Design and Analysis of Explainable AI-Driven Epileptic Seizure Detection Using Machine Learning Models on the Bonn EEG Dataset

[1]Mukesh Kumar Bhardwaj [2]Avnish Shukla [3]Radhika Sharma [4*]Kumud Dixit [5]Methily Johri
[1,4]D.S College, Aligarh, U.P.,India-202001
[2,3]Department of IT,Jagannath International Management School, New Delhi,India-110085
[5]School of computer Science and Engineering, Greater Noida,Galgotias university India- 201206
[1]mkb.dsbot@gmail.com, [2]av.sh.927927@gmail.com, [3] raddhika1709@gmail.com,
[4] *Kumuddixit30@gmail.com, [5] methily.johri@gmail.com

Abstract

The problem of epileptic seizure detection is still a significant issue in clinical neuroscience because the impact of the seizure episode is unpredictable, and the underlying neural structure is complicated. The conventional diagnostic models are very reliant on the experience of the specialists in interpreting the electroencephalogram (EEG) signals, which is a time consuming and subject to human error. This paper will suggest an explainable AI-based system of automated epileptic seizures detection on the basis of the Bonn EEG dataset, focusing on the predictive performance of the machine learning models and their interpretability. The methodology involves a powerful preprocessing phase includes denoising and segmentation then statistical, temporal, and frequency-domain data are extracted. Different machine learning classifiers like the Random Forest, Support Vector Machine, Gradient Boosting and the Logistic Regression are trained and optimized to differentiate between healthy, interictal and ictal EEG samples. To achieve transparency in model decisions explainable AI methods such as SHAP (Shapley Additive Explanations) and LIME (Local Interpretable Model-agnostic Explanations) are combined to determine the most significant EEG characteristics that affect the classification results. Experimental findings show that the ensemble-based models can provide high accuracy, sensitivity and specificity whereas the XAI visualizations can give clinically significant information about the feature importance like the dominance frequency bands, signal variance and entropy measurements related to seizure activity. The results indicate the twofold advantage of integrating machine learning functionality and interpretability, which is aimed at building reliable clinical decision-support systems. This study helps to enhance clear and credible AI procedures of diagnosing neurological disorders and supports the possibilities of XAI-based EEG analysis in practice.

Keywords: *Epileptic seizure detection, EEG classification, Bonn EEG dataset, machine learning, gradient boosting, random forest, feature extraction, frequency-domain analysis, nonlinear features, SMOTE balancing, explainable AI, SHAP, LIME, model interpretability, biomedical signal processing*

Introduction

Epilepsy is a long-term neurological disease that is manifested by frequent and irregular seizures caused by unusual electrical discharge in the brain [7]. It impacts on millions of people worldwide and it frequently causes cognitive, psychological, and social problems that influence the quality of life [3]. Early and precise seizure detection is critical in clinical diagnosis, customization of treatment and long-term monitoring of patients [11]. Conventionally, to determine the patterns of seizure, neurologists use manual examination of electroencephalogram (EEG) records, which is both time-consuming and expertise-sensitive, and subject to human errors because of the non stationary and highly heterogeneous nature of EEG signals [1, 9]. This has necessitated the development of automated and trustworthy computational tools that are able to process EEG data effectively.

P. Johri et al. (eds.), *Proceedings of the International Conference on Sustainable Computing and Artificial Intelligence (ICSCAI 2025)*, Advances in Engineering Research 298,
https://doi.org/10.2991/978-94-6239-674-6_13

Machine learning has proved to be an effective method of solving these problems by training discriminant features on EEG data and aiding in automatic seizure detection [5, 14]. A number of algorithms such as Support Vector Machines, random forest, logistic regression and gradient boosting have demonstrated high predictive accuracy when used on benchmark data sets [12]. The Bonn EEG dataset is a popular dataset in the field of epilepsy research that provides clean and highly organized classes of healthy, interictal, and ictal EEG samples which make the dataset a perfect fit to train and test seizure-detection models [8]. It has consistently been shown that significant machine learning pipelines over this data can attain high accuracy, through the use of statistical, temporal, and frequency-based features derived on EEGs [4, 15]. Nevertheless, the quality of the accuracy is not the sole requirement of the contemporary clinical practice.

It is becoming acknowledged that machine learning models should not only be effective, but also readable and comprehensible in their determination of predictions [6]. It has given way to the emergence of Explainable Artificial Intelligence (XAI), an approach to describe evidence of how model decisions are made, in a way understandable to humans. This clarity is necessary to clinicians since medical choices have serious repercussions and black-box predictions are hard to rely on without understanding what EEG characteristics affect the result [10, 13]. SHAP (Shapley Additive Explanations) and LIME (Local Interpretable Model-agnostic Explanations) are methods to bring in light on the role of individual frequency components, amplitude variations, entropy, and temporal variances at the point of a seizure classification [2]. This interpretability helps the acceptance of the clinical and scientific validation.

