Detecting Epileptic Seizures in Electroencephalogram Data

Nick Hershey

Collaborators: Dr. Chris Lee-Messer, MD, PhD; Jared Dunnmon, PhD; Prof. Christopher Ré, PhD

CS230: Deep Learning

I) Task

Overview: Affecting fifty million people worldwide, epilepsy is a chronic disorder of the central nervous system characterized by recurrent seizures. During a seizure, aberrations in the brain's electrical activity produce physical symptoms ranging from convulsion to loss of memory to unconsciousness. An electroencephalogram (EEG) is a record produced by electrophysiological monitoring of the electrical activity of the brain. Electrodes are placed on the scalp and measure voltage fluctuations between the nodes as the net effect of millions of neurons in the brain. EEGs are used for diagnosis of a number of neurological disorders, including epilepsy, sleep disorders, comas and more.

Our goal task is to <u>train a single neural network to classify an epoch of EEG data from any patient as being seizure or non-seizure</u>. Previous attempts at EEG classification have fallen short by

1. Building separate models for each patient
2. Using data sets of only certain seizures and non-seizure activities.
3. Yielding high false positive rates

We hope to overcome these shortcomings by using vastly more data than previous attempts. For example, Shoeb (2009) used 23 patients and 844 EEG files, whereas we have access to 12,385 patients and 136,363 EEG files.

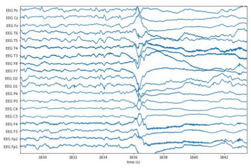

Figure 1: An example EEG where a seizure begins around t=1836 seconds

II) Dataset

Put briefly, our data set is <u>huge and heterogeneous</u>. The data set consists of 136,363 electroencephalograms: 99,721 from adults measured at the Stanford Hospital and 36,642 from children measured at the Lucile Packard Children's Hospital. Each electroencephalogram is stored in a hierarchical data format (HDF) containing anonymized data about the patient, metadata about the EEG read, the raw signals of the EEG read (a matrix with shape *number of channels* by *length of EEG*), and accompanying annotations with timestamps for the EEG. The data is remarkably heterogeneous:

- *Type*: intracranial and scalp
- *Reason*: routine, long-term EEG, ambulatory study, etc.
- *Length*: 5 seconds to 24 hours
- *Frequency*: 200 Hz to 500 Hz
- *Age*: 0 to 100+
- *Channels*: 3 to 142

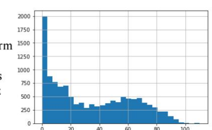

Figure 2: Histogram of patient ages

Because the data is not explicitly labeled as seizure and non-seizure, we use the nurse annotations as a proxy for seizure labels.

III) Data Pre-Processing

Given the heterogeneity of the data, we filtered our data by:

- Only including scalp EEG (i.e., excluding intracranial reads).
- Taking reads only from the nodes in the International 10-20 System.
- Limiting length to minimum time of seizure: 10 seconds.
- Only including files with a sampling rate of 200 Hz.
- Standardizing each waveform to have mean 0 and standard deviation 1.

With these filters in place, we defined our seizure and non-seizure data:

- <u>**Seizures**</u>: Every time a file contains a seizure annotation, slice the following 10 seconds. There are 11,641 unique files with 25,850 labeled seizures fitting the above filtering.
- <u>**Non-seizures**</u>: For each file without a seizure annotation, randomly sample a 10-second slice. Randomly sample from all such slices to get 25,850 non-seizures occurrences.

The result is 51,700 matrices shaped 25x2000 with a binary label.

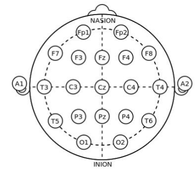

Figure 3: The nodes used in the International 10-20 System

IV) Proposed Architectures

We tested four different model architectures:

1. **Baseline**: Flatten the matrix and run logistic regression via a one-layer neural network with no hidden layers.
2. **Dense Network**: Flatten the matrix and run it through two hidden layers with ReLU activation functions before outputting a single neuron with a sigmoid activation function.
3. **Convolutional Network**: Run matrix through five 1-dimensional convolutional layers each with 10x1 filters, 3x1 max-pooling, and exponential linear units. Flatten the remaining matrix and run it through one layer with batch norm, dropout, and ReLU before a final layer with a sigmoid activation function. This model is based off Schirrmeister's work (2017) on EEG decoding.
4. **Recurrent Network**: Process each of the 2000 inputs into two bi-directional LSTM units followed by mean pooling, a dense layer with batch norm and ReLU, and a final layer with a sigmoid activation. This is based on many-to-one recurrent networks on temporal data.

