Seizure Prediction Based on EEG Data

6th January 2017

Kumar Saurabh (CS1522)
Ranjan Mishra (CS1519)

Introduction

- Epilepsy afflicts about 1% of world’s population.
- It is characterised by spontaneous seizure.
- Seizures are known to occur in clusters or groups.
- For 20-40% patients with epilepsy, medications are not effective.
- Even after surgical removal of epilepsy, medications are not effective.

Temporal Dynamics of Brain Activity

For a epilepsy patient can be classified into four states:

- Preictal - Before seizure
- Interictal - Between seizure
- Ictal - During seizure
- Postictal - After seizure

Recording EEG data of ictal state is near to impossible and also postictal data would not be any help. Focus remains on the interictal and preictal states.

Brain Activity Based Seizure Forecasting

- EEG data captures the activity of brain, with high temporal resolution.
- Using EEG data, it might be possible
  - To identify increased probability of seizure occurrence.
  - Design devices to warn seizure.
- It would help patients to avoid dangerous activity such as driving, swimming etc.
- Medication could be administered when needed to prevent impending seizure.

Objective

To distinguish between 10 min long preictal EEG data clip and 10 min long interictal EEG data clip, and to demonstrate the existence of classification of the preictal brain state in humans with epilepsy.

Data Acquisition

We gathered data from the online public competition hosted by Kaggle.com, which is one of the most popular online machine learning and data science competition hosting website, the competition was sponsored by Mathworks, National Institute of Health, University of Melbourne.

More details about competition is available on:
https://www.kaggle.com/c/melbourne-university-seizure-prediction

Data Description

- Human brain activity was recorded in the form of EEG data, which involves electrode positioning on cerebral cortex and recording the electrical signals.
- Data was sampled from 16 electrodes at 400 Hz.
- These are long duration recording, spanning multiple months to multiple years and recording many seizures.
- Each data clip is given as 10 min of EEG signals.
- Each clip is marked with preictal and interictal in training set.

Preprocessing of Data

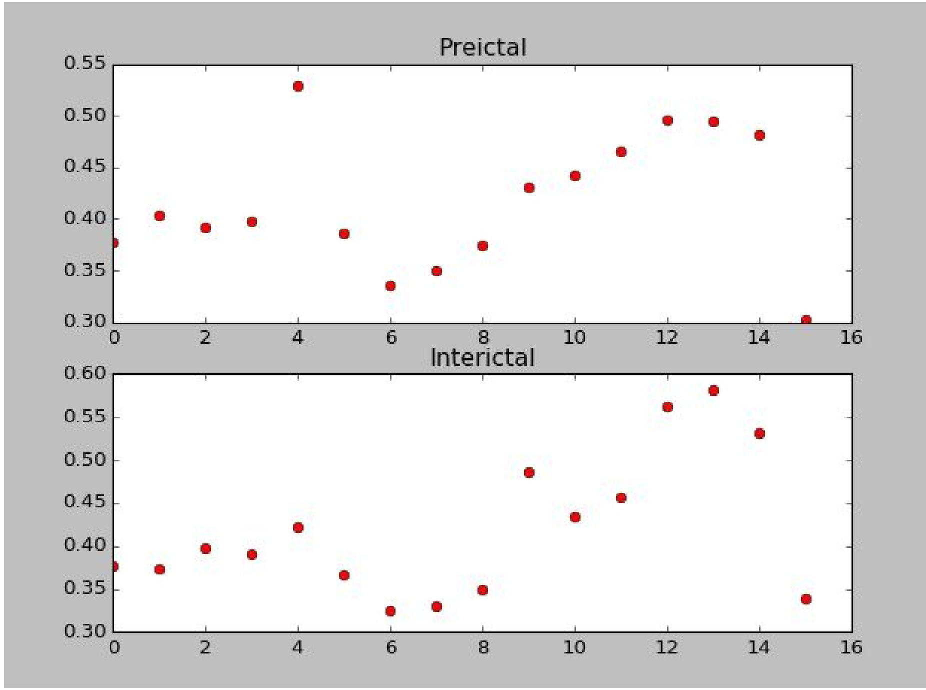
- ➢ All the EEG data clips are 10 min of duration.
- ➢ Each clip was broken further into 1 min window.
- ➢ Different frequency band (0.1, 4, 8, 12, 30, 70, 180 Hz) which include alpha to gamma band.
- ➢ FFT was used to transform the data from time to frequency domain.

Feature Extraction from Data

- Shannon’s Entropy
- Spectra Edge Frequency
- Correlation Between Channel
- Correlation Between Frequencies
- Spectral Entropy Dyadic Band
- Correlation Between Channel Dyadic Band
- Hjorth Parameters
    - Activity
    - Mobility
    - Complexity
- Statistical Measures
    - Skewness
    - Kurtosis

Shannon Entropy

- ➢ Shannon entropy is measure of uncertainty, which is nonlinear measure.
- ➢ EEG signals observed during during mental and physical activities are highly complex nonlinear and nonstationary.
- ➢ Hence Shannon entropy is best suited to measure the uncertainty associated with the different channels.

