Epileptic Seizure Detection

Domain Background

Epilepsy is among the most common neurological disorders in the world, second only to stroke, and affects over 50 million people around the globe (with 3 million just in the US). Unpredictable seizures are one of the most debilitating aspects of epilepsy and when uncontrolled, can cause severe limitations to an individual's ability to lead a normal life. During the event of a seizure, a subject may become confused or lose consciousness, resulting in complete loss of memory of the event. Stronger seizures can even cause spasms and uncontrollable muscle twitches which combined with nonconsciousness may result in severe injuries to the sufferer or even death in the worst case. In such scenarios, a seizure detection device which enables a caretaker to quickly aid the subject or triggers real time therapy that limits the complications of the seizure, could have huge life-saving potential. While many current methods of seizure detection have high levels of sensitivity (true positive rates >90%), they also suffer primarily from low levels of specificity resulting in a high number of false alarms. As a result, a significant amount of research in recent years has been focused on exploring new methodologies to detect seizures using EEG (brain), ECG (heart), and EMG (muscle) signals (1).

Problem Statement

The goal of this project is to compare and analyze the performance of various machine learning methods in their ability to detect seizures from EEG signals which can be obtained from commercially available wearable devices. Both traditional classification (e.g. naïve bayes, logistic regression, and random forest) and deep learning (neural networks) methods will be assessed to determine the strengths and weaknesses of each at recognizing seizure events from time series EEG data of 500 epileptic patients collected from the UCI Machine Learning Repository. Both models would be expected to perform a binary classification of seizure or no seizure and their test accuracy scores would be evaluated against the labels provided in the dataset. This research would hopefully aid in the development of wearable devices that more accurately signal the event of seizures in epileptic patients which could greatly aid in their mobility and reduce their risk of injury or death.

Datasets and Inputs

For this experiment, we use data obtained from Kaggle under the title Epileptic Seizure Recognition (2), which is a preprocessed/re-structured version of a commonly used dataset featuring epileptic seizure detection. The original dataset consists of a total of 500 EEG recordings, each from a different subject, containing time series data lasting a duration of 23.6 seconds and sampled into 4097 data points. In the Kaggle repository, each of these EEG recordings are further subdivided into 23 chunks containing 178 data points representing 1 second intervals of time. As a result a total of 500 * 23= 11,500 rows of data, each with 178 time series data points of EEG values and 1 response variable (marked as classes 1 – 5) was obtained and then stored in a csv file for easy accessibility to researchers. Note: Only subjects in class 1 are experiencing a seizure. Classes 2, 3, 4, and 5 represent different states of subjects who did not have an epileptic seizure. Each class contains an equal 2300 samples. For the purposes of this study, subjects in classes 2-5 will be marked as

0 to represent a state of no seizures while labels of 1 will be left alone to represent a state of seizures. As a result, there will a 4:1 class imbalance between non seizure and seizure patients.

The original dataset is collected from the UCI Machine Learning Repository

Solution Statement

Two distinct classes of machine learning algorithms will be used for training on the EEG dataset for seizure detection: traditional classification models and deep learning models. To create a proper training and testing set, the 11,500 samples will be split at a ratio of 80 : 20 (training : testing), at random using the train_test_split function embedded in sci-kit learn. These data sets will then be used consistently throughout the rest of the protocol to train and validate the aforementioned models under two distinct paradigms, one suited for each class (more details in project design). Parameters suited for each model will be adjusted to allow each model to perform optimally and a comparison of all models will be conducted at the conclusion using a set number of evaluation metrics (defined below).

Benchmark Model

For the benchmark comparison, we will use data collected from a scientific paper written by Ramgopal et al in 2014 (1) which summarizes a number of different machine learning methods (e.g. neural networks, linear regression, SVM, etc.) used to train EEG data for seizure detection and prediction as well as their level of sensitivity (TP/P rate) and specificity (TN/N rate). Both these metrics can be calculated for the models that we train in our experiment. Although there might be discrepancies in the preprocessing of EEG data and the source of the datasets, the benchmarks outlined in the paper will at least give us a coarse baseline to compare our results. (See table 1 in Appendix below for model performance)

Evaluation Metrics

Accuracy score can be used as one of the evaluation methods for optimization as it takes into account the ability of the model to identify both labels of seizure and no seizure correctly. In this score, both true positives and true negatives get equal weighting, and the higher the ratio, the fewer false positives and false negatives there are to trigger false alarms or fail to alarm at all.

$$accuracy = \frac{true\ positives + \ true\ negatives}{Total\ dataset}$$

F- beta score can be used to optimize the precision ($\frac{\text{true positives}}{\text{all positives}}$) and recall ($\frac{\text{true positives}}{\text{true positive + false negative}}$) of the models. (Precision refers to the percentage of results that are relevant and recall gives the percentage of relevant results that are correctly classified). For this application, a higher recall may be necessary to reduce the number of false negatives which may result in injury to epileptic patient. Since the data set is imbalanced towards non-seizure samples, F-score may be crucial for ensuring the model accurately classifies seizure patients correctly (A high accuracy model may simply classify patients as non-seizure correctly).

