Information Sciences 709 (2025) 122080

Contents lists available at ScienceDirect

Information Sciences

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

Approaches to attribute reduction of metric-fuzzy decision systems based on information theory

Guirong Peng [a], Fei Li [a,*], Wei Yao [b]

[a] College of Science, Beijing Forestry University, Beijing 100083, China
[b] School of Mathematics and Statistics, Nanjing University of Information Science and Technology, Nanjing 210044, China

ARTICLE INFO

Keywords:
Fuzzy rough set
Metirc fuzzy information systems
Fuzzy rough mutual information
Attribute reduction

ABSTRACT

Fuzzy rough sets and information theory are both effective tools for processing large-scale data. This study combines the advantages of both to establish the attribute reduction theory and the method of metric fuzzy information systems based on information theory. First, it defines the concepts of metric fuzzy rough entropy, fuzzy joint rough entropy and fuzzy rough mutual information, explores their properties and relationships and constructs a method for evaluating attribute importance. Secondly, based on this theoretical foundation, two efficient attribute reduction algorithms are designed: the first algorithm doesn't rely on decision attributes, and its advantage is that it can effectively improve the computational efficiency; the second algorithm combines decision attributes, and its advantage is that it can optimize the reduction effect. Both algorithms use the strategies of forward selection and backward elimination to eliminate redundant attributes. Finally, this paper compares these two reduction algorithms with five commonly used reduction algorithms on 20 datasets and uses the average classification accuracy of 14 classifiers to evaluate the reduction effects of these algorithms. Experimental results show that the two algorithms proposed in this paper perform well in classification tasks, with their average accuracy ranking among the highest compared to other algorithms, thus verifying the efficiency and advantages of the reduction algorithms of metric fuzzy information systems in large-scale data processing.

1. Introduction

With the rapid development of data science, the ability to process and analyze large-scale and high-dimensional datasets has become a critical issue. Particularly in the background of big data, attribute reduction techniques have emerged as a key research focus due to their significant role in reducing data dimensions and improving processing efficiency [1–3]. These techniques work by eliminating redundant or irrelevant attributes while retaining the essential features necessary for analysis tasks, thereby making algorithms more efficient and interpretable.

Rough set theory, proposed by Pawlak in 1982 [4], is a powerful tool for handling uncertainty and incomplete information without the need for external information. By using the concepts of lower and upper approximation sets, rough sets effectively characterize uncertainties within sets, demonstrating outstanding advantages in attribute reduction and rule extraction. This has led to widespread applications in data analysis, granular computing, big data mining and control systems, showcasing strong adaptability

* Corresponding author.
E-mail address: feifei_1004@bjfu.edu.cn (F. Li).

https://doi.org/10.1016/j.ins.2025.122080
Received 7 January 2024; Received in revised form 1 February 2025; Accepted 8 March 2025

Available online 13 March 2025

and practicality [5–7]. However, the strict limitation that classical rough set theory relies only on equivalence relations, exhibits certain deficiencies, particularly in numerical data processing and uncertainty reduction. To overcome these problems, researchers have proposed extended information system reduction theories based on covering relations [8], neighborhood relations [9] and fuzzy relations [10]. These extensions not only enhance the ability of rough set theory to handle complex and fuzzy data but also significantly improve its flexibility in managing uncertainty and fuzziness. In particular, Dubois and Prade combine fuzzy theory and rough set theory to propose the fuzzy rough set theory [11]. This theory substantially enhances the management of complex and fuzzy data and further develops fuzzy information systems by introducing fuzzy relationships. By integrating fuzzy set theory with information systems, fuzzy information systems excel in handling data fuzziness and uncertainty. Unlike traditional information systems, attribute values in fuzzy information systems are no longer precise but represented through fuzzy membership functions. That is, the value of a certain attribute of an object can partially belong to a set. This representation is more in line with the data characteristics of the real world, thus more accurately describing the fuzzy similarity between objects. Fuzzy information systems have demonstrated strong adaptability and practicality when dealing with complex and fuzzy data, playing an important role especially in data classification and pattern recognition.

