

Research Article

A New Knowledge Characteristics Weighting Method Based on Rough Set and Knowledge Granulation

Zhenquan Shi^{1,2} and Shiping Chen ¹

¹Business School, University of Shanghai for Science and Technology, Shanghai 200093, China

²Nantong University, Nantong, Jiangsu 226017, China

Correspondence should be addressed to Shiping Chen; 56254268@qq.com

Received 12 November 2017; Revised 14 April 2018; Accepted 26 April 2018; Published 31 May 2018

Academic Editor: Paolo Gastaldo

Copyright © 2018 Zhenquan Shi and Shiping Chen. 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.

The knowledge characteristics weighting plays an extremely important role in effectively and accurately classifying knowledge. Most of the existing characteristics weighting methods always rely heavily on the experts' a priori knowledge, while rough set weighting method does not rely on experts' a priori knowledge and can meet the need of objectivity. However, the current rough set weighting methods could not obtain a balanced redundant characteristic set. Too much redundancy might cause inaccuracy, and less redundancy might cause ineffectiveness. In this paper, a new method based on rough set and knowledge granulation theories is proposed to ascertain the characteristics weight. Experimental results on several UCI data sets demonstrate that the weighting method can effectively avoid subjective arbitrariness and avoid taking the nonredundant characteristics as redundant characteristics.

1. Introduction

In data mining, in order to effectively classify the knowledge, we need to make proper assessment on the knowledge characteristics sets. Therefore, it is very important to compute the weights of characteristics sets. Weights reflect the role of characteristics in the classification process and directly affect the validity and accuracy of the classifier. The common weighting methods include experts scoring method, fuzzy statistics method [1–3], Analytic Hierarchy Process (AHP) method [4–6], and Principal Component Analysis (PCA) method [7, 8]. In these methods, the a priori knowledge must be used.

Rough set theory was firstly proposed by Pawlak in 1982 [9]. It has become an extremely useful tool to handle the imprecision and uncertainty knowledge [9, 10]. Rough set theory can be used to analyze and process the fuzzy or uncertain data without the a priori knowledge [11–17]. Now, the rough set theory has been widely used in pattern recognition [18–20], data mining [21–23], machine learning [24–29], and other fields [30–36].

In recent years, the rough set method has been studied to calculate the characteristics weight. For instance, based on the concepts of characteristics importance, Wang et al. proposed a method to determine the characteristics weights. However, this method did not consider the influence of decision characteristics on conditional characteristics [37]. Cao and Liang combined the characteristics importance of the rough set and the experts' a priori knowledge to determine the characteristics weight [38]. This method achieved the unity of the subjective a priori knowledge with the objective situations, but it ignored the internal difference in the equivalent partitions. Therefore, some nonredundant characteristics would be handled by redundant characteristics. Bao et al. proposed a method ascertaining characteristics weight based on rough set and conditional information entropy. It avoids some nonredundant characteristics to be handled by redundant characteristics. But in this method the characteristics importance obtained by redundant characteristics was higher than that got by nonredundant characteristics [39]. Zhu and Chen constructed the priority queue of characteristics importance to improve Bao's research. They presented a weighting

method based on the conditional information entropy and rough set, but that method also involved additional costs [40].

In this paper, a new knowledge characteristics weighting method based on the rough set and knowledge granulation theory is proposed. The accuracy of equivalent partitions in knowledge characteristics is studied and the difference in equivalence classes is analyzed. Experimental results on several UCI data sets confirm our theoretical results. By comparing the numerical results with those of the AHP method, the PCA method, and two rough set based methods, we can draw the conclusion that our new method can effectively avoid taking nonredundant characteristics as redundant characteristics and can improve classification accuracy.

The rest of the paper is structured as follows. Some basic concepts about rough set are briefly introduced in Section 2. In Section 3, a new knowledge characteristics weighting method is proposed and studied. Some experimental results are given in Section 4 to show the effectiveness of the proposed weighting method. Finally, we end this paper with some conclusions in Section 5.

2. Basic Concepts

2.1. Rough Set. Rough set theory takes knowledge as a partition of the objects domain. The equivalence relations and equivalence classes produced by the equivalence relations are valid information or knowledge about the objects domain. Let U denote the universe of objects, which is a nonempty set. $R \subseteq U \times U$ is the equivalence relation on U , called the knowledge on the universe U . The equivalence relation R divides U into the disjoint subsets; it is denoted as U/R or $[U]_R$, representing all the equivalence classes. For the subset X of the universe U , there are the equivalence classes $[X]_R$. In general, there are two approximation sets—the lower approximation (set) $\underline{R}(X) = \{x \mid [x]_R \subseteq X\}$ and the upper approximation (set) $\overline{R}(X) = \{x \mid [x]_R \cap X \neq \phi\}$. The lower approximation (set) of the set X is also defined as the positive region $\text{POS}(X) = \underline{R}(X)$. The set $\text{BND}_R(X) = \overline{R}(X) - \underline{R}(X)$ will be referred to as the R -boundary region of X . Obviously, when the border area is larger, the set X divided by R is rougher. Therefore, the roughness of rough set X about the equivalence relation R can be achieved; it is denoted by

$$D_R(\overline{R}(X), \underline{R}(X)) = \frac{|\text{BND}(X)|}{|\overline{R}(X)|}. \quad (1)$$

The accuracy of rough set X about the equivalence relation R is defined as

$$\rho_R(X) = 1 - D_R(\overline{R}(X), \underline{R}(X)) = \frac{|\underline{R}(X)|}{|\overline{R}(X)|}, \quad (2)$$

where $|\cdot|$ represents the number of the elements in the collection, $0 \leq \rho_R(X) \leq 1$. When $\rho_R(X) = 1$, X is defined as the accuracy set about the equivalence relation R . When $\rho_R(X) < 1$, X is defined as the rough set about the equivalence relation R .

