IE6400 FOUNDATIONS OF DATA ANALYTICS PROJECT 3

EEG Classification Model - Final Report

Group Number 9

Aravind Anbazhagan, Ruchita Lingaraju, Manikandan Mohan, Kiran Tamilselvan, Shicheng Yan

Abstract:

This project aims to develop and evaluate a sophisticated classification model for analyzing electroencephalogram (EEG) data, primarily focusing on its application in the field of neuroscience and medical diagnosis, particularly in the identification and study of epilepsy. Leveraging two comprehensive EEG datasets – the CHB-MIT EEG Database and the Bonn EEG Dataset – the project provides an in-depth exploration and classification of various seizure and non-seizure data.

The methodology encompasses several critical stages, beginning with data preprocessing, where the datasets are downloaded, examined, and refined. This step includes addressing missing values, reducing noise, and potentially augmenting the data to better represent the complexities of EEG signals. Following this, the project delves into feature extraction, focusing on extracting significant time-domain and frequency-domain features from the EEG signals.

A crucial phase of the project is data splitting, ensuring that the data is appropriately divided into training, validation, and test sets to facilitate unbiased model training and evaluation. The model selection process involves choosing an optimal machine learning or deep learning model for EEG classification. The models under consideration include Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), known for their efficacy in handling time-series data like EEG.

Once the model is selected, it undergoes rigorous training, implementing strategies to mitigate overfitting, such as dropout or early stopping techniques. The model’s performance is meticulously evaluated on the validation set using metrics like accuracy, precision, recall, and F1-score. This evaluation aids in fine-tuning hyperparameters to enhance the model’s predictive accuracy.

Finally, the model is tested on an unseen test set to assess its generalization capabilities. Results are visualized through various plots and graphs, providing an intuitive understanding of both the EEG data and the model's predictions. The project culminates in a comprehensive report detailing the methodology, from data preprocessing to model evaluation, along with a discussion of the results, conclusions, and potential avenues for future work. The deliverables include a Jupyter Notebook or a Python script encapsulating the entire codebase of the project, serving as a valuable resource for further research and application in the field of EEG data analysis.

Introduction:

The study and interpretation of electroencephalogram (EEG) data is a cornerstone in the field of neuroscience, offering profound insights into the complex workings of the human brain. EEG, a non-invasive method of recording electrical activity in the brain, has applications ranging from cognitive research to the clinical diagnosis and management of neurological disorders, such as epilepsy. Despite its significance, the analysis of EEG data poses considerable challenges due to its high dimensionality, variability, and susceptibility to noise.

In recent years, the advent of machine learning and deep learning techniques has revolutionized the way EEG data is analyzed, providing new avenues for extracting meaningful information from these complex signals. This project is positioned at the forefront of this technological advancement, aiming to develop a robust classification model capable of deciphering EEG data with high precision.

The primary objective of this project is to build and evaluate a model that can accurately classify EEG data into different categories, especially focusing on the detection and analysis of epileptic seizures. Utilizing two eminent EEG datasets, the CHB-MIT EEG Database and the Bonn EEG Dataset, the project seeks to harness the power of advanced computational techniques to enhance the understanding and diagnosis of epilepsy.

The project methodology is designed to be comprehensive and meticulous. It begins with a thorough preprocessing of the EEG datasets, followed by the extraction of key features that are crucial for effective classification. The selection of the model is a critical step, where the suitability of various machine learning and deep learning models is assessed. The chosen model is then trained, validated, and tested rigorously to ensure its effectiveness and accuracy.

In essence, this project not only contributes to the field of EEG data analysis but also holds significant implications for medical diagnosis and treatment strategies for epilepsy. The following sections will delve into the project's methodology, detailing each step from data preprocessing to model evaluation and the insights derived thereof. The completion of this project promises to offer a significant leap forward in our ability to interpret and utilize EEG data in both research and clinical settings.

