Knowledge-Based Systems 260 (2023) 110093

Contents lists available at ScienceDirect

Knowledge-Based Systems

journal homepage: www.elsevier.com/locate/knosys

Incremental concept-cognitive learning approach for concept classification oriented to weighted fuzzy concepts

Chengling Zhang [a], Eric C.C. Tsang [a,*], Weihua Xu [b], Yidong Lin [c], Lanzhen Yang [a]

[a] School of Computer Science and Engineering, Macau University of Science and Technology, Taipa, Macao Special Administrative Region of China
[b] College of Artificial Intelligence, Southwest University, Chongqing, 400715, China
[c] School of Mathematics and Statistics, Minnan Normal University, Zhangzhou, 363000, China

ARTICLE INFO

Article history:
Received 14 June 2022
Received in revised form 22 October 2022
Accepted 4 November 2022
Available online 13 November 2022

Keywords:
Fuzzy entropy
Weighted fuzzy concept space
Concept classification
Progressive weighted fuzzy concept
Incremental learning mechanism

ABSTRACT

Fuzzy formal concept analysis (FFCA) is generally used to describe the processes of concept-cognitive learning (CCL). However, for fuzzy formal contexts, each attribute has the same weight (i.e, the same degree of importance) before constructing fuzzy concepts, which limits mining interesting knowledge and affects its application promotion. On the other hand, this model is hard to resist the influence of noise hidden in data, which results in poor classification learning. Moreover, the existing incremental CCL algorithms still face some challenges that the previously acquired knowledge is not fully utilized to improve the classification accuracies for dynamic data. To address these issues, we introduce different weights into fuzzy formal contexts and propose a novel incremental CCL mechanism in dynamic environment. Firstly, weight values of attributes from different decisions based on fuzzy entropy are established to measure the significant degree of attributes. Then, to comprehensively explicate the hierarchical relationships of fuzzy concepts, we construct the weighted fuzzy concept lattice and the weighted fuzzy concept space. Secondly, we design an algorithm to update the weighted fuzzy concepts for facilitating concept classification. To overcome the individual cognitive limitation, we put forward the progressive weighted fuzzy concept to remove repeated information. Furthermore, the classification prediction label and dynamic updating mechanism after adding objects are systematically discussed. Finally, we perform an experimental evaluation on ten data sets which explicate the feasibility and efficiency of our proposed approach.

1. Introduction

Cognitive computing is treated as an emerging computer system paradigm on the human brain by attempting to solve uncertain, imprecise and incomplete problems in biological system [1]. As is well known, it is the process of thinking, learning and perceiving by simulating the human brain using a computer. With years of development, cognitive learning, as a useful mathematical tool to realize cognitive computing, has been universally applied in cognitive psychology [2–5], machine learning [6–9], information science [10–12], and classification performance [13, 14].

Formal concept analysis (FCA) [15,16], proposed by Wille in 1980s, is a data analysis tool for discovering the uncertain knowledge. Actually, concept, as the central notion of human thinking, is the basic reflection of the concrete characteristics of objective reality. Commonly, formal context, a fundamental foundation of FCA, contains numerical values 1 and 0, whereas crisp relation between objects and attributes is endued merely. A classical concept can be categorized into two components: extent and intent among which they can be determined with each other. Since then, in some practical applications, with the increase of data types, various concepts from different relationships have been investigated, such as fuzzy concept [17], three-way concept [18], multi-scale concept [19], weighted concept [20–22]. Generally speaking, by weighted concepts, one can choose useful information according to their preferences and requirements. Information entropy, introduced by Shannon [23], as an uncertainty measure, improves the ability to resolve information into granules, which has been successfully extended to fuzzy entropy [24,25], non probabilistic entropy [26], and hybrid entropy [27]. Some of its expansions have been generally used to measure the fuzziness of rough set and FCA, which facilitate data classification. Concretely, Zhang et al. [21] developed a method to obtain weighted concepts based on information entropy in the absence of prior knowledge. In order to reduce the size of concepts, Singh et al. [28] explored the correlation between weights assigned to attributes before

* Corresponding author.
E-mail addresses: chengling_zhang@126.com (C. Zhang), cctsang@must.edu.mo (E.C.C. Tsang), chxuwh@gmail.com (W. Xu), yidong_lin@yeah.net (Y. Lin), lzyang@hbu.edu.cn (L. Yang).

https://doi.org/10.1016/j.knosys.2022.110093
0950-7051/

computing fuzzy concepts. Although these methods have indicated the importance of selected knowledge, that is, the weights of attributes are different during constructing concepts, they do not mine the weights of conditional attributes from the consistency of decision information granules. In other words, motivated by this problem, we find that fuzzy entropy provides a viable approach for the research of FFCA, and our work will discuss the shortage of the correlation description by utilizing conditional attributes and decision attributes.

With the high-speed development and rapid alteration of information technology, the available data has grown explosively. As an effective method to collect concepts from big data or information, CCL can reveal the cognitive process of a human brain. Incremental learning can directly deal with the acquired continuous data and extract valuable information, which can preserve previous knowledge rather than computing from scratch so as to increase computational efficiency. For instance, for improving efficiency greatly, Kumar [29] proposed the concept learning framework to analyze functionalities of bidirectional associative memory. Li et al. [30] investigated concept learning mechanism by granular computing from philosophy and cognitive psychology. Considering that retraining from a high-dimensional data is computationally expensive in the worst case, scholars introduced incremental learning systems [5,6,31,32]. To achieve uncertain classification task, Shi et al. [6,7] first discussed a novel CCL model that all samples are mapped to different subspaces, which is accompanied by dynamic concept learning under incremental cognitive process. Subsequently, Yan et al. [32] implemented an incremental CCL algorithm to concept extraction based on a three-way object partial order structure diagram when handling dynamic data. Inspired by Li's work [30], Zhao et al. [33] put forward cognitive concept learning in a sense of incomplete information. Additionally, to address fuzzy conceptual clustering analysis and state-of-the-art classification ability, Mi et al. [9] studied a fuzzy-based concept learning mechanism. Combining positive information and negative information, Yuan et al. [13] developed an incremental learning mechanism for implementing object classification in progressive fuzzy three-way concept. As a matter of fact, except for these fuzzy concepts, there are still other types of concepts, such as weighted concept. Therefore, how to identify the classification rule in weighted concepts is still a problem that worths exploring. And we will propose the solution in Section 3.

The fuzzy formal context considers only the same degree of importance with respect to attributes, while overlooking the differences. For the sake of solving this limitation, we in this article explore a novel approach to calculate the weight of fuzzy concept in a given fuzzy formal context. Furthermore, it should be noted that concept classification could be beneficial for comprehending and representing concept learning processes. Based on the above discussion, we construct a progressive weighted fuzzy concept space for removing the repeated concept information. In so doing, we achieve the concept classification by similarity measurement. Subsequently, to fully collect more valuable information about newly added objects, we propose a dynamic updating mechanism with respect to further concept learning. The flow chart of the proposed approach is shown in Fig. 1.

The remaining of this paper is organized as follows. We introduce some corresponding notions about FFCA and entropy, and motivation in Section 2. Section 3 studies the cognitive learning process of weighted fuzzy concept space. In addition, Section 4 discusses how to classify the class label of the new objects and how to update the weighted fuzzy concept space dynamically. Then the dynamic updating mechanism algorithm based on the progressive weighted fuzzy concept (for short DMPWFC) is analyzed in Section 4. Section 5 conducts numerical experiments and evaluates the effectiveness of our algorithm. Final summary and further research are drawn in Section 6.

Table 1
A fuzzy formal decision context.

G	a_1	a_2	a_3	d
x_1	0.3	0.5	0.7	1
x_2	0.7	0.5	0.6	2
x_3	0.2	0.8	0.9	1
x_4	0.9	0.6	0.8	2

2. Preliminaries

In this section, we briefly review some basic notions about FFCA and entropy theory. More specific information is available in [17,23–26].

2.1. FFCA

In FFCA, formal fuzzy context, introduced by S. Yahia [17], is a mathematical tool of data analysis and knowledge representation. Practically, it includes nonempty finite sets of objects and attributes, and a fuzzy binary relation between the two sets that are needed in the later discussion.

Assume U is a universe, and a fuzzy set, or more precisely a fuzzy set $\tilde{A}$ of U which is defined as a membership function $\tilde{A}(\cdot) : U \to [0, 1]$. For an arbitrary $x \in U$, the value $\tilde{A}(x)$ is said to be the fuzzy membership degree of x to $\tilde{A}$. Then we denote the set of all fuzzy subsets of U as $\mathcal{F}(U)$.

Let A and B be two fuzzy sets on U. If $\tilde{A}(x) \le \tilde{B}(x)$, $x \in U$, then A is a subset of B, i.e., $A \subseteq B$. Especially, we denote by $\mathcal{P}(U)$ the set of crisp sets on U.

A triplet $(G, M, \tilde{R})$ is called a fuzzy formal context, where $G = \{x_1, x_2, \ldots, x_n\}$ and $M = \{a_1, a_2, \ldots, a_m\}$ are all the sets of objects and attributes, respectively. $\tilde{R}$ is a fuzzy relation between G and M (i.e. $\tilde{R} : G \times M \to [0, 1]$), and each $\tilde{R}(x, a)$ reflects the membership degree of object x to attribute a.

Definition 1 *([17])*. Let $(G, M, \tilde{R})$ be a fuzzy formal context. For $X \subseteq G$ and $\tilde{B} \in \mathcal{F}(M)$, two operators $F : \mathcal{P}(G) \to \mathcal{F}(M)$ and $Q : \mathcal{F}(M) \to \mathcal{P}(G)$ are given as follows:

$$F(X)(a) = \bigwedge_{x \in X} \tilde{R}(x, a), a \in M, \tag{1}$$

$$Q(\tilde{B}) = \left\{x \in G : \forall a \in M, \tilde{B}(a) \le \tilde{R}(x, a)\right\}, \tag{2}$$

where a pair $(X, \tilde{B})$ is fuzzy concept satisfying $F(X) = \tilde{B}$ and $Q(\tilde{B}) = X$. In general, X and $\tilde{B}$ are called extent and intent, respectively.

Let $(G, M, \tilde{R})$ and (G, D, J) be two fuzzy formal contexts, $\tilde{R} : G \times M \to [0, 1]$ and $J : G \times D \to \{0, 1\}$. Then a fuzzy formal decision context is called $(G, M, \tilde{R}, D, J)$ where $M \cap D = \emptyset$, in which M is represented as the conditional attribute set and D is the decision attribute set.

Example 1. A fuzzy formal decision context $(G, M, \tilde{R}, D, J)$ is depicted in Table 1, where G and M are a set of students and a set of courses, respectively. a_1, a_2, a_3 represent the English, Mathematics, and Physical, respectively. $\tilde{R}(x_1, a_2)$ means the fuzzy membership value of the score of student x_1 in the percentile system of Mathematics(a_2) examination. Then $\{d\}$ is the decision attribute set that can divide G into two decision classes $D_1 = \{x_1, x_3\}$ and $D_2 = \{x_2, x_4\}$. The generated fuzzy concepts are revealed in Table 2. Remarkably, for convenience, we clearly abbreviate $\left(\{x_1, x_2, x_4\}, (a_1^{0.3}, a_2^{0.2}, a_3^{0.3})\right)$ as $\left(124, (0.3, 0.2, 0.3)\right)$.

