Pamukkale Univ Muh Bilim Derg, 32(4), 718-728, 2026

Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi

Pamukkale University Journal of Engineering Sciences

Intuitionistic fuzzy any relation clustering algorithm based on similarity matrix integration with intuitionistic fuzzy C-means and differential evolution optimization

Benzerlik matrisi entegrasyonu ile sezgisel bulanık C-ortalamalar ve diferansiyel evrim optimizasyonu tabanlı sezgisel bulanık herhangi ilişki kümeleme algoritması

Fatih Kutlu[1*], *Kübra Göleli*[1]

[1]Department of Analysis and Functions, Faculty of Science, Van Yüzüncü Yıl University, Van, Türkiye.
fatihkutlu@yyu.edu.tr, kubragoleli@yyu.edu.tr

Received/Geliş Tarihi: 24.06.2025 Revision/Düzeltme Tarihi: 26.10.2025 doi: 10.65206/pajes.76350
Accepted/Kabul Tarihi: 10.11.2025 Research Article/Araştırma Makalesi

Abstract

Data clustering, as a cornerstone technique in machine learning and data mining, plays a pivotal role in partitioning unlabeled datasets into distinct clusters based on inherent similarities. This study proposes the Intuitionistic Fuzzy Any Relation Clustering Algorithm (IF-ARCA) algorithm, a novel hybrid method that integrates the intuitionistic fuzzy C-means (IFCM) algorithm with the Any Relation Clustering Algorithm (ARCA). The IF-ARCA algorithm employs intuitionistic fuzzy similarity matrices (IFSM) constructed using cosine similarity (COS) and fuzzy metrics (FM), alongside dissimilarity and hesitation matrices, to enhance clustering precision. To address the inherent challenges of computational complexity and manual parameter tuning in traditional methods, the algorithm incorporates Differential Evolution (DE) optimization for automatic parameter adjustment, significantly improving performance in high-dimensional datasets. Experimental validation on UCI benchmark datasets demonstrates the superior efficacy of IF-ARCA in terms of clustering accuracy and scalability. The effectiveness of the proposed algorithm is rigorously evaluated using metrics such as F1 score, accuracy, precision, and recall, highlighting its potential for handling complex and ambiguous data.

Keywords: Intuitionistic fuzzy C-means, Differential evolution, Similarity matrices, Data clustering.

Öz

Veri kümeleme, makine öğrenimi ve veri madenciliğinin temel tekniklerinden biri olarak, etiketlenmemiş veri kümelerini içsel benzerliklerine göre farklı kümelere ayırmada kritik bir rol oynar. Bu çalışma, sezgisel bulanık C-means (IFCM) algoritması ile Any Relation Clustering Algorithm (ARCA) yönteminin entegrasyonunu içeren yeni bir melez yöntem olan Sezgisel Bulanık Herhangi-İlişki Kümeleme Algoritması'nı (IF-ARCA) önermektedir. IF-ARCA, kosinüs benzerliği ve bulanık ölçütlerle oluşturulan sezgisel bulanık benzerlik matrislerinin (IFSM) yanı sıra ayrışma ve tereddüt matrislerini kullanarak kümeleme hassasiyetini artırır. Geleneksel yöntemlerdeki yüksek hesaplama karmaşıklığı ve manuel parametre ayarlama zorluklarını gidermek amacıyla algoritma, parametreleri otomatik olarak ayarlayan Diferansiyel Evrim (DE) optimizasyonunu içerir; bu sayede yüksek boyutlu veri kümelerinde performansı önemli ölçüde iyileştirir. UCI benchmark veri kümeleri üzerindeki deneysel doğrulamalar, IF-ARCA'nın kümeleme doğruluğu ve ölçeklenebilirlik açısından üstün etkinliğini göstermektedir. Önerilen algoritmanın başarısı, F1 skoru, doğruluk, keskinlik (precision) ve geri çağırma (recall) gibi ölçütlerle titizlikle değerlendirilmiş olup, karmaşık ve belirsiz verileri işleme potansiyelini vurgulamaktadır.

Anahtar kelimeler: Sezgisel bulanık C-Ortalamalar, Diferansiyel evrim, Benzerlik matrisleri, Veri kümeleme.

1 Introduction

Data clustering, as a fundamental technique in machine learning and data mining, has garnered significant attention over recent decades. Its applications span various domains, including image segmentation, bioinformatics, customer segmentation, and anomaly detection. The primary aim of data clustering is to categorize an unlabeled dataset into distinct clusters, ensuring that data within the same cluster are similar, while data across different clusters remain distinct [1],[2]. Among many techniques, Fuzzy C-Means (FCM) stands out for its adaptability in soft clustering. Unlike other methods that categorically assign data to specific clusters, FCM allows data to belong to multiple clusters with varying degrees of membership, making it particularly effective in handling ambiguities inherent in real-world scenarios [3],[4].

Furthermore, fuzzy logic systems utilizing various inference Methods such as Mamdani, Larsen, and Tsukamoto-have shown practical success in nonlinear system control applications, including the speed control of permanent magnet synchronous motors [5]. However, FCM still faces several challenges.

A significant limitation is its exclusive reliance on membership values, which can lead to difficulties when clustering ambiguous data with unclear boundaries. Moreover, FCM tends to be sensitive to initial conditions, which increases the likelihood of the algorithm converging to local optima rather than the global optimum. These drawbacks highlight the need for more robust approaches that can handle such uncertainties more effectively. Researchers have looked into synergizing FCM with optimization techniques to mitigate these issues [6]-[8]. A novel clustering approach, Any Relation Clustering Algorithm (ARCA), was introduced in [9],[10], which focuses on a fuzzy relational matrix representing inter-relationships

*Corresponding author/Yazışılan Yazar

between data points, offering an alternative to conventional clustering approaches.

Addressing the limitations of conventional fuzzy set theory in handling ambiguities, Atanassov introduced intuitionistic fuzzy set (IFS) theory [11]. This framework offers a nuanced approach for understanding and modeling ambiguous and imprecise data. This enhanced approach incorporated additional degrees of non-membership and hesitancy, leading to the inception of intuitionistic fuzzy C-means (IFCM) clustering methods [12]-[16]. In [12] the potency of IFSs for characterizing and managing uncertain and ambiguous data was accentuated. Furthermore, an intuitionistic fuzzy C-means algorithm tailored for the clustering of IFSs was unveiled. This technique was subsequently broadened to accommodate clustering of interval-valued intuitionistic fuzzy sets (IVIFSs). The efficacy of the proposed algorithms was validated through their application to both authentic and synthetic datasets. Notably, data are represented in the form of IFS, implying that their prototypes are likewise depicted as IFS. Consequently, data affiliation with clusters is represented by fuzzy values, highlighting the inherent vagueness.

