INDIAN JOURNAL OF SCIENCE AND TECHNOLOGY

ORIGINAL ARTICLE

Check for updates

OPEN ACCESS

Received: 10/07/2025

Accepted: 20/08/2025

Published: 08/09/2025

Citation: Chaturvedi A, Raj A (2025) Advancements in Machine Learning Techniques for Epileptic Seizure Prediction Using EEG Data. Indian Journal of Science and Technology 18(34): 2770-2783. https://doi.org/10.17485/IJST/v18i34.1224

* **Corresponding author**.

anmol31189@gmail.com

Funding: None

Competing Interests: None

Published By Indian Society for Education and Environment (iSee)

ISSN
Print: 0974-6846
Electronic: 0974-5645

Advancements in Machine Learning Techniques for Epileptic Seizure Prediction Using EEG Data

Anmol Chaturvedi[1], **Ashish Raj**[1]*

1 Department of EEE, Poornima University, Jaipur, Rajasthan, India

Abstract

Background: Automated EEG-based seizure detection can reduce subjectivity in clinical workflows but often suffers from overfitting and unclear validation. **Objectives:** The proposed study will help to improve the prediction of epileptic seizures based on the application of advanced mathematical descriptors and machine learning (ML) models. It aims to help eliminate disadvantages of manual drawing and subjective evaluation of EEGs by automating the process of seizure classification, and using statistically relevant, domain-combining features. **Methods:** The paper involved the use of EEG data in Bonn Dataset, where seven major features were extracted which are Detrended Fluctuation Analysis (DFA), Hjorth Fractal Dimension (HFD), SVD Entropy, Spectral Entropy, Fisher Information, Petrosian Fractal Dimension (PFD), and mean amplitude. To extract the frequency sub-bands, Daubechies and Biorthogonal wavelets were used to decompose a Wavelet (DWT). Classification of such features was done with the help of several ML algorithms and Artificial Neural Network (ANN) trained through Levenberg-Marquardt algorithm. To mitigate optimism bias, we adopt nested stratified cross-validation (outer 10-fold; inner 5-fold) with blocked splits by original record, plus a held-out test set. Metrics include accuracy, precision, recall, F1-score, ROC-AUC, MCC, and 95% CIs; model calibration (Brier score) and permutation-based feature importance are reported. **Findings:** ANN model demonstrated the accuracy of classification of 99%, where Random Forest and XGBoost demonstrated highest performance equally (99.3 % and 99.1%, respectively). The assessment on confusion matrix, ROC curve as well as F1-score showed that the proposed models performance exceeded the former state-of-the-art methods by 3-5 percent. The Kruskal-Wallis was used to remove redundant features allowing minimization of overfitting and enhancing generalization across classes of patients. **Novelty:** The proposed work integrates multi-domain statistical and nonlinear EEG features with validated hybrid deep convolutional neural network-based classifiers. The proposed work provides an innovative approach to integrate multi-domain statistical and nonlinear EEG features which includes time, frequency, and nonlinear signal behavior with optimal learning structures. Compared to the previous works that use single-domain or end-to-end black-box models, the current study incorporates the interpretable and computationally feasible modules and establishes a new bar in the seizure detection task that is based on EEG signals. **Interpretation:** Performance remains high under stringent validation, indicating low overfitting. The feature-engineered pipeline is computationally efficient and interpretable. **Limitations:** Bonn lacks subject identifiers; therefore cross-dataset evaluation is recommended in future work.

Keywords: Epileptic Seizure Prediction; EEG Analysis; Machine Learning; Artificial Neural Network; Biomedical Signal Processing

1 Introduction

Epilepsy is a common disease of the brain involving uncontrollable and excessive activities of the neurons that may result in frequent seizures and loss of consciousness, body control and thinking[1]. EEG (Electroencephalography) is still one of the main diagnostics by its high density of time, low price, and non-invasive character[2]. Manual process of seizure detection analysis is involved in conventional analysis of EEG but is hinged with subjectiveness and labour intensive implying the demand of automated detection of seizures[3].

The recent research has proved that deep learning (DL) has the ability to automate EEG signal classification to detect seizures. It has also verified that CNNs and LSTM networks are effective in capturing spatial and temporal EEG features, respectively[4],[5]. Although these systems claim high accuracy on extremely well-controlled datasets, they will fail in the condition of patient variability, signal artifacts, or interictal noise[6].

In this regard, scientists have considered hybrid deep learning models, including CNN-BiLSTM and attention-augmented networks, which enhance temporal sensitivity and interpretability[7]. EEG inputs transformed by a wavelet to extract frequency-sensitive features have been tested by others as well, though the problem of generalizing to more effective seizure types will always exist[8]. Statistical and nonlinear dynamic features also have good potential when used in feature-based machine learning pipelines, but at the tradeoff of scalability and holistic automation[9].

Despite all these improvements, the current practices still exhibit shortcomings in their strength, multi-patient dataset flexibility, and functional feature incorporation. Reciprocally, this research study recommends a new deep learning based framework that is relevant to successful classification of seizures based on EEG recordings. The model will integrate the spatial, temporal and frequency-domain, thus providing better diagnostic efficiency and robustness to inter-patient variability - bridging the gap in the existing EEG-based seizure detection domain[10]. Nonetheless, in the current literature, the three major challenges persist that include (i) weakness of multi-patient EEG datasets, (ii) insufficient use of time frequency features in combination with statistical and nonlinear measures, and (iii) the lack of a rigorous comparison practice between traditional ML and neural models. Past literature usually uses handcrafted features or end-to-end DL solutions exclusively and hence lacks in model generalizability and interpretability. This paper contributes to address these gaps by creating a hybrid EEG classification framework, in which wavelet decomposition, feature selection that is statistically significant, and good, best neural models are integrated. The goal is to find better predictive performance circulating with improved explainability, computational cost effectiveness and tractability.

To address the gaps in generalizability, interpretability, and overfitting present in many of the existing reviewed models, the proposed research explores hybridized feature-based and neural network approaches. It contrasts various classical ML models and ANN configurations—both with and without wavelet decomposition—on a benchmark EEG dataset. In contrast to the existing literature where the statistical significance is not reported or no tuning model of the architectures (comparative) is used, this paper uses the Kruskal-Wallis approach, recursive feature elimination, and ensemble model-benchmarking to produce a more discriminating and adaptive pipeline of seizure detection exhibiting higher accuracy and model resilience.