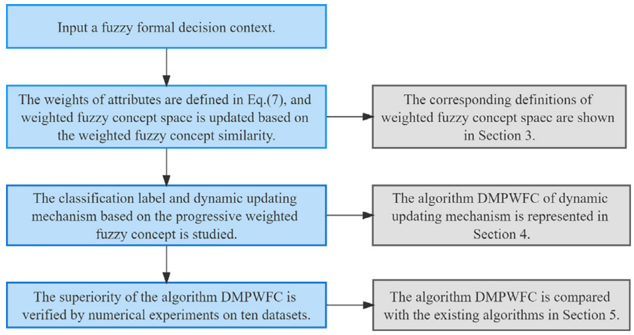

Fig. 1. The overall procedure of the proposed method.

Table 2
Fuzzy concepts of Table 1.

Fuzzy concepts	Nomenclature
$(U, (0.2, 0.5, 0.6))$	C_1
$(124, (0.3, 0.5, 0.6))$	C_2
$(134, (0.2, 0.5, 0.7))$	C_3
$(14, (0.3, 0.5, 0.7))$	C_4
$(24, (0.7, 0.5, 0.6))$	C_5
$(34, (0.2, 0.6, 0.8))$	C_6
$(4, (0.9, 0.6, 0.8))$	C_7
$(3, (0.2, 0.8, 0.9))$	C_8
$(\emptyset, (1, 1, 1))$	C_9

2.2. Entropy

Entropy is a measure of the uncertainty of a random experiment S. In other words, it is also a measure of the information obtained when observing results. Next, we first propose Shannon's entropy [23] based on the probability information. Eventually, a fuzzy entropy measure is also introduced which is an important extension of Shannon's theory.

Let $S = \{s_1, s_2, \dots, s_t\}$ be a finite discrete random variable set. If the probability distribution of s_i is $p(s_i)$, then the information generated by s_i is denoted as:

$$I(s_i) = -log_2 p(s_i).$$

Then the Shannon's entropy $H(S)$ of a discrete random variable S is given as:

$$H(S) = -\sum_{i=1}^{t} p(s_i) log_2 p(s_i). \quad (3)$$

As shown above, entropy is considered to be a probability distribution function of S.

Furthermore, in 1972, inspired by Shannon's entropy theory, De Luca [26] proposed a new parameterized fuzzy entropy to measure the fuzziness uncertainty of a random variable. Given a set-to-point mapping $H : \mathcal{F}(2^S) \to \Re^+$, the measure of fuzzy entropy is expressed as:

$$H(A) = -k \sum_{i=1}^{t} \left(\tilde{A}(s_i) log \left(\tilde{A}(s_i) \right) + \left(1 - \tilde{A}(s_i) \right) log \left(1 - \tilde{A}(s_i) \right) \right). \quad (4)$$

where $k > 0$ is an undetermined constant and $\tilde{A}(s_i)$ is the fuzzy value with respect to s_i. Meanwhile, this fuzzy entropy estimates the global deviations from ordinary sets. That is, any non-fuzzy set A indicates $H(A) = 0$, and a fuzzy set A with $\tilde{A}(s_i) = 0.5$ for each $i = 1, 2, \dots, t$ represents the role of maximum element of ordering defined by H.

Actually, different from the Shannon's entropy, the fuzzy entropy does not depend on probability of variable s_i, but depends on its membership degree.

2.3. Motivation

In fuzzy formal contexts, each attribute has the same weight in premise of constructing the fuzzy concept lattice. That is to say, the intents of fuzzy concepts are considered to have the same degree of importance, and all the nodes would be generated. However, in practical situations, human are not interested in all the original fuzzy concepts. On the contrary, they prefer to select the interesting learning process according to their preferences for some attributes. For instance, the total score of the final evaluation of undergraduate students in universities is 100 points, which is generally divided into three parts in the ratio of 3:3:4 to calculate their final scores. The Covid 2019 has caused a pandemic in more than 200 nations, thereby affecting billions of people. It is important to consider that the confirmed symptoms of patients frequently have a fever, pulmonary infection, cough, headache, muscle aches and so forth. Nevertheless, at first, doctors attach more importance to fever, pulmonary infection, and cough symptoms to some extent. Analogously, the evaluation indicators, used in the scholarship evaluation of postgraduates, include professional scores, scientific research achievements, personal moralities, and academic activities, etc. What we do know, however, is that scientific research achievements can largely determine whether a postgraduate could receive a scholarship or not. The reason is that the proportion of this indicator is higher than that of the others. To sum up, it should be noted that the above weight values are discussed according to the subjective experience of experts combined with different events, while ignoring the objective evidences. Consequently, it is necessary to assign different weights to different attributes to capture the degree of importance in a fuzzy formal decision context, which is consistent with human cognition.

In reality, from the above discussion, it should be pointed out that cognitive learning considers conceptual uncertainty according to the different degree of importance, which can better reflect the actual cognitive situation. For instance, by considering Example 1, the following Fig. 2 describes the fuzzy concept lattice where each dark pink node is a fuzzy concept. And what is particularly worth mentioning is that three light green nodes mean two students x_3, x_4 whose the scores of Mathematics(a_2) and Physical(a_3) are both more than 60 percent, and only students x_3, x_4 have all achieved more than 60 percent with respect to the scores of Mathematics(a_2) and Physical(a_3). Now, what we are concerned about is how to collect these three fuzzy concepts (three light green nodes). Additionally, assume that each fuzzy concept has its own weight, then it is convenient to select the interesting fuzzy concept information according to the weighted values.

In the process of cognitive learning, the intents formed by certain weighted attributes imply the uncertain knowledge reasoning. The results of cognitive learning of weighted attributes are

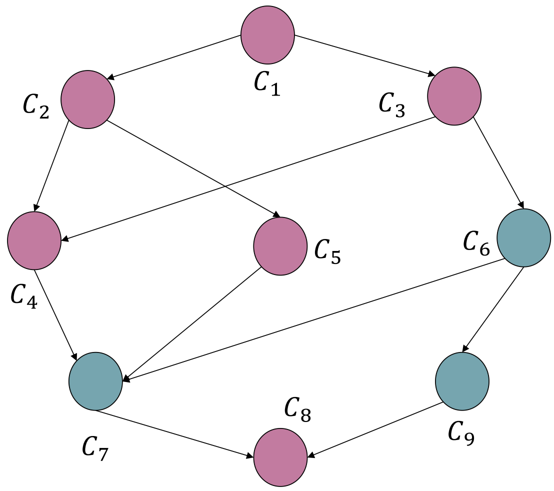

Fig. 2. Fuzzy concept lattice.

different from those of traditional cognitive learning [13,30,31]. To resolve the aforementioned issues, we introduce an approach to calculate the weights of attributes which quantifies the influence of different importance. It is evident that a weighting scheme can be distinguished from the global viewpoint of the individual classes not from the entire data and thereby implements sample classification. Thus, weighted fuzzy concept lattice is proposed to explicate the above-mentioned challenges.

3. The cognitive learning process of weighted fuzzy concept space

In this section, we propose the cognitive learning approach to weighted fuzzy concept for meeting actual cognition.

3.1. The weighted fuzzy concept

Given a fuzzy formal decision context $(G, M, \tilde{R}, D, J)$, where $G = \{x_1, x_2, \dots, x_n\}$ and $M = \{a_1, a_2, \dots, a_m\}$ are the sets of objects and attributes, respectively. $D = \{d_1, d_2, \dots, d_r\}$ is the set of decision attributes to mark the object classification. And $G/D = \{D_1, D_2, \dots, D_t\}$ is seen as a decision partition on G to D.

Definition 2. Let $(G, M, \tilde{R}, D, J)$ be a fuzzy formal decision context. For $x \in G$ and $a \in M$, $\tilde{R}(x, a)$ denotes the fuzzy membership value of object x to attribute a. Then the positive match degree of a provided by D_i is defined as follows:

$$E_i = \frac{\sum_{x \in D_i} \tilde{R}(x, a)}{\sum_{x \in G} \tilde{R}(x, a)}, \tag{5}$$

and the negative match degree of a of D_i is constructed as:

$$N_i = \frac{\sum_{x \in D_i} \left(1 - \tilde{R}(x, a)\right)}{\sum_{x \in G} \left(1 - \tilde{R}(x, a)\right)}, \tag{6}$$

Then the fuzzy entropy of a induced by D is denoted by:

$$H(a) = -\frac{1}{log_2 t} \sum_{i=1}^{t} \left(E_i log_2 E_i + N_i log_2 N_i \right). \tag{7}$$

where t is the number of decision partition G/D, and $H(a)$ indicates the uncertainty of a assigned by D and it is a non-probabilistic entropy. Conversely, E_i and N_i are measured via the fuzzy membership degree of each element in a certain decision class. Subsequently, the fuzzy entropy of each attribute is attained by summing fuzzy entropy of individual intervals in every feature dimension. In particular, $H(a) = 0$ if $E_i = 0$ or $E_i = 1$. Actually, the larger the output value of the fuzzy entropy is, the smaller the contribution value of the element to the fuzzy formal decision context is. Evidently, the weight of a is given by:

$$\omega(a) = \frac{1}{|M| - 1} \left(1 - \frac{H(a)}{\sum_{a \in M} H(a)} \right). \tag{8}$$

$\omega(a)$ represents the significance degree of attribute a to D. In fact, the larger the $\omega(a)$ is, the stronger the significance ability of a is. Obviously, it is apparent that $0 \le \omega(a) \le \frac{1}{|M|-1}$.

For an arbitrary $a \in M$, the weight vector of attribute is denoted as $W = (\omega(a_1), \omega(a_2), \dots, \omega(a_m))$, where $\omega(a_i) \in W$ is described as a weight vector of each attribute in M via the fuzzy information entropy and $\sum_{a_i \in M} \omega(a_i) = 1$.

Definition 3. Let $(G, M, \tilde{R}, D, J)$ be a fuzzy formal decision context. W is a weight vector of attributes in M. For $X \subseteq G$ and $\tilde{B} \in \mathcal{F}(M)$, if $F(X) = \tilde{B}$ and $Q(\tilde{B}) = X$, then the pair $h_\omega = (X, \tilde{B}, \omega)$ is called a weighted fuzzy concept. Practically, X and $\tilde{B}$ are referred to as the extent and intent of h_ω, respectively. ω is said to be the weight value of multi-attribute intent $\tilde{B}$ and is given by:

$$\omega = \frac{1}{|M|} \sum_{a_i \in M} \tilde{B}(a_i) \omega(a_i). \tag{9}$$

where $\tilde{B} = \left(\tilde{B}(a_1), \tilde{B}(a_2), \dots, \tilde{B}(a_m)\right)$ for all $a_i \in M$ and m is the number of attributes. In fact, it also measures the degree of importance with respect to extent X from the average information weight. Therefore, it provides the average weight of each fuzzy concept to determine their degree of importance.

