# Identification of functional neuroimaging markers for epilepsy using fMRI


By

**Atif Riaz**
**2011-NUST-MS-CS-36**

Supervised By
**Dr. Kashif Mahmood Rajpoot**
**Assistant Professor**
**NUST-SEECS**

This thesis is submitted in partial fulfillment of the requirements for the degree of Masters of Science in Computer Science (MS CS)

School of Electrical Engineering and Computer Science,
National University of Sciences and Technology (NUST),
Islamabad, Pakistan.

(May 2014)

# Approval

It is certified that the contents of this thesis titled "**Identification of functional neuroimaging markers for epilepsy using fMRI**" submitted by **Atif Riaz** have been found satisfactory for the requirement of the degree.

Advisor: ___Dr. Kashif Mahmood Rajpoot_

Signature:_____

Date: _____

Committee Member 1: __Dr. Nasir Rajpoot ___

Signature _____

Date _____

Committee Member 2: __Dr. Sohail Iqbal ___

Signature _____

Date _____

Committee Member 3: __Dr. Shahzad Younas_

Signature _____

Date _____

**To my parents, family and kids.**

# CERTIFICATE OF ORIGINALITY

I declare that the work titled **"Identification of functional neuroimaging markers for epilepsy using fMRI"** is my own work to the best of my knowledge. It contains no materials previously published or written by another person, nor material which to a substantial extent has been accepted for the award of any degree or diploma at National University of Sciences and Technology (NUST) or any other education institute, except where due acknowledgment, is made in the thesis. Any contribution made to the research by others, with whom I have worked at NUST SEECS, University of Warwick or elsewhere, is explicitly acknowledged in the thesis.

I also declare that the intellectual content of this thesis is the product of my own work, except to the extent that assistance from others in the project's design and conception or in style, presentation and linguistic is acknowledged. I also verified the originality of contents through plagiarism software.

Author Name: Atif Riaz

Signature: _____

# Acknowledgement

I am grateful to Allah Almighty who gave me courage and strength to complete this thesis. I am really thankful to Dr. Kashif Mahmood Rajpoot for his kind attention, precious time and guidance during this thesis. I specially thank Dr. Nasir Rajpoot for providing data and continuous supervision and guidance in this work. I am also thankful to my worthy committee members, Dr. Shahzad Younas and Dr. Sohail Iqbal, for their support and becoming a part of this work. I am gratified to Dr Waqas Majeed, LUMS, for his guidance in this work.

I am thankful to SEECS management for providing excellent environment and facilities and all those who have helped me.

# ABSTRACT

Identification of functional brain connectivity differences induced by certain neurological disorders from resting state functional magnetic resonance imaging (rfMRI) is generally considered a difficult task. This challenging task requires the identification of discriminant neuroimaging markers. In this work, we propose a two-stage algorithm to identify functional connectivity differences that can discriminate epileptic patients and healthy subjects. In the first stage, we determine the functional connectivity matrix between brain cortical regions for identification of potentially discriminant neuroimaging markers using a novel affinity propagation clustering method. Next, we propose a difference statistic to select the most discriminant connections between the cortical regions. Using selected connections and a support vector machine classifier, we achieve classification accuracy of 93.08% (specificity: 91.1%; sensitivity: 95.4%) on unseen dataset. We find that default mode network is impaired the most in epileptic subjects as compared to other resting state networks. The results demonstrate that the proposed algorithm is capable of determining functional connections between brain regions which aid in discrimination of epileptic patients versus healthy subjects. The methodology is expected to have broad applications for classification of other neurological diseases also.

# List of Contents

# List of Figures

# List of Tables


# Chapter 1

# Introduction

This chapter introduces motivation and contribution of the thesis. It also introduces the document breakdown in terms of its chapters.

# 1.1 Motivation

Neurological disorders have emerged as one of the greatest threats to the human health [2]. After cardiac diseases, neurological disorders are ranked as the biggest threat to human health [1]. The neurological disorder studied in this work is epilepsy. Epilepsy is a neurological disorder that affects the nervous system. Epilepsy causes disorder in brain function: neural activity behaves abnormally and abnormal signals cause seizures. Epilepsy is also known as seizures disorder and is characterized by these repeated and unpredictable seizures. About 50 million people suffer from epilepsy and nearly 80% of seizure cases occur in developing countries [2]. Risk of epilepsy is correlated with age: aged people have high risk of epilepsy as compared to young ones. While medication is often used to control seizures, epilepsy is not fully curable. These medicines are known as anti-epileptic drugs (AEDs). Even with the best available AEDs, epilepsy is not fully cured and seizure is not controllable in about 30% cases and the usage of AEDs is life-long. Diagnosis of epilepsy is critical as other neurological diseases may also share similar symptoms. Various diagnosis tests are used for detection of epilepsy, a few are described below.

**Electroencephalogram (EEG):** EEG measures electrical activity of the brain. For diagnosis of epilepsy, EEG should be conducted within 24 hours of a seizure, though EEG is not always reliable. Often, repeated EEGs may be needed to confirm a diagnosis.

**Computerized tomography (CT) scans:** CT can detect anatomical anomalies like tumors. The tumors can also be responsible for abnormal electrical activity of brain i.e. seizures. CT has harmful effects due to high exposure of X-ray radiations to the body.

**Positron emission tomography (PET):** PET is a medical imaging technique that uses radioactive material as source to map functional activity in brain. PET is considered to have harmful effects due to exposure of radioactive material to the body.

**Magnetic resonance imaging (MRI):** MRI is well suited for soft brain tissues of brain and is considered as less harmful than CT or PET scan. It is employed to reveal anatomical details of the brain.

**Functional MRI (fMRI):** fMRI is widely employed for studying functional activity of brain. Studies have effectively employed fMRI for studying brain function and neurological diseases and their effects on functional activity of the brain. In research community, it is also being employed for diagnosis of neurological disorders like epilepsy, schizophrenia. The main advantages of fMRI are:

i. Noninvasive.
ii. Avoids harmful radiations.
iii. Comparatively high spatial resolution.

fMRI experimental paradigms can be divided in following two broad categories:

**Event related fMRI** is used to study functional activity of brain in response to some predefined activity. Typically subjects are presented with some stimulus, like visualizing a photograph, or asked to perform some mathematical calculations or just move fingers. In some experiments

feedback from subjects is also required by selection through buttons, clicks, etc. Primary purpose of event related fMRI is to identify the brain networks that are active and responsible for a particular activity.

**Resting state fMRI**: In the early era of brain research, neuro-scientists believed that brain circuits are turned off when the person is at rest. However, fMRI studies have shown that brain is functionally active even when a person is not active [3]. For the first time Biswal et al. [3] studied that brain is functionally active while body is at rest. Authors demonstrated that brain regions belonging to primary motor cortex are active during rest and they have high correlation between their fMRI time series data. This domain of fMRI is termed as resting state fMRI (rfMRI). In a typical rfMRI experiment, the subject is asked to do nothing in particular, close his/her eyes and is asked not to focus at any particular thought. From the experimental point of view, resting state fMRI is comparatively easy to conduct than event related fMRI. Resting state fMRI does not require any specific experimentation setup as in event related fMRI. Event related fMRI data may be affected from various subject level attributes like random thoughts, timing mismatch etc.

A number of studies have explored functional connectivity of brain in epilepsy employed resting state fMRI. Most current research on epilepsy fMRI has been limited in scope by performing analysis of a particular brain network or region of interest with a specific category of epilepsy. In this thesis, the motivation is to identify the functional connectivity anomalies induced by different kinds of epilepsy without limiting the scope to a particular brain network. Recently Zhang et al. [1] have conducted similar study but they have employed *k*-means clustering for estimation of functional connectivity between brain regions. However, *k*-means clustering algorithm suffers from two major drawbacks: prior information about number of clusters and

arbitrary initial selection of cluster centroids. To overcome these short comings, we propose to employ a novel clustering algorithm called affinity propagation that does not suffer from these shortcomings. A feature selection algorithm called difference statistics is proposed for selection of discriminant connections.

## 1.2 Thesis contributions

In this thesis we explore the use of classification methods based upon resting state fMRI data. We have addressed the problem of identification of discriminant connections between healthy and epileptic subjects and pattern classification based upon these selected connections. This thesis presents following contributions:

(i) A novel clustering algorithm called **affinity propagation** has been employed for estimation of functional connectivity between brain regions. Main advantage of affinity propagation clustering is that the clustering algorithm does not require information about number of clusters in advance.

(ii) A difference statistic has been proposed as feature selection algorithm that can select discriminant features.

(iii) We employed support vector machines (SVM) classifier on functional data to predict subject being tested as normal or epileptic.

A part of thesis work has been published in the following peer-review conference:

- A connectivity difference measure for identification of functional neuroimaging markers for epilepsy. *IEEE EMBS conference on neural engineering, San Diego, California,* 2013 [4].

## 1.3 Thesis organization

The rest of this thesis is organized as following:

**Chapter 2** provides background information about fMRI and reviews related works on fMRI.

**Chapter 3** provides description of proposed methods in the study. We explain the methods employed for functional connectivity estimation, feature selection and classification. Affinity propagation and difference statistic have been proposed in this study for estimation of functional connectivity and feature selection.

**Chapter 4** describes the experimental setup and presents the results of the study. The chapter also presents a discussion of results.

**Chapter 5** concludes the thesis and provides future research dimensions.


# Chapter 2

# Background and literature review

This chapter presents background information related to fMRI data analysis and reviews the literature related to study of epilepsy employing resting state fMRI.

## 2.1 Background

Functional magnetic resonance imaging (fMRI) is an imaging technology that provides opportunity to study functional activity in the brain. fMRI has emerged as a popular imaging tool in neuroimaging experiments to study functional activity in the brain. Glucose, a primary source of energy for human body, is not stored in the brain. Energy required for any brain region is drawn from oxygen and glucose rich blood. The brain regions that are active for a particular time span require more energy and therefore are supplied more oxygenated blood as compared to inactive brain regions. fMRI can detect and differentiate between magnetic properties of oxygen rich and deoxygenated blood. With the help of this measure, fMRI provides an opportunity to estimate functional activities of brain.

fMRI scanning procedure is primarily based upon MRI technology. An MRI scanner comprises of a cylindrical tube with a powerful electric magnet. Typically a scanner has very high magnetic field of 3.0 Tesla (approximately 50,000 times the magnetic field of earth). Normally the atomic nuclei of brain are randomly oriented. Due to the effect of this powerful external magnetic field, these nuclei become aligned with the external field. After removing the magnetic field, nuclei re-align them to their original states, causing release of magnetic signals. The tiny magnetic signals from multiple nuclei coherently add to form strong signals which are measured by detectors and form an image after processing. In fMRI, magnetic signal from hydrogen nuclei of water ($H_2O$) is detected.

Hemoglobin in red blood cells transports oxygen to brain neurons. With the increase in neuronal activity in any brain region there is increase in demand of oxygen resulting in increase of blood flow to region of increased neuronal activity as visualized in Figure 1.

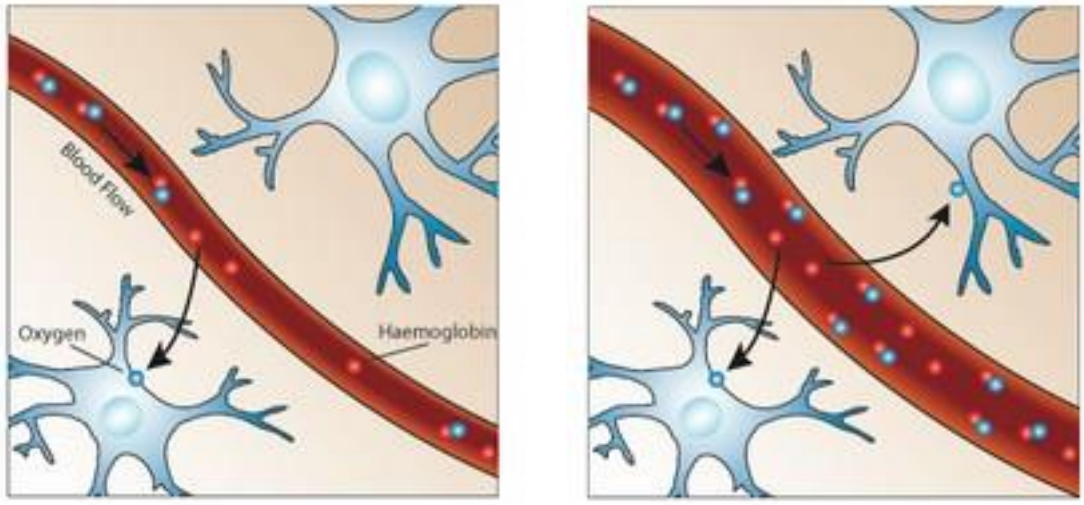


**Figure 1: Visualization of blood flow as an increase in neuronal activity (from [5]). Left image is presenting normal state and right is presenting activated state. Activated region (right image) is supplied with more blood as compared with normal state which is represented as thick blood vessel.**

Magnetic properties of oxygenated and deoxygenated hemoglobin are different: hemoglobin is diamagnetic when it is oxygenated and it is paramagnetic when deoxygenated. The differences in magnetic properties can be detected by small differences in MR signals of blood. Blood oxygenation is dependent on the neural activity, so by detecting these magnetic differences, we can estimate neural activity in brain regions. This method is known as blood oxygenation level dependent (BOLD) imaging.

A typical fMRI session constitutes of a number of time slices. MRI volume is acquired in each time slice and these 3d volumes are stacked together to generate 4d fMRI data. In one experiment, approximately 200 MRI scans or 3d volumes are acquired.

Data generated by fMRI scanner is referred as raw data and is typically affected with noise. The noise in fMRI scan can be grouped in two broad categories:

**System noise:** This is the noise introduced by experimental setup. Thermal noise is associated with temperature of apparatus. With the increase in temperature, heat causes to attenuate electron movement which distorts the current in fMRI detector. Thermal noise depends upon the range of the frequencies detected by receiver of the system. All voxels are equally affected by this noise. Scanner drift is another source of noise which is caused by drift in strength of the static magnetic field.

**Physiological noise:** Noise introduced by subject itself lies in this category. Main source of noise in this category is head movement. fMRI is time consuming and it is difficult to maintain head in same position for the whole course of scan. Even with carefully administered subjects, head motion is unavoidable and is introduced due to effects of breathing, heart beats, tension and nervousness. In some cases, depending upon experiment tasks, noise is introduced due to physical movements like pressing or clicking of some buttons in order to provide feedback. Even with the best possible training of the subjects, noise is introduced, due to certain inherent factors like random thoughts, scanner noise, different physical sensations, etc.

The introduced noise can affect the analysis of fMRI data. For example, head motion during scan can change region to signal mapping so that particular voxel or region may incorrectly be linked with other signal. In order to mitigate the effects of noise, preprocessing of the raw data is essential prior to analysis.

### 2.1.1 Preprocessing

The scanner generates 3D volume of MRI scans at multiple time points. In total, $t$ volumes of brain are scanned and joined together to form 4D scan of a session per subject where $t$ is the number of time intervals per session. Prior to analysis of the data, preprocessing is applied on this 4D data. Main steps involved in pre-processing are listed here.

#### I. Slice timing correction

A variety of scanners may acquire MRI slices in interleaved manner (1,3,5 … and 2,4,6,…) instead of sequential order in order to avoid contaminating neighboring slices. In such scenarios, incorrect placements of slices may alter analysis as each slice might have different state of activations. In this step, slices are placed in correct order.

#### II. Head motion correction

Head motion can introduce a major artifact in fMRI data. Error can be induced by even a very small head movement. Head motion correction is vital in preprocessing as head movement cannot be prevented in experimentation setup. The approach typically adopted involves applying rotation and translation of head part within image. Depending upon the algorithm implemented, a particular slice (first, last, middle or some other) is taken as reference image. In most of the software, first slice is taken as reference and all other images are aligned to this reference. In these scenarios, rotations and translations are applied to target image and sum of the squares of the difference (SSD) between pixels in reference and target image is calculated. By varying parameters, multiple translations and rotations are calculated to attain minimum SSD. The technique is commonly used in various softwares.

## III. Spatial smoothing

Random noise in image might cause errors in analysis phase. Spatial smoothing is carried out to smooth the image and get rid of random noise. Spatial smoothing improves signal to noise ratio but can reduce the image resolution also. A set of balanced parameters is required to balance this trade off. Typically a matrix of dimensions $3 \times 3$ is employed for smoothing (Figure 2).

$$\frac{1}{9} \times \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

**Figure 2: Filter for spatial smoothing.**

The filter can applied to the whole image by convolution as:

$$\boldsymbol{g = h * f} \tag{2-1}$$

where $*$ represents convolution operator , $h$ represent the mask, $f$ is the source image and $g$ is the resultant image.

## IV. Temporal smoothing

fMRI data can be interpreted as $n$ regions or voxels (depending upon mode of analysis, voxel based or region based) having time series of $t$ points. Similar to spatial smoothing, smoothing in time domain also improves the signal to noise ratio. The time series signal is determined by BOLD signals, which is based upon blood flow. The rate at which the signal changes is limited as blood flows in a moderate manner, applying temporal smoothing will improve signal to noise ratio. Typically a 1D Gaussian filter is applied to all time series to attain smoothing, which is expressed as:

$$G(x) = e^{-\frac{x^2}{2\sigma^2}} \quad (2\text{-}2)$$

where $\sigma$ is the standard deviation. The filter can also be illustrated by the following figure:

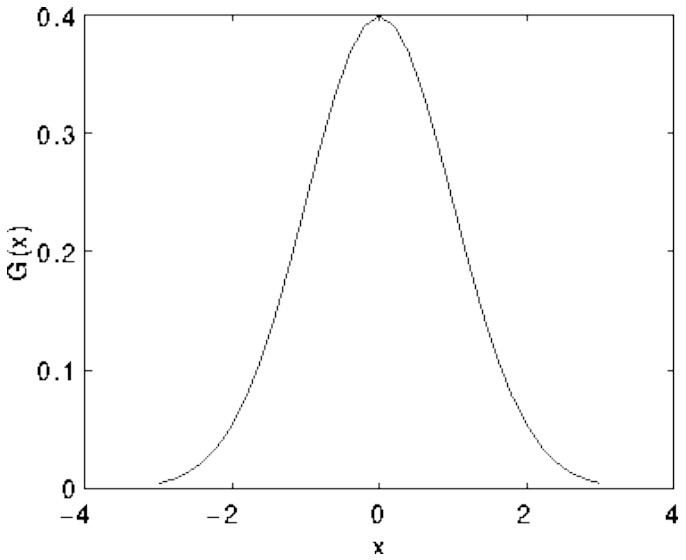


**Figure 3: Gaussian 1D filter with mean=0 and sigma=1.**

The effect of applying temporal smoothing can be visualized from Figure 4 and Figure 5. It can be seen that abrupt fluctuations are removed by filtering.

