INTERNATIONAL JOURNAL FOR ADVANCE RESEARCH IN ENGINEERING AND TECHNOLOGY

WINGS TO YOUR THOUGHTS.....

Application of Fuzzy Clustering Technique Using Medical Image Processing

TARUN DHAR DIWAN[1], KAMLESH LEHRE [2], BHARTI DWIVEDI[3]

[1]Dr.C.V.RAMAN UNIVERSITY, BILASPUR, INDIA
ASSISTANT PROFESSOR
DEPTT.OF ENGINEERING (CSE)
taruncsit@gmail.com
[2]DR.C.V.RAMAN UNIVERSITY, BILASPUR, INDIA,
ASSISTANT PROFESSOR
DEPTT.OF ENGINEERING (CSE)
lahrekamlesh@gmail.com
[3]DR.C.V.RAMAN UNIVERSITY, BILASPUR, INDIA,
MTECH SCHOLAR
b17.indian@gmail.com

Abstract: *This paper is based on Fuzzy Logic Techniques and clustering method, It comprises of study of different clustering algorithm and there Comparative Study. Fuzzy clustering techniques are widely and effectively used in, Pattern recognition, Image processing and fuzzy modeling. The best known method of fuzzy clustering is the Fuzzy K-Means, Fuzzy C-Means (FCM), Subtractive and many other algorithms. Data clustering is considered as an interesting approach for finding out similarities in data and putting similar data into groups.Clustering partitions of data set is done into several groups in such a way that the similarity within a group is larger than that among the cluster. The idea of data grouping, or clustering, is simple in its nature and is closely associated to the human way of thinking; whenever we are presented with a large sum of data, we usually tend to summarize this huge number of data into small number of groups or categories in order to further facilitate its analysis. Moreover, most of the data collected by several problems seem to have some inherent properties that lead themselves to natural groupings. Nevertheless, finding these groups or trying to categorize the data is not a simple task for humans unless the data is of low dimension (two or three dimensions at maximum.) this is why some methods in soft computing are proposed to solve this kind of problems. Those methods are called "data clustering methods" and they are the subject of this paper.*

Keywords: *clustering algorithm, Comparative Study, Data set, Data clustering, natural groupings, fuzzy modeling.*

1. INTRODUCTION

Clustering involves the task of dividing data points into homogeneous classes or clusters so that the items in same class are as like as possible and items in different classes are as different as possible. Clustering can also be thought as a form of data compression, where a large number of samples are converted into a smaller number of representative prototypes or clusters. Depending on the data and the application, different types of similarity measures may be used to identify classes, where the similarity measures controls the clusters that are formed. Some examples of values that can be used as similarity measures include distance connectivity, and intensity.

In non-fuzzy or hard clustering, data is divided into crisp clusters where each data point belongs to exactly one cluster. In fuzzy clustering, the data points can belong to more than one cluster, and can be associated with each of the points.The membership grades indicate the degree to which the data points belong to the different clusters [1].Data Clustering is considered as an interesting approach for finding similarities in data and putting similar data into groups. Clustering partitions is grouping data set into several groups such that the similarity within a group is larger than that of clusters [2]. The idea of data grouping, or clustering, is simple in its nature and is close to the human way of thinking; whenever we are presented with a large amount of data, we usually tend to summarize this

INTERNATIONAL JOURNAL FOR ADVANCE RESEARCH IN ENGINEERING AND TECHNOLOGY
WINGS TO YOUR THOUGHTS.....

huge number of data into smaller number of groups or categories in order to further facilitate its analysis. Moreover, most of the data collected in many problems seem to have some inherent properties that lead themselves to natural groupings. Nevertheless, finding these groupings or trying to categorize the data is not a simple task for humans unless the data is of low dimensionality (two or three dimensions at maximum.) This is why some methods in soft computing have been proposed to solve this kind of problem. Those methods are called "Data Clustering Methods" and they are the subject of this paper. Clustering algorithms are used extensively not only to organize and categorize data, but are also useful for data compression and model construction. By finding similarities in data, one can represent similar data with fewer symbols for example. Also if we can find groups of data, we can build a model of the problem based on those groupings [3].

