Int. J. Fuzzy Syst. (2024) 26(3):943–977
https://doi.org/10.1007/s40815-023-01644-5

Weighted Intuitionistic Fuzzy *C*-Means Clustering Algorithms

Meenakshi Kaushal[1] · Q. M. Danish Lohani[1] · Oscar Castillo[2]

Received: 7 April 2023 / Revised: 29 September 2023 / Accepted: 6 November 2023 / Published online: 5 February 2024

Abstract Atanassov intuitionistic fuzzy set (AIFS)-based *C*-means algorithms are successful in clustering uncertain or vague real-world datasets. The AIFS-based clustering algorithms are classified into adaptive class and non-adaptive class. An algorithm from the adaptive class computes its feature weight distribution with the help of the given dataset. On the other side, the algorithm belonging to the non-adaptive class mostly computes the feature weight distribution by employing an equally likely approach. The guarantee to reach up to the mark clustering performance is missing within this approach. Simultaneously, the performance gets deteriorated if the datasets showcase noises/irrelevant features. The irrelevant features in the datasets add to the computational cost. So, a feature reduction-equipped clustering algorithm called uni-weighted intuitionistic fuzzy *C*-means (uW-IFCM) is introduced in the paper. Moreover, the probabilistic weights-based adaptive clustering algorithm, namely bi-weighted probabilistic intuitionistic fuzzy *C*-means (bW-PIFCM) is proposed under the AIFS environment. The parametric analysis for uW-IFCM is provided to comprehend and compare its performance with bW-PIFCM, PIFCM, IFCM, and FCM algorithms. Here, an intuitionistic data fuzzification technique transforms the real-valued dataset into AIFS dataset, therefore bW-PIFCM and uW-IFCM algorithms cluster the real-valued datasets. The research proposal of Yang and Nataliani in [IEEE Transactions on Fuzzy Systems, 26(2), 817–835] motivates us to introduce a feature reduction-equipped uW-IFCM algorithm. We have considered synthetic datasets and some UCI machine learning datasets for the experimental study of uW-IFCM and bW-PIFCM. The efficacy and the precision of proposed algorithms are tested in terms of some popular benchmark indexes as well.

Keywords Atanassov intuitionisitic fuzzy set (AIFS) · Intuitionistic fuzzy *C*-means · Feature reduction · PIFCM

1 Introduction

In the paper, we explore multi-dimensional/multi-featured datasets with the help of clustering. Some of the well-known procedures of clustering are subspace clustering (see [1, 2]), density-based clustering [3], fuzzy clustering [4], and hierarchical clustering [5]. Fuzzy *C*-means (FCM) [6] is an example of a fuzzy partitional clustering technique, and its functioning has been widely observed over multiple types of real-world multi-dimensional datasets. The variants of FCM obtained in the purview of other fuzzy sets are known as intuitionistic fuzzy *C*-means (IFCM) [7], interval type-2 fuzzy *C*-means [8, 9], rough fuzzy *C*-means [10, 11], and hesitant fuzzy *k*-means [12]. The FCM algorithm uses Euclidean distance measure. Researchers exploited measures like Mahalanobis distance in place of the Euclidean distance measure to obtain clusters that are not spherical (see [13]). Further such

O. Castillo have contributed equally to this work.

✉ Oscar Castillo
ocastillo@tectijuana.mx

Meenakshi Kaushal
meenakshikaushal247@gmail.com

Q. M. Danish Lohani
danishlohani@cs.sau.ac.in

1 Department of Mathematics, South Asian University, Chanakyapuri, New Delhi 110021, India

2 Division of Graduate Studies and Research, Tijuana Institute of Technology, Mexico, Tijuana, Mexico

Table 1 Mathematical symbols

Symbols	Details
α	Tuning parameter in Yager'complement function
β	Tuning parameter in membership function
η	Weighing exponent
D	Number of features/dimensions
m	Fuzziness index
C	Number of clusters
P	Number of data points
J_m	Criterion function
u_{li}	Membership degree of ith data point in lth cluster
x_i	ith data point
v_l	Centroid of lth cluster
U	Membership matrix
V	Centroid matrix
W	Weight matrix
w	Fixed weight collection

distance-based variants of FCM are found less sensitive towards noises and outliers (see [14]). The weighted fuzzy C-mean clustering algorithms are successful in satisfactory clustering of the multi-dimensional/multi-featured datasets (see [15–18]). The weighted FCM assigns an appropriate role to each feature in a multi-featured dataset during its clustering. It involves a feature weight learning technique that decides the role of each feature (see [19–23]). A feature weight learning technique is proposed in [20] based on the gradient descent procedure. For an efficient clustering of the multi-dimensional datasets, an entropy-based weight learning technique is proposed by the [21]. Some of the recent studies on feature learning are [24, 25, 17, 26, 27, 28, 29], and [30].

The AIFS proposed by [31] is an effective tool to model the ambiguity and uncertainty present within real-world phenomena. The ambiguity and uncertainty present in the multi-dimensional datasets motivate us to cluster them under the AIFS environment using the hesitancy component of the AIFS. An AIFS-based clustering algorithm efficiently solves the problems of various fields discussed by [32–38]. [34] proposes a clustering technique for image representations, whereas medical image segmentation is done by an entropy-based technique (see [32]). [33] extends the clustering technique of [32] to cope with the issues of noise/outliers by selecting neighborhood pixels from different regions of the magnetic resonance imaging for its tuning. The natural extension of the FCM under the AIFS environment is the IFCM algorithm [7], which is further studied by [38]. It is improvised for the noisy medical image segmentation while using the rough set theory (see [39]).

During the clustering of a multi-dimensional/featured dataset, FCM and IFCM assign equal importance to each dimension/feature. Moreover, the derivation of FCM/IFCM or their variants involves calculus-based optimization. It is observed that many of these variants are using weighted Euclidean distance measure. Here, we categorize these variants on the basis of an equally likely approach as follows: (1) In the IFCM variants (see [32, 33, 34, 40, 36, 37]), an equally likely approach is used, that is, all the components of Euclidean distance are assigned same weight. (2) There are some variants (see [41, 39]) in which unequal weights are allocated to the components of Euclidean distance in contrast to the equally likely approach. Moreover, the variants of the FCM employing an equally likely approach traverse their distance along a circular path, whereas the IFCM counterparts compute distances along spherical paths. A single weighing exponent is sufficient to control those distance measures which track along a circular or spherical/hyper-spherical path. Now, if variants of FCM/IFCM do not employ an equally likely approach, then elliptical/ellipsoidal path-based distances are measured. To monitor an ellipsoidal path, three independent weighing exponents are exploited, here, the same path is monitored using a weighing exponent containing two variables. We show that the improvisation of the clustering algorithm depends on the computation of an accurate feature weight distribution. The clustering algorithms discussed in the paper are either uni-parametric, bi-parametric, tri-parametric, or quad-parametric (see Table 2). Let us discuss the major contributions of the paper:

- An equally likely approach often fails in the dealing of a multi-dimensional/featured dataset because all features of the dataset are not equally relevant. To address this problem, we have proposed two clustering techniques: (1) a singly weighted data-driven algorithm called uni-weighted intuitionistic fuzzy C-means (uW-IFCM), (2) two variables-based weight triplets introduce bi-weighted probabilistic intuitionistic fuzzy C-means algorithm (bW-PIFCM) (see Sect. 2.2 and 2.3).
- The initialization of uW-IFCM and bW-PIFCM algorithms requires a transformation of the real-valued multi-dimensional/featured dataset to an AIFS dataset. For this, we have utilized two parameters (α and β)-based novel data fuzzification technique (see Sect. 3).
- In uW-IFCM algorithm, the weighing exponent η and fuzzy factor m are used as parameters. We further tune the feature weights using the exponent η, which gives better feature weight distribution for uW-IFCM in comparison to bW-PIFCM. A study about the optimal choices of η and m at which uW-IFCM delivers its best clustering has been carried out in Sect. 4. Here, we

Table 2 Description of C-means algorithms

	Clustering Algorithm	Path of distance measure	Nature of algorithm	Parameters used
1	IFCM	Spherical path	Equally likely approach	m, α, β, U or (V)
2	weighted-IFCM	Ellipsoidal path	Random weight selection approach	m, β, U or (V)
3	wIFCM	Spherical path	Data-driven weight selection approach	m, β, U or (V)
4	FCM	Circular path	Equally likely approach	m, U or (V)
5	wFCM	Circular path	Data-driven weight selection approach	m, U or (V)
6	PIFCM	Ellipsoidal path	Adaptive weight selection approach	m, α, β, U or (V)
7	Proposed uW-IFCM	Spherical path	Data-driven non-adaptive weight selection approach	$m, \alpha, \beta, \eta, V, W$
8	Proposed bW-PIFCM	Ellipsoidal path	Data-driven adaptive weight selection approach	m, α, β, V

perform both experimental analysis and convergence analysis of the uW-IFCM and bW-PIFCM algorithms.

- The uW-IFCM detects irrelevant features in a multi-dimensional/featured dataset during its clustering. We have shown that uW-IFCM is a mechanized feature reduction technique (see Sect. 5.1), whereas an automatized feature reduction technique called FRT-equipped uW-IFCM algorithm is also proposed for nullification of irrelevant features (see Sect. 5.2).

The remaining paper is organized into five sections. In Section 2, we provide a mathematical solution for a fuzzy clustering problem. Section 3 discusses the procedural detailing of proposed uW-IFCM and bW-PIFCM algorithms. In Sect. 4, an experimental analysis of synthetic and some UCI machine learning datasets is carried out. We provide two uW-IFCM-based feature reduction techniques in Sect. 5. Finally, the conclusion is stated in Sect. 6.

2 Some Important Mathematical Results for C-Means Clustering Algorithms

We have divided this section into three subsections.

2.1 Description of Clustering Problem I

Intuitionistic fuzzy C-means (IFCM) is a well-known clustering algorithm that uses Euclidean distance measure to cluster the datasets.

Clustering problem I: Let $\{\hat{x}_1, \hat{x}_2, \ldots, \hat{x}_P\}$ be a set of multi-valued AIFS data items. Here, the dimension of each data item is D, that is, $\hat{x}_i = (\tilde{x}_{id})_{d=1}^{D}$, and $\tilde{x}_{id} = (\mu_{id}, \nu_{id}, \pi_{id})$. We have to group $\{\hat{x}_1, \hat{x}_2, \ldots, \hat{x}_P\}$ into C clusters. Let $V = \{\hat{v}_1, \hat{v}_2, \ldots, \hat{v}_C : \hat{v}_l = (\bar{\mu}_{ld}, \bar{\nu}_{ld}, \bar{\pi}_{ld})_{d=1}^{D}, 1 \le l \le C\}$ be a set of initial cluster centroids.

IFCM-based solution procedure: The criterion function J_m is defined as

$$J_m(U,V) = \sum_{l=1}^{C}\sum_{i=1}^{P} u_{li}^m D_1^2(\hat{x}_i, \hat{v}_l), \quad \text{where} , \tag{1}$$

$$D_1^2(\hat{x}_i, \hat{v}_l) = \frac{1}{2P}\sum_{d=1}^{D}\{(\mu_{id} - \bar{\mu}_{ld})^2 + (\nu_{id} - \bar{\nu}_{ld})^2 + (\pi_{id} - \bar{\pi}_{ld})^2\} \tag{2}$$

such that

$$\sum_{l=1}^{C} u_{il} = 1, 1 \le i \le P \tag{3a}$$

$$u_{il} \in [0,1] 1 \le i \le P, \ 1 \le l \le C \tag{3b}$$

$$\sum_{i=1}^{P} u_{il} > 0, 1 \le l \le C\,. \tag{3c}$$

The membership value u_{il} of ith data item in lth cluster and the centroid $(\mu_{\bar{v}_{ld}}, \nu_{\bar{v}_{ld}}, \pi_{\bar{v}_{ld}})_{d=1}^{D}$ are updated using Eqs. (4) and (5).

$$u_{il} = \frac{1}{\sum_{l=1}^{C}\left(\frac{D_1(\hat{x}_i, \hat{v}_l)}{D_1(\hat{x}_i, \hat{v}_l)}\right)^{\frac{2}{m-1}}} \tag{4}$$

$$\mu_{\bar{v}_{ld}} = \frac{\sum_{i=1}^{P} u_{il}^m \mu_{\tilde{x}_{id}}}{\sum_{i=1}^{P} u_{il}^m}, \nu_{\bar{v}_{ld}} = \frac{\sum_{i=1}^{P} u_{il}^m \nu_{\tilde{x}_{id}}}{\sum_{i=1}^{P} u_{il}^m}, \pi_{\bar{v}_{ld}} = \frac{\sum_{i=1}^{P} u_{il}^m \pi_{\tilde{x}_{id}}}{\sum_{i=1}^{P} u_{il}^m}. \tag{5}$$

In IFCM algorithm, all the features of a multi-dimensional/featured dataset are equally emphasized during clustering. The IFCM transforms into weighted-IFCM, if $D_1^2(\tilde{x}_i, \tilde{v}_l)$ is replaced with $D_w^2(\tilde{x}_i, \tilde{v}_l) = \sum_{d=1}^{D} w_d((\mu_{id} - \bar{\mu}_{ld})^2 + (\nu_{id} - \bar{\nu}_{ld})^2 + (\pi_{id} - \bar{\pi}_{ld})^2)$. In weighted-IFCM, we randomly select its weights (i.e., $w_d, 1 \le d \le D$) such that $\sum_{d=1}^{D} w_d = 1$.

2.2 A Proposal of Novel Probabilistic Intuitionistic Fuzzy *C*-Means Clustering Algorithm

Let us solve the Clustering problem I probabilistically. So, we propose bi-Weighted Probabilistic Intuitionistic Fuzzy *C*-Means algorithm (bW-PIFCM), and this variant of PIFCM uses two variable-based weight triplets. The proposed algorithm computes data-driven feature weights such that it results to an optimal feature weight distribution at the convergence. It means that the proposed algorithm is adaptive in nature. The weights portray the importance of features in a multi-dimensional/featured dataset. Here, a membership matrix $M = (u_{il})_{P\times C}$ of order $P \times C$ is obtained. The mathematical results which correspond to bW-PIFCM are given in the form of four theorems.

Theorem 2.1 *Let S be a collection that contains AIFS. If* $I_1 = [p', p'']$ *and* $I_2 = [q', q'']$ *be the weight intervals assigned to membership and non-membership components of an AIFS, which is an element of S. The weight interval I assigned to the hesitancy component of the AIFS is*

$$I = [1 - max(p',q',p'',q'') + max(p',q',p'',q'')\pi, 1 - min(p',q',p'',q'') + min(p',q',p'',q'')\pi]. \tag{6}$$

Proof Let $A = (\mu, v, \pi)$ be an AIFS belonging to S. As,

$$\min(p',q')(\mu+v) \le p'\mu + q'v \le \max(p',q')(\mu+v) \tag{7}$$

$$\min(p'',q'')(\mu+v) \le p''\mu + q''v \le \max(p'',q'')(\mu+v). \tag{8}$$

From (7) and (8), we have,

$$\min(p',q',p'',q'')(\mu+v) \le p'\mu+q'v, p'\mu+q'v \le \max(p',q',p'',q'')(\mu+v) \tag{9}$$

$$\min(p',q',p'',q'')(\mu+v) \le p''\mu+q''v, p''\mu+q''v \le \max(p',q',p'',q'')(\mu+v). \tag{10}$$

Therefore, from (9) and (10), we get,

$$I = [1-\max(p',q',p'',q'')+\max(p',q',p'',q'')\pi, 1-\min(p',q',p'',q'') + \min(p',q',p'',q'')\pi].$$

□

Theorem 2.2 *Let an AIFS, say* A_1 *belonging to S has been assigned weight intervals* $I_1 = [p', p'']$, $I_2 = [q', q'']$ *, and I to their membership, non-membership, and hesitancy components, respectively. Here,* $I_1 = [p', p'']$, $I_2 = [q', q'']$ *are independent of* A_1*, but the interval I depends on* A_1*. Let* $d : S \times S \to \mathbb{R}$ *be a mapping, such that,*

$$d(A_1,A_2) = \left[\frac{1}{2n}\sum_{i=1}^{n} p_{12}(\mu_{A_1}(x_i) - \mu_{A_2}(x_i))^2 + q_{12}(v_{A_1}(x_i) - v_{A_2}(x_i))^2 + r_{12}(\pi_{A_1}(x_i) - \pi_{A_2}(x_i))^2\right]^{1/2}. \tag{11}$$

The mean of the intervals $I_1, I_2, I_3 = I \cap I'$ gives the coefficients p_{12}, q_{12}, r_{12}, respectively. Here, the weight interval associated with the hesitancy component of A_2 is I'. The mapping d is a distance measure.

Proof We show that d satisfies all the properties of distance measure:

1. It is trivial to see $0 \le d(A_1, A_2) \le 1$, where A_1 and A_2 are any two AIFS of S.

2. Let, $d(A_1, A_2) = 0$

$$\Leftrightarrow \left[\frac{1}{2n}\sum_{i=1}^{n} p_{12}(\mu_{A_1}(x_i) - \mu_{A_2}(x_i))^2 + q_{12}(v_{A_1}(x_i) - v_{A_2}(x_i))^2 + r_{12}(\pi_{A_1}(x_i) - \pi_{A_2}(x_i))^2\right]^{1/2} = 0. \tag{12}$$

Equation (12) implies

$$\mu_{A_1}(x_i) = \mu_{A_2}(x_i), \quad v_{A_1}(x_i) = v_{A_2}(x_i).$$

3. It can be easily shown that $d(A_1, A_2) = d(A_2, A_1)$.

4. To establish this condition, it is equivalent to prove the property of transitivity (see [42, 43] and [41]). To do so, let A_1, A_2, A_3 be the three AIFSs. Now,

$$d(A_1,A_3) = \left[\frac{1}{2n}\sum_{i=1}^{n} p_{12}(\mu_{A_1}(x_i) - \mu_{A_3}(x_i))^2 + q_{12}(v_{A_1}(x_i) - v_{A_3}(x_i))^2 + r_{12}(\pi_{A_1}(x_i) - \pi_{A_3}(x_i))^2\right]^{1/2}. \tag{13}$$

Using Minkowski's inequality, we get

$$\le \left[\frac{1}{2n}\sum_{i=1}^{n} p_{12}(\mu_{A_1}(x_i) - \mu_{A_2}(x_i))^2 + q_{12}(v_{A_1}(x_i) - v_{A_2}(x_i))^2 + r_{12}(\pi_{A_1}(x_i) - \pi_{A_2}(x_i))^2\right]^{1/2} + \left[\frac{1}{2n}\sum_{i=1}^{n} p_{12}(\mu_{A_2}(x_i) - \mu_{A_3}(x_i))^2 + q_{12}(v_{A_2}(x_i) - v_{A_3}(x_i))^2 + r_{12}(\pi_{A_2}(x_i) - \pi_{A_3}(x_i))^2\right]^{1/2} \tag{14}$$

$$\Rightarrow d(A_1, A_3) \le d(A_1, A_2) + d(A_2, A_3)$$

Thus, $d(A_1, A_2)$ is a distance measure. □

Theorem 2.3 *Let* $\theta_1 : \mathbb{M}_{si} \times \mathbb{V}_{sd} \to \mathbb{R}$ *be a mapping, such that* $\theta_1(U, V) = J_m(U, V)$*, where cluster matrix,* $V \in \mathbb{V}_{sd}$ *and membership matrix,* $U \in \mathbb{M}_{si}$*. Then* U^* *is a strict* *local minima if U^* is calculated from Eq. (23). The coefficients p_{12}, q_{12}, r_{12} are known (see Theorem 2.2).*

Proof The criterion function and constraint are given as

$$J_m(U,V)=\sum_{s=1}^{C}\sum_{i=1}^{P}u_{si}^m\sum_{d=1}^{D}\left[p_{12}(\mu_{id}-\bar{\mu}_{sd})^2+q_{12}(v_{id}-\bar{v}_{sd})^2+r_{12}(\pi_{id}-\bar{\pi}_{sd})^2\right] \quad (15)$$

such that

$$\sum_{s=1}^{C}u_{si}=1, \qquad \forall\ 1\le i\le P. \quad (16)$$

The Lagrangian $G(U, V)$ is constructed based on the criterion function and constraint as follows:

$$G(U,V)=\sum_{i=1}^{P}\left(\sum_{s=1}^{C}u_{si}^m\sum_{d=1}^{D}(p_{12}(\mu_{id}-\bar{\mu}_{sd})^2+q_{12}(v_{id}-\bar{v}_{sd})^2+r_{12}(\pi_{id}-\bar{\pi}_{sd})^2)-\sum_{i=1}^{P}\lambda_i\left[\sum_{s=1}^{C}u_{si}-1\right]\right. . \quad (17)$$

In Eq. (17), we have used Lagrange's multipliers $\lambda_i, 1\le i\le P$. Here, the fuzzy factor $m\in(0,1)\cup(1,\infty)$ operates over the elements of the membership matrix, U. The element $u_{li}\in[0,1]$ denotes the membership grade of the ith element of the universe of discourse within the lth cluster. It is a minimization problem, so for its solution, we differentiate Eq. (17) with respect to Lagrange multipliers λ_i, and set them equal to zero as follows:

$$\frac{\partial G(U,\mathbf{V},\mathbf{W})}{\partial\lambda_i}=-\left[\sum_{s=1}^{c}u_{si}-1\right]=0, \quad \forall 1\le i\le P. \quad (18)$$

Similarly, derivatives of the Lagrangian condition (17) are set equal to zero with respect to membership parameter u_{si}, where $1\le i\le P, 1\le s\le C$ as

$$\frac{\partial G(U,V)}{\partial u_{si}}=mu_{si}^{m-1}\sum_{d=1}^{D}((p_{12}(\mu_{id}-\tilde{\mu}_{sd})^2+q_{12}(v_{id}-\tilde{v}_{sd})^2+r_{12}(\pi_{id}-\tilde{\pi}_{sd})^2))-\lambda_i=0 . \quad (19)$$

Solving Eq. (19), we get

$$u_{si}=\left(\frac{\lambda_i}{m}\right)^{\frac{1}{m-1}}\sum_{d=1}^{D}(p_{12}((\mu_{id}-\bar{\mu}_{sd})^2+q_{12}(v_{id}-\bar{v}_{sd})^2+r_{12}(\pi_{id}-\bar{\pi}_{sd})^2))^{\frac{1}{1-m}} . \quad (20)$$