Preprocessed fMRI data is utilized for any subsequent analysis.

## 2.1.2 fMRI Analysis

The primary goal of fMRI is to analyze activity in local areas of the brain. The analysis is typically carried out through two approaches:

**Voxel based analysis:** In this approach, voxel wise functional activity of brain is studied [6]. Voxel is the smallest addressable volume element.

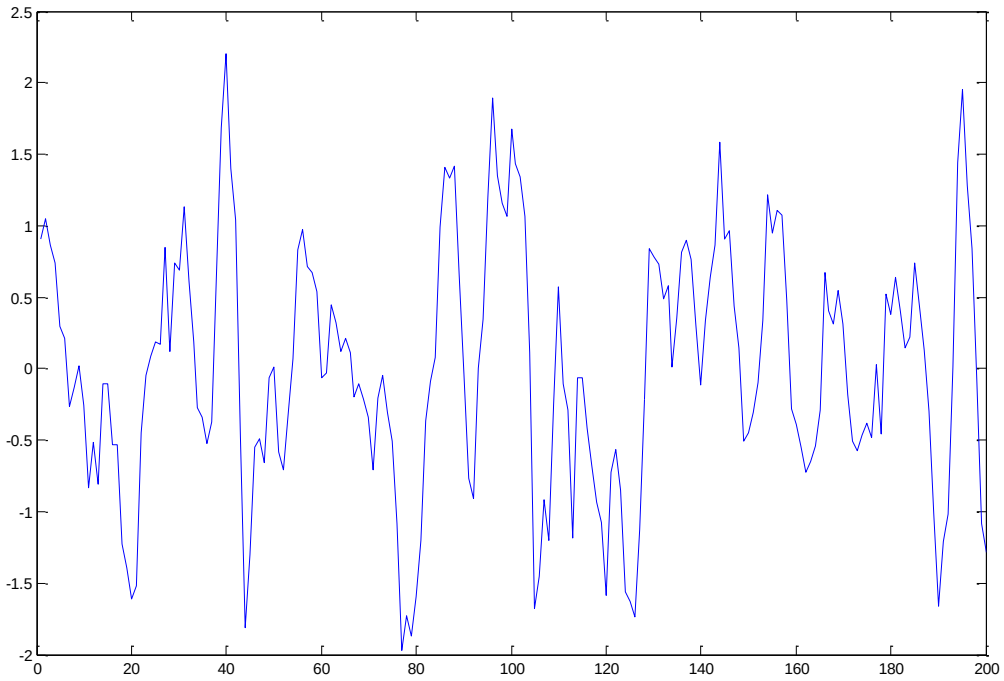


**Figure 4: Temporal signals of a region.**

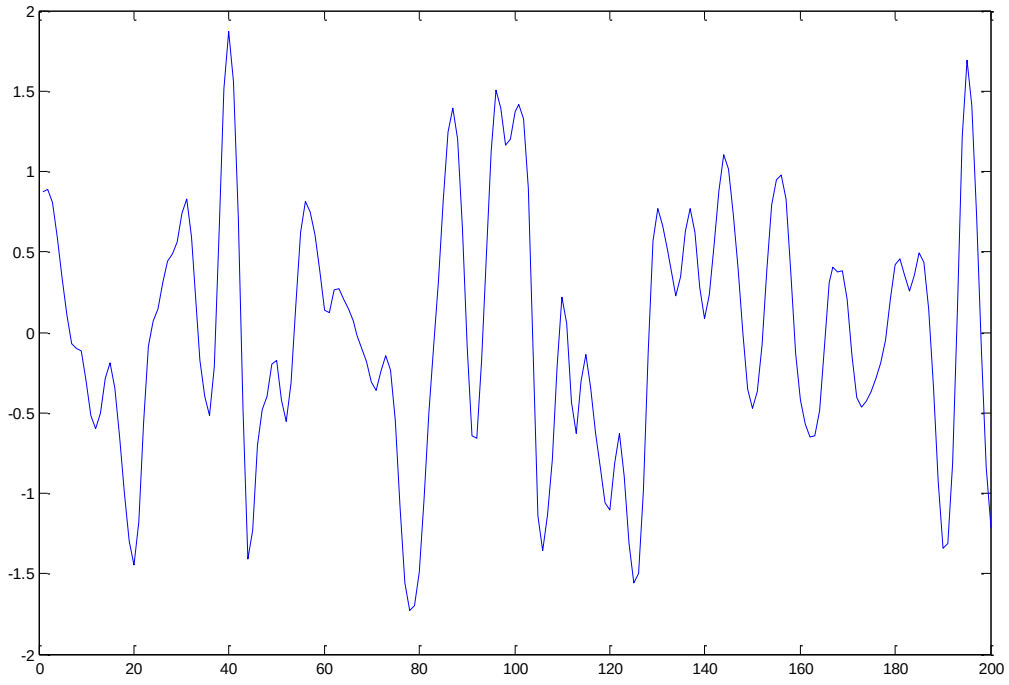


**Figure 5: Effect of applying temporal smoothing to signal of Figure 4.**

**Region based analysis:** In these methods, brain is segmented into specific predefined regions and functional activity of these regions is studied. Segmentation of brain regions is typically carried out in the following way.

i. **Anatomical locations based segmentation**

In anatomical segmentation, brain is segregated into regions based upon their anatomical locations. Typically predefined templates are used for segmentation, like in automatic anatomical labeling (AAL) template [7] where whole brain cortex is segmented in 116 regions.

ii. **Functional group segmentation**

In functional group segmentation, regions are identified based upon their functionality. Examples include visual cortex network, where all brain regions belonging to this network are segregated in one network.

### 2.1.3 Functional connectivity

Functional connectivity is often estimated by temporal correlations between brain regions. Human brain can be envisaged as a complex and efficient network [8] where nodes represent brain regions and edges represent their connectivity. Each region has its own task and different regions also share information between each other. Any two or more regions that show similar pattern of activity are assumed to be functionally connected. Compared with other neuro-imaging techniques like positron emission tomography (PET) and magnetoencephalographic (MEG), fMRI being non-invasive and high spatial resolution is considered most suitable towards determining functional connectivity [9]. In recent studies analysis of functional connectivity of brain regions is considered very important as it is assumed that the connectivity plays a key role

in cognitive processes of brain [8]. The functional connectivity might also be important for understanding complex organization of the brain. The connectivity analysis methods can be divided into two main types: model based methods and data driven methods [9]. Various methods are categorized and presented in Figure 6.

# I. Model based methods

In studies employing model based methods some predefined regions are selected which are also called seeds or regions of interest (ROIs), certain metrics are employed to determine whether other regions are functionally linked to these ROIs or not. All the linked regions are considered to be functionally connected. Selection of predefined regions is typically based upon some strong prior neuroscience knowledge. For example, if some regions are known to be in a visual network, these regions will be used as seed or ROIs and all other regions will be probed for functional connectivity with the group being studied. Different metrics can be employed for analyzing connectivity and model based methods can be sub categorized based upon these metrics.

## A. Cross correlation analysis

Cross correlation analysis (CCA) is well known measure for symmetry of two signals in digital signal processing domain. Cao et al. [10] introduced cross correlation in fMRI. If BOLD time series of two brain regions have high correlation coefficient, they are assumed to be functionally connected with each other. For any two brain regions $x$ and $y$ their cross correlation is calculated as:

$$Corr_{x,y} = \frac{Cov_{x,y}}{\sqrt{Var(x) \times Var(y)}} \tag{2-3}$$

where $Var(x)$ and $Var(y)$ is variance of $x$ and $y$ and $Cov$ is cross variance.

## B. Coherence analysis

CCA has been extensively employed for functional connectivity analysis but it can be prune to drawbacks. For example, two regions having no blood flow can depict high correlation value. Physiological noises in signal like cardiac activity can also mislead the analysis by having a high correlation value. A new metric measure called coherence was introduced by Sun et al. [11] to avoid such errors. Coherence between two regions $x$ and $y$ is calculated as:

$$Coh_{x,y}(\lambda) = \frac{|F_{x,y}(\lambda)|^2}{F_{x,x}(\lambda)F_{y,y}(\lambda)} \tag{2-4}$$

where $F_{x,y}(\lambda)$ is the cross spectrum defined by the Fourier transform of cross covariance as:

$$F_{x,y}(\lambda) = \sum_u Cov_{x,y}(u) \times e^{-j\lambda u} \tag{2-5}$$

## C. Statistical parameter mapping

Statistical parameter mapping [12] is a model based methods that has been employed to study functional connectivity between brain regions. The method averages brain regions in certain seed or ROI and are compared to all other regions and probed for functional connectivity.

# II. Data-driven methods

Main draw-back of model driven methods is that they require prior knowledge of neuroscience and pre selection of seed or ROI. These limitations can be overcome by data driven methods. These methods do not require prior neuroscience knowledge about brain regions. These methods are explorative methods and provide an opportunity to explore functional connectivity of whole brain. There are two sub categories of these methods.

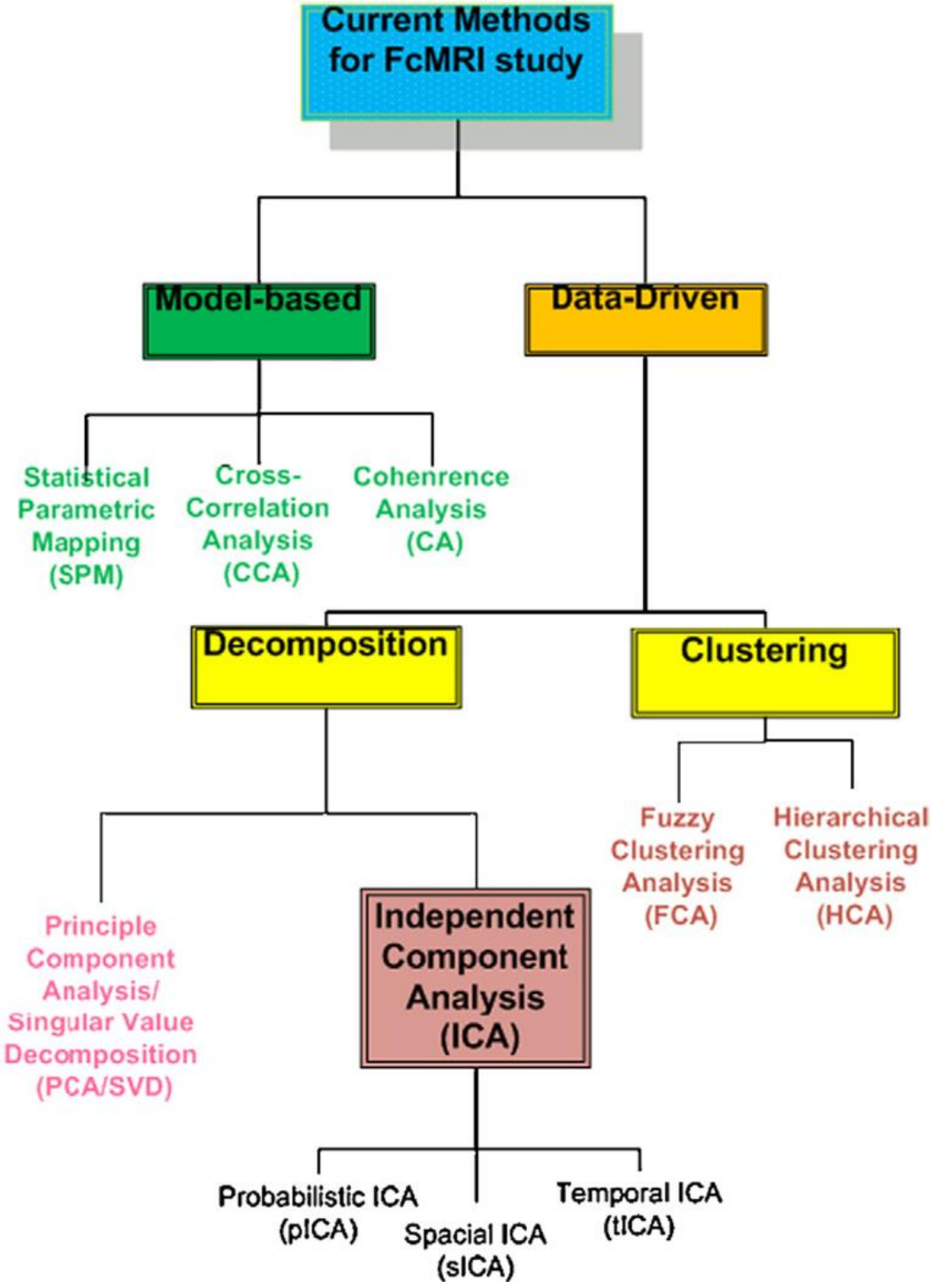


Figure 6: Methods for functional connectivity (from [6]). FcMRI represents functional connectivity from fMRI.

### A. Decomposition based methods

These methods are based upon decomposition techniques that decompose the fMRI signals into sub groups. Techniques like principal component analysis and independent component analysis fall in these methods. These methods try to represent the fMRI dataset as linear combination of basis vectors.

### B. Clustering based methods

In these methods, traditional clustering algorithms like k-means clustering, fuzzy clustering and hierarchical clustering are applied on fMRI data sets. Clustering can be based upon different metrics like distance measure, correlation value etc. Typically, the brain regions that are grouped in same cluster are assumed to be functionally connected.

## 2.2 Literature review

Over the years, research community around the globe has performed many interesting and useful research investigations for studying brain functionality that depend mainly on the fMRI data analysis. Researchers have applied a number of techniques in analyzing the resting state fMRI data.

### 2.2.1 Effect of lateralized epilepsy on default mode network

Haneef et al. [13] have studied default mode network (DMN) with lateralized temporal lobe epilepsy. DMN is related with resting state cognition and conscious. Authors have performed resting state connectivity analysis using region of interest (ROI) to compare DMN in temporal lobe epileptic subjects to normal subjects. Possibility of involvement of DMN in temporal lobe epilepsy (TLE) has been indicated by some prior studies [14, 15]. Reduction in blood flow to the

DMN in seizures has been revealed by SPECT and EEG-fMRI has shown decreased electrical activity within DMN. In this study, authors have employed ROI-based functional connectivity MRI (fcMRI) to better understand the impact of TLE on DMN. FSL (fMRIB software library) [16] is employed for pre-processing of fMRI data. Important steps of pre-processing include head motion correction, non-brain tissue elimination, regression and spatial smoothening. Group analysis was performed using ANOVA and SPSS (IBM) was used for analysis of group differences in the demographic and epilepsy related measures. Results of the study include:

i. Regardless of lateralization, results found reduced connectivity between posterior DMN to anterior portions of DMN.
ii. Left TLE indicated increased connectivity of the posterior DMN to bilateral opercular areas.
iii. Right TLE indicated reduced connectivity of posterior DMN to adjacent areas.

### 2.2.2 Effect of epilepsy on dorsal attention network

Zhang et al. [17] have studied alteration of dorsal attention network (DAN) in patients with temporal lobe epilepsy. Study involving behavioral data have suggested decline of attention in mesial temporal lobe epilepsy (mTLE) [18]. Abnormality within the frontal lobe detected by positron emission tomography (PET) has indicated attention impairments [19]. Previous fMRI studies on mTLE have analyzed deficits of memory and language but alteration of the attention network has not been studied before. Authors have employed independent component analysis (ICA) based resting state fMRI studies on dorsal attention network in mTLE patients. Also, scores of trail making test (TMT) in the patients were correlated to z-scores of the voxels within DAN to validate the results. Statistical parametric mapping2 (SPM2) [12] is used as a

preprocessing tool. Main steps in preprocessing are: slice timing correction, head motion correction, normalization using SPM template and smoothing. After preprocessing, independent component analysis (ICA) was applied on fMRI data. Studies have evaluated ICA-based analysis in extracting the dorsal attention network in resting state fMRI [20, 21]. Authors have chosen 40 (1/5 the number of time-points) independent components (IC). After the ICA separation, standard template of DAN provided by Mantini et al. [22] was used to extract data related to DAN. Decreased functional connectivity in almost all regions of DAN in temporal lobe epileptic patients was observed by this study concluding the alteration of DAN in mTLE patients. This result is supported by negative correlation between the TMT scores and *z*-scores of voxels within regions.

### 2.2.3 fMRI study on epilepsy patients with depressive symptoms

Chen et al. [23] have studied and compared treatment-naïve temporal lobe epilepsy patients with depressive symptoms, patients without depressive symptoms and healthy controls. The depression level of patients was assessed with a self-rating depression scale (SDS). Previously, it is suggested that common pathogenic pathways might be shared by depressive symptoms and epilepsy [24] and these common pathways may cause the development of one disease in presence of other. Previous studies on epilepsy were conducted on patients receiving anti-epilepsy drugs (AEDs) thus there might be strong possibility of impact of these drugs on brain functions and networks. Authors have studied treatment-naïve epileptic patients to diminish such possibilities. Preprocessing of functional MRI data was carried by data processing assistant for resting state fMRI (DPRASF) [25] employing these main steps: conversion of digitial imaging and communication in medicine (DICOM) data to neuro-imaging informatics technology initiative (NIFTI) images, removal of first ten points, slice timing correction, spatial

normalization to Montreal Neurological Institute (MNI) template. Voxel comparison was performed using analysis of covariance (ANCOVA) in REST [26]. Based upon SDS scores, epileptic patients were divided into two groups: patient with depressive symptoms (SDS score over 50) and patients without depressive symptoms (remaining patients). Following results are obtained by the study:

- Hyperactivity is found in TLE patients in the prefrontal-limbic, striatal areas and anterior cingulate in patients with depressive symptoms (PDS) group.
- TLE patients of PDS group showed regional brain activity alterations and disruption of the mood regulation network at the onset of seizures.

