2011 International Conference on Recent Trends in Information Systems

An Improved Fuzzy Clustering Method Using Modified Fukuyama-Sugeno Cluster Validity Index

Sailik Sengupta*, Soham De†, Amit Konar‡, R. Janarthanan§
*†‡Jadavpur University, Calcutta 700032, India
* Email: link2sailik@gmail.com
† Email: soham20491@gmail.com
‡ Email: konaramit@yahoo.co.in
§ Jaya Engineering College, Chennai 602024, India

***Abstract*—The objective of clustering algorithms is to group similar patterns in one class and dissimilar patterns in disjoint classes. This article proposes a novel algorithm for fuzzy partitional clustering with an aim to minimize a composite objective function, defined using the Fukuyama-Sugeno cluster validity index. The optimization of this objective function tries to minimize the separation between clusters of a data set and maximize the compactness of a certain cluster. But in certain cases, such as a data set having overlapping clusters, this approach leads to poor clustering results. Thus we introduce a new parameter in the objective function which enables us to yield more accurate clustering results. The algorithm has been validated with some artificial and real world datasets.**

I. Introduction

A pattern refers to a set of distinct attributes of an object or a phenomenon, which together is sufficient to recognize it. The objective of clustering is to group similar and segregate dissimilar patterns. The group of similar patterns is usually called a cluster. A cluster center is an ideal member of the class located at the central region of the cluster.

A clustering algorithm is expected to discover the natural grouping that exists in a set of patterns or data points. It aim at minimizing the distance between the data points with their respective cluster centers. K-means clustering, for example, is one such algorithm, which minimizes the Euclidean distance between the given set of data points with their respective cluster centers. Moreover, Clustering algorithms can be classified into two types, hierarchical and partitional.

Clustering can be broadly classified into two main categories, hard clustering and fuzzy clustering. The former indicates methods that classify any data point in a given set absolutely, i.e., a certain point in a data set is allowed to be classified into one and only one cluster center. This is a disadvantage in cases where a point maybe equidistant from more than one cluster. In fact, such points should be classified into more than one cluster with varying degrees of membership to each cluster. Hence the concept of fuzzy sets [13] was introduced in clustering resulting in introduction of Fuzzy Clustering [4]. Fuzzy C-Means clustering and its variants are examples.

Clustering algorithms are usually validated using a set of standard cluster validity indices (CVIs). These indices have been designed to study the performance of a new clustering algorithm and compare it with the existing algorithms.

Occasionally, CVIs are used as the objective function of a new clustering algorithm with an aim to ensure certain characteristics of the algorithm. In this paper, we propose a novel clustering algorithm by employing the Fukuyama-Sugeno [5] CVI and integrating a new fuzzy parameter in the objective function of the algorithm. The cluster validity index measures two attributes of clustering, cohesion and separation, as described in Section II. These attributes are optimized by minimizing the objective function and yields accurate clustering results in case of various data sets for certain values of the newly introduced parameter.

In this paper, we also describe a new measure for comparing clustering results obtained by different clustering algorithms. If the value of an individual attribute of a cluster center obtained is greater than its ideal value, we term it as positive error. And similarly, if the value is lower, we term it as negative error. Then we define the new measure as an absolute *error percentage* (formally introduced in Section V).

The paper is organized as follows. In Section II, we will discuss a popular method for fuzzy clustering, Fuzzy C-Means (FCM) Algorithm. We also define and explain the use of cluster validity indices and how it can be used as an objective function for obtaining clustering algorithms. In Section III, we propose our new clustering index. Section IV deals with the application of our Proposed Index for clustering well-known data-sets and tabulation of its results alongside the results of other fuzzy clustering algorithms. An analysis of these results in presented in Section V.

II. Background

A. Fuzzy C-Means Clustering Algorithm

A frequently used algorithm for fuzzy clustering is the Fuzzy C-Means (FCM) clustering algorithm. The method classifies a set of n p-dimensional points into c-fuzzy clusters through the use of a membership value. The membership value of a point in a specific cluster ranges from 0 to 1 and the summation of it's membership in the c-clusters is unity.

978-1-4577-0792-6/11/$26.00 ©2011 IEEE

Mathematically for any 'j'th point,

$$\sum_{i=1}^{c} \mu_{ij} = 1 \tag{1}$$

where μ_{ij} denotes the membership value of the j-th data point to lie in the i-th cluster.