To compute the iterative formula for u_{si}, we combine Eqs. (18) and (20) to yield

$$\sum_{s=1}^{C}\left(\frac{\lambda_i}{m}\right)^{\frac{1}{m-1}}\left(\sum_{d=1}^{D}(p_{12}(\mu_{id}-\bar{\mu}_{sd})^2+q_{12}(v_{id}-\bar{v}_{sd})^2+r_{12}(\pi_{id}-\bar{\pi}_{sd})^2)\right)^{\frac{1}{1-m}}=1 \quad (21)$$

$$\left(\frac{\lambda_i}{m}\right)^{\frac{1}{m-1}}\sum_{s=1}^{C}\left(\sum_{d=1}^{D}(p_{12}(\mu_{id}-\bar{\mu}_{sd})^2+q_{12}(v_{id}-\bar{v}_{sd})^2+r_{12}(\pi_{id}-\bar{\pi}_{sd})^2)\right)^{\frac{1}{1-m}}=1 . \quad (22)$$

Therefore, with the division of Eq. (20) by Eq. (22), the iterative formula for membership value u_{si} is obtained as follows:

$$u_{si}=\frac{\left\{\left(\sum_{d=1}^{D}p_{12}(\mu_{id}-\bar{\mu}_{sd})^2+q_{12}(v_{id}-\bar{v}_{sd})^2+r_{12}(\pi_{id}-\bar{\pi}_{sd})^2\right)^{\frac{1}{1-m}}\right\}}{\left\{\sum_{s=1}^{C}\left(\sum_{d=1}^{D}p_{12}((\mu_{id}-\bar{\mu}_{sd})^2+q_{12}(v_{id}-\bar{v}_{sd})^2+r_{12}(\pi_{id}-\bar{\pi}_{sd})^2)\right)^{\frac{1}{1-m}}\right\}} \quad (23)$$

□

Theorem 2.4 *The optimal minima U^* is obtained at a point $V=V^*$.*

Proof We define the Lagrangian $G(U, V)$ as

$$G(U,V)=\sum_{i=1}^{P}\left(\sum_{s=1}^{C}u_{si}^m\sum_{d=1}^{D}(p_{12}(\mu_{id}-\bar{\mu}_{sd})^2+q_{12}(v_{id}-\bar{v}_{sd})^2+r_{12}(\pi_{id}-\bar{\pi}_{sd})^2)-\sum_{l=1}^{P}\lambda_l\left[\sum_{s=1}^{C}u_{si}-1\right]\right) . \quad (24)$$

To find the cluster center matrix V, we equate derivatives of the Lagrangian (see Eq. (24)) equal to zero with respect to cluster center $\bar{\mu}_{sd}, \bar{v}_{sd}, \bar{\pi}_{sd}$, where $1\le s\le C, 1\le d\le D$. It yields,

$$\frac{\partial G(U,V)}{\partial\bar{\mu}_{sd}}=0 \quad (25)$$

$$\frac{\partial G(U,V)}{\partial\bar{v}_{sd}}=0 \quad (26)$$

$$\frac{\partial G(U,V)}{\partial\bar{\pi}_{sd}}=0 . \quad (27)$$

Solving Eq. (25), we get

$$\frac{\partial G(U,V)}{\partial \bar{\mu}_{sd}} = \sum_{i=1}^{P}\left(u_{si}^{m} p_{12}[-2(\mu_{id}-\bar{\mu}_{sd})]\right) = 0 \tag{28}$$

$$\bar{\mu}_{sa} = \frac{\sum_{i=1}^{P} u_{si}^{m}\mu_{id}}{\sum_{i=1}^{P} u_{si}^{m}}. \tag{29}$$

Similarly, we obtain the following solutions from Eqs. (26) and (27), respectively,

$$\bar{v}_{sa} = \frac{\sum_{i=1}^{P} u_{si}^{m} v_{id}}{\sum_{i=1}^{P} u_{si}^{m}} \tag{30}$$

$$\bar{\pi}_{sa} = \frac{\sum_{i=1}^{P} u_{si}^{m}\pi_{id}}{\sum_{i=1}^{P} u_{si}^{m}}. \tag{31}$$

Here, eqns. (29), (30), and (31) derive the point $V = V^*$. □ □

2.3 A Proposal of Novel Intuitionistic Fuzzy *C*-Means Clustering Algorithm

This section establishes some mathematical results to propose the uni-weighted intuitionistic fuzzy C-means clustering (uW-IFCM) algorithm. A feature weight matrix (namely, $W = (w_d)_{1\times D}$)-dependent distance measure $D_2^2(\hat{x_i}, \hat{v_l})$ is used here. We have

$$D_2^2(\hat{x_i},\hat{v_l}) = \sum_{d=1}^{D} w_d^{\eta}\left((\mu_{id}-\bar{\mu}_{ld})^2+(v_{id}-\bar{v}_{ld})^2+(\pi_{id}-\bar{\pi}_{ld})^2\right). \tag{32}$$

The set containing C cluster centroids against dth feature is $V = \{v_1, v_2, \ldots, v_C; v_l = (\bar{\mu}_{ld}, \bar{v}_{ld}, \bar{\pi}_{ld}), 1 \le l \le C\}$. Here, the theorems derive membership matrix U, cluster centroids $(\bar{\mu}_{ld}, \bar{v}_{ld}, \bar{\pi}_{ld})$, and weight matrix W.

Theorem 2.5 *Let $\theta_1 : \mathbb{M}_{li} \to \mathbb{R}$ be a mapping, such that $\theta_1(U) = J(U, \mathbf{V}, \mathbf{W})$, where cluster matrix $\mathbf{V} \in \mathbb{V}_{ld}$ and weight matrix $\mathbf{W} \in \mathbb{W}_{ld}$ are fixed. Then U^* is a strict local minima if U^* is calculated from Eq. (41).*

Proof As $\mathbf{V}$ and $\mathbf{W}$ are kept fixed, they are marked bold. Let us take cluster center matrix as $\mathbf{V} = [\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_C]^T$, where $\mathbf{v}_l = (\bar{\mu}_{lj}, \bar{v}_{lj}, \bar{\pi}_{lj})_{j=1}^{D}$. The weight matrix $\mathbf{W}$ is of order $1 \times D$, and $\sum_{d=1}^{D} w_d = 1$, for all $1 \le i \le P$. Here, $m, \eta \in (0,1) \cup (1,\infty)$. Now, we define the criterion function:

$$J_m(U,\mathbf{V},\mathbf{W}) = \sum_{l=1}^{C}\sum_{i=1}^{P} u_{li}^{m} \sum_{d=1}^{D} w_d^{\eta} \left[(\mu_{id}-\bar{\mu}_{ld})^2+(v_{id}-\bar{v}_{ld})^2+(\pi_{id}-\bar{\pi}_{ld})^2\right] \tag{33}$$

the following constraint is used for defining Lagrangian:

$$\sum_{l=1}^{C} u_{li} = 1, \qquad \forall\ 1 \le i \le P. \tag{34}$$

Equations (33) and (34) define the Lagrangian:

$$G(U,\mathbf{V},\mathbf{W}) = \sum_{i=1}^{P}\left(\sum_{l=1}^{C} u_{li}^{m}\sum_{d=1}^{D} w_d^{\eta}((\mu_{id}-\bar{\mu}_{ld})^2 + (v_{id}-\bar{v}_{ld})^2 + (\pi_{id}-\bar{\pi}_{ld})^2)\right. - \sum_{i=1}^{P}\lambda_i\left[\sum_{l=1}^{C} u_{li} - 1\right]. \tag{35}$$

The derivatives of the Lagrangian condition (35) are set equal to zero with respect to Lagrange's multipliers λ_i:

$$\frac{\partial G(U,\mathbf{V},\mathbf{W})}{\partial \lambda_i} = -\left[\sum_{l=1}^{c} u_{li} - 1\right] = 0, \quad \forall 1 \le i \le P. \tag{36}$$

Similarly, derivatives of the Lagrangian condition (35) are set equal to zero with respect to membership parameter u_{si}, where $1 \le i \le P, 1 \le s \le C$. So,

$$\frac{\partial G(U,\mathbf{V},\mathbf{W})}{\partial u_{si}} = m u_{si}^{m-1}\sum_{d=1}^{D} w_d^{\eta}(((\mu_{ia}-\bar{\mu}_{sa})^2 + (v_{ia}-\bar{v}_{sa})^2 + (\pi_{ia}-\bar{\pi}_{sa})^2)) - \lambda_i = 0. \tag{37}$$

Solving Eq. (37), we get

$$u_{si} = \left(\frac{\lambda_i}{m}\right)^{\frac{1}{m-1}} \sum_{d=1}^{D} w_d^{\eta}(((\mu_{id}-\bar{\mu}_{sd})^2 + (v_{id}-\bar{v}_{sd})^2 + (\pi_{id}-\bar{\pi}_{sd})^2))^{\frac{1}{1-m}}. \tag{38}$$

Equations (36) and (38) together yield:

$$\sum_{l=1}^{C}\left(\frac{\lambda_i}{m}\right)^{\frac{1}{m-1}}\left(\sum_{d=1}^{D} w_d^{\eta}((\mu_{ia}-\bar{\mu}_{sd})^2+(v_{id}-\bar{v}_{sd})^2+(\pi_{id}-\bar{\pi}_{sd})^2)\right)^{\frac{1}{1-m}} = 1 \tag{39}$$

$$\left(\frac{\lambda_i}{m}\right)^{\frac{1}{m-1}}\sum_{l=1}^{C}\left(\sum_{d=1}^{D} w_d^{\eta}((\mu_{ia}-\bar{\mu}_{sd})^2+(v_{id}-\bar{v}_{sd})^2+(\pi_{id}-\bar{\pi}_{sd})^2)\right)^{\frac{1}{1-m}} = 1. \tag{40}$$

Dividing Eq. (38) by Eq. (40) results in an iterative formula for membership value u_{si} as follows:

$$u_{si} = \frac{\left\{\left(\sum_{d=1}^{D} w_d^{\eta}(\mu_{id}-\bar{\mu}_{sd})^2 + (v_{id}-\bar{v}_{sd})^2 + (\pi_{id}-\bar{\pi}_{sd})^2\right)^{\frac{1}{1-m}}\right\}}{\left\{\sum_{l=1}^{c}\left(\sum_{d=1}^{D} w_d^{\eta}((\mu_{ia}-\bar{\mu}_{sd})^2 + (v_{id}-\bar{v}_{sd})^2 + (\pi_{id}-\bar{\pi}_{sd})^2)\right)^{\frac{1}{1-m}}\right\}} \tag{41}$$

□

Theorem 2.6 *Let $\theta_2 : \mathbb{W}_D \to \mathbb{R}$ be a mapping such that $\theta_2(W) = J(\mathbf{U}, \mathbf{V}, W)$, where membership matrix $\mathbf{U} \in \mathbb{U}_{li}$*

and cluster matrix $\mathbf{V} \in \mathbb{V}_{ld}$ *remain constant. Then* W^* *is a strict local minima if Eq. (50) calculates* W^*.

Proof The membership matrix $\mathbf{U}$ and cluster center matrix $\mathbf{V}$ are kept constant and we mark them bold. The Lagrangian $G(\mathbf{U}, \mathbf{V}, W)$ is defined using a weight matrix $W = (w_d^\eta)_{1\times D}$ as follows:

$$G(\mathbf{U},\mathbf{V},W)=\sum_{i=1}^{P}\left(\sum_{l=1}^{C}u_{li}^{m}\sum_{d=1}^{D}w_d^{\eta}((\mu_{id}-\bar{\mu}_{ld})^2+(\nu_{id}-\bar{\nu}_{ld})^2+(\pi_{id}-\bar{\pi}_{ld})^2)\right.\\ \left.-\sum_{l=1}^{P}\lambda_l\left[\sum_{d=1}^{D}w_d-1\right]\right). \quad (42)$$

The derivatives of the Lagrangian (see Eq. (42)) with respect to Lagrange multiplier λ_l are set equal to zero. Hence,

$$\frac{\partial G(U,V,W)}{\partial \lambda_l} = -\left[\sum_{d=1}^{D} w_d - 1\right] = 0, \quad \forall 1 \le l \le P \quad (43)$$

Similarly, the derivatives of the Lagrangian (see Eq. (42)) with respect to weight parameter w_a, $1 \le a \le D$ is set equal to zero. Hence,

$$\frac{\partial G(\mathbf{U},\mathbf{V},W)}{\partial w_a}=\sum_{i=1}^{q}u_{si}^{m}\eta w_{sd}^{(\eta-1)}(((\mu_{ia}-\bar{\mu}_{sa})^2+(\nu_{ia}-\bar{\nu}_{sa})^2+\\ (\pi_{ia}-\bar{\pi}_{sa})^2))-\lambda_l=0. \quad (44)$$

We simplify Eq. (44) and it derives feature weight w_a as follows:

$$\sum_{i=1}^{P}u_{si}^{m}\eta w_a^{(\eta-1)}(((\mu_{ia}-\bar{\mu}_{sa})^2+(\nu_{ia}-\bar{\nu}_{sa})^2+(\pi_{ia}-\bar{\pi}_{sa})^2))-\lambda_l=0 \quad (45)$$

$$\sum_{i=1}^{P}u_{si}^{m}\eta w_{sa}^{(\eta-1)}(((\mu_{ia}-\bar{\mu}_{sa})^2+(\nu_{ia}-\bar{\nu}_{sa})^2+(\pi_{ia}-\bar{\pi}_{sa})^2))=\lambda_l. \quad (46)$$

For $\eta \neq 1$:

$$w_a=\left(\frac{\lambda_l}{\eta}\right)^{\frac{1}{\eta-1}}\left[\sum_{i=1}^{P}u_{si}^{m}\eta w_a^{(\eta-1)}((\mu_{ia}-\bar{\mu}_{sa})^2+(\nu_{ia}-\bar{\nu}_{sa})^2+(\pi_{ia}-\bar{\pi}_{sa})^2)\right]^{\frac{1}{1-\eta}}. \quad (47)$$

Equations (46) and (47) are solved, so we have

$$\sum_{d=1}^{D}\left(\frac{\lambda_l}{\eta}\right)^{\frac{1}{\eta-1}}\left[\sum_{i=1}^{P}mu_{si}^{m}\eta w_d^{(\eta-1)}((\mu_{id}-\bar{\mu}_{sd})^2+(\nu_{id}-\bar{\nu}_{sd})^2\right.\\ \left.+(\pi_{id}-\bar{\pi}_{sd})^2)\right]^{\frac{1}{1-\eta}}=1 \quad (48)$$

$$\left(\frac{\lambda_l}{\eta}\right)^{\frac{1}{\eta-1}}\sum_{d=1}^{D}\left[\sum_{i=1}^{P}mu_{si}^{m}\eta w_d^{(\eta-1)}((\mu_{id}-\bar{\mu}_{sd})^2+(\nu_{id}-\bar{\nu}_{sd})^2\right.\\ \left.+(\pi_{id}-\bar{\pi}_{sd})^2)\right]^{\frac{1}{1-\eta}}=1. \quad (49)$$

Let us divide Eq. (47) by Eq. (49), and it results in the iterative formula for weight w_a as follows:

$$w_a=\frac{\left\{\left[\sum_{i=1}^{P}u_{si}^{m}((\mu_{ia}-\bar{\mu}_{sa})^2+(\nu_{ia}-\bar{\nu}_{sa})^2+(\pi_{ia}-\bar{\pi}_{sa})^2)\right]^{\frac{1}{1-\eta}}\right\}}{\left\{\sum_{d=1}^{D}\left[\sum_{i=1}^{P}u_{si}^{m}((\mu_{id}-\bar{\mu}_{sd})^2+(\nu_{id}-\bar{\nu}_{sd})^2+(\pi_{id}-\bar{\pi}_{sd})^2)\right]^{\frac{1}{1-\eta}}\right\}} \quad (50)$$

□

Corollary 2.1 *Equations. (55), (56), and (57) give the iterative formula for cluster center matrix V.*

For proof, we set derivatives of the Lagrangian (see Eq. (42)) with respect to cluster center $\bar{\mu}_{sa}, \bar{\nu}_{sa}, \bar{\pi}_{sa}$, *where* $1 \le s \le C, 1 \le a \le D$ *equal to zero is as follows:*

$$\frac{\partial G(U,V,W)}{\partial \bar{\mu}_{sa}} = 0 \quad (51)$$

$$\frac{\partial G(U,V,W)}{\partial \bar{\nu}_{sa}} = 0 \quad (52)$$

$$\frac{\partial G(U,V,W)}{\partial \bar{\pi}_{sa}} = 0. \quad (53)$$

From Eq. (51), we have

$$\frac{\partial G(U,V,W)}{\partial \bar{\mu}_{sa}} = \sum_{i=1}^{P}\left(u_{si}^{m}w_a^{\eta}[-2(\mu_{ia}-\bar{\mu}_{sa})]\right) = 0 \quad (54)$$

$$\bar{\mu}_{sa} = \frac{\sum_{i=1}^{P} u_{si}^{m}\mu_{id}}{\sum_{i=1}^{P} u_{si}^{m}}. \quad (55)$$

Similarly, from Eqns. (52) and (53), we get

$$\bar{\nu}_{sa} = \frac{\sum_{i=1}^{P} u_{si}^{m}\nu_{id}}{\sum_{i=1}^{P} u_{si}^{m}} \quad (56)$$

$$\bar{\pi}_{sa} = \frac{\sum_{i=1}^{P} u_{si}^{m}\pi_{id}}{\sum_{i=1}^{P} u_{si}^{m}} \quad (57)$$

□

3 Procedural Details of uW-IFCM and bW-PIFCM

The implementation of the uW-IFCM algorithm involves seven steps. In the first step, intuitionistic data fuzzification technique transforms the given dataset into an AIFS dataset. Here, the functions μ_{id}, ν_{id}, and π_{id} calculate membership, non-membership, and hesitancy values, respectively. The optimized iterative values of membership matrix U, centroid matrix V, weight matrix W, and convergence criteria are determined from step 2 to step 7. A brief procedure to implement uW-IFCM and bW-PIFCM has been given in Algorithm III and Algorithm IV.

Step 1. Intuitionistic data fuzzification technique: Let us consider a real-valued multi-dimensional/featured dataset $T = \{x_1^{'}, x_2^{'}, \ldots, x_P^{'}\}$, where $x_i^{'} = (x_{id})_{d=1}^{D}$ and $x_{id} \in R$ for all $1 \leq d \leq D$. The bi-parametric formula (a and b are the parameters) normalizes T as follows:

$$N_{id} = a + \frac{x_{id} - x^{i}_{\substack{min\\d}}}{x^{i}_{\substack{max\\d}} - x^{i}_{\substack{min\\d}}}(b - a)\,. \tag{58}$$

The symbols $x^{i}_{\substack{\min\\d}}$ and $x^{i}_{\substack{\max\\d}}$ denote the minima and maxima of the set $(x_{id})_{d=1}^{D}$. Here, $a = 0$ and $b = 1$.

1. *Membership function,* μ_{id}: Equation. (58) transforms T into a normalized data matrix N. The `norm` function of MATLAB computes Euclidean distance between each pair $(x_i^{'}, x_j^{'})$, and its resultant is a distance matrix $[dis_{ij}]$. Let us multiply N and $[dis_{ij}]$, and then we obtain the matrix $[P_{id}]$. The feature weight allocation to x_{id} is done with a factor M_{id} computed as follows:

$$M_{id} = \frac{1}{P_{id}}. \tag{59}$$

We have calculated $\sum_{i=1}^{P}\sum_{d=1}^{D}(x_{id} + M_{id})$, and it results to a new matrix called *normalized-* $(x_{id} + M_{id})$ $= \frac{x_{id}+M_{id}}{\sum_{i=1}^{n}\sum_{d=1}^{D}(x_{id}+M_{id})}$. Now, a data-centric membership function is computed as follows:

$$\mu_{id} = (normalized\text{-}\ (x_{id} + M_{id}))^{\beta}, \qquad \beta > 0\,. \tag{60}$$

The tuning parameter $\beta \in (0, 10)$ tunes the membership values as per the requirement of the dataset T.

2. *Non-membership function,* ν_{id}: The generalized intuitionistic fuzzy generator given by [40] derives the non-membership function ν_{id} as

$$\nu_{id} = (1 - \mu_{id}^{\alpha\beta})^{\frac{1}{\alpha}}\,. \tag{61}$$

The domain of the tuning parameter α is (0, 1].

3. *Hesitancy function,* π_{id}: The mathematical formulation of the hesitancy function is given as follows:

$$\pi_{id} = 1 - \mu_{id} - (1 - \mu_{id}^{\alpha\beta})^{\frac{1}{\alpha}} \tag{62}$$

Equations. (60), (61), and (62) together transform x_{id} into an AIFS $\tilde{x}_{id} = (\mu_{id}, \nu_{id}, \pi_{id})$. This novel data fuzzification technique is implemented with the help of Algorithm I.

Now, we discuss the outlines of Algorithm II:

Step 2. Initialization of centroid matrix V: To initialize uW-IFCM, and bW-PIFCM algorithms, it is necessary to fix the prior number of cluster centroids. We randomly select initial cluster centroids, $\hat{v}_l$, $1 \leq l \leq C$ with the help of *rand* function of MATLAB.

Step 3. Initialization of weight matrix, W: The feature information provided in the multi-dimensional dataset is used to compute initial feature weights in uW-IFCM (see Eqn 50), whereas bW-PIFCM algorithm uses data-driven weight triplets for its initialization (see Theorem 2.2).

Step 4. Computation of membership matrix, U: The membership matrix U contains the membership value of each data point in every cluster. The higher the membership value, the more the associativity of the data points in the particular cluster. Here, the matrix $U = [u_{kl}]$ is of order $P \times C$, where P is the number of data points and C is the number of clusters. Eq. (41) and Eq. (23) provide mathematical formulas with which we calculate the membership matrix corresponding to uW-IFCM and bW-PIFCM algorithms, respectively.

Step 5. Updation of cluster centroid matrix, V: We have arbitrarily selected C number of initial cluster centroids $(\hat{v}_l = (\mu_{ld}, \nu_{ld}, \pi_{ld})_{d=1}^{D}, 1 \leq l \leq C)$ in step 2. The same weightage has been given to each dimension d of $\hat{s}_l$ (lth cluster centroid). The uW-IFCM uses Eqns. (55-57) to update V, whereas Eqns. (29-31) update the cluster centroids of bW-PIFCM. In some algorithms, we see an explicit role of hesitancy during the updation of centroid matrix [32]).

