Information Sciences 625 (2023) 593–619

Contents lists available at ScienceDirect

Information Sciences

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

Incremental feature selection approach to interval-valued fuzzy decision information systems based on λ-fuzzy similarity self-information

Xiaoyan Zhang *, Jirong Li

College of Artificial Intelligence, Southwest University, Chongqing 400715, PR China

ARTICLE INFO

Article history:
Received 8 March 2022
Received in revised form 4 January 2023
Accepted 6 January 2023
Available online 11 January 2023

Keywords:
Dynamic object set
Fuzzy similarity relation
Incremental feature selection
Interval-valued fuzzy rough set
Self-information

ABSTRACT

The relative decision self-information is a crucial evaluation function of feature selection in information system. It encapsulates classification information in upper and lower approximations and pays attention to the boundary region of samples. Nevertheless, with the frequent replacement of data, the static feature selection neglects the previous information of samples, which diminishes the computational efficiency. With the purpose of adapting to the evolution of the era, incremental learning is widely exerted in the field of data mining. In combination with incremental technique, it is not cumbersome to update the reduct in time. Enlightened by this, our work focuses on the mechanism of incremental feature selection due to the variation of objects in IvFDIS. Firstly, we construct λ-fuzzy similarity relation and introduce λ-fuzzy similarity self-information into IvFDIS based on relative decision self-information. Besides, with the assistance of matrix operation, we recommend static feature selection according to λ-fuzzy similarity self-information. Furthermore, two relevant incremental algorithms involving the insertion and removal of objects in IvFDIS are made a research. Finally, some comparative experiments are conducted on twelve public data sets to certify the validity of our incremental algorithms. Experimental results show that comparable to three tested algorithms, the proposed incremental algorithms lessen the computation time greatly, and they select fewer features in most instances without decreasing classification accuracy in IvFDIS.

1. Introduction

The era of big data makes abundant data within reach. Large amounts of data are sometimes mixed with imprecise and ambiguous information. Single-valued data can hardly express the characteristics of things precisely, while interval-valued data better describe the uncertainty in the form of intervals. As an extension of single-valued data, interval-valued data characterize knowledge through interval values with a certain range, which are broadly applied in reality [1–4].

Fuzzy set theory was proposed by Zadeh [5] in 1965, which expressed the concept of fuzziness by means of appropriate membership function. Rough set theory was put forward by Pawlak [6,7] for the first time in 1982, providing an effective tool for dealing with uncertain data. In 1990, Dubois and Prade [8] developed the concept of fuzzy rough set by combining fuzzy set and rough set, which is capable of handling continuous data instead of being restricted to discrete data. Fuzzy similarity

* Corresponding author.
E-mail address: zxy19790915@163.com (X. Zhang).

https://doi.org/10.1016/j.ins.2023.01.058
0020-0255/

relation is extremely important for fuzzy rough set, which describes the similarity degree between different objects. On the basis of fuzzy similarity relation, the notion of fuzzy upper and lower approximations can be obtained, which describe continuous attribute values. Afterwards, a series of improved fuzzy rough set models were discussed successively [9–13], and fuzzy rough set was applied in pattern recognition [14], decision analysis [15], machine learning [16] and so on. Fuzzy rough set theory has been gradually perfected by scholars. Sun et al. [17] came up with the interval-valued fuzzy rough set model based on the interval-valued fuzzy information systems. Xu et al. [18] established two types of fuzzy rough sets models on tolerance relations and raised the optimistic and pessimistic multi-granulation fuzzy rough sets with the viewpoint of granular computing. Wang and Hu [19] developed the granular variable precision fuzzy rough sets with general fuzzy relations and defined the equivalent expressions of approximation operators.

High-dimensional data is ubiquitous in daily life. In machine learning, some attributes may be irrelevant and contain redundant information which is supposed to be deleted. It is these inessential information that enlarges the computation amount and reduces the classification accuracy. For the sake of avoiding dimension disaster, it is necessary to reduce the dimension of data while guaranteeing the fundamental information of original data, namely picking out significant attribute subset from multiple attributes by feature selection approach. With the gradual progress of technology, a variety of feature selection methods have been born [20–22], among which the method based on fuzzy rough set has attracted extensive attention. The inconsistent fuzzy decision system was defined, and discernibility matrix-based algorithms were advanced to find reducts by He et al. [23]. Lin et al. [24] used different fuzzy relations to measure the similarity between samples under different labels, then proposed a fuzzy rough set model for attribute reduction in multi-label learning. A feature selection approach based on fuzzy neighborhood multigranulation rough set in neighborhood decision systems was studied by Sun et al. [25].

Setting up an evaluation function is a vital procedure in feature selection, which is used to measure the classification ability of attribute subsets and determines the classification accuracy of attribute subsets. In recent years, plenty of researchers have constructed many feature evaluation functions. Wang et al. [26] introduced distance measures into fuzzy rough set and researched an iterative computation model based on a variable distance parameter. Employing the entropy to measure the uncertainty, Zhang et al. [27] used the information entropy for feature selection. A label distribution feature selection algorithm using mutual information was brought up by Qian et al. [28]. Information is generated and transmitted in a variety of forms. As the founder of information theory, these concepts of information entropy, mutual information, and self-information were given by Shannon [29,30]. Bringing information theory into fuzzy rough set, Wang et al. [31] constructed four kinds of uncertainty measures by combining fuzzy rough approximations with self-information. However, the notion of self-information has not been applied to the interval-valued information system. Motivated by this, we will introduce the self-information into the interval-valued information system, and utilize λ-fuzzy similarity self-information as the evaluation function of feature selection to estimate the classification ability of attribute subsets in IvFDIS.

As the latest data enters the information system, outdated and redundant data ought to be removed in a timely manner. In the process of data updating, repeated feature selection over and over again will increase calculation amount and consume a great deal of time. Taking advantage of the previous knowledge, the incremental feature selection method may select features from dynamic data and eliminate the need for repeated calculation based on the reduct at the last point in time. In recent years, incremental feature selection has engaged the attention of numerous scholars [32–44].Thereinto, Ni et al. [38] proposed incremental mechanisms of information measure by analyzing the basic concepts of fuzzy rough set on incremental datasets. Considering data with a preference-order relation, a matrix-based method was adopted to study incremental heterogeneous feature selection based on neighborhood rough set by Sang et al. [39]. What is more, Sang et al. [40] investigated incremental feature selection approaches based on a fuzzy dominance neighborhood rough set for dynamic interval-valued ordered data. In order to improve the classification accuracy and speed up the calculation time, we attempt to research the incremental feature selection approach based on λ-fuzzy similarity self-information in IvFDIS.

In this paper, some research on two incremental feature selection approaches based on λ-fuzzy similarity self-information are made in dynamic IvFDIS with time-evolving objects. As shown in Fig. 1, the main contributions of our work are as follows:

- A new relation is constructed named λ-fuzzy similarity relation, then the λ-fuzzy lower and upper approximations are defined. In terms of λ-fuzzy similarity relation, we present a novel model of fuzzy rough set in IvFDIS.
- The uncertainty metrics called λ-fuzzy similarity self-information based on λ-fuzzy similarity rough set is recommended. According to this evaluation function, a static feature selection algorithm is designed in IvFDIS.
- Based on static feature selection algorithm, two incremental feature selection algorithms with respect to the variation of object set are further analyzed, including deleting or adding multiple objects in IvFDIS.
- Some comparative experiments are performed on twelve data sets. Experimental results show that the proposed incremental feature selection approach is highly efficient and the classification of selected features are considerable.

The arrangement of remanent paper is as following statements. To promote subsequent comprehension, Section 2 starts with these introductions of some requisite and foundational knowledge about fuzzy rough set, interval-valued fuzzy decision information system and self-information. In Section 3, a new binary relation termed λ-fuzzy similarity relation is defined, whereafter an λ-fuzzy similarity rough set is constructed in IvFDIS. In the light of the above, the concept of λ-fuzzy similarity self-information (λ-FSSi) is presented. Section 4 explores the matrix operation of λ-FSSi and proposes the

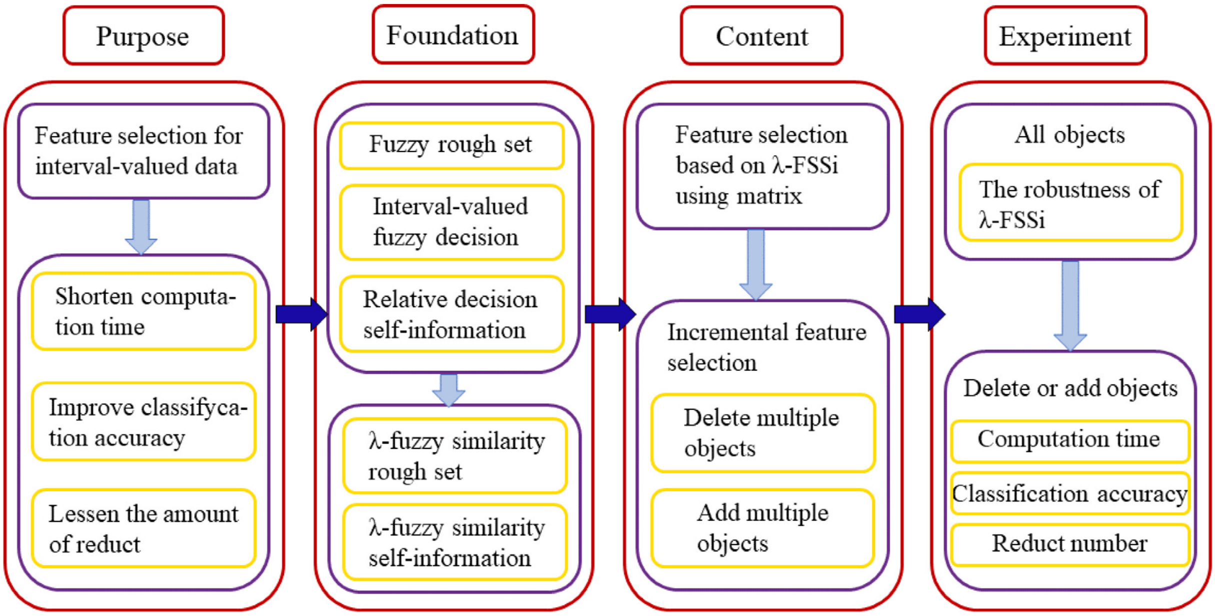


Fig. 1. The framework of our work.

corresponding feature selection algorithm. In Section 5, two types of incremental feature selection algorithms caused by the alteration of objects are studied in dynamic IvFDIS. The final experimental data are listed and the experimental results are plotted in Section 6. Eventually, Section 7 ends with a retrospection of the whole text and prospects for the future.

2. Preliminaries

To make the sequel of our essay easier to comprehend, we will make an introduction of some notions concerning fuzzy rough set, self-information and interval-valued fuzzy decision information systems in this section.

2.1. Fuzzy rough set

Equipped with these typical characteristics of rough set and fuzzy set concurrently, fuzzy rough set came into being. It is a influential instrument for tackling the uncertainty and ambiguity, which copes with real data rather than just discrete data effectively. The following are some statements of fuzzy rough set.

Given a mapping A over the universe $U, A: U \to [0,1],\ u \mapsto A(U) \in [0,1]$, then A is referred as the fuzzy set over U and $A(u)$ is referred as the membership degree of u to A. The whole fuzzy set over U is marked as $\mathscr{F}(U)$, namely $\mathscr{F}(U) = \{A | A: U \to [0,1]\}$.

Proposition 1. (See [8]) *Let U be a universe with finite elements, then a fuzzy similarity relation R over U is required to conform with the following properties:*

(1) *Reflexivity:* $R(x,x) = 1$,
(2) *Symmetry:* $R(x,y) = R(y,x)$.

Taking the fuzzy similarity relation as a starting point, Dubois and Prade [8] constructed the fuzzy rough set for the first time.

Definition 1. (See [8]) Let R be a fuzzy similarity relation over U. For any $A \in \mathscr{F}(U)$ and $x \in U$, the fuzzy lower and upper approximations of A concerning R are defined as

$$\begin{aligned} \underline{R}A(x) &= \inf_{y \in U} \max\{1 - R(x,y), A(y)\}, \\ \overline{R}A(x) &= \sup_{y \in U} \min\{R(x,y), A(y)\}, \end{aligned} \tag{1}$$

where $\underline{R}A(x)$ indicates the certain degree that object x belongs to A, and $\overline{R}A(x)$ indicates the possible degree that object x belongs to A. For any $y \in U$, the fuzzy similarity class of x induced by R is formulated as $[x]_R = R(x,y)$.

2.2. Interval-valued fuzzy decision information system

A quadruple (U,A,V,f) is specified as an information system (IS) if the non-empty finite object set $U = \{x_1,x_2,\cdots,x_n\}$, the non-empty finite attribute set $A = \{a_1,a_2,\cdots,a_m\}$, the domain of attribute set $V = \underset{a\in A}{\cup} V_a$, and the information function $f : U \times A \rightarrow V$. Noticeably, we describe each object over U by attribute set. When the attribute set consists of condition attributes C and decision attributes D, namely $A = C \cup D$ and $C \cap D = \varnothing$, $(U, C \cup D, V, f)$ is termed as decision information system (DIS). In particular, if D is fuzzy decision marked by $\widetilde{D}$, then $(U, C \cup \widetilde{D}, V, f)$ is termed as fuzzy decision information system (FDIS). Provided that the value of object x under attribute a is an interval number for each $x \in U, a \in C$, namely $f(x,a) = \{[a(x)^L, a(x)^U] | a(x)^L, a(x)^U \in R,\ a(x)^L \leqslant a(x)^U\}$, then $(U, C \cup \widetilde{D}, V, f)$ is stipulated as interval-valued fuzzy decision information system (IvFDIS).

Definition 2. (See [42]) In any $IvFDIS = (U, C \cup \widetilde{D}, V, f)$, for any $x_i, x_j \in U$ and $a \in C$, if two interval numbers $f(x_i,a) = [v^L, v^U]$ and $f(x_j,a) = [w^L, w^U]$, the similarity degree between x_i and x_j regarding attribute a is specified as

$$\delta_{ij}^a = \frac{l(f(x_i,a) \cap f(x_j,a))}{l(f(x_i,a) \cup f(x_j,a))}, \tag{2}$$

where l signifies the length of interval, $\cap$ and $\cup$ represent the intersection and union operation of two interval numbers. Particularly, if either of the intervals is real number, then $\delta_{ij}^a = 0$. If both of the intervals are identical, then $\delta_{ij}^a = 1$.

It is distinct that $0 \leqslant \delta_{ij}^a \leqslant 1$ and $\delta_{ij}^a = \delta_{ji}^a$. As far as condition attribute set C is concerned, the similarity degree between x_i and x_j regarding C is $\delta_{ij}^C = \frac{1}{|C|}\sum_{a\in C}\delta_{ij}^a$.

Definition 3. (See [42]) Given an $IvFDIS = (U, C \cup \widetilde{D}, V, f)$, for any $B \subseteq C$, the δ-similarity relation R_B^δ is formulated as

$$R_B^\delta = \{(x_i,x_j) \in U \times U | \delta_{ij}^B \geqslant \delta\}, \tag{3}$$

and the δ-similarity class of x_i with respect to R_B^δ is formulated as

$$[x_i]_B^\delta = \{x_j \in U | (x_i,x_j) \in R_B^\delta\}, \tag{4}$$

which contains a collection of objects similar to x_i at the δ-similarity level. δ can be made an adjustment in accordance with actual status, on condition that δ ranges from 0 to 1. From the foregoing, a conclusion that the δ-similarity relation is reflexive and symmetrical can be reached.

Definition 4. (See [42]) Let R_B^δ be a δ-similarity relation induced by B for $B \subseteq C$. Suppose that U is partitioned into s crisp decision classes, namely $U/D = \{D_1, D_2, \cdots, D_s\}$, where each decision class belongs to the same category. For any $x \in U$, the fuzzy decision of x is formulated as

$$\widetilde{D}_k = \frac{|[x_i]_B^\delta \cap D_k|}{|[x_i]_B^\delta|},\ (k = 1,2,\cdots,s) \tag{5}$$

indicating the membership degree of object x_i with respect to class D_k.

Definition 5. (See [8]) Let R_B^δ be a δ-similarity relation induced by B for $B \subseteq C$. For any $x \in U$, the δ-fuzzy lower and upper approximations of fuzzy decision $\widetilde{D}_k$ concerning R_B^δ are defined as

$$\begin{aligned} \underline{R_B^\delta}\widetilde{D}_k(x) &= \inf_{y\in U} \max\{1 - R_B^\delta(x,y), \widetilde{D}_k(y)\}, \\ \overline{R_B^\delta}\widetilde{D}_k(x) &= \sup_{y\in U} \min\{R_B^\delta(x,y), \widetilde{D}_k(y)\}. \end{aligned} \tag{6}$$

For any $y \in U$, the δ-fuzzy similarity class of x induced by R_B^δ is formulated as $[x]_B^\delta = R_B^\delta(x,y)$.

2.3. Self-information

Information is intangible, whereas it is measurable. As information is acquired, its uncertainty is lessened, whose result shows the information is deeply linked to the uncertainty. From the perspective of uncertainty and probability, Shannon advanced a viewpoint of self-information.

Definition 6. (See [29]) Suppose that the probability of random variable x is $p(x)$, the self-information of x is specified as

$$I(x) = -\log p(x), \tag{7}$$

that is the negative of logarithm about the probability of variable x.

Proposition 2. (See [29]) *The self-information $I(x)$ of random variable x conforms with the following properties:*

(1) *$I(x)$ is strictly monotonic decreasing function with respect to $p(x)$, namely the smaller the probability is, the larger the uncertainty of x is.*
(2) *Under the condition of limitation, if $p(x) \to 0, I(x) \to \infty$, and if $p(x) \to 1, I(x) \to 0$.*
(3) *The joint self-information of random variables x and y is equal to the sum of self-information of x and y, namely $I(xy) = I(x) + I(y)$.*

Definition 7. (See [31]) Given an $IvFDIS = (U, C \cup \widetilde{D}, V, f)$, R_B is a fuzzy similarity relation induced by B for $B \subseteq C$. For $D_k \in U/D$, the relative decision self-information of D_k is defined as

$$I_B(D_k) = -\beta_B(D_k) \log \alpha_B(D_k), \tag{8}$$

where $\alpha_B(D_k)$ and $\beta_B(D_k)$ are the relative decision precision and roughness of D_k. $\alpha_B(D_k)$ can be formulated as $\alpha_B(D_k) = \frac{\sum_{j=1}^{n} \underline{R_B} D_k(x_j)}{\sum_{j=1}^{n} \overline{R_B} D_k(x_j)}$, and $\beta_B(D_k) = 1 - \alpha_B(D_k)$.