Combining information entropy and rough set theory, researchers construct a powerful data analysis framework capable of effectively evaluating the importance of attributes [12–14]. However, challenges persist when processing large datasets, particularly in fuzzy information systems with uncertain and fuzzy data. To address these problems, fuzzy entropy and knowledge granulation have become key research directions. By integrating information entropy with fuzzy rough set theory, a more robust data analysis framework is formed. The enhanced framework not only improves the ability to evaluate attribute importance but also better addresses the inherent fuzziness and uncertainty in fuzzy information systems [15,16]. Recently, the applications of fuzzy conditional entropy, fuzzy joint entropy and mutual information have further highlighted the importance of fuzzy entropy in fuzzy rough set theory [17–19]. For example, Dai et al. enhance the performance of models to handle complex data using a maximum relevance and independence method based on fuzzy entropy [20]. Based on pessimistic multi-granulation fuzzy neighborhood conditional entropy, Sun et al. develop a feature selection algorithm, which significantly improves classification performance on various datasets, demonstrating the potential of the information entropy in large-scale complex data processing [21].

Generally, in attribute reduction using fuzzy rough sets based on information entropy, logical connectives are commonly used to compute fuzzy similarity relationships. However, research has shown that this method may become inaccurate when handling large attribute sets. The main issue is the use of logical connectives, which may cause the logarithmic parameters to approach zero [22–24], leading to information loss during the entropy computation. Recent studies, including [25–27] in 2023-2024, have introduced new methods, such as advanced similarity measures and entropy-based models, to enhance the robustness of fuzzy attribute reduction. These methods aim to optimize fuzzy similarity relationships through novel measures and refine entropy calculations, showing promise in handling high-dimensional and complex data. However, these approaches often face challenges related to computational complexity and scalability when dealing with large attribute sets. To address this issue, Yao et al. introduce a method based on distance measures to avoid information loss, demonstrating the effectiveness of metric fuzzy rough set theory [28–30]. Unlike the aforementioned methods, our approach directly incorporates distance-based entropy calculations, making it computationally efficient and more stable when processing large-scale fuzzy data. Therefore, in many fuzzy rough set models, fuzzy information granules based on distance measures are more stable and effective, offering significant advantages in handling complex and large datasets while maintaining computational feasibility. To provide a clearer overview of previous research, this paper has summarized the following key references and presented them in Table 1 in tabular form to better understand the research methods, contributions, and applicable domains of each study.

Based on the background mentioned above, this paper first introduces a novel entropy calculation method based on distance measures within the fuzzy rough set framework. This method overcomes the problem of information loss in traditional methods, enhances the accuracy and stability of entropy calculations and significantly improves the precision of information granulation. Furthermore, we extend the fuzzy rough set theory by proposing new entropy concepts, including fuzzy joint rough entropy, fuzzy conditional rough entropy and fuzzy rough mutual information. These concepts enrich the theory and enhance the ability to characterize attribute correlations and uncertainty, thereby increasing the framework's capacity of handling complex datasets. On the basis of these theories, we develop two attribute reduction algorithms based on metric fuzzy rough entropy and metric fuzzy rough mutual information. These algorithms effectively handle large datasets, enhancing feature selection efficiency and model classification performance while preserving essential data. Experimental results demonstrate their effectiveness and advantages in handling large datasets, highlighting their great potential in practical classification applications.

To provide a clear and concise summary of the core contributions of this paper, a graphical abstract is included in Fig. 1. This graphical abstract highlights the research objective, methodology, key findings, and conclusions of the study. It outlines the primary goal of solving the information loss in traditional fuzzy rough set attribute reduction methods and the development of new algorithms for improving large-scale data processing efficiency and accuracy. Additionally, it provides a visual summary of the experimental results, comparing the performance of the proposed algorithms with several commonly used algorithms across various datasets.

2. Preliminaries

In the field of fuzzy rough sets and computational science, the concepts of distance and similarity are both of high importance, with similarity usually characterized by equivalence relations, and distance usually regarded as the opposite of similarity. This view not only enriches the theory but also provides an innovative method for analyzing the relationship between objects [28].

Table 1
Literature Review and Research Methods.

