FSSIKM: a Novel Approach for Brain Interaction Patterns

K.Vidhyadevia[1], M.Beema Mehraj[2], Dr.K.P.Kaliyamurthie[3]

Dept of CSE, Bharath University, 173, Agaram Road, Selaiyur, Chennai, India[1,2,3.]

ABSTRACT*:* Functional magnetic resonance imaging (FMRI) patterns provides the prospective to study brain function in a non-invasive way. The FMRI data are time series of 3-dimensional volume images of the brain. The data is traditionally analyzed within a mass-univariate framework essentially relying on classical inferential statistics. Handling of feature selection and *clustering* is a complicated process in Interaction patterns of brain datasets. To understand the complex interaction patterns among brain regions our system proposes a novel clustering technique. Our system models each subject as multivariate time series, where the single dimensions represent the FMRI signal at different anatomical regions. In our proposed system, there are three algorithms are used to mining the brain interaction pattern such as FSS, IKM and Dimension Ranking Algorithm. Feature subset selection (FSS) is a technique to pre-process the data before performing any data mining tasks, e.g., classification and clustering. This technique was used to choose a subset of the original features to be used for the subsequent processes. Hence, only the data generated from those features need to be collected. After that, select the key features in the preprocessed dataset based on the threshold values. Interaction K-means (IKM), a partitioning clustering algorithm used to detect clusters of objects with similar interaction patterns classification and clustering. Finally, Dimension Ranking algorithm was used to select the best cluster for assuring best result.

KEYWORDS: Data mining, Clustering, Interaction K-means Algorithm, Feature Subset Selection, Dimension Ranking, FMRI.

I.   INTRODUCTION

Feature selection for interaction based clustering in interaction patterns among brain images by using feature subset selection algorithm. Feature selection for interaction based clustering in interaction patterns among brain images by using feature subset selection algorithm to provide efficient selection process. The project is to analyze the brain interaction patterns for identifying the state of the brain. Functional magnetic resonance imaging (FMRI) opens up the opportunity to study human brain function in a noninvasive way. The basic signal of FMRI relies on 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. Recently, resting-state FMRI has attracted considerable attention in the neuroscience community. In resting state FMRI, subjects are instructed to just close their eyes and relax while in the scanner. FMRI data are time series of 3-dimensional volume images of the brain. FMRI produce images at higher resolution than other scanning techniques. Data from FMRI experiments are massive in volume with more than hundred thousands of voxels and hundreds of time points. Since these data represent complex brain activity, also the information content can be expected to be highly complex.

Recent findings suggest a modular organization of the brain into different functional modules. In tracking cognitive processes with functional MRI mental chronometry it says that Functional magnetic resonance imaging (FMRI) is used widely to determine the spatial layout of brain activation associated with specific cognitive tasks at a spatial scale of millimeters. Recent methodological improvements have made it possible to determine the latency and temporal structure of the activation at a temporal scale of few hundreds of milliseconds. Despite the sluggishness of the hemodynamic response, FMRI can detect a cascade of neural activations - the signature of a sequence of cognitive processes. Decomposing the processing into stages is greatly aided by measuring intermediate responses. By combining event-related FMRI and behavioral measurement in experiment and analysis, trial-by-trial temporal links can be

established between cognition and its neural substrate. To obtain a better understanding of complex brain activity, it is essential to understand the complex interplay among brain regions during task and at rest. Inspired by this idea, a novel technique is proposed for mining the different interaction patterns in healthy and diseased subjects by clustering. At the core of method is a novel cluster notion: A cluster is defined as a set of subjects sharing a similar interaction pattern among their brain regions.

In addition, time series are a special type of data set where elements have a temporal ordering. Therefore clustering of such data stream is an important issue in the data mining process. Numerous techniques and clustering algorithms have been proposed earlier to assist clustering of time series data streams. The clustering algorithms and its effectiveness on various applications are compared to develop a new method to solve the existing problem. After standard pre-processing including parcellation into anatomical regions, each subject is modeled as a data object which is represented by a multivariate time series. The nature of the considered images and the objective of the segmentation being multiple, there is no unique technique for image segmentation and segmenting an image into meaningful regions remains a real challenge The algorithm is used to evaluate the values within a regional time and group them together based on the merging criteria, resulting in a smaller list and more number of information is collected and compared with the database and finally abnormality can be detected.

The layout of the paper is as follows. In section II, address the above mentioned techniques and also give a brief on the literature being reviewed for the same. Section III, presents a comparative study of the various research works explored in the previous section. Section IV, describes about future work. Section V gives the conclusion in and lastly provides references.

