Passer Journal

Volume 3 | Issue 1 Article 2

11-3-2021

A new approach to the fuzzy c-means clustering algorithm by automatic weights and local clustering

Yadgar Sirwan Abdulrahman
University of Garmian

Available online at www.garmian.edu.krd

Passer Journal

Passer 3 (2021) 95- 101

http://passer.garmian.edu.krd/

A new approach to the Fuzzy c-means Clustering Algorithm by Automatic Weights and Local Clustering

Yadgar Sirwan Abdulrahman *

IT Department, University of Garmian, Kalar, Kurdistan Region, Iraq

Received 01 November 2020; revised 12 January 2021;
accepted 13 January 2021; available online 11 March 2021

doi:10.24271/psr.18

ABSTRACT

Clustering is one of the essential strategies in data analysis. In classical solutions, all features are assumed to contribute equally to the data clustering. Of course, some features are more important than others in real data sets. As a result, essential features will have a more significant impact on identifying optimal clusters than other features. In this article, a fuzzy clustering algorithm with local automatic weighting is presented. The proposed algorithm has many advantages such as: 1) the weights perform features locally, meaning that each cluster's weight is different from the rest. 2) calculating the distance between the samples using a non-euclidian similarity criterion to reduce the noise effect. 3) the weight of the features is obtained comparatively during the learning process. In this study, mathematical analyzes were done to obtain the clustering centers well-being and the features' weights. Experiments were done on the data set range to represent the progressive algorithm's efficiency compared to other proposed algorithms with global and local features.

Keywords: Clustering, fuzzy clustering, the sum of square errors, Local feature weighting, Non-Euclidean metric, Fuzzy C-Means clustering.

1. Introduction

Clustering is an unsupervised machine learning method and the primary technique in data mining, which seeks to discover the structure of data by finding similarities or differences between data points. Based on similarities, data points devide into different categories. These categories called clusters. Therefore, a cluster is acollection of input data, and some of these data are similar within and some are not to the data in the other clusters.[1]

Clustering is the process of dividing a set of data sets x into subcategories or clusters, based on whether they are similar or not. One of the most critical problems with clustering algorithms is being sensitive to the primary points and getting stuck in the local optima. Researchers have used many metaheuristic algorithms in the last two decades to solve these problems and reach optimal global points. This study presents a combined method using fuzzy C-Means and an artificial immune system algorithm for data clustering. In this method, instead of randomly selecting the primary points from the data set, they are selected from the artificial immune system algorithm's best points. Experiments showed that by selecting these points as the primary points in the fuzzy C-Means algorithm, the C-Means algorithm converges in few iterations. The proposed algorithm has been tested on a standard data set. Comparing the results obtained with other methods shows that the proposed algorithm performs well [2]. Clustering aims to find similar clusters of objects among the input samples, but how can we say that one cluster is appropriate and the other is not? It can be shown that there are no absolute criteria for the best clustering, but it depends on the user's problem and opinion that he should decide whether the samples are correctly clustered or not. Clustering, the most critical question of unsupervised learning, deals with the data structure partition in a new area and is the basis for further learning. [3] The complete definition for clustering, however, has not agreed, and a classic one is described as follows:

(1) I Instances, in different clusters, must be different as much as possible;[4]

(2) Measurement for similarity must be precise and have practical meaning;[4]

(3) nstances in the same cluster must be similar as much as possible[4];

In this study, we are using non-Euclidean distance as a parameter of similarity.To understand fuzzy clustering, it is necessary first to get acquainted with the concept of fuzzy sets and their

* Corresponding author
E-mail address: yadgarsirwan89@gmail.com (Instructor).
Peer-reviewed under the responsibility of the University of Garmian.

differences from classic sets. In classical collections, a reference collection member is either a collection (A) member or not a collection (A) member. In classical collections, the value of belonging to each collection is either 0 or 1. Now consider the collection of young and old humans. The question here is whether a person over the age of 25 is part of this collection. How about age 30 or 35? Anybody may have guessed; borderline certainty cannot be considered for young and old. The reason is, if 35 is supposed to be young, 36 can be young, as well as 37, 38, and so on. We are dealing with the concept of uncertainty here. We have used uncertainty many times in our daily lives, such as cold weather, hot water, expensive car, etc. All of the above are examples of fuzzy sets. The function of fuzzy sets affiliation is not two values (0 or 1) but can take any value between 0 and 1. [5]

