Nile University
جامعة النيل الأهلية

EEG Epilepsy Seizure Recognition Using Machine Learning and Deep Learning

Name: Raghad Mohamed Abdelkader

Course Code: CBIO313

Supervisor: Dr. Mohamed Mahmoud Elsayeh

T.A: Malak Soliman

1.0 Introduction

Epilepsy is a chronic neurological disorder affecting approximately 50 million people worldwide, making it one of the most common neurological conditions globally. It is characterized by recurrent, unprovoked seizures resulting from sudden bursts of abnormal electrical activity in the brain. These seizures can vary in severity and presentation, ranging from brief lapses of attention to prolonged convulsions, significantly impacting a patient's quality of life. Timely and accurate diagnosis is critical for effective treatment and management.

Electroencephalography (EEG) is the primary clinical tool used to monitor and assess brain activity in epilepsy patients. EEG captures electrical signals from the brain through electrodes placed on the scalp, generating complex waveforms that reflect brain function. Seizures often cause specific, identifiable disruptions in these waveforms. However, interpreting EEG data manually is a labor-intensive and highly specialized task. It requires experienced neurologists to analyze lengthy recordings, which can be prone to human error and subjectivity, especially when subtle patterns are involved.

With the rapid growth of healthcare data and computational power, artificial intelligence has emerged as a powerful tool for automating and enhancing clinical decision-making. This project explores the application of machine learning (ML) and deep learning (DL) techniques to automate the detection of epileptic seizures using raw EEG data. By leveraging data-driven models, the aim is to identify seizure events with greater speed, accuracy, and consistency than traditional methods. Such a system could support neurologists in clinical settings by providing early warnings, reducing diagnostic delays, and ultimately improving patient outcomes.

This research involves a complete pipeline, including data cleaning, exploratory analysis, feature selection, and model training using both conventional ML algorithms and modern DL architectures. A comparison of performance, speed, and interpretability between the two approaches is presented, highlighting the potential of deep learning in handling complex biomedical signals like EEG. The ultimate goal is to contribute to the development of intelligent, real-time seizure detection systems for use in hospitals and diagnostic centers.

2.0 Materials and Methods

This study was conducted using a publicly available EEG dataset consisting of 33,047 samples with 180 signal features each, along with a target column (y) indicating seizure class labels ranging from 1 to 5. The raw data required extensive preprocessing. Initially, unnecessary columns such as those labeled 'Unnamed' were removed. Missing values were handled using mean imputation, and all features were standardized using StandardScaler to ensure uniform feature scaling. Exploratory Data Analysis (EDA) was performed using visual tools like histograms, boxplots, correlation heatmaps, and PCA plots to better understand the data's structure and identify outliers or patterns. Feature selection was applied to enhance model efficiency and reduce noise, including the removal of constant and low-variance features, as well as the use of correlation and statistical tests to retain the most informative signals. For modeling, both traditional machine learning algorithms (Logistic Regression, SVM, Random Forest) and deep learning were used. Machine learning models were optimized using GridSearchCV, while ensemble techniques like voting, bagging, boosting, and stacking were also explored. For deep learning, a neural network was built using LSTM and ANN algorithm, and trained on the same scaled features. Finally, the best-performing model was prepared for deployment, highlighting the potential for real-time clinical use.

