heart of XGBoost is optimising the value of the goal function. XGBoost includes a regularised model [27].The model in this study was tuned using GridSearchCV giving the accuracy of 97.01% on 4 bandpowers obtained from 19 channels.

2.3.6. Recurrent Neural networks

RNN is best suited for datasets that include details that is time-series or sequential in nature. RNN may retrieve information from previous stages' inputs and incorporate its influence on the present stage's input as well as its output. The information carried by the components in the sequence affect the outputs of recurrent neural networks. The connections between the nodes create a directed graph, which provides temporal dynamic behaviour [28]. By providing each output as feed to the following hidden layer, it may decrease the intricacies of rising parameters and memorise each previous detail. Eventually, all previous layers are combined to make a single recurring layer. In this study, the model consists of two RNN layers giving the accuracy of 96.15% on4 bandpower obtained from 19 channels.

2.3.7. Gated Recurrent Unit

GRU consists of reset gate and update gate, and the training parameters are much smaller than LSTM, so it needs less memory and runs faster. The Gated recurrent units may extract significant characteristics from time series input EEG signals [22]. There is no distinction between the internal and exterior states in a GRU, but in explicitly introducing a linear dependence between the current state as well as the state at the previous time, the gradient vanishing and gradient explosion are solved [29]. In this study, there are two layers of GRU that are giving an accuracy of 95.68% on 4 bandpower obtained from 19 channels.

2.3.8. Gradient Boosting

Gradient boosting has been applied successfully in classification, training to rank, structured prediction, and other fields. To avoid overfitting, it trains an ensemble of simple models [27]. The model in this study was tuned using GridSearchCV giving the accuracy of 92.42% on 4 bandpower obtained from 19 channels.

2.4. Parameters

The goodness of model can be evaluated on the basis of few parameters. Following are few of the parameters generally used for evaluating a model.

2.4.1. Accuracy

The ratio of correctly identified instances to the total number of cases being evaluated can also be used to determine accuracy. Accuracy has a greatest value of 1, and a worst value of 0.

$$Accuracy = \frac{Numberofcorrectlycases}{Totalnumberofcasesunderevaluation}$$

2.4.2. Precision

Both of the classes can be used to determine precision. The classifier's capability to refuse to classify a negative sample as positive is known as the precision of the negative class. The precision of the positive class refers to the classifier's ability to avoid categorising a positive sample as negative. Precision has a greatest value of 1, and a worst value of 0.

$$Precision = \frac{TruePositive/Negative}{Numberofcasesunderpredicted}$$

2.4.3. Sensitivity or recall of positive class

It is also possible to define recall in terms of either class. The recall of positive class is described as the ratio of True Positive to total number of positive cases. It is also known as sensitivity. It can be described simply as the classifier's capacity to identify every successful case.

$$Sensitivity = \frac{TruePositive}{Numberofactualpositivecases}$$

2.4.4. F1-score

Regardless of class imbalance, F1-score are regarded as one of the finest measures for classification models [30]. The weighted average of the class's precision and recall is the F1-score. The worst value is 0, while its best value is 1.

$$F1 - score = \frac{2*Precision*Recall}{Precision + Recall}$$

2.4.5. ROC and AUC score

The Receiver Operating Curve assists in determining the best threshold value for categorization. The area under the ROC curve is referred to as AUC and is used as a statistic to assess the classification model.

2.4.6. Hamming loss

Target misclassification is measured by the Hamming loss. The hamming loss has a best value of 0 and 1 is the worst value.

2.4.7. Jaccard score

It is determined by comparing the dimensions of all the anticipated labels and the actual labels' intersection and union with each other [30]. When comparing the projected classes with the actual classes, it is referred to as a similarity coefficient. The classification with a value of 1 is the best, while a value of 0 is the worst.

2.4.8. Cross-entropy loss

Deep neural networks become famous for using cross-entropy loss to solve vanishing gradient issues. It is also known as log loss. It quantifies the contamination brought on by misclassification. The cross-entropy loss is calculated by adding the logarithmic values of the prediction probability distribution for improperly identified data points.

2.5. Feature matrix

It is formed of r-rows and c-columns, where r rows represent the number of EEG data and c columns represent number of features extracted from EEG data. In this study 4 band power of each channel has been used.

Thus total number of EEG feature = 19 channel × 4 = 76 features.

