Iranian Journal of Fuzzy Systems

Volume 19, Number 5, (2022), pp. 167-181

Original paper

University of Sistan and Baluchestan

HFC: Data clustering based on hesitant fuzzy decision making

L. Aliahmadipour[1], M. Eftekhari[2] and V. Torra[3]

[1] *Department of Computer Science, Faculty of Mathematics and Computer, Shahid Bahonar University of Kerman, Kerman, Iran*

[2] *Department of Computer Engineering, Shahid Bahonar University of Kerman, Kerman, Iran*

[3] *Department of Computing Science, UmeåUniversity, Umeå, Sweden*

l.Aliahmadipour@uk.ac.ir, m.eftekhari@uk.ac.ir, vtorra@cs.umu.se

Abstract

In a clustering task, choosing a proper clustering algorithm and obtaining qualified clusters are crucial issues. Sometimes, a clustering algorithm is chosen based on the data distribution, but data distributions are not known beforehand in real world problems. In this case, we hesitate which clustering algorithm to choose. In this paper, this hesitation is modeled by a hesitant fuzzy multi criteria decision making problem (HFMCDM) in which some clustering algorithms play the role of experts. Here, we consider fuzzy C-means (FCM) and agglomerative clustering algorithms as representative of two popular categories of clustering algorithms partitioning and hierarchical clustering methods, respectively. Then, we propose a new clustering procedure based on hesitant fuzzy decision making approaches (HFC) to decide which of the FCM family or hierarchical clustering algorithms is suitable for our data. This procedure ascertains a good clustering algorithm using neutrosophic FCM (NFCM) through a two phases process. The HFC procedure not only makes a true decision about applying partitioning clustering algorithms, but also improves the performance of FCM and evolutionary kernel intuitionistic fuzzy c-means clustering algorithm (EKIFCM) with construction hesitant fuzzy partition (HFP) conveniently. Experimental results show that the clustering procedure is applicable and practical. According to HFC procedure, it should be mentioned that it is possible to replace the other clustering algorithms that belong to any partitioning and hierarchical clustering methods. Also, we can consider other categories of clustering algorithms.

Keywords: FCM clustering algorithm, hierarchical clustering algorithm, hesitant fuzzy decision making, hesitant fuzzy clustering.

1 Introduction

Our goal is to define a finite set of clusters according to similarities between objects by means of clustering [14]. Clustering methods are used in various contexts and disciplines like data mining, document retrieval, image segmentation and pattern classification. Clusters are considered sets of similar patterns, and clustering methods are used to gather patterns based on similarity (or dissimilarity) criteria [48]. Clustering algorithms could be divided according to different points of view. Based on [22, 43] we can classify the crisp clustering methods into five categories:

- *Hierarchical methods*: In this clustering method, data are grouped into three typed clusters. Agglomerative (bottom-up) and divisive (top-down) are two different types of hierarchical clustering approaches. agglomerative nesting (AGNES) and divisive analysis (DIANA) [25] are in this category.
- *Partitioning methods:* Many partitioning techniques of cluster analysis exist. The most typical approaches are K-means, fuzzy c-means and evolutionary FCM [24].
- *Density-based methods:* In this approach the local density of the samples is used to determine the clusters. So, non-convex and arbitrarily shaped clusters by DBSCAN and OPTICS can be detected [25].

Corresponding Author: L. Aliahmadipour
Received: July 2021; Revised: June 2022; Accepted: August 2022.
https://doi.org/10.22111/IJFS.2022.7163

- *Grid-based methods:* These methods are used for dividing the given data space into cells of a given grid and merge them so that clusters are constructed. Statistical information grid approach STING is a grid based clustering algorithm [28].
- *Model-based methods:* This approach applies maximum likelihood algorithms for evaluating parameters such as expectation maximization (EM) and self organizing map (SOM) [31], and models election techniques for certifying a model structure.

An alternative classification based on [16] distinguishes the two most popular categories: *hierarchical* and *partitioning* clustering methods because they work on opposite form of data distribution. Hierarchical clustering methods can be classified into either agglomerative or divisive types. Agglomerative hierarchical clustering applies each object in its own cluster and then joins these atomic clusters into larger ones. It continues until either all data are in a single cluster or certain termination conditions are satisfied. *Hierarchical* clustering methods are mostly belonging to this type for example linkage family clustering algorithm (single-linkage, complete-linkage, average-linkage and ward-linkage). They are only different in their definition of intercluster similarity [22, 29]. *Partitioning* clustering methods attain a single partition of data without other sub partitions such as hierarchical algorithms and in accordance with the optimization of an appropriate objective function. For instance, standard partitioning methods are K-means, FCM, and its improvements. The clustering algorithm can be chosen based on the type of available data and its particular purpose [22].

Since L.Zadeh proposed fuzzy set theory [47], this concept is using in clustering tasks [39] widely. Later, K.T. Atanassov introduced an extension of fuzzy set called intuitionistic fuzzy set (A-IFS) [4]. A-IFS is characterized by a membership function and a non-membership function. Recently, Torra introduced hesitant fuzzy sets (HFSs) [38, 35] to consider the importance of hesitation and uncertainty in the nature of some problems which were not well managed by the previous tools. A HFS is defined in terms of a function that returns a set of membership values for each element in the domain. Hesitant fuzzy sets can model situations in which membership of elements to a given set have different values. For instance, to have a rational decision result, a decision organization, with several decision makers, it is allowed to evaluate the degree of the attribute that an alternative should satisfy. If there are three causes, a group of decision makers can provide 0.4, a second group provide 0.3, and the others provide 0.6. These three can't persuade each other, so the degree that the alternative should satisfy the criterion is represented with this hesitant fuzzy set $\{0.4, 0.3, 0.6\}$ [42]. Hesitant fuzzy sets are applied to areas such as multi criteria decision making and clustering problems. We use the hesitant fuzzy decision making method for clustering data.

Based on [2] Fuzzy sets, intuitionistic fuzzy sets and hesitant fuzzy sets can be applied in the clustering tasks from two points of view:

1. Considering uncertain data and, in particular, fuzzy data (FD) [32], intuitionistic fuzzy data (IFD)[44] and hesitant fuzzy data (HFD) [1].
2. Considering crisp data set but uncertain clusters and, in particular, fuzzy partition (FP) [5], intuitionistic fuzzy partition (IFP)[6, 37] and hesitant fuzzy partition (HFP) [3].

Recently, Eftekhari et al. [12] review some new trends of research regarding the use of hesitant fuzzy sets, in machine learning. Also, They show that HFS through construction HFP can improve the result of clustering algorithms. Moreover, their research shows that applying the ability of HFSs in clustering tasks requires more studies. So, we focus on the ability of HFS and use it for some goals presented below. In this paper, we consider the problem of how to choose a suitable clustering algorithm among some other clustering algorithms without data processing. To this aim, we introduce HFC procedure based on hesitant fuzzy decision making approach. In the following, we present the major contributions of this paper:

- The first major contribution of the paper is considering some clustering algorithms as experts instead of users to choose the appropriate clustering algorithm types. To this aim, we propose hesitant fuzzy clustering (HFC) procedure which can select the proper algorithm from some categories e.g. partitioning and hierarchical methods on a data set.
- The second major contribution is estimating the number of clusters based on the fuzzy cluster validity indices.
- The last, since we focus on partitioning clustering algorithms, we improve the accuracy of fuzzy C-means clustering family by constructing hesitant fuzzy partitions.