3.0 Results

3.1 Heamap

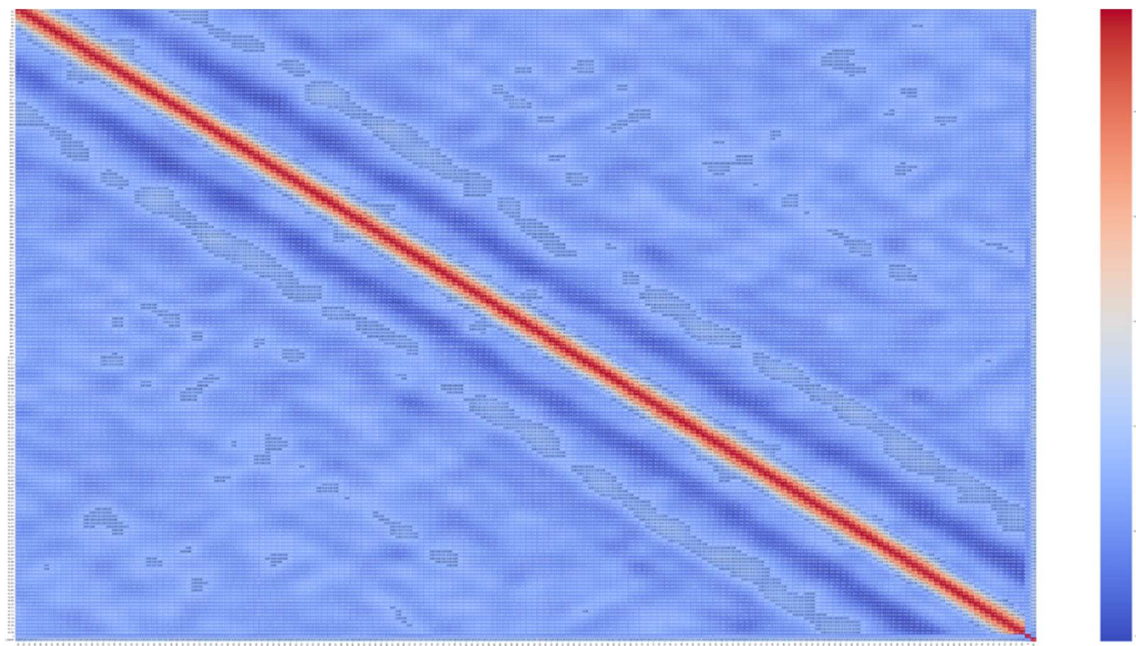

Figure 1: Correlation heatmap of the EEG signal features. The strong diagonal indicates perfect self-correlation for each feature, while the low off-diagonal values show minimal correlation between different features. This is expected in EEG data, where each channel records independent brain regions. The lack of multicollinearity supports the use of all features in the initial model training phase