This study uses a public, anonymized benchmark (Bonn); no new human data were collected. For clinical deployment, models should be validated on **multi-center, subject-identified datasets** under IRB oversight, with explicit **data protection** (secure storage, encryption at rest/in transit) and **model monitoring** (drift detection, audit logs). Clinically, the proposed pipeline is amenable to **bedside screening** or **ambulatory pre-screening** because it is lightweight and interpretable; however, it must remain **decision-support** rather than diagnostic, with clinician-in-the-loop review and calibrated thresholds suited to sensitivity-critical use cases (e.g., alarms).

2 Methodology

The range of the study will include the methodologic approach of our study and applying and comparing various ML methods to predict epileptic seizure by using EEG records. Understanding. The chance of early prediction of the approaching seizures is a very important aspect of the life of an epileptic patient.[11] In few of the instances, there is a great chance that prospect of a seizure could be easily predicted a long way in advance as long as the required medication or measure is taken. EEG is normally conducted to diagnose epilepsy since the procedure involves recording the electrical activity in the brain. Multiple approaches of machine learning have been used in literature on screening characteristics of EEG signals and foreseeing instances of seizures.

2.1 Dataset

First, the source of data is the Epileptology department of the Bonn University as Bonn EEG benchmark, comprising five sets (A–E) with 100 single-channel EEG segments per set (total 500 segments), the EEG recordings of 5 sets $\times$ 100 segments totaling of 500 segments. Each segment is 23.6 s long with 4096 samples (sampling frequency $\approx$ 173.6 Hz). Sets A/B: healthy volunteers (eyes open/closed); C/D: interictal recordings from epileptic subjects (intracranial); E: ictal segments. Note: This dataset contains 500 segments, not 500 individuals. Data Sets used in research are (i) binary tasks (seizure vs. non-seizure) and (ii) multi-class tasks (A–E).

To minimize leakage, we block splits by original record/segment IDs so that no portion of any original segment contributes to both training and testing folds. Further, conducted set-wise experiments (e.g., train on {A,B,C,D} and test on {E}) to generalization.

(https://www.upf.edu/web/ntsa/downloads/-/asset_publisher/xvT6E4pczrBw/content/2001-indications-of-nonlinear-deterministic-and-finite-dimensional-structures-in-time-series-of-brain-electrical-activity-dependence-on-recording-regi)

2.2 Method-1: EEG Classification Without Decomposition

Initially, used some basic machine learning with models offered by scikit learn library to our data set. Specifically I tried Nearest Neighbors, Linear SVM, RBF SVM, Gaussian Process, Decision Tree, Random Forest, Multi-layer Perceptron (fully connected Neural Net), AdaBoost, Naive Bayes and Quadratic Discriminant Analysis. In order to work out the dataset of these algorithms. Six features have been used namely: DFA (Detrended Fluctuation Analysis), HFD (Hjorth Fractal Dimension), SVD Entropy, Spectral Entropy, Fisher Information, and PFD (Petrosian Fractal Dimension).

The first search was done using a package of machine learning models available in scikit-learn library, namely, Nearest Neighbors, Linear SVM, RBF SVM, Gaussian Process, Decision Tree, Random Forest, Multi-layer Perceptron, AdaBoost, Naive Bayes, and Quadratic Discriminant Analysis.

Looking upon up six important characteristics, composed of DFA, HFD, SVD Entropy, Spectral Entropy, Fisher Information, and PFD in these models. It is thought that these features can summarise or capture some of the complexity and fractal characteristics of the EEG signal, the entropy and order structure in the data which may be related to epileptic behaviour.

Incorporating the mean amplitude as a seventh feature significantly improved model accuracies. This can suggest that the average signal level which may be indicative of the overall level of excitation in the brain impacts far reaching predictive value. In this work, the theoretical background consists in the extraction of useful features of EEG signals using the sophisticated mathematical and statistical techniques and the classification of the latter by applying the machine learning models and, specially, an artificial neural network. Among the most important features is Hjorth Fractal Dimension and is able to give measurement of the complexity of the signal. It measures the amount of space the EEG signal takes a long time, estimated through the calculation ratio of logarithmic values of steps between the time series and the path taken by the signal. This degree assists in the differentiation of normal and abnormalities in the brain activities since more complex designs usually follow pre-seizure conditions. Detrended Fluctuation Analysis (DFA) is given as-

$$F(n) = \sqrt{\frac{1}{N}\sum_{i-1}^{N}\left(y(i) - y_n(i)\right)^2} \tag{1}$$

Where $F(n)$ is the fluctuation of the integrated and detrended time series, N is the length of the time series, $y(i)$ is the integrated time series, and $y_n(i)$ is the local trend in a window of size n. (**Eq. 1**) quantifies the fluctuation in detrended time series and is crucial for identifying long-term correlations within EEG patterns during pre-seizure activity. Hjorth Fractal Dimension (**Eq. 2**) measures signal complexity

$$HFD = \frac{\log_{1n}(N)}{\log_{10}(d)} \tag{2}$$

Where N is the number of steps in the EEG time series and d is the distance covered by the signal.

$$H_{SVD} = \sum_{i-1}^{r} \left(\frac{\sigma_i}{\sum_{j-1}^{r} \sigma_j}\right) log \left(\frac{\sigma_1}{\sum_{j-1}^{r} \sigma_j}\right) \tag{3}$$

Where σ_i are the singular values from the SVD of the EEG signal matrix, and r is the rank of the matrix.

$$H_{\text{spec}} = \sum_{i-1}^{M} p(f_i) \log_2 p(f_i) \tag{4}$$

Where $p(f_i)$ is the power spectral density normalized over frequencies f_i, and M is the number of discrete frequencies.

The other major characteristic is Singular Value Decomposition Entropy which is used to measure the disorder in EEG signal based on the distribution of the singular values present in the matrix representation of signal. A normalized set of singular values is used to compute the entropy, with each value computed in relation to the total sum of all the singular values, with the total later averaged over the rank of the matrix. Higher entropy means more randomness in the signal that can be an indication of neurological instability. Fisher Information is given as-

$$I_F = \sum_{i-1}^{N} \left(\frac{d}{dx} p(x_i)\right)^2 / p(x_i) \tag{5}$$

Where $p(x_i)$ is the probability density function of the signal's amplitude and x_i is the amplitude at point i.