In order to do these, the paper is organized as follows: Section 2 presents a brief introduction of (FCM, NFCM, linkage family) clustering algorithms. Section 3 reviews some concepts related to hesitant fuzzy sets and their applications

Algorithm 1. Pseudo code of the FCM algorithm.

Input: Data set $X = (x_1, x_2, x_3, \cdots, x_N)$ and the number of clusters C.

Output: Optimal cluster centers, V^*, and optimal membership values, U^*.

Step 1. Generate initial cluster centers, V, randomly.

Step 2. Solve $\min_{\mu \in U} J_{FCM}(U, V, X)$ computing Equation. 3.

Step 3. Solve $\min_{V} J_{FCM}(U, V, X)$ computing Equation. 4.

Step 4. If the stop condition does not hold, go to step 2; otherwise, stop.

to decision making problems. Section 4 explains HFC procedure. Section 5 provides the experimental results of the clustering procedure in the context of a machine learning data set. Section 6 offers conclusions and some suggestions for further research.

2 Clustering methods

According to the purpose of this study, we consider two types of clustering methods that work in remarkably different ways such as the FCM algorithm and the Linkage family. They are representatives of the partitioning method and hierarchical method, respectively. In this section, we present a brief introduction to FCM, NFCM, and the linkage family. *Fuzzy C-mean* (FCM) introduced by Bezdek et al. [5] is an unsupervised algorithm for data clustering. The method is defined by the minimization of the objective function as follows:

$$J_{FCM}(U, V, X) = \sum_{i=1}^{N} \sum_{j=1}^{C} \mu_{ij}^{m} d^2(x_i, v_j), \tag{1}$$

subject to the constraints $\mu_{ij} \in [0, 1]$ and $\sum_{j=1}^{C} \mu_{ij} = 1$ for all $x_i \in X$. In the above equation, $X = \{x_1, \cdots, x_N\}$ is a set of input objects, N is the number of objects and C is the number of clusters. Then, μ_{ij} represents the degree of belongingness of the ith object to the jth cluster and $U_{N \times C}$ is the fuzzy membership function matrix that contains membership degrees μ_{ij}. $V = \{v_1, \cdots, v_C\}$ is the set of centers of the clusters and m is the fuzzy index that governs the influence of membership grades. In general, m is any real number greater than 1, m is frequently set to 2, see e.g. [36]. In Eq. (1), d^2 is the square of a distance. It is usual to use $d^2(x_i, v_j)$ to measure the dissimilarity between x_i and v_j:

$$d^2(x_i, v_j) = \|x_i - v_j\|^2. \tag{2}$$

Usually, the problem is solved by means of an algorithm that begins by randomly initializing the center vectors. Then, an iterative process is used which consists of two steps. Algorithm 1 presents the process of FCM. The first, where μ_{ij} are computed using the following equation:

$$\mu_{ij} = \frac{1}{\sum_{k=1}^{C} \left(\frac{d^2(x_i, v_j)}{d^2(x_i, v_k)} \right)^{\frac{2}{m-1}}} \qquad \sum_{k=1}^{C} \mu_{ij} = 1 \quad \forall i = 1, \cdots, N. \tag{3}$$

Next, the centers are updated as follows:

$$v_j = \frac{\sum_{i=1}^{N} (\mu_{ij})^m x_i}{\sum_{i=1}^{N} (\mu_{ij})^m} \qquad j = 1, \cdots, C, \quad 1 < m < N. \tag{4}$$

The iterative process is repeated with new memberships and centers until $\mid v_j^{t+1} - v_j^t \mid < \varepsilon$ and $\mid \mu_{ij}^{t+1} - \mu_{ij}^t \mid < \varepsilon$ for all i and j. Here ϵ is a small value (ε is a termination criterion between 0 and 1) and t denotes the iteration. Despite its efficiency, the FCM algorithm is sensitive to selecting initial values and takes long convergence time for large datasets. Usage of different kernels also causes different clustering. In our procedure to confirm the selected clustering algorithm, we apply the neutrosophic objective function [20] in the FCM and called it *neutrosophic fuzzy C-means* (NFCM). We detail this method below. *Neutrosophic evidential c-means (NECM)*, introduced by Guo et al. [20], is a new clustering method based on the neutrosophic sets (NSs) and evidence theory. Neutrosophic set describes each data using three membership values: T, F, and I, where they define the degree of truth, falsity and indeterminacy respectively. So, we utilized T, I, and F to represent the membership degrees to the cluster centers, ambiguity center and outlier center, respectively. Considering the effect of indeterminacy, a new objective function is defined (see [20] for details):

Algorithm 2. Pseudo code of NFCM clustering algorithm.

Input: Data set $X = (x_1, x_2, x_3, \cdots, x_N)$ and the number of clusters C.

Output: Optimal cluster centers, V^*, and optimal membership values, U^*.

Step 1. Initialize $T^{(0)}$, $I^{(0)}$, and $F^{(0)}$. Set C, m, δ, ϵ, w_1, w_2, and w_3 parameters.

Step 2. Compute the cluster centers, $v^{(k)}$ at kth step using Equation. 8.

Step 3. Calculate v_{jmax}.

Step 4. Update $T^{(k)}$, $F^{(k)}$, and $F^{(k)}$ using Equation. 9,10, and 11, respectively.

Step 5. If the stop condition does not hold, go to step 2; otherwise, stop.

$$J_{NFCM}(T, I, F, V) = \sum_{i=1}^{N}\sum_{j=1}^{C}(w_1 T_{ij})^m \parallel x_i - v_j \parallel^2 + \sum_{i=1}(w_2 I_i)^m \parallel x_i - v_{jmax} \parallel^2 + \sum_{i=1}^{N}\delta^2(w_3 F_i)^m. \tag{5}$$

$$v_{jmax} = \frac{v_{p_i} + v_{q_i}}{2}, \quad p_i = \underbrace{argmax}_{j=1,\cdots,C}(T_{ij}), \quad q_i = \underbrace{argmax}_{j\neq p_i \cap j=1,\cdots,C}(T_{ij}). \tag{6}$$

T_{ij} is the membership value of data i belonging to the jth cluster. Also, m is a constant and usually selected as 2. $w_k (k = 1, 2, 3)$ is their weighting factor. The number of objects referred as outliers are controlled by using δ. p_i and q_i are the cluster indices with the biggest and second biggest value of T. T_{ij}, I_i, and F_i are the membership values belonging to the true, indeterminacy and false set. They satisfy $0 < T_{ij}; I_i; F_i < 1$ and the following constrain:

$$\sum_{j=1}^{C} T_{ij} + I_i + F_i = 1. \tag{7}$$

Applying the method of Lagrange multipliers to the minimization of the objective function in Eq. (5) we obtain the following equations:

$$v_j = \frac{\sum_{i=1}^{N}(w_1 T_{ij})^m x_i}{\sum_{i=1}^{N}(w_1 T_{ij})^m}. \tag{8}$$

$$T_{ij} = \frac{K}{w_1}(x_i - v_j)^{\frac{-2}{m-1}}. \tag{9}$$

$$I_i = \frac{K}{w_2}(x_i - v_{jmax})^{\frac{-2}{m-1}}. \tag{10}$$

$$F_i = \frac{K}{w_3}\delta^{\frac{-2}{m-1}}. \tag{11}$$

$$K = \left[\sum_{j=1}^{C}(x_i - v_j)^{\frac{-2}{m-1}} + (x_i - v_{jmax})^{\frac{-2}{m-1}} + \delta^{\frac{-2}{m-1}}\right]^{-1}. \tag{12}$$