Suppose P and Q are two equivalence relations about the universe U , if $P \subseteq Q$, for $\forall x \in U$, there is $[x]_P \subseteq [x]_Q$. Thus, the equivalence classes U/P can be considered finer than the equivalence classes U/Q and the knowledge (U, P) is more accurate than the knowledge (U, Q) ; see [37–40] for details.

2.2. Knowledge Granularity. By the rough set theory, people learn that knowledge is related to the equivalence classes, which shows that knowledge is granular. That is why some scholars also identify the structure of knowledge granularity by the equivalence classes and calculate the size of the knowledge granularity [39].

Suppose that $K = (U, R)$ is a knowledge base, and R is an equivalence relation, also known as knowledge. Knowledge granularity is defined as

$$\text{GD}(R) = \frac{|R|}{|U|^2}. \quad (3)$$

If the granularity of R reaches its minimum, then $\text{GD}(R) = |U|/|U|^2 = 1/|U|$. If R reaches the universe U , i.e., the granularity reaches its maximum, then $\text{GD}(R) = |U|^2/|U|^2 = 1$. If $(u, v) \in R$, it indicates that the objects u and v belong to the same equivalence class with the equivalence relation R ; they are indiscernible. Obviously the smaller $\text{GD}(R)$ is, the stronger the discernibility of R becomes.

Assume that R is an equivalence relation, $K = (U, R)$ is a knowledge base, and $U/R = \{X_1, X_2, \dots, X_n\}$ is the equivalence class. According to (3), the knowledge granularity can be expressed as

$$\text{GD}(R) = \sum_{i=1}^n \frac{|X_i|^2}{|U|^2}. \quad (4)$$

And the discernibility of R is defined as

$$\text{Dis}(R) = 1 - \text{GD}(R). \quad (5)$$

According to (4), there is $\text{Dis}(R) = 1 - \sum_{i=1}^n (|X_i|^2/|U|^2)$. Therefore, we have $0 \leq \text{Dis}(R) \leq 1 - 1/|U|$.

3. Knowledge Characteristics Weighting Based on Rough Set and Knowledge Granulation

Cao and Liang calculated the characteristics weights by the cardinality of the positive region set over the cardinality of the discourse set, but the results may be inaccurate [38]. For example, on the field $U = \{1, 2, 3, 4, 5, 6, 7, 8, 9\}$. Let $X = \{1, 2, 3, 4, 5, 8\}$, and let R_1 and R_2 be defined as the equivalence relation on U . Then the following equivalence classes can be obtained:

$$\begin{aligned} \frac{U}{R_1} &= \{\{1, 2, 3, 4\}, \{5, 6, 7\}, \{8, 9\}\}, \\ \frac{U}{R_2} &= \{\{1, 2\}, \{3\}, \{4\}, \{5, 6, 7\}, \{8, 9\}\}. \end{aligned} \quad (6)$$

Their positive areas about X on R_1 and R_2 are $\overline{R_1}(X) = \overline{R_2}(X) = \{1, 2, 3, 4\}$. The weight of the knowledge characteristics $w_{R_1}(X) = \text{Card}(\text{POS}_{R_1}(X))/\text{Card}(U) = 4/9$, in which $\text{Card}(X)$ represents the number of the elements in the collection X . And the weight is also shown in $w_{R_2}(X) = \text{Card}(\text{POS}_{R_2}(X))/\text{Card}(U) = 4/9$. Thus $w_{R_2}(X) = w_{R_1}(X)$. It is obvious that the characteristics weights are the same, but the equivalence classes of these two characteristics are different.

In order to solve the problems above, we use the knowledge granularity to study the relationship between the various subsets in the complex sets of the equivalence classes and propose a method based on the knowledge granularity to compute the discernibility of knowledge characteristics. Then, the knowledge characteristics weights according to the relationship between the discernibility and the weights of knowledge characteristics will be determined.

3.1. The Discernibility of Knowledge Characteristics. We first give a definition about the discernibility of the knowledge characteristics.

Definition 1. Suppose that $K = (U, R)$ is a knowledge base, R is the equivalence relation, and $r \in R$ is a characteristic. Let $U/R = \{X_1, X_2, \dots, X_n\}$ and $U/(R - \{r\}) = \{Y_1, Y_2, \dots, Y_m\}$. Then, the discernibility of r is denoted by

$$\text{Dis}(r) = \text{Dis}(R) - \text{Dis}(R - \{r\}). \quad (7)$$

By Definition 1, we know that the larger $\text{Dis}(r)$ is, the more discernible the ability of r becomes. When we select two objects randomly on U , there are $|U|^2$ ways. After adding characteristic r into $(R - \{r\})$, the characteristic discernibility increases from $|R - \{r\}|$ to $|R|$. Thus, the number of equivalence classes is more than or equal to the original set. Thus, the ability of such discernibility is improved, and the discernibility increases.

