

Article

Forgetting-Based Concept-Cognitive Learning for Classification in Fuzzy Formal Decision Context

Chuanhong Sun ¹ , Xuewei Ling ² and Chengling Zhang ^{3,*}

¹ Department of Mathematics, Shantou University, Shantou 515063, China; 21chsun@stu.edu.cn

² School of Mathematics and Statistics, Shaanxi Normal University, Xi'an 710119, China; lxw930628@163.com

³ College of Mathematics and Computer Science, Quanzhou Normal University, Quanzhou 362000, China

* Correspondence: chengling_zhang@126.com

Abstract

Concept-cognitive learning reveals the principle of human cognition by simulating the brain's process of learning and processing concepts. Nevertheless, for neighborhood similarity granules, the average information of objects regarding all attributes is not considered, which may lead to unbalanced acquisition of knowledge. On the other hand, there are some unnecessary concepts in the extension of fuzzy concepts, which results in poor classification learning. To tackle these challenges, we present a forgetting-based concept-cognitive learning model for classification in a fuzzy formal decision context. Firstly, the fuzzy concept space is established based on the correlation coefficient matrix. Then, to delete unnecessary objects that are in the zone of proximal development, we construct the forgetting fuzzy concept space by selecting the concept corresponding to the maximum similarity. Subsequently, a forgetting-based fuzzy concept model (FCCLM) mechanism is proposed. In the end, experimental results on eight datasets validate the feasibility and efficiency of the proposed learning mechanism through classification performance assessment.

Keywords: concept-cognitive learning; fuzzy formal concept analysis; correlation coefficient; forgetting fuzzy concept; knowledge fusion

MSC: 03C45



Academic Editor: Hsien-Chung Wu

Received: 29 May 2025

Revised: 18 July 2025

Accepted: 25 July 2025

Published: 1 August 2025

Citation: Sun, C.; Ling, X.; Zhang, C. Forgetting-Based Concept-Cognitive Learning for Classification in Fuzzy Formal Decision Context. *Axioms* **2025**, *14*, 593. <https://doi.org/10.3390/axioms14080593>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Cognitive computing [1], as an important applied branch of cognitive science, can employ computational approaches to develop innovative solutions for complex real-world challenges [2,3]. Concepts are created by humans via the generalization and abstraction of the fundamental attributes of objects. As fundamental building blocks of human knowledge systems, they provide a structured approach to knowledge representation and organization, which is crucial for cognitive computing.

Wille [4,5] introduced formal concept analysis (FCA for short) to illustrate the formal context and formal concepts from a mathematical standpoint. With the development of FCA, some extended models such as fuzzy concepts, interval set concepts, three-way concepts, and multi-granularity concepts have been proposed by scholars [6–9]. It should be noted that fuzzy concepts generated from a fuzzy formal context are a special case of Chu space [10,11]. The basic form of a fuzzy Chu space is a triple (A, X, r) , where A and X are sets, $r : A \times X \rightarrow k$ is a mapping, and k is a truth value set. When k takes values in the unit interval $[0, 1]$, the fuzzy Chu space is reduced to fuzzy formal concept analysis. In

recent years, the combination of FCA and cognitive computing has led to the emergence of a popular research direction, namely concept-cognitive learning (CCL for short). CCL denotes the acquisition of concepts from provided clues via specific cognitive models, aiming to uncover the systematic regularity of how the human brain learns concepts. For example, a novel concept-cognitive model based on granular computing was first studied by Zhang et al. [12], and the experiments show that the model is feasible. Yao [13] put forward a three-level framework of concept cognition from three aspects: the abstract layer, the brain layer, and the machine layer. Qiu et al. [14] proposed a concept-cognitive system that can realize multiple concept cognitions. Additionally, Kumar et al. [15] investigated the concept learning approach to analyze functionalities of bidirectional associative memory. Thereafter, Li et al. [16] discussed the concept learning mechanism from the perspective of philosophy and cognitive psychology. Therefore, the above-mentioned results in concept-cognitive systems have mainly been obtained by means of granular computing.

Subsequently, common concept-cognitive learning models mainly include information similarity matching [17–19]. Xu et al. [17] designed a two-way learning method of fuzzy concepts from the necessary or sufficient information granules in fuzzy datasets. Then Xu et al. [18] further proposed the dynamic learning of necessary and sufficient concepts in a dynamic environment. Meanwhile, many scholars have presented achievements in the combination of CCL and machine learning, which mainly address the problems of cognitive classification of data and knowledge fusion. For example, Mi et al. [20–22] introduced an incremental CCL model [20], fuzzy-based CCL model [21], and concurrent CCL model [22]. Furthermore, Yuan et al. [23] presented an incremental learning mechanism based on progressive fuzzy three-way concepts. Zhang et al. [24] discussed a knowledge discovery and concept classification method by assigning different weights to the fuzzy formal context. Next, Wang et al. [25] studied an innovative multi-view fuzzy CCL model to integrate concepts through multiple views. Liang et al. [26] investigated an incremental CCL method based on concept reduction for dynamic classification in the formal context. In [23,24], note that neighborhood similarity granules induced by the Euclidean distance or cosine theorem do not consider the average information of objects. To this end, the correlation coefficient matrix is discussed to improve classification performance, which is a motivation for this paper.

The advantage of CCL is that it can integrate past experiences into itself to describe dynamic data, and some achievements have already been made. For instance, Guo et al. [27] introduced a memory-based CCL model to mine knowledge by combining the recalling and forgetting mechanisms. It is worth noting that the forgetting mechanism in the work of Guo is to forget the maximum concept in the inclusion relationship while retaining the remaining concepts. However, the principle of forgetting is that short-term memories are usually forgotten. This also means that objects are easily forgotten that are in the zone of proximal development. Therefore, the formulation of forgetting rules in fuzzy concepts remains an open research question requiring further investigation. We address this challenge in Section 3 through a novel proposed solution.

To overcome the aforementioned constraints, we propose a forgetting-based concept-cognitive learning model (FCCLM for short) for classification. Concretely, the fuzzy concept space is constructed with the correlation coefficient matrix. It is known that objects are easily forgotten that are in the zone of proximal development, which can improve knowledge discovery classification performance. Afterwards, to remove objects that are easily forgotten, we introduce a forgetting-based concept-cognitive learning mechanism for fusing concepts and class prediction. The overall workflow diagram is shown in Figure 1. Meanwhile, this paper makes the following innovations.

- The fuzzy concept based on the correlation coefficient matrix describes the average information embedded in data, which further decreases cognitive biases.
- It designs a concept-forgetting method via removing some unnecessary objects, which can improve the classification performance of concept recognition.

The remainder of this article is organized as follows. We recall some related notations about fuzzy formal concept analysis in Section 2. Section 3 studies the construction of the forgetting fuzzy concept space with the correlation coefficient and concept classification. The experimental results are presented and analyzed in Section 4. Finally, Section 5 concludes this work and outlines directions for future research.

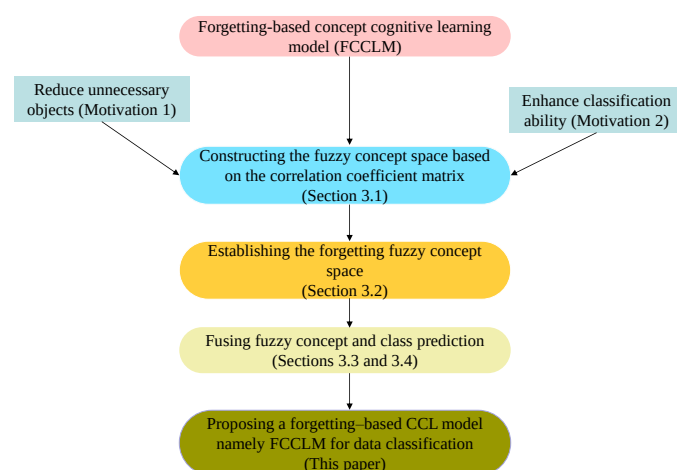


Figure 1. The overall workflow diagram.

2. Preliminaries

This section provides a concise overview of fundamental notions in fuzzy formal concept analysis. More detailed information can be found in [4–6].

Fuzzy Formal Concept Analysis

Consider a universe G ; a fuzzy set $\tilde{F}$ of G is characterized by a mapping $\tilde{F}(\cdot) : G \rightarrow [0, 1]$. For any $g \in G$, the value $\tilde{F}(g)$ represents the degree of membership of g belonging to $\tilde{F}$. We denote $\mathcal{F}(G)$ by the collection of all fuzzy sets on G .

Let E and F be two fuzzy sets on G . If $\tilde{E}(x) \leq \tilde{F}(x)$, $x \in G$, then E is a subset of F ; that is to say, $E \subseteq F$. Specifically, the collection of all crisp sets defined on G is denoted as $\mathcal{P}(G)$.

A triplet $(G, A, \tilde{R})$ is referred to as a fuzzy formal context, where $G = \{x_1, x_2, \dots, x_n\}$ and $A = \{a_1, a_2, \dots, a_m\}$ are the sets of objects and attributes, respectively. Let $\tilde{R}$ be a fuzzy relation between G and A (i.e., $\tilde{R} : G \times A \rightarrow [0, 1]$), with each $\tilde{R}(x, a)$ indicating the degree to which object x possesses attribute a .