For two weighted fuzzy concepts $(X_1, \tilde{B}_1, \omega_1)$ and $(X_2, \tilde{B}_2, \omega_2)$, the hierarchy order relation is described by the subconcept–superconcept relation:

$$(X_1, \tilde{B}_1, \omega_1) \le (X_2, \tilde{B}_2, \omega_2) \Leftrightarrow X_1 \subseteq X_2 \Leftrightarrow \tilde{B}_2 \le \tilde{B}_1 \left(or\ \omega_2 \le \omega_1\right). \tag{10}$$

Moreover, the set of all weighted fuzzy concepts forms a weighted fuzzy concept lattice (or abbreviated as WFCL) with respect to "$\le$", which is represented as $L_\omega(G, M, \tilde{R}, D, J)$. For convenience, we rewrite $F(\{x\})$ as $F(x)$ for short when no confusion exists. Moreover, the pair $\left(Q\left(F(x)\right), F(x), \omega\right)$ is referred to as a weighted fuzzy granular concept.

Proposition 1. *Let $(G, M, \tilde{R}, D, J)$ be a fuzzy formal decision context. W is a weight vector of attributes in M. For an arbitrary $X_1 \subseteq G$, then $\left(Q\left(F(X_1)\right), F(X_1), \omega_1\right)$ is a weighted fuzzy concept.*

Proof. It is straightforward from Definition 3. □

Therefore, the effectiveness of such a weighting approach can be illustrated with the assistance of the following example.

Example 2 (*Continued with Example 1*)**.** To understand the calculation process of weighted fuzzy concepts, for each attribute in Table 1, the fuzzy entropies and weights of attributes are calculated in Table 3. Furthermore, we can see that fuzzy concept-intents are used to compute weight values using Eq. (9) with respect to the attributes available described in Table 4.

From Table 4, we can see that the intent of each node of WFCL is represented by some assigned weights which quantify the corresponding importance of the node. Clearly, in reality, the human can focus on some certain nodes and deduce information

Table 3
Weight of each attribute.

U	$H(a)$	$\omega(a)$
a_1	1.5343	0.3602
a_2	1.9837	0.3192
a_3	1.9677	0.3207

Table 4
Weighted fuzzy concepts in Table 1.

Nomenclature	Weighted fuzzy concepts
WC_1	$(U, (0.2, 0.5, 0.6), 0.1413)$
WC_2	$(124, (0.3, 0.5, 0.6), 0.1533)$
WC_3	$(134, (0.2, 0.5, 0.7), 0.1520)$
WC_4	$(14, (0.3, 0.5, 0.7), 0.1640)$
WC_5	$(24, (0.7, 0.5, 0.6), 0.2014)$
WC_6	$(34, (0.2, 0.6, 0.8), 0.1734)$
WC_7	$(4, (0.9, 0.6, 0.8), 0.2574)$
WC_8	$(3, (0.2, 0.8, 0.9), 0.2053)$
WC_9	$(\emptyset, (1, 1, 1), 0.3333)$

according to their preferences and requirements. For instance, assume that one is more interested in those information whose weight values are greater than 0.1640. In such case, the nodes WC_1, WC_2, and WC_3 are not generated. In other words, we only consider the nodes WC_4, WC_5, WC_6, WC_7, WC_8 and WC_9. Therefore, WFCL can provide flexible ways for human-like problem solving and cognitive learning. We can quickly collect the interested knowledge and save storage space significantly.

Based on the above discussion, we now obtain the weighted fuzzy concepts that we need. Obviously, in Example 1, we know that learning fuzzy concept by exhaustive information granules is completed exponentially with the dimension of objects. Moreover, considering that information granules are the basic notion in the theory of granular computing, which plays a fundamental role in human cognition. It is natural for us to integrate information granules into cognitive learning for decreasing the time consumption. Therefore, granular computing (Grc for short) should be introduced into the process of learning fuzzy concept to decrease the amount of calculation. Next, how to construct concept space from the given weighted fuzzy granular concept is the key problem of cognitive learning.

3.2. Construction of weighted fuzzy concept space

In this subsection, we analyze a new classification task framework of weighted fuzzy concepts based on Grc, which comprises two aspects: initial concept space and updating concept space.

Let $(G, M, \tilde{R}, D, J)$ be a fuzzy formal decision context, where $G/D = \{D_1, D_2, \ldots, D_t\}$. W is a weight vector of attributes in M. For any D_i, the weighted fuzzy concept space $\mathcal{C}_i$ under D_i is denoted as follows:

$$\mathcal{C}_i = \left\{ \left(Q\left(F(x)\right), F(x), \omega \right) \middle| x \in D_i \right\}. \quad (11)$$

In addition, we denote by $\mathscr{C} = \{\mathcal{C}_1, \mathcal{C}_2, \ldots, \mathcal{C}_t\}$ the weighted fuzzy concept space (for short WFCS) in which $\mathcal{C}_i$ is called a weighted fuzzy subspace of $\mathscr{C}$. It should be noted that each object could be learned comprehensively for achieving classification performance. Based on the aforementioned discussion, we propose the procedure of constructing WFCS in Algorithm 1.

Example 3. Table 5 depicts a fuzzy formal decision context $(G, M, \tilde{R}, D, J)$, and nine objects are divided into two classes based on decision attribute d, i.e. $D_1 = \{x_1, x_2, x_3, x_4, x_5\}$ and $D_2 = \{x_6, x_7, x_8, x_9\}$.

Algorithm 1: Constructing weighted fuzzy concept space (CWFCS)

Input: A fuzzy formal decision context $(G, M, \tilde{R}, D, J)$.
Output: Weighted fuzzy concept space $\mathscr{C}$.
Compute the decision partition $G/D = \{D_1, D_2, \cdots, D_t\}$ and the weight W;
for each $D_i \in U/D$ **do**
  Set $\mathcal{C}_i \leftarrow \emptyset$;
  for each $x \in D_i$ **do**
    Compute the weight value ω of multi-attribute;
    Get the weighted fuzzy concept $\left(Q\left(F(x)\right), F(x), \omega\right)$;
    $\mathcal{C}_i \leftarrow \left(Q\left(F(x)\right), F(x), \omega\right)$;
  end for
  $\mathscr{C} \leftarrow \mathcal{C}_i$;
end for
Return $\mathscr{C} = \{\mathcal{C}_1, \mathcal{C}_2, \cdots, \mathcal{C}_t\}$.

From Eq. (8), the weight of attribute is $W = (0.5587, 0.4413)$. Therefore, the weighted fuzzy concept subspace $\mathcal{C}_1$ generated by class D_1 is represented as follows:

$$\begin{aligned}\mathcal{C}_1 = \Big\{ & \big(\{x_1, x_5\}, (0.08, 0.72), 0.1812\big), \\ & \big(\{x_2, x_4, x_5\}, (0.11, 0.56), 0.1543\big), \\ & \big(\{x_1, x_2, x_3, x_4, x_5, x_6\}, (0.04, 0.47), 0.1149\big), \\ & \big(\{x_4, x_5\}, (0.32, 0.65), 0.2328\big), \\ & \big(\{x_5\}, (0.55, 0.86), 0.3434\big) \Big\}.\end{aligned}$$

Analogously, the weighted fuzzy concept subspace $\mathcal{C}_2$ induced by class D_2 is given by:

$$\begin{aligned}\mathcal{C}_2 = \Big\{ & \big(\{x_6\}, (0.90, 0.47), 0.3551\big), \\ & \big(\{x_6, x_7, x_8, x_9\}, (0.68, 0.14), 0.2208\big), \\ & \big(\{x_8\}, (0.91, 0.36), 0.3336\big), \big(\{x_6, x_9\}, (0.75, 0.4), 0.2978\big) \Big\}.\end{aligned}$$

Evidently, the correlation analysis between two weighted fuzzy concepts in $\mathcal{C}_i$ may be influenced by the noise data besides decision class D_i. In such case, it is necessary to delete the weighted fuzzy concept caused by the noise. Inspired by this issue, we propose the concept similarity in the following discussion.

Definition 4. Let $(G, M, \tilde{R}, D, J)$ be a fuzzy formal decision context. W is a weight vector of attributes. For a weighted fuzzy subspace $\mathcal{C}_i$ and threshold δ, if $(X_1, \tilde{B}_1, \omega_1)$ is a weighted fuzzy concept in $\mathcal{C}_i$, and $(X_2, \tilde{B}_2, \omega_2)$ is its subconcept, then the weighted fuzzy concept similarity is represented as follows:

$$\theta_{1,2}^{\mathcal{C}_i} = \frac{|X_1 \cap X_2|}{|X_1 \cap X_2| + 2\Big(\alpha|X_1 - X_2| + (1-\alpha)|X_2 - X_1|\Big)}. \quad (12)$$

where $\alpha = |\omega_1 - \omega_2|$, and $|X_1|$ is the cardinality with respect to $(X_1, \tilde{B}_1, \omega_1)$. In fact, since $(X_2, \tilde{B}_2, \omega_2)$ is a subconcept of $(X_1, \tilde{B}_1, \omega_1)$, then $|X_2 - X_1| = 0$. Therefore, the above Eq. (12) can be simplified as follows:

$$\theta_{1,2}^{\mathcal{C}_i} = \frac{|X_1 \cap X_2|}{|X_1 \cap X_2| + 2\alpha|X_1 - X_2|}. \quad (13)$$

where $\theta_{1,2}^{\mathcal{C}_i}$ reflects the degree of similarity of $(X_1, \tilde{B}_1, \omega_1)$ with respect to $(X_2, \tilde{B}_2, \omega_2)$. The larger the value of $\theta_{1,2}^{\mathcal{C}_i}$ is, the stronger the ability of similarity is. Hence, when $\theta_{1,2}^{\mathcal{C}_i} > \delta$, the degree of importance of two weighted fuzzy concepts will be increased; when $\theta_{1,2}^{\mathcal{C}_i} \leq \delta$, the degree of importance of two weighted fuzzy concepts will be decreased. Practically, the weak relationship between the above weighted fuzzy concepts may be caused by the noise hidden in $\mathcal{C}_i$. Hence, $(X_2, \tilde{B}_2, \omega_2)$ should be removed during the learning process of weighted fuzzy concept space.

From the above statement, it can be seen that threshold δ controls the size of weighted fuzzy concept space. And the larger the threshold δ is, the smaller the size of weighted fuzzy concept space is. Hence, the specific process of updating weighted fuzzy concept space (UWFCS) is shown in Algorithm 2.

