Enhanced Predictor of Human Cognitive State

By

Tanay Chowdhury

B.Tech. West Bengal University of Technology, 2005

THESIS

Submitted as partial fulfillment of the requirements
for the degree of Master of Science in Computer Sciences
in the Graduate College of the
University of Illinois at Chicago, 2014
Chicago, Illinois

Defense Committee:

Dr Brian Ziebart, Chair and Advisor,
Computer Science
Dr Bing Liu, Computer Science
Dr Bhaskar Dasgupta, Computer Science

ACKNOWLEDGMENTS

I would like to thank my thesis committee--(*Dr Bing Liu and Dr Bhaskar Dasgupta*)--for their unwavering support and assistance. They provided guidance in all areas that helped me accomplish my research goals and enjoy myself in the process. I am also grateful to Professor Brian Ziebart for helpful discussions and detailed comments on various drafts of this paper.

I am thankful to Professor Tom Mitchell and his team from Carnegie Mellon University for making the fMRI data and corresponding work publicly available.

TC

TABLE OF CONTENTS

CHAPTER PAGE

LIST OF TABLES

LIST OF FIGURES

SUMMARY

Human brain is considered the most complicated and developed one among all the mammals living in the world. Physicist Sir Roger Penrose said "If you look at the entire physical cosmos, our brains are a tiny, tiny part of it. But they're the most perfectly organized part. Compared to the complexity of a brain, a galaxy is just an inert lump." Hence naturally it is considered an active research area. With the emergence of functional magnetic resonance imaging the research to detect unique pattern of brain activation and comprehend important meaning out of that got momentum. As part of this paper I propose a modified and a new methods to improve the accuracy of predictive modelling to predict cognitive state of human brain created earlier.

Keywords : Scientific data analysis, functional Magnetic Resonance Imaging, high dimensional data, feature selection, Adaboost classifier, Support Vector Machine, nearest neighbor, brain image analysis.

1. Introduction

Functional Magnetic Resonance Imaging (fMRI) is a newer three dimensional radiation free safe brain imaging technique that takes high spatial resolution image of the brain. The large volume data collected from the entire brain using this method help detect neural activity in the brain with the help of proper data analysis technique. As part of a recent research work fMRI data was collected from different subjects after showing them interleaved pictures and sentences in single time frame. Afterwards machine learning technique was used on that fMRI to do fairly accurate prediction on the cognitive state of the target subject. This paper summarizes the lessons learned about improving the machine learning methods applied to train classifiers in such settings.

2. Problem Definition

This paper tries to resolve the problem of classifying high dimensional, noisy fMRI data with modified and new methods with better accuracy.

2.1 Functional Magnetic Resonance Imaging (fMRI)

Brain imaging has recently caught the attention of researchers as it has the enormous power to explore the inner working pattern of the human brain.it has application in numerous fields such as psychology, political science and neuroscience. It has mainly two variants 1) Structural brain imaging and 2) Functional brain imaging. Structural brain imaging mainly deals with the structural aspect of the brain such as injury and diseases. It has few variants such as computed axial tomography (CAT), magnetic resonance imaging (MRI), and positron emission tomography (PET).However functional brain imaging is mainly dynamic in nature and is good for capturing cognitive processes. Variants of this are positron emission tomography (PET), functional magnetic resonance imaging (fMRI), electroencephalography (EEG), and magnetoencephalography (MEG).Although all types of brain imaging has their own usage functional magnetic resonance imaging is most widely used out of those because it strikes a nice balance nice balance among high quality spatial resolution, temporal resolution and invasiveness. Also this technique is incredibly user friendly.

The MRI scanner uses a highly powerful electro-magnet which is typically 50000 times stronger than the Earth's magnetic field. Under this highly powerful magnetic field minuscule magnetic signals from each nuclei sums up to create a measurable signal.fMRI process detects magnetic signal from hydrogen nuclei in water. Since fluid amount and density varies between gray matter, white matter and cerebrospinal fluid hence the signal from hydrogen nuclei also varies. **(Lindquist, Martin (2014) Introduction to fMRI from https://class.coursera.org/fmri-001/lecture)**

Hemoglobin in the blood carries oxygen, and hence when certain part in the brain gets active that requires more blood. However oxygenated hemoglobin is diamagnetic but deoxygenated hemoglobin is paramagnetic. **(Mitchell et**

al 2004) This certain magnetic property causes a difference in MRI signal if the level of oxygenation is good enough. This type of MRI is named as blood oxygenation level dependent (BOLD) imaging. Also the blood flow increase to a certain part of the brain follows a certain pattern in which after initial decrease ("initial dip") the blood flow picks up after 4-6 seconds

An fMRI scanner captures the value of fMRI signal in a three dimensional grid. In the actual research a three dimensional image was captured in each 1, 1.5 or 0.5 seconds. The smallest unit within this three-dimensional image are referred as voxels. It is almost a few tens of cubic millimeters in dimension and a general three-dimensional brain image contains 10000-15000 voxels. These voxels captures fMRI BOLD response during the interval when blood flow picks up after a stimulation within 4-5 seconds and comes back to baseline another 5-7 seconds.