Definition 1. Given a fuzzy formal context $(G, A, \tilde{R})$, for $Y \subseteq G$ and $\tilde{C} \in \mathcal{F}(A)$, two learning operators $\tilde{\mathcal{H}} : \mathcal{P}(G) \rightarrow \mathcal{F}(A)$ and $\mathcal{G} : \mathcal{F}(A) \rightarrow \mathcal{P}(G)$ are described as follows:

$$\begin{aligned} \tilde{\mathcal{H}}(Y)(a) &= \bigwedge_{x \in Y} \tilde{R}(x, a), a \in A, \\ \mathcal{G}(\tilde{C}) &= \left\{ x \in G : \forall a \in A, \tilde{C}(a) \leq \tilde{R}(x, a) \right\}, \end{aligned} \quad (1)$$

in which the pair $(Y, \tilde{C})$ is a fuzzy concept satisfying $\tilde{\mathcal{H}}(Y) = \tilde{C}$ and $\mathcal{G}(\tilde{C}) = Y$. In general, Y and $\tilde{C}$ are named extent and intent, respectively.

For any $Y, Y_1, Y_2 \subseteq G, \tilde{C}, \tilde{C}_1, \tilde{C}_2 \in \mathcal{F}(A)$, we derive the following characteristic properties:

1. $Y_1 \subseteq Y_2 \Rightarrow \tilde{\mathcal{H}}(Y_2) \subseteq \tilde{\mathcal{H}}(Y_1);$
2. $\tilde{\mathcal{C}}_1 \subseteq \tilde{\mathcal{C}}_2 \Rightarrow \mathcal{G}(\tilde{\mathcal{C}}_2) \subseteq \mathcal{G}(\tilde{\mathcal{C}}_1);$
3. $Y \subseteq \mathcal{G}\tilde{\mathcal{H}}(Y), \tilde{\mathcal{C}} \subseteq \tilde{\mathcal{H}}\mathcal{G}(\tilde{\mathcal{C}});$
4. $Y = \tilde{\mathcal{H}}\mathcal{G}\tilde{\mathcal{H}}(Y), \tilde{\mathcal{C}} = \mathcal{G}\tilde{\mathcal{H}}\mathcal{G}(\tilde{\mathcal{C}});$
5. $Y \subseteq \mathcal{G}(\tilde{\mathcal{C}}) \Leftrightarrow \tilde{\mathcal{C}} \subseteq \tilde{\mathcal{H}}(Y);$
6. $\tilde{\mathcal{H}}(Y_1 \cup Y_2) = \tilde{\mathcal{H}}(Y_1) \cap \tilde{\mathcal{H}}(Y_2), \mathcal{G}(\tilde{\mathcal{C}}_1 \cup \tilde{\mathcal{C}}_2) = \mathcal{G}(\tilde{\mathcal{C}}_1) \cap \mathcal{G}(\tilde{\mathcal{C}}_2).$

In addition, suppose $(G, A, \tilde{R})$ and (G, D, J) are two fuzzy formal contexts, where $\tilde{R} : G \times A \rightarrow [0, 1]$ and $J : G \times D \rightarrow \{0, 1\}$. Then a fuzzy formal decision context (FFDC) is characterized by $(G, A, \tilde{R}, D, J)$, where $A \cap D = \emptyset$. $G/D = \{D_1, D_2, \dots, D_t\}$ is regarded as a decision division based on decision attribute set D with $D = D_1 \cup D_2 \cup \dots \cup D_t$.

3. The Construction of a Fuzzy Concept with the Correlation Coefficient

3.1. Constructing the Fuzzy Concept Space

This section introduces a correlation coefficient matrix to construct the correlation similarity granule for achieving concept characterization.

Definition 2. Given a fuzzy formal context $(G, A, \tilde{R})$, for any $x_i, x_j \in G$, the correlation coefficient between x_i and x_j is defined as follows:

$$r(x_i, x_j) = \frac{\sum_{k=1}^m |\tilde{R}(i, k) - \bar{R}(i)| |\tilde{R}(j, k) - \bar{R}(j)|}{\sqrt{\sum_{k=1}^m (\tilde{R}(i, k) - \bar{R}(i))^2} \sqrt{\sum_{k=1}^m (\tilde{R}(j, k) - \bar{R}(j))^2}}, \quad (2)$$

where $\tilde{R}(i, k)$ denotes the membership degree of object x_i with respect to attribute a_k , while $\bar{R}(i)$ represents the average membership degree of x_i across the attribute set A . That is to say, $\bar{R}(i) = \frac{1}{m} \sum_{k=1}^m \tilde{R}(i, k)$. The value of $r(x_i, x_j)$ explains the similarity degree between x_i and x_j . The pairwise correlation coefficients form a symmetric matrix Q , where $Q(i, j) = r(x_i, x_j)$ quantifies the linear relationship between objects x_i and x_j . Meanwhile, this correlation coefficient matrix can be used to construct the correlation similarity granule.

Definition 3. Given an FFDC $(G, A, \tilde{R}, D, J)$, for any $D_i \in G/D$, the correlation similarity granule of object x is given by

$$CG(x) = \{y \in D_i | r(x, y) \geq \beta\}, \quad (3)$$

where β is a threshold and $CG(x)$ is the correlation similarity degree between x and y .

Definition 4. Given an FFDC $(G, A, \tilde{R}, D, J)$, where $G/D = \{D_1, D_2, \dots, D_t\}$, for any D_i , the fuzzy concept subspace $\mathcal{S}_i$ about D_i is denoted as follows:

$$\mathcal{S}_i = \left\{ \left(\mathcal{G}\tilde{\mathcal{H}}(CG(x)), \tilde{\mathcal{H}}(CG(x)) \right) \middle| x \in D_i \right\}. \quad (4)$$

where β is a threshold and $CG(x)$ is the correlation similarity degree between x and y .

Subsequently, we denote by $\mathcal{S} = \{\mathcal{S}_1, \mathcal{S}_2, \dots, \mathcal{S}_t\}$ the fuzzy concept space, where $\mathcal{S}_i$ is named a fuzzy subspace of $\mathcal{S}$. It should be emphasized that comprehensive feature learning of each individual object serves as a prerequisite for achieving optimal classification performance. Then we propose the procedure of constructing the fuzzy concept space in Algorithm 1 with the time complexity of $O(|G|^2|A|)$.

Algorithm 1: Constructing fuzzy concept space (CFCS).

```

1 Input: A FFDC  $(G, A, \tilde{R}, D, J)$  and  $\beta$ ;
2 Output: Fuzzy concept space  $\mathcal{S}$ .
3 Compute the decision partition  $G/D = \{D_1, D_2, \dots, D_t\}$ ;
4 for each  $D_i \in G/D$  do
5   Set  $\mathcal{S}_i \leftarrow \emptyset$ ;
6   for each  $x \in D_i$  do
7     Compute the correlation similarity granule  $CG(x)$  of object  $x$ ;
8     Get the fuzzy concept  $(\mathcal{G}\tilde{\mathcal{H}}(CG(x)), \tilde{\mathcal{H}}(CG(x)))$ ;
9      $\mathcal{S}_i \leftarrow (\mathcal{G}\tilde{\mathcal{H}}(CG(x)), \tilde{\mathcal{H}}(CG(x)))$ ;
10  end
11   $\mathcal{S} \leftarrow \mathcal{S}_i$ ;
12 end
13 Return  $\mathcal{S} = \{\mathcal{S}_1, \mathcal{S}_2, \dots, \mathcal{S}_t\}$ .

```

Example 1. Table 1 is an FFDC $(G, A, \tilde{R}, D, J)$, where $G = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8\}$ and $A = \{a_1, a_2, a_3, a_4\}$. d is the decision attribute that partitions all objects into two classes, with $D_1 = \{x_1, x_2, x_3, x_4\}$ and $D_2 = \{x_5, x_6, x_7, x_8\}$. The correlation coefficient matrix Q is computed as follows:

$$Q = \begin{bmatrix} 1.00 & 0.87 & 0.74 & 0.83 & 0.48 & 0.83 & 0.98 & 0.87 \\ 0.87 & 1.00 & 0.95 & 0.72 & 0.67 & 0.79 & 0.90 & 0.94 \\ 0.74 & 0.95 & 1.00 & 0.54 & 0.70 & 0.83 & 0.76 & 0.96 \\ 0.83 & 0.72 & 0.54 & 1.00 & 0.69 & 0.65 & 0.90 & 0.59 \\ 0.48 & 0.67 & 0.70 & 0.69 & 1.00 & 0.65 & 0.90 & 0.59 \\ 0.83 & 0.79 & 0.83 & 0.65 & 0.68 & 1.00 & 0.80 & 0.92 \\ 0.98 & 0.90 & 0.76 & 0.90 & 0.59 & 0.80 & 1.00 & 0.84 \\ 0.87 & 0.94 & 0.96 & 0.59 & 0.60 & 0.92 & 0.84 & 1.00 \end{bmatrix}.$$

Given $\beta = 0.65$, for the objects in decision class D_1 , we can have the correlation similarity granule $CG(x_1) = \{x_1, x_2, x_3, x_4\}$, $CG(x_2) = \{x_1, x_2, x_3, x_4\}$, $CG(x_3) = \{x_1, x_2, x_3\}$, and $CG(x_4) = \{x_1, x_2, x_4\}$. Next, fuzzy concept subspace $\mathcal{S}_1 = \{(\{x_1, x_2, x_3, x_4\}, (0.1, 0.3, 0.4, 0.4)), (\{x_1, x_2, x_3\}, (0.1, 0.3, 0.4, 0.5))\}$.

