Information Fusion 100 (2023) 101962

Contents lists available at ScienceDirect

Information Fusion

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

Full length article

M-FCCL: Memory-based concept-cognitive learning for dynamic fuzzy data classification and knowledge fusion

Doudou Guo [a], Weihua Xu [a,*], Yuhua Qian [b], Weiping Ding [c]

[a] College of Artificial Intelligence, Southwest University, Chongqing, 400715, PR China
[b] Institute of Big Data Science and Industry, Shanxi University, Taiyuan, 030006, PR China
[c] School of Information Science and Technology, Nantong University, Nantong 226019, PR China

ARTICLE INFO

Keywords:
Concept-cognitive learning
Dynamic data classification
Knowledge fusion
Granular computing
Three-way decision

ABSTRACT

Concept-cognitive learning (CCL) is an emerging field for studying the representation and processing of knowledge embedded in data. Many efforts are focused on this field due to the interpretability and effectiveness of the formal concept (not pseudo concept). However, the standard CCL methods cannot tackle continuous data directly. Although the current fuzzy-based CCL (FCCL) is a straightforward approach to discovering the knowledge embedded in continuous data, it does not sufficiently utilize the native advantage of concepts in simulating the cognitive mechanism. Then it causes it to be incomplete and complex cognition. Inspired by the memory mechanism, this paper combines the recalling and forgetting mechanisms with CCL, called memory-based concept-cognitive learning (M-FCCL). Specifically, a cosine measure is introduced to describe the relationship of samples and construct cosine-similar granules to learn the concept. Subsequently, a fuzzy three-way concept based on the cosine similar granules is defined to represent and discover knowledge. Furthermore, two memory mechanisms are borrowed for the process of concept cognition for dynamic data classification and knowledge fusion: concept-recalling can enhance the effectiveness of concept learning, and concept-forgetting can effectively reduce the complexity of concept cognition. Finally, some experiments are compared with other methods on 16 benchmark datasets to show that M-FCCL achieves superior performance. Specifically, on these datasets, the proposed M-FCCL method achieves 17.02% and 18.54% classification accuracy gain compared with some advanced CCL mechanisms and popular classification methods.

1. Introduction

As an emerging computing paradigm, cognitive computing is modeled on the human brain that implements cognitive intelligence by trying to simulate the cognitive mechanisms of the human brain, such as perception, reasoning, learning, and recognition [1]. The human brain is viewed as, by far, the most advanced and efficient cognitive, essential for realizing artificial intelligence. A fundamental concept of artificial intelligence covers the data-information-knowledge-wisdom (DIKW) hierarchy [2–4], in which knowledge, the prerequisite for wisdom, is viewed as information related to wisdom formation. Meanwhile, a value conceptual model in data science is the symbols-meaning-value (SMV) space [5], where data are regarded as a resource, and the power of data is the knowledge embedded in data, the wisdom views as the value of data. From an interpretability viewpoint, the representation and learning of concepts are critical topics for studying cognitive computing and artificial intelligence [6,7]. An influential theory of interpretability is granular computing(GrC), which plays an essential role in studying concerning information and knowledge processing when humans use concepts, symbols, and models to describe the objective world [8–10].

Concept, a fundamental carrier of knowledge representation, is the basic cognitive unit of the human brain that can effectively describe the general and objective essence of knowledge [11]. The intent and extent of a concept are uniquely determinable, and the concept by which they refer to each other is unique and can be used to describe different types of things. In 1982, Wille [12] proposed the formal concept analysis (FCA) theory to describe a formal concept from a mathematics perspective. Inspired by FCA and cognitive computing, a research topic about cognitive concept learning is gradually coming into view. A cognitive concept method based on granular computing is proposed by Zhang in his seminal paper [13]. The paper [14] discusses a mathematic algebraic model for concept learning from a cognitive viewpoint. A three levels framework of concept learning based on cognitive informatics and granular computing is presented by Yao in paper [15], which is necessary research for investigating concept

* Corresponding author.
E-mail addresses: doudou876517690@126.com (D. Guo), chxuwh@gmail.com (W. Xu), jinchengqyh@126.com (Y. Qian), dwp9988@163.com (W. Ding).

https://doi.org/10.1016/j.inffus.2023.101962
Received 18 April 2023; Received in revised form 25 July 2023; Accepted 31 July 2023
Available online 3 August 2023

cognition and concept learning. Note that three-way decision based on three-level thinking is another widely recognized methodology of GrC for investigating the conceptual model in papers [16–18], and its advantage lies in the granulation thinking paradigm about trichotomous (i.e., triadic thinking, triadic computing, triadic processing). Inspired by this theory, some scholars [19,20] combine three-way decisions with the formal concept to study formal concept analysis from positive and negative information, i.e., three-way concept analysis. There is no denying that three-way concept analysis is a fruitful marriage for study concept knowledge presentation and learning, which is one of the main reasons why this article characterizes knowledge in a framework of fuzzy three-way concept.

The research mentioned above mainly focuses on the concept learning framework from the perspective of the cognitive mechanism of the human brain. Note that human-level concept cognition can often generalize successfully from just a single instance, yet machine learning generally requires tens or hundreds of instances to accomplish with similar accuracy [6,21]. Hence, concept-cognitive learning theory emerges, via a unified view of the cognitive mechanism and machine learning, simulating the human cognitive process. As an essential topic of concept learning, knowledge representation, and cognitive computing, CCL has been investigated from different aspects. For example, Li [22] studies the granule concept learning method from a cognitive viewpoint. Li [23] proposes a three-way cognitive concept learning based on multi-granularity. Xu [24] combines GrC theory with two-way learning to study the fuzzy-based concept learning method. Xu and Guo [25,26] investigate two-way concept-cognitive learning from a fuzzy-based progressive learning viewpoint and a concept movement viewpoint. Zhang [27] propose a concept-cognitive learning method for multi-source information fusion. Qian [28] gives three-way concept lattices based on apposition and sub-position of formal contexts. Shi and Mi [29,30] discuss concept-cognitive learning for incremental concept learning. Zhang [31] constructed a causal asymmetry analysis using an incremental concept tree.

Concept-cognitive learning is the science of cognition and learning things via concepts [20]. Note that two advantages of CCL in dealing with knowledge discovery: (1) different concepts have the native character of recognizing different ontologies, and (2) the process of concept learning has the character of incremental learning that can integrate past experiences into itself to deal with dynamic data. Therefore, Mi [32] proposes a fuzzy-based CCL method to exploit data via a fuzzy conceptual clustering method. Yuan [33] establishes an incremental learning mechanism based on the progressive fuzzy three-way concept to deal with dynamic data. Zhang [34] uses weighted fuzzy concepts to design an incremental mechanism for the fuzzy-based CCL, etc. Since the weak learning ability of fuzzy concepts in classification tasks, the pseudo concept has been proposed in the above Ref. [32–34], which differs from regular fuzzy concepts. The pseudo concept, a derivative of some concepts, strengthens the influences of fuzzy concepts in the concept-cognitive process. Although these methods are significantly better than other classification algorithms in terms of rationality, they still have some problems.

- Existing incremental concept-cognitive learning models adopt a gradual strategy to incorporate new concepts. However, this strategy will result in the unsatisfactory performance of the model for inappropriate incremental learning.
- Most existing dynamic concept-cognitive learning methods focus on incremental learning mechanisms while ignoring the knowledge accumulation in a dynamic environment, which could disturb the subsequent cognitive learning and degrade performance.
- The current fuzzy concept-cognitive learning method does not sufficiently utilize the native advantage of fuzzy concepts in ontology recognition, which causes it to have weak interpretability due to only representing knowledge via pseudo concept.

Inspired by the current work of granular computing and machine learning [10,35–37], the main thread adopted in the current article is the memory-based CCL approach for dynamic data classification and knowledge fusion in a fuzzy formal context to address the issues mentioned above. This article aims to model novel functionalities of concept-cognitive learning to deal with dynamic data efficiently, namely forgetting the unnecessary knowledge in the original concept space and recalling part of the necessary knowledge to fuse knowledge for concept update. The main contributions of this paper are as follows.

- It presents a memory-based concept-cognitive learning method in a fuzzy formal context for dynamic data classification and knowledge fusion, and one can considerably enhance classification performance in most of the existing FCCL models and other classification algorithms.
- It defines the fuzzy three-way concept based on a cosine similarity granule to characterize and discover the knowledge embedded in data. One can simultaneously study the information and knowledge in formal concepts from positive and negative angles to ensure the integrity of knowledge depiction and reduce cognitive bias.
- It designs a concept-recalling mechanism for dynamic knowledge updating, integrating the past experience into itself by recalling related knowledge to reduce the complexity of the learning system. Meanwhile, some experiments verify that the M-FCCL performs better than other methods.
- It provides a novel thought of cognitive concept based on concept forgetting via forgetting some unnecessary concepts in concept space, which can enhance the performance of concept recognition and omit the step for pseudo-concept learning [32–34]. Moreover, in this way, one can sufficiently utilize the native advantage of fuzzy concepts in ontology recognition and integrability.

This article designs a novel concept-cognitive learning method using memory mechanisms to enhance concept learning ability (motivation I) and reduce cognitive concept complexity (motivation II). Fig. 1 describes the block diagram of the proposed M-FCCL. As shown in Fig. 1, the memory-based concept-cognitive learning method includes three main stages: a novel fuzzy three-way concept for concept learning is established in Section 3.1, a concept recalling mechanism based on the fuzzy three-way concept is constructed in Section 3.2, and then a concept forgetting mechanism to reduce the cognitive concept complexity is proposed in Section 3.3. Moreover, this paper is arranged as follows: Section 2 briefly reviews some notions related to concept-cognitive learning. Section 3 details the memory-based concept-cognitive learning method in four parts: concept learning, concept forgetting, concept recalling, and our procedure. Furthermore, the experimental results and analysis are presented in Section 4. Finally, some concluding remarks and future work are summarized in Section 5.

2. Preliminaries

This section briefly reviews some notions related to fuzzy formal context, fuzzy three-way concept, and granule concept learning. A detailed discussion of them can be found in the corresponding papers [33, 38–40]. Before starting with this section, it is necessary to claim that the current article investigates the concept-cognitive learning model in a fuzzy-classical decision formal context not a fuzzy–fuzzy or classical-fuzzy decision formal context, that is, conditional attributes are fuzzy data, and decision attributes are discrete data.

2.1. Fuzzy formal context

The fuzzy formal context, a database with fuzzy relations, mainly focuses on introducing the fuzzy member degree of the fuzzy set into the description of the binary relations between object and attribute.

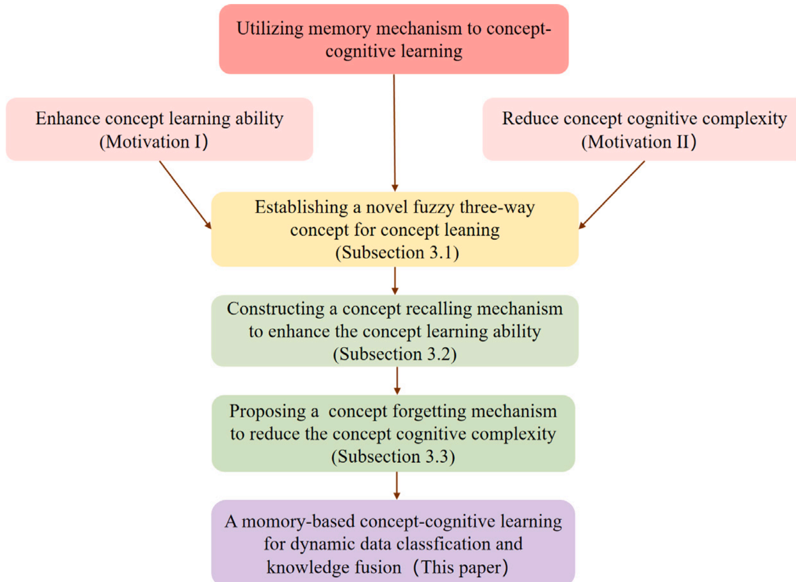


Fig. 1. Block diagram of the proposed M-FCCL.

One can effectively solve the limitation that classical formal concept analysis is only suitable for discrete data.

Let Ω be a universe and a fuzzy set $\widetilde{X}$ on Ω has the following form:

$$\widetilde{X} = \{\langle x, \mu_{\widetilde{X}}(x)\rangle | x \in \Omega\}, \tag{1}$$

where $\mu_{\widetilde{X}} : \Omega \to [0,1]$, $\mu_{\widetilde{X}}(x)$ is referred to as the membership degree of object x with respect to $\widetilde{X}$, and $\mu^c_{\widetilde{X}}(x) = 1 - \mu_{\widetilde{X}}(x)$ is the non-membership degree.

As is well known, a triplet $(\Omega, \Psi, \widetilde{I})$ can be represented as a fuzzy formal context, where $\Omega = \{x_1, x_2, \ldots, x_n\}$ and $\Psi = \{a_1, a_2, \ldots, a_m\}$ are, respectively, called the object set and the attribute set. Meanwhile, $\widetilde{I} = \{\langle (x,a), \mu(x,a)\rangle | (x,a) \in \Omega \times \Psi\}$ is a fuzzy relation, where $\mu : \Omega \times \Psi \to [0,1]$, $\mu(x,a)$ denotes the membership degree of object x with respect to a. For convenience, it is denoted by $\widetilde{I}(x,a) = \mu(x,a)$. For any $\widetilde{I}(x,a)$ and $\widetilde{I}(x',a)$, the $\widetilde{I}(x,a) \geqslant \widetilde{I}(x',a) \Leftrightarrow \mu(x,a) \geqslant \mu(x',a)$.

Definition 1. A triplet $(\Omega, \Psi, \widetilde{I})$ is a fuzzy formal context, for any $E \subseteq \Omega$, $T \subseteq \Psi$ and $\widetilde{T} \in \Gamma^{\Psi}$, the derivation operator $(\cdot)^*$ can be defined as follows:

$$E^*(a) = \bigwedge_{x \in E} \widetilde{I}(x,a), a \in \Psi, \tag{2}$$

$$\widetilde{T}^* = \{x \in \Omega | \forall a \in T, \widetilde{T}(a) \leqslant \widetilde{I}(x,a)\}, \tag{3}$$

where Γ^{Ψ} is the union of all fuzzy sets in Ψ.

Then, a pair $(E, \widetilde{T})$ is fuzzy concept if $E^* = \widetilde{T}$ and $\widetilde{T}^* = E$ hold, where E is the extent and $\widetilde{T}$ is the intent of the fuzzy concept $(E, \widetilde{T})$. Obviously, (E^{**}, E^*) and $(\widetilde{T}^*, \widetilde{T}^{**})$ are two fuzzy concepts. The fuzzy concept lattice $\widetilde{\mathcal{L}}(\Omega, \Psi, \widetilde{I})$ is the union of all fuzzy concept in $(\Omega, \Psi, \widetilde{I})$. For any $(E_1, \widetilde{T}_1), (E_2, \widetilde{T}_2) \in \widetilde{\mathcal{L}}(\Omega, \Psi, \widetilde{I})$, the ordered by $(E_1, \widetilde{T}_1) \leqslant (E_2, \widetilde{T}_2) \Leftrightarrow E_1 \subseteq E_2 \Leftrightarrow \widetilde{T}_2 \subseteq \widetilde{T}_1$. For any $E_1, E_2, E \subseteq \Omega, \widetilde{T}_1, \widetilde{T}_2 \subseteq \widetilde{T}$, the following properties hold:

(1) $E_1 \subseteq E_2 \Rightarrow E_2^* \subseteq E_1^*, \widetilde{T}_1 \subseteq \widetilde{T}_2 \Rightarrow \widetilde{T}_2^* \subseteq \widetilde{T}_1^*$;
(2) $E \subseteq E^{**}, \widetilde{T} \subseteq \widetilde{T}^{**}$;
(3) $E = E^{***}, \widetilde{T} = \widetilde{T}^{***}$;
(4) $E \subseteq \widetilde{T}^* \Leftrightarrow \widetilde{T} \subseteq E^*$;
(5) $(E_1 \cup E_2)^* = E_1^* \cap E_2^*, (\widetilde{T}_1 \cup \widetilde{T}_2)^* = \widetilde{T}_1^* \cap \widetilde{T}_2^*$;
(6) $(E_1 \cap E_2)^* \supseteq E_1^* \cup E_2^*, (\widetilde{T}_1 \cap \widetilde{T}_2)^* \supseteq \widetilde{T}_1^* \cup \widetilde{T}_2^*$.

Generally, the fuzzy formal context with decision labels is more suitable for classification scenarios. Hence, it is necessary to claim that this paper discusses the CCL model in a fuzzy decision formal context. A fuzzy decision formal context is a quintuple $(\Omega, \Psi, \widetilde{I}, D, J)$, where $(\Omega, \Psi, \widetilde{I})$ is a fuzzy formal context and (Ω, D, J) is a classical formal context. A classical fuzzy formal context (Ω, D, J) is a triple, where $\Omega / D = \{D_1, D_2, \ldots, D_l\}$ are referred to as a decision division based on decision class D, $\Omega = D_1 \cup D_2 \cup, \ldots, \cup D_l$, $J : \Omega \times D \to \{D_1, D_2, \ldots, D_l\}$ is a binary relation.

