Chandigarh University Journal of Undergraduate Research and Innovation

# UNIVARIATE ANALYSIS OF FMR DATA USING MACHINE LEARNING

Pankaj Pandey
Chandigarh University
Mohali, India

Neelam Sinha
International Institute of Infomation Technology
Bangalore, India

***Abstract*—Human Brain is always a fascinating area of study among researchers, and increasing research due to the increasing awareness and advancement in technology such as functional Magnetic Resonance Imaging (fMRI) to quantify functionality of the brain. fMRI has been the widely popular powerful tool to determine the cognitive states. The vibrant importance of fMRI are non-invasive process and found to be effective tool to diagnose medical disorders. Data is captured during fMRI does have high dimension and noise, hence employment of Machine Learning would be the productive choice to see the patterns in the data. Here we have 3D time-series data of subjects those are engaged in well-defined cognitive tasks. We design methods for classification between the cognitive tasks. Two methods have been carried out for finding the discriminating voxels for classification. (a) Pivotal voxels: Classification strength of each voxel is obtained and the best performing voxels are identified based on an empirical threshold, labelled here as pivotal voxels. (b) t-discriminating voxels: Two sample t-test is implemented for each voxel and those voxels with the lowest p-value, tagged here as t-discriminating voxels. We have used Public datasets: Probid, in which three cognitive states[5 subjects] (i) Viewing pleasant images (ii) Unpleasant (iii) Neutral images, we have 10 brain volume for each subject.**

***Keywords—fMRI; SVM; Gaussian NaiveBayes;***

## I. INTRODUCTION

Brain function is interesting subject for neuroscientist, psychologist, thanks to the state of the art research that has pushed fMRI to reach the potential so that it can capture images of neural activity with high spatial and temporal resolution. It has reached far distance since when it was pioneered, it is huge significant encouragement over the last decade from the advancement in Imaging techniques , a brain imaging method that has done the remarkable job to capture the neural brain activity of human subjects at high spatial resolution (several millimeters), across the brain volume. The benefits of fMRI are non-invasive and found to be effective tool to diagnose medical disorders.Data is captured during fMRI does have high dimension and noise, hence employment of Machine Learning would be the productive choice to see the patterns in the data. This has beed used to state a variety of studies on human brain, like frontal cortex is involved during working memory that could help in storing digits for very short duration and could help to improve mathematical ability. There are other medical challenges such as Autism and Alzheimers which is growing and pose a serious dedication among doctors, hence fMRI is the one of the powerful tool to analyze and find the biomarkers of these diseases.

## II. FUNCTIONAL MAGNETIC RESONANCE IMAGING

Functional Magnetic Resonance Imaging (fMRI) is a technique using Magnetic Resonance scanner that combines magnetic fields,magnetic gradients and radio waves to captures the 3D volume of brain. The priniciple behind fMRI is to obtain the Blood Oxygen Level Dependent (BOLD) contrast. It means, fMRI measures the ratio of oxygenated blood to deoxygenated blood with respect to a control baseline at relevant points on a 3 dimensional grid.

Local neural activity is captured by change in blood flow. It has been observed that activation of a brain area produces an increase in neuronal metabolism, accompanied by an increase in blood fow and blood volume. To obtain the neuronal metabolism in a local area (depending upon the activity being performed), the brain has to maintain an adequate supply of oxygen and glucose.

## III. STATE OF THE ART

There is enough work done in the literature of fMRI, the beginning era of investigating the brain using fMRI study [2], had shown to distinguish cognitive states that represent distinct tasks, such as observing a sentence/picture, reading an ambiguous/non ambiguous sentence etc. Variety of problems in machine learning require deep analysis because of high dimension and noise in the data, and very few availablility. In [3], the employment of classifying brain states is understood as mechanism for feedback in real time to control the stimulus . In this [4], identifies the challenges for finding the optimal features selection using analysis of algorithms. In [5], the researchers grouped a variety of classifiers for optimal performance. In [6], a new group of classifiers called Generalized Sparse Classifiers to address the issue of over-fitting, due to the large-dimensional data. In this study [7], the authors have designed in such a way that the phases of feature selection and classification are collected as one step. Support Vector Decomposition Machine is prominent machine learning algorithm that is employed so that reduction of dimensional could help to acheive the best classification. In [9], Corelation between features are detected by approach as Random forest which helps to get the cloud of informative regions providing patterns for predictive voxels that can be interpreted.In [8], Reducing the no of features that are having low importance, Recursive Feature Elimination plays a important contribution in which Support Vector Machine is behind the scenes. Several research groups have studied to localize the Region of Activation [2].