For the objects in decision class D_2 , their correlation similarity granule and fuzzy concept subspace are as follows: $CG(x_5) = \{x_5, x_6\}$, $CG(x_6) = \{x_5, x_6, x_7, x_8\}$, $CG(x_7) = \{x_6, x_7, x_8\}$ and $CG(x_8) = \{x_6, x_7, x_8\}$. $\mathcal{S}_2 = \{(\{x_5, x_6\}, (0.3, 0.4, 0.3, 0.4)), (\{x_5, x_6, x_7, x_8\}, (0.1, 0.4, 0.3, 0.3)), (\{x_6, x_7, x_8\}, (0.1, 0.5, 0.4, 0.3))\}$.

Table 1. An FFDC.

U	a_1	a_2	a_3	a_4	d
x_1	0.7	0.3	0.4	0.5	1
x_2	0.1	0.8	0.5	0.9	1
x_3	0.3	0.5	0.6	0.7	1
x_4	0.4	0.3	0.4	0.4	1
x_5	0.6	0.4	0.3	0.9	2
x_6	0.3	0.5	0.6	0.4	2
x_7	0.2	0.5	0.4	0.3	2
x_8	0.1	0.7	0.8	0.8	2

3.2. Constructing the Forgetting Fuzzy Concept Space

Section 3.1 discusses the construction process of the fuzzy concept space through the correlation similarity granule and a pair of cognitive operators $\tilde{\mathcal{H}}$ and $\mathcal{G}$. However,

since $X \subseteq \mathcal{GH}(X)$, some of the objects in extent $\mathcal{GH}(X)$ may be forgotten because of knowledge forgetting.

Definition 5. Given an FFDC $(G, A, \tilde{R}, D, J)$, for arbitrary $X \subseteq G$, the inner fringe of X is denoted as

$$\hbar(X) = \left\{ \mathcal{GH}(X) \setminus Y_i \mid Y_j \not\subseteq Y_i, \text{ for any } Y_j \in \Psi (i \neq j) \right\}. \quad (5)$$

where $\Psi = \left\{ L(\tilde{\mathcal{H}}(X), a) \mid a \in A, \tilde{\mathcal{H}}(X)(a) \neq 0 \right\} = \left\{ Y_1, Y_2, \dots, Y_{|\Psi|} \right\}$ and $L(\tilde{\mathcal{H}}(X), a) = \left\{ g \in \mathcal{GH}(X) \mid \tilde{R}(g, a) = \tilde{\mathcal{H}}(X)(a) \right\}$ for $a \in A$.

Definition 5 introduces a method for forgetting the minimal object set that is proximate to a certain object in the extent $\mathcal{GH}(X)$. Then the above forgetting minimal object set can be represented by a Boolean matrix, shown as Proposition 1.

Proposition 1. Let $(G, A, \tilde{R}, D, J)$ be an FFDC. $\hbar(X)$ is an inner fringe of $X \subseteq G$. Then, we know the following:

$$\hbar(X) = \left\{ \mathcal{GH}(X) \setminus Y_i \mid \sum_{j=1}^{|\Psi|} d_{ij} = 1, Y_i \in \Psi \right\}. \quad (6)$$

where $\Psi = \left\{ L(\tilde{\mathcal{H}}(X), a) \mid a \in A, \tilde{\mathcal{H}}(X)(a) \neq 0 \right\} = \left\{ Y_1, Y_2, \dots, Y_{|\Psi|} \right\}$ and

$$d_{ij} = \begin{cases} 1, & Y_j \subseteq Y_i, \\ 0, & \text{otherwise.} \end{cases}$$

for $i, j = 1, 2, \dots, |\hbar|$.

Proof. For any $Y_i \in \Psi$, $\sum_{j=1}^{|\hbar|} d_{ij} = 1$ means that

$$d_{ij} = \begin{cases} 1, & j = i, \\ 0, & j \neq i. \end{cases}$$

That is to say, for any $Y_j \in \Psi (j \neq i)$, it implies that $Y_j \not\subseteq Y_i$. Hence, for any $Y_i \in \Psi$, if $\sum_{j=1}^{|\hbar|} d_{ij} = 1$, then it is obvious that $\mathcal{GH}(X) \setminus Y_i \in \hbar(X)$. $\square$

Proposition 2. Let $(G, A, \tilde{R}, D, J)$ be an FFDC. $\hbar(X)$ is an inner fringe of $X \subseteq G$. For any $V_i \in \hbar(X)$, $(\mathcal{GH}(V_i), \tilde{\mathcal{H}}(V_i))$ is a fuzzy concept.

Proof. It is obvious from Definition 1. $\square$

In fact, there are multiple elements in $\hbar(X)$. After applying knowledge forgetting, the obtained forgetting fuzzy concept $(\mathcal{GH}(V_i), \tilde{\mathcal{H}}(V_i))$ is defined, which is maximally similar to fuzzy concept $(\mathcal{GH}(X), \tilde{\mathcal{H}}(X))$. For two concepts $(\mathcal{GH}(V_i), \tilde{\mathcal{H}}(V_i))$ and $(\mathcal{GH}(X), \tilde{\mathcal{H}}(X))$, the similarity degree is $K((\mathcal{GH}(V_i), \tilde{\mathcal{H}}(V_i)), (\mathcal{GH}(X), \tilde{\mathcal{H}}(X))) = \frac{|\mathcal{GH}(V_i) \cap \mathcal{GH}(X)|}{|\mathcal{GH}(V_i) \cup \mathcal{GH}(X)|} + \sqrt{\sum_{a \in A} (\tilde{\mathcal{H}}(V_i) - \tilde{\mathcal{H}}(X_i))^2}$.

Definition 6. Let $(G, A, \tilde{R}, D, J)$ be an FFDC. For $(\mathcal{GH}(X), \tilde{\mathcal{H}}(X)) \in S_i$, the collection of forgetting fuzzy concepts is defined as follows:

$$\mathcal{F}_X^i = \left\{ (\mathcal{GH}(V_i), \tilde{\mathcal{H}}(V_i)) \mid K((\mathcal{GH}(V_i), \tilde{\mathcal{H}}(V_i)), (\mathcal{GH}(X), \tilde{\mathcal{H}}(X))) = \text{Max0 for } V_i \in \hbar(X) \right\}, \quad (7)$$

where $Max0 = \max_{V_i \in \tilde{h}(X)} K((\mathcal{G}\tilde{\mathcal{H}}(V_i), \tilde{\mathcal{H}}(V_i)), (\mathcal{G}\tilde{\mathcal{H}}(X), \tilde{\mathcal{H}}(X)))$.

In a fuzzy formal decision context $(G, A, \tilde{R}, D, J)$ and $X \subseteq G$, there are some fuzzy subconcepts of $(\mathcal{G}\tilde{\mathcal{H}}(X), \tilde{\mathcal{H}}(X))$; that is, certain fuzzy concepts can be obtained from $(\mathcal{G}\tilde{\mathcal{H}}(X), \tilde{\mathcal{H}}(X))$ after knowledge forgetting through the inner fringe $\tilde{h}(X)$. From Proposition 2, for any $V_i \in \tilde{h}(X)$, $(\mathcal{G}\tilde{\mathcal{H}}(V_i), \tilde{\mathcal{H}}(V_i))$ is a fuzzy concept obtained from $(\mathcal{G}\tilde{\mathcal{H}}(X), \tilde{\mathcal{H}}(X))$ after knowledge forgetting. It is obvious that $(\mathcal{G}\tilde{\mathcal{H}}(V_i), \tilde{\mathcal{H}}(V_i)) \leq (\mathcal{G}\tilde{\mathcal{H}}(X), \tilde{\mathcal{H}}(X))$. The higher the similarity between $(\mathcal{G}\tilde{\mathcal{H}}(V_i), \tilde{\mathcal{H}}(V_i))$ and $(\mathcal{G}\tilde{\mathcal{H}}(X), \tilde{\mathcal{H}}(X))$, the higher the probability that objects in $\mathcal{G}\tilde{\mathcal{H}}(X) \setminus Y_i$ are forgotten.