Although ARCA and IFCM algorithms have made significant contributions to addressing uncertainties and complex relationships in clustering processes, they still face some inherent limitations. Chief among these is the issue of computational efficiency, particularly when dealing with high-dimensional and large datasets. While IFCM takes into account uncertainty and hesitation values through the theory of intuitionistic fuzzy sets (IFS), balancing the parameters of the algorithm remains a challenge. Achieving a fine balance between membership, non-membership, and hesitation values requires precise parameter tuning, which increases the dependence on manual intervention. This approach not only increases computational complexity but also significantly limits the algorithm's applicability across different data types. In complex and ambiguous datasets, the necessity of effective parameter optimization can hinder IFCM's overall usability. Similarly, while ARCA represents a strong alternative by modeling relationships between data points through fuzzy relational matrices, its scalability is limited for large datasets. The calculation and storage of fuzzy relational matrices impose significant computational costs, especially for large and high-dimensional datasets. This increasing computational burden hinders the efficient application of ARCA, adversely affecting its performance on larger datasets. Thus, although ARCA is more flexible than traditional methods, its ability to handle large-scale and multi-dimensional datasets is constrained, increasing the need for optimization processes. In this context, the proposed Intuitionistic Fuzzy Any Relation Clustering Algorithm (IF-ARCA) algorithm seeks to combine the advantages of both IFCM and ARCA to address these limitations. IF-ARCA integrates the strengths of intuitionistic fuzzy sets and ARCA's fuzzy relational matrix framework, thereby enhancing the efficiency and scalability of uncertainty management. This integration is supported by the construction of intuitionistic fuzzy similarity matrices, along with corresponding dissimilarity and hesitation matrices Notably, the Sugeno negation is employed to more accurately measure degrees of dissimilarity and uncertainty. Differential Evolution (DE) optimization, which is increasingly adopted due to its ease of implementation and strong optimization capability [17], provides a solution to parameter adjustment challenges encountered in IFCM and ARCA algorithms by enabling automatic parameter tuning without the need for manual intervention. DE-based parameter optimization enhances clustering accuracy while improving the efficiency of the algorithm, particularly in high-dimensional datasets. As a result, the IF-ARCA algorithm not only reduces computational complexity but also significantly improves performance on large datasets, making it more applicable across a broader range of data.

Recent studies in the field of intuitionistic fuzzy clustering have demonstrated remarkable progress toward modeling uncertainty more effectively and automating parameter optimization through hybrid approaches. Saladi et al. proposed the Gaussian-Kernelized Enhanced Intuitionistic Fuzzy C-Means (GKEIFCM) for brain MRI tissue segmentation, integrating kernel-based distances with intuitionistic fuzzy membership representation to enhance robustness against noise [26]. Kaur et al. conducted a comprehensive review of meta-heuristic optimization strategies—including Differential Evolution (DE), Genetic Algorithms (GA), and Particle Swarm Optimization (PSO)—highlighting their increasing role in fuzzy and intuitionistic clustering [27]. Priya et al. introduced a Modified Intuitionistic Fuzzy Clustering Method (MIFCM) for microarray image segmentation, further extending IFCM's applicability to complex biological data [28]. Furthermore, Xie et al. introduced the concept of "intuitionistic neighborhood" to refine three-way decision-making models within fuzzy relational data [29], while Alcantud et al. developed decision-making and clustering algorithms based on the "scored-energy" of hesitant fuzzy soft sets, providing a refined quantitative treatment of hesitation and uncertainty [30]. In addition, Kutlu, Ayaz, and Garg integrated fuzzy metrics and Sugeno negation operators into the FCM framework using a genetic optimization strategy for enhanced MRI image segmentation [31], and Kutlu, Atan, and Castillo proposed a dual-stage optimization scheme combining fuzzy metrics and cosine similarity within IF-ARCA and Intuitionistic Fuzzy K-Nearest Neighbor (IF-KNN) frameworks, optimized via the Harris Hawks algorithm [32]. More recently, Sethia et al. presented an effective imputation framework for missing data using intuitionistic fuzzy clustering algorithms, demonstrating strong robustness across incomplete datasets [33]. Koçoğlu et al. proposed a hybrid model based on fuzzy clustering for optimizing facility locations in post-disaster management scenarios [34], while Alharbe et al. introduced a fuzzy clustering-based scheduling algorithm that improves computational efficiency in large-scale optimization environments [35]. Collectively, these works indicate a strong research trend toward integrating intuitionistic fuzzy similarity modeling with meta-heuristic optimization and uncertainty representation—trends that align closely with the methodological foundations and objectives of the proposed IF-ARCA algorithm.

In this study, IF-ARCA algorithm is introduced as a result of hybridizing IFCM algorithm, as described in with cosine and fuzzy metric-based similarity matrices [12]. The cornerstone of this algorithm lies in the derivation of the similarity matrices. The proposed approach employs cosine similarity and fuzzy metrics to construct matrices that are compatible with both the ARCA and IF-ARCA algorithms. In addition to generating similarity matrices, IF-ARCA mandates the creation of dissimilarity and hesitation matrices. At this juncture, the Sugeno negation is employed to quantify the degrees of non-similarity and uncertainty. To enhance the efficacy of the methodology, parameters identified are optimized using the DE

approach. Given that the datasets possess known labels, evaluation metrics, including the F1 score, accuracy, precision, and recall, are utilized to gauge the effectiveness of the clustering outcomes. The structure of this paper is as follows: Section 2 provides a comprehensive overview of prior algorithms and delves into the intricacies of the proposed IF-ARCA clustering algorithm. Section 3 details the experimental outcomes of FCM, ARCA, and IF-ARCA algorithms when tested on prominent University of California (UCI) datasets. Conclusions and insights are discussed in Section 4.

2 Methodological foundations and algorithm introduction

In fuzzy set theory, a fuzzy set A in a non-empty domain X is uniquely characterized by its membership function, μ_A, which maps every element in X to a degree of membership in the interval [0,1]. This degree reflects how much an element belongs to the set A, breaking away from the classical binary representation of set membership. In the realm of clustering, the notion of graded membership stands as a cornerstone for FCM algorithm. Drawing upon these graded degrees, FCM offers a nuanced and refined clustering paradigm, distinguishing itself from conventional hard clustering methodologies by permitting data points to associate with multiple clusters, each with distinct membership intensities, rather than being rigidly confined to a singular cluster. The mechanism of FCM algorithm [3], and its derivative clustering methods, can be fundamentally elucidated through the following four stages:

- **Initialization:** Decide on the number of clusters (c). Then, randomly determine the initial positions for these c cluster centers,
- **Membership Assignment:** For each data point in the dataset, compute its membership grade for all cluster centers. This grade is determined by the inverse of its distance to each center and indicates the degree to which the data point belongs to a specific cluster,
- **Update Centers:** Recalculate the position of each cluster center. This is done by taking the weighted average of all data points, where the weights are the membership grades computed in the previous step,
- **Convergence Check:** The algorithm checks if the cluster centers have changed significantly from the previous iteration. If they haven't (or if a pre-defined maximum number of iterations is reached), the algorithm stops. Otherwise, it returns to the membership assignment step.