(Mitchell et al 2004)

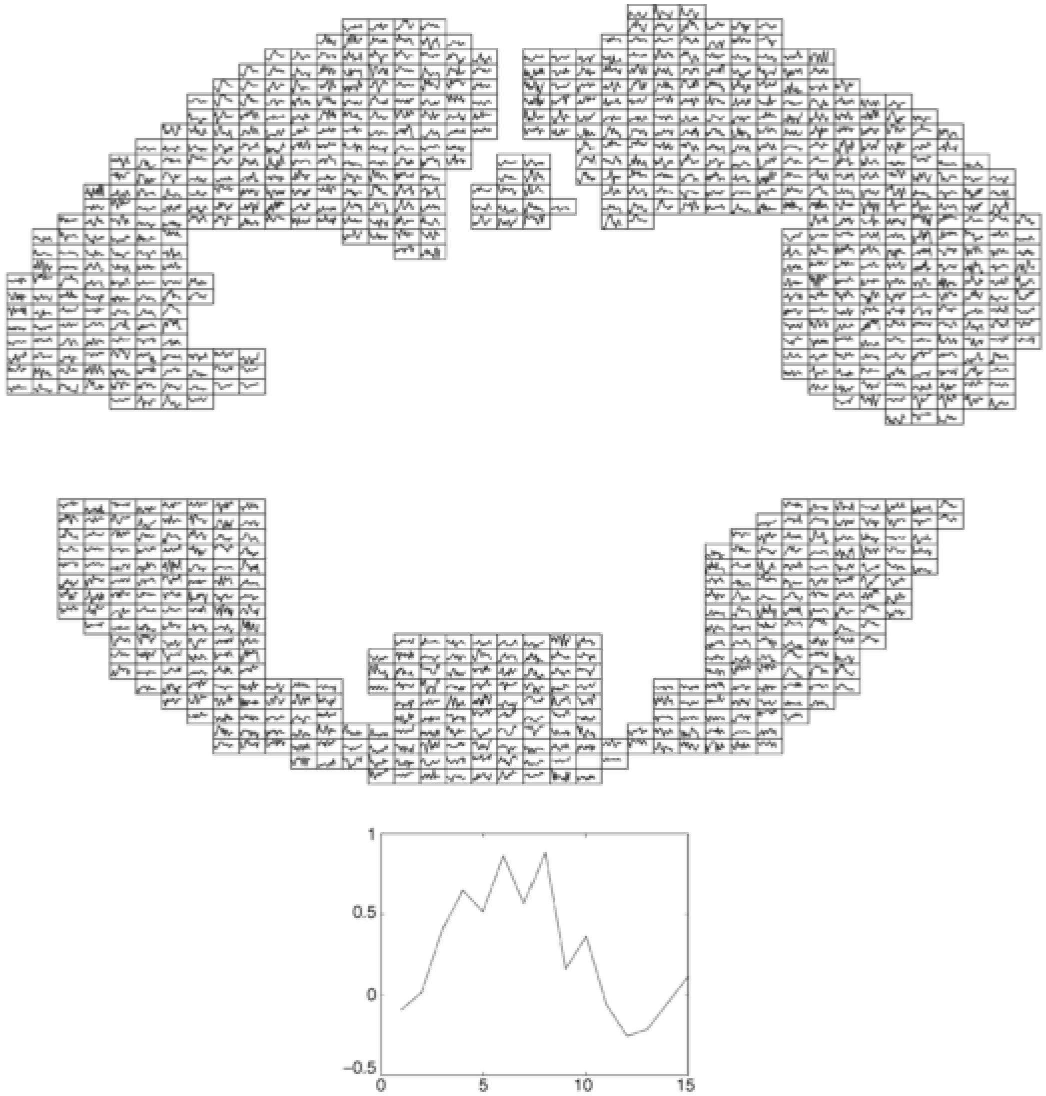

Figure 1: *"Typical fMRI data. The top portion of the figure shows fMRI data for a selected set of voxels in the cortex, from a two-dimensional image plane through the brain. A fifteen second interval of fMRI data is plotted at each voxel location. The anterior portion of the brain is at the top of the figure, posterior at bottom. The left side of the brain is shown on the right, according to standard radiological convention. The full three-dimensional brain image consists of sixteen such image planes. The bottom portion of the figure shows one of these plots in greater detail."* **(Mitchell et al 2004)**

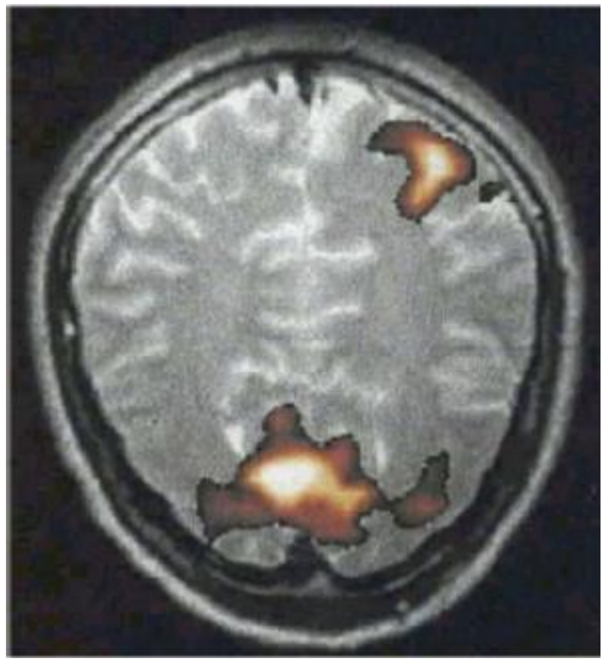

Figure 2: fMRI determines what regions in the brain are active during a certain task. (**Lindquist 2008)**

