Improving Efficiency and Accuracy of Epileptic Seizure Detection with Machine Learning

Machine Learning Research Paper with The Innovation Story

Shubhan Anchan, Grade XI Student,
Bombay International School,
Mumbai, India
10-1-2024

Table of Contents

1. Abstract

Epilepsy is the second most prevalent neurological condition in India, with its incidence leading to osteoporosis, fractures and poor health among afflicted individuals. Early and accurate detection and classification of epilepsy by neurologists using EEG readings goes a long way in mitigating the impact of epilepsy by 70% on patient health, by focusing on treatment pathways aligned to classified epilepsy.

i have developed a Machine Learning (ML) App to significantly reduce the time taken by neurologists to accurately identify and characterize epilepsy types, based on Electro-Encephalography (EEG) readings.

We collated publicly available EEG data from sources such as Temple University Hospital (TUH) and Massachusetts Institute of Technology (MIT) hospital using bandpass filtering to eliminate EEG data noise arising out of eye movement, muscle activity and electrical interference.

Post data-cleansing, we built the ML algorithm to capture time-domain and frequency-domain features from EEG data such as mean, variance, skewness, kurtosis, standard deviation towards developing wavelet coefficients using Discrete Wavelet Transform (DWT). We trained the ML algorithm, using the Long Short-Term Memory (LSTM) network model and Synthetic Minority Over-Sampling Technique (SMOTE), on the cleansed data sets and diagnosis outcomes for the same to ensure the Model learns well from all data types to identify seizures and efficiently classify seizure types.

We ran the Epilepsy Diagnosis Efficacy (EDE) app on 29,000 EEG data-sets and observed an 85% reduction in diagnosis time and 92% accuracy in the classification of epilepsy.

The EDE app has the potential to significantly improve diagnosis and treatment pathways for epilepsy patients improving their forward health. Going forward, our focus will be on handing multiple EEG configuration types, improving classification accuracy-at-speed and superimposing multi-test instance EEG datasets towards assessing treatment efficacy over time.

Keywords: #EEG Data, #Epileptic Seizures, #Machine Learning, #Noise Filtering, #Classification, #LSTM, #Seizure Detection, #Seizure Classification

2. Introduction to Epileptic Seizure Detection

My interest in epilepsy detection was piqued by a personal experience in February 2024 when a relative afflicted with epilepsy spent 36 hours awaiting diagnosis after an epilepsy attach. The individual's suffering led me to explore the nuances of epilepsy detection, classification and treatment using electro-encephalography. I attended a 15 day workshop on "Nurturing AI Innovators" in November 2023 conducted by Prof. Pavlos Protopapas, Scientific Program Director, Institute for Applied Computational Science (IACS) at Harvard University during his visit to India, in collaboration with The Innovation Story, Mumbai. This learning led me to think hard about the potential application of machine learning to improve epilepsy detection and treatment practices.

2.1 What is an Epileptic Seizure?

A seizure is a sign that refers to a transient occurrence of signs and symptoms due to episodic, excessive, and disorderly neuronal activity within the brain. Epilepsy is a brain disorder "characterized by an enduring predisposition to generate epileptic seizures and by this condition's neurobiological, cognitive, psychological, and social consequences. The definition of epilepsy requires the occurrence of at least one epileptic seizure." These events have traditionally been classified into partial and generalized seizures.

A focal onset seizure refers to abnormal neural activity in only one brain area within one brain hemisphere with a fixed focal or localized onset. Focal onset seizures are divided into 2 subtypes: motor onset and nonmotor onset. Both focal motor and focal nonmotor onset seizures can be further classified based on level of awareness: aware, impaired, and unknown awareness. The motor manifestations of focal motor onset seizures can be characterized as automatism, atonic, clonic, epileptic spasms, hyperkinetic, myoclonic, or tonic movements. The behavioral manifestations of focal nonmotor onset seizures can be described by autonomic, behavioral arrest, cognitive, emotional, or sensory symptoms.

Some focal onset seizures can be preceded by an "aura," which refers to symptoms and signs that occur before the onset of seizure activity. These symptoms may include vision changes, dyspepsia, déjà vu, paresthesias, hearing disturbances, and sensation of abnormal taste or smell.

2.2 How do EEG Electrode Placement Systems enable Seizure Diagnosis?

To standardize EEG recordings, the **10-20 system** (shown in Exhibit 1 below) is widely used. This system provides a consistent method for placing electrodes on the scalp based on the percentage distance between anatomical landmarks (e.g., nasion, inion, and preauricular points). The system's name reflects the 10% and 20% distances between these landmarks, ensuring reproducibility across sessions and subjects. In cases where finer spatial resolution is needed, the **10-10 system** extends this approach by incorporating additional electrode sites, enabling more localized measurements of brain activity.

