Generative Modeling and Prediction of Spontaneous Epileptic Seizures

Christine Liu
cliu99@stanford.edu

Peter Maldonado
phm@stanford.edu

Keegan Mehall
kmehall@stanford.edu

1 Introduction

Epilepsy is the 4th most common neurological condition affecting more than 65 million people (Carney, 2011). Absence epilepsy, a particular form of epilepsy, is a neurological disorder affecting children between the ages of 4 and 12, and accounts for approximately 10% of all patients with epilepsy (Jacoby, 2013). Although seizures have traditionally been described as stochastic events, more recent studies have shown that the brain undergoes notable preictal (pre-seizure) electrophysiological changes, which may predict seizure severity (Mormann, 2007). These preictal changes demonstrate a dynamic shift in brain state associated with an evolution into a seizure, and are clinically useful as patients may be warned of an impending seizure prior to its initiation. Currently, there is no model that can accurately predict epileptic seizures with sufficient warning time to administer anti-convulsant medications or relocate the patient to a safe location (Jacoby, 2013).

Current literature has focused on a binary classification approach between pre-seizure and seizure signals of an electroencephalogram (EEG), which is an electrophysiological monitoring method consisting of electrodes used to record electrical activity of the brain (Usman, 2019). For our binary task, the input is EEG signals divided into two categories for seizure and non-seizure. We then used logistic regression and support vector machines to output a binary prediction of whether or not the patient would have a seizure. For the generative modeling task, we inputted the pre-seizure portion of the EEG signal, and used LSTMs in order to output a prediction of the shape and magnitude of the seizure signal. The ultimate goal of the project was to use both of these predictive tasks in tangent, by first predicting if an EEG signal is a seizure, and then trying to use the preictal signal to predict the nature of the seizure itself.

2 Related Work

Previous work with general seizure prediction has focused on a binary prediction approach. Petrosian et al. was one of the first studies that focused on investigating the existence of a preictal stage before the seizure through the use of wavelet transformations and found that they were able to detect the presence of preictal stages before the occurrence of the seizure (Petrosian, 1999). These findings opened up the way for future predictive methods and studies based on classification with the preictal period. A study done by Li et al. in 2013 showed that seizures could be predicted up to 10 seconds prior to seizure onset with sensitivity of 0.758, despite previous studies basing prediction off of much longer timescales (Li, 2013). This allowed future studies to be based off of shoter time scales, which helped simplify the prediction problem and amount of data.

Two additional studies in 2017 used support vector machines to classify EEG seizure signals. One of the studies, Sharif et al. from 2017, was able to achieve sensitivities between 91.8-96.6% (Sharif 2017). The other study, Direito et al., was the first study to implement a realistic seizure prediction approach by using multi-channel high-dimensional datasets as opposed to the typical dimensionality reduction techniques used in previous papers (Direito, 2017). By using multiclass support vectors machines on 1206 seizures, Direito et al. was able to achieve sensitivity of 38.47%. While this number was rather low, it showed that seizure prediction could be applied to realistic clinical datasets that have much more data than previous studies. This gave our group the confidence to try and use a large, and realistic EEG dataset with 1913 seizures.

Finally, Tsiouris et al. was the first to use long short-term memory (LSTM) networks for EEG signal prediction and was able to achieve sensitivity over 99% (Tsiouris, 2018). We chose to mimic this use of LSTMs for seizure prediction, but rather than use it in the binary prediction methods that Tsiouris et al. used, we chose to implement it for seizure signal generation based of the raw preictal signal.

Our project specifically focused on absence seizures, and only used an additional type of generalized seizure data as validation. Absence seizures have traditionally been thought of as completely unpredictable events with no defined correlation between different network states (Danober, 1998). Behaviorally, absence seizures are characterized by periods of quiet wakefulness, when delta waves, brain wave oscillations recorded in an EEG between 1-4 Hz, are most prevalent in the brain (Marescaux, 1992). Absence epilepsy is characterized by abnormally synchronous electrical activity within two mutually connected brain regions, the thalamus and cortex (Fabri, 1991). This absence seizure dataset our group has allowed us to explore novel approaches for a difficult type of seizure prediction.