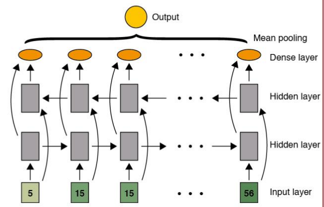

Figure 4: In a bi-directional recurrent neural network with hidden layers, each time-step of input is fed to each LSTM unit which then passes information forward and backward through time.

V) Results

Model	Data	Accuracy	F1 Score
Logistic	Train	56.5%	0.535
Logistic	Test	54.6%	0.533
Dense	Train	90.8%	0.921
Dense	Test	90.2%	0.897
Convolutional	Train	96.5%	0.947
Convolutional	Test	50.2%*	0.522*
Recurrent	Train	60.1%	0.621
Recurrent	Test	58.3%	0.592

- As expected, the worst performing architecture was the logistic regression model. With just 50,000 parameters, it couldn't fit much beyond guessing.
- Our top performing architecture on the test set was the densely connected feed-forward neural network, achieving an accuracy and F1 score of 90.8% and 0.897, respectively. The architecture's flexibility likely was its key to success with over 1 billion parameters to tune.
- The convolutional neural network also performed impressively. It achieved the highest training accuracy and F1 score at 96.5% and 0.947, respectively. However, this model may have over fit given its poor test set performance. *From preliminary error analysis, this difference seems to be an error in the convolutional evaluation code.
- The bi-directional LSTM network performed okay with an accuracy around 60%. Its lackluster performance is expected given the known challenge of training an LSTM with such a large number of time steps.

These results are at apparent convergence after 50 epochs of training.

VI) Discussion and Next Steps

Our top-performing architectures were the densely-connected and the convolutional networks. Both achieved accuracies and F1 scores over 90%, well-beyond the results of published work for a general EEG classifier used on any patient to classify any of the types of seizures. This is especially remarkable given the heterogeneity of the data, which contains ample non-seizure activity, many types of seizures, and many patients.

That said, we see four major areas of improvement for this task:

1. **Improve Labels**: The current seizure annotations are not completely trustworthy nor complete. We can apply labeling functions to generate sound training data as described in Ratner et al (2017).
2. **Improve Embeddings**: We currently use the raw waveform. However, Bashivan et al (2016) proposes an EEG "video" as discussed in Figure 5 that may better capture temporal and spatial information.
3. **Improve Network**: Our proposed models have been relatively shallows compared to what Dai et al (2016) suggest is necessary to capture all information from raw waveforms.
4. **Segmentation**: We would love to expand our slices to other time points of a seizure or segmentation of an EEG file into seizure and non-seizure activity. Unfortunately, the annotations usually only mark seizure start, so this is difficult.

This work is already clinically useful as a tool to identify seizures in EEGs. An ambitious additional goal is to forecast a seizure *t* seconds ahead of time. We plan to try this but do not expect great results as even the best clinicians cannot forecast seizures from EEG data.

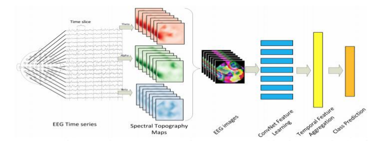

Figure 5: Bashivan et al propose representing an EEG as a video by (1) projecting the nodes into two-dimensions, (2) taking the Fourier intensity of three discretized buckets over a t-second interval, (3) combining these three channels to create an image, and (4) combining images temporally to make a video.

VII) References and Code

My code can be found at https://github.com/nhershey/cs230eeg.

References cited in the poster are (see report for all sources):

1. Bashivan, Pouya, et al. "Learning representations from EEG with deep recurrent-convolutional neural networks." *arXiv preprint arXiv:1511.06448* (2015).
2. Dai, Wei, et al. "Very deep convolutional neural networks for raw waveforms." *Acoustics, Speech and Signal Processing (ICASSP), 2017 IEEE International Conference on. IEEE*, 2017.
3. Ratner, Alexander J., et al. "Snorkel: Fast training set generation for information extraction." *Proceedings of the 2017 ACM International Conference on Management of Data*. ACM, 2017.
4. Schirrmeister, Robin Tibor, et al. "Deep learning with convolutional neural networks for EEG decoding and visualization." *Human brain mapping* 38.11 (2017): 5391-5420.
5. Shoeb, Ali Hossam. "Application of machine learning to epileptic seizure onset detection and treatment." *Diss. Massachusetts Institute of Technology*, 2009.

Stanford