Each data item may belong to different clusters with different degrees of affiliations. Hence the classical clustering methods have been extended using the concepts of fuzzy set and rough set theory to handle such fuzzy clustering problems. Both theoretical and empirical studies in the literature have proved that the existing clustering methods are doing exceptionally well on the data of suitable types. In classical clustering, each input sample belongs to one and only one cluster and cannot be a member of two or more clusters. Sometimes, each sample is a member of a cluster, and any sample is not a member of two clusters; in other words, the clusters do not overlap. Now consider a situation where the similarity of a sample of two or more clusters is the same. In clustering, it must be decided to which cluster this sample belongs to. The main difference between classical clustering and fuzzy clustering is that a specimen can belong to more than one cluster. To unsrestand fuzzy clustering, consider Figure 1. [6]

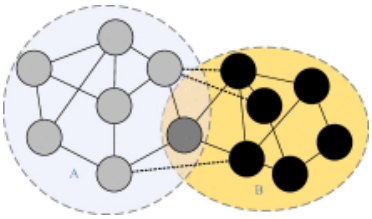

Figure 1: A schema of two local communities (A) and (B) with a standard node between them

If input samples are, as shown above, it is clear that the data can be divided into two clusters, but the problem is that the data specified in the middle can be a member of both clusters, so it must be decided that the data between (A) ,and (B) clusters, belongs to which cluster, the (A) cluster, or the (B) cluster. However, in fuzzy clustering, the questioned data belongs to the (A) cluster or the (B) cluster, a float between 0,1. Another difference is the example of the input samples on the right side of Figure 1 can also be a member of the left cluster with a shallow degree of affiliation, which is also true for the examples on the left. In classical clustering, each input sample belongs to one and only one cluster and cannot be a member of two or more clusters.

Each data item may belong to different clusters with different degrees of affiliations. Hence the classical clustering methods have been extended using fuzzy set and rough set theory to handle such fuzzy clustering problems. Both theoretical and empirical studies in the literature have proved that the existing clustering methods are doing exceptionally well on the data of suitable types.[7]

Nowadays, data mining science-especially clustering-has found complete application in various sciences. From geological studies to space studies, they all use a kind of clustering technique. Various convergences and research in social sciences such as social crime, students' academic failure and fight against crimes in cyberspace security, cyberspace banking affairs for clustering and how to treat each customer, etc., are done Various conferences are offered. The followings are some of these articles and researches.[8]

In this study, a new fuzzy clustering method with local weighting is presented. In this method, each of the weighting features is locally weighted. Thus, each feature weight's effect is different in various clusters. On the other hand, probability weighting is automatically performed through learning and clustering.

Also, in this algorithm, inspired by the idea used in the paper [9], the non-euclidean distance is used. The advantage of using this distance measure is that they assign less weight to outliers and less impact clustering. To evaluate the proposed algorithm's efficiency, we used different methods performed on five UCI data warehouses. The proposed algorithm is compared with the fuzzy c - means algorithm with global automatic weighting ,and hard thresholding with local automatic weighting. The obtained results showed high efficiency of the proposed algorithm in compare to the other algorithms.

The other parts of the paper are organized, as described in the second section of the necessary fuzzy c-means algorithm details. in the third second, the proposed algorithm is introduced, and the fourth second includes the results of the proposed algorithm, and finally, we will discuss what we can do in the future in the next section.