3 Dataset and Features

We worked with 2 datasets. First, an electroencephalogram (EEG) dataset that Christine Liu has access to from the Huguenard Lab. This dataset has seizure data from 9 rats and mice with genetic absence epilepsy (WAGRij) consisting of hours of extra-cellular recordings of individual cortical and thalamic neurons from rodents with genetic absence epilepsy. We chose to use 4 thalamic channels in order to reduce dimensionality, as the EEG data is consistent among channels.

The data is segmented into 12 second epochs broken into 3 components. The first window consists of the interictal signal, which is associated with no seizures and lasts 5 seconds. The interictal period has been described in literature to be consistent across animals and across time, so we felt confident in defining the direct period prior to the preictal period to be the interictal period (Danober, 1998). The second window consists of the preictal signal which lasts for 3 seconds. We chose this number to define the preictal period based off of current literature with absence epilepsy (Sorokin, 2016). The third window consists of the seizure itself, lasting for 4 seconds.

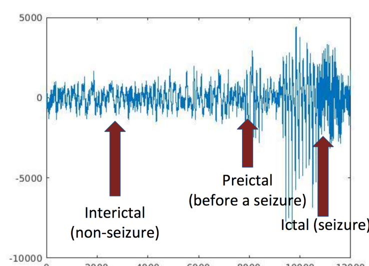

Figure 1: Pre-ictal to ictal EEG epoch (Huguenard dataset). Horizontal axis is time in ms from the start of the 12 second segment and the vertical axis is voltage in microvolts

The second dataset is hosted on Kaggle and contains data from intracranial EEGs from both dogs and humans. The EEG is sampled from 16 electrodes at 400 Hz and 5000 Hz, and contains the recorded voltages. Absence epilepsy has been noted to be highly stochastic, and in light of that, we chose to also try the binary prediction task with Kaggle data which focuses on generalized tonic-clonic seizures which are less stochastic in nature. The Kaggle data is pre-processed, as opposed to the raw recordings from the first lab dataset. In addition, the inclusion of three total types of data (human, dog, and rodent), allows the model to both improve generalizability, as well as offers a potential translational aspect of the project with the goal being to predict epileptic seizures in humans.

In addition to treating treating our data as time-series data, we also performed feature extraction by transforming each signal with a discrete fast fourier transform (FFT). This transposes our problem out of the time-domain into the frequency domain. The use of frequency domain analysis is well established in epilepsy prediction (Carney 2011).

Each output of the discrete Fourier transform is defined by the equation

$$F_n = \sum_{k=0}^{N-1} f_k \exp(-2\pi i n k / N)$$

for N data points, with f_k the value at point k in the time series data. This results in a complex vector with elements whose magnitude represents the amplitude of the signal at that frequency and the angle in the complex plane represents the phase of the signal.

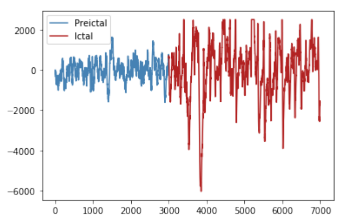

Figure 2: An example EEG signal

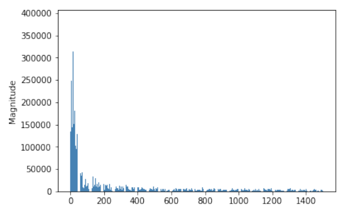

Figure 3: The FFT of the preictal signal. The x axis is the output index from the FFT; corresponds to units of $\frac{1}{3}$Hz

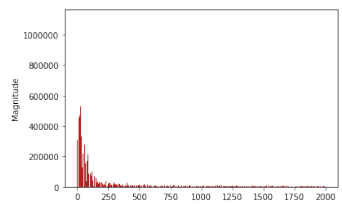

Figure 4: The FFT of the ictal signal. The x axis is the output index from the FFT; corresponds to units of $\frac{1}{4}$Hz

4 Methods

4.1 Step 1: Classify as Seizure or Non-Seizure

This task focuses on classifying a time-series segment as being either preictal with a positive label, or interictal with a negative label. We use a few metrics to evaluate performance on this task, including classification accuracy and sensitivity.