3.2Boxplot of EEG Features

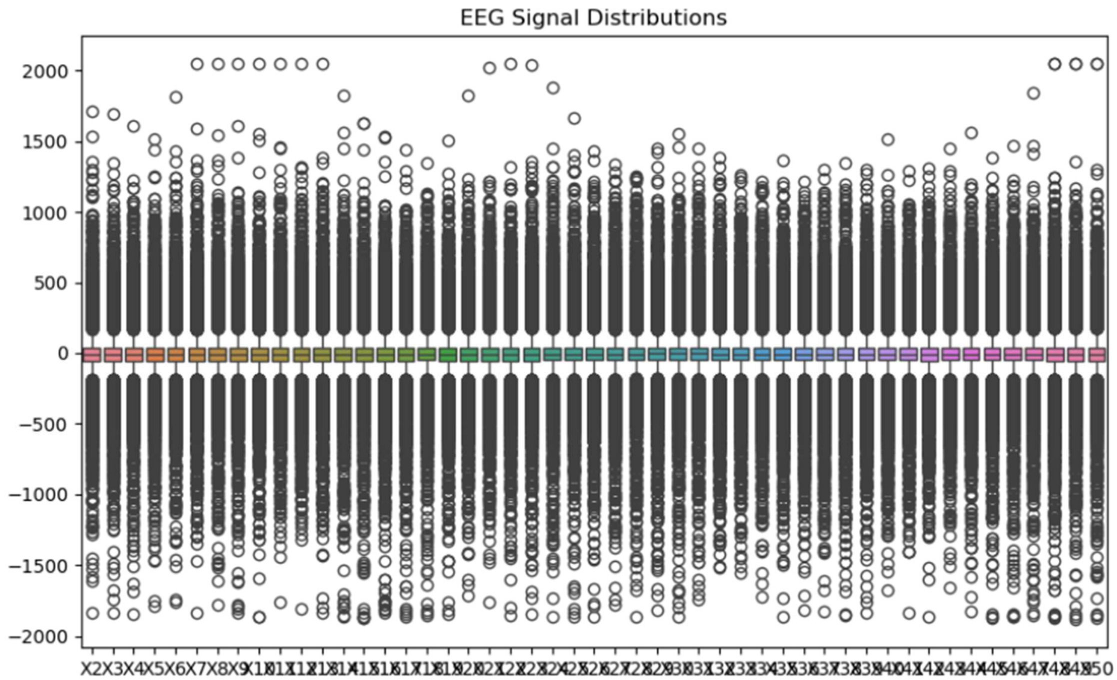

Figure 2: Boxplot showing the distribution of EEG signal values across selected features. A high number of outliers are present, represented by dots beyond the whiskers. This reflects the noisy nature of EEG data, where irregular spikes or fluctuations are common. Such variability can affect model performance and highlights the importance of robust preprocessing and feature scaling.

3.3PCA Scatter Plot

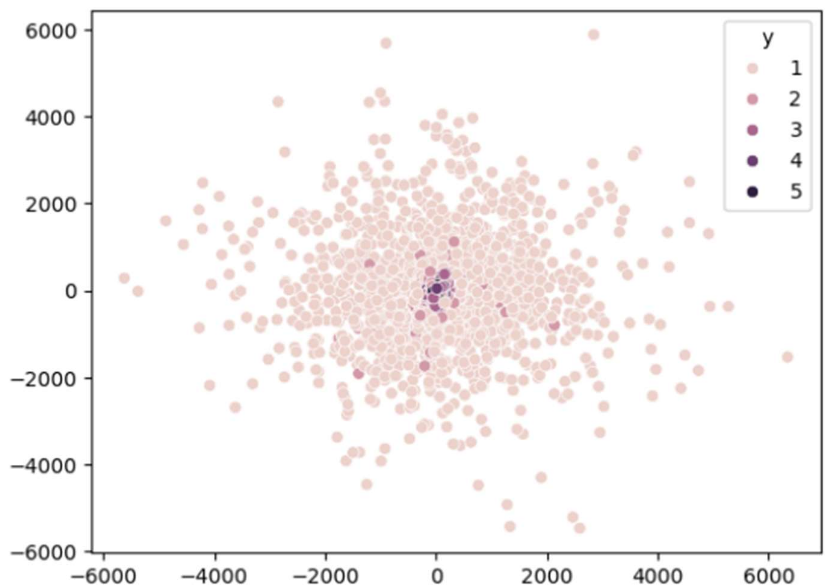

Figure 3: Principal Component Analysis (PCA) scatter plot displaying the projection of high-dimensional EEG data onto two principal components. Most data points cluster near the center, with minor separation among classes. This suggests class overlap and explains why simple machine learning models may struggle to differentiate seizure types. However, it also confirms some separable structure that complex models can exploit.

3.4EEG Signal Waveform

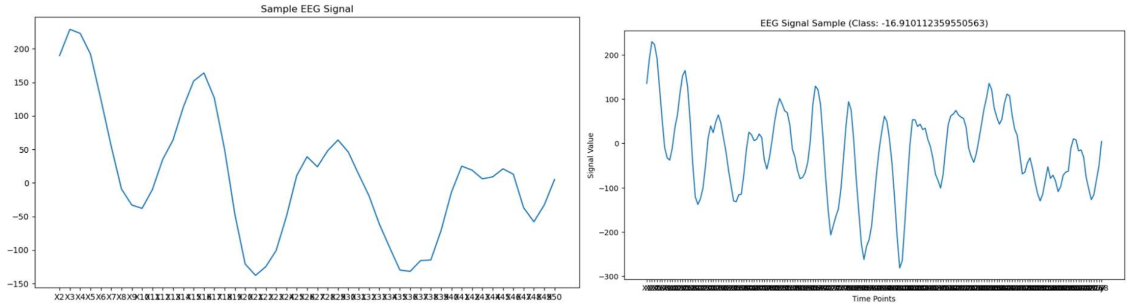

Figure 4: Raw EEG signal plot showing a time series of brainwave activity. The oscillatory behavior, including distinct peaks and troughs, reflects typical brain rhythms. Sudden changes in amplitude or frequency often indicate seizure onset, which is what the model is trained to detect. Understanding this structure is key for developing effective feature extraction or end-to-end models.