2. Basic Fuzzy C-means

In this section, we briefly explain the basic fuzzy C-means algorithm [10]. The objective function of this algorithm is defined In the following equation:

$$J = \sum_{i=0}^{N} \sum_{j=1}^{C} \mu_{ij}^{m} \; d^2 \, (xi - vj) \qquad (1)$$

The μij determines the degree to which the i-th sample belongs to the center of the cluster j, and m determines the fuzzy degree. Here $d^2 \, (xi - vj)$ is the non-Euclidean distance equal to $(xi - vj)^2$. As xi is the i-th sample and vj is the center of the j-th cluster. For this objective function, there are constraints $0 < \sum_{i=1}^{N} \mu_{ij} < N$ and $\mu_{ij} \epsilon [0 - 1]$ The values of the variables i and j are in the range of $1 \leq i \leq N$ and $1 \leq j \leq C$ so on, respectively, change. [10].

Based on the objective function introduced in equation (1), equations for improving the centers and functions of the affiliation will be as follows: [10].

$$\mu_{ij} = 1 / \sum_{l=1}^{C} \left(\frac{d^2 (x_i - v_j)}{d^2 (x_i - v_l)} \right)^{\frac{1}{(m-1)}} \qquad (2)$$

$$v_j = \left(\sum_{i=1}^{N} \mu_{ij}{}^{m} x_i\right) \Big/ \sum_{i=1}^{N} \mu_{ij}{}^{m} \quad (3)$$

3. Proposed algorithms

The proposed local weighting fuzzy C-means clustering algorithm (LWFCM) is formulated and then explained in a psuedu-code form with full details. In the following, the proof of convergence and its computational complexity will be explained.

3. 1 Formulation of the proposed algorithm

Clustering requires a need for an objective function is set according to which parameters of the Clustering algorithm are set. One of the most famous functions in this field is the sum of in-cluster distances. On the other hand, in many literature works, the Euclidean similarity measure is used to compute the objective function, but this criterion is highly sensitive to noise. In these metrics, all features are considered equal to the same weight.

In the proposed algorithm, a new objective function using a non-euclidean similarity metric is presented in [8]. Also, in this objective function, a sample's attachment rate to a cluster is a float between 0 and 1; for any features in each cluster, the specific weight of feature is considered. The objective function of the proposed algorithm is as follows:

$$J = \sum_{i=1}^{N} \sum_{j=1}^{C} \sum_{k01}^{M} \mu_{ij}^{m} w_{jk}^{\beta} d^2(x_{ik} - v_{jk}) \quad (4)$$

Here μ_{ij} is the i-th sample's attachment rate to the center of the j-th cluster, The weight of the k-th feature in the j-th cluster is $w_{jk,}$ And $d^2(x_{ik}\text{-}v_{jk})$ is a non-Euclidean distance type. It is often defined as $1\text{-}\exp(-\Upsilon_k(x_{ik}\text{-}v_{jk})^2)$ [8] For this objective function, the constraints are also $0 < \sum_{i=1}^{N} \mu_{ij} < N$ and $\mu_{ij} \epsilon [0-1]$, $\sum_{j=1}^{C} \mu_{ij} = 1$, $\sum_{k=1}^{M} w_{jk} = 1$ that the values of the variables ii, j, and k are changing in the range of $1 \le i \le N$, $1 \le j \le C$ and $1 \le k \le M$ respectively.

Here C is the number of clusters, and M is the number of features. Due to the constraints mentioned above, the Lagrangian equation of the objective function of the proposed algorithm can be written as follows:

$$J = \sum_{i=1}^{N} \sum_{j=1}^{C} \sum_{k01}^{M} \mu_{ij}^{m} w_{jk}^{\beta} d^2(x_{ik} - v_{jk}) \text{-} \lambda \left[\sum_{j=1}^{c} \mu_{ij} - 1\right] - \alpha(\sum w_{ij} - 1) \quad (5)$$

λ and α are the parameters of the Lagrange equation.