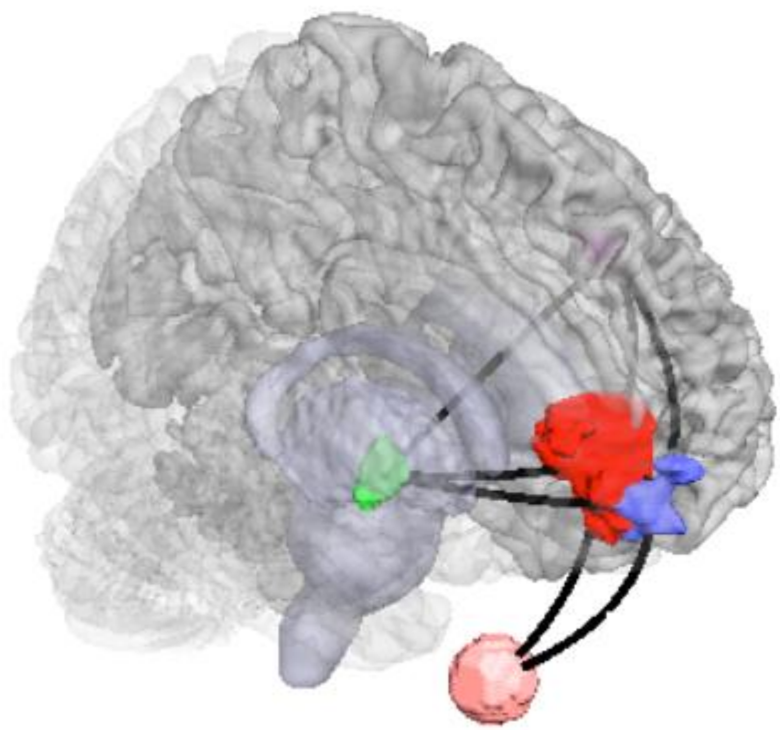

Figure 3: fMRI determines how each regions are connected to one another dynamically. (**Cribben et al. 2012)**

2.2 Problem Description

The problem hovers around a basic question. If human brain decides its course of action after processing it itself, then is it possible at least in principle to determine what a person is doing or going to do by measuring inner activity of his/her brain. The whole idea is not new but it could not be examined in the absence of proper technique to measure dynamic brain activity. However with the advent of MRI technique it became certainty in practical also. Although there are two popular methods for doing that but either way both conventional and decoding-based techniques are based on human neuroimaging.

2.2.1 Conventional approach

Conventional neuroimaging methods try to find a part of the brain that is responsible for a task to determine how a cognitive state is registered in brain. This is done by capturing activity from many locations in the brain across time and then analyzing each location for a significant change. If the difference between two mental states is sufficiently large then that place is considered. However in practice it is very difficult to find such individual locations.

2.2.2 Modern Approach

In comparison to earlier rigid conventional analysis the modern approach increases the sensitivity of neuroimaging by using the full spatial pattern of brain activity map. This multivariate simultaneous analysis at many locations has many advantages over univariate approaches that works on one location at a time. Firstly it can aggregate sparse spatial features and can use it effectively that could have got removed in the pre-processing spatial smoothing in conventional analysis. Also conventional analysis needs large amount of brain activity sample but during averaging the details gets lost. However better sensitivity of modern approach fares better in both these measures.

This paper uses the earlier work by **Tom Mitchell et al 2004** where they worked on training machine learning classifiers to automatically decode the subject's cognitive state at a single time instant or interval. The goal was to make it possible to detect transient cognitive states. While as part of this paper the existing supervised predictive classifier algorithm is enhanced so that it uses optimum number of active voxacts while using fMRI data for training or test data formation. Also a new solution has been formed which uses optimum number of active region of interest (ROI) to extract training and test data from fMRI data. This paper also exposes the fMRI data to some new classifiers and Meta classifier technique such as Adaboosting.

3. Solution Formulation

3.1 Picture versus sentence study

In this research the data from fMRI study of **(Keller, Just, & Stenger, 2001)** is used. In that study every subject faced a series of trial. During those trials subjects were shown a sentence and a picture subsequently. The data received from this experiment was used to train a classifier to identify whether a subject is seeing a sentence or a picture during a certain time interval. The pictures that were shown was of simple in nature and comprised of symbols. Also the sentences were related to those pictures. The sequence of showing the sentence and picture were also changed, by showing pictures ahead of sentences in half of the cases and vice versa in rest half of the cases. Each of the stimulus (sentence or picture) was presented for 4 seconds until the subject indicated whether the sentence correctly describes the picture. After each trial 15 seconds of rest period was given. Hence each trial continued for close to 27 seconds. Those test pictures comprised of geometrical signs such as +, $ and *.

(Mitchell et al 2004)

Sentences were descriptions about the picture such as “It is true that the plus is below the dollar”. Also half of the sentences were negative (e.g., “It is not true that the star is above the plus.”) mainly to make sure that subject was paying proper attention in trial. Rest half of the picture were affirmative sentences. Other than the rest period between each trial an additional rest period was provided after every 4 trials. Every subject faced altogether 40 trials and fMRI images were being collected in every 500 msec.The objective was to train a classifier in such a way that it can predict in a 8 seconds interval of fMRI data whether the subject is watching a sentence or a picture. They tried to learn a classifier of the following format for each subject.

f: fMRI-sequence(t_0, t_0 +8)→{Picture, Sentence} **(Mitchell et al 2004)**

t_0 is the start time of the stimulus, which is test picture or sentence. So as seen in the image input variable for the classifier is 8 seconds of fMRI data. Although the stimulus is shown to the subject for only 4 seconds but an 8 seconds interval is needed to capture full brain activity. Since as mentioned in earlier section fMRI bold signal starts picking up 4-5 seconds after the stimulus. Altogether there were 40 examples per class. As the fMRI was done every 500 msec, so for an 8 seconds interval altogether 16 images were captured. Since there were 10,000 voxels per subject that result an input feature vector of overwhelming 160,000 features.