Theorem 2. Let $r \in R$, $U/R = \{X_1, X_2, \dots, X_n\}$, $U/(R - \{r\}) = \{Y_1, Y_2, \dots, Y_m\}$, and denote $\text{Dis}(r)$ as discernibility of r ; then there is $0 \leq \text{Dis}(r) \leq 1 - 1/|U|$.

Proof. From (4) and (5), we have

$$\begin{aligned} \text{Dis}(r) &= \text{Dis}(R) - \text{Dis}(R - \{r\}) \\ &= (1 - \text{GD}(R)) - (1 - \text{GD}(R - \{r\})) \\ &= \sum_{j=1}^m \frac{|Y_j|^2}{|U|^2} - \sum_{i=1}^n \frac{|X_i|^2}{|U|^2} \\ &= \frac{(\sum_{j=1}^m |Y_j|^2 - \sum_{i=1}^n |X_i|^2)}{|U|^2}. \end{aligned} \quad (8)$$

After adding characteristic r into $(R - \{r\})$, the characteristic discernibility increases from $|R - r|$ to $|R|$, and the number of

equivalence classes increases. Thus, there exists $Y_j \in U/(R - \{r\})$ ($1 \leq j \leq m$) such that $Y_j = \bigcup_{k=1}^n X_k$. And we have

$$\begin{aligned} |Y_j|^2 &= \left| \bigcup_{k=1}^n X_k \right|^2 = \left(\sum_{k=1}^n |X_k| \right)^2 \\ &\geq \sum_{k=1}^n |X_k|^2, \end{aligned} \quad (9)$$

$$\sum_{j=1}^m |Y_j|^2 - \sum_{i=1}^n |X_i|^2 \geq 0,$$

which shows $\text{Dis}(r) \geq 0$.

When the granularity of $[X]_R$ attains its minimum, there is only one element in X_i . When $U/(R - \{r\}) = \{Y_1, Y_2, \dots, Y_m\}$ reaches the universe U , $\text{Dis}(r)$ reaches its maximum. Then we obtain

$$\begin{aligned} \text{Dis}(r) &= \frac{(\sum_{j=1}^m |Y_j|^2 - \sum_{i=1}^n |X_i|^2)}{|U|^2} = \frac{(|U|^2 - |U|)}{|U|^2} \\ &= 1 - \frac{1}{|U|}. \end{aligned} \quad (10)$$

Thus, $0 \leq \text{Dis}(r) \leq 1 - 1/|U|$ is proved. \square

3.2. Method to Determine Characteristics Weight. To propose our new characteristics weight method, we further give two definitions.

Definition 3. Suppose that $K = (U, R)$ is a knowledge base and $R = C \cap D$, where C denotes the condition characteristics and D denotes the decision characteristics. $[X]_R = U/D = \{X_1, X_2, \dots, X_n\}$ identifies the equivalence classes on the universe U equivalence partitioned by the decision characteristics D . $\text{Dis}(C)$ is the discernibility of C on the universe U . The discernibility of the knowledge characteristics on $[X]_R$ is defined as

$$\text{KCDis}(C) = \rho_c(X) \text{Dis}(C). \quad (11)$$

According to (2) and (5), we have the following formulation of $\text{KCDis}(C)$:

$$\begin{aligned} \text{KCDis}(C) &= \rho_c(X) (1 - \text{GD}(C)) \\ &= \frac{|R_c(X)|}{|\overline{R_c}(X)|} \left(1 - \frac{|\overline{R_c}(X)|^2}{|U|^2} \right) \\ &= \frac{[\sum_{i=1}^n |X_i|^2 (|R_c(X_i)| / |\overline{R_c}(X_i)|) (1 - |\overline{R_c}(X_i)|^2 / |U|^2)]}{U}. \end{aligned} \quad (12)$$

Definition 4. Suppose that $K = (U, R)$ is a knowledge base and $R = C \cap D$, where C is the condition characteristics and D is the decision characteristics. $[X]_R = U/D = \{X_1, X_2, \dots, X_n\}$ identifies the equivalence classes on the universe U equivalence partitioned by the decision characteristics D . For condition characteristics $c \in C$, the

Input: The knowledge base $K = (U, R)$, $R = C \cap D$;
Output: the weight of the characteristic, $W(c_i)$;
(1) compute the equivalence class $[X]_C = U/C$
(2) **for** $i = 1$ to n **do**
(3) compute the equivalence class $[Y_i]_D = U/D$
(4) **end for**
(5) **for** $j = 1$ to m **do**
(6) compute the equivalence class $[X_j]_{(C-c_j)} = U/(C - c_j)$
(7) **for** $i = 1$ to n **do**
(8) compute the upper approximation on the set $[Y_i]_D, \overline{R_{(C-c_j)}}$
(9) compute the lower approximation on the set $[Y_i]_D, \underline{R_{(C-c_j)}}$
(10) **end for**
(11) compute the discernibility of the knowledge characteristics $KCDis(c_j)$
(12) **end for**
(13) **for** $j = 1$ to m **do**
(14) compute the discernibility of the knowledge characteristics $W(c_j)$
(15) **end for**

ALGORITHM 1: Method to determine characteristics weight.

discernibility of C is $KCDis(C)$ and the discernibility of $(C - \{c\})$ is $KCDis(C - \{c\})$. Then the discernibility of the c ($\forall c \in C$) is defined as

$$KCDis(c) = KCDis(C) - KCDis(C - \{c\}). \quad (13)$$

According to Definitions 3 and 4, we present a new formula to compute the weight of characteristic in the following definition. Detailed computation process is shown in Algorithm 1.