In this paper, three of the most representative off-line clustering techniques are reviewed:

1. K-means (or Hard C-means) Clustering,
2. Fuzzy C-means Clustering,
3. Subtractive Clustering.

These techniques are usually used in conjunction with radial basis function networks (RBFNs) and Fuzzy Modeling. The results are presented with a comprehensive comparison of the different techniques and the effect of different parameters in the process [4].

2. LITERATURE SURVEY

As mentioned earlier, data clustering is concerned with the partitioning of a data set into several groups such that the similarity within a group is larger among groups. This implies that the data set to be partitioned has to have an inherent grouping to some extent; otherwise if the data is uniformly distributed, trying to find clusters of data will fail, or will lead to artificially introduced partitions. Another problem that may arise is the overlapping of data groups[5].Overlapping groupings may sometimes reduce the efficiency of the clustering method, and this reduction is proportional to the amount of overlap between groupings.In hard clustering, data is divided into distinct clusters, where each data element belongs exactly to one cluster. In fuzzy clustering, data elements can belong to more than one cluster, and associated with each element is a set of membership levels. These indicate the strength of the association between that data element and a particular cluster. Fuzzy clustering is a process of assigning these membership levels, and then using them to assign data elements to one or more clusters [6].Usually the techniques presented in this paper are used in conjunction with other sophisticated neural or fuzzy models. In particular, most of these techniques can be used as preprocessors for determining the initial locations for radial basis functions or fuzzy if then rules. The common approach of all the clustering techniques presented here. It is to find cluster centers that will represent each cluster. A cluster center is a way to tell where the heart of each cluster is located, so that later when presented with an input vector, by measuring similarity metric the clustering can tell that which vector belongs to which group between the input vector and all the cluster centers, and can determe which cluster is the nearest or most similar one. Some of the clustering techniques rely on knowing the number of clusters apriori. In that case the algorithm tries to partition the data into the given number of clusters [7]. K-means and Fuzzy C-means clustering are of that type. In other cases it is not necessary to have the number of clusters known from the beginning; instead the algorithm starts by finding the first large cluster, and then follows second, and so on. Mountain and Subtractive clustering are of that type [8]. In both cases a problem of known cluster numbers can be applied; however if the number of clusters is not known, K-means and Fuzzy C-means clustering cannot be used. Another aspect of clustering algorithms is their ability to be implemented in on-line or offline mode. On-line clustering is a process in which each input vector is used to update the cluster centers according to this vector position. The system in this case learns where the cluster centers are introduced by new input every time. In off-line mode, the system is presented with a training data set, which is used to find the cluster centers by analyzing all the input vectors in the training set. Once the cluster centers are found they are fixed, and then they are used later to classify new input vectors. The techniques presented here are of the off-line type. A brief overview of the three techniques is represented here. Full detailed discussion will follow in the next section. The first technique is K-means clustering (or Hard C-means clustering, as compared to Fuzzy C-means clustering) [9]. This technique has been applied to a variety of

INTERNATIONAL JOURNAL FOR ADVANCE RESEARCH IN ENGINEERING AND TECHNOLOGY
WINGS TO YOUR THOUGHTS.....

areas, including image and speech data compression, data preprocessing for system modeling using radial basis function networks, and task decomposition in heterogeneous neural network architectures. This algorithm finds cluster centers by trying to minimize cost function of dissimilarity (or distance) measure.The second technique is Fuzzy C-means clustering, which was proposed [4] as an improvement over earlier Hard C-means clustering. In this technique each data point belongs to a cluster to a degree specified by a membership grade. As in K-means clustering, Fuzzy C-means clustering minimizes cost function of dissimilarity measure. The third technique is Subtractive clustering, proposed by Chiu [10]. This technique is similar to mountain clustering, except that instead of calculating the density function at every possible position in the data space, it uses the positions of the data points to calculate the density function, thus reducing the number of calculations significantly. Characteristics of Fuzzy Clustering Algorithm

a. In real applications there is often no sharp boundary between clusters so that fuzzy clustering is often better suited for the data.

b. Membership degrees between zero and one are used in fuzzy clustering instead of crisp assignment of the data to clusters.

c. The resulting data partition improves data understanding and reveals its internal structure.