Step 6. Updation of weight matrix, W: To update W, the updated membership matrix U and the updated cluster centroids V of uW-IFCM are utilized. We observe that after each iteration, there is an increase in the weights of relevant features and a decrease in the weights of irrelevant features or noisy features. The data-specific appropriate selection of the exponent η and fuzzy factor m delivers an optimal feature weight distribution. The formula (see Eq. (50)) works for the weight matrix updation. In the case of bW-PIFCM, we update the probability weight intervals I_1, I_2 and I with the help of Algorithm III.

Step 7. Convergence criterion: The convergence of the uW-IFCM and bW-PIFCM algorithms are decided with an error margin ϵ equal to 10^{-6}. The convergence is met for uW-IFCM if the termination criterion $\sum_{d=1}^{D}\frac{d_2(W_d(t), W_d(t+1))}{D} < \epsilon$ is reached else the algorithm repeats themselves from Step 4. The termination criterion used for bW-PIFCM is $\sum_{l=1}^{C}\frac{d_2(V_l(t), V_l(t+1))}{C} < \epsilon$ till the convergence is reached.

Input:	Dataset (X), Number of clusters (C), Fuzzy factor (m), Intuitionistic fuzzy parameter (α), Tuning parameter for membership function (β), Tolerance level (ϵ), Weighting exponent η.
Output:	Intuitionistic fuzzy set (μ_{id}, ν_{id}, π_{id}) corresponding to x_{id}
Procedure:	**1**. Normalize the dataset, X using the Eqn. (58) to obtain matrix N.
	2. Deduce the Euclidean distance $[dis_{ij}]$ between x_i and x_j.
	3. Deduce the product of the distance matrix $[dis_{ij}]$ with normalized dataset N.
	4.Compute the membership value μ_{id} using Eqn. (60), non-membership value ν_{id} using Eqn. (61) and hesitancy value π_{id} using Eqn. (62), respectively.
End Procedure	AIFS dataset corresponding to X;

Input:	X: Matrix of data items, N: Number of iterations, C: Number of clusters, η: Weighting exponent, m: Fuzzy factor, α: Complement parameter, β: Tuning parameter of fuzzification, Tolerance level ($\epsilon = 0.006$), $\mathbb{T}_{\mathbb{C}}$ (Threshold condition), $\mathbb{I}_{\ltimes}$ (number of iterations).
Output	Weight matrix W, Membership matrix U, Cluster matrix V, Optimal α, β, m
Procedure:	**1**. Initialize cluster matrix V randomly at $t = 0$
	2. Initialize feature weight matrix W with random numbers and normalize with constraint $\sum_{d=1}^{D} w_d = 1$;
	For $m = 1$ to 30
	For $\alpha = 1$ to 20
	For $\beta = 1$ to 10
	Intuitionistic fuzzification of a given dataset using Algorithm I,
	fuzzify data to get (μ_{id}, ν_{id}, π_{id}) corresponding to each data point x_{id}.
	While $\mathbb{I}_{\ltimes} > 100$ & $\mathbb{T}_{\mathbb{C}} > \epsilon$ Update $U_{m,\alpha,\beta}$ with Eqn. (41) and normalize with Eqn. (34);
	Update feature weight matrix $W_{m,\alpha,\beta}$ with Eqn. (50) and normalize with constraint $\sum_{d=1}^{D} w_d = 1$;
	Update centroid matrix $V_{m,\alpha,\beta} = (\bar{\mu}_{ld}, \bar{\nu}_{ld}, \bar{\pi}_{ld})_{d=1}^{D}$ using (55 -57):
	End
	End
	End
	End
End Procedure	U^{t+1}, ($\bar{\mu}_l^{t+1}$, $\bar{\nu}_l^{t+1}$,$\bar{\pi}_l^{t+1}$) and W^{t+1} corresponding to optimal m, α, β;

Input	Intuitionistic fuzzified dataset $\tilde{X} = \{\hat{x}_i\}, \hat{x}_i = (\tilde{x}_{id})_{d=1}^{D}$, where $\tilde{x}_{id} = (\mu_{id},\ \nu_{id},\ \pi_{id})$ using Algorithm I.
Output	Probabilistic weight interval p_{ij}, q_{ij}, r_{ij}.
Process	1. Compute minimum of membership value $\mu_{\min}(x_i) = \min\limits_i \mu_{id}$ with respect to each feature of x_i. Then, compute minimal sum $\{\mu_{\min}(x_1) + \mu_{\min}(x_2) + \cdots + \mu_{\min}(x_D)\}$.
	2. Normalize $\{\mu_{\min}(x_i)\}$ of each feature d with $\{\mu_{\min}(x_1) + \mu_{\min}(x_2) + \cdots + \mu_{\min}(x_D)\}$ and hence probabilistic value p_i with respect to d of x_i is obtained.
	3. Compute minimum probability $p^i_{\min} = (p_i * \mu_{id})$ and maximum probability $p^i_{\max} = p_{\min} + (p_i * \pi_{id})$ for each d of x_i.
	4. Calculate $p'_{ij} = \max(p^i_{\min}, p^j_{\min})$, $p''_{ij} = \min(p^i_{\max}, p^j_{\max})$ to obtain $[p', p'']$.
	5. Compute minimum of membership value $\nu_{\min}(x_i) = \min\limits_i \nu_{id}$ with respect to each feature of x_i. Then, compute minimal sum $\{\nu_{\min}(x_1)+\nu_{\min}(x_2)+\cdots+\nu_{\min}(x_D)\}$.
	6. Normalize $\{\nu_{\min}(x_i)\}$ of each feature d with $\{\nu_{\min}(x_1) + \nu_{\min}(x_2) + \cdots + \nu_{\min}(x_D)\}$ and hence probabilistic value q_i with respect to d of x_i is obtained.
	7. Compute minimum probability $q^i_{\min} = (q_i * \mu_{id})$ and maximum probability $q^i_{\max} = q_{\min} + (q_i * \pi_{id})$ for each d of x_i.
	8. Calculate $q'_{ij} = \max(q^i_{\min}, q^j_{\min})$, $q''_{ij} = \min(q^i_{\max}, q^j_{\max})$ to obtain $[q', q'']$.
	9. Using theorem 2.1, we compute the weight interval corresponding to hesitancy component of an IFS (μ_{id}, ν_{id}, π_{id}).
Return:	Probabilistic weights p_{ij}, q_{ij}, r_{ij}

4 An Experimental Study of Synthetic and UCI Machine Learning Datasets

This section is divided into four subsections as follows:

Input:	X: Matrix of data items, N: Number of iterations, C: Number of clusters, η: Weighting exponent, m: Fuzzy index, α: Complement parameter, β: Tuning parameter of fuzzification, Tolerance level ($\epsilon = 0.006$), $\mathbb{T}_{\mathbb{C}}$ (Threshold condition), $\mathbb{I}_{\ltimes}$ (number of iterations)
Output	Membership matrix U, Cluster matrix V
Procedure:	1. Initialize cluster matrix V randomly at $t = 0$;
	For $m = 1$ to 30
	For $\alpha = 1$ to 20
	For $\beta = 1$ to 10
	Intuitionistic fuzzification of a given dataset using Algorithm I,
	fuzzify data to get $(\mu_{id},\ \nu_{id},\ \pi_{id})$ corresponding to each data point x_{id}.
	Using Algorithm III, compute the probabilistic weights $p_{ij,q_{ij},r_{ij}}$.
	While $\mathbb{I}_{\ltimes} > 100$ & $\mathbb{T}_{\mathbb{C}} > \epsilon$ **do**
	Update $U_{m,\alpha,\beta}$ with Eqn. (41) and normalize with Eqn. (34);
	Update centroid matrix $V_{m,\alpha,\beta} = \{(\bar{\mu}_{ld}, \bar{\nu}_{ld}, \bar{\pi}_{ld})\}_{d=1}^{D}$ using (55-57):
	End
	End
	End
	End
End Procedure	U^{t+1}, $(\bar{\mu}_l^{t+1},\ \bar{\nu}_l^{t+1}, \bar{\pi}_l^{t+1})$;

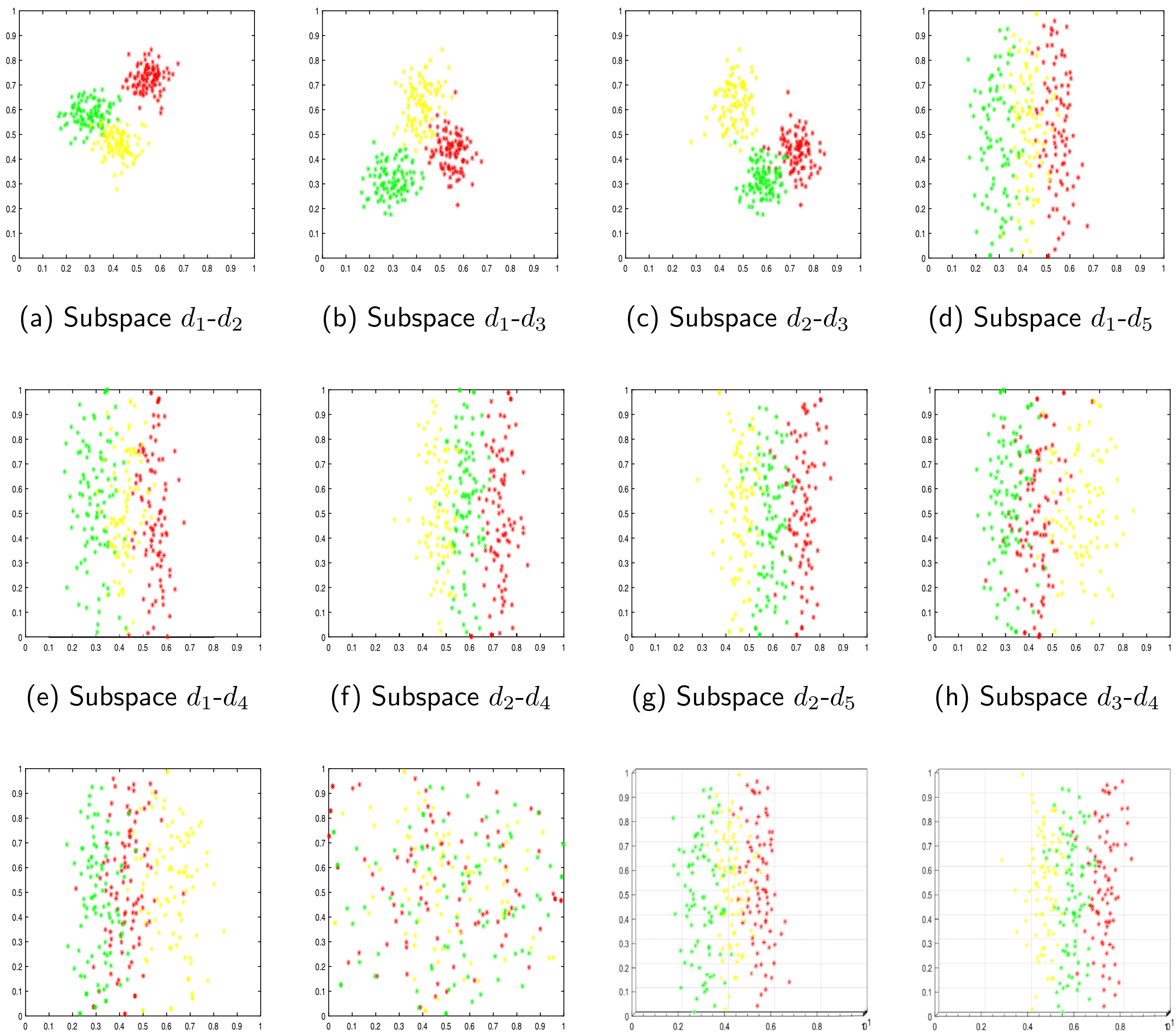
Input:	Dataset (X), number of centroids (C), weighing exponent (η), fuzzy factor (m), intuitionistic fuzzy parameter (α), tuning parameter for membership function (β), tolerance level (ϵ)
Output:	Fuzzy partition U, centroids $\{(\bar{\mu}_{ld}, \bar{\nu}_{ld}, \bar{\pi}_{ld})\}_{d=1}^{D}$, Weight matrix W
Procedure:	**1**. 1. Data fuzzification using Algorithm I.
	2. 2. Initialize centroid $\hat{V}$, U at $t = 0$.
	3. 3. Initialize weight matrix W with $\frac{1}{D}$'s.
	Repeat
	4. Update $(U = u_{il})^{t+1}$ by calculating the fuzzy partition using Eqn. (41).
	5. Update centroid $\{(\bar{\mu}_{ld}, \bar{\nu}_{ld}, \bar{\pi}_{ld})\}_{d=1}^{D}$ using Eqns. (55-57).
	6. Update weight matrix, W using Eqn. (50).
	7. 7. Optimize weight matrix with threshold value $1/(\sqrt{D})^{\chi}$.
	Until
	8. $\sum_{i=1}^{C} \frac{d_2(W_i(t), W_i(t+1))}{D} < \epsilon$ is satisfied.
Return	U^{t+1}, $(\bar{\mu}_l^{t+1},\ \bar{\nu}_l^{t+1}, \bar{\pi}_l^{t+1})$ and W^{t+1}.

4.1 Study of Synthetic Dataset I Using IFCM and Weighted-IFCM

Description of Synthetic Dataset I: For the experimentation, we use a Gaussian distribution function-based five-dimensional Synthetic Dataset I of sample size 300 (see [44]). Here, we denote the mean and the standard deviation with θ and σ, respectively. The mean value of the jth cluster along lth dimension (feature) is given as θ_{jl} with standard deviation σ_{jl} and distribution $\sum_{j=1}^{3}\sum_{l=1}^{5} N(\theta_{jl}, \sigma_{jl})$. The Synthetic Dataset I consists of three well-separated clusters and each cluster has 100 data items (see Fig. 1). In Table 3, the mean values and standard deviations, which we use for the three clusters are provided.

Claim: The features relevant to clustering are d_1, d_2, d_3, whereas d_4, d_5 are irrelevant features in the given dataset.

Strategy: This weighted-IFCM-based data analysis is bifurcated on the basis of the evaluation process of feature weight distribution as follows: (1) Firstly, we arbitrarily use four fixed feature weight distributions $W_i = \{w_{ik}\}_{k=1}^{5}$, $1 \leq i \leq 4$ (see Table 4). Here, w_{ik} is assigned to feature d_k, where $1 \leq k \leq 5$. (2) Secondly, we use an equally likely approach to calculate the feature weight distributions (see Table 5).



(a) Subspace d_1-d_2 (b) Subspace d_1-d_3 (c) Subspace d_2-d_3 (d) Subspace d_1-d_5

(e) Subspace d_1-d_4 (f) Subspace d_2-d_4 (g) Subspace d_2-d_5 (h) Subspace d_3-d_4

(i) Subspace d_3-d_5 (j) Subspace d_4-d_5 (k) Subspace $d_1 - d_4$-d_5 (l) Subspace $d_2 - d_4$-d_5

Fig. 1 Plotting of a three-cluster 5D Synthetic Dataset I with three normally distributed features d_1, d_2 and d_3 and two noisy features d_4 and d_5

Experimental analysis:

Table 3 Description of Synthetic Dataset I

Clusters	Parameters	d_1	d_2	d_3	d_4	d_5	Data items
	θ_{1l}	0.55	0.73	0.42	0.50	0.55	
Cluster1							100
	σ_{1l}	0.05	0.04	0.07	0.29	0.31	
	θ_{2l}	0.30	0.58	0.31	0.55	0.45	
Cluster2							100
	σ_{2l}	0.61	0.04	0.69	0.26	0.27	
	θ_{3l}	0.42	0.45	0.63	0.52	0.53	
Cluster3							100
	σ_{3l}	0.05	0.05	0.07	0.26	0.27	

Table 4 Clustering results of weighted-IFCM over Synthetic Dataset I

W	d_1	d_2	d_3	d_4	d_5	CA	R_I
W_1	**0.4656**	**0.2755**	**0.2000**	**0.0344**	**0.0245**	**0.9933**	**0.9911**
W_2	0.0245	0.2755	0.0344	0.2000	0.4656	0.6800	0.7039
W_3	0.0245	0.2755	0.2000	0.4656	0.0344	0.7333	0.7712
W_4	0.0245	0.2755	0.2000	0.0344	0.4656	0.7033	0.7572

Case 1. The collections W_1, W_2, W_3, and W_4 contain weights that are assigned to features d_1, d_2, d_3, d_4 and d_5 (see Table 4). The benchmark measuring indexes

Table 5 Clustering results over Synthetic Dataset I with an equally likely approach-based weighted-IFCM

W	d_1	d_2	d_3	d_4	d_5	CA	R_I
W_5	1/5	1/5	1/5	1/5	1/5	0.3333	0.3311
W_6	1/4	1/4	1/4	1/4	0	0.5566	0.6271
W_7	**1/3**	**1/3**	**1/3**	**0**	**0**	**0.9966**	**0.9955**
W_8	1/2	1/2	0	0	0	0.9800	0.9740
W_9	1	0	0	0	0	0.8600	0.8377

validating the clustering performance of weighted-IFCM are clustering accuracy CA and rand index R_I (for their details see Sect. 4.4). Using a random set of initial cluster centroids and the two parameters m, α (see Table 4), we initialize the weighted-IFCM algorithm. Here, 30 values of m are selected from interval [1, 4], and 20 values of α are selected out of the interval (0, 1) which results in 600 pairs of (m, α) for the algorithm. The use of W_1 in the weighted-IFCM yields better clustering in comparison to W_2, W_3, W_4, in this case, CA and R_I are obtained as 0.9933 and 0.9911, respectively (see Table 4). In W_1, the emphasis is given to features d_1, d_2, d_3 , whereas very less importance is assigned to d_4 and d_5. It is observed that a considerable weightage is given to d_4 or d_5 in W_2, W_3, W_4, so the weighted-IFCM does not perform well under these distributions (see Table 4). We further explore the role of features d_4 and d_5 in Case 2.

Case 2. The collection computed using an equally likely approach is W_5 = (1/5,1/5,1/5,1/5,1/5). We nullify features one by one, then the weight distributions obtained by applying this approach over four feature-based sets, three feature-based sets, two feature-based set, and single feature-based set are W_6 = (1/4,1/4,1/4,1/4,0), W_7 = (1/3,1/3,1/3,0,0), W_8 = (1/2,1/2,0,0,0), and W_9 = (1,0,0,0,0), respectively (see Table 5). The algorithm delivers its worst clustering performance at W_5 (CA = 0.3333 and RI = 0.3311), and delivers its best clustering performance at W_7 (CA = 0.9966 and RI = 0.9911). We have discarded features d_4 and d_5 in W_7. Both cases verify that d_4 and d_5 are irrelevant features. Please refer to Table 4 and Table 5 for all the results obtained using weighted-IFCM.

Through the experiment, we observe a marginal gap between the performance of weighted-IFCM with random allocation of weights and relevancy-based weights.

4.2 Study of Synthetic Dataset I Using bW-PIFCM

The weight distributions used in weighted-IFCM algorithm are randomly selected. In this algorithm, each feature is allocated a single weight, whereas the algorithms like PIFCM (see [41]) employ a weight triplet to cluster the multi-dimensional dataset. Here, we analyze Synthetic Dataset I with the bi-weighted variant of PIFCM called bW-PIFCM. The bW-PIFCM algorithm converges after seven iterations (see Fig. 3d). The bW-PIFCM performs well over synthetic datasets (see Table 6). It verifies that algorithm is adaptive in nature. Here, none of the features were neglected. Therefore, it cannot be used as a feature reduction technique.

Table 6 Clustering of synthetic datasets using bW-PIFCM

	Synthetic Dataset	CA	R_I
1	Synthetic Dataset I	0.9933	0.9920
2	Synthetic Dataset II	0.9980	0.9986
3	Synthetic Dataset III	0.9890	0.9874
4	Synthetic Dataset IV	0.9998	0.9999

4.3 Study of Synthetic Dataset I Using uW-IFCM

The research work of [21] motivates us to propose uW-IFCM clustering algorithm. The mathematical study of convergence of uW-IFCM becomes a trivial exercise due to the results established by [21]. So, the focus of the paper is on the experimental study of uW-IFCM algorithm. Our study on Synthetic Dataset I shows that uW-IFCM and IFCM take five and ten iterations, respectively, to converge (see Fig. 3a, b). The criterion value of uW-IFCM differs from that of wIFCM (see Table 2) by a large margin (3c). In the figure, the number of iterations is kept along the horizontal axis, whereas the criterion function values are placed on the vertical axis.

Role of weighting exponent, η: The weight w_d is iteratively calculated in the proposed uW-IFCM algorithm (see Eq. (50) in Theorem 2.6). Then, the algorithm assigns a weight w_d^η to the feature $\hat{x}_{id}$, where $1 \leq d \leq D$. Here, the feature obtaining the minimum sum of cluster distances is allocated a larger weight in comparison to other features. This weight assignment approach involves a weighing exponent η, so it distinguishes between relevant and irrelevant features. The weight w_d^η helps in the selection of relevant features. The set of positive real numbers ($\eta > 0$) is the domain of weighing exponent η.