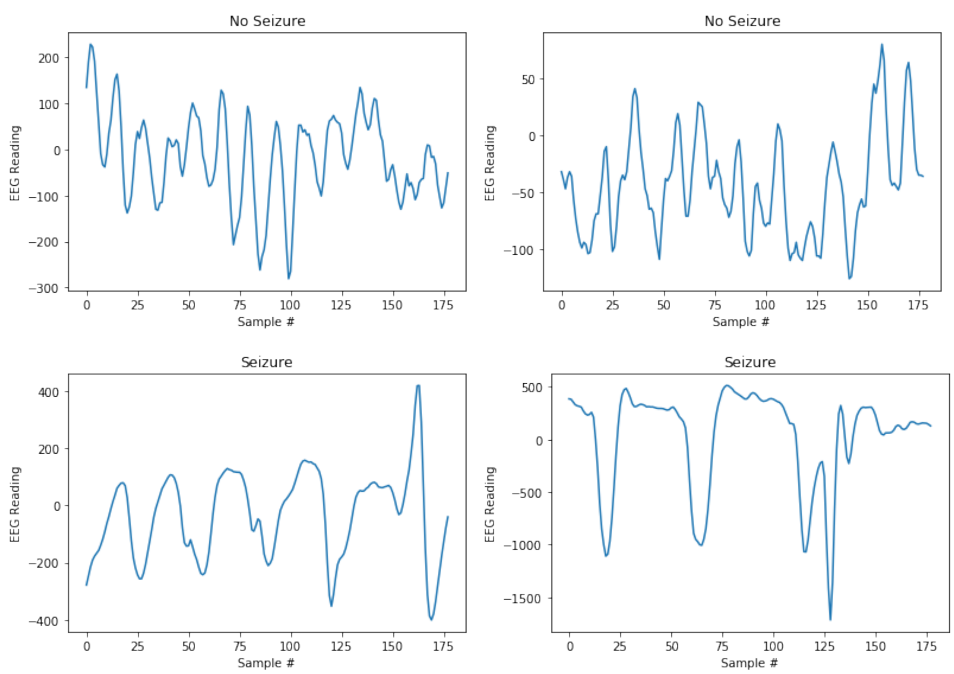
$$F_\beta = (1 + \beta^2) \cdot \frac{\text{precision} \cdot \text{recall}}{(\beta^2 \cdot \text{precision}) + \text{recall}}.$$

Project Design

The classification models will be trained on a series of features that can be extracted from the time series data since individual EEG sample values have little relationship with the final output. These include max and min values, average signal frequency, standard deviation, and median among other basic descriptive statistics. Four key classification models will be tested for their effectiveness – random forest, logistic regression, support vector machines, and Naïve Bayes. The most effective model out of the four will have its training parameters optimized using grid search and cross validation with the ultimate goal of comparing its performance to that of neural networks.

Contrarily, since convolutional neural networks have been shown to extract relevant features from image data on their own accord, it will be passed time series EEG data that has not been preprocessed as inputs with the expectation that certain layers in the model would autonomously filter out relevant information. Neural networks will assign a probability for each class, which will allow the setting of a threshold to optimize true positive rates while minimizing false positive rates. This classification threshold along with the overall neural network architecture (layers types, activation functions, and number of layers) and other training parameters (epochs, solver type, learning rate, etc.) will be altered until an optimal solution is reached. A TowardsDataScience tutorial (3), that creates a convolutional neural network in Keras for sleep stage classification will be used to establish the initial base algorithm for the deep learning seizure detection model.

Below are plots of 4 samples of the dataset, with the upper two being patients experiencing no seizure and the lower two of subjects experiencing seizure. From an early preliminary analysis, it seems that the frequency of the EEG signal is much higher in the subjects who don't experience any seizure compared to the ones that are in a state of seizure. As a result, a simple Fourier transform analysis measuring the predominant frequency component in the EEG signal may be the best feature for seizure detection (input for traditional classification models).



Appendix

Table 1
Selected seizure detection systems.