In order to more clearly explain the fuzzy formal context and fuzzy decision formal context in the current article, two examples are given in Fig. 2 to show the differences between them, respectively. Specifically, Fig. 2(a), including four objects and eight conditional attributes, is an example of fuzzy formal context $(\Omega, \Psi, \widetilde{I})$, where $\widetilde{I}$ represents the fuzzy relation between Ω and $\Psi = \{a, b, c, \ldots, h\}$. In addition, Fig. 2(b), including four objects, five conditional attributes, and three decision classes induced by decision attribute D, is an example of fuzzy decision formal context $(\Omega, \Psi, \widetilde{I}, D, J)$, where $\widetilde{I}$ represents the fuzzy relation between Ω and $\Psi = \{a, b, c, d, e\}$ and J is the binary relation between objects set and decision attribute D. Any dataset with decision labels can be processed to a fuzzy decision formal context for constructing concept subspace.

2.2. Fuzzy three-way concept analysis

Let $\widetilde{I}^- = \{\langle (x,a), 1 - \mu(x,a)\rangle | (x,a) \in \Omega \times \Psi\}$ be the complement of $\widetilde{I}$, where $1 - \mu(x,a)$ reflects the non-membership degree of object x with respect to a. For conveniencel, it is denoted by $\widetilde{I}^-(x,a) = 1 - \mu(x,a)$.

Definition 2. A triplet $(\Omega, \Psi, \widetilde{I})$ is a fuzzy formal context. For any $X \subseteq \Omega$, $T \subseteq \Psi$ and $\widetilde{T} \in \Gamma^{\Psi}$, the positive cognitive operators $\widetilde{\mathcal{L}} : 2^{\Omega} \to \Gamma^{\Psi}$ and $\mathcal{H} : \Gamma^{\Psi} \to 2^{\Omega}$ of $(\Omega, \Psi, \widetilde{I})$ are defined by:

$$\widetilde{\mathcal{L}}(X)(a) = \bigwedge_{x \in X} \widetilde{I}(x,a), a \in \Psi, \tag{4}$$

$$\mathcal{H}(\widetilde{T}) = \{x \in \Omega | \forall a \in T, \widetilde{T}(a) \leqslant \widetilde{I}(x,a)\}. \tag{5}$$

Similarly, the negative cognitive operators $\widetilde{\mathcal{L}}^- : 2^{\Omega} \to \Gamma^{\Psi}$ and $\mathcal{H}^- : \Gamma^{\Psi} \to 2^{\Omega}$ of $(\Omega, \Psi, \widetilde{I}^-)$ are defined by:

$$\widetilde{\mathcal{L}}^-(X)(a) = \bigwedge_{o \in X} \widetilde{I}^-(o,a), a \in \Psi, \tag{6}$$

$$\mathcal{H}^-(\widetilde{T}) = \{o \in \Omega | \forall a \in T, \widetilde{T}(a) \leqslant \widetilde{I}^-(o,a)\}, \tag{7}$$

Ω/Ψ	a	b	c	d	e	f	g	h
x_1	0.3	0.2	0.6	0.7	0.3	0.1	0.9	0.2
x_2	0.6	0.5	0.4	0.3	0.2	0.3	0.5	0.1
x_3	0.2	0.4	0.5	0.2	0.6	0.4	0.6	0.5
x_4	0.8	0.3	0.9	0.8	0.8	0.3	0.2	0.6
x_5	0.3	0.5	0.2	0.7	0.3	0.5	0.1	0.6

(a) Fuzzy formal context

Ω/Ψ	a	b	c	d	e	D_1	D_2	D_3
x_1	0.3	0.2	0.6	0.7	0.3	1	0	0
x_2	0.6	0.5	0.4	0.3	0.2	0	1	0
x_3	0.2	0.4	0.5	0.2	0.6	0	0	1
x_4	0.8	0.3	0.9	0.8	0.8	1	0	0
x_5	0.3	0.5	0.2	0.7	0.3	0	1	0

(b) Fuzzy decision formal context

Fig. 2. Illustration of two different forms of fuzzy contexts. (a) fuzzy formal context and (b) fuzzy decision formal context.

where 2^{Ω} is the power set of Ω and $\widetilde{T}(a)$ is the membership degree of object with respect to a on $\widetilde{T}$.

The above definition shows that the positive and negative cognitive operators can characterize the cognitive process between objects and attributes in two ways. Moreover, $\widetilde{\mathcal{L}}(X)(a)$ and $\widetilde{\mathcal{L}}^{-}(X)(a)$ are referred to as the cognitive process from object to attribute, and $\mathcal{H}(\widetilde{T})$ and $\mathcal{H}^{-}(\widetilde{T})$ are referred to as the cognitive process from attribute to object, which means that one can learn object or attribute information from the given information.

Definition 3. A triplet $(\Omega, \Psi, \widetilde{I})$ is a fuzzy formal context. For any $X \subseteq \Omega$ and $\widetilde{T_1}, \widetilde{T_2} \in \Gamma^{\Psi}$, the fuzzy three-way concept cognitive operator $\widetilde{\mathcal{L}}^{\nabla} : 2^{\Omega} \to \Gamma^{\Psi} \times \Gamma^{\Psi}$ and $\mathcal{H}^{\nabla} : \Gamma^{\Psi} \times \Gamma^{\Psi} \to 2^{\Omega}$ are defined by:

$$\widetilde{\mathcal{L}}^{\nabla}(X) = (\widetilde{\mathcal{L}}(X), \widetilde{\mathcal{L}}^{-}(X)), \tag{8}$$

$$\mathcal{H}^{\nabla}(\widetilde{T_1}, \widetilde{T_2}) = \mathcal{H}(\widetilde{T_1}) \cap \mathcal{H}^{-}(\widetilde{T_2}). \tag{9}$$

Then, $(X, (\widetilde{T}_1, \widetilde{T}_2))$ is a fuzzy three-way concept if $\widetilde{\mathcal{L}}^{\nabla}(X) = (\widetilde{T}_1, \widetilde{T}_2)$, $\mathcal{H}^{\nabla}(\widetilde{T}_1, \widetilde{T}_2) = X$, and X and $(\widetilde{T}_1, \widetilde{T}_2)$ are, respectively, known as the extent and intent of $(X, (\widetilde{T}_1, \widetilde{T}_2))$. Moreover, $(X, (\widetilde{T}_1, \widetilde{T}_2))$ is the subconcept of $(X', (\widetilde{T}'_1, \widetilde{T}'_2))$ if $X \subseteq X'$ or $(\widetilde{T}'_1, \widetilde{T}'_2) \geq (\widetilde{T}_1, \widetilde{T}_2)$, denoted by $(X, (\widetilde{T}_1, \widetilde{T}_2)) \leq (X', (\widetilde{T}'_1, \widetilde{T}'_2))$.

Furthermore, to describe both positive and negative information simultaneously, the positive and negative cognitive operators are combined to form a unique cognitive operator called the three-way operator, and the corresponding fuzzy concept is the fuzzy three-way concept. More discussion of the fuzzy three-way concept and the fuzzy three-way operator can be found in Ref. [33].

Definition 4. Let $(\Omega, \Psi, \widetilde{I})$ be a fuzzy formal context. For any $X \subseteq \Omega$ and $\widetilde{T}_1, \widetilde{T}_2 \in \Gamma^{\Psi}$, $(\mathcal{H}^{\nabla}\widetilde{\mathcal{L}}^{\nabla}(X), \widetilde{\mathcal{L}}^{\nabla}(X))$ and $(\mathcal{H}^{\nabla}(\widetilde{T}_1, \widetilde{T}_2), \widetilde{\mathcal{L}}^{\nabla}\mathcal{H}^{\nabla}(\widetilde{T}_1, \widetilde{T}_2))$ are called a pair of fuzzy three-way concepts.

According to Definition 4, a pair of fuzzy three-way concepts can be learned, where $(\mathcal{H}^{\nabla}\widetilde{\mathcal{L}}^{\nabla}(X), \widetilde{\mathcal{L}}^{\nabla}(X))$ and $(\mathcal{H}^{\nabla}(\widetilde{T}_1, \widetilde{T}_2), \widetilde{\mathcal{L}}^{\nabla}\mathcal{H}^{\nabla}(\widetilde{T}_1, \widetilde{T}_2))$ are the object-oriented fuzzy three-way concept and the attribute-oriented fuzzy three-way concept. Nevertheless, learning the attribute-oriented fuzzy three-way concept in a fuzzy context is sometimes immensely challenging without giving the initial clues $\widetilde{T}_1, \widetilde{T}_2 \in \Gamma^{\Psi}$. Hence, the object-oriented fuzzy three-way concept is usually utilized in concept learning to solve a particular problem, such as classification or recognition.

3. Memory-based concept-cognitive learning method

Note that forgetting and recalling are two crucial cognitive mechanisms of human memory, especially in learning new knowledge. One can be used to forget some irrelevant or insignificant knowledge, and another can help humans spontaneously recall some relevant or similar knowledge. However, the current CCL method easily ignores the cognitive properties of the cognitive subject (e.g., forgetting and recalling), and then resulting in some new issues in the cognitive system, such as the cumulative cognition of concepts and pseudo-concept learning with poor interpretability leading to poor performance.

Inspired by the memory mechanism, this section integrates humans' memory mechanisms into CCL and studies a novel concept-cognitive learning model via forgetting unnecessary concepts and recalling necessary concepts to enhance the performance of a cognitive system. In this model, the fuzzy concept is regarded as the knowledge acquired via the cognitive system, and the concept space of the system is used to store knowledge. Consequently, once new data arrives, the most similar fuzzy concept from the concept space is recalled for knowledge fusion, called concept recall. Moreover, the fuzzy concept in the concept space is evaluated according to specific rules, and the fuzzy concepts that do not meet the evaluation criteria are forgotten in the subsequent concept recognition process, called concept forgetting.

3.1. Concept learning for fuzzy three-way concept

This subsection introduces a cosine similarity degree to construct a similarity granule for concept learning. One of the primary motivations is that objects in similar granules tend to influence each other, especially in fuzzy decision formal contexts.

Definition 5. Let $(\Omega, \Psi, \widetilde{I}, D, J)$ be a fuzzy decision formal context, for any $x_1, x_2 \in \Omega$ and $a \in \Psi$, the cosine similarity degree between x_1 and x_2 as follows:

$$cs(x_1, x_2) = \frac{\sum_{a_i \in \Psi} \widetilde{I}(x_1, a_i) \cdot \widetilde{I}(x_2, a_i)}{\sqrt{\sum_{a_i \in \Psi} \widetilde{I}(x_1, a_i)^2} \cdot \sqrt{\sum_{a_i \in \Psi} \widetilde{I}(x_2, a_i)^2}}, \tag{10}$$

where $\widetilde{I}(x, a_i)$ is the membership degree of object x with respect to a_i.

Definition 5 can be utilized to construct a similarity granule, where $cs(x_1, x_2) \in [0, 1]$ and the valve of $cs(x_1, x_2)$ reflects the similarity degree between x_1 and x_2.

Definition 6. Let $(\Omega, \Psi, \widetilde{I}, D, J)$ be a fuzzy decision formal context, for any $x \in D_i$ and $a \in \Psi$, the cosine similarity granule of object x can be defined as follows:

$$CS(x) = \{x' \in D_k | cs(x, x') \geq \delta\}, \tag{11}$$

where δ is a threshold and $cs(x, x')$ is the cosine similarity degree between the object x and x'.

Generally speaking, learning the fuzzy concept in a fuzzy context is an effective way to deal with uncertainty and vagueness in fuzzy attributes, as a fuzzy membership relation between object and attribute constructs it [41,42]. However, constructing a classical fuzzy concept mainly considers the positive element of the commonality of the objects, ignoring negative information about the attributes, which means it describes information often incomplete. Hence, in this article, the classical fuzzy concept with the fuzzy three-way concept is replaced to represent the concept.

According to Definitions 5–6, a cosine similarity granule is obtained to describe similar objects in Ω. The threshold δ is vital in the subsequent concept-learning process. And then, a notion of the fuzzy three-way concept based on a cosine similarity granule is as follows.

Theorem 1. *Let $(\Omega, \Psi, \widetilde{I}, D, J)$ be a fuzzy decision formal context and $\Omega/D = \{D_1, D_2, \ldots, D_l\}$ be a decision division. For any $CS(x) \subseteq D_k$, $(\mathcal{H}\widetilde{\mathcal{L}}(CS(x)) \cap \mathcal{H}^{-}\widetilde{\mathcal{L}}^{-}(CS(x)), (\widetilde{\mathcal{L}}(CS(x)), \widetilde{\mathcal{L}}^{-}(CS(x))))$ is a fuzzy three-way concept.*

Proof. To prove this theorem, one can divide it into two steps:
(1) prove $\mathcal{H}^{\nabla}(\widetilde{\mathcal{L}}(CS(x)), \widetilde{\mathcal{L}}^{-}(CS(x))) = \mathcal{H}\widetilde{\mathcal{L}}(CS(x)) \cap \mathcal{H}^{-}\widetilde{\mathcal{L}}^{-}(CS(x))$ is valid;
(2) prove $\widetilde{\mathcal{L}}^{\nabla}(\mathcal{H}\widetilde{\mathcal{L}}(CS(x)) \cap \mathcal{H}^{-}\widetilde{\mathcal{L}}^{-}(CS(x))) = (\widetilde{\mathcal{L}}(CS(x)), \widetilde{\mathcal{L}}^{-}(CS(x)))$ is valid.

- For (1), it is immediate from Definition 3.
- For (2), according to Definition 2, for any $a \in \widetilde{T}$, then

$$\begin{aligned}\mathcal{H}\widetilde{\mathcal{L}}(CS(x)) \cap \mathcal{H}^{-}\widetilde{\mathcal{L}}^{-}(CS(x)) &= \{y \in D_k | \widetilde{I}(y,a) \geq \bigwedge_{o \in CS(x)} \widetilde{I}(o,a)\} \\ &\quad \cap \{y \in D_k | \widetilde{I}^{-}(y,a) \geq \bigwedge_{o \in CS(x)} \widetilde{I}^{-}(o,a)\} \\ &= \{y \in D_k | \widetilde{I}(y,a) \geq \bigwedge_{o \in CS(x)} \widetilde{I}(o,a)\} \\ &\quad \cap \{y \in D_k | 1 - \widetilde{I}^{-}(y,a) \leq 1 \\ &\quad - \bigwedge_{o \in CS(x)} \widetilde{I}^{-}(o,a)\} \\ &= \{y \in D_k | \widetilde{I}(y,a) \geq \bigwedge_{o \in CS(x)} \widetilde{I}(o,a)\} \\ &\quad \cap \{y \in D_k | \widetilde{I}(y,a) \leq \bigvee_{o \in CS(x)} \widetilde{I}(o,a)\} \\ &= \{y \in D_k | \bigwedge_{o \in CS(x)} \widetilde{I}(o,a) \leq \widetilde{I}(y,a) \\ &\quad \leq \bigvee_{o \in CS(x)} \widetilde{I}(o,a)\}.\end{aligned}$$

Therefore, for any $a \in \widetilde{T}$, the following relation holds

$$\widetilde{\mathcal{L}}(\mathcal{H}\widetilde{\mathcal{L}}(CS(x))) \cap \mathcal{H}^{-}\widetilde{\mathcal{L}}^{-}(CS(x))(a) = \bigwedge_{o \in CS(x)} \widetilde{I}(o,a) = \widetilde{\mathcal{L}}(CS(x))(a).$$

At the same time, the following equation holds

$$\begin{aligned}\mathcal{H}\widetilde{\mathcal{L}}(CS(x)) \cap \mathcal{H}^{-}\widetilde{\mathcal{L}}^{-}(CS(x)) &= \{y \in \Omega_i | \bigwedge_{o \in CS(x)} \widetilde{I}^{-}(o,a) \\ &\quad \leq \widetilde{I}^{-}(y,a) \leq \bigvee_{o \in CS(x)} \widetilde{I}^{-}(o,a),\end{aligned}$$

$$\widetilde{\mathcal{L}}^{-}(\mathcal{H}\widetilde{\mathcal{L}}(CS(x)) \cap \mathcal{H}^{-}\widetilde{\mathcal{L}}^{-}(CS(x)))(a) = \bigwedge_{o \in CS(x)} \widetilde{I}^{-}(o,a) = \widetilde{\mathcal{L}}^{-}(CS(x))(a).$$

Hence, $\widetilde{\mathcal{L}}^{\nabla}(\mathcal{H}\widetilde{\mathcal{L}}(CS(x)) \cap \mathcal{H}^{-}\widetilde{\mathcal{L}}^{-}(CS(x))) = (\widetilde{\mathcal{L}}(CS(x)), \widetilde{\mathcal{L}}^{-}(CS(x)))$.

By combining (1) and (2), this theorem is proven.