Wang et al. [31,45] evolved the relative decision self-information to measure the uncertainty of fuzzy set. The relative decision self-information (RDSi) not only considers the consistent classification of samples, but also emphasizes the boundary information of samples, which is a pretty effective indicator for feature selection. To carry forward this advantage continually, the RDSi will be apply to IvFDIS in what follows.

3. λ-fuzzy similarity rough set and self-information in IvFDIS

In this section, a novel rough set on the basis of λ-fuzzy similarity relation is constructed. Whereafter, an uncertainty metric using λ-fuzzy similarity self-information is investigated.

3.1. λ-fuzzy similarity rough set

In supervised learning, every sample is labeled with crisp integer which is an emblem of its classification. In contrast with previous decision-making methods, we will obfuscate all labels according to the consistency of condition and decision attributes. Making use of the λ-fuzzy similarity class of Definition 3, the fuzzy decision between sample x_i and its decision class which contains all samples with the same label as x_i is figured up. Compared with the fuzzy decision mentioned in Definition 4, the similarity lies in that identical calculation formula is adopted, while the difference is reflected from that only this fuzzy decision about the decision class of sample x_i is calculated. The fuzzy decision of sample x_i with respect to its decision class is marked as $\mathscr{D}_f$.

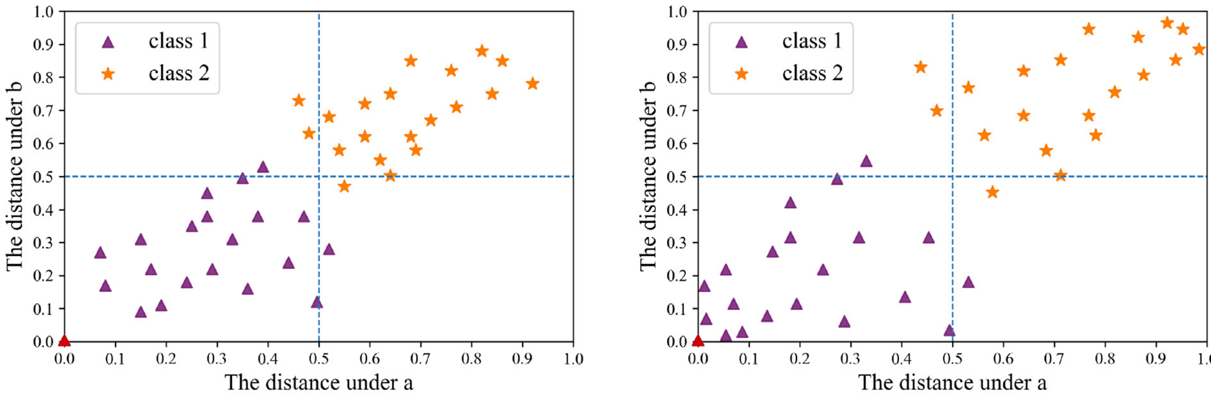
Definition 8. Let $(U, C \cup \widetilde{D}, V, f)$ be an *IvFDIS*. For any $x_i, x_j \in U, B \subseteq C$ and $a \in B$, the fuzzy similarity degree between x_i and x_j is specified as

$$S_a(x_i, x_j) = \frac{1}{2}\sin(\delta_{ij}^{a}\pi - \frac{\pi}{2}) + \frac{1}{2}, \tag{9}$$

where δ_{ij}^{a} is the similarity degree between x_i and x_j regarding a. For the attribute subset $B, S_B(x_i, x_j) = \frac{1}{|B|}\sum_{a \in B} S_a(x_i, x_j)$.

There are two classes (class 1 and class 2) containing 20 samples respectively. Given an sample x^{Δ} in class 1 and attribute subset $B = \{a, b\}$, similarity degree δ_{ij}^{B} and fuzzy similarity degree S^a between these 40 samples and x^{Δ} are computed. Considering that the distance between two samples is inversely proportional to its similarity degree, we mark the distance based on δ_{ij}^{B} or S^a as $1 - \delta_{ij}^{B}$ or $1 - S^a$. Therefore, the function of S^a is to make these samples whose similarity degrees among them are greater than 0.5 closer, while these samples whose similarity degrees among them are less than 0.5 farther, so as to narrow the intra-class distance and enlarge the inter-class distance. As shown in Fig. 2, (a) and (b) represent the distance between 40 objects in two classes and x^{Δ} under attributes a and b respectively, and x^{Δ} is marked in red. In comparison with (a), it is distinct that adopting the distance based on S^a, purple triangles are close to the origin of coordinates and orange stars are far away from the origin of coordinates in (b). In other words, the distance between these objects in class 1 and the red triangle lessens significantly. Meanwhile, the distance between these objects in class 2 and the red triangle augments significantly.

Taking $\lambda = 0.1$ as an example, Fig. 3 is the distribution of fuzzy similarity relation among 10 samples. For one thing, if the fuzzy similarity degree of x_i and x_j is not greater than the value of λ (blue area), it indicates that the fuzzy similarity relation between two samples is extremely approximate to 0. Thereupon, x_i and x_j can be regarded as completely dissimilar samples. For another, if the fuzzy similarity degree of x_i and x_j is not less than the value of $1-\lambda$ (red area), it manifests that the sim-



(a) The distance expressed as $1-\delta_{ij}^B$ (b) The distance expressed as $1-S^a$

Fig. 2. Two types of distance between x^{Δ} and 40 samples.

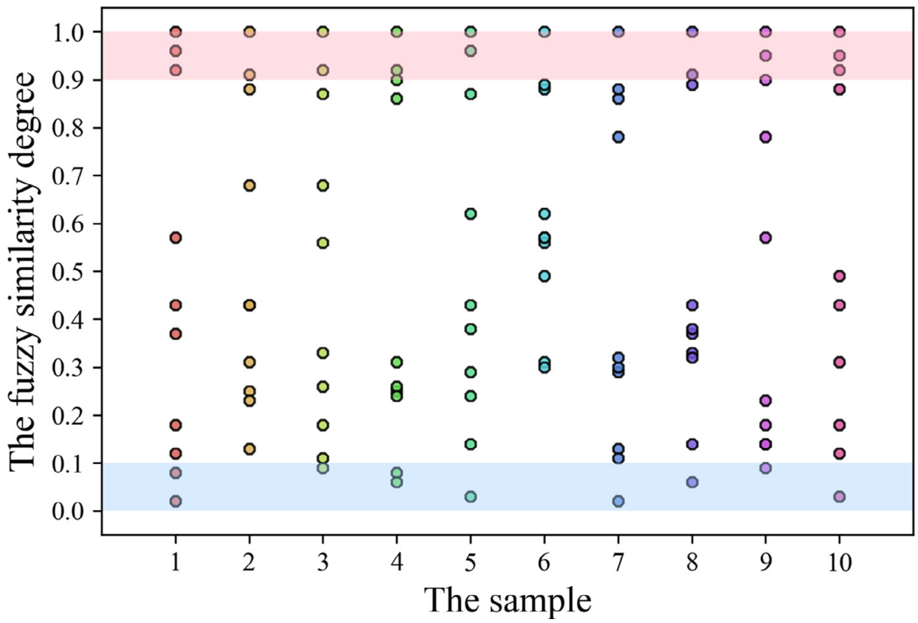


Fig. 3. The fuzzy similarity degree among 10 samples.

ilarity relation between two samples is pretty close to 1. Thereupon, we infer that there is hardly any difference between x_i and x_j. Accordingly, an improvement of fuzzy similarity relation is discussed.

Definition 9. Let $(U, C \cup \widetilde{D}, V, f)$ be an *IvFDIS*. For any $x_i, x_j \in U, B \subseteq C$ and $a \in B$, the λ-fuzzy similarity relation between x_i and x_j regarding attribute a is specified as

$$S_a^\lambda(x_i, x_j) = \begin{cases} 0, & S_a(x_i, x_j) \leqslant \lambda, \\ 1, & S_a(x_i, x_j) \geqslant 1 - \lambda, \\ S_a(x_i, x_j), & otherwise, \end{cases} \tag{10}$$

where λ is a threshold satisfied $\lambda \in (0, 0.3]$. Analogously, the λ-fuzzy similarity degree between x_i and x_j regarding attribute set B is $S_B^\lambda(x_i, x_j) = \frac{1}{|B|}\sum_{a \in B} S_a^\lambda(x_i, x_j)$.

Definition 10. Given an $IvFDIS = (U, C \cup \widetilde{D}, V, f), S_B^\lambda$ is a λ-fuzzy similarity relation induced by B for $B \subseteq C$. For any $x \in U$, the λ-fuzzy lower and upper approximations of fuzzy decision $\mathscr{D}_f$ concerning S_B^λ are defined as

$$\begin{aligned} \underline{S}_B^\lambda \mathscr{D}_f(x) &= \inf_{y \in U} \max\{1 - S_B^\lambda(x, y), \mathscr{D}_f(y)\}, \\ \overline{S}_B^\lambda \mathscr{D}_f(x) &= \sup_{y \in U} \min\{S_B^\lambda(x, y), \mathscr{D}_f(y)\}. \end{aligned} \tag{11}$$

For any $y \in U$, the λ-fuzzy similarity class of x induced by S_B^λ is formulated as $[x]_B^\lambda = S_B^\lambda(x, y)$.

3.2. λ-fuzzy similarity self-information

With the purpose of extracting attribute subsets qualified with powerful classification ability, feature selection accomplishes data dimension reduction while maintaining the classification accuracy unaltered or improved. Consequently, we advance λ-fuzzy similarity precision and roughness so as to construct λ-fuzzy similarity self-information, which provides a metric approach to estimate the classification ability of attribute subsets with respect to fuzzy decision $\mathscr{D}_f$.

Definition 11. Let $(U, C \cup \widetilde{D}, V, f)$ be an *IvFDIS*. For any $B \subseteq C$, the λ-fuzzy similarity precision α_B^λ and roughness β_B^λ of $\mathscr{D}_f$ are formulated as

$$\alpha_B^\lambda(\mathscr{D}_f) = \frac{\sum_{j=1}^{n} \underline{S}_B^\lambda \mathscr{D}_f(x)}{\sum_{j=1}^{n} \overline{S}_B^\lambda \mathscr{D}_f(x)}, \ \beta_B^\lambda(\mathscr{D}_f) = 1 - \alpha_B^\lambda(\mathscr{D}_f). \tag{12}$$

Apparently, $\alpha_B^\lambda \in [0, 1]$ and it embodies the degree that samples are classified correctly. $\beta_B^\lambda \in [0, 1]$ and it embodies the degree that samples belong to the boundary region.

Definition 12. Let S_B^λ be a λ-similarity relation induced by B for $B \subseteq C$. The λ-fuzzy similarity self-information (λ-FSSi) of $\mathscr{D}_f$ is defined as

$$I_B^\lambda(\mathscr{D}_f) = -\beta_B^\lambda(\mathscr{D}_f) \log \alpha_B^\lambda(\mathscr{D}_f). \tag{13}$$

It is noted that when $B = \emptyset$, we assume $I_B^\lambda(\mathscr{D}_f) \to \infty$.

Proposition 3. *For any $B_1, B_2 \subseteq C$, if $B_1 \subseteq B_2$, then $I_{B_1}^\lambda(\mathscr{D}_f) \geqslant I_{B_2}^\lambda(\mathscr{D}_f)$.*

Proof. For any $x \in U$, we can achieve that $\underline{S}_{B_1}^\lambda \mathscr{D}_f(x) \leqslant \underline{S}_{B_2}^\lambda \mathscr{D}_f(x)$ and $\overline{S}_{B_1}^\lambda \mathscr{D}_f(x) \geqslant \overline{S}_{B_2}^\lambda \mathscr{D}_f(x)$ according to $B_1 \subseteq B_2$. That indicates $\alpha_{B_1}^\lambda(\mathscr{D}_f) \leqslant \alpha_{B_2}^\lambda(\mathscr{D}_f)$ and $\beta_{B_1}^\lambda(\mathscr{D}_f) \geqslant \beta_{B_2}^\lambda(\mathscr{D}_f)$. Thereby, what can be reasoned out is that $0 \leqslant -\log \alpha_{B_2}^\lambda(\mathscr{D}_f) \leqslant -\log \alpha_{B_1}^\lambda(\mathscr{D}_f)$. Connected with $0 \leqslant \beta_{B_2}^\lambda(\mathscr{D}_f) \leqslant \beta_{B_1}^\lambda(\mathscr{D}_f) \leqslant 1$, it is noticeable to obtain $I_{B_1}^\lambda(\mathscr{D}_f) \geqslant I_{B_2}^\lambda(\mathscr{D}_f)$. □

Definition 13. For any $B \subseteq C$, a λ-similarity reduct B of C is required to conform with the following conditions:

(1) *Sufficiency:* $I_B^\lambda(\mathscr{D}_f) \leqslant I_C^\lambda(\mathscr{D}_f)$,
(2) *Necessity:* $I_{B-\{a\}}^\lambda(\mathscr{D}_f) > I_B^\lambda(\mathscr{D}_f)$ *for every* $a \in B$.

Here, (1) manifests that the information contained in attribute subset B is no less than that contained in attribute set C, namely B and C are equally significant to $\mathscr{D}_f$. As revealed from (2), all attributes in B are necessary. That is to say, B is a minimum attribute subset without any redundant attributes.

Example 1. Table 1 is a complete IvFDIS depending on a λ-similarity relation S_A^λ. Collected from a botanical garden, these data display the growth of three flower cultivars measured by staffs in centimeters. $U = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7\}$ representative of seven flower samples. $A = \{a_1, a_2, a_3, a_4\}$ representative of four indicators, among which $a_i (i = 1, 2, 3, 4)$ represent sepal length, sepal width, petal length and petal width in sequence. D is a decision attribute which involves three flower cultivars.

Through observation, we may discover that seven samples are partitioned into three decision classes, including $D_1 = \{x_1, x_3\}, D_2 = \{x_2, x_5\}$ and $D_3 = \{x_4, x_6, x_7\}$.

In the first place, we may acquire the similarity degree among all the objects by computation according to Definition 2:

$$\delta_{ij}^A = \begin{pmatrix} 1 & 0.405 & 0.608 & 0.358 & 0.278 & 0.125 & 0.306 \\ 0.405 & 1 & 0.357 & 0.575 & 0.464 & 0.594 & 0.617 \\ 0.608 & 0.357 & 1 & 0.304 & 0.292 & 0.127 & 0.312 \\ 0.358 & 0.575 & 0.304 & 1 & 0.637 & 0.567 & 0.846 \\ 0.278 & 0.464 & 0.292 & 0.637 & 1 & 0.587 & 0.667 \\ 0.125 & 0.594 & 0.127 & 0.567 & 0.587 & 1 & 0.61 \\ 0.306 & 0.617 & 0.312 & 0.846 & 0.667 & 0.61 & 1 \end{pmatrix}_{7\times7}.$$

Provided that $\delta = 0.5$, the δ-similarity class is figured out by Definition 3:

$$[x_1]_A^\delta = \{x_1, x_3\}, [x_2]_A^\delta = \{x_2, x_4, x_6, x_7\}, [x_3]_A^\delta = \{x_1, x_3\}, [x_4]_A^\delta = \{x_2, x_4, x_5, x_6, x_7\},$$

$$[x_5]_A^\delta = \{x_4, x_5, x_6, x_7\}, [x_6]_A^\delta = \{x_2, x_4, x_5, x_6, x_7\}, [x_7]_A^\delta = \{x_2, x_4, x_5, x_6, x_7\}.$$

Furthermore, the fuzzy decision $\mathscr{D}_f$ of x_i about its decision class is calculated:

$$\mathscr{D}_f = [1, 0.25, 1, 0.6, 0.25, 0.6, 0.6]^T$$

Conforming to Definition 9, we compute the fuzzy similarity degree:

$$S_A = \begin{pmatrix} 1 & 0.353 & 0.667 & 0.285 & 0.179 & 0.038 & 0.214 \\ 0.353 & 1 & 0.283 & 0.617 & 0.444 & 0.645 & 0.679 \\ 0.667 & 0.283 & 1 & 0.211 & 0.196 & 0.039 & 0.221 \\ 0.285 & 0.617 & 0.211 & 1 & 0.708 & 0.604 & 0.942 \\ 0.179 & 0.444 & 0.196 & 0.708 & 1 & 0.635 & 0.75 \\ 0.038 & 0.645 & 0.039 & 0.604 & 0.635 & 1 & 0.669 \\ 0.214 & 0.679 & 0.221 & 0.942 & 0.75 & 0.669 & 1 \end{pmatrix}_{7\times7}.$$

Table 1
A interval-valued fuzzy decision information system (IvFDIS).

U	a_1	a_2	a_3	a_4	D
x_1	[5.7, 6.2]	[3.1, 3.3]	[2.9, 3.3]	[1.1, 1.5]	1
x_2	[6.1, 6.3]	[3, 3.3]	[2.7, 3.2]	[1.3, 1.8]	2
x_3	[5.6, 6.2]	[3.1, 3.2]	[3, 3.4]	[1.2, 1.7]	1
x_4	[5.9, 6.3]	[2.7, 3.3]	[2.8, 3.2]	[1.5, 1.9]	3
x_5	[5.8, 6.4]	[2.6, 3.2]	[2.6, 3.1]	[1.4, 2]	2
x_6	[6.1, 6.3]	[2.8, 3.1]	[2.7, 3.1]	[1.5, 2.1]	3
x_7	[6, 6.3]	[2.7, 3.2]	[2.8, 3.2]	[1.4, 1.9]	3

λ is assigned as 0.1 by us, thus it is straightforward to get the λ-fuzzy similarity relation in the following step:

$$S_A^{\lambda} = \begin{pmatrix} 1 & 0.353 & 0.667 & 0.285 & 0.179 & 0 & 0.214 \\ 0.353 & 1 & 0.283 & 0.617 & 0.444 & 0.645 & 0.679 \\ 0.667 & 0.283 & 1 & 0.211 & 0.196 & 0 & 0.221 \\ 0.285 & 0.617 & 0.211 & 1 & 0.708 & 0.604 & 1 \\ 0.179 & 0.444 & 0.196 & 0.708 & 1 & 0.635 & 0.75 \\ 0 & 0.645 & 0 & 0.604 & 0.635 & 1 & 0.669 \\ 0.214 & 0.679 & 0.221 & 1 & 0.75 & 0.669 & 1 \end{pmatrix}_{7\times 7}.$$

4. Feature selection based on λ-FSSi using matrix operation in IvFDIS

After analysis, we conclude that λ-FSSi contains a large amount of information, which can be used to evaluate the importance of attribute subsets to $\mathscr{D}_f$. If the λ-FSSi is smaller, the attribute subset will provide more information and has a strong classification ability. About subsequent details, We exert λ-FSSi as an evaluation function to select feature in IvFDIS. Aimed at accelerating the computation speed, we define fuzzy similarity matrix to calculate λ-FSSi initially.