In addition, given D_i , the forgetting fuzzy concept subspace is denoted as $\mathcal{F}^i = \{\mathcal{F}_X^i | (\mathcal{G}\tilde{\mathcal{H}}(X), \tilde{\mathcal{H}}(X)) \in S_i\}$. At the same time, the process of constructing the forgetting fuzzy concept space is shown as Algorithm 2, where Ψ is run in Steps 3–11, which can be taken in $O(|G|^2|A|)$, and some unnecessary objects are reduced in Steps 12–25, which can be measured in $O(|\Psi|^2|G|)$, where Ψ means the zone of proximal development of an object. Hence, the time complexity of Algorithm 2 is $O(|\mathcal{S}|(|G|^2|A| + |\Psi|^2|G|))$.

Example 2. Continue with Example 1. With respect to class D_1 , we now discuss the forgetting fuzzy concept of $(\{x_1, x_2, x_3, x_4\}, (0.1, 0.3, 0.4, 0.4))$ in S_1 . In fact, $\Psi = \{\{x_2\}, \{x_1, x_4\}, \{x_4\}\}$, which implies that $\tilde{h}(X) = \{\{x_1, x_2, x_3, x_4\} \setminus \{x_2\}, \{x_1, x_2, x_3, x_4\} \setminus \{x_4\}\} = \{\{x_1, x_3, x_4\}, \{x_1, x_2, x_3\}\}$. Then we have two fuzzy concepts, $(\{x_1, x_3, x_4\}, (0.3, 0.3, 0.4, 0.4))$ and $(\{x_1, x_2, x_3\}, (0.1, 0.3, 0.4, 0.5))$. Subsequently, $K((\{x_1, x_3, x_4\}, (0.3, 0.3, 0.4, 0.4)), (\{x_1, x_2, x_3, x_4\}, (0.1, 0.3, 0.4, 0.4))) = 0.95$ and $K((\{x_1, x_2, x_3\}, (0.1, 0.3, 0.4, 0.5)), (\{x_1, x_2, x_3, x_4\}, (0.1, 0.3, 0.4, 0.4))) = 0.85$. Hence, we know that $\mathcal{F}_X^1 = \{(\{x_1, x_3, x_4\}, (0.3, 0.3, 0.4, 0.4))\}$. Similarly, $\mathcal{F}_Y^1 = \{(\{x_2, x_3\}, (0.1, 0.5, 0.5, 0.7))\}$. Finally, we have the forgetting fuzzy concept subspace $\mathcal{F}^1 = \{(\{x_1, x_3, x_4\}, (0.3, 0.3, 0.4, 0.4)), (\{x_2, x_3\}, (0.1, 0.5, 0.5, 0.7))\}$.

In addition, the forgetting fuzzy concept subspace in D_2 is $\mathcal{F}^2 = \{(\{x_5\}, (0.6, 0.4, 0.3, 0.9)), (\{x_6, x_7, x_8\}, (0.1, 0.5, 0.4, 0.3)), (\{x_6, x_8\}, (0.1, 0.5, 0.6, 0.4))\}$.

Meanwhile, the process of learning the forgetting fuzzy concept is shown in Figure 2.

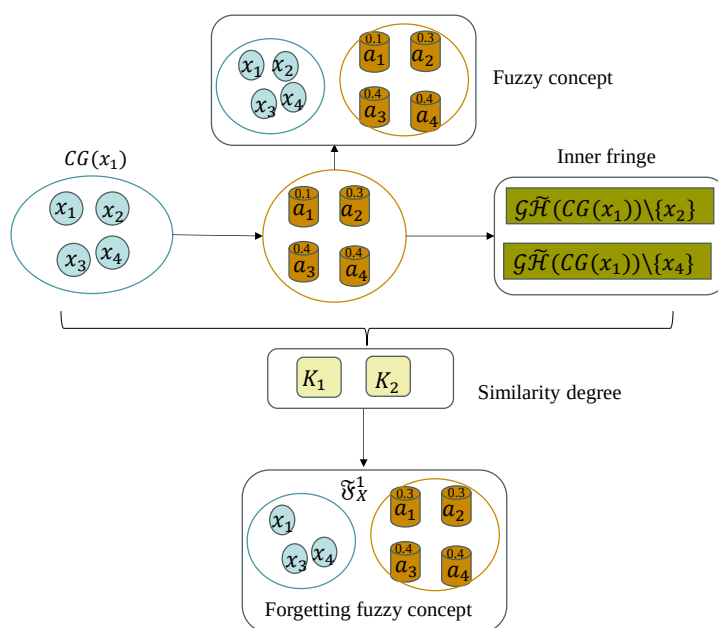


Figure 2. The process of learning the forgetting fuzzy concept.

Algorithm 2: Constructing forgetting fuzzy concept space (CFFCS).

```

1 Input: A FFDC  $(G, A, \tilde{R}, D, J)$ ;
2 Output: Forgetting fuzzy concept space  $\mathcal{F}$ .
3 Compute the decision partition  $G/D = \{D_1, D_2, \dots, D_t\}$ ;
4 for each  $(\mathcal{G}\tilde{H}(X), \tilde{H}(X)) \in \mathcal{S}_i$  do
5   Set  $\mathcal{F}_X^i \leftarrow \emptyset$ ;
6   for each  $a \in A$  do
7     for each  $g \in G$  do
8       if  $\tilde{R}(g, a) = \tilde{H}(X)(a)$  then
9          $L \leftarrow g$ ;
10      end
11    end
12     $\Psi \leftarrow L$ ;
13  end
14  for each  $Y_i \in \Psi$  do
15    for each  $Y_j \in \Psi$  do
16      if  $Y_j \subseteq Y_i$  then
17         $d_{ij} = 1$ ;
18      else
19         $d_{ij} = 0$ ;
20      end
21      if  $\sum_{j=1}^{|\Psi|} d_{ij} = 1$  then
22         $h(X) \leftarrow \mathcal{G}\tilde{H}(X) \setminus Y_i$  and compute the fuzzy concept
           $(\mathcal{G}\tilde{H}(\mathcal{G}\tilde{H}(X) \setminus Y_i), \tilde{H}(\mathcal{G}\tilde{H}(X) \setminus Y_i))$ ;
23        Compute the similarity degree
           $K((\mathcal{G}\tilde{H}(\mathcal{G}\tilde{H}(X) \setminus Y_i), \tilde{H}(\mathcal{G}\tilde{H}(X) \setminus Y_i)), (\mathcal{G}\tilde{H}(X), \tilde{H}(X)))$ ;
24         $\mathcal{F}_X^i \leftarrow \operatorname{argmax}_{V_i \in h(X)} K((\mathcal{G}\tilde{H}(V_i), \tilde{H}(V_i)), (\mathcal{G}\tilde{H}(X), \tilde{H}(X)))$ ;
25      end
26    end
27  end
28   $\mathcal{F}^i \leftarrow \mathcal{F}_X^i$ ;
29 end
30 Return  $\mathcal{F} = \{\mathcal{F}^1, \mathcal{F}^2, \dots, \mathcal{F}^t\}$ .

```

3.3. Fusing Concept

Mutual information exists among forgetting fuzzy concepts, manifested through their bidirectional influence mechanisms. In order to overcome the limitations of individual cognition and the incomplete cognitive environment [28], we propose a new method to fuse concepts through the forgetting fuzzy concept space.

Definition 7. Given an FFDC $(G, A, \tilde{R}, D, J)$, for a forgetting fuzzy concept subspace $\mathcal{F}^i$, if there exist forgetting fuzzy concepts $(X_1, \tilde{B}_1), (X_2, \tilde{B}_2), \dots, (X_n, \tilde{B}_n)$ satisfying $X_1 \subseteq X_2 \subseteq \dots \subseteq X_n$, then $(X_n, \tilde{B}_n)$ is regarded as the supremum fuzzy concept, and then the fusing forgetting fuzzy pseudo-concept $(Y_{i,j}, \tilde{B}_{i,j})$ is defined as follows:

$$\begin{aligned}
 Y_{i,j} &= X_1 \cup X_2 \cup \dots \cup X_n, \\
 \tilde{B}_{i,j} &= \frac{1}{2^n - 1} (\tilde{B}_1 + \tilde{B}_2 + 2\tilde{B}_3 + 4\tilde{B}_4 + \dots + 2^{n-1}\tilde{B}_n).
 \end{aligned}
 \tag{8}$$

where n means the number of forgetting fuzzy concepts.

Generally speaking, the fusing forgetting fuzzy concept space is denoted as $\mathcal{K} = \{\mathcal{K}_1, \mathcal{K}_2, \dots, \mathcal{K}_t\}$, where $\mathcal{K}_i = \{\mathcal{K}_{i,j} | j = 1, 2, \dots, n\} = \{(\mathbf{Y}_{i,j}, \tilde{\mathbf{B}}_{i,j}) | j = 1, 2, \dots, n\}$, in which n is the number of pseudo-concepts in $\mathcal{K}_i$. The intent $\tilde{\mathbf{B}}_{i,j}$ explicitly characterizes the magnitude of the pseudo-concept, in which the intents of subconcepts have been assigned different weights according to their corresponding extents. In other words, the greater the extent X_i is, the larger the weight of its corresponding intent $\tilde{B}_i$ is. Finally, Algorithm 3 describes the clustering process of the fusing forgetting fuzzy concept space with the worst-case time complexity of $O(|G|^2|A|)$.