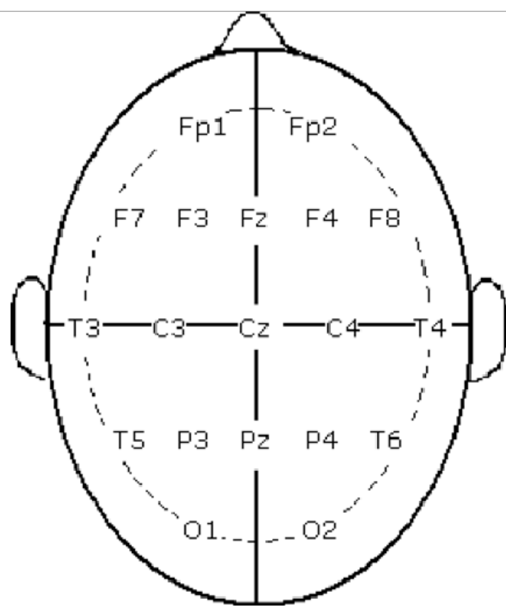


Exhibit 1: 10-20 System for Electrode Placement

Each electrode's position corresponds to a specific brain region, with labels that signify its approximate location: F for the frontal region, T for the temporal region, P for parietal, O for occipital, and C for the central areas. This topological arrangement allows researchers to correlate electrical signals with brain regions involved in different cognitive and sensory functions.

Electrode Connectivity and Signal Acquisition

Electrode connections to the EEG recording device are organized to capture potential differences in brain activity. Common configurations include:

1. **Monopolar (or Referential) Configuration**: In this setup, each electrode measures the potential difference relative to a common reference electrode. This reference electrode is often placed at a neutral site, allowing for clearer, consistent readings of each location's brain activity.
2. **Bipolar Configuration**: Here, signals are recorded as the difference between two neighboring electrodes. This arrangement is particularly useful for detecting localized brain activity and is commonly applied in epilepsy monitoring.

2.3 What are the different types of Epileptic Seizures?

Focal non specific seizure - occur due to an epileptogenic lesion on the contralateral frontal lobe. They usually originate from the supplementary motor area. The excitatory focus is generally around the Rolandic (motor) cortex. Temporal lobe seizures can also have motor symptoms. These symptoms include turning the head and neck to the opposite side and sometimes tonic contractions of the limbs and trunk. The motor manifestations of focal motor onset seizures can be characterized by tonic, clonic, atonic, myoclonic, hyperkinetic, epileptic spasms, and automatisms.

Complex Partial Seizure - Focal seizures that cause altered awareness are called focal unaware seizures or complex partial seizures.

Simple Partial Seizure - The electrical activity of the seizure can remain in one sensory or motor area of the brain, resulting in a focal aware seizure (also called simple partial seizure)

Tonic Seizure - A tonic seizure causes a sudden stiffness or tension in the muscles of the arms, legs or trunk. The stiffness lasts about 20 seconds and is most likely to happen during sleep. Tonic seizures that occur while the person is standing may cause them to fall. After the seizure, the person may feel tired or confused.

Clonic seizure - Clonic seizures are characterized by repeated jerking movements of the arms and legs on one or both sides of the body, sometimes with numbness or tingling. If it is a focal (partial) seizure, the person may be aware of what’s happening. During a generalized seizure, the person may be unconscious.

Tonic clonic seizure - a seizure that has a tonic phase followed by clonic muscle contractions and are usually associated with impaired awareness or complete loss of consciousness.

Atonic seizure - Atonic seizures are a type of seizure that causes sudden loss of muscle strength. These seizures are also called akinetic seizures, drop attacks or drop seizures. The sudden lack of muscle strength, or tone, can cause the person to fall to the ground. The person usually remains conscious, and may not always fall down.

Myoclonic seizure - Myoclonic (pronounced “my-oh-CLON-ick”) seizures are a type of seizure that cause a quick, uncontrollable muscle movement with no change in your level of awareness or consciousness. These usually affect either one muscle or a group of related muscles, but can sometimes affect wider areas of your body.

3. Project Research Methodology

I met with three neurologists at reputed Mumbai hospitals (Kokilaben Hospital, PD Hinduja Hospitals, NH-SRCC) to understand the challenges that the neurologist community face in diagnosing epilepsy. The neurologists cited key challenges of time consumed in reading EEG datasets and accurately identifying & classifying epilepsy types – accentuated in remote areas of the country by the lack of access to experienced neurologists (causing extended delays in diagnosis). They cited the need to have a technology solution that reads EEG datasets, identifies anomalies and suggests an epilepsy classification (once identified).

