

Research Article

Fuzzy Covering-Based Three-Way Clustering

Dandan Yang 

College of Mathematics and Physics, Inner Mongolia University for Nationalities, Tongliao 028000, China

Correspondence should be addressed to Dandan Yang; yangdandan@hrbeu.edu.cn

Received 17 April 2020; Revised 28 June 2020; Accepted 11 July 2020; Published 31 July 2020

Guest Editor: Yi Su

Copyright © 2020 Dandan Yang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper investigates the three-way clustering involving fuzzy covering, thresholds acquisition, and boundary region processing. First of all, a valid fuzzy covering of the universe is constructed on the basis of an appropriate fuzzy similarity relation, which helps capture the structural information and the internal connections of the dataset from the global perspective. Due to the advantages of valid fuzzy covering, we explore the valid fuzzy covering instead of the raw dataset for RFCM algorithm-based three-way clustering. Subsequently, from the perspective of semantic interpretation of balancing the uncertainty changes in fuzzy sets, a method of partition thresholds acquisition combining linear and nonlinear fuzzy entropy theory is proposed. Furthermore, boundary regions in three-way clustering correspond to the abstaining decisions and generate uncertain rules. In order to improve the classification accuracy, the k -nearest neighbor (kNN) algorithm is utilized to reduce the objects in the boundary regions. The experimental results show that the performance of the proposed three-way clustering based on fuzzy covering and kNN-FRFCM algorithm is better than the compared algorithms in most cases.

1. Introduction

Three-way decisions (3WD) proposed by Yao [1, 2] is a hot topic in various fields in recent years. Since it was put forward, the idea of tripartition has attracted many scholars to do research. Especially recently, great progress has been made in the theoretical research and model building of three-way decisions based on rough sets. For example, Liang and Liu et al. [3–6] proposed fuzzy three-way decision models and stochastic three-way decision models to deal with real-valued or linguistic-valued decision-making problems. Qian et al. [7] established multigranulation decision-theoretic rough set model based on granular computing theory. Hu [8, 9] introduced the concept of three-way decision space and established a three-way decision model based on partially ordered sets. Qi et al. [10] investigated the 3WD model in the framework of lattice theory. Li et al. [11] have constructed a cost-sensitive sequential three-way decision model to simulate the decision-making process from coarse granularity (high cost) to fine granularity (low cost) and please refer [12–14] for further generalizations and applications of this model. Yao et al. [15] construct an

optimization-based framework for three-way approximations of fuzzy sets. In the meanwhile, for dynamic objects and attributes, some algorithms and incremental 3WD models are designed for classification of dynamic data [16, 17]. From the viewpoint of application, three-way decisions have been widely used in research fields such as pattern recognition [18, 19], artificial intelligence [20–22], engineering, managements [23], and social communities [24].

Based on the above backgrounds and work in three-way decisions, a novel method for three-way clustering based on fuzzy covering is discussed. First, the fuzzy covering of the dataset according to the reasonable fuzzy similarity relation is constructed. The fuzzy covering of the universe requires that the more similar the objects in the universe are, the more similar the corresponding fuzzy classes are. The fuzzy covering established in this way can better reflect the intrinsic relationship between objects in the universe. Therefore, clustering results will have more accuracy with valid fuzzy covering. One of the inevitable problems of clustering is threshold calculation. As is well known, for most of the three-way decision models mentioned above, we

first need to obtain the pair of partition thresholds α and β . Different thresholds lead to different decision results. The appropriate partition thresholds make the decision more accurate, whereas the inappropriate thresholds distort the decision. Traditionally, the partition thresholds are usually selected according to the experts experience in advance [25–27]. According to the loss function, Yao et al. [1] proposed a method to determine the thresholds by Bayesian risk decision theory. By using Shannon entropy as a measure of uncertainty, Deng et al. [28] present an information-theoretic approach to explain and calculate the thresholds. Zhou et al. [29] explore the shadowed set to automatically obtain the partition thresholds of the three-way decisions but cannot theoretically give a reasonable semantic explanation. To address this issue, inspired by the idea of balancing the uncertainty change of fuzzy sets, a threshold calculation method combining linear fuzzy entropy with nonlinear fuzzy entropy is proposed. This method provides a new scientific explanation for the generation of thresholds. And then, the boundary regions of three-way clustering are processed by the kNN algorithm to reduce uncertainty and improve decision accuracy.

The structure of the rest of this paper is as follows: Section 2 briefly introduces the necessary notions of three-way decisions. Section 3 focuses on constructing the fuzzy covering of the raw dataset according to the fuzzy similarity relation and some necessary conditions and discusses its related properties. In Section 4, a novel rough fuzzy C-means (FRFCM) algorithm based on valid fuzzy covering is established. Then, we investigate the partition thresholds by combining the linear and nonlinear fuzzy entropy. Furthermore, the framework for processing the boundary region of three-way clustering using the kNN algorithm is introduced. In Section 5, the validity and practicability of the algorithm are evaluated by experiment. Concluding remarks are given in Section 6.

2. Preliminaries

The basic concepts on three-way decisions are briefly reviewed in this section.

An information system is defined as a 4-tuple $S = (U, At = C \cup D, V, f)$, where $U = \{x_1, x_2, \dots, x_n\}$ denotes a finite nonempty universe, C is a nonempty finite of condition attributes, D is a nonempty finite of decision attributes, and $V = \cup_{a \in At} V_a$, where V_a is a domain of attribute a ; $f: U \times C \rightarrow V$ is an information function such that $f(x, a) \in V_a$ for every $x \in U, a \in At$. If V_a is a membership function value, then the value of object x under attribute a can be expressed as $\mu_a(x) \in [0, 1]$.

The trisecting-and-acting framework of three-way decisions is an extension of binary decision in order to overcome some shortcomings of binary decision. The traditional binary decision model only has acceptance and rejection options, which can easily lead to errors when the information available is insufficient to make an accurate judgment. Sometimes, the cost of wrong decisions is very high. Therefore, deferment decision is necessary, which allows decision makers to collect more information and

make more accurate judgment. This is a strategy that people often adopt in the decision-making process, and deferment decision is consistent with human cognition. A three-way decision model based on the evaluation function and a pair of thresholds is shown as follows.