By solving the above Lagrangian equation, according to KKT(Karush-Kuhn-Tucker) conditions, the following update formulas for μ_{ij}, v_{jk}, and w_{jk} Are obtained. In fact, by deriving from the function of the above objective concerning μ_{ij}, v_{jk}, and w_{jk} and equating zero with the product of the derivative and solving the equation, the following formulas are obtained:

$$\mu_{ij} = 1 \Big/ \sum_{l=1}^{C} \left(\frac{\sum_{k=1}^{M} w_{jk}^{\beta} d^2(x_{ik} - v_{jk})}{\sum_{k=1}^{M} w_{lk}^{\beta} d^2(x_{ik} - v_{lk})}\right)^{1/(m-1)} \quad (6)$$

$$v_{jk} = \sum_{i=1}^{N} \mu_{ij}^{m} \exp\left(-\gamma_k (x_{ik} - v_{jk})^2\right) x_{ik} \Big/ \sum_{i=1}^{N} \mu_{ij}^{m} \exp\left(-\gamma_k (x_{ik} - v_{jk})^2\right) \quad (7)$$

$$w_{jk} = \begin{cases} \frac{1}{m-1} & if\ D_{jk} = 0, m_k = \left|\{l: D_{jl} = 0\}\right| \\ 0 & if\ D_{jk} \neq 0, but\ \exists l\ s.t.\ D_{jl} = 0 \\ \frac{1}{\sum_{l=1}^{M} \left(D_{jk} / D_{jl}\right)^{1/(\beta-1)}} & if\ D_{jl} \neq 0 \end{cases} \quad (8)$$

Djk The sum of the distances within the cluster is according to the k-th attribute, and the non-euclidian distance, which is defined as follows:

$$D_{jk} = \sum_{i=1}^{N} \mu_{ij}^{m} \left(1 - \exp(-\Upsilon_k (x_{ik} - v_{jk})^2)\right) \quad (9)$$

The steps of the proposed algorithm are summarized in pseudo-code 1. It should be noted that the elements of the weight matrix are the same in the initial set-up stage, so that w_{jk} =1/k, where k = 1,…, M, and M, the number of features in the data set. Then, the primary centers are randomly selected from the samples

pseudo-code.1 Proposed clustering algorithm is the number of samples and M is the number of features, and xik is the k-th features of the i-th sample.

Output: attachment matrix μ, the matrix of centers V, and weight matrix W.

Step 1: Primary value of the number of clusters C, degree of phasing m, the threshold value in the condition of stopping ε, T, the number of iterations, weight matrix W, and V, the centers matrix.

Step 2: Update the μ attribute matrix from the formula) (6)

Step 3: Update the V Center matrix from the formula (7)

Step: 4 Update the W weights matrix from the formula (8)

Step 5: If the condition $\left|v_i^t - v_j^{t-1}\right| < \varepsilon$ is not met as the vjt center of cluster j-th is in t repeating t or the number of iterations is less than T, Go to step2. Otherwise, go to step 6.

Step: 6 final center V matrix, attachment matrix μ, and features weight matrix.

3. 2 Computational complexity

Given the pseudo-code of the proposed algorithm, it can be seen that the computational complexity of the second stage is equal to NMC^2, which is N the number of samples, C is the number of clusters, and Is the number of features. In the second step, we encounter two sets of nests, the first of which is repeated twice and the second twice, and this operation is performed for each sample in each cluster. On the one hand, the computational complexity of the third stage is equal to NMC because we are dealing with a sum that is repeated N times, and on the other hand, this operation is for each feature in each cluster.

As a result, the computational complexity of the whole is equal to $T(C^2+NMC)=TNCM(C+1)$, which is almost the same as $O(TNC^2M)$, where T is the number of repetitions of the algorithm.

3. 3 Proof of convergence

In this section, the proof of proposed algorithm convergence is explained. Our purpose is proving the relation (5) function is

ultimately minimized or, in other words, converged, according to the formulas obtained to improve its components, given that the objective function is the sum of the distances within the pleasures, the smaller the value, the better.

Theorem 1. Assuming that the center matrix and the weight matrix are constant, the affiliation matrix is the local least related relation (5). The correlation is defined only if for any element (6) because the objective function has a direct relationship with the affiliation matrix.

Proof 1. As the relationship (6)'s μ_{ij} from deriving from the relationship (5) to μ_{ij} and equal to zero is obtained; therefore, it is either minimal or maximum. The relationship (5) derives from μ_{ij} And shows that the second partial derivative is positive (5). It can be proved that μ_{ij} Is defined by relation(6) is a local minimum of relation(5). Therefore, in the following, we take the second derivative from the relation (5.).