### 2.2.4 Identification of neuroimaging markers for epilepsy

Zhang et al. [1] have investigated pattern classification approach in classification of a subject scan as healthy or epileptic. There are two main steps of the algorithm: identification of neuroimaging markers and classification based upon these neuroimaging markers. The research identifies the significantly changed functional connectivity pathways in epileptic patients. The asymmetry of brain function in certain regions is also identified, like in temporal lobe (involved in auditory functions) and limbic system (involved with emotional behavior). Research employs resting state fMRI and concludes that resting state fMRI can provide an effective way in diagnosing epilepsy and possibly other neurological diseases.

The main achievements of the research are identification of discriminant neuroimaging markers and classification. Neuroimaging markers are markers that can discriminate between two groups i.e. healthy and epileptic. There are two neuroimaging markers identified by the research.

## I. Community matrix - $K$:

Community matrix $K$ is constructed to estimate the functional connectivity between brain regions. In this work, brain is segmented into 90 regions: 45 in left hemisphere and 45 in right hemisphere. Both the hemispheres are spatially symmetrical. Correlation of these 90 regions can depict the functional connectivity between brain regions. However, the correlation yields a dense network. This high density of network typically reduces the performance of pattern classification method. To overcome this problem, authors have proposed a neuroimaging marker termed as the community matrix $K$, which is constructed from $k$-means clustering. The proposed marker can reveal the functional connectivity of brain regions in a sparse manner. $K$-matrix is a 2-D matrix of $90 \times 90$ dimensions where each entry $K(i,j)$ represents the measure of functional connectivity of $i^{th}$ and $j^{th}$ brain regions in terms of probability. Figure 7 provides visualization of the community matrix. The figure compares the community matrix with the correlation matrix. First row (a and b) shows the dense structure in correlation matrix: it is quite difficult to identify distinct features/network from the matrix. While in second row (c and d) of $k$-matrix, network is sparse as compared to correlation matrix and can yield distinct features between healthy and epileptic subjects. It can be observed that the pixels lying near the diagonal have higher values and are much brighter in healthier scan as compared to that of patient, which is not observed in the correlation matrix.

The $K$-matrix obtained has high dimensions (4005 edges for 90 brain regions). These 4005 features may include both discriminatory and non-discriminatory features which may hamper the performance of classifier. Therefore, sparse regression is applied on $K$-matrix to reduce the dimensions and extract discriminant features. The extracted features are used subsequently in pattern classification.

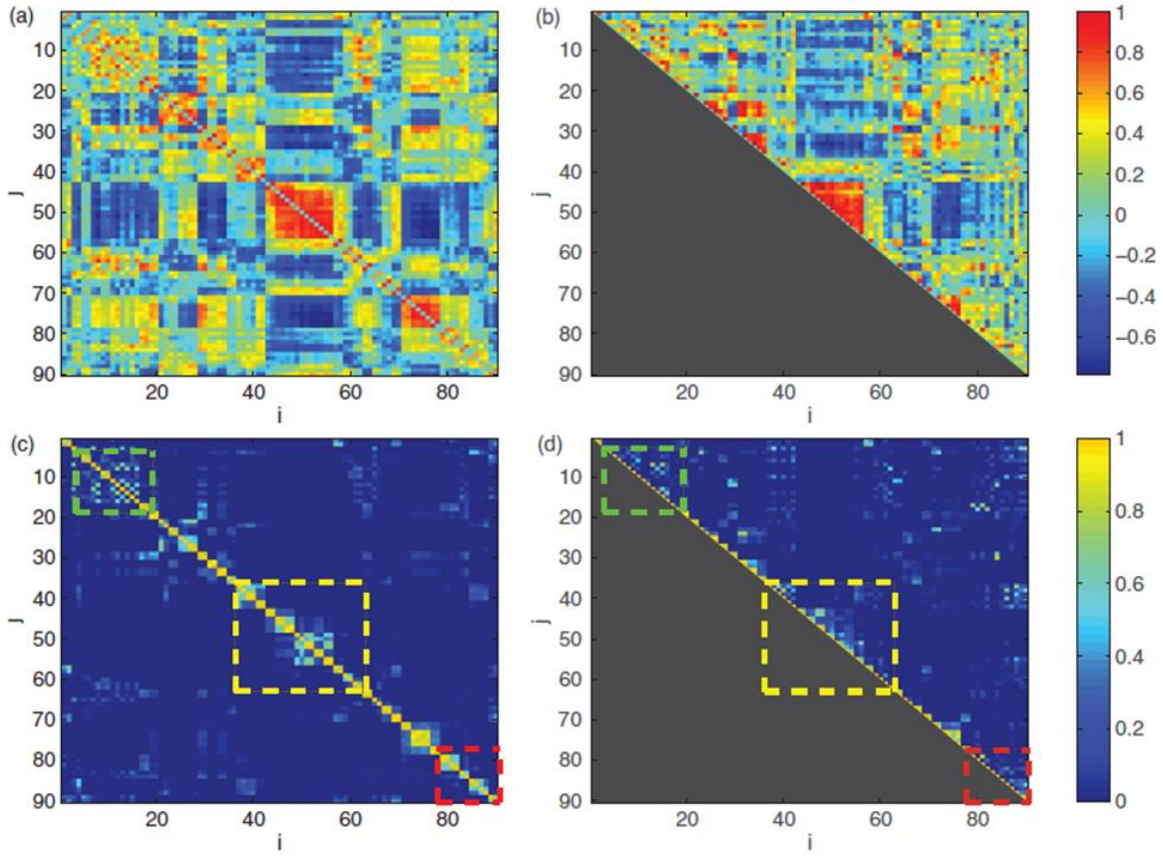


**Figure 7: Comparison of correlation and community matrices [1]. First row is correlation matrix, second row is community matrix – *K*. First column is for healthy subject and second is for epileptic.**

## II. Global connectivity asymmetry (GCA)

Brain cortex is divided into two spatially symmetrical portions: left hemisphere and right hemisphere. These two portions show marked structural symmetry but are clearly asymmetric with respect to their functionalities. Each region is responsible for different tasks. Brain is segmented in 90 regions with 45 regions in each hemisphere: there are 45 pairs of bilaterally homologous regions. Zhang et al. [1] have proposed a quantitative measure termed as global connectivity asymmetry (GCA), which is defined as the measure of dis-similarity between any

two bilaterally homologous brain regions. The dis-similarity is measured in terms of connectivity profiles of brain regions. Figure 8 illustrates the connectivity profiles of two regions.

Asymmetry index between region $i$ and $j$ is defined as:

$$\rho = 1 - \frac{\mathrm{cov}[K(i,:),K(j,:)]}{\sqrt{\mathrm{var}[K(i,:)]}\sqrt{\mathrm{var}[K(j,:)]}}. \quad (2\text{-}6)$$

i.e. the high the regions are correlated, the weaker the asymmetry between them and vice versa. This asymmetry measure is 45 dimensions for the whole cortex.

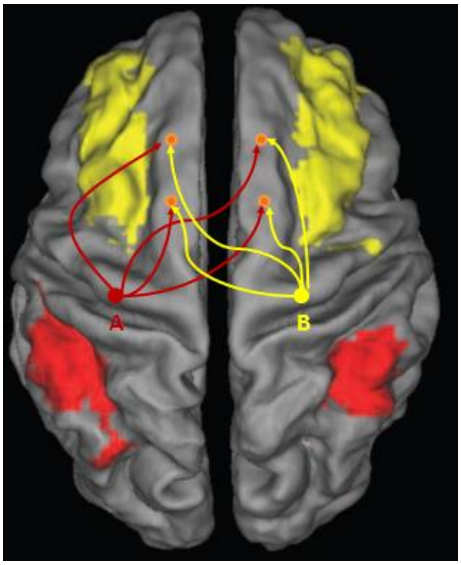


**Figure 8: Global connectivity asymmetry (from [1]). Connectivity profiles of bilaterally homologous regions A and B. For simplicity only 6 regions are shown here (out of 90 regions). Connectivity profile of region A is shown with red arrows while profile of B is drawn with yellow arrows. Global connectivity asymmetry measures is based on these profiles for bilaterally homologous regions.**

## III. Classification

Support vector machines (SVM) classifier is used as classification algorithm in this research and fifty percent of dataset is used as training and remaining for testing purpose.

Summary of accuracy achieved is:

- 77.6% accuracy with $K$- matrix (by using 400 edges selected by sparse regression).
- 75.5% accuracy with global connectivity asymmetry.
- 80.2% accuracy by both $K$- matrix and global connectivity asymmetry.

## 2.3 Summary

This chapter introduced the background information about fMRI analysis and reviewed the literature involving rfMRI for study of epilepsy. Most of the studies on epilepsy employing rfMRI have focused on a particular sub category of epilepsy or a particular network of brain. In this work, motivation is to identify functional connectivity anomalies induced by epilepsy without limiting the scope to a particular sub type of epilepsy or a particular brain region. Zhang et al. [1] conducted similar study but they employed *k*-means clustering for estimation of functional connectivity between brain regions. *k*-means clustering suffers from two major drawbacks: requirement of number of clusters in advance and arbitrary initial selection of cluster centroids. In this work, we addressed these major shortcomings and propose affinity propagation clustering algorithm that does not require number of clusters in advance and does not involve arbitrary initial selection of cluster centroids.


# Chapter 3

# Methodology

This chapter presents the methodology employed in this work. The proposed methodology can be divided into three main steps:

- Estimation of functional connectivity
- Selection of discriminant connections
- Classification

## 3.1 Functional connectivity

Functional connectivity between brain regions can be defined as temporal correlation between anatomically separate brain regions. Analysis of functional connectivity has emerged as an active domain, during the last decade. A number of studies have been presented which involve estimation of functional connectivity from neuroimaging experiments [27]. Among these studies, resting state analysis has become a dominant experimentation paradigm. Correlation is a common approach employed in neuroimaging experiments [28].

Correlation is a popular metric in domain of statistics that yields statistical relationship between two variables. Due to this property, it is effectively employed in the domain of functional connectivity analysis. In model driven approaches, as discussed in previous chapter, correlation can be employed to estimate the functional connectivity of brain regions with respect to a predefined or well established seed or ROI. However, prior strong knowledge of neuroscience is required in this approach. For functional connectivity analysis, data-driven approaches are typically preferred over model driven approaches because data driven approaches do not require prior information.

Functional connectivity analysis is often carried out via temporal correlation between brain regions [29]. In most approaches [28], correlation is applied on time series of brain regions and a 2D matrix $C$ of dimensions $n \times n$ is obtained where $n$ is the total number of brain regions defined in the template applied. The matrix $C$ is constructed such that each entry $C(i,j)$ in matrix is the correlation coefficient between region $i$ and $j$. The value of correlation coefficient lies between 0 and 1 where value of 1 can be assumed as highly functionally connected regions. An example of correlation matrices obtained by this approach is shown in Figure 9. The figure shows correlation matrices for two subjects: normal and epileptic. There are 90 brain regions defined by the template for each subject. The network thus obtained can be quite dense. This kind of dense network makes it difficult to identify the discriminant connections between the two classes. With the dense network, feature selection may result in poor selection of discriminant features. The inappropriate selection of features affects the classification accuracy yielding poor performance of the whole algorithm.

To overcome this problem, clustering under data driven methodologies can be adopted. In these methodologies, typically a matrix is constructed that reflects the functional connectivity of brain regions. For example, Zhang et al. [1] employed $k$-means clustering of brain regions to derive functional connectivity between brain regions. The $k$-means algorithm is a well-known clustering algorithm employed in a number of machine learning problems.

The main steps of $k$-means clustering algorithm are:

1. Initialize $k$ centroids by choosing $k$ data points (regions) arbitrarily.

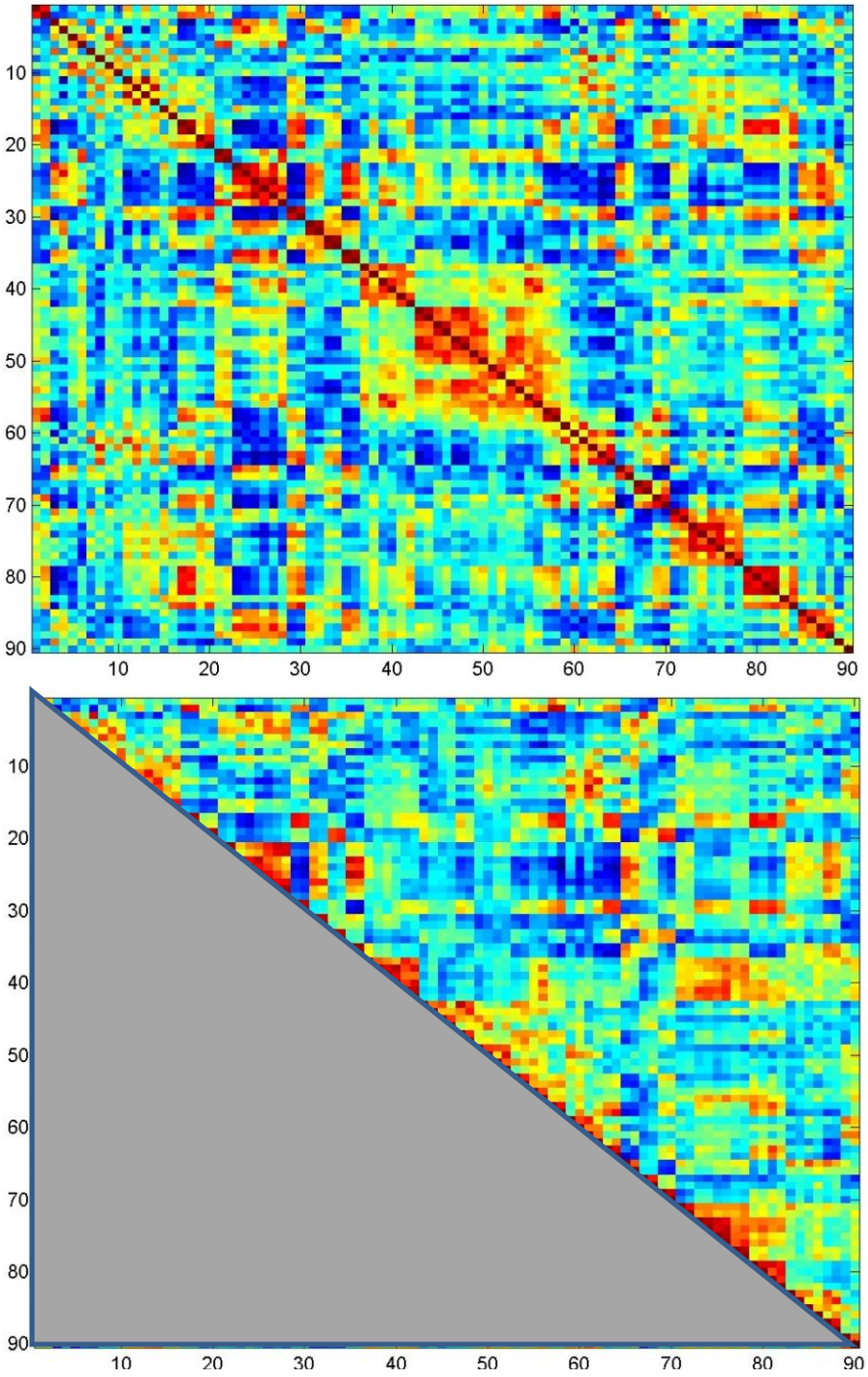


**Figure 9: Correlation matrix (from [1]). First row is of normal subject and second row is of patient subject. Each matrix is 2 dimensional with dimension of each matrix being 90 × 90 as there are 90 regions in the template.**

2. Assign all other data points to the one of the closest centroid (selected in step 1). Distance can be measured by Euclidean distance given by:

$$\boldsymbol{D} = \sqrt{(\boldsymbol{x} - \boldsymbol{y})^2} \tag{3-1}$$

where $D$ is distance between two points $x$ and $y$.

3. For each cluster, calculate its mean for making it the new centroid.
4. Repeat steps 2 and 3 until the stable centroids are achieved.

The $k$-means clustering algorithm requires information about number of clusters prior to clustering the data. In many domains, including functional connectivity analysis, number of clusters is not known in advance and typically an arbitrary number is used. One can observe the change in results of functional connectivity analysis by opting different number of initial clusters.

Another major drawback of clustering algorithms similar to *k*-means clustering is the initial selection of cluster centroids (Step 1 above). The algorithm can be quite sensitive to the initial selection of cluster centroids [30] and typically the initial clusters are chosen randomly. Typically, multiple runs of $k$-means are applied to attain stable results. Zhang et al. [1] employed 500 runs of $k$-means in order to attain stable results. However, it cannot guarantee to converge towards optimum solution. To overcome this shortcoming, we employ a recently proposed clustering algorithm called affinity propagation clustering.

Affinity propagation (AP) clustering was proposed by Frey and Dueck [30] as a novel clustering method that considers measure of similarity between each pair of data points and an initial preference given to each point as an exemplar (i.e. cluster centroid). AP clustering has been shown to perform well in multiple domains like clustering images of faces, gene selection and

others [30]. Many recent studies have effectively employed AP clustering in fMRI to study functional brain connectivity [31, 32]. Various studies have shown that AP clustering algorithm achieved low error rate compared to $k$-means [32, 33]. AP clustering has an important property that number of clusters do not need to be known in advance and there are no random centroid initializations needed.

