# Classification and Statistical Analysis of Auditory FMRI Data Using Linear Discriminative Analysis and Quadratic Discriminative Analysis

**Pritesh G. Shah, Bharti W. Gawali**

***Abstract* — Functional magnetic resonance imaging (fMRI) has the ability to not only get insight into how human brain functions but also to study the human brain of normal and diseased subjects. One of the methods to analyze the fMRI data is univariate approach, another approach is Multivariate discriminative approach. However for classification, there exists number of statistical techniques. In this paper, we perform classification of fMRI data using Fishers Linear Discriminative Analysis (LDA). The re-substitution error for the LDA calculated to be 0.1875. It is concluded that, the data are consistent with the multivariate normal distribution.**

***Index Terms* — Functional magnetic resonance imaging (fMRI), Linear Discriminative Analysis (LDA), Quadratic Discriminative Analysis (QDA), Statistical Parametric Mapping (SPM).**

## I. Introduction

Functional magnetic resonance imaging (fMRI) technology is primarily used to perform brain activations by measuring the neural activity. It is non-invasive technique to study brain activity. In fMRI experiment, a series of brain images are acquired while the subject is inside the scanner. Inferences regarding task related activations in the brain are made on the basis of changes in the measured signal between individual images.

Designing fMRI experiment expects adequate balance between spatial resolution and temporal resolution; where the temporal resolution determines the ability to separate brain events in time. This is determined by how quickly each individual image is acquired; and spatial resolution determines ability to distinguish changes in an image at various spatial locations.

In the fMRI terminology, the radio frequency signal can be considered in two ways known as T1 and T2.

**Manuscript received November 21, 2014.**

**Pritesh G. Shah**, Department of Computer Science and Information Technology, Dr. Babasaheb Ambedkar Marathwada University, Aurangabad, Maharashtra, India.

**Bharti W. Gawali**, Department of Computer Science and Information Technology, Dr. Babasaheb Ambedkar Marathwada University, Aurangabad, Maharashtra, India.

A structural T1 image possesses important characteristics which include [1]:

i. It has high spatial resolution.
ii. It has low temporal resolution.
iii. Able to distinguish among various types of tissues

A functional T2* images do have important characteristics which include [1]:

i. It has low spatial resolution.
ii. It has high temporal resolution.
iii. It can relate changes in signal to experimental changes.

For T1-weighted versus T2-weighted images, white matter will be brighter than gray matter in the T1-weighted images but vice versa in the T2-weighted image.

However, in fMRI, the large number of dimensions also poses a more fundamental problem in modeling. To adequately model a dataset, we need to measure systematically across the range of each variable being measured.

One way to do this is, to reduce the number of dimensions. For example, if there is redundant information across many dimensions, then one can use dimension reduction techniques such as principal component analysis or independent component analysis to identify a smaller set of alternative dimensions that give rise to the data [2].

Literature review suggests that, there has been growing interest in the use of machine learning classifiers for analyzing fMRI data. Machine learning is concerned with developing programs that automatically learn with experience. Classifiers learn how to perform prediction task. In this paper, we perform classification of fMRI data using Fishers Linear Discriminative Analysis (LDA).

The rest of the paper is organized as follows: Second section describes the methodology, Third sections gives us, the details of the database. Fourth section describes the classification task whereas fifth section shows classification results. Finally, the concluding remarks are given.

## II. Methodology

Here, we describe the computational steps of analysis of fMRI data which starts with, pre-processing the fMRI data in Statistical Parametric Mapping (SPM). The experimental workflow to classify fMRI data using linear discriminative analysis (LDA) technique is as shown in Figure 1.

