On the Detection and Prediction of Seizures using EEG

Fatima Ansari and Aren Wang
Mentor: Joanna Sands

Abstract:

Seizures are abrupt, rapid bursts of electrical activity within the brain. Those with epilepsy, a central nervous system disorder, suffer repeated seizures that appear to occur randomly and without warning. Frequent seizures may cause physical injury or even death [2]. A device that can quickly detect and respond to the onset of a seizure may lessen these risks.

The most commonly used instrument to detect such an event is an electroencephalogram (EEG), which is noninvasive and contains graphs of multiple channels. These graphs reveal the brain's electrical activity. EEG can be used to distinguish different seizure types and epilepsy types (focal or generalized, idiopathic or symptomatic, or a symptom of a larger epilepsy syndrome), and thus the choice of antiepileptic treatment and prognosis prediction [6].

Background:

The montage of an EEG refers to how the electrodes, also known as channels, are connected. In a monopolar montage, one electrode is active and the other two, which are connected, serve as a reference. In a bipolar montage, all electrodes are linked by wires. Most EEGs use a bipolar montage, also known as the standard 10-20 System [4]. In each chain, each electrode's voltage is subtracted from the one behind it, and these results are depicted on the EEG graph.

One problem with this type of EEG (and EEGs in general) is that it requires many channels, which renders it uncomfortable, not very portable (the user is often immobile while wearing it) and of questionable accuracy (the channels near FP2 and A2 are easily jostled.) It also cannot reveal where a seizure originated. There is thus a need for research on creating a system with fewer channels. The objectives for this project were to visualize the data within an EEG dataset from a live person, detect and characterize the variation of seizures while utilizing the reference and channels, and ultimately develop an algorithm that predicts the onset of seizures with as few channels as possible.

Methods:

In an attempt to determine whether an EEG system with fewer channels could be effective, we created a Python program that could analyze data from real EEG channels. The intent of this program was to pass data where the location of the seizures was known and to verify whether

these seizures could be detected in the system's output. Below are the steps taken to create the program.

First, we created a function to generate a hypothetical, random waveform in Python. This function could be set to produce waveforms of different frequencies, amplitudes, sampling rates, and noise levels. This function did not actually generate real seizure data; rather, it served to test whether our program could correctly identify key attributes of input data such as frequency and bandpower. Once we determined that it could via running multiple test waveforms through it, we progressed to the next step.

Second, we created a Python program that uses helper functions to annotate an EEG dataset. The functions include compressing EDF files that contain the seizures, plotting the seizures, reading the seizure annotations, and indexing into channels of interest. There are several public EEG datasets that include seizure data from epileptics, who range from neonates to adults. We analyzed several public datasets, exploring the various features and characterizing them. One feature of interest, for example, was short term continuous data (as in recordings under twenty four hours) so that the non seizure segments were not overly long in comparison to the seizure segments.

The aforementioned system also allows for analysis of any public EEG dataset and the distinguishing qualities. Its outputs include amplitude along the waveform vs. time, bandpowers of the five main frequency bands (delta, theta, alpha, beta, gamma) vs. time, and z-scores of those bandpowers vs. time. This was helpful since it is said that one of the hallmarks of a seizure is a marked increase in bandpowers, especially delta and gamma [3]. We also hypothesized that another hallmark would be larger peaks and troughs in the amplitude vs. time waveform, though this was less certain. We derived this hypothesis from the fact that a seizure involves an uncontrolled burst of electrical activity.

One of the challenges we faced was in finding the right dataset. We initially attempted to use the Sienna Scalp dataset; however, the seizure data did not match the annotations and when we ran it through our system, the time domain did not materialize properly. The second dataset, from an MIT study, was better, and it consisted of several sub-datasets that represented time samples for each participant. However, within this dataset we encountered the additional challenge of selecting the optimal time sample and the right patient to pass onto our system.

Results:

We first chose a one-hour sample that contained six seizures, thinking that having this would give us an idea of how a seizure presents itself when graphed. It could also have confirmed what we might have predicted about a seizure's signature based on articles. For the sample, we ran the

data from every possible channel through the program. This allowed us to find those that showed the most contrast between seizure and nonseizure data and would thus be most effective in constructing a short-channel EEG system.

Several unexpected results appeared. The first was a series of very high peaks near 3400 s (both in amplitude vs. time and bandpower vs. time), despite that the data from the study indicated no seizure occurred around this time. The second was a very noticeable *lack* of activity near 3570 s, even when the data indicated there *was* a seizure here. These unusual results occurred for the majority of the channels we tested, and we hypothesized that this dataset was an anomaly, where the frequent seizures affected the nonseizure data and prevented us from making a conclusion.