We divided our absence seizure dataset into discrete epochs which are labeled as either seizure (positive class) or non-seizure (negative class). We chose to apply logistic regression with scikit-learn. This decision treats each signal as a stochastic process, where we maximize the likelihood that a given signal is properly classified. This is a reasonable assumption, as the EEG data is collected across a number of noisy channels during each time step.

We also chose to apply an SVM with a RBF kernel function. The SVM learns the maximum margin decision boundary, which in the context of our problem means we focus on dividing the mass of seizure signals from that of the non-seizure signals. The RBF kernel can be interpreted both as a similarity metric, which helps our model classify similar signals consistently, and as a projection into an infinite dimensional feature space, which is helpful as our EEG data is representing significantly higher-dimensional neural activity in the subjects. We also chose to up-sample and down-sample the minority and majority classes respectively to address class imbalance.

4.2 Step 2: Longer Timescale Seizure Prediction

This task takes a time series EEG signal (classified as pre-ictal from the previous task) and predicts how it evolves forwards in time–effectively a time-series prediction problem.

Recent literature suggests that seizure prediction is a longer time-scale problem than previously thought, so this approach builds off of new knowledge in the field (3). Ideally, performing well on this task allows us to give a seizure onset time, which has not been successfully demonstrated before, along with predicting other clinically-useful properties of a seizure.

Here our task can be framed as sequence-to-sequence translation, where the input is a pre-ictal signal and the output is the corresponding future signal during the seizure.

Deep learning methods have proved successful for this class of task, so we experiment with the application of long short-term memory (LSTM) models. These are empirically well-adapted for longer time-series data (compared to vanilla recurrent neural networks (RNNs) or gated recurrent units (GRUs).

For our network architecture, we train a single-layer LSTM with four input channels, and a hidden size of 256, which is decided to four output channels by a single dense layer. We train our model with a batch size of 32 for 100 epochs. The learning rate is initially set to 1e-5, and dropped to 1e-6 after 50 epochs in order to enable the network to better fine-tune.

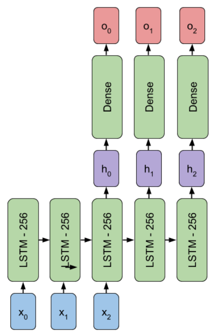

Figure 5: The Encoder-Decoder LSTM

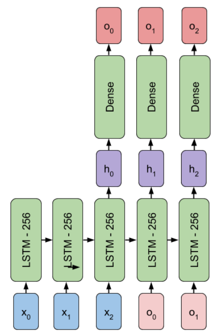

Figure 6: The Repeated Prediction LSTM

Furthermore, we apply our LSTM in two different ways. The first is an encoder-decoder model, where we encode the input time series into a hidden representation that is decoded over all desired time steps. Under this interpretation, the model tries to learn the dynamics of the seizure in some high-dimensional space, this is then decoded by the single dense layer. The second application is to repeatedly predict a single time step, which is fed back into the LSTM as the input during the subsequent time step (in contrast to the first application, where the inputs during decoding are zeros). We implement these models in PyTorch that are available with our accompanying source code.

We also train out network using two different loss functions, both mean-squared error (MSE) and mean absolute error (MAE). This is because the squared term in MSE tends to have a smoothing effect between multiple values in a vector, causing the model to favor many medium deviations rather than small deviations and a few large ones. In the context of a time-series prediction problem this corresponds to have mediocre results among at many times steps, rather than poor results at a few.