Definition 1 (see [30]). Let U be a finite nonempty universe, v be an evaluation function, and (α, β) a pair of thresholds, $0 \leq \beta < \alpha \leq 1$, then the positive, negative, and boundary regions of any subset $A \subseteq U$ are defined as follows:

$$\begin{aligned} \text{POS}_{(\alpha, \beta)}(A) &= \{x \in U \mid v_A(x) \geq \alpha\}, \\ \text{NEG}_{(\alpha, \beta)}(A) &= \{x \in U \mid v_A(x) \leq \beta\}, \\ \text{BND}_{(\alpha, \beta)}(A) &= \{x \in U \mid \beta < v_A(x) < \alpha\}. \end{aligned} \quad (1)$$

Evaluation function is the key of decision. The result of decision-making is different with different evaluation functions. There are various evaluation functions that can be adopted. If a fuzzy membership function μ_A is used as an evaluation function, then the induced three regions are defined by the following equations [31]:

$$\begin{aligned} \text{POS}_{(\alpha, \beta)}(\mu_A) &= \{x \in U \mid \mu_A(x) \geq \alpha\}, \\ \text{NEG}_{(\alpha, \beta)}(\mu_A) &= \{x \in U \mid \mu_A(x) \leq \beta\}, \\ \text{BND}_{(\alpha, \beta)}(\mu_A) &= \{x \in U \mid \beta < \mu_A(x) < \alpha\}. \end{aligned} \quad (2)$$

The three-valued approximations of a fuzzy set is described by Zadeh [32] as follows: (1) x belongs to A , if $\mu_A(x) \geq \alpha$; (2) x does not belong to A , if $\mu_A(x) \leq \beta$; (3) and x has an indeterminate status relative to A , if $\beta < \mu_A(x) < \alpha$. These three cases correspond to the three-way decisions of the above fuzzy set. When $\alpha = 1$ and $\beta = 0$, we obtain the qualitative three-way decisions of a fuzzy set. However, the qualitative decision model of fuzzy set is very restrictive, and we generally do not select these two thresholds.

3. Fuzzy Covering and Its Validity

The focus of this section is on the method of constructing valid fuzzy covering of raw data and discusses the properties of the fuzzy covering. Let us first recall some concepts that help us to better understand fuzzy covering.

Definition 2 (see [33, 34]). Let $U = \{x_1, x_2, \dots, x_N\}$ be a finite universe and $F(U)$ be the fuzzy power set of U . For each $\gamma \in (0, 1]$, we call $P = (P_1, P_2, \dots, P_m)$ with $P_i \in F(U) (i = 1, 2, \dots, m)$, a fuzzy γ -covering of U , if $(\cup_{i=1}^m P_i)(x) \geq \gamma$ for each $x \in U$. (U, \mathbf{P}) is called a fuzzy γ -covering approximation space. If $\sum_{i=1}^m P_i(x) \geq 1$ for each $x \in U$, then \mathbf{P} is called a fuzzy covering of U . (U, \mathbf{P}) is called a fuzzy covering approximation space. $\sum_{i=1}^m P_i(x) = 1$ for each $x \in U$, then \mathbf{P} is called a fuzzy partition of U . We call (U, \mathbf{P}) a fuzzy partition approximation space.

Definition 3 (see [35]). Let σ be a mapping $\sigma: F(U) \times F(U) \rightarrow [0, 1]$. $\sigma(A, B)$ is called the degree of similarity between fuzzy sets A and B , if $\sigma(A, B)$ satisfies the following properties:

- (1) $\sigma(A, A) = 1, \forall A \in F(U)$
- (2) $\sigma(A, B) = \sigma(B, A)$
- (3) if $A \subseteq B \subseteq C$, then $\sigma(A, C) \leq \sigma(A, B) \wedge \sigma(B, C)$

Some similarity measures are listed as follows:

$$\begin{aligned}\sigma_1(A, B) &= \frac{\sum_{i=1}^N (\mu_A(x_i) \wedge \mu_B(x_i))}{\sum_{i=1}^N (\mu_A(x_i) \vee \mu_B(x_i))}, \\ \sigma_2(A, B) &= \frac{2 \sum_{i=1}^N (\mu_A(x_i) \wedge \mu_B(x_i))}{\sum_{i=1}^N (\mu_A(x_i) + \mu_B(x_i))}, \\ \sigma_3(A, B) &= 1 - \frac{1}{N} \sum_{i=1}^N |\mu_A(x_i) - \mu_B(x_i)|.\end{aligned}\quad (3)$$

The fuzzy set in this paper is constructed by fuzzy similarity relation R which satisfies the following properties. For any $x, y \in U$,

- (1) $0 \leq R(x, y) \leq 1$
- (2) $R(x, y) = R(y, x)$

For a fuzzy similarity relation $R \in F(U \times U)$, $\forall x_i \in U$, and $[x_i]_R \in F(U)$, the membership of x_j belonging to fuzzy set $[x_i]_R$ is denoted as

$$[x_i]_R(x_j) = R(x_i, x_j), \quad x_j \in U. \quad (4)$$

Obviously, if $R(x_i, x_j) = 1$, it means that x_j certainly belongs to $[x_i]_R$. Conversely, if $R(x_i, x_j) = 0$, it indicates that x_j certainly does not belong to $[x_i]_R$. $[x_i]_R$ is also called a fuzzy similarity class associated with R on U . Therefore, the set of fuzzy similarity classes $\{[x_i]_R: i = 1, 2, \dots, |U|\}$ constructed by relation R is a fuzzy covering of universe U .

In the following, we investigate the validity and related properties of the fuzzy covering of the raw dataset.

Definition 4. Let $U = \{x_1, x_2, \dots, x_N\}$ be a universe. R is the fuzzy similarity relation on U , and σ is the similarity relation on $F(U)$. $\mathbf{P} = \{[x_1], [x_2], \dots, [x_N]\}$ is a fuzzy covering of U constructed by fuzzy similarity relation R . For any $x_i \in U$, $M_i = \{x_j | R(\sigma([x_i], [x_j]), R(x_i, x_j)) \geq \varphi, x_j \in U, j \geq i, \varphi \in (0.5, 1]\}$ is the set of similarity objects with x_i . \mathbf{P} is defined as a valid fuzzy covering of U with respect to θ , if the following condition holds:

$$I(\mathbf{P}) = \frac{2 \cdot \sum_{i=1}^N \text{card}(M_i)}{N(N+1)} \geq \theta, \quad (5)$$

where $\theta \in (0.5, 1]$.

It is easy to know that the value of $I(\mathbf{P})$ depends on σ and R and the choice of φ . φ is generally assigned no less than 0.8. The closer the $I(\mathbf{P})$ is to 1, the more relation the R expresses the structure of sample space. If θ is less than 0.5, the fuzzy covering of the universe is invalid. The fuzzy covering $\mathbf{P} = \{[x_1], [x_2], \dots, [x_N]\}$ satisfies that similar objects in U have corresponding similar fuzzy classes, so the fuzzy covering \mathbf{P} more fully reflects the original distribution of objects in U .

Proposition 1. Let $\varphi_1 < \varphi_2$, then $I_{\varphi_1}(C) \geq I_{\varphi_2}(C)$.

Proof. It can be easily verified by the definition. \square

Remark 1. Let \mathbf{P}_1 and \mathbf{P}_2 be two valid fuzzy coverings of U with respect to the same θ . We choose fuzzy covering with a larger validity index $I(\cdot)$ as research data.

4. Three-Way Clustering

4.1. Rough Fuzzy C-Means Algorithm Based on Fuzzy Covering. In this section, we discuss the rough fuzzy C-means algorithm with fuzzy covering. The reason for clustering with fuzzy covering is that each fuzzy similarity class can reflect the relationship with the whole dataset, avoiding the disadvantage of excessive loss of clustering information with raw data.

