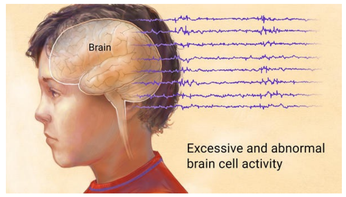
The 5 W's of Epilepsy

Xu Si, Hung-Yi Wu and James Zatsiorsky

What is epilepsy?

Chronic neurological disease

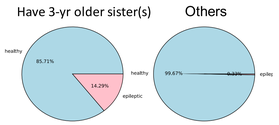


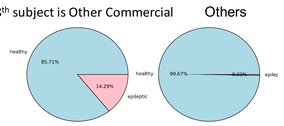
" According to the latest estimates, about 1% of children aged 0-17 years have had a diagnosis of epilepsy or seizure disorder, about 1.8% of adults aged 18 years or older have had a diagnosis of epilepsy or seizure disorder "

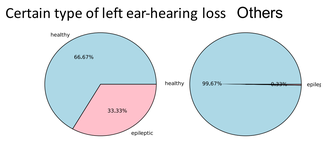
Who has epilepsy?

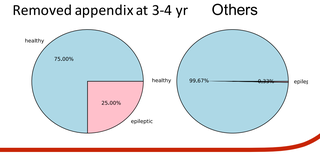
Genetically inherited or acquired

From the NCDS (National Child Development Study), we look at those who do not report having epilepsy at 7 and develop epilepsy at 16. Using logistic regression model, we found the following important factors.









When will they have epileptic seizures?

Seizure forecasting

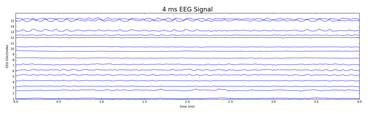
1. Background Knowledge

Preictal prior to seizure

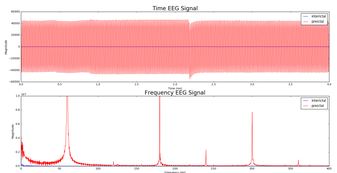
Post-ictal after seizures

Interictal between seizures

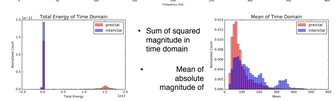
Ictal seizure



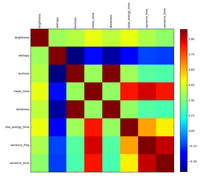
Attempt Two: Extract Features



Preictal EEG signal show more energy at both time and frequency domain. Based on this finding, we extract 8 features below.



- Sum of squared magnitude in time domain
- Mean of absolute magnitude of
- Variance of absolute magnitude of
- Variance of spectrum magnitude, upper limit = 200Hz domain
- Fourth standardized moment domain
- Third standardized moment domain
- Percentage of energy above
- 0Hz, 0-10Hz, 10-50Hz, 50-100Hz, 100-200Hz, 200-300Hz, 300-400Hz.

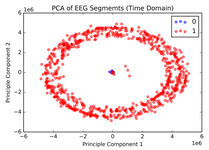


Correlation matrix of 8 features, most of them are somehow correlated.

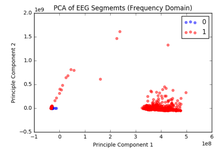
2. Classification Problem

We try to distinguish between preictal and intrictal EEG signals, this will help to forecast seizure in epilepsy patients. It will be really helpful to prevent severe consequences, e.g. danger while swimming.

Attempt One: PCA



Class	Accuracy
Preictal	0.3
Interictal	1

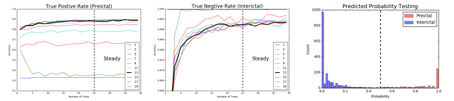


Class	Accuracy
Preictal	0.3
Interictal	1

Apply PCA on both time and frequency domain. Used top principle components (plot first two) and unable to separate them well. Move on to extracting features.