This algorithm uses an objective function J_m subject to the above mentioned condition. The performance criterion is to minimize this function J_m over the i-th cluster center V_i (for fixed membership values) and μ_{ij} (for fixed V_i).

$$J_m = \sum_{i=1}^{c} \sum_{j=1}^{n} (\mu_{ij})^m ||x_j - V_i||^2 \tag{2}$$

This function on constrained and unconstrained minimization yield formulas for the membership (μ_{ij}) of all the n-points and the cluster centers (V_i) of the c-clusters respectively. They are mathematically expressed as follows,

$$\mu_{ij} = \left[\sum_{k=1}^{c} \left(\frac{||x_j - V_i||^2}{||x_j - V_k||^2} \right)^{\frac{1}{m-1}} \right]^{-1} \tag{3}$$

$$V_i = \frac{\sum_{j=1}^{n} (\mu_{ij})^m x_j}{\sum_{j=1}^{n} (\mu_{ij})^m} \tag{4}$$

Initial membership values (μ_{ij}) of all the n-points are given for all the c-clusters. Naturally, the first set of cluster centers (V_i) is determined using these initial membership values and (4). In the consecutive iteration, new membership values are computed by using (3) and made the present μ_{ij}. The cluster centers are then updated according to the newly generated membership values. The process goes on repeating itself and every iteration generates a new set of values for the cluster centers (V_i) and the membership (μ_{ij}). Termination of this process occurs when the cluster center values do not show any further change.

B. Cluster Validity Indices

Cluster validity indices are mathematical functions which are used to evaluate the quality of clustering of a clustering algorithm. Ideally, cluster validity indices should examine of the following aspects of partitioning:

1) Cohesion: The patterns within the same cluster should be as similar as possible for good clustering. This is a measure of the compactness of the data points within a cluster.

2) Separation: Good clustering indicates that the clusters formed should be well separated. The distance between the cluster centers is an effective measure of the separation of two clusters.

3) Partitional Stability: The main idea behind such an approach is that clustering solutions for two different data sets that have been generated by the same source should be similar. An estimate of the agreement of clustering solutions generated by a clustering algorithm and by a classifier trained using a second (clustered) data set has been used for measuring stability in [14].

In addition, Cluster validity indices can also be used to evaluate the optimal number of clusters in a data set.They can be of two types, hard [7] or fuzzy [11]. In this paper, we will deal with only fuzzy clustering validity indices.

As a fuzzy clustering validity function, the partition coefficient F was designed [4] to measure the amount of overlap between clusters.

$$F = \frac{1}{n} \sum_{i=1}^{c} \sum_{j=1}^{n} (\mu_{ij})^2 \tag{5}$$

where μ_{ij} denotes the membership value of the j-th point to lie in the i-th cluster. In this form, F is inversely proportional to the overall overlap between pairs of fuzzy subsets. Disadvantages of the partition coefficient are the lack of a direct connection to a geometrical property and its monotonic decreasing tendency with c, the number of clusters.

Another popular validity index was proposed by Fukuyama and Sugeno by exploiting cohesion and separation. Here the first term is a compactness measure and the second term is a degree of separation between each cluster and the mean ($\overline{V}$) of cluster centroids.

$$J_m = \sum_{j=1}^{n} \sum_{i=1}^{c} (\mu_{ij})^m ||x_j - V_i||^2 - \sum_{i=1}^{c} ||V_i - \overline{V}||^2 \tag{6}$$

where $\overline{V} = \frac{1}{c} \sum_{i=1}^{c} V_i$.

The more the separation and the lesser the compactness measure, the better is the clustering. Xie and Beni also proposed a validity index which focused on compactness and separation [11]. The objective function S that they used was:

$$S = \frac{\sum_{j=1}^{n} \sum_{i=1}^{c} (\mu_{ij})^2 ||x_j - V_i||^2}{n \left[\min_{i \neq k} ||V_i - V_k||^2 \right]} \tag{7}$$

Kwon later introduced a new validity index [8] which eliminated the monotonically decreasing tendency of the Xie-Beni index by introducing a punishing function, leading to better results as the number of clusters increased.

$$v_K = \frac{\sum_{j=1}^{n} \sum_{i=1}^{c} (\mu_{ij})^m ||x_j - V_i||^2 + \frac{1}{c} \sum_{i=1}^{c} ||V_i - \bar{x}||^2}{\left[\min_{i \neq k} ||V_i - V_k||^2 \right]} \tag{8}$$

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

C. Use of cluster validity indices in clustering methods

Cluster validity indices are basically objective functions, which can be used to obtain clustering indices, and obtain a clustering solution. This approach was first used by Xie and Beni as they showed that when minimizing their objective function over their cluster center and membership values, we got the functions of the Fuzzy C-means algorithm. In this subsection we have shown how the Kwon index can be minimized to form a clustering function, which partitions data sets into clusters and produces comparable clustering results to the fuzzy c-means algorithm.