One of the notable extensions of FCM is ARCA [9]. Instead of applying the four-stage algorithm directly to the data, ARCA operates on a relational matrix, capturing the intricacies of relationships amongst data points. The objective function, which ARCA seeks to minimize, is represented in Eqn. (1) as follows:

$$J_m(U,V) = \sum_{i=1}^{C}\sum_{k=1}^{N} u_{ik}^{m}\, d(x_k, v_i)^2 \qquad (1)$$

Where the conditions are $\forall i,k, u_{i,k} \in [0,1]$ and $\sum_{i=1}^{C} u_{i,k} = 1$ for each k. Within this formulation, the set $\{x_1, x_2, \dots, x_N\}$ symbolizes the dataset to be clustered, while $\{v_1, v_2, \dots, v_C\}$ represents the prototypes of each cluster. The value $u_{i,k}$ denotes the degree of membership of the object x_k to the cluster depicted by v_i. Though it might resemble the standard FCM's objective function, the distinction arises in the context of distance measurement. In ARCA, the distance, defined as $d(x_k, v_i) = \sqrt{\sum_{s=1}^{N}(r_{ks} - v_{is})^2}$, relies on the relationship values between x_k and x_s, represented by r_{ks}, and between the prototype v_i and x_s, depicted as v_{is}. After this, the algorithm mirrors the traditional FCM. A salient feature of ARCA is the portrayal of prototypes not merely as data feature representatives but as descriptions based on their relationships with other data points, transcending discernible features. This approach is vital for clustering datasets where points are defined more by their inherent relationships than explicit descriptors. ARCA's ability to effectively yield results on any relational structure offers a wide array of options for gauging data relationships. Alongside classic similarity metrics, fuzzy and intuitionistic fuzzy similarity measures can be employed at this juncture. IFS theory extends the concept of membership to include non-membership and hesitation degrees, which allows for a more nuanced representation of uncertainty in data. This extension is particularly useful when dealing with ambiguous or conflicting datasets, as it captures the complexity of data relationships more effectively than traditional methods. It is well-known that an intuitionistic fuzzy set A over a universe of discourse X is given by the set $A = \{(x, \mu_A(x), \eta_A(x)) : x \in X\}$, where $\mu_A : X \to [0,1]$ and $\eta_A : X \to [0,1]$ are mappings such that for every $x \in X$, the condition $\mu_A(x) + \eta_A(x) \leq 1$ holds. In this context $\mu_A(x)$ and $\eta_A(x)$ represent the membership and non-membership degrees of x in A respectively, while the value defined by $\pi_A(x) = 1 - \mu_A(x) - \eta_A(x)$ is termed the hesitation degree. This hesitation captures the ambiguity when it is unclear whether a data point fully belongs to or is excluded from a cluster, making it a critical factor in situations where data is inherently ambiguous or contradictory. The IFCM algorithm, introduced in [12], modifies the traditional FCM algorithm by incorporating IFS. In this adaptation, both the data and the prototypes are characterized as IFS, maintaining the core four-stage structure of FCM. In the IFCM algorithm, both the data points and the cluster prototypes are represented as IFS. This extension beyond classical fuzzy sets allows for a more flexible representation of uncertainty. While traditional fuzzy sets only account for membership, IFS incorporates non-membership and hesitation degrees, providing a more nuanced view of how data points relate to clusters. This added complexity is crucial when dealing with ambiguous data, as it better captures the underlying uncertainty and conflicting information. The most significant of these modifications, as seen in ARCA, pertains to the concept of distance. In IFCM, the normalized Euclidean distance is used, defined in Eqn. (2) as follows:

$$d_1(A,B) = \frac{1}{2n}\sqrt{\sum_{i=1}^{n} \Delta(\mu_{AB}(x_i))^2 + \Delta(\eta_{AB}(x_i))^2 + \Delta(\pi_{AB}(x_i))^2} \qquad (2)$$

Where A and B are IFSs defined over X and n denotes the dimensionality of X. In another distinctive divergence from conventional methods, IFCM algorithm intricately updates the prototypes by taking weighted averages of both membership and non-membership degrees, as well as the degrees of uncertainty, using the partition matrix. The innovations introduced by IFCM algorithm underscore its potential for achieving precise clustering for complex datasets, especially when dealing with ambiguous and conflicting data. The added insights from IFS can enhance performance in real-world data

science applications, offering a significant edge in clustering challenges. Our primary motivation for this study is to meld the advantages of both IFCM and ARCA algorithms by implementing IFCM algorithm over intuitive fuzzy similarity matrices. A novel approach is introduced, which can be aptly termed IF-ARCA.

The derivation of the intuitionistic fuzzy similarity matrix (IFSM) forms a central aspect of this research, aiming to encapsulate the degrees of similarity, dissimilarity, and hesitation among data points. To construct this matrix, two key methodologies are proposed, which enable a comprehensive representation of relational dynamics. These methodologies include the cosine similarity metric and a fuzzy metric-based approach, both of which are adapted to intuitionistic fuzzy environments. Furthermore, the Sugeno negation is employed to derive the dissimilarity matrix, ensuring a mathematically robust transformation of similarity values into their dissimilar counterparts.

2.1 Calculation of similarity matrices

The calculation of similarity matrices is a fundamental step in understanding the relational dynamics between data points in intuitionistic fuzzy systems. Similarity measures capture how closely related two data points are, which is crucial for constructing IFSM and for tasks such as clustering, classification, and decision-making. Two primary methodologies are commonly employed for calculating similarity: cosine similarity, which is widely used for its efficiency in high-dimensional spaces, and fuzzy metric-based similarity, which incorporates uncertainty into the calculation of proximity. Each of these approaches offers unique advantages in different contexts, ensuring that the IFSM is both flexible and robust when modeling complex relationships between data points. The following sections will delve into each method, outlining their mathematical formulations and applicability.

2.1.1 Cosine similarity methodology

The first methodology utilizes cosine similarity, a well-known metric often applied in high-dimensional spaces where the orientation of vectors is more significant than their magnitudes. Cosine similarity measures the cosine of the angle between two vectors, reflecting their directional alignment [18]. The similarity between two vectors $x = (x_1, x_2, \dots, x_n)$ and $y = (y_1, y_2, \dots, y_n)$ is given by the formula in Eqn. (3) as follows:

$$S_{cos}(x,y) = \frac{\sum_{i=1}^{n} x_i y_i}{\sqrt{\sum_{i=1}^{n} {x_i}^2 \sum_{i=1}^{n} {y_i}^2}} \tag{3}$$

This formula yields values within the range [−1,1], where 1 signifies maximum similarity (i.e., the vectors are perfectly aligned), and −1 indicates complete dissimilarity (i.e., the vectors point in opposite directions). The cosine similarity metric is particularly suitable for applications where the scale or magnitude of the data points is less relevant than their relative orientations, such as in text mining or information retrieval systems.

2.1.2 Fuzzy metric-based similarity calculation

The second methodology relies on fuzzy metric spaces to determine the proximity between data points in a manner that incorporates uncertainty and imprecision inherent in real-world data. Unlike traditional metrics, which rely on exact distances, fuzzy metrics account for the degree of fuzziness in the relationship between points, thus allowing a more flexible and realistic similarity assessment. The fuzzy metric idea was first presented in [19] and later revisited in [20] to obtain the Hausdorff topology. Instead of zeroing in on the exact distances between points, the fuzzy metric emphasizes the level of fuzziness that determines such distances. In such a space, represented by three components $(X, M, *)$, X is a set, while $*$ is a special operation called a continuous t-norm. M, on the other hand, is a fuzzy set defined over combinations of pairs from X and positive real numbers. The 3-tuple $(X, M, *)$ is said to be a fuzzy metric space if X is an arbitrary set, $*$ is a continuous t-norm and M is a fuzzy set on $X^2 \times (0, \infty)$ satisfying following conditions:

1. $M(x,y,t) > 0$
2. $M(x,y,t) = 1$ if and only if $x = y$
3. $M(x,y,t) = M(y,x,t)$
4. $M(x,y,t) * M(y,z,s) \leq M(x,z,t+s)$
5. $M(x,y,.): (0,\infty) \to [0,1]$ is continuous.

