EAI Endorsed Transactions
on Scalable Information Systems
Research Article EAI.EU

A New Suppression-based Possibilistic Fuzzy c-means Clustering Algorithm

J. Arora[1,*], M. Tushir[2] and S. K. Dadhwal[1]

[1]Department of Information Technology, MSIT, India
[2]Department of Electrical & Electronics Engineering, MSIT, India

Abstract

Possibilistic fuzzy c-means (PFCM) is one of the most widely used clustering algorithm that solves the noise sensitivity problem of Fuzzy c-means (FCM) and coincident clusters problem of possibilistic c-means (PCM). Though PFCM is a highly reliable clustering algorithm but the efficiency of the algorithm can be further improved by introducing the concept of suppression. Suppression-based algorithms employ the winner and non-winner based suppression technique on the datasets, helping in performing better classification of real-world datasets into clusters. In this paper, we propose a suppression-based possibilistic fuzzy c-means clustering algorithm (SPFCM) for the process of clustering. The paper explores the performance of the proposed methodology based on number of misclassifications for various real datasets and synthetic datasets and it is found to perform better than other clustering techniques in the sequel, i.e., normal as well as suppression-based algorithms. The SPFCM is found to perform more efficiently and converges faster as compared to other clustering techniques.

Keywords: fuzzy c-means; FCM; possibilistic c-means; PCM; possibilistic fuzzy c-means; suppression possibilistic fuzzy c-means

Received on 19 July 2022, accepted on 22 December 2022, published on 03 January 2023

doi: 10.4108/eetsis.v10i3.2057

*Corresponding author. Email:joy.arora@gmail.com

1. Introduction

Data mining is the technique of processing through enormous data sets to find patterns and associations that can be used to classify and solve problems related to data analysis. It focuses on predicting the future and discovering patterns in the data using specialised machine learning and statistical models [1-3]. In the realm of data mining, clustering is one of the most essential strategy for assigning most-similar objects in one cluster and dissimilar objects in the other cluster based on some user defined similarity metrics. Clustering analysis is employed in various applications such as pattern recognition, image segmentation, sentiment analysis, etc [4-7]. These clustering methods can be used in different routing based algorithms in the field of wireless sensor networks and helps to improve the energy efficiency by clustering the WSN into different regions [8]. There are two basic clustering approaches that have been developed, based on the related pattern between a data vector and distinct clusters: crisp clustering methods [9] and fuzzy clustering methods [10-13]. Crisp clustering approaches, such as hard c-means (HCM), are based on hard partition (hard division) of the data. These approaches show the membership degree of a data vector to a class by utilising '1' to symbolise that it belongs to a set and '0' to represent that it does not belong to a set, based on the classic set theory. As a result, they assign each data vector into exactly one cluster, which has the advantages of being simple in concept and execution. However, because the membership assignment is based on the two-value logic function in classic set theory, they have the problem of producing unsatisfactory clustering results. On the other hand, fuzzy clustering approach allows data points to belong to more than one cluster simultaneously by using multiple membership values. One of the most widely used fuzzy clustering algorithms is Fuzzy c-means (FCM) by

Dunn [11] and Bezdek [12] where one data point can belong to more than one cluster based on the degree of membership. The fuzzy theory of clustering can more objectively reflect the real world than the crisp clustering analysis, and it has been used in a variety of applications. However, FCM employs a membership constraint where sum of membership of each data point is one which makes FCM more sensitive to noise and outliers. The constraint mandates that the sum of a data point's membership degrees to all clusters be 1, which allows the FCM to generate memberships that emphasise the relative property over the absolute property between a data vector and the clusters. However, in FCM noise vectors/ outliers, are allocated large membership degrees, having a negative impact on cluster centre computation.

Krishnapuram and Keller [13] proposed possibilistic c-means (PCM) clustering by relaxing the constraints on the membership degree and introducing possibilistic membership known as typicality, that better describe the absolute distances between data points and clusters. As a result, noise vectors/outliers, will be given very modest probabilistic membership degrees, drastically reducing the impact of noise points on the outcomes and overcoming FCM's noise sensitivity. However, the PCM pays a price for its weaker membership constraint as it tends to produce coincident clusters. It has other disadvantages of highly sensitive to initialization and problem of setting parameters. Fan et al. [14] proposed an improved model of FCM by incorporating a suppressed competitive learning mechanism into FCM, known as suppressed fuzzy c-means (SFCM) clustering. This method modifies the memberships of a data point to all the clusters in such a way that the highest membership (the winner) is prized and all others are suppressed, without disturbing the original order among them. Simultaneously, the SFCM utilizes a suppression rate $\alpha \in [0,1]$ to control the suppression strength. The SFCM with a reasonable parameter setting has a higher convergence speed and better performance than that of the FCM. A similar suppression based approach is employed in the typicality matrix in the PCM algorithm to generate the SPCM algorithm [15], which in turn has also achieved similar results as the SFCM with better convergence speed and performance than PCM.

To overcome the drawbacks of FCM and PCM, Pal et al. [16-17] proposed Possibilistic Fuzzy c-means (PFCM) clustering approach that generates membership and typicality values while clustering the unlabelled data. PFCM is a combination of both the FCM algorithm and the PCM algorithm. PFCM fixes the noise sensitivity of FCM and the coincident clusters problem of PCM. In the PFCM algorithm, the noise data has an influence on the estimation of centroids and converges slowly over the large datasets. To overcome these problems, this paper proposes a suppression-based possibilistic fuzzy-c means (SPFCM) clustering algorithm that combines the suppression approach with the possibilistic and probabilistic part of PFCM. SPFCM is obtained by applying suppressed competition learning mechanism into PFCM. The suppression mechanism is applied to both the membership matrix (FCM part) and the typicality matrix (PCM part). Further, we compare the performance of the proposed SPFCM algorithm on synthetic and several real datasets with different clustering approaches. The proposed method strengthens the susceptibilities present in the SFCM and SPCM.

Further, the paper is organized as follows; section 2 describes preliminary work. Section 3 discusses the proposed SPFCM algorithm, and Section 4 consists of the experimental results and discussion, followed by conclusions in section 5.

2. Preliminary Work

The behaviour of the clustering approaches is governed by the mathematical formulation of the objective function. This section discusses different clustering approaches related to the proposed work.