We are to minimize the objective function v_K over V_i (for fixed partitions) and μ_{ij} (for fixed V_i). The minimization technique is shown in Sec III, when we derive our proposed clustering algorithm, where the process of minimization is the same as this. The functions for membership and cluster centers obtained after minimization is as follows:

$$\mu_{ij} = \left[\sum_{k=1}^{c} \left\{ \frac{||x_j - V_i||^2}{||x_j - V_k||^2} \right\}^{\frac{1}{m-1}} \right]^{-1} \quad (9)$$

$$V_i = \frac{\sum_{j=1}^{n} (\mu_{ij})^m x_j + \frac{\bar{x}}{c}}{\sum_{j=1}^{n} (\mu_{ij})^m + \frac{1}{c}} \quad (10)$$

Let us name this fuzzy clustering method as Kwon-based Fuzzy Clustering Method (KFCM). This clustering algorithm has been compared to the FCM method as shown in Sec IV and Sec V, and was found to produce comparable results.

III. The Proposed Clustering Method

The proposed objective function is:

$$J_m = \sum_{j=1}^{n} \sum_{i=1}^{c} (\mu_{ij})^m ||x_j - V_i||^2 - \frac{1}{c} \sum_{i=1}^{c} ||V_i - \overline{V}^{\gamma}||^2 \quad (11)$$

where,

$$\overline{V} = \frac{1}{c} \sum_{i=1}^{c} V_i \quad (12)$$

$$\overline{V}^{\gamma} = \begin{bmatrix} \overline{V}_1^{\gamma} \\ \overline{V}_2^{\gamma} \\ \vdots \\ \overline{V}_k^{\gamma} \end{bmatrix} \quad (13)$$

and γ is the new parameter for a k dimensional data-set.

The function is mathematically the difference of two distinct terms. The former term that is added measures the compactness of a certain clusters, where as, the latter term that is subtracted evaluates the separation among the clusters. Hence minimizing the objective function results in decreasing the compactness measure (hence increasing the compactness) of a certain cluster and increasing the separation among the clusters, resulting in better clustering of data. However, in case of certain data sets, solely this criteria may not lead very accurate clustering and results in an increase in the number of outliers. This was observed in case of data sets having overlapping clusters such as the breast cancer and the iris. Thus we integrated a new parameter γ to the objective function. By varying γ, we get better clustering results in case of such data sets while still optimizing the afore-mentioned attributes of clustering.

We minimize J_m over V_i (for fixed partitions μ) and μ_{ij} (for fixed V_i),

subject to $\sum_{i=1}^{c} \mu_{ij} = 1$.

First let us consider the above constrained optimization problem for fixed V_i. Then we need to satisfy (11) and (12) together. Using Lagrange′s multiplier method, we obtain that the problem is equivalent to minimizing

$$L(\mu, \lambda) = \sum_{j=1}^{n} \sum_{i=1}^{c} (\mu_{ij})^m ||x_j - V_i||^2 - \frac{1}{c} \sum_{i=1}^{c} ||V_i - \overline{V}^{\gamma}||^2 - \sum_{j=1}^{n} \lambda_k \Big(\sum_{i=1}^{c} \mu_{ij} - 1 \Big) \quad (14)$$

without constraints. Here we take the $\min$ term and $\overline{V}$ to be constants. This is taken as both the terms are assumed to be evaluated by the cluster centers of the previous iteration, and thus they remain constant while solving this optimization problem. The necessary condition for the problem is

$$\frac{\partial L}{\partial \mu_{ij}} = m\, (\mu_{ij})^{m-1} \Big(||x_j - V_i||^2 \Big) = 0 \quad (15)$$

$$\text{and} \quad \frac{\partial L}{\partial \lambda_k} = \sum_{i=1}^{c} \mu_{ij} - 1 = 0. \quad (16)$$