Definition 5. Suppose that $K = (U, R)$ is a knowledge base and $R = C \cap D$, where C denotes the condition characteristics and D denotes the decision characteristics. $[X]_R = U/D = \{X_1, X_2, \dots, X_n\}$ identifies the equivalence classes on the universe U equivalence partitioned by the decision characteristics D . $KCDis(C)$ is the discernibility of the knowledge characteristics on $[X]_R$ equivalence partitioned by the condition characteristics C . For any conditional characteristics $c \in C$, the weight of the characteristic is defined as

$$W(c) = \frac{KCDis(c)}{\sum_{c \in C} KCDis(c)}. \quad (14)$$

Theorem 6. Assume that $X = U/D = \{X_1, X_2, \dots, X_n\}$ is the equivalence class on the universe U equivalence partitioned

by the characteristics D . For any condition characteristics $c \in C$, $KCDis(c)$ is the discernibility of c to U , and it satisfies $0 \leq KCDis(c) \leq 1 - 1/|U|$.

Proof. By the rough set theory, we know that

$$0 \leq \rho_c(X) = \frac{|R_C(X)|}{|\overline{R_C(X)}|} \leq 1. \quad (15)$$

According to Theorem 2, we have $0 \leq Dis(C) \leq 1 - 1/|U|$. Thus it is easy to check that $0 \leq KCDis(C) = \rho_C(X)Dis(C) \leq 1 - 1/|U|$. \square

Theorem 7. Assume that $X = U/D = \{X_1, X_2, \dots, X_n\}$ is the equivalence class on the universe U equivalence partitioned by the decision characteristics D . For any condition characteristics $c \in C$, $KCDis(c)$ is the discernibility of c to U . Then

- (1) if P and Q are two equivalence relations on U and $P \subseteq Q$, then $KCDis(P) \leq KCDis(Q)$;
- (2) $0 \leq KCDis(c) \leq 1 - 1/|U|$.

Proof. According to Definition 3, there are $Y = \overline{R_C(X)} = \{Y_1, Y_2, \dots, Y_n\}$ and $Z = \overline{R_{C-\{c\}}(X)} = \{Z_1, Z_2, \dots, Z_m\}$. According to (11), there is

$$\begin{aligned} KCDis(c) &= \rho_C(X) Dis(C) - \rho_{C-\{c\}}(X) Dis(C - \{c\}) \\ &= \frac{\left\{ \sum_{i=1}^n |X_i| \left[\left(\frac{|R_C(X_i)|}{|\overline{R_C(X_i)}|} \right) \left(1 - \frac{|\overline{R_C(X_i)}|^2}{|U|^2} \right) - \left(\frac{|R_{C-\{c\}}(X_i)|}{|\overline{R_{C-\{c\}}(X_i)}|} \right) \left(1 - \frac{|\overline{R_{C-\{c\}}(X_i)}|^2}{|U|^2} \right) \right] \right\}}{|U|} \\ &= \frac{\left\{ \sum_{i=1}^n |X_i| \left[\left(\frac{|R_C(X_i)|}{|\overline{R_C(X_i)}|} \right) \sum_{j=1}^n \left(1 - \frac{|Y_j|^2}{|U|^2} \right) - \left(\frac{|R_{C-\{c\}}(X_i)|}{|\overline{R_{C-\{c\}}(X_i)}|} \right) \sum_{k=1}^m \left(1 - \frac{|Z_k|^2}{|U|^2} \right) \right] \right\}}{|U|}. \end{aligned} \quad (16)$$

- (1) For the universe U , P and Q are two equivalence relations on the universe U . Let $P = Q - \{q\}$ ($\forall q \in Q$). There are $Y = \overline{R_Q}(X) = \{Y_1, Y_2, \dots, Y_n\}$ and $Z = \overline{R_{Q-\{q\}}}(X) = \{Z_1, Z_2, \dots, Z_m\}$. There exists $Z_j \in X/(Q - \{q\})$ ($1 \leq j \leq m$) such that $Z_j = \bigcup_{k=1}^n Y_k$.

For the universe U , there are $\overline{R_Q}(X) \supseteq \overline{R_{Q-\{q\}}}(X) = \overline{R_P}(X)$ and $\overline{R_Q}(X) \subseteq \overline{R_{Q-\{q\}}}(X) = \overline{R_P}(X)$. So the following is satisfied:

$$\rho_P(X) = \frac{|\overline{R_P}(X)|}{|U|} \leq \rho_Q(X) = \frac{|\overline{R_Q}(X)|}{|U|}. \quad (17)$$

When $Z_j = \bigcup_{k=1}^n Y_k$, we have $|Z_j|^2 = |\bigcup_{k=1}^n Y_k|^2 = (\sum_{k=1}^n |Y_k|^2) \geq \sum_{k=1}^n |Y_k|^2$ and

$$\sum_{j=1}^m \left(1 - \frac{|Y_j|^2}{|U|^2}\right) \geq \sum_{k=1}^m \left(1 - \frac{|Z_k|^2}{|U|^2}\right). \quad (18)$$

Substituting (18) into (16), we obtain

$$\sum_{i=1}^n \frac{|X_i|}{|U|} \frac{|\overline{R_Q}(X_i)|}{|\overline{R_Q}(X_i)|} \sum_{j=1}^n \frac{|Y_j|}{|U|^2} \geq \sum_{i=1}^n \frac{|X_i|}{|U|} \frac{|\overline{R_P}(X_i)|}{|\overline{R_P}(X_i)|} \sum_{k=1}^m \frac{|Z_k|}{|U|^2}. \quad (19)$$

Therefore, $\text{KCDis}(P) \leq \text{KCDis}(Q)$.