II.   RELATED WORK

In this paper [1] IKM achieves good results on synthetic data and on real world data. It is scalable and robust against noise. Ranking Algorithm improves the efficiency of clustering result there is no separate algorithm for feature selection process Complexity is high. Auto class technique is not applicable if the number of time points varies among the objects, a case frequently occurring in FMRI data. There is no separate mechanism for feature selection

In this paper [2] The k-means method is in fact a perfect candidate for smoothed analysis: it is extremely widely used, it runs very fast in practice, and yet the worst-case running time is exponential. It did not make a huge effort to optimize the exponents as the arguments are intricate enough even without trying to optimize constants. The smoothed analyses so far are unsatisfactory as the bounds are still super-polynomial in the number n of data points

In this paper [3] Classification decisions are supported by class-specific interaction patterns within the time series of a data object. It would be also interesting to design 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 kind of object representation is very natural and straightforward in many applications, not much research on data mining methods for objects of this particular type**.**

In this paper [4] This leads to a new view about the neural basis of human perceptual decision-making processes. It is also tempting to speculate that the general principles derived from the studies of simple perceptual decision processes reviewed here extend to other settings. Seen in this light, it is not surprising that motor structures seem to have a role in decision formation. It is not yet clear how these structures contribute to decisions that are not linked to particular actions.

In [5] author has monitored 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. None of these in the paper which included in this survey handle multivariate time series data with different length for each variable. Clustering seasonality patterns of retail data, Cluster analysis of country's energy consumption, discovering consumer power consumption patterns for the segmentation of markets

In this paper [6] 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). 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 [1] or one of its variants, this index will perform rather badly due to the well-known curse of dimensionality.

In this paper [7] By applying Support Vector Machines (SVM) to the feature vectors, we can efficiently classify and recognize real world multi-attribute motion data using only a single motion pattern in the database to recognize similar motions allows for less variations in similar motion real time recognition of individual isolated motions accurately and efficiently.

In [8] author has monitored the shape deformation patterns in Anatomical structures show up evidently. Shrinkage of the hippocampus and Cortical and sub-cortical gray matter along the directions.

In this paper [9]Provide a systematic approach to learn the nature of such time warps while simultaneously allowing for the variations in descriptors for actions. Activity recognition is not very robust to intra- and inter-personal changes of the same action, and are extremely sensitive to warping of the temporal axis due to variations in speed profile activity recognition are not very robust to intra- and inter-personal changes.

In this paper [10] This should provide enough information to allow anyone to reproduce Auto Class, or to use the same evaluation functions in other contexts where these models might be relevant. A related problem is unsupervised classification, where preparation cases are also unlabeled.

III.   ISSUES IN CURRENT SYSTEM

In the existing system they have merge all dimensions of multivariate time series into a long univariate time and supplying it to the Auto class algorithm. This technique is not applicable if the number of time points varies among the objects, a case frequently occurring in FMRI data. They further compare their method to structure-based Statistical Features clustering (SF) which has been especially designed for clustering multivariate time series. SF represents each attribute of the multivariate time series by a fixed-length vector whose components are statistical features of the time series which capture the global structure.

IV.   PROPOSED METHOD

In our proposed system, we are presenting a novel technique to overcome on the existed issues. Here, we are proposing FSS (Feature Subset Selection) method chooses a subset for the original features for using in the subsequent processes. Therefore, the generated data has needed to collect from those features. We propose interaction K-means (IKM), an efficient algorithm for partitioning clustering. An extensive experimental evaluation on benchmark data demonstrates the effectiveness and efficiency of our approach. The results on two real FMRI studies demonstrate the potential of IKM to contribute to a better understanding of normal brain function and the alternations characteristic for psychiatric disorders. Our approach Interaction K-means (IKM) simultaneously clusters the data and discovers the relevant cluster-specific interaction patterns. The algorithm IKM is a general technique for clustering multivariate time series and not limited to FMRI data. Our cluster notion requires a similarity measure, which is very different to LP metric distances. The similarity measures applied in IKM are the errors with respect to the set models of a cluster. Interaction K-means (IKM), a partitioning clustering algorithm suitable to detect clusters of objects with similar interaction patterns. IKM outperforms state-of-the-art techniques for clustering multivariate time series on synthetic data as well as on benchmark data sets from different applications.