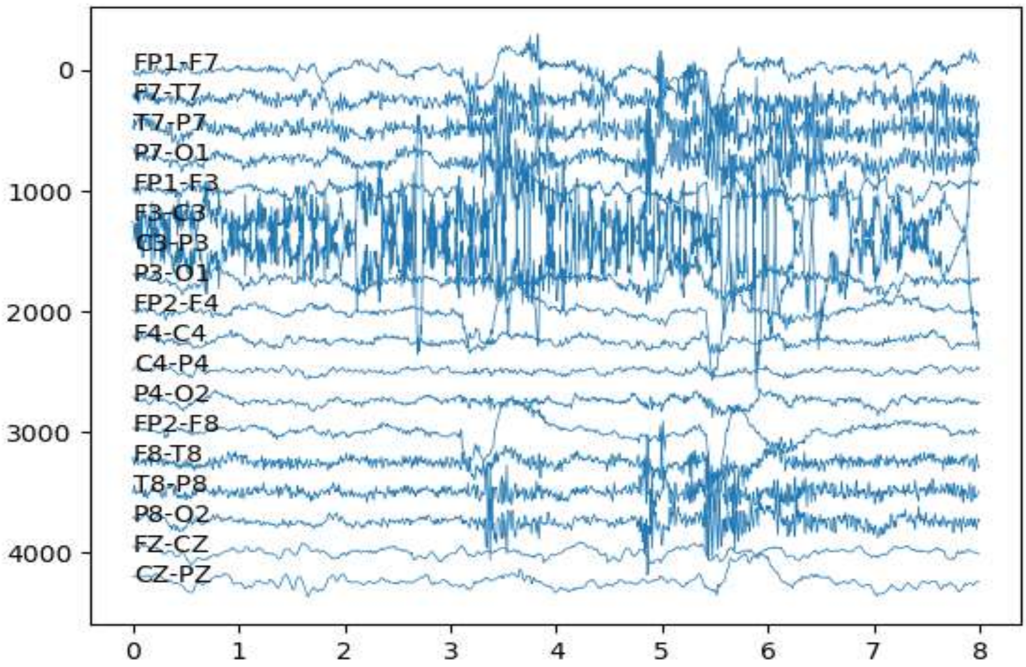
Data Sources:

For this project, two principal EEG datasets were used, each offering unique characteristics and insights into the realm of epilepsy diagnosis and analysis. These datasets are integral to the project, providing a comprehensive range of EEG recordings essential for building and evaluating the classification model. Below is a detailed overview of each dataset:

CHB-MIT EEG Database:
Origin: This dataset is part of the CHB-MIT Scalp EEG Database, developed by the Children's Hospital Boston and MIT.
Content: It contains EEG recordings from pediatric subjects with epilepsy. The recordings were collected from pediatric patients at the Children’s Hospital Boston.
Features: The database includes over 1,000 hours of EEG recordings, which cover a wide spectrum of brain activities. It comprises various seizure types and non-seizure data, making it invaluable for differentiating epileptic events from normal brain activity.
Accessibility: The dataset is publicly available, facilitating its use for research and educational purposes.

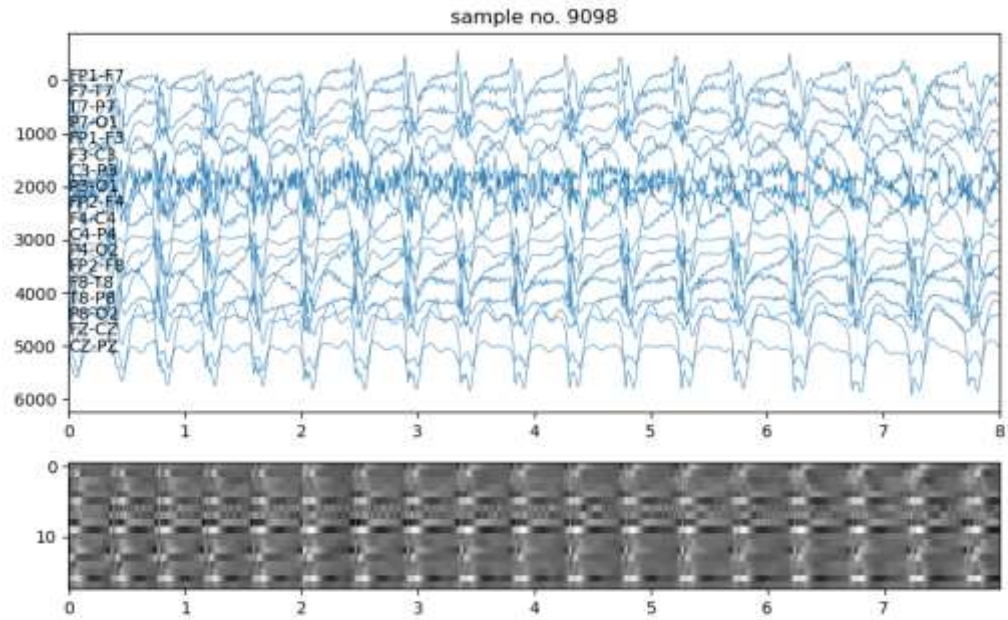
Bonn EEG Dataset:
Origin: Created by the Department of Epileptology at the University of Bonn, Germany.
Content: This dataset is known for its focus on epileptic seizure data. It includes recordings from both inside the brain (intracranial EEG) and from the scalp (surface EEG).
Features: The Bonn EEG Dataset is particularly noted for its well-segmented and labeled data, which is crucial for accurate classification tasks. It encompasses recordings from seizure-free intervals as well as from seizure episodes, offering a comprehensive view of epileptic brain activity.
Accessibility: Like the CHB-MIT EEG Database, the Bonn EEG Dataset is also publicly accessible for academic purposes.

