International Journal of Emerging Trend in Engineering and Basic Sciences (IJEEBS)
ISSN (Online) 2349-6967 *Volume 2 , Issue 1(Jan-Feb 2015), PP1-5*

Identification Of Brain Disorders Using Data Clustering Technique

Amruta Gadekar[1], Prof.Ravi Patki[2]
[1 2] (Computer Department,Dr.D.Y.Patil College of Engg., Pune University , India)

Abstract:- — Brain is a central as well as important part of a human body. Brain activities are very complicated and difficult to understand. Functional Magnetic Resonance Imaging (FMRI) helps to study human brain functions. To better understanding of complex brain activities, we have to first understand complex interplay among the brain regions so for that we propose clustering a technique. The purpose of Clustering Technique is to understand the complex interaction patterns among brain regions as well as Identify Brain disorders. Based on this Clustering Technique, we propose Interaction K-means (IKM), a partitioning clustering algorithm suitable to detect clusters of objects with similar interaction patterns and The Expectation Maximization(EM) algorithm with Gaussian Mixtures is very similar to k-Means algorithm. The algorithm begins with an initial guess to the cluster centers, and iteratively refines them an efficient Algorithm for partitioning clustering.

Keywords :- Clustering, Interaction K-means Clustering (IKM), Interpretation of Clustering, Interaction Regions of Brain.

I. INTRODUCTION

Brain is an important part of human body which performs many functions simultaneously. The brain functions are very complex and difficult to understand. To understand the complex functions and the psychiatric disorders of the brain, we have to understand the different brain activities. Brain activity often is the only resource to understand psychiatric disorders. Functional magnetic resonance imaging (FMRI) helps to study human brain function in a non-invasive way. The basic signal of FMRI depend upon the blood-oxygen-level-dependent (BOLD) effect, which allows indirectly imaging brain activity by changes in the blood flow related to the energy consumption of brain cells (see Figure2).

FMRI data are time series of 3-dimensional volume images of the brain. A typical statistical analysis involves comparing groups of subjects or different experimental conditions based on univariate statistical tests on the level of the single 3-d pixels called voxels. These voxels data from FMRI experiments are massive in volume, since these data represent complex brain activity. To make more effective and efficient multivariate data mining methods we need more potential information. To get a better understanding of complex brain activity, it is essential to understand the complex interplay among brain regions during task and at rest.

A cluster is a collection of objects which are “similar” between them and are “dissimilar” to the objects belonging to other clusters.So mainly we use clustering techinqic to detect and make a group of similar patterns.To find the similar interaction patterns among the brains regions we use this clustering technique.With the help of clustering techniques we capture the different interaction patterns in healthy and diseased subjects.

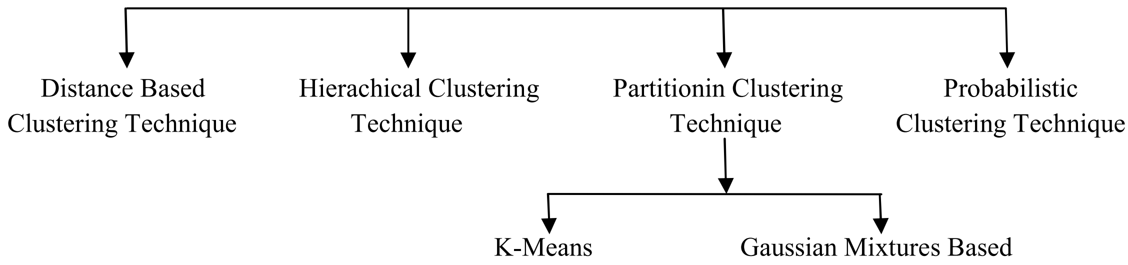


Fig.1: Classification of Clustering Technique

There are many clustering techniques are present such as Distance-based,Hierarchical,Partitioning, Probabilistic.Partitioning clutering techniqic divide the data into proper subset and recursively go through each subset and relocate points between clusters.In the partitioning clutering techniqic there are mainly k-means method and Gaussian Mixtures Based method.