Algorithm 3: Cognitive process of fusing forgetting fuzzy concept space.

```

1 Input: Forgetting fuzzy concept space  $\mathcal{F}$ ;
2 Output: The pseudo-concept space  $\mathcal{K} = \{\mathcal{K}_1, \mathcal{K}_2, \dots, \mathcal{K}_t\}$ .
3 for each  $\mathcal{F}^i \in \mathcal{F}$  do
4   Set  $P_i = \emptyset$ ,  $\hat{P}_i = \emptyset$ , and  $\hat{S}_i = \emptyset$ ; //Find the supremum concept and sub-concept;
5   for  $(X_j, \tilde{B}_j) \in \mathcal{F}^i$  do
6     for  $(X_k, \tilde{B}_k) \in \mathcal{F}^i$  do
7       Set  $S_{i,j} = \emptyset$ ; //Find the sub-concept of  $(X_j, \tilde{B}_j)$ ;
8       if  $(X_k, \tilde{B}_k)$  is a sub-concept of  $(X_j, \tilde{B}_j)$  then
9          $P_i = P_i \cup (X_j, \tilde{B}_j)$ , and  $S_{i,j} = S_{i,j} \cup \{(X_k, \tilde{B}_k)\}$ ;
10      end
11    end
12  end
13   $\hat{S}_i \leftarrow S_{i,j}$ ;
14  for  $S_{i,m} \in \hat{S}_i$  do
15    if there exists only one concept  $(X_t, \tilde{B}_t)$  in  $P_i$  such that each concept of  $S_{i,m}$  is a
      sub-concept of  $(X_t, \tilde{B}_t)$  then
16       $S_{i,m} = S_{i,m}$  and  $\hat{P}_i \leftarrow \{(X_t, \tilde{B}_t)\}$ ;
17    end
18  end
19  Calculate the fusing forgetting fuzzy pseudo-concept  $(\mathbf{Y}_{i,j}, \tilde{\mathbf{B}}_{i,j})$  from
    Definition 7;
20   $\mathcal{K}_i \leftarrow (\mathbf{Y}_{i,j}, \tilde{\mathbf{B}}_{i,j})$ ;
21 end
22 Return  $\mathcal{K} = \{\mathcal{K}_1, \mathcal{K}_2, \dots, \mathcal{K}_t\}$ .

```

Example 3. Continuing to Example 2, according to Definition 7, fusing forgetting fuzzy concept spaces are represented as follows:

$$\begin{aligned}
 \mathcal{K}_{1,1} &= (\{x_1, x_3, x_4\}, (0.3, 0.3, 0.4, 0.4)); \\
 \mathcal{K}_{1,2} &= (\{x_2, x_3\}, (0.1, 0.5, 0.5, 0.7)); \\
 \mathcal{K}_{2,1} &= (\{x_5\}, (0.6, 0.4, 0.3, 0.9)); \\
 \mathcal{K}_{2,2} &= (\{x_6, x_7, x_8\}, (0.1, 0.5, 0.5, 0.35)).
 \end{aligned}
 \tag{9}$$

It should be noted that there are two fusing forgetting fuzzy concepts in each decision class. It is evident that these concepts preserve the initial information while eliminating the redundant forgetting fuzzy concepts, which significantly enhances the efficiency of concept-cognitive learning.

3.4. Class Prediction

It should be noted that the class prediction of a testing sample is primarily determined based on the Euclidean distance between the testing sample and existing fuzzy concept clustering space [27,29,30].

Definition 8. Let x_{new} be a testing sample. $(x_{new}, \tilde{\mathcal{H}}(x_{new}))$ is a new fuzzy concept; then the distance between $\tilde{\mathcal{H}}(x_{new})$ and pseudo-concept $(\mathbf{Y}_{i,j}, \tilde{\mathbf{B}}_{i,j})$ in $\mathcal{K}_i$ is defined as follows:

$$Dis(\tilde{\mathcal{H}}(x_{new}), \tilde{\mathbf{B}}_{i,j}) = \sqrt{\sum_{a \in A} \left(\tilde{\mathcal{H}}(x_{new})(a) - \tilde{\mathbf{B}}_{i,j}(a) \right)^2}. \quad (10)$$

As is well known, $Dis(\tilde{\mathcal{H}}(x_{new}), \tilde{\mathbf{B}}_{i,j})$ is the distance similarity. The smaller the distance value $Dis(\tilde{\mathcal{H}}(x_{new}), \tilde{\mathbf{B}}_{i,j})$, the stronger the correlation between two fuzzy concepts. Consequently, class determination for a new sample can be achieved by computing their similarity. Algorithm 4 describes the class prediction with the time complexity of $O(|\mathcal{K}| |A|)$.

Algorithm 4: Class prediction of testing sample.

```

1 Input: The pseudo-concept space  $\mathcal{K}$  and the fuzzy membership degree  $\tilde{\mathcal{H}}(x_{new})$ 
   of testing sample  $x_{new}$ ;
2 Output: Class prediction  $L$  of testing sample  $x_{new}$ .
3 for  $\mathcal{K}_i \in \mathcal{K}$  do
4   for  $(\mathbf{Y}_{i,j}, \tilde{\mathbf{B}}_{i,j}) \in \mathcal{K}_i$  do
5     Compute  $Dis(\tilde{\mathcal{H}}(x_{new}), \tilde{\mathbf{B}}_{i,j})$  according to Definition 8 and
        $l_i \leftarrow Dis(\tilde{\mathcal{H}}(x_{new}), \tilde{\mathbf{B}}_{i,j})$ ;
6   end
7 end
8 Have  $L \leftarrow \arg \min(\min(l_i))$ .
9 Return  $L$ .
```

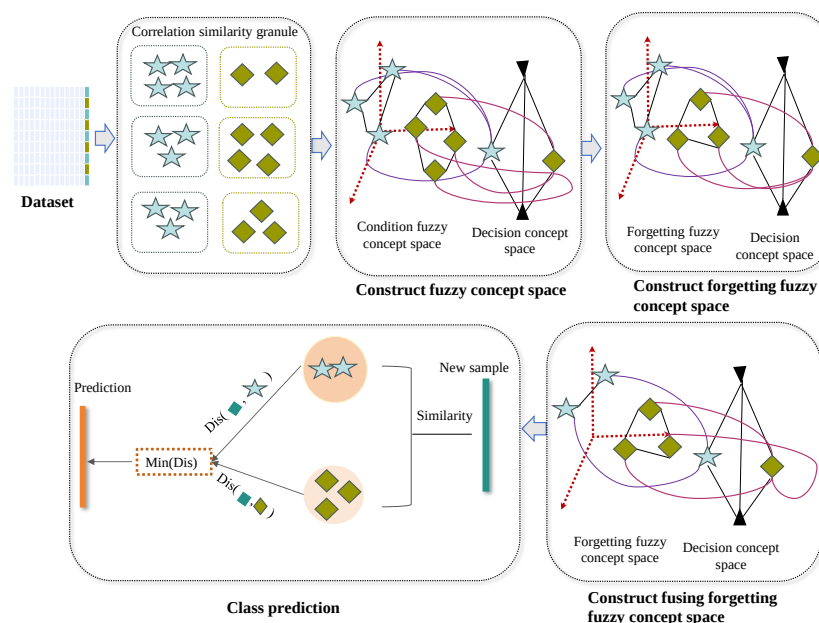
Example 4. In the FFDC of Table 2, where $x_1 - x_8$ are from Example 1, $x_9 - x_{10}$ are two new testing objects.

Regarding the object x_9 , the membership degree about $\tilde{\mathbf{R}}$ is $\tilde{\mathcal{H}}(x_{new}) = (0.6, 0.5, 0.1, 0.6)$, while its true label is 2. Subsequently, the Euclidean distance between x_9 and the existing fusing forgetting fuzzy concept space $\mathcal{K}$ is computed as follows: $Dis(\tilde{\mathcal{H}}(x_{new}), \tilde{\mathbf{B}}_{1,1}) = 0.5099$, $Dis(\tilde{\mathcal{H}}(x_{new}), \tilde{\mathbf{B}}_{1,2}) = 0.6481$, $Dis(\tilde{\mathcal{H}}(x_{new}), \tilde{\mathbf{B}}_{2,1}) = 0.3742$, and $Dis(\tilde{\mathcal{H}}(x_{new}), \tilde{\mathbf{B}}_{2,2}) = 0.6874$. Indeed, the distance between $\tilde{\mathcal{H}}(x_{new})$ and $\tilde{\mathbf{B}}_{2,1}$ attains its minimum in $\mathcal{K}_2$. Consequently, x_9 should be classified into decision class D_2 , which perfectly matches its real label 2.

Based on the above discussion, Figure 3 describes the overall flowchart of the FCCLM, which includes four parts, namely (1) constructing a fuzzy concept space; (2) constructing a forgetting fuzzy concept space; (3) constructing a fusing forgetting fuzzy concept space; and (4) class prediction. In summary, the total time complexity of our model FCCLM is $O(\max\{|\mathcal{S}|(|G|^2|A| + |\Psi|^2|G|), |\mathcal{K}| |A|\})$.