The partitioning is carried out through an iterative optimization of the objective function. The pseudo code of the NFCM algorithm is illustrated in Algorithm. 2. The cluster centers v_j, the memberships T_{ij}, I_i and F_i are updated as follows. (8) - (12), respectively. The iteration stops when $|T_{ij}^{n+1} - T_{ij}^{n}| < \epsilon$ where ϵ is a termination criterion between 0 and 1, and n is the iteration step [20] Algorithm 2 illustrates its iterative process. In this paper we apply NFCM to achieve two goals. First to confirm whether FCM family is appropriate for the data set or not, and then to improve FCM family through the proposed hesitant fuzzy decision making method. *Linkage family clustering:* In hierarchical clustering methods, we organize data points in a nested sequence of groups rather than simple partitions. We can classify these aforementioned techniques into two groups which are agglomerative and divisive. In these methods, hierarchy is created by recursively merging the two clusters at minimum distance. This procedure is described in Algorithm 3. In the initial state, we estimate the distance among data with Euclidean distance, Manhattan distance or any other distance [34]. So, the only difference among these algorithms is the determination of the distance between two clusters of data. Different distances lead to a different set of clusters. There is a lot of research on improving hierarchical

Algorithm 3. Pseudo code of the hierarchical clustering algorithm (Agglomerative).

Input: Data set $X = (x_1, x_2, x_3, \cdots, x_N)$.

Output: A dendogram showing the structure of clusters.

Step 1. Start with N clusters each holds one data point.

Step 2. Compute the distance matrix for clusters.

Step 3. **Repeat**
Merge two clusters with closest distance.
Update the cluster distance matrix.

Step 4. **Until** there is only one cluster.

clustering. For example, Rashedi et al. [29] discusses that the Bob-Hic clustering algorithm is a new boosting-based hierarchical cluster ensemble method that tries to establish a consensus hierarchical cluster (h-clustering) on a data set and is good for enhancing the clustering accuracy. We use all of the linkage algorithms on the data sets to evaluate the clustering procedure in experimental results section.

3 Hesitant fuzzy decision making

In this section, we describe some concepts related to hesitant fuzzy sets and hesitant fuzzy decision making methods which are used in the paper.

3.1 Hesitant fuzzy sets

In many real world problems having doubt between a set of different values may lead to difficulty in verifying the membership of an element. In this situation, the concept of HFSs as a new extension of fuzzy sets [38] was introduced. Several possible values in $[0, 1]$ can present the membership degree of HFS. The definition is as follows.

Definition 3.1. [35] *Let X be a fixed set. Then a hesitant fuzzy set (HFS) on X in terms of a function h is that when applied to X returns a subset of $[0, 1]$, i.e. , $h : X \to \mathcal{P}([0, 1])$.*

Furthermore, given a set of fuzzy sets, a HFS could be defined in accordance with the union of their memberships as follow:

Definition 3.2. [35] *Let $M = \{\mu_1, \mu_2, \ldots, \mu_n\}$ be a set of n membership functions and $x \in X$. The HFS associated to M, h_M is defined as:*

$$h_M(x) = \bigcup_{\mu \in M} \{\mu(x)\}.$$

Xia and Xu [41] called $h(x)$ a hesitant fuzzy element (HFE). A hesitant fuzzy element is a set of values in $[0, 1]$, and a HFS is a set of HFEs, for each $x \in X$.

Definition 3.3. [35] *Given a HFE, h, the intuitionistic fuzzy value (IFV) $Aenv(h)$ is defined as the envelope of h, where $Aenv(h)$ can be represented as $(h^-, 1 - h^+)$, with $h^- = \min\{\gamma | \gamma \in h\}$ and $h^+ = \max\{\gamma | \gamma \in h\}$, where γ is a membership value.*

For example, we have a hesitant fuzzy element, $h = \{0.2, 0.5, 0.6\}$, then the intuitionstic fuzzy envelope of h is $Aenv(h) = (0.2, 0.4)$.
Based on the relationship between the HFEs and IFVs, Xu and Xia in [46] defined some new operations on the HFEs h, h_1 and h_2:

1. $h^\lambda = \cup_{\gamma \in h} \{\gamma^\lambda\}$,
2. $\lambda h = \cup_{\gamma \in h} \{1 - (1 - \gamma)^\lambda\}$,
3. $h_1 \oplus h_2 = \cup_{\gamma_1 \in h_1, \gamma_2 \in h_2} \{\gamma_1 + \gamma_2 - \gamma_1\gamma_2\}$,
4. $h_1 \otimes h_2 = \cup_{\gamma_1 \in h_1, \gamma_2 \in h_2} \{\gamma_1\gamma_2\}$,
5. $h^c = \cup_{\gamma \in h} \{1 - \gamma\}$,

6. $h_1 \cup h_2 = \cup_{\gamma_1 \in h_1, \gamma_2 \in h_2} \max\{\gamma_1, \gamma_2\}$,
7. $h_1 \cap h_2 = \cap_{\gamma_1 \in h_1, \gamma_2 \in h_2} \min\{\gamma_1, \gamma_2\}$.

Rodriguez et al. in [31] present an in depth overview about HFSs considering both theoretical concepts and their application.

3.2 Hesitant fuzzy decision making

Multi attribute decision making (MADM) has usually been applied to evaluate a finite number of decision alternatives having multiple attributes. So experts widely apply this issue to diverse scientific fields [41], [23] such as manufacturing systems, environmental impact assessment and location selection. In many real world problems, decision makers (DMs) have a lot of difficulties to give their assessments on performance ratings and refer weights with decisive values. Since the discovery of the fuzzy sets [47], evaluating decision options for MADM problems is possible by using fuzzy sets to describe the ambiguities that decision makers face. Accordingly, various fuzzy MADM methods based on fuzzy sets, interval-valued fuzzy sets and intuitionistic fuzzy sets have been suggested to handle fuzzy information [27], [7]. But all the aforementioned methods do not cope with the Hesitant fuzzy assessments of DMs. To define and deal with uncertain information in the process of MADM, HFS allow the membership of an element to a given set have different values [45]. Consider an MADM problem with n alternatives $Y_i (i = 1, 2, \ldots, n)$ and m attributes $A_j (j = 1, 2, \cdots, m)$ with the attribute weight vector $w = (w_1, w_2, \cdots, w_n)^T$ such that $w_j \in [0, 1], (j = 1, 2, ..., n)$, and l decision makers $E_k (k = 1, 2, \cdots, l)$ that provide several values for the alternative Y_i under the attribute A_j. These values can be considered as a hesitant fuzzy element h_{ij}. A basic scheme of a MADM consists of two phases, aggregation and exploitation [31]. In the aggregation phase, the information is grouped to reflect a collective value for each alternative or attribute. In the exploitation phase, the best alternative is selected according to the collective values obtained in the previous phase. Experts use different aggregation techniques to fulfill decision making process in which they use HFE to indicate their evaluation. Xia and Xu presented in [41] some aggregation operators, such as hesitant fuzzy weighted averaging (HFWA) and hesitant fuzzy weighted geometric (HFWG). They are defined as follows.

