

# Machine Learning for Epileptic Seizure Detection and Prediction: Non-Invasive Approaches Using Wearable Technology

Rebecca Jacob

*Received June 25, 2025*

*Accepted September 21, 2025*

*Electronic access October 15, 2025*

Epilepsy is a neurological disorder that affects over 65 million people globally, often resulting in unpredictable seizures that severely impair daily functioning, increase the risk of injury, and elevate premature death rates. Current seizure monitoring relies on EEG readings in clinical settings, which are impractical for real-time prevention, and invasive intracranial implants, which are unsuitable for widespread use due to patient discomfort, surgical risks, and high costs. This research addresses these gaps by exploring whether machine learning algorithms trained on EEG data can be implemented in wearable, non-invasive devices to detect and predict seizures in real time. The study hypothesized that these algorithms could be extended to provide timely warnings to patients and caregivers. Using two datasets—CHB-MIT and Siena Scalp EEG—comprising 1160 hours of data from patients aged 1.5 to 71 (CHB-MIT: 1.5–22 years; Siena Scalp EEG: 20–71 years), four models were trained for seizure detection (K-Nearest Neighbors, Logistic Regression, Random Forest Classifier, Support Vector Machine) and one for prediction (Long Short-Term Memory). Evaluation metrics, including accuracy, recall, precision, and F1-score, showed a detection accuracy of 98.67%. Subsequently, the model was integrated with a wearable EEG headset (Muse) and a web app was developed that successfully predicts seizures 5 minutes ahead of time, with 84.52% accuracy. This study shows that clinical EEG data, paired with machine learning and wearable technology, enables accurate real-time seizure detection and early warnings. By providing accessible, non-invasive tools, this approach helps individuals take preventive action—reducing injury risk, easing anxiety, and enhancing autonomy.

**Keywords:** machine learning, artificial intelligence, seizures, epilepsy

## Introduction

Currently, over 65 million people have been diagnosed with epilepsy, a brain disease associated with chronic seizures, making it one of the most common neurological diseases globally. In the Global Burden of Disease Study 2010, severe epilepsy was assigned a disability-adjusted life year (DALY) value of 0.657, indicating a substantial reduction in quality of life, where 1 represents full disability or death<sup>1</sup>. Epilepsy also poses an economic burden, with significant healthcare costs and lost productivity associated with the condition.

The condition poses serious risks, including a significantly higher likelihood of falls leading to severe head injuries, a 13 to 19 times greater risk of drowning, and an increased chance of car accidents<sup>2</sup>. Additionally, individuals with epilepsy face a death rate up to three times higher than that of the general population due to premature death<sup>3</sup>. These unpredictable seizures pose a significant problem for individuals with the disease. Daily life can be filled with anxiety and challenges when planning activities, as it is difficult to anticipate when the next seizure will occur.

Epilepsy detection and prediction is usually performed on

EEG data, which are signal patterns from electrical activity in the brain. Although EEG readings can be obtained from doctors in a clinical setting, this isn't practical for real-time prevention of risks associated with seizures. Machine learning offers the possibility of processing EEG readings without requiring a clinical visit. Therefore, the research question aims to examine the extent to which machine learning algorithms can effectively predict seizures in real time using EEG data, allowing individuals to take preventative actions before harm occurs.

One approach that has been used to address the research problem is attempting to identify the pro-ictal stage of epilepsy, as demonstrated in a study conducted by researchers at the University of Alabama at Birmingham<sup>4</sup>. A “pro-ictal” stage in epilepsy refers to a period prior to seizure onset, where brain activity shows signs of impending seizure activity. The researchers successfully identified the onset of seizures 30–45 minutes prior. An important limitation of this study was its use of intracranial EEG readings, presumably due to their minimal noise interference. Intracranial readings require thalamus implants, which are invasive because they involve surgical procedures to place electrodes inside the brain. This makes it an impractical solution for daily applications due to high costs and potential patient

---

discomfort. In contrast, this research focuses on non-invasive methods that use external EEG devices, making the model more accessible and feasible for real-world applications.