4.1. Matrix operation about λ-fuzzy similarity relation

At the beginning, the λ-fuzzy similarity relation is converted into a matrix, and the $\mathscr{D}_f$ is converted into a vector.

Definition 14. Given an $IvFDIS = (U, C \cup \widetilde{D}, V, f), S_B^{\lambda}$ is a λ-fuzzy similarity relation induced by B for $B \subseteq C$. The λ-fuzzy similarity matrix concerning S_B^{λ} is specified as

$$\mathscr{M}^{S_B^{\lambda}} = [m_{ij}^{S_B^{\lambda}}]_{n\times n}, \tag{14}$$

where $m_{ij}^{S_B^{\lambda}} = S_B^{\lambda}(x_i, x_j)$, and n is the cardinality of U.

Definition 15. Given an $IvFDIS = (U, C \cup \widetilde{D}, V, f), U$ is partitioned into $\{D_1, D_2, \cdots, D_s\}$. For any $x \in U$ and $x \in D_k$ $(1 \leqslant k \leqslant s)$, the fuzzy decision of x with respect to D_k is formulated as

$$\mathscr{D}_f^x = \frac{|[x]_B^{\delta} \cap D_k|}{|[x]_B^{\delta}|}. \tag{15}$$

Furthermore, the fuzzy decision $\mathscr{D}_f$ of U concerning D is specified as:

$$\mathscr{D}_f = \frac{\mathscr{D}_f^{x_1}}{x_1} + \frac{\mathscr{D}_f^{x_2}}{x_2} + \frac{\mathscr{D}_f^{x_3}}{x_3} + \cdots + \frac{\mathscr{D}_f^{x_n}}{x_n} = \sum_{i=1}^{n} \frac{\mathscr{D}_f^{x_i}}{x_i}, \tag{16}$$

namely a vector $\mathscr{V}^{\mathscr{D}_f} = [\mathscr{D}_f^{x_1}, \mathscr{D}_f^{x_2}, \mathscr{D}_f^{x_3}, \cdots, \mathscr{D}_f^{x_n}]^T$.

Definition 16. Given an $IvFDIS = (U, C \cup \widetilde{D}, V, f), S_B^{\lambda}$ is a λ-fuzzy similarity relation induced by B for $B \subseteq C$. $\mathscr{M}^{S_B^{\lambda}}$ is the λ-fuzzy similarity matrix concerning S_B^{λ}, and $\mathscr{V}^{\mathscr{D}_f}$ is the fuzzy decision vector concerning D. For any $x \in U$, the λ-fuzzy lower and upper approximations of fuzzy decision $\mathscr{D}_f$ concerning S_B^{λ} using matrix operation are defined as

$$\begin{aligned} \underline{S}_B^{\lambda}\mathscr{D}_f(x) &= (1 - \mathscr{M}^{S_B^{\lambda}}) \odot \mathscr{V}^{\mathscr{D}_f}, \\ \overline{S}_B^{\lambda}\mathscr{D}_f(x) &= \mathscr{M}^{S_B^{\lambda}} \otimes \mathscr{V}^{\mathscr{D}_f}, \end{aligned} \tag{17}$$

where $\underline{S}_B^{\lambda}\mathscr{D}_f(x) = [\underline{s}_{ij}]_{n\times 1} = \bigwedge_{k=1}^{n} \{(1 - m_{ik}^{S_B^{\lambda}}) \vee \mathscr{D}_f^{x_{kj}}\}(i = 1, 2, \cdots, n, j = 1)$, and $\overline{S}_B^{\lambda}\mathscr{D}_f(x) = [\overline{s}_{ij}]_{n\times 1} = \bigvee_{k=1}^{n} \{m_{ik}^{S_B^{\lambda}} \wedge \mathscr{D}_f^{x_{kj}}\}$ $(i = 1, 2, \cdots, n, j = 1)$.

Even though Defining 16 and Defining 10 are in diverse forms, they are equivalent. Hence the λ-fuzzy similarity precision α_B^{λ} and roughness β_B^{λ} of $\mathscr{D}_f$ are still formulated in the shape of Definition 11.

Example 2. (Continued from Example 1) In compliance with formula (17), the calculation results about the λ-fuzzy lower and upper approximations of fuzzy decision $\mathscr{D}_f$ are shown:

$$\underline{S}_A^{\lambda}\mathscr{D}_f(x) = (1-\mathscr{M}^{S_A^{\lambda}})\odot\mathscr{V}^{\mathscr{D}_f} = \begin{pmatrix} 0 & 0.647 & 0.333 & 0.715 & 0.821 & 1 & 0.786 \\ 0.647 & 0 & 0.717 & 0.383 & 0.556 & 0.355 & 0.321 \\ 0.333 & 0.717 & 0 & 0.789 & 0.804 & 1 & 0.779 \\ 0.715 & 0.383 & 0.789 & 0 & 0.292 & 0.396 & 0 \\ 0.821 & 0.556 & 0.804 & 0.292 & 0 & 0.365 & 0.25 \\ 1 & 0.355 & 1 & 0.396 & 0.365 & 0 & 0.331 \\ 0.786 & 0.321 & 0.779 & 0 & 0.25 & 0.331 & 0 \end{pmatrix}_{7\times 7} \odot \begin{pmatrix} 1 \\ 0.25 \\ 1 \\ 0.6 \\ 0.25 \\ 0.6 \\ 0.6 \end{pmatrix}_{7\times 1} = \begin{pmatrix} 0.647 \\ 0.25 \\ 0.717 \\ 0.292 \\ 0.25 \\ 0.355 \\ 0.25 \end{pmatrix}_{7\times 1}.$$

$$\overline{S}_A^{\lambda}\mathscr{D}_f(x) = \mathscr{M}^{S_A^{\lambda}}\otimes\mathscr{V}^{\mathscr{D}_f} = \begin{pmatrix} 1 & 0.353 & 0.667 & 0.285 & 0.179 & 0 & 0.214 \\ 0.353 & 1 & 0.283 & 0.617 & 0.444 & 0.645 & 0.679 \\ 0.667 & 0.283 & 1 & 0.211 & 0.196 & 0 & 0.221 \\ 0.285 & 0.617 & 0.211 & 1 & 0.708 & 0.604 & 1 \\ 0.179 & 0.444 & 0.196 & 0.708 & 1 & 0.635 & 0.75 \\ 0 & 0.645 & 0 & 0.604 & 0.635 & 1 & 0.669 \\ 0.214 & 0.679 & 0.221 & 1 & 0.75 & 0.669 & 1 \end{pmatrix}_{7\times 7} \otimes \begin{pmatrix} 1 \\ 0.25 \\ 1 \\ 0.6 \\ 0.25 \\ 0.6 \\ 0.6 \end{pmatrix}_{7\times 1} = \begin{pmatrix} 1 \\ 0.6 \\ 1 \\ 0.6 \\ 0.6 \\ 0.6 \\ 0.6 \end{pmatrix}_{7\times 1}.$$

For an attribute subset $A_1 = \{a_1, a_2, a_4\}$, we have $I_{A_1}^{\lambda}(\mathscr{D}_f) = 0.351$. By computation, we achieve $I_A^{\lambda}(\mathscr{D}_f) = 0.384$, which implies that a_3 cannot provide additional information despite combining with A_1. Since A_1 is equipped with more classification information than A in line with Definition 13 (1), it is of great essence to examine whether A_1 covers redundant attributes to determine whether it is a reduct. Thereby, we get $I_{A_1-\{a_1\}}^{\lambda}(\mathscr{D}_f) = 0.526, I_{A_1-\{a_2\}}^{\lambda}(\mathscr{D}_f) = 0.459, I_{A_1-\{a_4\}}^{\lambda}(\mathscr{D}_f) = 0.367$. Due to $I_{A_1-\{a_1\}}^{\lambda}(\mathscr{D}_f) > I_{A_1-\{a_2\}}^{\lambda}(\mathscr{D}_f) > I_{A_1-\{a_4\}}^{\lambda}(\mathscr{D}_f) > I_{A_1}^{\lambda}(\mathscr{D}_f), A_1 = \{a_1, a_2, a_4\}$is a reduct.

4.2. A static algorithm for feature selection based on λ-FSSi

The subsequent statements are two measures to distinguish the significance of an attribute in an attribute subset.

Definition 17. Let $(U, C\cup\widetilde{D}, V, f)$ be an *IvFDIS*. For any $B\subseteq C$ and $a\in C-B$, the outer significance of a to B with respect to $\mathscr{D}_f$ is defined as

$$SIG_{out}(a, B, \mathscr{D}_f) = I_B^{\lambda}(\mathscr{D}_f) - I_{B\cup\{a\}}^{\lambda}(\mathscr{D}_f). \tag{18}$$

If $SIG_{out}(a, B, \mathscr{D}_f) = 0$, it indicates that a does not contain additional information, namely a is a unnecessary attribute for feature selection.

Choosing λ-FSSi as an evaluation indicator and combining the above two measures of significance, Algorithm 1 is designed for feature selection in a static information system. Three parameters are introduced into the static algorithm. δ is set for obtaining the δ-similarity class, which is a preparative course to compute the fuzzy decision $\mathscr{D}_f$. λ is a threshold of λ-fuzzy similarity relation, which determines the transformation of fuzzy similarity degree. μ is used to terminate the selection of *Red*, which implies only a slight decrement occurs in λ-FSSi. In steps 3–7, we conduct the process of obfuscation in order to calculate the vector $\mathscr{V}^{\mathscr{D}_f}$. Steps 8–12 accomplishes the computation of λ-fuzzy similarity relation under all the condition attributes. In steps 13–26, different attributes are combined and their λ-FSSi is compared to gain an attribute subset with the greatest outer significance in each loop about B. Thereinto, steps 8–12 estimate whether to stop the selection. If there is no conspicuous variation in λ-FSSi, the calculation is completed. In steps 27–31, those redundant attributes will be removed from the achieved reduct. Finally, a result of feature selection is output. The time complexity of major steps in Algorithm 1 are shown in Table 2.

Table 2
The time complexity of Algorithm 1

Steps	Timecomplexity	Steps	Timecomplexity
$3-7$	$O(\lvert C\rvert\lvert U\rvert^2)$	$13-26$	$O(\lvert B\rvert^2\lvert U\rvert^2)$
$8-12$	$O(\frac{1}{2}(\lvert U\rvert+\lvert U\rvert^2))$	$27-31$	$O(\lvert Red\rvert^2\lvert U\rvert^2)$

Algorithm 1 A static algorithm for feature selection based on λ-FSSi (FSSi-FS)

Input: An $IvFDIS = (U, C \cup \widetilde{D}, V, f)$, parameters δ, λ and μ.
Output: A reduct Red.

begin
  Initialize: $Red \leftarrow \emptyset$;
  for $i = 1 : |U|$ **do**
    Calculate $[x_i]_C^{\delta}$;
    Calculate $\mathcal{D}_f^{x_i}$ by formula (15);
  end
  Calculate $\mathcal{V}^{\mathcal{D}_f} \leftarrow [\mathcal{D}_f^{x_i}]_{|U|\times 1}$;
  for $i = 1 : |U|$ **do**
    for $j = i : |U|$ **do**
      Calculate $m_{ij}^{S_C^{\lambda}}$ by formula (10) and $m_{ji}^{S_C^{\lambda}} \leftarrow m_{ij}^{S_C^{\lambda}}$;
    end
  end
  while $C - Red \neq \emptyset$ **do**
    $B \leftarrow C - Red$;
    for *each* $a \in B$ **do**
      Calculate $\mathcal{M}^{S^{\lambda}_{Red\cup\{a\}}} \leftarrow [m_{ij}^{S^{\lambda}_{Red\cup\{a\}}}]_{|U|\times|U|}$;
      Calculate $I^{\lambda}_{Red\cup\{a\}}(\mathcal{V}^{\mathcal{D}_f}) \leftarrow -\beta^{\lambda}_{Red\cup\{a\}}(\mathcal{V}^{\mathcal{D}_f}) \log \alpha^{\lambda}_{Red\cup\{a\}}(\mathcal{V}^{\mathcal{D}_f})$;
    end
    Find a_k with minimum value of $I^{\lambda}_{Red\cup\{a_k\}}(\mathcal{V}^{\mathcal{D}_f})$;
    Calculate $SIG_{out}(a_k, Red, \mathcal{V}^{\mathcal{D}_f}) \leftarrow I^{\lambda}_{Red}(\mathcal{V}^{\mathcal{D}_f}) - I^{\lambda}_{Red\cup\{a_k\}}(\mathcal{V}^{\mathcal{D}_f})$;
    if $SIG_{out}(a_k, Red, \mathcal{V}^{\mathcal{D}_f}) \leq \mu$ **then**
      break;
    else
      $Red \leftarrow Red \cup \{a_k\}$;
    end
  end
  for *each* $a \in Red$ **do**
    if $I^{\lambda}_{Red-\{a\}}(\mathcal{V}^{\mathcal{D}_f}) \leq I^{\lambda}_{Red}(\mathcal{V}^{\mathcal{D}_f})$ **then**
      $Red \leftarrow Red - \{a\}$;
    end
  end
  return Red;
end

5. The incremental mechanism for feature selection in dynamic IvFDIS

In an IvFDIS, information is ever-changing over time. Confronted with time-evolving objects, incremental technique is exerted to dispose of feature selection with effect. These objects possibly emerge two main variations for a dynamic IvFDIS, including an increase or a decrease. Relying on the incremental technique, it is not neccessary for us to recalculate the reduct, which is an excellent time-saving measure. This section considers two incremental algorithms for feature selection aimed at inserting or removing multiple objects.

5.1. The incremental feature selection about deleting some objects

Since the uncertainty metric plays a pivotal role in feature selection, the computation time of a reduct is dependent upon λ-FSSi to a large extent. In addition, it is not hard to perceive that the updating of λ-FSSi is closely associated with the fuzzy decision vector $\mathscr{V}^{\mathscr{D}_f}$ and the λ-fuzzy similarity matrix $\mathscr{M}^{S_B^{\lambda}}$. The subsequent research has to do with the incremental mechanism when multiple objects are removed from original IvFDIS. The object set composed of deleted objects is denoted by ${}^{-}U$.

Proposition 4. *Let* $(U, C \cup \widetilde{D}, V, f)$ *be an IvFDIS with n finite objects. For any* $B \subseteq C$, *the object set* ${}^{-}U = \{x_{o_1}, x_{o_2}, \cdots, x_{o_{n'}}\}$ $(1 \leqslant o_1 < o_2 < \cdots < o_{n'} \leqslant n)$ *is deleted from original IvFDIS, then* $U' = U - {}^{-}U$ *in the new IvFDIS. The new fuzzy decision vector is formulated as* ${}^{-}\mathscr{V}^{\mathscr{D}_f} = [{}^{-}\mathscr{D}_f^{x_i}]_{(n-n')\times 1}$, *and its updating mechanism is as follows:*

$$
{}^{-}\mathscr{D}_f^{x_i} = \begin{cases} \mathscr{D}_f^{x_{i+k-1}}, & o_{k-1} - k + 1 < i < o_k - k + 1 \text{for any} 1 \leqslant k \leqslant n', {}^{-}[x_i]_B^{\delta} = [x_{i+k-1}]_B^{\delta} \wedge {}^{-}D(x_i) = D(x_{i+k-1}), \\ \mathscr{D}_f^{x_{i+n'}}, & o_{n'} - n' + 1 \leqslant i \leqslant n - n', {}^{-}[x_i]_B^{\delta} = [x_{i+k-1}]_B^{\delta} \wedge {}^{-}D(x_i) = D(x_{i+k-1}), \\ \frac{|{}^{-}[x_i]_B^{\delta} \cap {}^{-}D(x_i)|}{|{}^{-}[x_i]_B^{\delta}|}, \text{otherwise}. \end{cases} \tag{19}
$$

Here, $[x]_B^{\delta}$ denotes the δ-similarity class of x and $D(x)$ denotes the decision class of x in original IvFDIS. ${}^{-}[x]_B^{\delta}$ signifies the δ-similarity class of x and ${}^{-}D(x)$ signifies the decision class of x in new IvFDIS.