The combination of fuzzy set and rough set provides an important direction for uncertain reasoning. Lingras [36] developed rough C-means (RCM) by combining the C-means clustering algorithm with rough set theory. The new clustering center is only related to the positive region and the boundary region, unlike fuzzy C-means (FCM) [37], which is related to all objects. Since there is no membership involved, rough C-means (RCM) cannot effectively deal with the uncertainty caused by overlapping boundaries. In such circumstances, Mitra et al. [25] proposed a rough fuzzy C-means (RFCM) algorithm in which it combines the advantages of both fuzzy set and rough set into the framework of the C-means clustering algorithm. When dividing objects into approximation regions, replacing the absolute distance with a fuzzy membership is the innovation of the rough fuzzy C-means. This adjustment enhances the robustness of the clustering to deal with overlapping situations. Maji et al. [26] modified the calculation of the new clustering center in the RFCM model by assuming that the objects in the lower approximation have definite weights and the objects in the boundary have fuzzy weights. In what follows, we discuss the rough fuzzy C-means of fuzzy covering (FRFCM) algorithm, which is an RFCM algorithm based on fuzzy covering of the universe.

Suppose $S_A = [[x]_1, [x]_2, \dots, [x]_N]^T$ is a valid fuzzy covering of $U = \{x_1, x_2, \dots, x_N\}$. The cluster centers are denoted as $V = \{v_1, v_2, \dots, v_C\} \subset \mathfrak{R}^N$. In the FRFCM algorithm, S_A is divided into C clusters Q_1, Q_2, \dots, Q_C . The membership of $[x_j]$ to the cluster i is

$$\mu_{ij} = \frac{1}{\sum_{k=1}^C (d_{ij}/d_{kj})^{2/(m-1)}}, \quad (6)$$

where d_{ij} is the distance between $[x_j]$ and v_i , $\mu_{ij} \in [0, 1]$, and $\sum_{i=1}^C \mu_{ij} = 1$. The parameter m is the fuzzifier greater than 1.

A two-category dataset is taken to explain the influence of different parameters m on classification. The membership degree of each object belonging to each cluster can be considered as a function which is related to relative distances and the fuzzifier parameter. Then, formula (6) translates to the following form:

$$\mu(a, m) = \begin{cases} \frac{1}{1 + (a/(1-a))^{2/(m-1)}}, & a \in [0, 1), \\ 0, & a = 1, \end{cases} \quad (7)$$

where a denotes the relative distance of an object with respect to one of the clusters.

The uncertainty caused by different fuzzifier parameter m can be illustrated in Figure 1.

It is easily to obtain that if the value of m tends to 1, the memberships are most crisp, as well as the uncertainty of the system is reduced which is suitable for three-way clustering. In this circumstance, only objects that are approximately the same distance from each cluster center are divided into boundary regions. In addition, the parameter m cannot be assigned with a very large value because as the value increases, the memberships of objects around the center of the cluster will be assigned to 1 and most objects are divided into boundary region which will increase the uncertainty of the system and the error rate of decision-making. Furthermore, the positive region of cluster may become empty.

The center vectors v_1, v_2, \dots, v_C are updated as follows:

$$v_i = \begin{cases} w_{il}A_i + w_{ib}B_i, & \text{if } \underline{R}Q_i \neq \emptyset \wedge R_bQ_i \neq \emptyset, \\ B_i, & \text{if } \underline{R}Q_i = \emptyset \wedge R_bQ_i \neq \emptyset, \\ A_i, & \text{if } \underline{R}Q_i \neq \emptyset \wedge R_bQ_i = \emptyset, \end{cases} \quad (8)$$

where $A_i = \sum_{[x_j] \in \underline{R}Q_i} \mu_{ij}^m [x_j] / \sum_{[x_j] \in \underline{R}Q_i} \mu_{ij}^m$ and $B_i = \sum_{[x_j] \in R_bQ_i} \mu_{ij}^m [x_j] / \sum_{[x_j] \in R_bQ_i} \mu_{ij}^m$ can be considered as the contributions to the center v_i by the fuzzy lower region and fuzzy boundary region, respectively. $R_bQ_i = \overline{R}Q_i - \underline{R}Q_i$ denotes the boundary region of cluster G_i , where $\overline{R}Q_i$ and $\underline{R}Q_i$ are the lower and upper approximations of cluster Q_i with respect to relation R , respectively. The weighted values w_{il} and w_{ib} usually satisfy $w_{il} + w_{ib} = 1$ and $w_{il} > w_{ib}$. In this paper, we take $w_{il} = (\text{card}(\underline{R}Q_i)) / (\text{card}(\underline{R}Q_i) + \text{card}(R_bQ_i))$ and $w_{ib} = (\text{card}(R_bQ_i)) / (\text{card}(\underline{R}Q_i) + \text{card}(R_bQ_i))$.

The approximation regions are determined by the FRFCM algorithm with the following principles: if $\mu_{pj} - \mu_{qj} \leq \delta$, where $p = \min_l d(v_l, [x_j])$ and $q = \min_{l \neq p} d(v_l, [x_j])$, then $[x_j] \in \overline{R}Q_p$ and $[x_j] \in \overline{R}Q_q$. It also means $x_j \in \overline{R}Q_p$ and $x_j \in \overline{R}Q_q$. In this case, x_j cannot be divided into the positive region of any clusters. Otherwise, $[x_j] \in \underline{R}Q_p$ and $x_j \in \underline{R}Q_p$. Due to the particularity structure of the fuzzy covering of U , the results of fuzzy covering clustering can well reflect the clustering results of the raw dataset through the above FRFCM algorithm.

4.2. Acquisition of Thresholds for Three-Way Clustering.

In this section, we firstly review the shadowed set model for computing thresholds. Then, a novel method of calculating thresholds is proposed by combining the linear and nonlinear fuzzy entropy.

The FRFCM algorithm is an important tool to deal with imprecise, incomplete, and inconsistent data. The thresholds in FRFCM which determines the formation of approximation regions should be carefully selected. The

unreasonable thresholds may cause the partition of approximate regions to be distorted, and clustering centers may deviate from the expected locations. Therefore, we should compute the partition thresholds scientifically according to some principles.

There are many methods to obtain the thresholds, and the most popular method is the shadowed set [38]. In fact, the shadowed set adopts the method of elevating and reducing membership degree, which divides the domain of fuzzy set into three regions. The corresponding membership function is as follows:

$$S_A(x) = \begin{cases} 1, & \mu_A(x) \geq \alpha, \\ 0, & \mu_A(x) \leq \beta, \\ [0, 1], & \beta < \mu_A(x) < \alpha, \end{cases} \quad (9)$$

where μ_A is the membership function of fuzzy set A .

In the following study, only discrete fuzzy systems are considered, and similar models and conclusions can be obtained for continuous fuzzy systems. According to shadowed sets theory, the following formula is proposed to calculate the minimum V value to obtain the optimal thresholds α and β :

$$V = \left| \sum_{\mu_A(x) \leq \beta} \mu_A(x) dx + \sum_{\mu_A(x) \geq \alpha} (1 - \mu_A(x)) dx - \text{card}\{x \mid \beta < \mu_A(x) < \alpha\} \right|. \quad (10)$$

However, the semantic interpretation of obtaining threshold pairs by using the above method is not very clear. Because the shadowed set model can not reasonably explain the relationship between the obtained shadowed set and the fuzziness of the raw fuzzy set, further research is needed. Various methods for measuring uncertainty are described in the literature [39]. Fuzzy entropy is an important tool to measure the uncertainty of fuzzy set and meets the following requirements.