3.2 Learning Method

This paper uses following learning method.

f: fMRI-sequence (t_1, t_2)→Cognitive State **(Mitchell et al 2004)**

Here fMRI-sequence (t_1, t_2) is the sequences of fMRI images that were collected during time duration [t_1, t2], and cognitive state is corresponding cognitive state for watching sentence or picture. In almost all the cases a classifier was learned for each subject. The input was nothing but feature vector in the form of voxel activities. Since there were altogether 160,000 voxels involved proper dimension reduction and feature selection technique was required for an accurate classification. Some of the methods tried to replace multiple features with their mean.

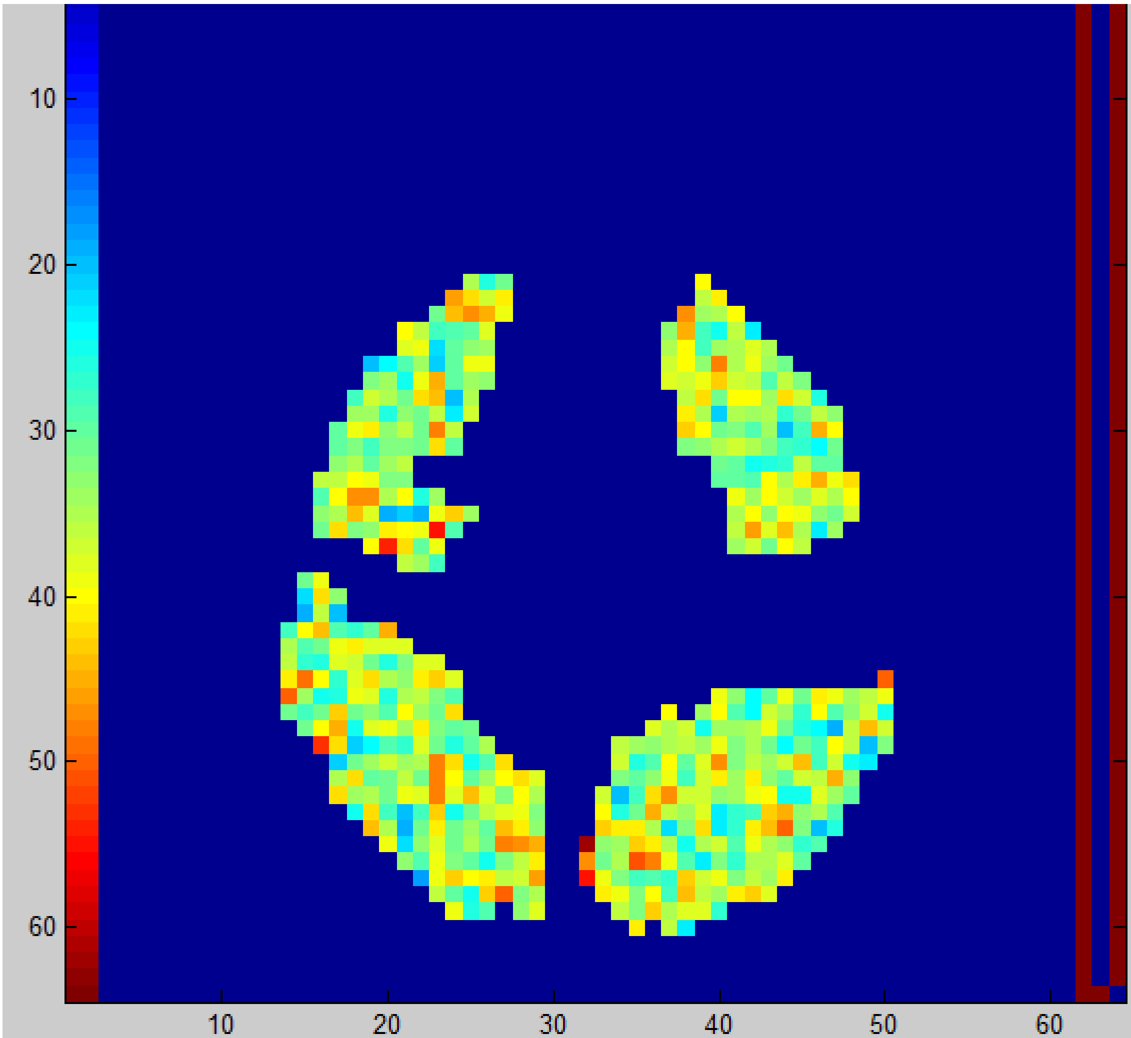

Figure 4: A picture of activation for the z=4 slice of the brain for trial 4, at time 8