The result of this derivative will be as follows:

$$\sum_{k=1}^{M} m(m-1)\,\mu_{ij}^{m-2} w_{jk}^{\beta}\, d^2(x_{ij}-v_{jk}) \qquad (10)$$

Since a, b, m> 1 can be shown $\sum_{k=1}^{M} m(m-1)\,\mu_{ij}^{m-2} w_{jk}^{\beta}\, d^2(x_{ij}-v_{jk})$ As a result of the relation (6) is a local minimum of relation (5).

Theorem 2. Assume that the affiliation matrix μ and the matrix of weights W are constant. W centers' matrix is then defined as a local minimum of the relation (5) if the correlation defines v_{jk} (7).

Proof 2. Proof 2 is the same as proof 1. It is sufficient to show that the second derivative of the relation (5) is positive for v_{jk} because v_{jk} derives from the relation (5) concerning v_{jk} and placing it equal to zero. Relation (5) is a positive expression for As a result of the relationship (7) is also a local minimum of the relationship (5).

Theorem 3. Assume that the affiliation matrix μ and the matrix of centers V are constant. The weight matrix W is defined as a local minimum of the relation (5) if the relation (8) defines w_{jk}.

Proof 3. Proof of Theorem 3 is similar to Proof 1 and 2. That is, it is sufficient to show that the second derivative of the relation (5) is positive for w_{jk} because w_{jk} is a deriving from the relation (5) concerning w_{jk} and placing it equal to zero. The second derivation of relation (5) is as same as w_{jk} compared to $\sum_{i=1}^{N} \mu_{ij}^{m-2}\,\beta w_{jk}^{\beta-1}\, d^2(x_{ij}-v_{jk})$ (which is a positive statement). As a result, the relationship is (8) is also a local minimum of the relationship (5).

Since the goal is minimizing the relationship (5), according to theorems 1-3, and the formulas for improving the center matrix v the relationship (6); and the weight matrix W, the relationship (7) and the affiliation matrix m, the relationship (8), we can be assured the objective function ultimately reaches its minimum value or converges.

4. Results and Discussions

In this section, the performance of the proposed algorithm (PROPOSED ALGO) is investigated, and its results are combined with C-Means fuzzy clustering algorithms with global weighting [9] (RFWFCM and C-Means hard drive with local buffer) [7]. The evaluation of the proposed algorithm in comparison with the two methods RFWFCM and RLFWHCM on five real data sets including Breast-cancer, Bupa, Glass, Wine, and Iris, from the UCI data warehouse. These data sets are summarized in Table 1. The standard parameters in all algorithms, including the threshold value ε and the number of T iterations, are set to $\varepsilon = 10^{-5}$ and T = 200.

Table 1: Five real datasets from UCI data wearhouse used in experiments

Dataset name	Number of samples	number of dimensions	Number of classes
Iris	150	4	3
Wine	178	13	3
Glass	214	10	6
Bupa	345	6	2
Breast	683	10	2

Also, the parameter m in the RFWFCM algorithm and the proposed algorithm are set as m = 2. The number of clusters per data set is the same as the number of classes in the data set Was used. the objective function's value and the error rate is used to measure the efficiency of the algorithms.

Table 2 shows the mean error rate and execution time and the proposed algorithm's objective function after 200 iterations on the five data sets used for β= 2. According to Table 2, it can be seen that The proposed algorithm's time is longer than the RLWHCM algorithm for all data sets. On the other hand, the proposed algorithm's execution time is even longer than the execution time of RFWFCM.

However, for all data sets used, the objective function and the proposed algorithm's error rate are less than the objective function and the error rate of both other algorithms. Respectively, this is a strong reason for increasing its efficiency compared to the other two algorithms after running on real data sets.

Table 3 shows the mean error rate of RLWHCM, RFWFCM, and PROPOSED ALGO algorithms after 200 iterations on five data sets used for different values of β, and Table 4 averages the objective function of the repetition algorithm over 200 iterations after PROPOSED ALGO and RFWFCM,

The five sets of data used for different values of β represent.