I defined a project along with The Innovation Story to enable Neurologists to improve the efficacy of Epilepsy detection and treatment, by providing an Epilepsy Detection App that achieved the following:

- Reduces the time taken to detect epilepsy
- Improves the accuracy of epilepsy type detection (complex partial seizure, simple partial seizure, myoclonic seizure, absence seizure, etc)
- Provides for remote diagnosis of Seizure Types by expert neurologists for EEG test data generated in remote locations

by utilizing Machine Learning methodologies on Electroencephalography (EEG) data.

3.1 Objective of My Project

I laid down the following objectives for developing the functionality of the Epileptic Seizure Detection App:

<u>Automate Seizure Classification</u>: Create a system that automatically classifies seizures from EEG signals, reducing the need for manual checks by doctors, which can lead to errors.

<u>Improve Accuracy</u>: Enhance the accuracy by using Machine Learning techniques to better analyze the EEG signals.

<u>Build a User-Friendly Web Application</u>: Develop an easy-to-use web application where users can upload EEG data and receive quick results regarding seizure classification.

<u>Use Deep Learning</u>: Utilize deep learning techniques, specifically Long Short-Term Memory (LSTM) networks, to effectively identify patterns in the EEG data for better classification results.

3.2 Innovation in Design for Epileptic Seizure Detection app

The innovation in my project comes from combining two methods for seizure detection using EEG data. Rather than relying only on extracted features like time-domain, frequency-domain, or wavelet features, I use the entire EEG signals as input alongside these features. This allows the model to learn both the broader patterns in the raw EEG signals and the specific characteristics captured by the features.

By incorporating an LSTM model, which is designed to handle sequential data, I can leverage its ability to detect time-based patterns related to seizures. Combining this with the additional features gives the model a more complete understanding, making it more accurate at detecting subtle signals that might otherwise go unnoticed.

Overall, this approach lets me analyze EEG data more comprehensively by using both raw signals and engineered features together, leading to improved seizure classification.

3.3 Data Sets for Epileptic Seizure Detection app

We used the Temple University Hospital's EEG Seizure Corpus(TUSZ) provided under the Neural Engineering Data Collaboration mechanism for annotated data for seizure events (start time, stop time, channel). This dataset comprised 29,000 EEG recordings from patients with various types of epilepsy and seizure episodes.

The dataset consists of two types of files

- EDF (European Data Format) files - These files store EEG recordings captured from multiple electrodes placed on the scalp, following standard clinical EEG protocols. The EEG data is recorded continuously and contains several channels representing different brain regions.

- CSV files: Each EDF file is accompanied by a CSV file that contains the annotations of seizure events. These annotations mark the start and end times of seizures and they also include labels that identify the type of seizure recorded.

3.4 Machine Learning Model for Epileptic Seizure Detection & Classification

3.4.1. EEG Signal Filtering -

EEG data is often noisy, containing various artifacts such as muscle activity, eye movement, and electrical interference which can distort the model's ability to detect patterns related to seizures. To address this, I used a bandpass filter with a frequency range of 0.5 to 45 Hz. I've used an FIR filter with a hamming window to preserve the essential brainwave components.

3.4.2. Analyses of EEG Signal Data for Wavelet Feature Extraction

Raw EEG Signals:

In addition to extracted features, the raw EEG signals were fed directly into the model to enable it to learn broader, time-dependent patterns within the EEG data. This approach allowed the model to gain insights from the complete signal sequence, capturing subtle variations that might otherwise be lost when relying solely on extracted features.

Time-Domain Features:

These features capture essential statistical properties of the signal as it evolves over time. The following features were calculated from each segment of EEG data for different channels:

- **Mean**: The average amplitude of the EEG signal within the segment.
- **Variance**: A measure of how much the signal values fluctuate around the mean.
- **Skewness**: Quantifies the asymmetry of the EEG signal distribution.
- **Kurtosis**: Describes the shape of the EEG signal distribution, particularly how sharply it peaks.
- **Standard Deviation**: Another measure of variability or dispersion within the EEG signal.