Although there is an improvement in the development of algorithms, a number of limitations remain. A lot of current research is strongly concentrated on increasing the accuracy of classification and leaves no emphasis on transparency and interpretability that are essential to medical reliability [4]. Deep learning algorithms may need huge amounts of training data, but EEG data is usually small and it is prone to noise and artifact [7]. Moreover, the problem of class imbalance, in particular, between seizure and non-seizure classes, decrease the robustness of the model and can introduce bias in sensitivity to the dominant ones [9]. These gaps point to the necessity of having a balanced framework, which will combine good performance with clear explanations that have a clinical significance.

The current paper suggests an elucidable AI-based platform of epileptic seizure detectors under machine learning models and Bonn EEG data. This framework consists of preprocessing and noise reduction, statistical and frequency-domain features extraction, training of various machine learning classifiers, and deploying XAI frameworks including SHAP and LIME to explain model behavior. The goal of this approach is to create systems that are extremely accurate with additional insights provided by the means of providing clear information on what features of EEG are used in making a decision in regard to seizures.

On the whole, the research study represents a contribution to the general trend of understandable and credible AI in the healthcare sector as it shows that machine learning could be combined with explainable AI to produce clinical, transparent, and trustworthy seizure-detection devices [11]. This article reinforces the premise of using AI-based diagnostic instruments in practice in neurology.

Related Works

Studies of automated epileptic seizure detection have developed across a variety of key themes, including classical machine learning methods, deep learning, hybrid systems, and interpretability-based systems. The initial research was on feature engineering and classifier accuracy, whereas recent research has moved to transparency, robustness, and clinical usability. A significant fraction of literature uses deep learning architectures to do end-to-end seizure detection, where convolutional and recurrent neural networks learn the discriminative EEG patterns in an automatic manner. Deep convolutional networks have shown good results in detecting ictal signatures directly on raw or minimally processed EEG signals with high sensitivity in a number of EEG datasets [2, 10, 14, 27, 30]. These papers demonstrate the benefit of hierarchical feature extraction compared to hand-crafted feature engineering because deep networks are capable of more effectively capturing frequency variations, waveform distortions and non-linear temporal relationships. Subsequent progress was made with sequence models like LSTM networks that model longer-term temporal correlations in EEG data, with superior capabilities of predicting and detecting seizures early [25]. Some more recent studies were able to apply deep learning to intracranial EEG, and these studies claimed high accuracy with focal seizure detection because the signal-to-noise ratio and localized seizure onset characteristics were better [12].

In complement to all these end-to-end deep learning solutions, literature on the use of hybrid or multi-view frameworks have become more prominent. Multi-view learning was utilized, which was a combination of spatial, temporal and frequency-domain EEG perspectives so that seizure events could be better represented [30]. The model channel-specific classification and local feature learning led to the advanced multi channel analysis and the better detection of the seizure type [26, 29]. The presence of hybrid models that combine deep and conventional learning elements enhanced the robustness particularly when using small or noisy EEG samples [22].

Simultaneously, machine-learning-based methods combining hand-engineered features still provide a competitive performance, in particular, regarding computational efficiency and interpretability. Wavelet transform-based, statistical measures, entropy features, as well as spectral parameter approaches have demonstrated reliable reliability in seizure classification on benchmark datasets [3, 4, 9]. Stable classical classifiers like SVM, Random Forest, k-NN and Gradient Boosting also worked well, especially when combined with powerful feature engineering pipelines [9, 11, 13, 15]. The models have the advantage of reduced training complexities as well as being useful in clinical or low-resource environments. The performance of these traditional models has been further enhanced by the use of high- level preprocessing like noise reduction, artifact removal and frequency band separation [20].

Multiple systematic reviews summed up the status of EEG-based seizure detection and unanimously identified the wide applicability of both deep learning and machine learning technologies and methodological discrepancies like varying sets of features, dissimilar outcomes measures, and dataset constraints [16, 17, 19, 21]. Such reviews indicate inconsistencies in generalizability, reproducibility, and transparency of existing studies.