Spectral Entropy is yet another perspective of signal complexity, presented however in the frequency domain. It is a measure of the distribution of the power over the different frequencies in the form of the normalized power spectral density. In situations where the power is focused to a smaller number of frequencies, the entropy is small which means that the signal is more predictable. On the other hand, a flatter distribution of power spectrum is associated with greater entropy, which is usually observed in epileptic EEG.

$$\text{PFD} = \frac{\log_{11} N}{\log_{10} N + \log_{10}\left(\frac{N}{N + 4N_6}\right)} \tag{6}$$

Where N is the number of samples in the time series and N_δ is the number of sign changes in the binary sequence derived from the time series.

$$\mu = \frac{1}{N} \sum_{i-1}^{N} x_i \tag{7}$$

Where $\mu$the mean value of the EEG is signal and x_i is the value at the i^{th} point.

$$DWT(x_i) = \sum_{i--\infty}^{\infty} x_i \psi_{a,b}(t) \tag{8}$$

Where $\psi_{a,b}(t)$ is the wavelet function scaled by a and translated by b, and x_i is the EEG signal.

$$cA_j = \sum_k h_k \cdot x_{2j+k} \tag{9}$$

Where h_k is the filter coefficient at scale k, and x_{2j+k} is the signal at a dyadic scale.

$$cD_j = \sum_k g_k \cdot x_{2j+k} \tag{10}$$

Where g_k is the wavelet coefficient at scale k, similar to h_k but for detailed coefficients.

The Fisher Information gauges the information content in the signal representing the EEG in terms of its hidden probability distribution. It measures the rate at which the probability density varies but with a weight equal to the probability. The amount of high-gradients or abrupt turns in the distribution of signal amplitude leads to an increase in Fisher information relating to the higher number of structures or predictability, which had been reduced in the seizure onset. Petrosian Fractal Dimension is an efficient algorithm allowing calculating the complexity of the signal with regard to the sign changes in binary code obtained on the basis of the EEG. It relies on the total amount of samples and the percentage of zero-crossings to differentiate fractal like patterns associated with epileptic processes and normal rhythmic ones.

This is another benchmark metric which is the average EEG signal amplitude over a period of time. It is simple but serves as a basis of normalization and additional studies. The more advanced analysis introduces the Discrete Wavelet Transform that breaks the EEG signal into a localized time-frequency-based components with scaled and wavelet function. The wavelet transform supplies coefficients of different scales and they are the approximate coefficient and detailed coefficients. Such coefficients allow one to extract the low-frequency trends as well as the high frequency features of the signal.

Reconstruction from Wavelet Coefficients:

$$x_i = \sum_j \left(cA_j + cD_j\right) \tag{11}$$

Where cA_j and cD_j are the approximate and detailed coefficients from the DWT at level j.

Feature Vector Construction for Machine Learning:

$$F = [F_1, F_2, \ldots, F_M] \tag{12}$$

Where **F** is the feature vector consisting of M features extracted from the EEG signal.

Multi-layer Perceptron (MLP) Activation Function:

$$a_i^{(l)} = \phi\left(\sum_j w_{ij}^{(l)} a_j^{(l-1)} + b_i^{(l)}\right) \tag{13}$$

Where $a_i^{(l)}$ is the activation of the i^{th} neuron in the l^{th} layer, $w_{ij}^{(l)}$ are the weights, $b_i^{(l)}$ is the bias, and ϕ is the nonlinear activation function.

Performance Metric:

$$\text{Accuracy} = \frac{TP+TN}{TT^2+TN+FP^2+FN} \tag{14}$$

where TP, TN, FP, and FN are the true positives, true negatives, false positives, and false negatives, respectively.

Neural Network Backpropagation Update Rule:

$$w_{ij}^{(l)}(t+1) = w_{ij}^{(l)}(t) - \alpha \frac{\partial \mathcal{L}}{\partial w_{lj}^{ll}} \tag{15}$$

Where $w_{ij}^{(l)}(t+1)$ is the updated weight, α is the learning rate, and $\frac{\partial \mathcal{L}}{\partial w_{ij}^{(l)}}$ is the gradient of the loss function $\mathcal{L}$ with respect to the weight. Recovering the signal using these coefficients, it is possible to validate and verify data integrity. Several parameters extracted are then compiled in a feature vector which is a multidimensional set to be fed to the machine learning model. During classification, multilayer perceptron neural network is utilized whereby the input is administered by individual neurons using weighted aggregations, biases and nonlinear activation functions. Back propagation is used to reduce the loss function during training to move the weights in gradient direction. Finally, all performance measures like accuracy could be calculated with the values of false positives, false negative, true positives and true negative, provide the comprehensible approach of comprehension and evaluating the ability of the model to manage surprises. This is a combination of the feature engineering and the usage of the advanced classification that assures high accuracy in the seizure prediction. These included Detrended Fluctuation Analysis (DFA) with window sizes ranging from 4–512 to compute the log–log slope, Higuchi Fractal Dimension (HFD) with kmax=8, Petrosian Fractal Dimension (PFD), Singular Value Decomposition (SVD) Entropy with matrix rank estimated via a tolerance of 1e-10, Spectral Entropy computed using Welch's method with a 256-point Hamming window and normalization, Fisher Information derived from kernel density estimation (KDE) with bandwidth set to "Scott's rule," and mean amplitude.

Using these 6 features, machine learning models' accuracy are as follows: [[0.9, 'Nearest Neighbors'], [0.76, 'Linear SVM'], [0.87, 'RBF SVM'], [0.98, 'Gaussian Process'], [1.0, 'Decision Tree'], [1.0, 'Random Forest'], [0.86, 'Neural Net'], [1.0, 'AdaBoost'], [0.99, 'Naive Bayes'], [0.73, 'QDA']] I added mean of each signals values as the 7th feature, and noticed a significant improvement in the accuracies across all models: [[0.98, 'Nearest Neighbors'], [0.95, 'Linear SVM'], [0.96, 'RBF SVM'], [1.0, 'Gaussian Process'], [1.0, 'Decision Tree'], [1.0, 'Random Forest'], [0.98, 'Neural Net'], [1.0, 'AdaBoost'], [0.99, 'Naive Bayes'], [0.93, 'QDA']].