Developing a machine learning model to predict seizures in advance would be highly practical to the field of seizure management, especially when integrated into a wearable device. The purpose of the research is to assess the feasibility of such systems and determine if they can be developed into reliable mechanisms to warn patients or caregivers of oncoming seizures. Elizabeth Stirling, a researcher in the Department of Biomedical Engineering at the University of Melbourne, and her colleagues state that wearable devices are the “perfect candidates to develop non-invasive, accessible forecasts”<sup>5</sup>. Such technology would enable individuals to anticipate seizures and take necessary precautions, such as “taking fast-acting medications or modifying activities”<sup>6</sup>, potentially saving lives. One example, as mentioned earlier, is the elevated risk of car accidents among individuals with epilepsy. Even just a brief warning could give a driver enough time to pull over safely, potentially saving lives. Thus, the critical implications of this technology make the research important to pursue. While both this study and Stirling et al. (2021) employed LSTM networks in combination with standard classifiers, the key distinction lies in the input data. Stirling et al. used physiological signals such as heart rate and circadian rhythms to predict seizure risk, emphasizing systemic patterns external to the brain. In contrast, this study uses direct EEG signals from the brain to capture epileptic activity.

Compared to Bhattacharya<sup>7</sup>, who achieved an average recall of 97.46% using a more complex CNN-LSTM architecture, this study employs a simpler yet effective model. Similarly, Esmaeilpour<sup>8</sup> reported a recall of 90.76% using random forest and support vector machines. Unlike many prior studies, this study evaluates both seizure detection and prediction across two distinct datasets (CHB-MIT and Siena Scalp) and demonstrates integration with a wearable EEG device, highlighting its potential for real-world, non-invasive use.

The objective of this research is to develop and integrate accurate machine learning algorithms into wearable technology, enabling the detection and prediction of seizures far enough in advance for individuals to take necessary precautions. I hypothesize that time-series machine learning models, like Long Short-Term Memory (LSTM) networks, can effectively identify abnormal neural patterns—such as spikes and sharp waves—that precede seizures. Then, these models could be implemented into wearable devices that emit warnings when a seizure is predicted. By recognizing complex, multi-dimensional EEG patterns, the models would provide patients with enough time to take preventative action and reduce the risk of injury.

The scope of this study and its applicability are defined by the range of demographics represented in the EEG data. The study excludes the development of models dependent on clinical or implant-based EEG systems, focusing instead on applications

that can be adapted to consumer-grade, wearable technology. Some limitations of this work include reliance on publicly available datasets, as clinical EEG data is often restricted. Additionally, as a student running these models on a personal laptop, I face limitations in computational power compared to high-end systems. Finally, there may be minor latency between EEG signal input and detection or prediction due to preprocessing steps.

This study employed a machine learning-based approach using publicly available EEG datasets to train seizure detection and prediction models. Data was preprocessed to reduce noise, segmented into time-based epochs, and analyzed using classification and time-series algorithms. The final models were evaluated based on several criteria and integrated into a wearable device system with real-time alert functionality.

### **Ethical Considerations**

This study involved no direct interaction with human subjects besides four healthy volunteers, and ethical concerns were minimal. The EEG datasets used—CHB-MIT and Siena Scalp EEG—were publicly available and already anonymized, with only limited demographic information (age and gender) retained. As such, informed consent, confidentiality, and ethical compliance were ensured by the original data providers.

This study acknowledges the risks associated with both false positives and false negatives. If the model incorrectly predicts an impending seizure (false positive), users may experience anxiety, social disruption, or unnecessarily restrict activities (e.g., avoiding driving or public outings). Conversely, a missed prediction (false negative) could result in users being unprepared for an actual seizure, possibly leading to injury, property damage, or even life-threatening events.

## **Methods**

### **Dataset Selection and Patient Demographics**

The overall structure of this study utilized a cross-sectional approach. To conduct this research, various publicly available EEG datasets were evaluated to identify the most suitable one for seizure prediction using machine learning. Six different datasets were evaluated—CHB-MIT, BERN, BONN/UCI, Siena Scalp, TUH EEG, and EPILEPSIAE—and reviewed based on several factors in each of them, including the type of EEG readings, overlap in channels used, demographics of patients, as well as size and accessibility of data for each dataset.