Algorithm 2: Updating weighted fuzzy concept space (UWFCS)

Input: An initial weighted fuzzy concept space $\mathscr{C}$ and threshold δ.
Output: An updating weighted fuzzy concept space $\mathscr{C}^\delta$.

```
for C_i ∈ 𝒞 do
    for each (Q(F(x_l)), F(x_l), ω_l) ∈ C_i do
        Set C^δ_{i,l} ← ∅ ;
        if (Q(F(x_j)), F(x_j), ω_j) is a fuzzy subconcept of
        (Q(F(x_l)), F(x_l), ω_l) in C_i then
            Compute θ^{C_i}_{l,j} from Definition 4;
            if θ^{C_i}_{l,j} > δ then
                C^δ_{i,l} ← (Q(F(x_j)), F(x_j), ω_j);
            end if
        end if
        C^δ_i ← ⋃_{Q(F(x_l)) ∈ C_i} C^δ_{i,l};
    end for
    𝒞^δ ← C^δ_i;
end for
Return 𝒞^δ = {C^δ_1, C^δ_2, ··· , C^δ_t}.
```

Example 4 (*Continued with Example 3*)**.** Given the threshold $\delta = 0.59$, from Definition 4, we can further compute the weighted fuzzy concept similarity based on $\mathcal{C}_1$ as follows:

$$\theta_{1,5}^{\mathcal{C}_1} = 0.76 > 0.59, \theta_{2,4}^{\mathcal{C}_1} = 0.93 > 0.59,$$
$$\theta_{2,5}^{\mathcal{C}_1} = 0.57 < 0.59, \theta_{3,1}^{\mathcal{C}_1} = 0.79 > 0.59,$$
$$\theta_{3,2}^{\mathcal{C}_1} = 0.93 > 0.59, \theta_{3,4}^{\mathcal{C}_1} = 0.68 > 0.59,$$
$$\theta_{3,5}^{\mathcal{C}_1} = 0.30 < 0.59, \theta_{4,5}^{\mathcal{C}_1} = 0.82 > 0.59.$$

Then $\mathcal{C}_1$ is updated to $\mathcal{C}_1^\delta$ as follows:

$$\mathcal{C}_1^\delta = \Big\{ \big(\{x_1, x_5\}, (0.08, 0.72), 0.1812\big), \big(\{x_2, x_4, x_5\}, (0.11, 0.56), 0.1543\big), \big(\{x_4, x_5\}, (0.32, 0.65), 0.2328\big), \big(\{x_5\}, (0.55, 0.86), 0.3434\big) \Big\}.$$

Analogously, the weighted fuzzy concept similarity based on $\mathcal{C}_2$ is calculated as:

$$\theta_{2,1}^{\mathcal{C}_2} = 0.55 < 0.59, \theta_{2,3}^{\mathcal{C}_2} = 0.60 > 0.59,$$
$$\theta_{2,4}^{\mathcal{C}_2} = 0.87 > 0.59, \theta_{4,1}^{\mathcal{C}_2} = 0.90 > 0.59.$$

Then $\mathcal{C}_2$ is updated to $\mathcal{C}_2^\delta$ as follows:

$$\mathcal{C}_2^\delta = \Big\{ \big(\{x_6\}, (0.90, 0.47), 0.3551\big), \big(\{x_8\}, (0.91, 0.36), 0.3336\big), \big(\{x_6, x_9\}, (0.75, 0.4), 0.2978\big) \Big\}.$$

Table 5
A fuzzy formal decision context.

G	a_1	a_2	d
x_1	0.08	0.72	1
x_2	0.11	0.56	1
x_3	0.04	0.47	1
x_4	0.32	0.65	1
x_5	0.55	0.86	1
x_6	0.90	0.47	2
x_7	0.68	0.14	2
x_8	0.91	0.36	2
x_9	0.75	0.40	2

3.3. Construction of the progressive weighted fuzzy concept space

In the previous section, we have researched how to construct the weighted fuzzy concept space and update weighted fuzzy concept space from operators F and Q. In fact, there is repeated information between weighted fuzzy concepts, and they interact with each other. To overcome individual cognitive limitations and incomplete cognitive environments [31], we propose a new approach to construct the progressive weighted fuzzy concept based on the weighted fuzzy concept space.

Definition 5. Let $(G, M, \tilde{R}, D, J)$ be a fuzzy formal decision context. W is a weight vector of attributes. For a weighted fuzzy subspace $\mathcal{C}_i^\delta$, if there exist weighted fuzzy concepts $(X_1, \tilde{B}_1, \omega_1)$, $(X_2, \tilde{B}_2, \omega_2), \ldots, (X_n, \tilde{B}_n, \omega_n)$ satisfying $X_1 \subseteq X_2 \subseteq \cdots \subseteq X_n$, then $(X_n, \tilde{B}_n, \omega_n)$ is referred to as the supremum concept. And the progressive weighted fuzzy concept is given by:

$$X_{i,j}^\delta = X_1 \cup X_2 \cup \cdots \cup X_n,$$
$$\tilde{B}_{i,j}^\delta = \frac{1}{2^{n-1}}(\tilde{B}_1 + \tilde{B}_2 + 2\tilde{B}_3 + 4\tilde{B}_4 + \cdots + 2^{n-2}\tilde{B}_n).$$

Then $(X_{i,j}^\delta, \tilde{B}_{i,j}^\delta, \omega_{i,j}^\delta)$ is a progressive weighted fuzzy concept, where $\omega_{i,j}^\delta = \frac{1}{|M|}\sum_{a_i \in M} \tilde{B}_{i,j}^\delta(a_i)\omega(a_i)$.

In a general sense, the progressive weighted fuzzy concept space is expressed as $\mathscr{K}^\delta = \{\mathcal{K}_1^\delta, \mathcal{K}_2^\delta, \ldots, \mathcal{K}_t^\delta\}$ where $\mathcal{K}_i^\delta = \{\mathcal{K}_{i,j}^\delta | j = 1, 2, \ldots, s\} = \big\{(X_{i,j}^\delta, \tilde{B}_{i,j}^\delta, \omega_{i,j}^\delta) | j = 1, 2, \ldots, s\big\}$. The s is used to denote the number of fuzzy concept in subspace $\mathcal{C}_i$. The intent $\tilde{B}_{i,j}^\delta$ concretely reflects the size of progressive weighted fuzzy concept where intents of subconcept have assigned different weights in accordance with their corresponding extents. That is to say, the larger the extent X_i is, the larger the weight of its relevant intent $\tilde{B}_i$ is. Accordingly, it is known that all weights of intents is 1. At last, algorithm 3 shows how to compute the progressive weighted fuzzy concept space $\mathscr{K}^\delta$.

Example 5 (*Continued with Example 4*)**.** From Definition 5, the progressive weighted fuzzy concepts are shown as follows:

$$\mathcal{K}_{1,1}^\delta = \big(\{x_1, x_5\}, (0.3150, 0.7900), 0.2623\big);$$
$$\mathcal{K}_{1,2}^\delta = \big(\{x_2, x_4, x_5\}, (0.2725, 0.7575), 0.2212\big);$$
$$\mathcal{K}_{2,1}^\delta = \big(\{x_8\}, (0.9100, 0.3600), 0.3336\big);$$
$$\mathcal{K}_{2,2}^\delta = \big(\{x_6, x_9\}, (0.8250, 0.4350), 0.3264\big).$$

There are two progressive weighted fuzzy concepts under each decision class. It can be seen that these concepts continue to retain the original information, and remove the redundant weighted fuzzy concepts for greatly facilitating the CCL efficiency.

4. Incremental cognitive learning based on the progressive weighted fuzzy concept

Given a fuzzy formal decision context $(G, M, \tilde{R}, D, J)$, the progressive weighted fuzzy concept space has a good performance

Algorithm 3: Constructing the progressive weighted fuzzy concept space

Input: An updating weighted fuzzy concept space $\mathscr{C}^\delta = \{\mathcal{C}_1^\delta, \mathcal{C}_2^\delta, \cdots, \mathcal{C}_t^\delta\}$ and threshold δ.
Output: The progressive weighted fuzzy concept space $\mathscr{K}^\delta = \{\mathcal{K}_1^\delta, \mathcal{K}_2^\delta, \cdots, \mathcal{K}_t^\delta\}$.

```
for C_i^δ ∈ 𝒞^δ do
  Set P_i = ∅, P̂_i = ∅, and Ŝ_i = ∅; //Find the supremum concept and sub-concept.
  for (Q(F(x_j)), F(x_j), ω_j) ∈ C_i^δ do
    for (Q(F(x_k)), F(x_k), ω_k) ∈ C_i^δ do
      Set S_{i,j} = ∅; //Find the sub-concept of (Q(F(x_j)), F(x_j), ω_j).
      if (Q(F(x_k)), F(x_k), ω_k) is a sub-concept of (Q(F(x_j)), F(x_j), ω_j) then
        P_i = P_i ∪ (Q(F(x_j)), F(x_j), ω_j), and
        S_{i,j} = S_{i,j} ∪ {(Q(F(x_k)), F(x_k), ω_k)};
      end if
    end for
  end for
  Ŝ_i ← S_{i,j};
  for S_{i,m} ∈ Ŝ_i do
    if there exists only one concept (Q(F(x_t)), F(x_t), ω_t) in P_i such that each concept of S_{i,m} is a sub-concept of (Q(F(x_t)), F(x_t), ω_t) then
      S_{i,m} = S_{i,m} and P̂_i ← {(Q(F(x_t)), F(x_t), ω_t)};
    end if
  end for
  Calculate the progressive weighted fuzzy concept (X_{i,m}^δ, B̃_{i,m}^δ, ω_{i,m}^δ) from Definition 5;
  K_i^δ ← (X_{i,m}^δ, B̃_{i,m}^δ, ω_{i,m}^δ);
end for
Return 𝒦^δ = {K_1^δ, K_2^δ, ···, K_t^δ}.
```

on data classification. When a new object $\triangle x$ is added, how to distinguish its class label is a problem that deserves exploring. Meanwhile, $\triangle x$ will cause the change of WFCS generated by the original context. In this section, the incremental learning during the construction of the progressive weighted fuzzy concept space will be considered.

4.1. Classification label prediction after adding objects

In weighted fuzzy concept space, the similarity of object could be described by the Euclidean distance using their attributes, or equivalently, the shorter the distance is, the greater the similarity between objects is. Hence, we can figure out the distance between the new object $\triangle x$ and weighted fuzzy concepts in $\mathscr{K}^\delta$ to determine the class label of $\triangle x$.

Definition 6. Let $(G, M, \tilde{R}, D, J)$ be a fuzzy formal decision context. W is a weight vector of attributes. For a newly added object $\triangle x$ whose membership value with respect to $\tilde{R}$ is $\tilde{B}$, the Euclidean distance between $\triangle x$ and the jth progressive concept $(X_{i,j}^\delta, \tilde{B}_{i,j}^\delta, \omega_{i,j}^\delta)$ in $\mathcal{K}_i^\delta$ is defined as:

$$ED(\triangle x, X_{i,j}^\delta) = \sqrt{\sum_{a\in M}\Big(\omega(a)\big(\tilde{B}(a) - \tilde{B}_{i,j}^\delta(a)\big)\Big)^2}. \tag{14}$$

The value of $ED(\triangle x, X_{i,j}^\delta)$ reflects the degree of similarity of $\triangle x$ with respect to $(X_{i,j}^\delta, \tilde{B}_{i,j}^\delta, \omega_{i,j}^\delta)$. The smaller the $ED(\triangle x, X_{i,j}^\delta)$ is, the stronger the similarity is; otherwise, the weaker the similarity is. $\triangle x$ should be classified based on the principle of minimum distance. If there exist multiple values of the minimum distance, then $\triangle x$ could be classified according to the priority principle of recognition.

