International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Impact Factor (2012): 3.358

A Fuzzy C – Means Clustering Based on Density Sensitive Distance Metric with a Novel Penalty Term

[1]S. Omprakash, T. Senthamarai[2], M. Hemalatha[3]

[1]Assistant Professor, Department of Computer Science, Kovai Kalaimagal College of Arts & Science, Coimbatore-109, India
[2]Assistant Professor, Department of Computer Applications, Pioneer College of Arts & Science, Jothipuram , Coimbatore – 47
[3]Assistant Professor Department of Computer Applications, Pioneer College of Arts & Science, Jothipuram , Coimbatore – 47

Abstract: ***A cluster is a group of objects which are similar to each other within a cluster and are dissimilar to the objects of other clusters. The major objective of clustering is to discover collection of comparable objects based on similarity metric. A similarity metric is generally specified by the user according to the requirements for obtaining better results. The distance between the measures of two objects in a particular cluster should be well defined using effective distance measures. There are several approaches available for clustering objects. The clustering approaches are, Penalty Fuzzy C-Means. But these techniques are not suitable for all applications and huge data collections. In the proposed approach an effective fuzzy clustering technique is used. Fuzzy Possibilistic C-Means (FPCM) is the effective clustering algorithm available to cluster unlabeled data that produces both membership and typicality values during clustering process. Penalized and Compensated terms are embedded with the Modified fuzzy positivistic clustering method's objective function to construct the Penalized based FPCM (PFPCM). In order to improve the clustering accuracy, third proposed approach uses the Improved Penalized Fuzzy C-Means (IPFCM). The penalty term takes the spatial dependence of the objects into consideration, which is inspired by the Neighborhood Expectation Maximization (NEM) algorithm and is modified according to the criterion of FCM. The proposed Improved Penalized for Fuzzy C-Means (IPFCM) clustering algorithm, uses improved penalized constraints which will help in better calculation of distance between the clusters and increasing the accuracy of clustering. The performance of the proposed approaches is evaluated on the University of California, Irvine (UCI) machine repository datasets such as Iris, Wine, Lung Cancer and Lymphograma. The parameters used for the evaluation is Clustering accuracy, Mean Squared Error (MSE), Execution Time and Convergence behavior.***

Keywords: Clustering, FCM, PFCM, MPFCM, Dataset

1. Introduction

Due to the remarkable development of the modern information systems, the quantity of data that is collected has increased massively. In order to examine these enormous collections of data, the interdisciplinary field of Knowledge Discovery in Databases (KDD) has emerged. The fundamental step of KDD is called Data Mining [1]. Data Mining utilizes effective techniques to obtain interesting patterns and regularities from the data. In addition to the sheer size of available data sources, the complexity of data objects has also increased. As a result, novel data mining techniques are essential to draw greatest benefit from this extra information.

1.1. Overview of Data Mining

Data mining is one of the major steps in the KDD method and consists of applying data investigation and discovery algorithms that, under adequate computational efficiency restrictions, generate a particular enumeration of patterns over the data. Data mining is the step that is accountable for the real knowledge discovery. To emphasize the requirement that data mining approaches need to process huge amount of data, the required patterns has to be found under adequate computational efficiency limitations. The most significant data mining techniques related to the kind of knowledge they mine are as follows.

Classification: Classification (also known as supervised learning) is the task of learning a function that points data objects to one or more classes in a predefined class set. In order to learn this function, classification techniques require a training set, has data objects that are previously mapped to the class they belong to.

After examining the training set, classification techniques can map new unidentified objects to the classes. A second function of classification is, deriving class representations to make clear why the objects are mapped in this way.

Clustering: Clustering (also known as unsupervised learning) is the task of recognizing a finite group of categories (or clusters) to illustrate the data. Therefore, similar objects are clustered to the similar category and dissimilar objects to different clusters. Clustering is also known as unsupervised learning since the data objects are pointed to a collection of clusters which can be interpreted as classes additionally.

1.2. Clustering and Classification

The major contribution of this thesis is the development of new techniques for clustering analysis. Therefore, it is very important to discuss the relation between clustering and classification.