Definition 7. Let $(\Omega, \Psi, \widetilde{I}, D, J)$ be a fuzzy decision formal context and $\Omega/D = \{D_1, D_2, \ldots, D_l\}$ be a decision division. For any $x \in D_k$, the object-oriented fuzzy three-way concept subspace $\widetilde{S}^{D_k}$ about D_k can be defined as follows:

$$\widetilde{S}^{D_k} = \{(\mathcal{H}\widetilde{\mathcal{L}}(CS(x)) \cap \mathcal{H}^{-}\widetilde{\mathcal{L}}^{-}(CS(x)), (\widetilde{\mathcal{L}}(CS(x)), \widetilde{\mathcal{L}}^{-}(CS(x)))) | x \in D_k\}. \tag{12}$$

Algorithm 1 Concept learning for fuzzy three-way concept

Input: A formal context $(\Omega, \Psi, \widetilde{I}, D, J)$, parameter δ.
Output: Fuzzy three-way concept space $\widetilde{S}$;
Initial $\widetilde{S} = \emptyset$;
for $D_k \subseteq \Omega/D$ **do**
  Initial $\widetilde{S}^{D_k} = \emptyset$;
  for all $x \in D_k$ **do**
    for $y \in D_k$ **do**
      Compute the cosine similarity degree $cs(x,y)$ according to Definition 5
    end for
    Get the cosine similarity granule $CS(x)$ according to Definition 6;
    Learn the fuzzy three-way concept $(\mathcal{H}\widetilde{\mathcal{L}}(CS(x)) \cap \mathcal{H}^{-}\widetilde{\mathcal{L}}^{-}(CS(x)), (\widetilde{\mathcal{L}}(CS(x)), \widetilde{\mathcal{L}}^{-}(CS(x))))$;
    $\widetilde{S}^{D_k} \leftarrow (\mathcal{H}\widetilde{\mathcal{L}}(CS(x)) \cap \mathcal{H}^{-}\widetilde{\mathcal{L}}^{-}(CS(x)), (\widetilde{\mathcal{L}}(CS(x)), \widetilde{\mathcal{L}}^{-}(CS(x))))$;
  end for
end for
$\widetilde{S} \leftarrow \widetilde{S}^{D_k}$
return $\widetilde{S}$.

Definition 7 shows how an object-oriented fuzzy three-way concept subspace is constructed. According to this definition, the fuzzy three-way concept space $\widetilde{S} = \{\widetilde{S}^{D_1}, \widetilde{S}^{D_2}, \ldots, \widetilde{S}^{D_l}\}$ is produced. The concept learning process for the fuzzy three-way concept is shown in Algorithm 1, and its time complexity is $O(|\Omega|^2|\Psi|)$.

3.2. Cognitive mechanism based on concept recalling

Section 3.1 discussed the concept learning method for fuzzy three-way concept via the cosine similarity granule and two pairs of cognitive operators $\mathcal{H}$, $\widetilde{\mathcal{L}}$, and $\mathcal{H}^{-}$, $\widetilde{\mathcal{L}}^{-}$. Then, this subsection will propose a cognitive mechanism based on concept recalling to concern how the object-oriented fuzzy three-way concept space $\widetilde{S}^{D_k}$ are timely updated as time goes on, which is not only an incremental concept-cognitive learning system but also good at knowledge fusion in concept space.

Let $(\Omega, \Psi, \widetilde{I}, D, J)$ be a fuzzy decision formal context, $\Omega/D = \{D_1, D_2, \ldots, D_l\}$. Then, the $\Omega_i^D = \{\Omega_i^{D_1}, \Omega_i^{D_2}, \ldots, \Omega_i^{D_l}\}$ is referred to as the object set under the ith cognitive state. For brevity, for any $D_k \subseteq D$, the $\Omega_1^{D_k}, \Omega_2^{D_k}, \ldots, \Omega_s^{D_k}$ with $\Omega_1^{D_k} \subseteq \Omega_2^{D_k} \subseteq \ldots \subseteq \Omega_s^{D_k}$ are denoted by $\{\Omega_t^{D_k}\}\uparrow$.

Definition 8. Let $\Omega_{i-1}^{D_k}$ and $\Omega_i^{D_k}$ be object sets of $\{\Omega_t^{D_k}\}\uparrow$, where $\{\Omega_t^{D_k}\}\uparrow$ is a nondecreasing sequence subset of Ω^{D_k}, that is, $\Omega_1^{D_k} \subseteq \Omega_2^{D_k} \subseteq \cdots \subseteq \Omega_s^{D_k}$. Denote by $\Delta\Omega_{i-1}^{D_k} = \Omega_i^{D_k} - \Omega_{i-1}^{D_k}$. Suppose

(1) $\widetilde{\mathcal{L}}_{i-1} : 2^{\Omega_{i-1}^{D_k}} \to \Gamma^{\Psi}$, $\quad \mathcal{H}_{i-1} : \Gamma^{\Psi} \to 2^{\Omega_{i-1}^{D_k}}$,
(2) $\widetilde{\mathcal{L}}_{\Delta\Omega_{i-1}^{D_k}} : 2^{\Delta\Omega_{i-1}^{D_k}} \to \Gamma^{\Psi}$, $\quad \mathcal{H}_{\Delta\Omega_{i-1}^{D_k}} : \Gamma^{\Psi} \to 2^{\Delta\Omega_{i-1}^{D_k}}$,
(3) $\widetilde{\mathcal{L}}_i : 2^{\Omega_i^{D_k}} \to \Gamma^{\Psi}$, $\quad \mathcal{H}_i : \Gamma^{\Psi} \to 2^{\Omega_i^{D_k}}$;
(4) $\widetilde{\mathcal{L}}^{-}_{i-1} : 2^{\Omega_{i-1}^{D_k}} \to \Gamma^{\Psi}$, $\quad \mathcal{H}^{-}_{i-1} : \Gamma^{\Psi} \to 2^{\Omega_{i-1}^{D_k}}$,
(5) $\widetilde{\mathcal{L}}^{-}_{\Delta\Omega_{i-1}^{D_k}} : 2^{\Delta\Omega_{i-1}^{D_k}} \to \Gamma^{\Psi}$, $\quad \mathcal{H}^{-}_{\Delta\Omega_{i-1}^{D_k}} : \Gamma^{\Psi} \to 2^{\Delta\Omega_{i-1}^{D_k}}$,
(6) $\widetilde{\mathcal{L}}^{-}_i : 2^{\Omega_i^{D_k}} \to \Gamma^{\Psi}$, $\quad \mathcal{H}^{-}_i : \Gamma^{\Psi} \to 2^{\Omega_i^{D_k}}$;

are six pairs of cognitive operators satisfying the following properties:

$$\widetilde{\mathcal{L}}_i(CS(x)) = \begin{cases} \widetilde{\mathcal{L}}_{i-1}(CS(x)) \cap \widetilde{\mathcal{L}}_{\Delta\Omega_{i-1}^{D_k}}(y), & if \ \ y \in CS(x) \\ \widetilde{\mathcal{L}}_{\Delta\Omega_{i-1}^{D_k}}(CS(x)), & otherwise \end{cases} \tag{13}$$

$$\widetilde{\mathcal{L}}^{-}_i(CS(x)) = \begin{cases} \widetilde{\mathcal{L}}^{-}_{i-1}(CS(x)) \cap \widetilde{\mathcal{L}}^{-}_{\Delta\Omega_{i-1}^{D_k}}(y), & if \ \ y \in CS(x) \\ \widetilde{\mathcal{L}}^{-}_{\Delta\Omega_{i-1}^{D_k}}(CS(x)), & otherwise \end{cases} \tag{14}$$

then $\widetilde{\mathcal{L}}_i$ and $\widetilde{\mathcal{L}}_i^-$ are extended cognitive operators of $\widetilde{\mathcal{L}}_{i-1}$ and $\widetilde{\mathcal{L}}_{i-1}^-$ with the update information $\Delta\Omega_{i-1}^{D_k}$.

Definition 8 describes the knowledge fusion process with updated information by defining six positive and negative cognitive operator pairs. Note that concept-cognitive learning was often considered incremental due to the whole being something else than the sum of its part [22,25,43]. A novel cognitive mechanism based on concept recalling according to Definition 8 is discussed in what follows.

Theorem 2. *Let $(\Omega, \Psi, \widetilde{I}, D, J)$ be a fuzzy decision formal context and $\Omega_i^{D_k}$ be a object set about D_k under ith cognitive state. For any new object $y \in \Omega_i^{D_k}$ (i.e $y \notin \Omega_{i-1}$), then the following statements hold:*

(1) for any $CS(x) \subseteq \Omega_i^{D_k}$, if $y \in CS(x)$, then

$$\begin{aligned}&(\mathcal{H}_i\widetilde{\mathcal{L}}_i(CS(x)) \cap \mathcal{H}_i^-\widetilde{\mathcal{L}}_i^-(CS(x)), (\widetilde{\mathcal{L}}_i(CS(x)), \widetilde{\mathcal{L}}_i^-(CS(x))))\\ &= (\mathcal{H}_i(\widetilde{\mathcal{L}}_{i-1}(CS(x)) \cap \widetilde{\mathcal{L}}_{\Delta\Omega_{i-1}^{D_k}}(y)) \cap \mathcal{H}_i^-(\widetilde{\mathcal{L}}_{i-1}^-(CS(x)) \cap \widetilde{\mathcal{L}}^-_{\Delta\Omega_{i-1}^{D_k}}(y)),\\ &\quad(\widetilde{\mathcal{L}}_{i-1}(CS(x)) \cap \widetilde{\mathcal{L}}_{\Delta\Omega_{i-1}^{D_k}}(y), \widetilde{\mathcal{L}}_{i-1}^-(CS(x))\\ &\quad\cap \widetilde{\mathcal{L}}^-_{\Delta\Omega_{i-1}^{D_k}}(y))).\end{aligned} \tag{15}$$

(2) for any $CS(x) \subseteq \Omega_i^{D_k}$, if $y \notin CS(x)$, then

$$\begin{aligned}&(\mathcal{H}_i\widetilde{\mathcal{L}}_i(CS(x)) \cap \mathcal{H}_i^-\widetilde{\mathcal{L}}_i^-(CS(x)), (\widetilde{\mathcal{L}}_i(CS(x)), \widetilde{\mathcal{L}}_i^-(CS(x))))\\ &= (\mathcal{H}_{i-1}\widetilde{\mathcal{L}}_{i-1}(CS(x)) \cap \mathcal{H}_{i-1}^-\widetilde{\mathcal{L}}_{i-1}^-(CS(x)), (\widetilde{\mathcal{L}}_{i-1}(CS(x)), \widetilde{\mathcal{L}}_{i-1}^-(CS(x)))).\end{aligned} \tag{16}$$

Proof. The proof is immediate from Definition 8.

Definition 9. Let $(\Omega, \Psi, \widetilde{I}, D, J)$ be a fuzzy decision formal context and $\Omega_i^{D_k}$ be a object set about D_k under ith cognitive state. For any object $x \in \Omega_i^{D_k}$, the object-oriented fuzzy three-way concept space about D_k under ith cognitive state can be defined as follows:

$$\widetilde{S}_i^{D_k} = \{(\mathcal{H}_i\widetilde{\mathcal{L}}_i(CS(x)) \cap \mathcal{H}_i^-\widetilde{\mathcal{L}}_i^-(CS(x)), (\widetilde{\mathcal{L}}_i(CS(x)), \widetilde{\mathcal{L}}_i^-(CS(x)))) | x \in \Omega_i^{D_k}\}. \tag{17}$$

According to the above definition, one can learn the fuzzy three-way concept space under ith cognitive states (i.e., $\widetilde{S}_i = \{\widetilde{S}_i^{D_1}, \widetilde{S}_i^{D_2}, \dots, \widetilde{S}_i^{D_l}\}$) through the concept fusion mechanism based on concept recalling. The details are shown in Algorithm 2, and its time complexity is $O(|\Delta\Omega_{i-1}^{D_k}\| D_k \|\Psi|)$.

So far, a concept fusion mechanism has been built based on concept recalling and the object-oriented fuzzy three-way concept space under various cognitive states. Different fuzzy concepts influence each other in concept space with much repetitive and interfering information between them. Moreover, accumulating concepts will influence the cognitive system's performance over time. Hence, the following subsection will discuss concept recognition via concept forgetting for the proposed concept-cognitive learning system.

3.3. Concept recognition via concept forgetting

According to the procedure in Section 3.2, the cognitive mechanism based on concept recalling in M-FCCL can directly update the concept space under different cognitive states for new objects without retaining the whole model. Nevertheless, as mentioned above, the previous two subsections mainly introduced how to carry out concept learning and concept cognition. However, it still needs to offer an effective manner of concept recognition, identifying which concept space needs to update with new objects.

Algorithm 2 Cognitive mechanism based on concept recalling

Input: A formal context $(\Omega, \Psi, \widetilde{I}, D, J)$, fuzzy three-way concept space $\widetilde{S}_{i-1}^{D_k}$, the added object set $\Delta\Omega_{i-1}^{D_k}$.
Output: Fuzzy three-way concept space $\widetilde{S}_i^{D_k}$.

```
for all y ∈ ΔΩ_{i-1}^{D_k} do
    Get the fuzzy three-way concept (y, (L̃_{ΔΩ_{i-1}^{D_k}}(y), L̃^-_{ΔΩ_{i-1}^{D_k}}(y)));
    for all CS(x) ∈ D_k do
        if y ∈ CS(x) then
            L̃_i(CS(x)) = L̃_{i-1}(CS(x)) ∩ L̃_{ΔΩ_{i-1}^{D_k}}(y);
            L̃_i^-(CS(x)) = L̃_{i-1}^-(CS(x)) ∩ L̃^-_{ΔΩ_{i-1}^{D_k}}(y);
        else L̃_i(CS(x)) = L̃_{i-1}(CS(x));
            L̃_i^-(CS(x)) = L̃_{i-1}^-(CS(x));
        end if
        Learn the fuzzy three-way concept (H_i L̃_i(CS(x)) ∩ H_i^- L̃_i^-(CS(x)), (L̃_i(CS(x)), L̃_i^-(CS(x))));
        S̃_i^{D_k} ← (H_i L̃_i(CS(x)) ∩ H_i^- L̃_i^-(CS(x)), (L̃_i(CS(x)), L̃_i^-(CS(x))));
    end for
end for
return S̃_i^{D_k}.
```

Meanwhile, note that various concepts carry different information values and importance in fuzzy-based concept learning for concept recognition. In order to evolve suitable fuzzy ontologies for concept recognition, various pseudo-concept learning ways have been researched, such as fuzzy concept clustering [32], progressive fuzzy concept learning [33], and progressive weighted fuzzy concept [34]. Accordingly, a concept recognition strategy based on the concept forgetting mechanism is proposed in this subsection, which can identify objects effectively and avoid learning pseudo-concepts.

Definition 10. Let $(\Omega, \Psi, \widetilde{I}, D, J)$ be a fuzzy decision formal context. For $\widetilde{S}_i^{D_k} \in \widetilde{S}_i$, $\widetilde{\mathcal{G}}_i^{D_k}$ is an object-oriented fuzzy three-way concept space based on concept forgetting about $\widetilde{S}_i^{D_k}$ under ith cognitive state if the following statements hold:

(1) For any $(X_i, (\widetilde{T_1}, \widetilde{T_2})) \in \widetilde{\mathcal{G}}_i^{D_k}$, there does not exist a fuzzy three-way concept $(X_i', (\widetilde{T_1'}, \widetilde{T_2'})) \in \widetilde{\mathcal{G}}_i^{D_k}$ that is the sub-concept of $(X_i, (\widetilde{T_1}, \widetilde{T_2}))$;
(2) For any $(X_i, (\widetilde{T_1}, \widetilde{T_2})) \in \widetilde{S}_i^{D_k}$, there exist at least one fuzzy three-way concept $(X_i', (\widetilde{T_1'}, \widetilde{T_2'})) \in \widetilde{\mathcal{G}}_i^{D_k}$ that makes $X_i \subseteq X_i'$;
(3) For any $(X_i, (\widetilde{T_1}, \widetilde{T_2})) \in \widetilde{\mathcal{G}}_i^{D_k}$, the statement $(X_i, (\widetilde{T_1}, \widetilde{T_2})) \in \widetilde{S}_i^{D_k}$ is hold.

Theorem 3. *Let $\widetilde{\mathcal{G}}_i^{D_k}$ be an object-oriented fuzzy three-way concept space based on concept forgetting about $\widetilde{S}_i^{D_k}$ under ith cognitive state. The following property holds*

$$1 \le |\widetilde{\mathcal{G}}_i^{D_k}| \le |\widetilde{S}_i^{D_k}|. \tag{18}$$

where $|\cdot|$ denotes the cardinality.