2.1. Fuzzy c- means (FCM) algorithm

FCM is one of the most widely used fuzzy clustering algorithms [10]. The objective function of the FCM algorithm is defined as follows:

$$J_{FCM} = \sum_{i=1}^{c}\sum_{k=1}^{n}\mu_{ik}^{m} d\left(x_k, v_i\right)^2, 1 \prec m \prec \infty \tag{1}$$

where, the total number of the pattern in the dataset is denoted by n and c defines the number of clusters, μ_{ik} represents membership function, $d\left(x_k, v_i\right)$ represents the Euclidean distance between data points x_k and center of the cluster v_i and m defines the degree of the fuzziness of the resulting partitions. The membership function and center are given by minimizing the objective function of FCM under the probability constraint:

$$\sum_{i=1}^{c}\mu_{ik} = 1 \forall k \tag{2}$$

Minimizing the objective function with respect to μ_{ik} and setting it to zero, we get the equation for membership value, i.e.,

$$\mu_{ik} = \frac{1}{\sum_{i=1}^{c}\left(\frac{d\left(x_k, v_i\right)}{d\left(x_k, v_j\right)}\right)^{\frac{2}{m-1}}} \tag{3}$$

Minimizing the objective function with respect to cluster center v_i, the equation for cluster centers i.e.,

$$v_i = \frac{\sum_{k=1}^{n} \mu_{ik}^m x_k}{\sum_{k=1}^{n} \mu_{ik}^m} \tag{4}$$

The main drawback of FCM is the presence of membership constraint that makes it sensitive to noise points and outliers. Also, it is ineffective in detecting clusters of different shapes other than spherical [15-16]. Since the degree of belongingness for data is not always represented best using FCM membership, a possibilistic approach called PCM was introduced.

2.2. Possibilistic c-means (PCM) algorithm

Krishnapuram and Keller [10] relaxed the column sum constraint of FCM and offered a possibilistic approach to clustering to overcome the noise sensitivity problem of FCM. The objective function of Possibilistic c-means (PCM) is defined as follows:

$$J_{PCM} = \sum_{i=1}^{c}\sum_{k=1}^{n} t_{ik}^m d(x_k, v_i)^2 + \sum_{i=1}^{c}\gamma_i \sum_{k=1}^{n}(1 - t_{ik})^m \tag{5}$$

Where γ_i is the parameter with the positive value and the membership of FCM is replaced by typicality t_{ik}. The first term requires that the distance between data points and the cluster centres be as small as possible, while the second term requires that t_{ik} be as high as possible, avoiding the simple solution. The cluster centres of PCM are updated in the same way as in FCM, but PCM typicality metrics is changed as follows:

$$t_{ik} = \frac{1}{1 + \left(\|x_k - v_i\|^2 / \gamma_i\right)^{1/(m-1)}} \tag{6}$$

2.3. Supressed Fuzzy c-means (SFCM)

Fan et al. [14] proposed the suppressed fuzzy c-means clustering (SFCM) technique to solve the slow convergence rate problem of FCM for large datasets. By changing the membership matrix in each iteration, the SFCM model brings a suppressed competitive learning mechanism into FCM as follows:

The membership μ_{wk} with the highest value of the given data point x_k is selected among all c clusters. The cluster w is considered as the closest cluster to the point x_k and declared as the winner. The membership μ_{wk} is referred to as winner membership, while the other memberships μ_{ik}, where $i \neq w$ are referred to as non-winning memberships as shown in Eq. 7.

$$\mu_{wk} = \begin{Bmatrix} \alpha_{fcm}\mu_{wk}, i \neq w \\ 1 - \alpha_{fcm}\sum_{i\neq w}\mu_{ik}, i = w \end{Bmatrix} \tag{7}$$

The strength of the suppression is controlled by the parameter α_{fcm}, which ranges from 0 to 1. When $\alpha_{fcm} = 0$, all non-winner memberships are set to 0, and the SFCM algorithm becomes identical to the hard clustering algorithm. There is no suppression when $\alpha_{fcm} = 1$, and the SFCM algorithm is the same as the FCM algorithm. The SFCM with a reasonable suppression rate α_{fcm} can improve the convergence speed of the FCM while maintaining good clustering accuracy.

In SFCM, for each iteration, each input vector x_k goes through competition, the cluster whose prototype is situated at the shortest distance from x_k wins. Fuzzy membership of any data point x_k with respect to any non-winner cluster is suppressed, while all suppressed parts are given to the winner cluster to preserve the probabilistic constraint: $\mu_{wk} = 1 - \alpha_{fcm}\sum_{i\neq w}\mu_{ik}$. Fig. 1 shows the effect of suppression caused on the distance of the winner data point from the cluster centre. While the distance of the non-winner from the cluster centre remains the same, the distance of the winner data point is shortened.

In Fig. 1, v_i are the data points around a particular cluster centre x_k and d_{ik} is the individual Euclidean distance between the data point and cluster centre. Since v_4 was the point closest to the cluster centre x_k as d_{4k} was the smallest, it was subjected to suppression, i.e, the Euclidean distance between v_4 and x_k was further shortened to δ_{4k} and the distance of other data points will remain same. However FCM and SFCM suffer from the noise sensitivity problem due to presence of membership constraint.

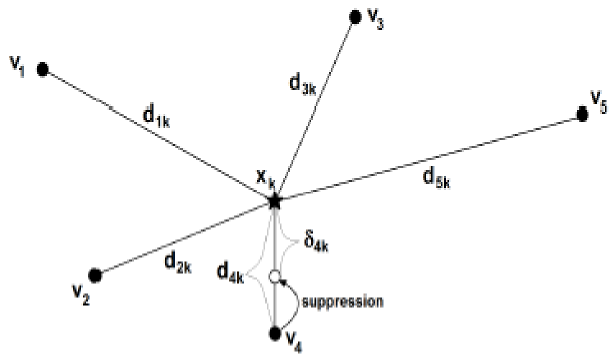


Figure 1: Effect on distance caused by suppression

To overcome the problem associated with the SFCM, Yu et. al. [15] introduced the suppression mechanism with the possibilistic approach of the clustering as Suppressed possibilistic c-means algorithm (SPCM). Here suppression constraints are applied on the typicality matrix of the PCM rather than the membership matrix of the FCM algorithm. This approach proved to perform better as compared to SFCM but suffers from the problem of coincident cluster.