d. Clustering used in pattern reorganization and image segmentation.

e. The detection of special geometrical shapes like circles and ellipses can be achieved. Fuzzy inference (or simply "fuzzy logic") is a powerful problem-solving methodology with wide applications in industrial control and information processing. It provides a simple way to draw definite conclusions from vague, ambiguous or imprecise information. It resembles human decision making with its ability to work from approximate data and find precise solutions [3]. Unlike classical logic which requires a deep understanding of a system, exact equations and precise numeric values, fuzzy logic incorporates an alternative way of thinking, which allows modeling complex systems using a higher level of abstraction originating from our knowledge and experience. Fuzzy Logic allows expressing this knowledge with subjective concepts such as "very good" and "a little bit satisfied" which are mapped into exact numeric ranges.

How does fuzzy logic work - it uses 3 simple steps defined below

1. Fuzzification - to convert numeric data (e.g., $24.50 in sales) in real-world domain to fuzzy numbers in fuzzy domain
2. Aggregation (rule firing) - computation of fuzzy numbers (all between 0.0 and 1.0) in fuzzy domain
3. De-fuzzification - convert the obtained fuzzy number back to the numeric data (e.g. 150.34% in total profitability) in the real-world domain.

2.1 PURPOSE of RESEARCH

Research and Application Challenges

The objective of this study is to make the huge data set according to the nature of the data using the different fuzzy techniques (FCM, Subtractive) clustering.in a comparative manner on the basis of different factors like cluster center. A comparative study is made in between FCM and Subtractive clustering.

3. EXPERIMENT DESIGN SPECIFICATION

To start clustering the data:

1. Choose the clustering function fcm (fuzzy C-Means clustering) or subtractive (subtractive clustering) from the drop-down menu under this Method [4,5].
2. Set options for the selected method using the Influence Range, Squash, Aspect Ratio, and Reject Ratio fields. For more information on these methods and their options, refer to fcm, and subclust respectively.
3. Begin clustering by clicking Start. After clustering gets completed, the cluster centers appear in black as shown in the next figure.

Vol. 1, Issue VII, Aug 2013
ISSN 2320-6802

INTERNATIONAL JOURNAL FOR ADVANCE RESEARCH IN ENGINEERING AND TECHNOLOGY

WINGS TO YOUR THOUGHTS.....

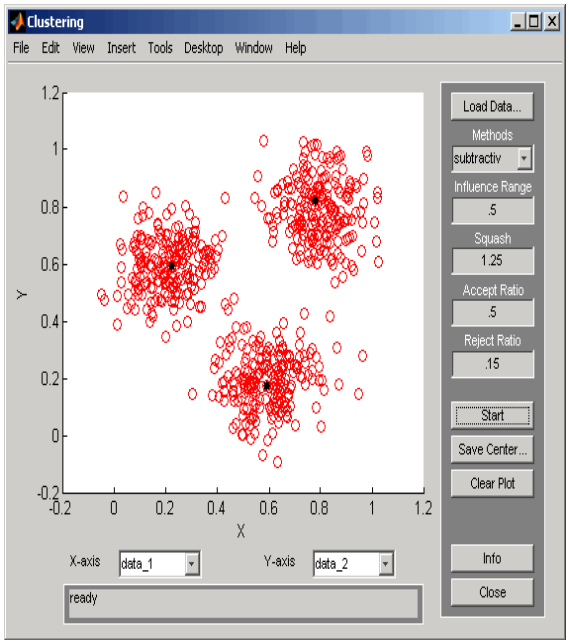

Figure 2: Saving the Cluster Center

Clustering with Fuzzy C-Means Clustering

This demo illustrates performing fuzzy c-means clustering on 2-dimensional data. Contents Interactive FCM Clustering Demo Performing FCM Clustering on Your Own Data Clustering of numerical data forms the basis of many classification and system modeling algorithms. The purpose of clustering is to identify natural groupings of data from a large data set to produce a concise representation of a system's behavior [1-2].

