IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Epileptic Seizure Onset Prediction: Evaluating Predictive Models With Eeg Analysis

Name: Meghna Indraprakash Singh

Research Scholar, Shri Jagdishprasad jhabarmal tibrewala university, computer science, Jhunjhunu, Rajasthan, India

Abstract

Epileptic seizure detection using electroencephalography (EEG) plays a crucial role in neurology, facilitating early diagnosis and timely intervention. Traditional EEG analysis is time-consuming and susceptible to inconsistencies, highlighting the need for automated systems that can enhance both efficiency and accuracy. This study evaluates various machine learning models to determine the most effective approach for seizure prediction, focusing on predictive accuracy and computational efficiency. Among the models tested, the Artificial Neural Network (ANN) demonstrated the highest AUC (0.98) and accuracy (0.9814), followed by XGBoost, CatBoost, and Random Forest. However, ANN's long training time (24.61s) restricts its suitability for real-time applications. In contrast, CatBoost emerged as the optimal model, delivering strong classification performance with 96.87% accuracy and a 0.9224 F1 score, while maintaining a much shorter training time (6.93s) and fast inference (0.1028s). These results suggest that CatBoost is well-suited for practical deployment in clinical environments, offering both high accuracy and computational efficiency for epileptic seizure detection.

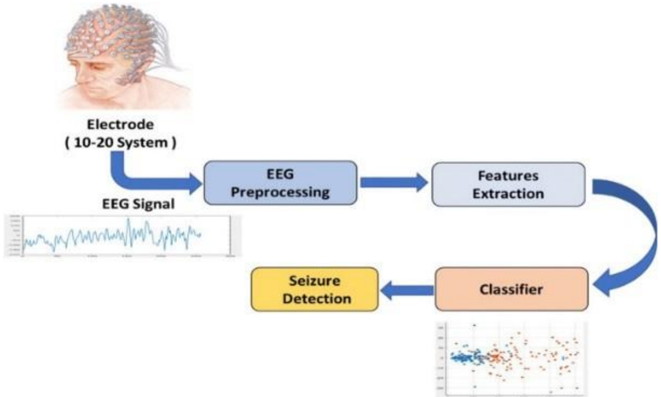
Keywords

Epileptic Seizure Detection, EEG Signal Analysis, Machine Learning Models, CatBoost, Computational Efficiency, Seizure Prediction

Introduction

Electroencephalography (EEG) is a non-invasive technique for assessing brain function and diagnosing neurological disorders like epilepsy. It captures crucial brain activity data with high temporal resolution, ideal for time-series analysis. Traditionally analyzed manually, EEG now benefits from automated algorithms, improving the detection and management of neurological conditions. Epilepsy, affecting about 1% of the global population, is the second most prevalent neurological disorder. Despite treatments, a cure remains elusive, and effectiveness varies. Automated EEG-based seizure detection systems offer promising solutions but often face challenges like high computational demands and lack of patient-specific tailoring. EEG remains the gold standard for epilepsy diagnosis due to its ability to monitor both ictal and inter-ictal activities. While visual interpretation by neurologists is effective, it is time-consuming and prone to inconsistencies. Computer-aided techniques provide more efficient and cost-effective solutions, highlighting the need for advanced automated detection methods. This study aims to develop a machine learning model for epileptic seizure detection using EEG signals. By evaluating multiple algorithms, the research seeks to identify the model with the highest accuracy and computational efficiency, focusing on both predictive performance and real-time applicability for clinical settings.

Fig 1: EEG-Based Seizure Detection Process



Literature Review:

This study develops a machine learning model for predicting epileptic seizures using EEG signals, focusing on preprocessing, feature extraction, and classification. The model applies Empirical Mode Decomposition (EMD) for noise reduction and common spatial pattern (CSP) filtering to enhance signal quality. Statistical and spectral features are extracted to characterize preictal and interictal states, and Support Vector Machines (SVM) are used for classification. Evaluating the CHB-MIT dataset, the model achieves a seizure prediction time of 23.61 minutes with 92.23% sensitivity. Compared to previous studies using SVM (Usman, Usman, & Fong, 2017), this research also incorporates advanced models like XGBoost, CatBoost, and Random Forest, offering a more comprehensive performance evaluation. Additionally, while existing literature often focuses on accuracy (Esha, Begum, & Rahman, 2024), this study emphasizes computational efficiency, with CatBoost emerging as the most practical model for real-time implementation, balancing high accuracy and low training times. This work addresses gaps in current methods by prioritizing both predictive performance and computational feasibility for seizure detection in clinical settings.

Methodology

In this section, we describe the methodology used in this research, detailing the processes involved in data preparation, model selection, and evaluation for epileptic seizure prediction using EEG signals. We applied various machine learning algorithms, including Logistic Regression, Random Forest, Support Vector Machines (SVM), XGBoost, and CatBoost, to develop predictive models.