Proof. The updating about vector ${}^{-}\mathscr{V}^{\mathscr{D}_f}$ is divided into three parts. If neither the δ-similarity class nor the decision class of object has changed after the deletion, we are about to study two situations. The first is to transform the position of other objects between two deleted objects whose subscripts are consecutive. Assume that the subscripts of any two deleted objects with consecutive subscripts are o_{k-1} and o_k, o_{k-1} indicates that there is $k-1$ objects to be deleted in front of certain object x_i, namely the $\mathscr{D}_f^{x_i}$ is supposed to move $k-1$ positions forward. That means when $o_{k-1} - k + 1 < i < o_k - k + 1$ and $1 \leqslant k \leqslant n'$, ${}^{-}\mathscr{D}_f^{x_i} = \mathscr{D}_f^{x_{i+k-1}}$ in new *IvFDIS*. Secondly, the position of other objects at the back of last deleted object are converted. If certain object x_i is behind the object x_{o_n}, it signifies that there are n' objects in front of x_i to be removed, indicating that x_i should be moved forward by n' positions. That is when $o_{n'} - n' + 1 < i < n - n'$, ${}^{-}\mathscr{D}_f^{x_i} = \mathscr{D}_f^{x_{i+n'}}$ in new *IvFDIS*. If the δ-similarity class or the decision class of object has changed after the deletion, the fuzzy decision of object is supposed to recalculate by formula (15). That is to say, ${}^{-}\mathscr{D}_f^{x_i} = \frac{|{}^{-}[x_i]_B^{\delta} \cap {}^{-}D(x_i)|}{|{}^{-}[x_i]_B^{\delta}|}$. □

Proposition 5. *Let $(U, C \cup \widetilde{D}, V, f)$ be an IvFDIS with n finite objects. For any $B \subseteq C$, the object set ${}^{-}U = \{x_{o_1}, x_{o_2}, \cdots, x_{o_n}\}$ $(1 \leqslant o_1 < o_2 < \cdots < o_{n'} \leqslant n)$ is deleted from original IvFDIS, then $U' = U - {}^{-}U$ in the new IvFDIS. The new λ-fuzzy similarity matrix is formulated as ${}^{-}\mathscr{M}^{S_B^{\lambda}} = [{}^{-}m_{ij}^{S_B^{\lambda}}]_{(n-n')\times(n-n')}$, and its updating mechanism is as follows:*

$$
{}^{-}m_{ij}^{S_B^{\lambda}} = \begin{cases} m_{i+k-1,j+k-1}^{S_B^{\lambda}}, & o_{k-1} - k + 1 < i,j < o_k - k + 1 \text{for any} 1 \leqslant k \leqslant n', \\ m_{i+n',j+n'}^{S_B^{\lambda}}, & o_{n'} - n' + 1 \leqslant i,j \leqslant n - n'. \end{cases} \tag{20}
$$

Proof. According to the proof of Proposition 4, it is analogous to reach the conclusion of Proposition 5. □

Example 3. (Continued from Example 2) The object set ${}^{-}U = \{x_4, x_6\}$ is removed from Table 1, then $U' = \{x_1, x_2, x_3, x_5, x_7\}$ in new *IvFDIS*, which is exhibited in Table 3. Firstly, the new fuzzy decision vector ${}^{-}\mathscr{V}^{\mathscr{D}_f}$ is updated in accordance with Proposition 4, which is computed that ${}^{-}\mathscr{V}^{\mathscr{D}_f} = [1, 0.5, 1, 0.5, 0.333]^T$. Next, we may update the new λ-fuzzy similarity matrix ${}^{-}\mathscr{M}^{S_A^{\lambda}}$ in accordance with Proposition 5:

$$
{}^{-}M^{S_A^{\lambda}} = \begin{pmatrix} 1 & 0.353 & 0.667 & 0.179 & 0.214 \\ 0.353 & 1 & 0.283 & 0.444 & 0.679 \\ 0.667 & 0.283 & 1 & 0.196 & 0.221 \\ 0.179 & 0.444 & 0.196 & 1 & 0.75 \\ 0.214 & 0.679 & 0.221 & 0.75 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0.353 & 0.667 & 0.179 & 0.214 \\ 0.353 & 1 & 0.283 & 0.444 & 0.679 \\ 0.667 & 0.283 & 1 & 0.196 & 0.221 \\ 0.179 & 0.444 & 0.196 & 1 & 0.75 \\ 0.214 & 0.679 & 0.221 & 0.75 & 1 \end{pmatrix}_{5\times5}.
$$

Subsequently, an incremental feature selection algorithm in regard to deleting multiple objects is recommended on the basis of Algorithm 1. The detailed designs are displayed in Algorithm 2. Step 2 deletes multiple objects from original *IvFDIS*. Step 3 updates the original fuzzy decision vector and λ-fuzzy similarity matrix by Proposition 4 and 5. Step 4 Calculates the λ-FSSi of original reduct in current *IvFDIS*. In steps 5–19, we distinguish whether the new λ-FSSi derived from the original

Table 3
A new IvFDIS after removing some objects.

U	a_1	a_2	a_3	a_4	D
x_1	[5.7, 6.2]	[3.1, 3.3]	[2.9, 3.3]	[1.1, 1.5]	1
x_2	[6.1, 6.3]	[3, 3.3]	[2.7, 3.2]	[1.3, 1.8]	2
x_3	[5.6, 6.2]	[3.1, 3.2]	[3, 3.4]	[1.2, 1.7]	1
x_5	[5.8, 6.4]	[2.6, 3.2]	[2.6, 3.1]	[1.4, 2]	2
x_7	[6, 6.3]	[2.7, 3.2]	[2.8, 3.2]	[1.4, 1.9]	3

reduct is greater than the λ-FSSi under all attributes in new *IvFDIS*. If so, it is indispensable to insert these remaining attributes to the previous reduct. Thereinto, steps 6–10 establish a ascending sequence of λ-FSSi aiming to remanent attributes. Steps 11–18 add every attribute in the ascending sequence to the new reduct until the variation of λ-FSSi is not apparent anymore. In steps 20–24, these redundant attributes from the selected reduct are removed. In the end, a new reduct is obtained. Table 3 shows the time complexity of major steps in Algorithm 2.

Algorithm 2 An incremental algorithm for feature selection about deleting some objects (FSSi-FSD)

Input:

(1) The original $IvFDIS = (U, C \cup \widetilde{D}, V, f)$, parameters δ, λ and μ.

(2) The original fuzzy decision vector $\mathcal{V}^{\mathcal{D}_f} = [\mathcal{D}_f^{x_i}]_{n\times 1}$, λ-fuzzy similarity matrix $\mathcal{M}^{S_B^\lambda} = [m_{ij}^{S_B^\lambda}]_{n\times n}$.

(3) The deleted object set $^{-}U = \{x_{o_1}, x_{o_2}, \cdots, x_{o_n}\}$.

(4) The original reduct Red.

Output: A new reduct Red'.

1 **begin**
2 Initialize: $Red' \leftarrow Red$, $U' = U - {}^{-}U$;
3 Update the fuzzy decision vector $^{-}\mathcal{V}^{\mathcal{D}_f} = [^{-}\mathcal{D}_f^{x_i}]_{(n-n')\times 1}$, λ-fuzzy similarity matrix $^{-}\mathcal{M}^{S_C^\lambda} = [^{-}m_{ij}^{S_C^\lambda}]_{(n-n')\times(n-n')}$;
4 Calculate the new $I_{Red'}^\lambda(^{-}\mathcal{V}^{\mathcal{D}_f})$ and $I_C^\lambda(^{-}\mathcal{V}^{\mathcal{D}_f})$;
5 **if** $I_{Red'}^\lambda(^{-}\mathcal{V}^{\mathcal{D}_f}) > I_C^\lambda(^{-}\mathcal{V}^{\mathcal{D}_f})$ **then**
6 **for** *each* $a \in C - Red'$ **do**
7 Calculate $^{-}M^{S^\lambda_{Red'\cup\{a\}}} \leftarrow [^{-}m_{ij}^{S^\lambda_{Red'\cup\{a\}}}]_{|U'|\times|U'|}$;
8 Calculate $I_{Red'\cup\{a\}}^\lambda(^{-}\mathcal{V}^{\mathcal{D}_f}) \leftarrow -\beta_{Red'\cup\{a\}}^\lambda(^{-}\mathcal{V}^{\mathcal{D}_f}) \log \alpha_{Red'\cup\{a\}}^\lambda(^{-}\mathcal{V}^{\mathcal{D}_f})$;
9 **end**
10 Sort these attributes $a_k \in C - Red'$ according to an ascending order of $I_{Red'\cup\{a_k\}}^\lambda(^{-}\mathcal{V}^{\mathcal{D}_f})$ and mark the sequence $A_r = \{a_{r_1}, a_{r_2}, \cdots, a_{r_{|C-Red'|}}\}$;
11 **for** *each* $a \in A_r$ **do**
12 Calculate $SIG_{out}(a, Red', {}^{-}\mathcal{V}^{\mathcal{D}_f}) \leftarrow I_{Red'}^\lambda(^{-}\mathcal{V}^{\mathcal{D}_f}) - I_{Red'\cup\{a\}}^\lambda(^{-}\mathcal{V}^{\mathcal{D}_f})$;
13 **if** $SIG_{out}(a, Red', {}^{-}\mathcal{V}^{\mathcal{D}_f}) \le \mu$ **then**
14 break;
15 **else**
16 $Red' \leftarrow Red' \cup \{a\}$;
17 **end**
18 **end**
19 **end**
20 **for** *each* $a \in Red'$ **do**
21 **if** $I_{Red'-\{a\}}^\lambda(^{-}\mathcal{V}^{\mathcal{D}_f}) \le I_{Red'}^\lambda(^{-}\mathcal{V}^{\mathcal{D}_f})$ **then**
22 $Red' \leftarrow Red' - \{a\}$;
23 **end**
24 **end**
25 **return** Red';
26 **end**

Example 4. (Continued from Example 3) According to Algorithm 2, after the updating of the fuzzy decision vector $^{-}\mathscr{V}^{\mathscr{D}_f}$ and λ-fuzzy similarity matrix $^{-}\mathscr{M}^{S_A^\lambda}$ are accomplished, we need to compare $I_{Red'}^\lambda(^{-}\mathscr{V}^{\mathscr{D}_f})$ and $I_A^\lambda(^{-}\mathscr{V}^{\mathscr{D}_f})$ in order to determine whether to go steps 5–19. By computation, it can be obtained that $I_{Red'}^\lambda(^{-}\mathscr{V}^{\mathscr{D}_f}) = 0.152$ and $I_A^\lambda(^{-}\mathscr{V}^{\mathscr{D}_f}) = 0.184$. Due to $I_{Red'}^\lambda(^{-}\mathscr{V}^{\mathscr{D}_f}) < I_A^\lambda(^{-}\mathscr{V}^{\mathscr{D}_f})$, we continue to conduct steps 20–24 so that redundant attributes are taken out from $Red' = \{a_1, a_2, a_4\}$. Since $I_{Red'-\{a_1\}}^\lambda(^{-}\mathscr{V}^{\mathscr{D}_f}) = 0.247, I_{Red'-\{a_2\}}^\lambda(^{-}\mathscr{V}^{\mathscr{D}_f}) = 0.194$ and $I_{Red'-\{a_4\}}^\lambda(^{-}\mathscr{V}^{\mathscr{D}_f}) = 0.125$, we achieve that $I_{Red'-\{a_4\}}^\lambda(^{-}\mathscr{V}^{\mathscr{D}_f}) < I_{Red'}^\lambda(^{-}V^{D_f}) < I_{Red'-\{a_2\}}^\lambda(^{-}\mathscr{V}^{\mathscr{D}_f}) < I_{Red'-\{a_1\}}^\lambda(^{-}\mathscr{V}^{\mathscr{D}_f})$, and $I_{Red'-\{a_4\}}^\lambda(^{-}\mathscr{V}^{\mathscr{D}_f}) < I_{\{a_2\}}^\lambda(^{-}\mathscr{V}^{\mathscr{D}_f}) = 0.225 < I_{\{a_1\}}^\lambda(^{-}\mathscr{V}^{\mathscr{D}_f}) = 0.294$. Consequently, the new reduct is $\{a_1, a_2\}$ after deleting object set.

5.2. The incremental feature selection about adding some objects

In a similar way, an incremental mechanism when adding multiple objects to original *IvFDIS* will be investigated in following part. The object set composed of added objects is denoted by ^{+}U.

Proposition 6. *Let $(U, C \cup \widetilde{D}, V, f)$ be an IvFDIS with n finite objects. For any $B \subseteq C$, the object set ${}^{+}U = \{x_{n+1}, x_{n+2}, \cdots, x_{n+n'}\}$ is added to original IvFDIS, then $U' = U + {}^{+}U$ in the new IvFDIS. The new fuzzy decision vector is formulated as ${}^{+}\mathscr{V}^{\mathscr{D}_f} = [{}^{+}\mathscr{D}_f^{x_i}]_{(n+n')\times 1}$, and its updating mechanism is as follows:*

$$
{}^{+}\mathscr{D}_f^{x_i} = \begin{cases} \mathscr{D}_f^{x_i}, & 1 \leqslant i \leqslant n, {}^{+}[x_i]_B^{\delta} = [x_i]_B^{\delta} \wedge {}^{+}D(x_i) = D(x_i), \\ \frac{|{}^{+}[x_i]_B^{\delta} \cap {}^{+}D(x_i)|}{|{}^{+}[x_i]_B^{\delta}|}, & \text{otherwise}. \end{cases} \tag{21}
$$

Here, $[x]_B^{\delta}$ denotes the δ-similarity class of x and $D(x)$ denotes the decision class of x in original IvFDIS. ${}^{+}[x]_B^{\delta}$ signifies the δ-similarity class of x and ${}^{+}D(x)$ signifies the decision class of x in new IvFDIS.

Proof. The updating about vector ${}^{+}\mathscr{D}_f^{x_i}$ is divided into two parts. In the first place, if neither the δ-similarity class nor the decision class of object has changed after adding some objects, the addition of new objects has no influence on the fuzzy decision $\mathscr{D}_f$ in regard to these original n objects, hence they maintain previous $\mathscr{D}_f$ unchanged. That is, for certain object x_i, if $1 \leqslant i \leqslant n$, then ${}^{+}\mathscr{D}_f^{x_i} = \mathscr{D}_f^{x_i}$. Moreover, aimed at all new objects added to *IvFDIS*, their fuzzy decisions are expected to be calculated in the light of Definition 15. Meanwhile, for any $1 \leqslant i \leqslant n$, if the δ-similarity class or the decision class of x_i has changed after the addition, the fuzzy decision of x_i should be recalculated equally. Namely for any x_i, when $n+1 \leqslant i \leqslant n+n'$ or ${}^{+}[x_i]_B^{\delta} \neq [x_i]_B^{\delta} \vee {}^{+}D(x_i) \neq D(x_i), {}^{+}\mathscr{D}_f^{x_i} = \frac{|{}^{+}[x_i]_B^{\delta} \cap {}^{+}D(x_i)|}{|{}^{+}[x_i]_B^{\delta}|}$ in new *IvFDIS*. □

Proposition 7. *Let $(U, C \cup \widetilde{D}, V, f)$ be an IvFDIS with n finite objects. For any $B \subseteq C$, the object set ${}^{+}U = \{x_{n+1}, x_{n+2}, \cdots, x_{n+n'}\}$ is added to original IvFDIS, then $U' = U + {}^{+}U$ in the new IvFDIS. The new λ-fuzzy similarity matrix is formulated as ${}^{+}\mathscr{M}^{S_B^{\lambda}} = [{}^{+}m_{ij}^{S_B^{\lambda}}]_{(n+n')\times(n+n')}$, and its updating mechanism is as follows:*

$$
{}^{+}m_{ij}^{S_B^{\lambda}} = \begin{cases} m_{ij}^{S_B^{\lambda}}, & 1 \leqslant i,j \leqslant n, \\ S_B^{\lambda}(x_i, x_j), & (n+1 \leqslant i \leqslant n+n') \vee (n+1 \leqslant j \leqslant n+n'). \end{cases} \tag{22}
$$

Proof. The updating about matrix ${}^{+}m_{ij}^{S_B^{\lambda}}$ is divided into two parts. At first, these previous n objects will not be affected by the addition of new objects, thereby for any $1 \leqslant i,j \leqslant n, {}^{+}m_{ij}^{S_B^{\lambda}}$ is still identical, namely ${}^{+}m_{ij}^{S_B^{\lambda}} = m_{ij}^{S_B^{\lambda}}$. Furthermore, when $i \in [n+1, n+n']$ or $j \in [n+1, n+n']$, the λ-fuzzy similarity relation between x_i and x_j is required to figure out in accordance with Definition 9. That means ${}^{+}m_{ij}^{S_B^{\lambda}} = S_B^{\lambda}(x_i, x_j)$ for any $(n+1 \leqslant i \leqslant n+n') \vee (n+1 \leqslant j \leqslant n+n')$. □

Example 5. (Continued from Example 2) The object set ${}^{+}U = \{x_8, x_9\}$ is inserted into Table 1, then $U' = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9\}$ in new *IvFDIS*, which is displayed in Table 5. Firstly, in accordance with Proposition 6, we update the new fuzzy decision vector ${}^{+}\mathscr{V}^{\mathscr{D}_f}$ as ${}^{+}\mathscr{V}^{\mathscr{D}_f} = [1, 0.25, 1, 0.6, 0.25, 0.6, 0.6, 1, 1]^T$. Next, in accordance with Proposition 7, the new λ-fuzzy similarity matrix ${}^{+}\mathscr{M}^{S_A^{\lambda}}$ can be updated as follows:

$$
{}^{+}M^{S_A^{\lambda}} = \left(\begin{array}{ccccccc|cc}
1 & 0.353 & 0.667 & 0.285 & 0.179 & 0 & 0.214 & 0.282 & 0.43 \\
0.353 & 1 & 0.283 & 0.617 & 0.444 & 0.645 & 0.679 & 0.217 & 0.247 \\
0.667 & 0.283 & 1 & 0.211 & 0.196 & 0 & 0.221 & 0.52 & 0.44 \\
0.285 & 0.617 & 0.211 & 1 & 0.708 & 0.604 & 1 & 0 & 0.118 \\
0.179 & 0.444 & 0.196 & 0.708 & 1 & 0.635 & 0.75 & 0 & 0 \\
0 & 0.645 & 0 & 0.604 & 0.635 & 1 & 0.669 & 0 & 0 \\
0.214 & 0.679 & 0.221 & 1 & 0.75 & 0.669 & 1 & 0.143 & 0 \\
\hline
0.282 & 0.217 & 0.52 & 0 & 0 & 0 & 0.143 & 1 & 0.15 \\
0.43 & 0.247 & 0.44 & 0.118 & 0 & 0 & 0 & 0.15 & 1
\end{array}\right)_{9\times 9}
$$

Taking advantage of the updating of fuzzy decision vector and λ-fuzzy similarity matrix, Algorithm 3 presents an incremental feature selection approach in regard to adding multiple objects. In Algorithm 3, step 2 adds multiple objects to original *IvFDIS*. The original fuzzy decision vector and λ-fuzzy similarity matrix are updated by Proposition 4 and 5 in step 3. Then, we calculate the λ-FSSi of original reduct in current *IvFDIS* in step 4. Steps 5–19 distinguish whether the new λ-FSSi derived from the original reduct is greater than the λ-FSSi under all attributes in new *IvFDIS*. If so, steps 6–10 will establish a ascending sequence of λ-FSSi about remanent attributes. In steps 11–18, we insert every attribute in the ascending sequence into the previous reduct in turn. Once the variation of λ-FSSi is not obvious anymore, the loop will be broken out. In steps 20–24, these redundant attributes from the selected reduct are removed. Lastly, a new reduct is output. The time complexity of major steps about Algorithm 3 is analyzed in Table 4.

Table 4
The time complexity of Algorithm 2.

Steps	Timecomplexity	Steps	Timecomplexity
3	$O(\lvert U'\rvert)$	5 – 19	$O(\lvert A_r\rvert\lvert U'\rvert^2)$
4	$O((\lvert Red'\rvert + \lvert C\rvert)\lvert U'\rvert^2)$	20 – 24	$O(\lvert Red'\rvert^2\lvert U'\rvert^2)$

Table 5
A new IvFDIS after inserting some objects.