3. Proposed Supressed Possibilistic Fuzzy c-means (SPFCM)

In this paper, we apply Fan et. al.[14] and Yu et. al. [15] suppression concept to the PFCM clustering algorithm and build a new suppressed possibilistic fuzzy c-means (SPFCM) clustering algorithm. Pal et. al. [16] proposed PFCM algorithm to obtain a stronger candidate for fuzzy clustering. This method combines the Fuzzy and Possibilistic approaches, resulting in two types of memberships: (1) a Fuzzy membership μ_{ik} that measures the relative degree of sharing of a point among clusters and (2) a possibilistic t_{ik} membership that measures the absolute degree of typicality of a point in a given cluster. PFCM algorithm is less sensitive to outliers and at the same time can avoid overlapping clusters. The objective function of the PFCM algorithm is as follows:

$$J_{PFCM} = \sum_{i=1}^{c}\sum_{k=1}^{n}\left(a\mu_{ik}^{m} + bt_{ik}^{\eta}\right)d\left(x_k, v_i\right)^2 + \sum_{i=1}^{c}\gamma_i\sum_{k=1}^{n}\left(1-t_{ik}\right)^{\eta} \tag{8}$$

Subject to membership constraint as given in Eq. 2 and $0 \le \mu_{ik}, t_{ik} \le 1$.Here the constants $a > 0, b > 0$ are used to determine the relative significance of the fuzzy and the possibilistic membership in the objective function. The constants $m > 1, \eta > 1$ represents the degree of the fuzziness. The minimization of the objective function defines membership, typicality and cluster centres as follows:

$$\mu_{ik} = \frac{1}{\sum_{i=1}^{c}\left(\frac{d\left(x_k, v_i\right)}{d\left(x_k, v_j\right)}\right)^{\frac{2}{m-1}}} 1 \le i \le c, 1 \le k \le n \tag{9}$$

$$t_{ik} = \frac{1}{1 + \left(\left\|x_k - v_i\right\|^2 b/\gamma_i\right)^{1/\eta-1}} 1 \le i \le c, 1 \le k \le n \tag{10}$$

$$v_i = \frac{\sum_{k=1}^{n}\left(a\mu_{ik}^{m} + bt_{ik}^{\eta}\right)x_k}{\sum_{k=1}^{n}\left(a\mu_{ik}^{m} + bt_{ik}^{\eta}\right)} \tag{11}$$

Further, the suppressed mechanism is introduced with the PFCM approach. The proposed SPFCM algorithm competes across clusters for typicality degrees and membership degrees and suppresses subordinate factors during the update phase.

The typicality metrics is updated as:

$$t_{wk} = \begin{Bmatrix} \alpha_{pfcm} t_{ik}, i \ne w \\ 1 - \alpha_{pfcm}\sum_{i \ne w} t_{ik}, i = w \end{Bmatrix} \tag{12}$$

The typicality t_{wk} with the highest value of the given data point x_k is selected among all the c clusters. The cluster w is considered as the closest cluster to the point x_k and declared as the winner. The typicality t_{wk} is referred to as winner membership, while the other typicalities t_{ik}, where $i \ne w$ are referred as non-winning typicalities as shown in Eq. 12.

The membership matrix is updated as

$$\mu_{wk} = \begin{Bmatrix} \alpha_{pfcm}\mu_{ik}, i \ne w \\ 1 - \alpha_{pfcm}\sum_{i \ne w}\mu_{ik}, i = w \end{Bmatrix} \tag{13}$$

The membership μ_{wk} with the highest value of the given data point x_k is selected among all c clusters is selected. The cluster w is considered as the closest cluster to the point x_k and declared as the winner. The membership μ_{wk} is referred to as winner membership, while the other memberships μ_{ik}, where $i \neq w$ are referred as non-winning memberships as shown in Eq. 13. $\alpha_{pfcm}, 0 \leq \alpha \leq$ is the suppression control parameter that controls the learning strength. When $\alpha_{pfcm} = 1$, then there is no suppression and the algorithm will behave like PFCM. The basic steps of the SPFCM algorithm are given below:

Algorithm: SPFCM

Input: Dataset X, Initial Parameters

Output: Membership matrix μ_{wk}, Optimized Cluster centers v_i

Algorithm:

S1: Initialize the number of clusters c, the partition matrix, such that U (0), the typicality matrix T (0), the termination tolerance ε > 0 and the user-defined constants.

S2: Calculate the cluster prototypes using Eq. (11)

S3: Update the partition matrix by using Eq. (10)

S4: Update the typicality matrix using Eq. (9)

S5: Introduce suppression in typicality matrix using Eq. (12)

S6: Introduce suppression in partition matrix using Eq. (13)

S7: Repeat the steps from S2 until the improvement of the objective function between two consecutive iterations is less than the termination tolerance ϵ.

4. Experimental results and discussion

In this section, comprehensive experiments are conducted to test the performance of proposed SPFCM. For the experiments, six datasets including two synthetic data sets and four real data sets from the UCI repository are used. Firstly, the clustering results of SPFCM are compared with several classical and state-of-the-art fuzzy clustering methods. Finally, a discussion part is given to summary the superiority, stability, and reliability of proposed method. The following typical parameters have been considered: m = 2, ϵ = 0.00001, maximum iterations= 100.

4.1 Artificial Datasets

Example 1: Dunn Dataset Algorithms: FCM, PCM, SFCM, SPCM, PFCM, SPFCM

Number of clusters: 2 clusters (2-dimensional data, 130 data points)

True cluster center $V_{true} = v^1 = [16,0] \quad v^2 = [5,0]$

The first simulation experiment involves the DUNN dataset which consists of one small and one big cluster of square shape. Fig. 2 displays the original DUNN data set and Table 1 lists the centroid locations of different clustering methods employed along with the E^* values that refer to the square mean error which is calculated as the square of difference of V_{true} and V^* where, V_{true} is the actual cluster centres for the DUNN dataset and V^* is the calculated centre using the above mentioned algorithms, ℓ represents the iteration number.

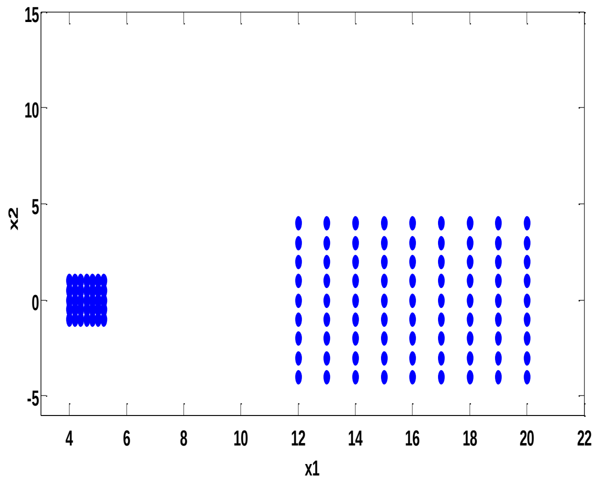


Figure 2. DUNN Dataset

As can be seen from table 1, the proposed algorithm produces the lowest mean error E^* taking least number of iterations (ℓ), thus showing its superiority over other algorithms.