Definition 3.4. [25] *Let $h_i (i = 1, \cdots, n)$ be a collection of $HFEs$, $h_i \in H$, the (HFWA) operator is a mapping $H^n \to H$ such that*

$$HFWA(h_1, \cdots, h_n) = \oplus_{i=1}^{n} (w_i h_i) = \bigcup_{\gamma_1 \in h_1, \cdots, \gamma_n \in h_n} \{1 - \prod_{i=1}^{n} (1 - \gamma_i)^{w_i}\}. \tag{13}$$

where $w = (w_1, \cdots, w_n)^T$ is a weighting vector with $w_i \in [0, 1]$ and $\sum_{i=1}^{n} w_i = 1$. In case of $w = (1/n, \cdots, 1/n)^T$, then the HFWA operator reduces to the hesitant fuzzy averaging (HFA) operator:

$$HFA(h_1, \cdots, h_n) = \oplus_{i=1}^{n} (\frac{1}{n} h_i) = \bigcup_{\gamma_1 \in h_1, \cdots, \gamma_n \in h_n} \{1 - \prod_{i=1}^{n} (1 - \gamma_i)^{\frac{1}{n}}\}. \tag{14}$$

Definition 3.5. *Let $h_i (i = 1, \cdots, n)$ be a collection of $HFEs$, $w = (w_1, \cdots, w_n)^T$ be a weighting vector with $w_i \in [0, 1]$ and $\sum_{i=1}^{n} w_i = 1$. A generalized hesitant fuzzy weighted geometric (GHFWG) operator is a mapping $H^n \to H$, and*

$$GHFWG_\lambda (h_1, ..., h_n) = \left(\oplus_{i=1}^{n} (w_i h_i)^\lambda\right)^{\frac{1}{\lambda}} = \bigcup_{\gamma_1 \in h_1, \gamma_2 \in h_2, ..., \gamma_n \in h_n} \left\{1 - \left(1 - \prod_{i=1}^{n} \left(1 - \gamma_i^\lambda\right)^{w_i}\right)^{1/\lambda}\right\}. \tag{15}$$

Other aggregation operators for hesitant fuzzy sets have been defined in the literature, see e.g. [31], [45]. After aggregation, we apply the exploitation phase. We choose the most desirable alternative(s) out of a set of alternatives in this phase. Which is possible by comparing HFEs of previous phase with a score function like the following [41]:

Definition 3.6. *For a HFE h, $S(h) = \frac{1}{\sigma(h)} \sum_{\gamma \in h} \gamma$ is called the score function of h, where $\sigma(h)$ is the number of the elements in h. For two HFEs h_1 and h_2,if $S(h_1) > S(h_2)$, then $h_1 > h_2$; if $S(h_1) = S(h_2)$, then $h_1 = h_2$.*

Farhadinia in [15] defined a series of score functions for hesitant fuzzy sets which rank HFSs. In our hesitant fuzzy decision making problem, two clustering algorithms play the role of experts. Each clustering algorithm designates (different) membership degrees to each cluster for each data set. To do this, we propose the following procedure.

4 HFC: Clustering procedure based on hesitant fuzzy decision making

In this section, we utilize the hesitant fuzzy decision making approach to cluster a given data set. In a clustering task, we consider the following three sets, X, C and E. $X = \{x_1, x_2, \cdots, x_N\}$ is a set of data points, $\mathcal{C} = \{C_1, C_2, \cdots, C_c\}$ is a set of m clusters, and $E = \{e_1, e_2, \cdots, e_l\}$ is a set of l clustering algorithms. In this study $E = \{FCM, NFCM\}$. To model the clustering problem by a hesitant fuzzy decision making problem, we consider that the elements of X, $\mathcal{C}$, and E play the role of alternatives, attributes, and experts in MADM problem, respectively. We consider that all clusters have the same priority. We do not consider a weighting vector on the cluster set, but we could consider a weighting vector $\omega = (\omega_1, \omega_2)$ for E. Our proposed clustering procedure to be applied to a given data set X has two phases:

Phase 1. Determining a suitable cluster number, m.

Phase 2. Selecting the appropriate clustering algorithm through a hesitant fuzzy decision making method and then clustering the data points.

4.1 Phase 1

Generally, in most clustering algorithms, the number of clusters is a predefined parameter. First, we should determine the initial number of clustersThe problem of estimating the number of clusters has been extensively studied in the literature [30], [11]. Since we are using fuzzy clustering algorithms, then we apply a cluster number estimation method for fuzzy clustering. In particular, we use a cluster validity index proposed in [30] for fuzzy clustering. The index $V_{SC}(c, U)$ is defined as the sum of the compactness measure, $Comp(c, U)$, and the separation measure, $Sep(c, U)$. It is formulated as follow:

$$V_{SC} = \frac{Sep(c, U)}{max_c \quad Sep(c, U)} + \frac{Comp(c, U)}{max_c \quad Comp(c, U)}. \tag{16}$$

$$Sep(c, U) = \{Sep(i, U)\}_{i=2,\cdots,c_{max}}, \tag{17}$$

$$Copm(c, U) = \{Copm(i, U)\}_{i=2,\cdots,c_{max}}, \tag{18}$$

where $c = 2, 3, \cdots, c_{max}$ is a variable for cluster number and U is a membership matrix in a fuzzy clustering such as FCM. The total compactness of the fuzzy partition, $Comp(k, U)$, is given by

$$Comp(c, U) = \sum_{i=1}^{c} \sum_{j=1}^{n} u_{ij}^2 \|x_j - v_i\|, \tag{19}$$

where u_{ij} is a membership value and indicates the degree that sample x_j belongs to the cluster center vector v_i, the separation measure is described as the average value of similarity of $c(c-1)/2$ pairs of clusters:

$$Sep(c, U) = \frac{2}{c(c-1)} \sum_{p \neq q}^{c} S(C_p, C_q). \tag{20}$$

The similarity of fuzzy clusters C_p and C_q is defined as

$$S(C_p, C_q) = \sum_{j=1}^{n} S(x_j : C_p, C_q), \tag{21}$$

where $S(x_j : C_p, C_q) = min(u_{C_p}(x_j), u_{C_q}(x_j))$ is the similarity of fuzzy clusters C_p and C_q at x_j and $h(x_j) = -\sum_{i=1}^{c} u_{C_q}(x_j) \log_a u_{C_q}(x_j)$ is a weight that adjusts the emphasis on the overlapped data points depending on how ambiguously an overlapped data point is shared between fuzzy clusters. A low degree of overlap and a small value for the compactness are two characteristics of a desirable partition. Therefore, the optimum value of the number of cluster, c, is obtained by minimizing $V_{SC}(c, U)$ in Eq. (16) over $c = 2, 3, ..., c_{max}$ [30]. The main reason for choosing the proposed method in [30] is that this validity index includes two measures: the compactness and separation measures. The compactness measure combines the fuzziness in U with the geometrical compactness of the representation of X via the c prototypes V. Moreover, considering the appropriate cluster numbers for a given data set improves the accuracy of the clustering procedure in the next phase.