To initiate the research, the CHB-MIT dataset was selected as it was freely available and included EEG data from frontal and temporal lobe epilepsy, which would integrate well with the channels in the wearable headset. It contained a total of 1032 hours of EEG recordings from 23 patients. These patients

**Table 1** Dataset Selection

No.	Source	Dataset Name	Channel Count	Mode of Signal	Sampling Rate (Hz)	Seizure Count	Description / Comments
1	Physionet	CHB-MIT Dataset	23 (18 commonly used)	Scalp EEG (10–20 system)	256	198	Recordings from 23 subjects aged 1.5–22 years. Includes 23 channels, 198 seizures, and non-seizure samples for binary classification.
2	Universitat Pompeu Fabra	BERN Dataset	Varies (6 per sample)	Intracranial EEG (iEEG)	512	375	Data from epileptic patients producing multichannel EEG signals of 32-second durations.
3	University of Bonn / Data World	BONN / UCI Dataset	1 per sample	Intracranial EEG (10–20 derivations)	173.61	100 ictal files	Dataset contains 5 folders with 100 files each, representing recordings from individual subjects.
4	Physionet	Siena Scalp Epilepsy Dataset	20–34 (depending on sample)	Scalp EEG (10–20 system)	512	Not specified	Data from 5 patients aged 15–71 years, stored in EDF format similar to CHB-MIT dataset.
5	Temple University (TUH EEG)	TUH EEG Dataset	21–36	Scalp EEG (10–20 system)	250–512	~2700	Access requires signing a data usage agreement form.
6	European Database	EPILEPSIAE	19–128 (varies)	Scalp + Intracranial EEG	256 or 512	1000+	Widely used in epilepsy research; access requires a license.

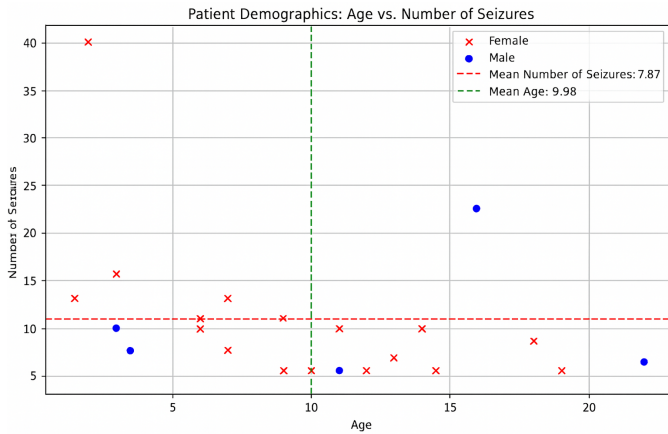
included both male and female subjects ranging from the age of one and a half years old to 22 years old. The data was collected using the 10-20 electrode scalp placement system, including 23 EEG channels.

The data was stored as edf files, or European Data Format files, which are commonly used to store medical time series, making it appropriate for the data. In total, 686 edf files were available for use. Out of these, 141 files contained one or more seizures, with a total of 198 seizures being recorded. This extensive amount of data from a variety of different patients made it ideal for the research being conducted.

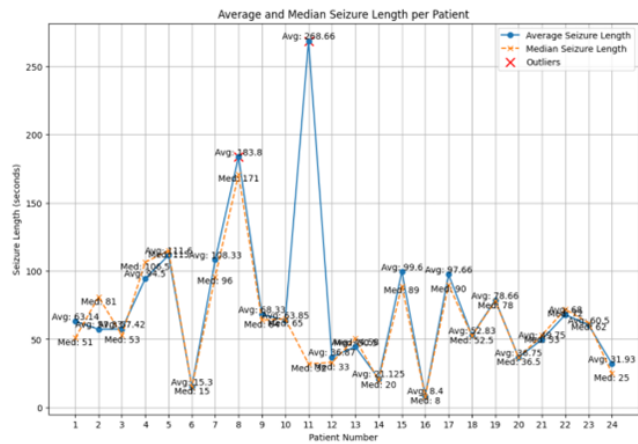
This was later extended to include Siena Scalp Dataset to get a wider patient demographic and since this had structure to CHB-MIT dataset and included the channels from the Muse headset.