We pictorially describe and analyze each feature of Synthetic Dataset I independently in Fig. 1. In Sect. 4.1, we have already discussed about relevant features d_1, d_2, d_3, and irrelevant/noisy features d_4, d_5. The inseparable clusters are obtained in the presence of noisy features d_4 and d_5 (see Fig. 1). Here, we interrelate an appropriate selection of η and fuzzy factor m with good clustering. Let us study the domain of η in two parts, $\eta \in (0, 1)$ and $\eta > 1$. Total of eight values of η are used for the experimentation. It is reasonable to select the domain equal to [1, 4] for m (see [45]). The experimental analysis is done in two parts:

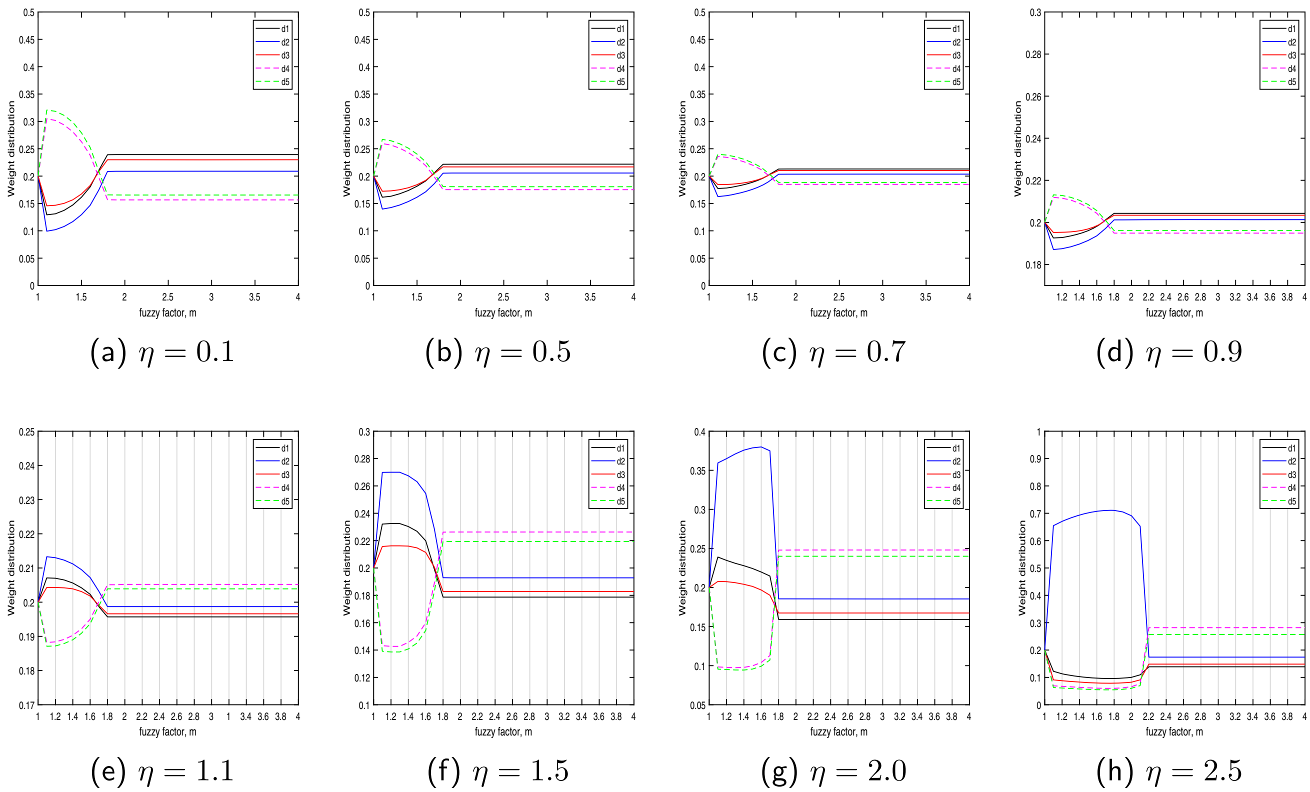


Fig. 2 Distribution of feature weights using uW-IFCM clustering algorithm over Synthetic Dataset I

1. *A case study showing impact of weighing exponent η on feature weight distribution*:

We implement uW-IFCM algorithm over Synthetic Dataset I for the study. The clustering depends on parameters η and m (see Fig. 2). Here, uW-IFCM has initialized under the consideration that all features are equally likely. If $\eta = 0.1$, then w_d^η allocates a weight distribution to the first three features that are very different to the distribution assigned to the last two features, provided m is suitably chosen. If $m < 1.8$, then $w_d^{0.1}$ assigns high weightages to d_4 and d_5, that is, $w_{d_4}^{0.1}, w_{d_5}^{0.1} > > 0.2$. Moreover, $w_d^{0.1}$ assigns weightages to the features d_1, d_2, d_3 of value lesser than 0.2 (see Fig. 2a). Now, if $m \geq 1.8$, then w_d^η assigns higher weights to relevant features d_1, d_2 and d_3 in comparison to noisy features d_4 and d_5. Here, we do not observe many changes in the weight distribution on changing η, where $\eta \in \{0.5, 0.7, 0.9\}$. From this discussion, we conclude that $m \geq 1.8$ and $0 < \eta < 1$ is the domain of uW-IFCM algorithm.

In Table 7, the initial centroids and five randomly generated weight distributions initializing the algorithm are given. We have $\eta \in \{0.1, 0.3, 0.5, 0.7, 0.9, 1.5, 2.0, 2.5, 5.0, 8.0, 10\}$, now the best performance of the algorithm is adjudged after experimenting upon all the values of η. Let $\eta = 1.1$ and $m < 1.8$. In this domain, w_d^η assigns high weights to d_1, d_2, d_3 in comparison to noisy features d_4 and d_5, here a good weight distribution operationalizes uW-IFCM. We vary η over the set $\{1.5, 2, 2.5\}$ while fixing $m < 1.8$, the distributions obtained slightly differ with the distribution at $\eta = 1.1$. If $m \geq 1.8$ and $\eta > 1$, then noisy features d_4 and d_5 are assigned high weights due to virtue of w_d^η. If $\eta > 2$, then the second feature d_2 is assigned a very high weightage in comparison to the remaining features. Here, corresponding to each η, we select an optimal m and the algorithm results in efficient clustering using the optimal (η, m) (see Table 8). It is concluded that for $\eta < 1$, we have $m \geq 1.8$, and $\eta > 1$ implies $m < 1.8$. The complete performance details of the uW-IFCM are provided in Table 7. Here, the two benchmark indexes CA and rand index R_I verify the performance of the uW-IFCM algorithm. The large deviations in the relevant feature weights and irrelevant feature weights are observed at $\eta = 2$. So, there are bright chances that the algorithm delivers an efficient performance at $\eta = 2$, and this case is analyzed separately in detail.

2. *A case study showing connection between the weighing exponent $\eta = 2$ and optimal feature weight distribution:*

Here, the role of $\eta = 2$ in obtaining efficient clustering is discussed. We randomly select five initial cluster centroids (see Table 7) and five initial weight distributions (see w_o rows in Table 9) for the experimentation. Here, all the

Table 7 Tuning of weighing exponent η with fuzzy factor m in uW-IFCM over Synthetic Dataset I

	Initial weights					Final weights					Optimal parameters			
	d_1	d_2	d_3	d_4	d_5	d_1	d_2	d_3	d_4	d_5	CA	R_I	η	m
Initial weight 1	0.3561	0.2646	0.0513	0.0779	0.2500	0.2218	0.2056	0.2168	0.1751	0.1806	0.9900	0.9868	0.5	2
						0.2542	**0.3451**	**0.2144**	**0.0960**	**0.0904**	**1.0000**	**0.9999**	**2**	**1.1**
						0.9740	0.0138	0.0042	0.0046	0.0035	0.8500	0.8281	5	1.1
Initial weight 2	0.1254	0.2768	0.2298	0.2007	0.1673	0.2216	0.2053	0.2169	0.1753	0.1808	0.9866	0.9824	0.5	2.1
						0.2537	**0.3451**	**0.2148**	**0.0960**	**0.0904**	**0.9900**	**0.9868**	**2**	**1.1**
						0.0076	0.9835	0.0046	0.0024	0.0019	0.9033	0.8830	5	1.1
Initial weight 3	0.2976	0.0363	0.3597	0.2005	0.1059	0.2218	0.2056	0.2168	0.1751	0.1806	0.9833	0.9781	0.5	2.1
						0.2537	**0.3451**	**0.2149**	**0.0960**	**0.0904**	**0.9900**	**0.9868**	**2**	**1.1**
						0.0012	0.0056	0.9865	0.0037	0.0029	0.7566	0.7431	5	1.1
Initial weight 4	0.3480	0.2276	0.3384	0.0415	0.0446	0.2216	0.2053	0.2169	0.1753	0.1808	0.9866	0.9824	0.5	2.1
						0.2537	**0.3451**	**0.2148**	**0.0960**	**0.0904**	**0.9900**	**0.9868**	**2**	**1.1**
						0.0012	0.0057	0.9865	0.0037	0.0029	0.7566	0.7431	5	1.1
Initial weight 5	0.1498	0.0214	0.2894	0.3403	0.1990	0.2216	0.2053	0.2169	0.1753	0.1808	0.9866	0.9824	0.5	1.8
						0.2536	**0.3451**	**0.2150**	**0.0960**	**0.0904**	**0.9900**	**0.9868**	**2**	**1.1**
						0.0005	0.0009	0.0006	0.9955	0.0025	0.3933	0.5410	5	1.1

Table 8 Final weights of Synthetic Dataset I with CA and R_I by uW-IFCM clustering algorithm when $\eta = 2$ with random initial weights

	d_1	d_2	d_3	d_4	d_5	(η, m)	CA	R_I
1	0.2218	0.2056	0.2168	0.1752	0.1806	(0.3,2.1)	0.9833	0.9781
2	0.2321	0.2699	0.2157	0.1432	0.1391	(1.5,1.1)	0.9933	0.9911
3	0.2536	0.3451	0.2150	0.0960	0.0904	(2.0,1.1)	0.9900	0.9968
4	0.0001	0.9998	0.0000	0.0000	0.0000	(8.0,1.3)	0.9033	0.8831
5	0.0000	1.0000	0.0000	0.0000	0.0000	(10,1.3)	0.8433	0.8270

Table 9 Distribution of feature weights of Synthetic Dataset I with random weight initialization using uW-IFCM clustering algorithm with $\eta = 2$

		d_1	d_2	d_3	d_4	d_5	CA	R_I
1	w_o	0.0432	0.1387	0.0631	0.5353	0.2197	0.6133	0.6548
	w_f	0.2536	0.3451	0.2149	0.0960	0.0904	0.9900	0.9868
2	w_o	0.0432	0.1387	0.0631	0.2197	0.5353	0.6400	0.6767
	w_f	0.2572	0.3407	0.2165	0.0965	0.0891	0.9933	0.9912
3	w_o	0.0432	0.1387	0.0631	0.2197	0.5353	0.6633	0.6967
	w_f	0.2569	0.3422	0.2177	0.0952	0.0880	0.9933	0.9912
4	w_o	0.2197	0.0432	0.1387	0.5353	0.0631	0.7400	0.7575
	w_f	0.2537	0.3469	0.2157	0.0946	0.0892	0.9900	0.9868
5	w_o	0.1197	0.0432	0.1387	0.3353	0.3631	0.6400	0.9868
	w_f	0.2536	0.3451	0.2149	0.0960	0.0904	0.9900	0.9868

w_o rows are initializing uW-IFCM algorithm one by one, still, the output remains the same because in all the five instances, the point of convergence is not changing (see Fig. 7b). Let us deal with the Synthetic Dataset I using the equally likely approach in uW-IFCM. The multiple convergence patterns are observed for feature weights on varying η, where $\eta \in \{0.3, 1.5, 2.0, 8.0, 10\}$ (see Fig. 7d). In Table 10, the performance of $\eta = 2$ based uW-IFCM is compared with the IFCM and weighted-IFCM. An optimal feature weight distribution appears at $\eta = 2$, and we see very different distributions at $\eta = 0.5$ and $\eta = 5$ (see Fig. 6a–d). In the case of $\eta = 5$, a high feature weight is

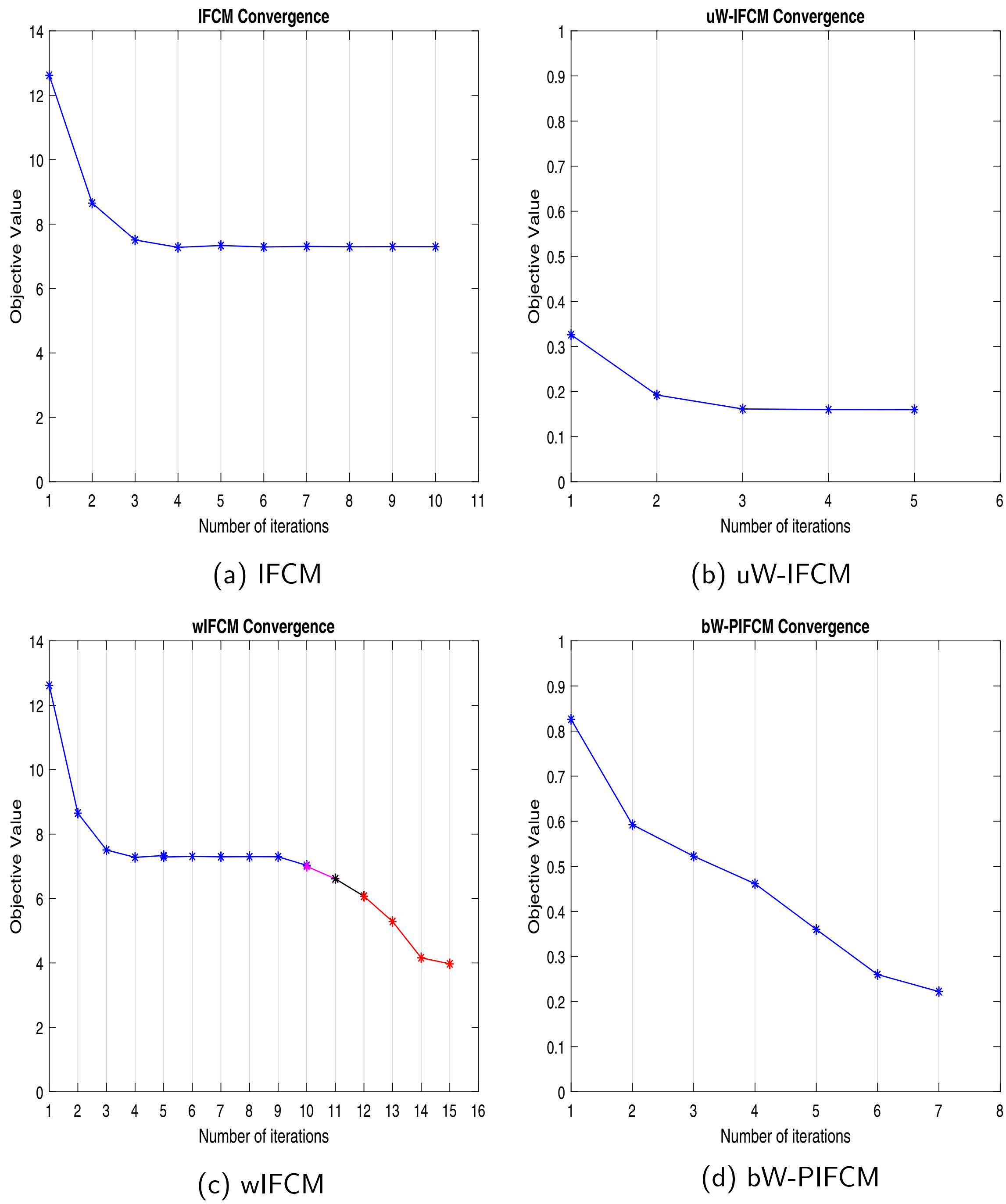


Fig. 3 A study of convergence for **a** IFCM, **b** uW-IFCM, **c** wIFCM, and **d** bW-PIFCM over Synthetic Dataset I

allocated to d_1 in comparison to d_2 and d_3 (see Fig. 6a). In Fig. 6b, the weight allocated to feature d_2 is very high, whereas d_2 and d_3 are assigned very low weights. In other words, the algorithm singles out feature d_2 for the clustering at $\eta = 5$. An optimal weight distribution is obtained using $\eta = 2$, therefore w_d^2 delivers the best clustering accuracy and rand index value also.

4.4 A Comparative Analysis of the Proposed Algorithms with some *C*-Means Algorithms Over UCI Machine Learning Datasets

Here, we explore the functioning of proposed uW-IFCM and bW-PIFCM algorithms on some real-valued UCI machine learning datasets. In addition, the clustering

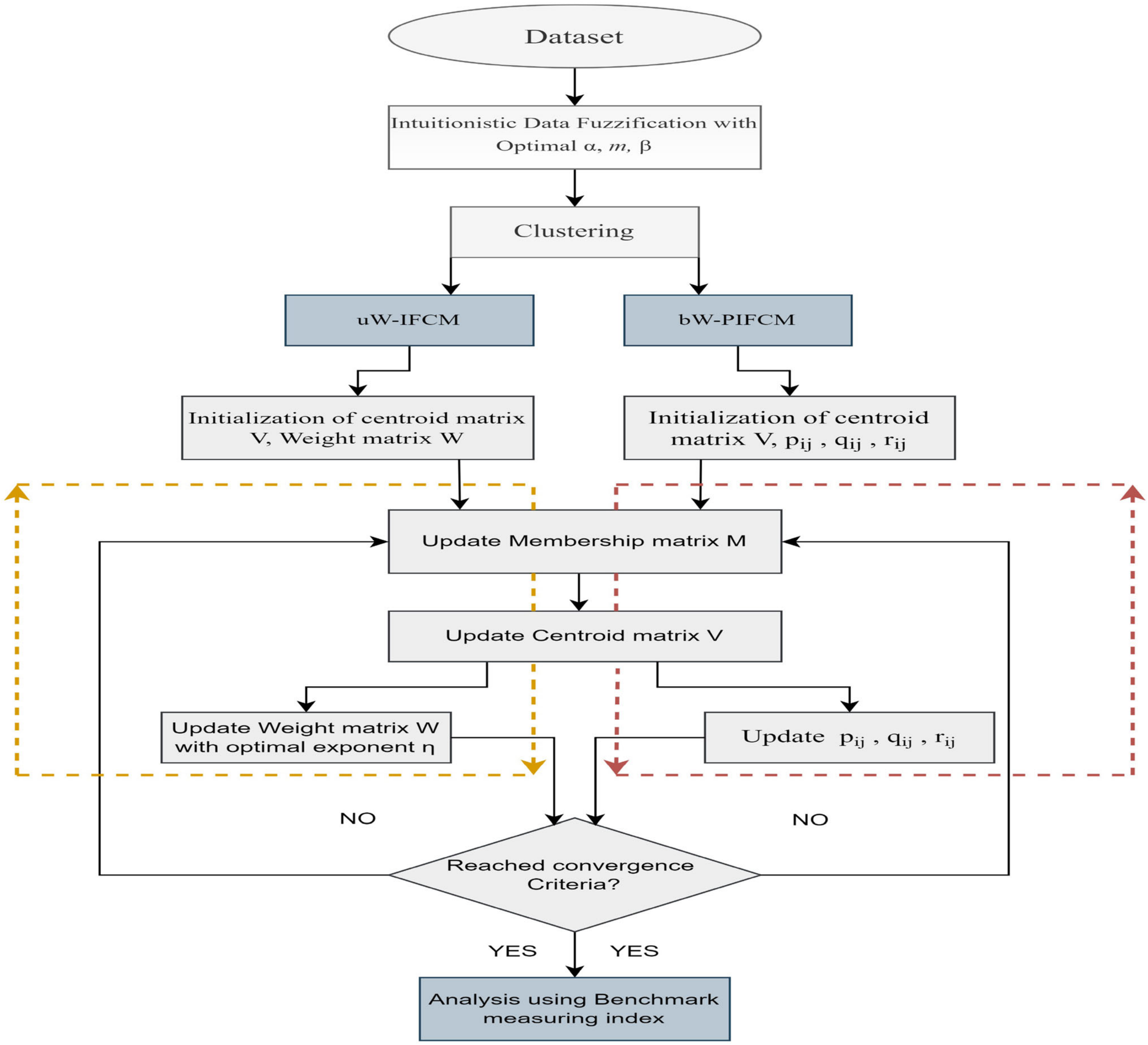


Fig. 4 Proposed Architecture of uW-IFCM and bW-PIFCM algorithms

performance of uW-IFCM is compared with bW-PIFCM, PIFCM, IFCM, and FCM clustering algorithms.

1. UCI Machine Learning Datasets

The seven machine learning datasets, namely, IRIS, thyroid, Bupa, zoo, heart, WDBC, and Ecoli are considered for experimentation (see [47]). Out of these seven datasets, five are disease-based datasets concerning to thyroid, liver, heart, breast, and cancer. IRIS is a plant dataset and zoo belongs to the category of the animal dataset. The description of the datasets is provided in Table 15. MATLAB version 8.1 running on a PC with 3.40 GHz frequency and RAM 16 GB is used for the computational tasks.

2. Some popular benchmark measuring indexes:

Here, the five benchmark measuring indexes, namely, clustering accuracy (*CA*), rand index (R_I), partition coefficient (*PC*), partition index (*PI*), and Dunn index (*DI*) are used to compare the performances of the algorithms. Now, we define the indexes as follows:

(1) Clustering accuracy (*CA*) [48]: It is the ratio of the number of correctly classified elements p_c and a total number of elements p, that is, $CA = \frac{p_c}{p}$. It is a popular index that works over labeled datasets.

(2) Rand index (R_I) [44]: Rand index is also used for labeled datasets. Let $\mathbb{C}$ be the set of clusters being provided in the original dataset. The cluster set resulting after the

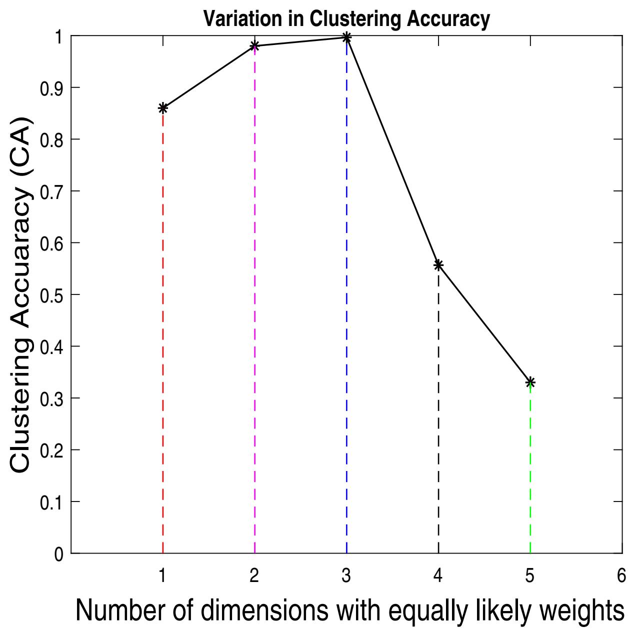


Fig. 5 Impact of equally likely approach on clustering Synthetic Dataset I by wighted-IFCM

clustering of the dataset is $\mathbb{C}'$. The elements in $\mathbb{C} \times \mathbb{C}'$ are categorized as follows: (1) *SD*, (2) *DS*, (3) *SS*, (4) *DD* (here *S* stands for same cluster and *D* stands for different cluster). Mathematically, we have,

$$R_I = \frac{t_1 + t_4}{R} \tag{63}$$

, where t_1 = number of *SS*, t_2 = number of *DS*, t_3 = number of *SD* and t_4 = number of *DD* with $R = a + b + c + d$.