Proof. To prove this theorem, one can divide it into three steps:

(1) For any $(X_i, (\widetilde{T_1}, \widetilde{T_2})) \in \widetilde{S}_i^{D_k}$, if there exist a fuzzy three-way concept $(X_i', (\widetilde{T_1'}, \widetilde{T_2'})) \in \widetilde{S}_i^{D_k}$ that makes $X_i \subseteq X_i'$, from item (1) of Definition 10, then $|\widetilde{\mathcal{G}}_i^{D_k}| = 1$;
(2) For any $(X_i, (\widetilde{T_1}, \widetilde{T_2})) \in \widetilde{S}_i^{D_k}$, if there exist a fuzzy three-way concept $(X_i', (\widetilde{T_1'}, \widetilde{T_2'})) \in \widetilde{S}_i^{D_k}$ that makes $X_i \nsubseteq X_i'$, from item (2) of Definition 10, then $|\widetilde{\mathcal{G}}_i^{D_k}| > 1$

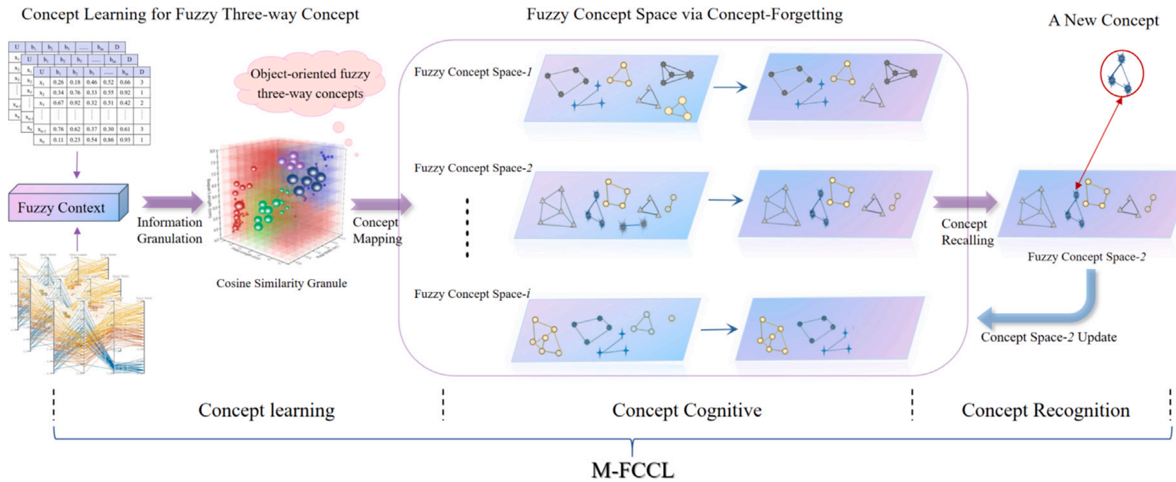


Fig. 3. Illustration of the overall procedure for M-FCCL, taking the update of fuzzy three-way concept space-2 as an example.

(3) For $(X_i, (\widetilde{T_1}, \widetilde{T_2})) \in \widetilde{S}_i^{D_k}$, if there does not exist a fuzzy three-way concept $(X_i', (\widetilde{T_1'}, \widetilde{T_2'})) \in \widetilde{S}_i^{D_k}$ that makes $X_i \subset X_i'$, from item (3) of Definition 10, then $|\widetilde{G}_i^{D_k}| \le |\widetilde{S}_i^{D_k}|$.

By combining (1), (2) and (3), this theorem is proven.

From a memory perspective, Definition 10 can completely characterize a forgetting process of concept: forgetting some unnecessary knowledge (i.e., sub-concept) and retaining some necessary knowledge in knowledge space(i.e., concept space). For brevity, $\widetilde{\mathcal{G}}_i^{D_k}$ is referred to as forgetting space of object-oriented fuzzy three-way concept about $\widetilde{S}_i^{D_k}$, then $\widetilde{\mathcal{G}}_i = \{\widetilde{\mathcal{G}}_i^{D_1}, \widetilde{\mathcal{G}}_i^{D_2}, \dots, \widetilde{\mathcal{G}}_i^{D_l}\}$.

According to the discussion of the concept recalling mechanism in Section 3.1, it only needs recalling some related knowledge (i.e., cosine similarity granule $CS(x)$) connected to the new information and updating the corresponding fuzzy three-way concepts. Furthermore, a concept recognition based on the concept forgetting method is defined as follows.

Definition 11. Let x be a new object and $(x, (\widetilde{\mathcal{L}}_{\Delta\Omega_{i-1}^{D_k}}(x), \widetilde{\mathcal{L}}^-_{\Delta\Omega_{i-1}^{D_k}}(x)))$ be a new fuzzy three-way concept, for any $(X_{h,k}, (\widetilde{T}_{h,k}, \widetilde{T}^-_{h,k})) \in \widetilde{\mathcal{G}}_i^{D_k}$ ($h \in \{1, 2, \dots, |\widetilde{\mathcal{G}}_i^{D_k}|\}$), the concept recognition degree between two concepts can be defined by:

$$\mathcal{R}_{h,D_k} = \sqrt{\|\widetilde{T}_{h,k} - \widetilde{\mathcal{L}}_{\Delta\Omega_{i-1}^{D_k}}(x)\|^2 + \|\widetilde{T}^-_{h,k} - \widetilde{\mathcal{L}}^-_{\Delta\Omega_{i-1}^{D_k}}(x)\|^2}. \tag{19}$$

where $||\cdot||^2$ denotes the 2-norm.

The smaller the value of $\mathcal{R}_{h,D_k}$, the stronger the relationship between the two concepts. Thus, for any added object x, according to Definition 11, one can compute the recognition degree between x and any concept in $\widetilde{\mathcal{G}}_i^{D_k}$. Additionally, one can compute the global minimum recognition degree between the added object x and $\widetilde{\mathcal{G}}_i$, denoted by $\mathcal{R}^*_{h,D_k} = argmin_{h,D_k} \mathcal{R}_{h,D_k}$.

Definition 12. Let x be a new object and $(x, (\widetilde{\mathcal{L}}_{\Delta\Omega_{i-1}^{D_k}}(x), \widetilde{\mathcal{L}}^-_{\Delta\Omega_{i-1}^{D_k}}(x)))$ be a new fuzzy three-way concept. Then, $(x, (\widetilde{\mathcal{L}}_{\Delta\Omega_{i-1}^{D_k}}(x), \widetilde{\mathcal{L}}^-_{\Delta\Omega_{i-1}^{D_k}}(x)))$ can be fused into $\widetilde{\mathcal{G}}_i^{D_k}$, if $\mathcal{R}^*_{h,D_k}$ is the global minimum recognition degree.

From Definitions 10–12, concept recognition via concept forgetting consists of three steps: (1) learning fuzzy three-way concept space based on concept forgetting; (2) computing concept recognition degree between different concepts; (3) getting the global minimum recognition. The details are shown in Algorithm 3, and its time complexity is $O(|\Omega|(|\Omega|^2 + |\Psi|))$.

Algorithm 3 Concept recognition via concept forgetting

Input: A formal context $(\Omega, \Psi, \widetilde{I}, D, J)$, fuzzy three-way concept space $\widetilde{S}_i = \{\widetilde{S}_i^{D_1}, \widetilde{S}_i^{D_2}, \dots, \widetilde{S}_i^{D_l}\}$, the added object y.
Output: $\mathcal{R}^*_{h,D_k}$ and D^*_k;
Get the fuzzy three-way concept space based on concept forgetting $\widetilde{\mathcal{G}}_i = \{\widetilde{\mathcal{G}}_i^{D_1}, \widetilde{\mathcal{G}}_i^{D_2}, \dots, \widetilde{\mathcal{G}}_i^{D_l}\}$ according to Definition 10;
Get the fuzzy three-way concept $(y, (\widetilde{\mathcal{L}}_{\Delta\Omega_{i-1}^{D_k}}(y), \widetilde{\mathcal{L}}^-_{\Delta\Omega_{i-1}^{D_k}}(y)))$;
for all $\widetilde{\mathcal{G}}_i^{D_k} \subseteq \widetilde{\mathcal{G}}_i$ **do**
  for all $(X_{h,k}, (\widetilde{T}_{h,k}, \widetilde{T}^-_{h,k})) \in \widetilde{\mathcal{G}}_i^{D_k}$ **do**
    Compute the recognitive degree $\mathcal{R}_{h,D_k}$ according to Definition 11;
    $\mathcal{R}^*_{h,D_k} \leftarrow min\{\mathcal{R}_{h,D_k}\}$;
    $D^*_k \leftarrow argmin_{D_k \in \Omega/D} R^*_{h,D_k}$;
  end for
end for
return $\mathcal{R}^*_{h,D_k}$ and D^*_k.

3.4. Overall procedure

In this subsection, Fig. 3 is utilized to show the overall procedure of M-FCCL (i.e., memory-based concept-cognitive learning), which includes three main stages: concept learning, concept cognition, and concept recognition. Take the update of the fuzzy three-way concept space-2 for brevity as an example. The concept learning stage involves learning the object-oriented fuzzy three-way concept via cosine similarity granule based on information granulation. According to the fuzzy three-way concept in the first stage, the fuzzy three-way concept space is constructed, and then the stage of concept cognition is to forget some unnecessary concepts in the fuzzy concept space.

In the stage of concept recognition, for any new object and its corresponding fuzzy three-way concept, the global minimum recognition degree is further used to compute the recognition degree between the new concept and the fuzzy three-way concept spaces. Then there is only a need to recall the related knowledge with the new concept in this fuzzy three-way concept space to fuse it. For any new concept, the above process can be repeated.

Combined with the theoretical analysis of M-FCCL in Sections 3.1–3.3, the memory-based concept-cognitive learning method details are

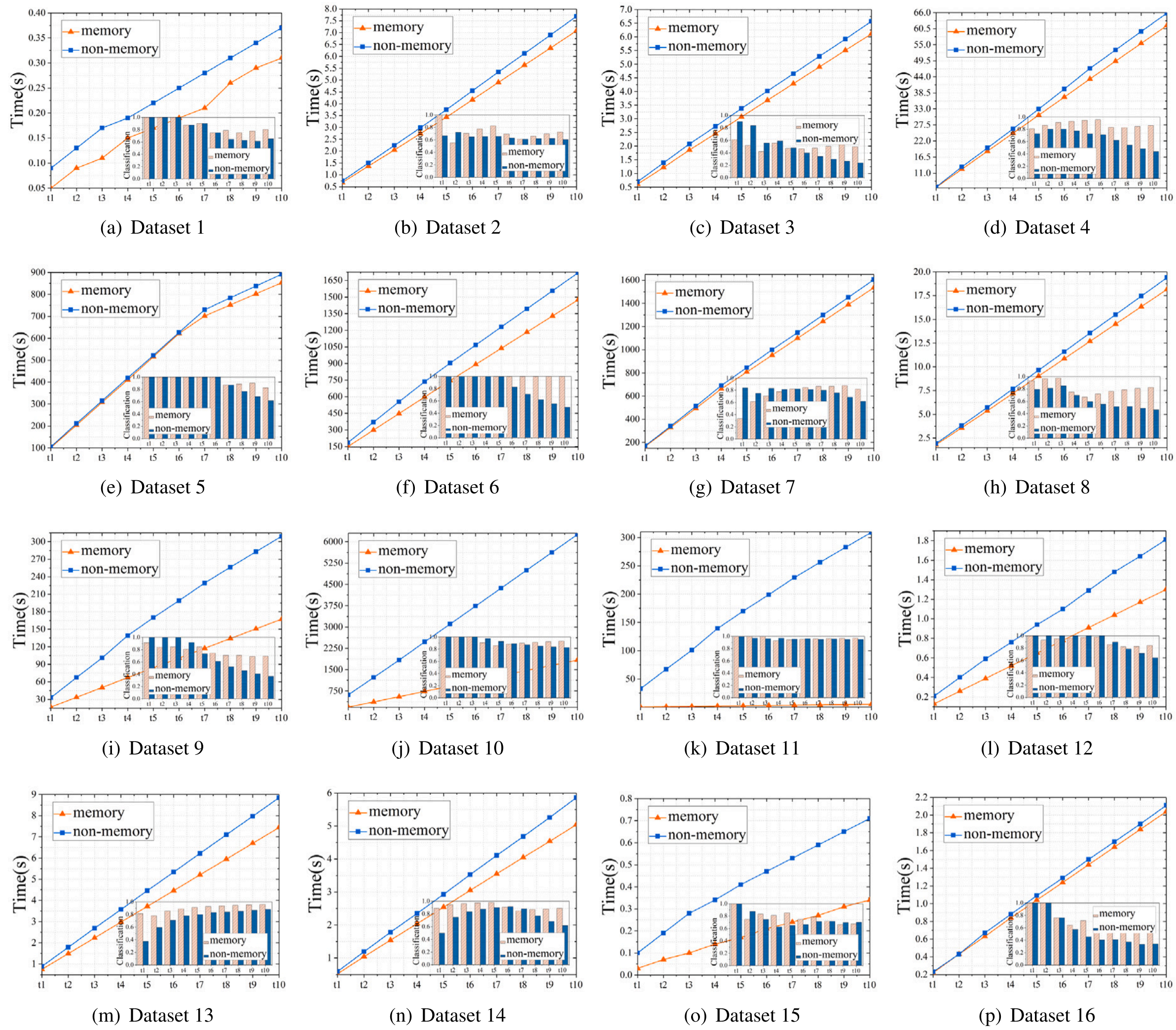


Fig. 4. The performance comparison of memory and non-memory mechanisms on 16 datasets.

Table 1
Detailed information of 16 selected datasets.

No.s	Datasets	Objects	Features	Classes
1	BreastTissue	106	10	6
2	Hill	606	101	2
3	Mice Protein Expression	1077	68	8
4	Cardiotocography	2126	23	3
5	Spambase	4597	57	2
6	Ring	7400	20	2
7	DryBean	13 611	17	6
8	South German Credit	1000	21	2
9	Abalone	4177	9	3
10	EGS	10 000	14	2
11	Wdbc	569	31	2
12	Heart	270	14	2
13	Climate Model Simulation Crashes	540	18	2
14	Balance	625	5	3
15	Urbanlandcover	168	148	9
16	Spectfheart	267	45	2

shown in Algorithms 1–3. Meanwhile, according to the introduction of the overall procedure of M-FCCL in this subsection, the process can be briefly described into four steps as follows.

- **Step 1** Given a fuzzy decision formal context $(\Omega, \Psi, \widetilde{I}, D, J)$. For any $x \in \Omega$ and $D_k \in \Omega/D$, get the cosine similarity granule $CS(x)$ and fuzzy three-way concept $(\mathcal{H}\widetilde{\mathcal{L}}(CS(x)) \cap \mathcal{H}^{-}\widetilde{\mathcal{L}}^{-}(CS(x)), (\widetilde{\mathcal{L}}(CS(x)), \widetilde{\mathcal{L}}^{-}(CS(x))))$, and then construct fuzzy three-way concept space $\widetilde{S} = \{\widetilde{S}^{D_1}, \widetilde{S}^{D_2}, \ldots, \widetilde{S}^{D_l}\}$.
- **Step 2** For any $\widetilde{S}^{D_k} \in \widetilde{S}$, construct object-oriented fuzzy three-way concept space based on concept forgetting $\widetilde{\mathcal{G}}^{D_k}$ according the method in Definition 10, and then get the fuzzy three-way concept space via knowledge fusion, i.e., $\widetilde{\mathcal{G}} = \{\widetilde{\mathcal{G}}^{D_1}, \widetilde{\mathcal{G}}^{D_2}, \ldots, \widetilde{\mathcal{G}}^{D_l}\}$.
- **Step 3** For a new fuzzy three-way concept $(x, (\widetilde{\mathcal{L}}_{\Delta\Omega^{D_k}}(x), \widetilde{\mathcal{L}}^{-}_{\Delta\Omega^{D_k}}(x)))$, the concept recognition degree $\mathcal{R}_{h,D_k}$ with each granular concept $(X_{h,k}, (\widetilde{T}_{h,k}, \widetilde{T}^{-}_{h,k})) \in \widetilde{\mathcal{G}}^{D_k}$ $(h \in \{1, 2, \ldots, |\widetilde{\mathcal{G}}^{D_k}|\})$, and then get the fuzzy three-way concept space and its corresponding class label according to $D_k{}^{*} = argmin_{D_k} \mathcal{R}^{*}_{h,D_k}$.
- **Step 4** For different cognitive states, the fuzzy three-way concept space $\widetilde{\mathcal{G}}_i = \{\widetilde{\mathcal{G}}_i^{D_1}, \widetilde{\mathcal{G}}_i^{D_2}, \ldots, \widetilde{\mathcal{G}}_i^{D_l}\}$ needs to be updated by six pairs of cognitive operators in Definition 8 and the cognitive mechanism in Theorem 2, and then repeat steps 2-3. The final fuzzy three-way concept space can be obtained using the recursive approach.

4. Experiments

In this section, the effectiveness of M-FCCL is validated for the performance of concept learning based on a fuzzy formal context. Specifically, M-FCCL with some popular classification methods is compared. All experimental setup of comparison methods is consistent with corresponding references. The experimental computing program on a public

Table 2
Accuracy comparison of two concept learning mechanisms on 16 datasets.