Table 1. Terminal centroids produced by FCM, PCM, PFCM, SFCM, SPCM, SPFCM on DUNN data set

Algorithms:	FCM	PCM	SFCM	SPCM	PFCM	SPFCM
DUNN	(m=2)	(η = 2)	(α=0.9)	(α=0.3)	(a=1,b=1 η=2)	(a=1,b=1 η=2,α=0.9)
V*	[16.34 0 4.98 0]	[15.67 0 4.65 0]	[15.97 0 4.60 0]	[15.97 0 4.60 0]	[16.27 0 4.79 0]	[16.16 0 4.90 0]
$E^* = \|\|V_{true} - V^*\|\|^2$	0.116	0.231	0.1609	0.1609	0.086	**0.0356**
ℓ (No of Iterations)	12	39	11	11	19	**11**

Example 2: GAUSSIAN Dataset

Algorithms: FCM, PCM, SFCM, SPCM, PFCM, SPFCM

Number of clusters: 2 clusters (2-dimensional data)

The second simulation experiment involves the GAUSSIAN dataset where a Gaussian random number generator is used to create a dataset containing two clusters with outliers. The effect of noise/outliers can be seen in the Fig. 3 on different clustering algorithms. The outlier points shifts the partition space towards the outliers in case of FCM, SFCM and PFCM. The results shown by PCM are good when given good initial values from the FCM results with one misclassified data point. The results shown by our proposed method SPFCM is better than FCM, SFCM, PFCM and SPCM as it properly partitions the feature space and yields good clusters just like PCM.

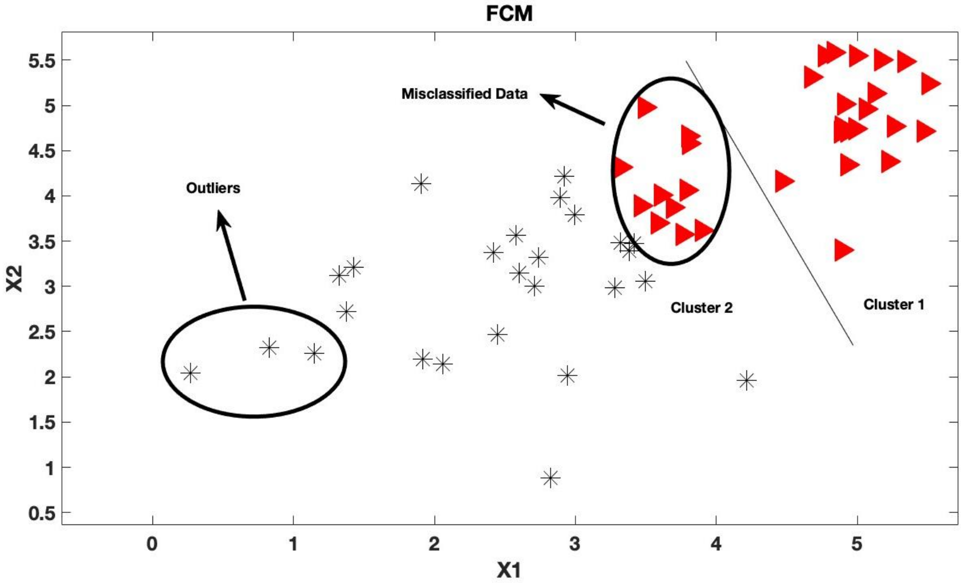
In addition to the producing good quality clusters, we have also compared the convergence rate (No of Iterations=ℓ) of different clustering algorithms. It can be seen from Table 2, though FCM takes the least number of iterations, but the misclassified data points are the highest. Although the misclassifications are the same in PCM and our proposed method, but our proposed algorithm converges in 20 iterations as compared to 36 iterations in PCM. Also as compared to PFCM, our proposed method produces good partitions of two clusters using lesser number of iterations.

.

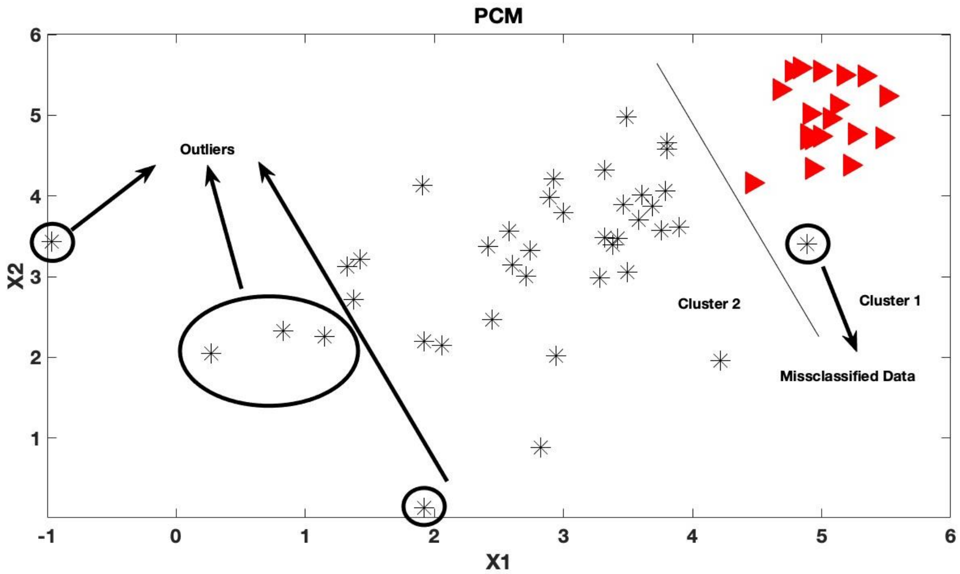
Table 2. Convergence of different clustering algorithms on GAUSSIAN data set

Algorithms:	FCM	PCM	SFCM	SPCM	PFCM	SPFCM
GAUSSIAN	(m=2)	(η = 2)	(α=0.9)	(α=0.3)	(a=1,b=1 η=2)	(a=1,b=1 η=2,α=0.9)
ℓ (No of Iterations)	17	36	19	39	30	**20**
Misclassified Data	11	**01**	11	02	04	**01**

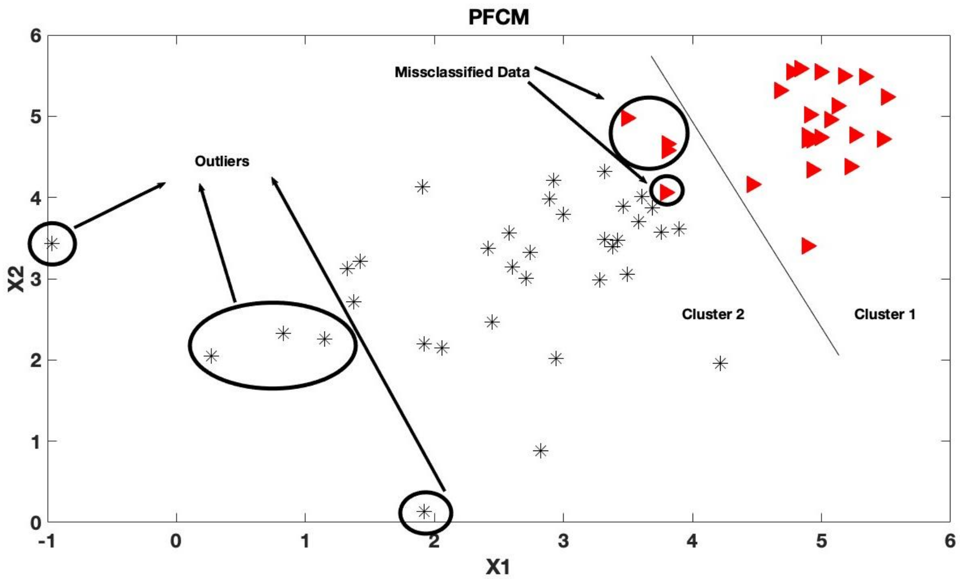
(a)