From (15) we have

$$\mu_{ij} = \left\{ \frac{1}{m\, \|x_j - V_i\|^2} \right\}^{\frac{1}{m-1}} \quad (17)$$

Substituting (17) in (16),

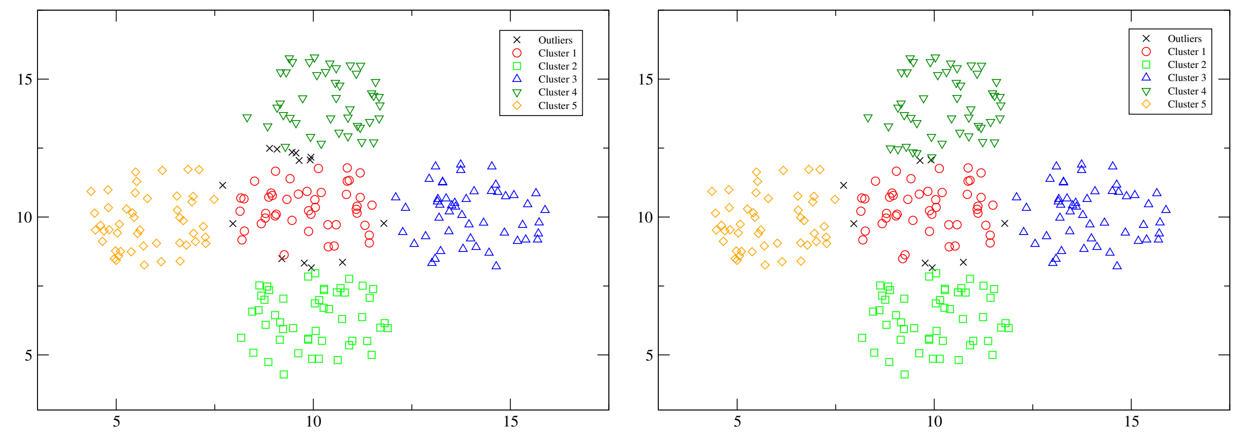
$$\left(\frac{\lambda_k}{m} \right)^{\frac{1}{m-1}} = \left[\sum_{i=1}^{c} \left\{ \frac{1}{||x_j - V_i||^2} \right\}^{\frac{1}{m-1}} \right]^{-1} \quad (18)$$

Substituting (18) in (17), we get

$$\mu_{ij} = \left[\sum_{k=1}^{c} \left\{ \frac{||x_j - V_i||^2}{||x_j - V_k||^2} \right\}^{\frac{1}{m-1}} \right]^{-1} \quad (19)$$

for $1 \le i \le c$ and $1 \le j \le n$

Now, suppose μ_{ij} is fixed. Then this is an unconstrained minimization problem, and the necessary condition is



(a) By FCM

(b) By Proposed Algorithm

Fig. 1. Clustering of the AD_5_2 Data Set with $m = 1.1$ and $\gamma = 1.1$.

$$\frac{\partial J_m}{\partial V_i} = -2 \sum_{j=1}^{n} (\mu_{ij})^m (x_j - V_i) - \frac{2}{c}(V_i - \overline{V}^{\gamma}) = 0 \quad (20)$$

which yields,

$$V_i = \frac{\sum_{j=1}^{n} (\mu_{ij})^m x_j - \frac{\overline{V}^{\gamma}}{c}}{\sum_{j=1}^{n} (\mu_{ij})^m - \frac{1}{c}} \quad (21)$$

A. *Algorithm of the Proposed Clustering Method*

The algorithm to be used for the proposed clustering method is as follows:

Input : Initial pseudo memberships μ_{ij} for $i = 1$ to c and $j = 1$ to n. Values of fuzzy factors m and γ.
Output : Final cluster centers.
Begin
Initialize $\overline{V}$ with 0
Repeat
For $i = 1$ to c
Evaluate V_i by (20)
End For
Calculate $\overline{V}$ by (12)
For $j = 1$ to n
For $i = 1$ to c
Put $old\_\mu_{ij} = \mu_{ij}$
Evaluate μ_{ij} by (18)
Let $current\_\mu_{ij} = \mu_{ij}$
End For
End For
Until $|current\_\mu_{ij} - old\_\mu_{ij}| \leq \epsilon$
where ϵ is a very small positive real number.
End