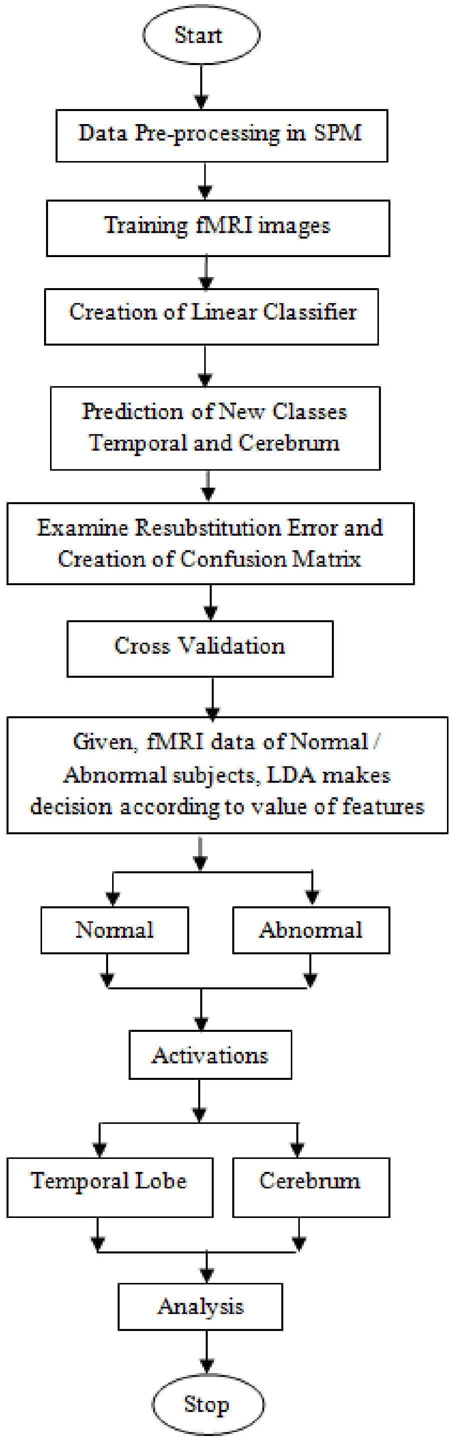


Figure 1: Experimental workflow for LDA Classifier

## III. Database

The Experiment uses the dataset of the Functional Imaging Biomedical Informatics Research Network (FBIRN) [3] which is multi-site dataset of subjects with schizophrenia and control subjects. Dataset includes functional MRI images, behavioral data, demographic, and clinical assessments on 253 subjects from around the US. Subjects, evaluated with a standardized clinical assessment. Auditory Oddball task (AO) was used for image acquisition with controlled variation. AO task consisted of two practice runs and 4 experimental runs each having duration of 280 seconds (4.67 minutes). Each run consisted of a standard tone (1000 Hz) and an oddball tone (1200 Hz). For the duration of each run, subjects viewed a gray screen with a black fixation cross in the middle of the screen. The subject was instructed to focus on the fixation cross while listening to the tones and to press button 1 each time they heard a deviant tone.

The functional scans were T2*-weighted gradient echo EPI sequences, with TR = 2, flip angle 90 degree, acquisition matrix 64x64, 4 mm thick slices with 1 mm gap, oblique axial anterior commissure(AC) - posterior commissure (PC) aligned. With this 140 acquisitions were made.

Each scan session consisted of a brief training session to familiarize the subject with the paradigms, placement in the scanner for about one and a half hours during which structural and functional images were collected. These images are stored in analyze format and are available at FBIRN site.

## IV. Classification

### *A.* Spatial Pre-processing

#### *1) Realignment:*

The goal of motion correction which is also known as realignment is to reduce the misalignment between images in an fMRI time series that occurs due to head motion [2]. This writes all realigned images into the directory where the functional images are stored. It eliminates unimportant voxels; registers the images and then computes available voxels followed by reslicing. SPM then plots the estimated time series of translations and rotations. This data gives us the regressors which are useful while fitting GLMs.

In this step, each successive scan of a subject is spatially registered to usually the first or the average scan, using a rigid body transformation model.

#### *2) Coregistration:*

It is essential to check the quality of any registration operation, including coregistration and spatial normalization. SPM then implements a coregistration between the structural and functional data that maximizes the mutual information.