5 Results

5.1 Step 1: Classify as Seizure or Non-Seizure

We include a table of results from our binary classification experiments.

	Dataset 1 Accuracy	DS 1 False Positives	DS 1 False Negatives	DS 2 Accuracy
LR	0.552	751	964	0.765
SVM	0.633	7	1396	0.878
LR w/ FFT	0.575	444	1183	x
SVM w/ FFT	0.620	6	1453	x

The above results are were our best for accuracy, but we also obtained this confusion matrix: $\begin{bmatrix} 1339 & 1031 \\ 800 & 656 \end{bmatrix}$ with both upsampling and Fourier transform for features, which performed best in terms of balancing accuracy with false positives/negatives. Most notably we observe an imbalance between the false positives and negatives from our logistic regression and SVM models. This is likely because logistic regression is a calibrated model, while SVMs are not, while leads to higher underfitting and overprediction the majority class.

5.2 Step 2: Longer Timescale Seizure Prediction

Below we include sample outputs from our generative modeling task. As mentioned before, our predictive task involves for thalamic channels of EEG signal, although because our model performs similarly on all of them, we demonstrate only a single channel in the figures below.

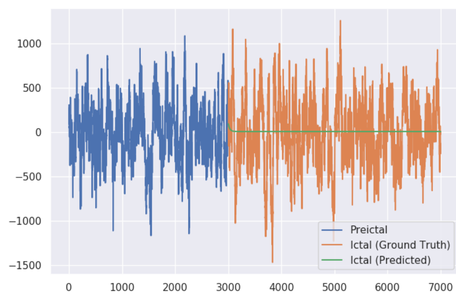

Figure 7: Sample seizure signal generated by the encoder-decoder LSTM model.

We observe that the encoder-decoder model works does not work well at predicting time steps far from the end of the pre-ictal signal. We attribute this to a number of factors, namely the stochasticity of the data and the high cost of data collection. It is hard to learn the complex dynamics of a seizure without a large network, but training a model with that many parameters would require significantly more data. Because our model is not expressive enough, it simply optimizes to the local minimum of predicting the mean value of the signal over time.

Further, we notice that the repeated prediction architecture outperforms the encoder-decoder model, likely because it is continuously given its past output as input. Specifically we note that while many of the encoder-decoder predictions look near-identical, the two predictions from the repeated prediction LSTM appear to better capture the starting trend of their respective signals.

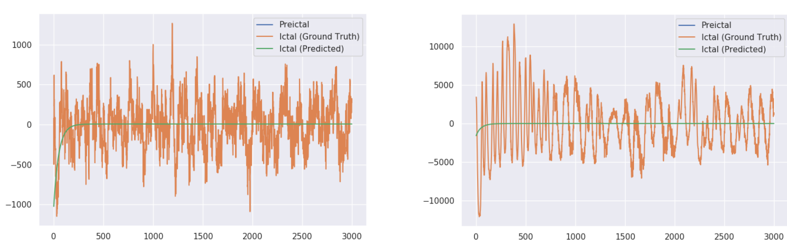

Figure 8: Predictions for two ictal signals from the Repeated Prediction LSTM model.

6 Conclusions and Future Work

This study aimed at predicting epileptic seizures and seizure signals using both binary classification and a generative modeling methods. Through the use of logistic regression and support vector machines, our binary prediction models were able to achieve accuracies between 55-63% on the absence seizure dataset, and between 76-87% for the Kaggle dataset. The absence seizure prediction rates were lower than the already-cleaned Kaggle data because the absence seizure dataset was raw and mostly unprocessed. For the generative task, none of the results were statistically significant. We attribute this to the highly stochastic nature of the data. An extension of the project could be to use LSTMs to generate the shape of the preictal signal as opposed to the seizure itself, as the voltage of the preictal signal is more closely correlated with the voltage of the interictal signal than the voltage of the seizure itself. In addition, we could also spend more time cleaning and processing the raw data through low-pass filtering or tensor decomposition to reduce dimensionality.