2.6. Analysis of algorithm

The size of the dataset is = 60x76. The dataset is divided into 70:30 ratio i.e. 70% for training and 30% for testing.100 iteration of 10CV has been applied in the study for validation.

3. Results and discussion

Each of the band power feature i.e A) delta, theta, alpha and beta of 19 channels has been used B) Temporal region channel has been used C) All 4 bandpower of 19 channels has been used along with the classifiers Random Forest, Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), Bidirectional Long-Short Term Memory (Bi-LSTM), Gradient Boosting, Extreme Gradient Boosting (XGBoost).

3.1. Classifier results with delta, theta, alpha and beta band power of 19 channels.

Each of the four band power feature has been applied to all the 8 classifiers separately. The results are shown from Table 1.

Table 1 show that the highest accuracy attained is 89.85% in XG Boost model using alpha band power, 85.03% using Beta power using Gradient Boosting, 87.86% in Random Forest using Theta power,. accuracy of 84.76% in RNN using Delta band power. Table 1 shows that the highest among all the band power, accuracy of 84.75% in RNN model using alpha band power Table 2..

3.2. Result of temporal region

This section will discuss the result of temporal region regarding alpha and theta power band. Alpha and theta power of temporal region i.e. channel T7 and T8. In the previous studies it has been found that the depression affects the temporal region of the brain [18,31]. So the channels from the temporal region have been selected. From previous studies it has been also found that the alpha and theta power shows higher classification accuracy which can also be observed in the Fig. 6 which shows that the average of accuracy of using theta and alpha power is higher. Table 2 and Fig. 7 shows higher accuracy of 79.5% has been achieved using only Theta and Alpha bandpower.

3.3. Results based on combination of all band power of 19 channels

Band power of all 19 channels was used as feature in the 8 classifiers model i.e. Decision Tree, Random Forest, Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), Bidirectional Long-Short Term Memory (Bi-LSTM), Gradient Boosting, Extreme Gradient Boosting (XGBoost). Classification accuracy of more than 92% has been achieved in all the 8 classifiers with highest classification of 98.13% obtained using CNN as shown on Table 3 and Fig. 8.

The study confirms that EEG signal can be used as an adjunct tool for detection of depression as highest classification accuracy of 98.13% is obtained using CNN and all band power of 19 channels. The study also reconfirms the fact that the depression effects the theta and alpha band power based on the higher average accuracy achieved as shown in Fig. Also that depression affects in the temporal region of the brain can be confirms that depression affects the temporal region of the brain which also supported in the previous studies [1,2,18,31].

4. Conclusion

Highest classification accuracy of 98.13% is obtained using CNN and all band power of 19 channels. Also high accuracy of 89.85% is achieved using only Alpha band power shows the significance of alpha band power for detection of depression. The study also confirms that the depression affects the temporal region of the brain since a high accuracy of 79.5% is achieved using Theta and Alpha band power using only data from temporal Region channels i.e. 2 channel T7 and T8. Since a high accuracy of 98.13% is obtained using CNN, it can act as an adjunct tool to detect depression with high accuracy. This could in terms help in correct and accurate diagnosis of depression and act as a biomarker.

Table 1
Result of different classifier on Alpha, Beta Theta and Delta band power.

Classifier	Accuracy of Alpha band power (%)	Accuracy of Beta band power (%)	Accuracy of Theta band power (%)	Accuracy of Delta band power (%)
Random Forest	82.33	84.22	**87.86**	81.78
CNN	72.04	74.88	83.86	76.17
RNN	83.55	86.98	87.60	**84.76**
Bidirectional LSTM	83.35	77.33	86.59	76.24
GRU	78.03	70.53	81.92	77.88
LSTM	79.27	65.20	87.64	80.29
Gradient Boosting	81.96	**85.03**	87.68	82.15
XG Boost	**89.85**	76.24	83.32	78.89

Table 2
Classification using Theta and Alpha band power using only data from temporal Region channels i.e. 2 channels T7 and T8.