Clustering: Clustering is the process of assembling the data records into significant subclasses (clusters) in a way that increases the relationship within clusters and reduces the similarity among two different clusters [2]. Other names for clustering are unsupervised learning (machine learning) and segmentation. Clustering is used to get an overview over a given data set. A set of clusters is often enough to get insight into the data distribution within a data set. Another important use of clustering algorithms is the preprocessing for some other data mining algorithm.

Classification: Classification is the process of learning a function that maps data objects to a subset of a particular class set. As a result, a classifier is trained with a labeled set of training objects, identifying each class. There are two goals of classification:

Identifying a better general mapping that can predict the class of unidentified data objects with more accuracy. For this purpose, the classifier is a simple function. In order to accomplish this objective, the classifier has to decide the features of the particular training instances which are typical for the entire class and features which are definite for single objects in the training set.

The additional objective of classification is to discover a condensed and comprehensible class model for all the classes. A class model should provide an explanation why the particular objects belong to a certain class and what is distinctive for the members of a given class. The class model should be condensed as far as possible since the more compact model is, the more general it is. In addition, small and simple class models are uncomplicated to understand and have less distracting information.

Classification and clustering are strongly associated. Classification attempts to learn the distinctiveness of a given set of classes, whereas clustering discovers a set of classes inside a given data set. A significant characteristic of clustering is that it is not essential to determine a set of specimen objects.

Consequently, clustering can be exploited in applications where there is no or some previous knowledge regarding the groups or classes in a database. On the other hand, the effectiveness of a found clustering is often subject to individual interpretation and robustly based on the selection of an appropriate similarity measure. In applications for which the existence of a dedicated set of classes is previously identified, the use of classification is more adequate. In these cases providing instant objects for every class is typically much simpler than assembling a feature space in which the predefined classes are grouped into delimited clusters. In addition, the performance of a classifier can simply be measured by the amount of accurate class predictions it realizes. Thus, it is found that clustering and classification are associated with data mining tasks that are used in various situations. Figure 1.1 displays class separation by a classifier on the left side and the clustering of two clusters in a noisy data set on the right side.

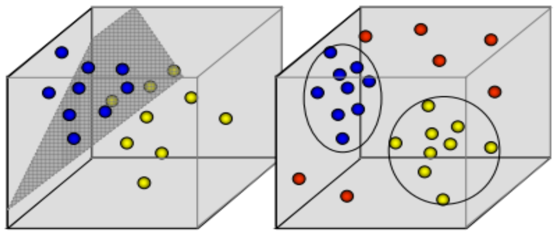

Figure 1.1: Classification separates the data space (left) and clustering groups data objects (right)

1.3. Essential Requirements of Clustering

Scalability: Several clustering approaches execute well with small amount of data objects. However, if the amount of objects is in the range of millions then clustering techniques will not perform very effectively but provide unnecessary or irrelevant results.

Variety of attributes should be handled: Large numbers of clustering approaches are developed to manage with numerical data. On the other hand, several applications are in need of clustering various existing types of data, for instance binary, categorical and ordinal data, or collection of these types.

Discovering arbitrary shaped clusters: Many clustering algorithms find clusters based on Euclidean or Manhattan distance measures. Algorithms based on such distance measures tend to find clusters with spherical shape with similar size and density. However, a cluster can be of any shape. It should be very important to develop algorithms that find arbitrary shaped clusters.

Minimum requirement of domain knowledge: Many clustering algorithms require users to initially give the input parameters like number of desired clusters etc, The clustering results are often very sensitive to input parameters. It is difficult to determine many parameters by the user especially for datasets that contain high dimensional data.

Ability to deal with noisy data: Most of the real time large databases have outliers, missing, unknown and erroneous data. Some of the clustering algorithms are sensitive to that kind of data and may lead to clusters of poor quality.

Insensitivity to the order of input records: Some clustering methods are sensitive to the order of input records passed to them. However the order in which the input records are given to the clustering algorithms they should be able to produce same clusters in any of the way.

High Dimensionality: Many clustering algorithms are good at handling low dimensional data, involving only two to three dimensions. The clustering algorithms should be able to cluster data objects in high-dimensional space, especially considering the fact that data in high-dimensional space can be very sparse and highly skewed.

Interpretability and Usability: Users may expect the clustering results to be interpretable, comprehensible and usable. Clustering may need to be tied up with some specific semantic interpretations and applications. It is important to know how an application goal may influence the selection of clustering methods.