Dataset	Mechanisms	Parameter	t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8	t_9	t_{10}	Ave. ± SD
1	Memory	0.05	1.0000	1.0000	1.0000	0.8750	0.9000	0.7500	0.7857	0.7500	0.7778	0.8000	**0.8639 ± 0.1057**
	Non-memory	0.05	1.0000	1.0000	1.0000	0.8750	0.9000	0.7500	0.6429	0.6250	0.6111	0.6500	0.8054 ± 0.1671
2	Memory	0.30	1.0000	0.5556	0.7037	0.7778	0.8222	0.6944	0.6111	0.6597	0.6975	0.7278	**0.7250 ± 0.1229**
	Non-memory	0.50	0.6667	0.7222	0.6481	0.6528	0.6556	0.6296	0.6111	0.6181	0.6235	0.6056	0.6433 ± 0.0345
3	Memory	0.15	0.6129	0.5161	0.4194	0.5565	0.4710	0.4624	0.4747	0.5040	0.5520	0.4968	**0.5066 ± 0.0557**
	Non-memory	0.05	0.9032	0.8387	0.5591	0.5887	0.4774	0.3978	0.3410	0.2984	0.2652	0.2387	0.4908 ± 0.2329
4	Memory	0.50	0.8095	0.8651	0.9101	0.9325	0.9460	0.9550	0.8299	0.8274	0.8466	0.8619	**0.8784 ± 0.0533**
	Non-memory	0.35	0.7302	0.8016	0.8042	0.7738	0.7270	0.7116	0.6213	0.5437	0.4832	0.4349	0.6632 ± 0.1345
5	Memory	0.40	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.8717	0.8878	0.9002	0.8241	**0.9484 ± 0.0694**
	Non-memory	0.15	1.0000	1.0000	1.0000	1.0000	0.9985	0.9988	0.8717	0.7655	0.6813	0.6161	0.8932 ± 0.1514
6	Memory	0.05	0.9910	0.9955	0.9970	0.9977	0.9928	0.9932	0.9942	0.9949	0.9955	0.9959	**0.9948 ± 0.0020**
	Non-memory	0.05	1.0000	1.0000	1.0000	1.0000	0.9946	0.8288	0.7104	0.6216	0.5525	0.4973	0.8205 ± 0.2074
7	Memory	0.00	0.3309	0.6127	0.7026	0.7763	0.8206	0.8435	0.8648	0.8640	0.8739	0.8142	0.7504 ± 0.1689
	Non-memory	0.00	0.8382	0.7475	0.8301	0.8113	0.8211	0.8105	0.7997	0.7558	0.6863	0.6176	**0.7718 ± 0.0715**
8	Memory	0.50	0.9310	0.9655	0.9770	0.7500	0.6690	0.7184	0.7586	0.7845	0.8084	0.8241	**0.8187 ± 0.1061**
	Non-memory	0.15	0.7931	0.8103	0.8506	0.6983	0.5931	0.5517	0.5123	0.5172	0.4866	0.4621	0.6275 ± 0.1473
9	Memory	0.20	0.9120	0.8400	0.8453	0.8060	0.8448	0.7427	0.7063	0.7070	0.6889	0.6952	**0.7788 ± 0.0801**
	Non-memory	0.25	1.0000	1.0000	1.0000	0.9160	0.7328	0.6107	0.5234	0.4580	0.4071	0.3664	0.7014 ± 0.2609
10	Memory	0.25	1.0000	0.9983	0.9989	0.9064	0.8582	0.8818	0.8987	0.9114	0.9212	0.9291	**0.9304 ± 0.0514**
	Non-memory	0.50	1.0000	1.0000	1.0000	0.9749	0.9251	0.8907	0.8696	0.8512	0.8369	0.8278	0.9176 ± 0.0713
11	Memory	0.15	1.0000	1.0000	1.0000	0.9265	0.9412	0.9510	0.9580	0.9632	0.9673	0.9706	**0.9678 ± 0.0257**
	Non-memory	0.00	1.0000	0.9706	0.9608	0.9706	0.9529	0.9608	0.9496	0.9559	0.9477	0.9471	0.9616 ± 0.0160
12	Memory	0.50	1.0000	0.9286	0.9524	0.9643	0.9714	0.9762	0.8571	0.8214	0.8254	0.8429	**0.9140 ± 0.0695**
	Non-memory	0.15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.8980	0.7857	0.7143	0.6429	0.9041 ± 0.1388
13	Memory	0.40	0.8125	0.7812	0.8542	0.8906	0.9125	0.9271	0.9375	0.9453	0.9514	0.9563	**0.8969 ± 0.0615**
	Non-memory	0.00	0.3750	0.5938	0.7083	0.7812	0.8000	0.8333	0.8393	0.8516	0.8681	0.8812	0.7532 ± 0.1587
14	Memory	0.25	0.8889	0.9444	0.9630	0.9722	0.9778	0.9074	0.8492	0.8681	0.8765	0.8889	**0.9136 ± 0.0469**
	Non-memory	0.45	0.5000	0.7500	0.8333	0.8750	0.9000	0.9167	0.8810	0.7708	0.6852	0.6167	0.7729 ± 0.1373
15	Memory	0.15	1.0000	0.7500	0.8333	0.8125	0.8500	0.7500	0.7857	0.7188	0.6667	0.6750	**0.7842 ± 0.0980**
	Non-memory	0.00	1.0000	0.8750	0.7500	0.6250	0.6500	0.6667	0.7143	0.7188	0.6944	0.7000	0.7394 ± 0.1142
16	Memory	0.05	1.0000	1.0000	0.7619	0.6429	0.7143	0.7619	0.7755	0.6964	0.6190	0.6286	**0.7601 ± 0.1384**
	Non-memory	0.10	1.0000	1.0000	0.7619	0.5714	0.4571	0.4048	0.4082	0.3750	0.3333	0.3429	0.5655 ± 0.2621

computer with OS: Microsoft Win10; Processor: Intel(R) Core(TM) i7-6800K CPU @ 3.4 GHz×12; Memory: 62.7 GB; Programming language: Matlab.

To extensively verify the performance of different approaches, 16 datasets with different scales are randomly selected from UCI (see https://www.uci.edu/), and the detailed information about them is shown in Table 1. For convenience, 16 datasets are denoted Dataset 1–16. The source data have been fuzzified into values ranging from 0 to 1 using the following method in [44].

$$\widetilde{I}(x, a_j) = \frac{f(x, a_j) - min(f(a_j))}{max(f(a_j)) - min(f(a_j))}, \tag{20}$$

where $f(x, a_j)$ represents the value of object x in feature a_j, the $max(f(a_j))$ and $min(f(a_j))$ denote the maximum and minimum value of all objects in a_j.

To clearly show the performance of the proposed method in the dynamic concept learning process, 70% data of each dataset as training samples set is selected to construct the initial concept space. The remaining data is divided into ten chunks and added to the testing samples set at different times in dynamic concept learning. The ten different times are denoted $t_1, t_2, \ldots, t_{10}$ for convenience. Note that the cosine similarity degree in the proposed method is an important parameter influencing the construction of the cosine similarity granule. Hence, this experiment sets the parameter δ from 0 to 0.5 in steps of 0.05 to learn the cosine similarity granule. For a fair comparison, ten trials were conducted for each method on the same training and testing samples. The average accuracy(Ave.) and the standard deviation(SD) are reported in this section.

4.1. Evaluating the performance of memory mechanism

Based on the proposed method in Section 2, we can know that M-FCCL can naturally achieve dynamic concept learning. Hence, the concept classification accuracy and running time of two concept learning mechanisms (i.e., memory and non-memory mechanisms) on 16 datasets for different time nodes to demonstrate the significance of the memory mechanism in M-FCCL method. The more intuitive remarkable comparison is shown in Fig. 4. It can be found that the performance of the memory mechanism of M-FCCL is significantly better than the non-memory mechanism in most cases.

(1) Concept classification accuracy: This part mainly demonstrates the concept classification accuracy of two concept learning mechanisms. Subsequently, the details of average accuracy and the standard deviation of ten-time results on 16 datasets are shown in Table 2. As seen from this table, the memory mechanism performs well for dynamic data classification on all datasets as time goes on. Table 2 shows that the dynamic data classification performance of the memory mechanism is almost more significant than the non-memory mechanism, except for dataset 7. Consequently, the classification results of two memory mechanisms in this table show that the proposed method performs reasonably efficiently for the 16 selected datasets from two aspects: (a) the average overall accuracy and (b) dynamic classification accuracy as the times go. In addition, the concept-cognitive learning mechanism is similar to human cognitive processes, the concept-learning results can be stored, and the subsequent concept-learning does not need to start from scratch.

(2) Runing time: This part mainly verifies the running time of two concept learning mechanisms on different datasets. Meanwhile, to reduce the randomness of the experiment, this part still ran ten

Table 3
Runing time of two concept learning mechanisms on 16 datasets.

Dataset	Mechanisms	t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8	t_9	t_{10}	Ave.
1	Memory	0.05	0.09	0.11	0.15	0.17	0.19	0.21	0.26	0.29	0.31	**0.18**
	Non-memory	0.09	0.13	0.17	0.19	0.22	0.25	0.28	0.31	0.34	0.37	0.24
2	Memory	0.68	1.37	2.05	2.74	3.43	4.17	4.90	5.63	6.35	7.08	**3.84**
	Non-memory	0.75	1.50	2.24	2.99	3.75	4.55	5.34	6.12	6.90	7.69	4.18
3	Memory	0.61	0.62	0.63	0.60	0.61	0.61	0.61	0.61	0.60	0.60	**0.61**
	Non-memory	0.70	0.70	0.69	0.64	0.65	0.64	0.64	0.63	0.63	0.65	0.66
4	Memory	6.17	12.36	18.53	24.70	30.88	37.05	43.27	49.37	55.47	61.56	**33.94**
	Non-memory	6.26	13.16	13.36	13.01	13.33	13.67	13.94	13.38	12.65	12.58	12.53
5	Memory	101.67	203.94	307.67	410.12	515.88	621.89	701.55	752.10	802.65	853.27	**527.07**
	Non-memory	104.51	211.55	314.68	419.15	521.88	626.91	730.08	784.54	838.12	891.95	544.34
6	Memory	149.08	298.06	447.19	596.07	745.37	891.30	1037.19	1183.03	1328.84	1474.74	**815.09**
	Non-memory	184.88	369.89	553.39	735.05	904.04	1067.11	1230.05	1393.28	1556.23	1719.18	971.31
7	Memory	164.68	329.29	493.91	661.99	807.55	953.05	1099.29	1244.56	1389.81	1535.06	**867.92**
	Non-memory	169.66	339.38	512.18	690.12	845.21	1000.12	1148.00	1299.72	1452.21	1605.14	906.17
8	Memory	1.79	3.57	5.36	7.21	9.03	10.86	12.69	14.51	16.34	18.16	**9.95**
	Non-memory	1.91	3.81	5.71	7.68	9.65	11.60	13.56	15.50	17.45	19.40	10.63
9	Memory	16.76	33.52	50.29	67.36	84.04	100.70	117.52	134.13	150.74	167.35	**92.24**
	Non-memory	32.92	67.31	100.98	139.19	169.95	199.16	229.27	256.29	282.93	309.24	178.73
10	Memory	182.40	364.80	547.17	735.45	918.98	1102.50	1286.03	1469.54	1653.08	1836.59	**1009.65**
	Non-memory	612.35	1224.28	1835.86	2476.07	3106.56	3736.50	4366.37	4996.18	5625.92	6255.58	3423.57
11	Memory	0.54	1.09	1.63	2.20	2.76	3.31	3.86	4.41	4.96	5.51	**3.03**
	Non-memory	32.92	67.31	100.98	139.19	169.95	199.16	229.27	256.29	282.93	309.24	178.73
12	Memory	0.13	0.26	0.39	0.52	0.65	0.78	0.91	1.04	1.17	1.30	**0.72**
	Non-memory	0.21	0.40	0.59	0.76	0.94	1.10	1.29	1.48	1.64	1.81	1.02
13	Memory	0.75	1.49	2.23	2.98	3.72	4.46	5.21	5.95	6.70	7.44	**4.09**
	Non-memory	0.89	1.79	2.69	3.57	4.46	5.34	6.22	7.10	7.97	8.85	4.89
14	Memory	0.52	1.04	1.54	2.04	2.54	3.05	3.55	4.05	4.54	5.04	**2.79**
	Non-memory	0.59	1.19	1.78	2.35	2.93	3.53	4.11	4.68	5.26	5.86	3.23
15	Memory	0.10	0.19	0.18	0.15	0.13	0.13	0.12	0.12	0.12	0.12	**0.14**
	Non-memory	0.03	0.07	0.10	0.14	0.17	0.21	0.24	0.27	0.31	0.34	0.19
16	Memory	0.22	0.43	0.63	0.84	1.04	1.24	1.44	1.64	1.84	2.04	**1.14**
	Non-memory	0.23	0.43	0.67	0.88	1.09	1.29	1.50	1.70	1.90	2.11	1.18

times on each time node to obtain the average results. The performance of dynamic data classification has been verified in the above discussion. This part will mainly compare the running time of two mechanisms for dynamic data to explain the advantage of memory in the proposed method. Next, the running time on 16 datasets is shown in Table 3, where results of the concept learning mechanism with less time are displayed in bold. As seen from the comparison of running time and classification for two mechanisms in Fig. 4, compared with the non-memory mechanism, the memory mechanism performs better classification accuracy and running time when facing dynamic data updates.

4.2. Evaluating the performance of CCL

In Section 4.2, the superior performance of the memory mechanism is analyzed for concept-cognitive learning in dynamic fuzzy data classification. To further verify the performance of memory-based concept-cognitive learning in the fuzzy formal context, the proposed method (M-FCCL) with two advanced concept-cognitive learning mechanisms [33,34], including IPFC and IWFC methods, is compared. IPFC is an incremental learning mechanism based on the progressive fuzzy three-way concept, and IWFC is an incremental CCL based on weighted fuzzy concepts. In addition, this subsection mainly validates the superior performance of M-FCCL from two aspects: classification performance evaluation with other CCL methods and running time comparison at different times.

This part mainly verifies the performance of dynamic data classification for the M-FCCL method compared with other fuzzy-based CCL methods. Therefore, the classification accuracy and average accuracy at different time nodes on 16 datasets are selected to compare the performance for different methods. Table 4 records the optimal threshold δ and M-FCCL, IPFC, and IWFC accuracy in these datasets. The last column shows the average accuracy and standard deviation, with excellent results in bold. As shown in this table, M-FCCL has higher accuracy and lower variance (indicating system stability) than other methods, which means that the classification permanence of the proposed method is superior to two CCL mechanisms. Moreover, the average classification accuracy of M-FCCL and the other CCL methods on 16 datasets are 0.8408, 0.7185, and 0.5128. Then the M-FCCL achieves a 17.02% gain in classification accuracy compared with the existing best CCL method, i.e., IPFC. All in all, these results indicate the superiority of the M-FCCL compared with the other two fuzzy-based CCL mechanisms in dynamic fuzzy data classification.

4.3. Evaluating the performance with other methods

In Section 4.3, the advantages of M-FCCL with the other CCL methods are compared for dynamic classification problems. Moreover, considering that M-FCCL is constructed based on the fuzzy concept, it is a fuzzy classifier based on concept recognition degree. Thus, to further verify the classification performance with other machine learning methods, this subsection needs to compare it with some popular distance-based classification methods [45–47]: IF_KNN, MFuzzKNN, FENN, KNN, TreeBigger, and Decision Tree.

In addition, note that these methods cannot directly achieve dynamic data classification. Therefore, the average classification performance (i.e., accuracy, precision, recall, and F1-score) between these methods is used to demonstrate the advantage of the proposed method on 16 datasets. The more intuitive remarkable comparison is shown in Fig. 5. It can be found clearly that the performance of the proposed

Table 4
Accuracy (Ave. ± SD) comparison with other concept-cognitive learning mechanisms at different time on 16 datasets.