Table 2. A FFDC.

U	a_1	a_2	a_3	a_4	d
x_1	0.7	0.3	0.4	0.5	1
x_2	0.1	0.8	0.5	0.9	1
x_3	0.3	0.5	0.6	0.7	1
x_4	0.4	0.3	0.4	0.4	1
x_5	0.6	0.4	0.3	0.9	2
x_6	0.3	0.5	0.6	0.4	2
x_7	0.2	0.5	0.4	0.3	2
x_8	0.1	0.7	0.8	0.8	2
x_9	0.6	0.5	0.1	0.6	2
x_{10}	0.2	0.6	0.4	0.9	2


Figure 3. The overall procedure of the proposed method.

4. Experimental Results

To evaluate the classification efficacy of model FCCLM, a comparative analysis will be conducted between it and concept-cognitive learning as well as machine learning classification algorithms. A total of eight datasets are selected from the UCI (dataset source: <http://archive.ics.uci.edu/>, accessed on 14 March 2015) and Gene (Dataset source: <https://jundongli.github.io/scikit-feature/datasets.html>, accessed on 14 March 2015) data set repositories, as detailed in Table 3.

Table 3. Data description.

ID	Dataset	Object	Attribute	Class
1	Appendicitis	106	7	2
2	Dermatology	358	34	6
3	Movement	360	90	15
4	Air	359	64	3
5	Austra	690	14	2
6	Sonar	208	60	2
7	Wdbc	569	30	2
8	Wine	178	13	3

4.1. Experimental Setting

As a preprocessing step, Max–Min normalization is applied to all attributes to guarantee dataset compatibility with fuzzy processing requirements, specifically given by

$$\tilde{R}(u_i, a_j) = \frac{f(u_i, a_j) - \min(f(a_j))}{\max(f(a_j)) - \min(f(a_j))},$$

where $f(u_i, a_j)$ is the initial value of object u_i for attribute a_j . Additionally, $\max(f(a_j))$ and $\min(f(a_j))$ are the maximum and minimum values of attribute a_j across all objects, respectively.

In the experiment, we compare FCCLM with three concept-cognitive learning classification algorithms, namely, DMPWFC [23], ILMPFTC [24], and FCLM [21]. Furthermore, we also compare FCCLM with four machine learning classification algorithms, including Complex Tree (CT), Fuzzy-Root-Sum-Square (F-RSS), Decision Tree (DT), and Fuzzy Classification and Regression Tree (F-CART). Notice that parameter β plays an important role in constructing correlation similarity granules. In this investigation, parameter β is varied systematically from 0.55 to 1.0 with a step size of 0.05 to influence correlation similarity granules. All analyses are conducted using five-fold cross-validation to guarantee unbiased performance estimation. Final performance metrics are computed as the mean values obtained from the five cross-validation test partitions. All results are generated in MATLAB 2021a using standardized hardware (Intel(R) Core(TM) i7-4790 CPU @ 3.6 GHz, 16 GB (Lenovo, Quanzhou, China)) to ensure computational consistency.

4.2. Results and Analyses

The results of accuracy and the optimal parameter β of the proposed algorithm FCCLM and seven classification algorithms on eight datasets are shown in Table 4. The bottom row displays the mean classification accuracy with the standard deviation, where bold–underline formatting highlights the algorithm’s statistically superior performance compared to all baseline methods. The results of the table indicate that ILMPFTC and CT achieve optimal accuracy on one dataset, respectively, while FCCLM demonstrates superior performance across six datasets. Furthermore, FCCLM has a higher average accuracy and lower standard deviation than the other algorithms, which indicates that the performance of FCCLM outperforms the other seven algorithms. The average accuracy of FCCLM is increased by 9.48% when compared to F-CART on all selected datasets. In summary, the above results indicate that FCCLM is superior to the other models. Figure 4 depicts the accuracy comparison between CCL and the other seven algorithms. This figure describes the accuracy of FCCLM, which possesses a smaller fluctuation range than the other algorithms in most datasets.

Table 4. Accuracy comparison (mean $\pm$ standard deviation%) of FCCLM and seven other algorithms.

ID	β	FCCLM	DMPWFC	ILMPFTC	FCLM	CT	F-RSS	DT	F-CART
1	0.80	86.75 $\pm$ 5.29	73.14 $\pm$ 12.03	57.80 $\pm$ 23.06	63.76 $\pm$ 19.02	83.89 $\pm$ 5.56	86.70 $\pm$ 8.85	74.50 $\pm$ 2.84	83.89 $\pm$ 4.43
2	0.95	96.36 $\pm$ 1.16	95.16 $\pm$ 2.29	95.92 $\pm$ 2.68	40.52 $\pm$ 3.00	94.68 $\pm$ 2.31	53.93 $\pm$ 10.06	61.77 $\pm$ 7.97	94.13 $\pm$ 2.09
3	1.00	81.67 $\pm$ 3.85	82.76 $\pm$ 4.86	87.29 $\pm$ 5.08	66.69 $\pm$ 3.23	61.67 $\pm$ 5.07	38.33 $\pm$ 4.77	14.44 $\pm$ 5.07	57.70 $\pm$ 11.31
4	1.00	95.83 $\pm$ 3.11	95.56 $\pm$ 1.95	82.61 $\pm$ 4.80	54.49 $\pm$ 5.80	85.52 $\pm$ 3.59	45.41 $\pm$ 7.73	50.09 $\pm$ 10.52	80.50 $\pm$ 3.83
5	1.00	80.14 $\pm$ 3.53	84.53 $\pm$ 3.32	53.05 $\pm$ 3.97	65.23 $\pm$ 4.54	84.78 $\pm$ 3.12	78.99 $\pm$ 1.85	72.46 $\pm$ 6.62	84.20 $\pm$ 4.45
6	1.00	86.09 $\pm$ 7.80	60.89 $\pm$ 14.58	82.57 $\pm$ 7.54	66.79 $\pm$ 12.77	71.18 $\pm$ 12.17	68.78 $\pm$ 4.86	65.82 $\pm$ 10.01	72.13 $\pm$ 10.35
7	1.00	93.32 $\pm$ 1.01	41.71 $\pm$ 16.69	83.07 $\pm$ 3.70	70.65 $\pm$ 1.15	78.09 $\pm$ 1.47	77.72 $\pm$ 0.90	88.93 $\pm$ 4.09	92.79 $\pm$ 1.34
8	0.95	96.63 $\pm$ 1.23	27.16 $\pm$ 37.90	89.05 $\pm$ 6.40	68.87 $\pm$ 26.22	92.67 $\pm$ 5.83	93.84 $\pm$ 2.29	66.81 $\pm$ 7.83	89.35 $\pm$ 6.61
Ave. $\pm$ SD		89.60 $\pm$ 3.37	70.11 $\pm$ 11.70	78.92 $\pm$ 7.15	62.13 $\pm$ 9.47	81.56 $\pm$ 4.89	67.96 $\pm$ 5.16	61.85 $\pm$ 6.87	81.84 $\pm$ 5.55
Rank		1.63	4.88	4.25	6.38	3.56	5.25	6.13	3.94
Win/tie/loss		6/0/8	0/0/8	1/0/8	0/0/8	1/0/8	0/0/8	0/0/8	0/0/8

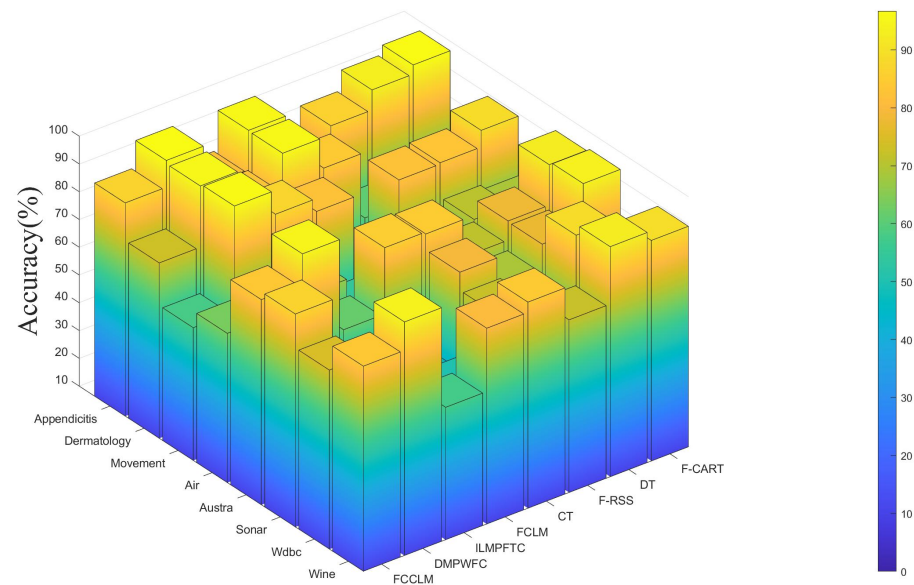


Figure 4. Accuracy comparison with CCL and other classification algorithms.