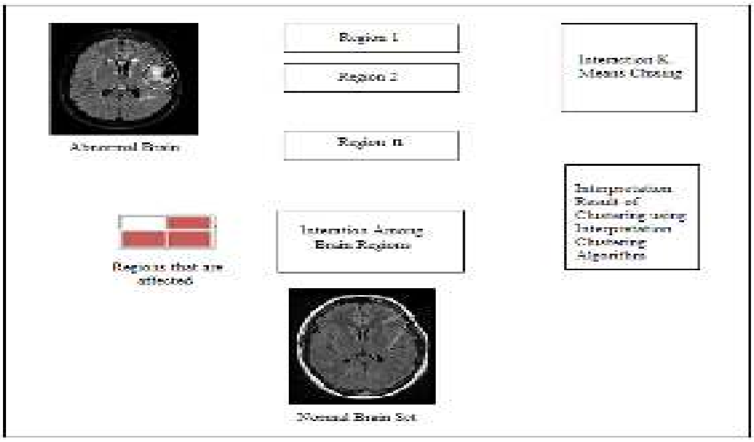


Fig.2. Flow of System Architecture

The Interaction K-means (IKM) is base on k-means techniqic.IKM is a partitioning clustering used to detect clusters of objects with similar interaction patterns. The algorithm IKM is a general technique for clustering multivariate time series and not limited to FMRI data. A cluster analysis of motion stream data potentially identifies clusters with similar movements, usually performed by different persons. Most of the techniques consider either the time series as whole as the data objects to be subjected to a cluster analysis or perform clustering on subsequence's which allows defining more meaningful similarity measures in many applications. Many of the univariate methods mentioned so far can be straightforwardly extended to the multivariate case. However, by doing so, information is lost: Data which is inherently multivariate often contains interactions between the different time series.

II. INTERACTION K-MEANS CLUSTERING

The Interaction K-means (IKM) a partitioning clustering used to detect clusters of objects with similar interaction patterns. IKM is iterative algorithm which minimizes the clustering objective.

Algorithm IKM:

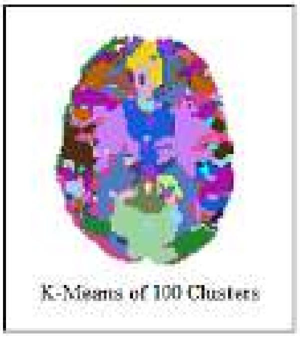


Fig.3: K-Means Clusters

IKM is iterative partitioning clustering algorithm; IKM follows a similar algorithmic paradigm as K-means. In the IKM, first step is initialization. For initialization we partition Data set (DS) into K cluster. For IKM it is favourable that the initial clusters are balanced in size to avoid over fitting. Therefore, we partition the data set into K equally sized random clusters and find a set of models for each cluster. After initialization, IKM iteratively performs mainly two steps that are: In the assignment step, each object O is assigned to the cluster w.r.t. which the error is minimal, i.e. O.cid = minC∈C EO,C.It is easy to see that this minimizes the objective function After assignment, in the update step, the models of all clusters are reformulated. Pseudo code of IKM is provided (see Figure 4).

```
algorithm IKM (data set DS, integer K):
Clustering C
  Clustering best Clustering;
//initialization
     for init := 1...maxi nit do
         C := randomInit(DS,K);
     for each C ∈C do
         MC := find Model(C);
     while not converged or iter < maxIter do
//assignment
     for each O ∈ DS do
         O.cid = minC∈CEO,C
//update
     for each C ∈C do
        MC := findModel(C);
     if improvement of objective function
     best Clustering :=C;
     end while
     end for
return best Clustering;
```

Fig.4: Algorithm Interaction K-means

III. INTERPRETATION OF THE CLUSTERING

For interpretation of the clustering result, we consider, a subset of the models which best differentiates among the clusters. For each pair of clusters, the best discriminating models are selected by leave-one-out validation using objects of the corresponding clusters. To find the discriminate among the cluster, we consider error with the correct cluster of test object. Correct cluster of test object with positive sign and error object cluster with negative sign (see Figure 5). Finally the Clustering result together with the information about which models best discriminate among cluster is a good basis for user interaction.