We then chose four different samples from other patients that each contained fewer seizures. As expected, most channels for these samples showed a bandpower peak where the seizures were reported (The amplitude peaks were sometimes discernible, sometimes not.) However, similarly to the first sample, some channels showed anomalous peaks at various periods, even though the data collected directly from the study indicated no seizure. It is possible these peaks manifested from interference, such as someone walking by.

In an attempt to solve this problem, we removed low frequency bands (delta and theta) from our analysis. These bands appeared the most likely to be affected by interference (such as the electrodes being disturbed or someone walking by) and anomalous non-seizure peaks in brain activity. Then we graphed the remaining bands on 3 separate graphs. We ensured these graphs were normalized using z-scores, which our system did not initially include, so we could compare them accurately.

After we made this change, it unfortunately remained that numerous peaks occurred that were, according to the data, not seizures. This revealed that our system was still inaccurate, or else the seizure time samples provided by the study were incorrect. However, we did make two important observations. The FZ-CZ channel and CZ-PZ channels were the most effective at distinguishing between seizures and non-seizures, and contained minimal interference and anomalous peaks. In contrast, the channels containing frontal electrodes, such as FP1-F7, were least effective and the most susceptible to noise or non-seizure brain signals. This could be because such channels are located at the front of the brain and may be particularly easily jostled or disturbed. Our findings are consistent with previous research on this topic [5], and they may prove useful to us or other researchers in the future in finding the optimal channels for a short-channel system.

Below are images of the contrast between a channel that contains the CZ electrode and a channel that does not contain it. The top images show the alpha, beta, and gamma band powers from left to right along the FZ-CZ channel for a particular patient. Whereas the bottom ones show the

same band powers along the FP1-F7 channel. The blue bars represent where a seizure occurred as indicated in the sample dataset.

