Applying Machine Learning for Human Seizure Prediction

Ehsan Dadgar-Kiani, Cagan Alkan, Ali Shameli
{dadgarki,calkan,shameli}@stanford.edu

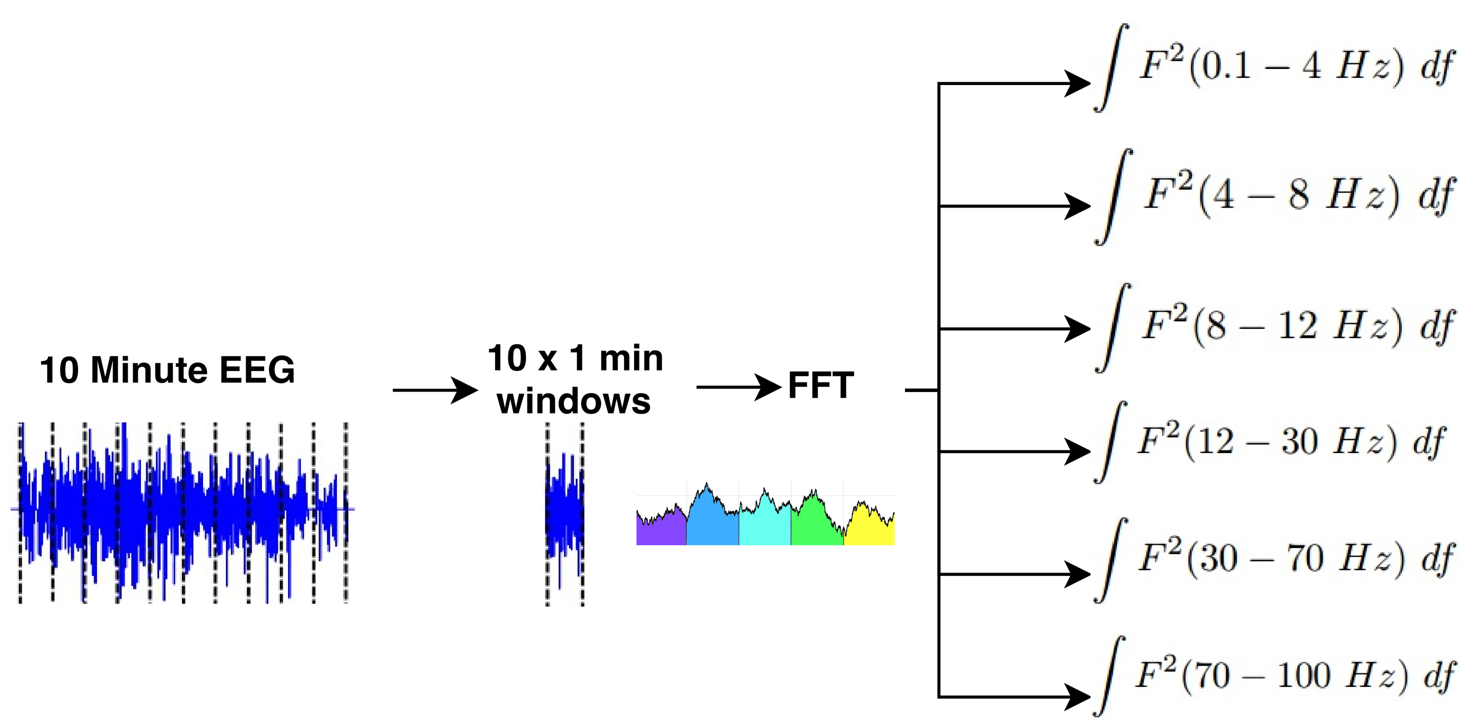
Overview

Epilepsy is a neurological condition associated with sporadic abnormal electrical activity ('seizures') in the brain. It is important to identify the periods when patients are more likely to have seizures, so they can avoid dangerous activities or take medications only when it is necessary. The proposed method is to use established machine learning methods to train a model for predicting seizures from 10 minute human iEEG (intracranial electroencephalography) recordings and classify it as either 'preictal' (pre-seizure) or 'interictal' (non-seizure).

Data

Data was recorded in the form of iEEG, which is obtained by positioning electrodes on the surface of cerebral cortex and measuring electrical signals. The data was available as part of an online competition through Kaggle.com [1]. We used data from a single patient (out of three possible patients). Each 10 minute period can be represented by a matrix where the row represents time, and column represents electrode channel (out of 16), and is also labeled as one of the two classes, preictal (1) or interictal (0).

Features



- Each 10 minute recording (sampled at 400Hz) was divided into 1 minute consecutive windows
- For each window, the Fast Fourier Transform (FFT) was taken from each channel
- Power in each frequency band of the signal is calculated by summing up squares of the magnitudes of FFT coefficients corresponding to each frequency band shown in the figure above.