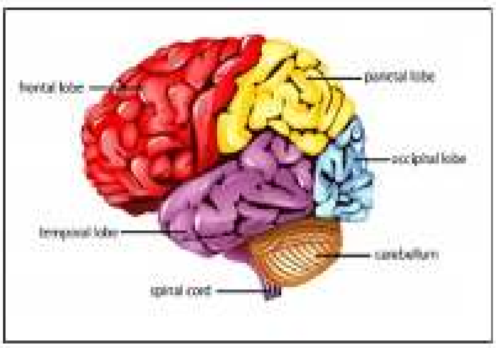
```
algorithm dimension Ranking
(Cluster Ci, Cluster Cj): ranking
      error in models := new ARRAY [d];
//leave-one-out-validation
     for each O ∈OCi ∪OCj do
         test := O
         OCi :=OCi \test; OCj :=OCj \test;
         findModel(Ci); findModel(Cj)
      for each cluster ∈{Ci,C j} do
                     if O.cid = cluster.id then
                sign := 1;
           else
                sign :=−1;
           end if
           for i := 1...d do
error in models[d]+ = sign∗ cluster.models[d]
 calculateErrorFrom(O. getTimeSeries(d));
           end for
        end for
   end for
   sort(error in models);
 return error in models;
```

Fig.5: Algorithm Interpretation Of The Results.

IV. Interaction among the Brain Regions

4.1. Functional magnetic resonance imaging (FMRI)

Functional Magnetic Resonance Imaging (FMRI) help to study human brain functions. Functional MRI generates a series of 3-D volume images of the brain. Each image consists of about 60,000 voxels and the interval between time points is about 2-3 seconds. Each voxel can use time series from the images. However, for neighbouring voxels signal activity is very similar.



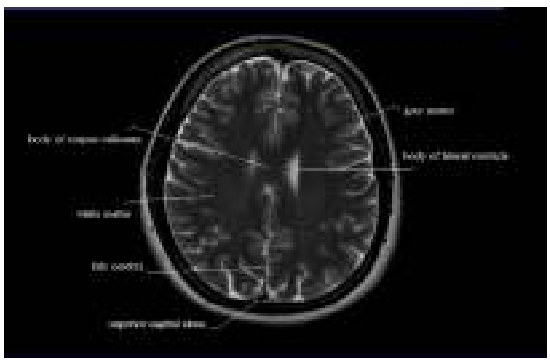


Fig.6: Brain Regions **Fig.7: MRI of Brain**

4.2. Somatoform Pain Disorder

Somatoform Pain Disorder has severe impact on the quality of living of the affected persons since the main symptom is severe and prolonged pain for which there is no medical explanation. The causes of this psychiatric disorder are not fully understood but the hypothesis is that patients have alternating mechanisms of observing and processing pain. Therefore subjects underwent alternating blocks of pain and non-painful stimulation while in the scanner. After pre-processing segment the data of each subject into 90 anatomical regions of interest (ROIs).

4.3. Schizophrenia

Schizophrenia is characterized by the impaired interaction between distributed brain regions particularly the striatum. Increased dopamine activity in the striatum is essential for schizophrenia and anti-dopaminergic treatment the main therapy of the disorder. Intrinsic brain networks are characterized by synchronous brain activity at rest.

5. CONCLUSIONS

Here, we propose an Interaction–based K-Means (IKM) clustering technique. We define cluster as a set of subjects sharing a similar interaction pattern among their brain regions with interaction k-means. IKM achieves good result on data from various domains. IKM algorithm is especially excellent results on FMRI data. IKM also scalable, robust against noise and dimensions. With the interpretation result we will obtain best discriminate among clusters which is good for user interaction. So with the IKM we understand and identify the complex interaction patterns among brain regions by data clustering technique which will help to find psychiatric disorders of the brain.

6. ACKNOWLEDGMENT

we would like to thanks to the publishers, researchers for making their resources available and teachers for guidance. We also thank to the college authority for supporting to us and providing required information. We would also like to thank our friends and family members.

REFERENCES

[1] M. D. Fox and M. E. Raichle, “Spontaneous fluctuations in brain activity observed with functional Rev. Neurosci., vol. 8, no. 9, pp. 700–711, 2007.

[2] X. Wang, A. Wirth, and L. Wang, “Structure-based statistical feature and multivariate time series clustering,” in Proc. ICDM, Omaha, NE, USA, 2007, pp. 351–360.

[3] X. Ge and P. Smyth, “Deformable Markov model templates for time-series pattern matching,” in Proc. KDD, New York, NY, USA,2000, pp. 81–90.

[4] L. Owsley, L. Atlas, and G. Bernard, “Automatic clustering of vector time-series for manufacturing machine monitoring,” in Proc. IEEE ICASSP, vol. 4. Munich, Germany, 1997, pp. 3393–3396.