(3) Partition coefficient (*PC*) [49]: It is the mean of the summation of the square of membership value of *i*th element in *j*th cluster, $1 \leq j \leq C$. Mathematically, $PC = \frac{1}{P}\sum_{i=1}^{P}\sum_{j=1}^{C} u_{ij}^2$, where *P* is the number of elements in the dataset. It informs about the trade-off between the clusters. If *PC* limits to unity, then the number of elements appearing in common clusters is very few. Therefore, higher the value of $PC \in [\frac{1}{C}, 1]$ (where *C* is the number of clusters) results in better clustering performance.

(4) Partition index (*SC*) [50]: The ratio between the compactness of a cluster and its separation from other clusters calculates the partition index. Mathematically,

$$SC = \frac{\sum_{i=1}^{P}\sum_{j=1}^{C} u_{ij} d^2(x_k, v_i)}{\sum_{k=1}^{P} u_{ik} \sum_{t=1}^{C} d^2(v_i, v_t)}. \tag{64}$$

A lower value of *SC* means better clustering.

(5) Dunn index (*DI*) [51]: It is a well-established hard cluster validity index. Dunn defines the clusters to be "compact, separate (CS)" relative to the metric iff the following condition is satisfied: for all p, q, r, where $p \neq q$, any pair of elements $u, v \in A_r$ are closer together (with respect to metric), then any pair of elements $x \in A_p$ and $y \in A_q$. Let $A_1, A_2, \ldots, A_C$ be *C*-partition of *A* and let *U* be the partition matrix. Dunn index is defined as follows:

$$DI = \min_{1 \leq i \leq C}\left\{\min_{i+1 \leq j \leq C-1}\left\{\frac{dis(u_i, u_j)}{\max\limits_{1 \leq c \leq k} dia(u_k)}\right\}\right\} \tag{65}$$

$$\text{where} \quad dis(u_i, u_j) = \min_{A_i \in u_i, A_j \in u_j} d(A_i, A_j), \tag{66}$$

$$dia(u_k) = \min_{A_i, A_j \in u_k} d(A_i, A_j) \tag{67}$$

$$DI = \min_{1 \leq i \leq C}\left\{\min_{i+1 \leq j \leq C-1}\left\{\frac{dis(u_i, u_j)}{\max\limits_{1 \leq c \leq k} dia(u_k)}\right\}\right\} \tag{68}$$

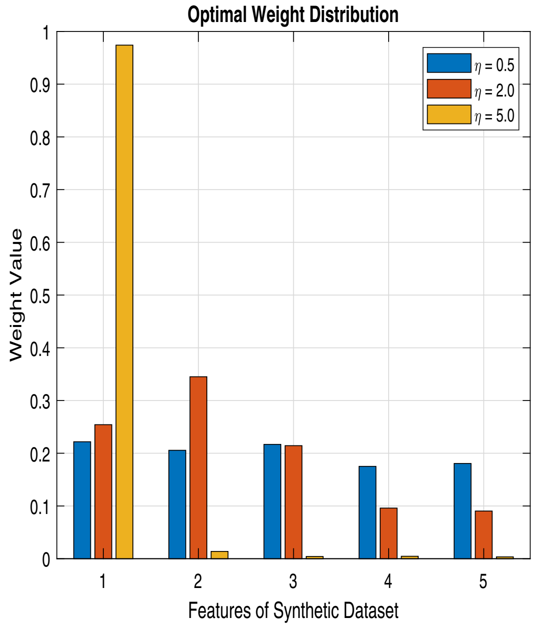
$$\text{where} \quad dis(u_i, u_j) = \min_{A_i \in u_i, A_j \in u_j} d(A_i, A_j), \tag{69}$$

$$dia(u_k) = \min_{A_i, A_j \in u_k} d(A_i, A_j). \tag{70}$$

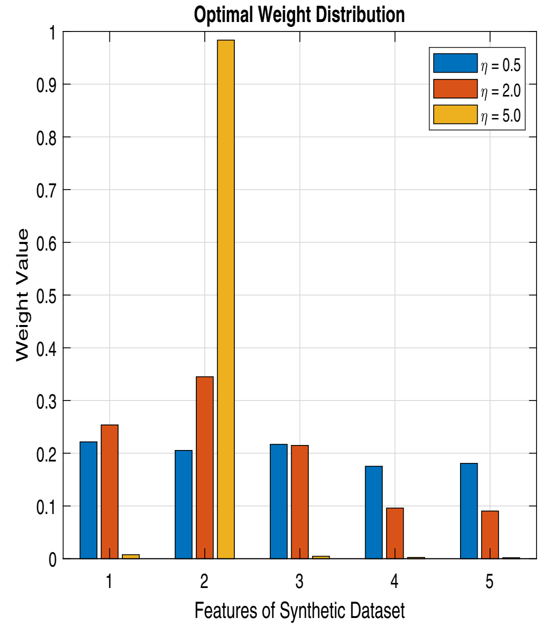
3. Experimental analysis of UCI machine learning datasets: In Table 17, we record three types of values for *CA*, R_I, *PC*, *SC*, *DI*, namely, worst, average, and best. The circular/spherical path-based algorithms, such as uW-IFCM, wIFCM, IFCM, wFCM, FCM, and elliptic/ellipsoidal path-based data-driven bW-PIFCM, PIFCM algorithms are implemented over the seven datasets (see Table 11-Table 14). In IFCM and FCM algorithms, the equally likely approach computes the weight distributions. The weight distributions in fixed-weighted data-driven algorithms, namely wIFCM and wFCM are calculated using Eq. (50). The uW-IFCM algorithm further tunes the weight distribution of wIFCM with the help of its weighing exponent η. Here, the values that frequently repeat themselves during 100 random initializations of the algorithms are termed as average values. The results of Table 17 are used for the following experimental analysis:

(1) The best *CA* obtained by uW-IFCM differs from its worst with a minimal value of 0.033 over the IRIS dataset, whereas the maximal deviation equal to 0.118 is observed on the zoo dataset. Hence, we find a *CA*-deviation interval [0.033, 0.118] for uW-IFCM. The *CA*-deviation intervals [0.032, 0.212], [0.007, 0.393], [0.033, 0.196], [0.031, 0.253] are obtained using wIFCM, IFCM, FCM, and wFCM, respectively. Since, in IFCM, the underlined requirements of the problem and weights are not correlated, therefore IFCM has the largest deviation interval. In uW-IFCM, η acts as a controlling parameter, so it deduces highly stable weight distributions (in other words, the deviation interval is least).

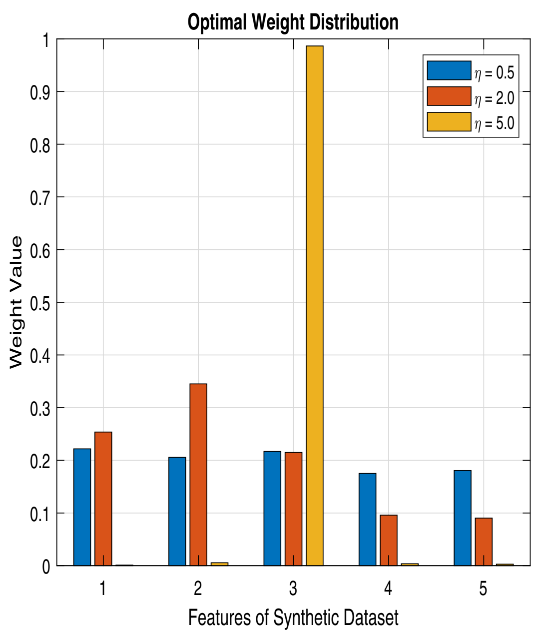
(2) In a similar fashion, the R_I-deviation intervals [0.007, 0.124], [0.002, 0.195], [0.008, 0.166], [0.004, 0.322], and [0.005, 0.317] are obtained using uW-IFCM, wIFCM, IFCM, FCM, and wFCM,



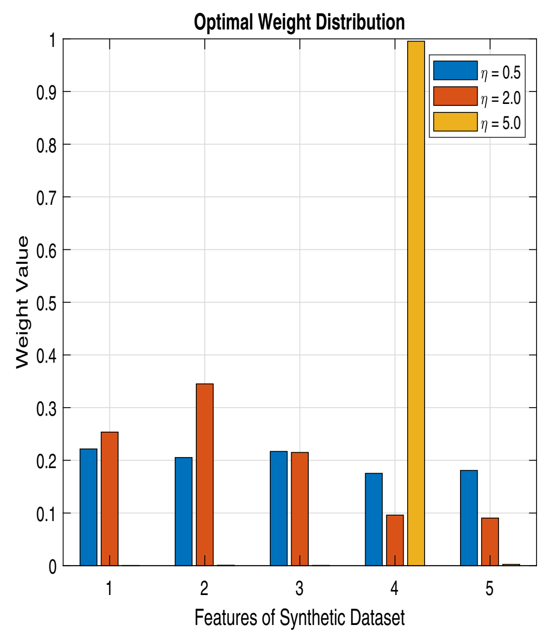
(a) Optimal distribution at $\eta = 2$



(b) Optimal distribution at $\eta = 2$



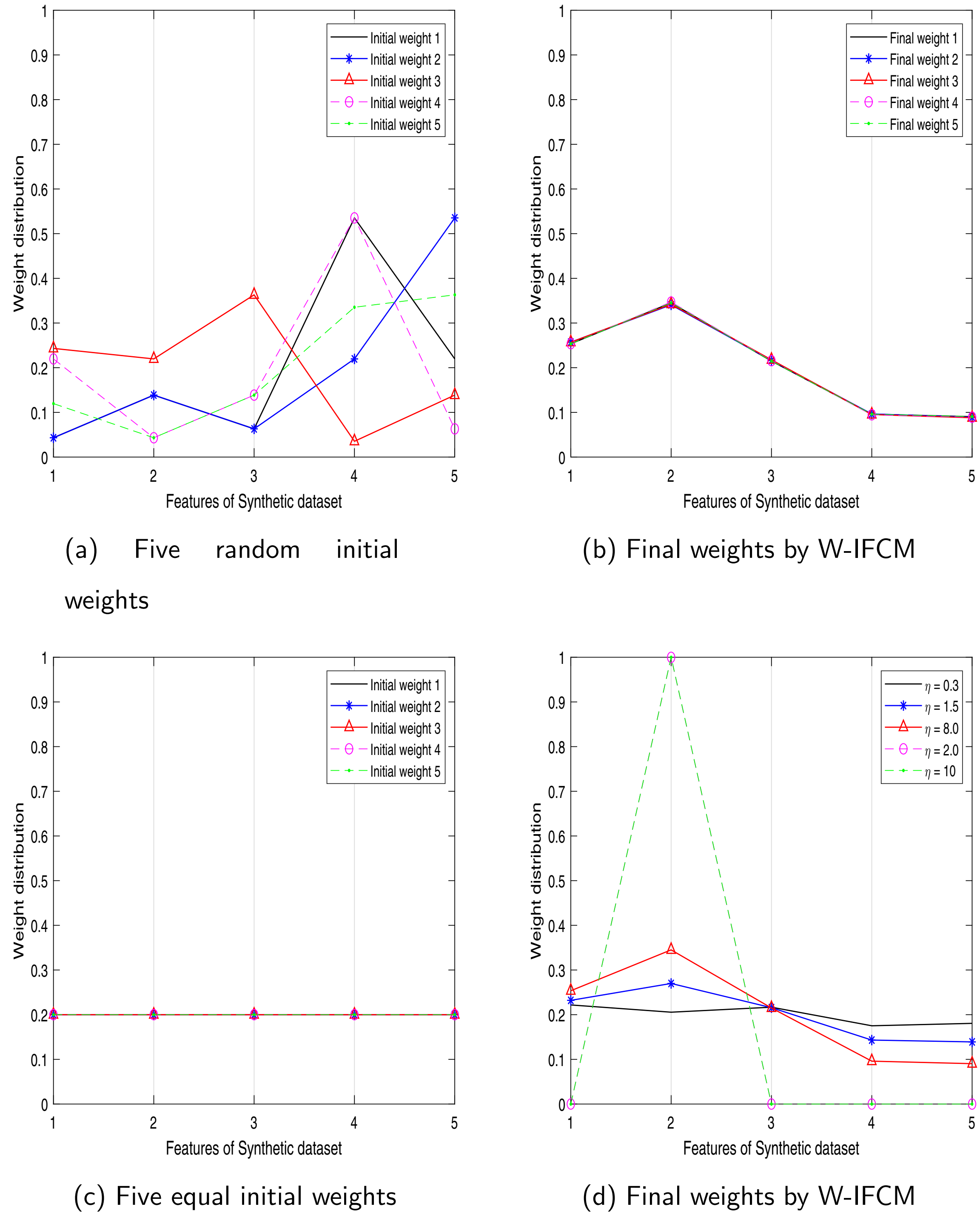
(c) Optimal distribution at $\eta = 2$



(d) Optimal distribution at $\eta = 2$

Fig. 6 Feature weight distribution over η and m for Synthetic Dataset I such that $m \geq 1.8$ for $\eta \in (0, 1)$ and $m < 1.8$ for $\eta > 1$

respectively. Since, in FCM also the weights and underlined requirement of the problem are not correlated, therefore FCM yields the largest deviation interval. The uW-IFCM uses a controlling parameter η-based weight distribution, and hence highly stable weights are obtained (in other words, the deviation interval is least). For all the algorithms, we have shown the patterns of R_I-deviations in Fig. 8.



(a) Five random initial weights

(b) Final weights by W-IFCM

(c) Five equal initial weights

(d) Final weights by W-IFCM

Fig. 7 **a** Depicting five random initial weights for the five features of Synthetic Dataset I. **b** The five final weights coinciding to same values when $\eta = 2$. **c** Five initial weights using the equally likely approach for Synthetic Dataset I (d) Graphical representation of variation in the distribution of final weights for Synthetic Dataset I with varying weighing exponent η in uW-IFCM

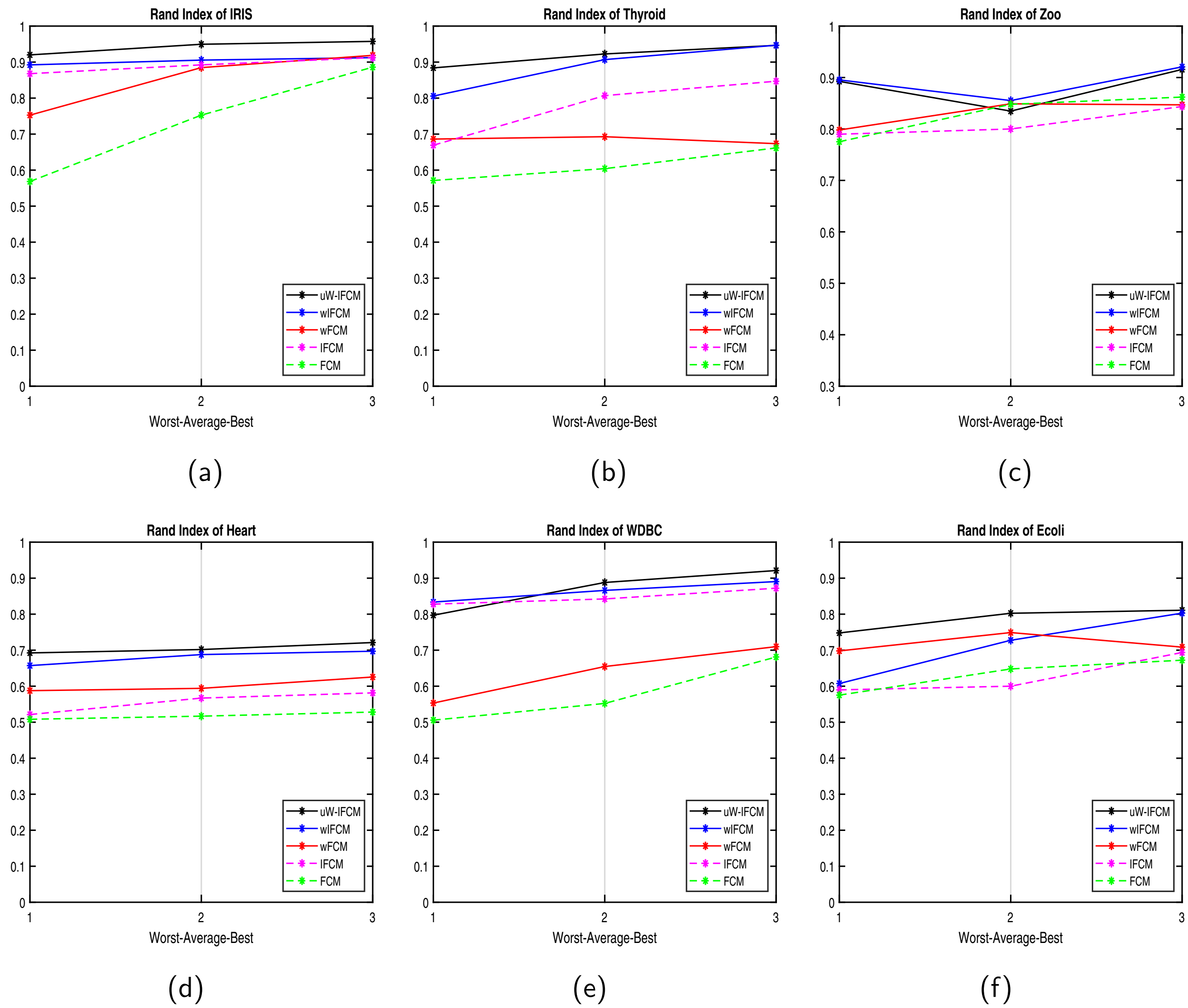


Fig. 8 Comparison of Rand Index, R_I over real datasets with uW-IFCM, wIFCM, wFCM, IFCM, and FCM

(3) Similar type of analysis can be done for the *PC*, *SC*, and *DI* indexes.

(4) Here, every algorithm shows its best and worst performances on IRIS and Bupa datasets, respectively.

(5) PIFCM is an adaptive algorithm (see [41]), and its clustering is improvised by using data-driven bW-PIFCM. Further, bW-PIFCM yields clustering better than wIFCM, IFCM, wFCM, FCM algorithms except for uW-IFCM. Hence, bW-PIFCM is also an adaptive algorithm.

(6) The worst, average, and best column of every benchmark measuring indexes clearly indicate that uW-IFCM outperforms bW-PIFCM, PIFCM, wIFCM, IFCM, wFCM, and FCM algorithms (see Table 17).

(7) The wFCM and wIFCM are data-driven non-adaptive fixed-weighted algorithms. These fixed weights are derived by Eq. (50). The weighing parameter tunes the weight provided in Eq. (50) during the implementation of uW-IFCM.

5 Feature Reduction Technique

We carry out this study in two sections. In the first section, the uW-IFCM algorithm is introduced as a mechanized feature reduction technique. The next section proposes automatized feature reduction technique called FRT-equipped uW-IFCM algorithm. Here, we try to establish a connection between computational cost and feature reduction.