In this paper, the parameter is important and affects the efficiency of algorithms. Here, considering the correct values β between -6 and +6 for the parameter ⧺ its effect on the algorithms' performance is studied. For this purpose, ten series of primary centers are produced for each data set, and each algorithm is executed ten times on each data set and each time during 200 iterations. The algorithms' average objective function is shown in Table 3, and the average error rate is shown in Table 3. In Figure 2, the vertical axis of the Error rate and the horizontal axis is the value of the parameter.

Table 2: Execution time, and objective function and average error rate of RLWHCM, RFWFCM, and PROPOSED ALGO algorithms

dataset	method	Time (MS)	Objective function	Error rate	Time (MS)	Objective function
		average	average	variance	average	average
Iris	RFWFCM	8242.6	2.59	0.44	12.153	4.33
	RLWHCM	199	3.21	0.007	6.93	0.68
	PROPOSED ALGO	848269.1	2.36	0.003	5.66	0.001
Wine	RFWFCM	12116	1.82	0.0001	27.76	0.0001
	RLWHCM	456.7	0.1	0.3	27.92	1.22
	PROPOSED ALGO	12130.1	0.001	0.003	27.58	0.002
Glass	RFWFCM	13256.6	1.82	0.82	0.05	2.63
	RLWHCM	1317.9	3.11	0.1	0.11	1.04
	PROPOSED ALGO	13298.5	0.0003	0.001	0.006	0.42
Bupa	RFWFCM	12977.3	0.05	8.67	0.0002	3.73
	RLWHCM	679.4	0.11	11.98	0.01	0.74
	PROPOSED ALGO	12985.1	0.006	8.62	0.0001	4.4
Breast	RFWFCM	50702.7	0.0002	8.02	0.19	3.09
	RLWHCM	1442.7	0.01	5.17	1.24	3.46
	PROPOSED ALGO	50707.2	0.0001	4.41	0.02	2.03

Table 3 shows that the value of the proposed algorithm's target function for all data sets used after 200 iterations for different values of β is less than the other two algorithms. Nevertheless, Table 4 shows that the proposed algorithm error rate for all Data sets for all different values of β are not less than the error rate of the other two algorithms. In fact, for the glass data set for β = 4, for the Bupa for β = -6, and the breast for β = -2, the error rate of the proposed algorithm is not less than the error rate of the other two algorithms.

Table 3: The average objective function of RLWHCM, RFWFCM, and PROPOSED ALGO after 200 iterations on five real data sets used for different values β

		β					
Data set	method	-6	-4	-2	2	4	6
Iris	RFWFCM	213701 .3	13459 .04	854 .75	2 .59	0 .18	0 .0119
	RLWHCM	296343 .3	18891 .69	1289 .76	3 .21	0 .24	0 .0167
	PROPOSED ALGO	208098 .6	13136 .62	838 .59	2 .36	0 .17	0 .0116
Wine	RFWFCM	$150* 10^7$	8882560	52569 .60	1 .82	0 .0108	0 .0001
	RLWHCM	$279* 10^7$	16581100	99147 .44	3 .11	0 .0194	0 .0001
	PROPOSED ALGO	$148* 10^7$	8759858	51983 .53	1 .75	0 .0106	0 .0001
Glass	RFWFCM	95255856	425330	11858 .64	8 .67	0 .25	0 .0069
	RLWHCM	NaN	NaN	NaN	0 .1	0.002	0.0003
	PROPOSED ALGO	88545798	896384	9188 .76	0 .001	0	0
Bupa	RFWFCM	15286659	425330	11858.68	8.67	0.25	0.0069
	RLWHCM	25234338	705509	19854.07	11.98	0.38	0.0109
	PROPOSED ALGO	15141407	421295	11748.20	8.62	0.24	0.0068
Breast	RFWFCM	$858* 10^6$	8600380	86588 .09	8 .02	0 .08	0 .0008
	RLWHCM	$1270 * 10^6$	12241971	124179 .57	5.17	0.15	0.0017
	PROPOSED ALGO	$806* 10^6$	8107069	85. 82503	4.41	0.07	0.0005