4.2 Phase 2

After estimating the cluster number c, we run the FCM and NFCM clustering algorithm on the data set. FCM clustering is sensitive to the selection of initial cluster centers and kernel. For example, Zhao in [49], supposed that achievement of good results by using Euclidean distance as Kernel when all clusters are spheroids with the same size or all clusters are well separated. Gustafson and Kessel [21] proposed Gustafson Kessel fuzzy C-means GKFCM algorithm which uses the Mahalanobis distance as the kernel in FCM [43]. They mentioned that the GKFCM is better than Euclidean distance based algorithms based on considering the shape of data. Therefore, for the data set without any prior knowledge, we hesitate which of the kernels and initial cluster centers can be used.

Here, we consider various kernels and initial cluster center selection methods to not miss the proper clusters. So, we use three kernels: Euclidean, Mahalanobis distance, and cosine similarity measure. Moreover, we choose initial cluster centers in three ways: randomly, based on the cumulative approach [13], and based on the subtractive clustering algorithm [9]. Then, we run FCM and NFCM clustering algorithm with the nine combinations of the above parameters before making a decision. We recall that the clustering results of FCM and NFCM with various kernels and initial cluster centers selection methods are different. We can model this diversity by a hesitant fuzzy decision making problem (HFDM) and each decision maker provides several values about the cluster of each data point. According to kernel types and initial cluster centers selection methods, FCM and NFCM are worked with 3×3 different parameter values. Therefore, FCM computes 9 membership matrix $U_{m\times N}$ and also NFCM computes 9 membership matrix $T_{m\times N}$, indeterminacy vector I_N and falsity vector F_N. In this situation for instance FCM for each data point x_i for $i = 1, 2, \cdots, N$ as different parametrization computes 9 membership values for each cluster C_j, these values considered as a hesitant fuzzy element h_{ij} for $i = 1, 2, \cdots, N$, and $j = 1, 2, \cdots, c$. Based on the above explanation, we proposed the following decision making model for data clustering:

Step 1. The decision makers (FCM and NFCM) provide their evaluations about the data point x_i under the cluster C_j, denoted by the hesitant fuzzy element h_{ij} for $i = 1, 2, \cdots, N$, and $j = 1, 2, \cdots, c$. $h_{ij}^{(1)}$ and $h_{ij}^{(2)}$ are hesitant fuzzy elements obtained by FCM and NFCM, respectively.

$$\begin{aligned} h_{ij}^{(1)} &= \{\mu_{ij}^{(k)} | 0 \le \mu_{ij}^{(k)} \le 1\}, \\ h_{ij}^{(2)} &= \{t_{ij}^{(k)} | 0 \le t_{ij}^{(k)} \le 1\} \quad k = 1, 2, \cdots, |M_1 \times M_2|. \end{aligned} \tag{22}$$

Therefore, $\mu_{ij}^{(k)}$ denotes that μ_{ij} is the membership value provided by FCM in the kth execution and $t_{ij}^{(k)}$ is a truth value provided by NFCM in the kth execution. Here, M_1 corresponds to the set of kernel types, $M_1 = \{Euclidean, Mahalanobis, cosine\}$ and M_2 is the set of initial cluster center selection methods, $M_2 = \{randomly, cumulative, subtractive\}$.

Step 2. Applying hesitant fuzzy weighted averaging (HFWA)operator, Eq. (13), to obtain the hesitant fuzzy elements h_{ij} for the decision makers $e_l (l = 1, 2)$, i. e.,

$$h_{ij} = HFWA(h_{ij}^{(1)}, h_{ij}^{(2)}), \tag{23}$$

where $h_{ij}^{(l)}$ is presented in Eq. (22). In this step, the hesitant fuzzy partitions are constructed which are used to improve fuzzy and intuitionistic fuzzy clustering algorithms.

Step 3. Computing the score values $S(h_{ij})$ of $h_{ij} (i = 1, 2, \cdots, N)$ and $(j = 1, 2, \cdots, c)$ by Def. (3.6).

Step 4. Determining cluster of each data point x_i based on ranking $S(h_{ij})$.

$$x_i \in C_{r_i} \quad \text{if} \quad r_i = \underbrace{argmax}_{j=1,2,\cdots,c} S(h_{ij}),$$

it means that $x_i (i = 1, 2, \cdots, N)$ belongs to C_{r_i} which has maximum score value $S(h_{ir_i})$ among score values.

Step 5. Deciding whether FCM family clustering algorithm is appropriate for the given data set or not. In this case, we handle some clustering results of NFCM, indeterminacy vectors I and falsity vectors F. We collect the results and have an indeterminacy matrix $I_{k\times N}$ and a falsity matrix $F_{k\times N}$, where N is the number of data points and k denotes that k times NFCM is performed on $M_1 \times M_2$. As previously mentioned, the implementation for NFCM is similarly to the one for FCM. So, if FCM family is suitable for the data set, NFCM confirms it with small values in I and F, because in this situation the algorithm works truly without any indeterminacy and falsity. Applying the validity index

Algorithm 4. Pseudo code of HFC procedure.

Input: Data set $X = (x_1, x_2, x_3, \cdots, x_N)$ and θ.

Output: Optimal clusters.

Step 1. Estimate the appropriate cluster numbers c for X based on *Phase. 1.*

Step 2. Initialize parameters of FCM and NFCM.

Step 3. Run FCM and NFCM on $\{M_1 \times M_2\}$ in the parallel way.

Step 4. Compute $test_{value}$using Equation. 24. 8.

Step 5. Solve the obtained HFDM problem.

Step 6. If $test_{value} \leq \theta$
Choose linkage family clustering algorithms, Stop
Else
Final clusters are the results of solving HFDM problem, Stop.

for fuzzy clustering in phase 1 confirms the above conclusion. On the other hands, since the values of I and F are bigger than T_{ij} of data point x_i for each clusters C_j, then it is obvious that the FCM family is not suitable due to high indeterminacy and falsity in NFCM. Therefore, it persuades us to apply linkage family clustering algorithms or other types of clustering methods.

Step 6. This is carried out based on $test_{value}$ computed as follow:

$$test_{value} = \sum_{k=1}^{9}\sum_{i=1}^{N} I(k,i) + \sum_{k=1}^{9}\sum_{i=1}^{N} F(k,i). \tag{24}$$

The decision is made based on the following rule:

$$\begin{aligned} \text{If} \quad test_{value} \leq \theta \quad & \text{Then FCM family is proper.} \\ & \text{Else linkage family is proper.} \end{aligned} \tag{25}$$

Where, θ is a threshold estimated based on experimental results and it is user defined.

Step 7. Stop.

In short, Algorithm 4 shows the pseudo code of HFC procedure. It should be mentioned that we use FCM and NFCM as experts. Then, we compute $test_{value}$ while we could use every hierarchical clustering algorithm when $test_{value} \leq 0.6$. Here, we use { single, complete, ward, average linkage, Bob-hic} in both experimental models, and also we could select other hierarchical algorithms such as [40]. We note that the main aim of this study is recommending the suitable clustering method for a given data set (specially partitioning or other clustering categories). Moreover, we improve the accuracy of FCM and EKIFCM [10] in an easy way. In the case of HFC procedure persuades us to apply the hierarchical clustering algorithm, we use one of the best improved algorithms of this clustering method or other meta clustering algorithms such as ensemble [29] or consensus approach. The suggested hesitant fuzzy decision making method can also be applied to various clustering algorithms, with the role of experts, to attain better results.

It should be noticed that HFC suffers from determining $test_{value}$ and cost for choosing partitioning or hierarchical clustering algorithms. For the first one, we estimate it based on the trial and error method. For the last, running several methods with several parameters have extra costs, plus the cost of deciding and processing all these results. To deal with these costs, we suggest applying simple clustering algorithms as experts can lead to reducing these costs. After a suitable clustering method is chosen, we can use each meta clustering to estimate the best solution.