Interactive FCM Clustering Demo

Using the fcm command, we can launch a GUI that lets we try out various parameter settings for the fuzzy c-means algorithm and observe the effect on the resulting 2-D clustering. We can choose a sample data set and an arbitrary number of clusters from the drop down menue on the right, and then click "Start" to start the fuzzy clustering process. The clustering itself is performed by the fcm function.

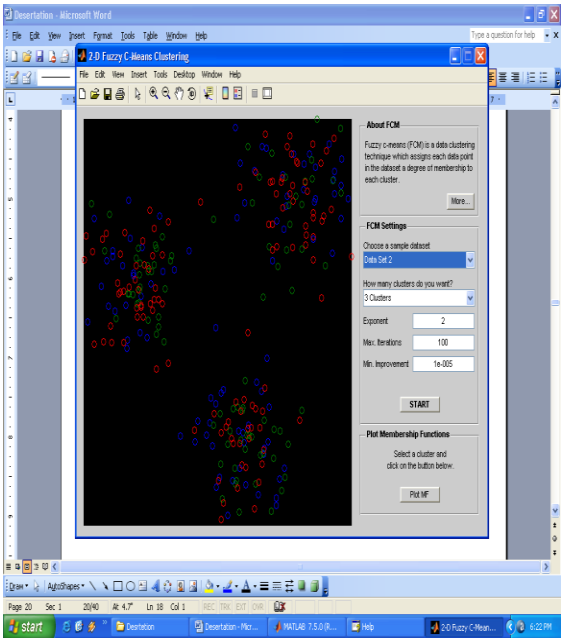

Figure 3: Demo GUI for Fuzzy C-Means Clustering.

4. IMPLEMENTATION AND RESULT

From this comparison we can conclude some remarks:

1. In FCM each data point belongs to a cluster of degree of membership grade, Subtractive Clustering is based on the same idea, and uses the data points as cluster centers candidates.If the number of clusters is not known, Fuzzy C-means clustering cannot be used. If we do not have a clear idea how many clusters are needed for a given set of data, Subtractive clustering can be used.

2. FCM clustering computates exponentially with the dimension of the problem, In Subtractive Clustering computation is now proportional to the problem size instead of the problem dimension.

3. FCM has a very poor performance regarding its requirement for high number of computation and low accuracy.

4. FCM is suitable only for problems with two or three dimensions.

5. FCM produces requires more computation time than Subtractive because of the fuzzy measures calculations involved in the algorithm.

6. Subtractive clustering, is fast, one-pass algorithm for estimating the number of clusters and the cluster

INTERNATIONAL JOURNAL FOR ADVANCE RESEARCH IN ENGINEERING AND TECHNOLOGY
WINGS TO YOUR THOUGHTS.....

centers in a set of data .There is more then one Pass in FCM Clustering algorithm.

No	Data (x, y)	Method	Center	X	Y	Z
3	(f1,f2)	FCM	C1	7.9391	7.8924	1.9682
			C2	2.0513	5.9600	8.0897
			C3	5.9626	1.9230	5.0572
	(f1,f3)		C1	2.0514	5.9600	8.0897
			C2	7.9391	7.8925	1.9682
			C3	5.9627	1.9230	5.0571
3	(f1,f2)	Subtractive	C1	2.2199	5.9368	8.1127
			C2	7.7971	8.1905	1.8012
			C3	5.9141	1.7214	4.8719
	(f1,f3)		C1	2.2199	5.9368	8.1127
			C2	7.7971	8.1905	1.8012
			C3	5.9141	1.7214	4.8719