3.5Distribution Plots of EEG Features

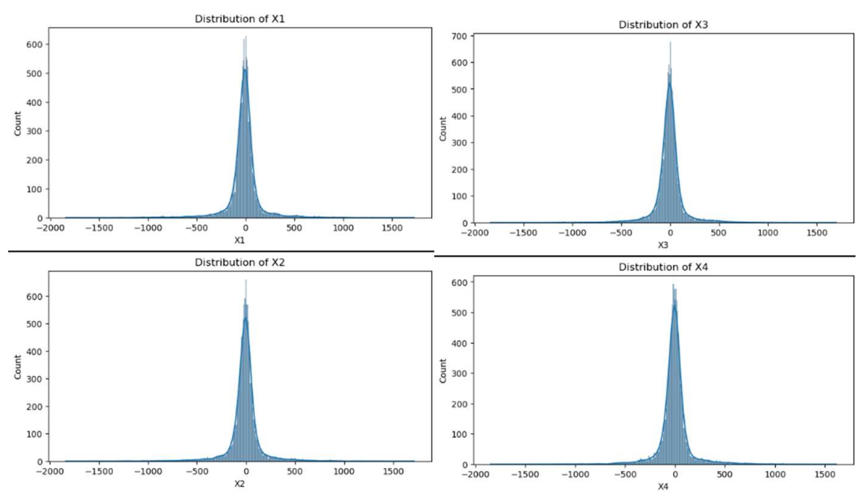

Figure 5: Distribution plots for selected EEG features. Most feature distributions are symmetric and bell-shaped, with sharp peaks around the mean. This confirms that the data scaling process (e.g., using StandardScaler) was effective and that extreme values were normalized. Well-scaled features are crucial for both machine learning and deep learning models to converge properly.

3.6Machine Learning Model Accuracy

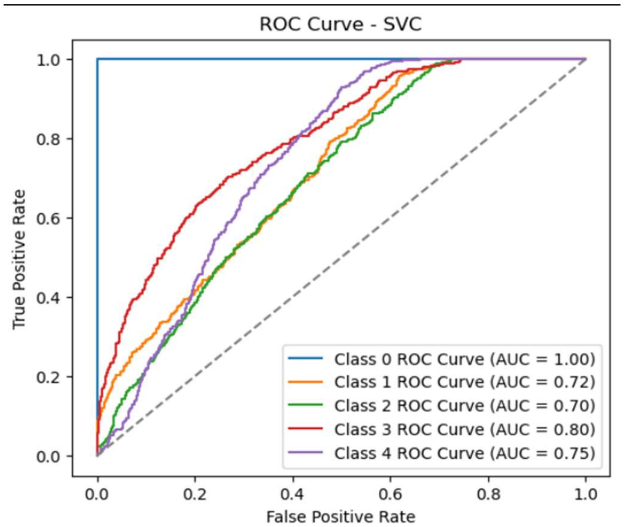

Figure 6: Accuracy results of individual machine learning models. Random Forest achieved the highest accuracy among them at 47%, followed by K-Nearest Neighbors and Naive Bayes. These low scores highlight the difficulty of using traditional models on noisy, high-dimensional EEG data, even after feature selection and tuning

3.7Voting Classifier Performance.