FCM
Misclassified Data
Outliers
Cluster 2
Cluster 1
X1
X2

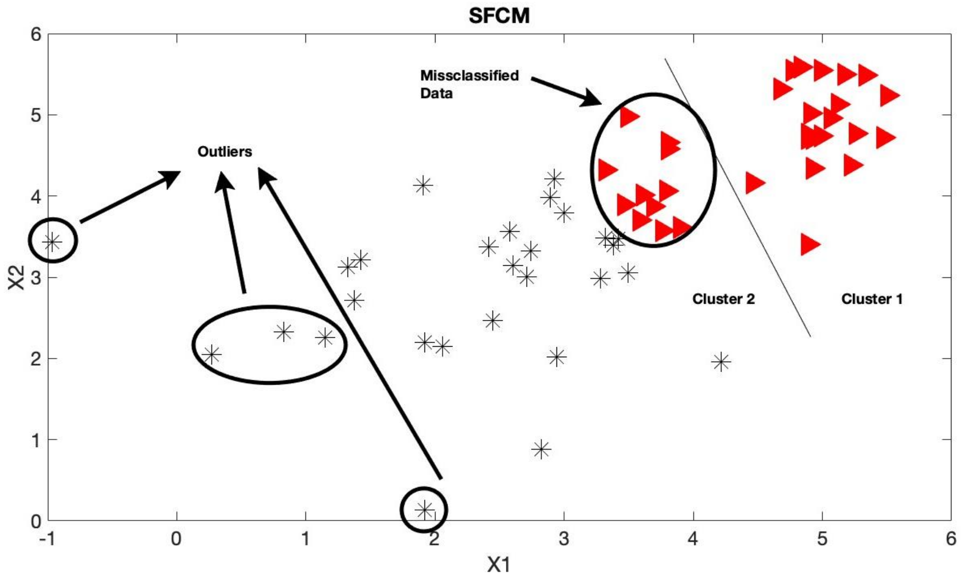
(b)


PCM
Outliers
Cluster 2
Cluster 1
Missclassified Data
X1
X2

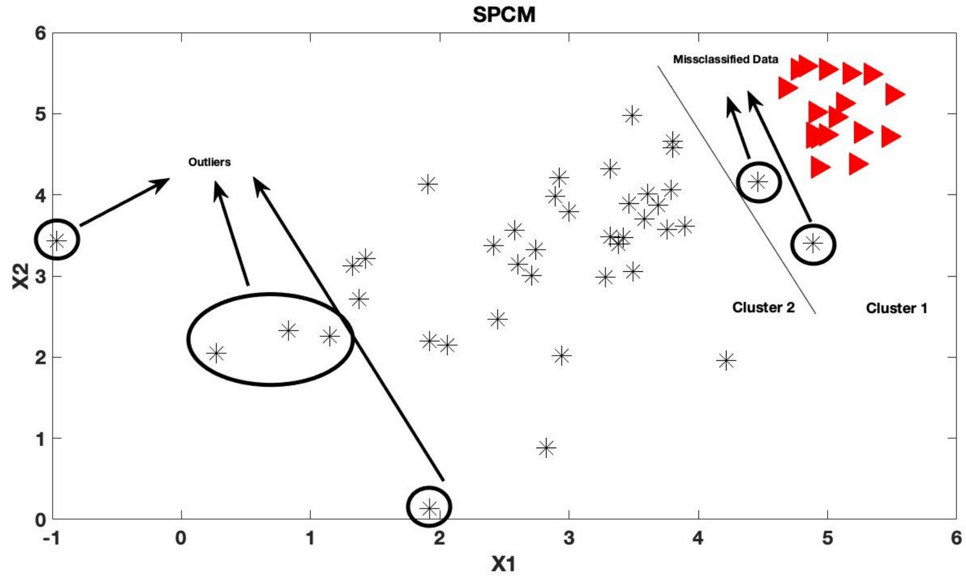
(c)


PFCM
Missclassified Data
Outliers
Cluster 2
Cluster 1
X1
X2

(d)



(e)



(f)

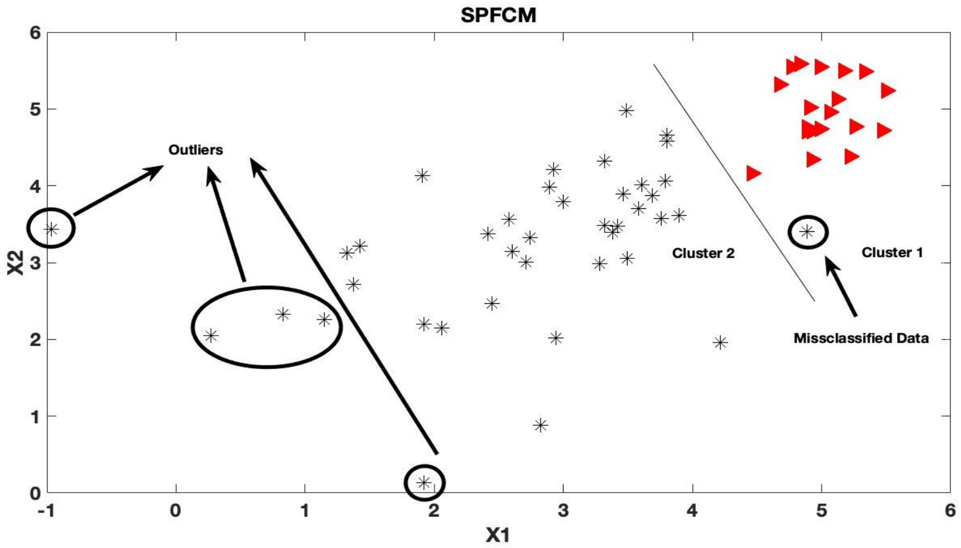


Figure 3. Clustering results produced by (a) FCM (b) PCM (c) PFCM (d) SFCM (e) SPCM (f) SPFCM o GAUSSIAN data