- (2) From (16) and (19), we have $\text{KCDis}(c) \geq 0$. When C becomes the universe U , it partitions the universe U into equivalence classes (one class comprises individual elements). For this case, $\text{KCDis}(c)$ reaches its maximum $\text{KCDis}(c) = 1/|U|$. Therefore, $0 \leq \text{KCDis}(c) \leq 1 - 1/|U|$ is obtained. \square

4. Experimental Results

In this section, some experiments are used to show the effectiveness of our new method. The data used in our experiments come from the Pima Indians Diabetes Data Set, which includes a total of 768 cases, of which 392 are valid, and the rest of the data cases' characteristics values are missing. Note that the Pima Indians Diabetes Data Set is no longer available due to permission restrictions.

In actual computations, we use these 392 cases for experimentation. The condition characteristics information includes "plasma glucose concentration at 2 hours in an oral glucose tolerance test", "diastolic blood pressure (mm Hg)", "triceps skin fold thickness (mm)", "2-hour serum insulin ($\mu\text{U/ml}$)", "body mass index (weight in kg/(height in m)²)". The data set is given in Table 1, where "c1", "c2", "c3", "c4", and "c5" denote the condition characteristics, respectively. "d" stands for the decision characteristics "class variable (0 or 1)". Then the condition characteristics values are discretized to different levels as "A, B, C" or "A, B, C, D"; see Table 2.

TABLE 1: The Pima Indians Diabetes Data Set.

U	$c1$	$c2$	$c3$	$c4$	$c5$	d
U1	89	66	23	94	28.1	0
U2	137	40	35	168	43.1	1
U3	78	50	32	88	31	1
U4	197	70	45	543	30.5	1
U5	189	60	23	846	30.1	1
U6	166	72	19	175	25.8	1
U7	118	84	47	230	45.8	1
U8	103	30	38	83	43.3	0
U9	115	70	30	96	34.6	1
U10	126	88	41	235	39.3	0
\vdots						
U384	100	84	33	105	30	0
U385	81	74	41	57	46.3	0
U386	187	70	22	200	36.4	1
U387	121	78	39	74	39	0
U388	181	88	44	510	43.3	1
U389	128	88	39	110	36.5	1
U390	88	58	26	16	28.4	0
U391	101	76	48	180	32.9	0
U392	121	72	23	112	26.2	0

According to Algorithm 1, the following characteristics weights can be obtained:

$$\begin{aligned} W(c_1) &= 0.3625, \\ W(c_2) &= 0.0451, \\ W(c_3) &= 0.2388, \\ W(c_4) &= 0.2848, \\ W(c_5) &= 0.0688. \end{aligned} \quad (20)$$

Two experiments are conducted to show the advantages of our new method. The first experiment is to compare different rough set based methods with our method. The second one is to compare the AHP and PCA methods with our method. Both comparisons can show that our new proposed method is more effective than those methods.

In the first experiment, we also choose two rough set-based methods. One is based on the dependence in rough set theory to calculate the characteristics weight. The other is based on rough sets and conditional information entropy.

In knowledge bases $K = (U, R)$ and $R = C \cap D$, the dependence of the characteristic is defined as $\gamma_C(D) = |\text{POS}_B(D)|/|U|$. The characteristics importance $\text{Sig}(c) = \gamma_C(D) - \gamma_{C-\{c\}}(D)$. Then the characteristics weight is $W_1(c) = \text{Sig}(c) / \sum_{a \in C} \text{Sig}(a)$ [39]. By calculation, we have

$$\begin{aligned} W_1(c_1) &= 0.5, \\ W_1(c_2) &= 0, \\ W_1(c_3) &= 0.3, \end{aligned}$$

TABLE 2: The discretized Pima Indians Diabetes Data Set.

U	$c1$	$c2$	$c3$	$c4$	$c5$	d
$U1$	A	B	B	B	A	0
$U2$	C	A	C	C	C	1
$U3$	A	A	C	B	B	0
$U4$	D	B	D	D	B	1
$U5$	D	B	B	D	B	1
$U6$	D	B	A	C	A	0
$U7$	B	A	D	C	C	0
$U8$	B	B	C	B	C	1
$U9$	B	B	C	B	B	0
$U10$	C	C	D	C	B	0
\vdots						
$U384$	D	B	C	B	B	0
$U385$	A	B	D	A	C	1
$U386$	D	B	B	C	B	0
$U387$	C	B	C	B	B	0
$U388$	D	B	D	D	C	0
$U389$	C	B	C	B	B	0
$U390$	A	A	B	A	A	1
$U391$	B	B	D	C	B	1
$U392$	C	B	B	B	A	0

$$W_1(c_4) = 0.2,$$

$$W_1(c_5) = 0.$$

(21)