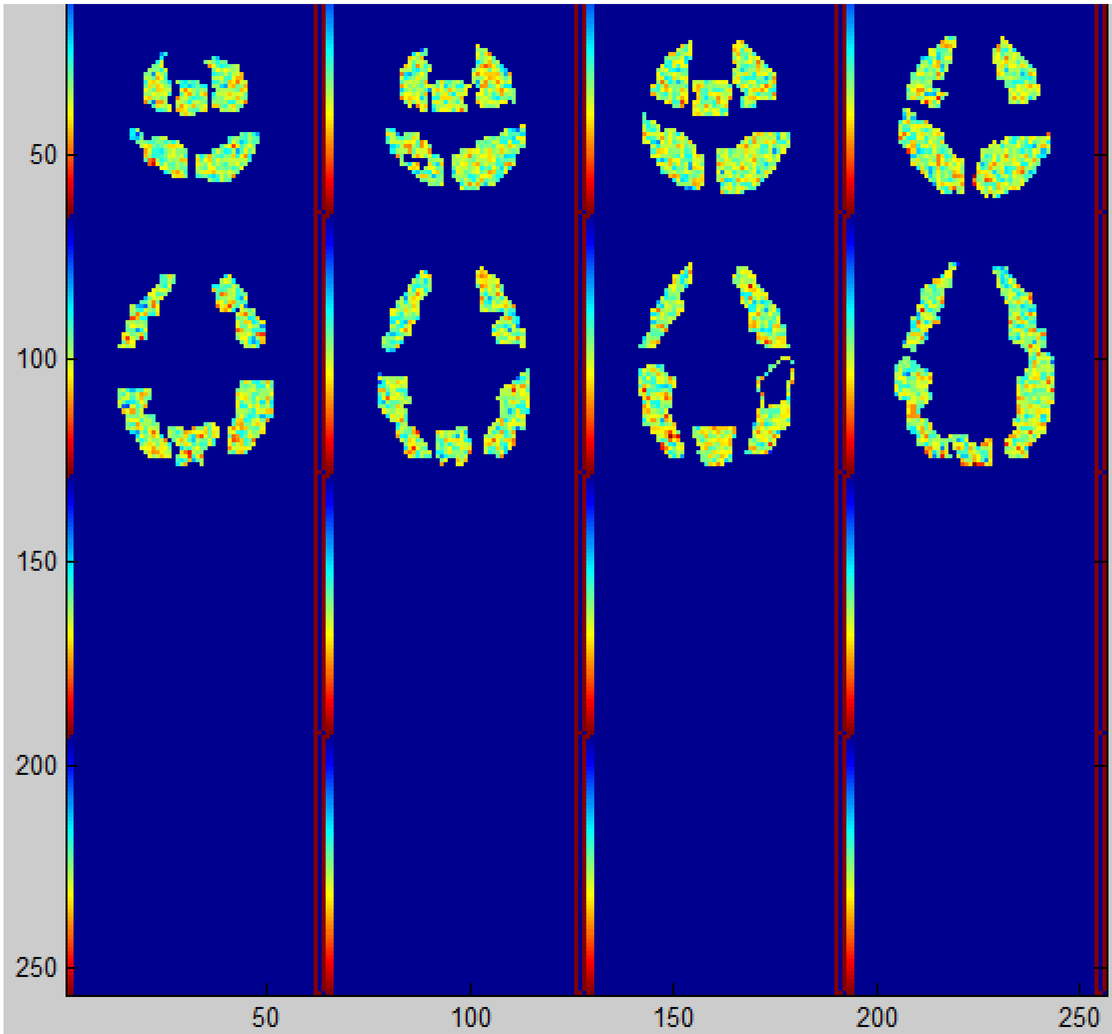

Figure 5: Picture of activation for all slice of the brain for trial 4, at time 8

3.3 Evaluating results

In this research **(Mitchell et al 2004)** all the classifiers for different subjects were classified using cross validation. There are processes that yield 2 classes and the classification result on that can be measured by building a confusion matrix; however there are cases where more than 2 classes are output. In that case success of the prediction is measured by the normalized rank of the correct class in the list which is already sorted. This normalized rank error principle works any number of classes. When a class is predicted properly the normalized rank error is 0 for that and when the rank is least likely the value is 1.Also random guessing results rank error value as 0.5 irrespective of number of classes. To make the k-fold cross validation work properly in this process classes are balanced by taking same number of example per class. This is done for the very nature of fMRI BOLD signal nature, which picks up or blurs out for several seconds. Hence if starting timestamp is t, an image occurring at time t+1 or t-1 can be largely correlated with the other image and if included in training set can cause optimistic bias in training data. This particular bias is removed by removing all images from the training set that occurs within 5 seconds of the test image that has been held out. In this validation of the experiment in leave one out cross validation the test set will have one example per class and the entire nearby image that can cause bias are removed from the training set.

3.4 Role of Feature selection

As found in the previous section an interval of 8 seconds creates a feature vector of dimension 160,000 hence that makes the data very high dimensional in nature. Also the data inherently comes up with noise in it. Hence to make the data more effective for using in classification, dimension reduction or feature selection technique is required. Although dimension reduction didn’t result betterment in classification, feature selection was useful.
In the following sections the present feature selection technique and improvement on that will be discussed.

3.4.1 Present Approach

Among the feature selection techniques the widely used technique is to choose those features that discriminate the target class most accurately. A general approach to learn a classification is to make a list of all features available and sort them on the basis of how well they can discriminate the target class and then select some of the well performing features.