5 Experimental results and comparison

In order to show how HFC procedure works, we consider two types of experiments. The aim of the fisrt experiments is to visualize the selection process between partitioning and hierarchical clustering algorithms. To this aim, we use two-dimensional labeled benchmark data sets for these types of algorithms which are shown in Fig 1. Data sets are named Jain, Aggregation, Spiral, Flam, Compound [33] and Longsquare [29]. The main characteristics of these data sets are presented in Table 1.

Also, we aim that HFC procedure improves the intuitionistic fuzzy clustering, EKIFCM, by constructing hesitant fuzzy partitions. So, the second set of experiments compares HFC procedure with EKIFCM. Lin in [10] proposed a

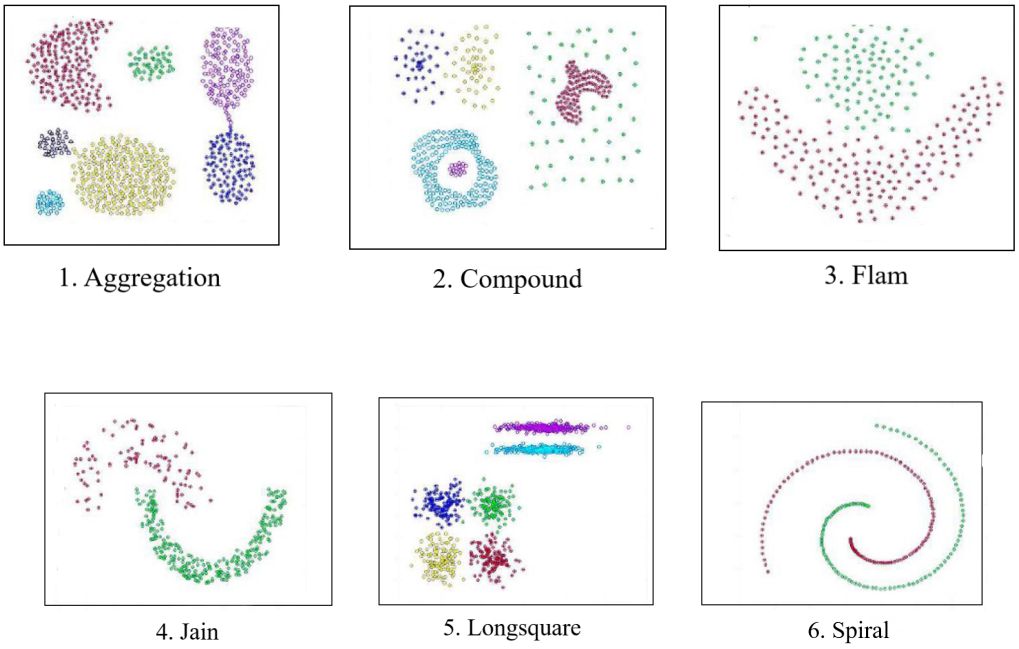


Figure 1: Two-dimensional datasets used to demonstrate the quality of the HFC procedure.

novel evolutionary kernel intuitionistic fuzzy c-mean clustering algorithm (EKIFCM). This algorithm applies Atanassov's intuitionistic fuzzy sets with kernel-based fuzzy c-mean (KFCM) and genetic algorithm (GA) to select the parameters. Lin showed that EKIFCM is more efficient than conventional FCM family algorithms such as the k-means (KM) [43], FCM, Gustafson Kessel (GKFCM) clustering algorithm [21], Gath-Geva (GG) clustering algorithm [18], Chairas intuitionistic fuzzy c-mean (IFCM) [6], and kernel-based fuzzy c-means with Gaussian kernel functions (KFCM(G)) [48]. Here, we consider UCI Machine learning data sets [17], including Iris, Wine, Contraceptive Method Choice (CMC), Wisconsin Prognostic Breast Cancer (WPBC), SPECT Heart (SH), and Glass identification Dataset. The main characteristics of these datasets are given in Table 4. At the starting point of the experiments, we implement HFC procedure using MATLAB(2017b) on a Windows version 10 platform powered with an i7 processor of 2.50GHz and a RAM of 8 GB. We set the following parameters for NFCM: $w_1 = 0.8$, $w_2 = 0.1$, $w_3 = 0.1$, $\delta = 0.5$, and $\epsilon = 10e - 5$. Also, we set $m = 2$ and $\epsilon = 10e - 5$ for FCM.

5.1 First experiments

The purpose of these experiments is to show that HFC procedure can choose the suitable clustering method between partitioning and hierarchical clustering algorithms. Since we use cluster validity indices in phase 1, then we use $Fscore$ measure to evaluate the accuracy of HFC procedure. $Fscore$ is proposed to assess the quality of the final resulting cluster [49]. Let c be the number of individual classes. Then, the total $Fscore$ is computed as the weighted sum of these classes's $Fscore$ according to their size. The $Fscore$ can be calculated using Eq. (26), in which cl_q is a class with the size of n_r and $F(C_r)$ is the $Fscore$ of the class cl_q.

$$Fscore = \sum_{r=1}^{c} \frac{n_r}{N} F(C_r). \tag{26}$$

For each class cl_q, $F(C_r)$ finds a cluster C_i that agrees with cl_q better than the other clusters. $F(C_r)$ is calculated using equation (27), where P_{C_i} is the

precision (the number of objects in the cluster C_i belonging to the class cl_q, divided by the number of objects in the cluster C_i) and R_{C_i} is the recall (the number of objects in the cluster C_i belonging to the class cl_q, divided by number of objects in the class cl_q),

$$F(C_r) = max_{C_i}\{\frac{2P_{C_i}R_{C_i}}{P_{C_i} + R_{C_i}}\}. \tag{27}$$

Table 1: Characteristics of the used datasets in the first experiments

Dataset	# of cluster	# of samples
Jain	2	373
Aggregation	7	787
Spiral	2	312
Flam	2	240
Compound	6	399
Longsquare	6	1000

The results of the first set of experiments are illustrated in Table. 2 and the best results are shown in bold style to highlight them. Some notes regarding the table and the results should be mentioned here:

- Columns 2-5 display our method's $Fscore$ values and FCM, GKFCM and linkage family, respectively. $test_{value}$ column stands for computed values using Eq. (24) for confirming the true selected clustering method. SM column stands for the clustering method chosen by HFC procedure.
- We have estimated heuristically the threshold value $\theta = 0.6$ for $test_{value}$.
- The $Fscore$ value is computed based on final decision in Step. 6 and θ value.
- The results of Table 2 show that $test_{value}$ is greater than 0.6 for the Jain data set, so, we conclude that the linkage family is appropriate. We run all linkage families and the result of the average linkage is better. As well as, for the spiral data set, the $test_{value}$ is greater than 0.6, and the single linkage is better.