U	a_1	a_2	a_3	a_4	D
x_1	[5.7, 6.2]	[3.1, 3.3]	[2.9, 3.3]	[1.1, 1.5]	1
x_2	[6.1, 6.3]	[3, 3.3]	[2.7, 3.2]	[1.3, 1.8]	2
x_3	[5.6, 6.2]	[3.1, 3.2]	[3, 3.4]	[1.2, 1.7]	1
x_4	[5.9, 6.3]	[2.7, 3.3]	[2.8, 3.2]	[1.5, 1.9]	3
x_5	[5.8, 6.4]	[2.6, 3.2]	[2.6, 3.1]	[1.4, 2]	2
x_6	[6.1, 6.3]	[2.8, 3.1]	[2.7, 3.1]	[1.5, 2.1]	3
x_7	[6, 6.3]	[2.7, 3.2]	[2.8, 3.2]	[1.4, 1.9]	3
x_8	[5.4, 5.8]	[3.1, 3.2]	[2.8, 3]	[1.2, 1.6]	1
x_9	[5.5, 6.1]	[3, 3.3]	[3.1, 3.5]	[1.3, 1.4]	2

Algorithm 3 An incremental algorithm for feature selection about adding some objects (FSSi-FSA)

Input:
(1) The original $IvFDIS = (U, C \cup \widetilde{D}, V, f)$, parameters δ, λ and μ.
(2) The original fuzzy decision vector $\mathcal{V}^{\mathcal{D}_f} = [\mathcal{D}_f^{x_i}]_{n\times 1}$, λ-fuzzy similarity matrix $\mathcal{M}^{S_B^\lambda} = [m_{ij}^{S_B^\lambda}]_{n\times n}$.
(3) The deleted object set $^-U = \{x_{o_1}, x_{o_2}, \cdots, x_{o_n}\}$.
(4) The original reduct Red.
Output: A new reduct Red'.

begin
  Initialize: $Red' \leftarrow Red$, $U' = U - {}^-U$;
  Update the fuzzy decision vector $^-\mathcal{V}^{\mathcal{D}_f} = [^-\mathcal{D}_f^{x_i}]_{(n-n')\times 1}$, λ-fuzzy similarity matrix $^-\mathcal{M}^{S_C^\lambda} = [^-m_{ij}^{S_C^\lambda}]_{(n-n')\times(n-n')}$;
  Calculate the new $I_{Red'}^\lambda(^-\mathcal{V}^{\mathcal{D}_f})$ and $I_C^\lambda(^-\mathcal{V}^{\mathcal{D}_f})$;
  if $I_{Red'}^\lambda(^-\mathcal{V}^{\mathcal{D}_f}) > I_C^\lambda(^-\mathcal{V}^{\mathcal{D}_f})$ **then**
    for *each* $a \in C - Red'$ **do**
      Calculate $^-M^{S_{Red'\cup\{a\}}^\lambda} \leftarrow [^-m_{ij}^{S_{Red'\cup\{a\}}^\lambda}]_{|U'|\times|U'|}$;
      Calculate $I_{Red'\cup\{a\}}^\lambda(^-\mathcal{V}^{\mathcal{D}_f}) \leftarrow -\beta_{Red'\cup\{a\}}^\lambda(^-\mathcal{V}^{\mathcal{D}_f}) \log \alpha_{Red'\cup\{a\}}^\lambda(^-\mathcal{V}^{\mathcal{D}_f})$;
    end
    Sort these attributes $a_k \in C - Red'$ according to an ascending order of $I_{Red'\cup\{a_k\}}^\lambda(^-\mathcal{V}^{\mathcal{D}_f})$ and mark the sequence $A_r = \{a_{r_1}, a_{r_2}, \cdots, a_{r_{|C-Red'|}}\}$;
    for *each* $a \in A_r$ **do**
      Calculate $SIG_{out}(a, Red', {}^-\mathcal{V}^{\mathcal{D}_f}) \leftarrow I_{Red'}^\lambda(^-\mathcal{V}^{\mathcal{D}_f}) - I_{Red'\cup\{a\}}^\lambda(^-\mathcal{V}^{\mathcal{D}_f})$;
      if $SIG_{out}(a, Red', {}^-\mathcal{V}^{\mathcal{D}_f}) \leq \mu$ **then**
        break;
      else
        $Red' \leftarrow Red' \cup \{a\}$;
      end
    end
  end
  for *each* $a \in Red'$ **do**
    if $I_{Red'-\{a\}}^\lambda(^-\mathcal{V}^{\mathcal{D}_f}) \leq I_{Red'}^\lambda(^-\mathcal{V}^{\mathcal{D}_f})$ **then**
      $Red' \leftarrow Red' - \{a\}$;
    end
  end
  return Red';
end

Example 6. (Continued from Example 5) According to Algorithm 3, after the updating of the fuzzy decision vector ${}^{+}\mathscr{V}^{\mathscr{D}_f}$ and λ-fuzzy similarity matrix ${}^{+}\mathscr{M}^{S_A^\lambda}$ are accomplished, we need to compare $I^{\lambda}_{Red'}({}^{+}\mathscr{V}^{\mathscr{D}_f})$ and $I^{\lambda}_{A}({}^{+}\mathscr{V}^{\mathscr{D}_f})$ in order to determine whether to go steps 5–19. By computation, it can be obtained that $I^{\lambda}_{Red'}({}^{+}\mathscr{V}^{\mathscr{D}_f}) = 0.266$ and $I^{\lambda}_{A}({}^{+}\mathscr{V}^{\mathscr{D}_f}) = 0.272$. Due to $I^{\lambda}_{Red'}({}^{+}\mathscr{V}^{\mathscr{D}_f}) < I^{\lambda}_{A}({}^{+}\mathscr{V}^{\mathscr{D}_f})$, we continue to conduct steps 20–24 so that redundant attributes are taken out from $Red' = \{a_1, a_2, a_4\}$. Since $I^{\lambda}_{Red'-\{a_1\}}({}^{+}\mathscr{V}^{\mathscr{D}_f}) = 0.565, I^{\lambda}_{Red'-\{a_2\}}({}^{+}\mathscr{V}^{\mathscr{D}_f}) = 0.24$ and $I^{\lambda}_{Red'-\{a_4\}}({}^{+}\mathscr{V}^{\mathscr{D}_f}) = 0.289$, we achieve that $I^{\lambda}_{Red'-\{a_2\}}({}^{+}\mathscr{V}^{\mathscr{D}_f}) < I^{\lambda}_{Red'}({}^{-}V^{D_f}) < I^{\lambda}_{Red'-\{a_4\}}({}^{+}\mathscr{V}^{\mathscr{D}_f}) < I^{\lambda}_{Red'-\{a_1\}}({}^{+}\mathscr{V}^{\mathscr{D}_f})$, and $I^{\lambda}_{Red'-\{a_2\}}({}^{+}\mathscr{V}^{\mathscr{D}_f}) < I^{\lambda}_{\{a_1\}}({}^{+}\mathscr{V}^{\mathscr{D}_f}) = 0.379 < I^{\lambda}_{\{a_4\}}({}^{+}\mathscr{V}^{\mathscr{D}_f}) = 0.424$. Consequently, the new reduct is $\{a_1, a_4\}$ after adding object set.

6. Experimental analysis

In this section, a series of experiments are designed to verify the feasibility and effectiveness of our incremental algorithms. Ten data sets are downloaded from UCI Machine Learning Repository and details are available in Table 7. All the experimental programs are executed on a computer with an Intel Core i7-9750H at 2.60 GHz, 8 GB RAM and Windows 10 (64-bit). These algorithms are accomplished by Python using an environment of Anaconda Navigator. Before numerical experiments are implemented, we proceed the data preprocessing. Subsequently, the proposed incremental algorithm is compared with three current feature selection algorithms, namely accelerated algorithm by fuzzy rough set-based information entropy (AFRI) [46], heuristic algorithm based on variable distance parameter (AVDP) [26], attribute reduction with fuzzy rough self-information measures (FSI) [31]. The main evaluation indicators about different algorithms are four aspects, including the robustness of evaluation function, the computation time of algorithm, the quantity of feature selection and the classification accuracy. In order to minimize the experimental error, we will carry out the program ten times for each data set to obtain the average value as the final result (Table 6).

6.1. Data preparation

For some non-numerical data in several data sets such as time data and textual data, they are processed with the way of erasing time data and partial textual data that are difficult to be converted into numerical data, and converting other textual data into numerical data.

What is noteworthy is that all data sets retrieved from UCI are single-valued, so it is of great essence to extend them for obtaining interval values. Firstly, the min–max normalization is adopted to normalize the single-valued data, that is for any $x_i \in U$ and $a \in C$,

$$f(x_i, a) = \frac{f(x_i, a) - \min_{x_i \in U}\{f(x_i, a)\}}{\max_{x_i \in U}\{f(x_i, a)\} - \min_{x_i \in U}\{f(x_i, a)\}}. \tag{23}$$

Table 6
The time complexity of Algorithm 3.

Steps	Timecomplexity	Steps	Timecomplexity
3	$O(\lvert{}^{+}U\rvert\lvert U\rvert)$	$5-19$	$O(\lvert A_r\rvert\lvert U'\rvert^2)$
4	$O((\lvert Red'\rvert + \lvert C\rvert)\lvert U'\rvert^2)$	$20-24$	$O(\lvert Red'\rvert^2\lvert U'\rvert^2)$

Table 7
The detailed description of data sets.

No.	Dataset	Abbreviation	Object	Attribute	Class
1	DARWIN	DW	174	451	2
2	Wine	W	178	14	3
3	Heart	H	270	14	2
4	Foresttypemapping	FTM	326	27	4
5	Turkishmusicemotion	TME	400	50	4
6	Australiancreditapproval	ACA	690	14	2
7	Parkinsonspeech	PS	1040	28	2
8	MiceProteinExpression	MPE	1080	82	8
9	Winequalityred	WQR	1599	12	6
10	Winequalitywhite	WQW	4898	12	7
11	Pageblocks	PB	5473	11	2
12	DryBean	DB	13611	17	7

In addition, an interval value $[f(x_i,a)^L, f(x_i,a)^U]$ is generated from a single value $f(x_i,a)$ by the following method:

$$\begin{aligned} f(x_i,a)^L &= (1-2\sigma)\times f(x_i,a),\\ f(x_i,a)^U &= (1+2\sigma)\times f(x_i,a), \end{aligned} \tag{24}$$

where σ represents the standard deviation of $f(x_i,a)$ for any $x_i \in U$ under attribute a. Particularly, if $f(x_i,a)^U > 1$, it can be reassigned to 1. Finally, there are three parameters (δ,λ,μ) in our experiments. By searching for appropriate parameters on twelve data sets, we set $\delta = 0.5, \lambda = 0.1$ and $\mu = 0.01$ preliminarily.

6.2. The robustness of uncertainty metrics

In this subsection, the robustness of evaluation function is tested on the data set No.1 to No.6. Three uncertainty metrics are λ-conditional entropy in algorithm AFRI, relative decision self-information in algorithm FSI and λ-fuzzy similarity self-information in the proposed algorithm FSSi-FS. After adding random noise to IvFDIS, different evaluation functions are computed on the basis of the same attributes. For any $x_i \in U$ and $a \in C$, the random noise is added as follows:

$$[f(x_i,a)^L, f(x_i,a)^U] = \begin{cases} [f(x_i,a)^L + r, f(x_i,a)^U + r], & 0 \leqslant r \leqslant 1,\\ [f(x_i,a)^L, f(x_i,a)^U], & \text{otherwise}, \end{cases} \tag{25}$$

It should be noted that $0 \leqslant f(x_i,a)^L + r \leq 1$ and $0 \leqslant f(x_i,a)^U + r \leq 1$.

For each data set, a certain percentage of random noise will be added to it each time, starting at 10% and ending at 50% of all objects, increasing by 10% at a time up to five additions. Fig. 4 and Fig. 5 display the final experimental results, among which Fig. 4 reflects the variation trend of three evaluation functions in different data sets and Fig. 5 reflects the standard deviation of three evaluation functions after five additions about each data set. The experimental results of Fig. 4 and Fig. 5 are presented in Table 8, from which we can achieve that the proposed λ-FSSi outperforms other evaluation functions on five data sets when different proportions of noise are added to original data set. In Fig. 4, it can be found that as the added noise continues to increase, the evaluation function of algorithm FSI fluctuates greatly, especially in the data sets H and TME. Moreover, the fluctuation of evaluation function about the algorithm ARFI is obvious on the data set H. For data set PS, the evaluation function of ARFI algorithm is slightly more stable than that of FSSi-FS algorithm, because of the instability caused by adding ten percent noise to λ-FSSi. Even though the standard deviation of λ-FSSi is not the smallest on data set PS, its fluctuation is unconspicuous. Thereby, it is deduced that λ-FSSi is robust comparable to the uncertainty metrics of algorithm ARFI and FSI.

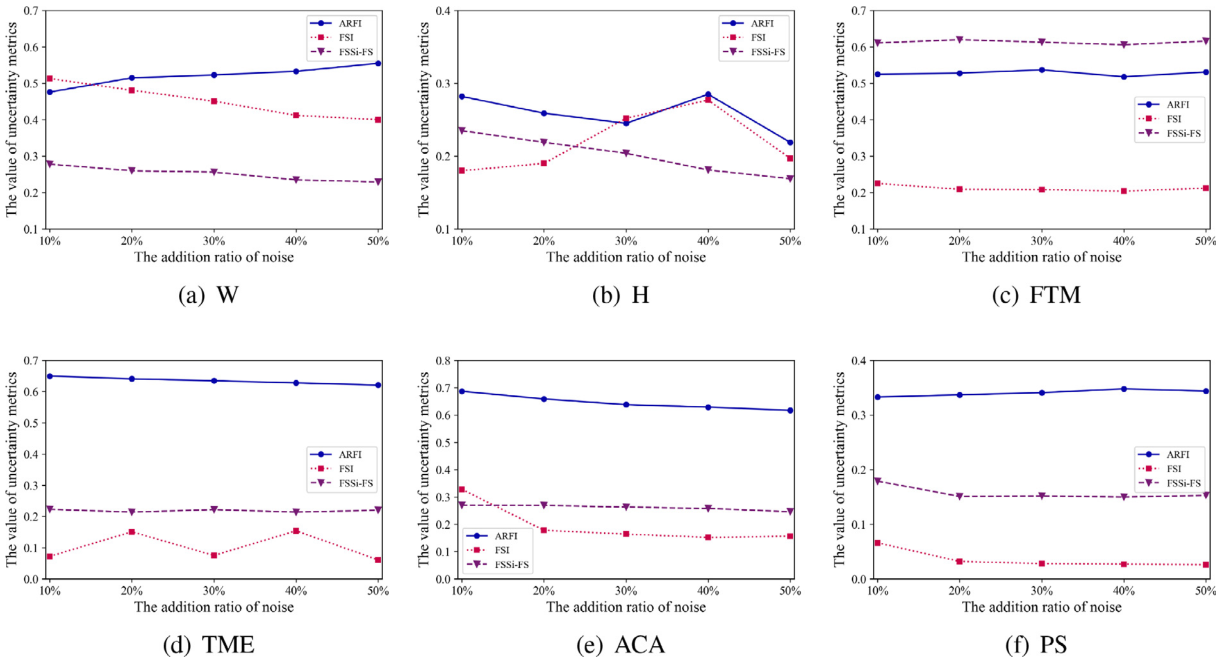


Fig. 4. The variation of evaluation function when adding a certain ratio of noise in IvFDIS.

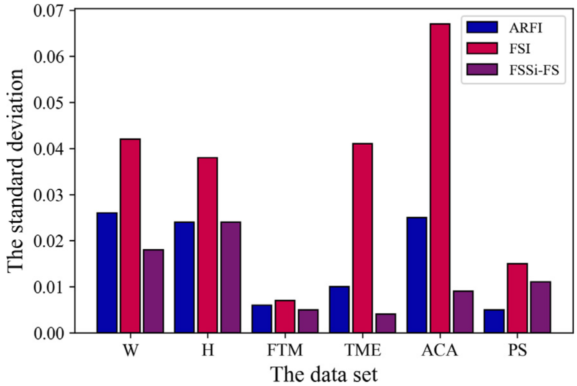


Fig. 5. The standard deviation of evaluation function about six data sets.

Table 8
The comparison of three evaluation functions on six data sets.

Data set	Algorithm	10%	20%	30%	40%	50%	Standard deviation
	ARFI	0.476	0.515	0.523	0.533	0.555	0.026
W	FSI	0.513	0.481	0.451	0.412	0.400	0.042
	FSSi-FS	0.278	0.260	0.256	0.235	0.229	**0.018**
	ARFI	0.282	0.245	0.285	0.219	0.312	**0.024**
H	FSI	0.180	0.190	0.252	0.277	0.197	0.038
	FSSi-FS	0.235	0.219	0.204	0.181	0.169	**0.024**
	ARFI	0.525	0.528	0.537	0.518	0.531	0.006
FTM	FSI	0.225	0.209	0.208	0.204	0.212	0.007
	FSSi-FS	0.611	0.620	0.613	0.606	0.616	**0.005**
	ARFI	0.650	0.641	0.635	0.628	0.621	0.010
TME	FSI	0.072	0.151	0.075	0.154	0.061	0.041
	FSSi-FS	0.223	0.214	0.222	0.214	0.220	**0.004**
	ARFI	0.687	0.659	0.638	0.629	0.617	0.025
ACA	FSI	0.328	0.178	0.164	0.152	0.156	0.067
	FSSi-FS	0.270	0.269	0.263	0.258	0.246	**0.009**
	ARFI	0.333	0.337	0.341	0.348	0.344	**0.005**
PS	FSI	0.066	0.032	0.028	0.027	0.026	0.015
	FSSi-FS	0.179	0.151	0.152	0.150	0.153	0.011

6.3. The computation time

This subsection implements these experiments of feature selection in regard to deleting multiple objects and adding multiple objects in IvFDIS on twelve data sets. Taking the computation time of feature selection as the main evaluation indicator, three feature selection algorithms (ARFI, AVPD, FSI) based on fuzzy rough set are contrasted with two proposed incremental algorithms (FSSi-FSD, FSSi-FSA). Because algorithms ARFI, AVPD and FSI play a part in single-valued data, we should make some changes and replace single-valued data in aforementioned algorithms with interval-valued data during our experiments.

6.3.1. The experiment with the deletion of some objects

In twelve data sets, a certain percentage of objects will be deleted randomly each time, starting at 10% and ending at 50% of original objects, increasing by 10% at a time up to five deletions. If the magnitude of deleted objects is not an integer, the floor function will be employed, namely for any percentage $\alpha, |^{-}U| = \lfloor \alpha \times |U| \rfloor$.

Taking the computation time of feature selection as an evaluation indicator, FSSi-FSD is compared with ARFI, AVPD and FSI on twelve data sets. The final experimental results are presented in Table 9, where the unit of computation time is seconds. To make these results of deleting objects among four algorithms more intuitive, some associative pictures are plotted in Fig. 6, where the x-coordinate represents the quantity ratio of deleted objects from original objects and the y-coordinate represents the computation time of four algorithms. As observed with the experimental results, it is fairly obvious that the computation time of feature selection for all algorithms decreases gradually when the ratio of deleted objects increases. Besides, in the process of deleting objects, FSSi-FSD is more efficient than other algorithms.