1.4. Clustering Algorithm

Clustering algorithms are general schemes, which use particular similarity measures as subroutines. The particular choice of clustering algorithms depends on the desired properties of the final clustering. Other considerations include the usual time and space complexity. A clustering algorithm attempts to find natural groups of components (or data) based on some similarity. The clustering algorithm also finds the centered of a group in data sets. To determine cluster membership, most algorithms evaluate the distance between a point and the cluster centroids. The output from a clustering algorithm is basically a statistical description of the cluster centroids with the number of components in each cluster.

1.5 .Objectives of the Research

The primary aim of this research is to propose the novel technique for the clustering. The other goals of this research include the following:

- To develop a clustering technique that is not sensitive to initial positions of cluster centers.
- Developing a novel technique for clustering by using effective approaches which provide very high classification accuracy.
- To develop a clustering technique with very less error rate and with minimized execution time.
- Developing a clustering technique with less convergence time and also with less number of iterations.
- To develop an effective clustering technique which provides better results in incomplete/noisy data.

2. Fuzzy Clustering Approaches

Five clustering algorithms taken from the literature are reviewed, assessed and compared on the basis of the selected properties of interest by Baraldi et al., [3]. These clustering models are

- Self-Organizing Map (SOM);
- Fuzzy Learning Vector Quantization (FLVQ);
- Fuzzy Adaptive Resonance Theory (fuzzy ART);
- Growing Neural Gas (GNG);
- Fully Self-Organizing Simplified Adaptive Resonance Theory (FOSART).

Although our theoretical comparison is fairly simple, it yields observations that may appear parodoxical. First, only FLVQ, fuzzy ART, and FOSART exploit concepts derived from fuzzy set theory (e.g., relative and/or absolute fuzzy membership functions). Secondly, only SOM, FLVQ, GNG, and FOSART employ soft competitive learning mechanisms, which are affected by asymptotic misbehaviors in the case of FLVQ, i.e., only SOM, GNG, and FOSART are considered effective fuzzy clustering algorithms.

The fuzzy clustering of fuzzy rules, here proposed, as well as clustering of data, leads to a fuzzy partition of the S space. The result is a set of fuzzy sub-systems, one for each cluster that will be conveniently linked in a new structure. Salgado et al., proposed a new recursive clustering algorithm for the partition of a fuzzy system into a hierarchical collaborative structure. The global response of the hierarchical collaborative structure is identical to the input fuzzy system.

The new modeling approach introduces three features: i) an Improved Fuzzy Clustering (IFC) algorithm, ii) a new structure identification algorithm, and iii) a nonparametric inference engine. The IFC algorithm yields simultaneous estimates of parameters of c-regression models, together with fuzzy c-partitioning of the data, to calculate improved membership values with a new membership function.

2.1 Penalty Fuzzy C-Means Clustering

In 1997, Pal et al., [4] proposed the Penalty fuzzy-possibilistic C-Means (PFCM) technique and algorithm that generated both membership and typicality values when clustering unlabeled data. PFCM constrains the typicality values so that the sum over all data points of topicalities to a cluster is one. For large data sets the row sum constraint produces unrealistic typicality values. In this approach, a new model is presented called Possibilistic-Fuzzy C-Means (PFCM) model. PFCM produces memberships and possibilities concurrently, along with the usual point prototypes or cluster centers for each cluster. PFCM is a hybridization of FCM and Possibilistic C-Means (PCM) that often avoids various problems of PCM, FCM and PFPCM.

The noise sensitivity defect of FCM is resolved in PFCM, overcomes the coincident clusters problem of PCM and eliminates the row sum constraints of PFCM. The first-order essential conditions for extreme of the PFCM objective function is driven, and used them as the basis for a standard alternating optimization approach to find local minima of the PFCM objective function. PFCM prototypes are less sensitive to outliers and can avoid coincident clusters; PFCM is a strong candidate for fuzzy rule-based system identification.