In knowledge bases $K = (U, R)$ and $R = C \cap D$, the dependence of the characteristic is defined as $I(D | C)$. $I(D | C) = \sum_{i=1}^m (|C_i^2|/|U^2|) \sum_{j=1}^k (|D_j \cap C_i|/|C_i|)(1 - |D_j \cap C_i|/|C_i|)$, and the characteristics importance $\text{Sig}(c) = I(D | C - \{c\}) - I(D | C)$. Then the characteristics weight is $W_2(c) = \text{Sig}(c) / \sum_{a \in C} \text{Sig}(a)$ [40]. By calculation, we have

$$W_2(c_1) = 0.5,$$

$$W_2(c_2) = 0,$$

$$W_2(c_3) = 0.2857, \quad (22)$$

$$W_2(c_4) = 0.2143,$$

$$W_2(c_5) = 0.$$

In Table 3, we list the weighting results of the three methods based on rough set. Figure 1 clearly shows their comparison. From Table 3 and Figure 1, it shows that when the methods based on the dependence of rough set and the method based on the rough set and conditional information entropy are used to calculate the characteristics weights, the weights of “ $c2$ ” and “ $c5$ ” are redundant. But when the proposed method is used to calculate the characteristics weights, the results were not redundant. There is a little relation between “diastolic blood pressure (mm Hg)”, “body mass index (weight in kg/(height in m)²)”, and diabetes, but

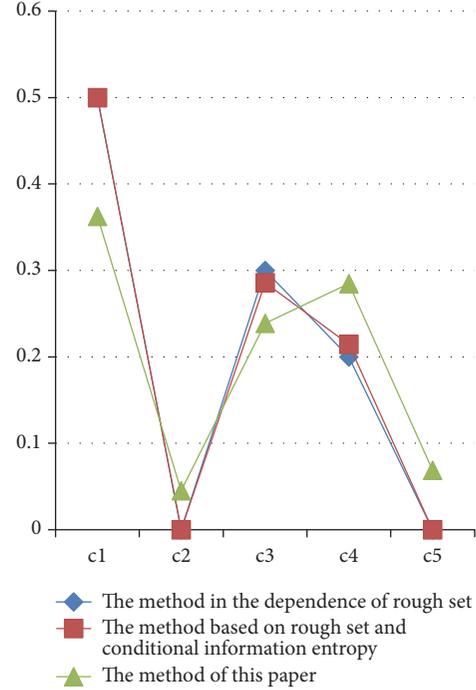


FIGURE 1: Comparison of three methods based on rough set.

they are related. So, from this point of view, the new method is more accurate than the other two rough set-based methods.

In the second experiment, the AHP method and the PCA method are used to calculate the characteristics weight. We also compare their results with ours.

For the AHP method, we construct the analytic hierarchy matrix according to the opinion of medical experts [41]. Then we obtain the weights:

$$W_3(c_1) = 0.0604,$$

$$W_3(c_2) = 0.1012,$$

$$W_3(c_3) = 0.3103, \quad (23)$$

$$W_3(c_4) = 0.1815,$$

$$W_3(c_5) = 0.3465.$$

For the PCA method, we select the representative variables through the transformation of multiple variables. Then the SPSS software is used to seek the explanation of the total variance and component of the matrix. We take principal components variance contribution rate as weight [41] and finally normalize them to get the weights:

$$W_4(c_1) = 0.432,$$

$$W_4(c_2) = 0.1114,$$

$$W_4(c_3) = 0.0978, \quad (24)$$

$$W_4(c_4) = 0.2568,$$

$$W_4(c_5) = 0.1019.$$

TABLE 3: Comparison of three methods based on rough set.

Method	c1	c2	c3	c4	c5
The method based on the dependence of rough set	0.5000	0	0.3000	0.2000	0
The method based on rough set and conditional information entropy	0.5000	0	0.2857	0.2143	0
The method of this paper	0.3625	0.0451	0.2388	0.2848	0.0688

TABLE 4: Comparison of different types of methods.

Method	c1	c2	c3	c4	c5
The AHP method	0.0604	0.1012	0.3103	0.1815	0.3465
The PCA method	0.432	0.1114	0.0978	0.2568	0.1019
The method of this paper	0.3625	0.0451	0.2388	0.2848	0.0688

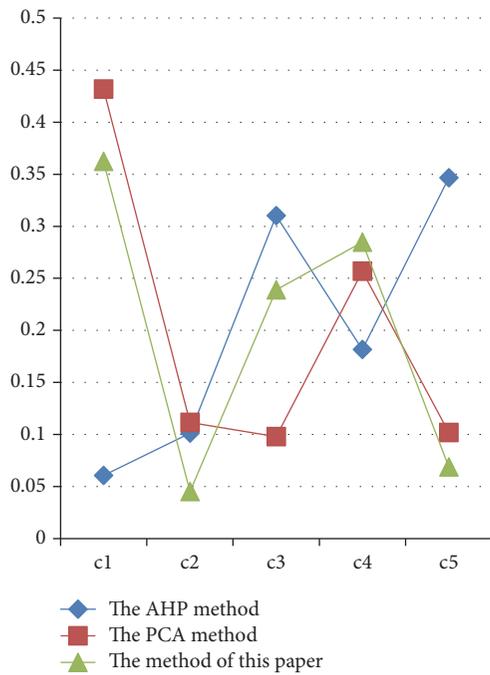


FIGURE 2: Comparison of three methods.

The weighting results are given in Table 4. Figure 2 shows the comparison between the proposed method and two well-known methods. From Table 4 and Figure 2, it is easy to check that the rank of the results calculated with our method is “c1” > “c4” > “c3” > “c5” > “c2”. It shows that there is a closed relation between “plasma glucose concentration at 2 hours in an oral glucose tolerance test” and diabetes, and there is a little relation between “diastolic blood pressure (mm Hg)” and diabetes. These results are synthetic optimization on the results calculated by AHP and PCA from Figure 2. By consulting the medical experts, the results calculated by our method are more accordant with the actual situation.