IV. Experiments and Simulation

A. *Data Sets Used*

Data Set	No. of Points	Clusters	Attributes
AD_5_2	250	5	2
Iris	150	3	4
Breast Cancer	569	2	30

B. *Results*

Figure 1 shows a graphical representation of the results obtained on clustering the artificial data set AD_5_2 by FCM and our proposed algorithm respectively. It shows that at a value of $m = 1.1$ and $\gamma = 1.1$, the number of points that are classified to each cluster. If a point does not have a membership of 0.6 or greater in its expected cluster we mark it as an outlier. Our proposed algorithm gives **43%** less outliers than FCM.

Our proposed fuzzy clustering method has also been compared to the fuzzy c-means clustering algorithm and to the fuzzy clustering method derived from the Kwon validity index KFCM, by clustering real world data sets such as the iris and the breast cancer.

Table I shows us some of the cluster centers obtained after the last run of the algorithm for the Breast Cancer Data Set. Here V_i denotes the i-th cluster and x_j denotes the j-th attribute of any cluster center. There are 4 columns denoting the ideal cluster centers, the cluster centers obtained from the fuzzy c-means algorithm, those obtained from KFCM, and the cluster centers obtained from our proposed algorithm. The results given were obtained keeping fuzzy factor $m = 2.5$ for our algorithm and KFCM, and $m = 4$ for FCM. For our algorithm, the optimal value of γ was found to be 1.47 to yield better

TABLE I
CLUSTER CENTERS OF BREAST CANCER DATA SET

x_j	V_1				V_2			
	Ideal	FCM	KFCM	Proposed	Ideal	FCM	KFCM	Proposed
x_1	**17.463**	**19.120**	**19.246**	**18.182**	**12.147**	**12.352**	**12.396**	**12.050**
x_3	**115.365**	**126.386**	**127.306**	**117.662**	78.075	79.618	79.989	76.864
x_7	**0.161**	**0.173**	**0.176**	**0.161**	**0.046**	**0.055**	**0.058**	**0.054**
x_{11}	**0.609**	**0.700**	**0.718**	**0.653**	0.284	0.293	0.299	0.294
x_{13}	**4.324**	**4.932**	**5.064**	**4.586**	**2.000**	**2.076**	**2.117**	**2.074**
x_{16}	0.032	0.033	0.032	0.031	0.021	0.022	0.023	0.023
x_{21}	**21.135**	**23.235**	**23.479**	**21.949**	**13.380**	**13.728**	**13.818**	**13.376**
x_{23}	**141.370**	**155.211**	**156.891**	**143.132**	**87.006**	**89.596**	**90.330**	**86.276**
x_{24}	**1422.286**	**1672.794**	**1719.014**	**1427.132**	**558.899**	**588.508**	**599.822**	**530.795**
x_{26}	**0.375**	**0.361**	**0.359**	**0.344**	**0.183**	**0.208**	**0.215**	**0.205**
x_{28}	**0.182**	**0.189**	**0.191**	**0.181**	**0.074**	**0.084**	**0.087**	**0.082**

clustering results. Attributes displayed in bold indicate those cases where our proposed algorithm gives the best result of the three.

Similarly, the Iris Data Set was clustered using the proposed algorithm for $m = 2.5$. More accurate results in comparison to the other algorithms mentioned above were obtained when γ made -0.11, its optimal value in this case.

V. PERFORMANCE ANALYSIS

A. Accuracy of Cluster Centers

From the results of Table I, we observe that FCM and the KFCM method yield cluster centers having high positive errors in most cases. As our proposed method depends on the value of each attribute of a cluster center, we get more accurate results.

For the analysis of clustering results in case of Breast Cancer and Iris data sets, we compute percentage error of the obtained cluster centers with respect to the ideal cluster centers.

We define the percentage error mathematically as:

$$E = \frac{\sum_{i=0}^{c}\sum_{k=0}^{d} |V_{ik} - V_{ideal}|}{\sum_{i=0}^{c}\sum_{k=0}^{d} V_{ideal}} * 100\%$$

where 'c' denotes the number of clusters and 'd' denotes the number of dimensions of the data set. V_{ik} denotes the k-th dimension of the i-th cluster center obtained from the clustering technique and V_{ideal} denotes the k-th dimension of the i-th ideal cluster center.