5. CONCLUSION AND FUTURE WORK

We have replicated a set of data to explain in both FCM and Subtractive clustering using MATLAB software. Comparative study is made in between FCM and Subtractive clustering. Fuzzy clustering is one of the methods used for datareplication. We have applied each algorithm in collection of data set. Clustering involves the task of dividing data points into homogeneous classes or clusters so that each items in the same class are as alike as possible and items in different classes are as unlike as possible. Depending on the data and the application, different types of similarity measures may be used to identify classes, where the similarity measure controls the clusters that are formed. Some examples of values that can be used as similarity measures include distance, connectivity, and intensity. We have applied different clustering method in large data set and present the result in comparative manner. The Two methods have been implemented and tested against a data set. The comparative study done here is concerned with the accuracy of each algorithm; great care is taken toward the efficiency in calculation and other performance measures. It was obvious that Subtractive clustering is not one of the good techniques for solving the problems with the high number of dimensions due to its exponential proportionality with relavance to the dimension of the problem. However in other problems where the number of clusters is not known, K-means and FCM cannot be used to solve this type of problems, leaving the choice only to subtractive clustering. Subtractive clustering seems to be a better alternative to FCM clustering. Since the above clustering is based on the same idea, and uses the data points as cluster centers candidates instead of grid points; however, FCM can lead to better results if the grid granularity is small enough to capture the potential cluster centers, but with the side effect of increasing computation needed for the larger number of membership grade. FCM produces high proportion of actives using the value. Subtractive clustering is calculated only at every data point with the difference of a density function, instead at every membership grad. So the data points themselves are the candidates for cluster centers. This will reduce the number of computations significantly, and makes it linearly proportional to the number of input data instead of being exponentially proportional to its dimension.

Each fuzzy clustering technique has its own characteristics. FCM and Subtractive algorithm used for simulation of data set. We found that Subtractive algorithm is better than the FCM algorithm. The following points are worth stressing.

1. The data simulation (analysis) result well than FCM.
2. Comparison ratio is less than the FCM.
3. Cluster center is better than FCM.
4. Less execution time is needed then FCM.

Finally, the clustering technique discussed here does not needed to be used as stand-alone approaches; they can be used in conjunction with other neural or fuzzy systems for further refinement of the overall system performances.

REFERENCES

[1] Fuzzy sets and Fuzzy logic Theory and

INTERNATIONAL JOURNAL FOR ADVANCE RESEARCH IN ENGINEERING AND TECHNOLOGY

WINGS TO YOUR THOUGHTS…..

Application-George J. Klir/Bo Yuan.

[2] Computer Study of India (CIS) Communications, page no.-14 to 18 Image Segmentatiom Using Reformed Fuzzy Clustering Technique-B Soumya, B Sheelarani , Volume no. 33, Issue 3, June 2009, ISSN 0970-647x.

[3] J.C.Dunn, A Fuzzy Relation of the ISODATA Process and Its Use in Detecting Compact Well-Separated Clusters, Journal of Cybernetics 3; 1973:32-57.

[4] Bezdek, J.C. Pattern Recognition with Fuzzy Objective function Algorithm, Plenum Press, NY, 1981.

[5] Hanifi GULDEMIR, Abdulkadir SENGUR Classification of Analog Modulated Communication Signals using Clustering Techniques: A Comparative Study, 23119, Elazig, TURKEY.

[6] Khaled Hammouda, Prof. Fakhreddine Karray University of Waterloo, Ontario, Canada A Comparative Study of Data Clustering Techniques.

[7]. E.Cox, Fuzzy Modeling And Genetic Algorithm For Data Mining And Exploration, Elsevier, 2005

[8]. Jang, J.-S. R., Sun, C.-T., Mizutani, E., "Neuro- Fuzzy and Soft Computing – A Computational Approach to Learning and Machine Intelligence," Prentice Hall.

[9] Lin, C., Lee, C., "Neural Fuzzy Systems," Prentice Hall, NJ, 1996.

[10] The MathWorks, Inc., "Fuzzy Logic Toolbox – For Use with MATLAB," The MathWorks, Inc., 1999.

AUTHOR BIOGRAPHIES

TARUN DHAR DIWAN RECEIVED HIS MASTER OF ENGINEERING (COMPUTER TECHNOLOGY AND APPLICATION) DEGREE FROM CHHATTISGARH SWAMI VIVEKANANDA TECHNICAL UNIVERSITY –BHILAI, INDIA, AND MASTER OF PHILOSOPHY (GOLD MEDAL LIST) FROM DR. C.V. RAMAN UNIVERSITY. HE IS CURRENTLY AN HOD & MTECH CO-ORDINATOR DEPTT. OF ENGINEERING AT THE DR.C.V.RAMAN UNIVERSITY-BILASPUR, INDIA. HIS CURRENT RESEARCH WORK ARTIFICIAL INTELLIGENT, IMAGE PROCESSING AND SOFTWARE ENGINEER