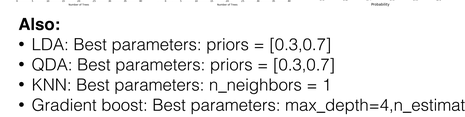
Classification Models on Extracted Features

- Logistic Regression



Tune two parameters: class weight and regularization parameter C.

- Random Forest Classifier

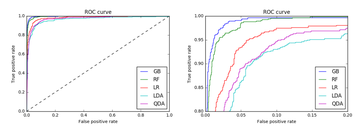


Tune two parameters: n_estimator (number of trees) and maximum depth.

Also:

- LDA: Best parameters: priors = [0.3,0.7]
- QDA: Best parameters: priors = [0.3,0.7]
- KNN: Best parameters: n_neighbors = 1
- Gradient boost: Best parameters: max_depth=4,n_estimators=300

ROC and Accuracy



We successively run logistic regression, random forest, LDA, QDA, KNN, and boosting and tuned the best parameters on the training set.

The best model is **boosting**. The ROC curve resides in tthe upper left corner, achieving both highest true positive rate and lowest false positive rate.

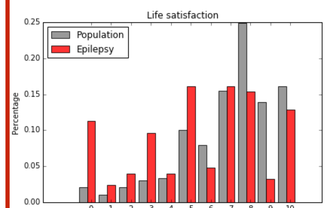
Accuracy	KNN	Logistic Regression	LDA	QDA	Random Forest	Boosting
Preictal	0.88	0.82	0.79	0.96	0.97	0.97
Interictal	0.94	0.98	0.97	0.81	0.96	0.98
Total	0.93	0.94	0.92	0.85	0.97	0.98

We have achieved, at highest, total classification accuracy of **0.98**, with 0.97 accuracy on preictal data and 0.98 accuracy on interictal data.

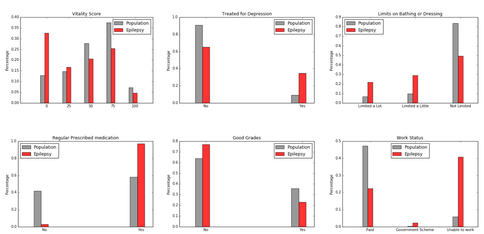
3. Further Work

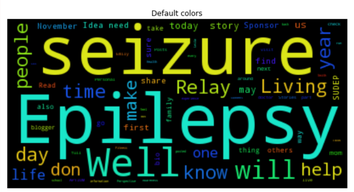
If we can have access to EEG data from many more different features, we can try to give a general seizure forecast metric. Also, we can explore on extracting more features from the EEG signal.

What kind of life do they lead?



They are generally **less satisfied** with life and may suffer from anxiety and **depression**. Some are even permanently **unable to work** and perform **poorly on school work.**



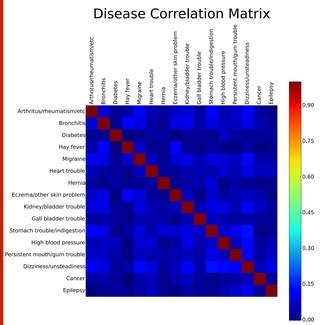


Blog. Epileptic people write about their life, positive and full of hope.

Why do they have epilepsy?

Correlation with other diseases

If epilepsy is correlated with other diseases, can we use this correlation to give a better-than-chance diagnosis of epilepsy in NCDS patients?



From the correlation matrix and random forest classification model, we can not see clear correlation of epilepsy with other disease. **Thus, there is no evidence that certain disease will cause epilepsy.**

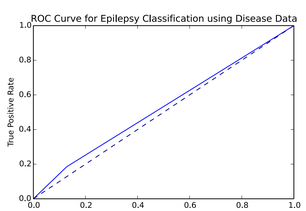


Figure: On average, a Random Forest classification model fit to the disease predictors barely outperforms a random coin flip in diagnosing epilepsy.

Conclusion: Provide better care to the **1%** of people who have epilepsy

With our understanding of how people acquire epilepsy, its correlation with other disease, the life of patients and how to help patients detect abnormality in EEG signal, we can together bring better care to epilepsy patients.