Where $x, y, z \in X$ and $t, s > 0$ [20]. In this context, the measure, $M(x,y,t)$, describes the degree to which point x is near to point y based on a value t. In essence, it quantifies the accuracy of claiming a certain closeness between x and y. Further studies offer insights on how traditional distance measurements can be transformed into these fuzzy ones. A commonly used fuzzy metric is the exponential decay function given in Eq. (4), which is defined based on the Euclidean distance as follows:

$$M(x,y,t) = e^{(-d(x,y)/t)} \tag{4}$$

where $d(x,y)$ represents the Euclidean distance between x and y, and $t > 0$ [21] is a scaling parameter that adjusts the sensitivity of the similarity measure to variations in distance. This formulation allows for a gradual decay of similarity as the distance between two points increases, while also incorporating fuzziness to account for uncertainties in the data. The sensitivity of the similarity measure to changes in distance can be controlled by adjusting the value of t, thus allowing for greater flexibility in analyzing data with heterogeneous or uncertain characteristics. The fuzzy metric-based similarity matrix M derived through this method offers a more nuanced representation of relationships in complex datasets.

2.2 Dissimilarity calculation using Sugeno's negation

Once the similarity matrix M has been computed, the corresponding dissimilarity matrix N is obtained by applying Sugeno's strong negation function. Sugeno's negation is defined in Eq. (5) as follows:

$$N(x) = \frac{1-x}{(1+\alpha x)} \tag{5}$$

Where $\alpha > -1$ is a parameter that controls the strength of negation and $x \in [0,1]$. This negation transforms similarity values into dissimilarity values in a manner that allows for adaptive sensitivity to changes in similarity based on the choice of α. By varying α, the transformation can be fine-tuned to match the specific characteristics of the data, making it a flexible and powerful tool for dissimilarity assessment.

2.3 Hesitation degree calculation

The final step in constructing the IFSM involves computing the hesitation degree, which represents the uncertainty or indecisiveness in assigning a definitive similarity or

dissimilarity value between two points. The hesitation degree $P(x, y)$ is calculated in Eq. (6) by subtracting the sum of the similarity and dissimilarity degrees from 1, as follows:

$$P(x, y) = 1 - M(x, y) - N(x, y) \tag{6}$$

This equation ensures that the total relationship between any pair of points is partitioned into similarity, dissimilarity, and hesitation. The hesitation degree captures the residual uncertainty and plays a critical role in intuitionistic fuzzy logic by acknowledging the potential for ambiguity in relational assessments.

To ensure clarity and reproducibility of the proposed methodology, the construction of the IFSM is outlined using a step-by-step pseudo-code, as presented in Table 1

Table 1. Calculation of IFSM.

Input: Dataset X containing n data points
Initialize the similarity matrix M
For each pair of data points (x, y) ∈ X:
- Calculate the cosine similarity S_cos(x, y)
- Alternatively, compute the fuzzy metric similarity M(x, y, t)
Construct the dissimilarity matrix N using Sugeno's negation:
Compute the hesitation degree matrix P:
Return the final IFSM matrix

2.4 Hesitation degree calculation

Once the necessary IFSM is established for the IF-ARCA framework, the focus shifts to a dataset denoted as $\{A_1, A_2, \dots, A_n\}$. IFSM, derived via the methods described above, is structured with rows symbolized as $X_j = \{(M_{jk}, N_{jk}, P_{jk}): k = 1 \dots n\}$. Here, M_{jk}, N_{jk} and P_{jk} provide insights into the respective degrees of similarity, dissimilarity, and hesitation between the elements A_j and A_k. Advancing this groundwork, IFCM algorithm is integrated into the matrix. Building on this foundation, the IFCM algorithm is integrated into the matrix. Within the context of IFSM, the algorithm modifies how the partition matrix is computed and redefines the cluster prototypes, marking a significant departure from its conventional implementation. The partition matrix in IF-ARCA is computed using Eq. (7) as follows:

$$u_{ij} = \frac{1}{\sum_{s=1}^{c} \left(\frac{d_1(X_j, V_i)}{d_1(X_j, V_s)} \right)^{\frac{2}{m-1}}} \tag{7}$$

Where the prototype vectors are defined by the notation $V_i = \{(VM_{ik}, VN_{ik}, PN_{ik}): k = 1 \dots n\}$. This representation reflects the relationship between the prototype V_i and A_k. Specifically, the values, VM_{jk}, VN_{jk} and VP_{jk} indicate the degrees of similarity, dissimilarity, and hesitancy between the prototype V_i and A_k, respectively. These values are calculated using Eqs. (8)–(10) as follows:

$$VM_{jk} = \frac{\sum_{i=1}^{n} {u_{ij}}^m M_{jk}}{\sum_{i=1}^{n} {u_{ij}}^m}, \tag{8}$$

$$VN_{jk} = \frac{\sum_{i=1}^{n} {u_{ij}}^m N_{jk}}{\sum_{i=1}^{n} {u_{ij}}^m} \tag{9}$$

$$VP_{jk} = \frac{\sum_{i=1}^{n} {u_{ij}}^m P_{jk}}{\sum_{i=1}^{n} {u_{ij}}^m}. \tag{10}$$

Thus, the distance is computed using Eq. (11) as follows:

$$d_1(X_j, V_i) = \frac{1}{2n} \sqrt{\sum_{k=1}^{n} \left(M_{jk} - VM_{ik}\right)^2 + \left(N_{jk} - VN_{ik}\right)^2 + \left(P_{jk} - VP_{ik}\right)^2} \tag{11}$$

The detailed steps of the IF-ARCA algorithm are presented in Table 2.

Table 2. IF-ARCA algorithm.

Begin
Compute IFSM
Fix the number of clusters c
Set a small positive number ε for convergence threshold
Set max_iter as the maximum number of iterations allowed
Set fuzziness parameter m as $m \geq 1$
Initialize membership matrix with random values, normalized row-wise
For iteration_count from 1 to max_iter Do:
Calculate cluster centers V_i for $j = 1, \dots, c$
Update the membership matrix U
Calculate the norm difference between the new membership matrix and the previous one
If the norm difference < ε Then:
Break out of the loop
Else:
Update the membership matrix and cluster centers with new values
End If
End For

FCM, along with its derivative methods, depends on several key parameters for optimal performance. Among these, the fuzzification parameter plays an indispensable role, heavily influencing the outcome of clustering. As the accuracy and effectiveness of clustering are directly tied to these parameters, their optimal selection becomes paramount. This has led researchers to delve into sophisticated optimization techniques to refine the parameters of FCM and its variants [8],[22],[23]. Differential Evolution (DE), an optimization strategy renowned for its robustness and adaptability, has emerged as a promising solution. Originally proposed by Storn and Price [24], DE has established its mark in global optimization tasks, especially those within continuous spaces. In the algorithm under consideration, DE has been expertly incorporated to adjust parameters during the formation of the similarity matrix and the ensuing clustering stage. This incorporation of DE is strategic; it harnesses the algorithm's capability to traverse expansive and complex parameter terrains. Leveraging DE's strengths, this holistic approach seeks to achieve unprecedented precision in clustering, delivering results that maximize both accuracy and reliability.