By rigorously testing, evaluating, and analyzing model performance, we aim to optimize classification accuracy and reliability in detecting epileptic seizures from EEG signals.

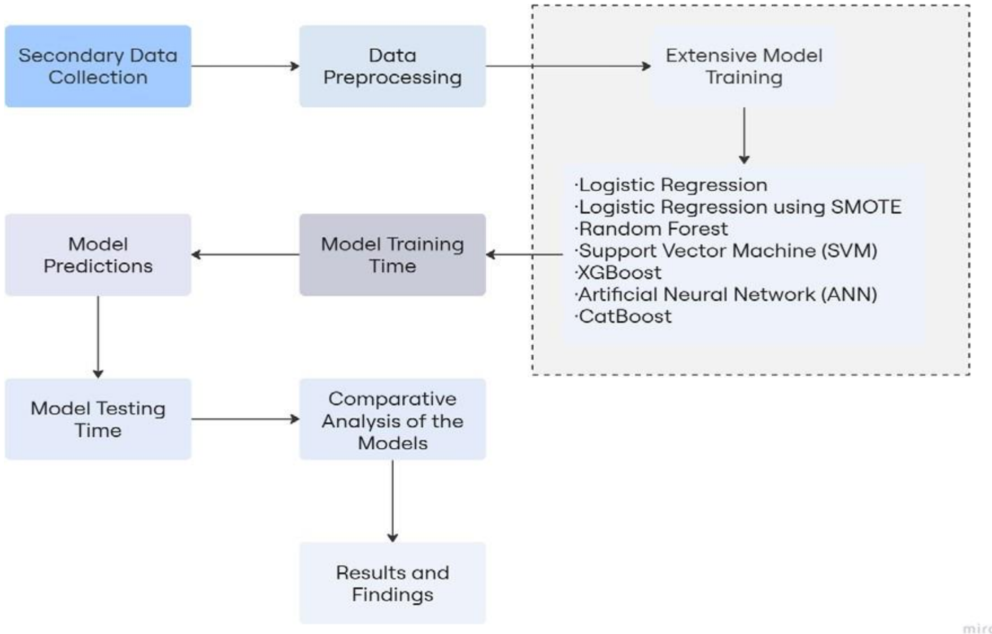


Fig 2: Workflow for Machine Learning-Based Seizure Detection

DataFramework

The dataset, sourced from Kaggle (accessed November 18, 2023), consists of 11,500 samples and 178 features of EEG signals from 500 individuals. Each 23.5-second EEG recording is split into 23 one-second segments, and the target variable 'y' contains values from {1, 2, 3, 4, 5}. The data was split into 80% training (9,200 examples) and 20% testing (2,300 examples) to optimize model training and evaluation.

ResearchFramework

This study introduces CatBoost for epileptic seizure prediction, comparing it with other machine learning models (Logistic Regression, SVM, Random Forest, XGBoost). The method includes data preprocessing, normalization, and feature selection to ensure optimal performance. CatBoost is highlighted for its high accuracy and efficiency, providing a computationally feasible solution for real-time seizure prediction.

Data Pre-Processing

Data preprocessing involves using Python 3.11.5 with pandas for cleaning, handling missing values, and standardizing features. The dataset is split into training and testing sets (80/20). Principal Component Analysis (PCA) reduces dimensionality, selecting 39 components to enhance model training and evaluation.

Model Training

- **Logistic Regression:** A binary classifier used for predicting seizures by applying a logistic function.
- **Random Forest:** An ensemble method using 150 decision trees, effective for managing high-dimensional EEG data.
- **SVM:** A classification method using the RBF kernel for non-linear data separation.
- **XGBoost:** An advanced ensemble technique using gradient boosting for high accuracy and speed.
- **ANN:** A neural network model designed to handle complex EEG data patterns.
- **CatBoost:** A gradient boosting method that excels at handling categorical data, offering a fast and accurate model for seizure prediction.

SMOTE

Synthetic Minority Over-sampling Technique (SMOTE) is applied to address class imbalance in logistic regression, improving sensitivity to seizure detection.

ClassificationMetrics

Models are evaluated using accuracy, precision, recall, F1 score, and ROC-AUC to assess performance, particularly in handling imbalanced data and predicting rare events like seizures.

ROC and AUC

The ROC curve and AUC measure a model's ability to distinguish between seizure and non-seizure events. AUC values above 0.7 are considered good, indicating effective classification.