6.3.2. The experiment with the addition of some objects

For twelve data sets, each data set is randomly divided into two parts, among which 50% of data set is regarded as the original data set and the remaining 50% of data set is regarded as the test data set. In our experiment, a certain percentage of objects originated from test data set will be added randomly each time, starting at 10% and ending at 50%, increasing by 10% at a time up to five additions. Equally, the floor function will be adopted when the magnitude of added objects is not an integer.

The comparison of computation time about these four algorithms is demonstrated in Table 10, and some relevant pictures are plotted as shown in Fig. 7. For each subgraph in Fig. 7, the x-coordinate stands for the quantity ratio of added objects from test data set and the y-coordinate stands for the computation time of four algorithms. It comes easy to discover that the computing time of all algorithms shows an upward trend when the ratio of added objects increases. In addition, the computation speed of FSSi-FSA on eleven data sets is faster than that of other three algorithms.

Table 9
The comparison of four algorithms on twelve data sets when deleting objects.

Data set	Algorithm	10%	20%	30%	40%	50%
DW	FSSi-FSD	61.446	35.915	31.581	33.103	19.233
	FSI	162.737	126.250	89.924	73.358	70.316
	ARFI	591.934	560.729	348.490	234.738	123.349
	AVPD	301.071	237.468	178.270	122.725	89.234
W	FSSi-FSD	2.270	1.673	1.440	0.956	1.001
	FSI	8.308	6.716	4.974	3.356	2.205
	ARFI	14.873	12.131	9.039	6.567	4.584
	AVPD	4.374	3.223	2.543	1.864	1.335
H	FSSi-FSD	0.470	0.845	0.605	0.750	0.556
	FSI	2.689	7.308	4.974	4.195	2.930
	ARFI	7.042	19.365	14.505	10.679	6.827
	AVPD	3.427	8.943	7.194	4.977	3.341
FTM	FSSi-FSD	5.509	4.393	3.545	2.522	1.721
	FSI	32.307	23.891	21.304	17.721	11.934
	ARFI	120.006	92.141	74.569	52.200	33.740
	AVPD	46.700	38.737	28.451	21.090	14.518
TME	FSSi-FSD	10.438	8.062	6.241	5.627	3.312
	FSI	515.576	394.719	313.574	213.109	134.244
	ARFI	501.585	393.582	301.464	251.745	133.923
	AVPD	233.183	187.309	136.877	124.901	66.133
ACA	FSSi-FSD	10.578	8.444	6.719	4.938	3.477
	FSI	55.687	43.952	31.320	24.240	17.551
	ARFI	168.666	123.558	97.422	71.229	47.840
	AVPD	74.479	69.026	52.641	37.477	26.425
PS	FSSi-FSD	12.925	9.842	7.828	6.327	4.226
	FSI	129.254	104.136	78.338	60.330	43.026
	ARFI	167.627	130.029	96.374	70.077	52.606
	AVPD	80.399	61.805	46.869	35.813	25.001
MPE	FSSi-FSD	14.355	11.977	9.982	6.953	4.873
	FSI	1230.231	915.415	723.305	600.119	358.556
	ARFI	5155.241	3320.064	2563.852	1864.281	1391.248
	AVPD	3272.916	2826.840	2320.269	1673.366	949.972
WQR	FSSi-FSD	140.430	110.869	86.588	61.409	42.059
	FSI	402.573	320.060	263.516	183.963	127.473
	ARFI	931.764	746.340	581.392	407.854	288.260
	AVPD	537.456	417.325	316.118	230.939	157.475
WQW	FSSi-FSD	1078.309	768.531	655.683	447.410	365.226
	FSI	2895.362	2389.150	1771.659	1366.639	972.849
	ARFI	8140.049	6776.826	5299.761	3632.221	2783.827
	AVPD	4944.336	3822.464	3210.661	2400.260	1559.032
PB	FSSi-FSD	1233.352	940.883	724.119	549.476	357.780
	FSI	4433.770	2762.386	2230.156	1751.021	1178.211
	ARFI	8412.603	6539.061	5274.567	3854.957	2652.457
	AVPD	3144.329	2609.903	1934.152	1573.839	950.309
DB	FSSi-FSD	4680.265	2841.157	2097.611	1474.534	947.924
	FSI	76023.554	63325.158	55215.124	46516.472	28517.052
	ARFI	93214.659	72255.548	61120.756	47832.565	35559.488
	AVPD	51128.284	42638.127	35562.188	26517.356	18221.878

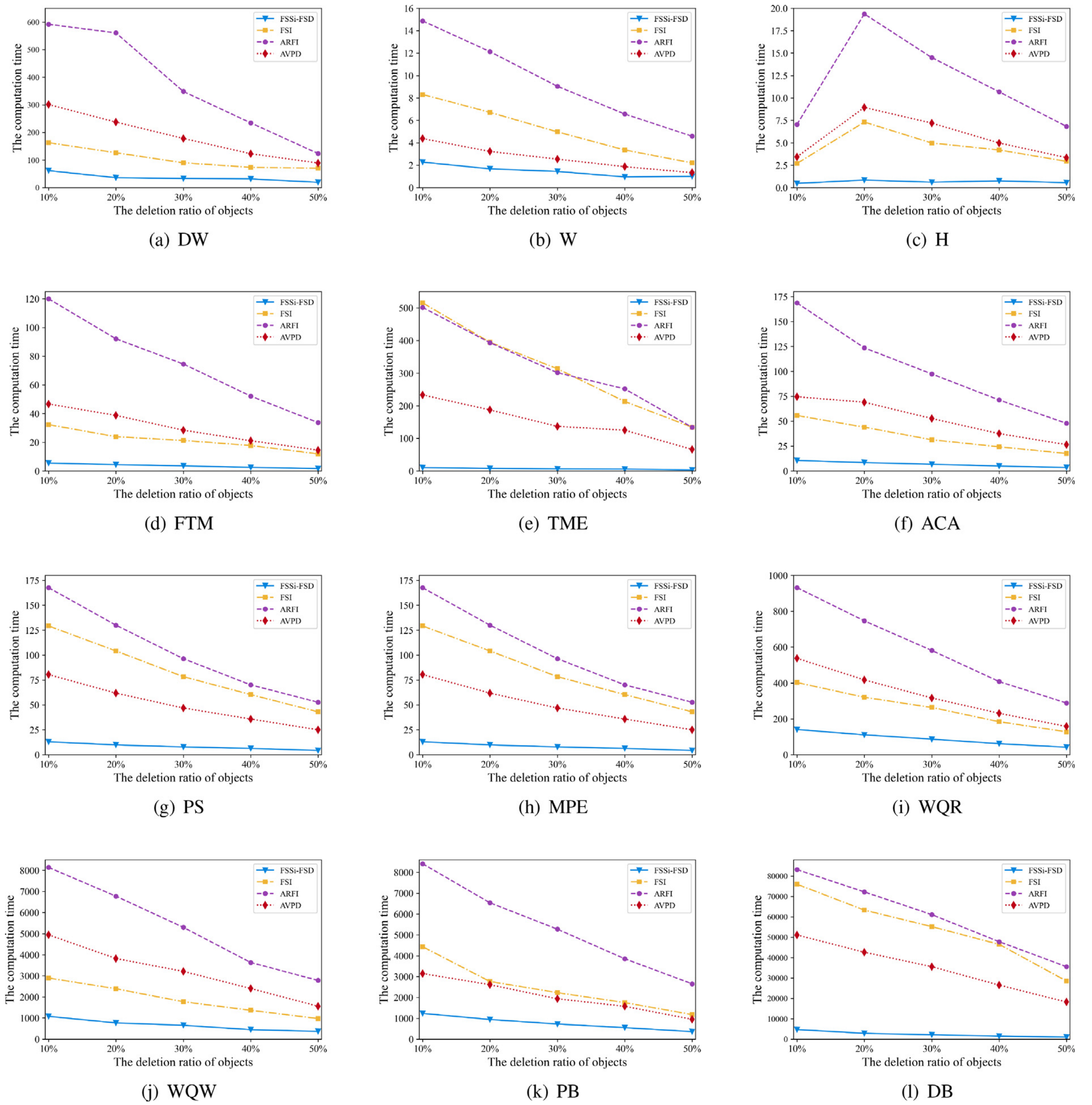


Fig. 6. The computation time among four algorithms with a certain ratio of deleting objects.

6.4. The accuracy of classification and the quantity of reduct

Dai et al. [47,48] have improved the existing classifiers and developed two classifiers capable of processing interval-valued data, that are Probabilistic Neural Network (PNN) and K-Nearest Neighbor (KNN). In both classifiers, the distance function between objects is modified. For any $x_i, x_j \in U$, attribute set $A = \{a_1, a_2, \cdots, a_m\}$, the distance between x_i and x_j is defined:

$$d(x_i, x_j) = \sqrt{\sum_{k=1}^{m} [P(f(x_i, a_k) \geqslant f(x_j, a_k)) - P(f(x_j, a_k) \geqslant f(x_i, a_k))]^2}, \quad (26)$$

where $P(f(x_i, a_k) \geqslant f(x_j, a_k) = \min\{1, \max\{\frac{v^U - w^L}{(v^U - v^L) + (w^U - w^L)}, 0\}\}, f(x_i, a_k) = [v^L, v^U]$ and $f(x_j, a_k) = [w^L, w^U]$. In following experiment, PNN and KNN are utilized to make an assessment on the classification accuracy in the process of removing and insert-

Table 10
The comparison of four algorithms on twelve data sets when adding objects.

Data set	Algorithm	10%	20%	30%	40%	50%
DW	FSSi-FSA	41.492	55.734	66.361	90.902	110.491
	FSI	112.688	135.778	131.637	163.460	174.399
	ARFI	264.103	327.195	457.926	530.232	593.028
	AVPD	140.140	183.661	249.921	343.670	428.648
W	FSSi-FSA	1.378	2.122	3.045	3.631	4.207
	FSI	3.077	5.151	6.797	8.715	9.712
	ARFI	6.220	8.750	12.108	15.003	18.865
	AVPD	1.811	2.207	3.504	4.175	5.830
H	FSSi-FSA	0.470	0.606	0.841	1.086	1.327
	FSI	3.759	6.494	8.013	9.481	11.322
	ARFI	12.079	15.269	18.652	26.890	32.079
	AVPD	5.413	7.105	8.936	10.405	15.586
FTM	FSSi-FSA	2.717	3.704	4.638	6.103	7.341
	FSI	20.110	21.536	26.165	33.574	38.603
	ARFI	54.399	70.335	83.611	118.867	155.351
	AVPD	24.470	30.960,	34.793	51.757	66.376
TME	FSSi-FSA	4.415	5.915	8.457	10.398	12.669
	FSI	224.642	307.602	383.015	463.282	767.953
	ARFI	218.735	266.196	367.434	450.015	616.948
	AVPD	104.230	146.538	178.382	218.344	290.034
ACA	FSSi-FSA	5.516	7.464	10.078	12.925	16.868
	FSI	27.360	35.492	44.798	53.163	66.410
	ARFI	73.866	107.919	129.863	157.943	209.751
	AVPD	38.828	52.719	70.162	66.120	117.027
PS	FSSi-FSA	78.754	121.285	146.266	201.358	245.338
	FSI	63.244	87.844	110.859	145.351	172.484
	ARFI	80.964	106.813	135.440	169.217	207.454
	AVPD	41.681	55.835	66.078	86.489	107.126
MPE	FSSi-FSA	46.837	60.974	78.779	87.835	96.798
	FSI	551.233	681.214	927.726	1192.377	1404.187
	ARFI	2091.413	2518.831	3855.042	5035.437	6297.512
	AVPD	1304.419	2021.576	2703.860	3618.754	4506.839
WQR	FSSi-FSA	66.623	94.705	121.874	143.571	180.578
	FSI	183.734	252.831	307.231	384.442	501.421
	ARFI	430.239	604.695	747.855	946.754	1178.817
	AVPD	233.317	323.921	407.335	533.863	636.529
WQW	FSSi-FSA	507.722	697.324	957.847	1233.592	1468.430
	FSI	1403.569	1647.775	2357.470	2753.028	3319.448
	ARFI	3885.340	5323.310	6477.873	8235.409	10204.255
	AVPD	2235.339	3024.981	3917.523	3518.280	3845.474
PB	FSSi-FSA	562.964	808.043	1070.237	1333.347	1619.066
	FSI	1321.005	2040.307	2969.312	3136.246	5233.427
	ARFI	3822.507	5228.343	6480.772	8241.516	10507.266
	AVPD	1197.275	1895.925	2633.034	3194.675	3834.865
DB	FSSi-FSA	8774.259	12548.215	16673.235	22548.322	30857.286
	FSI	41283.284	53628.515	65836.476	76873.826	91216.532
	ARFI	50247.658	66831.143	78526.852	98658.212	112786.455
	AVPD	26841.362	32588.665	42326.871	52366.121	78653.225

ing multiple objects. At the same time, 10-fold cross-validation is adopted on twelve data sets. The classification accuracy of twelve data sets and the quantity of reduct when deleting objects are shown in Tables 11 and 13. Simultaneously, The classification accuracy of twelve data sets and the quantity of reduct when adding objects are shown in Tables 12 and 13. Particularly, "Original" represents these results under all the attributes.

From Table 11, we come to the conclusion for deleting objects that FSSi-FSD reaches the highest average accuracy in ten data sets when PNN classifier is used, while it reaches the highest average accuracy on eleven data sets when KNN classifier is used. From Table 12, it is not hard to see that the average accuracy of FSSi-FSA reaches the maximum eleven times when adding objects in the PNN classifier. Meanwhile, when adding objects in the KNN classifier, the average accuracy of FSSi-FSA reaches the maximum ten times. To sum up, the proposed algorithms have decent classification accuracy in different classifiers. Table 13 demonstrates the quantity of reduct when objects are deleted and added in different proportions, from which it can be found that the proposed algorithms exceed the performance of algorithms ARFI and AVPD in most occasions, and outperform algorithm FSI in six data sets.

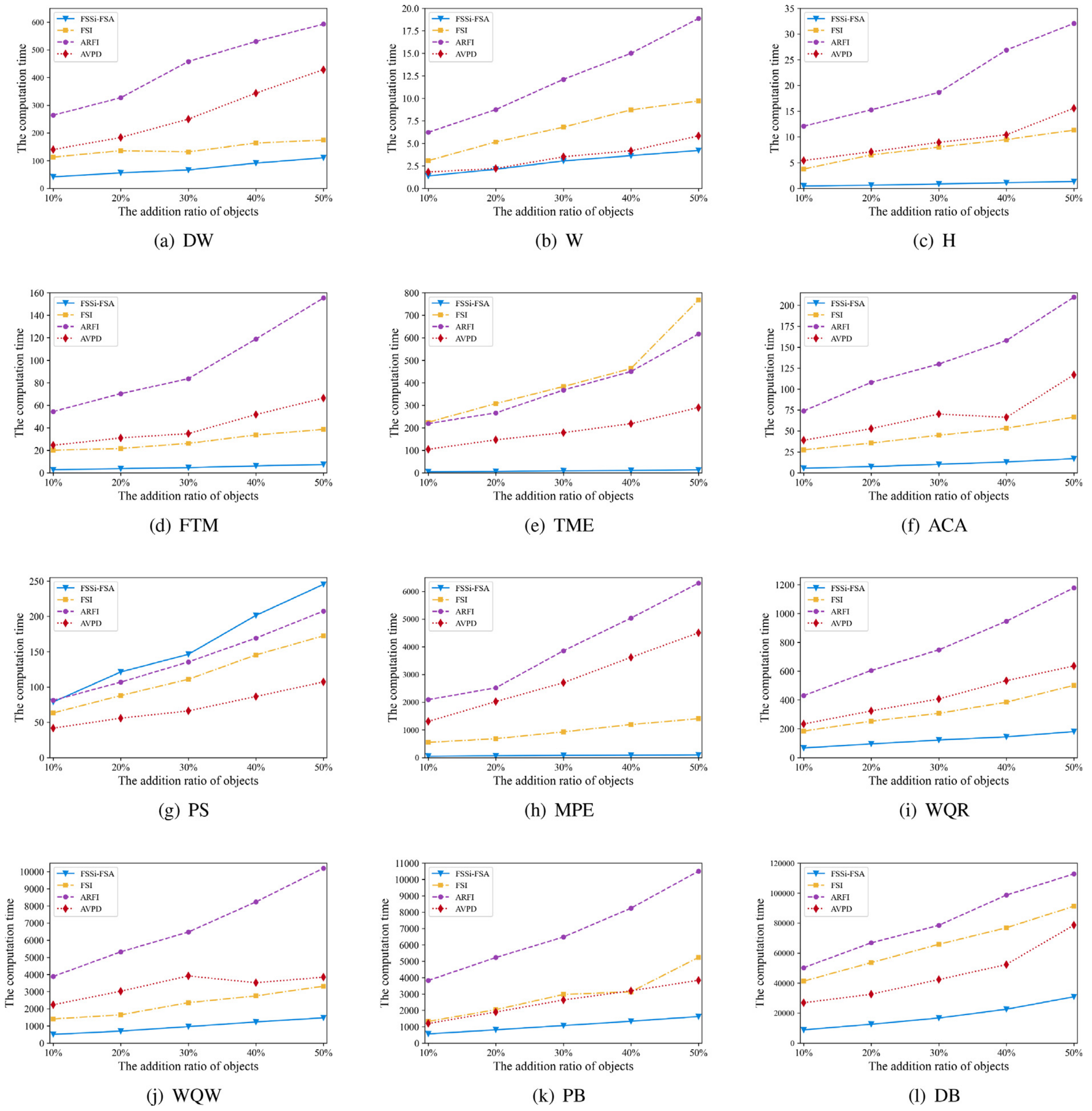


Fig. 7. The computation time among four algorithms with a certain ratio of adding objects.

6.5. The combination with the deep learning

Recently, classification problems centered on deep learning have been widely prevalent. In this field, convolutional neural network (CNN) [49,50] is utilized to cope with classification prediction about various types of images, while one-dimensional convolutional neural network (1D-CNN) is exploited to dispose of data similar to digital signals. Therefore, we attempt to adopt the proposed incremental feature selection algorithm, associated with one-dimensional convolutional neural network to predict the class of samples. Firstly, for an interval-valued information system whose objects are constantly changing, the incremental feature selection algorithm on the basis of fuzzy self-information is used to select these significant features in a data set, so as to effectively decrease the dimension of features and enhance the computational efficiency. Then, by establishing a neural network with three one-dimensional convolutional layers whose number of filters are 16, 32 and 64, these selected features are input to the neural network to classify the samples. A series of experiments, compared the classification accuracy of convolutional neural network before and after feature selection, are conducted on six data sets in Table 7. More details can be seen in Table 14. According to Table 14, it indicates that by feature selection based on fuzzy self-information,

Table 11
The classification accuracy of four algorithms on twelve data sets when deleting objects.