Author	Research Method	Key Contributions	Evaluation Metrics	Datasets	Applicable Domains
Pawlak (1982)	Rough Set Theory.	Introduced rough set theory, defining lower and upper approximations.	/	/	Approximation operations and tolerance theory, applied in equivalence and inclusion.
Dubois and Prade (1987)	Fuzzy Rough Set.	Proposed fuzzy rough set theory to enhance fuzzy data processing.	/	/	Applications in modal logic, fuzzy inference, uncertainty modeling.
Yao et al. (2019)	Metric-based L-fuzzy Rough Sets.	Introduced metric-based fuzzy rough sets to prevent information loss.	/	/	L-fuzzy rough set theory, particularly in fuzzy clustering and weighted graphs.
Dai et al. (2020)	Method Based on Maximum Relevance and Independence.	Proposed a method based on fuzzy entropy to improve relevance and independence.	Reduct size, Accuracy.	9 datasets from UCI, 2 from NCBI, and 4 from the Kent Ridge Biomedical Dataset.	Feature selection based on fuzzy conditional mutual information, especially in tumor classification.
Sun et al. (2021)	Feature Selection Algorithm Based on Fuzzy Neighborhood Conditional Entropy.	Developed a feature selection algorithm based on fuzzy neighborhood entropy.	Reduct size, Accuracy, Friedman Test, CD test.	7 datasets from UCI and 5 from gene expression datasets.	Feature selection and classification for heterogeneous datasets with fuzzy neighborhoods.
Xie et al. (2023)	Local Information Entropy to Improve Attribute Reduction Efficiency.	Enhanced entropy to improve efficiency while maintaining accuracy.	Reduct size, Precision, Running time.	6 datasets from UCI.	Pattern recognition, data mining, knowledge.
Xu et al. (2024)	Statistical Distribution and KL Divergence-Based Similarity Measurement.	Proposed a similarity measure to avoid loss, improving classification of interval data.	Average Precision, Average Quality, Friedman Test, CD test.	12 datasets from UCL.	Multi-source data fusion and interval-valued data analysis, decision support systems.

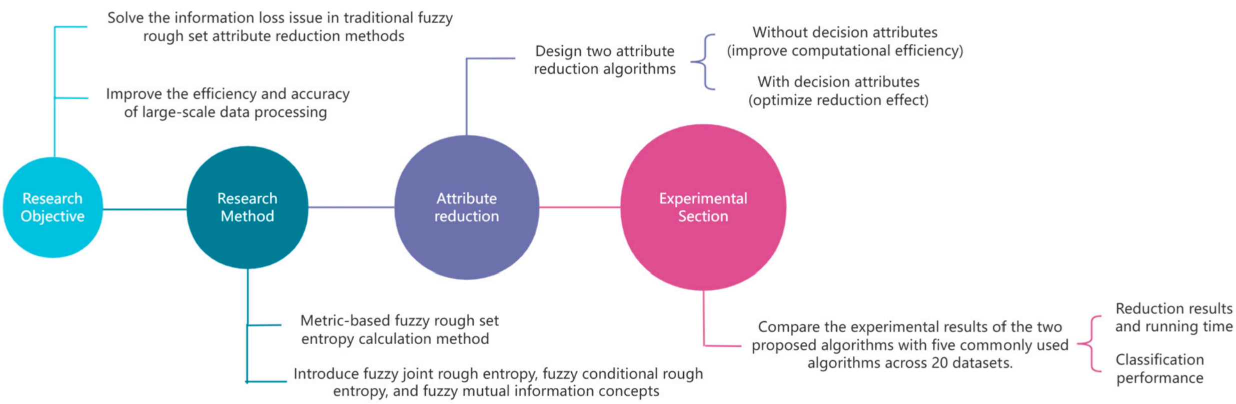


Fig. 1. Research Framework of the Study.

Definition 1. Let U be a nonempty set. A mapping $d : U \times U \to [0, +\infty)$ is called a hemimetric on U if

- $d(x,x) = 0$ for all $x \in U$.
- $d(x,z) \le d(x,y) + d(y,z)$ for all $x, y, z \in U$.

The pair (X,d) is called a hemimetric space.

Definition 2 *([29]).* Let (X,d) be an hemimetric space and $S \subseteq X$. Define two operators $\overline{Apr}_d$, $\underline{Apr}_d : \mathcal{F}(X) \to \mathcal{F}(X)$ respectively by, for x in X:

$$\underline{Apr}_d(S)(x) = \bigwedge_{y \in X} (S(y) + d(x,y)); \tag{1}$$

$$\overline{Apr}_d(S)(x) = \bigvee_{y\in X}(S(y) - d(y,x)); \tag{2}$$

The operators $\underline{Apr}_d$, $\overline{Apr}_d$: $\mathcal{F}(X) \to \mathcal{F}(X)$ are called the fuzzy upper rough approximation operator and the fuzzy lower rough approximation operator on X induced by the hemimetric d, respectively.

Rough entropy, a concept first introduced by Yao in Reference [31], is also referred to as co-entropy in Reference [32]. The concept of rough entropy is defined within the framework of an approximate space, as described below.