AP clustering algorithm falls under the category of message passing algorithm. To understand basic principle of AP clustering algorithm, imagine that sample data points in a given problem domain send messages to all other data points. Each data point considers every other data point as a potential centroid or exemplar of cluster of that point. Each point sends message to all other points informing its target about each target's relative attractiveness to sender. Each target then sends reply message to all the senders informing each sender of its availability to associate with the sender, taking into account of all other attractiveness messages it has received from all senders. Sender receives the message and reply to the target informing each target of the target's updated attractiveness to the sender taking into account the availability message from all data points it has received. The message passing continues until all the points decide the clusters and their centroids. The algorithm can be viewed as conversations between gathering of people where each person shares its views with all others, and based upon the views people identify their representative for a specific function.

AP clustering is a message passing algorithm where there is no initial random selection of centroid rather each data point is considered as a candidate for exemplar (centroid) and each data point is also considered as candidate for all the clusters. In iterations of algorithm, messages are passed between data points until robust exemplars (centroids) and clusters emerge i.e. there is no

change in cluster centroids and data points or strength of messages passed fall below a threshold. There are two kinds of messages exchanged between data points: responsibility and availability.

The responsibility, $r(i,j)$, sent from data point $i$ to candidate exemplar data point $j$, indicates how strongly data point $i$ favors the data point $j$ to be chosen as its exemplar over other candidate exemplars. The responsibility is represented in Figure 10 where a data point $i$ is shown to send the responsibility message to candidate exemplar $j$ based upon messages (availabilities) received from other candidate exemplars.

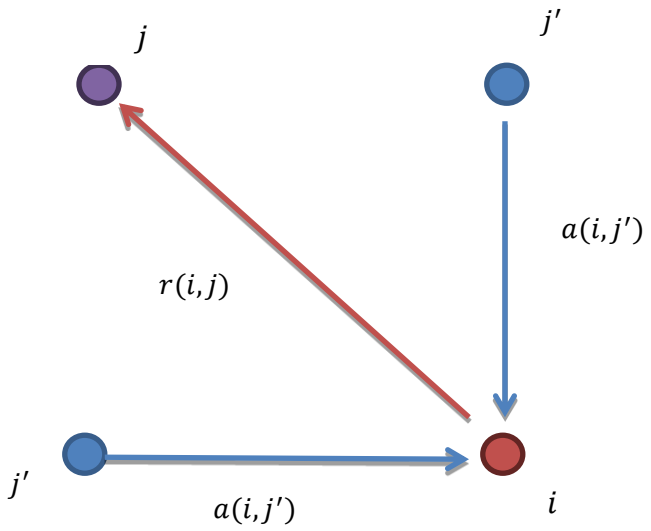


**Figure 10: Responsibility message. Responsibility is being sent from point *i* to point *j* taking into account the availability message received from all other data points represented here as *j′* [30].**

The availability message, $a(i,j)$, sent from candidate exemplar point $j$ to point $i$, indicates the extent to which data point $j$ is available as a cluster center (exemplar) for point $i$ taking into account responsibilities received from all other points. The availability message is presented in Figure 11, where point $j$ as being candidate exemplar sends its availability to point $i$ by considering all responsibilities it is receiving from other two points.

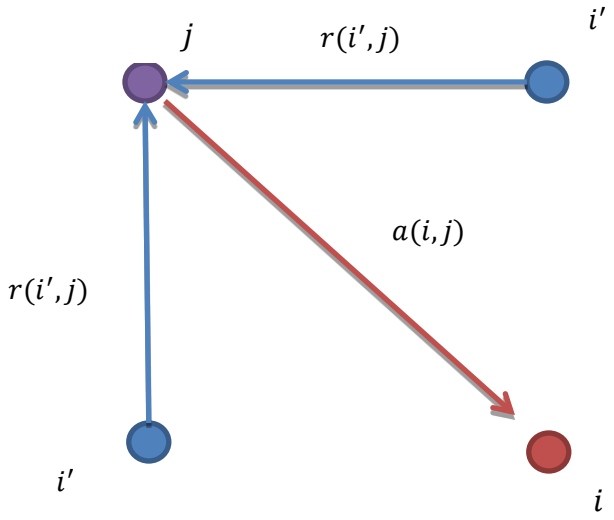


**Figure 11: Availability message. Availability is being sent from point *j* to point *i* taking into account the responsibility message received from all other data points represented here as *i'* [30].**

These messages are iteratively passed between data points and can be combined at any stage to determine cluster exemplars and clusters. Figure  represents illustration of AP algorithm for two-dimensional data points. Each point is colored according to its strength for being exemplar.

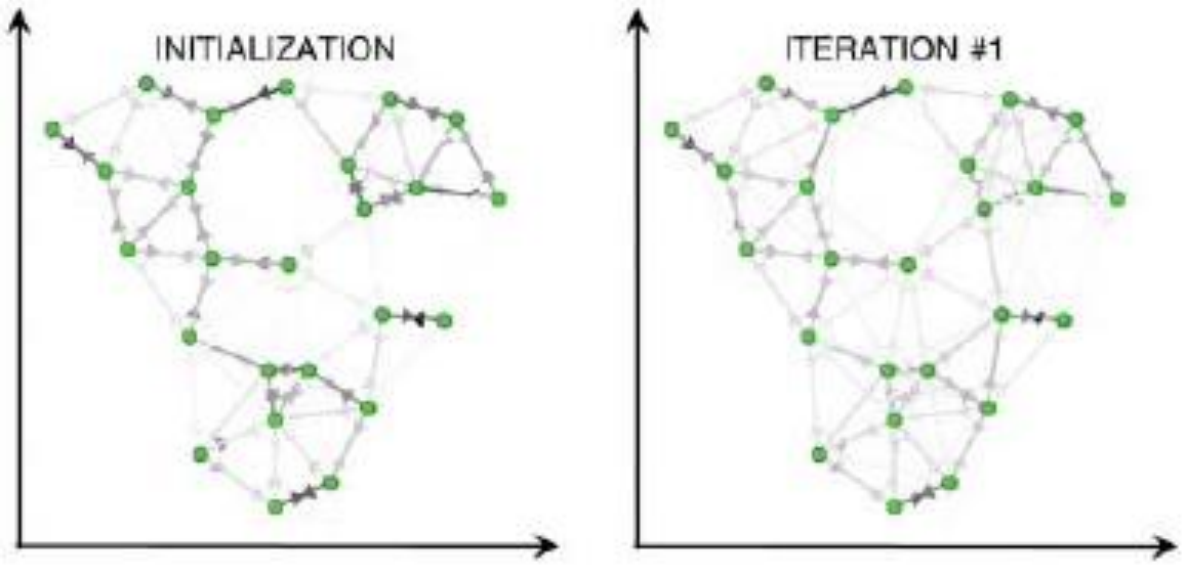


**Figure 12: Visualization of AP clustering for 2D data points. Initially all points are equally considered as centroid of clusters but as messages passes, clusters and their centroids gradually emerge [30].**

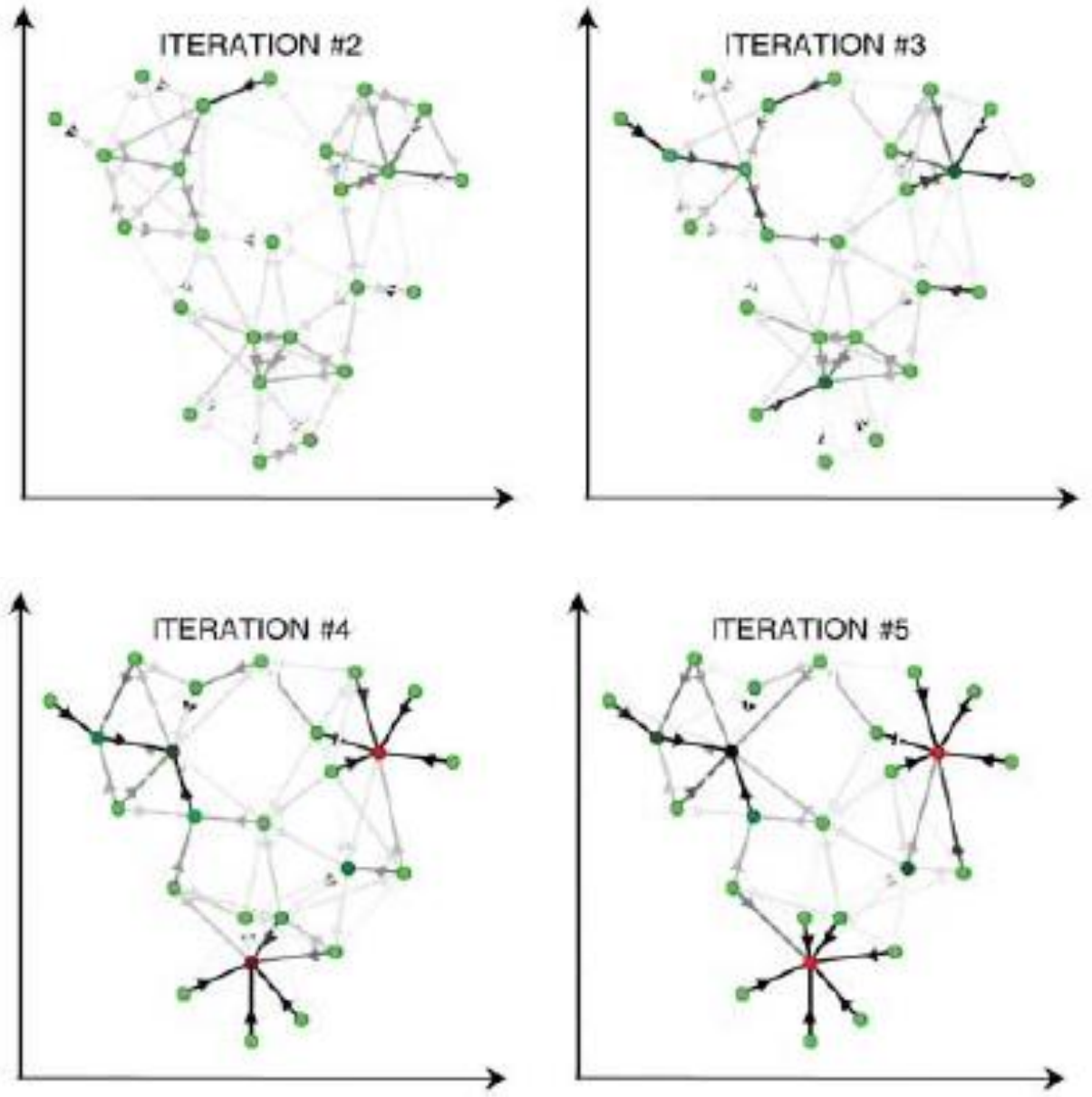


**Figure 12: (Continued).**

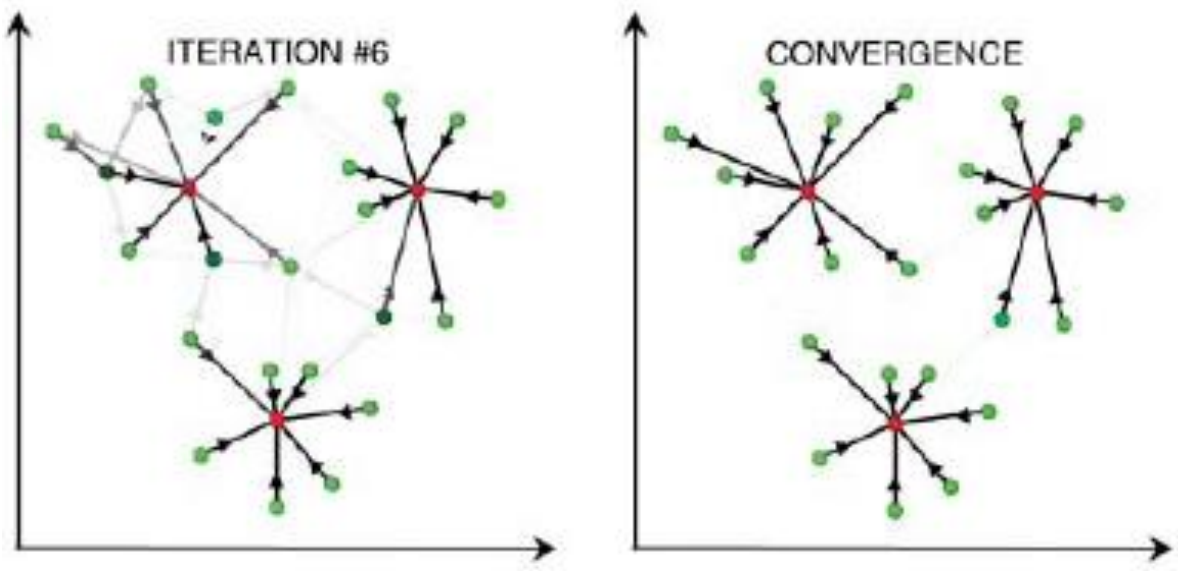


**Figure 12: (Continued).**

Procedurally three matrices are maintained by the algorithm: similarity matrix ($S$), responsibility matrix ($r$) and availability matrix ($a$). Similarity matrix ($S$) is maintained such that each value $S(i,j)$ in similarity matrix represents extent of similarity between points $i$ and $j$. It can be calculated by different metrics (correlation, distance based measure, etc.). In each iteration of algorithm, the responsibility and availability messages are maintained in their respective matrices by equations:

Responsibility message:

$$r(i,j) = s(i,j) - \max_{j',j' \neq j} a(i,j') + s(i,j') \tag{3-2}$$

Availability message:

$$a(i,j) = \min\left\{0, r(j,j) + \sum_{i',i' \neq \{i,j\}} \max\{0, r(i',j)\}\right\} \tag{3-3}$$

Self-availability:

$$a(j,j) = \sum_{i',i \neq j} \max\{0, r(i',j)\} \tag{3-4}$$

At any stage of algorithm, the messages can be combined in one $M$ matrix to determine number of clusters and their centroids.

$$\boldsymbol{M(i,j) = r(i,j) + a(i,j)} \tag{3-5}$$

In our study similarity matrix $S$ is constructed from time series of all brain regions where $S(i,k)$ is the similarity measure between brain regions$i$ and $j$ and is calculated as:

$$S(i,j) = 1 - \frac{< i.j >}{\sqrt{< i.i >}\sqrt{< j.j >}} \tag{3-6}$$

where <·> indicates vector dot product.

There are no random initializations for initial selection of centroids, rather the AP clustering algorithm requires a preference value as its input which depicts the initial preference given to each point as being an exemplar. The number of identified cluster centroids is influenced by the preference value and also emerges from message passing procedure. The change in preference value results in different number of clusters. Assigning a common preference value for all points assigns all data points equal probability for exemplar. Assigning median value of similarity will result in moderate clusters and minimum value for preference will result in minimum number of clusters. Mostly these common approaches for assignment of preference value are adopted and typically a common value is employed for setting preference values for all points.

In this work, we do not use a common value for all preference values but rather derive a data driven method for calculation of preference vector. Preference vector ($p$) is calculated from the $S$ matrix such that:

$$p(i) = \mu\left(\max_{n}\left(S(i,:)\right)\right) \tag{3-7}$$

where $\max_{n}$ denotes choosing $n$ maximum values from the vector, $S(i,:)$ denotes the $i^{th}$ row vector of matrix $S$ and μ denotes the mean.

After each run of AP clustering, a binary connectivity matrix $C$ is constructed, such that each entry $C(i,j)$ is calculated as:

$$\boldsymbol{C(i,j)} = \begin{cases} \mathbf{1}, \boldsymbol{if\ i\ and\ j\ grouped\ in\ same\ cluster} \\ \mathbf{0}, \qquad\qquad\qquad \boldsymbol{otherwise} \end{cases} \tag{3-8}$$

By varying $n$ in Equation (3-8) the preference value is changed which results in changing the number of clusters produced. Therefore we conducted experiments by varying n such that the consistency of functional connectivity among brain regions is tested across different $C$ matrices. In this respect AP is run multiple times to yield multiple connectivity matrices by varying $n$ from 5 to 30 in steps of 5. In this way we calculated six $C$ connectivity matrices to attain stable clustering results. From these connectivity matrices, community matrix $K$ is constructed such that each entry $K(i,j)$ is calculated as:

$$K(i,j) = \frac{1}{m}\sum_{l=1}^{m} C_l(i,j) \tag{3-9}$$

where $m$ = 6 and $C_l$ is the $l^{th}$ connectivity matrix.

We can consider entry $K(i,j)$ as an estimate of the probability of $i^{th}$ and $j^{th}$ brain regions belonging to the same functional community. The constructed community matrix $K$ represents functional brain network in a sparse manner which may not be possible with correlation matrix, see Figure 13. The connectivity differences between normal and epileptic subjects can be identified relatively easily from the sparse data (i.e. community matrix).

The resultant community matrix $K$ is a $90 \times 90$ matrix, corresponding to 90 cortical regions, and it contains 4005 distinct edges (i.e. connections). We believe that this community matrix contains crucial information which can be essentially employed to discriminate epilepsy subjects from normal subjects. We note that the number of 4005 distinct connections from the community matrix $K$ is a high number and may include both discriminatory and non-discriminatory connections which may hamper the performance of the classifier. Therefore, there is a need to select a smaller number of sparse connections having high discriminatory power.

## 3.2 Feature selection

A typical fMRI dataset contains far more features than number of subjects available. In our study the community matrix calculated from previous section has very high dimensions as compared to number of subjects studied. In our dataset numbers of features are of the order of thousands while number of subjects is of order of hundreds. In most of the scenarios, many of the features are irrelevant for a particular problem but cannot be identified because of lack of (neurological) knowledge about individual feature or feature set. The high dimensionality of features in fMRI domain offer a well-known problem called the *curse of dimensionality*. The performance of classifier typically depends upon number of subjects and number of features available.

The main motivation of feature selection is to reduce the feature set by removing irrelevant features and increase the classification accuracy of the classifier. The goal is to find the most relevant and discriminant features that are important for subsequent classification. A feature selection algorithm searches all possible subset of features to find the best informative ones. However, exhaustive search is not feasible in most of the cases due to high number of features.

Also, in most of the cases number of optimum features is not known in advance. Certain criteria are implemented in feature selection algorithms to find the optimum set of features.