3. An Efficient Penalized Fuzzy C-Means With Density-Sensitive Distance Metric

3.1. Fuzzy Clustering Algorithm

The fuzzified version of the K-Means algorithm is the Fuzzy C-Means (FCM). It is a method of clustering which allows one piece of data to belong to two or more clusters. This method was developed by Dunn in 1973 [5] this is frequently used in pattern recognition. The algorithm is an iterative clustering method that brings out an optimal c partition by minimizing the weighted within group sum of squared error objective function J_{FCM}:

$$J_{FCM}(V,U,X) = \sum_{i=1}^{c}\sum_{j=1}^{n} u_{ij}^{m}\, d^2(x_j, v_i), 1 < m < +\infty$$

In the equation $X = \{x_1, x_2, \ldots, x_n\} \subseteq R^p$ is the dataset in the p-dimensional vector space, where the number of data items is represented as p. c is the number of clusters with $2 \le c \le n-1$. $V = \{v_1, v_2, \ldots, v_c\}$ is the c centers or prototypes of the clusters, v_i represents the p-dimension center of the cluster i, and $d^2(x_j, v_i)$ represents a distance measure between object x_j and cluster center v_i. $U = \{u_{ij}\}$ represents a fuzzy partition matrix with $u_{ij} = u_i(x_j)$ is the degree of membership of x_j in the i[th] cluster; x_j is the j[th] of p-dimensional measured data. The fuzzy partition matrix satisfies:

$$0 < \sum_{j=1}^{n} u_{ij} < n, \forall i \in \{1, \ldots, c\}$$

$$\sum_{i=1}^{c} u_{ij} = 1, \forall j \in \{1, \ldots, n\}$$

where m is a weighting exponent parameter on each fuzzy membership and establishes the amount of fuzziness of the resulting classification; it is a fixed number greater than one. Under the constraint of U the objective function J_{FCM} can be minimized. Specifically, taking of J_{FCM} with respect to u_{ij} and v_i and zeroing them respectively is necessary but not sufficient conditions for J_{FCM} to be at its local extrema will be as the following:

$$u_{ij} = \left[\sum_{k=1}^{c} \left(\frac{d^2(x_j, v_i)}{d^2(x_j, v_k)}\right)^{1/(m-1)}\right]^{-1}, 1 \le i \le c, 1 \le j \le n.$$

$$v_i = \frac{\sum_{k=1}^{n} u_{ik}^{m} x_k}{\sum_{k=1}^{n} u_{ik}^{m}}, 1 \le i \le c.$$

In noisy environment memberships of FCM do not always correspond well to the degree of belonging of the data, and may be inaccurate, because the real data unavoidably involves some noises. To recover this weakness of FCM, relaxing the constrained condition (3.3) of the fuzzy c-partition to obtain a possibilistic type of membership function and PCM for unsupervised clustering is proposed. The component generated by the PCM corresponds to a dense region in the dataset; each cluster is independent of the other clusters in the PCM strategy.

3.2 Modified Penalized Fuzzy C-Means Technique (MPFCM)

To obtain clustering results with more accuracy, better objective function is required. A new algorithm was given by Wen-Liang Hung called Modified Suppressed Fuzzy C-Means (MS-FCM), which significantly improves the performance of FCM due to a prototype-driven learning parameter α [6]. Exponential separation strength between clusters is the base for the learning process of α and is updated at each of the iteration. The parameter α can be computed as

$$\alpha = exp\left[-\min_{i \ne k} \frac{\|v_i - v_k\|^2}{\beta}\right]$$

In the above equation β is a normalized term so that β is chosen as a sample variance. That is, β is defined:

$$\beta = \frac{\sum_{j=1}^{n} \|x_j - \bar{x}\|^2}{n}$$

Where

$$\bar{x} = \frac{\sum_{j=1}^{n} x_j}{n}$$

But the remark which must be mentioned here is the common value used for this parameter by all the data at each iteration, which may lead to error. A new parameter is added with this which suppresses this common value of α and replaces it by a new parameter like a weight to each vector. Or every point of the dataset has a weight in relation to every cluster. Consequently this weight permits to have a better classification especially in the case of noise data. The following equation is used to calculate the weight.

$$w_{ji} = exp\left[-\frac{\|x_j - v_i\|^2}{\left[\sum_{j=1}^{n} \|x_j - \bar{v}\|^2\right] * c/n}\right]$$