Each power calculation represented one feature, resulting in a feature size of 6 bands $\times$ 10 segments $\times$ 16 channels = 960 features per recording. This is a reduction from 4 million time points in the raw data.

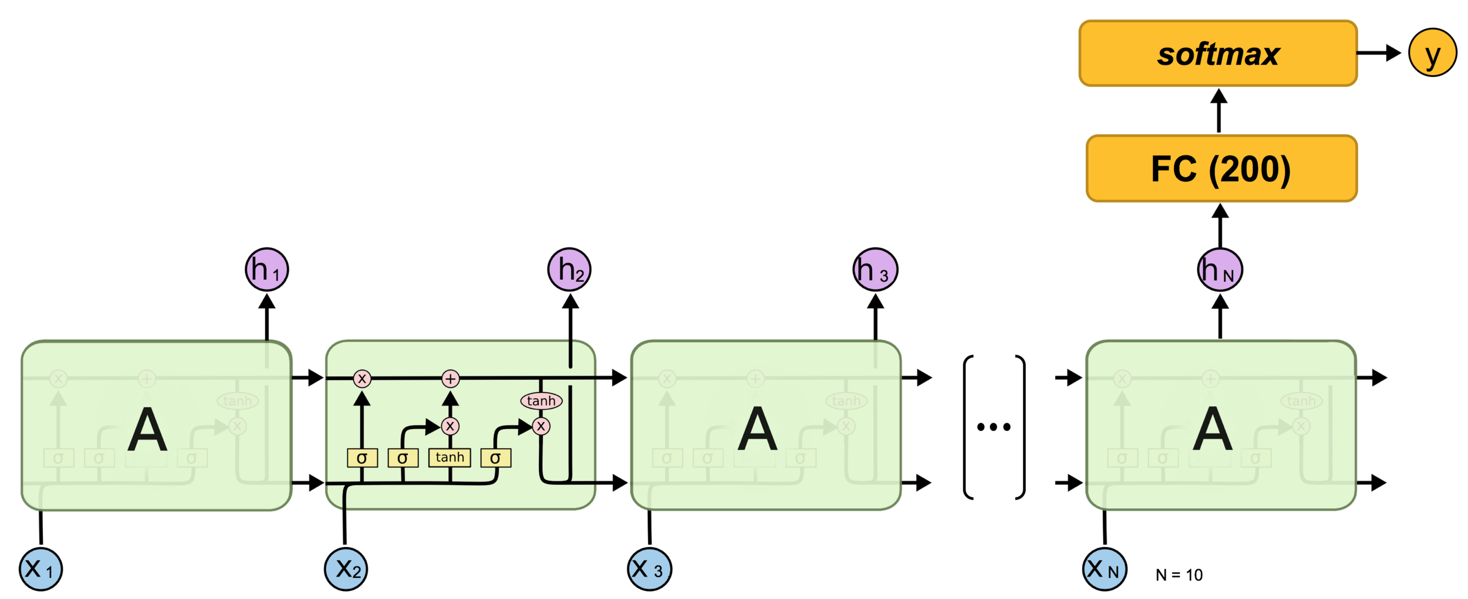
The main advantage of this feature extraction scheme is that it gives information about both the time course and the frequency spectrum of the iEEG signals.

Models

Using the features described above, we implemented Logistic Regression, SVM (Linear Kernel), SVM (RBF), and LSTM. These were implemented in Python using the scikit-learn and Keras packages. All models were trained using 70-30 hold-out cross validation. L2 regularization was used for both Logistic Regression and SVM models.

Description of LSTM Model

- Features are segregated by their minute window and fed into their corresponding LSTM unit
- Sigmoidal Activation Function
- LSTM Hidden State Size of 100
- Final hidden state connected to FC layer of size 200
- Binary cross-entropy objective used for optimization
- SGD Batch-Size of 100



The update equations for a single LSTM cell are given by [2]:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f)$$

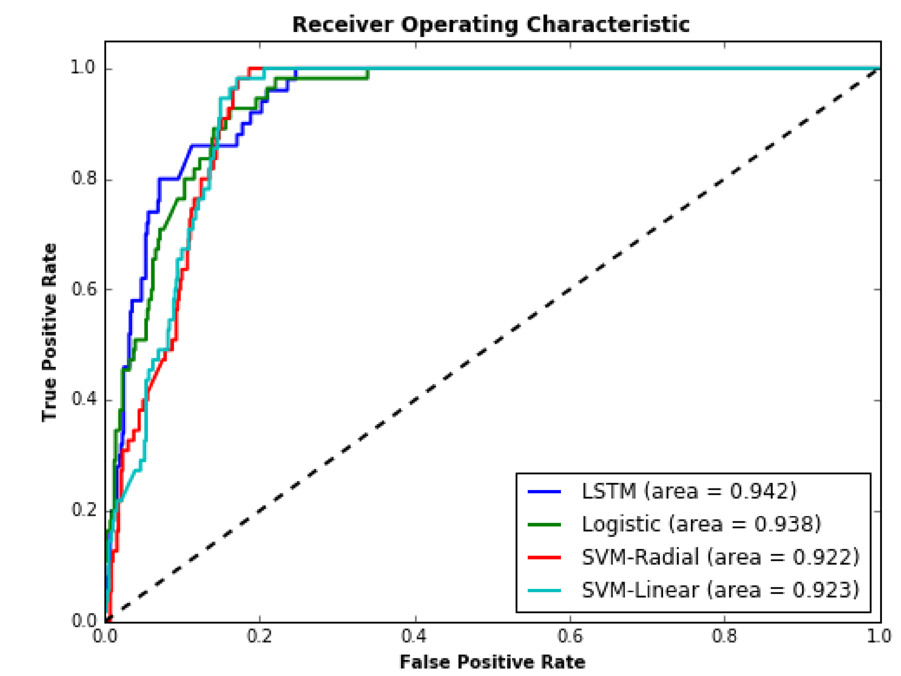
$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i)$$