### Dataset Preprocessing

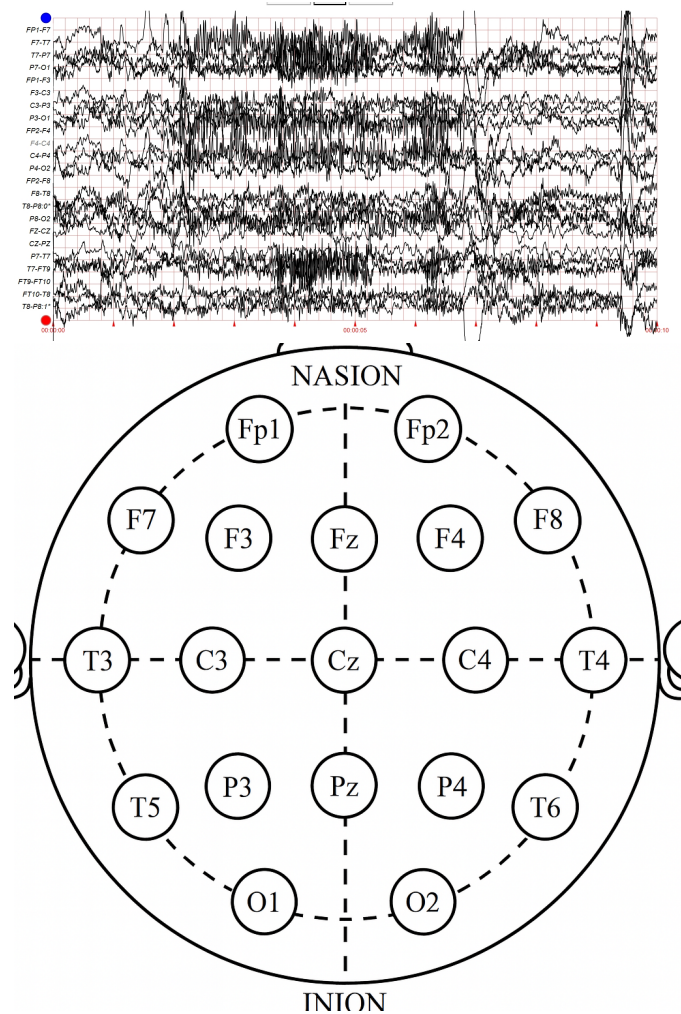
Google Colaboratory was used to code and build the models. While both TensorFlow and PyTorch are widely used in industry for Machine Learning, TensorFlow was used because it contains additional libraries such as Keras which further simplify the



**Fig. 1** Plot Patient Demographics with Number of Seizures from CHB-MIT Dataset



**Fig. 2** Plot Average and Median Seizure Length / Look for Outliers



**Fig. 3** EEG Signals CHB-MIT Dataset and 10-20 Electrode Scalp Placement<sup>9</sup>.

process of creating models and loading data. To extract the EEG data from the edf files, MNE was used — a popular library and open-source Python package for processing and analyzing EEG data. This library allowed for efficient loading, preprocessing, and creation of epochs (time-sliced segments) from continuous EEG signals. These epochs were 20 seconds each with a 4-second overlap and were saved as pickle files in Google Drive for efficient use, enabling easy loading for subsequent analysis. Computing power constraints was one of the factors for selection of data segment length and overlap. Shorter windows of 5-second segments with a 1-second overlap had a longer processing time, with a marginal increase in recall as it was more sensitive to seizure. However, the accuracy dropped due to more false positives.

A factor that can affect accurate detection and prediction models is the amount of noise contained in EEG Data. Thus, Noise Covariance was run on the data to assess the extent of the noise. To optimize data quality, comprehensive preprocess-

ing was implemented to reduce noise using three approaches. First, Independent Component Analysis (ICA) was used to isolate and remove noise, artifacts, or unwanted signals from EEG data, especially those due to eye movements, muscle activity, or other external interferences. Specifically, ICA worked to improve signal quality by isolating components representing true brain activity from those that were artifacts, followed by manual analysis and selective removal of noise components before reconstructing the EEG signals. The parameters used for ICA was `mne.preprocessing.ICA (n_components = 8, random_state = 97, max_iter = 800)`, where `n_components` matched the number of EEG channels used, `max_iter` ensured sufficient iterations and `random_state` ensured reproducibility. ICA is applied first to isolate and suppress temporal noise components without suppressing ictal (seizure) patterns, which is usually more structured.

Signal Space Projection (SSP) was used next to remove resid-

---

ual noise that was not removed by ICA by projecting signal components associated with specific artifacts (like eye blinks or muscle movements) into a separate subspace, allowing those components to be removed from the original data. The function `mne.compute_proj_epochs()` was used on the EEG data and returned projection vectors which were applied to the signals to eliminate noise. Applying SSP after ICA helped sharpen the result by targeting spatially structured noise patterns.

Lastly, Source Power Comodulation (SPoC) was applied to the labeled data—a technique that represents data in a way that highlights essential features while reducing noise, making it easier to process and analyze while retaining key information. By employing these three approaches, effective removal of noise from the EEG signals was successfully ensured. Band filtering was then used to isolate frequencies in the 0-64 Hz range that are relevant for seizure activity. Finally, the preprocessed data was converted from an MNE Epochs Array into a 2D Numpy array for compatibility with the models.