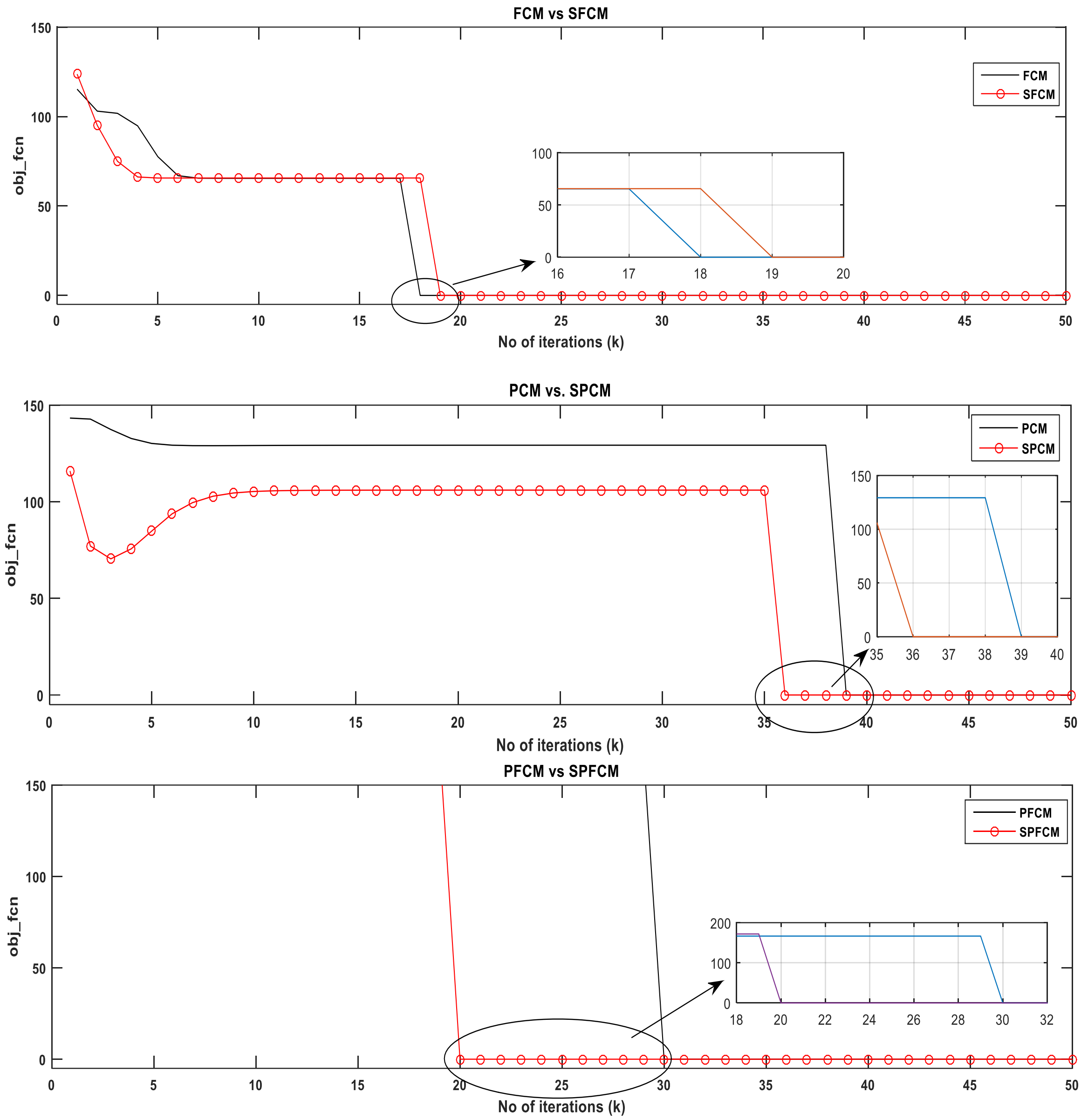


Figure 4. Comparison of the Convergence rate on the GAUSSIAN Dataset (a) FCM vs. SFCM (b) PCM vs SPCM (c) PFCM vs. SPFCM

Fig. 4 represents the graphs of convergence of different suppression based clustering algorithms as compared to their conventional counterparts

4.2 Real Datasets

Experiments were performed on a number of real datasets taken from UCI machine learning repository [20] including: wine, seed, glass and abalone. The datasets are selected with the variations in the features, shape of the clusters, and size of the dataset. Huang's accuracy metric [21] was used to evaluate the clustering results:

$$r = \frac{\sum_{i=1}^{k} n_i}{n} \tag{14}$$

where n_i represents the data occurring in the clusters, i^{th} cluster and the corresponding true cluster and n is the total number of data points in the data set. The greater the value of r indicates superior clustering results and yields $r=1$ perfect clustering.

Example 3: Wine dataset

Algorithms: FCM, PCM, PFCM, SFCM, SPCM, SPFCM
Number of clusters: 3 clusters (13-dimensional data)
Size of clusters: 59, 71, 48

The wine dataset consists of three types of wine found in Italy and their chemical analysis. The three categories are represented in 178 samples and consist of 13 attributes obtained as a result of the chemical analysis. Table 3 shows the accuracy percentage for the SPFCM above 72% with error percentage around 27% meanwhile all the other algorithms have accuracy percentage below the proposed method. The misclassification number of the data points is 49 which are the lowest among all the algorithms.

Example 4: Seed dataset

Algorithms: FCM, PCM, PFCM, SFCM, SPCM, SPFCM
Number of clusters: 3 clusters (7-dimensional data)
Size of clusters: 70, 70, 70

The seed dataset has measurements of seven geometric parameters of wheat kernels. These kernels are divided into three categories: Kama, Rosa and Canadian. It is observed in Table 3 that the accuracy percentage for the proposed SPFCM algorithm is the highest at 90.95% clearly showing that it surpasses other algorithm in performance. The error percentage is also recorded close to 9% being the lowest among all. The misclassification number is 19 representing the efficiency of the proposed SPFCM.

Example 5: Glass dataset

Algorithms: FCM, PCM, PFCM, SFCM, SPCM, SPFCM
Number of clusters: 7 clusters (10-dimensional data)
Size of clusters: 70, 17, 76, 0, 13, 9, 29

Glass dataset includes 214 instances, describing 7 categories of glass using 10 features. Table 3 shows the percentage accuracy of the proposed SPFCM algorithm for glass dataset to be over 88%, highest among the entire clustering algorithm used, both the error percentage and misclassifications are also minimal as compared to the other algorithms.

Example 6: Abalone dataset

Algorithms: FCM, PCM, PFCM, SFCM, SPCM, SPFCM
Number of clusters: 3 clusters (8-dimensional data)
Size of clusters: 1342, 1528, 1307

The abalone dataset is widely used to test the performance of the clustering. Abalone dataset consists of physical measurements of abalones, which are large, edible sea snails. The dataset is divided into three categories, 8 features and a total of 4,177 samples. Table 3 illustrates proposed SPFCM performs well on large dataset of abalone with the maximum accuracy and minimum misclassification error as compared with different clustering algorithms.

Table 3. Comparative analysis of SPFCM algorithm with other fuzzy clustering methods in terms of accuracy, misclassifications, and error.