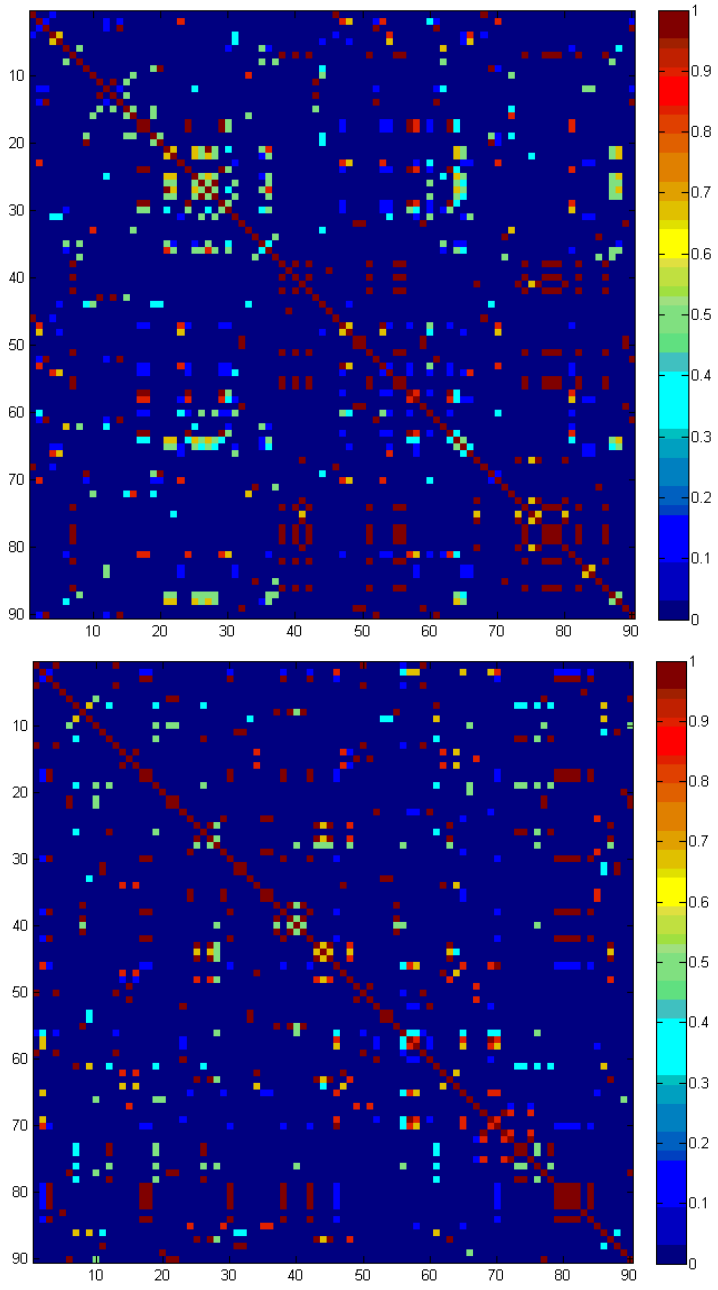


**Figure 13: Visualization of community matrices. Community matrix of normal (first row) and patient (second row). It can be easily seen that connectivity differences between the two can be easily identified due to sparse nature of network.**

### 3.2.1 Review of feature selection methodologies

Based upon the methods employed, feature selection methodology is divided into three types [34]: filter methods, wrapper methods and embedded methods.

**Filter methods:** Filter techniques estimate the importance or discriminative power of features by using intrinsic properties of data. Typically a relevance score for each feature is calculated. In order to select optimum features, a threshold value against relevance score or number of feature is calculated and features with relevance score lower than threshold are considered irrelevant and are removed. Features with higher score are considered most relevant and are presented to classifier for classification. In filter based methods, calculation of relevance score is independent from classifier. Therefore, these methods can be used for any classification algorithm. Filter based methods are typically univariate where relevance score of each feature is calculated independently from other features. These methods are therefore simple and fast and can easily fit to high dimensional data.

One main disadvantage of these methods is that these methods ignore feature dependencies which are important in many domains. In many machine learning and classification problems, a subset of features when grouped together is considered as important and can discriminate between multiple classes. But when the grouped features are viewed separately, they typically appear with low rank features. These methods typically lack to identify such interactions between features.

Some common examples of these methods are t-test, f-test and Gini coefficient.

**t-test:** Score for each feature is calculated based upon following equation. Features with high t value are selected.

$$t(v) = \frac{\mu_n(v) - \mu_p(v)}{\sqrt{\frac{\mu_n^2}{n_n} + \frac{\mu_p^2}{n_p}}} \tag{3-10}$$

where $\mu$ is mean value, $n$ and $p$ are normal and patient classes, respectively.

**Gini index:** Score for a feature $v$ for two classes $n$ and $p$ is calculated by the following equation.

$$GiniIndex(f) = 1 - \sum_{i=1}^{c} [p(i|f)]^2 \tag{3-11}$$

where $c$ is the number of classes, $f$ is the feature. Low value indicates high discriminative power.

**Wrapper methods:** Filter based methods lack the ability to interact with classifiers. To overcome this problem, wrapper methods are adopted. Wrapper methods improve the problem of feature selection by interacting feature selection algorithm with classifier. In these methods, feature subset selection algorithm is defined and various subsets of features are generated and evaluated. For evaluation of a particular subset of features, typically cross validation classification model is employed. Feature set with best accuracy are retained for further analysis. By this way, for selection of optimum subset of features, the search algorithm is *wrapped* around the classifier. These methods have a benefit of taking into account the dependencies of features, also they have advantage of interaction between classifier. One drawback of these methods is that they are sometimes viewed as brute force solution, and if the number of features increases, the search for subset feature selection may degrade the efficiency. To overcome this limitation, heuristics are applied for searching of subset of features.

Main advantages of these methods include consideration of dependencies of features and interaction with classifier to yield good accuracy. Disadvantage of these methods are that they

are computationally intensive, they may cause problem of over fitting as compared to filter methods.

In e**mbedded methods,** classification algorithm itself is employed for feature selection. Instead of using classifier as black box just for performing classification, parameters of classifier are employed for feature selection. In most of the classification algorithms, classifier itself ranks features with respect to discriminative ability of each feature towards classification problem. The features from these selection or ranking can be utilized for feature selection. These methods are typically less computationally intensive as compared to wrapper methods and feature selection is performed in closed interaction with the classifier.

### 3.2.2 Difference statistic

In this work we propose a feature selection technique called difference statistic (DS). This method lies somewhere in between filter and wrapper methods. Score of each feature is calculated independently following filter based methods approach and number of optimum selected features are determined with the help of classifier by employing cross-validation method.

We propose a technique to extract discriminant connections from the total 4005 connections. A difference matrix D is constructed from normal and patient community matrices such that each entry $D(i,j)$ reflects the discriminative power of the particular edge (between $i^{th}$ and $j^{th}$ brain regions). Each entry $D(i,j)$ is calculated as:

$$\boldsymbol{D(i,j) = |\mu K^{+}(i,j) - \mu K^{-}(i,j)|} \tag{3-12}$$

where $\mu K^{+}(i,j)$ is the mean of patient community matrices $K^{+}$ at index $(i,j)$ and $\mu K^{-}(i,j)$ is the mean of normal community matrices $K^{-}$ at index $(i,j)$.

Difference matrix $D$ reflects the discriminative power of each connection – the feature with larger value indicates higher discriminative power. Community matrix, after selection of connections, represents the discriminative functional connectivity between healthy and epileptic groups. This difference matrix $D$ can serve as a sensitive neuroimaging marker for pattern recognition.

Next, we propose two methods for the selection of discriminant features (i.e. connections) that are subsequently employed for classification. For each method, 100 iterations of feature selection are run such that, for each iteration, the data is randomly distributed by 50% split into training data and test data. In each iteration, following steps are performed: (i) the top $h$ number of features with largest absolute difference statistic value $D(i,j)$ are selected from the training data, (ii) the SVM classifier is trained on the training data, and (iii) the performance of these selected features is evaluated by validating the trained classifier on the test data. The two proposed feature selection methods are described below:

### i) High-accuracy feature selection

Let $f_i$ represent the set of $h$ number of features from $i^{\text{th}}$ iteration. In this method, those $h$ features are selected from a particular $i^{\text{th}}$ iteration which yielded the highest classification accuracy across all the 100 iterations. The proposed feature selection process is depicted below:

$$f_{high} = arg\,max\, Acc(f_i)_{i\in\{1,2,\ldots,100\}} \tag{3-13}$$

where $Acc(f_i)$ function denotes classification accuracy obtained with features $f_i$, $arg\,max$ represents choosing features $f_i$ that yield high classification accuracy, and $f_{high}$ denotes the set of selected features with high accuracy.

### ii) Consistent feature selection

Let $f_i$ represent the set of $h$ number of features from $i^{\text{th}}$ iteration. In this method, those $h$ features are selected which occurred most consistently (i.e. most frequently) across all the 100 iterations. The proposed feature selection process is depicted below:

$$f_{consistent} = arg\ max_h\ Cons(f_i)_{i\in\{1,2,\ldots,100\}} \quad (3\text{-}14)$$

where $Cons(f_i)$ function denotes extracting a subset of most consistent (i.e. frequent) features from a large set, $\arg\max_h$ represents selecting $h$ most consistent features, $f_{consistent}$ denotes the set of selected features which occurred most consistently.

## 3.3 Classification by support vector machines (SVM)

The next step in this study is the classification. To confirm the discriminative ability of selected connections, support vector machine (SVM) classifier has been explored in this work for classification of normal and patient groups. SVM classifier, with foundations in machine learning, is commonly used in literature [1, 28]. It seeks optimal margin between the different classes. In training mode, the region connection and classification label information (1 for normal, -1 for epileptic subject) is provided to classifier. The SVM classifier then learns the classification model between the two classes (normal or epileptic) from the training data presented to it. In testing mode, the region connections are provided and the classification label is predicted by the classifier. In our case, the selected connections from the previous section are provided to the classifier for training or testing.

SVM [35] is a supervised learning algorithm used to solve classification problems. It uses machine learning techniques to maximize prediction accuracy. It is widely employed in pattern classification applications like hand written recognition and speech recognition.

SVM uses nonlinear mapping to transform training data into higher dimensions. Within this new dimension, it searches for linear optimum hyper plane splitting the tuples of one class from other. Hyper plane is also regarded as decision boundary. With a suitable nonlinear mapping and appropriate high dimensions, a hyper plane can always be calculated to separate two classes [36]. SVM calculates this hyper plane by finding essential tuples from training data (called support vectors) and margins (between these support vectors).

In order to understand basic principle of SVM, consider an example of data with two classes that are linearly separable. Let us consider the data set as $(x_1, y_1), (x_2, y_2), \ldots (x_n, y_n)$, where $x_i$ is the set of training tuple and $y_i$ is the associated class label such that:

$$y_i \in \{+1, -1\}$$

where +1 and -1 are labels assigned to both classes. Figure 14 represents visualization of 2-D points based upon arbitrary attributes $A_1$ and $A_2$, the figure shows that both the classes are linear separable as all tuples can be separated by a straight line.

There can be infinite number of lines that can be drawn between the two classes to separate them. The objective of a classification algorithm would be to find a separating line that will yield minimum classification error on unseen dataset. If we consider 3D dataset (with three attributes), there would be a separating plane instead of separating line. Generally for $n$ dimensions, goal of classification algorithm is to find best separating hyper plane with minimum classification error.

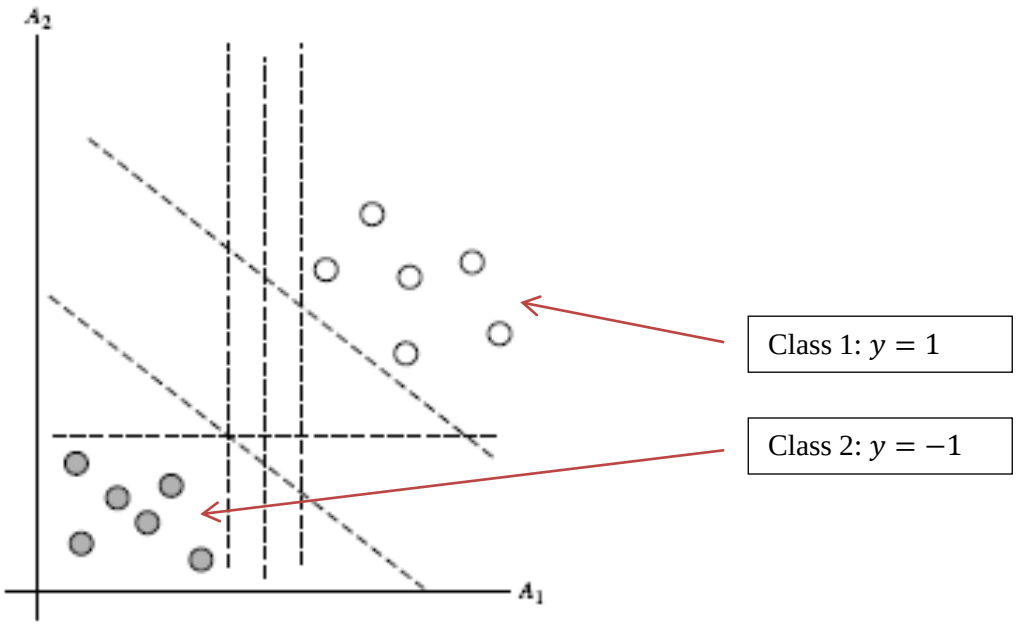


**Figure 14: Visualization of 2D data points (from [36]).**

SVM solves this problem of finding best separating hyper plane by searching for the maximum marginal hyper plane. Figure 15 shows same dataset of Figure 14 with two possible separating hyper planes and their associated margins. Both the planes can correctly classify the given unseen tuples to the system, however, hyper plane with larger margin is expected to be attain better classification accuracy than hyper plane with small margin. During training phase, SVM searches for maximum marginal hyper plane (MMH) that gives the largest margins between the given classes.

A separating hyper plane can be mathematically expressed as:

$$W.X + b = 0 \tag{3-15}$$

where $W$ is a weight vector:

$$W = \{w_1, w_2, \dots, w_n\} \tag{3-16}$$

$n$ is the number of attributes and $b$ is biasness.

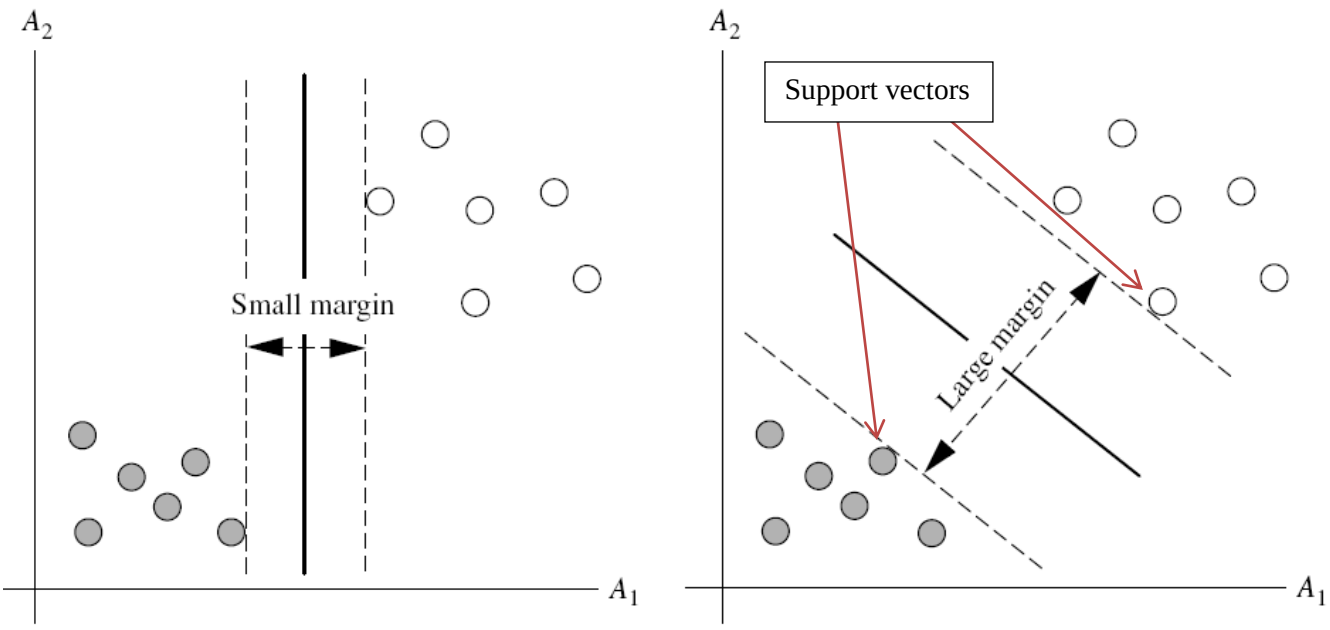


**Figure 15: Visualization of separating hyper plane. Demonstrating large margins and support vector (from [36]).**

Consider 2-D tuples as in Figure 14 with attributes $A_1$ and $A_2$ and $X$ is:

$$X = (x_1, x_2) \tag{3-17}$$

where $x_1, x_2$ are values of attributes $A_1$and $A_2$. If biasness is also considered as weight ($w_0$), equation 3.17 can be written as:

$$w_1 x_1 + w_2 x_2 + w_0 = 0 \tag{3-18}$$

The weights can be tuned so that hyper plane defining the sides (sides of margin are parallel to hyper plane) of the margins can be expressed as:

$$H_1: w_1 x_1 + w_2 x_2 + w_0 \geq 1 \; for \; y_i = 1 \tag{3-19}$$

$$H_2: w_1 x_1 + w_2 x_2 + w_0 \leq 1 \; for \; y_i = -1 \tag{3-20}$$

Any tuple falling above or on $H_1$ is assumed to belong to class 1 and any tuple falling below or on $H_2$ belongs to class -1.

Any tuple that falls on hyper planes $H_1$or $H_2$is called support vectors and can be expressed by equation (3-25). Support vectors are shown in Figure 15.The selection of support vector is most crucial as they provide most information regarding classification. These support vectors and MMH is calculated during training phase.