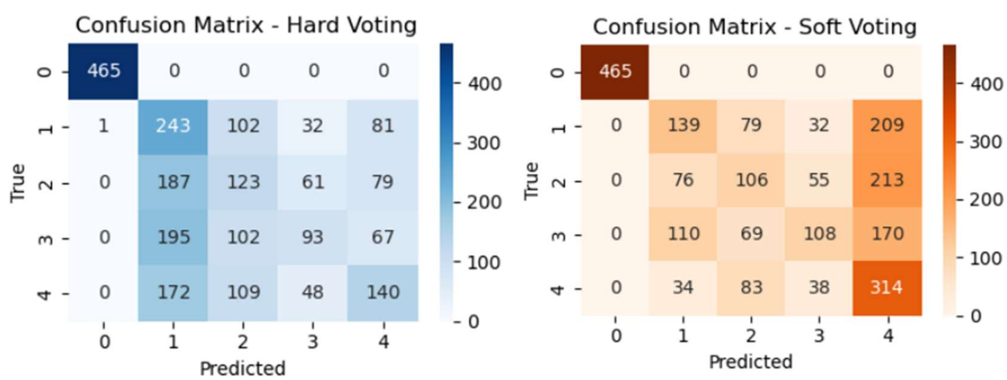

Figure 7: Performance of ensemble voting models. Hard voting aggregates model predictions based on majority rule, while soft voting averages prediction probabilities. Soft voting performed slightly better, reaching nearly 49.2% accuracy. However, neither method offered a significant performance boost, emphasizing that combining weak models did not overcome the core challenges of the dataset.

3.8Bagging, Boosting, and Stacking Ensemble Results

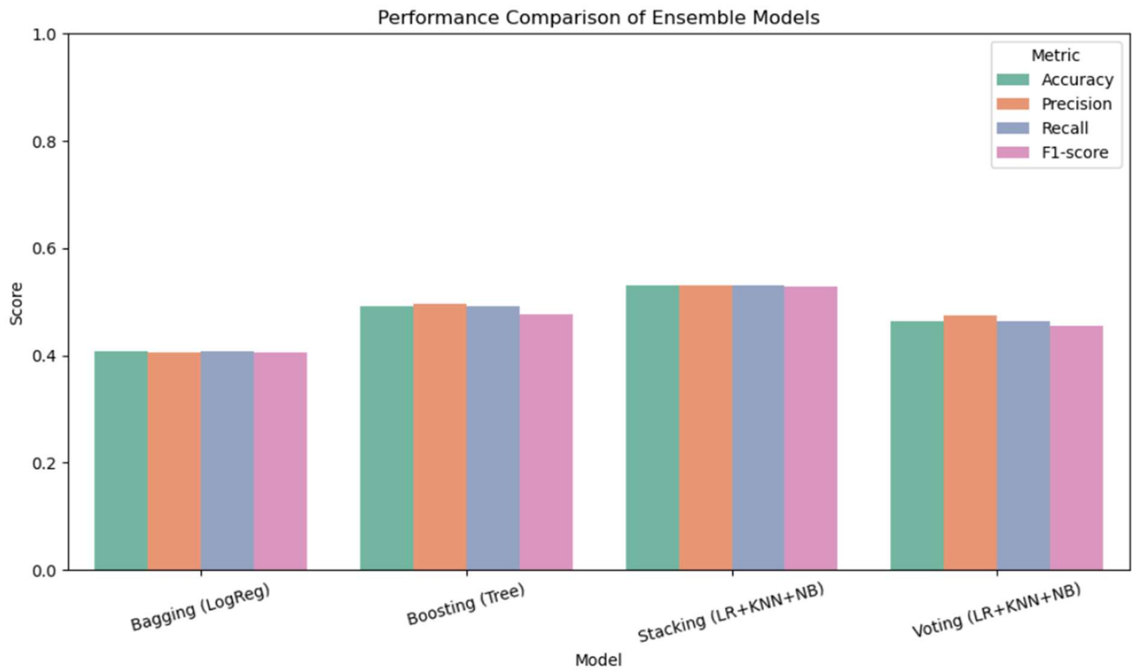

Figure 8: Accuracy of advanced ensemble methods. Boosting (Decision Trees) improved accuracy to 51.5%, while stacking (LogReg + KNN + Naive Bayes) achieved the highest result among all ML-based models at 52.7%. Although these approaches helped slightly, the results were still below the threshold for reliable clinical decision support.