1) Clustering the AD_5_2 Data Set: In case of Figure 1, the values of compactness and separation were measured separately for both the algorithms: the FCM and the proposed one. The compactness and separation measures were used as defined by Fukuyama and Sugeno. As the clusters are very close to one another, it was expected that a moderately high value of compactness and separation would produce accurate clustering results. On varying γ, an optimal value of S was found to be at $\gamma = 1.1$. Our proposed algorithm gave an error percentage of only **0.72%** as compared to **1.16%** given by FCM. It also resulted in lesser number of outliers (better by **43%**). Thus in this case, our proposed algorithm performs better than the FCM in all the measures.

2) Clustering the Real World Data Sets: The iris data set has 3 clusters, two of which are overlapping. Each data point has 4 attributes, thus giving a total of 12 values. Of these 12 values we observe that our proposed method gives most accurate results for 8 of them, FCM gives better results for 1, while the method using Kwon′s index gives better values for the remaining 3. This has been illustrated in Fig. 2.

Using (21), we see for the fuzzy c-means method,

$E = \mathbf{1.20\%}$

for the method using Kwon's index KFCM,

$E = \mathbf{2.90\%}$

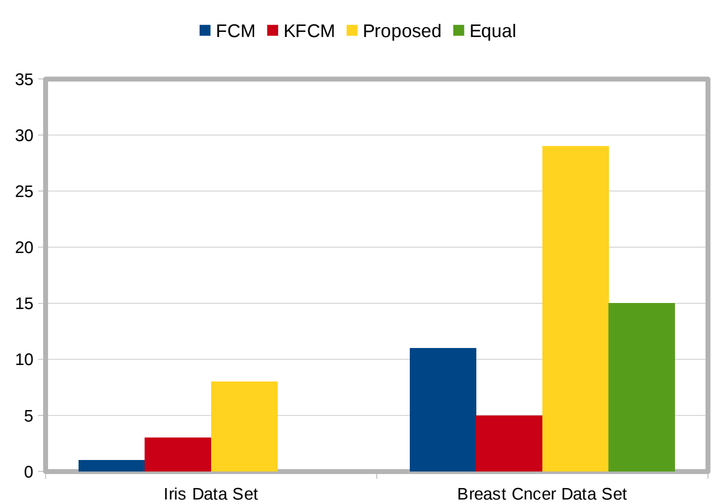


Fig. 2. Shows us the number of points that each clustering method gives best results for.

and for our proposed fuzzy clustering method,

$E =$ **0.99%**

The breast cancer data set has 2 clusters with each data point having 30 attributes, thus giving a total of 60 values. Our proposed clustering algorithm gives 29 values that are most accurate to the ideal ones, while KFCM gives 5 and the FCM gives 11.The most accurate results of the remaining 15 values were given by our proposed algorithm along with either the FCM or the KFCM. We denote this case as Equal in Fig 2.

Using (21), we compute the percentage error for the fuzzy c-means method,

$E =$ **12.60%**

for method derived from Kwon's index KFCM,

$E =$ **18.98%**

and for our proposed fuzzy clustering method,

$E =$ **2.28%**

As it can be seen, in both cases, the percentage of error in case of our proposed method is the lowest of the three algorithms and hence the values obtained are the most accurate of the three methods. Thus our clustering technique performs better than the FCM even when dimensions are high or two clusters overlap, as in case of the Breast Cancer and Iris data sets respectively.

B. Outliers elimination

In case of the Iris data set, let us define clustering a data point to be successful if the membership of the point in its correct cluster is more than 0.55. Thus we can calculated the percentage of data points an algorithm classified successfully. It was seen FCM successfully clusters **88.67%** of the data points. The proposed fuzzy clustering algorithm successfully clusters **90%** of the data points.

In case of the Breast Cancer data set, let us define clustering a data point to be successful if the membership of the point in its correct cluster is more than or equal to 0.6. Similarly calculating the percentage of successfully clustered data points in case of the breast cancer data set, we find FCM successfully clusters **86.11%** of the data points while the proposed fuzzy clustering algorithm successfully clusters **89.46%** of the data points.

Thus with the proposed clustering algorithm, we eliminate the outliers to a larger extent than the FCM algorithm.

VI. Conclusion

It has been already mentioned in Section II that the functions used for finding the cluster centers and membership values of the FCM algorithm can be obtained from the Xie-Beni Index for $m = 2$. However, the FCM and its variants did not provide any method to enhance clustering results for overlapping datasets for a given value of m.