Definition 3 *([33]).* Let U be a nonempty finite set, $X \subseteq U$, and $U/R = \{X_1, X_2, \ldots, X_k\}$, then the rough entropy of approximate space (U, R) is defined as

$$E(R) = -\sum_{i=1}^{k} \frac{1}{|U|}\log_2 \frac{1}{|X_i|}, \tag{3}$$

where $E : R \to [0,\infty)$ is a mapping from all partitions of U about equivalence relation R to nonnegative real numbers.

This definition represents the rough entropy of a rough space divided by a crisp equivalence relation. Additionally, the definition overlooks the metric operations between attributes.

3. The information entropy in metric fuzzy information systems

This section integrates the concept of fuzzy rough entropy into the metric fuzzy information systems, explores the concepts of fuzzy joint rough entropy, fuzzy conditional rough entropy and fuzzy rough mutual information, and the characteristics of fuzzy rough entropy based on measurement. To ensure the accuracy of data processing results, this study performs min-max normalization on the initial numerical attribute values, converting their range to values between 0 and 1.

To better handle data characterized by numerical, nominal, or a combination of these attributes, a hybrid similarity measure is established in the following manner.

Definition 4. For $a \in C$, where C is the set of attributes, and $u_i, u_j \in U$, with U being the universe of objects. The fuzzy binary relation $R_a(u_i,u_j)$ between u_i and u_j with respect to attribute a is defined as

$$R_a(u_i,u_j) = \begin{cases} 0, & \text{if } a(u_i) = a(u_j) \text{ and } a \text{ is nominal;} \\ 1, & \text{if } a(u_i) \neq a(u_j) \text{ and } a \text{ is nominal;} \\ |a(u_i) - a(u_j)|, & \text{if } a \text{ is numerical.} \end{cases} \tag{4}$$

The fuzzy binary relation initiated by a is symbolized as R_a, and the corresponding fuzzy relation matrix is represented by $M_{R_a} = (r_{ij})_{nxn}$, with $(r_{ij})_{nxn}$ denoting $R_a(u_i,u_j)$.

Suppose $F_1 = \{a_1, a_2, ..., a_{|F_1|}\} \subseteq C$, then $R_{F_1}(u_i,u_j) = \bigvee_{k=1}^{|F_1|} R_{a_k}(u_i,u_j)$. (5)

Definition 5. The fuzzy granule of $u \in U$ induced by $F_1 \subseteq C$ is defined as

$$[u]_{R_{F_1}} = \frac{R_{F_1}(u,u_1)}{u_1} + \frac{R_{F_1}(u,u_2)}{u_2} + \ldots + \frac{R_{F_1}(u,u_n)}{u_n}, \tag{6}$$

where $R_{F_1}(u,u_i) = 0$ means that u_i is indiscernible with respect to the relation R_{F_1} and belongs to the equivalence class, while $R_{F_1}(u,u_i) = 1$ means u_i does not belong to the equivalence class. The cardinality of $[u]_{R_{F_1}}$ is defined as $|[u]_{R_{F_1}}| = \sum_{j=1}^{n} R_{F_1}(u,u_j)$.

Proposition 6. *if $F_1 \subseteq F_2 \subseteq C$, then $R_{F_1} \subseteq R_{F_2}$.*

Proof. It is easy to prove that is true.

$$F_1 \subseteq F_2 \Rightarrow \forall u_i,u_j \in U, R_{F_1}(u_i,u_j) = \bigvee_{k=1}^{|F_1|} R_{a_k}(u_i,u_j) \le \bigvee_{k=1}^{|F_2|} R_{a_k}(u_i,u_j) = R_{F_2}(u_i,u_j)$$

and

$$\forall u_j \in U, R_{F_1}(u_i,u_j) = \bigvee_{k=1}^{|F_1|} R_{a_k}(u_i,u_j) \le \bigvee_{k=1}^{|F_2|} R_{a_k}(u_i,u_j) = R_{F_2}(u_i,u_k)$$

$\Rightarrow R_{F_1} \subseteq R_{F_2}$.

Expanding on the metric fuzzy rough set framework, the notion of rough entropy is integrated into the fuzzy rough set theory. This integration leads to the formulation of fuzzy rough entropy specifically for subsets of attributes.