Our focus is to identify the optimal feature selection and regions of activation.

## IV. METHODOLOGY

*A. Pivotal voxels : Pivotal voxels are those voxels that can be identified after the voxel wise analysis, each voxel is analysed so that it comes into best -classfying voxel. Here below is the method to find such voxels.*

(a) Single voxel classification of the two considered cognitive states is carried out. We employ the Leave-One-Example-Out validation scheme, where we use all available examples for training the classifier with the exception of one that is used for testing. This is repeated for all the examples in a particularsubject and the accuracy of the classification is calculated by the number of correct predictions.

(b) For each voxel, we measure the classifying accuracy with threshold , if it comes under the threshold category, we mark it as pivotal voxels, hence i voxels are selected from j voxels.

For finding an empirical threshold value of classification accuracy, we follow the below process.

(a) We have initiated with randomly chosen value and that value is increased and decreased so that we can identify a effective threshold value.

(b) In each iteration, for every distinct value of threshold, separate sets of pivotal voxels are obtained. The average misclassification error over all subjects is computed for each separate set of pivotal voxels.

*B. t -discriminating voxels :For each voxel, a two sample t-test is applied to compare the voxels fMRI activity of one cognitive task with respect to other cognitive task. We test the alternative hypothesis that the population means are not equal, and then n voxels are selected with the lowest p-value, and are called as t-discriminating voxels.For selecting the n voxels, we have performed the following iterative procedure*

(a) We have initiated with randomly chosen m voxels with lowest p-values and note the accuracy of classification with these voxels, repeat the process with both, increasing and decreasing range of m.

(b) The best performing set of voxels are set as n t-discriminating voxels.

## V. DATASETS AND EXPERIMENTAL RESULTS

*A. Dataset*

Probid - Real fMRI data sets available at the website [10] is the second dataset used to assess the performance of the proposed scheme.

In this data, there is three visual tasks performed by a Subject, the stimuli were presented in an event-related fashion. There were three different active conditions: viewing unpleasant (dermatological diseases), neutral (people) and pleasant (pretty girls in swimsuits) images, and a control condition (fixation) [11]. The stimuli were presented in a random order according to a randomized design. There were 10 image presentations (events) of each condition.Data is available as a time series of 3-D brain volumes, of size 79x95x69, collected over 121 time points.

We have used Probid toolbox [11] for pre-processing the data, after preprocessing ,we are having total of 10 brain volumes available from each subject (10 volumes per class). The number of voxels per subject was 2,50,000. Yielding an input feature matrix [10 × 250000] for each subject per class.We have performed classification by taking all the examples from all subjects per class and forming a matrix [50 × 250000] per class. Subject-wise classification is not carried out due to non-availability of required data.

*B. Classifiers and Statistical Techniques*

In this section, we have explained two classifiers and one statistical method that are used in our experiments and defining about cross validation method and data preparation.

(a) Gaussian Naive Bayes classifier(GNB) - A Gaussian Naive Bayes algorithm is a extended version of naive bayes algorithm. It’s used often when the features are posessing continuous values and also assume that all the features are supposed to have a gaussian distribution i.e, normal distribution.

(b) Support Vector Machine Support vector machine is most popular Machine learning algorithms and belongs to the class of supervised learning algorithms which is widely used in two class classification problems. The basic principle of Support Vector machines in its simplest form is that it projects the input vector to be classified as points in a high-dimensional space and finding a line that separates them.

The distance from the separating hyperplane to the nearest vector is the margin and SVM selects the separating hyperplane margin that is at maximum distance from the support vectors. This effectively maximizes SVMs capability to predict the correct classification of testing samples. There are cases where the data cannot be separated by a linear hyperplane, instead the separating hyperplane turns out to be a single point which is of no used for any type of classification or regression problem. Hence a mathematical function called the kernel function is used in SVM to perform a higher dimensional classification of an input vector which is of lower dimension. Different types of kernel functions are available based on which the separating hyperplanes can be obtained in higher dimension. Most popular kernel functions are Linear Kernel, Polynomial Kernel, Radial Basis Kernel and Gaussian Kernel.