Table 4: The mean error rate of RLWHCM, RFWFCM, and PROPOSED ALGO after 200 iterations on five real data sets used for different values β

		β					
Data set	method	-6	-4	-2	2	4	6
Iris	RFWFCM	13.13	13.7	13.7	12.53	12.53	12.53
	RLWHCM	7.8	7.86	12.66	6.93	7.93	11.96
	PROPOSED ALGO	6.73	6.73	6.73	5.66	6.93	6.93
Wine	RFWFCM	27.8	27.89	27.80	27.76	27.80	27.80
	RLWHCM	28.31	28.10	28.14	27.92	27.48	27.75
	PROPOSED ALGO	27.07	27.12	27.19	27.58	26.79	26.62
Glass	RFWFCM	23.29	24.19	22.53	30.70	27.71	22.71
	RLWHCM	23.29	27.71	30.70	22.53	24.49	23.19
	PROPOSED ALGO	22.93	29.12	21.47	23.31	24.38	30.33
Bupa	RFWFCM	32.98	32.13	20.23	22.18	23.10	18.64
	RLWHCM	33.36	33.33	32.32	33.73	33.82	32.98
	PROPOSED ALGO	32.86	32.86	33.37	33.63	34.38	33.36
Breast	RFWFCM	34.04	34.02	30.34	32.25	33.44	32.86
	RLWHCM	37.81	37.81	37.42	33.79	31.12	34.04
	PROPOSED ALGO	33.49	33.49	31.78	32.38	31.82	32.88

As a result, the proposed algorithm is not more efficient than the other two algorithms only when it is small. From Table 3 and Table 4, and Figure 2, it can be shown that the proposed algorithm is better for most different values of the parameter β than other compared algorithms. Besides, Figure 2 can quickly determine the appropriate β parameter value for each data set. For example, the appropriate value of the parameter β for the proposed algorithm is equivalent to the iris data set, 2 for the balanced wine data set, 6 for the equivalent glass data set, 6 for the Bupa data set, 2 for the breast-cancer data set, 2.