Definition 7. The fuzzy rough entropy of F_1 is defined as

$$E(F_1) = E(R_{F_1}) = -\frac{1}{n}\sum_{i=1}^{n} \log_2 \frac{1}{|[u_i]_{R_{F_1}}|}. \tag{7}$$

It is easy to get $0 \leq E(F_1) \leq \infty$, if and only if $\forall u_i, u_j \in U, R_{F_1}(u_i, u_j) = 0$,

$$|[u_i]_{R_{F_1}}| = 0, \text{ so } E(F_1) = \infty,$$

in this case all object pairs are indistinguishable. Therefore, the granulation space is the coarsest at this time. On the other hand, for any $u_i \in U$, there exists a unique $u_l \in U$ such that $R_{F_1}(u_i, u_l) = 1$, and for all other $u_j \in U$ where $u_j \neq u_i$ and $u_j \neq u_l$, $R_{F_1}(u_i, u_j) = 0$, that is $|[u_i]_{R_{F_1}}| = 1$. So

$$E(F_1) = -\frac{1}{n}(n \log_2 1) = 0,$$

at this time, the granulation space is the smallest.

Proposition 8. *If $R_{F_1} \subseteq R_{F_2}$, then $E(R_{F_1}) \leq E(R_{F_2})$.*

Proof. Since $R_{F_1} \subseteq R_{F_2}$, then $\forall u_i, u_j \in U, R_{F_1}(u_i, u_j) \leq R_{F_2}(u_i, u_j)$. So $|[u_i]_{R_{F_1}}| \leq |[u_i]_{R_{F_2}}|$, therefore, by Definition 7, $E(R_{F_1}) \leq E(R_{F_2})$.

In Proposition 6 and 8, they suggest that as the size of the attribute subset diminishes, the fuzzy rough entropy correspondingly decreases. Conversely, the fuzzy rough entropy increases with the expansion of the attribute subset. Furthermore, the fuzzy rough entropy as outlined in Definition 7 serves as a tool to quantify the inherent uncertainty in the fuzzy approximation space.

Definition 9. The fuzzy joint rough entropy of F_1 and F_2 is defined as

$$E(F_1, F_2) = E(R_{F_1 \cup F_2}) = -\frac{1}{n}\sum_{i=1}^{n} \log_2 \frac{1}{\left|[u_i]_{R_{F_1}} \cup [u_i]_{R_{F_2}}\right|}. \tag{8}$$

Proposition 10. *Suppose $F_1, F_2 \subseteq C$, then $E(F_1, F_2) \geq \max\{E(F_1), E(F_2)\}$.*

Proof. Since $\forall u_i \in U$, $|[u_i]_{R_{F_1}} \cup [u_i]_{R_{F_2}}| = \sum_{j=1}^{n} \max(R_{F_1}(u_i, u_j), R_{F_2}(u_i, u_j))$, then $|[u_i]_{R_{F_1}} \cup [u_i]_{R_{F_2}}| \geq |[u_i]_{R_{F_1}}|, |[u_i]_{R_{F_1}} \cup [u_i]_{R_{F_2}}| \geq |[u_i]_{R_{F_2}}|$. By Definition 9,

$$E(F_1, F_2) \geq E(F_1), E(F_1, F_2) \geq E(F_2).$$

Therefore, $E(F_1, F_2) \geq \max\{E(F_1), E(F_2)\}$.

Proposition 11. *if $F_1 \subseteq F_2 \subseteq C$, then $E(F_1, F_2) = E(F_2)$.*

Proof. Since $F_1 \subseteq F_2$, by Proposition 6 $R_{F_1} \subseteq R_{F_2}$. Then $|[u_i]_{R_{F_1}}| \subseteq |[u_i]_{R_{F_2}}|$. So $|[u_i]_{R_{F_1}} \cup [u_i]_{R_{F_2}}| = |[u_i]_{R_{F_2}}|$. By Definition 9, $E(F_1, F_2) = E(F_2)$.

In Proposition 11, the fuzzy joint rough entropy of a set and its subset is equal to the fuzzy rough entropy of the set.

Definition 12. The fuzzy conditional rough entropy of F_2 conditioned F_1 is defined as

$$E(F_2|F_1) = -\frac{1}{n}\sum_{i=1}^{n} \log_2 \frac{|[u_i]_{R_{F_1}}|}{|[u_i]_{R_{F_1}} \cup [u_i]_{R_{F_2}}|}. \tag{9}$$

Proposition 13. $E(F_2|F_1) = E(F_1, F_2) - E(F_1)$.