7 Contributions

All authors contributed equally to this work. Christine and Peter helped train the models. Peter and Christine and Keegan helped with the data preprocessing and data organization/infrastructure. All three members wrote the paper.

A zip file of our project code can be found here:
https://drive.google.com/file/d/1iWZolVr2K75TVCaxqdhMvDgmM5ybY0hu/view?usp=sharing

8 References

Carney, Paul R et al. “Seizure prediction: methods.” Epilepsy & behavior : EB vol. 22, Suppl 1 (2011): S94-101. doi:10.1016/j.yebeh.2011.09.001

Florian Mormann, Ralph G. Andrzejak, Christian E. Elger, Klaus Lehnertz, Seizure prediction: the long and winding road, Brain, Volume 130, Issue 2, February 2007

Usman Syed Muhammad, Shehzad Khalid, Rizwan Akhtar, Zuner Bortolotto, Zafar Bashir, Haiyang Qiu, "Using scalp EEG and intracranial EEG signals for predicting epileptic seizures: Review of available methodologies," Seizure, Volume 71, 2019, Pages 258-269

Petrosian A, Prokohorov D, Homan R, Dasheiff R, Wunsche D, "Recurrent neural network based prediction of epileptic seizures in intra- and extracranial EEG." Neurocomputing 2000l 30-201-18. 1999.

Li S, Zhou W, Yuan Q, Liu Y. "Seizure prediction using spike rate of intracranial EEG". IEEE Trans Neural Syst Rehabil Eng 2013; 21: 880-6. 2013.

Sharif B, Jafari AH. "Prediction of epileptic seizures from EEG using analysis of ictal rules on Poincare plane". Computing Methods Prob Biomed 2017; 145: 11-22. 2017.

Direito B, Teixeira CA, Sales F, Castelo-Branco M, Dourado A. "A realistic seizure prediction study based on multiclass SVM." International Journal of Neural System 2017; 27:1750006. 2017.

Kaggle Dataset. https://www.kaggle.com/c/seizure-prediction/data

Jacoby, A., Baker, G., Crossley, J., and Schachter, S. "Tools for assessing quality of life in epilepsy patients." *Expert review of neurotherapeutics*, pp 1355-1369, 2013.

Danober, L., Deransart, C., Depaulis, M., Vergnes, M., Marescaux, C. "Pathophysiological mechanisms of genetic absence epilepsy in the rat." *Progress in Neurobiology*, pp 27-57, 1998.

Sorokin, J., Paz, J., Huguenard, J. "Absence seizure susceptibility correlates with pre-ictal beta oscillations." *Journal of Physiology-Paris*, pp 372-381, 2016.

Marescaux, C., Vergnes, M., Depaulis, A. "Genetic absence epilepsy in rats from Strasbourg- A review." *Journal of Neural Transmission*, pp 37-69, 1992.

Fabri, M., Burton, H. "Ipsilateral cortical connections of primary somatic sensory cortex in rats." *Journal of Comparative Neurology*, pp 405-424, 1991.

Malik, A., Wu, A., Vasavada, V., Song, J., Mitchell, J., Goodman, N., Piech, C. "Generative Grading: Neural Approximate Parsing for Automated Student Feedback" - *arXiv preprint* arXiv:1905.09916, 2019

McGinley, M., Vinck, M., Reimer, J., Tolias, A., Cardin, J., McCormick, D. "Waking State: Rapid Variations Modulate Neural and Behavioral Responses." *Neuron*, pp 1143-1161, 2015.

Sorokin, J., Ganguli, S., Huguenard, J. "Identification of unique pre-ictal states via non-negative tensor decomposition of single unit recordings." *COSYNE Conference*, 2015.

Michielli, N., Acharya, U., Molinari, F. "Cascaded LSTM recurrent neural network for automated sleep stage classification using single-channel EEG signals." *Computers in Biology and Medicine*, 2019.