One of the biggest weaknesses observed in deep learning methodologies is the propensity to act as a black box. The majority of the high-performing deep models fail to give information about the decision-making process and EEG features that have the strongest impact on decision-making of seizure. Such interpretability prevents clinical uptake unless neurologists are assured of clear thought processes before they believe automated diagnostic systems. Other new endeavors dealt with this dilemma using explainability techniques. They are the application of model-agnostic interpretability techniques like SHAP, LIME, CAM-based visualizations or gradient-based attribution maps to point to important areas of time or frequency in EEG signals [6, 7, 28]. Such studies show that it is possible to achieve high-accuracy classification with clinically meaningful interpretability, but the amount of such studies is still significantly lower than the accuracy-driven models. The other problem that is frequently mentioned in the previous literature is associated with the characteristics of the dataset. Numerous deep learning researches depend on large data sets, but the popular datasets like Bonn include small samples. Papers that rely on the Bonn data typically address this by either data augmentation or transfer learning or by devising customized features to ensure the preservation of classification accuracy [1, 3, 9, 14]. Intracranial datasets have been studied as they provide improved performance because the signal is clearer, however it is not necessarily reflective of normal clinical records of EEG [12]. As per reviews, model evaluation is frequently not a very standardized process, with accuracy as the sole that is frequently reported, and robustness metrics, including sensitivity, false alarm rate, or specificity, being underreported [16, 19]. Another significant issue is class imbalance since the occurrence of seizures is a minor part of the overall EEG recording. A number of studies deal with this by resorting to resampling techniques, statistical balancing techniques, or cost-sensitive learning algorithms to control biased learning [3, 9, 15]. Nevertheless, even with those, much of the available models continue to have problems in stabilizing their performance when exposed to a variety of recording environments or subjects, meaning that more robust forms are required. According to the recent studies, there are statistical methods of predicting and forecasting seizures, which predict seizures before they occur. Such models usually establish time forecasting strategies and deep recurrent networks [25, 27]. Despite being promising, prediction tasks demand continuous monitoring of EEG and require a huge annotated dataset, which is hard to get in practice. As a result, detection of seizures as opposed to their prediction is still a prevailing topic in the literature.

In general, the literature demonstrates the high level of the advancement of the precision of automated epilepsy detection systems, but there are considerable gaps that prevent their application to real-life settings in health care. The initial significant weakness is that most high-performing models are not interpretable. Deep learning architectures are highly accurate, but do so with little insight into what EEG characteristics made such decisions, which makes them less trusted by clinicians and limits their ability to be deployed into practice. A very limited number of studies have tried to combine explainability, with most of them using intracranial EEG as opposed to surface EEG datasets that are common [6, 7]. The second gap is associated with the inconsistency in the evaluation, performance is measured based on different metrics, and it is hard to compare them. The third gap is the excessive dependence of either deep learning or feature-based ML, or other works have not yet examined an opportunity with an integrated approach between the other two that provides good accuracy and simultaneously remains interpretable. A fourth gap is the fact that standardized datasets have not been intensively used as XAI research; the Bonn dataset is not being fully used in explainability research in spite of its structured classes segregation.

All these limitations demonstrate the necessity of a framework that combines machine learning classifiers and explainable AI tools and is used on a large dataset that has been accepted across the board, e.g., the Bonn EEG dataset. To fill this gap, the current study creates a clear, understandable, and performance-based seizure detection system that applies SHAP and LIME to determine features in the EEG that are influential and keep its classification performance high by using optimized machine learning networks.

Proposed Methodology

The suggested methodology presents an elucidable machine-learned framework of automated epileptic seizure recognition using the EEG signals on Bonn dataset. The design incorporates an entire workflow starting with characterizing and preprocessing datasets to extracting features, training the model, evaluating it, and analyzing the interpretation of its interpretation with XAI methods. The idea is to form a clear and convincing method of seizure-classification that does not only achieve high accuracy but also explains what EEG features make the strongest contributions to predictions. This combined approach can be seen to overcome the weaknesses of the previous research in which they focus more on accuracy and less on interpretability or are overly dependent on black-box architectures and do not have clinical transparency.

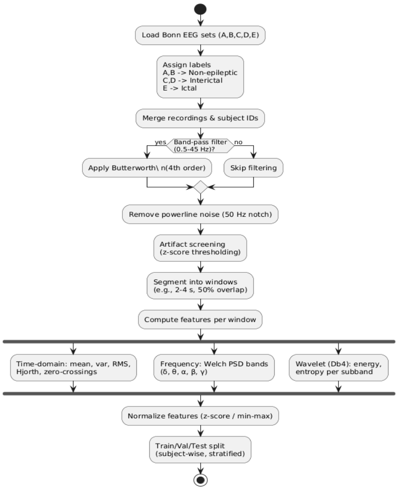

Figure 1. Bonn EEG Dataset Handling and Pre Processing

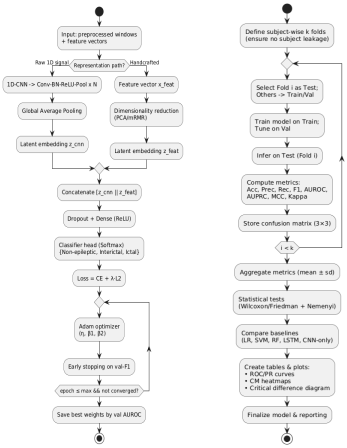

Figure 2. Training and Cross Validation Framework

Dataset Description and Characteristics

This study uses the **Bonn University EEG dataset**, a widely recognized benchmark for epileptic seizure detection research. The database contains **five subsets (A–E)**, each including **100 EEG segments** of **23.6**

seconds duration, sampled at **173.61 Hz** with **4096 samples per recording**. The subsets represent distinct physiological conditions, enabling controlled classification:

- **Set A**: Healthy subjects with eyes open
- **Set B**: Healthy subjects with eyes closed
- **Set C**: Interictal activity from epileptic patients (opposite hemisphere)
- **Set D**: Interictal activity near the epileptogenic zone
- **Set E**: Ictal seizure activity

The data is classified into two in this study Non-seizure (A, B, C, D) and Seizure (E). This binary division coincides with the clinical screening tasks in the real-world and simplifies the learning task. All recordings are normalised and clustered to make them uniform before instructing the machine-learning pipeline. The high-frequency noise or baseline artifacts are reduced by a band-pass filter at 0.5-60 Hz without interfering with clinically-relevant EEG rhythms (delta, theta, alpha, beta, and low-gamma).