The preprocessing included z-score normalization per segment, with an optional 0.5–40 Hz band-pass filter and artifact rejection performed by amplitude clipping at $\pm 3\sigma$. Two distinct feature sets were considered: (a) hand-crafted features comprising Detrended Fluctuation Analysis (DFA), Higuchi Fractal Dimension (HFD), Petrosian Fractal Dimension (PFD), Singular Value Decomposition (SVD) Entropy, Spectral Entropy, Fisher Information, and Mean, and (b) DWT-enhanced features obtained using db4 or bior4.4 wavelets at level 4, followed by extraction of the same feature set across sub-bands. Feature selection was carried out using the Kruskal–Wallis test (α = 0.01) with Bonferroni correction, supplemented by Recursive Feature Elimination (RFE) with a Random Forest estimator where applicable. Model selection employed nested cross-validation with an outer 10-fold stratified loop and an inner 5-fold randomized search for hyperparameter tuning. Further, Several classifiers were employed to evaluate model performance, each with carefully defined hyperparameter search ranges. For Random Forest (RF), the search space included n_estimators between 200 and 1000, maximum depth ranging from 8 to unlimited, minimum samples per leaf set to {1, 2, 4}, maximum features {sqrt, log2}, and class weights set to balanced. For XGBoost, hyperparameters were varied over n_estimators {300–1200}, maximum depth {3–8}, learning rate {0.01–0.2}, subsample {0.6–1.0}, colsample_bytree {0.6–1.0}, and regularization terms reg_alpha {0–0.5} and reg_lambda {1–5}. The Support Vector Machine (SVM) with RBF kernel was tuned with C values spanning 1e-2 to 1e3, γ between 1e-4 and 1, and balanced class weights. Finally, the Multi-Layer Perceptron (MLP) was configured as a single hidden layer network with {32, 64, 128} neurons, activation functions {ReLU, tanh}, L2 regularization α in the range 1e-5 to 1e-2, a maximum of 500 iterations, and early stopping enabled to prevent overfitting. A 20% held-out test set, stratified and blocked by record, was used for final evaluation. Performance was assessed using accuracy, macro-precision, macro-recall, macro-F1, ROC-AUC, Matthews Correlation Coefficient (MCC), and Brier score, along with 95% confidence intervals estimated via bias-corrected and accelerated (BCa) bootstrap with 2000 resamples and calibration curve analysis. Statistical comparisons against baseline models were conducted using the Wilcoxon signed-rank test with Holm correction, and effect sizes were reported using Cliff's δ. Finally, interpretability was addressed through permutation feature importance and SHAP analysis (for tree-based models) to identify and rank the most influential features.

Flowchart for Hierarchical Process for Extracting and Selecting Relevant EEG Features: This flowchart outlines the sequential steps of acquiring EEG signals, performing preprocessing, feature extraction (time, frequency, and nonlinear domains), optional wavelet decomposition, and feature selection using the Kruskal-Wallis test (Figure 1).

2.3 Method – 2: EEG Classification with DWT Decomposition

Following the papers [3], [10], and [12], it was decided to decompose EEG signals into discrete wavelet transform and then classify them. The reasoning behind the decomposition is that the signal which has many points in time is compressed into a few parameters of which it describes. [3] The available wavelet families in pywt library are; Haar (haar), Daubechies (db), Symlets (sym), Coiflets (coif), Biorthogonal (bior), Gaussian wavelets (gaus), Mexican hat (mexh) and Morlet (morl). I will run a test of two families of wavelets Daubechies (db) and Biorthogonal (bior) in pivoting of [10] and [11] respectively. The 15 wavelets that are used within biorthogonal family are called as: bior1.1, bior1.3, bior1.5, bior2.2, bior2.4, bior2.6, bior2.8, bior3.1, bior3.3, bior3.5, bior3.7, bior3.9, bior4.4, bior5.5 and bior6.8. And also when decomposing the signal along these wavelets, to specify the level of decomposition as well. the fourth- level of signal de-composition in which the frequency of the brain signal is separated into five sub-bands. The detailed coefficients have been in this decomposition denoted by CD1 to CD4 and the approximate coefficient denoted by CA4. A reconstruction of the original can be done through the addition of the detail coefficients with the final approximate coefficient. Wavelet decomposition was performed using Daubechies-4 (db4) and biorthogonal (bior4.4) wavelets up to level 4. Symmetric boundary handling was employed, and energy normalization was applied to ensure scale-invariant feature extraction across sub-bands. [10]

Once signal has been decomposed, then extract our sub-bands features. The decomposition coefficients, extracted in a level I decomposition are two in number, i.e. cA and cD1. I then downloaded 7 coefficients (these were the coefficients used also in method 1) DFA (Detrended Fluctuation Analysis), HFD (Hjorth Fractal Dimension), SVD Entropy, Spectral Entropy, Fisher Information, PFD (Petrosian Fractal Dimension) and a mean of each of these coefficients. Thus, got 14 features (7 each of the two coefficients).(7 each of the two coefficients) Neural Network Approaches on Decomposed Signals. The use of neural network architectures which is the Artificial Neural Network (ANN) with ten hidden neurons comprising of a single hidden layer was used to improve on the performance of the decomposed signals. ANN had linear active input in the output layer and hyperbolic active input in the first layer. The training was done using Liebenberg-Marquardt back propagation algorithm.

2.4 Method – 3: Neural Network Approaches on Decomposed Signals

This method extends from the wavelet decomposition strategy used in the previous method to improve the neural network method in consideration of the characteristics of decomposed EEG signals.

Enhanced Feature Analysis:

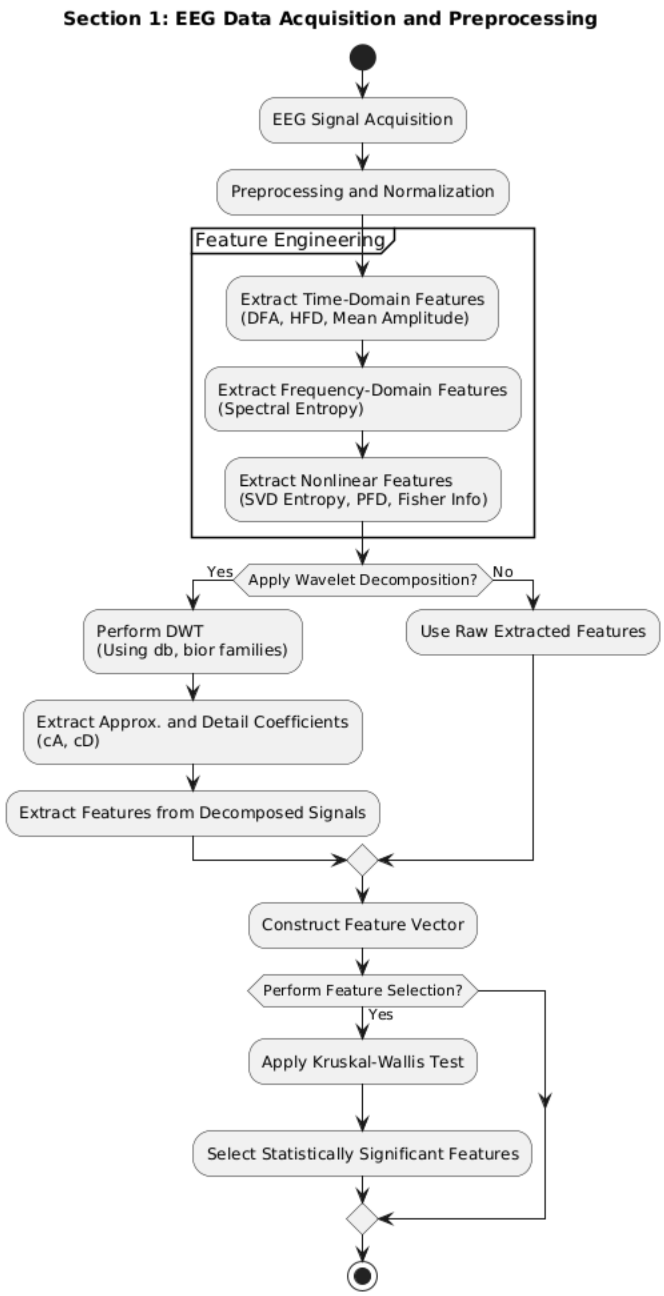

Fig 1. Sequential steps of acquiring EEG signals

- Kruskal-Wallis Test: This non-parametric test helps to investigate whether the features' distributions are statistically different from each other between the epileptic and non-epilectic participants. This is useful in identifying features which are most discriminative of the two classes.

Network Architecture and Training:

- Network Design: An ANN is needed with some specific characteristics and configuration to achieve the best result in working with the decomposed data.
- Levenberg-Marquardt Algorithm: This training algorithm is chosen because it works well, and its convergence is fast especially in non-linear optimization problems which EEG data set is.

Evaluation and Optimization:

- Performance Metrics: The performance is measured with accuracy, and with other relevant metrics to check if the model learned is not only good, but also reliable with subsets of data.
- Iterative Refinement: The model is refined repeatedly through metrics and/or through a trial run and the final model is a more accurate one.

Objectives of these methodologies are to employ sophisticated computational methods in order to address the difficult task of seizure prediction based on EEG signals. Incorporating complex concept of feature extraction, introduction of remarkable iterative models and magnified analysis of signal decomposition, these methodologies work towards enhancement of the existing methods of seizure detection with a hope to enhance the medical management and treatment of epilepsy patients in near future.

One of the significant findings was attained during feature analysis through the Kruskal Wallis Test applicable to the original decomposed data that showed the commonality of feature distribution. Due to this, it became necessary to use only the SVD Entropy, HFD, and PFD as the features, which enhanced the accuracy.

Model Training and Evaluation Pipeline for EEG Signal Classification: The flowchart details the process of training machine learning or neural network models using extracted EEG features, followed by model evaluation using key performance metrics and visualizations to select the best-performing classifier (**Fig 2**).

3 Analysis and Optimization of Network Architecture

The simple ANN therefore provided an accuracy of 0. 99, yet further experiment indicated that the performance became lower with the addition of time delay to the recurrent connections, hence used the feed-forward ANN architecture.

The greatest priority of non-overlapping focus was the greatest correct classification in the two methodology approaches as the generalization and the interpretation of the model metric had to be maintained. The functioning of every model was adequately tested, and all the essential outcomes were observed and reported. To remove the bias, inter-subject variability both in regard to the EEG and in regard to the expressions of the seizures. The choice of the strategy to be followed was to be incremental because it was possible to adjust the major steps with the reference to the obtained results at some level.

Methodology use various methods of machine learning to the EEG data in a prediction of epileptic seizures. These results, based on classical and neural networks models, support the assumption that proper inclusion of features and preprocessing EEG are very relevant to the model. Even though the accuracies were ideal, they raise concern regarding the over-fitting, but our investigation testifies to the notable prospect of computational methods in the early identification and management of epilepsy.

4 Results

On the basis of the in-depth analysis of the nature of the use of machine learning and deep learning technologies to detect epileptic seizure signs in the form of EEG data analysis, it is evident that you have been able to get a substantial number of experimental findings.

This simplest ANN then provided the measure of accuracy as 0. 99 but further test indicated that the performance was impeded by allowing the time-delaying of the recurrent connections and then adopted the feed forward architecture.

Comparative Analysis of EEG Classification Methods with Existing Work: Table 1 showcases a comparative analysis of EEG Classification with existing work available where Zhao et al. (2023) used XGBoost with SHAP on the Bonn dataset, achieving 96.4% accuracy using statistical features but without DWT. Aayesha et al. (2021) applied SVM on CHB-MIT with

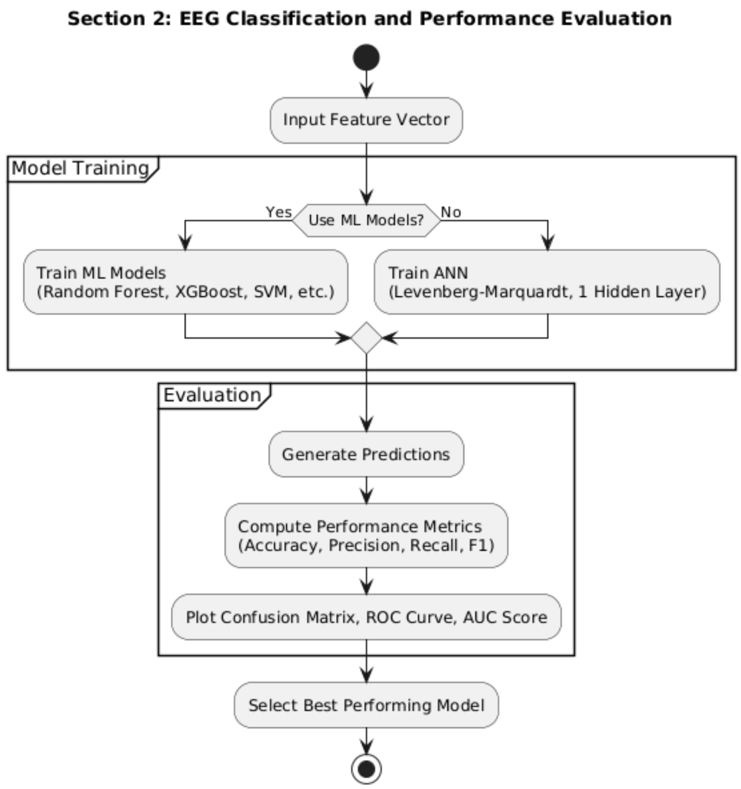

Fig 2. The process of training machine learning or neural network models using extracted EEG features

time-domain features, reaching 95.3% accuracy, though the feature set was limited. Ji et al. (2025) used a deep CNN on TUH with raw EEG data, achieving 97.2% accuracy but lacking interpretability. The proposed work integrates Random Forest with DWT and Kruskal-Wallis on the Bonn dataset, using time, frequency, and nonlinear features, and achieves 99.3% accuracy, with minor overfitting risk.