Table 10 Comparison of clustering performance between IFCM, weighted-IFCM and proposed uW-IFCM with *CA* and R_I when $\eta = 2$

Sr. No	IFCM		weighted-IFCM		uW-IFCM	
	CA	R_I	CA	R_I	CA	R_I
1	0.3800	0.5300	0.6400	0.6800	0.9000	0.9100
2	0.3633	0.5333	0.7000	0.7255	0.9333	0.9011
3	0.3700	0.5400	0.7255	0.7000	**1.0000**	**0.9999**
4	0.4000	0.5600	0.4800	0.5400	0.8800	0.8778
5	0.3900	0.5700	0.5800	0.6011	0.8900	0.8856
Average	0.38066	0.54666	0.62510	0.64932	**0.91932**	**0.89408**

Table 11 Clustering results based on bW-PIFCM with fixed initial cluster centroids over IRIS, Thyroid, and Bupa dataset

Dataset	Indexes	bW-PIFCM	Optimal parameters	Running time (seconds)
IRIS	CA	0.9533	$m = 1.1, \alpha = 0.35, \beta = 10$	0.074412
	RI	0.9417		
	PC	0.9793		
	SC	0.2676		
	DI	0.0347		
Thyroid	CA	0.9813	$m = 2.4, \alpha = 0.6, \beta = 1.5$	0.093567
	RI	0.9693		
	PC	0.5033		
	SC	0.7137		
	DI	0.0563		
Bupa	CA	0.6057	$m = 3.5, \alpha = 0.05, \beta = 10$	0.098364
	RI	0.5209		
	PC	0.5000		
	SC	0.3010		
	DI	0.0134		

Table 12 Clustering results based on bW-PIFCM with fixed initial cluster centroids over Zoo, Heart, WDBC, and Ecoli dataset

Dataset	Indexes	bW-PIFCM	Optimal parameters	Running time (seconds)
Zoo	CA	0.8910	$m = 1.4, \alpha = 0.1, \beta = 5$	0.090445
	RI	0.9540		
	PC	0.8135		
	SC	0.2906		
	DI	0.3896		
Heart	CA	0.8370	$m = 1.4, \alpha = 0.2, \beta = 7.5$	0.101622
	RI	0.6924		
	PC	0.6950		
	SC	3.3093		
	DI	0.1701		
WDBC	CA	0.9455	$m = 1.9, \alpha = 0.4, \beta = 1$	0.063069
	RI	0.8967		
	PC	0.6428		
	SC	1.2786		
	DI	0.0750		
Ecoli	CA	0.8654	$m = 1.3, \alpha = 0.4, \beta = 1$	0.121369
	RI	0.9040		
	PC	0.8534		
	SC	0.4478		
	DI	0.0430		

Table 13 Clustering results based on PIFCM with fixed initial cluster centroids over IRIS, Thyroid, and Bupa dataset

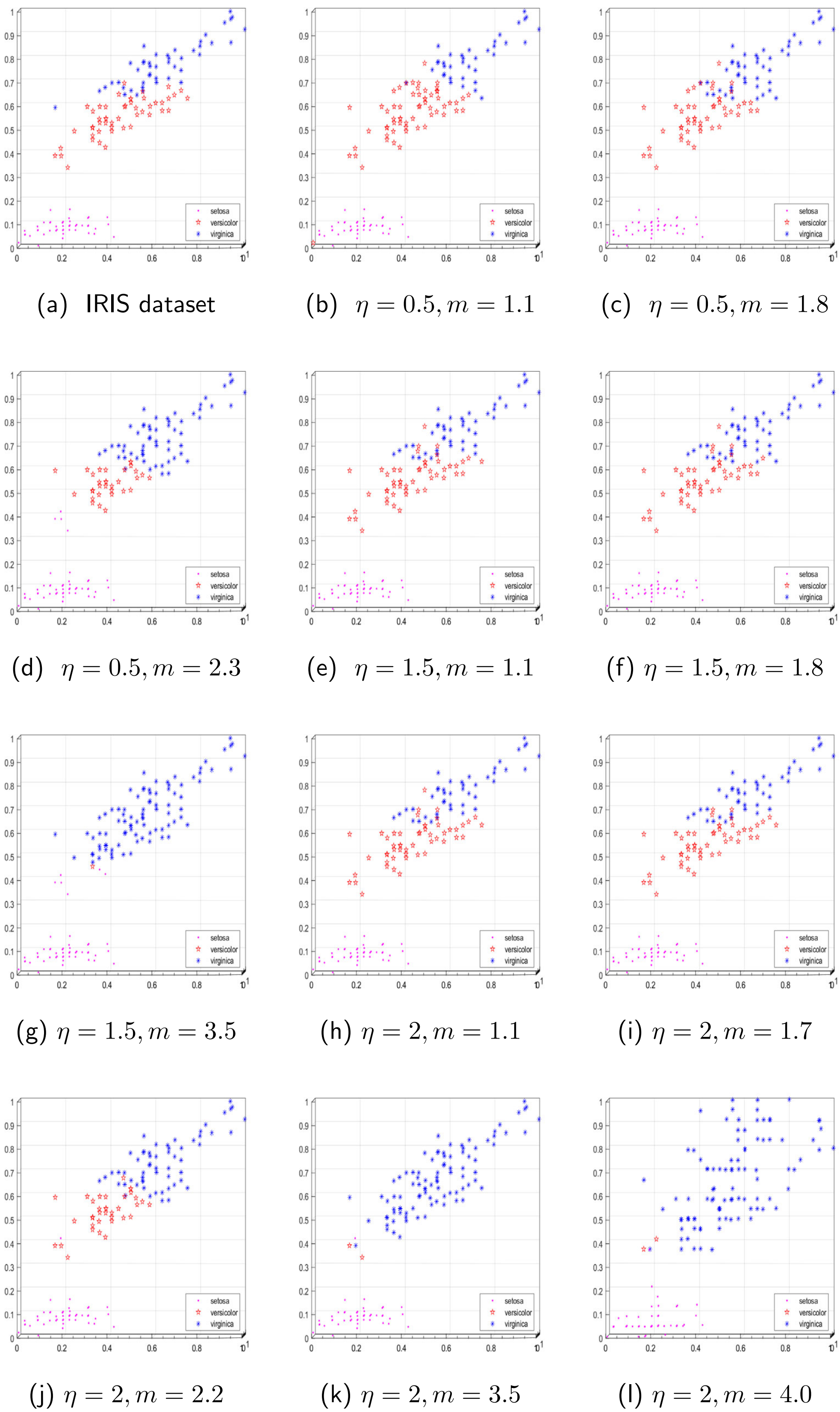
Dataset	Indexes	PIFCM	Optimal parameters	Running time (seconds)
IRIS	CA	0.9333	$m = 3.2, \alpha = 0.40$	0.074607
	RI	0.9195		
	PC	0.4049		
	SC	0.5679		
	DI	0.0701		
Thyroid	CA	0.9767	$m = 2.4, \alpha = 0.75$	0.067314
	RI	0.9616		
	PC	05219		
	SC	0.6283		
	DI	0.0421		
Bupa	CA	0.5797	$m = 2, \alpha = 0.4$	0.109408
	RI	0.5113		
	PC	0.5804		
	SC	0.2039		
	DI	0.0354		

Table 14 Clustering results based on PIFCM with fixed initial cluster centroids over Zoo, Heart, WDBC, and Ecoli dataset

Dataset	Indexes	PIFCM	Optimal parameters	Running time (seconds)
Zoo	CA	0.8712	$m = 1.2, \alpha = 0.15$	0.116985
	RI	0.9442		
	PC	0.9603		
	SC	0.2732		
	DI	0.4082		
Heart	CA	0.8111	$m = 1.5, \alpha = 0.4$	0.088025
	RI	0.6924		
	PC	0.6752		
	SC	2.5419		
	DI	0.1832		
WDBC	CA	0.9420	$m = 1.9, \alpha = 0.35$	0.64306
	RI	0.8905		
	PC	0.6752		
	SC	1.2302		
	DI	0.0759		
Ecoli	CA	0.84440	$m = 1.3, \alpha = 0.2$	0.091093
	RI	0.9003		
	PC	0.8501		
	SC	0.5188		
	DI	0.0425		

5.1 A discussion of uW-IFCM clustering algorithm as a feature reduction technique:

Here, we analyze IRIS dataset deeply with the help of the uW-IFCM algorithm. The dataset consists of four features $\{d_1, d_2, d_3, d_4\}$, namely, sepal length, sepal width, petal length, and petal width, respectively. The 150 instances present in the dataset are divided into three classes: setosa, virginica, versicolor, and each class contains 50 instances. We use twenty feature weight distributions to initialize the algorithm. Here, the first ten feature weight distributions are calculated using the equally likely approach, whereas the next ten feature weight distributions are randomly generated with the help of `rand` function of MATLAB (see Table 19). The initial cluster centroids given in Sect. 4.3 are used here also. The importance of weighing exponent



(a) IRIS dataset (b) $\eta = 0.5, m = 1.1$ (c) $\eta = 0.5, m = 1.8$

(d) $\eta = 0.5, m = 2.3$ (e) $\eta = 1.5, m = 1.1$ (f) $\eta = 1.5, m = 1.8$

(g) $\eta = 1.5, m = 3.5$ (h) $\eta = 2, m = 1.1$ (i) $\eta = 2, m = 1.7$

(j) $\eta = 2, m = 2.2$ (k) $\eta = 2, m = 3.5$ (l) $\eta = 2, m = 4.0$

Fig. 9 Study of clustering of IRIS by uW-IFCM with variation in η and m

$\eta = 2$ in uW-IFCM is highlighted in Sect. 4.4. Let us use $\eta = 2$, $m = 1.7$, and equally likely approach-based initial weights (see rows 1–10 of Table 19) to initialize the uW-IFCM, the algorithm assigns higher weights to the third feature and fourth feature in comparison to other two features. Further, a good clustering accuracy $CA = 0.9867$, rand index $RI = 0.9875$, $PC = 0.9205$ are obtained with $SC = 0.00007$ and $DI = 0.05123$. These values validate that initial weight distribution converges to optimal feature weight distribution. Now, we initialize the uW-IFCM algorithm using initial weight distribution (0.3119, 0.2371, 0.3183, 0.1327), $\eta = 2$, and $m = 1.7$, its result is an optimal weight distribution (0.1622, 0.1661, 0.3527, 0.3191) (See Table 19). The experimental findings confirm that uW-IFCM yields optimal weight distribution over the IRIS dataset at $\eta = 2$ (see Fig. 10a, e, f). The algorithm assigns higher weights to the third and fourth features in comparison to the first two features, while using $\eta = 2$ and $m = 1.1$. A graphical comparison of feature weight distribution obtained over the IRIS dataset is also presented (see Fig. 10). At $\eta = 0.5$ and $\eta = 1.1$, uW-IFCM gives almost equal weightage to all the four features (see Fig. 10b, c). The clustering completely depends on the third feature whenever $\eta > 2.5$ is explored (see Fig. 10g–j). Here, convergence is achieved at the ninth iteration. Further, if $\eta > 2$, the clustering performance of uW-IFCM deteriorates which can be improvised by tuning η with m. The domain of m depends on η (see Sect. 4.4). Therefore, from an η-dependent domain of m, we have selected some values for the experimentation (see Table 19). The clustering results obtained using different $\eta > 2$ and $m = 1.1$ are recorded in Table 19, we exploit $m = 1.1$ as it suitably tunes the η. During the clustering of the IRIS dataset, we obtain higher relevancy for the third and fourth features in comparison to the remaining features. In Table 20, the relevancy of the features of the IRIS dataset is given.

Summary of this discussion: During the clustering of the IRIS dataset, the uW-IFCM allocates more weightage to features d_3, d_4 in comparison to d_1, d_2. Here, due to the absence of proper feature nullification criteria, the less relevant features are mechanically neglected. So, uW-IFCM mechanically observes that d_1, d_2 are irrelevant features. This algorithm can be used as a feature reduction technique. The uW-IFCM results in an efficient clustering of two featured versions of the IRIS dataset also. The algorithm converges in nine iterations while clustering the IRIS dataset, whereas the two featured versions of the IRIS dataset are clustered in five iterations (see Fig. 11). The total computational time taken by uW-IFCM over IRIS dataset and its two featured versions is 0.1485 and 0.1118, respectively.

5.2 Performance of the FRT-equipped uW-IFCM algorithm

Here, a feature reduction technique (FRT) is incorporated in the uW-IFCM algorithm for the automatic elimination of irrelevant features. The computational cost of the algorithm is low. In the FRT-equipped uW-IFCM, a collective threshold is decided for all features such that the weights equal and above the threshold value are accepted and weights having values lower than the threshold gets rejected. The dataset contains D numbers of features/dimensions, so for the removal of irrelevant features the threshold value equal to $1/(\sqrt{D})^{\chi}, \chi > 0$ is selected for computational work. The features having weight values lower than $1/(\sqrt{D})^{\chi}$ are not useful for clustering. Higher the number of dimensions/features, the lower the value of the threshold and vice-versa. If D is large, then $1/(\sqrt{D})^{\chi}$ becomes small, and in this case, the underlying iterative process of the algorithm converges after discarding low relevancy features. The small value of D results in high threshold value, and hence in this situation the features are rapidly eliminated. This threshold selection works well for all finitely valued D. In this section, we worked on two synthetic datasets of two dimensions and one six-dimensional synthetic dataset. Please refer to Table 16 for synthetic datasets used in the following examples.

Algorithm V:	FRT-equipped uW-IFCM algorithm
Input:	Dataset (X), number of centroids (C), weighing exponent (η), fuzzy factor (m), intuitionistic fuzzy parameter (α), tuning parameter for membership function (β), tolerance level (ϵ)
Output:	Fuzzy partition U, centroids $\{(\bar{\mu}_{ld}, \bar{\nu}_{ld}, \bar{\pi}_{ld})\}_{d=1}^{D}$, Weight matrix W
Procedure:	**1**. 1. Data fuzzification using Algorithm I.
	2. 2. Initialize centroid $\hat{V}$, U at $t = 0$.
	3. 3. Initialize weight matrix W with $\frac{1}{D}$'s.
	Repeat
	4. Update $(U = u_{il})^{t+1}$ by calculating the fuzzy partition using Eq. (41).
	5. Update centroid $\{(\bar{\mu}_{ld}, \bar{\nu}_{ld}, \bar{\pi}_{ld})\}_{d=1}^{D}$ using Eqns. (55-57).
	6. Update weight matrix, W using Eq. (50).
	7. 7. Optimize weight matrix with threshold value $1/(\sqrt{D})^{\chi}$.
	Until
	8. $\sum_{i=1}^{C} \frac{d_2(W_i(t), W_i(t+1))}{D} < \epsilon$ is satisfied.
Return	U^{t+1}, $(\bar{\mu}_l^{t+1}, \bar{\nu}_l^{t+1}, \bar{\pi}_l^{t+1})$ and W^{t+1}.

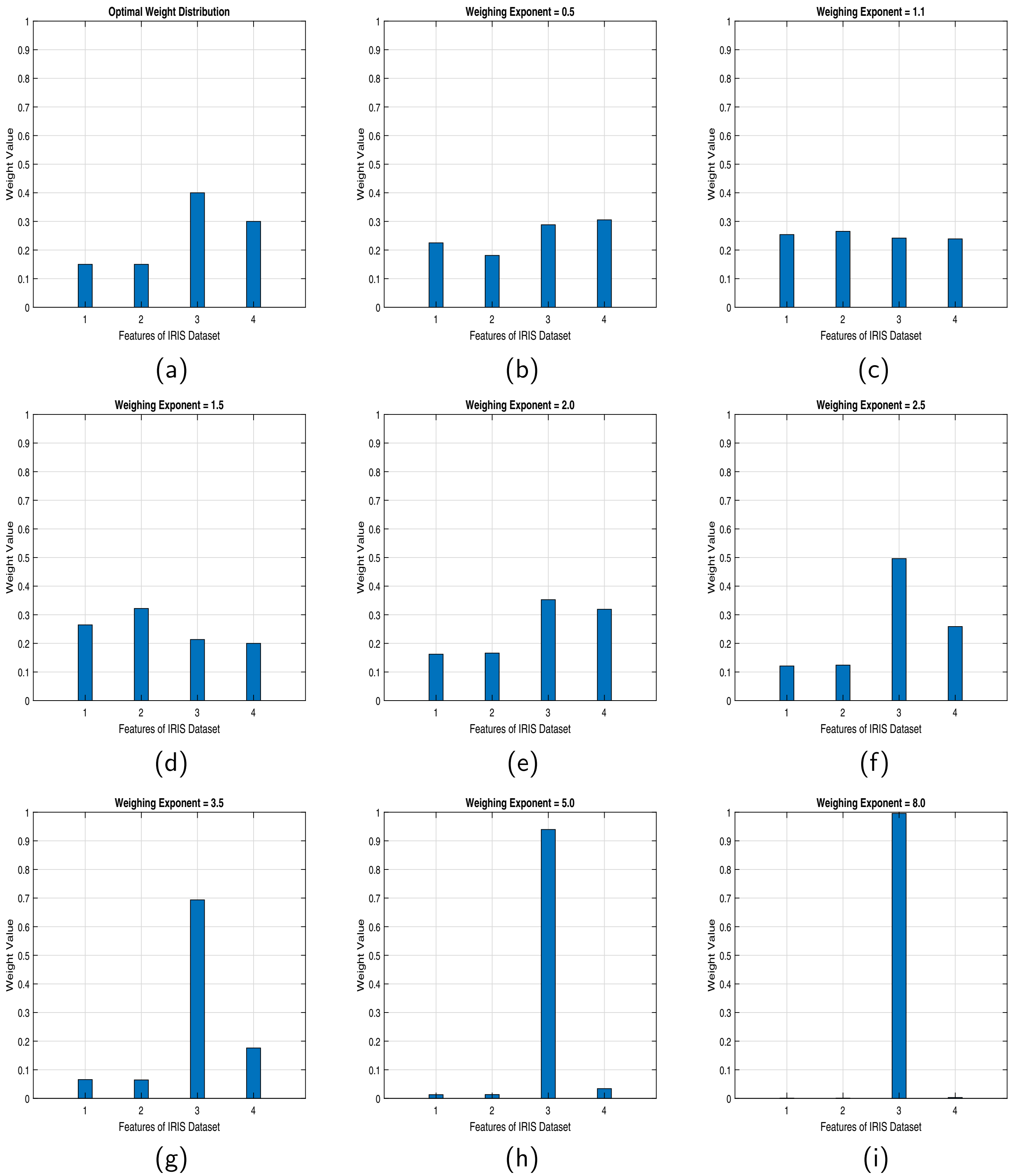


Fig. 10 Weight distribution for IRIS dataset among its features using uW-IFCM

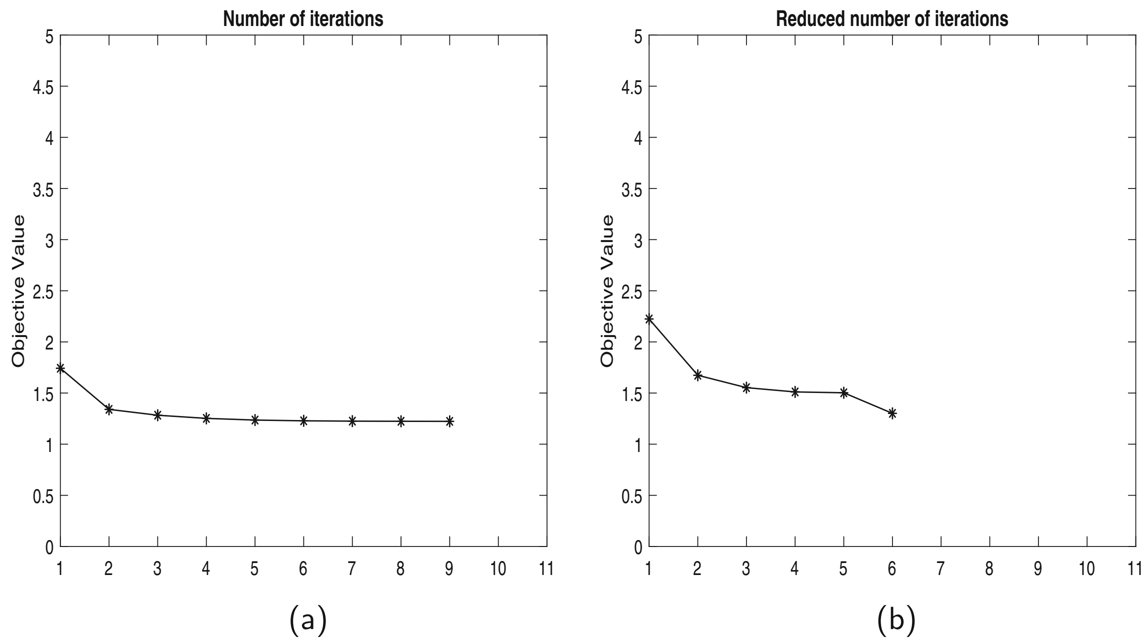


Fig. 11 Plots to show a reduced number of iterations for IRIS dataset clustering using uW-IFCM

Example 1 We generate a two-dimensional Synthetic Dataset II of sample size 1000 using a Gaussian distribution function, $\sum_{j=1}^{3}\sum_{l=1}^{2} q_k N(\theta_{jl}, \sigma_{jl})$. Here, we denote the mean and the standard deviation with θ and σ, respectively. The mean value of the jth cluster along lth dimension/feature is given as θ_{jl} with standard deviation σ_{jl}. Synthetic Dataset II consists of three well-separated clusters. Here, the parameter $q_k = 1/3, \forall 1 \le k \le 3$. The values of the mean and standard deviation of the three clusters are given as $\theta_{j1} = (0\ 2)^T, \theta_{j2} = (4\ 7)^T, \theta_{j3} = (8\ 12)^T$ and $\sigma_{j1} = \sigma_{j2} = \begin{pmatrix} 10 & 0.01 \\ 0.01 & 0.01 \end{pmatrix}$ and $\sigma_{j3} = \begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix}$.

The threshold limit $1/(\sqrt{D})^{\chi}$ of FRT-equipped uW-IFCM algorithm nullifies the first feature because the relevancy of the first feature is quite less in comparison to the second feature. Now, uW-IFCM computes the CA and R_I values using only the second feature. The clustering accuracy $CA = 0.9950$ and rand index $R_I = 0.9960$ are calculated over Synthetic Dataset II. Here, with reduced computational cost, we obtain similar results. These criteria easily eliminate irrelevant features from the dataset.

Example 2 We generate a two-dimensional/featured Synthetic Dataset III of sample size 1000 using a Gaussian distribution function, $\sum_{j=1}^{4}\sum_{l=1}^{2} q_k N(\theta_{jl}, \sigma_{jl})$. Here, we denote the mean and the standard deviation with θ and σ, respectively. The mean value of the jth cluster along lth dimension/feature is given as θ_{jl} with standard deviation σ_{jl}. The Synthetic Dataset III consists of four well-separated clusters. Here, the parameter $q_k = 1/4, \forall 1 \ \le k \le 4$. The values of the mean and standard deviation of the four clusters are given as $\theta_{j1} = (10\ 5)^T$, $\theta_{j2} = (10\ 10)^T, \theta_{j3} = (10\ 15)^T, \theta_{j4} = (10\ 20)^T$ and $\sigma_{jk} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \forall 1 \le k \le 4$.

We have transformed two-dimensional/featured dataset into a three-dimensional/featured dataset using the mapping $\chi(x, y) = (x\cos x, y, x\sin x)$. A two-dimensional cross-sectional view of the clustering is shown in Fig. 14, and in this case, we obtain clustering accuracy $CA = 0.5888$ and rand index $R_I = 0.4999$. The threshold limit of the dataset is given by $1/(\sqrt{3})^{\chi}$. The incorporation of feature reduction criteria $(1/(\sqrt{3})^{\chi})$ in uW-IFCM excludes the first feature during the clustering, and thus improved $CA = 0.9954$ and $R_I = 0.9960$ are obtained.

Table 15 Summary of UCI datasets

	Dataset	Instances	Features	Classes
1	IRIS	150	4	3
2	Thyroid	215	5	3
3	Bupa	345	6	2
4	Zoo	101	17	7
5	Heart	270	13	2
6	WDBC	569	30	2
7	E.coli	327	7	5

Example 3 Here, we use a two-cluster six-dimensional dataset named Synthetic Dataset IV. The dataset contains two types of data points. Each data point has six features. The Gaussian distribution selects the $1st$, $3rd$, and $5th$ feature values of the first 100 data points. This distribution has two parameters, mean = 3 and variance = 0.5. Using uniform distribution, the $2nd$, $4th$, and $6th$ feature values are assigned in the interval [0,10]. In a similar fashion, the $2nd$, $4th$, and $5th$ feature values of the next 100 data points are selected with the help of Gaussian distribution having mean = 7 and variance = 0.5. Uniform distribution assigns values to the remaining features, i.e., $1st$, $3rd$, and $6th$ in [0,10].