Proof. By (7)–(9),

$$\forall u_i \in U, E(F_1,F_2) - E(F_1) = \left(-\frac{1}{n}\sum_{i=1}^{n}\log_2\frac{1}{|[u_i]_{R_{F_1}} \cup [u_i]_{R_{F_2}}|}\right) - \left(-\frac{1}{n}\sum_{i=1}^{n}\log_2\frac{1}{|[u_i]_{R_{F_1}}|}\right)$$
$$= \frac{1}{n}\sum_{i=1}^{n}\log_2\frac{1}{|[u_i]_{R_{F_1}}|} - \frac{1}{n}\sum_{i=1}^{n}\log_2\frac{1}{|[u_i]_{R_{F_1}} \cup [u_i]_{R_{F_2}}|}$$
$$= \frac{1}{n}\sum_{i=1}^{n}\log_2\frac{|[u_i]_{R_{F_1}} \cup [u_i]_{R_{F_2}}|}{|[u_i]_{R_{F_1}}|} = -\frac{1}{n}\sum_{i=1}^{n}\log_2\frac{|[u_i]_{R_{F_1}}|}{|[u_i]_{R_{F_1}} \cup [u_i]_{R_{F_2}}|}$$
$$= E(F_2|F_1).$$

Proposition 14. *If $F_2 \subseteq F_1 \subseteq C$, then $E(F_2|F_1) = 0$.*

Proof. Since $F_2 \subseteq F_1$, by Proposition 6, $R_{F_2} \subseteq R_{F_1}$, then $\forall u_i, [u_i]_{R_{F_2}} \subseteq [u_i]_{R_{F_1}}$. Then $\forall u_i, [u_i]_{R_{F_2}} \cup [u_i]_{R_{F_1}} = [u_i]_{R_{F_1}}$. So by Definition 12,

$$E(F_2|F_1) = -\frac{1}{n}\sum_{i=1}^{n}\log_2\frac{|[u_i]_{R_{F_1}}|}{|[u_i]_{R_{F_1}}|} = 0.$$

Fuzzy rough mutual information can effectively evaluate not only the uncertainty within the fuzzy approximation spaces, but also the correlation between attributes and their respective classes.

Definition 15. The fuzzy rough mutual information of F_2 and F_1 is calculated by

$$I(F_2;F_1) = I(F_1;F_2) = -\frac{1}{n}\sum_{i=1}^{n}\log_2\frac{|[u_i]_{R_{F_1}} \cup [u_i]_{R_{F_2}}|}{|[u_i]_{R_{F_1}}||[u_i]_{R_{F_2}}|}. \tag{10}$$

Definition 16. The fuzzy rough mutual information between D and F_1 is calculated by

$$I(D;F_1) = -\frac{1}{n}\sum_{i=1}^{n}\log_2\frac{|[u_i]_{R_{F_1}} \cup [u_i]_{D}|}{|[u_i]_{R_{F_1}}||[u_i]_{D}|}. \tag{11}$$

In Definition 16, the fuzzy rough mutual information, denoted as $I(D;F_1)$, describes the correlation between the attribute subset F_1 and the decision attribute D. It implies that the higher value of fuzzy rough mutual information between D and F_1, the stronger the correlation between F_1 and D.

Proposition 17. $I(F_2;F_1) = E(F_2) - E(F_2|F_1) = E(F_1) - E(F_1|F_2)$.

Proof. By (7) and (9),

$$E(F_2) - E(F_2|F_1) = \left(-\frac{1}{n}\sum_{i=1}^{n}\log_2\frac{1}{|[u_i]_{R_{F_2}}|}\right) - \left(-\frac{1}{n}\sum_{i=1}^{n}\log_2\frac{|[u_i]_{R_{F_1}}|}{|[u_i]_{R_{F_1}} \cup [u_i]_{R_{F_2}}|}\right)$$
$$= \frac{1}{n}\sum_{i=1}^{n}\left[\log_2\frac{|[u_i]_{R_{F_1}}|}{|[u_i]_{R_{F_1}} \cup [u_i]_{R_{F_2}}|} - \log_2\frac{1}{|[u_i]_{R_{F_2}}|}\right]$$
$$= \frac{1}{n}\sum_{i=1}^{n}\log_2\frac{|[u_i]_{R_{F_1}}||[u_i]_{R_{F_2}}|}{|[u_i]_{R_{F_1}} \cup [u_i]_{R_{F_2}}|}$$
$$= -\frac{1}{n}\sum_{i=1}^{n}\log_2\frac{|[u_i]_{R_{F_1}}| \cup |[u_i]_{R_{F_2}}|}{|[u_i]_{R_{F_1}}||[u_i]_{R_{F_2}}|} = I(F_2;F_1).$$

So, by Definition 15, $I(F_2;F_1) = E(F_2) - E(F_2|F_1)$ holds. In a similar manner, $I(F_2;F_1) = E(F_1) - E(F_1|F_2)$.