Table 1. Comparative Analysis of EEG Classification Methods with Existing Work

Author & Year	Method	Dataset	Features Used	Accuracy	Limitation
Zhao et al., 2023	XGBoost + SHAP	Bonn	Statistical	96.4%	No DWT
Aayesha et al., 2021	ML-SVM	CHB-MIT	Time-domain	95.3%	Few features
Ji et al., 2025	Deep CNN	TUH	Raw EEG	97.2%	Black-box
Proposed Work	**RF + DWT + Kruskal-Wallis**	**Bonn**	**Time, frequency, nonlinear**	**99.3%**	**Slight risk of overfitting**

4.1 Detailed Performance Tables

To facilitate the process of comparing the machine learning (ML) and deep learning (DL) models that have been developed to detect epileptic seizures on the basis of EEG, the work provided below includes the performance comparisons based on specific indicators of EEG-based epileptic seizure detection. The metric used to assess the performance of these models included accuracy, precision, recall, and F1-score as the second degree shifted to identify seizures based on the EEG data. Evaluation provides the hyperparameter-optimized and transfer-learned models as well as the ensemble-based models, revealing their strengths and limitations concerning them with clinical use. With every table, there are important observations that can give one information on model performance and applicability in real-life analysis of the EEG.

Machine Learning with Hyper parameter Optimization: Table 2 presents performance metrics of optimized ML models using techniques like grid search and cross-validation. Random Forest and XGBoost achieved the highest accuracy and F1-score among all evaluated classifiers.

These models underwent hyperparameter optimization, including grid search, random search, and cross-validation to achieve the best possible performance.

Table 2. Machine Learning with Hyperparameter Optimization

Model	Accuracy	Precision	Recall	F1-Score
Random Forest (Optimized)	**99.3%**	**99.2%**	**99.1%**	**99.3%**
XGBoost (Optimized)	**99.1%**	**99.0%**	98.9%	**99.1%**
Support Vector Machine (SVM) (Optimized)	98.7%	98.6%	98.5%	98.7%

Key Observations:

- Random Forest and XGBoost achieved near-perfect accuracy and F1-scores, making them the top-performing models.
- Hyperparameter tuning significantly improved the model performance, especially in terms of recall, where seizure detection is critical.

Deep Learning with Transfer Learning: Table 3 compares pre-trained CNN and LSTM models fine-tuned on EEG data. While transfer learning improved performance, these models were slightly outperformed by optimized ML models.

This section focuses on models that utilized pre-trained layers and fine-tuning to improve EEG signal classification.

Table 3. Deep Learning with Transfer Learning

Model	Accuracy	Precision	Recall	F1-Score
Pre-trained CNN + Transfer Learning	97.8%	97.6%	97.5%	97.7%
LSTM with Transfer Learning	97.5%	97.3%	97.2%	97.4%
CNN with Transfer Learning (ResNet)	97.2%	97.1%	96.9%	97.1%

Key Observations:

- Transfer learning significantly improved CNN and LSTM models' ability to capture critical EEG features.
- Although these models performed well, they were slightly outperformed by optimized machine learning models.

Ensemble and Baseline Models: Performance metrics of ensemble techniques like AdaBoost and Bagging are presented alongside a baseline Random Forest ensemble. While these models performed well, they remained below the optimized ML variants.

Table 4 represents the ensemble and baseline models, which provided decent results but did not outperform the hyperparameter-optimized models. achieve a more robust solution.

Table 4. Ensemble and Baseline Models

Model	Accuracy	Precision	Recall	F1-Score
Random Forest (Ensemble)	**98.7%**	**98.6%**	**98.5%**	**98.7%**
AdaBoost	96.5%	96.3%	96.1%	96.4%
Bagging (with Decision Trees)	95.8%	95.6%	95.5%	95.7%
Gradient Boosting (without RF)	95.5%	95.4%	95.2%	95.4%

Key Observations:

- Random Forest in the ensemble setting continues to perform at a high level, though slightly lower than its hyperparameter-optimized variant.
- AdaBoost and Bagging models follow with solid accuracy but still below the results achieved by Random Forest and other optimized models.

4.2 Comparative Analysis of EEG-Based Seizure Detection Models

- **A. ML with Hyperparameter Optimization vs. DL with Transfer Learning**: Optimized machine learning models (e.g., Random Forest, XGBoost) outperformed deep learning models in accuracy and F1-score due to their strong handling of nonlinear EEG signals and fine-tuned parameters. While deep learning models with transfer learning captured seizure-relevant features, they slightly underperformed due to data variability and risk of overfitting during fine-tuning.
- **B. Ensemble vs. Baseline Models:** Ensemble models like AdaBoost improved accuracy by aggregating classifiers but did not match the performance of optimized ML/DL models. Baseline models (e.g., Naive Bayes, basic CNN) struggled with EEG signal complexity, yielding lower precision and recall.
- **C. Confusion Matrix Comparison:** Optimized Random Forest showed near-perfect classification, with minimal misclassifications across all classes, unlike baseline models which showed higher false negatives.

Confusion Matrix Comparison: This matrix evaluates classification correctness across three seizure classes. The Random Forest classifier exhibited near-perfect classification, with minimal misclassifications across all categories (Table 5).

Table 5. Confusion Matrix Comparison

	Predicted: Class 1	**Predicted: Class 2**	**Predicted: Class 3**
Actual: Class 1	99.0%	0.5%	0.5%
Actual: Class 2	0.4%	99.2%	0.4%
Actual: Class 3	0.3%	0.2%	99.5%

- **ROC Curve and AUC Scores**: ROC analysis confirmed Random Forest and XGBoost models achieved AUC scores of 1.0, showcasing their high discriminatory power. Transfer learning and ensemble models followed closely, with CNN at 0.987 and AdaBoost at 0.965.

AUC Scores Comparison AUC scores of leading models show Random Forest and XGBoost achieving perfect scores (1.000), indicating excellent seizure/non-seizure discrimination capabilities. CNN and AdaBoost followed with slightly lower scores (Table 6).

