Identify Brain Disorder Using Clustering Technique

[1]Priti Mithari, [2]Amruta Gadekar,
[1, 2] Computer Engineering Department
DYPIET, Pune, India

Abstract: The Brain activities are very complicated and difficult to understand. Many psychiatric disorders related to the brain are difficult to detect or identify. Functional magnetic resonance imaging (fMRI) provides the potential to study brain function. The basic signal of fMRI depends upon the blood-oxygen-level-dependent (BOLD) effect which helps to study human brain functions. The purpose of Clustering Technique is to understand the complex interaction patterns among brain regions as well as identify brain disorders. To detect clusters of objects with similar interaction patterns we proposed a partitioning clustering algorithm that is propose K-means (KM), an efficient algorithm for partitioning clustering. The K-means (KM), an efficient algorithm which begins with an initial guess to the cluster centers, and iteratively refines them an efficient algorithm for partitioning clustering..

IndexTerms **– Clustering techniques, K-means, fMRI technique.**

I. INTRODUCTION

Brain disorders or mental illnesses involve psychological or behavioral patterns that are usually associated with distress or disability, which are not part of the natural evolution of a person or culture. As such, mental disorders can be defined generally through a combination of features that reflect the feelings of a person or his actions and explain his thinking and perceptions. Many psychiatric disorders still neither be identified by biomarkers, nor by physiological or histological abnormalities of the brain. The brain functions are very complex and difficult to understand. To understand such complex functions and the psychiatric disorders of the brain, it is necessary to first understand the different brain activities of brain functions. Brain activity is the only resource to understand brain disorders. Functional Magnetic Resonance Imaging (FMRI) helps to study human brain function or activity. The Functional Magnetic Resonance Imaging (FMRI) measures the Blood Oxygen-Level Dependent (BOLD) signal. The main objective is to find out objects having a similar intrinsic interaction pattern to a common cluster as well as identify brain disorders by data clustering technique. To formalize this idea, this paper introduces a partitioning clustering - The K-means (KM)which is used to find out the similar interaction patterns among the brain and help to study brain related disorders.

Basic Concepts:
Clustering technique: Cluster is a collection of objects which are similar in their group and are dissimilar to the objects belonging to other groups. So mainly we use clustering technique to detect and make a group of similar patterns.Basically clustering technique is divided into partition method and hierarchical method.

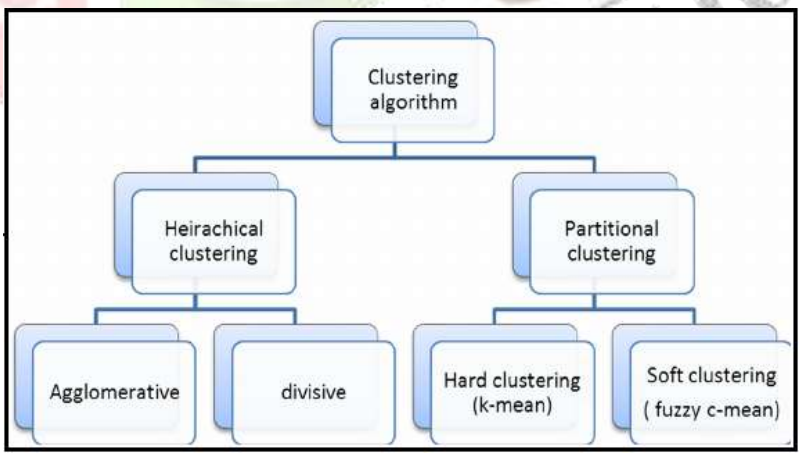

Fig. Clustering Techniques

EEG(Electro Encephalography)

Electroencephalography (EEG) is the measurement of the electrical activity of the brain by recording from electrodes placed on the scalp. The resulting traces are known as an electroencephalogram (EEG) and represent an electrical signal from a large number of neurons.

Functional Magnetic Resonance Imaging(FMRI)

Functional Magnetic Resonance Imaging (FMRI) helps to study human brain function in a noninvasive Way. The basic signal of FMRI depends upon the blood-oxygen-level-dependent (BOLD) effect, which allows indirect imaging brain activity by changes in the blood flow related to the energy consumption of brain cells. It detects the changes in blood oxygenation and flow that occur in response to neural activity. When a brain area is more active it takes more oxygen and to meet this increased demand blood flow increases to the active area. FMRI can be used to create activation maps showing which parts of the brain are involved in a particular mental process. FMRI data are time series of 3-dimensional volume images of the brain.