Frequency-Domain Features:

Since EEG signals are typically analyzed in terms of their frequency components, I used frequency-domain analysis to extract power in specific frequency bands. These frequency bands are known to be associated with different brain states, which are relevant for seizure detection:

- **Delta (0.5–4 Hz)**
- **Theta (4–8 Hz)**
- **Alpha (8–13 Hz)**
- **Beta (13–30 Hz)**
- **Gamma (30–45 Hz)**

I used power spectral density (PSD) to quantify the distribution of signal power across these frequency bands. These features provide insight into rhythmic activity that gives information about specific seizure types.

Wavelet Features:

These capture both time and frequency information, which is needed for detecting patterns associated with seizures. I applied discrete wavelet transform (DWT) to extract features that represent variations in the signal at different resolutions.

- **Wavelet Coefficients**: These coefficients capture localized changes in the EEG signal across different frequency scales.
- **Energy of Wavelet Coefficients**: Energy at different levels of decomposition reflects the activity across different frequency bands.

By utilizing wavelet features, we were able to capture seizure-specific transient events that are not as easily detected by time-domain or frequency-domain features alone, as shown in Exhibits 2a and 2b below.

3.4.3. Handling class Imbalance:

Certain seizure types were underrepresented compared to others, which could have negatively impacted the model’s performance by causing it to be biased toward the more frequent classes. To solve this, I used SMOTE (Synthetic Minority Over-sampling Technique), to generalize better and identify seizure types more effectively, even when they were underrepresented in the original dataset.

Exhibit 2a: EDF file with no seizure:

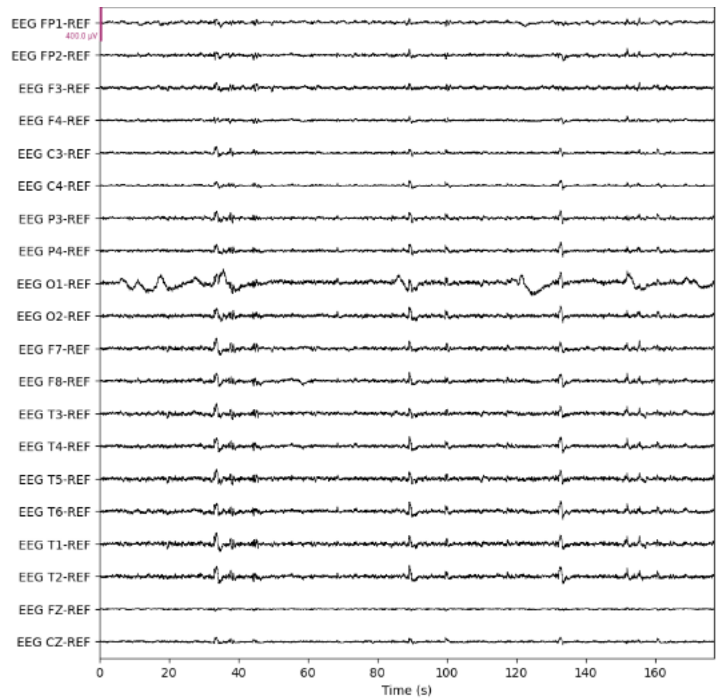
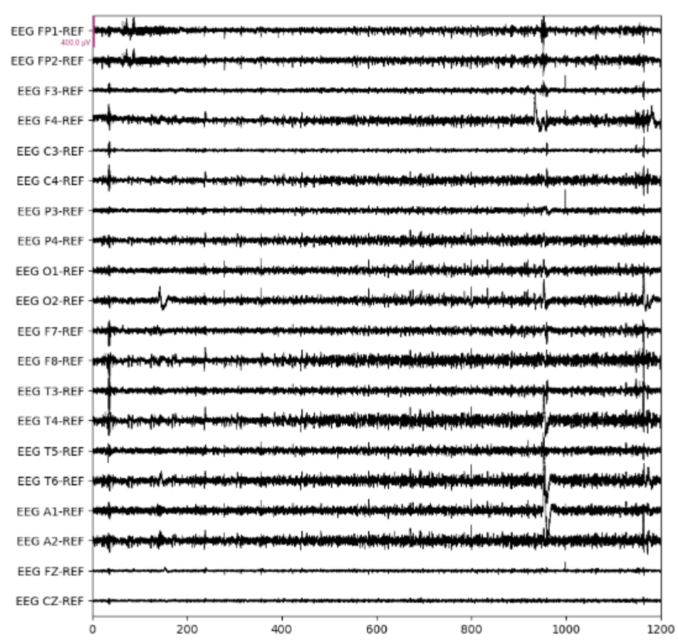


Exhibit 2b: EDF file with seizure:



3.4.4 Convolutional Neural Network Model for Classifying Wavelet Features

I used a Long Short-Term Memory (LSTM) network for classifying seizure types from EEG data, as LSTMs are particularly effective for time-series data like EEG signals. This is because LSTMs can capture long-term dependencies and patterns, which is essential for recognizing seizure events that unfold over time. By leveraging this sequential processing capability, the model can accurately classify different seizure types based on temporal patterns in the data.