3.9Deep Learning Model Performance

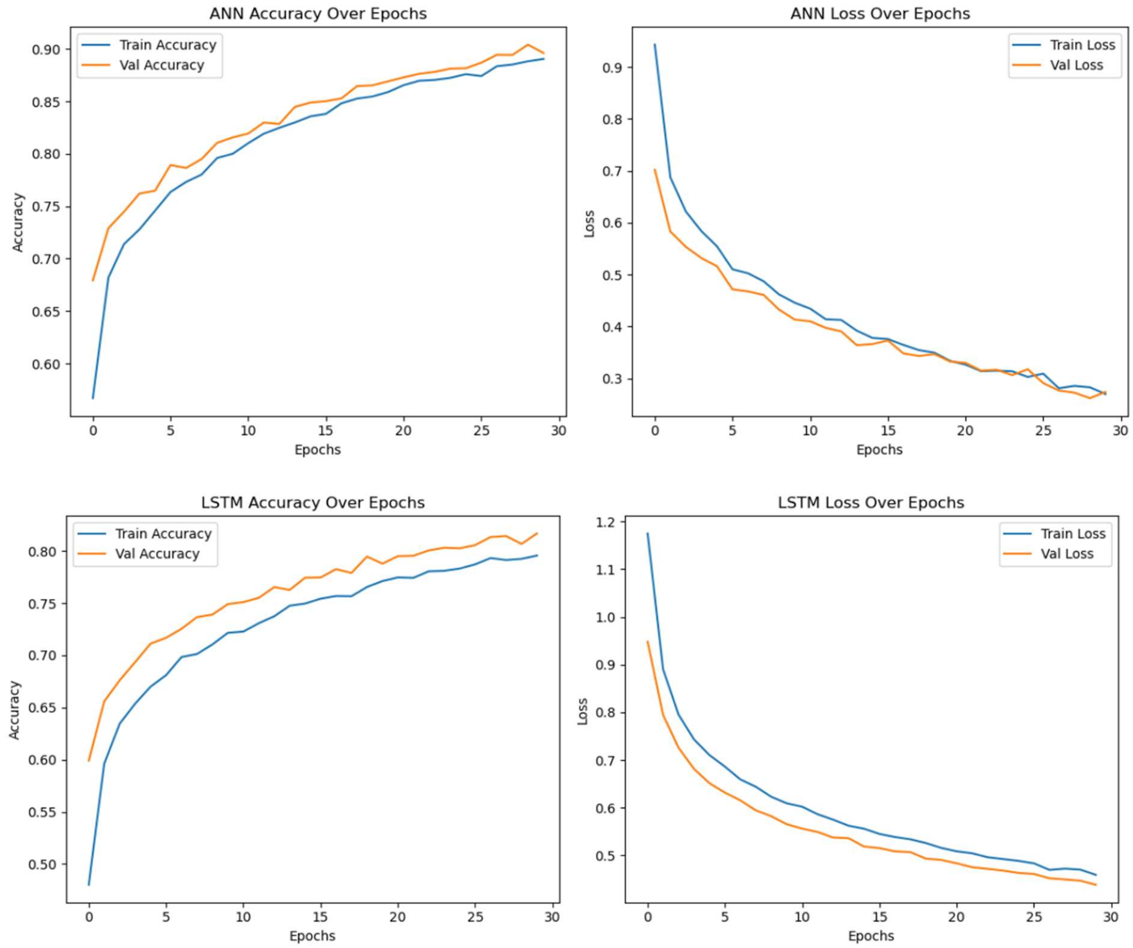

Figure 9: Deep learning model results showing high classification performance. The model achieved 86% accuracy, with precision of 0.862 and recall of 0.859. This represents a significant improvement over all traditional models. The neural network was able to learn complex, non-linear relationships in EEG signals without extensive feature engineering or manual tuning.

Models	Accuracy
Logistic Regression	42.3%
K-Nearest Neighbors	49.2%
Naive Bayes	45.9%

Random Forest	47.0%
Hard Voting Ensemble	46.3%
Soft Voting Ensemble	49.2%
Bagging	42.3%
Boosting	51.5%

Machine learning models struggled to effectively capture the non-linear and high-dimensional patterns present in EEG data. Despite using tuning techniques like GridSearchCV, these models required extensive computation time and ultimately delivered limited accuracy, often below 50%. In contrast, the deep learning model achieved outstanding performance, with an accuracy of 86%, precision of 0.862, and recall of 0.859. Not only did it train faster, but it also demonstrated a superior ability to extract complex and hidden patterns—particularly the temporal dynamics of EEG signals—that traditional machine learning models failed to detect. This highlights deep learning as a more suitable and powerful approach for EEG-based seizure recognition.