Methods	Accuracy of temporal region of Alpha and Theta power spectrum (in %)
Random Forest	**79.50**
CNN	76.40
RNN	79.40
LSTM	78.70
GRU	79.08
Bidirectional LSTM	78.38
Gradient Boosting	77.64
XG Boost	77.00

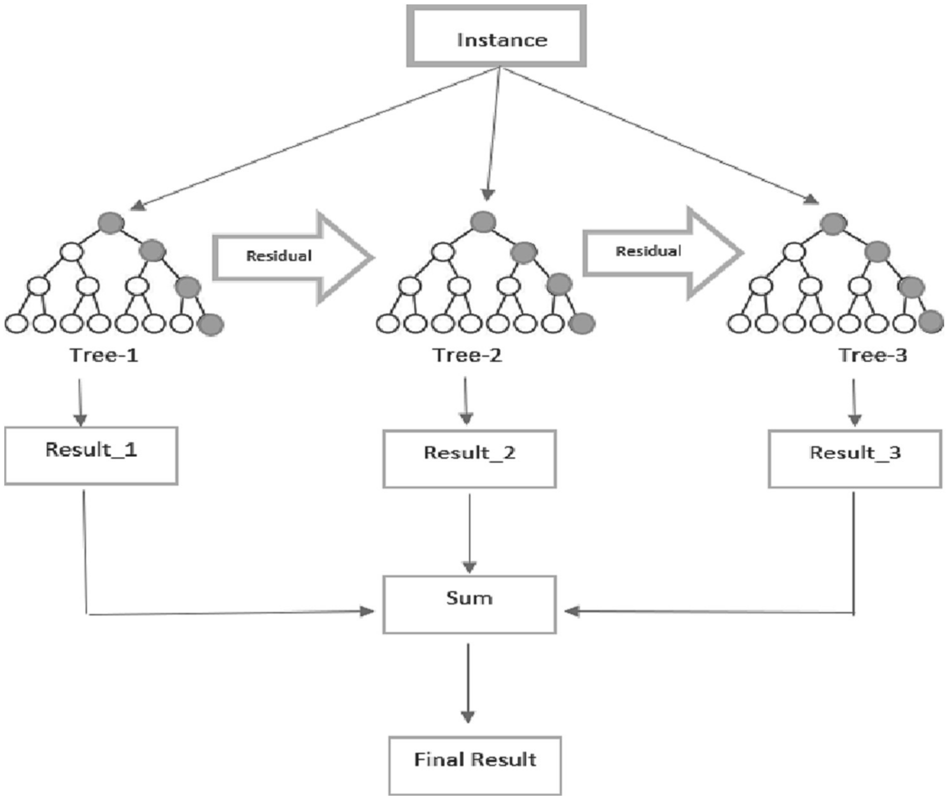

Fig. 6. Simplified XG Boost Architecture.

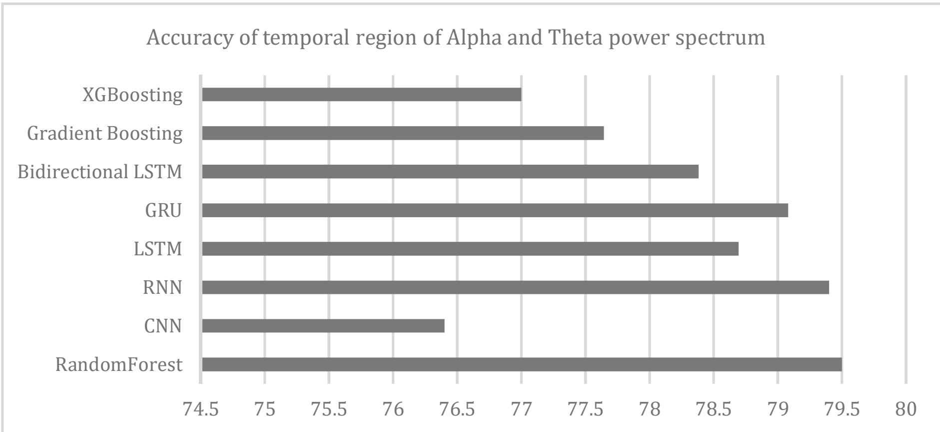

Fig. 7. Comparison of accuracy using only temporal region channel T7 and T8.

Table 3
Result of all 4 band power of 19 channels using 8 classifiers.