Definition 5 (see [40]). Let $A = \{(x_i, \mu_A(x_i)), x_i \in U\}$ be a fuzzy set on the universe of discourse $U = \{x_1, x_2, \dots, x_N\}$. The fuzzy entropy of fuzzy set A is the mapping $E: F(U) \rightarrow R^+$, which satisfies the following four conditions:

- (1) $E(A) = 0$ if $A \in P(U)$
- (2) $E([1/2]_U) = \max_{A \in F(U)} E(A)$
- (3) $\forall A, B \in F(U)$, if $\mu_B \geq \mu_A \geq 1/2$ or $\mu_B \leq \mu_A \leq 1/2$, then $E(A) \geq E(B)$
- (4) $E(A^c) = E(A)$, $\forall A \in F(U)$

It is easy to verify that, for any $x_i \in U$, $\mu_A(x_i) = 0$ or $\mu_A(x_i) = 1$, the value of corresponding entropy function is 0, then the fuzzy entropy of the fuzzy set equals 0; i.e., the uncertainty of the fuzzy set is the minimum. When $\mu_A(x_i) = 1/2$ holds for any $x_i \in U$, the value of corresponding entropy function is 1, then the fuzzy set has maximum uncertainty. The commonly used linear and nonlinear fuzzy entropy functions are listed as follows [41–43]:

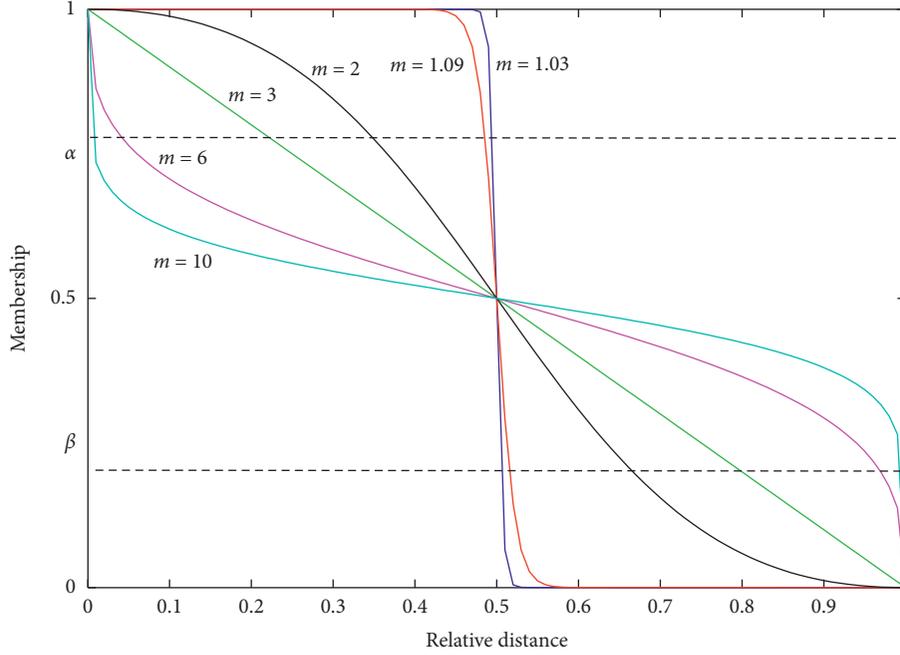


FIGURE 1: The approximate regions with different values of m .

$$\begin{aligned}
 E_A^1(x) &= 1 - |2\mu_A(x) - 1|, \quad x \in U, \\
 E_A^2(x) &= \sin \pi\mu_A(x), \quad x \in U, \\
 E_A^3(x) &= 2 \min(\mu_A(x), 1 - \mu_A(x)), \quad x \in U, \\
 E_A^4(x) &= -\frac{1}{\log 2} (\mu_A(x)\log \mu_A(x) + (1 - \mu_A(x))\log(1 - \mu_A(x))), \quad x \in U.
 \end{aligned} \tag{11}$$

With the above fuzzy entropy functions of fuzzy measure, the corresponding fuzzy entropy of the fuzzy set A can be easily obtained as follows:

$$E_j(A) = \sum_{x \in U} E_A^j(x), \quad j = 1, 2, 3, 4. \tag{12}$$

The basic idea of calculating the thresholds by fuzzy entropy is to reduce the uncertainty of the membership of the objects which are the elevating or reducing operation in

the shadowed set to 0, while the membership of objects corresponding to the middle part in the shadowed set is adjusted to the maximal uncertainty; i.e., the fuzzy degree increases to 1. In what follows, we propose a flexible fuzzy entropy method which combines the linear fuzzy entropy function $E_A^1(x_i)$ and nonlinear fuzzy entropy function $E_A^2(x_i)$ to obtain the clustering thresholds. Then, the calculation model is as follows:

$$\begin{aligned}
 (\alpha_{\text{opt}}, \beta_{\text{opt}}) &= \arg \min_{(\alpha, \beta)} \left| \lambda \left(\sum_{\substack{x_i \in U \\ \mu_A(x_i) \geq \alpha \\ \mu_A(x_i) \leq \beta}} E_A^1(x_i) - \sum_{\substack{x_i \in U \\ \beta < \mu_A(x_i) < \alpha}} (1 - E_A^1(x_i)) \right) \right. \\
 &\quad \left. + (1 - \lambda) \left(\sum_{\substack{x_i \in U \\ \mu_A(x_i) \geq \alpha \\ \mu_A(x_i) \leq \beta}} E_A^2(x_i) - \sum_{\substack{x_i \in U \\ \beta < \mu_A(x_i) < \alpha}} (1 - E_A^2(x_i)) \right) \right|,
 \end{aligned} \tag{13}$$

where $\lambda \in [0, 1]$ is a parameter adjusting the impacts of linear entropy and nonlinear entropy.

In equation (13), when $\lambda = 1$, only linear fuzzy entropy function $E_A^1(x_i)$ is used to calculate the thresholds. If $\lambda = 0$, only nonlinear fuzzy entropy function $E_A^2(x_i)$ is used to calculate the thresholds. The smaller the value of λ , the more the influence brought from the linear fuzzy entropy, and vice versa. In the subsequent experiments of this study, we assign $\lambda = 0.5$.

Figure 2 illustrates the increase and decrease in fuzzy degree of the fuzzy entropy function by taking the linear fuzzy entropy function $E_A^1(x)$, the nonlinear fuzzy entropy function $E_A^2(x)$, and the fuzzy entropy function $E_A^C(x)$ which is combined by $E_A^1(x)$ and $E_A^2(x)$ with equal weight as examples.

It can be seen from Figure 2 that the curve of flexible fuzzy entropy function lies between the curve of linear and nonlinear entropy functions. The method of using flexible fuzzy entropy to obtain the thresholds can prevent the uncertainty of fuzzy set measured by linear or nonlinear fuzzy entropy from being too small or too large, which leads to the partition thresholds unreasonable.