During testing phase, consider a tuple $X^T$ is supplied to trained SVM for prediction of its class label. The input tuple is used in equation:

$$d(X^T) = \sum_{i=1}^{l} y_i \propto_i X_i X^T + b_0 \tag{3-21}$$

where $X^T$ is the test tuple, $l$ is the number of support vector, $y_i$ is the label of support vector $X_i$, $\propto_i$ and $b_0$ are numeric parameters.

The sign of the above equation is used for label prediction. If it is positive, input tuple $X^T$ falls on or above MMH and is predicted to class 1, if sign is negative it falls on or below MMH and is assigned label class -1.

In the case of nonlinear data (where it is not possible to have a separating plane), SVM uses kernel function to map the data to higher dimensions and searches separating hyper plane within this high dimensional space.

## 3.4 Summary

This chapter describes our proposed methodology consisting of three main steps: functional connectivity estimation, feature selection and classification. We have proposed affinity propagation clustering for estimation of functional connectivity between brain regions. The clustering algorithm has benefit over certain other clustering algorithms like *k*-means as the

algorithm does not involve arbitrary initial selection of cluster centroids and does not require number of clusters in advance. Difference statistics is proposed for feature selection and support vector machines (SVM) is employed for classification.


# Chapter 4

# Results and discussion

This chapter presents the details of the dataset and discusses the results of the study. The chapter also provides discussion of these results and highlights neurological significance and findings of these results.

The resting state fMRI (rfMRI) data used in this study is same as used in the study of Zhang et al. [1]. The data acquisition and pre-processing details are provided below.

The data contained 180 subjects including 80 healthy controls (age 24.89 ± 8.63) and 100 epileptic patients (age 23.85 ± 5.66). All the subjects were right handed. The epileptic subjects suffered from different kinds of epilepsy (e.g., partial and global epilepsy, temporal lobe epilepsy). There were 18 epileptic patients with global seizure and 70 with partial seizure. Antiepileptic drugs (AEDs) were used by 82 patients and 18 were not using any AEDs.

Before scanning, all the subjects were asked to relax, close their eyes without falling into sleep, hold still and think nothing in particular. The rfMRI data was collected on 3 Tesla Siemens Trio Tim scanner machine with an eight channel array head coil. Echo-planar imaging (EPI) sequence was used to acquire resting state fMRI data. Parameters used were: TR/TE = 2000ms/30ms, FA = 90º, matrix = 64×64, FOV = 24×24 $cm^2$, slice thickness = 4 mm, slice gap = 0.4 mm and slices acquired = 30. Coronal T2-FLAIR-weighted image parameters: TR/TE = 7000ms/87ms, FA = 150º, matrix = 256×256, FOV = 24×19.5$cm^2$, slice thickness = 4mm, slice gap = 0 mm and acquired slices = 28.

Data-preprocessing was performed using data processing assistant for resting-State fMRI (DPARSF) software which is based on statistical parametric mapping (SPM) and resting-state fMRI data analysis toolkit (REST). Slice-timing adjustment and head motion correction was performed. Spatial normalization was performed by using standard Montreal neurological

institute (MNI) template provided by SPM2 (resampling voxel size: $3\times3\times3mm^3$). After smoothing, the BOLD signals were filtered to remove low-frequency drift and high frequency noises. Anatomically labeled template provided by Tzourio-Mazoyer et al. [7] was used to segment registered fMRI time series into 116 regions (90 from cortex and 26 from cerebellum). For each region, fMRI time series of all voxels lying in the region were averaged to obtain representative fMRI time series or BOLD signal of that region. After preprocessing, for each subject, there are 116 BOLD signals such that $x_i(t)=1,2,\ldots,116$ represents bold signal in the $i^{th}$ brain region. For this study, we look at 90 brain regions belonging to cortex.

For each subject there are 90 cortical regions and each region has 200 time points. By applying AP clustering as discussed in Chapter 3, community matrix is constructed for each subject having dimensions of $90 \times 90$. The community matrix is employed for further analysis of data. Community matrix is symmetric, having 4005 unique features.

For classification, the data was divided into two subsets: 50% for training and the remaining 50% for testing. Segregation of data is shown in Table 1.

| | **Total subjects** | **Number of training subjects** | **Number of test subjects** |
|---|---|---|---|
| **Normal** | 80 | 40 | 40 |
| **Epileptic** | 100 | 50 | 50 |
| **Total** | 180 | 90 | 90 |

**Table 1: Number of subjects and their distribution into testing and training subsets.**

For comparing our results with results of Zhang et al. [1] using same dataset and distribution for training and testing, result of their study are presented in Table 2.

| **Number of selected features** | 50 | 400 |
|---|---|---|
| **Prediction accuracy** | 73.2% | 77.6% |

**Table 2: Results of Zhang et al. [1] using *k*-means clustering.**

Our results are based upon two feature selection methods as discussed in Section 3.2: high-accuracy feature selection and consistent feature selection.

# 4.1 High-accuracy feature selection

We used the training data to extract discriminant connections and to train the classifier. The test data, not previously presented to the classifier, was then provided to the classifier.

## 4.1.1 Feature selection and classification

In order to make comparison with study of Zhang et al. [1] we used 400 most discriminative features selected from t-test, gini index, Wilcoxon and our proposed difference statistics feature selection. In initial experiment, we explored SVM and random forest (RF) classifier for classification. The results are presented in Table 3. These results show that best accuracy is achieved by our proposed difference statistics and SVM classifier. Our algorithm achieved better accuracy than Zhang et al. [1] (77.6% accuracy with 400 features while our study yield 80.79%).

| Sr. No | Feature selection | Classifier | Prediction accuracy |
|---|---|---|---|
| 1 | Difference statistic | SVM | 80.79% |
| 2 | Difference statistic | RF | 76.89% |
| 3 | Gini index | SVM | 80.17% |
| 4 | Gini index | RF | 78.34% |
| 5 | t-test | SVM | 78.94% |
| 6 | Wilcoxon | SVM | 72.44% |

**Table 3: Evaluation of different feature selection methods and classifiers. Number of selected connections is fixed to 400 to make comparison with Zhang et al. Different feature selection methods are employed and difference statistic is yielding best accuracy. Data split for training and testing is according to Table 1. We employed cross validation using 100 trials for assessing classifier performance.**

For selection of classifier, results of Table 3 shows that SVM yields better results than RF and best results are achieved with combination of difference statistic as feature selection and SVM as classifier.

### 4.1.2 Study of number of selected connections

In order to observe the dependency of classifier upon number of selected connections, we repeated the experiment by varying the number of features while feature selection method and classifier were fixed to difference statistic and SVM, respectively. In this experiment number of selected connections was varied from 350 to 600. The results are presented in Figure 16.

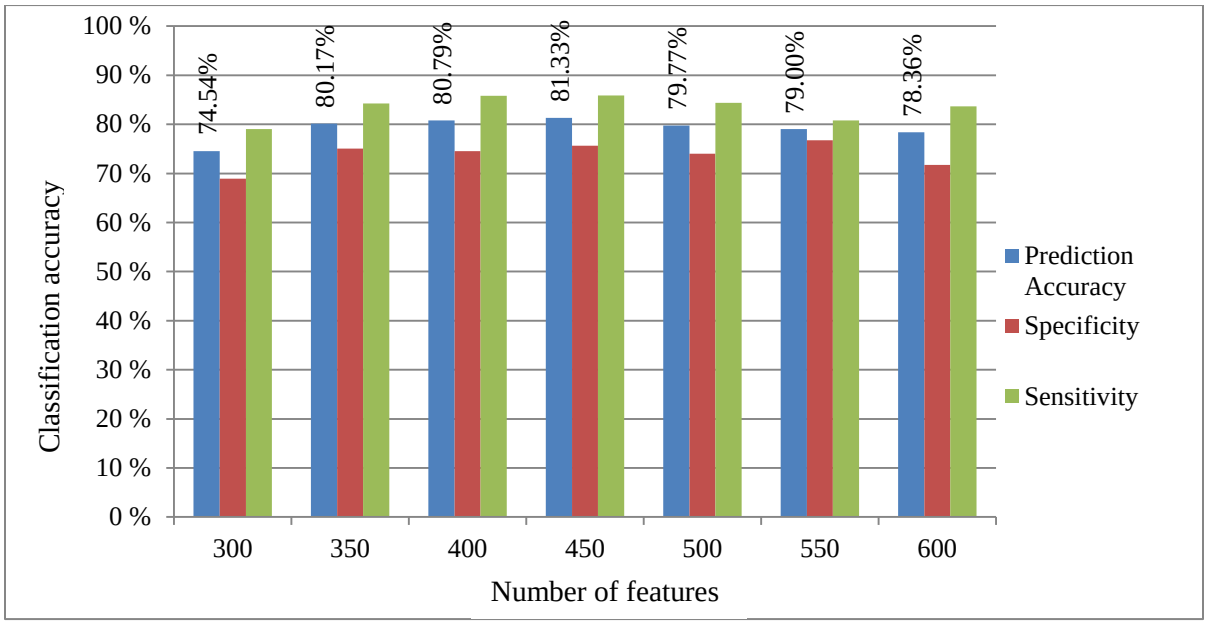


**Figure 16: Classification accuracy by varying number of features. Number of features is varied from 300 to 600 while classifier and features selection is fixed to SVM and difference statistics, respectively.**

The specificity and sensitivity shown in Figure 16 are defined as following:

$$Specificity = \frac{TN}{TN + FP} \times 100\%$$

$$Sensitivity = \frac{TP}{TP + FN} \times 100\%$$

where

TN (number of true negatives) = normal subject correctly identified as normal,

FP (number of false positives) = normal subject incorrectly identified as epileptic,

TP (number of true positives) = epileptic subject correctly identified as epileptic, and

FN (number of false negatives) = epileptic subject incorrectly identified as normal.

The results from Figure 16 show that the classification accuracy is not significantly affected by these variations in the number of selected connections and best accuracy of 81.33% is achieved with 450 selected features.

### 4.1.3 Leave one out validation

We conducted leave one out (LOO) validation employing 400 and 450 best selected features. In LOO validation, we trained SVM classifier using selected features on 179 subjects, and tested the classifier against the one unseen subject. The procedure was repeated 180 times so that each subject in dataset was treated once as test data while SVM trained on all remaining 179 subjects. Prediction accuracy from all 180 iterations was averaged to yield LOO validation and results are reported in Figure 17.

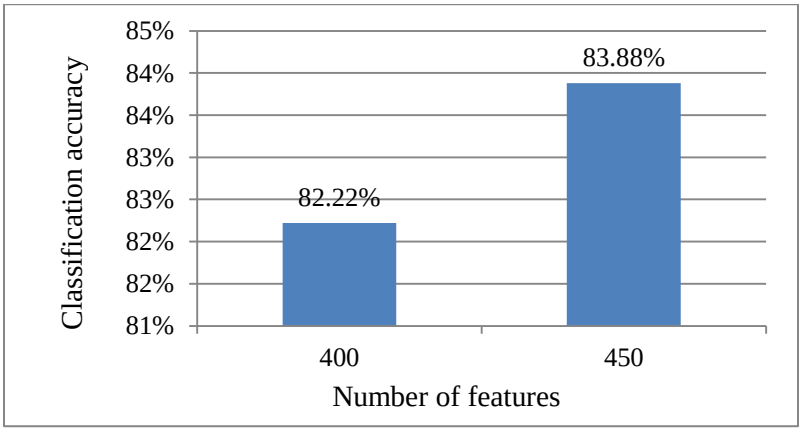


**Figure 17: LOO validation results. Features varied from 400 to 450.**

### 4.1.4 Comparison with *k*-means

In order to evaluate performance of feature selection and classifier, we studied the results by combination of our methodologies and Zhang et al. [1] We tried different combinations of clustering and feature selection and the results are presented in Figure 18. The results demonstrate that highest prediction accuracy is achieved by combination of our proposed methodology.

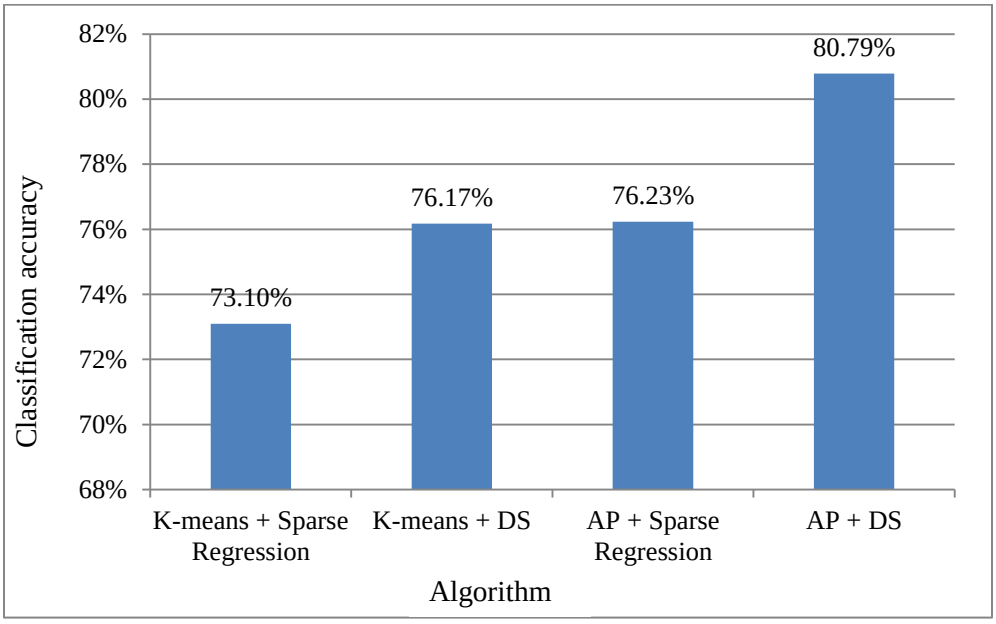


**Figure 18: Comparison of Zhang et al. [1] and our proposed methodology. Best accuracy is achieved with our proposed methodology. Number of selected features are 400.**

In next phase of analysis, we studied the neurological details of our selected connections using brain regions. Automated anatomical labeling (AAL) [7] template was employed to segment brain in 90 regions. Human brain is composed of two symmetric hemispheres: right hemisphere and left hemisphere. There are 90 regions with 45 regions in each hemisphere. The regions are listed in Table 4.

To better understand the neurological importance of selected connections, we analyzed thirty most discriminant connections that are listed in Table 5, the connections that have highest values in difference matrix ($D$). These connections are sorted based upon their absolute D matrix values.

| No | Region Name | No | Region Name |
|---|---|---|---|
| 1 | Precentral gyrus | 24 | Lingual gyrus |
| 2 | Superior frontal gyrus, dorsolateral | 25 | Superior occipital |
| 3 | Superior frontal gyrus, orbital part | 26 | Middle occipital |
| 4 | Middle frontal gyrus (lateral part) | 27 | Inferior occipital |
| 5 | Middle frontal gyrus, orbital part | 28 | Fusiform gyrus |
| 6 | Opercular part of inferior frontal gyrus | 29 | Postcentralgyrus |
| 7 | Triangularis part of inferior frontal gyrus | 30 | Superior parietal lobule |
| 8 | Orbital part of inferior frontal gyrus | 31 | Inferior parietal lobule |
| 9 | Rolandic operculum | 32 | Supramarginalgyrus |
| 10 | Supplementary motor area | 33 | Angular gyrus |
| 11 | Olfactory cortex | 34 | Precuneus |
| 12 | Superior frontal gyrus, medial part | 35 | Paracentral lobule |
| 13 | Superior frontal gyrus, medial orbital part | 36 | Caudate nucleus |
| 14 | Gyrus rectus | 37 | Putamen |
| 15 | Insula | 38 | Globuspallidus |
| 16 | Anterior cingulate gyrus | 39 | Thalamus |
| 17 | Middle cingulate | 40 | Transverse temporal gyri |
| 18 | Posterior cingulate gyrus | 41 | Superior temporal gyrus |
| 19 | Hippocampus | 42 | Superior temporal pole |
| 20 | Parahippocampalgyrus | 43 | Middle temporal gyrus |
| 21 | Amygdala | 44 | Middle temporal pole |
| 22 | Calcarine sulcus | 45 | Inferior temporal gyrus |
| 23 | Cuneus | | |

**Table 4: Names of the brain regions as per AAL [7].**

Among these thirty connections, seven connections are from bilaterally homologous regions (traverse temporal gyri, cuneus, fusiform, inferior occipital, precentral gyrus, middle occipital and supramarginal gyrus), which indicate that these bilateral homologous functional connections are impaired. Out of these seven, five (traverse temporal gyri, cuneus, fusiform, inferior occipital, precentral gyrus) have decreased connectivity and last two (middle occipital and supramarginal gyrus) have increased connectivity in epileptic subjects compared to normal subjects. In our study, number of inter-hemispheric functional connectivity anomalies are dominant over intra-hemispheric connectivity anomalies i.e. seventeen connections across both the hemispheres are disturbed while six connections within left and seven within right hemisphere are altered. Out of these seventeen inter-hemispheric connections, five have increased connectivity and twelve have decreased connectivity.