	Accuracy	Precision	Sensitivity	Specificity	f1_Score	ROC-AUC score	Hamming_loss	log_loss	jaccard_score
Random Forest	96.7	0.97	0.96	0.97	0.97	0.97	0.033	0.15	0.93
CNN	**98.13**	0.99	0.97	0.99	0.98	0.98	0.0187	0.05	0.96
RNN	96.15	0.96	0.96	0.97	0.96	0.96	0.0385	0.12	0.92
LSTM	96.7	0.96	0.97	0.96	0.97	0.97	0.033	0.09	0.94
GRU	95.68	0.97	0.94	0.97	0.96	0.96	0.0432	0.12	0.91
Bidirectional LSTM	96.93	0.97	0.97	0.97	0.97	0.97	0.0307	0.09	0.94
Gradient Boosting	92.42	0.92	0.92	0.92	0.92	0.92	0.0758	0.23	0.85
XG Boost	97.01	0.97	0.97	0.97	0.97	0.97	0.0299	0.09	0.94

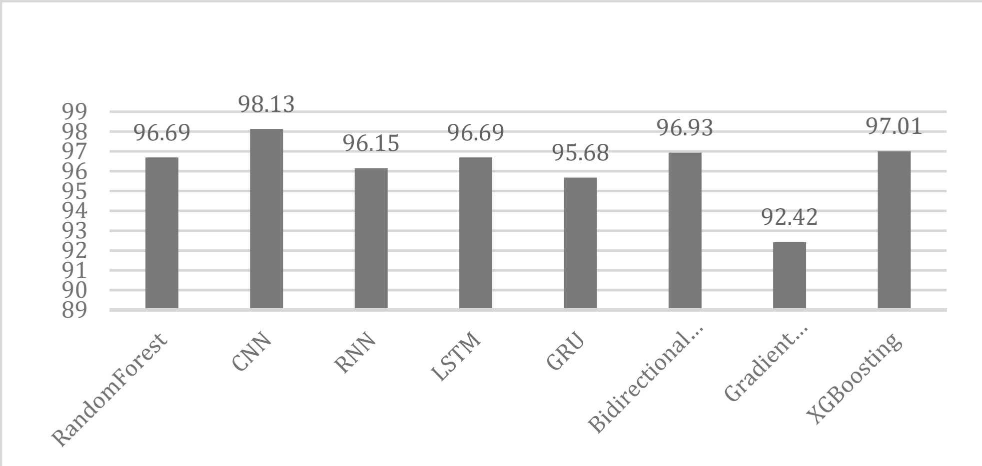

Fig. 8. Result of overall dataset.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgment

The authors gratefully acknowledge technical support and resources provided by the Indian Institute of Information Technology (IIIT), Ranchi, India.

The authors also would like to thank Dr. Rakesh Kumar Sinha, Professor and Head, Bioengineering and Biotechnology, Birla Institute of Technology (BIT), Mesra, India and Dr. Nishant Goyal, Professor of Psychiatry, Central Institute of Psychiatry (CIP) Ranchi, India for always supporting us with their immense knowledge and expertise.

The authors also would like to thank Mumtaz et al. [15] for the dataset contribution.

Funding: This research work was funded by Ministry of Micro, Small & Medium Enterprises (MSME) Innovative Scheme, Government of India under grant no. IDEAJH 00327.

Compliance with ethical standards

Conflict of Interests: The authors have no conflict of interests regarding the publication of this paper.

Ethical approval: This article does not contain any studies with human participants performed by any of the authors

References

[1] S. Mahato, S. Paul, Detection of major depressive disorder using linear and non-linear features from EEG signals, Microsyst. Technol. 25 (3) (2019) 1065–1076, https://doi.org/10.1007/s00542-018-4075-z.

[2] S. Mahato, S. Paul, Classification of depression patients and normal subjects based on electroencephalogram (eeg) signal using alpha power and theta asymmetry, J. Med. Syst. 44 (1) (2020), https://doi.org/10.1007/s10916-019-1486-z.

[3] Suganthi, V., Scholar, R., & Punithavalli, M. (2022). A Review of Deep Learning and Machine Learning Methods for Analyzing Covid-19 Stress. In *International Journal of Innovative Science and Research Technology* (Vol. 7, Issue 3). www.ijisrt.com.

[4] M. Liu, J. Tan, Y. Jiang, Y. Tian, Using deep learning to decode abnormal brain neural activity in MDD from single-trial EEG signals, Https://Doi.Org/10.1080/27706710.2022.2075242 1 (1) (2022) 28–37, https://doi.org/10.1080/27706710.2022.2075242.