Data	Algorithm	10%		20%		30%		40%		50%		Average	
		PNN	KNN	PNN	KNN	PNN	KNN	PNN	KNN	PNN	KNN	PNN	KNN
DW	FSSi-FSD	**0.667**	**0.697**	**0.680**	**0.702**	**0.683**	**0.703**	0.662	**0.714**	**0.681**	0.663	**0.675**	**0.696**
	FSI	0.554	0.631	0.566	0.690	0.587	0.697	0.643	0.667	0.605	0.620	0.591	0.661
	ARFI	0.547	0.652	0.557	0.691	0.562	0.695	0.570	0.696	0.596	0.642	0.566	0.675
	AVPD	0.656	0.679	0.662	0.698	0.679	0.693	**0.674**	0.701	0.656	0.654	0.665	0.685
	Original	0.526	0.602	0.482	0.624	0.503	0.635	0.546	0.627	0.529	0.606	0.619	0.517
W	FSSi-FSD	**0.945**	**0.944**	**0.935**	**0.941**	**0.941**	0.941	**0.943**	0.936	**0.935**	0.914	**0.940**	**0.935**
	FSI	0.938	0.938	0.923	0.924	0.937	**0.942**	0.939	**0.942**	0.919	**0.927**	0.931	**0.935**
	ARFI	0.935	0.930	0.928	0.925	0.943	0.928	0.928	0.922	0.921	0.912	0.931	0.923
	AVPD	0.939	0.937	0.922	0.929	0.932	0.930	0.938	0.931	0.924	0.924	0.931	0.930
	Original	0.926	0.925	0.919	0.926	0.926	0.915	0.915	0.912	0.912	0.893	0.920	0.914
H	FSSi-FSD	**0.533**	**0.548**	0.529	**0.519**	**0.549**	**0.528**	**0.515**	**0.522**	0.549	**0.559**	0.535	**0.535**
	FSI	**0.533**	0.481	**0.531**	0.468	0.548	0.471	**0.515**	0.477	**0.569**	0.478	**0.539**	0.475
	ARFI	0.497	0.521	0.499	0.475	0.511	0.484	0.484	0.483	0.528	0.490	0.504	0.491
	AVPD	0.525	0.479	0.529	0.483	0.519	0.490	0.510	0.482	0.524	0.474	0.521	0.482
	Original	0.491	0.506	0.495	0.485	0.501	0.504	0.489	0.477	0.494	0.508	0.494	0.496
FTM	FSSi-FSD	0.769	**0.822**	**0.778**	0.805	0.780	**0.813**	**0.777**	**0.807**	**0.767**	**0.808**	**0.774**	**0.811**
	FSI	0.675	0.694	0.675	0.687	0.691	0.714	0.680	0.701	0.682	0.714	0.681	0.702
	ARFI	0.750	0.804	0.764	0.798	**0.788**	0.791	0.754	0.782	0.738	0.775	0.759	0.790
	AVPD	**0.783**	0.795	0.765	**0.808**	0.762	0.800	**0.777**	0.780	0.759	0.775	0.769	0.792
	Original	0.753	0.790	0.744	0.792	0.745	0.783	0.737	0.778	0.735	0.763	0.743	0.781
TME	FSSi-FSD	0.494	0.515	**0.500**	0.510	**0.502**	0.505	**0.510**	0.512	**0.505**	0.495	**0.502**	0.507
	FSI	0.477	0.444	0.440	0.412	0.466	0.433	0.439	0.429	0.452	0.449	0.455	0.433
	ARFI	0.485	0.502	0.482	0.509	0.480	0.497	0.465	0.479	0.472	0.474	0.477	0.492
	AVPD	**0.497**	0.505	0.468	0.509	0.475	0.515	0.482	0.521	0.481	0.501	0.481	0.510
	Original	0.487	**0.601**	0.488	**0.585**	0.492	**0.585**	0.487	**0.567**	0.483	**0.526**	0.487	**0.573**
ACA	FSSi-FSD	0.703	**0.608**	0.701	**0.580**	0.707	**0.601**	0.692	**0.607**	**0.681**	**0.579**	0.697	**0.595**
	FSI	**0.774**	0.444	**0.775**	0.437	**0.754**	0.440	**0.715**	0.444	0.650	0.469	**0.734**	0.447
	ARFI	0.604	0.531	0.626	0.556	0.600	0.541	0.626	0.514	0.599	0.529	0.611	0.534
	AVPD	0.619	0.502	0.623	0.429	0.624	0.435	0.598	0.449	0.595	0.471	0.612	0.457
	Original	0.566	0.453	0.568	0.458	0.571	0.448	0.558	0.449	0.563	0.451	0.565	0.452
PS	FSSi-FSD	**0.537**	**0.504**	**0.515**	**0.540**	**0.545**	**0.536**	**0.537**	**0.503**	**0.540**	**0.507**	**0.535**	**0.518**
	FSI	**0.537**	**0.504**	**0.515**	**0.540**	**0.545**	**0.536**	**0.537**	**0.503**	**0.540**	**0.507**	**0.535**	**0.518**
	ARFI	**0.537**	**0.504**	**0.515**	**0.540**	**0.545**	**0.536**	**0.537**	**0.503**	**0.540**	**0.507**	**0.535**	**0.518**
	AVPD	**0.537**	**0.504**	**0.515**	**0.540**	**0.545**	**0.536**	**0.537**	**0.503**	**0.540**	**0.507**	**0.535**	**0.518**
	Original	0.453	0.379	0.439	0.351	0.448	0.364	0.429	0.356	0.431	0.335	0.440	0.357
MPE	FSSi-FSD	**0.876**	**0.836**	**0.871**	**0.843**	**0.871**	**0.851**	**0.876**	**0.843**	**0.870**	**0.841**	**0.873**	**0.843**
	FSI	0.771	0.738	0.767	0.731	0.745	0.728	0.741	0.716	0.767	0.730	0.758	0.729
	ARFI	0.807	0.797	0.819	0.832	0.811	0.820	0.819	0.820	0.826	0.814	0.816	0.817
	AVPD	0.842	0.815	0.859	0.828	0.840	0.817	0.843	0.817	0.844	0.820	0.846	0.819
	Original	0.773	0.725	0.750	0.736	0.755	0.735	0.746	0.735	0.753	0.720	0.755	0.730
WQR	FSSi-FSD	**0.612**	**0.567**	**0.601**	**0.573**	**0.591**	0.560	**0.573**	**0.570**	**0.565**	**0.550**	**0.588**	**0.564**
	FSI	0.557	0.528	0.537	0.521	0.541	0.537	0.523	0.518	0.515	0.500	0.535	0.521
	ARFI	0.586	0.552	0.579	0.554	0.572	**0.564**	0.548	0.537	0.533	0.521	0.564	0.546
	AVPD	0.578	0.541	0.575	0.549	0.561	0.533	0.545	0.525	0.546	0.511	0.561	0.532
	Original	0.593	0.555	0.584	0.561	0.569	0.556	0.554	0.537	0.534	0.523	0.567	0.546
WQW	FSSi-FSD	0.594	**0.537**	**0.590**	0.530	**0.569**	0.521	0.546	**0.523**	**0.540**	**0.512**	**0.568**	**0.525**
	FSI	0.488	0.429	0.506	0.442	0.501	0.448	0.485	0.419	0.480	0.437	0.492	0.435
	ARFI	0.591	0.535	0.585	0.535	0.568	**0.529**	0.541	0.513	0.525	0.499	0.562	0.522
	AVPD	0.594	0.526	0.567	0.531	0.547	0.505	0.538	0.501	0.530	0.498	0.555	0.512
	Original	**0.595**	0.517	0.584	**0.540**	0.568	0.526	**0.551**	0.518	0.535	0.505	0.567	0.521
PB	FSSi-FSD	0.926	0.931	**0.922**	**0.926**	**0.921**	**0.926**	**0.921**	**0.926**	**0.924**	**0.926**	**0.923**	**0.927**
	FSI	**0.932**	**0.936**	0.903	0.910	0.898	0.907	0.904	0.910	0.889	0.902	0.905	0.913
	ARFI	0.915	0.920	0.914	0.918	0.909	0.914	0.907	0.912	0.910	0.912	0.911	0.915
	AVPD	0.920	0.924	0.917	0.922	0.913	0.921	0.919	0.925	0.920	0.920	0.918	0.922
	Original	0.917	0.921	0.912	0.917	0.908	0.916	0.908	0.916	0.911	0.919	0.911	0.918
DB	FSSi-FSD	**0.901**	**0.904**	**0.902**	0.902	0.901	**0.903**	**0.898**	**0.903**	**0.903**	0.901	**0.901**	**0.903**
	FSI	0.898	0.901	**0.902**	**0.903**	**0.904**	0.900	0.897	0.901	0.901	**0.902**	0.900	0.901
	ARFI	0.896	0.902	0.890	0.902	0.898	0.900	0.892	0.898	0.894	0.901	0.894	0.901
	AVPD	0.870	0.873	0.868	0.876	0.873	0.877	0.871	0.878	0.874	0.878	0.871	0.877
	Original	0.887	0.899	0.898	0.896	0.891	0.900	0.886	0.892	0.891	0.895	0.891	0.896

Table 12
The classification accuracy of four algorithms on twelve data sets when adding objects.

Data set	Algorithm	10%		20%		30%		40%		50%		Average	
		PNN	KNN	PNN	KNN	PNN	KNN	PNN	KNN	PNN	KNN	PNN	KNN
DW	FSSi-FSA	**0.672**	**0.702**	**0.671**	**0.697**	**0.686**	**0.701**	**0.671**	**0.707**	**0.673**	**0.705**	**0.675**	**0.702**
	FSI	0.579	0.663	0.573	0.677	0.586	0.690	0.652	0.670	0.638	0.669	0.606	0.674
	ARFI	0.560	0.671	0.566	0.682	0.571	0.685	0.571	0.587	0.618	0.651	0.577	0.655
	AVPD	0.663	0.683	0.665	0.688	0.685	0.689	0.670	0.695	0.663	0.690	0.669	0.689
	Original	0.535	0.611	0.522	0.602	0.530	0.613	0.543	0.618	0.545	0.603	0.535	0.609
W	FSSi-FSA	**0.957**	**0.960**	**0.956**	**0.955**	**0.959**	**0.949**	0.950	**0.948**	**0.952**	0.941	**0.955**	**0.951**
	FSI	0.942	0.927	0.931	0.916	0.947	0.934	**0.953**	0.922	0.946	**0.946**	0.944	0.929
	ARFI	0.953	0.944	0.944	0.913	0.941	0.924	0.947	0.927	0.946	0.934	0.946	0.928
	AVPD	0.946	0.937	0.929	0.919	0.945	0.924	0.941	0.942	0.942	0.930	0.941	0.930
	Original	0.948	0.923	0.925	0.917	0.932	0.925	0.934	0.923	0.934	0.925	0.935	0.923
H	FSSi-FSA	0.581	**0.587**	0.584	**0.563**	**0.589**	**0.586**	**0.576**	**0.560**	**0.583**	**0.545**	**0.583**	**0.568**
	FSI	**0.585**	0.463	**0.586**	0.455	0.545	0.475	0.575	0.482	0.578	0.436	0.574	0.462
	ARFI	0.537	0.470	0.540	0.464	0.545	0.475	0.536	0.484	0.564	0.454	0.544	0.469
	AVPD	0.564	0.440	0.560	0.454	0.536	0.470	0.550	0.486	0.557	0.436	0.553	0.457
	Original	0.528	0.524	0.530	0.526	0.535	0.511	0.538	0.490	0.533	0.523	0.533	0.515
FTM	FSSi-FSA	**0.808**	**0.846**	**0.797**	**0.802**	**0.779**	**0.810**	**0.763**	**0.819**	0.776	**0.820**	**0.785**	**0.819**
	FSI	0.677	0.741	0.611	0.656	0.633	0.679	0.694	0.732	0.638	0.674	0.651	0.696
	ARFI	0.785	0.817	0.760	0.746	0.723	0.746	0.740	0.789	0.734	0.770	0.748	0.774
	AVPD	0.762	0.784	0.739	0.778	0.732	0.790	0.760	0.802	**0.781**	0.786	0.755	0.788
	Original	0.744	0.778	0.731	0.764	0.747	0.775	0.718	0.779	0.754	0.769	0.739	0.773
TME	FSSi-FSA	0.479	0.499	**0.504**	0.502	**0.495**	0.504	**0.494**	0.503	**0.497**	0.509	**0.494**	0.503
	FSI	0.474	0.450	0.483	0.464	0.399	0.387	0.489	0.430	0.489	0.431	0.467	0.432
	ARFI	0.476	0.478	0.477	0.467	0.466	0.504	0.415	0.419	0.419	0.425	0.451	0.459
	AVPD	0.434	0.415	0.486	0.498	0.470	0.505	0.492	0.500	0.490	0.501	0.474	0.484
	Original	**0.499**	**0.554**	0.499	**0.549**	0.488	**0.581**	0.490	**0.592**	0.493	**0.594**	**0.494**	**0.574**
ACA	FSSi-FSA	0.669	**0.576**	0.673	**0.572**	0.686	**0.570**	0.668	**0.590**	0.706	**0.573**	0.680	**0.576**
	FSI	**0.697**	0.457	**0.699**	0.471	**0.713**	0.451	**0.702**	0.470	**0.716**	0.446	**0.705**	0.459
	ARFI	0.605	0.506	0.570	0.538	0.589	0.562	0.572	0.552	0.614	0.542	0.590	0.540
	AVPD	0.587	0.448	0.592	0.472	0.609	0.471	0.570	0.458	0.599	0.441	0.591	0.458
	Original	0.563	0.440	0.541	0.466	0.563	0.472	0.567	0.463	0.580	0.446	0.563	0.457
PS	FSSi-FSA	**0.525**	**0.505**	**0.506**	**0.482**	**0.516**	**0.495**	**0.504**	**0.491**	**0.514**	**0.497**	**0.513**	**0.494**
	FSI	0.502	0.498	0.500	0.463	0.506	0.472	0.497	0.475	0.506	0.479	0.502	0.477
	ARFI	0.502	0.498	0.500	0.463	0.506	0.472	0.497	0.475	0.506	0.479	0.502	0.477
	AVPD	0.502	0.498	0.500	0.463	0.506	0.472	0.497	0.475	0.506	0.479	0.502	0.477
	Original	0.438	0.377	0.431	0.336	0.435	0.330	0.430	0.337	0.440	0.339	0.502	0.477
MPE	FSSi-FSA	**0.884**	**0.859**	**0.870**	**0.842**	**0.883**	**0.857**	**0.878**	**0.849**	**0.876**	**0.858**	**0.878**	**0.853**
	FSI	0.766	0.731	0.766	0.732	0.757	0.737	0.749	0.717	0.768	0.721	0.761	0.728
	ARFI	0.808	0.829	0.817	0.826	0.814	0.821	0.816	0.818	0.814	0.815	0.814	0.822
	AVPD	0.850	0.814	0.858	0.821	0.859	0.819	0.845	0.820	0.845	0.821	0.851	0.819
	Original	0.787	0.724	0.772	0.734	0.767	0.730	0.753	0.715	0.761	0.711	0.768	0.723
WQR	FSSi-FSA	**0.557**	**0.550**	**0.561**	**0.551**	**0.568**	**0.567**	**0.569**	**0.553**	**0.591**	**0.582**	**0.569**	**0.561**
	FSI	0.507	0.502	0.535	0.516	0.523	0.506	0.543	0.508	0.537	0.515	0.529	0.509
	ARFI	0.540	0.542	0.544	0.522	0.549	0.539	0.552	0.533	0.572	0.567	0.551	0.541
	AVPD	0.531	0.526	0.546	0.523	0.558	0.537	0.552	0.532	0.561	0.553	0.550	0.534
	Original	0.546	0.541	0.549	0.526	0.545	0.541	0.551	0.543	0.570	0.562	0.552	0.543
WQW	FSSi-FSA	**0.557**	**0.520**	**0.556**	0.513	**0.556**	0.519	0.566	**0.543**	**0.574**	0.546	**0.562**	0.528
	FSI	0.490	0.452	0.493	0.422	0.494	0.434	0.474	0.440	0.485	0.453	0.487	0.440
	ARFI	0.550	0.516	0.551	0.521	0.551	0.518	0.565	0.534	0.555	0.548	0.554	0.527
	AVPD	0.541	0.515	0.542	0.511	0.543	0.499	0.505	0.470	0.501	0.464	0.526	0.492
	Original	0.556	0.519	0.553	**0.526**	0.553	**0.532**	**0.572**	0.524	0.561	**0.551**	0.559	**0.530**
PB	FSSi-FSA	**0.925**	**0.931**	**0.925**	**0.930**	**0.924**	**0.929**	**0.925**	**0.929**	**0.925**	**0.927**	**0.925**	**0.929**
	FSI	0.909	0.909	0.908	0.910	0.912	0.914	0.914	0.914	0.922	0.926	0.913	0.915
	ARFI	0.911	0.921	0.914	0.920	0.915	0.923	0.919	0.921	0.914	0.920	0.915	0.921
	AVPD	0.916	0.923	0.923	0.927	0.919	0.926	0.921	0.924	0.916	0.919	0.919	0.924
	Original	0.915	0.919	0.915	0.919	0.914	0.918	0.914	0.918	0.917	0.921	0.915	0.919
DB	FSSi-FSA	**0.902**	**0.905**	0.901	**0.904**	**0.903**	0.902	**0.903**	**0.903**	**0.902**	**0.904**	**0.902**	**0.904**
	FSI	0.901	0.903	**0.904**	0.902	0.902	**0.903**	0.900	**0.903**	**0.902**	0.903	**0.902**	0.903
	ARFI	0.891	0.902	0.889	0.903	0.893	0.898	0.895	0.901	0.897	0.900	0.893	0.901
	AVPD	0.876	0.879	0.872	0.877	0.875	0.878	0.876	0.881	0.875	0.882	0.875	0.879
	Original	0.891	0.898	0.894	0.900	0.893	0.897	0.892	0.899	0.897	0.898	0.893	0.898

Table 13
The quantity of reduct about four algorithms on twelve data sets when deleting or adding objects.