#### *3) Segmentation:*

This method combines spatial normalization and bias field correction with tissue segmentation, so that the prior probability that any voxel contains gray or white matter can be determined using a probabilistic atlas of tissue types; this prior probability is then combined with the data from the image determine the tissue class [2].

In a, pre-processing step called spatial normalization, the brain scans are spatially matched to a standard anatomical template image usually using an affine transformation.

#### 4) *Smoothing:*

There are number of reasons to apply spatial smoothing to fMRI data [2].

i. By removing high frequency information i.e. small scale changes in the image, smoothing increases the signal-to-noise ratio for signals with larger spatial scales.
ii. When data are combined across individuals, it is known that there is variability in the spatial location of functional regions that is not corrected by spatial normalization. By blurring the data across space, spatial smoothing can help reduce the mismatch across individuals, at the expense of spatial resolution.

In this step, data is smoothed using a Gaussian kernel, to decrease the effect of inter-subject variation of the anatomical regions in the brain.

### B. *Model Specification, Review and Estimation:*

In SPM, with the help of specify 1st –Level functionality, calls specification of fMRI which in turn then plots the design matrix. To check, model specification, SPM provides Review facility.

### C. *Estimation:*

It implies that the interest is in estimating the precise shape of the hemodynamic response associated with the task referred to as estimation. In SPM, after clicking on estimate button, SPM writes a number of files into the selected directory including SPM.mat file.

### D. *Linear Discriminative Analysis (LDA)*

Linear discriminative analysis is also known as the Fisher discriminative, named for its inventor, Sir R. A. Fisher [4].

LDA is also closely related with principal component analysis (PCA) and factor analysis in which they both look for linear combinations of variables that best explains the data.

The principal objective of LDA is, to reduce dimensionality while preserving as much of the class discriminating information as possible.

Discriminative analysis is a classification method. It assumes that different classes generate data based on different Gaussian distributions. It is used to find the linear combination of features which separates two or more classes. This resulting combination may be used as a linear classifier.

The classification of fMRI data with linear discriminative analysis proceeds as follows:

Step 1] Initially, Load the fMRI data.

Step 2] To train a classifier, the fitting function estimates the parameters of a Gaussian distribution for each class. This function has two arguments; first argument is array containing multiple regression value and left template values.

Step 3] To classify fMRI data with average measurements.

Step 4] Once, the linear classifier is created; we can calculate the average measurements and can proceed for predicting the mean class.

Step 5] Same steps can be followed in order to create quadratic classifier, but here we will have to specify the 'discrimType', as 'quadratic'.

Classification using fitting function works such that:

i. For linear discriminative analysis, the model has the same covariance matrix for each class; only the means vary.
ii. For quadratic discriminative analysis, both means and covariance's of each class vary.
iii. For linear discriminative analysis, it computes the sample mean of each class. Then it computes the sample covariance by first subtracting the sample mean of each class from the observations of that class, and taking the empirical covariance matrix of the result.
iv. For quadratic discriminative analysis, it computes the sample mean of each class. Then it computes the sample covariance's by first subtracting the sample mean of each class from the observations of that class, and taking the empirical covariance matrix of each class.

Step 6] This step is responsible for creating and visualizing the classifier. Plot the data, showing classification.

Step 7] Plot the classification boundaries with the help of coeffients of linear classifier.

Step 8] Step 6 and Step 7 can also be repeated for quadratic classification.

Step 9] This step examines the resubstitution error and creates the confusion matrix.

The resubstitution error is the difference between the response training data and the predictions the classifier makes from the response based on the input training data. If the resubstitution error is high the predictions of the classifier cannot be good. However, having low resubstitution error does not guarantee good predictions for new data.

Findings in [5] shows that the prediction accuracy decreases as the complexity of the experimental task increases.

Resubstitution error is often an overly optimistic estimate of predictive error on new data.