In this paper, we thus propose an objective function along with a new parameter γ for the fuzzy clustering of artificial and real world data sets. This objective function is obtained from the Fukuyama-Sugeno Validity Index and it produces better results than the Fuzzy C-Means Algorithm by varying values of the new parameter 'γ' for a definite value of m. This is well illustrated when we see that the error percentage of the FCM algorithm in case of the Breast Cancer data set is **12.60%**, whereas, the proposed algorithm gives an error percentage of only **2.28%**.

But according to Pal and Bezdek's analysis [10], the Fukuyama-Sugeno index [5] is sensitive to both high and low values of weighing exponent m and may be sometimes unreliable because of this. Also all the fuzzy clustering methods referred to in this paper use Euclidean distance as a computational tool for measurement of distance between two points. This restricts the clustering to a definite shape. The fuzzification factors m and the new parameter γ also influence the shape of the clusters.

There are many other data-sets that will provide better results with the proposed fuzzy clustering method by adjusting the value of γ experimentally as well as mathematically by using the property of minimization of the objective function. We plan to extend our study on this topic.

References

[1] Anderson, E., *The IRISes of the Gaspe peninsula*, Bull. Amer. IRIS Soc. 59, 25, 1935.

[2] Baraldi, A., Blonda, P., *A survey of fuzzy clustering algorithms for pattern recognition. I.*, IEEE Trans. Systems, Man and Cybernetics, Part B, vol. 29, no. 6, pp.778 - 785, Dec. 1999.

[3] Bezdek, J. C., *Pattern Recognition with Fuzzy Objective Function Algorithms.*, Plenum, New York, 1981.

[4] Fukuyama, Y., Sugeno, M., *A new method of choos- ing the number of clusters for the fuzzy c-means method.*, Proceeding of fifth Fuzzy Syst. Symp., pp. 247 250, 1989.

[5] H. Roubos and M. Setnes, *Compact fuzzy models through complexity reduction and evolutionary optimization*, In Proc. of the Ninth IEEE International Conference on Fuzzy Systems, 2000.

[6] J. C. Dunn, *Indices of partition fuzziness and detection of clusters in large data sets*, Fuzzy Automata and Decision Processes. New York: Elsevier, 1977.

[7] Kwon, S.H., *Cluster validity index for fuzzy clustering*, Electron. Lett. 34 (22), 21762177, 1998.

[8] M. Halkidi and M. Vazirgiannis, *Clustering validity assessment: Finding the optimal partitioning of a data set*, in Proc. IEEE ICDM, San Jose, CA, 2001, pp. 187194.

[9] M. Hung and D. Yang, *An efficient fuzzy -means clustering algo- rithm*, in Proc. IEEE Int. Conf. Data Mining, 2001, pp. 225232.

[10] N. Pal and J. Bezdek, *On cluster validity for the fuzzy -means model*, IEEE Trans. Fuzzy Syst., vol. 3, no. 3, pp. 370379, Aug. 1995.

[11] Xie, X. L., Beni, G., *A validity measure for fuzzy clustering*, IEEE Trans. Pattern Anal. Machine Intell. 13, 841847, 1991.

[12] Yu, J., Yang, M.S., *Optimality test for a generalized FCM and its application to parameter selection*, IEEE Trans. Fuzzy Systems.

[13] Zadeh, L.A., 1965. Fuzzy sets. Inform. Control 8, 338353.

[14] Bandyopadhyay S., *Multiobjective simulated annealing for fuzzy clustering with stability and validity*, IEEE Transactions on Systems, Man, and Cybernetics – Part C: Applications and Reviews.

[15] Bandyopadhyay, S., Maulik, U., *Genetic Clustering for Automatic Evolution of Clusters and Application to Image Classification*, IEEE Trans., Pattern Recognition, vol.35, pp. 1197-1208, 2002.

[16] S. Bandyopadhyay, S. K. Pal, *Classification and Learning Using Genetic Algorithms: Applications in Bioinformatics and Web Intelligence*, Springer, Heidelberg, 2007.

[17] S. Bandyopadhyay, U. Maulik, *Nonparametric genetic clustering: Comparison validity indices*, IEEE Trans. SMC Part C, vol. 31, no. 1, pp. 120-125, 2001.