Thresholds used in RFCM and its related algorithms are usually user-defined. However, the threshold calculated by the above model can not only be interpreted from the change in fuzzy degree of fuzzy set but also be adjusted and optimized automatically.

According to α_{opt} and β_{opt} , the positive, boundary, and negative regions of each cluster Q_i can be expressed as

$$\begin{aligned} \text{POS}(Q_i) &= \left\{ x_j \mid \mu_{ij} \geq \alpha_{\text{opt}} \right\}, \\ \text{BND}(Q_i) &= \left\{ x_j \mid \beta_{\text{opt}} < \mu_{ij} < \alpha_{\text{opt}} \right\}, \\ \text{NEG}(Q_i) &= U - \text{POS}(Q_i) - \text{BND}(Q_i) = \left\{ x_j \mid \mu_{ij} \leq \beta_{\text{opt}} \right\}, \end{aligned} \quad (14)$$

where μ_{ij} is the membership degree of the j th object belonging to the i th class.

4.3. Boundary Region Processing of Three-Way Clustering Based on kNN Algorithm. Following the above discussion on automatically selecting the optimal partition thresholds based on fuzzy entropy theory, this section will present the object processing in the boundary regions of three-way clustering.

In the three-way clustering, the boundary region objects are rarely further processed. k -nearest neighbor (kNN) algorithm [44] is a well-known nonparametric classifier, which is considered as one of the simplest methods in data mining and pattern recognition. The principle of the kNN algorithm is to find k nearest neighbors of a query in dataset and then predicts the query with the major class in the k nearest neighbors. In this paper, the kNN algorithm will be utilized to process the objects in the boundary regions. If the object does not find a positive region, it is still classified to the boundary region. Therefore, the uncertainty of the boundary region decreases with the decrease in the number

of objects in the boundary region, and reclassifying the objects in the boundary region can improve the accuracy of the three-way clustering.

The details of updating the boundary region with the kNN algorithm are as follows.

Because the kNN algorithm mainly relies on limited adjacent objects for classification, it is more suitable than other methods for the overlap of class domain or the object set to be classified at the boundary region. Therefore, Algorithm 1 can handle the uncertainty arising from the boundary region. Of course, dealing with the boundary region with the k -nearest neighbor algorithm will add extra computing burden and may also face the risk of misclassification of objects.

In what follows, based on valid fuzzy covering, FRFCM and kNN algorithms, we proposed a three-way clustering algorithm, which is called the kNN-FRFCM algorithm, and it can be formed, as shown in Algorithm 2.

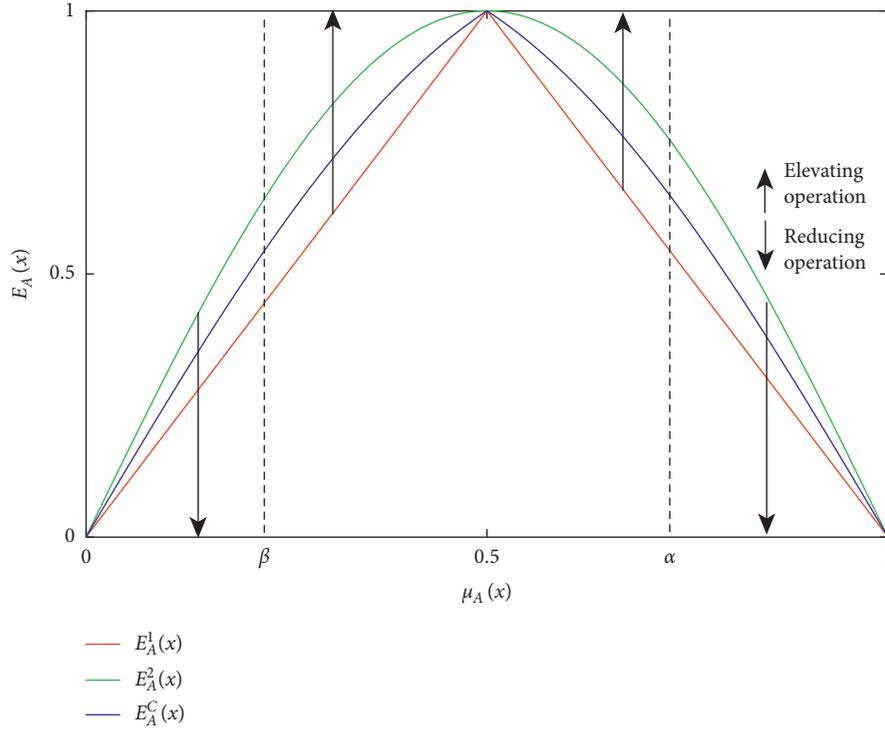
Thus, according to Algorithm 2, we obtain three-way clustering results of the original dataset by using the valid fuzzy covering.

5. Experiment Analysis

Three-way clustering method based on fuzzy covering proposed in this paper is suitable for dataset with less data and dimension or data with similar amount of data and dimension. Otherwise, clustering with the fuzzy covering constructing by the data with a large amount of data and few dimension will cause the curse of dimensionality. In this paper, six datasets include Iris, Breast Cancer Wisconsin (Original) (BCWO) which eliminates the missing data, New thyroid, Seeds, Forest-type mapping (FTM), and CT from UCI Machine Learning Repository [45] for empirical study. On these datasets and their corresponding fuzzy covering, the results of clustering methods including FCM, RCM, RFCM, kNN-RCM, and kNN-RFCM are compared. In order to distinguish the results of the raw dataset and the fuzzy covering with the same algorithm, the clustering algorithms of the fuzzy covering are expressed as FFCM, FRCM, FRFCM, kNN-FRCM, and kNN-FRFCM, respectively. Details of the six datasets are described in Table 1.

The partition threshold related to RCM and its related algorithms is set as 0.001. φ and θ involved in fuzzy covering are set as 0.8 and 0.9, respectively. The value of k in the kNN algorithm is assigned as 7, and the evaluation indexes such as the normalized mutual information (NMI) [47], ACC [48], and rand index (RI) [49] are utilized to investigate the validity of the algorithm. Furthermore, the reasonable values of fuzzifier m involved in all comparison algorithms are greater than 1. $m = 1.03$ and $m = 1.1$ are selected, and the experimental comparison results are listed in Tables 2–7.

From Tables 2–7, it can be easily concluded that the selected fuzzy parameters have a significant impact on the performance of all comparison algorithms when dealing with the same dataset. Since the boundary region is the main cause of system uncertainty, thus, too large boundary regions are not required for three-way clustering and we need to pay attention to the uncertainty caused by the fuzzifier m


 FIGURE 2: The operations of fuzzy entropy functions $E_A^1(x)$, $E_A^2(x)$, and $E_A^3(x)$.

Input: a set of objects $U = \{x_1, x_2, \dots, x_N\}$, the cluster centers $V = \{v_1, v_2, \dots, v_C\}$, the positive region $\text{POS} = \cup_{i=1}^C \text{POS}(Q_i)$, boundary region $\text{BND} = \cup_{i=1}^C \text{BND}(Q_i)$, and the optimal value of k .