However, the Analytical Hierarchy Process (AHP) method is based on the subjective judgment of the experts and the Principal Component Analysis (PCA) method needs to extract representative principal components and increase an additional a priori information and evaluation criteria.

Therefore, these two methods cannot objectively reflect the weight distribution. The new method does not need the prior knowledge, but the obtained weights are in line with the actual situation.

From the above discussion, the weighting method based on rough set can avoid the arbitrariness of subjective judgment. In addition, the weighting method with granularity theory can effectively avoid taking nonredundant characteristics as redundant characteristics. We can conclude that our new method reasonably distributes the weight for each characteristic. The weights basically reflect the importance of each characteristic and can also objectively reflect the actual situation of the patient’s body. Thus, the proposed method is a powerful method in knowledge classification.

5. Conclusions

Knowledge characteristics can help us have a good understanding of the knowledge base. The determination of knowledge characteristics weight can help us effectively classify the knowledge base, so as to achieve the purpose of knowledge management and decision making. In this paper, based on rough set theory and knowledge granularity theory, the weights of knowledge characteristics are determined. Experimental results show that the proposed method can effectively avoid taking nonredundant characteristics as redundant characteristics and can effectively determine the weights of knowledge characteristics.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work is supported by the National Natural Science Foundation of China (no. 61472256, no. 61771265) and the Natural Science Foundation of Jiangsu Province (no. BK20151272).

References

- [1] L. Fan, Y.-J. Lei, and S.-L. Duan, "Interval valued intuitionistic fuzzy statistic adjudging and decision-making," *Systems Engineering Theory and Practice*, vol. 31, no. 9, pp. 1790–1797, 2011.
- [2] R. Körner and W. Näther, "Linear regression with random fuzzy variables: extended classical estimates, best linear estimates, least squares estimates," *Information Sciences*, vol. 109, no. 1-4, pp. 95–118, 1998.
- [3] H. Zhu, J. G. Ibrahim, N. Tang, and H. Zhang, "Diagnostic measures for empirical likelihood of general estimating equations," *Biometrika*, vol. 95, no. 2, pp. 489–507, 2008.
- [4] L. Abdullah and L. Najib, "Sustainable energy planning decision using the intuitionistic fuzzy analytic hierarchy process: choosing energy technology in Malaysia," *International Journal of Sustainable Energy*, vol. 35, no. 4, pp. 360–377, 2016.
- [5] L. Abdullah and L. Najib, "A new preference scale mcdm method based on interval-valued intuitionistic fuzzy sets and the analytic hierarchy process," *Soft Computing*, vol. 20, no. 2, pp. 511–523, 2016.
- [6] A. D. Sutadian, N. Muttill, A. G. Yilmaz, and B. J. C. Perera, "Using the analytic hierarchy process to identify parameter weights for developing a water quality index," *Ecological Indicators*, vol. 75, pp. 220–233, 2017.
- [7] V. Radha and M. Pushpalatha, "Comparison of PCA based and 2DPCA based face recognition systems," *International Journal of Engineering Science and Technology*, vol. 2, no. 12, pp. 7177–7182, 2010.
- [8] V. Perlibakas, "Distance measures for PCA-based face recognition," *Pattern Recognition Letters*, vol. 25, no. 6, pp. 711–724, 2004.
- [9] Z. Pawlak and A. Skowron, "Rudiments of rough sets," *Information Sciences*, vol. 177, no. 1, pp. 3–27, 2007.
- [10] H.-C. Zhu and N.-H. Chen, "The Improved method of ascertaining weigh based on rough sets and conditional information entropy," *Statistics & Decisions*, vol. 322, no. 8, pp. 154–156, 2011.
- [11] R. Cong, X. Wang, K. Li, and N. Yang, "New method for discretization of continuous attributes in rough set theory," *Journal of Systems Engineering and Electronics*, vol. 21, no. 2, pp. 250–253, 2010.
- [12] J. J. Alpignini, J. F. Peters, A. Skowron, and N. Zhong, "Rough sets and current trends in computing," in *Proceedings of the 3rd International Conference, RSCTC 2002*, pp. 14–16, Malvern, Pa, USA, October 2002.
- [13] R. W. Swiniarski and A. Skowron, "Rough set methods in feature selection and recognition," *Pattern Recognition Letters*, vol. 24, no. 6, pp. 833–849, 2003.
- [14] H. Yu, Z. Liu, and G. Wang, "An automatic method to determine the number of clusters using decision-theoretic rough set," *International Journal of Approximate Reasoning*, vol. 55, no. 1, part 2, pp. 101–115, 2014.
- [15] J. F. Peters and A. Skowron, "A rough set approach to knowledge discovery, Selected papers of the international workshop on rough sets in knowledge discovery and soft computing," in *Proceedings of the Selected Papers of The International Workshop on Rough Sets in Knowledge Discovery and Soft Computing, RSDK*, pp. 5–13, Warsaw, Poland, April 2003.
- [16] J. Li, Y. Ren, C. Mei, Y. Qian, and X. Yang, "A comparative study of multigranulation rough sets and concept lattices via rule acquisition," *Knowledge-Based Systems*, vol. 91, pp. 152–164, 2016.
- [17] L. Polkowski, S. Tsumoto, and T.-Y. Lin, Eds., *Rough Set Methods and Applications: New Developments in Knowledge Discovery in Information Systems*, Physica-Verlag, Heidelberg, Germany, 2000.
- [18] "Rough Sets, Fuzzy Sets, Data Mining, and Granular Computing," in *Proceedings of the 9th International Conference, RSFDGrC 2003*, G. Wang, Q. Liu, Y. Yao, and A. Skowron, Eds., pp. 26–29, Chongqing, China, May 2003.
- [19] M. Beynon, "Reducts within the variable precision rough sets model: A further investigation," *European Journal of Operational Research*, vol. 134, no. 3, pp. 592–605, 2001.
- [20] J. R. Xia, "The granular accuracy of approximation for the rough sets," *Applied Mathematics. A Journal of Chinese Universities. Series A*, vol. 27, no. 2, pp. 248–252, 2012.
- [21] D.-Q. Miao and S.-D. Fan, "The calculation of knowledge granulation and its application," *Systems Engineering Theory and Practice*, vol. 33, no. 1, pp. 48–56, 2002.
- [22] B. W. Fang and B. Q. Hu, "Probabilistic graded rough set and double relative quantitative decision-theoretic rough set," *International Journal of Approximate Reasoning*, vol. 74, pp. 1–12, 2016.
- [23] M. Restrepo, C. Cornelis, and J. Gómez, "Partial order relation for approximation operators in covering based rough sets," *Information Sciences*, vol. 284, pp. 44–59, 2014.
- [24] T. Feng and J.-S. Mi, "Variable precision multigranulation decision-theoretic fuzzy rough sets," *Knowledge-Based Systems*, vol. 91, pp. 93–101, 2016.
- [25] Y. Yao and B. Zhou, "Two Bayesian approaches to rough sets," *European Journal of Operational Research*, vol. 251, no. 3, pp. 904–917, 2016.
- [26] L.-C. Yang, P.-F. Zhang, and D.-W. Wang, "Fuzzy comprehensive evaluation method of metabolic syndrome based on PCA," *Journal of Biomedical Engineering*, vol. 30, no. 1, pp. 67–70, 2013.
- [27] Z. Xu and X. Zhang, "Hesitant fuzzy multi-attribute decision making based on TOPSIS with incomplete weight information," *Knowledge-Based Systems*, vol. 52, pp. 53–64, 2013.
- [28] B. Zhang and L. Zhang, "Discussion on future development of granular computing," *Journal of Chongqing University of Posts and Telecommunications*, vol. 23, no. 5, pp. 538–540, 2010.
- [29] Y. Li, T. Feng, S. Zhang, and Z. Li, "A generalized model of covering rough sets and its application in medical diagnosis," in *Proceedings of the International Conference on Machine Learning and Cybernetics*, pp. 145–150, 2010.
- [30] N. Mac Parthaláin and Q. Shen, "On rough sets, their recent extensions and applications," *The Knowledge Engineering Review*, vol. 25, no. 4, pp. 365–395, 2010.
- [31] X. Ma, G. Wang, H. Yu, and T. Li, "Decision region distribution preservation reduction in decision-theoretic rough set model," *Information Sciences*, vol. 278, pp. 614–640, 2014.
- [32] D. Liang and D. Liu, "A novel risk decision making based on decision-theoretic rough sets under hesitant fuzzy information," *IEEE Transactions on Fuzzy Systems*, vol. 23, no. 2, pp. 237–247, 2015.
- [33] Y. Sang, J. Liang, and Y. Qian, "Decision-theoretic rough sets under dynamic granulation," *Knowledge-Based Systems*, vol. 91, pp. 84–92, 2016.
- [34] Y. Qian, X. Liang, G. Lin, Q. Guo, and J. Liang, "Local multigranulation decision-theoretic rough sets," *International Journal of Approximate Reasoning*, vol. 82, pp. 119–137, 2017.
- [35] D. Liu, T. Li, and D. Liang, "Incorporating logistic regression to decision-theoretic rough sets for classifications," *International*