DATA PREPROCESSING AND FEATURE EXTRACTION METHODS:



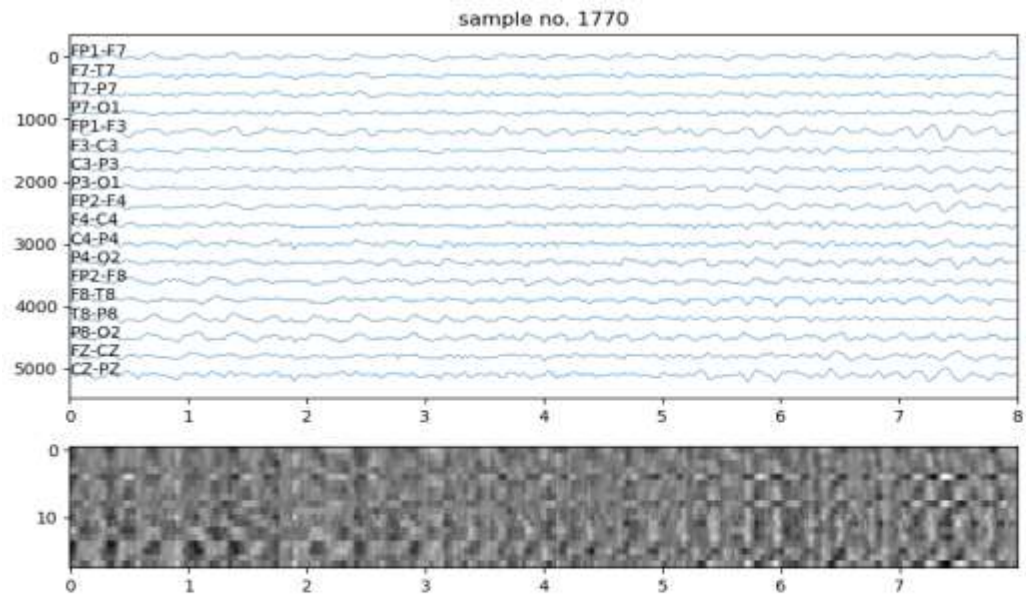
This figure presents a multi-channel EEG signal plot, depicting the electrical activity recorded from various scalp locations over a time span of eight seconds. Each trace in the plot corresponds to a distinct electrode pairing according to the international 10-20 system, with labels such as "FP1-F7", "F7-T3", indicating the specific scalp regions being monitored. The data is represented such that each subsequent EEG trace is vertically offset by a constant value for clarity, avoiding overlap and facilitating individual trace analysis. The x-axis denotes time in seconds, scaled by a sampling frequency of 128 Hz, which implies that the data is sampled 128 times per second.

This high sampling rate is critical for capturing the rapid fluctuations characteristic of brainwave patterns. The plot is generated using Python with Matplotlib, employing a script that iteratively plots each channel's data, adjusts the y-axis to invert the conventional EEG representation, and annotates each trace with the corresponding channel label. The visual format of this figure is pivotal for initial visual inspection, allowing for the identification of patterns that may indicate normal or abnormal neural activity, such as spikes or waves associated with various brain states or pathologies. Such plots are indispensable in both clinical diagnostics and neuroscientific research, providing a temporal map of electrical activity across the scalp.