[5] X. Z. Wang and C. Mc Greavy, “Automatic classification for mining process operational data,” Ind. Eng. Chem. Res., vol. 37, no. 6.

[6] F. Mörchen, “Time series feature extraction for data mining using DWT and DFT,” Dept. Math. Comput. Sci., University of Marburg, Marburg, Germany, Tech. Rep. 33, 2003.

[7] A. Sudjianto and G. S. Wasserman, “A nonlinear extension of principal component analysis for clustering and spatial differentiation,”IIE, vol. 28, no. 12, pp. 1023–1026, 1996.

[8] A. Trouve and Y. Yu, “Unsupervised clustering trees by nonlinear principal component analysis,” Pattern Recognit. Image Anal., vol. 2, pp. 108–112, 2001.

[9] E. H. C. Wu and P. L. H. Yu, “Independent component analysis for clustering multivariate time series data,” in Proc. ADMA, Wuhan, China, 2005, pp. 474–482..S. M. Metev and V. P. Veiko, Laser Assisted Microtechnology, 2nd ed., R. M. Osgood, Jr., Ed. Berlin, Germany: Springer-Verlag, 1998.

[10] S. Smith et al., “Correspondence of the brain’s functional architecture during activation and rest,” PNAS, vol. 106, no. 31, p. 13040, 2009.

[11] C. Li, L. Khan, and B. Prabhakaran, “Feature selection for classificationof variable length multiattribute motions,” in Multimedia Data Mining and Knowledge Discovery, V. A. Petrushin and L. Khan, Eds. London, U.K.: Springer, 2007.

[12] C. Pamminger, Bayesian Clustering of Categorical Time Series:An Approach Using Finite Mixtures of Markov Chain Models.Saarbrücken, Germany: VDM Publishing Group, 2008.

[13] T. W. Liao, “Clustering of time series data – A survey,” Pattern Recognit., vol. 38, no. 11, pp. 1857–1874, 2005.

[14] C. Plant, A. M. Wohlschläger, and A. Zherdin, “Interaction-based clustering of multivariate time series,” in Proc. ICDM, Miami, FL, USA, 2009, pp. 914–919.

[15] F. Mörchen, “Time series feature extraction for data mining using DWT and DFT,” Dept. Math. Comput. Sci., University of Marburg, Marburg, Germany, Tech. Rep. 33, 2003.

[16] A. R. McIntosh and B. Miic, “Multivariate statistical analyses for neuroimaging data,” Annu. Rev. Psychol., vol. 64, no. 1, pp. 499–525, 2013 [Online]. Available: http://www.annualreviews.org/doi/abs/ 10.1146/ annurev-psych-113011-143804.

[17] J. B. MacQueen, “Some methods for classification and analysis of multivariate observations,” in Proc. 5th Berkeley Symposium Mathematical Statistics and Probability, vol. 1, L. M. L. Cam and J. Neyman, Eds. Berkeley, CA, USA: University of California Press, 1967, pp. 281–297.

[18] A. Hyvärinen and E. Oja, “Independent component analysis: Algorithms and applications,” Neural Netw., vol. 13, no. 4-5, pp. 411–430, 2000.

[19] L. Xiao, H. Begleiter, B. Porjesz, W. Wenyu, and A. Litke, “Event related potentials during object recognition tasks,” Brain Res. Bull., vol. 38, pp. 531–538, 1995 [Online]. Available: http://www.ingentaconnect.com/content/els/03619230/1995/00 000038/00000006/art02023.

[20] A. Veeraraghavan and A. K. R. Chowdhury, “The function space of an activity,” in Proc. CVPR,New York,NY, USA, 2006, pp. 959–968.

[21] M. Kudo, J. Toyama, and M. Shimbo, “Multidimensional curve classification using passing-through regions,” Pattern Recognit.Lett., vol. 20, no. 11-13, pp. 1103–1111, 1999.

[22] B. Dom, “An information-theoretic external cluster-validity measure,”IBM Research Division, Tech. Rep. RJ 10219, May 2001 [Online]. Available: citeseer.nj.nec.com/dom01informationtheoretic.html.

[23] H. Heekeren, S. Marrett, and L. Ungerleider, “The neural systems that mediate human perceptual decision making,” Nat. Rev. Neurosci., vol. 9, no. 6, pp. 467–79, Jun. 2008.