Table 2: The results of first experiments

	Fscore of					
Dataset	HFC	FCM	GKFCM	Linkage	$test_{value}$	SM^1
Jain	**0.9442**	0.7800	0.7800	0.9442	0.6574	H^2
Aggregation	**0.9420**	0.8470	0.8349	0.2703	0.6000	P^3
Spiral	**1**	0.5000	0.5000	1	0.6608	H
Flam	**0.8824**	0.8721	0.8761	0.8620	0.5573	P
Compound	**0.6042**	0. 50	0.53	0.5653	0.6000	P
Longsquare	**0.8795**	0.71	0.66	0.779	0.5983	P

1 Selected method according to the proposal.

2 Hierarchical methods.

3 Partitioning methods.

Table 3: The results obtained by the WILCOXON test for the HFC procedure

Methods		**Exact P-value**
FCM	**Versus Proposed**	0.03126
GKFCM		0.03126
Linkage		0.07813

Table. 2 confirms that HFC procedure leads to good results because the $test_{value}$ of each data set persuades us to truly choose the clustering method. Also, in some cases, when the partitioning method is suitable, the proposed procedure improves the accuracy of clusters versus FCM and GKFCM. Furthermore, the Wilcoxon test [10] is carried out to analyze these results. The test results presented in Table 3 exhibits differences among the FCM, GKFCM and HFC procedures.

Table 4: Characteristics of the used datasets in the second experiments

Dataset	# of feature	# of cluster	# of samples	Attribute characteristics	Missing Values
Iris	4	3	150	Real	No
Wine	13	3	178	Integer, Real	No
CMC	9	3	1473	Categorical, Integer	No
WPBC	10	2	699	Integer	Yes
SH	22	2	267	Binary	No
Glass	10	6	214	Real	No

Table 5: The results os the second experiments

	CR of			
Dataset / methods	Proposed	EKIFCM	$test_{value}$	SM1
Iris	0.9268	**0.98**	0.4439	P2
Wine	**0.9771**	0.6910	0.8783	H3
CMC	**0.48**	0.4535	0.6142	H
WPBC	0.9642	**0.9742**	0.5496	P
SH	**0.8464**	0.7940	0.7230	H
Glass	0.6215	**0.9252**	0.5441	P

1 Selected Method according to HFC

2 Partitioning methods.

3 Hierarchical methods.

Moreover, the results of HFC procedure are significantly better than those of the FCM and GKFCM in the data sets which are clustered using partitioning methods based on the statistic test (significantly different at $\alpha = 0.05$). Since the considerable differences of $Fscore$ values are obtained using linkage family and HFC procedure for Jain and Spiral data sets, then the proposed procedure could not reject the linkage algorithm in the Wilcoxon test. However, HFC procedure by $test_{value}$ confirms that the partitioning methods are not suitable for these data sets. Therefore, based on the Wilcoxon test, it is confirmed that HFC procedure can be properly adapted to different datasets. The main reason is its suitable decision-making mechanism for applying various data clustering algorithms.

5.2 Second experiments

In the second set of experiments, we compare HFC procedure with the EKIFCM clustering algorithm. So, we apply these methods to Machine Learning Repository data sets illustrated in Table 4. We normalize the real attributes of each data set with linear transformation $\frac{x-\min}{\max-\min}$. Since 16 data points have a single missing (i.e., unavailable) attribute value in breast cancer Wisconsin WPBC, then we use mean imputation. We consider the classification rate CR as a measure for evaluating these methods which is formulated as follows:

$$CR = \frac{\text{Number of correctly classified samples}}{\text{Total number of samples}}. \tag{28}$$

We calculate the classification rate by converting the fuzzy partition matrix to a Boolean partition matrix and selecting the cluster with the maximum membership value for each pattern. The percentage of patterns in a correctly labeled cluster is called the classification rate. Better clustering results are represented by higher classification rates. We use the classification rate to assess the performance of fuzzy clustering algorithms[19]. The results of these evaluations are illustrated in Table. 5. The Table. 5 demonstrates the CR results of applying HFC procedure and EKIFCM on the data sets presented in Table 4. This table shows that it is possible to use our suggested clustering procedure to choose the best clustering algorithm for a given data set. For Iris, WPBC, and Glass data sets, the low values of the $test_{value}$ persuade us to apply one of the efficient partitioning methods and the obtained results of CR confirm this decision. For Wine, CMC and SH datasets based on the computed values of $test_{value}$ the procedure selects Bob-hic [29] clustering algorithm which leads to higher CR. We note that in some cases, the proposed procedure could not work better than EKIFCM. We could use other partition methods in the proposed hesitant fuzzy decision making method.

6 Conclusion and future work

In this paper, we proposed HFC procedure to choose the appropriate clustering algorithm (between partitioning clustering methods and hierarchical clustering methods). In our procedure, we focus on the most common types of clustering algorithms such as Fuzzy C-Means (FCM) and linkage clustering algorithms which are suitable for a particular type of data distribution.HFC procedure based on the neutrosophic objective function in the NFCM algorithm with various parameterizations helps us to select the suitable clustering algorithm. The clustering procedure works based on the hesitant fuzzy decision making method that does not depend on the data distribution. The HFC procedure includes two phases. In the first phase, we estimated the suitable cluster number based on the fuzzy cluster validity indices. In the second phase, we introduced a hesitant fuzzy decision making method from a new point of view. FCM and NFCM with various kernels and initial cluster center selection methods were modeled as two experts with diverse assessments of the cluster for each data point. Finally, the proper cluster is determined for each data based on the result of hesitant fuzzy partitions. Also, we introduce $test_{value}$ to determine whether the partitioning clustering methods are appropriate or not based on a threshold value. The results of two sets of experiments showed that HFC procedure not only makes a good decision of applying partitioning clustering algorithms but also improves the accuracy of some partitioning methods. The main aim of HFC procedure is modeling clustering algorithms to experts in MCDM approaches and shows that considering diverse points of view are efficient to make a proper decision. Also, HFC procedure illustrated the advantages of applying hesitant fuzzy decision-making method in the clustering tasks. Finally, we want to underline that this procedure, especially phase 2, can be used to various clustering algorithms or classification algorithms to achieve reliable results. Moreover, applying the other clustering algorithm methods and finding a new $test_{value}$ for HFC can be considered as future works.

Acknowledgement

The authors wish to express their appreciation for several excellent suggestions for improvements in this paper made by the referees.

References

[1] L. Aliahmadipour, A. Taghavi, E. Eslami, *An introduction to hesitant fuzzy data clustering*, 4th Iranian Joint Congress on Fuzzy and Intelligent Systems (CFIS), Zahedan, Iran, (2015), 1-4.

[2] L. Aliahmadipour, V. Torra, E. Eslami, *On hesitant fuzzy clustering and clustering of hesitant fuzzy data*, In Fuzzy Sets, Rough Sets, Multisets and Clustering, Springer, Cham, 157-168.

[3] L. Aliahmadipour, V. Torra, E. Eslami, M. Eftekhari, *A definition for hesitant fuzzy partitions*, International Journal of Computational Intelligent System, **9** (2016), 497-505.

[4] K. T. Atanassov, *Intuitionistic fuzzy sets*, Fuzzy Sets and Systems, **20** (1986), 87-96.

[5] J. C. Bezdek, *Pattern recognition with fuzzy objective function algorithms*, Plenum Press, New York, (1981), 191-203.

[6] T. Chaira, *A novel intuitionistic fuzzy C-means clustering algorithm and its application to medical images*, Applied Soft Computing, **11** (2011), 1711-1717.

[7] Y. H. Chang, C. H. Yeh, Y. W. Chang, *A new method selection approach for fuzzy group multicriteria decision making*, Applied Soft Computing, **13** (2013), 2179-2187.

