Detecting Epileptic Seizures from EEG Data: A Summary

By

Preonty Paul

Introduction

Epilepsy is a chronic neurological disorder affecting approximately 50 million people worldwide, characterized by recurrent seizures. These seizures occur due to abnormal bursts of electrical activity in the brain, which can lead to loss of consciousness, involuntary movements, and disruptions in behavior, memory, or emotions. The electroencephalogram (EEG) is a critical diagnostic tool used by healthcare professionals to detect epileptic seizures. EEG technology allows medical practitioners to analyze brain activity and identify seizure patterns, aiding in diagnosis and treatment planning.

This case study explores the use of machine learning (ML) and data science techniques to detect seizures using EEG data. The primary goal is to improve seizure detection accuracy and efficiency, which can significantly enhance clinical outcomes. However, analyzing EEG data presents challenges such as noise interference, variability in signal quality, and differences in sampling rates. This study aims to overcome these challenges by developing a robust ML model capable of distinguishing seizure and non-seizure events. The research emphasizes the importance of transitioning from clean, controlled datasets to real-world noisy data to test the model's effectiveness in practical healthcare settings.

Data Exploration and Preprocessing

The study utilized two datasets: a full-signal dataset containing 500 samples of 23-second EEG recordings and a short-signal dataset derived from segmenting each 23-second signal into 1-second windows, resulting in 11,500 samples. Each dataset was assigned binary labels (0 for non-seizure and 1 for seizure) to facilitate model training. Preprocessing steps included segmenting EEG signals for better temporal resolution, normalizing data to reduce inconsistencies, and converting datasets to Excel format for further analysis. These steps ensured uniformity, reliability, and suitability for feature extraction and model training.

Model Development and Key Findings

Initially, the study explored feature engineering and the Random Forest model, but due to a limited understanding of EEG data and low model accuracy, this approach was abandoned. Instead, a Convolutional Neural Network (CNN) model was adopted, yielding significantly better results. The CNN model demonstrated an overall accuracy of 98% in classifying seizure and non-seizure events.

- For non-seizure cases (0): Precision = 1.00, Recall = 0.98, F1-score = 0.99
- For seizure cases (1): Precision = 0.91, Recall = 0.98, F1-score = 0.95

These results indicate that the CNN model effectively minimizes false positives and false negatives, making it a highly reliable tool for seizure detection. The confusion matrix further validated the model's performance, showing that out of 460 actual non-seizure cases, 449 were correctly classified, with only 11 false positives. Similarly, out of 115 actual seizure cases, 113 were accurately identified, with just two false negatives.

The CNN model parameters included two Conv1D layers with 64 and 128 filters, two Max Pooling layers, ReLU and Sigmoid activation functions, a train-test split of 95%-5%, 50 training epochs, and a batch size of 32. These parameters optimized model performance, ensuring high accuracy and reliability in seizure detection.

Conclusion and Future Recommendations

Future research should focus on incorporating advanced feature engineering techniques, multi-channel EEG data, and improved model explainability. Testing the model on real-world, noisy datasets and developing deployment strategies for real-time seizure detection will enhance its practical applicability in healthcare. This study highlights the role of artificial intelligence in advancing epilepsy diagnosis and management, contributing to personalized and effective treatment strategies.