Binary classification of this data was used to analyze for the presence or absence of seizures. For each model, I split the data into 80% for training and 20% for testing. I applied cross-validation, splitting the training data into multiple subsets. The model was trained and validated on different subsets each time, helping reduce overfitting and providing a more robust performance evaluation.

After working extensively with the CHB-MIT dataset, the research was expanded by incorporating the Siena Scalp database. This dataset was also easily accessible and stored EEG signals in the same EDF format as CHB-MIT. A significant difference was that it contained data from 14 patients ranging from 20 to 71 years of age. This substantially increased the diversity of the research, as the data now spanned ages 1.5 to 71, making it applicable to patients of all ages. Other key features of this dataset included an additional 47 seizures from 128 hours of recording, as well as 29 different EEG channels, including 17 frontal and temporal channels. The same preprocessing and noise reduction techniques used on the CHB-MIT dataset was applied.

## Machine Learning Models

### A. Evaluation Metrics

To evaluate the performance of the models, standard metrics were used, including accuracy, precision, recall, and F1-score. Accuracy measures the proportion of correctly classified instances (both seizure and non-seizure). Precision measures the proportion of correctly predicted seizure instances out of all instances predicted as seizure. Recall measures the proportion of correctly predicted seizure instances out of all actual seizure instances. The F1-score is the harmonic mean of precision and recall. All in all, these metrics allowed for a balanced and

rigorous evaluation of the model's performance.

### B. Seizure Detection

**Classification Using K-Nearest Neighbors (kNN)**The first model trained was a K-Nearest Neighbors classifier. kNN is a simple, non-parametric algorithm that classifies data points based on the majority class of their k nearest neighbors in the feature space. A grid search was performed over k values ranging from 3 to 11 (odd values only), and accuracy was reviewed to define the optimal k value. A k value of 5 was used, meaning each data point was classified based on the 5 closest data points. The distance metric used was the Euclidean distance, a standard measure of distance between two points in multidimensional space.

**Classification Using Logistic Regression**Logistic Regression was also used, a linear model that is widely used for binary classification problems. Despite its name, it's used for classification, not regression. Logistic Regression models the probability that a given data point belongs to a certain class. The Logistic Regression model was fit to the training data using the liblinear solver, which is suitable for relatively small datasets.

**Classification Using Random Forest Classifier**The third model used was a Random Forest Classifier. Random Forest is an ensemble learning method that constructs a multitude of decision trees at training time and outputs the class that is the mode of the classes (classification) or the mean prediction (regression) of the individual trees. A grid search was performed over the number of trees (50, 100, 200) and maximum depth (5, 10, 15), and accuracy and was evaluated for the different parameters. The final Random Forest model was trained with 100 trees and a maximum depth of 10 for each tree. The maximum depth limits the number of nodes in each tree, which can help to prevent overfitting.

**Classification Using Support Vector Machine (SVM)**The final classification model implemented was a Support Vector Machine (SVM). SVM is a powerful and flexible algorithm that can be used for both classification and regression. SVM works by finding the hyperplane that best separates the different classes in the data. A linear kernel was used for SVM, which means that the decision boundary is a straight line. The regularization parameter C was tuned over a range of values (0.01, 0.1, 1, 10), and the model was evaluated using cross-validation on the training set. The best performance was achieved with C=1.

### C. Seizure Prediction Using Long Short-Term Memory (LSTM)

The phases of a seizure can be split into four different parts: the stage before the seizure, pre-ictal; the stage during the seizure, ictal; and the stage immediately after the seizure, or post-ictal;

---

the stage in between pre-ictal and post-ictal, interictal. My focus was on training a model to recognize the pre-ictal stage to be able to predict seizures.

For seizure prediction, Long Short-Term Memory (LSTM) network was used. LSTMs are a type of recurrent neural network (RNN) that are well-suited to time series data, such as EEG recordings. LSTMs are capable of learning long-term dependencies in the data, which is important for seizure prediction, as seizures are often preceded by changes in brain activity that occur over a period of several minutes.

The LSTM model consisted of several layers. The first layer was an LSTM layer with 64 units and a dropout rate of 0.3. Dropout is a technique that helps to prevent overfitting by randomly dropping out some of the units in the network during training. The second layer was another LSTM layer with 32 units and a dropout rate of 0.3. The third layer was a dense (fully connected) layer with a single output unit and a sigmoid activation function. The sigmoid function outputs a value between 0 and 1, which can be interpreted as the probability that a seizure will occur within the next 5 minutes. The model was compiled using the Adam optimizer with a learning rate of 0.0001. The model was trained with a batch size of 32 and an initial maximum of 100 epochs. Early stopping was used with a patience of 10 to prevent overfitting, and model checkpoints were saved based on best training accuracy. Class imbalance was handled using class weighting.