Before deciding the proper feature selection technique we need to know the nature of the data. Unlike typical boolean class problem this data actually has 3 class associated with it.Two of that are 1 (picture) and 2 (sentence) and the 3rd one is the rest state, when the subject is put to rest. The main difference of the resting state class with other 2 classes is there is no signal present in that class unlike other two classes and only background noise is present there. So one of the alternative feature selections that could be effective was to determine how a feature can differentiate class 1 and class 2 from rest state or zero signal. This technique grades each feature basing on how active it was during the class 1 time period or class 2 time period compared to rest interval. This approach actually ends up choosing voxels that has high signal-to- noise ratio. This whole setting was referred to as “zero signal” learning setting **(Mitchell et al 2004)**.This type of feature selection has application in other field as well.

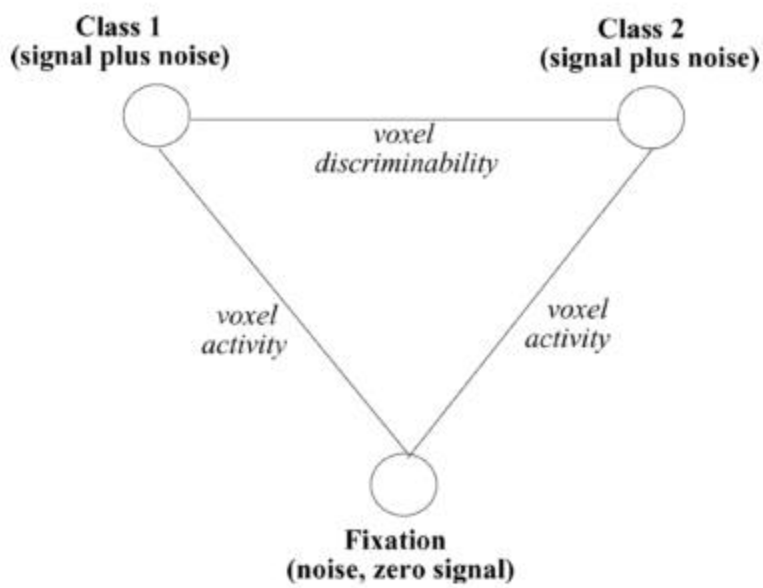

Figure 6. “*The ‘zero signal’ learning setting. Boolean classification problems in the fMRI domain naturally give rise to three types of data: data corresponding to the two target classes plus data collected when the subject is in the ‘fixation’ or ‘rest’ condition. We assume the data from class 1 and class 2 are composed of some underlying signal plus noise, whereas data from the fixation condition contains no relevant signal but only noise. In such settings, feature selection methods can consider both voxel discriminability (how well the feature distinguishes class 1 from*

*class 2), and voxel activity (how well the feature distinguishes class 1 or class 2 from the zero signal class).''***(Mitchell et al 2004)**

The target of this feature selection was to reduce the number of voxel to consider and select only a subset from the training sample and compare that with the corresponding subset from the test sample. Two of the most useful feature selection methods used were that select voxels basing on their ability to discriminate target classes from one another and that select voxels basing on their ability to discriminate target classes from the rest condition.

In the first method a classifier was trained for each voxel and the accuracy of the classifiers on a single voxel was considered its discriminating power and n highest scoring voxel were selected. In the second method for each voxel and each class a t-test was conducted to test the voxels performance for that class and its performance to the resting class. Then n voxels were chosen that have got highest t-statistics.

3.5 Improved Approach

In **Mitchell et al 2004** selecting n most active voxels were used as most differentiating factor while extracting the fMRI dataset that can be used to predict cognitive state of human. Other than that selecting n most active voxels per Region of Interest was also a good factor. In this paper the most successful active voxel selection method has been improved and a new method has been formulated using earlier selecting n most active voxels per Region of Interest method.

3.5.1 Selecting n most active voxels [1](Mitchell et al 2004) improvement

In this process from a certain dataset initially n numbers of voxels are selected and 10 fold cross validation is applied on that dataset. The numbers of voxels are incremented going forward and the accuracy received from later steps is compared with the accuracy received from earlier state. When for 3 consecutive steps there is no improvement in accuracy the process will stop and the minimum number of active voxels that resulted highest accuracy will be selected.

Algorithm

Step 1: Initialize n=5, where n=number of active voxel

Step 2: Store the accuracy received for that n in array A

Step 3: Increase n by 5, n=n+5

Step 4: Store the accuracy and compare with earlier stored accuracy

Step 5: Repeat step 3, 4 when new accuracy >= old accuracy

Step 6: If new accuracy = old accuracy for 3 cycle return n for highest accuracy

3.5.2 Select mean of n active ROI [1](Mitchell et al 2004) improvement

In this process using selectROIVoxels method the 10 fold cross validated accuracy for every ROI are collected and stored in an array where ROI name is the index of the array. The array is then sorted in decreasing accuracy. Then chose top 5 ROIs from the list and using avgROIVoxels method create the dataset. Using 10 fold cross validation calculate the accuracy and store it. Now successively add 6th, 7th and subsequent ROIs to the avgROIVoxels and calculate the accuracy and store it. If new accuracy is less than previous accuracy or unchanged for 3 consecutive cycles return the least number of ROIs.

Algorithm

Step 1: Use selectROIVoxels and store accuracy received from all ROI in array with ROI name as index

Step 2: Sort the array in descending order

Step 3: Select top 5 ROI from array

Step 4: Prepare dataset using avgROIVoxels on initial ROIs

Step 5: Store the accuracy received using 10 fold CV

Step 6: Add next ROI from the list and store accuracy using received using 10 fold CV

Step 7: Repeat step 6 as long as previous accuracy <= new accuracy

Step 8: If previous accuracy = new accuracy for 3 consecutive loop stop

Step 9: Return list of minimum number of ROI that return highest accuracy.