Subsequently, the results of recall for the proposed algorithm FCCLM and seven classification algorithms on eight datasets are presented in Table 5. The last row illustrates the average accuracy and standard deviation, with bold–underline formatting denoting statistically superior performance relative to comparator algorithms. This table indicates that ILMPFTC, CT, and F-CART achieve the best accuracy twice, once, and once, respectively, while FCCLM has the best performance on four datasets. Furthermore, FCCLM demonstrates statistically superior performance, exhibiting both a significantly higher average accuracy and lower standard deviation compared to the other algorithms, which illustrates that the performance of FCCLM outperforms the other seven algorithms.

Table 5. Recall comparison (mean ± standard deviation%) of FCCLM and seven other algorithms.

ID	β	FCCLM	DMPWFC	ILMPFTC	FCLM	CT	F-RSS	DT	F-CART
1	0.80	75.36 ± 11.83	74.71 ± 24.04	57.13 ± 26.77	63.33 ± 27.71	65.21 ± 9.71	71.00 ± 19.64	67.02 ± 8.87	68.96 ± 8.78
2	0.95	95.83 ± 2.08	92.58 ± 11.35	95.41 ± 9.56	27.72 ± 35.27	94.03 ± 2.81	36.76 ± 3.98	-	92.42 ± 1.92
3	1.00	80.43 ± 4.79	58.54 ± 39.32	84.29 ± 22.15	68.69 ± 18.01	62.05 ± 5.15	44.73 ± 2.23	-	67.70 ± 16.54
4	1.00	95.54 ± 3.46	95.29 ± 7.05	80.90 ± 10.11	56.12 ± 30.88	85.35 ± 3.75	38.65 ± 3.28	-	80.23 ± 4.47
5	1.00	79.62 ± 4.02	83.85 ± 3.32	54.40 ± 14.09	64.76 ± 18.61	84.75 ± 3.35	76.89 ± 1.06	62.36 ± 13.42	83.92 ± 3.94
6	1.00	85.86 ± 7.84	-	87.86 ± 6.04	77.77 ± 17.59	71.60 ± 13.00	66.81 ± 4.94	-	71.77 ± 10.35
7	1.00	90.90 ± 2.26	-	84.82 ± 10.02	76.85 ± 31.02	91.92 ± 2.80	82.58 ± 3.28	-	92.58 ± 1.37
8	0.95	97.19 ± 0.94	-	90.33 ± 11.09	33.33 ± 57.74	93.20 ± 5.42	93.95 ± 1.70	56.37 ± 14.56	88.28 ± 7.52
Ave. ± SD		87.59 ± 4.65	80.99 ± 17.01	79.39 ± 13.73	58.57 ± 29.60	81.01 ± 5.75	63.92 ± 5.01	61.91 ± 12.28	80.73 ± 6.86

Figure 5 describes the classification accuracy as β changes. From this figure, we can observe that the accuracy shows a slow upward trend with fluctuations as the parameter varies, among which the fluctuation range is the largest in the interval [0.75, 0.85]. At the same time, most datasets achieve optimal accuracy in the range of [0.85, 0.95].

Additionally, the statistical significance of the eight classification algorithms could be compared using the Friedman test [31] and Bonferoni–Dunn test [32]. For the Friedman test, a Fisher distribution F_F measuring the performance of different algorithms is given by

$$F_F = \frac{(N-1)\chi_F^2}{N(k-1) - \chi_F^2} \sim F(k-1, k-1(N-1)) \quad (11)$$

where $\chi_F^2 = \frac{12N}{k(k+1)} (\sum_{i=1}^k R_i^2 - \frac{k(k+1)^2}{4})$. N is the number of datasets, k is the cardinality of different algorithms, and R_i is the average rank of algorithm R_i across all datasets. We first assume the null hypothesis that no statistically significant differences exist in predictive accuracy among the compared classification methods. If $F_F > F(k-1, k-1(N-1))$, then the initial hypothesis will be rejected. The rank result of the eight models is represented in Table 6. In fact, $F_F = 4.5634$, and the critical value of $F(7, 49) = 2.203$ at level $\alpha = 0.05$. Therefore, we reject the null hypothesis and accept the alternative hypothesis that the classification performance of the eight models is remarkably different.

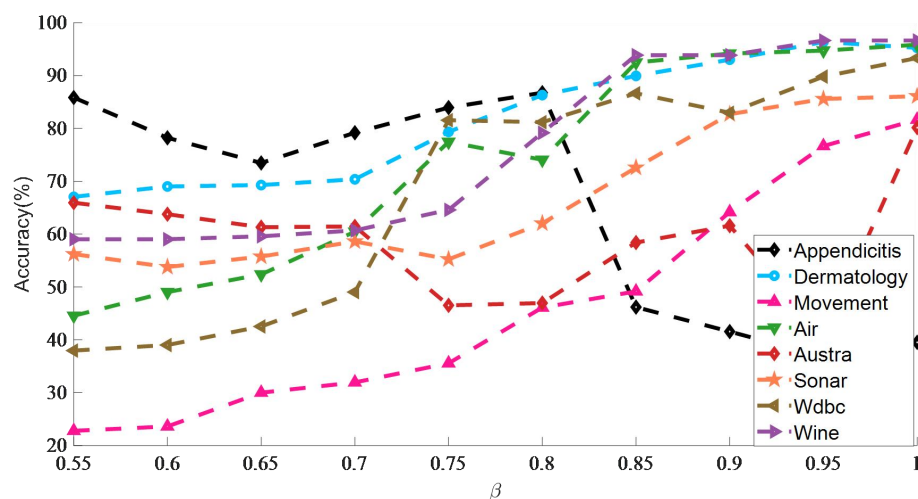


Figure 5. Accuracy comparisons with varying β on eight datasets.

Table 6. Rank of classification algorithms.

ID	FCCLM	DMPWFC	ILMPFTC	FCLM	CT	F-RSS	DT	F-CART
1	1	6	8	7	3.5	2	5	3.5
2	1	3	2	8	4	7	6	5
3	3	2	1	4	5	7	8	6
4	1	2	4	6	3	8	7	5
5	4	2	8	7	1	5	6	3
6	1	8	2	6	4	5	7	3
7	1	8	4	7	5	6	3	2
8	1	8	5	6	3	2	7	4
Average	1.63	4.88	4.25	6.38	3.56	5.25	6.13	3.94

Subsequently, the Bonferroni–Dunn test is applied to conduct pairwise comparisons of classification performance among the five top-performing models. Assume that the mean rank between two models exceeds the critical value

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

and then the above two models exhibit statistically significant differences in classification performance, in which q_{α} is denoted as the critical value in the test.

At level $\alpha = 0.05$, we notice that the critical value $q_{0.05} = 2.690$ in [32]. Subsequently, $CD_{\alpha} = 3.2946$ ($k = 8$ and $N = 8$). Therefore, we can see from Figure 6 that algorithm FCCLM outperforms algorithms F-RSS, FCLM, and DT at level $\alpha = 0.05$.

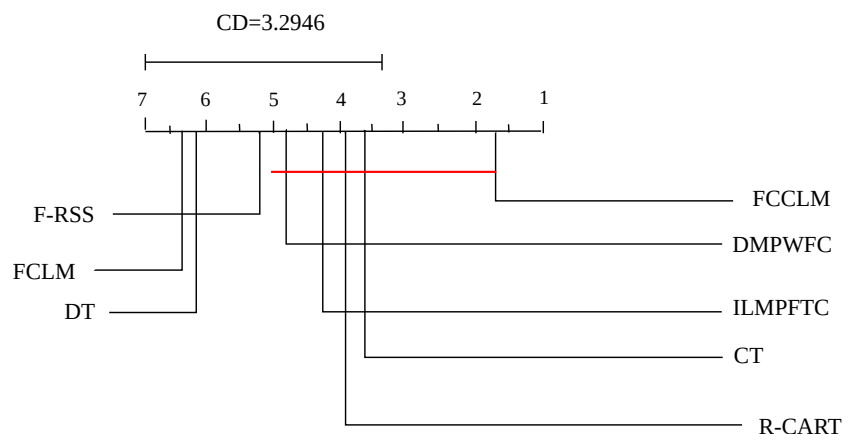


Figure 6. CD comparison of all classification algorithms with Bonferroni–Dunn test ($\alpha = 0.05$).

4.3. Discussion

In the above subsection, we compare the performance of FCCLM with other algorithms using eight datasets. Experimental results demonstrate the effectiveness and efficiency of our proposed algorithm FCCLM from the perspective of accuracy and statistical results. Specifically, some advantages are reflected on as follows: (1) We could counterpoise the sensitivity and specificity of Algorithm 1, named CFCs, to capture rich information by adjusting the threshold, thus facilitating the interpretability and understanding of the concept learning process. (2) Knowledge forgetting of the learned fuzzy concepts is performed through the inner fringe, thereby aligning with the logic of the Ebbinghaus forgetting curve and enhancing interpretability. (3) In the classification task, the fusing forgetting fuzzy pseudo-concept and class prediction can accelerate the acquisition of sample labels and reduce the time consumption of model FCCLM. In summary, a novel forgetting-based concept-cognitive learning model for classification in a fuzzy formal decision context is introduced to enhance the cognitive process.