Output: the updated positive region $\text{POS}(X)$ and boundary region $\text{BND}(X)$

Step 1: calculate the distance between x_i and other objects, where $x_i \in \text{BND}$;

Step 2: find the region where the k points with the smallest distance are located;

Step 3: n_{Q_i} is the number of k objects in the positive region of class Q_i , where $i = 1, 2, \dots, C$. $n_{Q_{C+1}}$ is the number of k objects in the boundary region, and $n_{Q_1} + n_{Q_2} + \dots + n_{Q_C} + n_{Q_{C+1}} = k$. If there is only one cluster Q_j , such that $n_{Q_j} = \max_{i \in \{1, 2, \dots, C+1\}} n_{Q_i}$, then $\text{POS}(Q_j) = \text{POS}(Q_j) \cup \{x_i\}$ and $\text{BND} = \text{BND} - \{x_i\}$ else $x_i \in \text{BND}$

Step 4: repeat Steps 1–3 until all boundary objects have been computed.

ALGORITHM 1: Processing the boundary regions of three-way clustering based on the kNN algorithm.

Input: the valid fuzzy covering of universe $[x_i] (i = 1, 2, \dots, N)$, the cluster centers $v_i (i = 1, 2, \dots, C)$, and the initial fuzzy membership degrees $\mu_{ij} (i = 1, 2, \dots, C, j = 1, 2, \dots, N)$;

Output: the positive, boundary, and negative regions of each cluster, respectively.

Step 1: compute the optimal partition thresholds $\alpha_{i\text{opt}}$ and $\beta_{i\text{opt}}$ for each cluster Q_i using formula (13);

Step 2: according to formula (14), determine the positive region $\text{POS}(Q_i)$, boundary region $\text{BND}(Q_i)$, and $\text{NEG}(Q_i)$ for each cluster Q_i by $\alpha_{i\text{opt}}$, $\beta_{i\text{opt}}$, and fuzzy partition matrix $(\mu_{ij})_{C \times N}$;

Step 3: update each clustering region by Algorithm 1;

Step 4: update the membership partition matrix $(\mu_{ij})_{N \times N}$ by formula (6);

Step 5: update the cluster center $v_i (i = 1, 2, \dots, C)$ with formula (8);

Step 6: repeat Step 1 to Step 5 until convergence is reached;

Step 7: the results of fuzzy covering clustering are replaced by the corresponding objects in the universe.

ALGORITHM 2: kNN-FRFCM algorithm-based three-way clustering.

TABLE 1: Description of datasets.

No.	Datasets	# objects	# attributes	# classes
1	Iris	150	4	3
2	BCWO	683	10	2
3	New thyroid	215	5	3
4	Seeds	210	7	3
5	FTM	326	27	4
6	CT	221	36	2

TABLE 2: The comparative validity results ($m = 1.03$).

	Iris			Seeds		
	NMI	ACC	RI	NMI	ACC	RI
FCM	0.7419	0.8867	0.8737	0.6949	0.8952	0.8744
RCM	0.7328	0.8400	0.8891	0.6670	0.8857	0.8666
RFCM	0.7419	0.8867	0.8737	0.6670	0.8857	0.8666
kNN-RCM	0.7777	0.9000	0.8859	0.6743	0.8905	0.8693
kNN-RFCM	0.7419	0.8867	0.8737	0.6743	0.8905	0.8693
FFCM	0.8226	0.9333	0.9195	0.6748	0.8952	0.8713
FRCM	0.7767	0.9133	0.9124	0.6748	0.8952	0.8713
FRFCM	0.8112	0.9267	0.9160	0.6777	0.8952	0.8742
kNN-FRCM	0.7919	0.9267	0.9124	0.6748	0.8952	0.8713
kNN-FRFCM	0.8226	0.9333	0.9195	0.6852	0.9000	0.8770

TABLE 3: The comparative validity results ($m = 1.03$).

	BCWO			New thyroid		
	NMI	ACC	RI	NMI	ACC	RI
FCM	0.7478	0.9605	0.9240	0.4945	0.8605	0.7908
RCM	0.7368	0.9502	0.9277	0.5585	0.8744	0.8203
RFCM	0.7585	0.9605	0.9312	0.5966	0.8884	0.8180
kNN-RCM	0.7368	0.9590	0.9277	0.5563	0.9023	0.7913
kNN-RFCM	0.7546	0.9619	0.9267	0.5966	0.8884	0.8180
FFCM	0.7759	0.9649	0.9321	0.6245	0.8977	0.8329
FRCM	0.7759	0.9649	0.9321	0.6501	0.9070	0.8531
FRFCM	0.7759	0.9649	0.9321	0.6448	0.9023	0.8523
kNN-FRCM	0.7759	0.9649	0.9321	0.6583	0.9116	0.8540
kNN-FRFCM	0.7759	0.9649	0.9321	0.6583	0.9116	0.8540

TABLE 4: The comparative validity results ($m = 1.03$).

	FTM			CT		
	NMI	ACC	RI	NMI	ACC	RI
FCM	0.7271	0.8939	0.9031	0.3118	0.8145	0.6964
RCM	0.7475	0.8990	0.9039	0.3296	0.8235	0.7080
RFCM	0.7411	0.8990	0.9018	0.3309	0.8190	0.7133
kNN-RCM	0.7475	0.8990	0.9039	0.3296	0.8235	0.7080
kNN-RFCM	0.7411	0.8990	0.9018	0.3550	0.8326	0.7234
FFCM	0.7823	0.9091	0.9153	0.4327	0.8371	0.7260
FRCM	0.7823	0.9091	0.9153	0.4327	0.8371	0.7260
FRFCM	0.7677	0.8990	0.9128	0.4267	0.8281	0.7234
kNN-FRCM	0.7823	0.9091	0.9153	0.4327	0.8371	0.7260
kNN-FRFCM	0.7906	0.9141	0.9200	0.4244	0.8326	0.7200

in the implementation of the algorithms. Moreover, the clustering results show that kNN-FRFCM algorithm has better performance than the other algorithms in most of

TABLE 5: The comparative validity results ($m = 1.1$).

	Iris			Seeds		
	NMI	ACC	RI	NMI	ACC	RI
FCM	0.7582	0.8933	0.8797	0.6949	0.8952	0.8744
RCM	0.7328	0.8400	0.8891	0.6670	0.8857	0.8666
RFCM	0.7360	0.8733	0.8714	0.6769	0.8857	0.8746
kNN-RCM	0.7777	0.9000	0.8859	0.6743	0.8905	0.8693
kNN-RFCM	0.7582	0.8933	0.8797	0.6728	0.8905	0.8740
FFCM	0.8024	0.9267	0.9124	0.6629	0.8905	0.8694
FRCM	0.7767	0.9133	0.9124	0.6748	0.8952	0.8713
FRFCM	0.7991	0.9200	0.9197	0.6345	0.8762	0.8643
kNN-FRCM	0.7919	0.9267	0.9124	0.6748	0.8952	0.8713
kNN-FRFCM	0.8136	0.9333	0.9197	0.6480	0.8857	0.8622

TABLE 6: The comparative validity results ($m = 1.1$).