Example 6. In a fuzzy formal decision context of Table 6, where $x_1 - x_9$ are from Example 2, $x_{10} - x_{11}$ are two newly added objects. For the object x_{10}, whose membership degree with respect to $\tilde{R}$ is $\tilde{B} = (0.30, 0.55)$, and its real label is 1. Next, the Euclidean distance between x_{10} and the existing progressive weighted fuzzy concept space $\mathscr{K}^\delta$ is computed as follows: $ED(x_{10}, X_{1,1}^\delta) = 0.1063$, $ED(x_{10}, X_{1,2}^\delta) = 0.0928$, $ED(x_{10}, X_{2,1}^\delta) = 0.3510$, and $ED(x_{10}, X_{2,2}^\delta) = 0.2977$. In fact, the distance between x_{10} and $X_{1,2}^\delta$ is minimal in $\mathscr{K}^\delta$ such that x_{10} should be classified into decision class D_1 which is in accordance with the real label 1.

In the following, the algorithm 4 is class label prediction when new objects are randomly added.

Algorithm 4: Class label prediction

Input: The progressive weighted fuzzy concept space $\mathscr{K}^\delta = \{\mathcal{K}_1^\delta, \mathcal{K}_2^\delta, \cdots, \mathcal{K}_t^\delta\}$ and the newly added object $\triangle x$.
Output: The class label of $\triangle x$.

```
for each K_i^δ ∈ 𝒦^δ do
  for each K_{i,j}^δ ∈ K_i^δ do
    Compute ED(△x, K_{i,j}^δ) from Definition 6;
  end for
  The minimal distance is s_i = min(ED(△x, K_{i,j}^δ)), where K_{i,j}^δ ∈ K_i^δ;
end for
Compute argmin_{i=1,2,···,t} s_i in 𝒦^δ;
Return the class label of △x.
```

4.2. Dynamic updating method of the progressive weighted fuzzy concept space

Generally speaking, when a new object $\triangle x$ is added, its class label prediction must be determined first from Algorithm 4. Assume that the class label is j, then $D_j' = D_j \cup \triangle x$. To avoid recalculate the weighted fuzzy concepts in D_j', we study a dynamic updating algorithm of the progressive weighted fuzzy concept space in Algorithm 5.

Based on the fact that the class label of $\triangle x$ is jth class, the central idea of the algorithm is lines 3–9. The first task is to renew the weight vector generated from the new data, and then compute the weight values of multi-attribute intents in $\mathcal{K}_i^\delta$ $(i \neq j)$. Next, it is obvious that $\mathcal{K}_i^\delta$ is updated to $\hat{\mathcal{K}}_i^\delta$. Assuming $|G|$ and $|M|$ are the cardinalities of object and attribute sets, respectively. $|C|$ is the cardinality of class label, and $|D_i|$ is the classification set to decision partition. Thus, in step 3, the time complexity is $O\big(\sum_{i=1}^{|C|}|C|(|D_i|^2 + |\mathcal{K}_i^\delta|)\big)$. Furthermore, the difference between the dynamic updating algorithm and the static updating algorithm is how to acquire new weighted fuzzy concept subspace $\hat{\mathcal{C}}_j$. When updating the jth weighted fuzzy concept, we are first concerned with the difference between $\triangle x$ and x_j, which can be obtained in $O(|M|)$. Subsequently, the computation of the

Table 6
A fuzzy formal decision context.

G	a_1	a_2	d
x_1	0.08	0.72	1
x_2	0.11	0.56	1
x_3	0.04	0.47	1
x_4	0.32	0.65	1
x_5	0.55	0.86	1
x_6	0.90	0.47	2
x_7	0.68	0.14	2
x_8	0.91	0.36	2
x_9	0.75	0.40	2
x_{10}	0.30	0.55	1
x_{11}	0.78	0.45	2

extent of new weighted fuzzy concept can be measured within $O\big(|M|(|G|-x_j^{**})\big)$. Hence, in steps 4–9, the time complexity is $O\Big(\sum_{j=1}^{|D_j|}|M|\big(1+(|G|-x_j^{**})\big)\Big)$. However, for the static updating algorithm, it is essential to recalculate the new weighted fuzzy concepts, with the complexity $O\big(|M|(1+|G|^2)\big)$. In a summary, the time complexity of Algorithm 5 is $O\Big(\sum_{i=1}^{|C|}|C|(|D_i|^2+|\mathcal{K}_i^{\delta}|)+\sum_{j=1}^{|D_j|}|M|\big(1+(|G|-x_j^{**})\big)\Big)$.

According to the above discussion, compared with the static updating algorithm, we see that dynamic updating method of progressive weighted fuzzy concept space improves learning efficiency and decreases unnecessary iteration.

Algorithm 5: Dynamic updating method of progressive weighted fuzzy concept when an object is added

Input: The weighted fuzzy concept space $\mathscr{C}=\{\mathcal{C}_1,\mathcal{C}_2,\cdots,\mathcal{C}_t\}$, the progressive weighted fuzzy concept space $\mathscr{K}^{\delta}=\{\mathcal{K}_1^{\delta},\mathcal{K}_2^{\delta},\cdots,\mathcal{K}_t^{\delta}\}$, the decision partition $G/D=\{D_1,D_2,\cdots,D_t\}$, the newly added data $\triangle x$, and threshold δ.
Output: The updated progressive weighted fuzzy concept space $\hat{\mathscr{K}}^{\delta}=\{\hat{\mathcal{K}}_1^{\delta},\hat{\mathcal{K}}_2^{\delta},\cdots,\hat{\mathcal{K}}_t^{\delta}\}$.

1: The membership degree of $\triangle x$ to $\tilde{R}$ is $\tilde{B}$;
2: Distinguish the class label of $\triangle x$ from Algorithm 4, and assume that the class label of $\triangle x$ is j;
3: Recalculate weight vector W, and update the weight value of multi-attribute intent $\tilde{B}$ in $\mathcal{K}_i^{\delta}$ $(i\neq j)$. Then the updated progressive weighted fuzzy concept is stored in $\hat{\mathcal{K}}_i^{\delta}$ and $\hat{\mathscr{K}}^{\delta}\leftarrow\hat{\mathcal{K}}_i^{\delta}$;
4: Set $\hat{\mathcal{C}}_j\leftarrow\emptyset$;
5: **for** each $x_j\in D_j$ **do**
6:   **if** $\tilde{R}(\triangle x,a)\geq\tilde{R}(x_j,a)$ for each $a\in M$ **then**
7:     $x_j^{**}\leftarrow\triangle x$ and update $\hat{\mathcal{C}}_j\leftarrow(x_j^{**},x_j^{*},\omega_j)$;
8:   **end if**
9: **end for**
10: Compute $(\triangle x^{**},\triangle x^{*},\omega)$ and update $\hat{\mathcal{C}}_j\leftarrow(\triangle x^{**},\triangle x^{*},\omega)$;
11: Calculate an updating weighted fuzzy concept space $\hat{\mathcal{C}}_j^{\delta}$ from Algorithm 2;
12: The updated progressive weighted fuzzy concept space $\hat{\mathcal{K}}_j^{\delta}$ based on Definition 6;
13: Return $\hat{\mathscr{K}}^{\delta}=\{\hat{\mathcal{K}}_1^{\delta},\hat{\mathcal{K}}_2^{\delta},\cdots,\hat{\mathcal{K}}_t^{\delta}\}$.

In fact, in actual incremental learning, how to achieve the dynamic concept learning between the newly added objects and the original progressive weighted fuzzy concept space is the key problem that worths researching. For an increased data without label, we need to judge its label first, and then replace the weighted fuzzy concept space for the incremental learning. In this process, dynamic updating mechanism with respect to concept learning makes the best of information granules of the newly added objects, which facilitates uncertain classification and thereby improves efficiency greatly. Next, dynamic updating mechanism of progressive weighted fuzzy concept with multiple objects is shown in Algorithm 6.

Algorithm 6: Dynamic updating mechanism of progressive weighted fuzzy concept when multiple objects are added (DMPWFC)

Input: The weighted fuzzy concept space $\mathscr{C}=\{\mathcal{C}_1,\mathcal{C}_2,\cdots,\mathcal{C}_t\}$, the progressive weighted fuzzy concept space $\mathscr{K}^{\delta}=\{\mathcal{K}_1^{\delta},\mathcal{K}_2^{\delta},\cdots,\mathcal{K}_t^{\delta}\}$, the decision partition $G/D=\{D_1,D_2,\cdots,D_t\}$, the newly added data block $X=\{X_1,X_2,\cdots,X_k\}$, and threshold δ.
Output: The updated progressive weighted fuzzy concept space $\hat{\mathscr{K}}^{\delta}=\{\hat{\mathcal{K}}_1^{\delta},\hat{\mathcal{K}}_2^{\delta},\cdots,\hat{\mathcal{K}}_t^{\delta}\}$ and the class label of added objects.

for $X_i\in X$ **do**
  for $x_j\in X_i$ **do**
    Distinguish the class label $l_{i,j}$ of x_j from Algorithm 4 and command $L(i,j)=l_{i,j}$;
    The updated progressive weighted fuzzy concept space $\hat{\mathscr{K}}^{\delta}=\{\hat{\mathcal{K}}_1^{\delta},\hat{\mathcal{K}}_2^{\delta},\cdots,\hat{\mathcal{K}}_t^{\delta}\}$ from Algorithm 5;
  end for
end for
Return the class label of X and $\hat{\mathscr{K}}^{\delta}$.

In addition, Fig. 3 shows the incremental learning mechanism of progressive weighted fuzzy concept when objects are added. As Fig. 3 illustrates, given a fuzzy formal decision context with three classes based on decision attribute set, it will generate corresponding weighted fuzzy concept subspaces and classical decision concept space. Subsequently, to promote the concept classification, the updated weighted fuzzy concept space is defined according to Definition 4. At present, it is also noted that each weighted fuzzy concept subspace has a unique class label. Then the progressive weighted fuzzy concept is constructed to remove repeated information based on human cognitive process. For several waiting predicted objects, we need to identify their class labels from the learned progressive weighted fuzzy concept space. Hence, this incremental learning mechanism improves the classification performance owing to the diversified learning of new information.

5. Experimental evaluation

In this section, we will implement a series of experimental analysis to verify the feasibility and superiority of our proposed algorithm DMPWFC. We compare it with two kinds of algorithms, that is, one is FCA-based algorithm and the other is non-FCA-based algorithm. Specifically, the former includes the latest progressive fuzzy three-way concept in object classification tasks, so we compare DMPWFC and ILMPFTC [13]. In addition, the latter also includes some incremental methods with the Naive Bayes (INB) [34], Decision Tree (DT) [35], Nearest Neighbour Classifiers (KNN) [36] with $k=3$, and Random Forest (RF) [37]. We also compare the classification accuracy of incremental learning mechanism between DMPWFC and ILMPFTC, and the superiority of classification mechanism between DMPWFC and the other four algorithms. Furthermore, the setting of different parameters in the experiment is discussed. For the sake of fairness, the above experiments are completed in MATLAB 2015b on a personal computer with Intel(R) Core(TM) i7-4790 CPU @ 3.6 GHz and 16 GB memory.