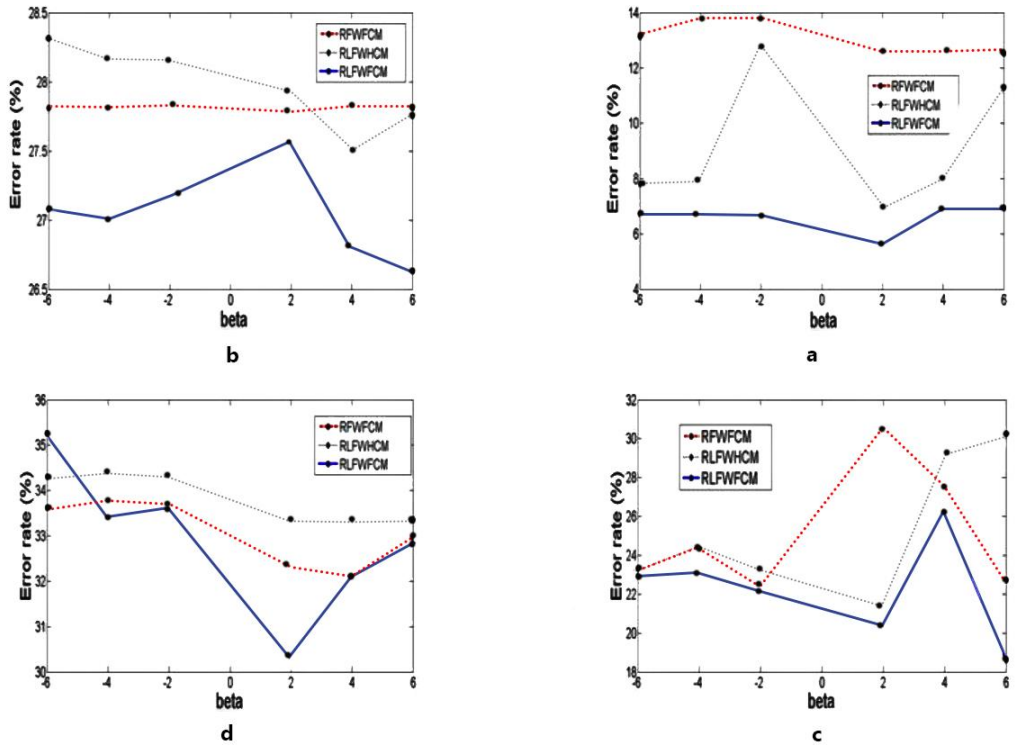

Figure 2: The average rate of error charts of "Robust local feature weighting hard c-means clustering algorithm" (RLWHCM), "Robust local feature weighting hard c-means clustering algorithm"(RFWFCM), and RLFWFC on real datasets after 200 iterations on different levels: a) iris, b) wine, c) glass, d) Bupa

5. Conclusion

So far, many kinds of research have been done to design a clustering algorithm with multi-stage clustering. However, in none of those, both fuzziness and local weighting are not applied at the same time, in addition to the euclidean similarity measure using, which is very sensitive to noise. In this paper, the proposed algorithm is proposed defining a new objective function and then defining the updating rules. In this paper, we prove the validity and the computational complexity of the proposed algorithm.

The results showed that the proposed algorithm's performance is more significant than the objective function, and the error rate decreased in the same algorithms. One of the tasks that can be done to improve the proposed algorithm's efficiency is the automatic determination of the number of partitions. A cluster-based classifier can be used—the next work to be done in developing the proposed algorithm to run over a nominal attribute set.

References

1. T. Soni Madhulatha Associate Professor, " AN OVERVIEW ON CLUSTERING METHODS ", IOSR Journal of Engineering Apr. 2012, Vol. 2(4) pp: 719-725 ISSN: 2250-3021 www.iosrjen.org 719 | P a g e, (2012)
2. Wang, Fu Zhong, et al. "Research on Transformer Fault Diagnosis Method Based on Artificial Immune Network and Fuzzy C-Means Clustering Algorithm." Applied Mechanics and Materials, vol. 574, Trans Tech Publications, Ltd., pp. 468–473. Crossref. https://doi.org/10.4028/www.scientific.net/AMM.574.468. (2014)
3. Jain A, Dubes R Algorithms for clustering data. Prentice-Hall, Inc, Upper Saddle River (1988)
4. Xu, D., Tian, Y. A Comprehensive Survey of Clustering Algorithms. Ann. Data. Sci. 2, 165–193. DOI: 10.1007/s40745-015-0040-1 (2015)
5. Yin D, Hong L, Davison BD "2111, Structural link analysis and prediction in microblogs the 21st ACM
   International Conference on Information and Knowledge Management (CIKM'11),(Glasgow, UK), DOI: https://doi.org/10.1145/3269206.3271740, (2018)
6. Kalantari, H., Ghazanfari, M., Fathian, M., & Shahanaghi, K. A new multi-objective optimization model in the one-layer weighted network through key nodes identification in overlapping communities. Computers & Industrial Engineering, 106413. doi:10.1016/j.cie.2020.106413 , (2020).
7. Reddy, Keshava & K., Mrudula. Hard And Fuzzy Clustering Methods: A Comparative Study. (2017).
8. XB Zhi, JL Fan, and F. Zhao, "Robust local feature weighting hard c-means clustering algorithm," Neurocomputing, vol. 134, pp. 20-29, https://doi.org/10.1007/978-3-642-31919-8_75 (2014)
9. KL. Wu and M.S. Yang, "Alternative c-means clustering algorithms," Pattern Recognition, vol. 35, pp. 2267-2278, https://doi.org/10.1016/S0031-3203(01)00197-2. (2002)
10. J.C. Bezdek, Pattern Recognition with Fuzzy Objective Function Algorithms. New York: Springer US, (1981)
11. Tabourier, L., Bernardes, D. F., Libert, A. S., Lambiotte, R., Rankmerging: A supervised learning-torank framework to predict links in the extensive social network, The Journal of Systems and Software, 2115, Vol. 88, No. 3, pp. 1-29. https://doi.org/10.1007/s10994-019-05792-4. (2019)