Results

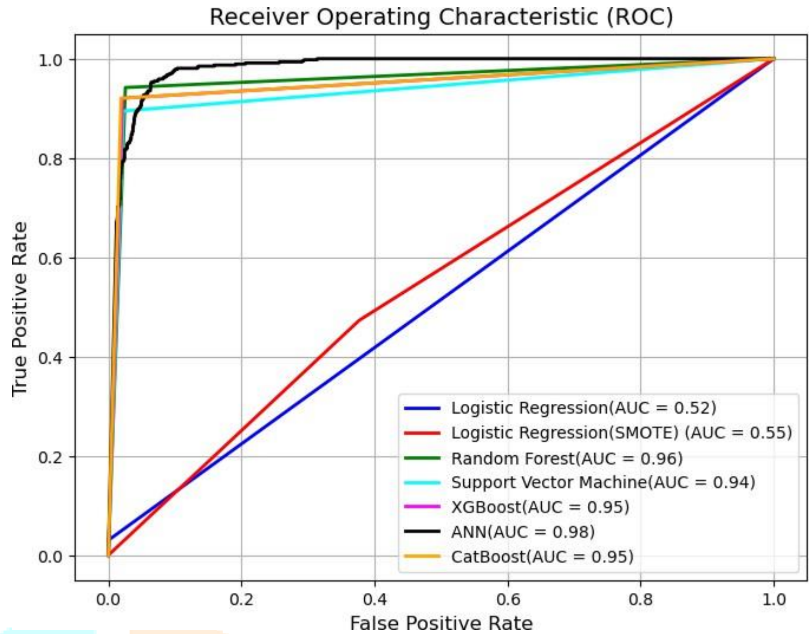
This section presents the results of machine learning models used to predict **seizure onset** based on EEG signal analysis. The models—Logistic Regression, Random Forest, Support Vector Machine (SVM), XGBoost, ANN, and CatBoost—are evaluated using performance metrics such as accuracy, precision, recall, and F1-score, with a focus on accurately distinguishing between seizure and non- seizure events. The impact of key features on seizure predictions is highlighted, emphasizing their role in improving the models' ability to detect seizures and enhance predictive performance.

Model Selection and Prediction

A thorough comparison of multiple machine learning algorithms was conducted to identify the best model for seizure onset prediction. The models evaluated include Logistic Regression, Random Forest, Support Vector Machine (SVM), XGBoost, Artificial Neural Networks (ANN), and CatBoost. To comprehensively assess each model's performance, a table of precision metrics is provided, along with a combined ROC-AUC graph to visualize the models' discriminative abilities. Additionally, a bar graph illustrating the training time of all models is included, offering insights into computational efficiency alongside predictive performance.

Model	Accuracy	F1 Score	Precision	Recall
Logistic Regression	0.8039	0.0585	1.0	0.0301
Logistic Regression using SMOTE	0.5926	0.3195	0.2412	0.4731
Random Forest	0.9639	0.9129	0.8914	0.9355
Support Vector Machine (SVM)	0.9609	0.9024	0.9103	0.8946
XGBoost	0.9670	0.9185	0.9165	0.9204
Artificial Neural Network (ANN)	0.9814	0.9516	0.9601	0.9433
CatBoost	0.9687	0.9224	0.9244	0.9204

Table 1: Performance Metrics of Machine Learning Models for Epileptic Seizure Detection



ROC Curve Comparison of Machine Learning Models for Seizure Prediction

References

[1] Esha, O. K., Begum, N., & Rahman, S. (2024). EpiNet: A hybrid machine learning model for epileptic seizure prediction using EEG signals from a 500 patient dataset. *International Journal of Advanced Computer Science and Applications*, 15(1), 116–125. https://doi.org/10.14569/IJACSA.2024.01501116

[2] Usman, S. M., Usman, M., & Fong, S. (2017). Epileptic seizures prediction using machine learning methods. *Computational and Mathematical Methods in Medicine*, 2017, 1–10. https://doi.org/10.1155/2017/9074759

[3] Altaf, Z., Unar, M. A., Narejo, S., Zaki, M. A., & N.-D. (2023). Generalized epileptic seizure prediction using machine learning method. *International Journal of Advanced Computer Science and Applications*, 14(1), 1–8. https://doi.org/10.14569/IJACSA.2023.0140155

[4] Costa, G., Teixeira, C., & Pinto, M. F. (2024). Comparison between epileptic seizure prediction and forecasting based on machine learning. *Scientific Reports*, 14(1), 5653. https://doi.org/10.1038/s41598-024-56019-z

[5] Kunekar, P., Gupta, M. K., & Gaur, P. (2024). Detection of epileptic seizure in EEG signals using machine learning and deep learning techniques. *Journal of Engineering and Applied Sciences*, 71(1), 21. https://doi.org/10.1186/s44147-023-00353-y

[6] Engel, J. (2006). ILAE classification of epilepsy syndromes. *Epilepsy Research*, 70(1), 5–10.

[7] Moghim, N., & Corne, D. W. (2014). Predicting epileptic seizures in advance. *PLoS ONE*, 9(6), e99334.