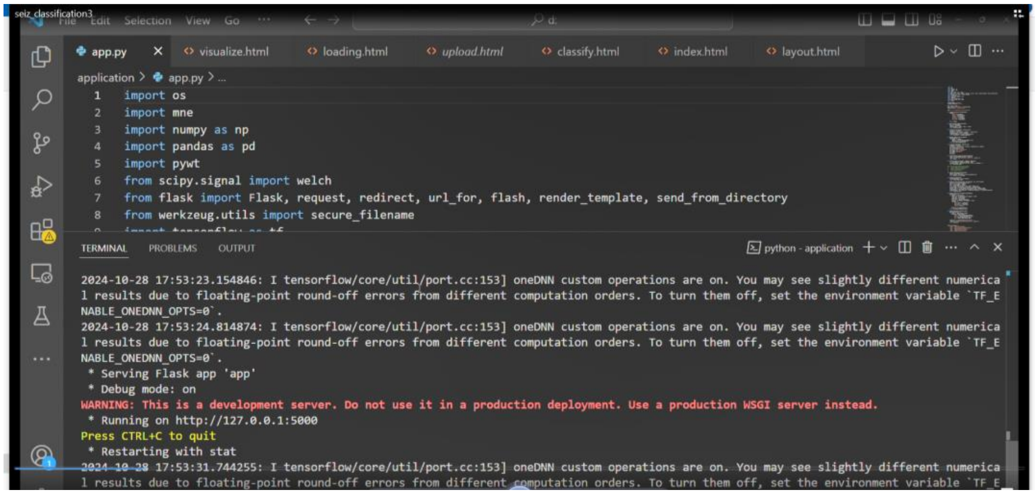


Exhibit 3: LSTM Modeling for Identifying and Classifying Epileptic Seizures

3.4.5 User App for Epileptic Seizure Detection and Classification:

I have developed a web-based interface using Flask as the backend framework. Flask was chosen for its simplicity and flexibility in handling server-side logic. The frontend was built using HTML, CSS, Bootstrap, and JavaScript.

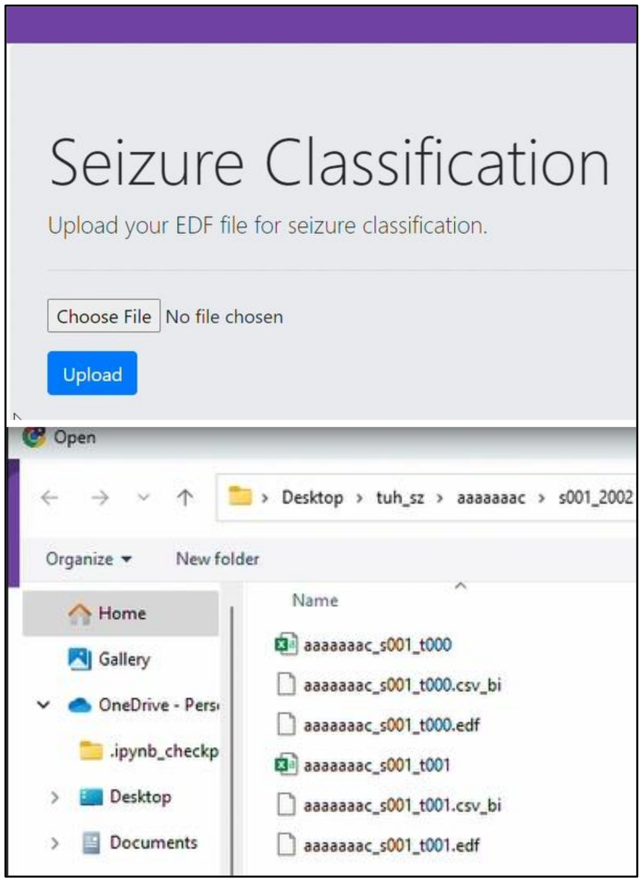
<u>Features</u>

1. **File Upload**: Users can upload their EEG data in EDF format through a web interface. The app ensures that only valid file formats are accepted.
2. **Seizure Classification**: Once a file is uploaded, the app processes the EEG data and classifies it into one of several types of seizures, such as generalized non-specific seizures, focal seizures, or absence seizures, using a machine learning model. If no seizure is detected, the app informs the user accordingly.
3. **Real-Time Feedback**: After processing, users receive immediate feedback on the classification results, including the predicted seizure type and a visualization showing the time of activity.

