Predict seizures in intracranial EEG recordings

Arnab Ghosh(arnabgho@iitk.ac.in)
Amitabha Mukerjee(amit@cse.iitk.ac.in)
Computer Science And Engineering
IIT KANPUR

November 21,2014

Abstract

Seizure forecasting systems hold promise for improving the quality of life for patients with epilepsy.Despite the fact that epilepsy is infrequent , the epileptic patients suffer from an constant paranoia of a seizure.

Seizure forecasting systems have the potential to help patients with epilepsy lead more normal lives. In order for EEG-based seizure forecasting systems to work effectively, computational algorithms must reliably identify periods of increased probability of seizure occurrence. If these seizure-permissive brain states can be identified, devices designed to warn patients of impeding seizures would be possible. Patients could avoid potentially dangerous activities like driving or swimming, and medications could be administered only when needed to prevent impending seizures, reducing overall side effects.

Research has shown that the process of seizure can be divided into 4 phases :

- Interictal:Baseline in-between seizures
- Preictal : Before Seizure
- Ictal:Seizure
- Postictal:After Seizure

The primary challenge in detecting seizures is to discriminate between preictal and interictal states and correctly classify preictal states before hand with some confidence.

The goal is to accurately detect these preictal stages from the EEG data of patients and dogs of naturally occurring epilepsy .

1 Introduction

Epilepsy is a brain disorder in which clusters of nerve cells, or neurons, in the brain sometimes signal abnormally for a brief time, causing strange sensations, emotions, and behavior, or sometimes convulsions, muscle spasms, and loss of

Figure 1: How an EEG Data looks like

consciousness(known as epileptic seizures).

The outward effect can vary from uncontrolled jerking movement (tonic-clonic seizure) to as subtle as a momentary loss of awareness (absence seizure). The disease of the brain characterized by an enduring predisposition to generate epileptic seizures is called epilepsy.

5–10% of people who live to 80 years old have at least one epileptic seizure and the chance of experiencing a second seizure is between 40% and 50%.About 50% of patients with an unprovoked apparent "first seizure" have had other minor seizures, so their diagnosis is epilepsy. Epilepsy affects about 1% of the population currently and affects about 4% of the population at some point in time. Most of those affected—nearly 80%—live in developing countries.

Epilepsy may develop because of an abnormality in brain wiring, an imbalance of neurotransmitters, or some combination of these factors.Some people with epilepsy have an abnormally high level of excitatory neurotransmitters that increase neuronal activity, while others have an abnormally low level of inhibitory neurotransmitters that decrease neuronal activity in the brain.

Seizures result in direct economic costs of about one billion dollars in the United States. Epilepsy results in economic costs in Europe of around 15.5 billion Euros in 2004. In India epilepsy is estimated to result in costs of 1.7 billion USD or 0.5% of the GDP. They make up about 1% of emergency department visits (2% for emergency departments for children) in the United States.

Many areas of the world require there to be a minimum of six months from the last seizure before people can return to driving.Scientific work into the prediction of epileptic seizures began in the 1970s. Several techniques and methods have been proposed, but evidence regarding their usefulness is still lacking.

Refer Figure 1 for an example EEG data.

2 Mechanism

Normally brain electrical activity is non synchronous. In epileptic seizures, due to problems within the brain, a group of neurons begin firing in an abnormal, excessive, and synchronized manner. This results in a wave of depolarization known as a paroxysmal depolarizing shift.

Normally after an excitatory neuron fires it becomes more resistant to firing for a period of time. This is due in part from the effect of inhibitory neurons, electrical changes within the excitatory neuron, and the negative effects of adenosine. In epilepsy the resistance of excitatory neurons to fire during this period is decreased. This may occur due to changes in ion channels or inhibitory neurons not functioning properly. This then results in a specific area from which seizures may develop, known as a "seizure focus". Another cause of epilepsy may be the up regulation of excitatory circuits or down regulation of inhibitory circuits follow an injury to the brain. These secondary epilepsies, occur through processes known as epileptogenesis. Failure of the blood–brain barrier may also be a causal mechanism.

Focal seizures begin in one hemisphere of the brain while generalized seizures begin in both hemispheres. Some types of seizures may change brain structure, while others appear to have little effect. Gliosis, neuronal loss, and atrophy of specific areas of the brain are linked to epilepsy but it is unclear if epilepsy causes these changes or if these changes result in epilepsy.