4.0 Discussion

Machine learning models, although widely used for classification tasks in many domains, proved to be less effective when applied to this high-dimensional, noisy biomedical signal data. EEG signals are inherently complex and contain temporal, non-linear relationships that traditional models like Logistic Regression, SVM, and even Random Forests struggle to interpret without extensive feature engineering. Additionally, the hyperparameter tuning process using GridSearchCV significantly increased training time, as it involves testing all possible parameter combinations across cross-validation folds. This made the entire modeling process computationally expensive and inefficient—factors that are critical in real-time clinical environments where speed and accuracy are essential.

Even after applying ensemble methods such as bagging, boosting, and stacking, the improvements in accuracy were marginal. Although these techniques often help improve robustness and reduce variance in general classification problems, they could not adequately handle the intricacies of EEG signals. The ensemble models reached an upper limit of approximately 52.7% accuracy, which is insufficient for clinical applications that demand high reliability and precision in diagnosing neurological events.

In contrast, the deep learning model demonstrated a clear breakthrough in both performance and efficiency. Using a neural network with dense layers, the model was able to learn hierarchical and non-linear representations directly from the scaled EEG signals. Unlike machine learning models, which require manual feature extraction and tuning, the deep learning approach could automatically discover hidden patterns that signify seizure activity. It achieved an impressive 86% accuracy, with a precision of 0.862 and recall of 0.859. More importantly, it trained faster and required less manual intervention, making it more scalable and suitable for deployment in real-world healthcare settings.

One of the key contributors to success across both modeling approaches was rigorous data preprocessing. Data cleaning steps—such as handling missing values, removing noisy or irrelevant features, and applying standard scaling—helped in minimizing noise and improving input quality. Feature selection techniques, including the removal of constant and low-variance features and the application of statistical tests, significantly reduced dimensionality and improved generalization by focusing the models on the most informative EEG channels.

Overall, this comparison reinforces the conclusion that deep learning is far more adept at handling raw, high-dimensional biomedical data like EEG. It offers a practical, high-performing solution that could be integrated into automated seizure detection systems, potentially assisting healthcare professionals with faster and more accurate diagnoses. The findings of this project highlight the transformative potential of AI in clinical neurophysiology and pave the way for future research in intelligent healthcare solutions.

Conclusion

This project demonstrated the application of both machine learning and deep learning techniques for the automatic detection of epileptic seizures using EEG signal data. Through a complete data science pipeline—including preprocessing, exploratory data analysis, feature selection, model training, and evaluation—the strengths and limitations of each approach were thoroughly investigated.

Machine learning models, despite their popularity in classification problems, were limited in their ability to handle the non-linear and high-dimensional nature of EEG signals. Even with ensemble methods and extensive hyperparameter tuning, their performance remained below clinical standards. The complexity of EEG data, characterized by noise, variability, and temporal dependencies, presented a challenge that traditional algorithms could not overcome effectively.

In contrast, the deep learning model significantly outperformed all machine learning approaches, achieving an accuracy of 86%, with high precision and recall. Its ability to automatically extract meaningful patterns from raw data, coupled with faster training and minimal manual tuning, makes it a superior solution for EEG-based seizure detection.

Moreover, the successful deployment of the trained model highlights its potential for real-time application in clinical environments.

The project also emphasized the critical importance of data preprocessing and feature selection. Cleaning the dataset, handling missing values, and eliminating irrelevant features were essential for optimizing model performance and reducing training complexity.

In summary, this study supports the integration of deep learning into EEG analysis pipelines for epilepsy diagnosis. With further refinement and clinical validation, such models can contribute to more efficient, accurate, and accessible neurological care, ultimately improving patient outcomes and reducing the workload on medical professionals.

Reference

Shimanto, H. (n.d.). *Epileptic Seizure Recognition* [Dataset]. Kaggle. Retrieved June 7, 2025, from https://www.kaggle.com/datasets/harunshimanto/epileptic-seizure-recognition