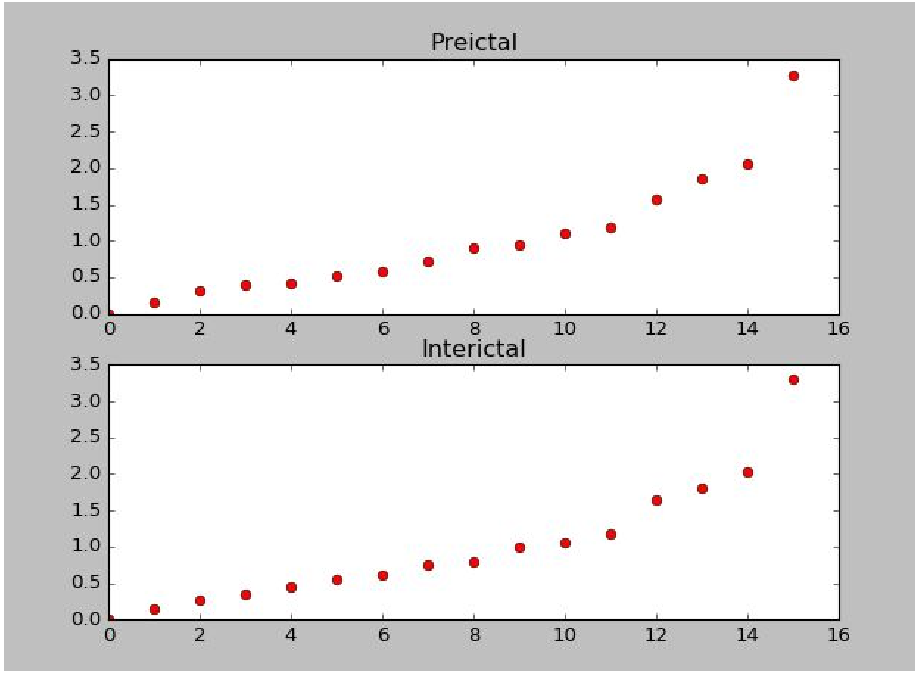


Spectral Edge Frequency

- Spectral edge frequency denoted by SEF x, is the frequency of the signal below which the x% of the total power of signal lies.
- Generally the SEF 70 to SEF 95 is taken.
- We have taken SEF 95.

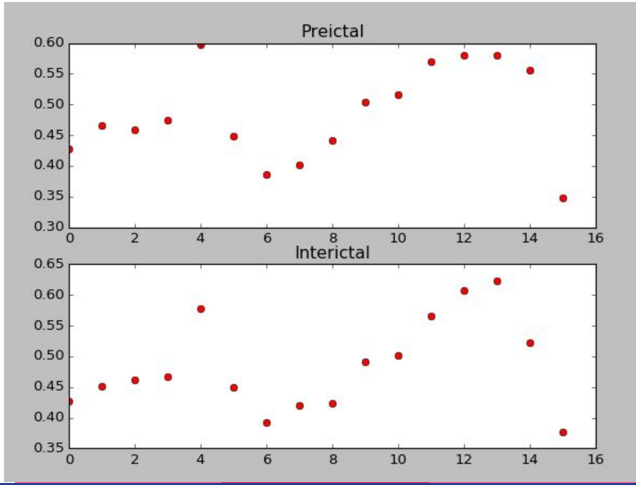
Correlation

- ➢ Correlation between 16 electrode channels were taken.
- ➢ Correlation between different frequency band were calculated.
- ➢ These two together corresponds to 120 features each.
- ➢ The eigenvalues of the correlation matrix was taken, since the eigenvalues are representative of variations, only eigenvalues were taken.



Spectral Entropy and correlation in Dyadic Band

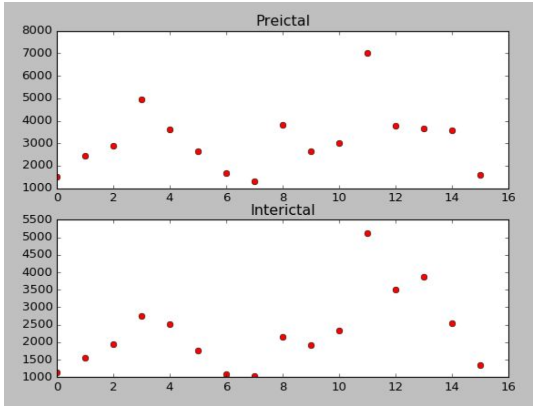
- Dyadic band decomposition signal is technique used in wavelet analysis.
- The signal is decomposed using various high and low pass filters,
- Such that the decomposed signal band is orthogonal to each other.
- These orthogonal frequency band is taken and fourier transform is applied in these band.
- Then the spectral entropy and correlation is taken.