[8] T. Y. Chen, H. P. Wang, Y. Y. Lu, *A multicriteria group decision-making approach based on interval-valued intuitionistic fuzzy sets: A comparative perspective*, Expert Systems and Application, **38** (2011), 7647-7658.

[9] S. L. Chiu, *Fuzzy model identification based on cluster estimation*, Journal of Intelligent and Fuzzy System, **2** (1994), 267-278.

[10] J. Demsar, *Statistical comparisons of classifiers over multiple data sets*, Journal of Machine Learning Research, **7** (2006), 1-30.

[11] R. Duwairi, M. Abu-Rahmeh, *A novel approach for initializing the spherical K-means clustering algorithm*, Simulation Modelling Practice and Theory, **54** (2015), 49-63.

[12] M. Eftekhari, A. Mehrpooya, F. Saberi-Movahed, V. Torra, *How fuzzy concepts contribute to machine learning*, Studies in Fuzziness and Soft Computing, Springer, **416**, (2022).

[13] M. Erisoglu, N. Calis, S. Sakallioglu, *A new algorithm for initial cluster centers in k-means algorithm*, Pattern Recognition Letters, **32** (2011), 1701-1705.

[14] B. S. Everitt, S. Landau, M. Leese, *Cluster analysis*, U. K.: Arnold, London, 2001.

[15] B. Farhadinia, *A series of score functions for hesitant fuzzy sets*, Information Sciences, **277** (2014), 102-110.

[16] M. Filipponea, F. Camastrab, F. Masulli, S. Rovetta, *A survey of kernel and spectral methods for clustering*, Pattern Recognition, **41** (2008), 176-190.

[17] A. Frank, A. Asuncion, *UCI machine learning repository*, University of California, School of Information and Computer Science, Irvine, CA, USA, 2010, http://archive.ics.uci.edu/ml.

[18] I. Gath, A. Geva, *Unsupervised optimal fuzzy clustering*, IEEE Transactions on Pattern Analysis and Machine Intelligence, **7** (1989), 773-781.

[19] D. Graves, W. Pedrycz, *Kernel-based fuzzy clustering and fuzzy clustering: A comparative experimental study*, Fuzzy Sets and Systems, **161** (2010), 522-543.

[20] Y. Guo, A. Sengur, *NECM: Neutrosophic evidential c-means clustering algorithm*, Neural Computing and Application, **26** (2014), 561-571.

[21] D. E. Gustafson, W. Kessel, *Fuzzy clustering with a fuzzy covariance matrix*, In: Proceedings of IEEE Conference on Decision Control, San Diego, (1979), 761-766.

[22] J. Han, M. Kamber, *Data mining: Concepts and techniques*, Morgan Kaufmann Publishers is an Imprint of Elsevier, 2009.

[23] A. Hatami-Marbini, M. Tavana, M. Moradi, F. Kangi, *A fuzzy group electre method for safety and health assessment in hazardous waste recycling facilities*, Safety Science, **51** (2013), 414-426.

[24] A. K. Jain, *Data clustering: 50 years beyond k-means*, Pattern Recognition Letters, **31** (2010), 651-666.

[25] L. Kaufman, P. J. Rousseeuw, *Finding groups in data: An introduction to cluster analysis*, John Wiley Sons, 1990.

[26] S. A. Mingoti, J. O. Lima, *Comparing SOM neural network with fuzzy c-means, K-means and traditional hierarchical clustering algorithms*, European Journal of Operational Research, **174** (2006), 1742-1759.

[27] W. Pedrycz, *Allocation of information granularity in optimization and decision-making models: Towards building the foundations of granular computing*, European Journal of Operational Research, **232** (2014), 137-145.

[28] A. H. Pilevar, M. Sukumar, *GCHL: Agrid-clustering algorithm for high-dimensional very large spatial data bases*, Pattern Recognition Letters, **26** (2005), 999-1010.

[29] E. Rashedi, A. Mirzaei, *A hierarchical clusterer ensemble method based on boosting theory*, Knowledge-Based Systems, **45** (2013), 83-93.

[30] B. Rezaee, *A cluster validity index for fuzzy clustering*, Fuzzy Sets and Systems, **161** (2010), 3014-3025.

[31] R. M. Rodriguez, L. Martinez, V. Torra, Z. S. Xu, F. Herrera, *Hesitant fuzzy sets: State of the art and future directions*, International Journal of Intelligent Systems, **29** (2014), 495-524.

[32] M. Sato, Y. Sato, *Fuzzy clustering model for fuzzy data*, Fuzzy Systems, International Joint Conference of the Fourth, (1995), 2123-2128.

[33] School of computing university of Eastern Finland P.O.Box 111 FIN-80101 Joensuu Finland, http://cs.joensuu.fi/sipu/datasets.

[34] S. Theodoridis, K. Koutroumbas, *Pattern recognition*, Elsevier, 2009.

[35] V. Torra, *Hesitant fuzzy sets*, International Journal of Intelligent Systems, **25** (2010), 529-539.

[36] V. Torra, *On the selection of m for fuzzy c-means*, Proceedings of the 2015 Conference of the International Fuzzy Systems Association and the European Society for Fuzzy Logic and Technology, (2015), 1571-1577.

[37] V. Torra, S. Miyamoto, *A definition for I-fuzzy partition*, Soft Computing, **15** (2011), 363-369.

[38] V. Torra, Y. Narukawa, *On hesitant fuzzy sets and decision*, IEEE International Conference on Fuzzy Systems, (2009), 1378-1382.

[39] J. Valente de Oliveira, W. Pedrycz, *Advances in fuzzy clustering and its applications*, Wiley, 2007.

[40] Y. K. Varshney, P. K. Muhuri, Q. M. Danish Lohani, *PIFHC: The probabilistic intuitionistic fuzzy hierarchical clustering algorithm*, Applied Soft Computing, **120** (2022), 108584.

[41] M. M. Xia, Z. S. Xu, *Hesitant fuzzy information aggregation in decision making*, International Journal of Approximate Reasoning, **52** (2011), 395-407.

[42] M. M. Xia, Z. S. Xu , N. Chen, *Some hesitant fuzzy aggregation operators with their application in group decision making*, Group Decision and Negotiation, **22** (2011), 259-279.

[43] H. J. Xing, M. H. Ha, *Further improvements in feature-weighted fuzzy C-means*, Information Sciences, **267** (2014), 1-15.

[44] Z. S. Xu, *Intuitionistic fuzzy aggregation and clustering*, Studies in Fuzziness and Soft Computing, Springer, *279* (2012), 1-284.

[45] Z. Xu, *Hesitant fuzzy sets theory*, Studies in Fuzziness and Soft Computing, **314** (2014), 474 pages.

[46] Z. S. Xu, M. M. Xia, *Distance and similarity measures for hesitant fuzzy sets*, Information Sciences, **181** (2011), 2128-2138.

[47] L. A. Zadeh, *Fuzzy sets*, Information and Control, **8** (1965), 338-353.

[48] D. Q. Zhang, S. C. Chen, *Clustering incomplete data using kernel-based fuzzy C-means algorithm*, Neural Processing Letters, **18** (2003), 155-162.

[49] Y. Zhao, G. Karypis, *Evaluation of hierarchical clustering algorithms for document datasets*, in Proceeding of 11th International Conference on Information and Knowledge Management, McLean, Virginia, USA, (2002), 515-524.