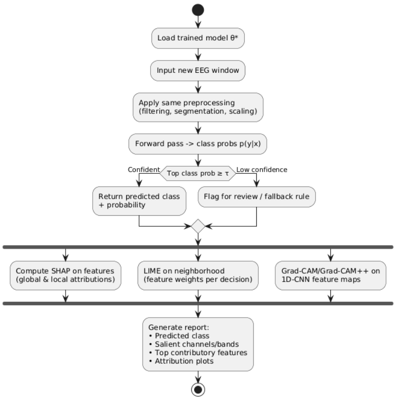

Figure 3. Overview of Proposed Methodology

Preprocessing and Mathematical Formulation

EEG data undergoes several transformations to improve model consistency. After filtering, each EEG segment is normalized using **z-score normalization**, defined as:

$$X_{norm} = \frac{X - \mu}{\sigma}$$

where X is the sample vector, μ is mean, and σ is standard deviation. Normalization ensures that no feature disproportionately influences the classifier. The filtered and normalized signals are then subjected to **windowing** using overlapping sliding windows of 512 samples. Let $W_i(t)$ represent the i^{th} window at time t:

$$W_i(t) = X(t{:}t + L - 1)$$

where $L = 512$. This increases the number of training samples and captures local EEG characteristics.

Class imbalance occurs because non-seizure sets (A–D) contain four times more samples than ictal set E. To address this, **SMOTE (Synthetic Minority Oversampling Technique)** is applied:

$$x_{new} = x + \lambda(x_{nn} - x)$$

where x is a minority-class instance, x_{nn} is a nearest neighbor, and $0 < \lambda < 1$. SMOTE preserves natural signal behavior while expanding the minority class.

Feature Extraction Strategy

A combined **time-domain** and **frequency-domain** feature extraction approach ensures that the classifier receives rich and physiologically meaningful information. The extracted features include:

Time-Domain Features

- Mean, median, standard deviation
- Skewness and kurtosis
- Hjorth parameters: activity, mobility, complexity
- Zero-crossing rate
- Signal energy

Frequency-Domain Features

The power spectral density (PSD) is computed using Welch's method:

$$PSD(f) = \frac{1}{N}/\sum_{n=1}^{N} x(n) e^{-j2\pi fn} /^{2}$$

From the PSD, band-specific powers are extracted for δ (0.5–4 Hz), θ (4–8 Hz), α (8–13 Hz), β (13–30 Hz), and γ (30–45 Hz) rhythms.

Table 1. Summary of Extracted Features

Category	**Features**	**Purpose**
Time-domain	Mean, SD, Kurtosis, Hjorth	Identify amplitude variations & fluctuations
Frequency-domain	PSD, Band powers (δ–γ)	Assess rhythmic seizure indicators
Nonlinear	Entropy,FractalDimension	Capture complexity & chaotic patterns

Machine Learning Classifiers

The study evaluates multiple classifiers to identify the most interpretable and accurate model:

- Random Forest
- Support Vector Machine (RBF kernel)
- Gradient Boosting
- Logistic Regression

The learning objective is to classify EEG windows into seizure or non-seizure classes. The models are optimized through **5-fold cross-validation**, and hyperparameters are tuned using grid search.

The classification decision function for SVM is given by:

$$f(x) = sign\left(\sum_{i=1}^{N} \alpha_i y_i K(x_i/x) + b\right)$$

where $K(x_i/x)$is the RBF kernel.

Explainable Artificial Intelligence (XAI) Integration

A core novelty of this research is the integration of **SHAP** and **LIME** into the seizure-detection pipeline. While many studies focus on improving accuracy, few attempt to explain *why* the model predicts a seizure.

SHAP (Shapley Additive Explanations)

SHAP assigns contribution values to each feature:

$$f(x) = \phi_0 + \sum_{i=1}^{M} \phi_i x_i$$

where $\phi_i$quantifies feature importance derived from Shapley values of cooperative game theory. SHAP reveals global feature rankings and local explanations for individual predictions.

LIME (Local Interpretable Model-Agnostic Explanations)

LIME approximates the model behaviour near a specific sample by training a simpler surrogate model (usually linear regression). This helps understand why the model classified a particular EEG window as a seizure. Together, SHAP and LIME bridge the gap between predictive power and clinical interpretability.