Dataset	Methods	Parameter	t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8	t_9	t_{10}	Ave. ± SD
1	M-FCCL	0.05	1.0000	1.0000	1.0000	0.8750	0.9000	0.7500	0.7857	0.7500	0.7778	0.8000	**0.8639 ± 0.1057**
	IPFC	0.30	1.0000	1.0000	1.0000	0.8750	0.8000	0.8333	0.7857	0.6875	0.6111	0.6000	0.8193 ± 0.1533
	IWFC	0.30	0.5000	0.7500	0.8333	0.7500	0.7000	0.7500	0.7143	0.6250	0.5556	0.5500	0.6728 ± 0.1091
2	M-FCCL	0.30	1.0000	0.5556	0.7037	0.7778	0.8222	0.6944	0.6111	0.6597	0.6975	0.7278	0.7250 ± 0.1229
	IPFC	0.00	1.0000	0.8611	0.8704	0.8750	0.8889	0.7685	0.6667	0.5903	0.5370	0.4889	**0.7547 ± 0.1733**
	IWFC	0.30	0.0556	0.0278	0.0185	0.0139	0.0111	0.1667	0.2857	0.3750	0.4444	0.4944	0.1893 ± 0.1938
3	M-FCCL	0.15	0.6129	0.5161	0.4194	0.5565	0.4710	0.4624	0.4747	0.5040	0.5520	0.4968	0.5066 ± 0.0557
	IPFC	0.50	0.7419	0.5806	0.5914	0.5242	0.4968	0.5108	0.4977	0.5161	0.4910	0.5355	**0.5486 ± 0.0761**
	IWFC	0.30	0.4516	0.2419	0.1613	0.1210	0.1032	0.1505	0.2304	0.2379	0.3190	0.3161	0.2333 ± 0.1073
4	M-FCCL	0.50	0.8095	0.8651	0.9101	0.9325	0.9460	0.9550	0.8299	0.8274	0.8466	0.8619	**0.8784 ± 0.0533**
	IPFC	0.15	0.6190	0.6984	0.6402	0.6468	0.6762	0.7037	0.6712	0.6865	0.6949	0.7032	0.6740 ± 0.0295
	IWFC	0.30	0.6032	0.6746	0.5344	0.5119	0.4952	0.5026	0.5057	0.5456	0.5785	0.5889	0.5541 ± 0.0572
5	M-FCCL	0.40	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.8717	0.8878	0.9002	0.8241	**0.9484 ± 0.0694**
	IPFC	0.45	0.9416	0.9489	0.9489	0.9124	0.9139	0.9148	0.8488	0.7856	0.7486	0.7241	0.8688 ± 0.0862
	IWFC	0.75	0.0949	0.0949	0.0657	0.0493	0.0394	0.0328	0.1564	0.2619	0.3439	0.4088	0.1548 ± 0.1359
6	M-FCCL	0.05	0.9910	0.9955	0.9970	0.9977	0.9928	0.9932	0.9942	0.9949	0.9955	0.9959	**0.9948 ± 0.0020**
	IPFC	0.30	0.9321	0.9434	0.9412	0.9480	0.9466	0.9118	0.8533	0.8150	0.7803	0.7502	0.8822 ± 0.0762
	IWFC	0.30	0.1357	0.1267	0.1463	0.1459	0.1493	0.2903	0.3911	0.4661	0.5239	0.5710	0.2946 ± 0.1784
7	M-FCCL	0.00	0.3309	0.6127	0.7026	0.7763	0.8206	0.8435	0.8648	0.8640	0.8739	0.8142	0.7504 ± 0.1689
	IPFC	0.00	0.8382	0.7659	0.8423	0.8388	0.8471	0.8358	0.8403	0.7914	0.7342	0.6618	**0.7996 ± 0.0620**
	IWFC	0.30	0.3309	0.4608	0.5948	0.5141	0.5990	0.6548	0.6800	0.6881	0.6895	0.6711	0.5883 ± 0.1193
8	M-FCCL	0.50	0.9310	0.9655	0.9770	0.7500	0.6690	0.7184	0.7586	0.7845	0.8084	0.8241	**0.8187 ± 0.1061**
	IPFC	0.35	0.4138	0.3621	0.4138	0.4828	0.5724	0.5747	0.5764	0.6121	0.6169	0.6310	0.5256 ± 0.0987
	IWFC	0.80	0.4138	0.3448	0.3793	0.4914	0.5586	0.5460	0.5616	0.5560	0.5824	0.6000	0.5034 ± 0.0915
9	M-FCCL	0.20	0.9120	0.8400	0.8453	0.8060	0.8448	0.7427	0.7063	0.7070	0.6889	0.6952	**0.7788 ± 0.0801**
	IPFC	0.00	0.4960	0.4760	0.4533	0.4520	0.4800	0.4960	0.5234	0.5050	0.4924	0.4896	0.4864 ± 0.0220
	IWFC	0.30	0.0000	0.0000	0.0000	0.0000	0.0192	0.0213	0.0434	0.1630	0.2560	0.3304	0.0833 ± 0.1223
10	M-FCCL	0.25	1.0000	0.9983	0.9989	0.9064	0.8582	0.8818	0.8987	0.9114	0.9212	0.9291	**0.9304 ± 0.0514**
	IPFC	0.50	0.8462	0.8144	0.8094	0.8110	0.8334	0.8428	0.8466	0.8499	0.8532	0.8565	0.8363 ± 0.0182
	IWFC	0.30	0.9365	0.9097	0.9186	0.8662	0.7652	0.6979	0.6536	0.6137	0.5831	0.5629	0.7507 ± 0.1475
11	M-FCCL	0.15	1.0000	1.0000	1.0000	0.9265	0.9412	0.9510	0.9580	0.9632	0.9673	0.9706	**0.9678 ± 0.0257**
	IPFC	0.00	1.0000	0.9706	0.9608	0.9706	0.9 529	0.9608	0.9496	0.9559	0.9477	0.9529	0.9622 ± 0.0155
	IWFC	0.30	0.8824	0.8529	0.8627	0.8676	0.8941	0.9118	0.9244	0.9338	0.9412	0.9471	0.9018 ± 0.0346
12	M-FCCL	0.50	1.0000	0.9286	0.9524	0.9643	0.9714	0.9762	0.8571	0.8214	0.8254	0.8429	**0.9140 ± 0.0695**
	IPFC	0.00	0.8571	0.7857	0.7143	0.7143	0.7429	0.7857	0.7959	0.7679	0.7619	0.7571	0.7683 ± 0.0420
	IWFC	0.30	1.0000	0.9286	0.9048	0.8929	0.9143	0.8810	0.8980	0.8750	0.8413	0.8143	0.8950 ± 0.0501
13	M-FCCL	0.40	0.8125	0.7812	0.8542	0.8906	0.9125	0.9271	0.9375	0.9453	0.9514	0.9563	**0.8969 ± 0.0615**
	IPFC	0.00	0.3750	0.5938	0.7083	0.7812	0.8000	0.8333	0.8393	0.8516	0.8681	0.8812	0.7532 ± 0.1587
	IWFC	0.30	0.6875	0.6562	0.6250	0.5781	0.5625	0.5833	0.5982	0.5781	0.5694	0.5625	0.6001 ± 0.0427
14	M-FCCL	0.25	0.8889	0.9444	0.9630	0.9722	0.9778	0.9074	0.8492	0.8681	0.8765	0.8889	**0.9136 ± 0.0469**
	IPFC	0.00	0.3889	0.6111	0.7407	0.8056	0.8444	0.8333	0.8095	0.8333	0.7840	0.7778	0.7429 ± 0.1417
	IWFC	0.30	0.9444	0.7778	0.6667	0.6111	0.5889	0.4907	0.4286	0.3750	0.3333	0.3000	0.5517 ± 0.2067
15	M-FCCL	0.15	1.0000	0.7500	0.8333	0.8125	0.8500	0.7500	0.7857	0.7188	0.6667	0.6750	**0.7842 ± 0.0980**
	IPFC	0.00	1.0000	0.8750	0.7500	0.6250	0.6500	0.6667	0.7143	0.7188	0.6944	0.7000	0.7394 ± 0.1142
	IWFC	0.30	1.0000	0.5000	0.5000	0.4375	0.4500	0.5000	0.5357	0.5625	0.5278	0.5250	0.5539 ± 0.1612
16	M-FCCL	0.05	1.0000	1.0000	0.7619	0.6429	0.7143	0.7619	0.7755	0.6964	0.6190	0.6286	**0.7601 ± 0.1384**
	IPFC	0.50	0.2857	0.5000	0.4762	0.5357	0.5143	0.5476	0.6122	0.6607	0.6508	0.6714	0.5455 ± 0.1155
	IWFC	0.30	0.2857	0.1429	0.2857	0.4643	0.5143	0.5476	0.5306	0.5357	0.5873	0.5857	0.4480 ± 0.1539

M-FCCL method is significantly better than the other six methods in most cases.

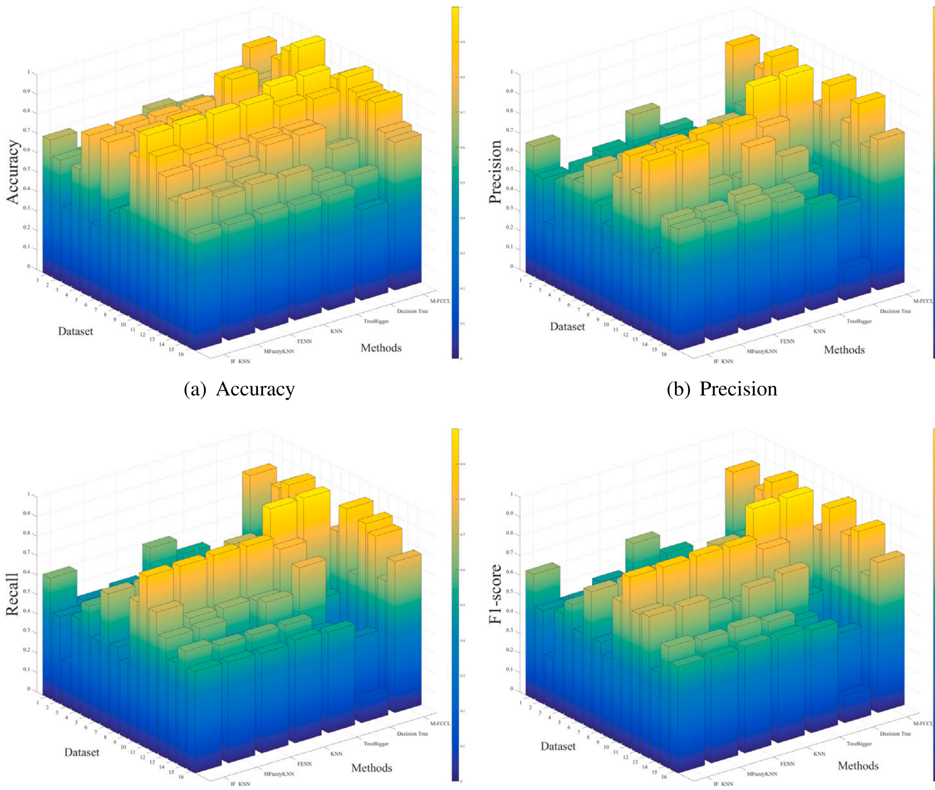
Table 5 records the average classification accuracy of ten-fold cross-validation for M-FCCL and six other classical methods. Table 5 shows that the M-FCCL achieves a maximum value of thirteen times on 16 datasets, except in datasets 7, 10, and 11. Meanwhile, M-FCCL's average accuracy and rank are 0.8395 and 1.56. In particular, the M-FCCL improves 18.54% of average classification accuracy compared to the TreeBigger method on 16 datasets.

Similarly, Tables 6–8 records the value of average precision, recall, and F1-score results for these methods, respectively. While the average results in 16 datasets are the highest among these methods, the classification performance of the proposed method is better than other methods as the number of objects increases. Among these tables, rank is the average order of sixteen datasets under different approaches, and the outstanding results are in bold and underlined. Overall, the results show that the proposed M-FCCL is an excellent dynamic fuzzy data classification method compared to some popular machine learning methods.

To evaluate whether M-FCCL is significantly different from these compared classical methods in classification performance, a Friedman's test is performed on 16 datasets at a significance level of $P = 0.1$. The resulting P-values are 1.21×10^{-7}, 4.98×10^{-8}, 9.96×10^{-9}, and 1.55×10^{-8}, which are lower than the significance level, indicating a statistically significant difference between M-FCCL and the six classical methods.

Consequently, a post hoc test using nemenyi's method is conducted to determine whether there is a significant difference between any two methods. Specifically, the two methods are considered significantly different when the difference between these average ranks is equal to or greater than the critical value. At a confidence level of $\alpha = 0.1$, the critical values for nemenyi's post hoc test are 1.7125 when comparing seven methods on the 16 datasets.

Fig. 6 indicates the nemenyi's test results of M-FCCL and six other machine learning methods in classification performance. This figure



(a) Accuracy (b) Precision

(c) Recall (d) F1-score

Fig. 5. Comparison of classification performance with other methods.

Table 5
Accuracy (Ave. ± SD) comparison with other classification methods on 16 datasets.

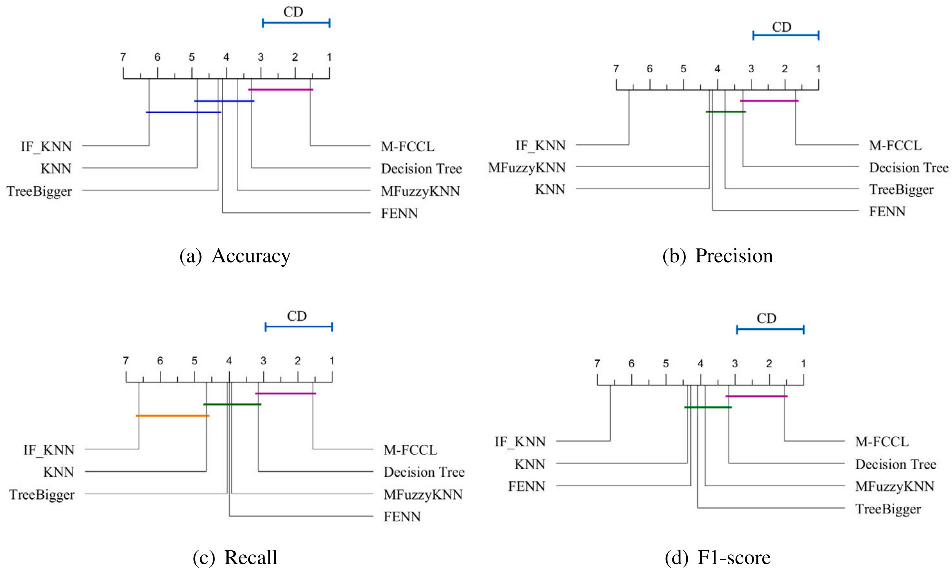
Dataset	IF_KNN	MFuzzyKNN	FENN	KNN	TreeBigger	Decision Tree	M-FCCL
1	0.7008 ± 0.1882	0.5877 ± 0.1021	0.6093 ± 0.1044	0.7008 ± 0.1882	0.6673 ± 0.2241	0.5869 ± 0.2513	**0.8639 ± 0.1057**
2	0.6214 ± 0.0390	0.6241 ± 0.0264	0.4468 ± 0.0686	0.6083 ± 0.0355	0.6879 ± 0.0834	0.4791 ± 0.0251	**0.7250 ± 0.1229**
3	0.3915 ± 0.1503	0.3822 ± 0.1416	0.3822 ± 0.1416	0.3915 ± 0.1503	0.4097 ± 0.0829	0.2025 ± 0.0573	**0.5066 ± 0.0557**
4	0.6270 ± 0.0270	0.6158 ± 0.0294	0.5965 ± 0.0369	0.6110 ± 0.0335	0.7003 ± 0.0667	0.6107 ± 0.0754	**0.8784 ± 0.0533**
5	0.8257 ± 0.0478	0.8236 ± 0.0494	0.8188 ± 0.0490	0.8155 ± 0.0481	0.9187 ± 0.0480	0.7801 ± 0.0513	**0.9484 ± 0.0694**
6	0.3847 ± 0.1570	0.3733 ± 0.1618	0.2375 ± 0.1951	0.5089 ± 0.1207	0.9317 ± 0.0327	0.8118 ± 0.0204	**0.9948 ± 0.0020**
7	**0.8402 ± 0.0666**	0.8401 ± 0.0674	0.8512 ± 0.0636	0.4766 ± 0.2037	0.4039 ± 0.2698	0.1163 ± 0.0477	0.7504 ± 0.1689
8	0.4991 ± 0.1059	0.5237 ± 0.1027	0.4967 ± 0.1334	0.5322 ± 0.0825	0.6113 ± 0.0658	0.3962 ± 0.0404	**0.8187 ± 0.1061**
9	0.5380 ± 0.0397	0.5370 ± 0.0343	0.5217 ± 0.0490	0.5317 ± 0.0259	0.5881 ± 0.0603	0.2898 ± 0.0221	**0.7788 ± 0.0801**
10	0.8429 ± 0.0357	0.8378 ± 0.0374	0.8281 ± 0.0450	0.8502 ± 0.0328	0.9985 ± 0.0013	**1.0000 ± 0.0000**	0.9304 ± 0.0514
11	0.9783 ± 0.0115	0.9783 ± 0.0115	**0.9836 ± 0.0074**	0.9783 ± 0.0115	0.9220 ± 0.0574	0.9088 ± 0.0704	0.9678 ± 0.0257
12	0.8957 ± 0.0457	0.8543 ± 0.0277	0.8267 ± 0.0409	0.8308 ± 0.0299	0.8201 ± 0.0510	0.6698 ± 0.0314	**0.9140 ± 0.0695**
13	0.8170 ± 0.1733	0.8170 ± 0.1733	0.7621 ± 0.2253	0.8232 ± 0.1559	0.8230 ± 0.0492	0.6989 ± 0.1978	**0.8969 ± 0.0615**
14	0.7121 ± 0.1856	0.7108 ± 0.1855	0.7041 ± 0.1841	0.6603 ± 0.1759	0.6258 ± 0.1696	0.5015 ± 0.2191	**0.9136 ± 0.0469**
15	0.7520 ± 0.1040	0.7091 ± 0.1300	0.7220 ± 0.1167	0.7161 ± 0.1102	0.6352 ± 0.1124	0.6133 ± 0.2124	**0.7842 ± 0.0980**
16	0.5975 ± 0.2059	0.5934 ± 0.1906	0.5984 ± 0.2301	0.5914 ± 0.1851	0.5870 ± 0.0513	0.4653 ± 0.2061	**0.7601 ± 0.1384**
Ave. ± SD	0.6890 ± 0.0990	0.6755 ± 0.0919	0.6491 ± 0.1057	0.6642 ± 0.0994	0.7082 ± 0.0891	0.5707 ± 0.0955	**0.8395 ± 0.0785**
Rank	3.13	4.13	4.81	4.00	3.69	6.25	**1.56**
Win/tie/loss	1/0/15	0/0/16	1/0/15	0/0/16	0/0/16	1/0/15	**13/0/3**

shows that the proposed M-FCCL ranks first on compared algorithms in all test indicators and is significantly higher than other methods. These results verify the effectiveness of M-FCCL in object classification.