(c) t-test It is a widely used statistical technique in neuroimaging that is used to detect brain areas which are relevant for a particular cognitive task. A t-test is any test where the quantity measured (derived from a set of samples) has a Student T distribution if the null hypothesis is true. It is applied when the population follows a Gaussian distribution but the sample size is too small for statistic purposes (quantitative measurement). It is often used to compare two small sets of quantitative data when samples are collected independent of one another.

*C. Experiment on Probid Data*

In this work, we classify cognitive datasets (i) a subject viewing pleasant image or Unpleasant image (ii) a subject viewing pleasant image or neutral image (iii) a subject viewing unpleasant image or neutral image.

(a) Implementation of different Methods

Following are the three different methods that help us to classify these cognitive classes

(i)Pivotal Voxels : Pivotal voxels are selected after setting a threshold value between two cognitive states in all three datasets of classification. For establish a threshold value, we have tried some values as shown in Fig 5.1. We have obtained .6 as the threshold that is consistent in all three sets. We have selected after observing the accuracy in both the classifier and number of voxels. Classification using SVM yielded than 90% in two datasets pleasant versus unpleasant and unpleasant versus neutral.

(ii) t-discriminating voxels : We run t-test on three sets and selected the voxels on the ascending order of p-values, we have tried 20 different random values as shown in Fig 5.2. We have successfully obtained accuracy more than 90 % in two sets.

(b) Discussion

Here, we classify three sets of cognitive states.

1. Pleasant versus Unpleasant

2. Pleasant versus Neutral

3. Unpleasant versus Neutral

We have classified (i) &(iii) with 90 % accuracy and (ii)with about 85%. We have obtained more than 90% accuracy in (i) and (iii) using SVM with pivotal voxels and t-discriminating voxels. Pivotal voxels outperforms t-discriminating voxels and SVM gives better result than GNB.

| Threshold | SVM | GNB | Voxels | Threshold | SVM | GNB | Voxels | Threshold | SVM | GNB | Voxels |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Plesant_Unpleasant | | | | Plesant_Neutral | | | | Unplesant_Neutral | | | |
| 0.5 | 0.84 | 0.7 | 121734 | 0.5 | 0.8 | 0.58 | 124474 | 0.5 | 0.8 | 0.63 | 123706 |
| 0.55 | 0.9 | 0.73 | 50607 | 0.55 | 0.84 | 0.68 | 50833 | 0.55 | 0.85 | 0.71 | 48905 |
| 0.6 | 0.92 | 0.79 | 6110 | 0.6 | 0.82 | 0.76 | 6706 | 0.6 | 0.93 | 0.84 | 5248 |
| 0.65 | 0.89 | 0.87 | 436 | 0.65 | 0.79 | 0.82 | 439 | 0.65 | 0.91 | 0.84 | 338 |
| 0.7 | 0.8 | 0.77 | 11 | 0.7 | 0.71 | 0.74 | 4 | 0.7 | 0.69 | 0.73 | 7 |