Wine Dataset				
Clustering	**Parameters**	**% Accuracy**	**% Error**	**Misclassification**
FCM	m=2	68.53%	31.46%	56
PCM	m=2	41.57%	58.42%	104
PFCM	m=2, a=2,b=2	70.78%	29.21%	52
SFCM	m=2,α=0.2	70.22%	29.77%	53

SPCM	m=2,α=0.8	62.92%	37.07%	66
S-PFCM	m=2,a=1,b=1,α=0.6	**72.47%**	**27.528090%**	**49**

Seed Dataset

Clustering	**Parameters**	**% Accuracy**	**% Error**	**Misclassification**
FCM	m=2	89.52%	10.47%	22
PCM	m=2	64.76%	35.23%	74
PFCM	m=2, a=2,b=2	89.52%	10.47%	22
SFCM	m=2,α=0.3	89.52%	10.47%	22
SPCM	m=2,α=0.22	89.04%	10.95%	23
S-PFCM	m=2,a=1,b=1,α=0.3	**90.952381%**	**9.047619%**	**19**

Glass Dataset

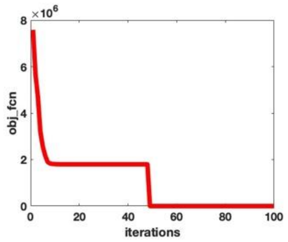
Clustering	**Parameters**	**% Accuracy**	**% Error**	**Misclassification**
FCM	m=2	83.17%	16.82%	36
PCM	m=2	87.38%	12.61%	27
PFCM	m=15,a=0.2,b=2	83.64%	16.35%	35
SFCM	m=2,α=0.3	83.64%	16.35%	35
SPCM	m=2,α=0.22	87.38%	12.61%	27
S-PFCM	m=15,a=0.2,b=1.7,α =0.22	**88.31%**	**11.68%**	**25**

Abalone Dataset

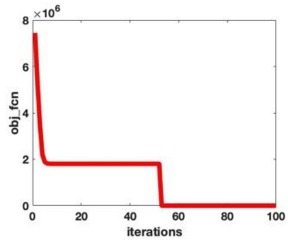
Clustering	**Parameters**	**% Accuracy**	**% Error**	**Misclassification**
FCM	m=2	51.66%	48.33%	2019
PCM	m=2	36.58%	63.41%	2649
PFCM	m=2, a=2,b=2	52.04%	47.95%	2003
SFCM	m=2,α=0.9	51.49%	48.50%	2026
SPCM	m=2,α=0.9	52.07%	47.92%	2002
S-PFCM	m=2,a=1,b=1, α=0.2	**52.74%**	**47.25%**	**1974**

4.3 Evaluation of convergence rate

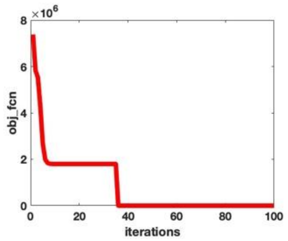
The convergence rate of the proposed SPFCM is compared with different clustering algorithms on the real datasets to determine the efficiency of the proposed technique. The objective values for various clustering algorithms are plotted on each set of iterations to show the number of iteration taken by an algorithm for attaining the convergence. In the case of proposed algorithm SPFCM,



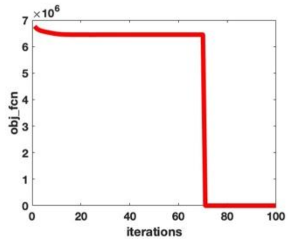
(a) FCM



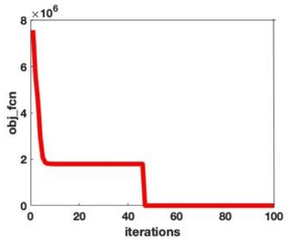
(b)SFCM



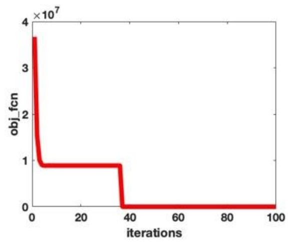
(c) PCM



(d) SPCM



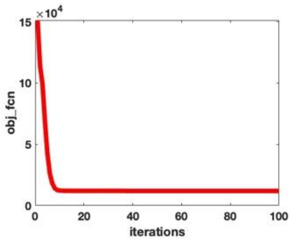
(e) PFCM



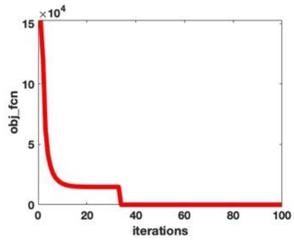
(f) SPFCM

Figure 5. Wine Dataset

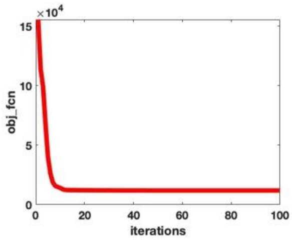
the objective value decreases monotonically and converges in lesser number of iterations. Fig. 5 shows convergence of the different clustering algorithms with the number of iterations on the wine dataset. It can be observed from Fig. 5 that SPCM and SFCM converge slowly as compared to their conventional counter-parts whereas only our proposed SPFCM converges faster than its conventional counter-part PFCM. Fig. 6 shows the convergence of the glass dataset with the number of iterations. As shown in Fig. 6(f), the objective function for SPFCM not only converges faster than other algorithms but the value of the objective function is also much less compared to others. This proves the efficiency of the proposed SPFCM over the other clustering algorithms in attaining the convergence with minimum number of iterations.