- Journal of Approximate Reasoning*, vol. 55, no. 1, part 2, pp. 197–210, 2014.
- [36] J. Liang, F. Wang, C. Dang, and Y. Qian, “An efficient rough feature selection algorithm with a multi-granulation view,” *International Journal of Approximate Reasoning*, vol. 53, no. 6, pp. 912–926, 2012.
- [37] H.-K. Wang, B.-X. Yao, and H.-Q. Hu, “The method of ascertaining weight based on rough sets theory,” *Computer Engineering and Applications*, vol. 36, no. 19, p. pp, 2003.
- [38] X. Y. Cao and J. G. Liang, “The Method of ascertaining attribute weight based on rough sets theory,” *Chinese Journal of Management*, vol. 10, no. 5, pp. 98–100, 2002 (Chinese).
- [39] X.-Z. Bao, J.-B. Zhang, and C. Liu, “A new method of ascertaining attribute weight based on rough sets conditional information entropy,” *Chinese Journal of Management Science*, vol. 6, no. 6, pp. 507–510, 2009 (Chinese).
- [40] H.-C. Zhu and H. Chen, “Rough set can be conditional information entropy weights improved method of determining,” *Statistics & Decisions*, vol. 32, no. 8, pp. 154–156, 2011.
- [41] Q. Chen, D. Shi, G. Feng, X. Zhao, and B. Luo, “On-line handwritten flowchart recognition based on grammar description language,” *Computer Science*, vol. 42, no. 11, pp. 113–117, 2015.



Hindawi

Submit your manuscripts at
www.hindawi.com