For DMPWFC, the threshold δ is a variable parameter which influences the classification accuracy from Definition 4. Hence,

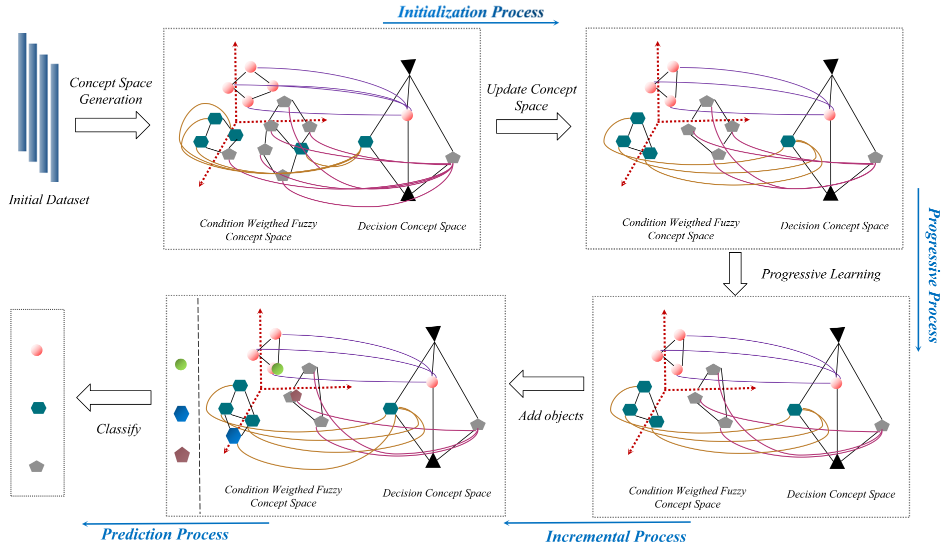

Fig. 3. The dynamic updating mechanism of DMPWFC.

Table 7
Data description.

ID	Datasets	Samples	Attributes	Class
1	Wpbc	198	33	2
2	SonarEW	208	60	2
3	Thyroid	215	5	3
4	HeartEW	270	13	2
5	Derm	366	34	6
6	Australian	690	14	2
7	Wifi_localization	2000	7	4
8	Spam	4601	57	2
9	German	1000	21	2
10	Sat	4435	36	6

the δ is set to a value between 0.3 and 0.8 with step 0.025. With respect to each dataset, 90% of the data is used to train the model, and the remaining 10% is separated into 8 portions and added to the test set to evaluate the classification accuracy between DMPWFC and ILMPFTC.

For the other four compared algorithms, the performances on these datasets are employed by 10-fold cross-validation. Each dataset is randomly separated into 10 parts, nine of them is used to train the model, and the other one is for testing. After ten cycles, the average performance and standard deviation of ten cycles are considered as the final evaluation indices.

5.1. Collection of datasets

In the experiment, ten datasets downloaded from UCI Machine Learning Repository [38] are depicted in Table 7. In the procedure of data pretreatment, value of all attributes are first normalized to obtain the membership degree into [0.1, 0.9], and the formula is given as follows:

$$\tilde{R}(x_i, a_j) = 0.8\frac{f(x_i, a_j) - min\big(f(a_j)\big)}{max\big(f(a_j)\big) - min\big(f(a_j)\big)} + 0.1 \tag{15}$$

where $f(x_i, a_j)$ is the membership value of x_i with respect to attribute a_j, and $max\big(f(a_j)\big)$ and $min\big(f(a_j)\big)$ are respectively denoted as the maximum and minimum of all objects in a_j. In Eq. (7), $H(a)$ is a formula with logarithmic function, in order to avoid such a situation $log_2 0$, then it is normalized to [0.1, 0.9]. In fact, $\tilde{R}(x_i, a_j)$ is the membership degree (x_i, a_j) to fuzzy relation $\tilde{R}$, that is, the larger the value of $f(x_i, a_j)$ is, the greater the degree of $\tilde{R}(x_i, a_j)$ is.

5.2. Results and analysis

The results of the optimal δ and classification accuracies of incremental learning mechanism under DMPWFC and ILMPFTC in the ten datasets are presented in Table 8, where the underlined boldface underlines the best accuracy performance over the other algorithm ILMPFTC. Apparently, we can demonstrate that the average accuracies of DMPWFC with respect to each dataset are better than those of ILMPFTC except for dataset Derm and Sat. Furthermore, the standard deviation of datasets Wpbc, SonarEW, Thyroid, Wifi_localization, Spam, and Sat under DMPWFC is less than those of ILMPFTC. From Fig. 3, under the optimal δ, as the adding objects increase, the classification accuracies of DMPWFC first increase and then decrease, and are higher than those of ILMPFTC (Wpbc, SonarEW, HeartEW, Australian, Wifi_localization, and German). In addition, the statistical performance comparison is shown in Table 9. The underlined boldface indicates the best average classification accuracy of DMPWFC in the ten datasets. When Comparing the statistical evaluation, we see that the critical value employed by Wilcoxon test is 0.05 and the test P-value is $0.0273 < 0.05$. Hence, the null hypothesis is rejected and we accept the alternative hypothesis that there is a significant difference under two algorithms. In summary, the classification performance of DMPWFC is evidently better than that of ILMPFTC.

Hereafter, the comparative results of non-FCA-based algorithms are presented in Table 10. The underlined boldface data sets indicate the best classification accuracy in the corresponding dataset. Concretely, of the total ten cases, DMPWFC gets the maximum accuracy in six cases, and INB and KNN achieve the maximum accuracy in one case and three cases, respectively. Furthermore, from Table 11, the average accuracy of DMPWFC is higher than the other four algorithms, and its standard deviation is less than those of algorithms DT, KNN, INB, and RF. In particular, we can see that as the objects increase, the classification accuracies increase first and then decrease to some extent from Fig. 4. Therefore, it should be pointed out that the more data are added, the classification accuracies will not necessarily be improved.

Additionally, the statistical significance of the five classification algorithms can be compared by the Friedman test [39] and Bonferroni–Dunn test [40]. With respect to Friedman test, a Fisher distribution F_F measured the performance of different algorithms is denoted as follows:

$$F_F = \frac{(N-1)\chi_F^2}{N(k-1) - \chi_F^2} \sim F\big(k-1, k-1(N-1)\big) \tag{16}$$

Table 8
The optimal δ and accuracy(%) comparison between DMPWFC and ILMPFTC.

ID	Model	δ	t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8	Average±std
1	DMPWFC	0.8000	100.00	100.00	95.24	96.43	97.14	95.24	87.76	76.79	**93.57 ± 7.7875**
	ILMPFTC	0.5000	85.71	64.29	61.90	67.86	62.86	61.90	65.31	62.50	66.54 ± 8.0065
2	DMPWFC	0.8000	85.71	92.86	95.24	96.43	80.00	78.57	67.35	58.93	**81.89 ± 13.5105**
	ILMPFTC	0.3000	85.71	50.00	33.33	46.43	48.57	54.76	46.94	48.21	51.75 ± 15.0180
3	DMPWFC	0.8000	100.00	100.00	100.00	100.00	100.00	100.00	95.92	92.86	**98.60 ± 2.7227**
	ILMPFTC	0.7500	100.00	100.00	100.00	100.00	100.00	100.00	95.92	92.86	98.60 ± 2.7227
4	DMPWFC	0.6000	88.89	88.89	88.89	91.67	88.89	85.19	80.95	81.94	**86.91 ± 3.8083**
	ILMPFTC	0.4250	77.78	77.78	70.37	77.78	80.00	77.78	76.19	76.39	76.76 ± 2.8272
5	DMPWFC	0.8000	100.00	100.00	94.87	86.54	89.23	91.03	91.21	92.31	93.15 ± 4.8525
	ILMPFTC	0.6250	100.00	100.00	100.00	94.23	95.38	96.15	93.41	94.23	**96.68 ± 2.8719**
6	DMPWFC	0.7250	88.00	90.00	90.67	90.00	88.00	86.00	83.43	79.00	**86.89 ± 3.9841**
	ILMPFTC	0.3000	72.00	78.00	81.33	78.00	75.20	72.67	72.57	72.00	75.22 ± 3.5265
7	DMPWFC	0.7500	62.50	76.56	84.38	88.28	86.25	73.44	70.98	73.83	**77.03 ± 8.7556**
	ILMPFTC	0.4250	53.13	73.44	82.29	86.72	82.50	72.40	65.63	59.38	71.93 ± 11.9001
8	DMPWFC	0.6250	100.00	97.30	96.85	95.95	95.95	96.17	96.72	97.13	**97.01 ± 1.3148**
	ILMPFTC	0.3000	95.95	95.95	97.30	96.96	94.05	91.44	91.51	89.36	94.06 ± 2.9643
9	DMPWFC	0.8000	89.19	91.89	93.69	91.89	90.27	86.49	79.54	72.64	**86.95 ± 7.2720**
	ILMPFTC	0.5500	70.27	78.38	81.08	81.76	82.16	80.18	76.06	69.26	77.39 ± 5.1101
10	DMPWFC	0.5750	93.33	89.09	75.96	73.64	70.18	68.18	67.27	67.12	75.60 ± 10.1843
	ILMPFTC	0.3000	98.79	96.97	92.32	76.06	67.52	64.85	62.86	61.67	**77.63 ± 15.9394**

Table 9
The results of average accuracy and Wilcoxon test between DMPWFC and ILMPFTC.