	BCWO			New thyroid		
	NMI	ACC	RI	NMI	ACC	RI
FCM	0.7478	0.9605	0.9240	0.4945	0.8605	0.7908
RCM	0.7368	0.9502	0.9277	0.5585	0.8744	0.8203
RFCM	0.7391	0.9517	0.9268	0.6058	0.8884	0.8250
kNN-RCM	0.7368	0.9590	0.9277	0.5563	0.9023	0.7913
kNN-RFCM	0.7347	0.9575	0.9186	0.6058	0.8884	0.8250
FFCM	0.7759	0.9649	0.9321	0.6245	0.8977	0.8329
FRCM	0.7759	0.9649	0.9321	0.6501	0.9070	0.8531
FRFCM	0.7778	0.9649	0.9340	0.6448	0.9023	0.8523
kNN-FRCM	0.7759	0.9649	0.9321	0.6583	0.9116	0.8540
kNN-FRFCM	0.7889	0.9678	0.9376	0.6583	0.9116	0.8540

TABLE 7: The comparative validity results ($m = 1.1$).

	FTM			CT		
	NMI	ACC	RI	NMI	ACC	RI
FCM	0.7271	0.8939	0.9031	0.3118	0.8145	0.6964
RCM	0.7475	0.8990	0.9039	0.3296	0.8235	0.7080
RFCM	0.7411	0.8990	0.9018	0.3023	0.7919	0.6954
kNN-RCM	0.7475	0.8990	0.9039	0.3296	0.8235	0.7080
kNN-RFCM	0.7411	0.8990	0.9018	0.3274	0.8190	0.7055
FFCM	0.7746	0.9040	0.9107	0.4327	0.8371	0.7260
FRCM	0.7823	0.9091	0.9153	0.4327	0.8371	0.7260
FRFCM	0.7632	0.8838	0.9110	0.4338	0.8100	0.7374
kNN-FRCM	0.7823	0.9091	0.9153	0.4327	0.8371	0.7260
kNN-FRFCM	0.8074	0.9192	0.9246	0.4725	0.8416	0.7350

cases. This is mainly because it can reduce the uncertainty of the system by reprocessing the objects in the boundary regions. From the clustering results, we can also obtain that the results of clustering based on fuzzy covering are mostly better than the results of clustering with raw data. Therefore, the valid fuzzy covering can replace the raw dataset for clustering, and the clustering results are better than the raw dataset. The premise that fuzzy covering can replace the raw dataset for clustering is to select the appropriate fuzzy similarity relation [46].

6. Conclusions

In this paper, a valid fuzzy covering of the raw dataset is constructed by some principles. Because the similarity between fuzzy similarity classes in the valid fuzzy covering can be used to measure the similarity between objects in the raw dataset, each fuzzy similarity class reflects the connection with the whole dataset, so valid fuzzy covering instead of the raw data for clustering can improve the precision of clustering. From the perspective of semantic explanation of uncertainty change in fuzzy sets, we investigate the method of combining linear fuzzy entropy with nonlinear fuzzy entropy to obtain decision threshold pairs. The advantage of calculating thresholds method in this paper not only objectively obtains the classification thresholds based on the objects intrinsic relations but also the formula is simple and easy to understand, as well as the method of calculating the thresholds avoids the inappropriate subjective assignment. Additionally, the objects in the boundary region obtained by the FRFCM algorithm are reprocessed by the kNN algorithm to reduce the uncertainty of the system.

Furthermore, we will continue to investigate the method of thresholds acquisition and the processing method of boundary region for three-way clustering following the idea of this paper. The three-way clustering in incremental information system is one of the future research directions too.

Data Availability

The experimental data supporting the findings of this study are available on the website provided in this article.

Conflicts of Interest

The author declares that there are no conflicts of interest.

Acknowledgments

This work was supported by the Science Research Project of Inner Mongolia University for Nationalities with the title “Research on three-way clustering methods of preference linguistic data” (no. NMDYB18030) and Natural Science Foundation of Inner Mongolia Autonomous Region (nos. 2018MS01008 and 2020MS07008).

References

- [1] Y. Yao, “Three-way decisions with probabilistic rough sets,” *Information Sciences*, vol. 180, no. 3, pp. 341–353, 2010.
- [2] Y. Y. Yao, “An outline of a theory of three-way decisions,” in *RSTC 2012, LNCS (LNAI)*, Springer, Berlin, Germany, 2012.
- [3] D. Liang, D. Liu, W. Pedrycz, and P. Hu, “Triangular fuzzy decision-theoretic rough sets,” *International Journal of Approximate Reasoning*, vol. 54, no. 8, pp. 1087–1106, 2013.
- [4] D. Liang and D. Liu, “Deriving three-way decisions from intuitionistic fuzzy decision-theoretic rough sets,” *Information Sciences*, vol. 300, pp. 28–48, 2015.
- [5] D. Liang, W. Pedrycz, D. Liu, and P. Hu, “Three-way decisions based on decision-theoretic rough sets under linguistic assessment with the aid of group decision making,” *Applied Soft Computing*, vol. 29, pp. 256–269, 2015.
- [6] D. Liu, T. R. Li, and D. C. Liang, “Three-way decisions in stochastic decision-theoretic rough sets,” in *Transactions on Rough Sets XVIII*, Springer, Berlin, Germany, 2014.
- [7] Y. Qian, H. Zhang, Y. Sang, and J. Liang, “Multigranulation decision-theoretic rough sets,” *International Journal of Approximate Reasoning*, vol. 55, no. 1, pp. 225–237, 2014.
- [8] B. Q. Hu, “Three-way decisions space and three-way decisions,” *Information Sciences*, vol. 281, pp. 21–52, 2014.
- [9] B. Q. Hu, “Three-way decision spaces based on partially ordered sets and three-way decisions based on hesitant fuzzy sets,” *Knowledge-Based Systems*, vol. 91, pp. 16–31, 2016.
- [10] J. Qi, T. Qian, and L. Wei, “The connections between three-way and classical concept lattices,” *Knowledge-Based Systems*, vol. 91, pp. 143–151, 2016.
- [11] H. X. Li, X. Z. Zhou, B. Huang, and D. Liu, “Cost-sensitive three-way decision: a sequential strategy,” in *Proceedings of the International Conference on Rough Sets and Knowledge Technology*, Halifax, NS, Canada, October 2013.
- [12] J. Yang, G. Wang, Q. Zhang, Y. Chen, and T. Xu, “Optimal granularity selection based on cost-sensitive sequential three-way decisions with rough fuzzy sets,” *Knowledge-Based Systems*, vol. 163, pp. 131–144, 2019.
- [13] H. Ju, W. Pedrycz, H. Li, W. Ding, X. Yang, and X. Zhou, “Sequential three-way classifier with justifiable granularity,” *Knowledge-Based Systems*, vol. 163, pp. 103–119, 2019.
- [14] Y. Fang, C. Gao, and Y. Yao, “Granularity-driven sequential three-way decisions: a cost-sensitive approach to classification,” *Information Sciences*, vol. 507, pp. 644–664, 2020.
- [15] Y. Yao, S. Wang, and X. Deng, “Constructing shadowed sets and three-way approximations of fuzzy sets,” *Information Sciences*, vol. 412–413, pp. 132–153, 2017.
- [16] X. Yang, T. Li, D. Liu, H. Chen, and C. Luo, “A unified framework of dynamic three-way probabilistic rough sets,” *Information Sciences*, vol. 420, pp. 126–147, 2017.
- [17] Q. Zhang, G. Lv, Y. Chen, and G. Wang, “A dynamic three-way decision model based on the updating of attribute values,” *Knowledge-Based Systems*, vol. 142, pp. 71–84, 2018.
- [18] H. Li, L. Zhang, B. Huang, and X. Zhou, “Sequential three-way decision and granulation for cost-sensitive face recognition,” *Knowledge-Based Systems*, vol. 91, pp. 241–251, 2016.
- [19] A. V. Savchenko, “Sequential three-way decisions in multi-category image recognition with deep features based on distance factor,” *Information Sciences*, vol. 489, pp. 18–36, 2019.
- [20] Y. Zhang, Z. Zhang, D. Miao, and J. Wang, “Three-way enhanced convolutional neural networks for sentence-level sentiment classification,” *Information Sciences*, vol. 477, pp. 55–64, 2019.
- [21] A. Campagner and D. Ciucci, “Three-way and semisupervised decision tree learning based on orthopartitions,” in *Proceedings of the International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems*, Cádiz, Spain, June 2018.
- [22] H.-R. Zhang and F. Min, “Three-way recommender systems based on random forests,” *Knowledge-Based Systems*, vol. 91, pp. 275–286, 2016.
- [23] D. Liu, T. Li, and D. Liang, “Three-way government decision analysis with decision-theoretic rough sets,” *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 20, no. 1, pp. 119–132, 2012.
- [24] Y. L. Liu, L. Pan, X. Y. Jia, C. J. Wang, and J. Y. Xie, “Three-way decision based overlapping community detection,” in *Proceedings of the International Conference on Rough Sets and Knowledge Technology*, Halifax, NS, Canada, October 2013.