3 Dataset

The dataset that was used was the Kaggle Dataset , obtained from the American Epileptic Society .

Preictal training and testing data segments were provided covering one hour prior to seizure with a five minute seizure horizon. (i.e. from 1:05 to 0:05 before seizure onset.) This pre-seizure horizon ensures that 1) seizures could be predicted with enough warning to allow administration of fast-acting medications.

See the Figure 2 for reference .

Similarly, one hour sequences of interictal ten minute data segments were provided. The interictal data were chosen randomly from the full data record, with the restriction that interictal segments be as far from any seizure as can be practically achieved, to avoid contamination with preictal or postictal signals. In the long duration canine recordings it was possible to maintain a restriction of one week before or after a seizure. However, in the human recordings (which may be less than a week in total duration) interictal data was restricted to be more than four hours before or after any seizure.

Additional annotated intracranial EEG data is freely available at the International Epilepsy Electrophysiology Portal, jointly developed by the University of Pennsylvania and the Mayo Clinic.

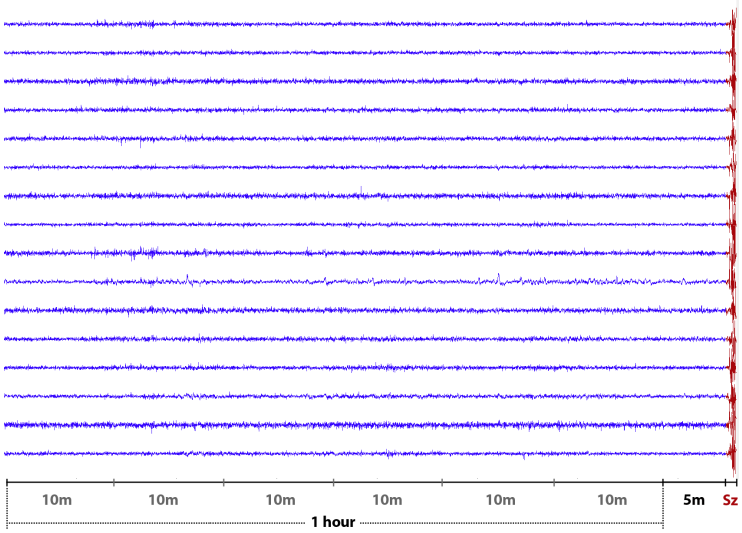

Figure 2: How the Dataset looks like

4 Observations

4.1 Observation 1

Some of the channels , particularly those near the Hippocampus face excessive amount of brain activity.
Generally it is seen that during a seizure, the signals of the brain work in a constructive interference pattern and reaches very high levels , ultimately culminating in a seizure. We used this observation to model the SVM and to tune the parameters for better prediction. Refer 3 for how experts in epilepsy perceive such data.

4.2 Observation 2

Not all channels are equally excited in the case of a seizure.

This is because a seizure generally doesn't impact the whole of the brain but rather concentrated around the Hippocampus , hence we cannot treat the data for all the channels as equivalent . Hence we find out the correlation among the various channels and use it as the Predictor Matrix for the Training of the Decision Trees of the Forest . Refer 4 for more details.