In (3.14) w_{ji} represents weight of the point j in relation to the class i. This weight is used to modify the fuzzy and typical partition. The objective function is composed of two expressions: the first is the fuzzy function and uses a fuzziness weighting exponent, the second is possibilistic function and uses a typical weighting exponent; but the two coefficients in the objective function are only used as exhibitor of membership and typicality. A new relation, enables a more rapid decrease in the function and increase in the membership and the typicality when they tend toward 1 and decrease this degree when they tend toward 0. This relation is to add Weighting exponent as exhibitor of distance in the two objective functions. The density-sensitive distance metric between two points is defined to be

$|p|$=1

$$D_{ij} = \min_{p \in P_{ij}} \sum_{k=1}^{|p|-1} L(p_k, p_{k+1})$$

Thus *Dij* satisfies the four conditions for a metric, i.e. $D_{ij} = D_{ji}$, $D_{ij} \ge 0$; $D_{ij} \le D_{ik} + D_{kj}$ for all x_i, x_j, x_k; and $D_{ij} = 0 \; iff \; x_i = x_j$. As a result, the density-sensitive distance metric can measure the geodesic distance along the manifold, which results in any two points in the same region of high density being connected by a lot of shorter edges while any two points in different regions of high density are connected by a longer edge through a region of low density.

This achieves the aim of elongating the distance among data points in different regions of high density and simultaneously shortening that in the same region of high density. Hence, this distance metric is data-dependent, and can reflect the data character of local density, namely, what is called density-sensitive.

4. Experimental Results

To evaluate the proposed Penalized Fuzzy C-Means (PFCM), against Fuzzy C-Means (FCM) and Penalized Fuzzy C-Means (FCM), experiments were carried out using the similar experimental setup and parameters.

4.1 Iris Dataset

Clustering Accuracy: Clustering accuracy is calculated for FCM and PFCM and the proposed MPFCM in iris dataset.

Mean Squared Error: The cluster centers found by proposed MPFCM are closer to the true centers, than the centers found by FCM and PFCM. The mean squared error of the Iris dataset for the two cluster centers of the three approaches are provided in table 4.1.

Table 4.1: Comparison of Mean Squared Error in Iris Dataset

Cluster	*FCM*	*PFCM*	*MPFCM*
Cluster 1	0.4852	0.4212	0.3895
Cluster 2	0.4956	0.4620	0.3915

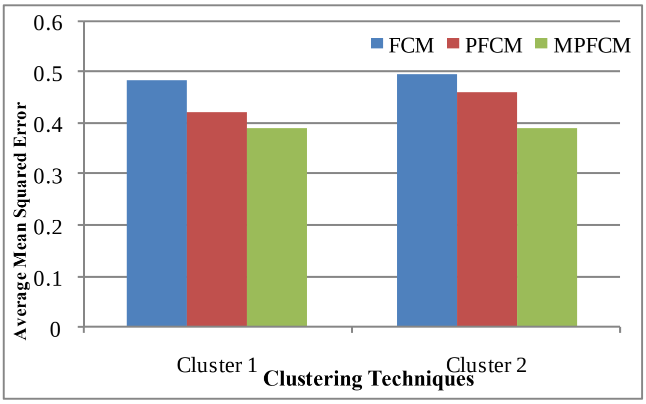

Figure 4.1: Comparison of Mean Squared Err or for Iris Dataset

It is observed from the figure 4.1 that the proposed MPFCM gives very low MSE values for both the clusters (0.3895 and 0.3915) than the FCM (0.4852 and 0.4956) and PFCM (0.4212 and 0.4620).

4.2 Wine Dataset

Clustering Accuracy: Accuracy of the clustering results is calculated for FCM, PFCM and the proposed MPFCM for wine dataset.

Mean Squared Error: The cluster centers found by proposed MPFCM are closer to the true centers, than the centers found by FCM, PFCM. The mean squared error of the wine dataset for the three cluster centers of the three approaches are provided in table 4.2.