## 4.2 Consistent features selection

In this method, features are selected based upon their frequency as discussed in Section 3.2. Difference statistic was used as feature selection method and SVM as the classifier. Number of selected frequently occurring features was varied from 300 to 600 to study the effect of number of features on classification accuracy. The results are presented in Figure 19. The results in Figure 19 show that their accuracy is increased from 89.51% to 93.08% as features are increased from 300 to 450, while accuracy is decreased from 93.08% to 89.79% as features are increased from 450 to 600 with best accuracy of 93.08% being achieved by 450 selected features.

| **Brain region** | **Brain region** | ***D* value** |
|---|---|---|
| Superior temporal pole-Right | Posterior cingulate gyrus-Left | -0.1967 |
| Transverse temporal gyri-Right | Transverse temporal gyri-Left | 0.1883 |
| Inferior temporal gyrus-Right | Middle frontal gyrus, orbital part-Right | -0.1833 |
| Cuneus-Right | Cuneus-Left | 0.1817 |
| Lingual gyrus-Left | Calcarine sulcus-Right | 0.1758 |
| Superior frontal gyrus, medial orbital part-Right | Superior frontal gyrus, orbital part-Right | -0.1617 |
| Gyrus rectus-Right | Olfactory cortex-Right | 0.1617 |
| Orbital part of inferior frontal gyrus-Left | Area triangularis-Left | -0.1608 |
| Amygdala-Right | Hippocampus-Right | 0.1592 |
| Superior temporal pole-Right | Rolandic operculum-Right | 0.1575 |
| Parahippocampalgyrus-Right | Hippocampus-Left | 0.1575 |
| Superior temporal gyrus-Right | Rolandic operculum-Left | 0.1508 |
| Globuspallidus-Left | Caudate nucleus-Left | 0.1492 |
| Middle temporal gyrus-Right | Superior frontal gyrus, medial part-Right | 0.1467 |
| Middle occipital-Right | Middle occipital-Left | -0.1467 |
| Supramarginalgyrus-Right | Supramarginalgyrus-Left | -0.1442 |
| Middle occipital-Left | Middle frontal gyrus, lateral part-Right | 0.1442 |
| Fusiform gyrus-Right | Fusiform gyrus-Left | 0.1417 |
| Inferior occipital-Right | Inferior occipital-Left | 0.1408 |
| Amygdala-Left | Superior frontal gyrus, orbital part-Left | -0.14 |
| Gyrus rectus-Left | Olfactory cortex-Right | 0.1392 |
| Orbital part of inferior frontal gyrus-Left | Superior frontal gyrus, dorsolateral-Left | -0.1383 |
| Superior temporal pole-Left | Superior temporal gyrus-Left | 0.1375 |
| Supramarginalgyrus-Left | Posterior cingulate gyrus-Left | -0.1375 |
| Middle frontal gyrus, orbital part-Left | Precentralgyrus-Right | -0.1367 |
| Area triangularis-Right | Opercular part of inferior frontal gyrus-Left | 0.1367 |
| Insula-Right | Rolandic operculum-Left | 0.1358 |
| Inferior occipital-Left | Calcarine sulcus-Right | -0.135 |
| Precentralgyrus-Right | Precentralgyrus-Left | 0.1342 |
| Middle temporal pole-Right | Middle frontal gyrus, lateral part-Right | -0.1333 |

Table 5. Thirty most discriminant connections. The connections are sorted with respect to the corresponding value in ***D*** matrix. (ranked with absolute values). First two columns represent the brain regions involved in connection and third column represent their value in ***D*** matrix. Positive sign of the value represents decreased connectivity (healthy minus epileptic) and negative sign represents increased connectivity. Among these thirty, seventeen are inter-hemispheric (between left and right hemispheres). Out of these seventeen, seven are between symmetric brain regions.

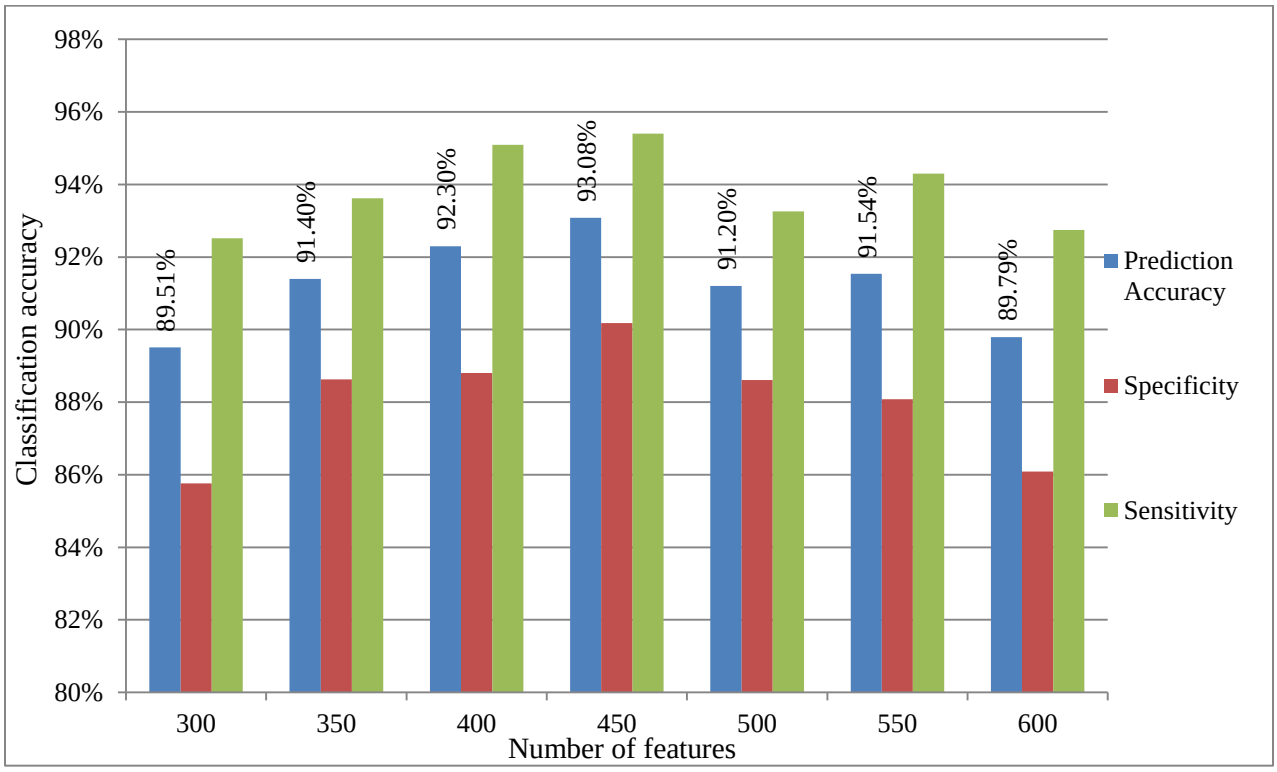


**Figure 19: Accuracy by consistent feature selection method. Number of features selected is varied from 300 to 600, and highest accuracy is achieved by 450 features.**

### 4.2.1 Study of biasness

One may expect that number of iterations for feature selection may introduce biasness in prediction accuracy produced by the algorithm. To study possible effects of biasness by feature selection iterations, we varied the number of $c$ iterations in Equation (3-15) of Section 3.2 from 100 to 10 while number of selected features was fixed to 450. The results are shown in Figure 20. As number of iterations $c$ in Equation (3-15) is dropped from 100 to 10, the classification accuracy shows a minute drop of 2.91% from 93.08% to 90.37%. The results suggest that feature selection is not introducing potential biasness in classification accuracy.

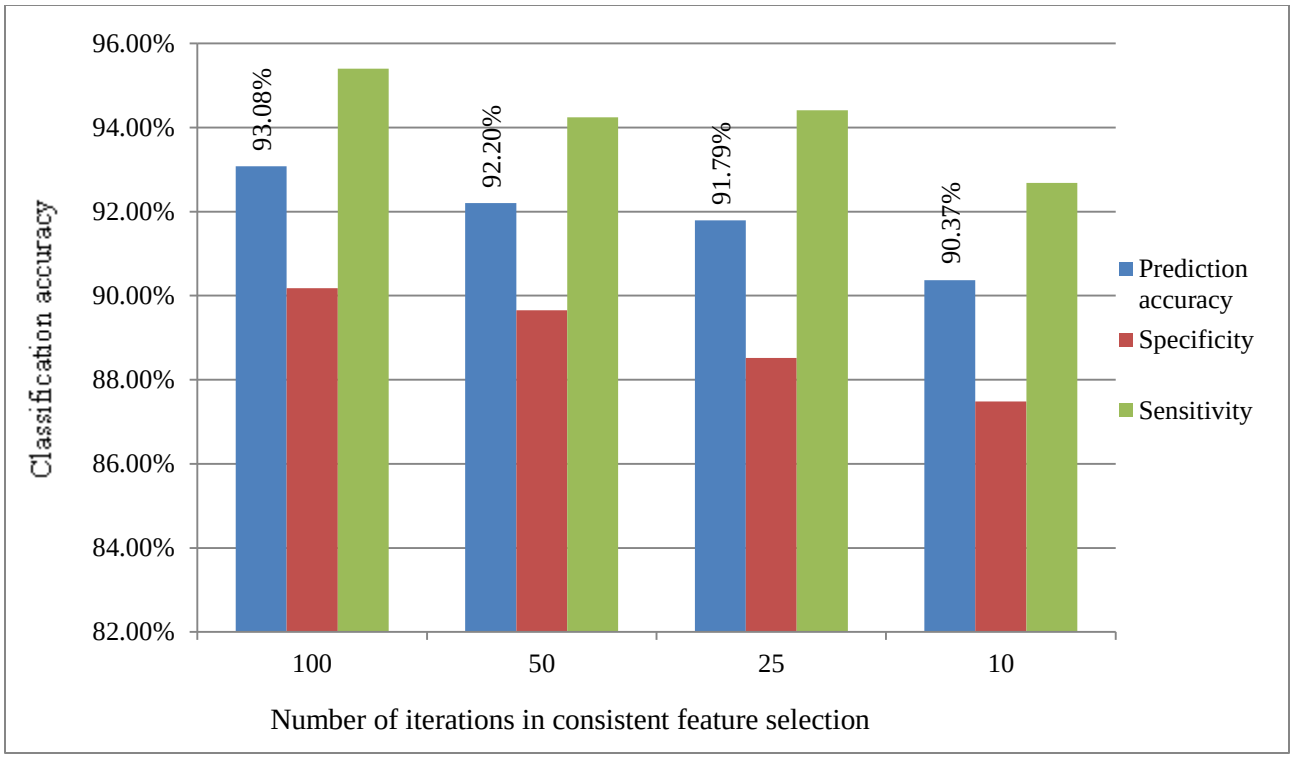


**Figure 20: Effect of biasness. Number of feature is fixed at 450 and c is changed from 100 to 10. As c is decreased 90% (from 100 to 10), there is a minute difference in classification accuracy (from 93.08% to 90.73%).**

These results demonstrate that the proposed method can be used to extract sensitive neuroimaging markers with the capability to discriminate between normal and epileptic subjects. Our results indicate that the proposed method outperforms Zhang et al. [1] method for the purpose of neuroimaging marker estimation. With the proposed method, we achieved classification accuracy of 80.79% with high-accuracy feature selection and 92.3% using consistent feature selection using 400 connections, while Zhang et al. [1] reported 77.60% accuracy. Note that both these studies are using the same rfMRI data.

### 4.2.2 Anatomical group wise study

To better understand neurological significance of our selected features with maximum accuracy (450 features with 93.08% accuracy using consistent feature selection), we performed analysis

based upon established six anatomical groups of brain defined by Salvador et al. [37]. Group names and their corresponding brain regions are listed in Table 6.

| 1 - Medial Temporal | | | |
|---|---|---|---|
| **Hem** | **Name** | **Hem** | **Name** |
| L+R | Hippocampus | L+R | superior temporal pole |
| L+R | parahippocampal gyrus | L+R | middle temporal pole |
| L+R | Amygdala | | |
| **2 – Subcortical** | | | |
| **Hem** | **Name** | **Hem** | **Name** |
| L+R | olfactory cortex | L+R | Globuspallidus |
| L+R | caudate nucleus | L+R | Thalamus |
| L+R | Putamen | | |
| **3 – Occipital** | | | |
| **Hem** | **Name** | **Hem** | **Name** |
| L+R | calcarine sulcus | L+R | middle occipital |
| L+R | Cuneus | L+R | inferior occipital |
| L+R | lingual gyrus | L+R | fusiform gyrus |
| L+R | superior occipital | | |
| **4 – Frontal** | | | |
| **Hem** | **Name** | **Hem** | **Name** |
| L+R | superior frontal gyrus, dorsolateral | L+R | orbital part of inferior frontal gyrus |
| L+R | superior frontal gyrus, orbital part | L+R | superior frontal gyrus, medial part |
| L+R | middle frontal gyrus (lateral part) | L+R | superior frontal gyrus, medial orbital part |
| L+R | middle frontal gyrus, orbital part | L+R | gyrus rectus |
| L+R | opercular part of inferior frontal gyrus | L+R | anterior cingulate gyrus |
| L+R | triangularis part of inferior frontal gyrus | | |
| **5 – Temporal** | | | |
| **Hem** | **Name** | **Hem** | **Name** |
| L+R | Insula | L+R | middle temporal gyrus |
| L+R | transverse temporal gyri | L+R | inferior temporal gyrus |
| L+R | superior temporal gyrus | | |

| 6 - Parietal-(pre)Motor | | | |
|---|---|---|---|
| **Hem** | **Name** | **Hem** | **Name** |
| L+R | precentral gyrus | L+R | superior parietal lobule |
| L+R | rolandic operculum | L+R | inferior parietal lobule |
| L+R | supplementary motor area | L+R | Supramarginal gyrus |
| L+R | middle cingulate | L+R | angular gyrus |
| L+R | posterior cingulate gyrus | L+R | Precuneus |
| L+R | Postcentralgyrus | L+R | paracentral lobule |
| | | | |

**Table 6: Anatomical groups and corresponding regions [37].**

We analyzed altered functional connectivity in terms of the six anatomical groups and categorized the alterations with respect to inter-group and intra-group alterations. We studied detailed analysis with respect to increased and decreased connectivity in patient group as compared to normal group. The results are presented in Figure 21 and Figure 22.

These figures yield following findings:

- Parietal-(pre) motor is affected with respect to both inter and intra group functional connectivity. In both categories the group is found to have reduced connectivity as compared to healthy ones. The group is known to be involved in movement intention and motor awareness [38].
- Functional connectivity of any particular group is decreased with all other groups in epileptic subjects.
- Intra group alterations: Connections within occipital and frontal groups have suffered from increased connectivity while others are suffered from decreased connectivity.

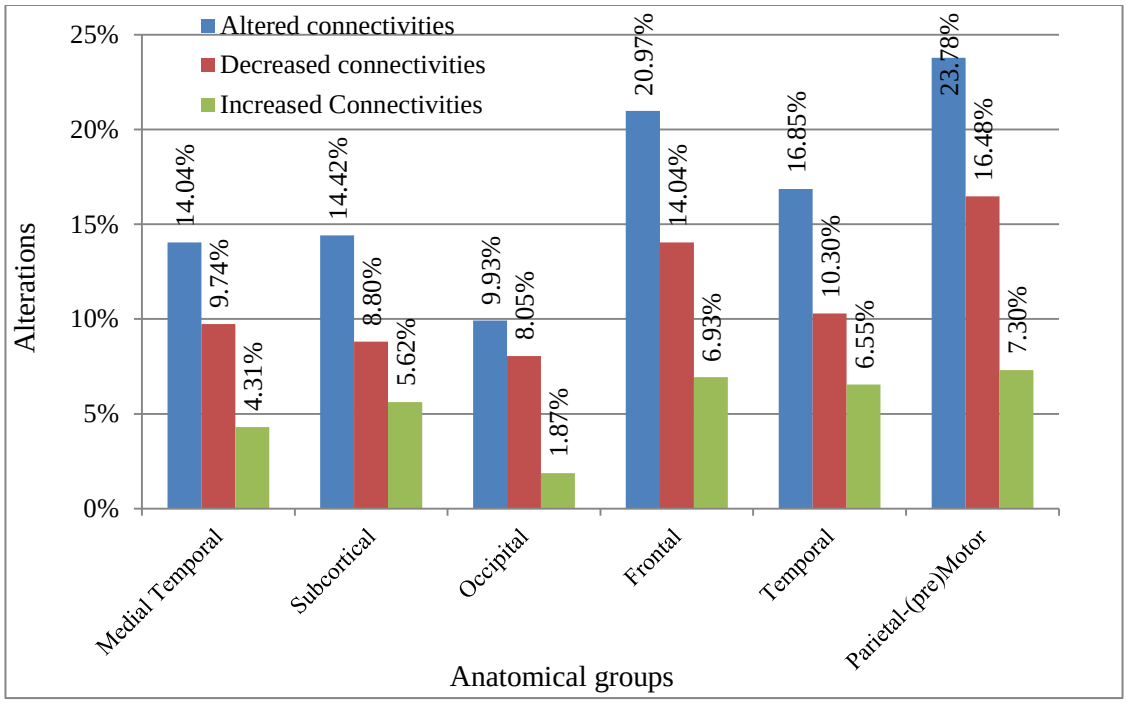


**Figure 21: Inter group alterations. Parietal-(pre)Motor group has highest intergroup alterations followed by frontal group. It is interesting to find that all six groups are affected by decreased functional connectivity in epileptic as compared to healthy subjects.**

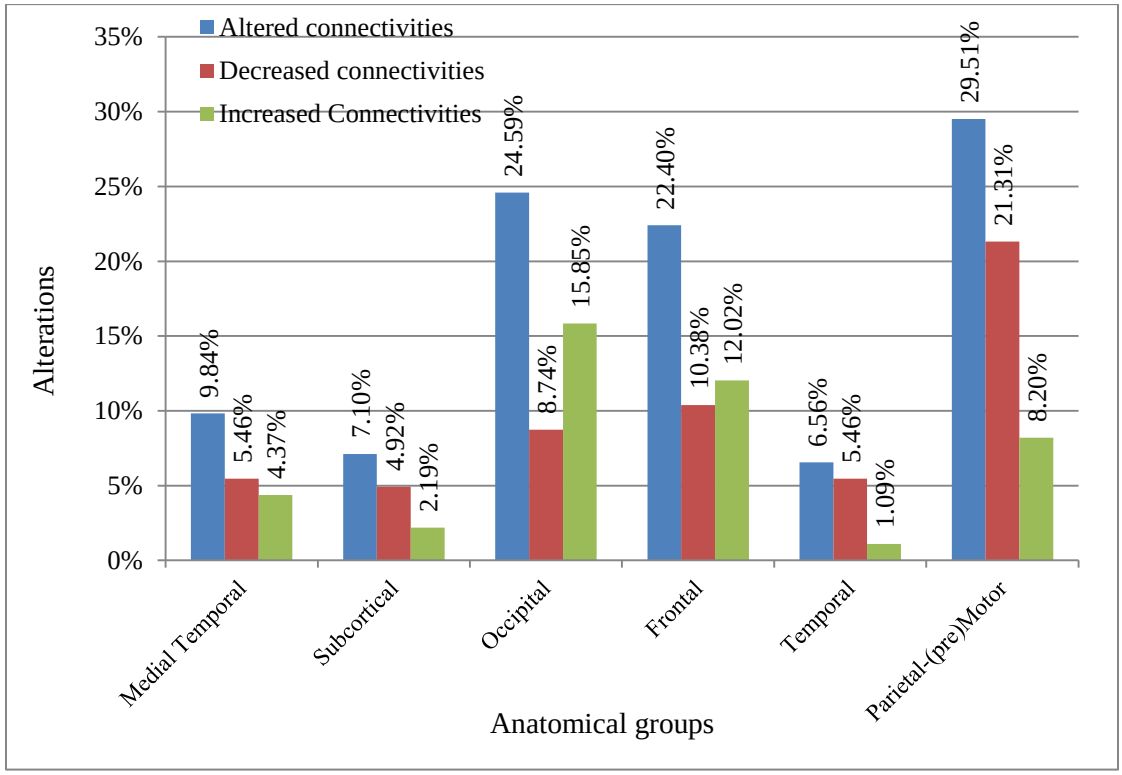


**Figure 22: Intra group alterations. Parietal-(pre)Motor group has highest intragroup alterations followed by occipital group. All the groups except Occipital and Frontal are affected by decreased connectivity.**

### 4.2.3 Resting state network analysis

Finally we present our selected connections in terms of established six resting state networks (RSNs) [22]. The RSNs are listed in Table 7.

| No | Name | Abbreviation | No | Name | Abbreviation |
|---|---|---|---|---|---|
| **1** | Default mode network | DMN | 4 | Auditory Network | AN |
| **2** | Dorsal attention network | DAN | 5 | Somato motor network | SMN |
| **3** | Visual network | VN | 6 | Self-referential network | SRN |

**Table 7: RSN names and their abbreviations [22].**

We studied the altered functional connectivity in terms of these six RSNs and categorized them with respect to inter-RSN and intra-RSN alterations and evaluated increased and decreased functional connectivity in both groups and present them in Figure 23 and Figure 24.

