# *Seizure Prediction using Online Learning and Anomaly Detection*


**Haidar Khan[1], Nimit Dhulekar[1], Lara Marcuse[2], Bülent Yener[1]**

**1. Computer Science Department, Rensselaer Polytechnic Institute, Troy, NY**

**2. The Mount Sinai Epilepsy Center, Mount Sinai Hospital, New York, NY**


## 1. Problem Overview

We propose a seizure prediction model that combines anomaly detection with online learning. We use anomaly detection to predict significant deviation from normal brain activity and use online learning to reduce the frequency of false-positive predictions.

## 2. Background

### a. Machine Learning

- Machine "learns" to distinguish pre-seizure and non-seizure from examples
- Seizure examples in EEG (electro-encephalogram) are rare.
- Seizure patterns differ between patients.
- Three general types of EEG activity for seizure prediction:
  - Ictal – activity during seizure
  - Inter-ictal – between seizure activity, "normal"
  - Pre-ictal – EEG activity before seizure

### b. Synchronization Graphs

- Learning is performed on features derived from the EEG signal
- Graphs are built based on synchronization between EEG electrodes
- To measure synchrony, Phase Locking Value (PLV) is a popular choice
- Features are extracted from synchronization graphs using graph mining techniques

## 3. Experimental Setup

### a. EEG Data

- 4 patients, >3 recordings per patient
- Extracranial EEG cap with 22 channels (10-20 system)
- ~1 hour long recordings, sampling frequency of 256 Hz
- Focal seizures that originate from different brain regions
- Recordings contain inter-ictal, pre-ictal, and ictal periods

### b. System Diagram

- System is tailored to individual patients. This is based on the belief that pre-ictal EEG signatures are patient specific
- This limits the amount of training data to recordings from the same patient.
- Due to this shortage, the best strategy is leave-one out cross-validation.
- Training and Validation Set: 1.5 hours of EEG activity
- Testing Set: 1-2 hours of EEG activity

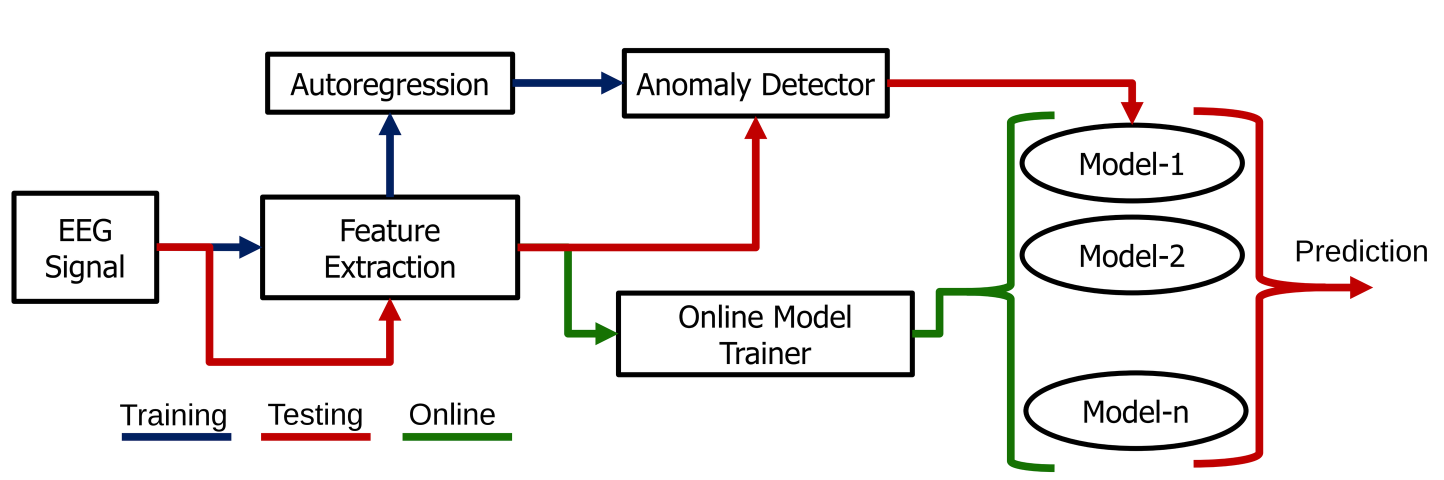


Figure 1. System overview

## 4. Methods : Synchronization Graph Features

- For each epoch (10s) of our filtered EEG signal, we create a synchronization graph $G = (V, E)$
- Vertices correspond to each electrode on EEG recording device
- Edges are placed in the graph based on phase locking between electrodes