#### D. Web App and Phone Alerts

To make the seizure prediction model more accessible and practical, a web application was developed. This web app is designed to provide real-time seizure predictions and alert users and their caregivers in the event of an impending seizure.

The web app was built using Streamlit, which is a lightweight Python framework. The app uses JavaScript to enhance interactivity and communicates with a Python backend, which runs the trained LSTM model. The web app provides a user interface that displays the seizure risk level.

In addition to the web app, a system was developed for sending SMS alerts to users and their caregivers. This system uses the Twilio API to send text messages when a seizure is predicted. This ensures that users and their caregivers are notified of an impending seizure, even if they are not actively using the web app.

In the current implementation, the integration of the Streamlit interface and Twilio alert system was primarily designed to demonstrate usability and real-time alerting. Privacy protections, latency optimizations, and offline resilience mechanisms were not implemented. Future enhancements can include encryption at rest, encryption in transit, user authentication and login for Streamlit app, and use more secure messaging for alerts to meet patient privacy protection and HIPAA needs. Twilio depends

on external API and internet for communication that can introduce latency for alerts. Future studies can review on-device notifications like a sound or vibration to reduce dependency on external APIs. If the user loses internet connection (e.g., in a remote area), the Streamlit app cannot send or receive data, and Twilio alerts will fail to deliver. Some mitigation steps can be to store predictions locally and retry alert transmission when connectivity is restored, and also inform users if the app detects offline status or failed alert attempts.

#### E. Integration with a Wearable Muse Headset

A key aspect of the research is the integration of the seizure prediction model with a wearable EEG headset. The Muse headset was chosen for this purpose due to its accessibility, affordability, and ease of use. The Muse headset is a non-invasive EEG device that can be used to record brain activity in real-world settings.

A system was developed for streaming EEG data from the Muse headset to the web app in real-time. This system uses the Muse SDK (Software Development Kit) to connect to the headset and stream data. The streamed data is then preprocessed and fed into the trained LSTM model, which generates a seizure risk prediction.

To evaluate the integration with the Muse headset, two tests were conducted. With the “Simulated Seizure Input Test”, an unseen EDF file from the CHB-MIT dataset, containing seizure events that were not used for model training, was processed through a Muse-simulated channel configuration. This allowed assessment of the model’s ability to detect new seizure instances. With The “Healthy Subject Test”, the Muse headset was worn by four healthy volunteers, and real-time EEG signals were streamed into the trained seizure detection pipeline.

#### F. Physical Embodiment and Real-Time Alerts:

The wearable EEG headset, such as the Muse, can be structured as a comfortable headband or similar headgear. This device would continuously monitor the user’s EEG data. The trained LSTM model processes this data in real-time to predict the likelihood of a seizure. If a seizure is predicted, the system will send a text alert notification to the user and their caregiver. This would allow the user to take immediate precautions, such as sitting down or taking medication. This integration of wearable technology, real-time data processing, and alerts can significantly improve the safety and quality of life for individuals with epilepsy.

### Results

#### Seizure Detection:

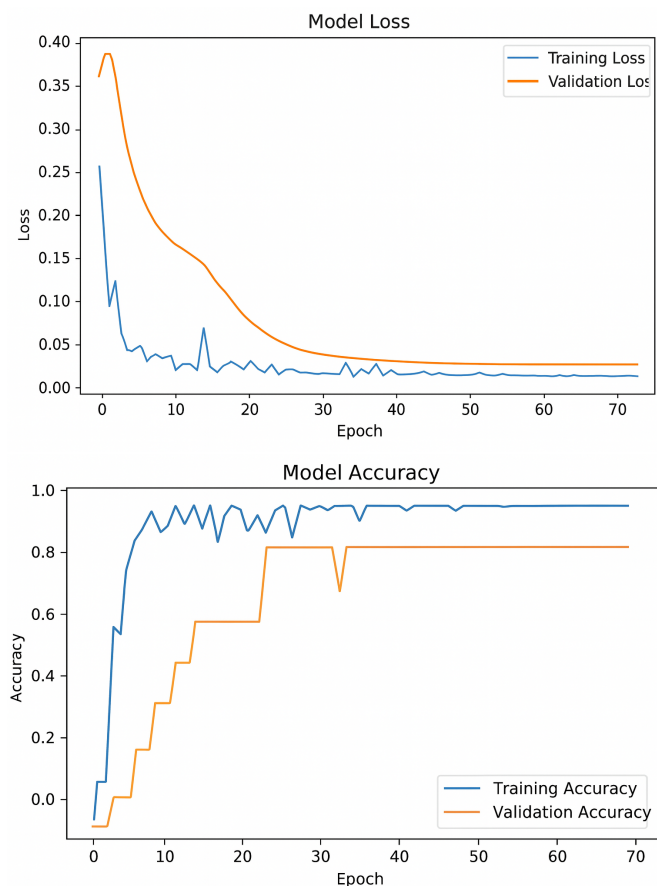
The results of the seizure detection models on the CHB-MIT dataset are shown in Table 2. The K-Nearest Neighbor (kNN)