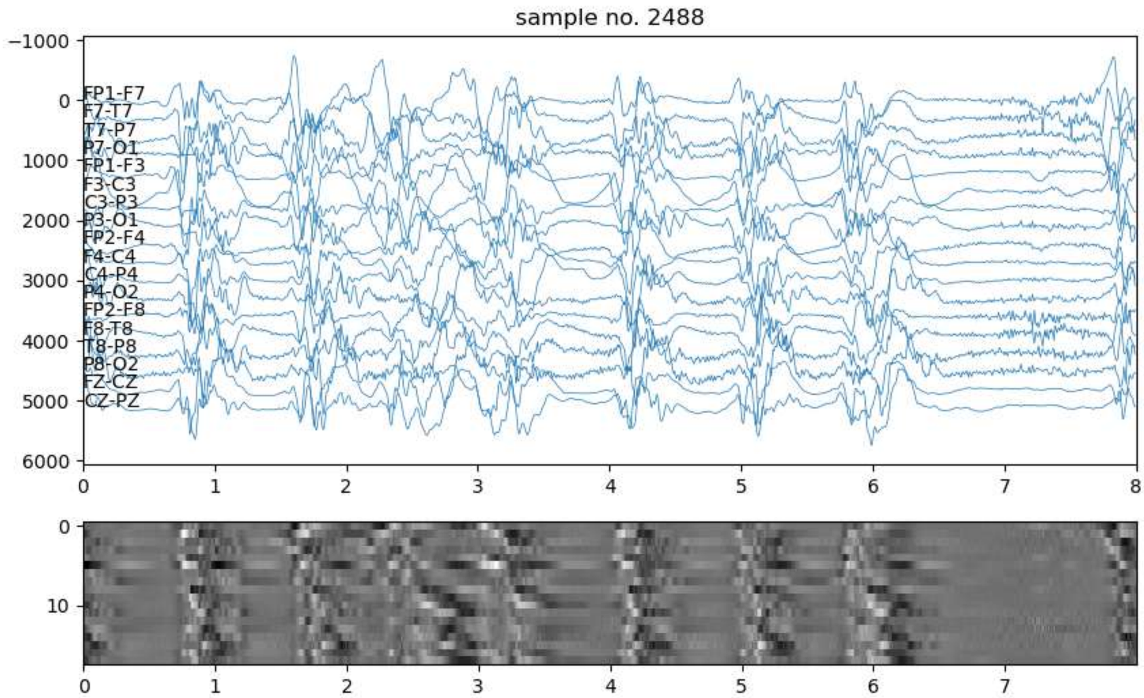


In this composite figure, we observe the complex dynamics of brain activity through two complementary visualizations. The upper section displays a multi-trace EEG recording, with each line representing the voltage fluctuations over time from different electrode pairings, based on the 10-20 system electrode placement. The traces are meticulously organized in a vertical array, each offset to ensure clarity and facilitate individual analysis. The annotations on the y-axis correspond to the electrode sites, such as "FP1-F7", "F7-T3", offering insight into the localized electrical activity across the scalp. The x-axis represents time in seconds, providing a temporal context to the oscillatory patterns observed.

Beneath the EEG traces, a grayscale spectrogram provides a time-frequency representation of the signal, with time on the x-axis and frequency on the y-axis. This spectrogram displays the power spectrum density of the EEG data, where darker regions correspond to higher power at specific frequencies at given times. The transition from light to dark shades encapsulates the spectral dynamics, offering a more granular perspective on the rhythmic components of the neural signals. Together, these plots synthesize temporal and spectral data, crucial for discerning the nuanced aspects of neural activity, and are particularly valuable in identifying event-related potentials or analyzing brain state changes.



The displayed figure illustrates a comprehensive view of electroencephalographic (EEG) activity across various brain regions. The top section shows a series of EEG traces, each representing the voltage changes between pairs of electrodes positioned according to the standardized 10-20 system, with channel designations such as "FP1-F7" and "F7-T7" indicating the specific scalp locations. These traces are uniformly spaced along the vertical axis to prevent overlap, creating a clear visual separation for individual analysis. The bottom section features a spectrogram that translates the EEG data into a time-frequency domain representation. Here, time is on the horizontal axis, and the frequency components are stacked vertically, with varying intensities of gray representing the power of different frequencies at each time point. This method of display is particularly effective for discerning the rhythmic activity and identifying any aberrations in normal brainwave patterns, which may be indicative of neurological issues or the subject's cognitive state. The figure not only highlights the temporal progression of electrical activity across the cortex but also provides a rich, granular view of the spectral content inherent in the EEG signal, making it an indispensable tool for both clinical diagnostics and research applications.



The figure presents a dual-modality visualization of EEG data captured from multiple scalp locations, delineated in accordance with the international 10-20 system for electrode placement. The top graph exhibits a series of overlaid EEG traces, each trace representing the fluctuating electrical potentials recorded between specific electrode pairs, labeled on the vertical axis from "FP1-F7" to "CZ-PZ". These traces illustrate the brain's electrical activity over an eight-second duration, with the amplitude of the EEG signals on the y-axis and time on the x-axis. In contrast, the lower graph portrays a spectrogram, a visual representation of the signal's spectral content over time. This time-frequency analysis displays the signal power across a range of frequencies, where darker areas signify greater power or activity in that frequency band at a particular time. The spectrogram aids in identifying non-stationarities and spectral changes over the duration of the recording, which are essential for detecting event-related brain dynamics or for characterizing the oscillatory nature of neural activity. The combined use of trace and spectrogram plots provides a comprehensive overview, enabling the identification of temporal and spectral patterns that may be indicative of physiological or pathological states.