The dataset is constructed such that the 1st, 3rd, and 5th dimensions/features reside in the first cluster, whereas the $2nd$, $4th$, and $6th$ dimensions reside in the second cluster. In such datasets, subspace clustering gives effective results. A detailed subspace view of the dataset is presented (see Fig. 12). The importance of the $5th$ feature in separating two clusters is clearly shown in Fig. 12c and h. The uW-IFCM algorithm also verifies that $5th$ feature is most important as it calculates weight distribution, $W = [0.1215, 0.1102, 0.0934,\ 0.0956, \mathbf{0.5252}, 0.0539]$. The CA and R_I are being recorded as 0.9989 and 0.9954, respectively.

The weight distributions obtained using the uW-IFCM algorithm over different datasets are compared in Fig. 15. The figure clearly shows the differences in the weights of relevant and irrelevant features. Once the number of features is reduced, the running time also decreases (see Table 22). A detailed analysis of the parameter χ is provided in Table 21.

Table 16 Summary of synthetic datasets used

	Dataset	Instances	Features	Classes
1	[44]	300	5	3
2	[21]	1000	2	3
3	[21]	1000	2	3
4	[46]	300	6	2

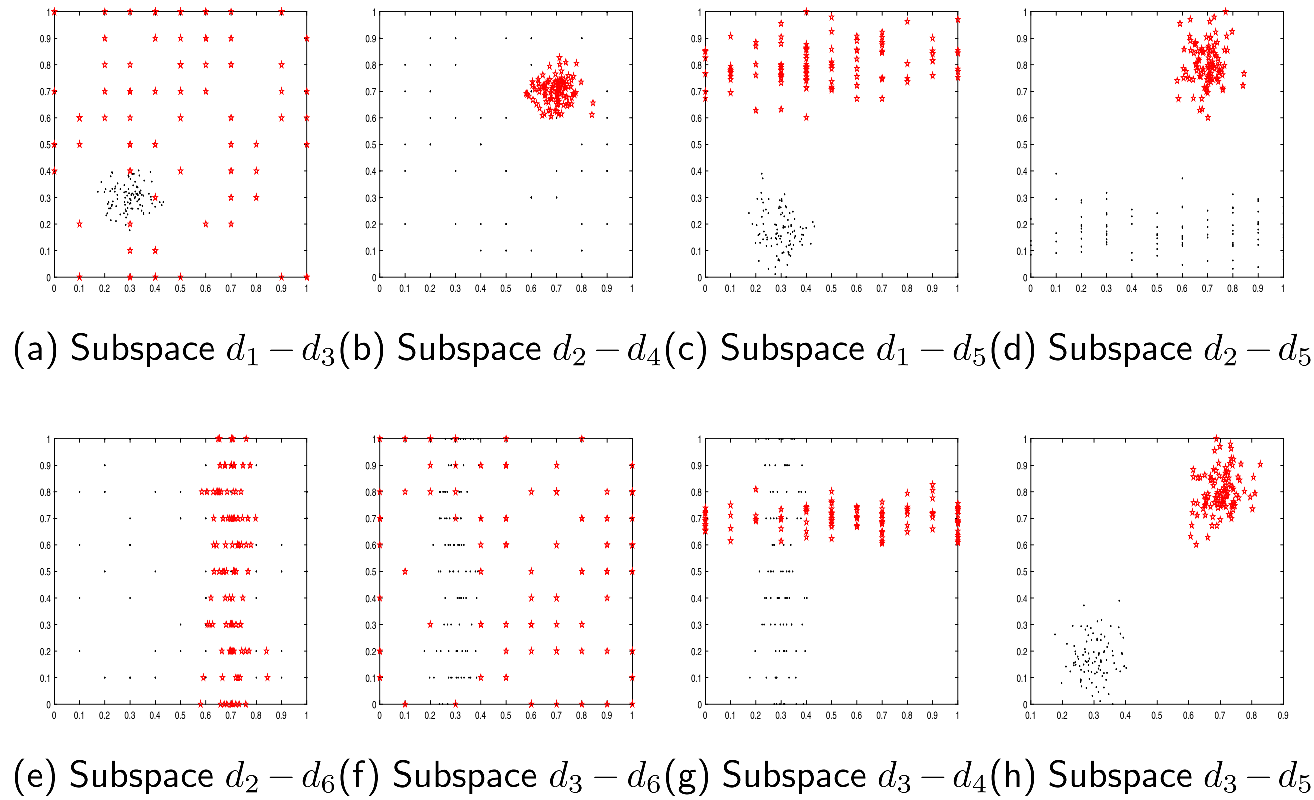
6 Conclusion

In this paper, we have introduced a spherical path-based clustering algorithm called uW-IFCM and an ellipsoidal path-based algorithm known as bW-PIFCM. The uW-IFCM algorithm has assigned a single weight to each feature, whereas bW-PIFCM is an adaptive algorithm that allocates data-driven weight triplet to each feature. The feature weights in uW-IFCM are further tuned with the help of a weighing parameter η. Here, the real-valued dataset has been transformed into an AIFS dataset using a novel intuitionistic fuzzification technique. The technique does not extract membership and non-membership functions from the dataset. This lacking of the extraction process leads to uncertainty for which the two parameters α and β have been introduced. The uW-IFCM is not an adaptive algorithm, so it is computationally expensive. In order to reduce the cost of the uW-IFCM algorithm, we have introduced the FRT-equipped uW-IFCM algorithm. The feature reduction technique eliminates irrelevant features from the dataset. The proposed uW-IFCM algorithm has been compared with the introduced bW-PIFCM algorithm, PIFCM, IFCM, wIFCM, wFCM, and FCM algorithms over UCI machine learning datasets and some synthetic datasets. The feature weight distribution used in uW-IFCM is appropriate, so it has effectively handled irrelevant and noisy features in comparison to other algorithms. The uW-IFCM has shrunk the search space of fuzzy factor m besides eliminating irrelevant features, so the running time of the algorithm is reduced.

In the future, the aim is to study the outliers and noises possessing real-world datasets using uW-IFCM. We will try to formulate an AIFS extraction technique involving a multi-valued function. The final objective is to suggest exponent η-based improvisation for the proposed bW-PIFCM and to tune the weight triplets with η. The FRT-

Table 17 WORST/AVERAGE/THE BEST *CA*, R_I, *PC*, *SC*, and *DI* with fixed initial feature weights and different initial cluster centroids over UCI datasets

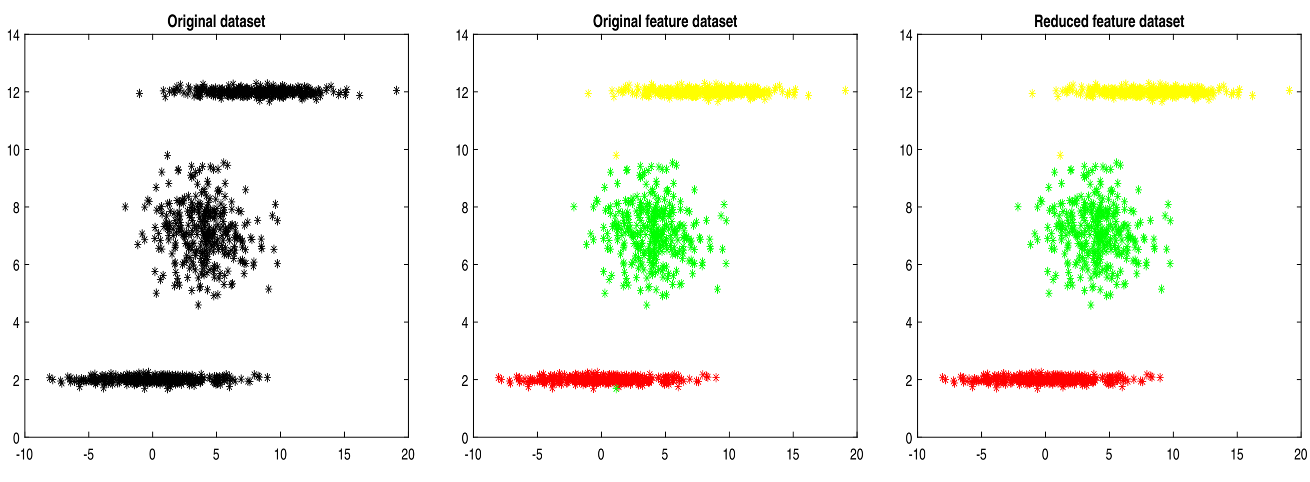
		uW-IFCM			wIFCM			IFCM			wFCM			FCM		
		Worst	Average	Best	Worst	Average	Best	Worst	Average	Best	Worst	Average	Best	Worst	Average	Best
IRIS	CA	0.9333	0.9467	**0.9677**	0.8867	0.9133	0.9333	0.5400	**0.9000**	**0.9333**	0.8800	0.9067	0.9233	0.6667	0.7533	0.9200
	RI	0.9200	0.9495	**0.9575**	0.8923	0.9055	0.9124	0.7527	0.8847	**0.9187**	0.8679	0.8923	0.9124	0.5684	0.7527	0.8859
	PC	0.8264	0.9053	**0.9926**	0.7818	0.8862	0.9932	0.3333	0.9935	**0.9554**	0.7366	0.8867	0.8860	0.3333	0.3333	0.7748
	SC	0.0001	0.0001	**0.0001**	0.6617	0.5315	0.1571	0.5509	0.0153	**0.0166**	0.7027	0.4201	0.1778	1.5904	0.8530	0.4720
	DI	0.0874	0.0504	**0.0452**	0.0701	0.0564	0.0347	0.0715	0.1385	**0.0347**	0.0701	0.0694	0.0347	0.1701	0.1408	0.0968
Thyroid	CA	0.9070	0.9227	**0.9674**	0.8930	0.9070	**0.9674**	0.5488	0.6233	0.7814	0.8837	0.9070	0.9228	0.5814	0.6279	0.7581
	RI	0.8839	0.9226	**0.9464**	0.8056	0.9069	**0.9469**	0.6862	0.6928	0.6735	0.6692	0.8069	0.8469	0.5713	0.6041	0.6616
	PC	0.8909	0.9131	**0.9945**	0.8800	0.9227	**0.9958**	0.3333	0.4196	**0.6934**	0.7517	0.8885	0.8879	0.3333	0.6057	0.6538
	SC	0.0017	0.0008	**0.0004**	1.1073	0.6644	**0.2528**	0.1130	0.0835	0.0127	1.4577	0.3429	0.2893	1.1058	0.7817	0.2630
	DI	0.0262	0.0139	**0.0068**	0.0834	0.0509	**0.0142**	0.1134	0.0821	0.0579	0.2996	0.2986	0.2421	0.2142	0.1423	0.1222
Bupa	CA	0.5072	0.5217	**0.5600**	0.5014	0.5275	0.5697	0.5217	0.5304	0.5681	0.5101	0.5304	0.5478	0.5072	0.5072	0.5536
	RI	0.4987	0.5026	0.5062	0.4986	0.4987	0.5008	0.4995	0.5004	0.5079	0.4987	0.5008	0.5031	0.4988	0.4987	0.5043
	PC	0.6421	0.7830	0.8804	0.5270	0.7988	0.8843	0.5000	0.5400	0.6000	0.6298	0.7073	0.8685	0.5000	0.5000	0.5189
	SC	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
	DI	0.0410	0.0206	0.0142	0.1267	0.0563	0.0379	0.1148	0.0279	0.0232	0.0797	0.3512	0.1402	0.2592	0.1592	0.1160
zoo	CA	0.7921	0.8812	**0.9109**	0.7327	0.8317	**0.9208**	0.7426	0.7624	0.7921	0.6733	0.7327	0.8020	0.6040	0.6238	0.7327
	RI	0.8925	0.8347	**0.9157**	0.8954	0.8552	**0.9207**	0.8644	0.8457	0.7500	0.8145	0.8903	0.7833	0.7202	0.6058	0.6088
	PC	0.8964	0.9037	**0.9803**	0.4019	0.8360	**0.9953**	0.3358	0.7999	**0.8527**	0.3457	0.8065	0.9894	0.1429	0.7400	0.7541
	SC	0.0292	0.0143	**0.0104**	0.0487	0.0309	**0.0122**	0.0188	0.0024	0.0071	0.7211	0.4270	0.2939	0.2712	0.2212	0.2212
	DI	0.2056	0.1213	0.1170	0.4061	0.3714	**0.0928**	0.2673	0.3015	0.0115	0.3522	0.3522	0.3417	0.9887	0.3123	0.8941
Heart	CA	0.8001	0.8185	**0.8333**	0.6963	0.7037	0.8122	0.5704	0.5963	0.6222	0.5926	0.7889	0.7848	0.5111	0.5185	0.5519
	RI	0.6924	0.7018	0.7212	0.5213	0.5670	0.5814	0.5081	0.5168	0.5281	0.6572	0.6878	0.6971	0.5876	0.5940	0.6255
	PC	0.8419	0.8902	0.9675	0.7642	0.8557	0.9991	0.7943	0.7411	0.8000	0.5630	0.7414	0.9611	0.5113	0.5113	0.5113
	SC	0.0015	0.0008	0.0004	0.3047	0.2305	0.1400	0.1150	0.1217	0.6655	2.6022	2.4132	1.8340	0.0002	0.0003	0.0002
	DI	0.0145	0.0133	0.0125	0.3715	0.2438	0.1042	0.0467	0.0414	0.0114	0.7547	0.6744	0.5873	0.8125	0.7125	0.6133
WDBC	CA	0.8830	0.9011	**0.9420**	0.9086	0.9244	**0.9402**	0.5571	0.7786	0.8594	0.9121	0.9279	0.8350	0.6643	0.6626	0.8243
	RI	0.7971	0.8880	**0.9212**	0.8336	0.8660	**0.8905**	0.5532	0.6546	0.7098	0.8279	0.8423	0.5057	0.5521	0.6811	0.7098
	PC	0.8951	0.9435	**0.9724**	0.5612	0.7874	**0.9801**	0.5000	0.5000	**0.9322**	0.6497	0.7941	0.9781	0.5000	0.5000	0.9240
	SC	0.0009	0.0006	**0.0004**	0.0500	0.0480	**0.0309**	9.2764	0.0354	0.0074	1.6470	1.4797	1.2826	0.5408	0.5258	0.1516
	DI	0.0149	0.0137	**0.0155**	0.0838	0.0800	**0.0759**	0.0023	0.0022	0.0194	0.0838	0.0800	0.0761	0.0476	0.0439	0.1121
Ecoli	CA	0.7676	0.8165	**0.8563**	0.6330	0.7706	0.8459	0.7339	0.7339	0.7409	0.6049	0.7047	0.7310	0.6820	0.7034	0.7131
	RI	0.7479	0.8025	**0.8109**	0.6074	0.7275	0.8029	0.6980	0.7489	0.7085	0.5898	0.6000	0.6936	0.5752	0.6479	0.6722
	PC	0.8537	0.9174	**0.9729**	0.6307	0.8229	0.9705	0.9778	0.9361	0.8640	0.4044	0.6154	0.9048	0.2013	0.4702	0.8695
	SC	0.3161	0.2391	**0.1971**	0.1616	0.1148	0.0775	0.0202	0.0213	0.0247	1.1374	0.9397	0.4827	3.0291	0.4132	0.1329
	DI	0.0927	0.0530	**0.0393**	0.0599	0.0543	0.0448	0.0706	0.0706	0.0962	0.3719	0.2627	0.2424	0.2553	0.1583	0.1793



(a) Subspace $d_1 - d_3$ (b) Subspace $d_2 - d_4$ (c) Subspace $d_1 - d_5$ (d) Subspace $d_2 - d_5$

(e) Subspace $d_2 - d_6$ (f) Subspace $d_3 - d_6$ (g) Subspace $d_3 - d_4$ (h) Subspace $d_3 - d_5$

Fig. 12 Synthetic Dataset IV illustrated in Example 3 viewed in different subspaces



(a) (b) (c)

Fig. 13 **a** 2D—Original Synthetic Dataset II, **b** clustering with uW-IFCM using original features, **c** clustering with uW-IFCM with final features

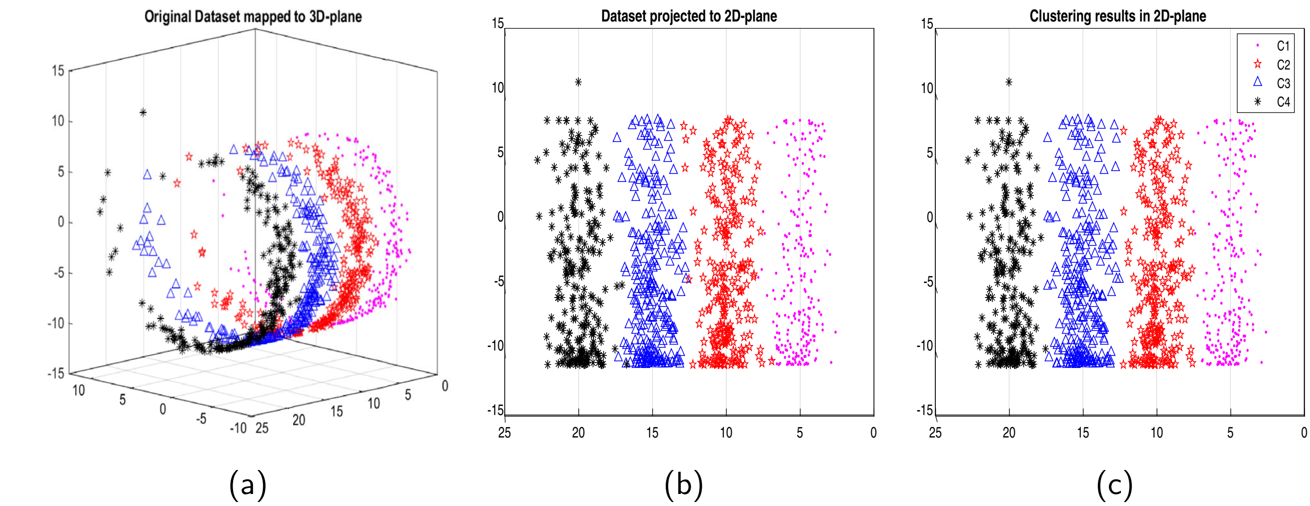


Fig. 14 **a** Original Synthetic Dataset III mapped to 3D, **b** dataset projected to x–y plane, **c** clustering with uW-IFCM

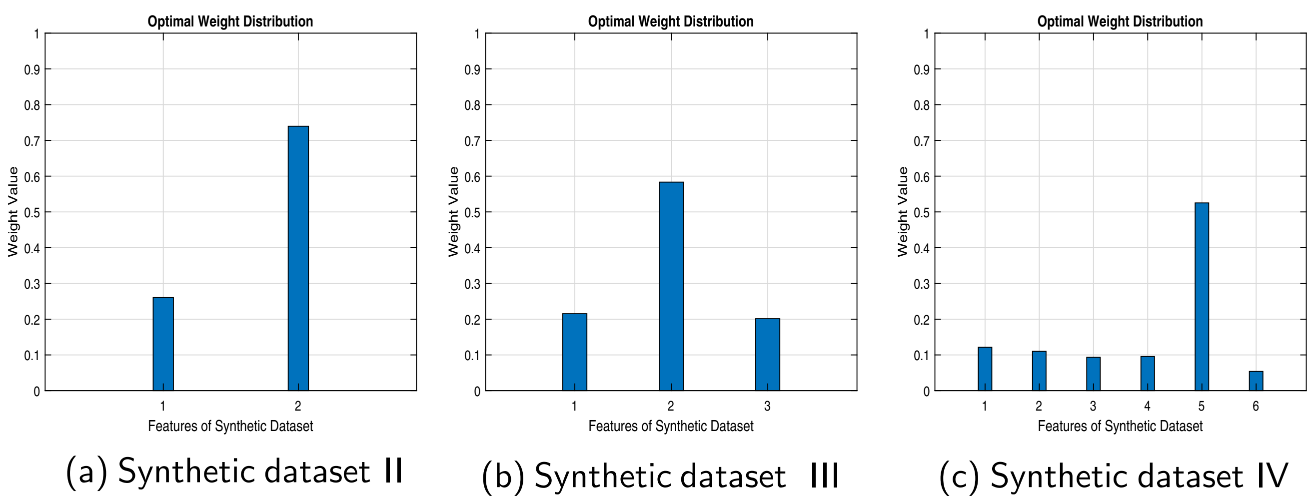


Fig. 15 Feature weight distribution of the three synthetic datasets using uW-IFCM

equipped uW-IFCM algorithm deduces a threshold-dependent weight distribution, and thereby relevant features get separated from irrelevant features. In the future, we will try to introduce the FRT-equipped bW-PIFCM algorithm. Here, the paper has established bW-PIFCM as an adaptive algorithm.