The results show that DMN is affected the most with respect to inter-RSN and intra-RSN in terms of both increased and decreased connectivity. This pattern of alterations is followed by DAN: second highest affected network both in inter and intra RSN with respect to increased and decreased connectivity.

DMN is known to be associated with internal thoughts [22, 39] such as day dreaming, retrieving memories and envisioning the future. DAN is known to be associated with goal-directed stimulus response [22, 40]. Activity in DAN increases after paying concentration towards a particular thought.

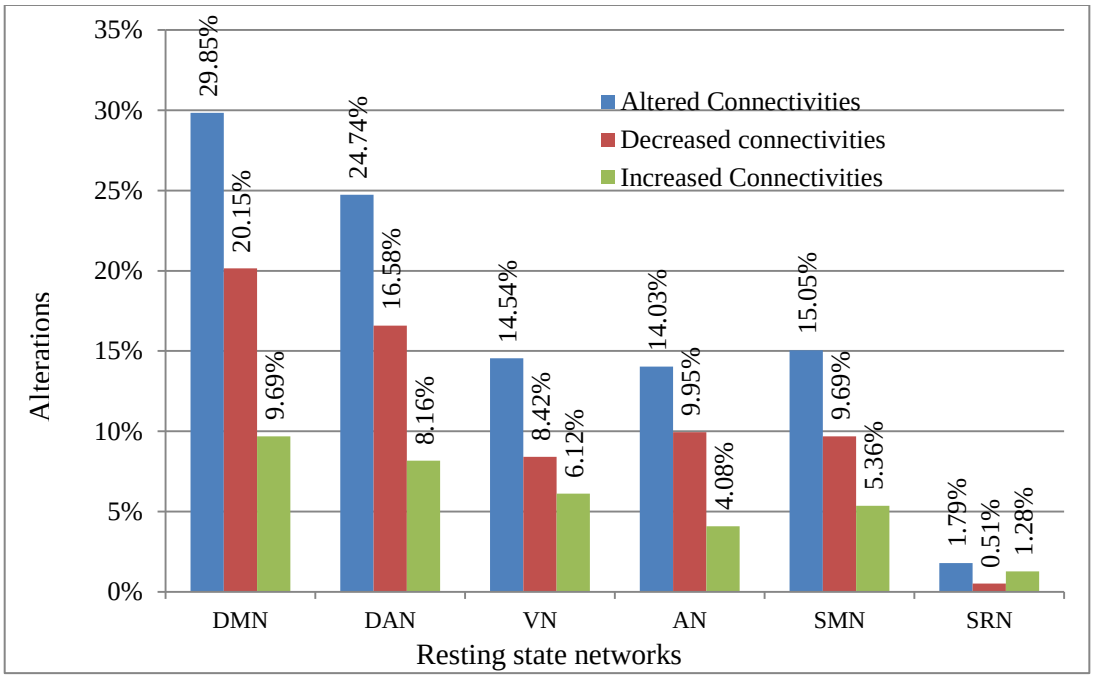


**Figure 23: Inter-RSN alterations. Increased, decreased and accumulated alterations for all six RSNs are presented. The decreased connectivity dominates in all RSNs as compared to increased connectivity. DMN is affected the most.**

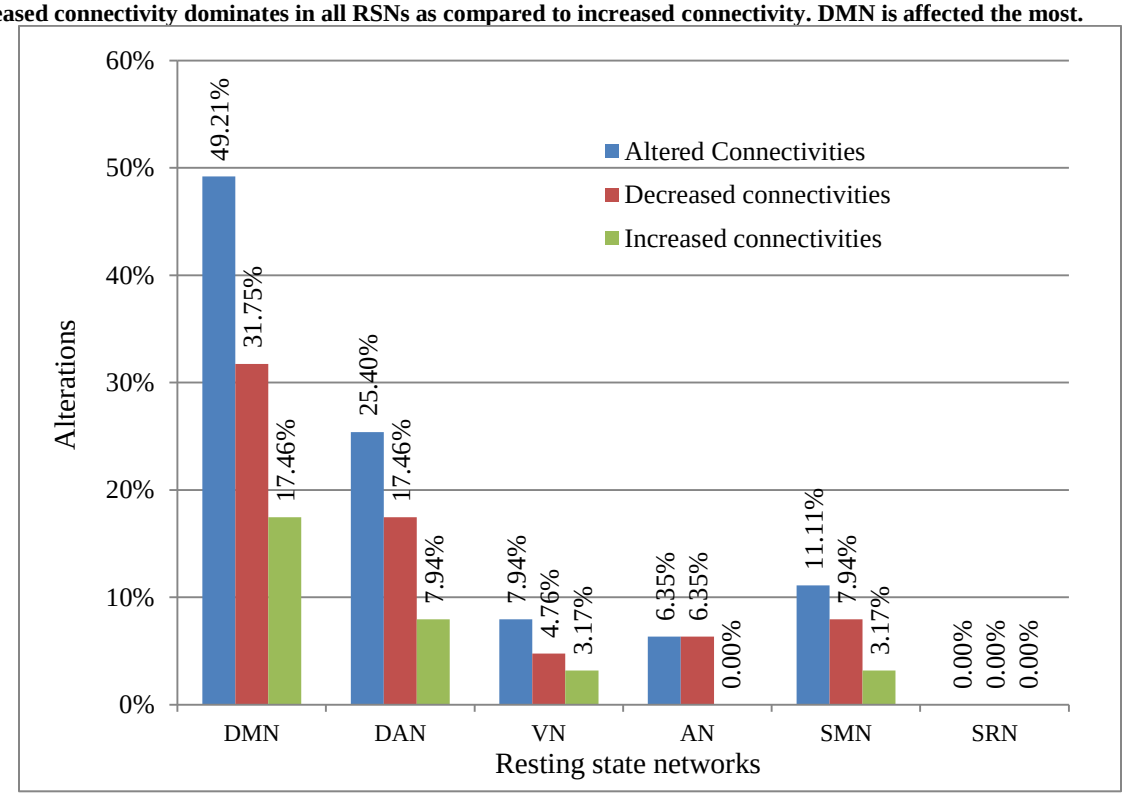


**Figure 24: Intra-RSN alterations. Increased, decreased and accumulated alterations for all six RSNs are presented. The decreased connectivity dominates in all RSNs as compared to increased connectivity. The DMN shows a dominant peak and is affected the most.**

## 4.3 Summary

The chapter presented the results and discussions of the study and explained the neurological findings. The proposed methodology is capable of achieving accuracy of 93.10%. The results suggest that default mode network is impaired the most as compared to other resting state networks in epileptic subjects. Results also demonstrate that parietal (pre) motor cortex is impaired the most as compared to other anatomical groups.


# Chapter 5

# Conclusions

This chapter concludes the thesis and suggests future work directions. The chapter is divided in two parts: i) conclusion ii) future work.

# 5.1 Conclusion

In this study, we addressed the problem of identifying functional connectivity differences induced by epilepsy and employing these discriminant connections to classify the brain state as healthy or epileptic. We proposed a new method for estimation of functional connectivity between brain regions and a new difference statistic which was shown to be able to discriminate between healthy and epileptic subjects. A novel clustering algorithm called affinity propagation was employed for estimation of functional connectivity between brain regions. A feature selection algorithm called difference statistics was proposed to identify discriminant connections. Using selected connections and support vector machine classifier, our methodology was able to achieve classification accuracy of 93.08%. Our results suggested that default mode network (DMN) is affected the most as compared to other resting state networks. Study also concluded that among all six anatomical groups, functional connectivity of parietal-(pre) motor group is affected the most. The group is known to be involved in movement intention and motor awareness.

The patients inducted in our study suffer from different sub types of epilepsy and we conducted experiments on all patients. Results of the study show that different kinds of epilepsy may share common altered functional connectivity pathways.

## 5.2 Future work

In our study, we are using automated anatomical labeling (AAL) template to segment brain in different regions. The AAL may be treated as having coarse resolution. All voxels lying within a region are averaged to yield value for that particular region. Studying with a higher dense resolution template may increase the classification accuracy by identifying individual regions more accurately. In dense resolution, number of voxels lying in a particular region will be reduced thereby reducing the possible effect of any outlier region. However, the dense resolution template may generate even high dimensions and may introduce *curse of dimensionality* problem. The problem can be addressed by applying two templates: coarse resolution and high resolution template. The regions of altered connections can be identified by using coarse template in first step. Based upon this information, a high resolution template may be applied to spot the changes more precisely on a small scale.

The proposed algorithm is expected to have broad applications in the study of other neurological disorders also. It remains a question of interest whether the methodology can be used to predict other neurological diseases also. Application of the methodology in neurological diseases like Alzheimer's disease, depression etc. may have some interesting findings where symptoms of these diseases are not obvious at early stages.


# References

1. Zhang, J., et al., *Pattern classification of large-scale functional brain networks: identification of informative neuroimaging markers for epilepsy*. PloS one, 2012. **7**(5): p. e36733.

2. Santé, O.m.d.l., *Neurological disorders: public health challenges*. 2006: World Health Organization.

3. Biswal, B.B., J.V. Kylen, and J.S. Hyde, *Simultaneous assessment of flow and BOLD signals in resting-state functional connectivity maps.* NMR in Biomedicine, 1997. **10**(45): p. 165-170.

4. Riaz, A., K. Rajpoot, and N. Rajpoot. *A connectivity difference measure for identification of functional neuroimaging markers for epilepsy*. in *Neural Engineering (NER), 2013 6th International IEEE/EMBS Conference on*. 2013. IEEE.

5. *Psych central*. Available from: http://psychcentral.com/.

6. Norman, K.A., et al., *Beyond mind-reading: multi-voxel pattern analysis of fMRI data.* Trends in cognitive sciences, 2006. **10**(9): p. 424-430.

7. Tzourio-Mazoyer, N., et al., *Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain.* Neuroimage, 2002. **15**(1): p. 273-289.

8. Van Den Heuvel, M.P. and H.E. Hulshoff Pol, *Exploring the brain network: a review on resting-state fMRI functional connectivity.* European Neuropsychopharmacology, 2010. **20**(8): p. 519-534.

9. Kaiming Li, L.G., Jingxin Nie, Gang Li, Tianming Liu, *Review of methods for functional brain connectivity detection using fMRI.* Computerized medical imaging ang graphics, 2009. **33**: p. 131-139.

10. Cao, J. and K. Worsley, *The geometry of correlation fields with an application to functional connectivity of the brain.* The Annals of Applied Probability, 1999. **9**(4): p. 1021-1057.

11. Sun, F.T., L.M. Miller, and M. D'Esposito, *Measuring interregional functional connectivity using coherence and partial coherence analyses of fMRI data.* Neuroimage, 2004. **21**(2): p. 647-658.

12. Friston, K.J., *Statistical parametric mapping*, in *Neuroscience Databases*. 2003, Springer. p. 237-250.

13. Haneef, Z., et al., *Effect of lateralized temporal lobe epilepsy on the default mode network.* Epilepsy & Behavior, 2012. **25**(3): p. 350-357.

14. Dupont, P., et al., *Dynamic perfusion patterns in temporal lobe epilepsy.* European journal of nuclear medicine and molecular imaging, 2009. **36**(5): p. 823-830.

15. Blumenfeld, H., et al., *Positive and negative network correlations in temporal lobe epilepsy.* Cerebral Cortex, 2004. **14**(8): p. 892-902.

16. Smith, S.M., et al., *Advances in functional and structural MR image analysis and implementation as FSL.* Neuroimage, 2004. **23**: p. S208-S219.

17. Zhang, Z., et al., *Impaired attention network in temporal lobe epilepsy: A resting FMRI study.* Neuroscience Letters, 2009. **458**(3): p. 97-101.

18. Stella, F. and J.A. Maciel, *Attentional disorders in patients with complex partial epilepsy.* Arquivos de neuro-psiquiatria, 2003. **61**(2B): p. 335-338.

19. Takaya, S., et al., *Prefrontal hypofunction in patients with intractable mesial temporal lobe epilepsy.* Neurology, 2006. **67**(9): p. 1674-1676.

20. Beckmann, C.F., et al., *Investigations into resting-state connectivity using independent component analysis.* Philosophical Transactions of the Royal Society B: Biological Sciences, 2005. **360**(1457): p. 1001-1013.

21. De Luca, M., et al., *fMRI resting state networks define distinct modes of long-distance interactions in the human brain.* Neuroimage, 2006. **29**(4): p. 1359-1367.

22. Mantini, D., et al., *Electrophysiological signatures of resting state networks in the human brain.* Proceedings of the National Academy of Sciences, 2007. **104**(32): p. 13170-13175.

23. Chen, S., et al., *Resting-state fMRI study of treatment-naïve temporal lobe epilepsy patients with depressive symptoms.* Neuroimage, 2012. **60**(1): p. 299-304.

24. Kanner, A.M., *Do epilepsy and psychiatric disorders share common pathogenic mechanisms? A look at depression and epilepsy.* Clinical neuroscience research, 2004. **4**(1): p. 31-37.

25. Chao-Gan, Y. and Z. Yu-Feng, *DPARSF: a MATLAB toolbox for "pipeline" data analysis of resting-state fMRI.* Frontiers in systems neuroscience, 2010. **4**.

26. Song, X.-W., et al., *REST: a toolkit for resting-state functional magnetic resonance imaging data processing.* PloS one, 2011. **6**(9): p. e25031.

27. Lang, E.W., et al., *Brain connectivity analysis: a short survey.* Computational intelligence and neuroscience, 2012. **2012**: p. 8.

28. Venkataraman, A., et al. *Robust feature selection in resting-state fMRI connectivity based on population studies*. in *Computer Vision and Pattern Recognition Workshops (CVPRW), 2010 IEEE Computer Society Conference on*. 2010.

29. Venkataraman, A., et al. *Robust feature selection in resting-state fmri connectivity based on population studies*. in *Computer Vision and Pattern Recognition Workshops (CVPRW), 2010 IEEE Computer Society Conference on*. 2010. IEEE.

30. Dueck, B.J.F.a.D., *Clustering by Passing Messages Between Data Points.* Science, 2007. **315**: p. 972-976.

31. Jiang, Z. and C. Huafu. *Analysis of activity in fMRI data for multitask experimental paradigm using affinity propagation clustering*. in *Computer and Automation Engineering (ICCAE), 2010 The 2nd International Conference on*. 2010.

32. Zhang, J., et al., *Analysis of activity in fMRI data using affinity propagation clustering.* Computer Methods in Biomechanics and Biomedical Engineering, 2011. **14**(03): p. 271-281.

33. Liu, D., W. Lu, and N. Zhong, *Clustering of fMRI Data Using Affinity Propagation*, in *Brain Informatics*, Y. Yao, et al., Editors. 2010, Springer Berlin Heidelberg. p. 399-406.

34. Saeys, Y., I. Inza, and P. Larrañaga, *A review of feature selection techniques in bioinformatics.* bioinformatics, 2007. **23**(19): p. 2507-2517.

35. Cortes, C. and V. Vapnik, *Support vector machine.* Machine learning, 1995. **20**(3): p. 273-297.

36. Han, J., M. Kamber, and J. Pei, *Data mining: concepts and techniques*. 2006: Morgan kaufmann.

37. Salvador, R., et al., *Neurophysiological architecture of functional magnetic resonance images of human brain.* Cerebral Cortex, 2005. **15**(9): p. 1332-1342.

38. Desmurget, M. and A. Sirigu, *A parietal-premotor network for movement intention and motor awareness.* Trends in Cognitive Sciences, 2009. **13**(10): p. 411-419.

39. Greicius, M.D., et al., *Functional connectivity in the resting brain: a network analysis of the default mode hypothesis.* Proceedings of the National Academy of Sciences, 2003. **100**(1): p. 253-258.

40. Corbetta, M. and G.L. Shulman, *Control of goal-directed and stimulus-driven attention in the brain.* Nature reviews neuroscience, 2002. **3**(3): p. 201-215.