In this study, the performance of FCM, ARCA, and IF-ARCA algorithms has been tested on UCI datasets. All datasets were normalized using z-score normalization to ensure that features with different scales contributed equally to the distance and similarity computations. For each feature x_j, normalization was performed as $x_j' = (x_j - \mu_j)/\sigma_j$. Initially, FCM algorithm was directly applied to the dataset using a traditional approach. In FCM, the only parameter that needs optimization is m, which is the fuzzification parameter. To derive a similarity matrix, two different approaches have been employed for ARCA and IF-ARCA algorithms. In the first approach, the similarity matrix was generated using cosine similarity measure. While this matrix is sufficient for applying ARCA algorithm, additional

calculations were necessary for IF-ARCA algorithm, specifically for the dissimilarity and uncertainty components of the IFSM. The Sugeno negation was used at this stage to make these calculations. However, this approach demands the determination of the optimal α, hence the dissimilarity and uncertainty matrices are recalculated at each optimization step. In the alternative approach, a fuzzy metric was utilized. When calculating the similarity matrix in this approach, there is also an additional parameter t that needs optimization. In this method, the similarity matrix is updated in every step of the optimization. The parameter optimizations for all algorithms were all conducted with F1 Score chosen as the performance metric. Pseudocodes for the methods based on cosine similarity for IF-ARCA (IF-ARCA+COS+DE) and those based on fuzzy metric similarity for IF-ARCA (IF-ARCA+FM+DE) are provided below.

The pseudocode of the enhanced variants, IF-ARCA+COS+DE and IF-ARCA+FM+DE, is presented in Tables 3 and 4, respectively.

Table 3. IF-ARCA+COS+DE algorithm.

```
Begin
  Fix the number of clusters c, positive number ε for convergence threshold and maximum iteration
  load dataset and normalize
  compute initial cosine similarity matrix, M
  define true labels from dataset
  define fitness function:
    compute N and P using M matrix
    start IF-ARCA:
      initialize membership matrix
      compute prototypes using update_prototypes_IF-ARCA
      for each iteration:
        compute new membership matrix using update_membership_IF-ARCA
        update prototypes
      end IF-ARCA
    assign data points to clusters using threshold
    evaluate F1 score
  end define
  Optimize alpha and m using differential evolution with the fitness function
  recalculate N and P with optimized alpha and m
  apply IF-ARCA with optimized parameters
  assign data points to clusters
  compute and display performance metrics: accuracy, F1 score, recall, precision
```

Table 4. IF-ARCA+FM+DE algorithm.

```
Begin
  Fix the number of clusters c, positive number ε for convergence threshold and maximum iteration
  load dataset and normalize
  define true labels from dataset
  define fitness function:
    compute M matrix using fuzzy metric
    compute N and P using M matrix
    start IF-ARCA:
      initialize membership matrix
      compute cluster centroids
      for each iteration:
        update membership matrix
        update cluster centroids
      end IF-ARCA
    assign data points to clusters using threshold
    for each cluster:
      find and assign dominant label
    evaluate F1 score
  end define
  define parameter bounds for alpha, m, and t
  optimize alpha, m, and t using differential evolution with the fitness function
  extract optimized parameters: alpha_opt, m_opt, t_opt
  compute new M, N, and P matrices using optimized parameters
  apply IF-ARCA with optimized parameters
  assign data points to clusters using threshold
  for each cluster:
    find and assign dominant label
  compute and display performance metrics: accuracy, F1 score, recall, precision
  end
```

3 Experimental result

The algorithms delineated previously were meticulously scrutinized to evaluate their efficacy across a range of scenarios. Within this evaluation context, the DE-optimized FCM, ARCA, and IF-ARCA algorithms were put under rigorous examination. It's pivotal to note that both the ARCA and IF-ARCA methodologies incorporated two divergent approaches in the constitution of their similarity matrices. The complexities of constructing similarity matrices in ARCA and IF-ARCA methods require deeper exploration. Specifically, these methodologies adopted bifurcated strategies to ascertain the degree of resemblance among the data points. The first of these strategies leveraged cosine similarity, which measures the cosine of the angle between two non-zero vectors, offering an indication of their orientation and thus their similarity. In contrast, the second approach employed a fuzzy metric-based proximity measure, capitalizing on the inherent fuzziness in real-world data interpretations. It is important to note that dissimilarity and uncertainty values were extracted from these similarity measures using Sugeno negation, which helps quantify the inverse relationships and ambiguity in the data.

For the empirical substantiation, six quintessential datasets from the UCI repository were selected: Iris, Wine, Breast Cancer, Glass, Seeds, and Ecoli [25]. These datasets, each presenting distinct challenges, are widely recognized benchmarks in the machine learning domain. They were chosen specifically to test various facets of the algorithm's performance, including its ability to handle noise, overlapping classes, and unbalanced distributions. The details and reasons for selecting each dataset are elaborated below:

- **Iris:** This classic dataset contains three classes of iris plants, with four numerical features: sepal length, sepal width, petal length, and petal width. Despite its simplicity, two of the classes are easily distinguishable, while the third class presents significant overlap with the others, testing the algorithm's capacity to handle ambiguous boundaries,
- **Wine:** This dataset involves a multi-class classification problem with three types of wine, represented by 13 chemical features such as alcohol content and malic acid. The high dimensionality and correlated features make it a good candidate for assessing the algorithm's ability to manage feature redundancy and perform effective feature selection,
- **Breast Cancer (Wisconsin):** Comprising two classes (malignant and benign tumors) with 30 real-valued features, this dataset contains noisy and irregular data, which makes it ideal for evaluating the algorithm's robustness against uncertainty and outliers. The presence of overlapping clusters further complicates the classification, challenging the model's precision in discriminating similar instances,
- **Glass:** This dataset involves a multi-class classification problem with six types of glass, described by nine chemical features. The challenge lies in its unbalanced class distribution and subtle feature differences between some classes. This makes it difficult for algorithms to separate the data cleanly, providing a test of the algorithm's sensitivity to minor feature variations,
- **Seeds:** Consisting of three classes of seeds, each described by seven morphological features, this dataset is characterized by its subtle inter-class variations. While the dataset is relatively small, it poses a challenge in distinguishing between groups with minimal feature differences, testing the algorithm's performance in handling fine-grained distinctions,
- **Ecoli:** This dataset contains eight classes and seven features, presenting an imbalanced class distribution where some classes are significantly underrepresented. The challenge is not only in distinguishing between similar classes but also in accurately identifying the minority classes, testing the algorithm's ability to work well with skewed data distributions.

By choosing these datasets, the evaluation covers a wide range of clustering scenarios, from overlapping classes to imbalanced distributions and high-dimensional data. Each dataset tests a unique aspect of the algorithm, ensuring a comprehensive evaluation of its capability to manage real-world complexities. To critically assess the outcomes, a series of pivotal performance metrics are provided to measure accuracy, precision, recall, and the F1-score, each offering valuable insights into different aspects of the algorithm's performance.