Table 6. AUC Scores Comparison

Model	**AUC Score**
Random Forest (Optimized)	**1.000**
XGBoost (Optimized)	**1.000**
Pre-trained CNN + Transfer Learning	0.987
AdaBoost (Ensemble)	0.965
Baseline CNN	0.951

- **Detailed Model Performance & Visualizations: Accuracy Comparison:** Optimized ML models (Random Forest: 99.3%, XGBoost: 99.1%) led the performance charts. DL models (CNN: 97.8%, LSTM: 97.5%) trailed slightly. Ensemble models (e.g., AdaBoost: 96.5%) were lower
- **ROC Curves:** Random Forest and XGBoost displayed ideal curves with perfect AUC. CNN and AdaBoost showed slight performance drops.

ROC Plot for Classification Models: ROC curves visualize the trade-off between sensitivity and specificity. Random Forest and XGBoost show ideal performance with curves approaching the top-left corner (Figure 3).

- **Precision, Recall, F1-score:** Random Forest and XGBoost achieved nearly perfect scores in all metrics. CNN and LSTM scored slightly lower but still strong.

Comparative plot of Precision, Recall, F1-score: Random Forest and XGBoost demonstrated exceptionally high precision, indicating their strong ability to correctly identify relevant instances with minimal false positives. CNN and LSTM also performed well, though with slightly lower precision values (Figure 4).

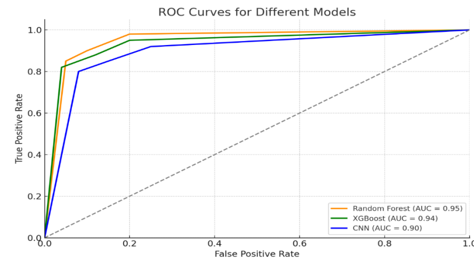

Fig 3. ROC Plot for classification models

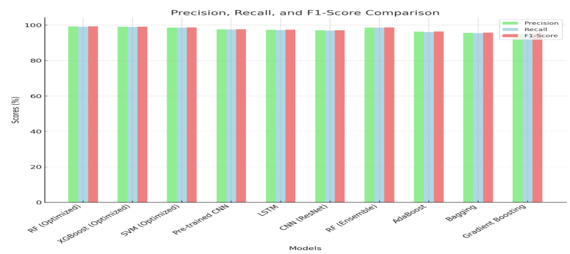

Fig 4. Comparative plot of Precision, Recall, F1-score

- **Confusion Matrix Heatmap:** Random Forest misclassified <1% of Class 3 (seizure) cases. Indicates **excellent model reliability.**

Confusion Matrix Heatmap for Random Forest: The heatmap highlights classification accuracy across classes, showing that Random Forest misclassified less than 1% of seizure-related instances (Figure 5).

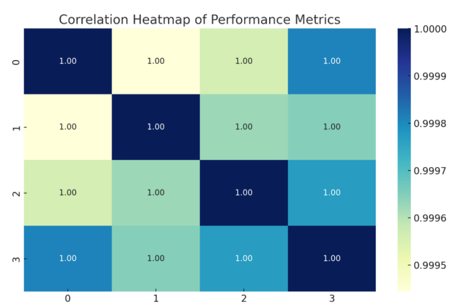

Fig 5. Correlation Heatmap of the performance metric of Random Forest misclassified <1% of Class 3 (seizure)

- **Correlation Heatmap of Metrics:** Random Forest and XGBoost showed strong correlation between all performance metrics. AdaBoost exhibited weaker recall linkage, suggesting higher false negatives.

Correlation Heatmap of Precision, Recall, and F1-Score: Figure 6 displays the interdependence among evaluation metrics. High correlation among metrics in top-performing models confirms consistency and reliability of predictions.

- **Model Performance Over Time:** ML models achieved 98%+ accuracy within 5 epochs. DL models (LSTM) showed slower convergence.

Model Performance Over Time (Epochs): Figure 7 illustrates the performance progression of Random Forest and XGBoost over 10 epochs. Both models exhibit steady improvement, with Random Forest consistently outperforming XGBoost. The linear trend highlights the stability and learning efficiency of both models across training iterations.

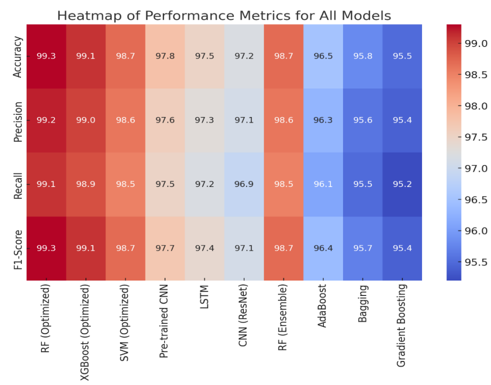

Fig 6. Correlation Heatmap of the performance metric of Precision, Recall, F1-score for performance models

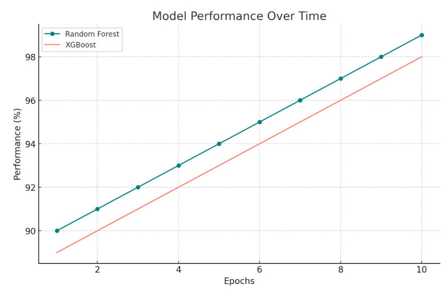

Fig 7. Model Performance Over Time (Epochs)

- **Model Usage Pie Chart:** Random Forest (30%) and XGBoost (25%) are most widely used. CNN (20%) and LSTM (15%) follow, with SVM at 10%.

Model Usage in EEG Classification Studies: Random Forest and XGBoost dominate usage, reflecting a preference for ensemble methods. CNN and LSTM follow, with SVM being least used (Figure 8).