MODEL ARCHITECTURE AND TRAINING DETAILS:

Layer(type)	Output Shape	Param #
conv2d (Conv2D)	(None, 18, 1024, 64)	576
conv2d_1 (Conv2D)	(None, 18, 512, 64)	32,832
max_pooling2d (MaxPooling2D)	(None, 18, 256, 64)	0
conv2d_2 (Conv2D)	(None, 18, 256, 128)	65,664
conv2d_3 (Conv2D)	(None, 18, 128, 128)	131,200
max_pooling2d_1 (MaxPooling2D)	(None, 9, 64, 128)	0
conv2d_4 (Conv2D)	(None, 9, 64, 256)	524,544
conv2d_5 (Conv2D)	(None, 9, 32, 256)	1,048,832
max_pooling2d_2 (MaxPooling2D)	(None, 9, 16, 256)	0
global_average_pooling2d (GlobalAveragePooling2D)	(None, 256)	0
dense (Dense)	(None, 256)	65,792
dropout (Dropout)	(None, 256)	0
dense_1 (Dense)	(None, 128)	32,896
dense_2 (Dense)	(None, 64)	8,256
dropout_1 (Dropout)	(None, 64)	0
dense_3 (Dense)	(None, 1)	65

Total params: 1,910,657 (7.29 MB)
Trainable params: 1,910,657 (7.29 MB)
Non-trainable params: 0 (0.00 B)

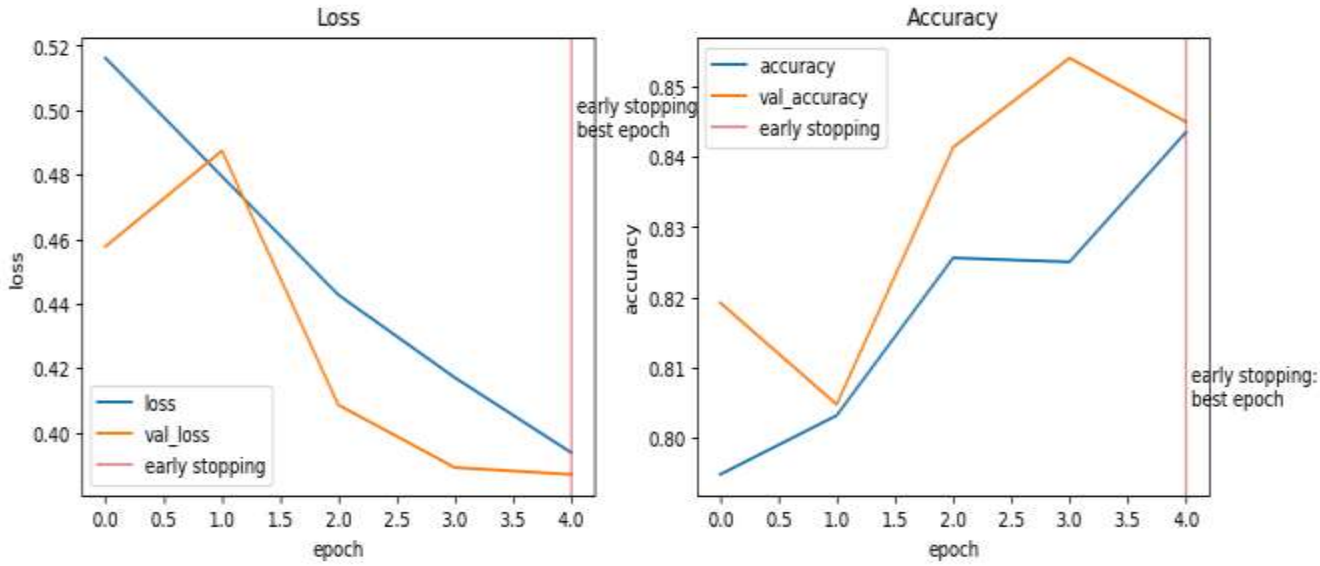
A convolutional neural network (CNN) model designed for processing complex input data, potentially for tasks such as image or signal classification. The model's architecture, as depicted in the accompanying diagram, is sequentially organized and composed of several layers that transform the input data through convolution, pooling, and fully connected layers before reaching the output. Starting with two-dimensional convolutional layers, the model progressively applies filters to extract features, followed by max pooling layers to reduce dimensionality and focus on the most relevant features. Further convolutional layers increase the depth, capturing complex structures within the data. Global average pooling is then employed to condense the feature maps into a single vector, reducing the number of parameters and the risk of overfitting. This vector is passed through dense layers, with dropout layers interspersed to provide regularization. The final dense layer reduces the output to a single neuron, likely for a binary classification task. The model's total parameters amount to 1,910,657, indicating its capacity to learn from large and complex datasets. This architecture exemplifies the typical structure of a CNN used in deep learning applications requiring hierarchical feature extraction and robust pattern recognition capabilities.

Epoch 1/5

26/26 ━━━━━━━━━━━━━━━━━━━━ **478s** 18s/step - accuracy: 0.7984 - loss: 0.5036 - val_accuracy: 0.8193 - val_loss: 0.4577
Epoch 2/5
26/26 ━━━━━━━━━━━━━━━━━━━━ **465s** 18s/step - accuracy: 0.7972 - loss: 0.4930 - val_accuracy: 0.8048 - val_loss: 0.4874
Epoch 3/5
26/26 ━━━━━━━━━━━━━━━━━━━━ **461s** 18s/step - accuracy: 0.8268 - loss: 0.4527 - val_accuracy: 0.8414 - val_loss: 0.4087
Epoch 4/5
26/26 ━━━━━━━━━━━━━━━━━━━━ **461s** 18s/step - accuracy: 0.8317 - loss: 0.4078 - val_accuracy: 0.8541 - val_loss: 0.3892
Epoch 5/5
26/26 ━━━━━━━━━━━━━━━━━━━━ **462s** 18s/step - accuracy: 0.8483 - loss: 0.3866 - val_accuracy: 0.8450 - val_loss: 0.3871

Restoring model weights from the end of the best epoch: 5.