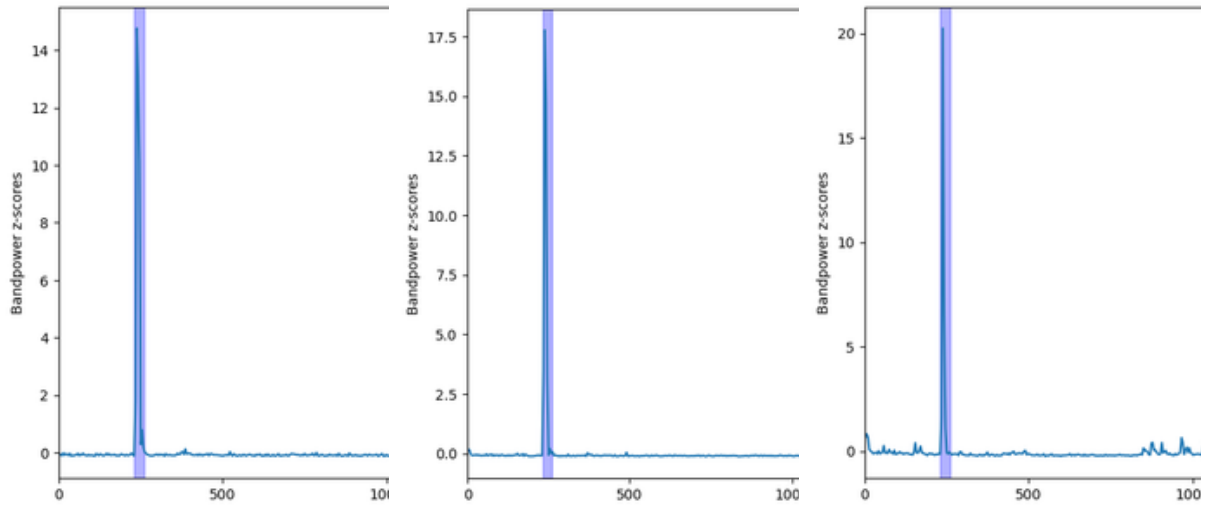

Figure 1. Channel FZ-CZ alpha, beta, and gamma band power z-score graphs

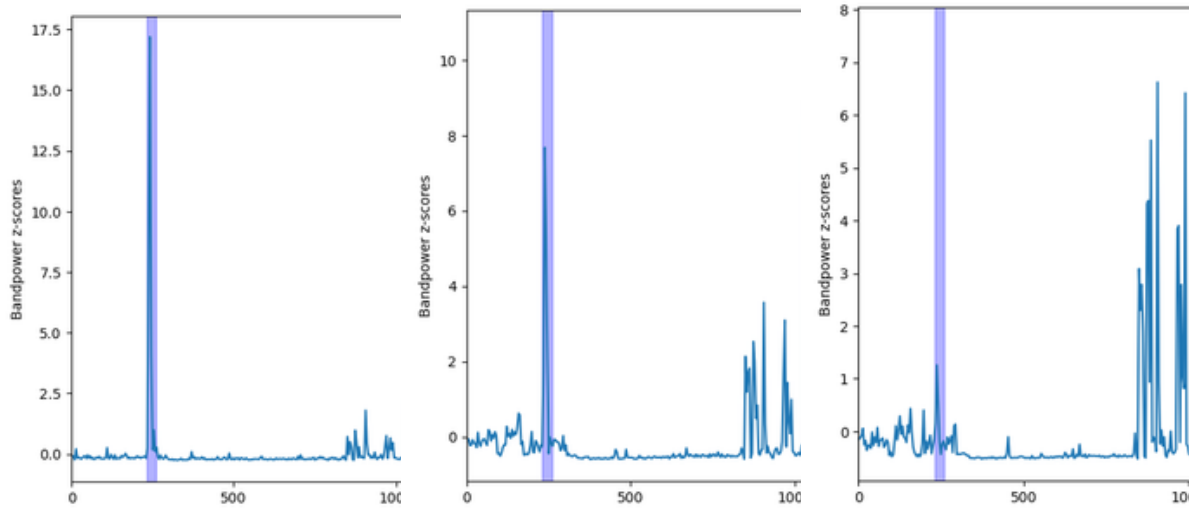

Figure 2. Channel FP1-F7 alpha, beta, and gamma band power z-score graphs

Conclusions:

Evidently the majority of seizures show as peaks in bandpower, though perhaps not always peaks in amplitude as we initially predicted. In addition, seizures show the greatest peaks in the FZ-CZ and CZ-PZ channels. While our program is a step in the right direction, it is imperfect. There were still some peaks that appeared on the three relevant z-score graphs that were not seizures.

Since accurate detection of seizures for epileptic patients is frequently a matter of life and death, it is vital to ensure our system is as accurate as possible.

Future Directions:

One possible next step would be for someone to improve our system and find more effective ways to contrast seizure and non-seizure data, as well as detect signs of seizures in the waveforms before they happen. This would allow the system to predict when an actual seizure will happen in a user, as well as weed out false negatives that may be caused by interference or other anomalous brain activity. In the end, this would help work toward the goal of creating an EEG system with fewer channels that is more comfortable and accurate.

Another possible next step would be for someone to modify our system and enable it to detect where a seizure originated in the brain in the first place. This would allow potential users of the system to not only prepare for an oncoming seizure, but also for them to work with doctors to select the best treatment plan. Overall, we hope that our research can inspire and aid others in their future work surrounding seizures.

References:

[1]Shoeb, Ali. Application of Machine Learning to Epileptic Seizure Onset Detection and Treatment. PhD Thesis, Massachusetts Institute of Technology, September 2009. Accessed 16 Aug. 2023.

[2]Bifftu, Berhanu Boru et al. “Seizure-Related Injuries among People with Epilepsy at the Outpatient Department of the University of Gondar Hospital, Northwest Ethiopia: Cross-Sectional Institutional-Based Study.” *Neurology research international* vol. 2017 (2017): 4970691. doi:10.1155/2017/4970691. Accessed 16 Aug. 2023.

[3]Haddad, Tahar, et al. “Temporal Epilepsy Seizures Monitoring and Prediction Using Cross-Correlation and Chaos Theory.” *Healthcare Technology Letters*, 21 Mar. 2014, www.ncbi.nlm.nih.gov/pmc/articles/PMC4611494/#:~:text=EEG%20signals%20are%20i. Accessed 16 Aug. 2023.

[4]Morley, Andrew. “ERS-Education.” *10-20 System EEG Placement*, 2016, www.ers-education.org/lrmedia/2016/pdf/298830.pdf. Accessed 16 Aug. 2023.

[5]Shah, Vinitshah. “ArXiv.Org e-Print archive OPTIMIZING CHANNEL SELECTION FOR SEIZURE DETECTION .” *Optimizing Channel Selection For Seizure Detection*, arxiv.org/ftp/arxiv/papers/1801/1801.02472.pdf. Accessed 16 Aug. 2023.

[6]Smith, Shelagh. “EEG in the Diagnosis, Classification, and Management of Patients with Epilepsy.” *Journal of Neurology, Neurosurgery & Psychiatry*, 1 June 2005, jnnp.bmj.com/content/76/suppl_2/ii2. Accessed 16 Aug. 2023.