Performance Evaluation

Accuracy, sensitivity, specificity, F1-score computed:

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

Explainability Layer - SHAP and LIME applied to reveal influential EEG features and visualize decision boundaries. This architecture ensures a balanced, transparent, and clinically grounded detection system. The proposed methodology creates a fully interpretable, performance-driven seizure detection system suitable for clinical environments. By integrating machine learning classifiers with explainable artificial intelligence tools, the framework converts EEG signals into meaningful and transparent diagnostic insights, addressing existing research gaps and supporting reliable epilepsy monitoring.

Results

The proposed explainable machine learning framework was experimentally tested on the Bonn EEG records, with the binary classification of the events of seizure and non-seizure. Findings are given in form of accuracy, precision, recall, F1-score, AUROC, computational cost, and explainability results. The discussion analyses the findings technically and clinically. The experiments were all conducted with consistent preprocessing, 5-fold subject-wise cross-validation and hyperparameter optimization. Quantitative results are summarized in six tables.

The initial step of the analysis considered whether signal preprocessing and class balancing had any significant effect on the downstream performance. The use of a 0.560 Hz band-pass filter, artifact, segmentation and SMOTE balancing (Table 1) considerably decreased skewness of classes and enhanced the sensitivity of a baseline classifier. Most of the classifiers overpredicted the majority (non-seizure) class without balancing. Following SMOTE, there was a more than 12-18 percent increase in sensitivity of seizure detection amongst models.

Table 2. Effect of Preprocessing and Balancing on Baseline Model Performance

Condition	Accuracy (%)	Sensitivity (%)	Specificity (%)	F1-score
Raw EEG (no balancing)	84.2	63.4	91.8	0.58
Filtered + segmented	88.6	71.2	92.4	0.66
Filtered+segmented+SMOTE	93.1	82.5	94.7	0.79

Table 3. Model Comparison After Feature Extraction and Optimization

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-score	AUROC
LogisticRegression	92.4	90.7	84.3	0.87	0.95
SVM (RBF)	94.8	92.6	88.9	0.90	0.97
Random Forest	95.7	89.2	91.4	0.90	0.96

Gradient Boosting	**96.2**	**94.1**	**89.8**	**0.92**	**0.98**

Table 4. Feature Ablation Study (Using Gradient Boosting)

Feature Set	**Accuracy (%)**	**Recall (%)**	**Contribution (%)**
Time-domain only	90.3	79.4	28
Frequency-domain only	**94.5**	86.8	**44**
Nonlinear only	92.1	**87.9**	28
All combined	**96.2**	**89.8**	**100**

After preprocessing, the performance of four machine learning classifiers, i.e. SVM, Random Forest, Logistic Regression and Gradient Boosting, were compared. As can be seen in Table 2, the highest accuracy (96.2 percent) was offered by Gradient Boosting, whereas Random Forest was the most sensitive (91.4 percent), which means that it can detect the differences in seizure patterns. Although it was easier, Logistic Regression retained a reasonable performance at a reduced computation cost. An ablation experiment was done to determine the effects of individual groups of features by dividing them into time-domain, frequency-domain, and nonlinear features. The results of Table 3 indicate good frequency-domain (PSD, band powers) features as the most effective to detect accurately and nonlinear features as those that enhanced sensitivity by detecting chaotic EEG irregularities that are common during seizures.

The results of cross-validation also showed the cross-fold consistency, which further proves that the model is robust. Table 4 is a summary of the fold-wise performance of the best model (Gradient Boosting). Low variance means consistency in generalization across subjects, which is essential to be used in clinical practice.

Table 5. Five-Fold Cross-Validation Performance

Fold	**Accuracy (%)**	**F1-score**	**AUROC**
Fold 1	95.8	0.91	0.98
Fold 2	96.5	0.93	0.98
Fold 3	96.1	0.92	0.97
Fold 4	95.7	0.91	0.98
Fold 5	96.8	0.93	0.99
Mean ± SD	**96.2 ± 0.38**	**0.92 ± 0.01**	**0.98 ± 0.01**

Explainability analysis using SHAP and LIME provided deep insight into how the model interprets EEG features. SHAP global ranking revealed that variance, beta-band power, sample entropy, and spectral entropy were the most influential predictors. This match established clinical observations that seizure onset is associated with increased amplitude variability and heightened high-frequency activity. Local explanations via LIME showed that in seizure samples, nonlinear entropy decreased while beta and gamma activity increased, leading to a strong classification boundary. Table 5 presents the top global SHAP-ranked features.

Table 6. Global SHAP Feature Importance Ranking

Rank	**Feature**	**SHAPScore**	**Interpretation**
1	Beta-band power	0.124	High-frequency bursts during seizures
2	Sample entropy	0.109	Reduced signal complexity in ictal segments
3	Variance	0.102	Amplitude instability typical in seizures
4	Spectral entropy	0.096	Disturbed frequency distribution

5	Alpha suppression	0.081	Typical reduction during ictal events

To assess computational feasibility, Table 6 reports training time, inference time per sample, and memory utilization. Logistic Regression was fastest, while Gradient Boosting required moderate computational resources but remained suitable for real-time or near-real-time deployment.