Fig 5.1 : Accuracy of classification in all three sets at different threshold

| Plesant vs Unpleasant | | | Pleasant vs Neutral | | | Unpleasant vs Neutral | | |
|---|---|---|---|---|---|---|---|---|
| SVM | GNB | Voxels | SVM | GNB | Voxels | SVM | GNB | Voxels |
| 0.88 | 0.85 | 1000 | 0.72 | 0.8 | 1000 | 0.97 | 0.82 | 1000 |
| 0.93 | 0.84 | 2000 | 0.8 | 0.77 | 2000 | 0.92 | 0.8 | 2000 |
| 0.92 | 0.85 | 3000 | 0.77 | 0.76 | 3000 | 0.93 | 0.82 | 3000 |
| 0.92 | 0.84 | 4000 | 0.8 | 0.74 | 4000 | 0.9 | 0.8 | 4000 |
| 0.94 | 0.82 | 5000 | 0.83 | 0.77 | 5000 | 0.93 | 0.8 | 5000 |
| 0.91 | 0.83 | 6000 | 0.82 | 0.71 | 6000 | 0.91 | 0.78 | 6000 |
| 0.91 | 0.8 | 7000 | 0.81 | 0.74 | 7000 | 0.91 | 0.79 | 7000 |
| 0.92 | 0.79 | 8000 | 0.84 | 0.72 | 8000 | 0.91 | 0.77 | 8000 |
| 0.93 | 0.78 | 9000 | 0.85 | 0.71 | 9000 | 0.92 | 0.78 | 9000 |
| 0.92 | 0.79 | 10000 | 0.84 | 0.7 | 10000 | 0.93 | 0.76 | 10000 |
| 0.93 | 0.79 | 11000 | 0.83 | 0.71 | 11000 | 0.92 | 0.77 | 11000 |
| 0.94 | 0.76 | 12000 | 0.83 | 0.73 | 12000 | 0.92 | 0.79 | 12000 |
| 0.93 | 0.77 | 13000 | 0.88 | 0.68 | 13000 | 0.91 | 0.75 | 13000 |
| 0.91 | 0.78 | 14000 | 0.83 | 0.69 | 14000 | 0.92 | 0.75 | 14000 |
| 0.9 | 0.77 | 15000 | 0.83 | 0.68 | 15000 | 0.95 | 0.79 | 15000 |
| 0.91 | 0.75 | 16000 | 0.85 | 0.69 | 16000 | 0.92 | 0.75 | 16000 |
| 0.94 | 0.79 | 17000 | 0.83 | 0.66 | 17000 | 0.93 | 0.76 | 17000 |
| 0.92 | 0.78 | 18000 | 0.82 | 0.66 | 18000 | 0.92 | 0.74 | 18000 |
| 0.91 | 0.78 | 19000 | 0.8 | 0.69 | 19000 | 0.93 | 0.76 | 19000 |
| 0.92 | 0.77 | 20000 | 0.86 | 0.67 | 20000 | 0.91 | 0.73 | 20000 |

Fig 5.2: Accuracy of classification in all three sets and voxels are selected with ascending order of p-values

## VI. Future work

In our work, we have done univariate analysis on fMR data, we can extend this work with multivariate analysis on voxels and could find the correlation involved among the regions.

We have worked on few subjects, there are many datasets available publicly in which subjects are more than 100. And futher we can work on medical disorders such as Autism, Alzheimer's.

## References

[1] V. M. Haughton J. S. Hyde B. Biswal, F. Z. Yetkin. Functional connectivity in the motor cortex of resting human brain using echo planar mri. Magnetic resonance in medicine, 34(4):537–541, Oct 1995.

[2] R. Niculescu F. Pereira X. Wang M. Just T. Mitchell, R. Hutchinson and S. Newman. Learning to decode cognitive states from brain images. Machine Learning, 1(2):145–175, 2004. Scott J. Peltier Stephen M. LaConte and Xiaoping P. Hu. Real-time fmri using brain-state classification. Human Brain Mapping, 28:1033–1044, 2007.

[3] Tom Mitchell b Matthew Botvinick Francisco Pereira a, . Machine learning classifiers and fmri: A tutorial overview. Neuroimage, 45:199–209, 2009.

[4] Margarida Silveiraa Carlos Cabrala [Author Vitae] and b [Author Vitae] Patricia Figueiredoa. Decoding visual brain states from fmri using an ensemble ofclassifiers. Pattern Recognition, 45(6):2064–2064, June 2012.

[5] Ghassan Hamarneh Bernard Ng, Arash Vahdat and Rafeef Abugharbieh. Generalized sparse classifiers for decoding cognitive states in fmri. MACHINE LEARNING IN MEDICAL IMAGING, 6357:108–115, 2010.

[6] Francisco Pereira and Geoffrey Gordon. The support vector decomposition machine. Proceedings of the 23rd International Conference on Machine Learning, Pittsburgh, pages 689–696, 2006.

[7] Gianarlo Valentea Nol Staerena John Ashburnerb Rainer Goebela Elia Formisanoa Federico De Martinoa. Combining multivariate voxel selection and support vector machines for mapping and classification of fmri spatial patterns. NeuroImage, 43(1):4458, Oct 2008.

[8] E. Eger R. Genuer, V. Michel and B. Thirion. Random forests based feature selection for decoding fmri data. IEEE International Symposium on Biomedical Imaging, 2009.

[9] Probid. http://www.brainmap.co.uk/.

[10] Probid toolbox http://www.kcl.ac.uk/ioppn/depts/neuroimaging/

[11] research/imaginganalysis/Software/PROBID.aspx.