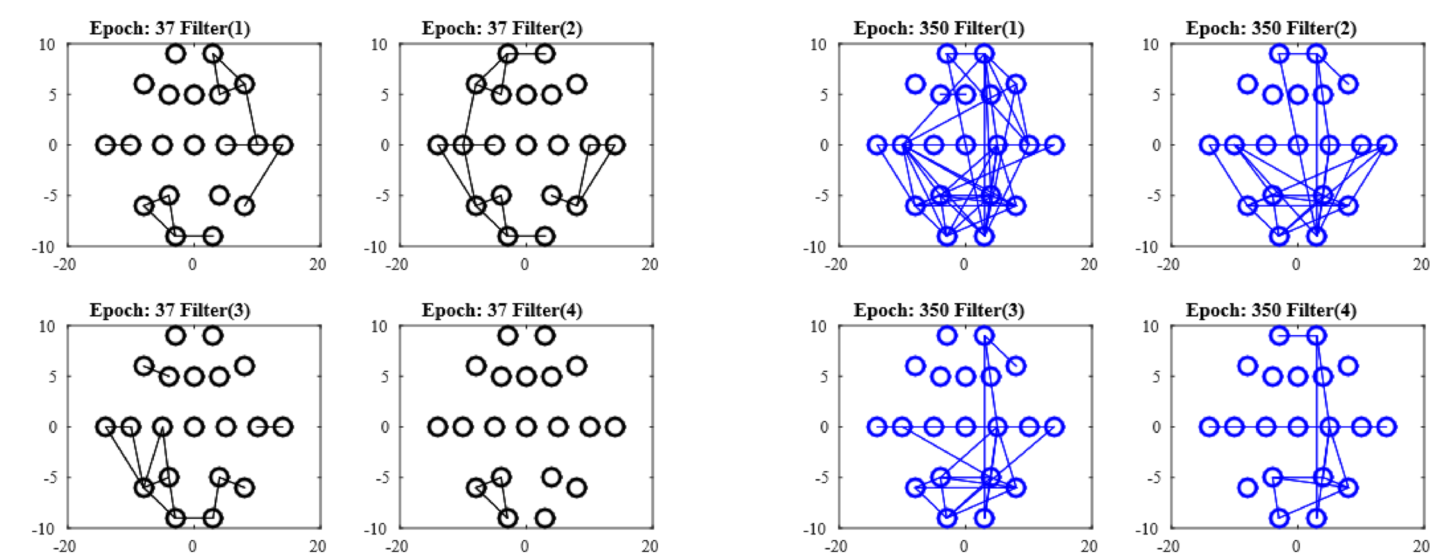


Figure 2. Synchronization graphs during inter-ictal period (left) and pre-ictal period (right). A separate graph is constructed for each frequency band (delta, theta, alpha, beta)

Table 1. Selected features extracted from each synchronization graph

| Feature | Definition |
|---|---|
| Average Degree | Average number of edges incident to a node |
| Clustering Coefficient C | Ratio between number of edges and number of possible edges |
| Average Eccentricity | Average maximum distance from a node to any other |
| Diameter | Maximum eccentricity |
| Spectral Radius | Largest eigenvalue of the adjacency matrix |
| Spectral Gap | Difference between largest and smallest eigenvalue of the adjacency matrix |
| Laplacian Trace | Sum of eigenvalues of the Lapacian matrix |

## 5. Methods : Anomaly Detection

- Model baseline EEG signal by training an autoregressive (AR) model on non-seizure (inter-ictal) data

$$X_t = c + \sum_{\{i=1\}}^{p} \varphi_i X_{\{t-i\}} + \epsilon_t$$

- Non-seizure data is plentiful, but represents many hidden states.
- Given values of the signal, a trained AR model can predict future signal values
- We calculate the error between our predicted values and the actual values
- To declare a segment of the signal an anomaly:
  - determine a threshold on the error using validation data (hyper-parameter)
  - estimate probability of being outlier