A IKM Algorithm Description:

Analogously to K-means, the first step of IKM is the initialization. As a common strategy for K-means, we propose to run IKM several times with different random initializations and keep the best overall result. For initialization, we randomly partition DS into K clusters. For IKM it is favorable 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 as described in the previous section. After initialization, IKM iteratively performs the following two steps until convergence: In the assignment step, each object O is assigned to the cluster which the error is minimal; IKM converges as soon as no object changes its cluster assignment during two consecutive iterations. Usually, a fast convergence can be observed, but there are some rare cases in which IKM does not converge. Analogously to standard K-means, it can be straightforwardly proven that the assignment and the update step strictly decrease the objective function. However, due to the greedy stepwise algorithm applied for model finding. O.cid = $\min_{C \in C}$ 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... As an iterative partitioning clustering algorithm, IKM follows a similar algorithmic paradigm as K-means. However, note that there are significant differences: Our cluster notion requires a similarity measure, which is very different to LP metric distances. The similarity measure applied in IKM is the errors with respect to the set models of a cluster. This similarity measure is always evaluated between an object and a cluster, and not between two objects, In contrast to K-means or K-medoid algorithms.

Algorithm
Input:
Data set DS,
 Integer K
Clustering C
Output:
Clustering bestClustering;
Method:
Step 1: randomly initialize the dataset into k- clusters.
   for init := 1 . . . maxInit do
   C := randomInit(DS,K);
Step 2: partition the data set into *K* equally sized random clusters and finds a set of models for each cluster.
   for each C ∈ C do
   MC := findModel(C);
   while not converged or iter < maxIter do
Step 3: assign objects for the minimal error rate clusters and find the minimal error rate for each cluster.
   for each O ∈ DS do
   O.cid = $\min_{C \in C}$ EO,C
Step 4: Repeat this process until achieve the best clustering.
   for each C ∈ C do
   MC := findModel(C);
   if improvement of objective function
   End

B ComponentIllustration

FSSIKM: a Novel Approach for Handling Brain Interaction Patterns is Implemented in four different module and they are:
- Dataset preparation and storage
- Feature selection
- Clustering (Interaction K-means Clustering)
- Dimension Ranking Algorithm

a Dataset preparation and storage

  Functional magnetic resonance imaging (FMRI) provides the potential to study brain function in a non-invasive way. FMRI data are time series of 3-dimensional volume images of the brain. The data is traditionally analyzed within a mass-univariate framework essentially relying on classical inferential statistics. 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. Collect the FMRI datasets in multi variate time series for different brain regions.

b Feature Selection

Select the features in preprocessed functional magnetic resonance image (FMRI) dataset before clustering. Feature Subset Selection method chooses a subset of the original features to be used for the subsequent processes. Hence, only the data generated from those features need to be collected. In this algorithm, initially collect the preprocessed dataset then find out the ranges of the respected dataset. After that, select the key features in the preprocessed dataset based on the threshold values.

c Clustering (Interaction K-means Clustering)

In this module, Interaction K-means (IKM), a partitioning clustering algorithm suitable to detect clusters of objects with similar interaction patterns. IKM outperforms state-of-the-art techniques for clustering multivariate time series on synthetic data as well as on benchmark data sets from different applications. Optimization of clustering process was depends on feature subset selection process. In brief, randomly initialize the dataset into k- clusters. Then partition the data set into K equally sized random clusters and finds a set of models for each cluster. After that, assign objects for the minimal error rate clusters and find the minimal error rate for each cluster. Repeat theses steps for finding the best clustering.

d Dimension Ranking Algorithm

In this module, Best Ranking algorithm was used to select the best cluster for assuring best result. For each pair of clusters, the best discriminating models are selected by leave-one-out validation using objects of the corresponding clusters. The top-ranked models are best discriminate among the clusters. In brief, this algorithm, initially generate the models of each individual cluster of from the training data. Then, compute the error of the test objects we know that all models and sum up all errors after that, To obtain a ranking of the models regarding their ability to discriminate among the clusters, we consider errors i.e., the correct cluster of the test object with a positive sign. Finally, sort all models ascending according to the error. The top-ranked models best discriminate among the clusters.

V. RESULT

This section shows the screen shots of the result. The results for proposed system i.e. FSSIKM: a Novel Approach for Handling Brain Interaction Patterns are obtained by Time taken Clustering Feature Selection
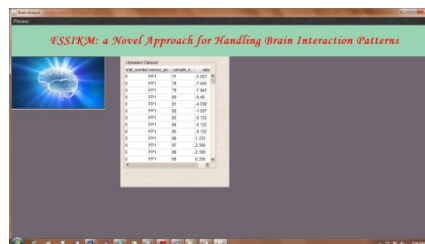

Fig. 1 Data sets are uploaded

Data sets are uploaded in Dataset preparation and storage collects the FMRI datasets in multi variants time series for different brain regions. FMRI data are time series of 3-dimensional volume images of the brain. Different experimental conditions based on univariate statistical tests on the level of the single 3-d pixels called voxels.
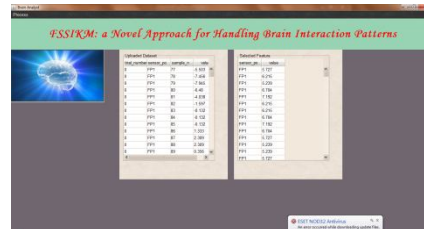
Fig. 2 Data sets are uploaded and Feature Selected data set.