- **F1 Score:** It computes the harmonic mean of precision and recall, providing a balance between the two when they diverge. It is particularly useful when there is an imbalance between classes, as it offers a single metric that balances precision and recall:

$$\frac{2TP}{2TP + FP + FN}$$

- **Accuracy:** Accuracy measures the proportion of correctly classified instances among the total instances, providing an overall view of the model's correctness. However, accuracy can be misleading in cases of imbalanced datasets, where the majority class dominates the prediction:

$$\frac{TP + TN}{TP + FP + FN + TN}$$

- **Recall (or Sensitivity):** Recall, also known as sensitivity or true positive rate, measures how well the algorithm identifies all positive instances within the dataset. High recall is essential in scenarios where missing a positive instance (false negatives) would be problematic, such as in medical diagnoses.

$$\frac{TP}{TP + FN}$$

- **Precision:** Precision focuses on the positive predictive value, indicating how many of the instances identified as positive are actually positive. This is crucial in datasets where false positives (incorrectly identified instances) could be costly or misleading.

$$\frac{TP}{TP + FP}$$

In these formulas, the terms represent:

- TP (True Positives): Correctly predicted positive instances.
- TN (True Negatives): Correctly predicted negative instances.
- FP (False Positives): Incorrectly predicted positive instances.
- FN (False Negatives): Incorrectly predicted negative instances

These metrics collectively ensure a comprehensive evaluation of the algorithm's performance across various dimensions, enabling a deeper understanding of how well the model performs in terms of both overall accuracy and its ability to minimize false positives and false negatives. Despite foundational roots in supervised learning, metrics such as the F1 Score, Accuracy, Recall, and Precision have been repurposed innovatively for the clustering paradigm in this research. Through simulating labeled datasets as if they lacked labels, the clustering mechanism endeavored to ascertain the most dominant label within each cluster. This methodological innovation transcends mere academic curiosity and adeptly simulates scenarios where datasets may lack definitive labels. Such an approach becomes crucial as it confers credibility upon the algorithms' resilience, juxtaposing derived cluster labels with their genuine counterparts, thus underlining the precision of the clustering process. Delving into the intricate specifics of the adopted approach, clustering parameters for each dataset were rigorously calibrated to mirror authentically the number of classes inherently present. This diligent synchronization sought to resonate with the data's inherent structural nuances and its natural distribution. A salient observation, particularly evident in Table 5, is the exceptional performance exhibited by the IF-ARCA+FM+DE algorithm.

Table 5. Performance of algorithms on selected UCI datasets.

	FCM+DE				
	Optimal Values	F1	Accuracy	Recall	Precision
Iris	m=2.62	0.49	0.59	0.59	0.45
Wine	m =1.10	0.60	0.70	0.71	0.54
Breast	m =2.74	0.49	0.62	0.63	0.40
Glass	m =2.78	0.47	0.54	0.54	0.46
Seeds	m =1.74	0.63	0.64	0.64	0.63
Ecoli	m =2.51	0.58	0.57	0.58	0.58
	ARCA+COS+DE				
	Optimal Values	F1	Accuracy	Recall	Precision
Iris	m =2.18	0.71	0.70	0.70	0.73
Wine	m =1.65	0.85	0.82	0.82	0.93
Breast	m =2.21	0.60	0.57	0.57	0.81
Glass	m =1.25	0.48	0.52	0.52	0.47
Seeds	m =1.27	0.83	0.83	0.84	0.85
Ecoli	m =2.44	0.69	0.70	0.69	0.69
	ARCA+FM+DE				
	Optimal Values	F1	Accuracy	Recall	Precision
Iris	m =1.30, t=9.59	0.90	0.90	0.90	0.92
Wine	m =1.11, t=1.30	0.94	0.94	0.94	0.95
Breast	m =1.17, t=8.47	0.86	0.87	0.87	0.88
Glass	m =1.25, t=14.60	0.86	0.88	0.88	0.86
Seeds	m =2.23, α =4.49	0.90	0.90	0.91	0.91
Ecoli	m =2.12, α =3.32	0.89	0.89	0.88	0.89
	IF-ARCA+COS+DE				
	Optimal Values	F1	Accuracy	Recall	Precision
Iris	m =2.72, α =3.35	0.73	0.73	0.73	0.74
Wine	m =2.51, α =3.73	0.91	0.91	0.92	0.91
Breast	m =2.97, α =3.84	0.71	0.71	0.75	0.75
Glass	m =1.48, α =5.84	0.51	0.53	0.53	0.54
Seeds	m =2.23, α =4.49	0.85	0.85	0.83	0.82
Ecoli	m =1.52, α =5.14	0.81	0.83	0.83	0.84
	IF-ARCA+FM+DE				
	Optimal Values	F1	Accuracy	Recall	Precision
Iris	m =1.30, t =2.19, α =2.57	0.96	0.96	0.96	0.95
Wine	m =1.14, t =1.14, α =9.35	0.97	0.97	0.97	0.97
Breast	m =1.49, t =2.19, α =5.9	0.89	0.89	0.88	0.90
Glass	m =1.96, t =5.89, α =6.4	1.00	1.00	1.00	1.00
Seeds	m =1.13, t =1.78, α =9.4	0.98	0.99	0.98	0.99
Ecoli	m =1.52, t =2.03, α =5.14	1.00	1.00	1.00	1.00

Upon deeper introspection, this unparalleled efficacy can be credited to its astute utilization of similarity matrices, notably those crafted through the nuanced application of the fuzzy metric. Intuitively, the finely-tuned IFSMs derived from the fuzzy metric elucidate inter-relational complexities between data points more effectively. Highlighting the innovative facet of this paper, attention is drawn to the profound success of the IF-ARCA+FM+DE methodology. Reasons for such standout performance can be attributed to its dexterous handling of similarity matrices and the inherent advantages of integrating the fuzzy metric.

Table 5 encapsulates the comprehensive results, showcasing the strengths and nuances of the tested strategies. It must be underscored that within the clustering domain, such distinguished scores are indicative of algorithmic excellence, diverging from misconceptions related to overfitting. Embedded within this comprehensive analysis is a staunch commitment to maintaining objectivity and fairness. By sourcing parameters for each algorithm from a unified optimization framework, not only was the evaluation terrain rendered uniform, but it was also robustly shielded against potential biases.

Table 6 summarizes the main parameters employed in the Differential Evolution (DE)-based optimization process. The parameter *m* represents the fuzzifier exponent that controls the degree of cluster fuzziness, while *c* corresponds to the number of clusters, which was set equal to the number of true classes since the experiments are classification-oriented. The parameters α and *t* play critical roles in defining the structure of the similarity and distance space: α determines the non-membership transformation strength in the Yager complement function applied in cosine-based models, whereas *t* acts as the softening coefficient of the fuzzy metric regulating the smoothness of distance decay. The iterative parameters *max_iter* and *error* define, respectively, the upper limit of optimization cycles and the convergence tolerance for membership updates. In the DE process, *F* (mutation factor), *CR* (crossover rate), and *pop_size* (population size) control the balance between exploration and exploitation and maintain the diversity of candidate solutions. Two additional DE settings are also defined: the **fitness function**, which minimizes the negative weighted F_1-score to maximize clustering performance, and the **termination criterion**, which stops optimization after a fixed number of iterations (*max_iter = 100*) or when no further improvement in F_1 is observed.