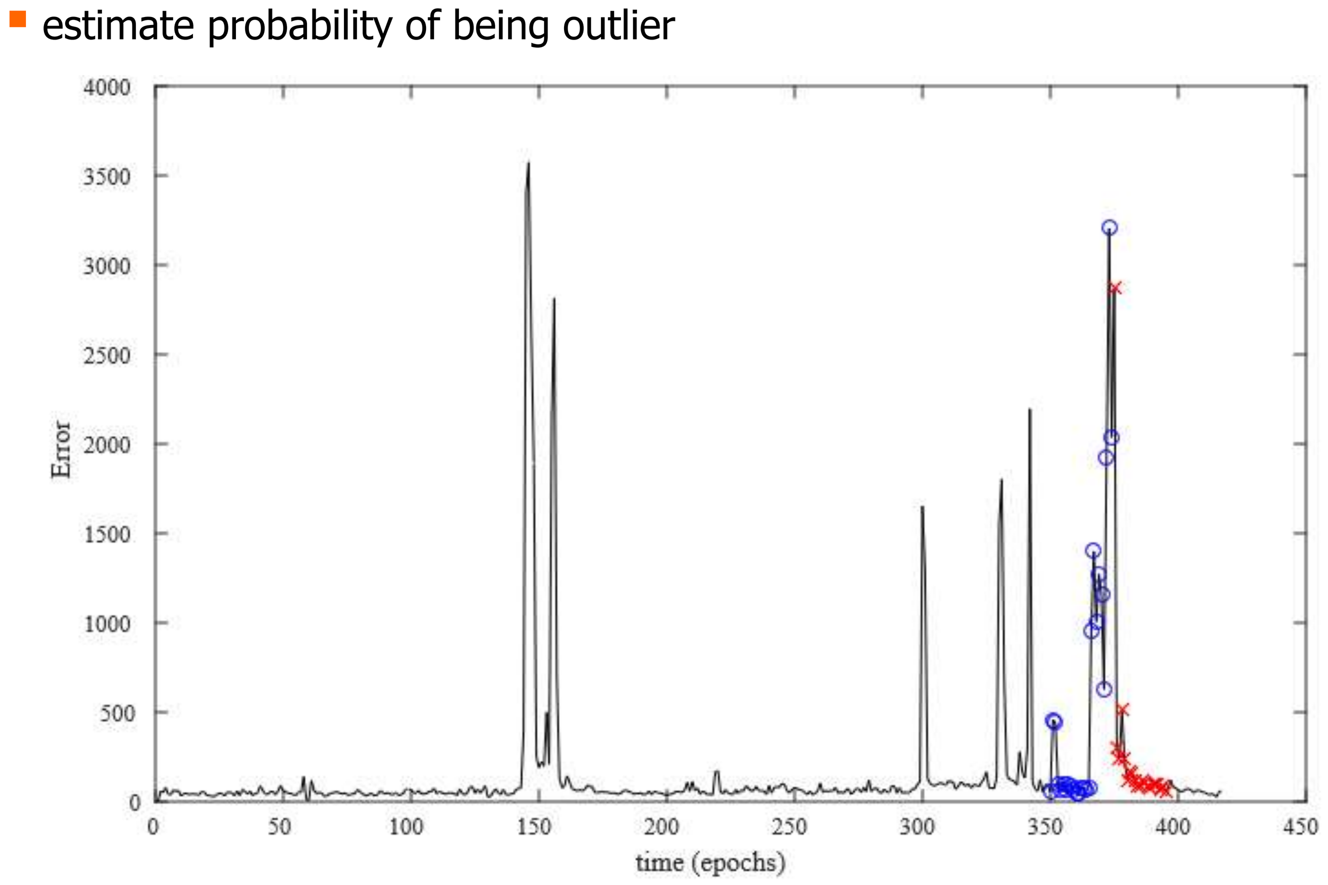


Figure 3. Prediction error over epochs.
Blue and red regions indicate pre-ictal and ictal periods respectively.

## 6. Methods : Online Learning

### a. False Positive Problem

- Brain is highly dynamical while AR is a simple linear model
- Results in a large number of anomalies detected from changing brain states
  - Sleep and waking effects captured as anomalies
- Cannot deal with this problem using only AR

### b. One-class Support Vector Machines

- Input: examples of a single class
- Output: hyper dimensional surface enclosing examples
- Distinguish between true pre-seizure anomalies and false positive anomalies

### c. Online Model Building

- Train a new one-class SVM for each anomaly detected based on past mistakes/successes
- Classify future anomalies using the whole set of models
- Ensemble learning, some resemblance to random forests