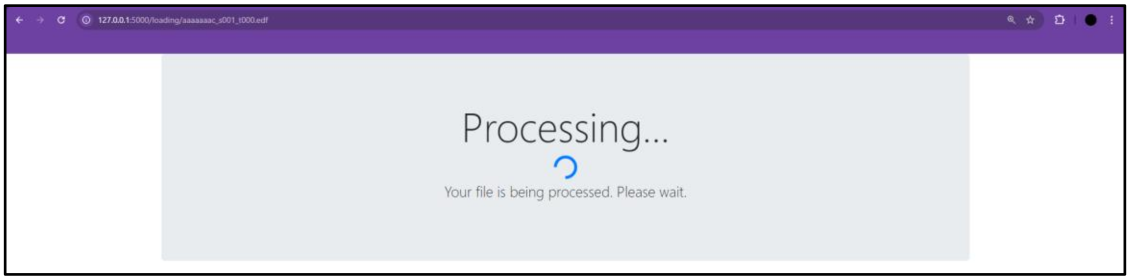
User Interface

The app offers a simple interface using Bootstrap for responsive design. There are different pages for file uploads, classification results, and a home page.

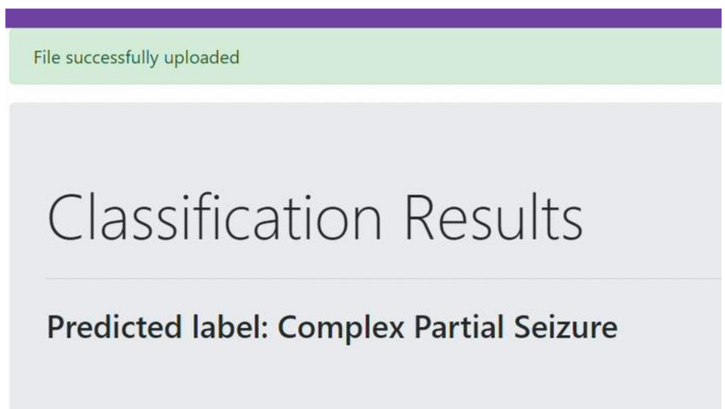
- **Homepage**: The homepage contains a simple form for uploading the EDF file.

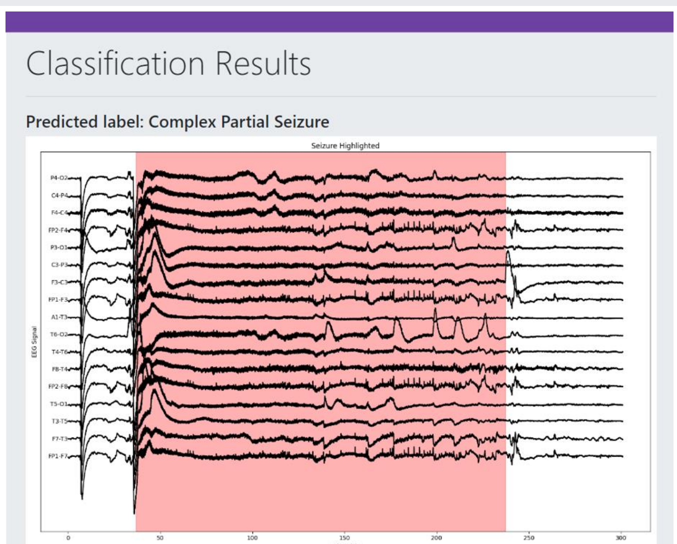


- **Processing Page**: Users receive confirmation of their file upload status, with any issues, such as unsupported file formats, flagged with clear error messages.



- **Classification & Visualization Results Page**: After the model analyzes the data, users are shown the predicted seizure type in a concise report format. This page also includes a link to return to the upload page for additional analysis.





4. Discussion on Research Outcomes

4.1 Findings for Current Epileptic Seizure Detection Model

The combination of time-domain, frequency-domain, and wavelet features provided a strong foundation for detecting seizures in EEG data. By using wavelet features, the model was able to capture sudden changes and more complex patterns that might not be easily identified by time or frequency analysis alone. The use of a bandpass filter (0.5 - 40 Hz) helped reduce noise from muscle movement and other interference, making the data cleaner for the model.