[5] K.J. Mathewson, M.K. Jetha, I.E. Drmic, S.E. Bryson, J.O. Goldberg, L.A. Schmidt, Regional EEG alpha power, coherence, and behavioral symptomatology in autism spectrum disorder, Clin. Neurophysiol. 123 (9) (2012) 1798–1809, https://doi.org/10.1016/j.clinph.2012.02.061.

[6] J. Zhu, Z. Wang, T. Gong, S. Zeng, X. Li, B. Hu, J. Li, S. Sun, L. Zhang, An improved classification model for depression detection using eeg and eye tracking data, IEEE Trans. NanoBiosci. 19 (3) (2020) 527–537, https://doi.org/10.1109/TNB.2020.2990690.

[7] H. Albaqami, G.M. Hassan, A. Subasi, A. Datta, Automatic detection of abnormal EEG signals using wavelet feature extraction and gradient boosting decision tree, Biomed. Signal Process. Control 70 (2021) 102957.

[8] M.S. Khan, N. Salsabil, M.G.R. Alam, M.A.A. Dewan, M.Z. Uddin, CNN-XGBoost fusion-based affective state recognition using EEG spectrogram image analysis, Sci. Rep. 12 (1) (2022) 14122, https://doi.org/10.1038/s41598-022-18257-x.

[9] Orgo, L., Bachmann, M., Kalev, K., Hinrikus, H., & Järvelaid, M. (n.d.). *Brain Functional Connectivity in Depression: Gender Differences in EEG**.

[10] U.M. Haque E. Kabir R. Khanam H. Wang Detection of child depression using machine learning methods PLoS ONE 16 12 e0261131.

[11] Ul, M., Arif, I., Camargo, M., Forkel, J., Holdack, G., Drumond, R. R., Schilling, N., Hensch, T., Hegerl, U., Schmidt-Thieme, L., Rafael, ·, & Drumond, R. (n.d.). *Depression Diagnosis using Deep Convolutional Neural Networks*. https://doi.org/10.5445/KSP/1000087327/29.

[12] M. Bachmann, L. Päeske, K. Kalev, K. Aarma, A. Lehtmets, P. Ööpik, J. Lass, H. Hinrikus, Methods for classifying depression in single channel EEG using linear and nonlinear signal analysis, Comput. Methods Programs Biomed. 155 (2018) 11–17, https://doi.org/10.1016/J.CMPB.2017.11.023.

[13] W. Liu, K. Jia, Z. Wang, Z. Ma, A Depression prediction algorithm based on spatiotemporal feature of eeg signal, Brain Sci. 12 (5) (2022), https://doi.org/10.3390/brainsci12050630.

[14] X. Li, D. Song, P. Zhang, G. Yu, Y. Hou, B. Hu, Emotion recognition from multi-channel EEG data through convolutional recurrent neural network, IEEE Int. Conf. on Bioinf. Biomed. (BIBM) 2016 (2016) 352–359, https://doi.org/10.1109/BIBM.2016.7822545.

[15] W. Mumtaz, L. Xia, M.A. Mohd Yasin, S.S. Azhar Ali, A.S. Malik, D. Hu, A wavelet-based technique to predict treatment outcome for major depressive disorder, PLoS One 12 (2) (2017), https://doi.org/10.1371/journal.pone.0171409.

[16] C. Cusin, H. Yang, A. Yeung, M. Fava, Rating Scales for Depression, in: L. Baer, M. A. Blais (Eds.), Handbook of Clinical Rating Scales and Assessment in Psychiatry and Mental Health, Humana Press, Totowa, NJ, 2010, pp. 7–35.

[17] Klem, G. H., Otto Lu Èders, H., Jasper, H., & Elger, C. (n.d.). *The ten±twenty electrode system of the International Federation*.

[18] S. Mahato, N. Goyal, D. Ram, S. Paul, Detection of depression and scaling of severity using six channel eeg data, J. Med. Syst. 44 (7) (2020), https://doi.org/10.1007/s10916-020-01573-y.

[19] K. Vijayalakshmi, M. Vinayakamurthy, Anuradha., A hybrid recommender system using multiclassifier regression model for autism detection, in: Proceedings of the International Conference on Smart Technologies in Computing, Electrical and Electronics, 2020, https://doi.org/10.1109/ICSTCEE49637.2020.9277034.