achieved the highest accuracy of 98.67%, followed by Support Vector Machine (SVM) with 98.38%, Random Forest with 98.19%, and Logistic Regression with 97.22%.

### Seizure Prediction:

A variety of times or prediction window was evaluated and ultimately five minutes yielded the most optimal results. As expected, longer windows of ten and twenty minutes were less accurate because there is likely less pre-seizure activity in the brain that long before the onset of the seizure.

The LSTM model's performance in predicting seizures 5 minutes ahead of time was evaluated on both the CHB-MIT and Siena Scalp datasets. On the CHB-MIT dataset, the LSTM model achieved an accuracy of 84.52%, with a precision of 69.56%, recall of 72.72%, and F1-score of 71.11%. On the Siena Scalp dataset, the LSTM model achieved an accuracy of 81.03%, with a precision of 72.72%, recall of 76.19%, and F1-score of 74.42%.



**Fig. 4** CHB-MIT Long Short Term Memory (LSTM) Results for Seizure Prediction

### Integration with Muse Headset:

The simulated seizure test using the Muse headset and an unseen EDF file achieved an accuracy of 76.12% and a recall of 86.66%.

As expected, testing the headset with healthy volunteers indicated no seizure activity, demonstrating a zero false positive rate in this limited trial. These tests demonstrate the initial feasibility of real-time Muse-based integration. However, future work will need to address volunteers with seizures to better assess latency and false-alarm rates under more valid conditions.

### Discussion and Conclusion

The research demonstrates the potential of machine learning to accurately detect and predict epileptic seizures using non-invasive EEG data. The K-Nearest Neighbors (kNN) model achieved the highest accuracy in seizure detection on the CHB-MIT dataset, with an accuracy of 98.67%, while the Long Short-Term Memory (LSTM) network showed promising results in predicting seizures 5 minutes ahead of time on both the CHB-MIT and Siena Scalp datasets, with accuracies of 84.52% and 81.03%, respectively. This study acknowledges the risks of false positives. Even with an accuracy of 84.52%, the LSTM model still carries an error rate of approximately 15–19%, meaning that roughly one in five alerts may be incorrect. For users receiving multiple alerts per day, this level of false positives could lead to alarm fatigue, reduced trust in the system, or heightened anxiety.

This study accomplished its original objectives by achieving accurate seizure detection and prediction, successfully forecasting seizures five minutes in advance. Furthermore, the models were successfully implemented into a wearable and commercially available EEG device, upholding principles of functionality, affordability, and accessibility. Finally, the development of a web application that seamlessly integrates the model with the wearable device, featuring live EEG recording and the ability to input phone numbers for both the user and their caregiver, enabled a real-time alert system designed for both practicality and ease of use.

However, a major limitation to this research was inability to test this on a patient with actual seizures, due to safety and ethical constraints. Furthermore, the model would likely need more noise cancellation when tested with actual patients, due to the nature of it being a headset instead of cranial implants. Nonetheless, to address this limitation as effectively as possible, the model was tested and proved capable of processing on previously unseen EDF files and personal EEG signals, indicating its potential effectiveness in real-world scenarios.