Algorithm	ID1	ID2	ID3	ID4	ID5	ID6	ID7	ID8	ID9	ID10	Average	P-value
DMPWFC	93.57	81.89	98.60	86.91	93.15	86.89	77.03	97.01	86.95	75.60	**87.76**	–
ILMPFTC	66.54	51.75	98.60	76.76	96.68	75.22	71.93	94.06	77.39	77.63	78.66	0.0273

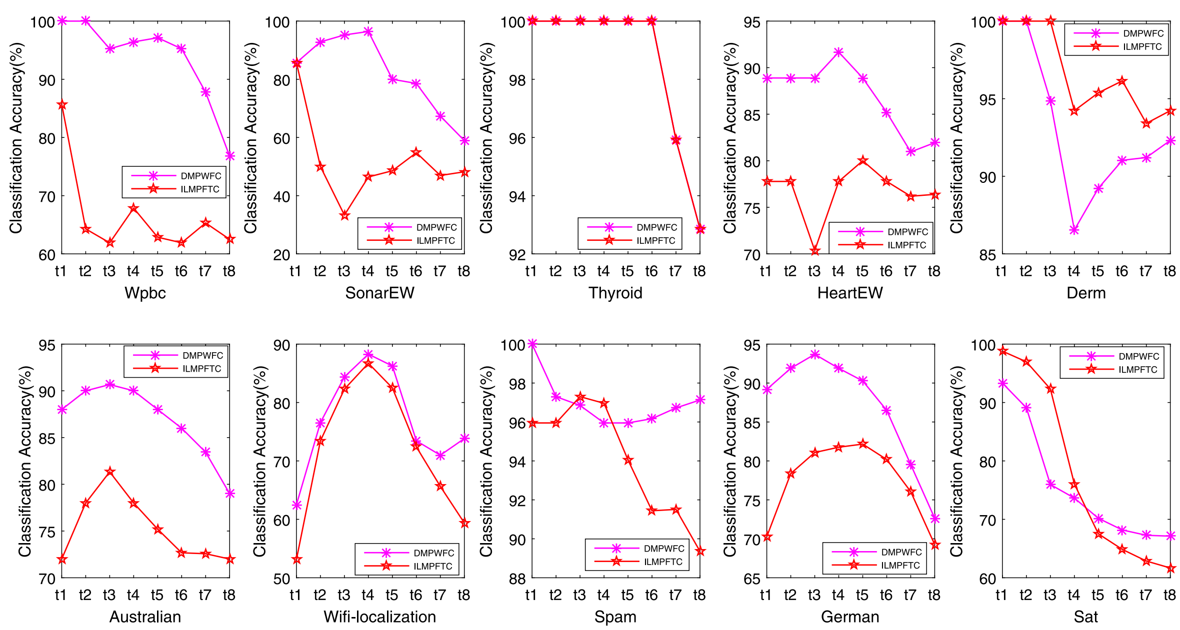

Fig. 4. Accuracy(%) comparison between DMPWFC and ILMPFTC.

where $\chi_F^2 = \frac{12N}{k(k+1)}\left(\sum_{i=1}^{k} R_i^2 - \frac{k(k+1)^2}{4}\right)$. N is the cardinality of datasets, k is the number of different algorithms, and R_i is the average sort of algorithm R_i about all datasets. At first, the initial hypothesis is that the accuracy performance of all algorithms is considered to be the same. If $F_F > F\big(k-1, k-1(N-1)\big)$, then the initial hypothesis will be rejected. In fact, $F_F = 3.4654$ and the critical value of $F(4, 36) = 2.11$ at level $\alpha = 0.1$. Hence, the initial hypothesis is rejected and we accept the alternative assumption that the classification performance of five models is remarkably different.

Furthermore, the classification performance of five models is further employed by Bonferroni–Dunn test. Suppose that the mean rank of between two models exceeds the critical value

$$CD_\alpha = q_\alpha \sqrt{\frac{k(k+1)}{6N}}$$

and then the classification performance of the above two models is evidently different, in which q_α is denoted as the critical value in the test. The rank result of five models is shown in Table 12.

At level $\alpha = 0.1$, we notice that the critical value $q_{0.1} = 2.241$ in [40], afterwards, $CD_\alpha = 1.58$ ($k = 5$ and $N = 10$). Therefore,

Table 10
The comparison of accuracy(%) performance about five algorithms.

ID	Model	t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8	Average±std
1	DMPWFC	100.00	100.00	95.24	96.43	97.14	95.24	87.76	76.79	**93.57 ± 7.7875**
	DT	65.00	60.00	63.33	65.00	66.00	64.17	65.71	66.25	64.43 ± 2.0345
	KNN	65.00	70.00	71.67	71.25	72.00	70.83	72.14	70.00	70.36 ± 2.3170
	INB	80.00	82.50	76.67	75.00	77.00	75.83	76.43	79.38	77.85 ± 2.5327
	RF	80.00	77.50	71.67	73.75	72.00	72.50	75.71	74.38	74.69 ± 2.9105
2	DMPWFC	85.71	92.86	95.24	96.43	80.00	78.57	67.35	58.93	**81.89 ± 13.5105**
	DT	60.00	67.50	71.67	71.25	72.00	72.50	75.00	74.38	70.54 ± 4.8208
	KNN	60.00	67.50	71.67	76.25	79.00	80.83	81.43	81.25	74.74 ± 7.7810
	INB	55.00	60.00	66.67	70.00	70.00	70.83	69.29	71.25	66.63 ± 5.9521
	RF	75.00	77.50	86.67	77.50	82.00	79.17	70.71	70.00	77.32 ± 5.5464
3	DMPWFC	100.00	100.00	100.00	100.00	100.00	100.00	95.92	92.86	98.60 ± 2.7227
	DT	85.00	90.00	90.00	90.00	92.00	91.67	91.43	92.50	90.32 ± 2.3672
	KNN	95.00	97.50	98.33	97.50	97.00	97.50	97.86	96.25	97.12 ± 1.0498
	INB	100.00	100.00	100.00	100.00	99.00	99.17	99.29	98.75	**99.53 ± 0.5297**
	RF	95.00	92.50	96.67	98.75	96.00	98.33	97.14	94.38	96.10 ± 2.0899
4	DMPWFC	88.89	88.89	88.89	91.67	88.89	85.19	80.95	81.94	**86.91 ± 3.8083**
	DT	76.67	80.00	76.67	80.83	82.00	80.56	79.52	79.17	79.43 ± 1.9092
	KNN	73.33	80.00	78.89	79.17	80.00	78.89	79.52	77.92	78.46 ± 2.1817
	INB	73.33	78.33	76.67	80.83	82.00	81.67	81.43	81.67	79.49 ± 3.1294
	RF	80.00	68.33	71.11	75.00	72.67	76.11	75.71	72.50	73.93 ± 3.5625
5	DMPWFC	100.00	100.00	94.87	86.54	89.23	91.03	91.21	92.31	93.15 ± 4.8525
	DT	92.50	88.75	90.00	91.88	93.00	93.75	94.29	93.13	92.16 ± 1.8964
	KNN	100.00	97.50	97.50	98.13	98.00	98.33	97.86	97.81	**98.14 ± 0.8039**
	INB	95.00	92.50	94.17	94.38	95.00	95.42	95.36	95.31	94.64 ± 0.9785
	RF	90.00	88.75	88.33	86.25	89.00	89.58	93.57	89.06	89.32 ± 2.0528
6	DMPWFC	88.00	90.00	90.67	90.00	88.00	86.00	83.43	79.00	**86.89 ± 3.9841**
	DT	86.25	85.63	82.92	82.81	83.25	83.75	83.39	83.59	83.95 ± 1.2775
	KNN	87.50	86.88	85.00	85.00	84.50	83.96	83.93	83.59	85.04 ± 1.4248
	INB	88.75	86.25	85.42	85.00	85.50	84.79	85.36	86.09	85.89 ± 1.2541
	RF	81.25	83.75	78.33	82.81	83.75	79.79	78.04	79.84	80.95 ± 2.3035
7	DMPWFC	62.50	76.56	84.38	88.28	86.25	73.44	70.98	73.83	77.03 ± 8.7556
	DT	96.00	96.80	96.27	96.40	95.84	95.93	96.34	96.70	96.28 ± 0.3500
	KNN	100.00	99.40	98.53	98.50	98.40	98.33	98.46	98.50	**98.77 ± 0.6020**
	INB	99.20	99.40	98.53	97.70	97.44	97.53	97.77	97.95	98.19 ± 0.7624
	RF	99.60	99.40	98.80	98.00	97.20	95.73	96.06	96.10	97.61 ± 1.5636
8	DMPWFC	100.00	97.30	96.85	95.95	95.95	96.17	96.72	97.13	**97.01 ± 1.3148**
	DT	90.53	90.96	90.47	91.10	91.68	91.98	92.43	91.86	91.38 ± 0.7186
	KNN	85.96	88.77	87.31	87.76	89.12	89.68	90.50	90.42	88.69 ± 1.5867
	INB	87.19	89.39	88.01	89.30	90.56	91.08	91.93	92.19	89.96 ± 1.8010
	RF	92.63	93.95	94.50	92.50	91.61	90.09	91.50	89.30	92.01 ± 1.7751
9	DMPWFC	89.19	91.89	93.69	91.89	90.27	86.49	79.54	72.64	**86.95 ± 7.2720**
	DT	70.83	65.00	68.06	68.13	66.83	67.50	67.62	67.08	67.63 ± 1.6281
	KNN	78.33	75.00	74.44	72.92	73.00	71.67	72.02	71.88	73.66 ± 2.2385
	INB	76.67	74.58	74.17	74.17	73.83	73.06	73.57	73.33	74.17 ± 1.1224
	RF	76.67	76.25	72.25	75.83	72.33	72.08	72.14	69.17	73.37 ± 2.6156
10	DMPWFC	93.33	89.09	75.96	73.64	70.18	68.18	67.27	67.12	75.60 ± 10.1843
	DT	83.82	86.55	85.09	83.14	82.58	83.06	83.48	82.73	83.81 ± 1.3590
	KNN	91.82	93.18	91.82	89.77	89.64	89.94	90.44	90.61	**90.90 ± 1.2522**
	INB	87.45	89.00	87.64	84.41	84.47	84.03	84.86	85.64	85.94 ± 1.8495
	RF	88.36	88.58	82.58	80.59	81.09	80.67	81.45	80.61	83.33 ± 2.4795

Table 11
The results of average accuracy and Wilcoxon test under five algorithms.

Algorithm	*ID*1	*ID*2	*ID*3	*ID*4	*ID*5	*ID*6	*ID*7	*ID*8	*ID*9	*ID*10	Average±std
DMPWFC	93.57	81.89	98.60	86.91	93.15	86.89	77.03	97.01	86.95	75.60	**87.76 ± 7.9265**
DT	64.43	70.54	90.32	79.43	92.16	83.95	96.28	91.38	67.63	83.81	81.99 ± 11.1789
KNN	70.36	74.74	97.12	78.46	98.14	85.04	98.77	88.69	73.66	90.90	85.59 ± 10.7750
INB	77.85	66.63	99.53	79.49	94.64	85.89	98.19	89.96	74.17	85.94	85.23 ± 10.7489
RF	74.69	77.32	96.10	73.93	89.32	80.95	97.61	92.01	73.37	83.33	84.01 ± 9.3082

we can see that algorithm DMPWFC outperforms algorithms DT and RF at level $\alpha = 0.1$. Nevertheless, there is no obvious evidence to demonstrate the statistical difference with algorithms KNN and INB (see Fig. 5).

6. Conclusion

With the rapid development of data, how to identify the dynamic classification learning is a crucial issue. Moreover, some classical CCL methods based on a fuzzy formal decision context mainly focus on the same weights of attributes before constructing the fuzzy concepts, ignoring the internal knowledge of attributes and decisions in advance. In this paper, we have

Table 12
Rank of classification algorithms.