Proposition 18. *If $F_2 \subseteq F_1 \subseteq C$, then $I(F_2;F_1) = E(F_2)$.*

Table 2
Data of Example.

Object	a_1	a_2
u_1	0.8	0.4
u_2	0.5	0.6
u_3	0.2	0.7
u_4	0.6	0.3
u_5	0.4	0.8

Proof. By Propositions 14 and 17, if $F_2 \subseteq F_1$ then $E(F_2|F_1) = 0$, and therefore:

$$I(F_2; F_1) = E(F_2) - E(F_2|F_1) = E(F_2) - 0$$

Thus, $I(F_2; F_1)$ is equal to $E(F_2)$ as required.

Proposition 19. $I(F_2; F_1) = I(F_1; F_2) = E(F_1) + E(F_2) - E(F_1, F_2)$.

Proof. By (7), (8) and (10),

$$\begin{aligned}
&E(F_1) + E(F_2) - E(F_1, F_2) \\
&= -\frac{1}{n}\sum_{i=1}^{n} \log_2 \frac{1}{|[u_i]_{R_{F_1}}|} - \frac{1}{n}\sum_{i=1}^{n} \log_2 \frac{1}{|[u_i]_{R_{F_2}}|} + \frac{1}{n}\sum_{i=1}^{n} \log_2 \frac{1}{|[u_i]_{R_{F_1}} \cup [u_i]_{R_{F_2}}|} \\
&= -\frac{1}{n}\sum_{i=1}^{n} \left[\log_2 \frac{1}{|[u_i]_{R_{F_1}}|} + \log_2 \frac{1}{|[u_i]_{R_{F_2}}|} - \log_2 \frac{1}{|[u_i]_{R_{F_1}} \cup [u_i]_{R_{F_2}}|} \right] \\
&= \frac{1}{n}\sum_{i=1}^{n} \log_2 \frac{|[u_i]_{R_{F_1}}||[u_i]_{R_{F_2}}|}{|[u_i]_{R_{F_1}} \cup [u_i]_{R_{F_2}}|} = -\frac{1}{n}\sum_{i=1}^{n} \log_2 \frac{|[u_i]_{R_{F_1}}| \cup |[u_i]_{R_{F_2}}|}{|[u_i]_{R_{F_1}}||[u_i]_{R_{F_2}}|} \\
&= I(F_1; F_2).
\end{aligned}$$

In Proposition 19, the fuzzy rough mutual information $I(D; F_1)$ represents the shared fuzzy rough information between the attribute subset F_1 and the decision attribute D. If the value of $I(D; F_1)$ equals $I(D; C)$, the correlation between the attribute subset F_1 and D equals the correlation between the complete attribute set C and D. Consequently, the attribute subset F_1 demonstrates effective classification capabilities in metric fuzzy decision information systems, particularly in systems where the complete attribute set C contains redundant attributes. Additionally, F_1 can be considered a more appropriate reduction subset in the metric fuzzy decision information system. Based on this proposition, an attribute reduction algorithm using fuzzy rough mutual information is proposed.

In this section, a small example with 5 objects and 2 attributes is provided to better illustrate the calculations and meanings of the definitions presented. The data set is shown in Table 2.

For attributes a_1 and a_2, the fuzzy binary relation is calculated as follows:

$$\begin{aligned}
&R_{a_1}(u_1, u_2) = |0.8 - 0.5| = 0.3 && R_{a_2}(u_1, u_2) = |0.4 - 0.6| = 0.2 \\
&R_{a_1}(u_1, u_3) = |0.8 - 0.2| = 0.6 && R_{a_2}(u_1, u_3) = |0.4 - 0.7| = 0.3 \\
&R_{a_1}(u_1, u_4) = |0.8 - 0.6| = 0.2 && R_{a_2}(u_1, u_4) = |0.4 - 0.3| = 0.1 \\
&R_{a_1}(u_1, u_5) = |0.8 - 0.4| = 0.4 && R_{a_2}(u_1, u_5) = |0.4 - 0.8| = 0.4 \\
&R_{a_1}(u_2, u_3) = |0.5 - 0.2| = 0.3 && R_{a_2}(u_2, u_3) = |0.6 - 0.7| = 0.1 \\
&R_{a_1}(u_2, u_4) = |0.5 - 0.6| = 0.1 && R_{a_2}(u_2, u_4) = |0.6 - 0.3| = 0.3 \\
&R_{a_1}(u_2, u_5) = |0.5 - 0.4| = 0.1 && R_{a_2}(u_2, u_5) = |0.6 - 0.8| = 0.2 \\
&R_{a_1}(u_3, u_4) = |0.2 - 0.6| = 0.4 && R_{a_2}(u_3, u_4) = |0.7 - 0.3| = 0.4 \\
&R_{a_1}(u_3, u_5) = |0.2 - 0.4| = 0.2 && R_{a_2}(u_3, u_5) = |0.7 - 0.8| = 0.1 \\
&R_{a_1}(u_4, u_5) = |0.6 - 0.4| = 0.2 && R_{a_2}(u_4, u_5) = |0.3 - 0.8| = 0.5
\end{aligned}$$