All parameter bounds and numerical settings were determined based on standard practices in DE-based optimization and fuzzy clustering research, ensuring stable convergence and effective global search behavior. These configurations collectively provide a robust optimization scheme that maintains consistency and reliability across different experimental scenarios. Since the experiments were classification-oriented, the number of clusters (c) was fixed to the actual number of classes in each dataset to enable direct performance comparison. Preliminary sensitivity tests with varying c values showed no substantial improvement in clustering performance, supporting the chosen configuration.

Upon analyzing the computational complexities of the five algorithms under review, distinct patterns and intricacies emerge in their respective methodological designs. The FCM+DE algorithm exhibits a baseline complexity of $O(\text{max\_iter} \times n \times c \times d)$. However, when integrated with Differential Evolution (DE), which iteratively explores a population of solutions across multiple generations G, the overall complexity increases to $O(G \times \text{max\_iter} \times n \times c \times d)$. Similarly, the ARCA+COS+DE method combines cosine similarity matrix construction with evolutionary optimization, resulting in a dual computational burden: $O(G \times \text{DE_iter} \times \text{max_iter} \times n \times c \times d)$from the optimization process and an additional $O(n^2 \times d)$from similarity computations.

The ARCA+FM+DE model, which integrates fuzzy metric computations within DE-driven optimization, achieves a more intricate complexity of $O(G \times NP \times \text{max\_iter} \times n \times c \times d)$. Similarly, the IF-ARCA-COS+DE approach maintains a comparable order, $O(G \times NP \times \text{max\_iter} \times n \times c \times d)$, while introducing an additional $O(n^2 \times d)$cost associated with the intuitionistic cosine similarity matrix. The IF-ARCA+FM+DE variant presents the most comprehensive formulation, characterized by $O(G \times NP \times n^2 \times d)$, owing to the intensive relational computations in the IFCM core coupled with DE's population-based optimization. Across all methods, the space complexity is dominated by distance or similarity matrix storage, typically $O(n^2)$. To ensure **fair and consistent evaluation**, all algorithms were executed under equivalent **computational time budgets** rather than fixed iteration counts.

Table 6. Differential evolution and clustering parameters.

Parameter	Meaning and Functional Role	Typical Range / Value	Used In
$\boldsymbol{m}$	Fuzzifier exponent controlling the degree of cluster fuzziness; higher values produce smoother membership transitions.	1.1 - 5.0	All variants
$\boldsymbol{c}$	Number of clusters corresponding to the number of true classes in the dataset, since the experiments involve classification-based clustering.	Equal to the number of actual classes	All variants
$\boldsymbol{\alpha}$	Yager complement parameter defining the non-membership transformation strength in cosine-based intuitionistic fuzzy space.	1.0 - 5.0	IF-ARCA+COS+DE, IF-ARCA+FM+DE
$\boldsymbol{t}$	Softening parameter of the fuzzy metric regulating the sensitivity of distance decay.	0.1 - 50.0	ARCA+FM+DE, IF-ARCA+FM+DE
$\boldsymbol{max\_iter}$	Maximum iteration count limiting optimization steps.	100	All variants
$\boldsymbol{error}$	Convergence tolerance; stops iteration when membership	1×10^{-3}	All variants
$\boldsymbol{F}$	Mutation factor in DE controlling global exploration strength.	0.5	All variants
$\boldsymbol{CR}$	Crossover rate defining the probability of parameter mixing in DE.	0.7	All variants
$\boldsymbol{pop\_size}$	Population size of candidate solutions in DE optimization.	30	All variants

This adjustment was made to account for the fact that each algorithm exhibits distinct per-iteration costs and convergence behaviors. Consequently, simulation times were matched across variants, allowing a more balanced assessment of efficiency and scalability under comparable computational loads. Empirical analyses confirmed that, despite their higher theoretical complexity, DE-enhanced variants achieved superior clustering accuracy and stability without incurring disproportionate runtime increases. This demonstrates that the optimization overhead introduced by DE is justified by its substantial contribution to global search effectiveness and convergence precision.

In summary, while incorporating DE increases theoretical computational demand due to its population-based iterative process, the time-controlled evaluation protocol adopted in this study ensures equitable comparison among all algorithms. This methodological alignment not only satisfies fairness criteria but also provides a more realistic view of the trade-off between computational cost and clustering performance.

4 Conclusion

In this comprehensive study, the nuanced intricacies and capabilities of DE-optimized FCM, ARCA, and IF-ARCA algorithms were meticulously dissected and scrutinized across diverse scenarios to gauge their intrinsic efficacy and adaptability. Relying on the rigor of experiments conducted using acclaimed datasets drawn from the UCI database, the intricate adaptability and versatility of the evaluated algorithms to cope with assorted complexities surfaced more clearly. The IF-ARCA+FM+DE algorithm stood out, consistently outperforming its counterparts across various performance metrics.

A deeper probe into the mechanics revealed that this performance surge stemmed from the integration of the fuzzy metric in crafting similarity matrices, a pivotal element in the clustering process. Pushing the envelope further, the study ventured into the realm of supervised learning metrics within the context of clustering. This approach was rigorously tested on ostensibly unlabeled datasets, challenging traditional norms. The ensuing results were nothing short of remarkable; the algorithm demonstrated the capability to derive precision-laden outcomes on datasets typically considered ambiguous, thereby shedding light on its immense potential for deployment in real-world, complex scenarios.

In a bid to uphold the highest standards of research ethics and rigor, this study unfailingly adhered to the principle of impartiality. Every algorithm was assessed in a standardized environment, with the optimization methodologies being uniformly applied across the board, ensuring parity in evaluations and the absence of any undue biases.

Drawing the curtains on this research, the findings cement the hypothesis that the paradigm of clustering, when fortified with similarity matrices-more so with the fuzzy metric-stands as a formidable approach, delivering impressive results on specific datasets. It's not mere hyperbole to earmark the IF-ARCA+FM+DE algorithm as a trailblazing solution for analogous problems that future research might grapple with. As this field expands, future studies aim to develop more powerful algorithms by using advanced intuitionistic fuzzy metrics and refined similarity measures. Beyond just clustering, the aspiration is to seamlessly integrate this foundational idea into a broader spectrum of machine learning algorithms, potentially transforming the field.

5 Author contributions statement

Fatih Kutlu conceptualized the study, designed the mathematical framework of the proposed method, and developed the core clustering algorithm. He also contributed to the preprocessing pipeline, parameter optimization strategies, and served as the corresponding author. Kübra Göleli implemented the algorithm, handled data preprocessing, and carried out the experimental evaluations and performance analyses. She also contributed to the development of the similarity and hesitation matrices, and co-wrote and edited the manuscript. Both authors collaboratively reviewed the results, contributed to the discussion and interpretation of findings, and approved the final version of the manuscript.

6 Ethics committee approval and conflict of interest statement

"Ethics committee approval was not required for the preparation of this manuscript". "There is no conflict of interest between any individual or institution in the preparation of this manuscript".