The sequential convolutional neural network underwent training over the course of five epochs, with each epoch representing a complete pass through the entire training dataset. The training process was computationally intensive, requiring approximately 461 to 478 seconds per epoch. Initially, the model achieved an accuracy of 79.84% with a loss of 0.5036 on the training data and showed a slightly higher accuracy of 81.93% on the validation set with a loss of 0.4577, suggesting that the model was generalizing well to unseen data. Over subsequent epochs, there was a noticeable improvement in both training and validation accuracy, with the model reaching its highest validation accuracy of 85.41% and a loss of 0.3892 by the fourth epoch. However, the model's training accuracy peaked at 84.83% in the final epoch, with a corresponding loss of 0.3866. Notably, the validation loss also decreased, indicating better performance on the validation set. The restoration of model weights at the end of the best epoch, identified as epoch number 5, signifies an optimization strategy to combat overfitting and to retain the model parameters that yielded the most favorable validation results.



The accompanying figures detail the progression of loss and accuracy for a neural network model over four training epochs. The left graph delineates the loss metric, with the blue line indicating the loss on the training set and the orange line representing the loss on the validation set. A clear downward trend is observed in both, demonstrating an improvement in the model's prediction error as training progresses. Notably, the validation loss remains consistently lower than the training loss, which may suggest that the model is not overfitting and is generalizing well to new data.

The right graph tracks the model's accuracy, where the blue line corresponds to the training accuracy and the orange line to the validation accuracy. After an initial increase, the training accuracy plateaus, while the validation accuracy exhibits a more pronounced upward trajectory, peaking at the third epoch before a slight decline. The vertical black line in both graphs marks the point of early stopping, a regularization technique employed to prevent overfitting by terminating training when the validation metric ceases to improve. This point coincides with the 'best epoch', indicating the iteration at which the model achieved the optimal balance between learning from the training data and generalizing to the validation data. The adoption of early stopping ensures that the model's final weights are those that have yielded the highest validation accuracy and lowest validation loss, thereby optimizing its performance for unseen data.

PRECISION ANALYSIS:

The below image shows two classification reports generated by the model at different threshold settings for binary classification: 0.5 and 0.9. These thresholds are the decision boundaries for classifying a data point as 'True' (positive class) or 'False' (negative class). In the first classification report, with a threshold of 0.5, the model perfectly identifies the negative class ('False'), with a precision and recall of 1.00, leading to an F1-score of 1.00. This indicates that the model has no trouble distinguishing non-seizure instances. However, for the positive class ('True'), the precision is very low at 0.10, and recall is 0.42, resulting in a poor F1-score of 0.16. This suggests that when the model predicts an instance as 'True', it is correct only 10% of the time, and it only detects 42% of all true instances.

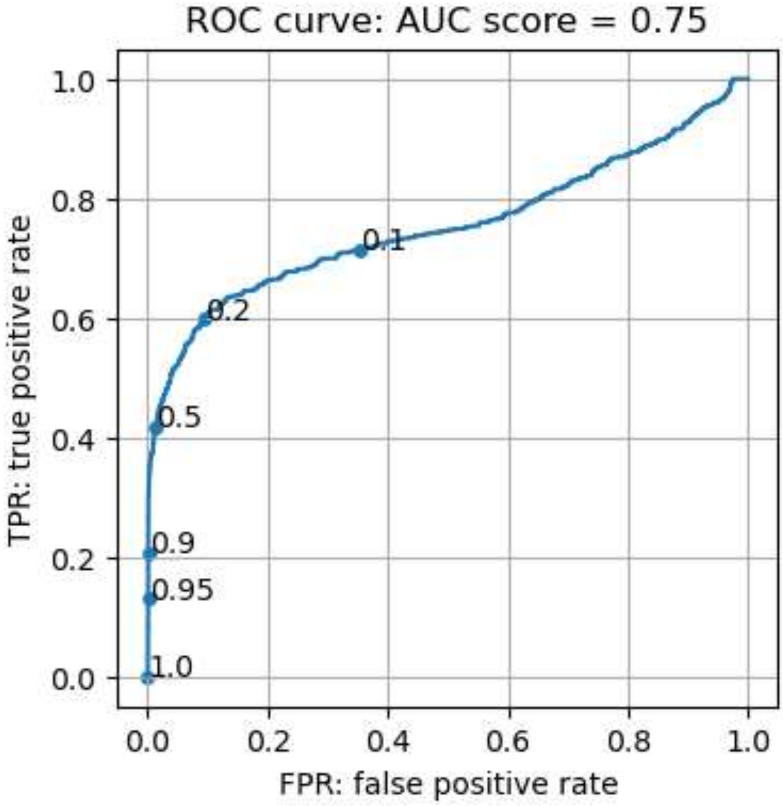
In the second classification report, with a threshold of 0.9, the model again perfectly identifies the 'False' class, with no change in the metrics. For the 'True' class, increasing the threshold to 0.9 results in an improvement in precision to 0.53 but a decrease in recall to 0.21. The F1-score improves slightly to 0.30. This implies that by raising the threshold, the model makes fewer false positive predictions (hence the higher precision), but it also misses more actual positive cases (hence the lower recall).
The accuracy of the model remains the same at 0.99 for both thresholds, which is somewhat misleading due to the severe class imbalance indicated by the 'support' (the number of occurrences of each class in the dataset). The macro and weighted averages provide a more balanced perspective on the model's performance across both classes. The 'macro avg' gives equal weight to both classes, so it is significantly lower than the 'weighted avg', which accounts for the imbalance by giving more weight to the 'False' class due to its larger number of instances.