[20] G. Devi S, S. Vairavasundaram, Y. Teekaraman, R. Kuppusamy, A. Radhakrishnan, S.A.C. Bukhari, A deep learning approach for recognizing the cursive tamil characters in palm leaf manuscripts, Comput. Intell. Neurosci. 2022 (2022) 1–15.

[21] U.R. Acharya, S.L. Oh, Y. Hagiwara, J.H. Tan, H. Adeli, D.P. Subha, Automated EEG-based screening of depression using deep convolutional neural network, Comput. Methods Programs Biomed. 161 (2018) 103–113, https://doi.org/10.1016/j.cmpb.2018.04.012.

[22] Qayyum, A., Mumtaz, W., & Razzak, I. (n.d.). *Hybrid Deep Shallow Network for Diagnosis of Patient Suffering from Depression using Electroencephalogram Signals Neurosignal Processing View project Depression Assessment using EEG based on Deep Learning Models View project Hybrid Deep Shallow Network for Assessment of*

Depression using Electroencephalogram Signals. https://www.researchgate.net/publication/344999628.
[23] B. Ay, O. Yildirim, M. Talo, U.B. Baloglu, G. Aydin, S.D. Puthankattil, U. R. Acharya, Automated depression detection using deep representation and sequence learning with eeg signals, J. Med. Syst. 43 (7) (2019), https://doi.org/10.1007/s10916-019-1345-y.
[24] S. Hochreiter, J. Schmidhuber, Long short-term memory, Neural Comput. 9 (8) (1997) 1735–1780, https://doi.org/10.1162/neco.1997.9.8.1735.
[25] N.A. Ali, A.R. Syafeeza, A.S. Jaafar, S. Shamsuddin, N.K. Nor, LSTM-based Electroencephalogram classification on autism spectrum disorder, Int. J. Integr. Eng. 13 (6) (2021) 321–329. https://doi.org/10.30880/ijie.13.06.028.
[26] Li, W., Li, Q., Wang, L., IEEE Engineering in Medicine and Biology Society, & Institute of Electrical and Electronics Engineers. (n.d.). Proceedings, 2018 11th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics : CISP-BMEI 2018 : 13-15 October 2018, Beijing, China.
[27] T. Chen, C. Guestrin, XGBoost: A scalable tree boosting system, in: Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2016, https://doi.org/10.1145/2939672.2939785.
[28] A. Sarkar, A. Singh, R. Chakraborty, A deep learning-based comparative study to track mental depression from EEG data, Neurosci. Inf. 2 (4) (2022), 100039, https://doi.org/10.1016/j.neuri.2022.100039.
[29] X. Wang, B. Guo, X. Ma, X. Bai, Research on emotional music reconstruction method based on DBN-GRU, J. Phys. Conf. Ser. 1966 (1) (2021) 012016.
[30] R. Yousef, G. Gupta, N. Yousef, M. Khari, A holistic overview of deep learning approach in medical imaging, Multimedia Syst. 28 (3) (2022) 881–914, https://doi.org/10.1007/s00530-021-00884-5.
[31] A.O. Khadidos, K.H. Alyoubi, S. Mahato, A.O. Khadidos, S.N. Mohanty, Computer aided detection of major depressive disorder (mdd) using electroencephalogram signals, IEEE Access 11 (2023) 41133–41141, https://doi.org/10.1109/ACCESS.2023.3262930.
[32] S. Mahato, L.K. Pathak, K. Kumari, Detection of Schizophrenia Using EEG Signals. Data Analytics in Bioinformatics: A Machine Learning, in: R. Satpathy, T. Choudhury, S. Satpathy, S.N. Mohanty, X. Zhang (Eds.), Data Analytics in Bioinformatics: A Machine Learning Perspective, Wiley, 2021, pp. 359–390.
[33] S. Mahato, A. Roy, A. Verma, S. Paul, in: Analysis of Electroencephalogram (EEG) Signals for Detection of Major Depressive Disorder (MDD) Using Feature Selection and Reduction Techniques, Springer Singapore, 2021, pp. 429–440.
[34] S. Mahato, S. Paul, Analysis of region of interest (RoI) of brain for detection of depression using EEG signal. Multimedia Tools and Applications. https://doi.org/10.1007/s11042-023-15827-7.