- [25] S. Mitra, H. Banka, and W. Pedrycz, "Rough-fuzzy collaborative clustering," *IEEE Transactions on Systems, Man and Cybernetics, Part B (Cybernetics)*, vol. 36, no. 4, pp. 795–805, 2006.
- [26] P. Maji and S. Pal, "RFCM: A hybrid clustering algorithm using rough and fuzzy sets," *Fundamenta Informaticae*, vol. 80, no. 4, pp. 475–496, 2007.
- [27] P. Maji and S. K. Pal, "Rough set based generalized fuzzy C-means algorithm and quantitative indices," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 37, no. 6, pp. 1529–1540, 2007.
- [28] X. Deng and Y. Yao, "An information-theoretic interpretation of thresholds in probabilistic rough sets," in *Proceedings of the International Conference on Rough Sets and Knowledge Technology*, Chengdu, China, August 2012.
- [29] J. Zhou, W. Pedrycz, and D. Miao, "Shadowed sets in the characterization of rough-fuzzy clustering," *Pattern Recognition*, vol. 44, no. 8, pp. 1738–1749, 2011.
- [30] Y. Yao, "Three-way decisions and cognitive computing," *Cognitive Computation*, vol. 8, no. 4, pp. 543–554, 2016.
- [31] X. Deng and Y. Yao, "Decision-theoretic three-way approximations of fuzzy sets," *Information Sciences*, vol. 279, pp. 702–715, 2014.
- [32] L. A. Zadeh, "Fuzzy sets," *Information and Control*, vol. 8, no. 3, pp. 338–353, 1965.
- [33] L. Ma, "Two fuzzy covering rough set models and their generalizations over fuzzy lattices," *Fuzzy Sets and Systems*, vol. 294, pp. 1–17, 2016.
- [34] C. Wang, D. Chen, and Q. Hu, "Fuzzy information systems and their homomorphisms," *Fuzzy Sets and Systems*, vol. 249, pp. 128–138, 2014.
- [35] G. Deng, Y. Jiang, and J. Fu, "Monotonic similarity measures between fuzzy sets and their relationship with entropy and inclusion measure," *Fuzzy Sets and Systems*, vol. 287, pp. 97–118, 2016.
- [36] P. Lingras and C. West, "Interval set clustering of web users with rough k -means," *Journal of Intelligent Information Systems*, vol. 23, no. 1, pp. 5–16, 2004.
- [37] J. C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*, Plenum Press, New York, NY, USA, 1981.
- [38] W. Pedrycz, "Interpretation of clusters in the framework of shadowed sets," *Pattern Recognition Letters*, vol. 26, no. 15, pp. 2439–2449, 2005.
- [39] G. J. Klir and T. A. Folger, *Fuzzy Sets, Uncertainty, and Information*, Prentice-Hall, Englewood Cliffs, NJ, USA, 1988.
- [40] X. C. Liu, "Entropy, distance measure and similarity measure of fuzzy sets and their relations," *Fuzzy Sets and Systems*, vol. 52, no. 3, pp. 305–318, 1992.
- [41] N. R. Pal and J. C. Bezdek, "Measuring fuzzy uncertainty," *IEEE Transactions on Fuzzy Systems*, vol. 2, no. 2, pp. 107–118, 1994.
- [42] G. J. Klir, U. H. St. Clair, and B. Yuan, *Fuzzy Set Theory: Foundations and Applications*, Prentice-Hall, Englewood Cliffs, NJ, USA, 1997.
- [43] K. Yao, "Sine entropy of uncertain set and its applications," *Applied Soft Computing*, vol. 22, pp. 432–442, 2014.
- [44] A. Mucherino, P. J. Papajorgji, and P. M. Pardalos, "K-nearest neighbor classification," *Data Mining in Agriculture*, pp. 83–106, Springer, Berlin, Germany, 2009.
- [45] D. Dua and C. Graff, *UCI Machine Learning Repository*, University of California, School of Information and Computer Science, Irvine, CA, USA, 2020, <http://archive.ics.uci.edu/ml>.
- [46] Q. Q. Gu, C. Ding, and J. W. Han, "On trivial solution and scale transfer problems in graph regularized NMF," in *Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence*, pp. 1288–1293, Barcelona, Catalonia, Spain, July 2011.
- [47] Y. Lei, J. C. Bezdek, J. Chan, N. X. Vinh, S. Romano, and J. Bailey, "Extending information-theoretic validity indices for fuzzy clustering," *IEEE Transactions on Fuzzy Systems*, vol. 25, no. 4, pp. 1013–1018, 2017.
- [48] W. Xu, X. Liu, and Y. Gong, "Document clustering based on non-negative matrix factorization," in *Proceedings of the 26th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 267–273, Toronto Canada, July 2003.
- [49] R. J. G. B. Campello, "A fuzzy extension of the rand index and other related indexes for clustering and classification assessment," *Pattern Recognition Letters*, vol. 28, no. 7, pp. 833–841, 2007.