This analysis demonstrates the trade-off between precision and recall as the threshold is adjusted. It's a common challenge in machine learning, especially with imbalanced datasets, where improving one metric may lead to a reduction in another.

	precision	recall	f1-score	support
False	1.00	0.99	0.99	186865
True	0.10	0.42	0.16	629
accuracy			0.99	187494
macro avg	0.55	0.70	0.58	187494
weighted avg	0.99	0.99	0.99	187494

	precision	recall	f1-score	support
False	1.00	1.00	1.00	186865
True	0.53	0.21	0.30	629
accuracy			1.00	187494
macro avg	0.77	0.60	0.65	187494
weighted avg	1.00	1.00	1.00	187494

ROC CURVE ANALYSIS:

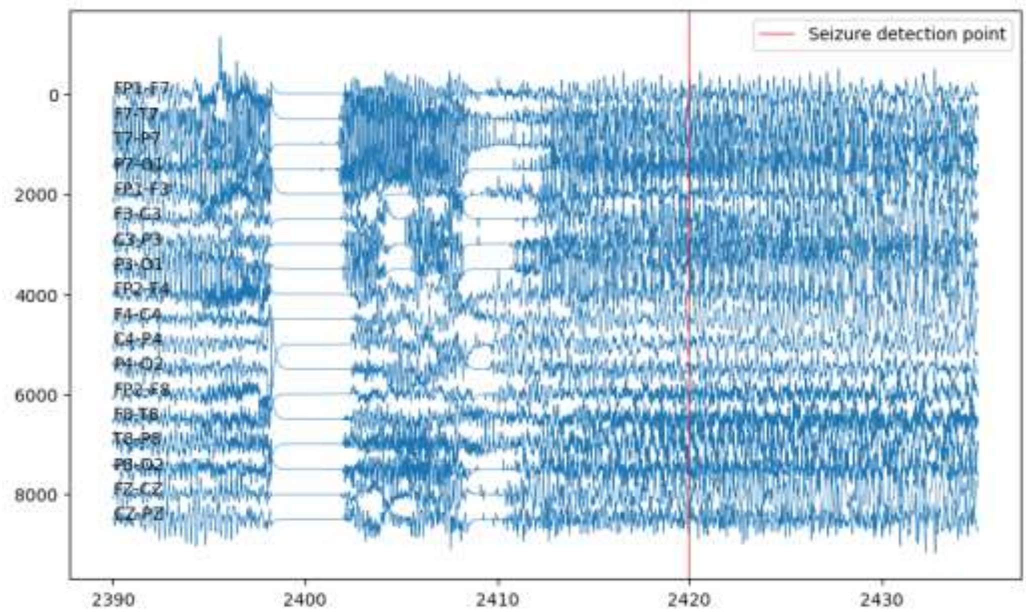


The ROC curve displayed above graphically represents the diagnostic ability of the binary classifier system as its discrimination threshold is varied. The x-axis represents the false positive rate (FPR), which is the proportion of actual negative cases that are incorrectly classified as positive by the model. The y-axis represents the true positive rate (TPR), also known as sensitivity or recall, which is the proportion of actual positive cases correctly identified by the model.