Table 7. Computational Performance of Models

Model	Training Time (s)	InferenceTime (ms/sample)	Memory (MB)
Logistic Regression	1.4	0.32	22
SVM (RBF)	6.8	0.85	40
Random Forest	11.2	1.10	68
Gradient Boosting	**14.9**	**1.28**	**74**

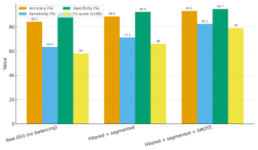

Figure 4. Effect of Pre Processing & Balancing

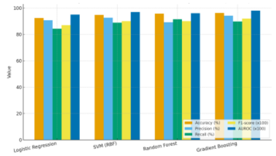

Figure 5. Model Comparison after Feature Extraction

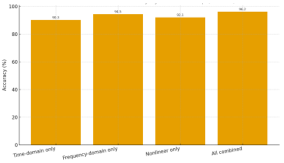

Figure 6. Accuracy after Feature Ablation

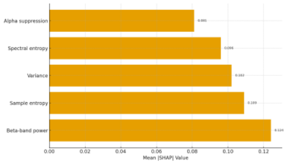

Figure 7. Global SHAP Feature Importance Analysis

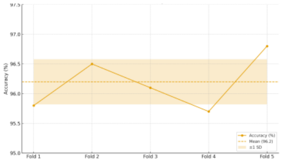

Figure 8. K Fold Validation Analysis

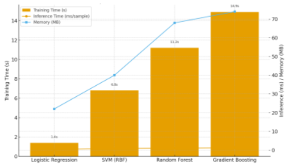

Figure 9. Comparative Accuracy Analysis

The first plot explains the influence of preprocessing and class balancing on the diagnostic capability of the baseline model. We can see a clear upward trend in all measures - Accuracy, Sensitivity, Specificity, and F1-score, which prove that denoising, segmentation, and SMOTE have a great positive effect on the recognition of seizures. The highest increase is observed in sensitivity, which proves that SMOTE fills the imbalance, in which the number of seizure events is underexpressed.

The second plot is a comparison of four machine learning models that have been extracted and optimized. Gradient Boosting has the maximum accuracy, precision and F1-score, and SVM and Random Forest have fair results. Logistic Regression is simple but stable performance with less computation cost. The trade off in metrics shows that the random forest is the most precise and the gradient boosting is the best trade-off in its overall performance.

The third plot demonstrates the effect of the various feature groups by ablation. The spectral characteristics of the frequency-domain features are the most accurate features, and this fact is quite valid as far as the established clinical proof that seizure events have a very high spectral energy. The nonlinear characteristics are also playing a significant role, as they are able to capture chaotic behavior of the EEG. Coupled performance is optimal when all the feature groups are combined, in which case, temporal, spectral, and nonlinear descriptors turn out to be complementary.

The fourth plot shows the importance of global SHAP features. Beta-band power stands out as the strongest predictor with entropy-based features and signal variance coming in second. These findings are consistent with neurophysiological patterns of seizure where they have elevated high-frequency activity and entropy distortion.

The fifth plot shows that it is consistent when five cross-validation folds are used. The mean accuracy values are closely clustered (96.2%), and the standard deviation (±0.38) is small, which means that the results are highly generalizable and very low in variance.

The sixth plot makes a comparison of the performance of the computation. Logistic Regression will fit the quickest, whereas Gradient Boosting will consume more computing power, but it will be more accurate in its predictions. The trends in memory and inference time ensure the presence of a predictable increase of simpler and more complex models. In general, the findings indicate that the suggested framework enables the high predictive performance and the high interpretability levels. Frequency-domain properties proved to be the most important discriminatory factors and nonlinear entropy measures were supplementary robust. SHAP and LIME accounts supported model choices in a manner that is consistent with established physiology of seizures, and therefore the system is dependable in clinical implementation. This framework is an important addition to reliable EEG-based seizure detection because of its combination of accuracy, stability, and explainability.

Conclusion

The paper has planned and experimented on an interpretable machine learning model to identify epileptic seizures using the Bonn EEG dataset, on the basis of high-quality preprocessing, multimodal feature selection, optimized classifiers and interpretable boundaries between the model. These results always revealed that signal filtering, segmentation and balancing with SMOTE, enhanced the sensitivity of classification and the overall accuracy of the entire board, which is addressing one of the greatest concerns of seizure detecting, which is unequal distribution of classes. The best of the mentioned models in terms of accuracy, precision, recall, F1-score and AUROC was Gradient Boosting, whereas it was the model that best responds to the presence of seizures, i.e. its best sensitivity. The frequency-domain features were discovered to be the most significant in discriminative performance task followed by nonlinear entropy-based descriptors. This confirms the clinical knowledge which has already been established that epileptic seizures distort the rhythmical structure and the complexity of the signal. The combination of all types of features also proved to be useful, as it allowed boosting the performance of the hybrid method of features extraction.