K-means (KM)

The K-means (KM) is the existing algorithm which is based on k-means technique. KM is a partitioning clustering used to detect clusters of objects with similar interaction patterns. The algorithm KM is a common technique for clustering multivariate time series using FMRI data. KM achieves good results on synthetic data and on real world data from various domains, but generally giving a good result on EEG and FMRI data. The interaction patterns detected by KM are easy to interpret. K Means clustering can handle big amount of data.

II. LITRATURE REVIEW

Sr. No.	Paper Title	Auther	Issues	Methods and Algorithm	Advantages and Disadvantage
1	Mining Interaction Patterns Among Brain Regions by Clustering	C.Plant,A. Zherdin, C.Sorg, A.Meyer-Baese, and A. M. Wohlschläger	There is no separate mechanism for feature selection	Interaction K-means Algorithm	**Advantages:** IKM achieves good results on synthetic data and on real world data. **Disadvantages:** It will take more time for clustering.
2	Feature Selection for Classification of Variable length Multiattribute Motions	Chuanjun C.Li, L. Khan, and B.Prabhakaran	Real time recognition of individual isolated motions accurately and efficiently.	Support Vector Machines (SVM)	**Advantages:** By applying Support Vector Machines (SVM) to the feature vectors, we can efficiently classify and recognize real world multi-attribute motion data. **Disadvantages:** Using only a single motion pattern in the database to recognize similar motions allows for less variation in a similar motion.
3	Clustering of time series data—a survey	T.Warren Liao	Clustering seasonality patterns of retail data, Cluster analysis of country's energy consumption, Discovering Consumer power consumption patterns for the segmentation of markets.	Relocation clustering, Relocation clustering, k-Means and fuzzy cmeans	**Advantages:** The goal of clustering is to identify structure in an unlabeled data set by objectively organizing data into homogeneous groups where the within group-object similarity is minimized and the between-group-object dissimilarity is Maximized. **Disadvantages:** None of these in the paper which included in this survey handle multivariate time series data with different length for each

					variable.
4	Approximate Clustering of Time Series Using Compact Model-based Descriptions	H.P.Kriegel, P.Kroger, A. Pryakhin, M.Renz, and A.Zherdin	The key issue for approximate clustering is of course to generate accurate results.	Mathematical Model, Efficient Approximative Clustering	**Advantages:** The great benefit is that the size of our approximation depends only on the number of coefficients of the model (i.e. the number of reference time series). **Disadvantages:** The distance computation requires rather high runtimes and, if the time series are indexed by a standard spatial indexing method such as the R-Tree or one of its variants, this index will perform rather bad due to the wellknown curse of dimensionality.
5	Model-based Classification of Data with Time Seriesvalued Attributes	C.Bohm, L.Laer, C. Plant, A. Zherdin	Kind of object representation is very natural and straightforward in many applications.	Model-based classifier, iterative stepwise Algorithm	**Advantages:** Classification decisions are supported by classspecific Interaction patterns within the time series of a data object. **Disadvantages:** It would be also Interesting to design a model-based classifiers for FMRI data and combined FMRI-EEG data sets which are very challenging because of the large number of time series in FMRI data.

III. PROPOSED SYSTEM

To understand the complex functions and the psychiatric disorders of the brain, we have to first understand the different brain activities. Brain activity is the only resource to understand psychiatric disorders. Functional magnetic resonance imaging (FMRI) helps to study human brain function.

PROBLEM DEFINITION

Many psychiatric disorders still neither be identified by biomarkers, nor by physiological or histological abnormalities of the brain. To identify such brain disorders as well as to understand the complex interaction patterns among brain regions, this paper introduces partitioning clustering K-Means(KM) algorithm. The KM algorithm will find out the similar interaction patterns among the brain which will help to study brain related disorders.

GOALS AND OBJECTIVES

Goals:

This System is developed with following goal:

- The system should find, the better the clusters, i.e having a better cluster purity.
- The system should use less memory.
- The system should form clusters time efficiently.
- It should help in identifying anatomical regions of brain involved in the abnormal behaviour of medical subject.

Objective:

- To study and understand the concept of identification of brain disorders.
- Understand and identify the complex interaction patterns among brain regions by data clustering technique.
- Reduce the time taken for clustering formation.
- Maintain accuracy in the final result.
- Implementation of clustering techniques KM (Interaction K-means).