The LSTM model effectively handled the sequential nature of EEG signals, capturing patterns over time that are important for classifying seizures. The model architecture, which included dropout layers and batch normalization, helped prevent overfitting. Although using SMOTE addressed class imbalance by generating synthetic samples for underrepresented seizure types, there’s a risk that synthetic data might not fully capture the variability seen in real-world cases. Testing the model on more datasets could improve its ability to generalize.

The Epileptic Seizure Detection app provides a user-friendly interface, allowing users to upload EEG files easily and receive results quickly. This makes the tool accessible for practical use in medical settings, offering fast feedback on seizure classification. Future iterations could also incorporate automatic noise removal or more sophisticated feedback on the detected seizure type, making the tool even more valuable for medical professionals.

The App provides for a significant reduction in the lead time for analyzing EEG data from EEG data sets to arriving at a preliminary recommendation on seizure type from 12-15 minutes to 60-75 seconds.

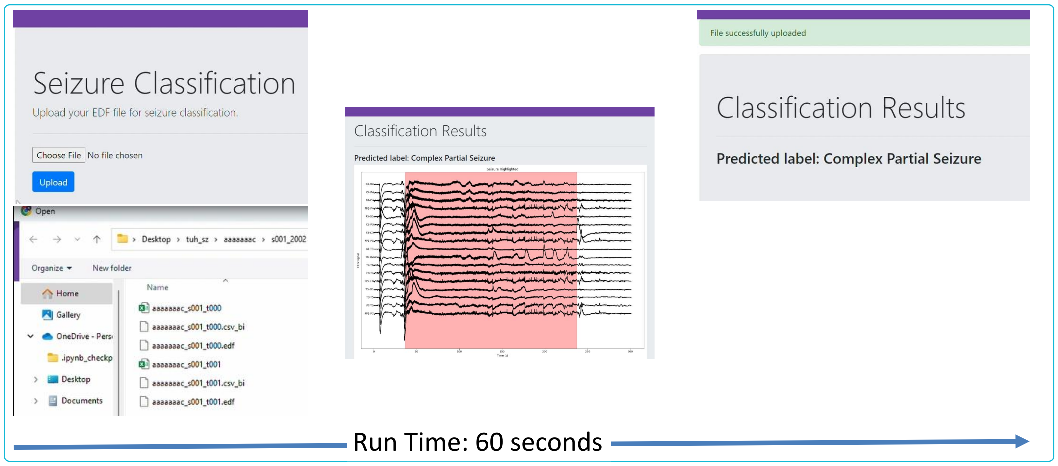


Exhibit 4: Machine Learning enabled optimization of Seizure Detection & Classification timeframe

The App provided for 92% accuracy on the 29,000 strong EEG data set with higher accuracies (96-98%) for the more statistically significant data sets.

Exhibit 5: Accuracy Matrix for the TUH NEDC EEG Data Set

4.2 Proposed Plan for Future Enhancement

Customized Electrode Configuration and Improved Accuracy: EEG systems can differ from one hospital to another, in terms of both the number of electrodes and their positions. To improve the tool's flexibility, I plan to explore whether the model requires signals from all electrodes or if it can work with fewer. I'll be consulting neurologists to understand which electrode signals are most critical for accurate detection. Based on their guidance, the tool might be adapted to work with different EEG setups. This will further improve the accuracy of classification outcomes, by enabling processing of multi-configuration data sets.

Remote EEG Diagnosis & Classification: One of the greatest utilities for the Detection App is allow improved collaboration between Primary & Secondary Healthcare institutions in relative remote areas (focused on Testing) with Expert Neurologists (in Metros and Tier1 towns). The App provides for a preliminary recommendation on identification and classification of seizure which can be shared with the Expert Neurologist for validation. This provides for a significant increase in the utilization of the Expert Capability, thanks to the 90-95% optimization of detection and classification lead time.

Symptom-Based Classification: To make the tool more helpful for doctors, I plan to add a feature where they can select symptoms related to seizures before analyzing the EEG data. By including this extra information, the tool can better understand the situation and might improve the accuracy of its predictions. This will be especially helpful when the EEG data alone isn't clear enough for a diagnosis.

Epilepsy Treatment Efficacy: I propose to super-impose tracking parameters over multiple EEG test instances and develop algorithm for visualization depicting progression (over time) of Wavelet parameters. This will enable Neurologists to assess the progression of epilepsy among patients and correlate with efficacy of epilepsy treatment pathways prescribed.

5. Conclusion

This project successfully developed an automated system for detecting and classifying epileptic seizures using EEG data. By combining raw EEG signals with extracted time-domain, frequency-domain, and wavelet features, the model gained a comprehensive understanding of seizure patterns.