ID	DMPWFC	DT	KNN	INB	RF
1	1	5	4	2	3
2	1	4	3	5	2
3	2	5	3	1	4
4	1	3	4	2	5
5	3	4	1	2	5
6	1	4	3	2	5
7	5	4	1	2	3
8	1	3	5	4	2
9	1	5	3	2	4
10	5	3	1	2	4
Average	2.10	4.00	2.80	2.40	3.70

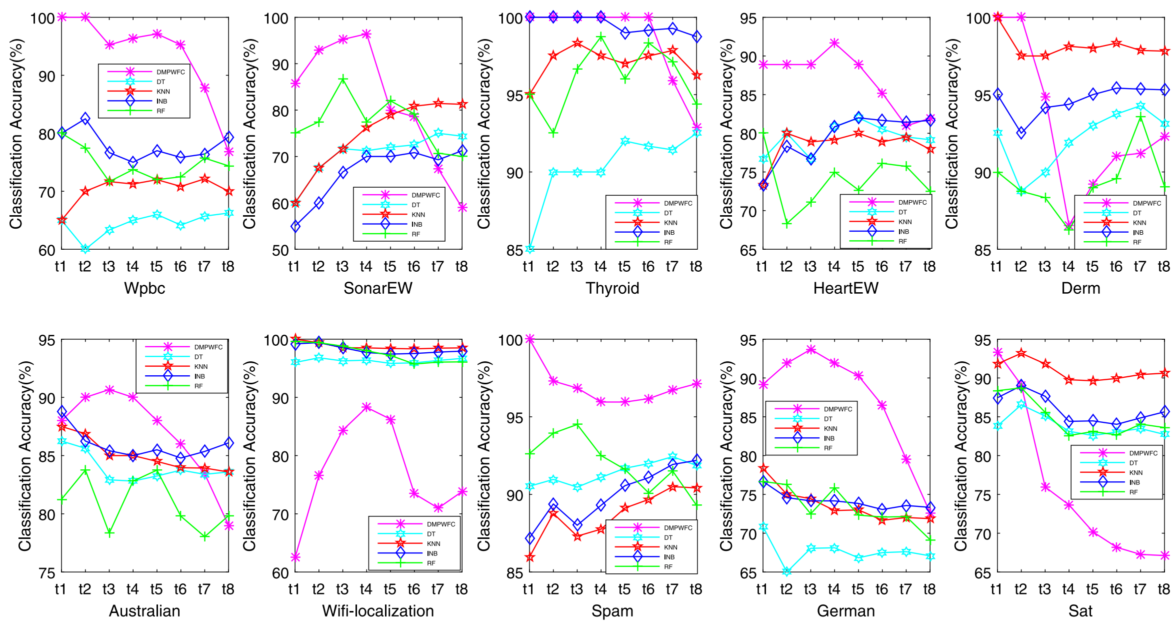

Fig. 5. Accuracy(%) comparison of five algorithms.

proposed a new model, named DMPWFC, to recognize different class labels under dynamic environments. Concretely, by denoting weight values of fuzzy concept, we first investigate weighted fuzzy concept space. Then certain algorithms of CCL are discussed to update the weighted fuzzy concepts and construct the progressive weighted fuzzy concepts for achieving concept classification based on human cognitive process. Furthermore, an incremental CCL approach for determining classification label and updating fuzzy concepts is researched in the sense of preserving information. At last, to better comprehend our approach, a series of comparative experiments on ten UCI datasets are performed to demonstrate that DMPWFC can achieve better classification performance.

This paper only studies the incremental CCL algorithm under the prerequisite of adding objects without bringing out attribute increments. In addition, the proposed DMPWFC algorithm cannot directly deal with numerical data when achieving incremental learning. Our future research will consider these issues, so as to improve the accuracies and efficiency of incremental learning approaches with respect to CCL.

CRediT authorship contribution statement

Chengling Zhang: Conceptualization, Methodology, Software, Writing – original draft. **Eric C.C. Tsang:** Formal analysis, Investigation, Supervision, Validation. **Weihua Xu:** Supervision, Validation. **Yidong Lin:** Software, Visualization. **Lanzhen Yang:** Investigation, Data Curation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The authors are unable or have chosen not to specify which data has been used.

Acknowledgments

This work is supported by the Macao Science and Technology Development Funds (0019/2019/A1, and 0075/2019/A2) and the Natural Science Foundation of China (No. 61976245, and No. 12201284), and Natural Science Foundation of Fujian Province, China (No. 2022J05169).

References

[1] Y. Wang, On cognitive computing, Int. J. Softw. Sci. Comput. Intell. 1 (3) (2009) 1–15.
[2] J. Feldman, The simplicity principle in human concept learning, Curr. Direct. Psychol. Sci. 12 (6) (2003) 227–232.
[3] B. Fan, E.C.C. Tsang, W. Xu, et al., Attribute-oriented cognitive concept learning strategy: A multi-level method, Int. J. Mach. Learn. Cybern. 10 (9) (2019) 2421–2437.

[4] E.C.C. Tsang, B. Fan, D. Chen, et al., Multi-level cognitive concept learning method oriented to data sets with fuzziness: A perspective from features, Soft Comput. 24 (5) (2020) 3753–3770.
[5] T. Zhang, H. Li, M. Liu, et al., Incremental concept-cognitive learning based on attribute topology, Internat. J. Approx. Reason. 118 (2020) 173–189.
[6] Y. Shi, Y. Mi, J. Li, et al., Concept-cognitive learning model for incremental concept learning, IEEE Trans. Syst. Man Cybern.: Syst. 51 (2) (2018) 809–821.
[7] Y. Shi, Y. Mi, J. Li, et al., Concurrent concept-cognitive learning model for classification, Inform. Sci. 496 (2019) 65–81.
[8] S. Kuznetsov, Machine learning and formal concept analysis, in: International Conference on Formal Concept Analysis, Springer, Berlin, Heidelberg, 2004, pp. 287–312.
[9] Y. Mi, Y. Shi, J. Li, et al., Fuzzy-based concept learning method: Exploiting data with fuzzy conceptual clustering, IEEE Trans. Cybern. (2020).
[10] Y. Wang, On concept algebra: A denotational mathematical structure for knowledge and software modeling, Int. J. Cogn. Inform. Nat. Intell. 2 (2) (2008) 1–19.
[11] W. Xu, J. Pang, S. Luo, A novel cognitive system model and approach to transformation of information granules, Internat. J. Approx. Reason. 55 (3) (2014) 853–866.
[12] D. Guo, C. Jiang, R. Sheng, et al., A novel outcome evaluation model of three-way decision: A change viewpoint, Inform. Sci. 607 (2022) 1089–1110.
[13] K. Yuan, W. Xu, W. Li, et al., An incremental learning mechanism for object classification based on progressive fuzzy three-way concept, Inform. Sci. 584 (2022) 127–147.
[14] C. Luo, T. Li, H. Chen, et al., Incremental rough set approach for hierarchical multicriteria classification, Inform. Sci. 429 (2018) 72–87.
[15] R. Wille, Restructuring lattice theory: An approach based on hierarchies of concepts, in: I. Rival (Ed.), Ordered Sets, Reidel, Dordrecht, The Netherlands, 1982, pp. 445–470.
[16] B. Ganter, R. Wille, Formal Concept Analysis: Mathematical Foundations, Springer, New York, 1999.
[17] S. Yahia, K. Arour, A. Slimani, A. Jaoua, Discovery of compact rules in relational databases, Inform. Sci. 4 (3) (2000) 497–511.
[18] J. Qi, L. Wei, Y. Yao, Three-way formal concept analysis, in: International Conference on Rough Sets and Knowledge Technology, Springer, Cham, 2014, pp. 732–741.
[19] D. Chen, J. Li, R. Lin, Formal concept analysis of multi-scale formal context, J. Ambient Intell. Humaniz. Comput. 11 (11) (2020) 5315–5327.
[20] J. Li, Z. He, Q. Zhu, An entropy-based weighted concept lattice for merging multi-source geo-ontologies, Entropy 15 (6) (2013) 2303–2318.
[21] S. Zhang, P. Guo, J. Zhang, et al., A completeness analysis of frequent weighted concept lattices and their algebraic properties, Data Knowl. Eng. 81 (2012) 104–117.
[22] S. Peng, A. Yang, L. Cao, et al., Social influence modeling using information theory in mobile social networks, Inform. Sci. 379 (2017) 146–159.
[23] C. Shannon, A mathematical theory of communication, Bell Syst. Tech. J. 27 (3) (1948) 379–423.
[24] B. Kosko, Fuzzy entropy and conditioning, Inform. Sci. 40 (2) (1986) 165–174.
[25] H. Lee, C. Chen, J. Chen, et al., An efficient fuzzy classifier with feature selection based on fuzzy entropy, IEEE Trans. Syst. Man Cybern. B 31 (3) (2001) 426–432.
[26] A. Luca, S. Termini, A definition of a nonprobabilistic entropy in the setting of fuzzy sets theory, Inf. Control 20 (4) (1972) 301–312.
[27] S. Al-Sharhan, F. Karray, W. Gueaieb, et al., Fuzzy entropy: A brief survey, in: 10th IEEE International Conference on Fuzzy Systems, Vol. 3, IEEE, 2001, pp. 1135–1139.
[28] P. Singh, A. Cherukuri, J. Li, Concepts reduction in formal concept analysis with fuzzy setting using Shannon entropy, Int. J. Mach. Learn. Cybern. 8 (1) (2017) 179–189.
[29] C. Kumar, M. Ishwarya, C. Loo, Formal concept analysis approach to cognitive functionalities of bidirectional associative memory, Biol. Inspired Cogn. Archit. 12 (2015) 20–33.
[30] J. Li, C. Mei, W. Xu, et al., Concept learning via granular computing: A cognitive viewpoint, Inform. Sci. 298 (2015) 447–467.
[31] J. Li, Y. Mi, W. Liu, Incremental cognition of concepts: Theories and methods, Chinese J. Comput. 42 (9) (2019) 1–19.
[32] E. Yan, C. Yu, L. Lu, et al., Incremental concept cognitive learning based on three-way partial order structure, Knowl.-Based Syst. 220 (2021) 106898.
[33] Y. Zhao, J. Li, W. Liu, et al., Cognitive concept learning from incomplete information, Int. J. Mach. Learn. Cybern. 8 (1) (2017) 159–170.
[34] V. Losing, B. Hammer, H. Wersing, Incremental on-line learning: A review and comparison of state of the art algorithms, Neurocomputing 275 (2018) 1261–1274.
[35] I. Witten, E. Frank, M. Hall, et al., Data mining: Practical machine learning tools and techniques, Morgan Kaufmann 2 (4) (2005).
[36] T. Cover, P. Hart, Nearest neighbor pattern classification, IEEE Trans. Inform. Theory 13 (1) (1967) 21–27.
[37] D. Cutler, J. Edwards, K. Beard, et al., Random forests for classification in ecology, Ecology 88 (11) (2007) 2783–2792.
[38] A. Asuncion, UCI Machine Learning Repository, University of California, Irvine, School of Information and Computer Sciences, 2007, http://www.ics.uci.edu/mlearn/MLRepository.
[39] M. Friedman, A comparison of alternative tests of significance for the problem of M rankings, Ann. Math. Stat. 11 (1) (1940) 86–92.
[40] O. Dunn, Multiple comparisons among means, J. Amer. Statist. Assoc. 56 (293) (1961) 52–64.