SYSTEM ARCHITECTURE

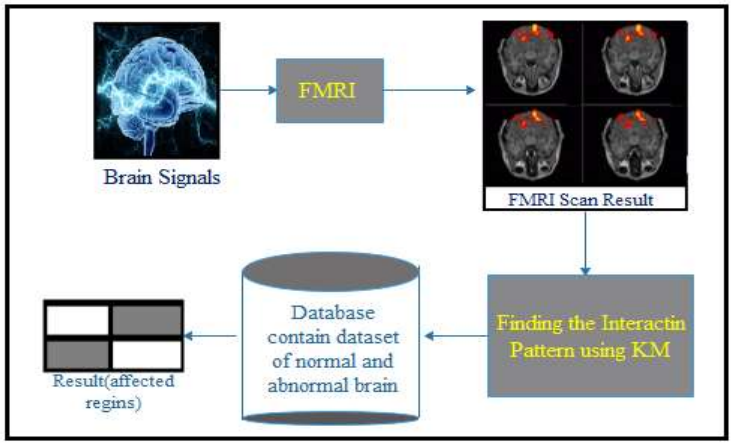

Fig. : System Architecture

The figure Shows the architecture of identification of brain disorders.The Architecture Explains how to find out whether the brain is affected or not.In the initial step brain will get scanned in the scanner machine. Functional magnetic resonance imaging, or FMRI, is a technique for measuring brain activity. It works by detecting the changes in blood oxygenation and flow that occur in response to neural activity. MRI detects changes in blood flow without using a radioactive tracer. When a particular site in the brain is more active, blood flows to that area. This blood brings oxygen to the hard-working brain cells. By tracking variations in blood flow, functional MRI can detect activity in the brain as it happens.

In this process the FMRI result (scan result) will be used to process. The K-Means algorithm is used for finding the similar interaction patterns among the brain regions. The database contains two kinds of cluster sets. One is a normal brain cluster and another one is the diseased brain cluster. The normal brain cluster contains the properties of normal human brain and the diseased brain cluster contain the properties of abnormal(diseased) human brain. If the patterns are matched with normal dataset means the brain is not affected by any disorder and if the patterns are matched with diseased data set means the patients is having the brain disorder. In this way the result will be given.

ALGORITHM

K means is an iterative clustering algorithm that aims to find local maxima in each iteration. This algorithm works in these 5 steps:

1. Specify the desired number of clusters K :
2. Randomly assign each data point to a cluster.
3. Compute cluster centroids
4. Re-assign each point to the closest cluster centroid: Note that only the data point at the bottom is assigned to the red cluster even though it's closer to the centroid of grey cluster.
5. Re-compute cluster centroids: Now, re-computing the centroids for both the clusters.

ADVANTAGES

- K Means clustering can handle big data well. This is because the time complexity of K Means is linear i.e. O(n).
- K Means is found to work well when the shape of the clusters is hyper spherical (like circle in 2D, sphere in 3D).
- K Means clustering requires prior knowledge of K i.e. no. of clusters you want to divide your data into. But, you can stop at whatever number of clusters you find appropriate in hierarchical clustering by interpreting the dendrogram

IV. CONCLUSION

The proposed system is responsible for finding the similar interaction patterns among the brain region as well as used to identify the brain disorders. The system uses K-Means algorithm which begins with an initial guess to the cluster centers, and iteratively refines them an efficient algorithm for partitioning clustering. The clustering technique-K-means used to define a cluster of interaction patterns of healthy and diseased object which is used to detect the brain disorder. K Means is a first clustering technique which handles big data.The K-mean clustering method achieves good result with high accuracy.

REFERENCES

[1] M. D. Fox and M. E. Raichle, Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging,Nat. Rev. Neurosci., vol. 8, no. 9, pp. 700711, 2007.
[2] C. Li, L. Khan, and B. Prabhakaran, Feature selection for classification of variable length multiattribute motions, in Multimedia Data Mining and Knowledge Discovery, V. A. Petrushin and L. Khan, Eds. London, U.K.: Springer, 2007.
[3] T. W. Liao, Clustering of time series data A survey, Pattern Recognit., vol. 38, no. 11, pp. 18571874, 2005.
[4] H.-P. Kriegel, P. Krger, A. Pryakhin, M. Renz, and A. Zherdin, Approximate clustering of time series using compact modelbased descriptions, in Proc. DASFAA, New Delhi, India, 2008, pp. 364379.
[5] C. Bhm, L. Ler, C. Plant, and A. Zherdin, Model-based classification of data with time series-valued attributes, in Proc. BTW,2009, pp. 287296.
[6] P. Cheeseman and J. Stutz, Bayesian classification (autoclass): Theory and results, in Proc. Adv. Knowl. Discov. Data Min., Menlo Park, CA, USA, 1996, pp. 153180.
[7] A. Veeraraghavan and A. K. R. Chowdhury, The function space of an activity, in Proc CVPR,New York,NY, USA, 2006, pp. 959968.