An LSTM neural network, chosen for its ability to handle sequential data, proved effective in identifying various seizure types accurately. A Flask-based web application was also created to facilitate user-friendly file uploads and rapid classification feedback, making the tool accessible for practical, real-world use.

Future enhancements include adapting the model for different EEG configurations, improving accuracy, enabling remote diagnosis and expert feedback loops and incorporating symptom-based analysis, delivering quantum improvements in epilepsy diagnosis and care.

6. Acknowledgements

- Dr. Charu Sankhla, Faculty Neurologist & Movement Disorders Specialist, PD Hinduja Hospital, Mumbai
- Dr. Tushar Raut, Neurologist, Kokilaben Hospital, Mumbai
- Dr. Pradnya Gadgil, Neurologist, Narayan Hrudalaya SRCC, Mumbai
- Mr. Jaikumar Patel and Ms. Latifa Zaman, mentors for Python coding, The Innovation Story

7. References

1. Temple University Hospital (TUH) Neural Engineering Data Consortium data and document repository
   Temple University EEG Corpus
2. Ali Shoeb. Application of Machine Learning to Epileptic Seizure Onset Detection and Treatment. PhD Thesis, Massachusetts Institute of Technology, September 2009.
3. Guttag, J. (2010). Children's Hospital of Boston – Massachusetts Institute of Technology Scalp EEG Database (version 1.0.0). PhysioNet.
   CHB-MIT Scalp EEG Database v1.0.0
4. Enhanced Epileptic Seizure Detection through Wavelet-Based Analysis of EEG Signal Processing (Applied Sciences)
   by Sebastián Urbina Fredes, Ali Dehghan Firoozabadi, Pablo Adasme, David Zabala-Blanco, Pablo Palacios Játiva and Cesar Azurdia-Meza
   Applied Sciences 2024, 14(13), 5783; https://doi.org/10.3390/app14135783
5. Epileptic seizure detection combining power spectral density and high frequency oscillations (IJAMEC)
   by Rabia Tutuk and Reyhan Zengin
   https://dergipark.org.tr/en/download/article-file/2874211
6. Evaluating CNN Methods for Epileptic Seizure Type Classification Using EEG Data (IEEE Access)
   by Manuel J. Rivera, Javier Sanchis, Oscar Corcho,Miguel A. Teruel, Juan Trujillo,

https://www.researchgate.net/publication/380952355_Evaluating_CNN_Methods_for_Epileptic_Seizure_Type_Classification_Using_EEG_Data

7. EEG epileptic seizure detection and classification based on dual-tree complex wavelet transform and machine learning algorithms by Itaf Ben Slimen, Larbi Boubchir, Zouhair Mbarki, Hassene Seddik, The Journal of Biomedical Research, 2020 34(3): 151–161 https://pmc.ncbi.nlm.nih.gov/articles/PMC7324280/
8. EEG seizure detection: concepts, techniques, challenges, and future trends By Ein Shoka, A.A., Dessouky, M.M., El-Sayed, A. et al. Multimedia Tools Applications 82, 42021–42051 (2023). https://doi.org/10.1007/s11042-023-15052-2
9. Epileptic Seizure Detection Technique Using EEG Signals with Mobile Application Development By Lasefr, Zakareya, Khaled Elleithy, Ramasani Rakesh Reddy, Eman Abdelfattah, and Miad Faezipour. 2023. Applied Sciences 13, no. 17: 9571. https://doi.org/10.3390/app13179571
10. Epileptic seizure detection with deep EEG features by convolutional neural network and shallow classifiers By Wei Zeng, Liangmin Shan, Bo Su and Shaoyi Du, Frontiers in Neuroscience, 22 May 2023, Section (Neuroprosthetics), Volume 17 - 2023 | https://doi.org/10.3389/fnins.2023.1145526
11. Automated Epileptic Seizure Detection Based on Wearable ECG and PPG in a Hospital Environment By Vandecasteele, Kaat, Thomas De Cooman, Ying Gu, Evy Cleeren, Kasper Claes, Wim Van Paesschen, Sabine Van Huffel, and Borbála Hunyadi. 2017. Sensors 17, no. 10: 2338. https://doi.org/10.3390/s17102338
12. Wearable Epileptic Seizure Prediction System Based on Machine Learning Techniques Using ECG, PPG and EEG Signals By David Zambrana-Vinaroz, Jose Maria Vicente-Samper, Juliana Manrique-Cordoba and Jose Maria Sabater-Navarro, Sensors 2022, 22, 9372. https://doi.org/10.3390/s22239372