7 References

[1] Jain AK. "Data clustering: 50 years beyond K-means". *Pattern Recognition Letters*, 31(8), 651-666, 2010.

[2] Gan G, Ma C, Wu J**.** *Data Clustering: Theory*, Algorithms, and Applications**.** 2nd ed. Philadelphia, USA, Siam, 2020.

[3] Bezdek JC, Ehrlich R, Full W. "FCM: The fuzzy c-means clustering algorithm". *Computers & Geoscience*, 10(2-3), 191-203, 1984.

[4] Ruspini EH, Bezdek JC, Keller JM. "Fuzzy clustering: A historical perspective". *IEEE Computational Intelligence Magazine*, 14(1), 45-55, 2019.

[5] Alışkan İ, Ünsal S. "Farklı çıkarım yöntemlerine sahip bulanık mantık denetleyicileri kullanarak kalıcı mıknatıslı senkron motorun hız denetimi". *Pamukkale University Journal of Engineering Sciences*, *24*(2), 185-191, 2016.

[6] Tongbram S, Shimray BA. Singh LS, Dhanachandra N. "A novel image segmentation approach using fcm and whale optimization algorithm". *Journal of Ambient Intelligence and Humanized Computing*, 2021.

[7] Sharma R, Vashisht V, Singh U. "EEFCM-DE: Energy-efficient clustering based on fuzzy C means and differential evolution algorithm in WSNs". *IET Communications*. 13(8), 996-1007, 2019.

[8] Katarya R, Verma OP. "Recommender system with grey wolf optimizer and FCM". *Neural Computing and Applications,* 30(5), 1679-1687, 2018.

[9] Corsini P, Lazzerini B, Marcelloni F. "A Fuzzy relational custering algorithm based on a dissimilarity measure extracted from data". *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics),* 34(1), 775-781, 2004.

[10] Corsini P, Lazzerini B, Marcelloni F. "A new fuzzy relational clustering algorithm based on the fuzzy C-means algorithm". *Soft Computing*, 9(6), 439-447, 2005.

[11] Atanassov K.T. *Intuitionistic Fuzzy Sets*. *In Intuitionistic Fuzzy Sets: Theory and Applications* (pp. 1-137). Heidelberg: Physica-Verlag HD, 1999.

[12] Xu Z, Wu J. "Intuitionistic fuzzy C-means clustering algorithms". *Journal of Systems Engineering and Electronics*, 21(4), 580-590, 2010.
[13] Chaira T. "A novel intuitionistic fuzzy C means clustering algorithm and its application to medical images". *Applied Soft Computing Journal*, 11(2), 1711-1717, 2011.
[14] Kumar D, Verma H, Mehra A, Agrawal RK. "A modified intuitionistic fuzzy c-means clustering approach to segment human brain MRI image". *Multimedia Tools and Applications*, 78(10), 12663-12687, 2019.
[15] Thao NX, Ali M, Smarandache F. "An intuitionistic fuzzy clustering algorithm based on a new correlation coefficient with application in medical diagnosis". *Journal of Intelligent & Fuzzy Systems*, 36(1), 189-198, 2019.
[16] Hou WH, Wang YT, Wang, JQ, Cheng PF, Li L. "Intuitionistic fuzzy c-means clustering algorithm based on a novel weighted proximity measure and genetic algorithm". *International Journal of Machine Learning and Cybernetics*. 12(3), 859-875, 2021.
[17] Ren C, Song Z, Meng Z. "Differential Evolution with fitness-difference based parameter control and hypervolume diversity indicator for numerical optimization". *Engineering Applications of Artificial Intelligence*, 133, 108081, 2024.
[18] Xia P, Zhang L, Li F. "Learning similarity with cosine similarity ensemble". *Information Science*, 307, 39-52, 2015.
[19] Kramosil I, Michalek J. "Fuzzy Metrics and Statistical Metric Spaces". *Kybernetika*. 11(5), 336-344, 1975.
[20] George A, Veeramani P. "On some results in fuzzy metric spaces". *Fuzzy Sets and Systems*. 64(3), 395-399, 1994.
[21] Gregori V, Morillas S, Sapena A. "Examples of fuzzy metrics and applications". *Fuzzy Sets Systems*. 170(1), 95-111, 2011.
[22] Shankar B, Murugan K, Obulesu A, Shadrach FDS, Anitha R. MRI image segmentation using bat optimization algorithm with fuzzy C means clustering". *Journal of Medical Imaging Health information*, 11(3), 661-666, 2020.
[23] Soppari K, Chandra NS. "Development of improved whale optimization based FCM clustering for image watermarking". *Computer Science Review*, 37, 100287, 2020.
[24] Storn R, Price K. "Differential evolution a simple and efficient heuristic for global optimization over continuous spaces". *Journal of Global Optimization*, 11(4), 341-359, 1997.
[25] University of California, Irvine. "UCI Machine Learning Repository". https://archive.ics.uci.edu (15.05.2025).
[26] Saladi, S., Yepuganti, K., Chinthaginjala, R., Kim, T. H., & Ahmad, S. "A unique unsupervised enhanced intuitionistic fuzzy C-means for MR brain tissue segmentation". *Scientific Reports*, 14(1), 29804, 2024.
[27] Kaur A, Kumar Y, Sidhu J. "Exploring meta-heuristics for partitional clustering: Methods, metrics, datasets, and challenges". *Artificial Intelligence Review*, 57(10), 287, 2024.
[28] Priya MP, Roopa CK, Harish BS. "Modified Intuitionistic Fuzzy Clustering Method (MIFCM) for Microarray Image Spot Segmentation", *Procedia Computer Science*, 235, 878-888, 2024.
[29] Xie Y, Yang J, Luo Y, Zhang X, Luo J. "Constructing intuitionistic neighborhood based on intuitionistic fuzzy sets for three-way clustering". *International Journal of Approximate Reasoning*, 186, 109512, 2025.
[30] Alcantud JCR, Stojanović N, Djurović L, Laković M. "Decision-making and clustering algorithms based on the scored-energy of hesitant fuzzy soft sets". *International Journal of Fuzzy Systems*, 28, 492–506, 2026.
[31] Kutlu F, Ayaz İ, Garg H. "Integrating fuzzy metrics and negation operator in FCM algorithm via genetic algorithm for MRI image segmentation". *Neural Computing & Applications*, 36, 17057-17077, 2024.
[32] Kutlu F, Göleli K, Castillo O. "Enhanced classification in IF-ARCA and IF-KNN with fuzzy metrics and cosine similarity through dual stage optimization using Harris Hawks algorithm". *Signal, Image & Video Processing*, 19(3), 1162, 2025.
[33] Sethia, K., Singh, J., & Gosain, A. "An effective imputation approach for handling missing data using intuitionistic fuzzy clustering algorithms". *Discover Computing*, 28(1), 133, 2025.
[34] Koçoğlu FÖ, Esnaf Ş. "Fuzzy clustering-based hybrid model proposal on location selection for disaster stations". *Journal of Industrial and Management Optimization*, 21(11), 6632–6656, 2025.
[35] Alharbe NR. "Fuzzy clustering based scheduling algorithm for minimizing the tasks completion time in cloud computing environment". *Scientific Reports*, 15(1), 19505, 2025.