The confusion matrix shows how many errors, and which types, arise in resubstitution. When there are K classes, the confusion matrix R is a K-by-K matrix with: R (i, j) = the number of observations of classes i that the classifier predicts to be of class j.

The confusion matrix shows that:
R (1,:) = [10 6] means obj classifies 10 Cerebrum data correctly;
R (2,:) = [3 13] means obj classifies 13 Temporal Lobe fMRI data correctly.

Step 10] Typically, discriminative analysis classifiers are robust and do not exhibit overtraining when the number of predictors is much less than the number of observations. Nevertheless, it is good practice to cross validate your classifier to ensure its stability.

Step 11] Cross validating a discriminative analysis qualifier consists of creating a quadratic discriminative analysis for the data; followed by finding the resubstitution error. The resubstitution error found to be 0.0313.

Step 12] Find the cross-validation loss for the model, meaning the error of the out-of-fold observations; which was found to be 0.1250.

## V. Results

Analysis confirms that, for Multiple regression value and parameters for left template; the resubstitution error of the linear discriminative analysis classifier for the fMRI data examined to be 0.1875 having 78.13% of accuracy; whereas resubstitution error for a quadratic discriminative analysis classifier for the fMRI data was 0.0313 having accuracy of 96.88%.

It is also analyzed that, for Multiple regression value and parameters for right template; the resubstitution error of the linear discriminative analysis classifier for the fMRI data examined to be 0.2188 having 81.25% of accuracy; whereas resubstitution error for a quadratic discriminative analysis classifier for the fMRI data was 0.0313 having accuracy of 96.88% that is same as that of linear classifier.

Cross-validation loss for the model, meaning the error of the out-of-fold observations was 0.1250.

Table I describes results obtained for fMRI images using LDA with Multiple Regression Value and Parameters for left template.

Table 1: Performance of LDA with multiple regression analysis and left template

| Classification Technique | Kernel Function | Resubstitution Error | Accuracy (%) |
|---|---|---|---|
| LDA | Linear | 0.1875 | 81.25 |
| | Quadratic | 0.0313 | 96.88 |

Table 2 describes results obtained for fMRI images using LDA with Multiple Regression Value and Parameters for right template.

Table 2: Performance of LDA with multiple regression analysis and right template

| Classification Technique | Kernel Function | Resubstitution Error | Accuracy (%) |
|---|---|---|---|
| LDA | Linear | 0.2188 | 78.13 |
| | Quadratic | 0.0313 | 96.88 |

### *A. Graphical Representation*

In this section, experimental results of the proposed technique are shown in graphical form for auditory fMRI data using different kernel functions in LDA.

Comparison between algorithms is made on the basis of accuracy.

For graphs of auditory fMRI images; red points denotes data points for activations in Temporal Lobe and black points denotes data points for activations in Cerebrum of a subject.

Figure 2 given below, shows classification of fMRI data using linear discriminative analysis (LDA) with multiple regression value and parameters for left template; which predicts components belonging to classes Cerebrum and Temporal Lobe.