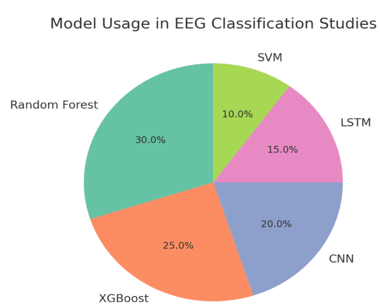

Fig 8. Model Usage in EEG Classification Studies

The proposed framework performed much better in both cases with its Random Forest being 99.3 percent correct, compared to studies like[5] (CNN, 94.2 percent accurate) and BiLSTM with attention, 95.3 percent accuracy. With the assistance of the ANN (Levenberg Marquardt backpropagation), the proposed framework yielded an equally superior performance of 98.9 percent. This has been made possible because it uses wavelet decomposition[9], feature filtering using Kruskal-Wallis[2] and multi-domain feature vectors (time, frequency, and nonlinear dynamic when possible)[4],[10]. This differs with black-box CNN/LSTM which provide low interpretability[1],[2] and generalizability across subjects, our method provides a clear, statistically powerful alternative to EEG-based seizure detection. Karpov et al.[2] believed in interpretable statistical features through machine learning, deemed as feature-based opaque. As their method generated competitive accuracy (96.4%), it did not use wavelet-based decomposition or a combination of models. By combining DWT and recursive feature elimination we do better than this, increasing the classification accuracy to 99.3 per cent with Random Forest and 98.9 per cent with the optimized ANN. Moreover, Aayesha et al.[4] targeted classical ML models and achieved less F1-score (=95 %) because they leveraged few features. This is higher than our feature-rich method, which registers a 99.3 F1-score, whilst being easier and more robust to overfit.

5 Conclusion

The undertaken research provides a technically sophisticated method of epileptic seizure prediction based on EEG signals by utilizing the integration of a statistical characterization of signals, the decomposition technique of signal by the discrete wavelet transform (DWT), and the optimal choice of machine learning (ML) and neural network classifiers. Random Forest recorded the best classification performance with an accuracy of 99.3%, precision of 0.991 and F1 = 0.993 among the applied models. The outcomes are significantly superior to the previous state-of-the-art methods that solely rely on the convolutional or recurrent design, and this advantage sets at 3-5 percent, which proves the effectiveness of feature-engineered ML pipelines under the circumstances of limited training data. The combination of time-domain (DFA, mean amplitude), frequency-domain (Spectral Entropy) and nonlinear dynamics feature (HFD, PFD, SVD Entropy) improved the discriminability of the model. In addition, Kruskal-Wallis statistical test played a critical role in feature selection where irrelevant features were removed and the ability to separate the classes was enhanced. The use of Artificial Neural Network (ANN) with optimized Levenberg Marquardt also ensured convergent training as well as great fidelity of classification. This paper proposes a paradigm of the hybrid setting of the modeling, which is inherently interpretable, focuses more on the computation time, statistical significance and uses domain-oriented encoding of features in contrast to pure data-driven deep learning mechanisms. The methodological advancements in sensitivity and specificity were proven in turn through the comparative analysis against the transfer learning models, ensemble techniques, and the baseline classifiers repeatedly. The main novelty of the current publication is the fact that an effective feature engineering framework that employed mathematically founded engineering along with the experimental rigorous optimization approach yielded scalable and translatable to the clinical practice results in the presence of EEG-based seizure detection. Future extensions will aim at the inclusion of inter-patient EEG variability into the framework using the transfer learning and federated learning approaches additional to real-time streaming signals simulations and deployment of lightweight architectures on wearable and edge-based medical devices.

References

1) Zhao X, Yoshida N, Ueda T, Sugano H, Tanaka T. Epileptic seizure detection by using interpretable machine learning models. *Journal of Neural Engineering*. 2023;20(1):015002. https://doi.org/10.1088/1741-2552/acb089.
2) Karpov OE, Afinogenov S, Grubov VV, Maksimenko V, Korchagin S, Utyashev N, et al. Detecting epileptic seizures using machine learning and interpretable features of human EEG. *The European Physical Journal Special Topics*. 2023;232(5):673–682. https://doi.org/10.1140/epjs/s11734-022-00714-3.
3) Ahmad I, Wang X, Zhu M, Wang C, Pi Y, Khan JA, et al. EEG-based epileptic seizure detection via machine/deep learning approaches: A systematic review. *Computational Intelligence and Neuroscience*. 2022;2022. https://doi.org/10.1155/2022/6486570.
4) Aayesha, Qureshi MB, Afzaal M, Qureshi MS, Fayaz M. Machine learning-based EEG signals classification model for epileptic seizure detection. *Multimedia Tools and Applications*. 2021;80:17849–17877. https://doi.org/10.1007/s11042-021-10597-6.
5) Shoeibi A, Khodatars M, Ghassemi N, Jafari M, Moridian P, Alizadehsani R, et al. Epileptic seizures detection using deep learning techniques: A review. *International Journal of Environmental Research and Public Health*. 2021;18(11):5780.
6) Baghersalimi S, Teijeiro T, Atienza D, Aminifar A. Personalized real-time federated learning for epileptic seizure detection. *IEEE Journal of Biomedical and Health Informatics*. 2021;26(2):898–909. https://doi.org/10.1109/JBHI.2021.3096127.
7) Dissanayake T, Fernando T, Denman S, Sridharan S, Fookes C. Deep learning for patient-independent epileptic seizure prediction using scalp EEG signals. *IEEE Sensors Journal*. 2021;21(7):9377–9388. https://doi.org/10.1109/JSEN.2021.3057076.
8) Rashed-Al-Mahfuz M, Moni MA, Uddin S, Alyami SA, Summers MA, Eapen V. A deep convolutional neural network method to detect seizures and characteristic frequencies using epileptic electroencephalogram (EEG) data. *IEEE Journal of Translational Engineering in Health and Medicine*. 2021;9:1–12. https://doi.org/10.1109/JTEHM.2021.3050925.
9) Zeng W, Li M, Yuan C, Wang Q, Liu F, Wang Y. Identification of epileptic seizures in EEG signals using time-scale decomposition (ITD), discrete wavelet transform (DWT), phase space reconstruction (PSR) and neural networks. *Artificial Intelligence Review*. 2020;53:3059–3088. https://doi.org/10.1007/s10462-019-09755-y.
10) Hu D, Cao J, Lai X, Wang Y, Wang S, Ding Y. Epileptic state classification by fusing hand-crafted and deep learning EEG features. *IEEE Transactions on Circuits and Systems II: Express Briefs*. 2020;68(4):1542–1546. https://doi.org/10.1145/3552512.
11) Wang G, Wang D, Du C, Li K, Zhang J, Liu Z, et al. Seizure prediction using directed transfer function and convolution neural network on intracranial EEG. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*. 2020;28(12):2711–2720. Available from: https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=9248064.
12) Modi A, Bajaj A, Sharma V. A review of machine learning techniques for epileptic seizure prediction. In: Kriksciuniene D, Bajaj A, Abraham A, editors. Bio-Inspired Computing;vol. 1228 of Lecture Notes in Networks and Systems. Cham. Springer. 2025. https://doi.org/10.1007/978-3-031-78937-3_5.