Data sets are uploaded and Feature Selected data set in Feature Selection collects the preprocessed dataset then find out the ranges of the respected dataset. After that, select the key features in the preprocessed dataset based on the threshold values.
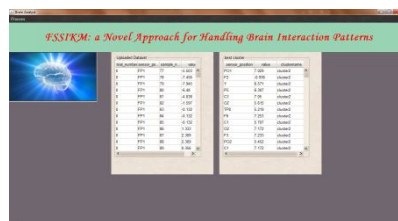

Fig. 3 Data sets are uploaded and clustered data set

Data set are uploaded and clustered data set in clustering. Interaction K-means (IKM), a partitioning clustering algorithm suitable to detect clusters of objects with similar interaction patterns. Assign objects for the minimal error rate clusters and find the minimal error rate for each cluster.
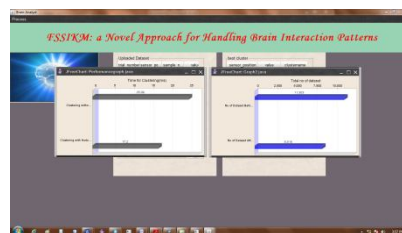

Fig.4. Data sets are uploaded and Ranked data set

Data sets are uploaded and Ranked data set in Dimension Ranking Algorithm Best Ranking algorithm was used to select the best cluster for assuring best result. Sort all models ascending according to the error. The top-ranked models best discriminate among the clusters.

VI.   CONCLUSION

The research, a new approach for the feature selection before clustering mechanism has been presented. Feature Subset Selection method used to choose a subset of the original features to be used for the subsequent processes. In this algorithm, initially collect the preprocessed dataset then find out the ranges of the respected dataset. After that, select the key features in the preprocessed dataset based on the threshold values. Interaction K-means (IKM), a partitioning clustering algorithm used to detect clusters of objects with similar interaction patterns classification and clustering. Interaction K-means (IKM), a partitioning clustering algorithm suitable to detect clusters of objects with similar interaction patterns classification and clustering. Finally, Best Ranking algorithm used to select the best cluster for assuring best result.

REFERENCES

[1] D. Arthur, B. Manthey, and H. Röglin, "Smoothed analysis of the k-means method," J. ACM, vol. 58, no. 5, pp. 19:1–19:31, Oct. 2011.

[2] Nikhil Singh, P. Thomas Fletcher, J. Samuel Preston, Linh Ha, Richard King, J. Stephen Marron2, Michael Wiener3, and Sarang Joshi : Multivariate Statistical Analysis of Deformation Momenta Relating Anatomical Shape to Neuropsychological Measures, Springer, MICCAI 2010.

[3] C. Böhm, L. Läer, C. Plant, and A. Zherdin, "Model-based classification of data with time series-valued attributes," in Proc. BTW, 2009, pp. 287–296.

[4] 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.

[5] H.-P. Kriegel, P. Kröger, A. Pryakhin, M. Renz, and A. Zherdin, "Approximate clustering of time series using compact modelbased descriptions," in Proc. DASFAA, New Delhi, India, 2008, pp. 364–379.

[6] I. Strigo, A. Simmons, S. Matthews, A. Craig, and M. Paulus, "Association of major depressive disorder with altered functional brain response during anticipation and processing of heat pain." Arch. Gen. Psychiatry, vol. 65, no. 11, pp. 1275–84, Nov. 2008.

[7] 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

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

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

[10] P. Cheeseman and J. Stutz, "Bayesian classification (autoclass): Theory and results," in Proc. Adv. Knowl. Discov. Data Min., Menlo Park, CA, USA, 1996, pp. 153–180.

BIOGRAPHY

K. Vidhyadevi received the B.E degree in Computer Science and Engineering from Arunai Engineering College, Tirunavanamalai in 1998. She Worked as Lecturer for 3 Years at Polytechnic College and Worked as MIS Co-Ordinator for 2 Years at Govt. SSA Chennai. Pursing M.Tech Computer Science and Engineering from Bharath University. Participated in workshops on Android PHP .Net C# programming language.

M. Beema Mehraj received M.Tech in Bharath university Now She is working as an Assistant professor at Bharath University for a years.

Dr. K.P. Kaliyamurthi received Ph.D and M.Tech from the Bharath University He had 23 years experience in teaching .Now he is working as a HOD in the Department of CSE at Bharath University. He has published more than 40 research papers in international journals and national journals.