Also before using the fMRI dataset in classification a supervised resample filtering method has been used on the datasets. Some of the condition for using that is the original dataset must fit entirely in memory and the dataset must have a nominal class attribute. The filter can be made to maintain the class distribution in the subsample, or to bias the class distribution toward a uniform distribution.

4. Solution Evaluation

For this evaluation two types of result have been collected. In Table 1,2,3 and 4 the result of 10 fold cross validation of a dataset (extracted from fMRI data using different methods) have been enlisted and in table 5,6 and 7 the list of accuracies have been listed where among 3 datasets one is used as training dataset when other 2 are used as test datasets.

While evaluating the result as seen between Table 1 and Table 2 the improved method of using optimum number of active voxacts cause on average 3.25%,1.25%,2.5% improvement in accuracy respectively for dataset 4799,4820 and 4847.The improvement is further enhanced by using supervised resampling filter to 4%,4.5% and 4% respectively for the dataset 4799,4820 and 4847 from the baseline as seen between table 1 and table 3.However given that the average accuracy for all baseline dataset already stands at 88% this could be considered a significant improvement. Also 1.75%, 1.5% and 4.25% improvement in accuracy respectively for datasets 4799, 4820 and 4847 can be seen while using the new active roi based method as seen between table 1 and table 4.

To evaluate the result seen in Table 5, 6 and 7 we can see that the improved activity based method and new ROI based method improves the base accuracy. For all 6 cases improved activity based method improves it by 5.75%, 8%,-3.25%, 1.5%,-1% and 3% respectively, on average 2.33% increase. The new active ROI based method improves it by 2.25%, 5.5%, 0.5%, 0%, 4.5% and 4.75% respectively, on average 2.91% increase. Considering the overall average for all processes is 86.25% the improvement is significant.

Also some additional findings are SMO and logistic regression are better than other classifiers on the high dimension data. Also as SVM works on the principle of creating a margin in high dimension place the presence of high dimension doesn't help its purpose. Another observation is Adaboost works well with tree based weak classifier SimpleCart, but in case of tree based strong classifier HoeffdingTree the improvement using boosting was marginal.

Datasets	Naive Bayes	Logistic	SMO	SVM(Sig)	Adaboost(H)	Adaboost(S)
4799	86	88	88	84	88/84	90/80
4820	92	92	94	86	96/96	88/80
4847	90	88	92	88	90/88	86/78

Table 1: Result of base method using few active voxacts using 10 fold CV.

Adaboost (H) stands for result of HoeffdingTree classifier with and without Adaboost.

Adaboost(S) stands for result of SimpleCart classifier with and without Adaboost.

SVM (Sig) stands for the version with sigmoid kernel.

Datasets	Naive Bayes	Logistic	SMO	SVM(Sig)	Adaboost(H)	Adaboost(S)
4799	90	92	92	86	90/88	86/78
4820	92	92	96	94	94/94	90/78
4847	92	92	96	90	92/92	80/82

Table 2: Result of improved method using optimum number of active voxacts using 10 fold CV.

Datasets	Naive Bayes	Logistic	SMO	SVM(Sig)	Adaboost(H)	Adaboost(S)
4799	94	92	96	82	94/94	90/78
4820	96	92	98	94	96/96	96/92
4847	92	92	96	90	96/96	92/88

Table 3: Result of improved method using optimum number of active voxacts using 10 fold CV After Supervised Resample Filtering.

Datasets	Naive Bayes	Logistic	SMO	SVM(Sig)	Adaboost(H)	Adaboost(S)
4799	**90**	**92**	**92**	**86**	**90/88**	**86/78**
4820	**94**	**96**	**96**	**94**	**94/94**	**90/78**
4847	**92**	**100**	**96**	**90**	**92/92**	**80/82**

Table 4: Result of active ROI based method using 10 fold CV

Training	Naive Bayes	Logistic	SMO	SVM(Sig)	Adaboost(H)	Adaboost(S)	Testing
4799	84	92	92	90	90/84	88/74	**4820**
4799	90	92	86	88	90/90	78/62	**4847**
4820	92	82	86	84	86/92	90/80	**4799**
4820	96	88	94	88	96/96	88/82	**4847**
4847	88	86	86	88	80/80	68/84	**4799**
4847	96	90	90	92	60/96	92/74	**4820**

Table 5: Baseline activity based classification result

Training	Naive Bayes	Logistic	SMO	SVM(Sig)	Adaboost(H)	Adaboost(S)	Testing
4799	96	94	96	92	96/96	90/80	**4820**
4799	96	94	92	92	96/96	94/80	**4847**
4820	84	86	84	84	84/82	82/80	**4799**
4820	96	90	96	90	96/96	96/80	**4847**
4847	80	88	86	86	80/80	68/84	**4799**
4847	90	92	96	94	90/90	88/74	**4820**

Table 6: Improved activity based classification result

Training	Naive Bayes	Logistic	SMO	SVM(Sig)	Adaboost(H)	Adaboost(S)	Testing
4799	92	90	94	94	94/94	88/86	**4820**
4799	92	92	90	90	96/96	92/88	**4847**
4820	86	88	88	88	88/86	86/84	**4799**
4820	86	96	92	90	94/94	94/90	**4847**
4847	86	88	90	88	88/84	78/82	**4799**
4847	92	90	94	94	92/89	86/78	**4820**

Table 7: New ROI Based classification result

5. Conclusion

This paper uses fMRI data that was collected from different subjects. It tries to improve the feature selection method uses the data more effectively by reducing number of feature variable (voxel) present in the dataset. It does so by introducing two algorithms to select better dataset using most active voxel and most active ROI. Also it exposes the dataset to some new set of supervised classifiers such as SMO and tree based classifiers HoeffdingTree and SimpleCart. It also shows that using a weak classifier with Adaboosting process help get good accuracy from this.