Another limitation is that the model was trained and tested using Google Colab, which offers access to high-performance GPUs and optimized runtime environments. However, real-time deployment on consumer hardware such as smartphones or personal laptops might present computational challenges since the

**Table 2** Seizure Detection Results on CHB-MIT Dataset

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	Cross-Validation Accuracy (% ± SD)
K-Nearest Neighbors	98.67	90.91	96.44	93.59	88.6 ± 2.7 (95% CI)
Logistic Regression	97.22	88.90	87.37	88.10	87.3 ± 3.1 (95% CI)
Random Forest Classifier	98.19	98.17	91.24	94.58	90.2 ± 2.4 (95% CI)
Support Vector Machine	98.38	1.00	90.68	95.11	89.7 ± 2.9 (95% CI)

**Table 3** Prediction Window Results

Prediction Window	Prediction Accuracy	Recall
1 minute	70.59%	71.43%
5 minutes	84.52%	77.27%
10 minutes	34.01%	33.33%
20 minutes	18.00%	11.76%

**Table 4** Seizure Prediction (LSTM) Results

Dataset	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
CHB-MIT	84.52	69.56	72.72	71.11
Siena Scalp	81.03	72.72	76.19	74.42

LSTM model with multiple layers might introduce computational overhead. While Colab GPUs can easily handle the overhead, consumer CPUs (especially on smartphones) may struggle with low-latency and continuous processing of EEG signals. Additionally, continuous EEG data acquisition and processing may lead to high battery drain, especially on smartphones.

The high performance of the seizure detection and prediction models, achieving over 80% accuracy, demonstrates that machine learning offers a viable path for real-time, non-invasive seizure monitoring. This contributes to the growing body of research suggesting that EEG-based prediction can help patients take proactive, even life-saving, measures in high-risk situations. The integration of the trained model into a lightweight, wearable headset (Muse) and a web application provides the groundwork for a practical, accessible solution that can be implemented on a large scale to users worldwide. Patients can simply open the app, tap to record, and receive alerts about impending seizures, allowing them to continue daily activities with greater confidence and safety. This research advances the field by bridging the gap between machine learning advancement and real-world implementation, illustrating how early seizure warning systems can be both technically feasible and user-friendly.

While the results are promising, several opportunities exist for future research and practical application. First, evaluating the models on larger and more diverse datasets could improve generalizability across patient populations. Second, future work could explore additional machine learning approaches, including more advanced deep learning architectures. Third, expanding

the prediction window, anticipating seizures further in advance, could enhance patient preparedness and intervention. On the practical side, the web application and phone alert system can be further refined to improve usability and reliability. For example, notifications could be updated to deliver a subtle yet noticeable alert, such as a vibration or soft tone, when a seizure is predicted. In public settings, the device could trigger a louder alarm upon detection to notify bystanders that the user may need assistance.

Together, these advancements could help make seizure forecasting more accurate, accessible, and actionable in real-world environments. With further development, this system could become a life-saving companion for millions living with epilepsy — offering not just alerts, but autonomy, safety, and peace of mind.

## Acknowledgments

I would like to thank my parents and sister for providing feedback on my research paper and for supporting and encouraging me from start to finish. Finally, I would like to thank the contributors of the various datasets that I used in my project.

## References

- 1 M. Mehndiratta and S. Wadhai, *International Epilepsy Day - A day notified for global public education awareness*.
- 2 M. C. L. Wong-Kisiel, *Epilepsy - Symptoms and Causes*, <https://www.mayoclinic.org/diseases-conditions/epilepsy/symptoms-causes/syc-20350093>.

- 
- 3 W. H. Organization, *Epilepsy*, <https://www.who.int/news-room/fact-sheets/detail/epilepsy>.
  - 4 A. Ilyas, O. Alamoudi, K. Riley and S. Pati, *Pro-Ictal State in Human Temporal Lobe Epilepsy*.
  - 5 R. Stirling, D. Grayden, W. D'Souza, M. Cook, E. Nurse, D. Freestone, D. Payne, B. Brinkmann, T. Attia, P. Viana, M. Richardson and P. Karoly, *Forecasting Seizure Likelihood With Wearable Technology*.
  - 6 B. Brinkmann, P. Karoly, E. Nurse, S. Dumanis, M. Nasser, P. Viana, A. Schulze-Bonhage, D. Freestone, G. Worrell, M. Richardson and M. Cook, *Seizure Diaries and Forecasting With Wearables: Epilepsy Monitoring Outside the Clinic*.
  - 7 A. Bhattacharya, *Automatic Seizure Prediction using CNN and LSTM*, arXiv.org.
  - 8 A. Esmailpour, S. Tabarestani and A. Niazi, *Deep learning-based seizure prediction using EEG signals: A comparative analysis of classification methods on the CHB-MIT dataset*.
  - 9 Physionet.org, *LightWAVE 0.71*, <https://physionet.org/lightwave/?db=chbmit/1.0.0>.