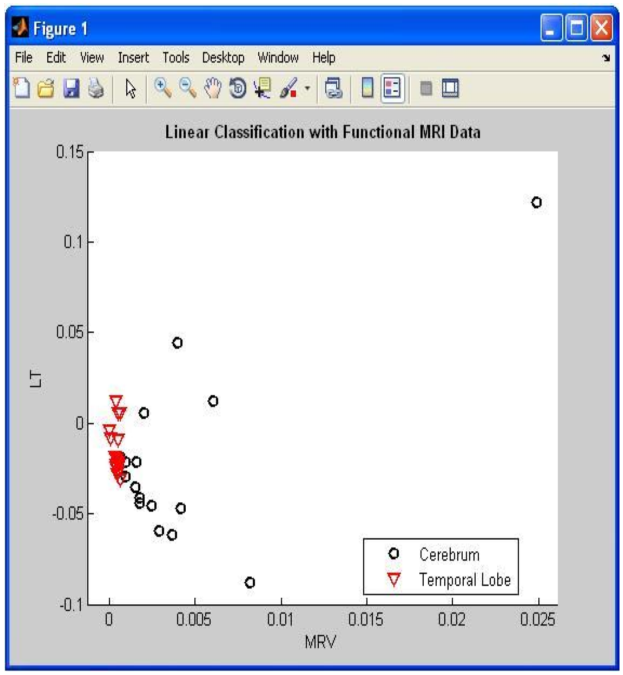


Figure 2: Classification of fMRI data of multiple regression (MR) values and Left Template parameter values with linear discriminative Analysis (LDA).

Following figure 3 shows, classification of fMRI data using quadratic discriminative analysis (QDA) with multiple regression value and parameters for left template;

which predicts components belonging to classes Cerebrum and Temporal Lobe.

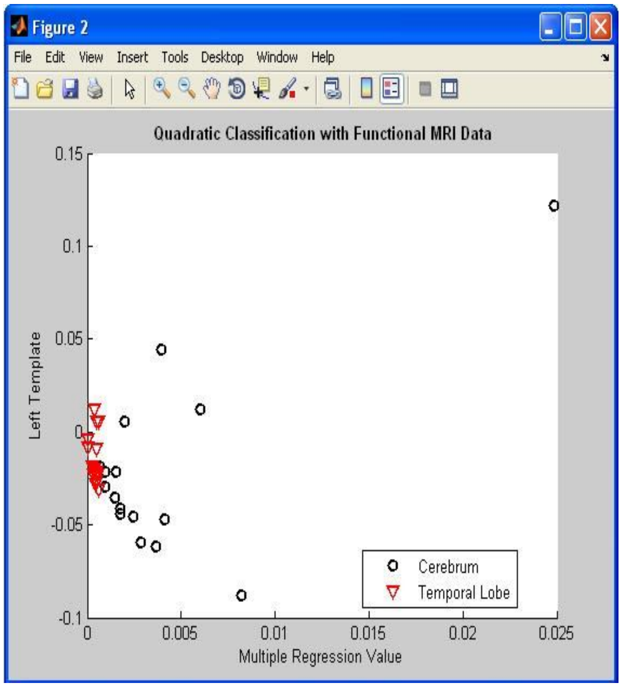


Figure 3: Classification of fMRI data of MR values and Left Template parameter values with Quadratic Discriminative Analysis (QDA.).

Figure 4 shows classification of fMRI data using linear discriminative analysis (LDA) with multiple regression value and parameters for right template; which predicts components belonging to classes Cerebrum and Temporal Lobe.

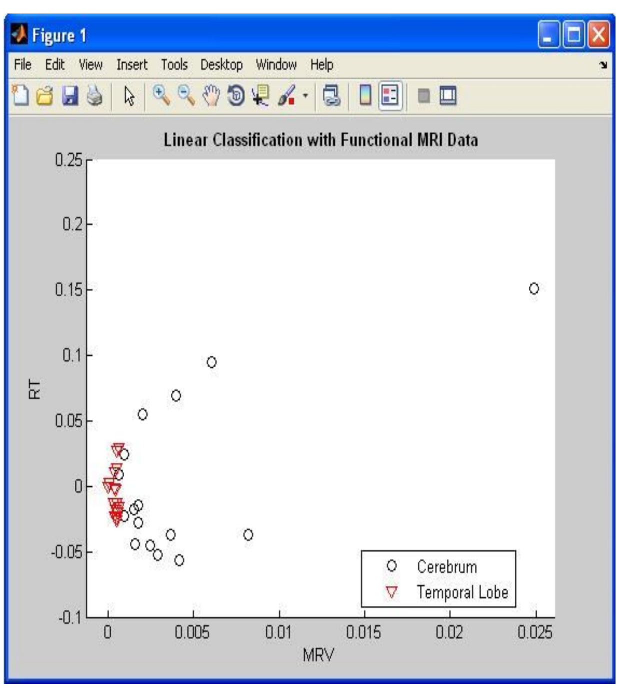


Figure 4: Classification of fMRI data of MR values and Right Template parameter values with linear discriminative Analysis (LDA).