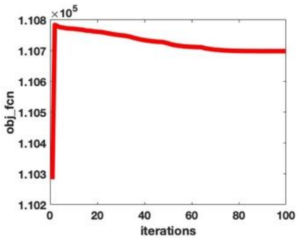
(a)FCM



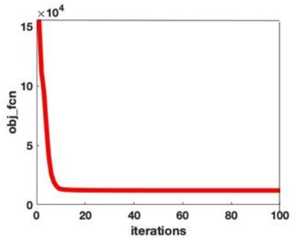
(b)SFCM



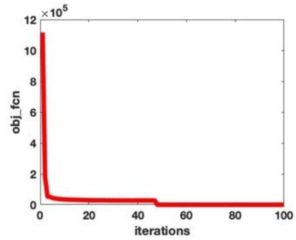
(c)PCM



(d) SPCM



(e)PFCM

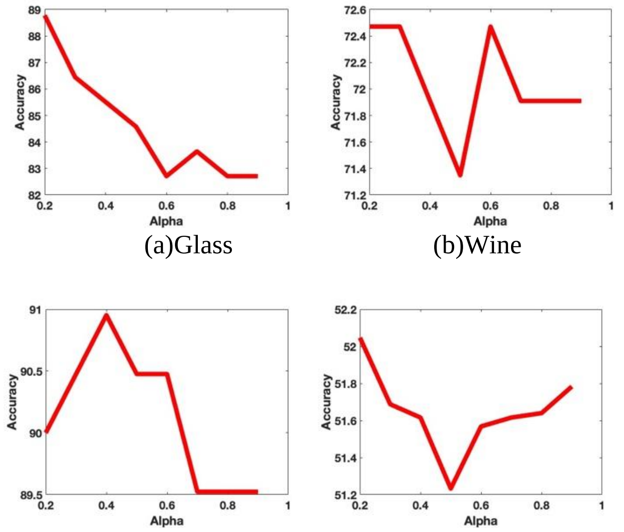


(f) SPFCM

Figure 6. Glass Dataset

4.4. Effect of Suppressed Parameter (α) on SPFCM Clustering

This section evaluates the performance of the proposed SPFCM algorithm on the different values of the suppressed parameter α. Fig. 7 shows the variation in the accuracy achieved with the change in the α value on the benchmarking real datasets. The value of α ranges between 0 and 1, and it handles the extent of suppression performed. The y-axis represents the percentage of accuracy achieved and the x- axis represents the value of α. In case of the glass dataset, the average accuracy obtained is decreasing as the value of α is increased from 0.2 to 0.9, Fig. 7(b) shows that in case of wine dataset high accuracy is obtained initially however, the value begins to decreases from the change in the value of α from 0.3 to 0.5 and then again increases and finally settles, Fig. 7(c) represents increase from lower accuracy at low values of α to maximum value around 0.4, finally Fig. 7(d) expresses the variations in the accuracy for abalone dataset, with a V-shaped plot, showing maximum accuracy at 0.2 mark.



(a)Glass (b)Wine

(c)Seed (d)Abalone

Figure 7. Effect of α value on the accuracy of SPFCM

5. Conclusions

In this paper, we proposed a suppressed possibilistic fuzzy c-means (SPFCM) clustering approach that incorporates suppression mechanism with possibilistic and fuzzy approach of PFCM. The non-winner typicalities and memberships are suppressed to improve the effect of winner membership and typicalities to develop the clustering approach with better performance. Experiments are conducted on several synthetic and real datasets to prove the effectiveness of the proposed algorithm. In terms of clustering accuracy, error and misclassifications, the proposed SPFCM shows appreciable improvements as compared to the existing clustering algorithm. The proposed SPFCM clustering proved to be more comprehensive and have important theoretical and application values for the research over the PFCM and the other clustering approaches.

References

[1] Zhang, Q., Yang, L.T., Yan, Z., Chen, Z., and Li., P. (2018) An efficient deep learning model to predict cloud workload for industry informatics. IEEE Transactions on Industrial Informatics 14(7): 3170-3178.

[2] Bezdek, J.C., Keller, J., Krishnapuram, R., and Pal, N. R. (1999) Fuzzy Models and Algorithms for Pattern Recognition and Image Processing. Springer.

[3] Sarki, R., Ahmed, K., Wang, H., Zhang, Y., and Wang, K. (2022) Convolutional neural network for multi-class classification of diabetic eye disease. EAI Endorsed Transactions on Scalable Information Systems 9(4):1-11.

[4] Fang, C., and Liu, H. (2021) Research and Application of Improved Clustering Algorithm in Retail Customer Classification. Symmetry 13(10).

[5] Mittal, H., Pandey, A.C., Saraswat, M., Kumar, S., Pal, R., and Modwel. G. (2021) A comprehensive survey of image segmentation: clustering methods, performance parameters, and benchmark datasets. Multimedia Tools and Applications.

[6] Zheng, H., He, J., Huang, G., Zhang, Y., and Wang, H. (2019) Dynamic optimisation based fuzzy association rule mining method. International Journal of Machine Learning and Cybernetics 10:2187–2198.

[7] Singh, R., Zhang, Y., Wang, H., Miao, Y., and Ahmed, K. (2020) Investigation of Social Behaviour Patterns using Location-Based Data – A Melbourne Case Study. EAI Endorsed Transactions on Scalable Information Systems 8(3): 1-18.

[8] Wang, H., Chen, Y., and Dong, S. (2017). Research on efficient-efficient routing protocol for WSNs based on improved artificial bee colony algorithm. IET Wireless Sensor Systems, 7(1), 15-20.

[9] Dong, S., Zhou, D., Ding, W., and Gong, J. (2014) Flow Cluster Algorithm Based on Improved K-means Method. IETE Journal of Research 59(4): 326-333.

[10] Bora, D.J., and Gupta., A.K. (2014) Clustering approach towards image segmentation: An Analytical Study. International Journal of Research in Computer Applications and Robotics 2(7): 115-124.

[11] Dunn, J.C. (1973) A fuzzy relative of the ISODATA process and its use in detecting compact well-separated cluster. Journal of Cybernetics 32–57.

[12] Bezdek, J.C. (1981) Pattern Recognition with Fuzzy Objective Function Algorithms, Plenum Press, New York.

[13] Krishnapuram, R., and Keller, J.M. (1993) A possibilistic approach to clustering IEEE Transactions on Fuzzy Systems, 1(2): 98–110.

[14] Fan, J.L., Zhen, W-Z., and Xie, W-X. (2003) Suppressed fuzzy c-means clustering algorithm. Pattern Recognition Letters 24:1607–1612.

[15] Yu, H., Fan, J., and Lan, R. (2019) Suppressed possibilistic c-means clustering algorithm. Applied Soft Computing, 80: 845-872.

[16] Pal, N.R., Pal, K., and Bezdek, J.C. (1997) A mixed c-means clustering model. in: Proceedings of the IEEE International Conference on Fuzzy Systems, Spain 11–21.

[17] Pal, N.R., Pal, K., Keller, J., and Bezdek, J.C. (2005) A possibilistic Fuzzy c-means clustering algorithm. IEEE Transactions on Fuzzy Systems, 13(4): 517–530.

[18] Ozdemir, O. and Kaya, A. (2018) Effect of parameter selection on fuzzy clustering. Journal of Applied Sciences of Mehmet Akif Ersoy University 2(1): 22-33.

[19] Ozdemir, O., and Kaya, A. (2018) K-medoids and fuzzy c-means algorithms for clustering CO2 emissions of Turkey and other OECD countries. Applied Ecology and Environmental Research, 16(3): 2513-2526.
[20] Blake, C., Keough, E., and Merz, C.J. UCI Repository of Machine Learning Database. [online] http://www.ics.uci.edu/mlearn/Mlrepository.html.
[21] Arora, J., and Tushir, M. (2017) A new kernel-based possibilistic intuitionistic fuzzy c-means clustering. International Journal of Artificial Intelligence and Soft Computing 6(4): 306-325.