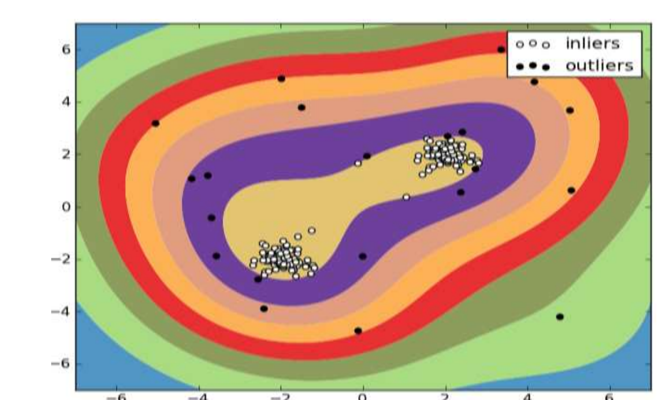


Figure 4. Example of a one-class SVM.
(image obtained from scikit-learn documentation)

## 7. Results & Future Work

- Demonstrated success in predicting seizures in some patients
- Reduced false positive rate with online model building

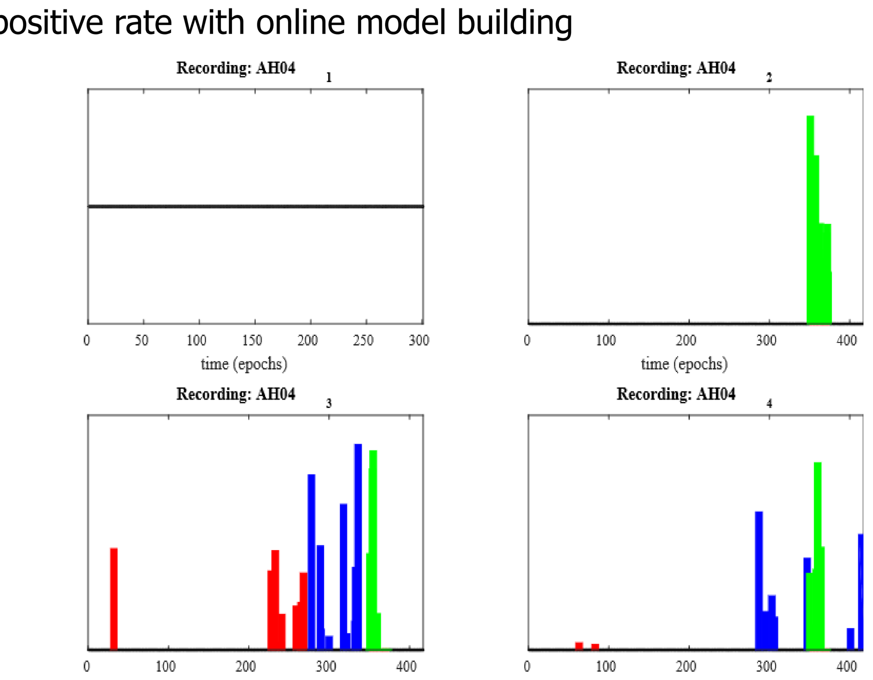


Figure 4. Prediction system run on patient AH04. Bars indicate detection of an anomaly.
(green - ictal region, blue - preictal, red - interictal) Height shows confidence level.

- Validate methods on long term EEG signals (days)
- Combine similar models created during online model building to reduce complexity
- Move from simple linear predictor (AR) to more complex Recurrent Neural Network predictor trained on long term non-seizure data

## 8. References


- Dhulekar, Nimit, Srinivas Nambirajan, Basak Oztan, and Bülent Yener. "Seizure Prediction by Graph Mining, Transfer Learning, and Transformation Learning." In *Machine Learning and Data Mining in Pattern Recognition*, pp. 32-52. Springer International Publishing, 2015.Reduced false positive rate with online model building
- Schölkopf, Bernhard, John C. Platt, John Shawe-Taylor, Alex J. Smola, and Robert C. Williamson. "Estimating the support of a high-dimensional distribution." *Neural computation* 13, no. 7 (2001): 1443-1471.
- Lachaux, Jean-Philippe, Eugenio Rodriguez, Jacques Martinerie, and Francisco J. Varela. "Measuring phase synchrony in brain signals." *Human brain mapping*8, no. 4 (1999): 194-208.
- Akaike, Hirotugu. "Fitting autoregressive models for prediction." *Annals of the institute of Statistical Mathematics* 21, no. 1 (1969): 243-247.