6. Future work

The availability of ground breaking technique like fMRI has made meaningful inside data of human brain more and more available for researcher. It has also made this field an active research area and already researches have been done to share classifiers across spatial and temporal domain and across different types of subject. However unlike as seen in this paper human thoughts are not always as simple as reading a sentence or a picture and also the aspect of outside thought crossing the test subject brain at the time of test makes this process more complex. So to make a process that will not be susceptible to outside thought and will be able to classify real world complex thought should be a good continuation of this work. Also with the advent of technology the possibility of making portable fMRI machine that can take the input from a human brain outside test condition is not far-fetched anymore. So a continued work to better this technology to apply it on subjects outside test condition to thought-read could be hugely useful in the field of criminology or in any other related fields

Cited Literature

Cribben, Ivor, Ragnheidur Haraldsdottir, Lauren Y. Atlas, Tor D. Wager, and Martin A. Lindquist. "Dynamic connectivity regression: determining state-related changes in brain connectivity." Neuroimage 61, no. 4 (2012): 907-920.

Datta, Susmita, Vasyl Pihur, and Somnath Datta. "An adaptive optimal ensemble classifier via bagging and rank aggregation with applications to high dimensional data." *BMC bioinformatics* 11, no. 1 (2010): 427.

Demichelis, Francesca, Paolo Magni, Paolo Piergiorgi, Mark A. Rubin, and Riccardo Bellazzi. "A hierarchical Naive Bayes Model for handling sample heterogeneity in classification problems: an application to tissue microarrays."*BMC bioinformatics* 7, no. 1 (2006): 514.

Haynes, John-Dylan, and Geraint Rees. "Decoding mental states from brain activity in humans." *Nature Reviews Neuroscience* 7, no. 7 (2006): 523-534.

Hutchinson, Rebecca A., Radu Stefan Niculescu, Timothy A. Keller, Indrayana Rustandi, and Tom M. Mitchell. "Modeling fMRI data generated by overlapping cognitive processes with unknown onsets using Hidden Process Models."*NeuroImage* 46, no. 1 (2009): 87-104.

Jin, Bo, Alvin Strasburger, Steven J. Laken, F. Andrew Kozel, Kevin A. Johnson, Mark S. George, and Xinghua Lu. "Feature selection for fMRI-based deception detection." *BMC bioinformatics* 10, no. Suppl 9 (2009): S15.

Keller, T. A., M. A. Just, and V. A. Stenger. "Reading span and the time-course of cortical activation in sentence-picture verification." In Annual Convention of the Psychonomic Society. 2001.

Kim, Kyung In, and Richard Simon. "Probabilistic classifiers with high-dimensional data." *Biostatistics* 12, no. 3 (2011): 399-412.

Lindquist, Martin A. "The statistical analysis of fMRI data." Statistical Science23, no. 4 (2008): 439-464.

McKeown, Martin J., Lars Kai Hansen, and Terrence J. Sejnowsk. "Independent component analysis of functional MRI: what is signal and what is noise?" Current opinion in neurobiology 13, no. 5 (2003): 620-629

Mitchell, Tom M., Rebecca Hutchinson, Radu S. Niculescu, Francisco Pereira, Xuerui Wang, Marcel Just, and Sharlene Newman. "Learning to decode cognitive states from brain images." *Machine Learning* 57, no. 1-2 (2004): 145-175.

Mitchell, Tom M., Rebecca Hutchinson, Marcel Adam Just, Radu S. Niculescu, Francisco Pereira, and Xuerui Wang. "Classifying instantaneous cognitive states from fMRI data." In *American medical informatics association annual symposium*. 2003.

Norman, Kenneth A., Sean M. Polyn, Greg J. Detre, and James V. Haxby. "Beyond mind-reading: multi-voxel pattern analysis of fMRI data." *Trends in cognitive sciences* 10, no. 9 (2006): 424-430.

Palatucci, Mark, and Tom M. Mitchell. "Classification in very high dimensional problems with handfuls of examples." In *Knowledge Discovery in Databases: PKDD 2007*, pp. 212-223. Springer Berlin Heidelberg, 2007.

Rustandi, Indrayana. "Feature Selection on Raw and Wavelet-Transformed fMRI Data."

Rustandi, Indrayana. "Classifying multiple-subject fMRI data using the hierarchical Gaussian Naïve Bayes classifier." In *13th Conference on Human Brain Mapping*. 2007.

Shen, Xilin, and François G. Meyer. "Nonlinear dimension reduction and activation detection for fmri dataset." In *Computer Vision and Pattern Recognition Workshop, 2006. CVPRW'06. Conference on*, pp. 90-90. IEEE, 2006.

Tillander, Annika. "Classification models for high-dimensional data with sparsity patterns." (2013).

VITA

NAME: Tanay Chowdhury

EDUCATION: B.Tech. Information Technology,

Kalyani Government Engineering College, 2005

EXPERIENCE: BI Consultant & Data Analyst, Cognizant Technology Solutions (2005-2013)