The future work of this paper is motivated by the limitations of our proposed approach, which involve

Table 18 WORST/AVERAGE/THE BEST *CA*, R_I, *PC*, *SC*, and *DI* of bW-PIFCM and PIFCM with fixed initial feature weights and different initial cluster centroids over UCI datasets

		bW-PIFCM			PIFCM		
		Worst	Average	Best	Worst	Average	Best
IRIS	CA	0.5	0.8584	**0.9633**	0.3400	0.8500	0.9398
	RI	0.7894	0.8989	**0.9317**	0.5120	0.8985	0.9200
	PC	0.4587	0.7852	**0.9755**	0.1254	0.4785	0.5871
	SC	0.2547	0.1248	**0.1582**	0.8546	0.6848	0.7997
	DI	0.2140	0.0547	**0.0384**	0.1245	0.0568	0.0458
Thyroid	CA	0.5684	0.9012	**0.9845**	0.5487	0.7845	0.9794
	RI	0.6845	0.8461	0.9493	0.6984	0.8946	**0.9696**
	PC	0.4569	0.5487	0.6033	0.5897	0.4879	0.5487
	SC	0.8794	0.6924	0.7237	0.8947	0.7894	0.6584
	DI	0.0846	0.5846	0.0564	0.7914	0.1249	0.0579
Bupa	CA	0.3845	0.3487	0.6897	0.3541	0.5100	0.5847
	RI	0.4824	0.4925	0.5487	0.2847	0.3541	0.4879
	PC	0.3841	0.4500	0.4879	0.4824	0.5100	0.5712
	SC	0.5479	0.3547	0.2801	0.8940	0.4580	0.2984
	DI	0.8946	0.5875	0.1356	0.4514	0.1058	0.0342
zoo	CA	0.5134	0.6548	0.8879	0.4876	0.6854	0.8874
	RI	0.6841	0.8497	0.9548	0.5640	0.8945	0.9395
	PC	0.6845	0.7985	0.8451	0.2546	0.8945	0.9587
	SC	0.4879	0.3546	0.2894	0.6547	0.3541	0.2748
	DI	0.7845	0.4578	0.3875	1.4650	0.5464	0.4798
Heart	CA	0.4578	0.6984	**0.8311**	0.2846	0.6548	0.8457
	RI	0.2345	0.6547	0.6987	0.5846	0.5266	0.6845
	PC	0.5487	0.6451	0.7024	0.3546	0.4879	0.6854
	SC	3.1541	2.2458	2.1540	4.4680	2.1546	2.1457
	DI	0.5456	0.4578	0.1478	2.1478	1.9000	1.8794
WDBC	CA	0.6489	0.8254	**0.9408**	0.5870	0.8465	0.9548
	RI	0.7815	0.8947	0.8845	0.8045	0.8874	0.8974
	PC	0.4897	0.5487	0.6584	0.7894	0.6745	0.6845
	SC	2.1640	1.8490	1.2846	2.1645	1.5469	1.6540
	DI	0.8794	0.1254	0.0659	1.2540	0.1479	0.0845
Ecoli	CA	**0.8074**	**0.8174**	**0.8514**	0.4581	0.7900	0.8057
	RI	0.9022	0.9200	0.9245	0.4100	0.9011	0.9015
	PC	0.8000	0.8512	0.8746	0.1100	0.6104	0.8745
	SC	0.5899	0.5478	0.5487	0.8010	0.8541	0.6547
	DI	0.8790	0.0468	0.0245	0.8914	0.1500	0.0456

Table 19 Clustering performance of uW-IFCM on IRIS dataset

Sr.No	Initial Feature Weights				Final Feature Weights				Measuring indexes with optimal parameters						
	d_1	d_2	d_3	d_4	D_1	D_2	D_3	D_4	CA	RI	PC	SC	DI	m	η
1	0.2500	0.2500	0.2500	0.2500	0.2659	0.2735	0.2168	0.2438	0.9000	0.8859	0.7674	0.00080	0.04342	1.8	0.5
2	0.2500	0.2500	0.2500	0.2500	0.2470	0.2461	0.2566	0.2503	0.9200	0.9055	0.7175	0.00095	0.04342	1.9	1.1
3	0.2500	0.2500	0.2500	0.2500	0.2107	0.2106	0.2863	0.2925	0.9533	0.9417	0.9737	0.00044	0.08123	1.3	1.5
4	0.2500	0.2500	0.2500	0.2500	**0.1688**	**0.1598**	**0.3765**	**0.2949**	**0.9867**	**0.9875**	**0.9205**	**0.00007**	**0.05123**	**1.7**	**2.0**
5	0.2500	0.2500	0.2500	0.2500	0.0893	0.0903	**0.6025**	0.2179	0.9467	0.9326	0.8823	0.00044	0.08123	1.1	3.0
6	0.2500	0.2500	0.2500	0.2500	0.0453	0.0442	**0.7647**	0.1458	0.9000	0.8949	0.9625	0.00441	0.04342	1.1	4.0
7	0.2500	0.2500	0.2500	0.2500	0.0200	0.0194	**0.8658**	0.0948	0.8900	0.8749	0.8962	0.00441	0.04342	1.1	5.0
8	0.2500	0.2500	0.2500	0.2500	0.0034	0.0032	**0.9587**	0.0347	0.8800	0.8892	0.7962	0.00048	0.04342	1.2	7.0
9	0.2500	0.2500	0.2500	0.2500	0.0013	0.0013	**0.9770**	0.0204	0.8940	0.8925	0.7966	0.00441	0.04342	1.1	8.0
10	0.2500	0.2500	0.2500	0.2500	0.0002	0.0002	**0.9928**	0.0068	0.9400	0.9249	0.8825	0.00441	0.04342	1.1	10.0
11	0.1541	0.1821	0.3943	0.2696	0.2594	0.2592	0.2260	0.2555	0.8833	0.8977	0.6670	0.01167	0.04342	2.0	0.5
12	0.1676	0.3710	0.1575	0.3039	0.2479	0.2478	0.2553	0.2489	0.9200	0.9045	0.6677	0.00115	0.04342	2.0	1.1
13	0.0992	0.3175	0.2733	0.3099	0.2107	0.2106	0.2863	0.2925	0.9533	0.9417	0.9737	0.00044	0.08123	1.4	1.5
14	0.3119	0.2371	0.3183	0.1327	**0.1622**	**0.1661**	**0.3527**	**0.3191**	**0.9700**	**0.9495**	**0.9782**	**0.00044**	**0.08123**	**1.5**	**2.0**
15	0.5649	0.0754	0.0016	0.3581	0.0893	0.0903	0.6025	0.2179	0.9467	0.9326	0.9123	0.00004	0.08123	1.1	3.0
16	0.5666	0.2759	0.0601	0.0974	0.0200	0.0194	0.8658	0.0948	0.8840	0.9249	0.9625	0.00008	0.04342	1.1	5.0
17	0.1095	0.5150	0.2241	0.1514	0.0034	0.0032	0.9587	0.0347	0.8740	0.8952	0.8962	0.00004	0.04342	1.1	7.0
18	0.2479	0.4323	0.0984	0.2214	0.0013	0.0013	0.9770	0.0204	0.8894	0.8249	0.8862	0.00441	0.04342	1.1	8.0
19	0.0401	0.4558	0.0880	0.4161	0.0005	0.0005	0.9871	0.0118	0.8666	0.8692	0.9625	0.00044	0.04342	1.1	9.0
20	0.0609	0.3043	0.3416	0.2931	**0.0002**	**0.0002**	**0.9928**	**0.0068**	0.8494	0.8925	0.9625	0.00441	0.04342	1.1	10.0

Table 20 Updation of feature weights for IRIS dataset using uW-IFCM when $\eta = 2$

Number of Iterations	Feature Weights			
	Sepal Length	Sepal Width	Petal length	Petal Width
Initialization of feature weights	0.2500	0.2500	0.2500	0.2500
Iteration 1	0.2059	0.2081	0.3275	0.2584
Iteration 2	0.2042	0.1881	0.3483	0.2594
Iteration 3	0.1833	0.1769	0.3701	0.2697
Iteration 4	0.1696	0.1702	0.3809	0.2792
Iteration 5	0.1627	0.1657	0.3833	0.2884
Iteration 6	0.1606	0.1649	0.3777	0.2969
Iteration 7	0.1608	0.1651	0.3693	0.3048
Iteration 8	0.1615	0.1656	0.3602	0.3127
Iteration 9	0.1622	0.1661	0.3527	0.3191

Table 21 Final number of features for synthetic datasets with varying χ

		Variation in exponent χ						
		0.5	0.8	1	1.5	2	3	5
	Exponent $(\chi/2)$	0.25	0.4	0.5	0.75	1	1.5	2.5
1	Synthetic Dataset I	–	–	–	1	3	5	5
2	Synthetic Dataset II	–	1	1	1	1	1	1
3	Synthetic Dataset III	–	–	–	1	1	3	3
4	Synthetic Dataset IV	–	–	–	–	1	6	6
5	IRIS dataset	–	–	–	–	2	4	4

Table 22 Running time (in seconds) of uW-IFCM with original and final features of dataset

Sr.No	Datasets	Original features	Final features	uW-IFCM with original features	uW-IFCM with final features
1	Synthetic Dataset II	2	1	0.0763	0.0698
2	Synthetic Dataset III	3	1	0.1989	0.1100
3	Synthetic Dataset IV	6	1	0.0733	0.0467

implementing the algorithm on real-valued datasets in a computationally efficient manner.

Author Contributions All the authors have equally contributed in the paper.

Funding The research work of the first author is funded by University Grants Commission (UGC), New Delhi, India under the grant number 19/06/2016(i)EU-V.

Declarations

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

Ethical Approval This manuscript does not contain any studies with human participants performed by any of the authors.

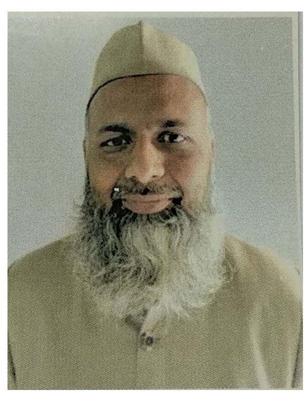
Informed Consent We do not have any associated data used in our manuscript.

References

1. Agrawal, R., Gehrke, J., Gunopulos, D., Raghavan, P.: Automatic subspace clustering of high dimensional data for data mining applications. In: Proceedings of the 1998 ACM SIGMOD International Conference on Management of Data, pp. 94–105 (1998)
2. Vidal, R.: Subspace clustering. IEEE Signal Process. Mag. **28**(2), 52–68 (2011)
3. Sander, J., Ester, M., Kriegel, H.-P., Xiaowei, X.: Density-based clustering in spatial databases: the algorithm gdbscan and its applications. Data Min. Knowl. Discov. **2**(2), 169–194 (1998)
4. Valente, D.O.J., Witold, P.: Advances in Fuzzy Clustering and Its Applications. Wiley, New York (2007)
5. Cohen-Addad, V., Kanade, V., Mallmann-Trenn, F., Mathieu, C.: Hierarchical clustering: objective functions and algorithms. J. ACM **66**(4), 1–42 (2019)
6. Bezdek, J.C., Ehrlich, R., Full, W.: Fcm: the fuzzy c-means clustering algorithm. Comput. Geosci. **10**(2–3), 191–203 (1984)
7. Zeshui, X., Junjie, W.: Intuitionistic fuzzy c-means clustering algorithms. J. Syst. Eng. Electron. **21**(4), 580–590 (2010)
8. Qiu, C., Xiao, J., Yu, L., Han, L., Iqbal, M.N.: A modified interval type-2 fuzzy c-means algorithm with application in mr image segmentation. Pattern Recogn. Lett. **34**(12), 1329–1338 (2013)
9. Chengmao, W., Xiaokang, G.: A novel interval-valued data driven type-2 possibilistic local information c-means clustering for land cover classification. Int. J. Approx. Reason. **148**, 80–116 (2022)
10. Ji, Z., Sun, Q., Xia, Y., Chen, Q., Xia, D., Feng, D.: Generalized rough fuzzy c-means algorithm for brain mr image segmentation. Comput. Methods Programs Biomed. **108**(2), 644–655 (2012)
11. Li, F., Ye, M., Chen, X.: An extension to rough c-means clustering based on decision-theoretic rough sets model. Int. J. Approx. Reason. **55**(1), 116–129 (2014)
12. Chen, N., Ze-shui, X., Xia, M.: Hierarchical hesitant fuzzy k-means clustering algorithm. Appl. Math. A **29**(1), 1–17 (2014)
13. Gustafson,D.E., Kessel, W.C.: Fuzzy clustering with a fuzzy covariance matrix. In: 1978 IEEE Conference on Decision and Control Including the 17th Symposium on Adaptive Processes, pp. 761–766. IEEE (1979)
14. Kuo-Lung, W., Yang, M.-S.: Alternative c-means clustering algorithms. Pattern Recognit. **35**(10), 2267–2278 (2002)
15. Lee, J.Y., Kim, D., Mun, J.Y., Kang, S., Son, S.H., Shin, S.: Texture weighted fuzzy c-means algorithm for 3d brain mri segmentation. In: Proceedings of the 2018 Conference on Research in Adaptive and Convergent Systems, pp. 291–295 (2018)
16. Liu, X., Li, X., Zhang, Y., Yang, C., Xu, W., Li, M., Luo, H.: Remote sensing image classification based on dot density function weighted fcm clustering algorithm. In: 2007 IEEE International Geoscience and Remote Sensing Symposium, pp. 2010–2013. IEEE (2007)
17. Miin-Shen, Y., Sinaga Kristina, P.: Collaborative feature-weighted multi-view fuzzy c-means clustering. Pattern Recognit. **119**, 108064 (2021)
18. Wenyuan, Z., Tianyu, H., Jun, C.: A robust bias-correction fuzzy weighted c-ordered-means clustering algorithm. Math. Probl. Eng. (2019). https://doi.org/10.1155/2019/5984649
19. Li, J., Gao, X., Ji, H.: A feature weighted FCM clustering algorithm based on evolutionary strategy. In: Proceedings of the 4th World Congress on Intelligent Control and Automation (Cat. No. 02EX527), Vol. 2, pp. 1549–1553. IEEE (2002)
20. Wang, X., Wang, Y., Wang, L.: Improving fuzzy c-means clustering based on feature-weight learning. Pattern Recognit. Lett. **25**(10), 1123–1132 (2004)
21. Yang, M.-S., Nataliani, Y.: A feature-reduction fuzzy clustering algorithm based on feature-weighted entropy. IEEE Trans. Fuzzy Syst. **26**(2), 817–835 (2017)
22. Stephan, T., Sharma, K., Shankar, A., Punitha, S., Varadarajan, V., Liu, P.: Fuzzy-logic-inspired zone-based clustering algorithm for wireless sensor networks. Int. J. Fuzzy Syst. **23**, 506–517 (2021)
23. Wang, G., Wang, J.-S., Wang, H.-Y.: Fuzzy c-means clustering validity function based on multiple clustering performance evaluation components. Int. J. Fuzzy Syst. **24**(4), 1859–1887 (2022)
24. Sinaga, K.P., Hussain, I., Yang, M.-S.: Entropy k-means clustering with feature reduction under unknown number of clusters. IEEE Access **9**, 67736–67751 (2021)
25. D'Urso, P., Leski, J.M.: Fuzzy clustering of fuzzy data based on robust loss functions and ordered weighted averaging. Fuzzy Sets Syst. **389**, 1–28 (2020)
26. He Yu-Lin, O., Gui-Liang, P.F.-V., Huang, J.Z., Suganthan, P.N.: A novel dependency-oriented mixed-attribute data classification method. Expert Syst. Appl. **199**, 116782 (2022)
27. Sun, L., Zhang, J., Ding, W., Jiucheng, X.: Feature reduction for imbalanced data classification using similarity-based feature

clustering with adaptive weighted k-nearest neighbors. Inf. Sci. **593**, 591–613 (2022)
28. Nha, V.P., The, P.L., Pedrycz, W., Ngo, L.T.: Feature-reduction fuzzy co-clustering approach for hyper-spectral image analysis. Knowl. Based Syst. **216**, 106549 (2021)
29. Siminski, K.: An outlier-robust neuro-fuzzy system for classification and regression. Int. J. Appl. Math. Comput. Sci. **31**(2), 303–319 (2021)
30. Khan, M.A., Arshad, H., Nisar, W., Javed, M.Y., Sharif, M.: An integrated design of fuzzy c-means and nca-based multi-properties feature reduction for brain tumor recognition. In: Signal and Image Processing Techniques for the Development of Intelligent Healthcare Systems, pp. 1–28. Springer, New York (2021)
31. Atanassov, K.T.: Intuitionistic fuzzy sets. Fuzzy Sets Syst. **20**(1), 87–96 (1986)
32. Chaira, T.: A novel intuitionistic fuzzy c means clustering algorithm and its application to medical images. Appl. Soft Comput. **11**(2), 1711–1717 (2011)
33. Huang, C.-W., Lin, K.-P., Ming-Chang, W., Hung, K.-C., Liu, G.-S., Jen, C.-H.: Intuitionisti fuzzy c-means clustering algorithm with neighborhood attration in segmenting medial image. Soft Comput. **19**(2), 459–470 (2015)
34. Iakovidis, D.K., Pelekis, N., Kotsifakos, E., Kopanakis, I.: Intuitionistic fuzzy clustering with applications in computer vision. In: International Conference on Advanced Concepts for Intelligent Vision Systems, pp. 764–774. Springer, New York (2008)
35. Lin, K.-P.: A novel evolutionary kernel intuitionistic fuzzy *c*-means clustering algorithm. IEEE Trans. Fuzzy Syst. **22**(5), 1074–1087 (2013)
36. Hanuman, V., Agrawal, R.K., Aditi, S.: An improved intuitionistic fuzzy c-means clustering algorithm incorporating local information for brain image segmentation. Appl. Soft Comput. **46**, 543–557 (2016)
37. Zhao, F., Chen, Y., Liu, H., Fan, J.: Alternate pso-based adaptive interval type-2 intuitionistic fuzzy c-means clustering algorithm for color image segmentation. IEEE Access **7**, 64028–64039 (2019)
38. Zhou, X., Zhao, R., Fengquan, Yu., Tian, H.: Intuitionistic fuzzy entropy clustering algorithm for infrared image segmentation. J. Intell. Fuzzy Syst. **30**(3), 1831–1840 (2016)
39. Namburu, A., Samayamantula, S.K., Edara, S.R.: Generalised rough intuitionistic fuzzy c-means for magnetic resonance brain image segmentation. IET Image Process. **11**(9), 777–785 (2017)
40. Kaushal, M., Solanki, R., Danish Lohani, Q.M., Muhuri Pranab, K.: A novel intuitionistic fuzzy set generator with application to clustering. In: 2018 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), pp. 1–8. IEEE (2018)
41. Danish Lohani, Q.M., Solanki, R., Muhuri, P.K.: Novel adaptive clustering algorithms based on a probabilistic similarity measure over atanassov intuitionistic fuzzy set. IEEE Trans. Fuzzy Syst. **26**(6), 3715–3729 (2018)
42. Szmidt, E., Kacprzyk, J.: Geometric similarity measures for the intuitionistic fuzzy sets. In: EUSFLAT Conference (2013)
43. Szmidt, E., Kacprzyk, J.: Distances between intuitionistic fuzzy sets. Fuzzy Sets Syst. **114**(3), 505–518 (2000)
44. Zhexue, H.J., Ng Michael, K., Hongqiang, R., Zichen, L.: Automated variable weighting in k-means type clustering. IEEE Trans. Pattern Anal. Mach. Intell. **27**(5), 657–668 (2005)
45. Danish Lohani, Q.M., Solanki, R., Muhuri, P.K.: A convergence theorem and an experimental study of intuitionistic fuzzy c-mean algorithm over machine learning dataset. Appl. Soft Comput. **71**, 1176–1188 (2018)
46. Siminski, K.: Fuzzy weighted c-ordered means clustering algorithm. Fuzzy Sets Syst. **318**, 1–33 (2017)
47. Asuncion, A., Newman, D.: Uci machine learning repository (2007)
48. Graves, D., Pedrycz, W.: Kernel-based fuzzy clustering and fuzzy clustering: a comparative experimental study. Fuzzy Sets Syst. **161**(4), 522–543 (2010)
49. Pal, N.R., Bezdek, J.C.: On cluster validity for the fuzzy c-means model. IEEE Trans. Fuzzy Syst. **3**(3), 370–379 (1995)
50. Bensaid Amine, M., Hall Lawrence, O., Bezdek James, C., Clarke Laurence, P., Silbiger Martin, L., Arrington John, A., Murtagh, R.F.: Validity-guided (re) clustering with applications to image segmentation. IEEE Trans. Fuzzy Syst. **4**(2), 112–123 (1996)
51. Bezdek, J.C., Pal, N.R.: Cluster validation with generalized dunn's indices. In: Proceedings 1995 Second New Zealand International Two-Stream Conference on Artificial Neural Networks and Expert Systems, pp. 190–190. IEEE Computer Society (1995)

Meenakshi Kaushal has received her Ph.D. degree in Mathematics from South Asian University, India. She has been funded Junior Research Fellowship (JRF) and Senior Research Fellowship (SRF) from the University Grant Commission, Government of India. She made an international visit to Rio de Janeiro, Brazil to present her paper in FUZZ IEEE 2018 which was funded by South Asian University, New Delhi. Her current research interests are in the fields of machine learning, optimization, data envelopment analysis, intuitionistic fuzzy sets, type-2 fuzzy sets, and operator theory. She has published eleven papers in reputed journals/conferences including Information Science, IEEE Access, and granular computing.



Q. M. Danish Lohani (Member, IEEE) has received his Ph.D. degree from the Department of Mathematics, Aligarh Muslim University, India, in 2009. He is currently working as an Associate Professor with the Department of Mathematics, at South Asian University, India. He has published more than 55 scientific articles in well-known journals, including IEEE Transactions on Fuzzy Systems, Applied Soft Computing, Chaos, Solitons and Fractals, and reputed conferences, such as Fuzz-IEEE, IEEE-CEC, and IJCNN. His current research interests include the study of both theoretical and application aspects of fuzzy, intuitionistic fuzzy, type-2 fuzzy sets, operators, and summability theory. He is a member of the IEEE Computational Intelligence.

Oscar Castillo holds the Doctor in Science degree (Doctor Habilitatus) in Computer Science from the Polish Academy of Sciences (with the Dissertation "Soft Computing and Fractal Theory for Intelligent Manufacturing"). He is a Professor of Computer Science in the Graduate Division, Tijuana Institute of Technology, Tijuana, Mexico. In addition, he is serving as Research Director of Computer Science and head of the research group on Hybrid Fuzzy Intelligent Systems. Currently, he is President of HAFSA (Hispanic American Fuzzy Systems Association) and Past President of IFSA (International Fuzzy Systems Association). He is also Chair of the Mexican Chapter of the Computational Intelligence Society (IEEE). He also belongs to the Technical Committee on Fuzzy Systems of IEEE and to the Task Force on "Extensions to Type-1 Fuzzy Systems." He is also a member of NAFIPS, IFSA, and IEEE. He belongs to the Mexican Research System (SNI Level 3). His research interests are in Type-2 Fuzzy Logic, Fuzzy Control, Neuro-Fuzzy, and Genetic-Fuzzy hybrid approaches. He has published over 300 journal papers, 10 authored books, 60 edited books, 300 papers in conference proceedings, and more than 300 chapters in edited books, in total more than 1000 publications. He has been Guest Editor of several successful Special Issues in the past, like in the following journals: Applied Soft Computing, Intelligent Systems, Information Sciences, Soft Computing, Non-Linear Studies, Fuzzy Sets and Systems, JAMRIS, and Engineering Letters. He is currently Associate Editor of the Information Sciences Journal, Applied Soft Computing Journal, Journal of Engineering Applications on Artificial Intelligence, Granular Computing Journal, and the IEEE Transactions on Fuzzy Systems. He has been elected as IFSA Fellow and MICAI Fellow member last year. Finally, he recently received the Recognition as Highly Cited Researcher in 2017 and 2018 by Clarivate Analytics and Web of Science.