Author, year	Measuring device/seizures detected	Detection algorithm	Results
Electroencephalography/electrocorticography			
Webber, 1996 [5]	EEG (24–40 channels)/seizures not stated	ANN classification system	SEN of 76% and FPR of 1 event/h
Pradhan, 1996 [6]	EEG (8 channels)/seizures not stated	Wavelet transformation feature acquisition, ANN classification	SEN of 97% and SPEC of 89.5%
Gabor, 1998 [7]	EEG (8 channels)/seizures not stated	Self-organizing neural network with unsupervised training	SEN of 92.8% and FPR of 1.35 events/h
Wilson, 2004 [8]	EEG (8–32 channels)/seizures not stated	Combined algorithm (utilizes matching pursuit, small neural networks, and clustering algorithm)	SEN of 76% and FPR of 0.11 events/h
Wilson, 2005 [9]	EEG (single channel selected)/CPS, secondary GS and primary GS	Used a trained probabilistic neural network for rapid detection of seizures	SEN of 89% and FPR of 0.56 events/h
Alkan, 2005 [10]	EEG (4 channels)/absence seizures	Comparison of linear regression systems and ANN classification systems	ANN-based systems found to be greater. ANN-based system provided greater accuracy compared with linear regression
D'Alessandro, 2005 [11]	Intracranial EEG/seizures not stated	Genetic algorithm for signal processing, probabilistic neural network for classification	100% prediction of seizures within 10 min prior to onset
Arabi, 2006 [12]	EEG/neonatal seizures	Used linear correlation feature selection methods and back propagation neural network for classification. Used in detection of neonatal seizures	SEN of 91% and FPR of 1.17 events/h
Casson, 2007 [13]	Ambulatory EEG	Continuous wavelet transform	Over 90% of spike detection
Chan, 2008 [14]	Intracranial EEG/PS	SVM system	SEN of 80–98%, FPR of 38%
Netoff, 2009 [15]	EEG (6 channels)/PS	Cost-sensitive SVM system	SEN of 77.8%, no false positives detected
Chua, 2009 [16]	EEG/PS	Data processing by higher-order spectra analysis followed by classification by the Gaussian mixture model or SVM	Accuracy of 92–93%
Mirowski, 2009 [17]	EEG/PS	Variable feature extraction methods used followed by patient-specific machine learning-based classifiers	Convolutional networks combined with wavelet coherence yielded sensitivity of 71% and no false positives
Sorensen, 2010 [18]	EEG (3 channels)/GTCS, SPS, CPS	Features classified by matching pursuit algorithm and classified by SVM	SEN of 78–100 and FPR of 0.16–5.31 events/h
Chisci, 2010 [19]	EEG (multichannel)/focal seizures	Least-squares parameter estimator for extraction followed by SVM classification	SEN of 100%
Peterson, 2011 [20]	EEG (single channel)/absence seizures	Wavelet transform followed by SVM classification used to detect absence seizures using single-channel EEG	SEN of 99.1% and PPV of 94.8%
Temko, 2011 [21]	EEG (8 bipolar)/neonatal seizures	Fast Fourier transform used for feature extraction followed by SVM classification. Used to detect neonatal seizures	SEN adjustable, with 89% SEN yielding one false detection/h
Acharya, 2011 [22]	EEG/seizures not stated	Higher-order spectra-based feature extraction followed by SVM	Detection accuracy of 98.5%
Kharbouch, 2011 [23]	Intracranial EEG/focal epilepsy	Multistep feature extraction system followed by SVM classifier, individualized for patients	Detected 97% of seizures, FPR of 0.6 events/day
Liu, 2012 [24]	Intracranial EEG/GTCS, SPS, CPS	Wavelet decomposition-based feature extraction followed by SVM classification	SEN of 94.5% and SPEC of 95.3%
Xie, 2012 [25]	EEG (6 channels)/focal seizures, others not stated	Feature extraction by wavelet-based sparse functional linear model and 1-NN classification method	Has 99–100% classification accuracy
Direito, 2012 [26]	EEG (multichannel)/focal seizures	Markov modeling classification system. Identified four states — preictal, ictal, postictal, and interictal	Point-by-point accuracy of 89.3%
Rabbi, 2012 [27]	Intracranial EEG/GTCS, SPS, CPS	Used fuzzy algorithms for feature extraction for classification	SEN of 95.8% and FPR of 0.26 events/h
Implanted advisory system			
Cook, 2013 [28]	Intracranial implanted device/partial-onset seizure	Cluster computing system at NeuroVista (one algorithm for each patient)	SEN of 65%–100%
Electromyography			
Conradsen, 2010 [29]	Features extracted from surface electromyography acceleration and angular velocity/seizure-like movements performed by healthy volunteers	Classification based on SVM	SEN of 91–100% and SPEC of 100%
Conradsen, 2012 [30]	Electromyography and motion sensor features/motor seizures, seizure-like movements performed by healthy volunteers	Discrete wavelet transformation/wavelet packet transform techniques used to extract features. SVM classification system	Evaluated healthy subjects simulating seizures. SEN of 91–100% and SPEC of 100%

Table of EEG seizure detection models and associated sensitivity/specificity.
Source: *Seizure detection, seizure prediction, and closed-loop warning systems in epilepsy.* **Sriram Ramgopal, Sigride Thome-Souza, et al.** 2014, Epilepsy & Behavior, pp. 292.

Bibliography

1. *Seizure detection, seizure prediction, and closed-loop warning systems in epilepsy.* **Sriram Ramgopal, Sigride Thome-Souza, et al.** 2014, Epilepsy & Behavior, pp. 291-307 .
2. **Harun-Ur-Rashid.** https://www.kaggle.com/harunshimanto/epileptic-seizure-recognition. [Online] October 11, 2018. [Cited: August 16, 2019.]
3. **Mansar, Youness.** https://towardsdatascience.com/sleep-stage-classification-from-single-channel-eeg-using-convolutional-neural-networks-5c710d92d38e. [Online] October 1, 2018. [Cited: August 15, 2019.]