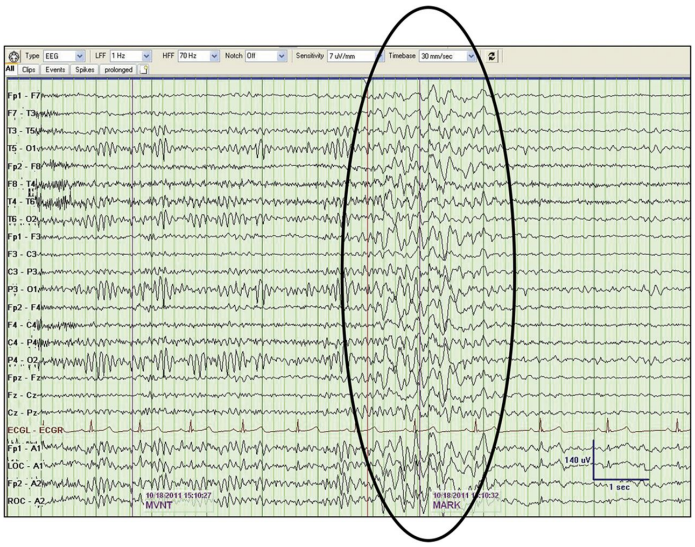

Figure 3: Experts perceive epileptic behavior from EEG data

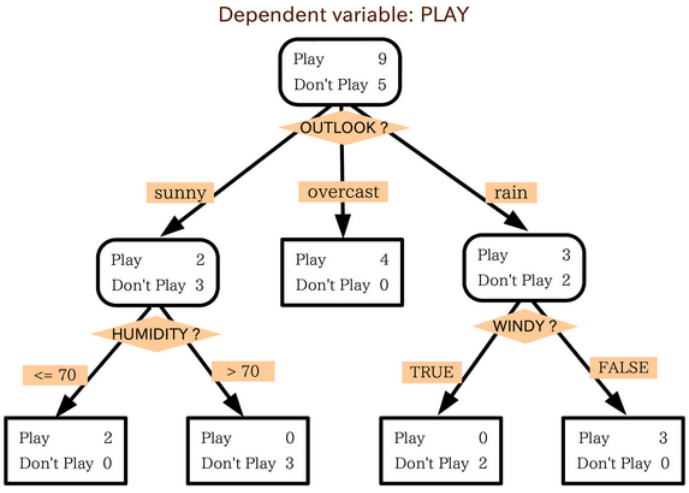

Figure 4: A toy Decision Tree Example.

5 Methods & Algorithms

5.1 Using Random Forests And Decision trees

Data: The EEG data
Result: Predict 1(Seizure) or 0(Otherwise)
for *all channels* **do**
  Compute the variance and the correlation between the channels
end
use the Correlation Coefficients to build the Predictor Matrix
while *learning from the training data* **do**
  for *each subject* **do**
    Learn a set of decision trees (Forest) using the Predictor Matrix.
  end
end
while *A new EEG sample is given* **do**
  for *all the decision trees in the subjects Forest* **do**
    Prediction of all trees on this sample
  end
  Output the majority of all Predictions
end

Algorithm 1: Using Random Forests

5.2 Using Support Vector Machines

Data: The EEG data
Result: Predict 1(Seizure) or 0(Otherwise)
while *learning from the training data* **do**
  for *each subject* **do**
    for *each data sample X(16*239766)* **do**
      Obtain a Random Projection Matrix RP(239766*2400);
      Compute the inner product of RP*X
      Pred= SVM with Gaussian Kernel to train the classifier.
    end
  end
end
while *A new EEG sample is given* **do**
  use RP to preject the data into the low dimensional space
  use Pred matrix and data to compute the Prediction
end

Algorithm 2: Using Support Vector Machines

Note: We also tried each method alongside a gaussian Kernel and once without a gaussian kernel.

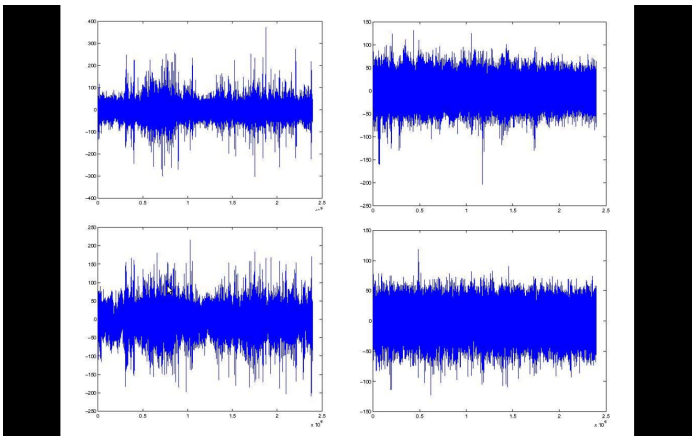

Figure 5: Stark Contrast in some channels between Preictal and Interictal Segments

6 Results

Algorithm	F1 Score
Random Forest	61%
Random Forest (Gaussian Kernel)	65%
SVM (no kernel)	70%
SVM (Gaussian kernel)	72%

6.1 Explanation Of the Results

Random Forests The decision tree model using the bootstrapping technique works quite good in practice but here because of the high dimensional embedding in the data takes the F1 score only till 61 . We tried with different number of trees in the forest but 1000 seemed to work the best considering the speed and accuracy .