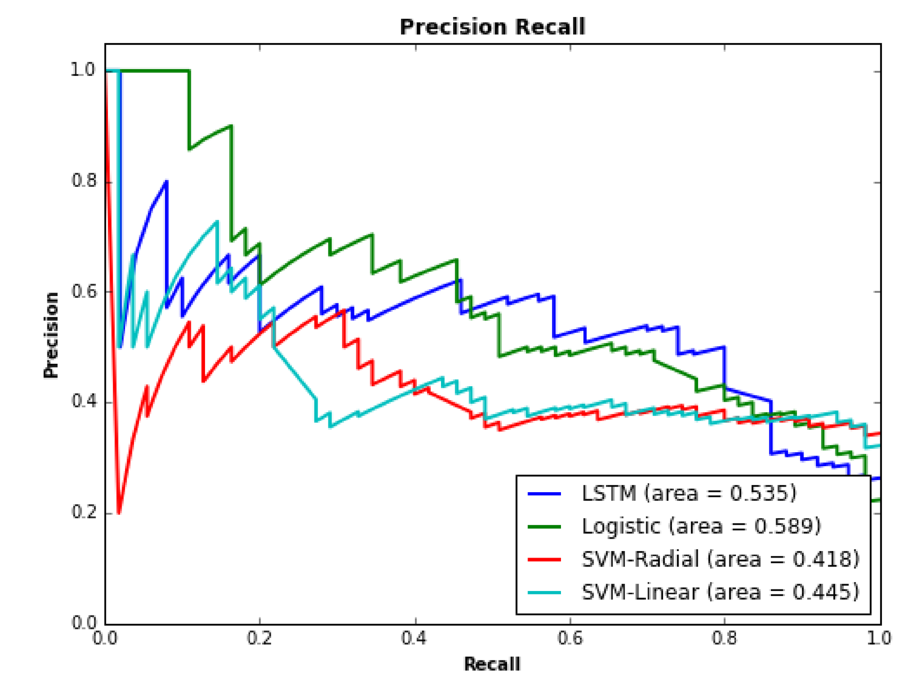
$$C_t = f_t * C_{t-1} + i_t * tanh(W_C \cdot [h_{t-1}, x_t] + b_C)$$

Results

The training set size was 1440, while the testing set size was 618. It is also important to note the class imbalance between pre-ictal examples (150) and inter-ictal examples (1908), which is why we relied primarily on AUROC and recall over accuracy.

Model	Training Acc.	Test Acc.	AUROC	F-Score	AUPRC	Recall
LSTM	0.852	0.854	0.942	0.489	0.534	0.860
Logistic Regression	0.909	0.872	0.938	0.533	0.589	0.818
SVM (RBF)	0.874	0.861	0.923	0.538	0.434	0.909
SVM (Linear)	0.899	0.867	0.923	0.445	0.506	0.764





Discussion

Our results demonstrate the feasibility of using spectral power across different bands to forecast a seizure event. For this application, it is vital to not miss a seizure, thus our emphasis on optimizing the recall rate. For most of our models, our recall was high but precision was low, which explains the low F-scores. All models had high AUROC, with the LSTM model having the highest. We were unable to train a reasonable model when using a data set with multiple patients simultaneously, so such a model would only be customized for a single patient.

Future Directions

- Incorporate more features (DWT, time-series statistics, etc.) into feature set
- Use autoencoders for nonlinear dimensionality reduction before performing supervised learning
- *Domain adaptation* techniques for applying trained model to multiple patients

References

[1] Kaggle.com, "Melbourne University AES/MathWorks/NIH Seizure Prediction." https://www.kaggle.com/c/melbourne-university-seizure-prediction, 2016. [Online; accessed October-22-2016].

[2] "Understanding lstm networks." http://colah.github.io/posts/2015-08-Understanding-LSTMs/, 2015. [Online; accessed October-22-2016].