Table 4.2: Comparison of Mean Squared Error for Wine Dataset

Cluster	*FCM*	*PFCM*	*MPFCM*
Cluster 1	0.5213	0.4925	0.4111
Cluster 2	0.4256	0.4102	0.3652
Cluster 3	0.5033	0.4910	0.4125

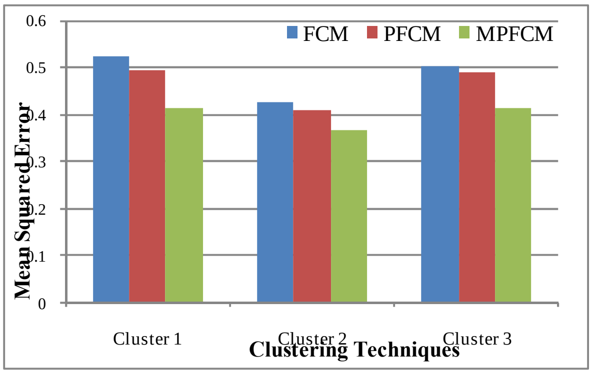

Figure 4.2: Comparison of Mean Squared Error in Wine Dataset

It is observed from the figure 4.2 that the proposed MPFCM gives very low MSE values for three clusters (0.4111, 0.3652 and 0.4125) than the FCM (0.5213, 0.4256 and 0.5033) and PFCM (0.4925, 0.4102 and 0.4910).

4.3 Lung Cancer Dataset

Clustering Accuracy: Accuracy of the clustering results is calculated for FCM, PFCM and the proposed MPFCM for lung cancer dataset.

Mean Squared Error: The cluster centers found by proposed MPFCM are closer to the true centers, than the centers found by FCM and PFCM. The mean squared error of the Lung Cancer dataset for the three cluster centers of the three approaches are provided in table 4.3.

Table 4.3: Comparison of Mean Squared Error for Lung Cancer Dataset

Cluster	*FCM*	*PFCM*	*MPFCM*
Cluster 1	0.5252	0.5002	0.4803
Cluster 2	0.4754	0.4512	0.4216
Cluster 3	0.4221	0.4051	0.3859

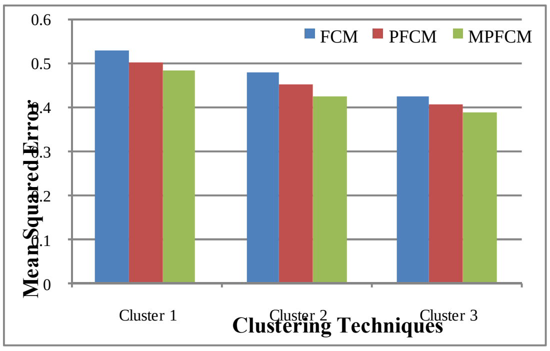

Figure 4.3: Comparison of Mean Squared Error for Lung Cancer Dataset

It is observed from the figure 4.3 that the proposed MPFCM gives very low MSE values for three clusters (0.4803, 0.4216 and 0.3859) than the FCM (0.5252, 0.4754 and 0.4221) and PFCM (0.5002, 0.4512 and 0.4051).

4.4 Lymphography Dataset

Clustering Accuracy: Accuracy of the clustering results is calculated for FCM, PFCM and the proposed MPFCM in lymphography dataset.

Mean Squared Error: The cluster centers found by the proposed MPFCM are closer to the true centers, than the centers found by FCM and PFCM. The mean squared error of the lymphography dataset for the four cluster centers of the three approaches are provided in table 4.4.

Table 4.4: Comparison of Mean Squared Error in Lymphography Dataset

Cluster	*FCM*	*PFCM*	*MPFCM*
Cluster 1	0.7149	0.6858	0.6302
Cluster 2	0.6145	0.5889	0.5148
Cluster 3	0.6525	0.6335	0.5882
Cluster 4	0.7412	0.7289	0.6910

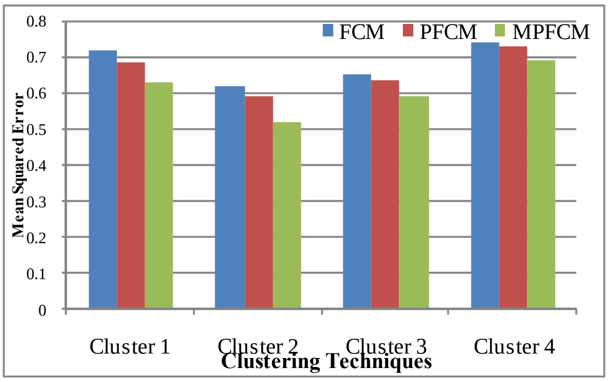

Figure 4.4: Comparison of Mean Squared Error for Lymphography Dataset

It is observed from the figure 4.4 that the proposed MPFCM gives very low MSE values in all the four clusters (0.6302, 0.5148, 0.5882 and 0.6910) than the FCM (0.7149, 0.6145, 0.6525 and 0.7412) and FCM (0.6858, 0.5889, 0.6335 and 0.7289).