5. Conclusions

In the context of rapidly expanding data, how to mine knowledge and fuse information is an important issue. Some existing CCL methods in the fuzzy formal decision context mainly focus on forgetting unnecessary concepts in constructing the concept space; ignoring the object is the key to constructing concepts. In this article, we have proposed a forgetting-based concept-cognitive learning model for classification tasks in a fuzzy formal decision context. Concretely, we firstly construct the fuzzy concept space with the correlation coefficient matrix. After that, in order to eliminate unnecessary objects in the zone of proximal development, we build a forgetting fuzzy concept space by choosing the concept with the highest similarity. Next, we propose a forgetting-based fuzzy concept model, which handles concept fusion and concept class prediction. Finally, to gain a better understanding of our proposed model, a series of comparative experiments on eight datasets are conducted to show that FCCLM can attain superior classification performance.

This article introduces a forgetting mechanism in the process of concept-cognitive learning, but does not consider recognition in the incremental learning process. Our future research will address this issue with the aim of enhancing the accuracy and efficiency of incremental learning approaches in the context of CCL.

Author Contributions: Conceptualization: X.L.; Software: C.Z.; Writing—original draft: C.S. All authors have read and agreed to the published version of the manuscript.

Funding: Work is supported by the Natural Science Foundation of Fujian Province (2024J01793) and the Natural Science Basic Research Program of Shaanxi Province (Grant No. 2024JC-YBQN-0025).

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Conflicts of 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.

References

1. Pylyshyn, Z. Computation and cognition: Issues in the foundations of cognitive science. *Behav. Brain Sci.* **1980**, *3*, 111–132. [\[CrossRef\]](#)
2. Wang, Y.; Howard, N.; Kacprzyk, J.; Frieder, O.; Sheu, P.; Fiorini, R.A.; Gavrilova, M.L.; Patel, S.; Peng, J.; Widrow, B. Cognitive informatics: Towards cognitive machine learning and autonomous knowledge manipulation. *Int. J. Cogn. Inform. Nat. Intell.* **2018**, *12*, 1–13. [\[CrossRef\]](#)
3. Feldman, J. Minimization of Boolean complexity in human concept learning. *Nature* **2000**, *407*, 630–633. [\[CrossRef\]](#) [\[PubMed\]](#)
4. Ganter, B.; Wille, R. *Formal Concept Analysis: Mathematical Foundations*; Springer: New York, NY, USA, 1999.
5. Wille, R. Restructuring lattice theory: An approach based on hierarchies of concepts. In Proceedings of the Formal Concept Analysis: 7th International Conference, ICFCA 2009, Darmstadt, Germany, 21–24 May 2009; Proceedings 7; Springer: Berlin/Heidelberg, Germany, 2009; pp. 314–339.
6. Yahia, S.; Arour, K.; Slimani, A.; Jaoua, A. Discovery of compact rules in relational databases. *Inf. Sci.* **2000**, *4*, 497–511.
7. Qi, J.; Qian, T.; Wei, L. The connections between three-way and classical concept lattices. *Knowl.-Based Syst.* **2016**, *91*, 143–151. [\[CrossRef\]](#)
8. Xu, F.; Xing, Z.; Yin, H. Attribute reductions and concept lattices in interval-valued intuitionistic fuzzy rough set theory: Construction and properties. *J. Intell. Fuzzy Syst.* **2016**, *30*, 1231–1242. [\[CrossRef\]](#)
9. Qi, J.; Wei, L.; Wan, Q. Multi-level granularity in formal concept analysis. *Granul. Comput.* **2019**, *4*, 351–362. [\[CrossRef\]](#)
10. Basil, K. Papadopoulos and apostolos syropoulos, fuzzy sets and fuzzy relational structures as Chu spaces. *Int. J. Uncertain. Fuzziness Knowl.-Based Syst.* **2000**, *8*, 471–479.
11. Papadopoulos, B.; Syropoulos, A. Categorical relationships between Goguen sets and “two-sided” categorical models of linear logic. *Fuzzy Sets Syst.* **2005**, *149*, 501–508. [\[CrossRef\]](#)
12. Zhang, W.; Xu, W. Cognitive model based on granular computing. *Chin. J. Eng. Math.* **2007**, *24*, 957–971.
13. Yao, Y. Interpreting concept learning in cognitive informatics and granular computing. *IEEE Trans. Syst. Man Cybern. Part B (Cybern.)* **2009**, *39*, 855–866. [\[CrossRef\]](#)
14. Qiu, G.; Ma, J.; Yang, H.; Zhang, W. A mathematical model for concept granular computing systems. *Sci. China Inf. Sci.* **2009**, *39*, 1239–1247. [\[CrossRef\]](#)
15. Kumar, C.; Ishwarya, M.; Loo, C. Formal concept analysis approach to cognitive functionalities of bidirectional associative memory. *Biol. Inspired Cogn. Archit.* **2015**, *12*, 20–33.
16. Li, J.; Mei, C.; Xu, W.; Qian, Y. Concept learning via granular computing: A cognitive viewpoint. *Inf. Sci.* **2015**, *298*, 447–467. [\[CrossRef\]](#) [\[PubMed\]](#)
17. Xu, W.; Li, W. Granular computing approach to two-way learning based on formal concept analysis in fuzzy datasets. *IEEE Trans. Cybern.* **2014**, *46*, 366–379. [\[CrossRef\]](#)
18. Xu, W.; Guo, D.; Qian, Y.; Ding, W. Two-way concept-cognitive learning method: A fuzzy-based progressive learning. *IEEE Trans. Fuzzy Syst.* **2022**, *31*, 1885–1899. [\[CrossRef\]](#)
19. Xu, W.; Guo, D.; Mi, J.; Qian, Y.; Zheng, K.; Ding, W. Two-way concept-cognitive learning via concept movement viewpoint. *IEEE Trans. Neural Netw. Learn. Syst.* **2023**, *34*, 6798–6812. [\[CrossRef\]](#)
20. Shi, Y.; Mi, Y.; Li, J.; Liu, W. Concept-cognitive learning model for incremental concept learning. *IEEE Trans. Syst. Man Cybern. Syst.* **2018**, *51*, 809–821. [\[CrossRef\]](#)
21. Mi, Y.; Shi, Y.; Li, J.; Liu, W.; Yan, M. Fuzzy-based concept learning method: Exploiting data with fuzzy conceptual clustering. *IEEE Trans. Cybern.* **2020**, *52*, 582–593. [\[CrossRef\]](#)
22. Shi, Y.; Mi, Y.; Li, J.; Liu, W. Concurrent concept-cognitive learning model for classification. *Inf. Sci.* **2019**, *496*, 65–81. [\[CrossRef\]](#)
23. Yuan, K.; Xu, W.; Li, W.; Ding, W. An incremental learning mechanism for object classification based on progressive fuzzy three-way concept. *Inf. Sci.* **2022**, *584*, 127–147. [\[CrossRef\]](#)
24. Zhang, C.; Tsang, E.C.; Xu, W.; Lin, Y.; Yang, L. Incremental concept-cognitive learning approach for concept classification oriented to weighted fuzzy concepts. *Knowl.-Based Syst.* **2023**, *260*, 110093. [\[CrossRef\]](#)
25. Wang, J.; Xu, W.; Ding, W.; Qian, Y. Multi-view fuzzy concept-cognitive learning with high-order information fusion of fuzzy attributes. *IEEE Trans. Fuzzy Syst.* **2024**, *32*, 6965–6978. [\[CrossRef\]](#)

26. Liang, T.; Lin, Y.; Li, J.; Lin, G.; Wang, Q. Incremental cognitive learning approach based on concept reduction. *Int. J. Approx. Reason.* **2025**, *179*, 109359. [\[CrossRef\]](#)
27. Guo, D.; Xu, W.; Qian, Y.; Ding, W. M-FCCL: Memory-based concept-cognitive learning for dynamic fuzzy data classification and knowledge fusion. *Inf. Fusion* **2023**, *100*, 101962. [\[CrossRef\]](#)
28. Li, J.; Mi, Y.; Liu, W. Incremental cognition of concepts: Theories and methods. *Chin. J. Comput.* **2019**, *42*, 1–19.
29. Ding, Y.; Xu, W.; Ding, W.; Qian, Y. IFCRL: Interval-intent fuzzy concept re-cognition learning model. *IEEE Trans. Fuzzy Syst.* **2024**, *32*, 3581–3593. [\[CrossRef\]](#)
30. Guo, D.; Xu, W. Fuzzy-based concept-cognitive learning: An investigation of novel approach to tumor diagnosis analysis. *Inf. Sci.* **2023**, *639*, 118998. [\[CrossRef\]](#)
31. Friedman, M. A comparison of alternative tests of significance for the problem of m rankings. *Ann. Math. Stat.* **1940**, *11*, 86–92. [\[CrossRef\]](#)
32. Dunn, O. Multiple comparisons among means. *J. Am. Stat. Assoc.* **1961**, *56*, 52–64. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.