Figure 5 shows classification of fMRI data using quadratic discriminative analysis (QDA) with multiple regression value and parameters for right template; which predicts components belonging to classes Cerebrum and Temporal Lobe.

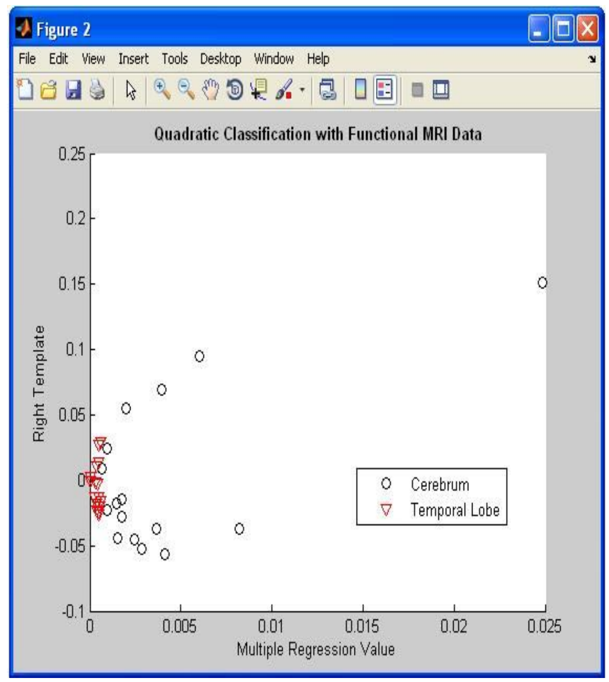


Figure 5: Classification of fMRI data of MR values and Right Template parameter values with Quadratic Discriminative Analysis (QDA).

## VI. CONCLUSION

In this paper, automatic classification technique is developed to classify activations in specific regions between normal and abnormal images. It is assumed that activations are in mainly two areas of the brain considering the kind of data; which in our case is auditory fMRI data. One area is Cerebrum; since its functions are movement, sensory, processing language, and thinking and memorizing, second area is, Temporal Lobe; since this area of the brain is involved in "Main Auditory Perception".

Initially the auditory fMRI data is pre-processed in SPM.

For auditory fMRI data classification with linear discriminative analysis with linear kernel function for multiple regression value and parameter for Left template gives accuracy of 81.25%; whereas LDA with same kernel function for multiple regression value and parameter for Right template gives accuracy of 78.13%.

For auditory fMRI data classification with quadratic discriminative analysis with quadratic kernel function for multiple regression value and parameter for Left template as well as for Right template gives accuracy of 96.88%.


## ACKNOWLEDGMENT

Author wishes to thank, Department of Computer Science and IT, Dr. Babasaheb Ambedkar Marathwada University, Aurangabad. Author also wishes to thank, the authorities of Biomedical Informatics Research Network for providing the fMRI data under the following support: U24-RR021992 Functional Imaging Biomedical Informatics Research Network (FBIRN).


## REFERENCES


[1] H. Mao and G. S. Berns, “MRI in the study of brain functions: clinical perspectives”, MedicaMundi, April 2002.

[2] Russell A. Poldrack, Jeanette A. Mumford, Thomas E. Nichols, “Handbook of functional MRI Data Analysis”, Cambridge Press, ISBN: 978-0-521-51766-9, 2011.

[3] Functional Imaging Biomedical Informatics Research Network (FBIRN): http://fbirnbdr.nbirn.net:8080/BDR/

[4] MATLAB R2012a Help Manual.

[5] S. Strother et al, “The quantitative evaluation of functional neuroimaging experiments: The NPAIRS data analysis framework”, NeuroImage, 15(4):747-771, 2002.