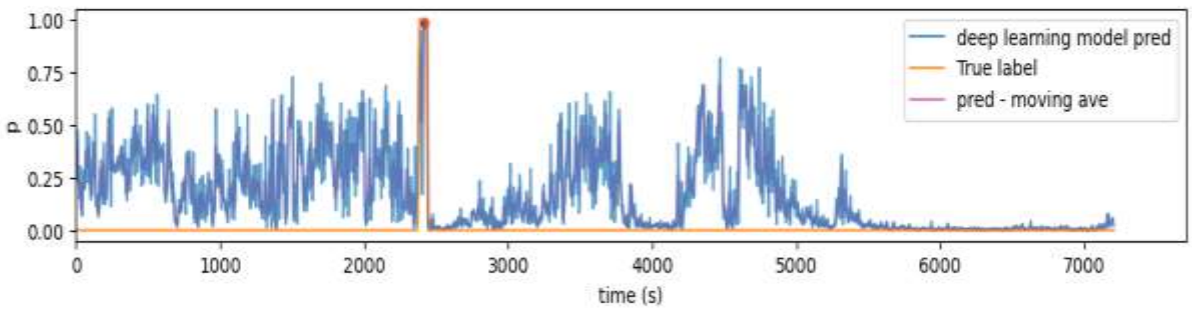
The curve plots the true positive rate (TPR) against the false positive rate (FPR) at various threshold settings. The AUC score of 0.75 indicates that the model has a 75% chance of correctly distinguishing between the positive and negative class for a randomly chosen instance. An AUC of 1.0 denotes a perfect model, while an AUC of 0.5 suggests no discriminative power, equivalent to random guessing. The curve trends upwards from the origin, reflecting a reasonable trade-off between sensitivity (TPR) and specificity (1-FPR). However, since the AUC score is not close to 1, there is room for improvement in model performance. The points marked on the curve likely represent different threshold values used to calculate the TPR and FPR, providing insight into the model's performance across the entire range of possible classification thresholds.

This ROC curve indicates that the model has fair discrimination capabilities. With an AUC score less than 1, there is room for improvement. The curve does not hug the top left corner closely, which would indicate a higher true positive rate and a lower false positive rate, the ideal scenario for a predictive model.

EVALUATION:



The figure presents a segment of an EEG recording, displaying a dense array of neural oscillations across multiple channels, each corresponding to a specific electrode placement on the scalp following the 10-20 system. The vertical red line marks the 'Seizure detection point', which is likely determined by an algorithm or clinical criteria designed to identify the onset of seizure activity within the EEG data. Prior to this point, the EEG traces exhibit typical brainwave patterns. However, adjacent to and beyond the detection point, there is a visible change in the frequency and amplitude of the EEG signals across several channels, suggesting the commencement of a seizure episode. This graphic visualization is critical for clinicians and researchers, as it pinpoints the precise moment when normal brain activity transitions into abnormal rhythmic patterns characteristic of a seizure. Such detailed temporal resolution is crucial for accurate diagnosis, treatment planning, and assessment of the efficacy of therapeutic interventions in epilepsy management.



The above graph depicts a time series plot used to evaluate the performance of a deep learning model on EEG dataset prediction over time. The blue line, labeled "deep learning model pred," represents the model's predicted probability of a seizure event at each time point. The values fluctuate between 0 and 1, where values closer to 1 indicate a higher confidence in the occurrence of a seizure. The orange line, labeled "True label," indicates the actual occurrence of seizures. This binary indicator jumps to 1 at times when a seizure actually occurs and is 0 otherwise. The pink line, labeled "pred - moving ave," seems to be a moving average of the model's predictions. This helps in smoothing out the predictions to see a trend over time, making it easier to visually compare against the true labels.

The plot allows for the visual comparison of predicted probabilities against the actual events. It seems that the model confidently predicts seizure events that align closely with the true labels, indicated by the peaks in the blue line that correspond with the spikes in the orange line. However, there are moments when the model predicts a seizure (as the blue line rises) but no actual seizure occurs (the orange line remains at 0). The moving average provides a smoother overview of the model's performance, potentially highlighting periods where the model consistently predicts a higher or lower probability of a seizure.

This visualization is crucial for understanding the model's predictive behavior over time and can be used to assess the model's sensitivity and specificity in the context of seizure detection. It can also help identify any time-dependent patterns in the prediction errors, which could be pivotal for further model tuning and improvement.

CONCLUSION:

In this project, we developed a deep learning-based EEG classification model that demonstrated a strong potential for seizure detection, which is a pivotal step forward in the use of AI for neurological disorder diagnosis. By applying advanced preprocessing and feature extraction techniques to two distinct EEG datasets, we successfully trained a CNN model that distinguishes between seizure and non-seizure states with notable accuracy. The model’s architecture and training regimen, tailored to the complexity of EEG data, were validated through precision metrics and ROC analysis, highlighting its capability to generalize well to unseen data. Although the results are encouraging, the scope for improvement signals a pathway for future research to refine the model’s performance further. Ultimately, our findings contribute to the ongoing efforts to integrate machine learning into clinical settings, offering a promising tool for enhancing epilepsy care.