Data set	Algorithm	10%		20%		30%		40%		50%		Average	
		Delete	Add	Delete	Add	Delete	Add	Delete	Add	Delete	Add	Delete	Add
	FSSi-FSD/A	**3.2**	**4.2**	**3.5**	**3.9**	**3.5**	**4.3**	**3.2**	4.9	**3.4**	**4.6**	**3.4**	**4.4**
	FSI	3.5	**4.2**	3.6	4	**3.5**	4.6	3.5	**4.4**	3.6	5.2	3.5	4.5
DW	ARFI	7	5.3	9	6	7.2	7.8	6.2	6.3	4.6	6.8	6.8	6.4
	AVPD	8.9	9.3	9	9.7	8.8	9.5	8.6	9.1	8.6	9.1	8.8	9.3
	Original	174	174	174	174	174	174	174	174	174	174	174	174
	FSSi-FSD/A	**4.7**	**3.5**	**4.6**	**3.1**	**4.6**	**4.8**	**4.3**	**4.9**	**4.9**	**5.4**	**4.6**	**4.3**
	FSI	7	5.7	7.1	6.2	6.8	7.2	6.5	7	5.7	6.8	6.6	6.6
W	ARFI	12	11.7	12	12	12	12	11.9	12	11	12	11.8	11.9
	AVPD	5.5	4.5	4.9	4.5	5	5.6	5	5.6	5.3	5.7	5.1	5.2
	Original	13	13	13	13	13	13	13	13	13	13	13	13
	FSSi-FSD/A	**1.4**	**1**	**1.4**	**1**	**1.2**	**1**	**1.6**	**1**	**1.4**	**1**	**1.4**	**1**
	FSI	3.4	3	3.7	4.2	4.1	3.7	3.9	3.5	3.8	4.5	3.8	3.8
H	ARFI	5.3	6.6	5.9	6	5.8	5.6	5.6	5.5	5	5.4	5.5	5.8
	AVPD	8.6	8.7	9.3	8.8	9.6	9	9	8.8	8.4	9.5	9	9
	Original	13	13	13	13	13	13	13	13	13	13	13	13
	FSSi-FSD/A	7	7	7	7	7	7	7	7	7	7	7	7
	FSI	**2**	**2.9**	**2**	**2.5**	**2.6**	**2.9**	**2.9**	**2.6**	**2.8**	**2**	**2.5**	**2.6**
FTM	ARFI	8.5	8.3	8.5	7.8	9.3	8.3	8.1	8.8	7.7	9.4	8.4	8.5
	AVPD	6.4	6.6	6.4	6.3	6.6	6.4	6.4	6.6	6.4	7	6.4	6.6
	Original	26	26	26	26	26	26	26	26	26	26	26	26
	FSSi-FSD/A	**9**	**6.9**	**9**	**7.7**	**9**	**7.4**	**9**	**7.6**	**9**	**7.5**	**9**	**7.4**
	FSI	15.2	14.5	14.9	13.5	15.5	14.5	13.3	14	12.9	17	14.4	14.7
TME	ARFI	13	10.9	13.3	10	13.1	12.6	13.1	13	10.8	12.1	12.7	11.7
	AVPD	10	9.6	10	10	10.1	10	10.1	10.5	9.3	10	9.9	10
	Original	49	49	49	49	49	49	49	49	49	49	49	49
	FSSi-FSD/A	**2**	**2**	**2**	**2**	**2**	**2**	**2**	**2**	**2**	**2**	**2**	**2**
	FSI	3	3	3	2.8	2.6	2.8	2.8	3	3	2.7	2.9	2.9
ACA	ARFI	5.7	4.6	5	5.1	4.9	5	4.8	5	4.6	5	5	4.9
	AVPD	8.2	8.6	9.4	8.3	9.1	9.2	8.5	9	8.6	10	8.8	9
	Original	13	13	13	13	13	13	13	13	13	13	13	13
	FSSi-FSD/A	**1**	2	**1**	2	**1**	2	**1**	2	**1**	2	**1**	2
	FSI	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**
PS	ARFI	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**
	AVPD	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**
	Original	27	27	27	27	27	27	27	27	27	27	27	27
	FSSi-FSD/A	6.6	7	6	6.5	6.4	5.5	5.9	6	5.8	6.8	6.1	6.4
	FSI	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**
MPE	ARFI	10.5	10.3	8.3	10.5	6.7	10.4	8.6	8	6.6	9	8.1	9.6
	AVPD	7.5	8	7.8	8	8	8	6.5	7.8	7.5	7.4	7.5	7.8
	Original	82	82	82	82	82	82	82	82	82	82	82	82
	FSSi-FSD/A	7	7.5	7	7.4	7	7.5	7	7.7	7	7.8	7	7.6
	FSI	**2**	**2**	**2**	**2**	**2.2**	**2**	**2**	**2**	**2**	**2**	**2**	**2**
WQR	ARFI	10	10	10	10	10	10	10	10	10	10	10	10
	AVPD	6	6	6	6	5.9	6	6	5.7	6	6	6	5.9
	Original	11	11	11	11	11	11	11	11	11	11	11	11
	FSSi-FSD/A	5	5.5	5	5.5	5	5.6	5	5.6	5	5.5	5	5.5
	FSI	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**
WQW	ARFI	10	10	10	10	10	10	10	10	10	10	10	10
	AVPD	5	5	5	5	5	5	5	3	5	2.6	5	4.1
	Original	11	11	11	11	11	11	11	11	11	11	11	11
	FSSi-FSD/A	5	5	5	5	5	5	5	5	5	5	5	5
	FSI	**3**	**1.8**	**2**	**1.8**	**2**	**2**	**2**	**2**	**2**	**2.4**	**2.2**	**2**
PB	ARFI	9	9	9	9	9	9	9	9	9	9	9	9
	AVPD	**3**	3	3	3	3	3	3	3	3	3	3	3
	Original	10	10	10	10	10	10	10	10	10	10	10	10
	FSSi-FSD/A	**4.6**	**4.8**	**4.7**	**4.8**	**4.7**	**4.7**	**4.6**	**4.8**	**4.8**	**5**	**4.7**	**4.8**
	FSI	5	5	5	5	5	5	5	5	5	5	5	5
DB	ARFI	7.7	8	7.8	8	8	8	8	8	8	8.2	7.9	8
	AVPD	4.8	**4.8**	4.8	5	4.8	5	5	**4.8**	**4.8**	5.1	4.8	4.9
	Original	17	17	17	17	17	17	17	17	17	17	17	17

Table 14
The accuracy and quantity on six data sets when deleting or adding objects

Data set	Method	Delete		Add		Data set	Method	Delete		Add	
		Acc	Num	Acc	Num			Acc	Num	Acc	Num
	FSSi-CNN	0.941	0.959	4	4.6		FSSi-CNN	0.870	0.898	7	7
W	CNN	0.928	0.939	14	14	FTM	CNN	0.852	0.864	27	27
	FSSi-CNN	0.731	0.738	10	11.2		FSSi-CNN	0.812	0.818	3	3.4
TME	CNN	0.722	0.734	50	50	ACA	CNN	0.810	0.814	14	14
	FSSi-CNN	0.981	0.984	6.8	7.2		FSSi-CNN	1.000	1.000	5	5.6
MPE	CNN	0.961	0.972	82	82	PB	CNN	0.960	0.968	11	11

data sets almost remain the original information, which means that the combination of incremental feature selection and one-dimensional convolutional network is effective. Other possibilities for both will be further explored in our future research.

7. Conclusions

Feature selection is a potential method for dimension reduction in data mining, which screens out the pivotal attributes that maintain the original information of data set. Making the best of previous knowledge, the incremental learning is pretty appropriate for time-evolving information system. By virtue of developing incremental feature selection approach, it can not only sort out significant attributes and filter redundant attributes, but also speed up the course of feature selection. In this paper, we look back on the indispensable notions of fuzzy rough set, interval-valued fuzzy decision system and self-information initially. Furthermore, the relative decision self-information is introduced into interval-valued fuzzy decision information system, λ-fuzzy similarity relation is constructed, and λ-fuzzy similarity self-information based on above relation is investigated. Moreover, two important updating media, λ-fuzzy similarity matrix and fuzzy decision vector, are defined for static feature selection algorithm in IvFDIS. As a consequence, the updating principles of before-mentioned matrix and vector are deeply explored, then two incremental algorithms in regard to the variation of object set are further researched. Finally, we compare the performance of different algorithms on twelve data sets from UCI. The experimental results reveal that apart from improving the computational efficiency, the proposed incremental algorithms update the reduct promptly in IvFDIS.

Although the incremental technique is beneficial for dynamic IvFDIS, as technology makes momentous progress, the forms of data become more diversified and the alterations are not limited to objects in information systems. Therefore, in our future direction of work, we will be committed to developing the incremental feature selection in interval-valued hesitant fuzzy information system. At the same time, in consideration of various elements that lead to a variation in the information system, we will study the impact of other changes on information system. Additionally, we intend to carry dynamic mechanics into multi-granulation spaces.

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.

Acknowledgement

This work is supported by the National Natural Science Foundation of China (No.61976245, 61772002).

References

[1] M. Boczeka, L.S. Jin, M. Kaluszkaa, Interval-valued seminormed fuzzy operators based on admissible orders, Inf. Sci. 574 (2021) 96–110.
[2] K. Kumara, S.M. Chen, Multiattribute decision making based on converted decision matrices, probability density functions, and interval-valued intuitionistic fuzzy values, Inf. Sci. 554 (2021) 313–324.
[3] L. Maciel, R. Ballini, A fuzzy inference system modeling approach for interval-valued symbolic data forecasting, Knowl.-Based Syst. 164 (15) (2019) 139–149.
[4] X.Y. Zhang, X.W. Chen, W.H. Xu, W.P. Ding, Dynamic information fusion in multi-source incomplete interval-valued information system with variation of information sources and attributes, Inf. Sci. 608 (2022) 1–27.
[5] A. Zadeh, Fuzzy sets, Inf. Control 8 (3) (1965) 338–356.
[6] Z. Pawlak, Rough sets, Int. J. Comput. Inform. Sci. 11 (5) (1982) 341–356.
[7] Z. Pawlak, Rough set: theoretical aspects of reasoning about data, Kluwer Academic Publishers, Dordrecht, 1991.
[8] D. Dubois, H. Prade, Rough fuzzy sets and fuzzy rough sets, Int. J. General Syst. 17 (2–3) (1990) 191–209.
[9] C.X. Jin, J.S. Mi, F.C. Li, M.S. Liang, A novel probabilistic hesitant fuzzy rough set based multi-criteria decision-making method, Inf. Sci. 608 (2022) 489–516.
[10] J.H. Wan, H.M. Chen, T.R. Li, X.L. Yang, B.B. Sang, Dynamic interaction feature selection based on fuzzy rough set, Inf. Sci. 581 (2021) 891–911.
[11] X.L. Yang, H.M. Chen, T.R. Li, P.F. Zhang, C. Luo, Student-t kernelized fuzzy rough set model with fuzzy divergence for feature selection, Inf. Sci. 610 (2022) 52–72.

[12] W.H. Xu, J.H. Yu, A novel approach to information fusion in multi-source datasets: A granular computing viewpoint, Inf. Sci. 378 (2017) 410–423.
[13] W.H. Xu, W.T. Li, Granular computing approach to two-way learning based on formal concept analysis in fuzzy datasets, IEEE Trans. Cybern. 46 (2) (2016) 366–379.
[14] S. Roy, P. Shivakumara, N. Jain, V. Khare, A. Dutta, U. Pal, T. Lu, Rough-fuzzy based scene categorization for text detection and recognition in video, Pattern Recogn. 80 (2018) 64–81.
[15] B.Z. Sun, W.M. Ma, H.Y. Zhao, Decision-theoretic rough fuzzy set model and application, Inf. Sci. 283 (1) (2014) 180–196.
[16] D.G. Chen, Q. He, X.Z. Wang, FRSVMs: fuzzy rough set based support vector machines, Fuzzy Sets Syst. 161 (4) (2010) 596–607.
[17] B.Z. Sun, Z.T. Gong, D.G. Chen, Fuzzy rough set theory for the interval-valued fuzzy information systems, Inf. Sci. 178 (13) (2008) 2794–2815.
[18] W.H. Xu, Q.R. Wang, X.T. Zhang, Multi-granulation fuzzy rough sets in a fuzzy tolerance approximation space, Int. J. Fuzzy Syst. 13 (4) (2011) 246–259.
[19] C.Y. Wang, B.Q. Hu, Granular variable precision fuzzy rough sets with general fuzzy relations, Fuzzy Sets Syst. 275 (2015) 39–57.
[20] W.H. Xu, K.H. Yuan, W.T. Li, W.P. Ding, An Emerging Fuzzy Feature Selection Method Using Composite Entropy-Based Uncertainty Measure and Data Distribution, IEEE Trans. Emerg. Top. Comput. Intell. doi: 10.1109/TETCI.2022.3171784.
[21] W.T. Li, H.X. Zhou, W.H. Xu, X.Z. Wang, W. Pedrycz, Interval Dominance-Based Feature Selection for Interval-Valued Ordered Data, IEEE Trans. Neural Networks Learn. Syst. doi: 10.1109/TNNLS.2022.3184120.
[22] J. Lee, D.W. Kim, Mutual Information-based multi-label feature selection using interaction information, Expert Syst. Appl. 42 (4) (2015) 2013–2025.
[23] Q. He, C.X. Wu, D.G. Chen, S.Y. Zhao, Fuzzy rough set based attribute reduction for information systems with fuzzy decisions, Knowl.-based Syst. 24 (5) (2011) 689–696.
[24] Y.J. Lin, Y.W. Li, C.X. Wang, J.K. Chen, Attribute reduction for multi-label learning with fuzzy rough set, Knowl.-based Syst. 152 (2018) 51–61.
[25] L. Sun, L.Y. Wang, J.C. Xu, Feature selection using fuzzy neighborhood entropy-based uncertainty measures for fuzzy neighborhood multigranulation rough sets, IEEE Trans. Fuzzy Syst. 29 (1) (2021) 19–33.
[26] C.Z. Wang, Y. Huang, M.W. Shao, X.D. Fan, Fuzzy rough set-based attribute reduction using distance measures, Knowl.-based Syst. 164 (2019) 205–212.
[27] X. Zhang, C.L. Mei, D.G. Chen, J.H. Li, Feature selection in mixed data: a method using a novel fuzzy rough set-based information entropy, Pattern Recogn. 56 (2016) 1–15.
[28] W.B. Qian, J.T. Huang, Y.L. Wang, W.H. Shu, Mutual information-based label distribution feature selection for multi-label learning, Knowl.-based Syst. 195 (2020) 105684.
[29] C.E. Shannon, A Mathematical Theory of Communication, Bell Syst. Tech. J. 27 (1948) 379–423.
[30] C.E. Shannon, Communication theory of secrecy systems, Bell Syst. Tech. J. 28 (4) (1949) 656–715.
[31] C.Z. Wang, Y. Huang, W.P. Ding, Z.H. Cao, Attribute reduction with fuzzy rough self-information measures, Inf. Sci. 549 (2021) 68–86.
[32] X. Yang, Y.X. Yang, J.F. Luo, D. Liu, T.R. Li, A unified incremental updating framework of attribute reduction for two-dimensionally time-evolving data, Inf. Sci. 601 (2022) 287–305.
[33] W.P. Ding, T.Z. Qin, X.J. Shen, H.R. Ju, H.P. Wang, J.S. Huang, M. Li, Parallel incremental efficient attribute reduction algorithm based on attribute tree, Inf. Sci. 610 (2022) 1102–1121.
[34] A.P. Zeng, T.R. Li, D. Liu, J.B. Zhang, H.M. Chen, A fuzzy rough set approach for incremental feature selection on hybrid information systems, Fuzzy Sets Syst. 258 (2015) 39–60.
[35] Y.Y. Huang, K.J. Guo, X.W. Yi, Z. Li, T.R. Li, Matrix representation of the conditional entropy for incremental feature selection on multi-source data, Inf. Sci. 591 (2022) 263–286.
[36] W. Wei, X.Y. Wu, J.Y. Liang, J.B. Cui, Y.J. Sun, Discernibility matrix based incremental attribute reduction for dynamic data, Knowl.-based Syst. 140 (2018) 142–157.
[37] M.J. Cai, G.M. Lang, H. Fujita, Z.Y. Li, T. Yang, Incremental approaches to updating reducts under dynamic covering granularity, Knowl.-based Syst. 172 (2019) 130–140.
[38] P. Ni, S.Y. Zhao, X.Z. Wang, H. Chen, C.P. Li, E.C.C. Tsang, Incremental feature selection based on fuzzy rough sets, Inf. Sci. 536 (2020) 185–204.
[39] B.B. Sang, H.M. Chen, T.R. Li, W.H. Xu, H. Y, Incremental approaches for heterogeneous feature selection in dynamic ordered data, Inf. Sci. 541 (2020) 475–501.
[40] B.B. Sang, H.M. Chen, L. Yang, T.R. Li, W.H. Xu, Incremental feature selection using a conditional entropy based on fuzzy dominance neighborhood rough sets, IEEE Trans. Fuzzy Syst. doi: 10.1109/TFUZZ.2021.3064686.
[41] B.B. Sang, H.M. Chen, L. Yang, T.R. Li, W.H. Xu, C. Luo, Feature selection for dynamic interval-valued ordered data based on fuzzy dominance neighborhood rough set, Knowl.-based Syst. 227 (2021) 107223.
[42] J.H. Yu, W.H. Xu, Incremental knowledge discovering in interval-valued decision information system with the dynamic data, Int. J. Mach. Learn. Cybern. 8 (2017) 849–864.
[43] K.H. Yuan, W.H. Xu, W.T. Li, W.P. Ding, An incremental learning mechanism for object classification based on progressive fuzzy three-way concept, Inf. Sci. 584 (1) (2022) 127–147.
[44] W.H. Xu, K.H. Yuan, W.T. Li, Dynamic updating approximations of local generalized multigranulation neighborhood rough set, Appl. Intell. 52 (8) (2022) 9148–9173.
[45] C.Z. Wang, Y. Huang, M.W. Shao, Q.H. Hu, D.G. Chen, Feature selection based on neighborhood self-information, IEEE Trans. Cybern. 50 (9) (2020) 4031–4042.
[46] X. Zhang, X. Liu, Y.Y. Yang, A fast feature selection algorithm by accelerating computation of fuzzy rough set-based information entropy, Entropy 20 (10) (2018) 788.
[47] J.H. Dai, W.T. Wang, J.S. Mi, Uncertainty measurement for interval-valued information systems, Inf. Sci. 251 (2013) 63–78.
[48] J.H. Dai, H. Hu, G.J. Zheng, Q.H. Hu, H.F. Han, H. Shi, Attribute reduction in interval-valued information systems based on information entropies, Front. Inf. Technol. Electron. Eng. 17 (2016) 919–928.
[49] Y. Lecun, Generalization and network design strategies, Connectionism in perspective 19 (1989) 143–155.
[50] A. Krizhevsky, I. Sutskever, G.E. Hinton, Imagenet classification with deep convolutional neural networks, Commun. ACM 60 (6) (2012) 84–90.