4.4. Parameters analysis

According to algorithm 1, the parameter δ plays an essential role in constructing cosine similarity granules and fuzzy three-way concepts for the following concept learning process. In order to further investigate the impact of this parameter in the proposed M-FCCL, this subsection record and analyze the classification accuracy of the memory mechanism on 16 datasets.

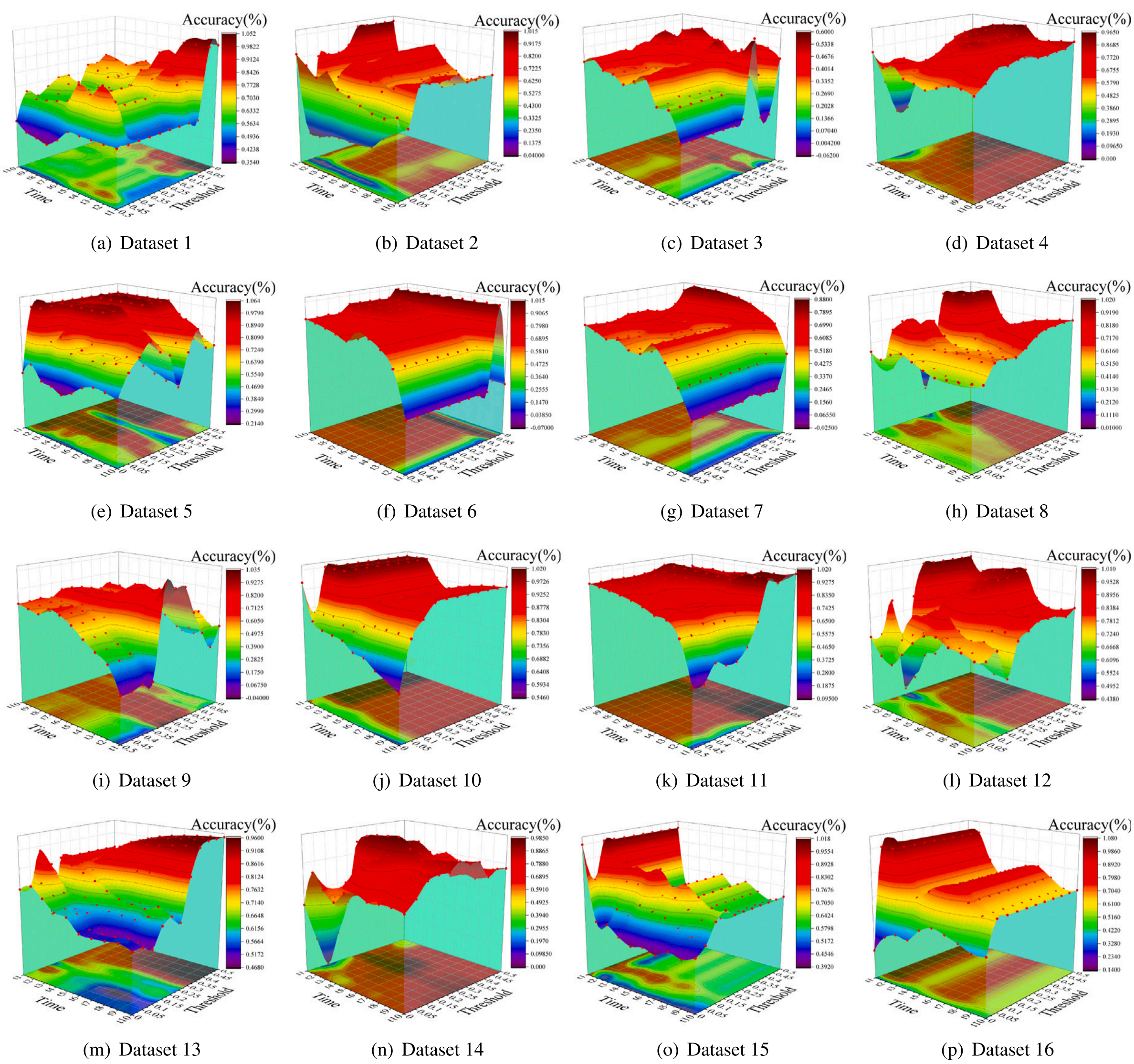
The classification accuracy of M-FCCL under different parameters with time goes is shown in Fig. 7. This figure shows that the accuracy changes significantly with threshold parameter changes, indicating that the memory mechanism is sensitive to this parameter. Consequently, it is necessary to determine an optimal parameter to enhance the concept learning performance. Meanwhile, the classification accuracy varies



(a) Accuracy (b) Precision

(c) Recall (d) F1-score

Fig. 6. Test of Nemenyi on M-FCCL and six classical methods. (a)–(d) show the test results of M-FCCL with the other six classical methods in Accuracy, Precision, Recall, and f1-score in classification.



(a) Dataset 1 (b) Dataset 2 (c) Dataset 3 (d) Dataset 4

(e) Dataset 5 (f) Dataset 6 (g) Dataset 7 (h) Dataset 8

(i) Dataset 9 (j) Dataset 10 (k) Dataset 11 (l) Dataset 12

(m) Dataset 13 (n) Dataset 14 (o) Dataset 15 (p) Dataset 16

Fig. 7. The accuracy comparison under different parameters on 16 datasets.

Table 6
Precision (Ave. ± SD) comparison with other classification methods on 16 datasets.

Dataset	IF_KNN	MFuzzyKNN	FENN	KNN	TreeBigger	Decision Tree	M-FCCL
1	0.6674 ± 0.1930	0.4537 ± 0.1219	0.5394 ± 0.0919	0.6755 ± 0.1927	0.5488 ± 0.2570	0.3899 ± 0.3374	**0.8735 ± 0.1482**
2	0.5280 ± 0.0311	0.5386 ± 0.0416	0.5039 ± 0.0107	0.5267 ± 0.0304	0.5289 ± 0.0344	0.0968 ± 0.0380	**0.5871 ± 0.1835**
3	0.3138 ± 0.0243	0.2939 ± 0.0265	0.2942 ± 0.0266	0.3138 ± 0.0243	0.2565 ± 0.0427	0.1411 ± 0.0163	**0.4513 ± 0.1156**
4	0.5564 ± 0.0569	0.5415 ± 0.0524	0.5082 ± 0.0558	0.5199 ± 0.0556	0.5244 ± 0.0752	0.1111 ± 0.0675	**0.8348 ± 0.1083**
5	0.6006 ± 0.1335	0.6004 ± 0.1335	0.5983 ± 0.1308	0.5974 ± 0.1296	0.6322 ± 0.1719	0.0914 ± 0.0156	**0.9196 ± 0.1746**
6	0.6067 ± 0.1208	0.6059 ± 0.1199	0.5946 ± 0.1083	0.6191 ± 0.1332	0.7043 ± 0.2042	0.0649 ± 0.0152	**0.7471 ± 0.2611**
7	0.7138 ± 0.1813	0.7115 ± 0.1815	**0.7245 ± 0.1846**	0.3470 ± 0.0638	0.2419 ± 0.2833	0.0343 ± 0.0146	0.6546 ± 0.2131
8	0.5395 ± 0.0377	0.5549 ± 0.0465	0.5475 ± 0.0393	0.5526 ± 0.0444	0.5594 ± 0.0491	0.0983 ± 0.0206	**0.6675 ± 0.1742**
9	0.4595 ± 0.0935	0.4584 ± 0.0923	0.4620 ± 0.0959	0.4485 ± 0.0855	0.4170 ± 0.0635	0.0264 ± 0.0029	**0.6450 ± 0.1384**
10	0.7350 ± 0.1757	0.7324 ± 0.1745	0.7319 ± 0.1745	0.7377 ± 0.1774	0.9984 ± 0.0016	**1.0000 ± 0.0000**	0.7960 ± 0.2168
11	0.8742 ± 0.1984	0.8742 ± 0.1984	0.8801 ± 0.2015	0.8742 ± 0.1984	0.8134 ± 0.1743	0.5201 ± 0.2566	**0.9243 ± 0.1629**
12	**0.7081 ± 0.1937**	0.6475 ± 0.1606	0.6415 ± 0.1594	0.6286 ± 0.1415	0.6152 ± 0.1280	0.2002 ± 0.0563	0.6388 ± 0.2004
13	**0.8978 ± 0.1082**	**0.8978 ± 0.1082**	0.3810 ± 0.1127	0.8128 ± 0.0783	0.7143 ± 0.0895	0.2276 ± 0.0442	0.7831 ± 0.1324
14	0.4493 ± 0.1573	0.4497 ± 0.1577	0.4470 ± 0.1555	0.4545 ± 0.1649	0.6026 ± 0.0527	0.2206 ± 0.1018	**0.9078 ± 0.1070**
15	0.6489 ± 0.1733	0.6371 ± 0.1722	0.6690 ± 0.1556	0.6237 ± 0.1785	0.5094 ± 0.1209	0.4505 ± 0.2854	**0.7094 ± 0.1324**
16	0.6231 ± 0.0769	0.6081 ± 0.0690	0.6425 ± 0.0821	0.6041 ± 0.0618	0.5437 ± 0.0410	0.1637 ± 0.0593	**0.7733 ± 0.1745**
Ave. ± SD	0.6201 ± 0.1222	0.6004 ± 0.1160	0.5729 ± 0.1116	0.5835 ± 0.1100	0.5757 ± 0.1118	0.2398 ± 0.0832	**0.7446 ± 0.1652**
Rank	3.13	3.69	4.25	4.06	4.25	6.63	**1.69**
Win/tie/loss	1/1/14	0/1/15	1/0/15	0/0/16	0/0/16	1/0/15	**12/0/4**

Table 7
Recall (Ave. ± SD) comparison with other classification methods on 16 datasets.

Dataset	IF_KNN	MFuzzyKNN	FENN	KNN	TreeBigger	Decision Tree	M-FCCL
1	0.6193 ± 0.2355	0.4521 ± 0.1762	0.4688 ± 0.1664	0.6193 ± 0.2355	0.5369 ± 0.2539	0.3832 ± 0.3313	**0.8363 ± 0.1356**
2	0.4458 ± 0.1260	0.4565 ± 0.1416	0.3541 ± 0.1624	0.4403 ± 0.1285	0.4650 ± 0.0955	0.0582 ± 0.0162	**0.5290 ± 0.2207**
3	0.2353 ± 0.0839	0.2169 ± 0.0907	0.2169 ± 0.0907	0.2353 ± 0.0839	0.2153 ± 0.0636	0.0527 ± 0.0184	**0.3808 ± 0.1037**
4	0.4883 ± 0.1074	0.4876 ± 0.1078	0.4463 ± 0.1031	0.4955 ± 0.1138	0.5634 ± 0.0845	0.1049 ± 0.0718	**0.8485 ± 0.1388**
5	0.5776 ± 0.2035	0.5775 ± 0.2050	0.5738 ± 0.2029	0.5717 ± 0.2027	0.6263 ± 0.1983	0.0798 ± 0.0213	**0.8883 ± 0.1687**
6	0.4244 ± 0.2586	0.4177 ± 0.2602	0.3346 ± 0.2593	0.4995 ± 0.2544	0.7115 ± 0.2015	0.0607 ± 0.0166	**0.7450 ± 0.2604**
7	0.6914 ± 0.1880	0.6908 ± 0.1873	**0.7006 ± 0.1918**	0.3134 ± 0.0749	0.3518 ± 0.2460	0.0086 ± 0.0020	0.5678 ± 0.2788
8	0.4415 ± 0.1835	0.4600 ± 0.1851	0.4411 ± 0.1953	0.4660 ± 0.1777	0.5175 ± 0.1263	0.0557 ± 0.0157	**0.6536 ± 0.1688**
9	0.3833 ± 0.1618	0.3816 ± 0.1585	0.3787 ± 0.1681	0.3649 ± 0.1447	0.3670 ± 0.1125	0.0123 ± 0.0044	**0.6316 ± 0.1783**
10	0.7286 ± 0.2256	0.7246 ± 0.2245	0.7205 ± 0.2291	0.7328 ± 0.2247	0.9981 ± 0.0016	**1.0000 ± 0.0000**	0.7783 ± 0.2039
11	0.8824 ± 0.2083	0.8824 ± 0.2083	0.8871 ± 0.2106	0.8824 ± 0.2083	0.8150 ± 0.1777	0.5028 ± 0.2689	**0.9192 ± 0.1514**
12	**0.7242 ± 0.2440**	0.6423 ± 0.2316	0.6407 ± 0.2495	0.6266 ± 0.2218	0.5800 ± 0.1729	0.1393 ± 0.0324	0.6062 ± 0.1860
13	0.6154 ± 0.0000	0.6154 ± 0.0000	0.5000 ± 0.0000	0.6458 ± 0.0075	0.7715 ± 0.0296	0.1982 ± 0.0436	**0.8936 ± 0.1466**
14	0.5143 ± 0.0225	0.5130 ± 0.0223	0.5099 ± 0.0253	0.4699 ± 0.0248	0.4596 ± 0.0831	0.1521 ± 0.0600	**0.8553 ± 0.1048**
15	0.6003 ± 0.2054	0.5796 ± 0.2150	0.5807 ± 0.2061	0.5774 ± 0.2107	0.4401 ± 0.1842	0.4084 ± 0.3018	**0.6518 ± 0.1404**
16	0.5225 ± 0.2144	0.5180 ± 0.2045	0.5117 ± 0.2165	0.5274 ± 0.2093	0.5172 ± 0.1169	0.1215 ± 0.0539	**0.7753 ± 0.1567**
Ave. ± SD	0.5559 ± 0.1668	0.5385 ± 0.1637	0.5166 ± 0.1673	0.5293 ± 0.1577	0.5585 ± 0.1343	0.2087 ± 0.0786	**0.7225 ± 0.1715**
Rank	3.00	3.94	4.63	3.88	3.94	6.63	**1.56**
Win/tie/loss	1/0/15	0/0/16	1/0/15	0/0/16	0/0/16	1/0/15	**13/0/3**

Table 8
F1-score (Ave. ± SD) comparison with other classification methods on 16 datasets.