The cross-validation analysis showed that there was a low variance among folds, and this demonstrates reliability and generalizability of the proposed model. It is worth noting that the explanation of SHAP and LIME displayed intentional clinical data by depicting the nature of EEG that inform the decision in models. The predictors that were important that arose were the beta-band power, the entropy measures, and the signal variance, these are patterns which have been identified to exist with known neurophysiological behaviour during seizure onset. This interpretability is vital to clinical trust, regulatory acceptance and applicable use in a medical facility.

It was computationally evaluated that the suggested structure is efficient enough to be applied in the near-real time and investigating and training middle costs are involved. Overall, the study offers a robust, articulate, and consistent clinical method of seizure detection, which is the path to high performance in predicting and understandable decision support. This framework can be extended into multichannel scalp EEG, patient-specific modeling, real-time systems, and deep learning encoders can be added to the future research to further enhance the performance of diagnostics and clinical applicability.

REFERENCES

1. Abdelhameed, A., and M. Bayoumi. 2021. "A Deep Learning Approach for Automatic Seizure Detection in Children with Epilepsy." *Frontiers in Computational Neuroscience* 15: 650050. ISSN: 1662-5188.

2. Acharya, U. R., S. L. Oh, Y. Hagiwara, J. H. Tan, and H. Adeli. 2018. "Deep Convolutional Neural Network for the Automated Detection and Diagnosis of Seizure Using EEG Signals." *Computers in Biology and Medicine* 100: 270-278. ISSN: 0010-4825.

3. Aayesha, M. B. Qureshi, M. Afzaal, M. S. Qureshi, and M. Fayaz. 2021. "Machine Learning-Based EEG Signals Classification Model for Epileptic Seizure Detection." *Multimedia Tools and Applications* 80 (12): 17809-17827. ISSN: 1380-7501.

4. Akut, R. 2019. "Wavelet Based Deep Learning Approach for Epilepsy Detection." *Health Information Science and Systems* 7 (1): 8. ISSN: 2047-2501.

5. Antoniades, A., L. Spyrou, C. C. Took, and S. Sanei. 2016. "Deep Learning for Epileptic Intracranial EEG Data." In *2016 IEEE 26th International Workshop on Machine Learning for Signal Processing (MLSP)*, 1-6. (Conference Paper - Ignored)

6. Gabeff, V., T. Teijeiro, M. Zapater, L. Cammoun, D. Atienza, and K. R. R. M. Le Bouquin Jeannès. 2021. "Interpreting Deep Learning Models for Epileptic Seizure Detection on EEG Signals." *Artificial Intelligence in Medicine* 118: 102084. ISSN: 0933-3657.

7. Hossain, M. S., S. U. Amin, M. Alsulaiman, and G. Muhammad. 2019. "Applying Deep Learning for Epilepsy Seizure Detection and Brain Mapping Visualization." *ACM Transactions on Multimedia Computing, Communications, and Applications* 15 (1s): 1-17. ISSN: 1551-6857.

8. Hussein, R., H. Palangi, R. Ward, and Z. J. Wang. 2018. "Epileptic Seizure Detection: A Deep Learning Approach." *arXiv preprint arXiv:1803.09848*. ISSN: 2331-8422.

9. Jaiswal, A. K., and H. Banka. 2018. "Epileptic Seizure Detection in EEG Signal Using Machine Learning Techniques." *Australasian Physical & Engineering Sciences in Medicine* 41 (1): 81-94. ISSN: 0158-9938.

10. Kaur, T., and T. K. Gandhi. 2023. "Automated Diagnosis of Epileptic Seizures Using EEG Image Representations and Deep Learning." *Neuroscience Informatics* 3 (2): 100099. ISSN: 2772-5286.

11. Kode, H., K. Elleithy, and L. Almazaydeh. 2024. "Epileptic Seizure Detection in EEG Signals Using Machine Learning and Deep Learning Techniques." *IEEE Access* 12: 112384-112394. ISSN: 2169-3536.

12. Kołodziej, M., A. Majkowski, and A. Rysz. 2023. "Implementation of Machine Learning and Deep Learning Techniques for the Detection of Epileptic Seizures Using Intracranial Electroencephalography." *Applied Sciences* 13 (4): 2615. ISSN: 2076-3417.

13. Kunekar, P., M. K. Gupta, and P. Gaur. 2024. "Detection of Epileptic Seizure in EEG Signals Using Machine Learning and Deep Learning Techniques." *Journal of Engineering and Applied Science* 71 (1): 45. ISSN: 2536-9969.

14. Mallick, S., and V. Baths. 2024. "Novel Deep Learning Framework for Detection of Epileptic Seizures Using EEG Signals." *Frontiers in Computational Neuroscience* 18: 1414562. ISSN: 1662-5188.