5. Conclusion

This research focuses on the effective clustering techniques for data clustering. Data clustering has become an important research area in the field of data mining. This research proposed three efficient clustering techniques for data clustering. Effective fuzzy clustering approaches are used in this research which improves the results of clustering.

In the approach, penalized FCM is improved by using NEM algorithm and it is combined with compensated constraints which is said to be Improved Penalized constraints for Fuzzy Possibilistic C-Means (IPFPCM) clustering algorithm. The usage of improved penalized constraints in MPFCM will help in better calculation of distance between the clusters and increasing the accuracy of clustering.

The performances of the proposed approaches are evaluated on UCI machine repository datasets namely Iris, Wine, Lung cancer and Lymphography. It is observed from the experimental results that the proposed MPFCM method outperforms the other proposed approaches in terms of accuracy, Mean Squared Error, Execution Time and Convergence Behavior. Thus, the proposed MPFCM approach is best suited for the data clustering applications.

The present research work can be applied to various specific applications in the field of data mining. Clustering plays an outstanding role in data mining applications such as Scientific Data Exploration, Information Retrieval and Text Mining, Spatial Database Applications, Web Analysis, Marketing, Medical Diagnostics especially Gene Classification, Computational Biology, Customer Relationship Management (CRM), etc.

6. Scope for Future Work

The proposed clustering approaches provide effective data clustering technique. MPFPCM provides very significant performance when compared with the other proposed clustering approaches. This research utilized a clustering algorithm to provide the best clustering results with greater clustering accuracy and reduced mean squared error and execution time, respectively with quick convergence.

The problem that still occurs in this clustering and also in the real world is how to determine exactly how many concepts are actually present in clustering the data. In order to solve this issue, future enhancement of proposed approach is necessary. The future enhancement would be to use the statistical method to find the optimal cluster number.

The further enhancement of the proposed approach is presented below for better performance and efficiency of the data clustering:

- Better optimization techniques like Ant Colony Optimization can also be used for the better performance of the data clustering approach.

- Accuracy of the clustering results of the proposed approach can still be improved. 17, 2005.
- Moreover, the time taken by the proposed approach should also be considered. The time taken for clustering data should be very less with high accuracy.

References

[1] G. Karypis, E.-H Han and V. Kumar. "Chameleon: Hierarchical Clustering using Dynamic Modelling", IEEE Computer, Pp. 68–75, 1999
[2] Filippone M, Masulli F and Rovetta S, "Applying the Possibilistic c-Means Algorithm in Kernel-Induced Spaces", IEEE Transactions on Fuzzy Systems, Vol. 18, No. 3, Pp. 572–584, 2010.
[3] Baraldi A and Blonda P, "A survey of fuzzy clustering algorithms for pattern recognition. II", IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, Vol. 29, No. 6, Pp. 786–801, 1999.
[4] Pal N.R, Pal K, Keller J.M and Bezdek J.C, "A new hybrid c-means clustering model", Proceedings IEEE International Conference on Fuzzy Systems, Vol. 1, Pp. 179–184, 2004.
[5] Chunhui Zhang, Yiming Zhou and Trevor Martin, "Similarity Based Fuzzy and Possibilistic c-means Algorithm", Proceedings of the 11th Joint Conference on Information Sciences, Pp. 1-6, 2008
[6] W.W. Melek, A.A. Goldenberg and M.R. Emami, "A fuzzy noise-rejection data partitioning algorithm", International Journal of Approximate Reasoning, Vol. 38, Pp.

Author Profile

Mr. S. Omprakash., M.Sc., M.Phil., Assistant Professor , Department of Computer Science.

Ms. T. Senthamarai., M.Sc., M.Phil., Assistant Professor , Department of Computer Applications.

Mrs. M. Hemalatha., M.Sc., M.Phil., Assistant Professor , Department of Computer Applications.