Dataset	IF_KNN	MFuzzyKNN	FENN	KNN	TreeBigger	Decision Tree	M-FCCL
1	0.6387 ± 0.2161	0.4458 ± 0.1431	0.4958 ± 0.1349	0.6430 ± 0.2164	0.5395 ± 0.2529	0.3848 ± 0.3346	**0.8532 ± 0.1391**
2	0.4778 ± 0.0874	0.4879 ± 0.1011	0.3966 ± 0.1197	0.4734 ± 0.0892	0.4924 ± 0.0689	0.0718 ± 0.0211	**0.5523 ± 0.2037**
3	0.2614 ± 0.0539	0.2420 ± 0.0622	0.2421 ± 0.0624	0.2614 ± 0.0539	0.2326 ± 0.0552	0.0753 ± 0.0194	**0.4101 ± 0.1051**
4	0.5158 ± 0.0761	0.5090 ± 0.0744	0.4715 ± 0.0749	0.5036 ± 0.0780	0.5427 ± 0.0778	0.1076 ± 0.0694	**0.8363 ± 0.1107**
5	0.5860 ± 0.1683	0.5857 ± 0.1691	0.5827 ± 0.1668	0.5811 ± 0.1662	0.6286 ± 0.1844	0.0849 ± 0.0188	**0.9022 ± 0.1691**
6	0.4675 ± 0.2346	0.4609 ± 0.2384	0.3775 ± 0.2729	0.5319 ± 0.2133	0.7059 ± 0.1998	0.0623 ± 0.0153	**0.7460 ± 0.2607**
7	0.7022 ± 0.1846	0.7009 ± 0.1843	**0.7122 ± 0.1881**	0.3276 ± 0.0649	0.2732 ± 0.2746	0.0134 ± 0.0029	0.6013 ± 0.2592
8	0.4674 ± 0.1439	0.4871 ± 0.1456	0.4678 ± 0.1578	0.4916 ± 0.1370	0.5338 ± 0.0936	0.0699 ± 0.0160	**0.6583 ± 0.1681**
9	0.4114 ± 0.1387	0.4103 ± 0.1359	0.4087 ± 0.1447	0.3971 ± 0.1257	0.3874 ± 0.0933	0.0164 ± 0.0042	**0.6350 ± 0.1505**
10	0.7287 ± 0.1992	0.7252 ± 0.1980	0.7224 ± 0.2004	0.7324 ± 0.1996	0.9983 ± 0.0014	**1.0000 ± 0.0000**	0.7859 ± 0.2080
11	0.8781 ± 0.2030	0.8781 ± 0.2030	0.8834 ± 0.2057	0.8781 ± 0.2030	0.8140 ± 0.1755	0.5109 ± 0.2629	**0.9215 ± 0.1567**
12	**0.7132 ± 0.2145**	0.6412 ± 0.1917	0.6355 ± 0.2006	0.6232 ± 0.1766	0.5948 ± 0.1492	0.1640 ± 0.0408	0.6216 ± 0.1918
13	0.7276 ± 0.0419	0.7276 ± 0.0419	0.4213 ± 0.0981	0.7182 ± 0.0347	0.7399 ± 0.0556	0.2119 ± 0.0440	**0.8341 ± 0.1374**
14	0.4649 ± 0.1187	0.4644 ± 0.1184	0.4616 ± 0.1171	0.4429 ± 0.1079	0.5156 ± 0.0514	0.1749 ± 0.0671	**0.8792 ± 0.0994**
15	0.6223 ± 0.1913	0.6044 ± 0.1972	0.6188 ± 0.1865	0.5982 ± 0.1963	0.4664 ± 0.1633	0.4275 ± 0.2946	**0.6782 ± 0.1344**
16	0.5479 ± 0.1942	0.5402 ± 0.1826	0.5490 ± 0.2030	0.5432 ± 0.1833	0.5260 ± 0.0842	0.1366 ± 0.0559	**0.7712 ± 0.1620**
Ave. ± SD	0.5757 ± 0.1542	0.5569 ± 0.1492	0.5279 ± 0.1584	0.5467 ± 0.1404	0.5619 ± 0.1238	0.2195 ± 0.0792	**0.7304 ± 0.1660**
Rank	3.06	4.00	4.38	4.19	3.88	6.63	**1.56**
Win/tie/loss	1/0/15	0/0/16	1/0/15	0/0/16	0/0/16	1/0/15	**13/0/3**

at different times due to the objective nature of the concept-learning process. Thus, the average accuracy in dynamic data is adopted to determine the optimal parameter.

5. Conclusion

M-FCCL is a practical extension of existing fuzzy-based CCL methods to characterize knowledge from the perspective of fuzzy three-way concepts. It can achieve the dynamic data classification task. The significance of the memory mechanism in M-FCCL contains two meanings: (1) recalling the necessary knowledge in the process of concept space update and (2) forgetting some unnecessary knowledge in the process of cognitive recognition. The essence of M-FCCL is learning concepts mainly from a cosine similarity granule and then knowledge representation via a fuzzy three-way concept. The theory of M-FCCL has been rigorously verified, and simulations in real datasets also validate the practicability of M-FCCL in this study.

Concept-cognitive learning theory is an efficacious data analysis tool for a study concerning knowledge representation and learning in artificial intelligence. Currently, however, fuzzy-based concept-cognitive learning represents and discovers knowledge mainly focused on pseudo-concept learning, which causes a series of issues, such as incomplete and complex cognition concept learning. Moreover, there is also a need to answer some research questions: Does the proposed method outperforms other FCCL methods? How to improve the interpretability of FCCL? What are the differences between these methods? This work mainly explores a novel Memory-based concept-cognitive learning method to answer the above questions.

The current article studies the cognitive-cognitive learning model by introducing recalling and forgetting mechanisms for dynamic data classification and knowledge representation. Hence, some limitations still need to be considered, such as how to apply the proposed method to high-dimensional data analysis or multi-source information fusion, especially in handling incomplete data. Although M-FCCL can significantly improve the efficiency of concept learning and save time for a dynamic data environment, it still cannot be learned for billions of data. Hence, how to combine machine learning and deep learning theory into CCL theory also deserves to be investigated. Our future work will persist in focusing on these topics.

CRediT authorship contribution statement

Doudou Guo: Investigation, Conceptualization, Methodology, Writing – original draft. **Weihua Xu:** Methodology, Investigation, Writing – review & editing. **Yuhua Qian:** Methodology, Writing – review & editing. **Weiping Ding:** Methodology, Writing – review & editing.

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

No data was used for the research described in the article.

Acknowledgments

This paper is supported by the National Natural Science Foundation of China (Nos. 61976245, 61976120). The authors would like to thank Editor-in-Chief, Associate Editor, and Reviewers for their insightful comments and suggestions.

References

[1] Y.X. Wang, N. Howard, J. Kacprzyk, et al., Cognitive informatics: Towards cognitive machine learning and autonomous knowledge manipulation, Int. J. Cogn. Inform. Nat. 12 (2018) 1–13.
[2] Y.Y. Yao, Tri-level thinking: models of three-way decision, Int. J. Mach. Learn. Cybern. 11 (2020) 947–959.
[3] D.D. Guo, C.M. Jiang, S.R. Sheng, et al., A novel outcome evaluation model of three-way decision: A change viewpoint, Inform. Sci. 607 (2022) 1089–1110.
[4] Y.Y. Yao, Three-way conflflict analysis: reformulations and extensions of the Pawlak model, Knowl.-Based Syst. 180 (2019) 26–37.
[5] Y.Y. Yao, Symbols-Meaning-Value (SMV) space as a basis for a conceptual model of data science, Internat. J. Approx. Reason. 144 (2022) 113–128.
[6] B.M. Lake, R. Salakhutdinov, J.B. Tenenbaum, Human-level concept learning through probabilistic program induction, Science 350 (2015) 1332–1338.
[7] Z. Yuan, H.M. Chen, C. Luo, et al., MFGAD: Multi-fuzzy granules anomaly detection, Inf. Fusion 95 (2023).
[8] P.M. Arunkumar, S. Kannimuthu, Mining big data streams using business analytics tools: a bird's eye view on MOA and SAMOA, Int. J. Bus. Intell. Data Min. 17 (2020) 226–236.
[9] V.S. Devi, S. Kannimuthu, Author profiling in code-mixed WhatsApp messages using stacked convolution networks and contextualized embedding based text augmentation, Neural Process. Lett. 55 (2023) 589–614.
[10] B.B. Sang, H.M. Chen, L. Yang, Incremental feature selection using a conditional entropy based on fuzzy dominance neighborhood rough sets, IEEE Trans. Fuzzy Syst. 30 (2022) 1683–1697.
[11] Y.X. Wang, On cognitive foundations and mathematical theories of knowledge science, Int. J. Cogn. Inform. Nat. 10 (2016) 1–25.
[12] R. Wille, Restructuring lattice theory: An approach based on hierarchies of concepts, in: I. Rival (Ed.), Ordered Sets, Springer, Amsterdam the Netherlands, 1982, pp. 445–470.
[13] W.X. Zhang, W.H. Xu, Cognitive model based on granular computing, Chin. J. Eng. Math. 24 (2007) 957–971.
[14] Y.X. Wang, On concept algebra: A denotational mathematical structure for knowledge and software modeling, Int. J. Cogn. Inform. Nat. 2 (2008).
[15] Y.Y. Yao, Interpreting concept learning in cognitive informatics and granular computing, IEEE Trans. Syst. Man Cybern. B 39 (2009) 855–866.
[16] Y.Y. Yao, Three-way granular computing, rough sets, and formal concept analysis, Internat. J. Approx. Reason. 116 (2020) 106–125.
[17] D.D. Guo, C.M. Jiang, P. Wu, Three-way decision based on confifidence level change in rough set, Internat. J. Approx. Reason. 143 (2022) 57–77.
[18] Y.Y. Yao, Three-way decisions and granular computing, Internat. J. Approx. Reason. 103 (2018) 107–123.
[19] J.J. Qi, T. Qian, L. Wei, The connections between three-way and classical concept lattices, Knowl.-Based Syst. 91 (2016) 143–151.
[20] D.D. Guo, W.H. Xu, Fuzzy-based concept-cognitive learning: An investigation of novel approach to tumor diagnosis analysis, Inform. Sci. 639 (2023) 118998.
[21] P.D. Mahendhiran, S. Kannimuthu, CLSA-CapsNet: Dependency based concept level sentiment analysis for text, J. Intell. Fuzzy Systems 43 (2022) 107–123.
[22] J.H. Li, C.L. Mei, W.H. Xu, et al., Concept learning via granular computing: A cognitive viewpoint, Inform. Sci. 298 (2015) 447–467.
[23] J.H. Li, C.C. Huang, J.J. Qi, et al., Three-way cognitive concept learning via multi-granularity, Inform. Sci. 378 (2017) 244–263.
[24] W.H. Xu, W.T. Li, Granular computing approach to two-way learning based on formal concept analysis in fuzzy dataset, IEEE Trans. Cybern. 46 (2016) 366–379.
[25] W.H. Xu, D.D. Guo, Y.H. Qian, et al., Two-way concept-cognitive learning method: A fuzzy-based progressive learning, IEEE Trans. Fuzzy Syst. 31 (2022) 1885–1899.
[26] W.H. Xu, D.D. Guo, J.S. Mi, et al., Two-way concept-cognitive learning via concept movement viewpoint, IEEE Trans. Neural Netw. Learn. (2023) http://dx.doi.org/10.1109/TNNLS.2023.3235800.
[27] X.Y. Zhang, D.D. Guo, W.H. Xu, Two-way concept-cognitive learning with multi-source, Cogn. Comput. (2023) http://dx.doi.org/10.1007/s12559-023-10107-w.
[28] T. Qian, L. Wei, J.J. Qi, Constructing three-way concept lattices based on apposition and subposition of formal contexts, Knowl.-Based Syst. 116 (2017) 39–48.
[29] Y. Shi, Y.L. Mi, J.H. Li, et al., Concept-cognitive learning model for incremental concept learning, IEEE Trans. Syst. Man Cybern.-Syst. 51 (2021) 809–821.
[30] Y.L. Mi, W.Q. Liu, Y. Shi, et al., Semi-supervised concept learning by concept-cognitive learning and concept space, IEEE Trans. Knowl. Data Eng. 34 (2022) http://dx.doi.org/10.1109/TKDE.2020.3010918.
[31] T. Zhang, M. Rong, H.R. Shan, et al., Causal asymmetry analysis in the view of concept-cognitive learning by incremental concept tree, Cogn. Comput. 13 (2021) 1274–1286.
[32] Y.L. Mi, Y. Shi, J.H. Li, et al., Fuzzy-based concept learning method: Exploiting data with fuzzy conceptual clustering, IEEE Trans. Cybern. 52 (2022) 582–593.
[33] K.H. Yuan, W.H. Xu, W.T. Li, et al., An incremental learning mechanism for object classification based on progressive fuzzy three-way concept, Inform. Sci. 584 (2022) 127–147.

[34] C.L. Zhang, C.C. Tsanga, W.H. Xu, et al., Incremental concept-cognitive learning approach for concept classification oriented to weighted fuzzy concepts, Knowl.-Based Syst. 260 (2023) 110093.
[35] W.J. Wang, J.M. Zhan, C. Zhang, et al., A regret-theory-based three-way decision method with a priori probability tolerance dominance relation in fuzzy incomplete information systems, Inf. Fusion 89 (2023) 382–396.
[36] S.T. Zhang, J.J. Zhu, X.D. Liu, et al., Regret theory-based group decision-making with multidimensional preference and incomplete weight information, Inf. Fusion 31 (2016) 1–13.
[37] P.F. Zhang, T.R. Li, Z. Yuan, et al., A data-level fusion model for unsupervised attribute selection in multi-source homogeneous data, Inf. Fusion 80 (2022) 87–103.
[38] L.A. Zadeh, Fuzzy sets, Inf. Control 8 (1965) 338–353.
[39] K. Chola Raja, S. Kannimuthu, Conditional generative adversarial network approach for autism prediction, Comput. Syst. Sci. Eng. 44 (2023) 741–755.
[40] B. Ganter, R. Wille, Formal Concept Analysis, Mathematical Foundations, Springer, Berlin, 1999.
[41] J.J. Niu, D.G. Chen, J.H. Li, et al., Fuzzy rule based classification method for incremental rule learning, IEEE Trans. Fuzzy Syst. http://dx.doi.org/10.1109/TFUZZ.2021.3128061.
[42] J. Yang, G. Wang, Q. Zhang, et al., Knowledge distance measure for the multigranularity rough approximations of a fuzzy Concept, IEEE Trans. Fuzzy Syst. 28 (2020) 706–717.
[43] K. Koffka, Principles of Gestalt Psychology, Harcourt, San Diego, 1967.
[44] A. Elhadad, A. Ghareeb, S. Abbas, A blind and high-capacity data hiding of DICOM medical images based on fuzzification concepts, Alex. Eng. J. 60 (2021) 2471–2482.
[45] Z.H. Zhou, Machine Learning, Tsinghua University Press, Bejing, 2016.
[46] M.N. Adnan, M.Z. Islam, P.A. Forest, Constructing a decision forest by penalizing attributes used in previous trees, Expert Syst. Appl. 18 (2017) 389–403.
[47] J. Derrac, S. Garcia, F. Herrera, Fuzzy nearest neighbor algorithms: Taxonomy, experimental analysis and prospects, Inform. Sci. 260 (2014) 98–119.

Doudou Guo received the M.Sc. degree from the School of Computer Science and Information Engineer, Harbin normal University, Harbin, China, in 2021. He is currently pursuing the Ph.D. degree from the College of Artificial Intelligence, Southwest University, Chongqing, China. His current research interests include concept learning, cognitive computing, granular computing, fuzzy sets, three-way decision. He has authored over 10 articles on these topics in international journals.

Weihua Xu received the Ph.D. degree from School of Sciences, Xi'an Jiaotong University, Xi'an, China, in 2007, and the M.Sc. degree from School of Mathematics and Information Sciences, Guangxi University, Nanning, China, in 2004. He is currently a professor with the College of Artificial Intelligence, Southwest University, Chongqing, China. He serves in the editorial boards of several international journals, and has published 4 monographs and over 180 articles in international journals. His current research interests include granular computing, cognitive computing, information fusion.

Yuhua Qian received the M.S. and Ph.D. degrees in computers with applications from Shanxi University, Taiyuan, China, in 2005 and 2011, respectively. He is currently a Director at the Institute of Big Data Science and Industry, Shanxi University, where he is also a Professor with the Key Laboratory of Computational Intelligence and Chinese Information Processing, Ministry of Education.He is involved in research on machine learning, pattern recognition, feature selection, granular computing, and artificial intelligence. He has authored over 100 articles on these topics in international journals. He served on the Editorial Board of the *International Journal of Knowledge-Based Organizations* and *Artificial Intelligence Research*. He has served as the Program Chair or Special Issue Chair of the Conference on Rough Sets and Knowledge Technology, the Joint Rough Set Symposium, and the Conference on Industrial Instrumentation and Control and a PC Member of many machine learning, data mining conferences.

Weiping Ding (M'16-SM'19) received the Ph.D. degree in Computer Science, Nanjing University of Aeronautics and Astronautics, Nanjing, China, in 2013. From 2014 to 2015, he is a Postdoctoral Researcher at the Brain Research Center, National Chiao Tung University, Hsinchu, Taiwan. In 2016, He was a Visiting Scholar at National University of Singapore, Singapore. From 2017 to 2018, he was a Visiting Professor at University of Technology Sydney, Ultimo, NSW, Australia. Now he is a professor of School of Information Science and Technology, Nantong University, China. His research interests include deep neural networks, multimodal machine learning, and medical images analysis. He has published over 200 articles, including over 90 IEEE journal papers, such as *IEEE Transactions on Fuzzy Systems*, *IEEE Transactions on Neural Networks and Learning Systems*, *IEEE Transactions on Cybernetics*, *IEEE Transactions on Evolutionary Computation*. His fifteen authored/co-authored papers have been selected as ESI Highly Cited Papers. He has co-authored four books. He has holds 27 approved invention patents, including one U.S. patent and one Australian patent. He served/serves as the Editorial Board Member of *Knowledge-Based Systems*, *Engineering Applications of Artificial Intelligence*, *Applied Soft Computing*, and serves as the Area Editor of *Information Fusion*. He serves as the Associate Editor of *IEEE Transactions on Neural Networks and Learning Systems*, *IEEE Transactions on Fuzzy Systems*, *IEEE/CAA Journal of Automatica Sinica*, *IEEE Transactions on Intelligent Transportation Systems*, *IEEE Transactions on Emerging Topics in Computational Intelligence*, *IEEE Transactions on Intelligent Vehicles*, *IEEE Transactions on Artificial Intelligence*, *Information Sciences*, *Neurocomputing*, and so on. He is the Leading Guest Editor of Special Issues in several prestigious journals, including *IEEE Transactions on Evolutionary Computation*, *IEEE Transactions on Fuzzy Systems*, and *Information Fusion*. He is the Co-Editor-in-Chief of both *Journal of Artificial Intelligence and Systems* and *Artificial Intelligence Advances*.