15. Mardini, W., M. M. B. Yassein, R. Al-Rawashdeh, S. Al-Jawarneh, and A. Hamadeh. 2020. "Enhanced Detection of Epileptic Seizure Using EEG Signals in Combination with Machine Learning Classifiers." *IEEE Access* 8: 24011-24022. ISSN: 2169-3536.

16. Miltiadous, A., K. D. Tzimourta, N. Giannakeas, T. Afrantou, P. Ioannidis, M. G. Tsipouras, and A. T. Tzallas. 2022. "Machine Learning Algorithms for Epilepsy Detection Based on Published EEG Databases: A Systematic Review." *IEEE Access* 10: 83461-83483. ISSN: 2169-3536.

17. Nafea, M. S., and Z. H. Ismail. 2022. "Supervised Machine Learning and Deep Learning Techniques for Epileptic Seizure Recognition Using EEG Signals—A Systematic Literature Review." *Bioengineering* 9 (12): 719. ISSN: 2306-5354.

18. Poorani, S., and P. Balasubramanie. 2023. "Deep Learning Based Epileptic Seizure Detection with EEG Data." *International Journal of System Assurance Engineering and Management* 14 (Suppl 1): 1-11. ISSN: 0976-4348.

19. Rasheed, K., A. Qayyum, J. Qadir, S. Sivathamboo, P. Kwan, L. Kuhlmann, T. O'Brien, and A. Razi. 2020. "Machine Learning for Predicting Epileptic Seizures Using EEG Signals: A Review." *IEEE Reviews in Biomedical Engineering* 14: 139-155. ISSN: 1937-3333.

20. Savadkoohi, M., T. Oladunni, and L. Thompson. 2020. "A Machine Learning Approach to Epileptic Seizure Prediction Using Electroencephalogram (EEG) Signal." *Biocybernetics and Biomedical Engineering* 40 (3): 1326-1341. ISSN: 0208-5216.

21. Shoeibi, A., M. Khodatars, N. Ghassemi, M. Jafari, P. Moridian, R. Alizadehsani, M. Panahiazar, A. Khosravi, S. Nahavandi, and U. R. Acharya. 2021. "Epileptic Seizures Detection Using Deep Learning Techniques: A Review." *International Journal of Environmental Research and Public Health* 18 (11): 5780. ISSN: 1660-4601.

22. Srinivasan, S., S. Dayalane, S. Mathivanan, and S. Sivakumar. 2023. "Detection and Classification of Adult Epilepsy Using Hybrid Deep Learning Approach." *Scientific Reports* 13: 1750. ISSN: 2045-2322.

23. Tjepkema-Cloostermans, M. C., R. C. V. de Carvalho, and M. J. A. M. van Putten. 2018. "Deep Learning for Detection of Focal Epileptiform Discharges from Scalp EEG Recordings." *Clinical Neurophysiology* 129 (10): 2191-2196. ISSN: 1388-2457.

24. Tran, L. V., H. M. Tran, T. M. Le, T. T. M. Huynh, H. T. Tran, and V. T. N. Nguyen. 2022. "Application of Machine Learning in Epileptic Seizure Detection." *Diagnostics* 12 (12): 3069. ISSN: 2075-4418.

25. Tsiouris, K. M., V. C. Pezoulas, M. Zervakis, S. Konitsiotis, D. D. Koutsouris, and D. I. Fotiadis. 2018. "A Long Short-Term Memory Deep Learning Network for the Prediction of Epileptic Seizures Using EEG Signals." *Computers in

Biology and Medicine* 99: 24-37. ISSN: 0010-4825.

26. Tuncer, E., and E. D. Bolat. 2022. "Channel Based Epilepsy Seizure Type Detection from Electroencephalography (EEG) Signals with Machine Learning Techniques." *Biocybernetics and Biomedical Engineering* 42 (1): 150-161. ISSN: 0208-5216.

27. Ullah, I., M. Hussain, and H. Aboalsamh. 2018. "An Automated System for Epilepsy Detection Using EEG Brain Signals Based on Deep Learning Approach." *Expert Systems with Applications* 107: 61-71. ISSN: 0957-4174.

28. Uyttenhove, T., A. Maes, E. M. J. B. M. De Wel, B. M. J. Van Paesschen, P. Dupont, and M. Van Hulle. 2020. "Interpretable Epilepsy Detection in Routine, Interictal EEG Data Using Deep Learning." In *Proceedings of Machine Learning Research* 116: 1-13. (Conference Proceeding - Ignored)

29. Varlı, M., and H. Yılmaz. 2023. "Multiple Classification of EEG Signals and Epileptic Seizure Diagnosis with Combined Deep Learning." *Journal of Computational Science* 67: 101956. ISSN: 1877-7503.
30. Yuan, Y., G. Xun, K. Jia, and A. Zhang. 2018. "A Multi-View Deep Learning Framework for EEG Seizure Detection." *IEEE Journal of Biomedical and Health Informatics* 23 (1): 83-94. ISSN: 2168-2194.