Random Forests With Gaussian Kernels

With the introduction of the gaussian Kernel method with ($\sigma : 0.01$ worked the best) introduced non linearity in the model and helped capture high dimensional embedding of the data sample . Thus it produced better results than just the Vanilla Random Forests method.

SVM With No Kernel

Support Vector machines work extremely well in practice for classification problems but the high dimensional data forced us to reduce the dimension to tractable sizes by using a random projection matrix . We tried reducing the data (16*239766) to one-tenth of its proportions ie(16*24000) and one-hundredth of its proportions(16*2400) but these gave roughly similar results(70% & 70.2%) hence the faster running one was reported .

SVM With Gaussian Kernel

Support Vector Machines with Gaussian Kernel seemed to work very well with an overall F1 score of 72% again hinting on the fact that the data was

instrinsically high dimensional .

We tried reducing the data (16*239766) to one-tenth of its proportions ie(16*24000) and one-hundredth of its proportions(16*2400) but these again gave roughly similar results(72% & 72.1%) hence the faster running one was reported .

Also here again the gaussian kernel with ($\sigma = 0.01$) seemed to work the best .

7 Achievements

1. We looked at this problem in a computational aspect using algorithms that had been previously unapplied .

2. This problem had been looked at from the medicinal aspect , and from the point of view of medical science and very few instance by computer scientists.

3. With the enhancement of Technology , a new contraption has been devised by the American Epileptic Society which measures the EEG data real time using a strap , which is kind of a mobile device which gathers the EEG data and can be sent to your mobile phone . It is still in its nascent stages and has been tested on canines . But we hope that if this device comes to use in the near future then using our study an app could be made which fathoms the severity of the preictal stage and give one early warning about an epileptic seizure.

8 Conclusion & Further Research

The methods we tried to apply proved good enough results but it is still far from state of the art. We have to look at slightly different mechanisms trying to exploit the time series component of the EEG data.

If experts well versed in this field can be consulted and understand their perception of the data detection of anomalies using their instinct , we could try some more fancy neural networks and deep learning techniques to solve this problem.

9 References

(1)Howbert JJ, Patterson EE, Stead SM, Brinkmann B, Vasoli V, Crepeau D, Vite CH, Sturges B, Ruedebusch V, Mavoori J, Leyde K, Sheffield WD, Litt B, Worrell GA (2014) Forecasting seizures in dogs with naturally occurring epilepsy. PLoS One 9(1):e81920.

(2)Cook MJ, O'Brien TJ, Berkovic SF, Murphy M, Morokoff A, Fabinyi G, D'Souza W, Yerra R, Archer J, Litewka L, Hosking S, Lightfoot P, Ruedebusch V, Sheffield WD, Snyder D, Leyde K, Himes D (2013) Prediction of seizure likelihood with a long-term, implanted seizure advisory system

in patients with drug-resistant epilepsy: a first-in-man study. LANCET NEUROL 12:563-571.

(3)Park Y, Luo L, Parhi KK, Netoff T (2011) Seizure prediction with spectral power of EEG using cost-sensitive support vector machines. Epilepsia 52:1761-1770.

(4)Park Y, Luo L, Parhi KK, Netoff T (2011) Seizure prediction with spectral power of EEG using cost-sensitive support vector machines. Epilepsia 52:1761-1770.

(5)Davis KA, Sturges BK, Vite CH, Ruedebusch V, Worrell G, Gardner AB, Leyde K, Sheffield WD, Litt B (2011) A novel implanted device to wirelessly record and analyze continuous intracranial canine EEG. Epilepsy Res 96:116-122.

(6)Andrzejak RG, Chicharro D, Elger CE, Mormann F (2009) Seizure prediction: Any better than chance? Clin Neurophysiol.

(7)Snyder DE, Echauz J, Grimes DB, Litt B (2008) The statistics of a practical seizure warning system. J Neural Eng 5: 392–401.

(8)Mormann F, Andrzejak RG, Elger CE, Lehnertz K (2007) Seizure prediction: the long and winding road. Brain 130: 314–333.

(9)Haut S, Shinnar S, Moshe SL, O'Dell C, Legatt AD. (1999) The association between seizure clustering and status epilepticus in patients with intractable complex partial seizures. Epilepsia 40:1832–1834.