[8] Ramoser, H., Müller-Gerking, J., & Pfurtscheller, G. (2000). Optimal spatial filtering of single trial EEG during imagined hand movement. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 8(4), 441–446.

[9] Vidaurre, C., Krämer, N., Blankertz, B., & Schlögl, A. (2009). Time domain parameters as a feature for EEG-based brain computer interfaces. *Neural Networks*, 22(9), 1313–1319.

[10] Osowski, S., Swiderski, B., Cichocki, A., & Rysz, A. (2007). Epileptic seizure characterization by Lyapunov exponent of EEG signal. *COMPEL - The International Journal for Computation and Mathematics in Electrical and Electronic Engineering*, 26(5), 1276–1287.

[11] Hebert, L. E., Scherr, P. A., Bienias, J. L., Bennett, D. A., & Evans, D. A. (2003). Alzheimer disease in the US population: Prevalence estimates using the 2000 census. *JAMA Neurology*, 60(8), 1119–1122.

[12] Xiang, J., Li, C., Li, H., Cao, R., Wang, B., Han, X., & Chen, J. (2015). The detection of epileptic seizure signals based on fuzzy entropy. *Journal of Neuroscience Methods*, 243, 18–25.

[13] Bergen, D. C. (2006). Do seizures harm the brain? *Epilepsy Currents*, 6(4), 117–118.
[14] Usman, S. M., Usman, M., & Fong, S. (2017). Epileptic seizures prediction using machine learning methods. *Computational and Mathematical Methods in Medicine*, 2017.
[15] Hart, Y. M. (2012). Epidemiology, natural history and classification of epilepsy. *Medicine*, 40(9), 471–476.
[16] Tran, L. V., et al. (2022). Application of machine learning in epileptic seizure detection. *Diagnostics*, 12(11), 2879.
[17] Cao, X., et al. (2021). Automatic seizure classification based on domain-invariant deep representation of EEG. *Frontiers in Neuroscience*, 15, 760987.
[18] Tripathi, P. M., et al. (2023). Automatic seizure detection and classification using super-resolution superlet transform and deep neural network - A preprocessing-less method. *Computer Methods and Programs in Biomedicine*, 107680.
[19] Ahmed, M. I. B., et al. (2022). A review on machine learning approaches in identification of pediatric epilepsy. *SN Computer Science*, 3(6), 437. https://doi.org/10.1007/s42979-022-01358-9
[20] Farooq, M. S., Zulfiqar, A., & Riaz, S. (2023). Epileptic seizure detection using machine learning: Taxonomy, opportunities, and challenges. *Diagnostics*, 13(6), 1058.
[21] Tharayil, J. J., Chiang, S., Moss, R., Stern, J. M., Theodore, W. H., & Goldenholz, D. M. (2017). A big data approach to the development of mixed-effects models for seizure count data. *Epilepsia*, 58(5), 835–844. https://doi.org/10.1111/epi.13727
[22] Jaiswal, A. K., & Banka, H. (2017). Epileptic seizure detection in EEG signal with GModPCA and support vector machine. *Bio-Medical Materials and Engineering*, 28(2), 141–157.
[23] Tiwari, A. K., Pachori, R. B., Kanhangad, V., & Panigrahi, B. K. (2017). Automated diagnosis of epilepsy using key-point-based local binary pattern of EEG signals. *IEEE Journal of Biomedical and Health Informatics*, 21, 888–896.
[24] Toraman, S. (2020). Preictal and interictal recognition for epileptic seizure prediction using pre-trained 2DCNN models. *Traitement du Signal*, 37(6).
[25] Satapathy, S., Jagadev, A., & Dehuri, S. (2017). Weighted majority voting based ensemble of classifiers using different machine learning techniques for classification of EEG signal to detect epileptic seizure. *Informatica*, 41, 99–110.
[26] Rogowski, Z., Gath, I., & Bental, E. (1981). On the prediction of epileptic seizures. *Biological Cybernetics*, 42(1), 9–15.
[27] Harrison, M. A. F., Osorio, I., Frei, M. G., Asuri, S., & Lai, Y. C. (2005). Correlation dimension and integral do not predict epileptic seizures. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 15(3), 033106.

- Salant, Y., Gath, I., & Henriksen, O. (1998). Prediction of epileptic seizures from two-channel EEG. *Medical and Biological Engineering and Computing*, 36(5), 549–556.

- van Drongelen, W., Nayak, S., Frim, D. M., Kohrman, M. H., Towle, V. L., Lee, H. C., McGee, A. B., Chico, M. S., & Hecox, K. E. (2003). Seizure anticipation in pediatric epilepsy: Use of Kolmogorov entropy. *Pediatric Neurology*, 29(3), 207–213.

- Lehnertz, K., & Litt, B. (2005). The first international collaborative workshop on seizure prediction: Summary and data description. *Clinical Neurophysiology*, 116(3), 493–505