Hjorth Parameters

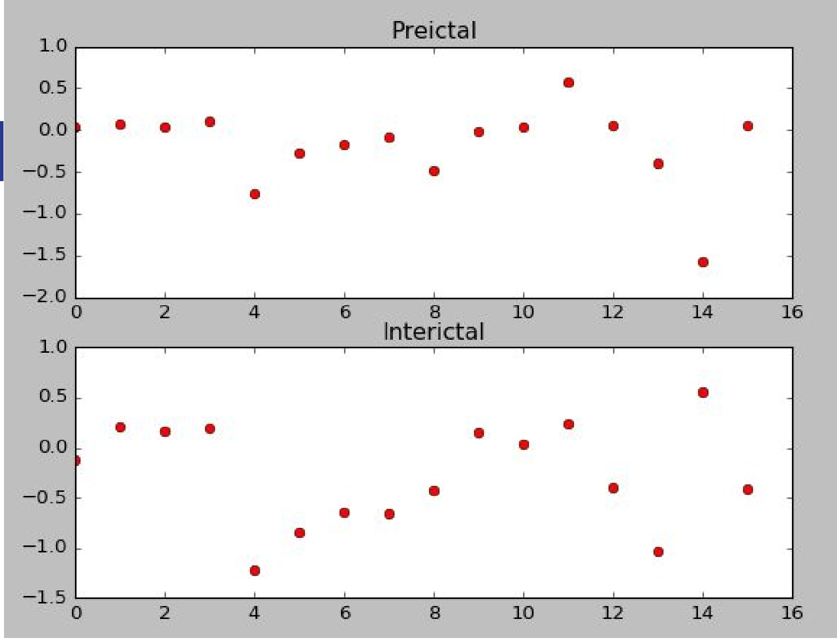
Hjorth Parameter are indicators of statistical properties of signals in time domain.

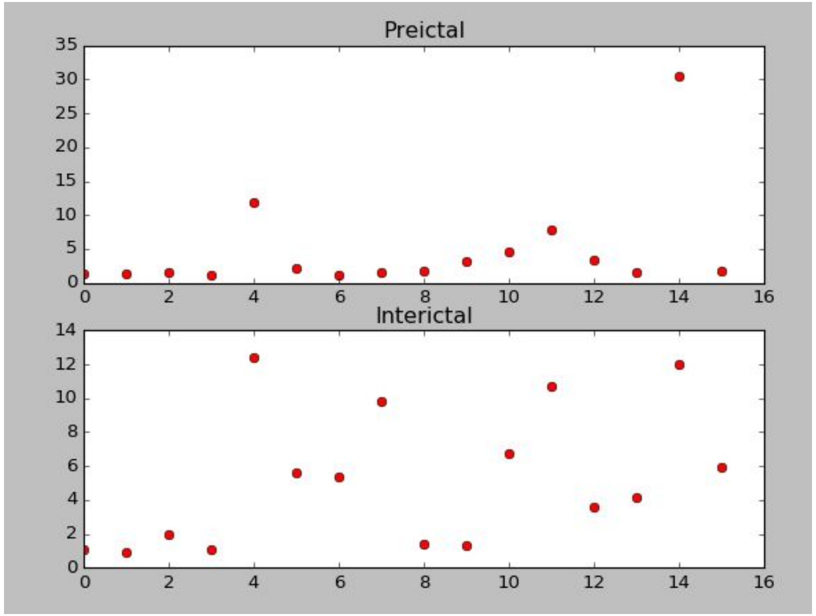
- Activity
  - Represents the signal power, the variance of a time function
- Mobility
  - Represents the mean frequency of the power spectrum.
- Complexity
  - Represents the change in frequency.
  - The parameter compares the signal's similarity to a pure sine wave



Statistical Measure of signal

- Skewness of the time series data.
- Kurtosis of the time series data.





Classification

We had total of 1610 features, and we used PCA to extract further only 600 features.

We used different classifiers such as

- Naive Bayes
- Support Vector Machine
- Random Forest
- Multilayer Perceptron
- Logistic Regression

Result and Observation

- Most of the complex algorithms such as Random Forest, MLP Gradient Boost algorithm failed to classify data correctly.
- We Observed that all of them were classifying the data to same class.
- Also the simplest algorithms such as Naive Bayes and Logistic Regresion performed better.

Result from the Random Forest

```
[[1356   46]
 [ 104    5]]
printing accuracy level
0.900727994705
             precision    recall  f1-score   support

        0.0       0.93      0.97      0.95      1402
        1.0       0.10      0.05      0.06       109

avg / total       0.87      0.90      0.88      1511
```

Result of Logistic Regression

```
[[804 598]
 [ 36  73]]
             precision    recall  f1-score   support

        0.0       0.96      0.57      0.72      1402
        1.0       0.11      0.67      0.19       109

avg / total       0.90      0.58      0.68      1511
```

Result of Naive Bayes

```
[[904 498]
 [ 48  61]]
             precision    recall  f1-score   support

        0.0       0.95      0.64      0.77      1402
        1.0       0.11      0.56      0.18       109

avg / total       0.89      0.64      0.73      1511
```

Thankyou!