This relation measures the similarity or dissimilarity between objects for each attribute. A smaller value indicates higher similarity, while a larger value indicates greater dissimilarity. This is crucial for computing fuzzy rough entropy and mutual information.

By using attribute a_1, we compute the fuzzy rough entropy. The cardinalities of the fuzzy granules are:

$$|[u_1]_{R_{a_1}}| = 1.5, \quad |[u_2]_{R_{a_1}}| = 0.8, \quad |[u_3]_{R_{a_1}}| = 1.5,$$
$$|[u_4]_{R_{a_1}}| = 0.9, \quad |[u_5]_{R_{a_1}}| = 0.9$$

Now, the fuzzy rough entropy for $F_1 = \{a_1\}$ is calculated as:

$$E(F_1) = -\frac{1}{5}\left(\log_2 \frac{1}{1.5} + \log_2 \frac{1}{0.8} + \log_2 \frac{1}{1.5} + \log_2 \frac{1}{0.9} + \log_2 \frac{1}{0.9}\right) = 0.109$$

Fuzzy rough entropy quantifies the uncertainty in the partition of objects under the fuzzy relation. A higher entropy value indicates greater uncertainty in the attribute set, while a lower value suggests a more certain partition. It is useful in evaluating how effective attributes are in distinguishing between objects.

The fuzzy conditional rough entropy of $F_2 = \{a_2\}$ conditioned on F_1 is calculated as:

$$E(F_2|F_1) = -\frac{1}{5}\sum_{i=1}^{5} \log_2 \frac{|[u_i]_{R_{a_1}}|}{|[u_i]_{R_{a_1}} \cup [u_i]_{R_{a_2}}|}$$

For object u_1, we can obtain:

$$|[u_1]_{R_{a_1}}| = 1.5, \quad |[u_1]_{R_{a_1} \cup a_2}| = \sum_{i=2}^{5} \max(R_{a_1}(u_1, u_i), R_{a_2}(u_1, u_i)) = 1.5$$

Now, we calculate the fuzzy conditional rough entropy:

$$E(F_2|F_1) = -\frac{1}{5}\left[\log_2 \frac{1.5}{1.5} + \log_2 \frac{0.8}{1.1} + \log_2 \frac{1.5}{1.5} + \log_2 \frac{0.9}{1.4} + \log_2 \frac{0.9}{1.3}\right] = 0.925$$

Fuzzy conditional rough entropy measures the remaining uncertainty in one attribute set (F_2) given another attribute (F_1). It is useful for analyzing dependencies between attributes and feature selection in multi-attribute systems.

Finally, the fuzzy rough mutual information between F_2 and F_1 is calculated as:

$$I(F_2; F_1) = -\frac{1}{5}\sum_{i=1}^{5} \log_2 \frac{|[u_i]_{R_{F_1}} \cup [u_i]_{R_{F_2}}|}{|[u_i]_{R_{F_1}}||[u_i]_{R_{F_2}}|}$$

For object u_1:

$$\log_2 \frac{|[u_1]_{R_{F_1}} \cup [u_1]_{R_{F_2}}|}{|[u_1]_{R_{F_1}}||[u_1]_{R_{F_2}}|} = \log_2 \frac{1.5}{1.5 \times 1.0} = 0$$

Substitute these values and perform the summation:

$$I(F_2; F_1) = -\frac{1}{5}\left(\log_2 \frac{1.5}{1.5 \times 1.0} + \log_2 \frac{1.1}{0.8 \times 0.8} + \log_2 \frac{1.5}{1.5 \times 0.9} + \log_2 \frac{1.4}{0.9 \times 1.3} + \log_2 \frac{1.3}{1.3 \times 0.9}\right) = -0.17$$

Fuzzy rough mutual information quantifies the degree of shared information between two attribute sets. It helps identify how much one attribute set reveals about another, which is important for feature selection and dimensionality reduction.

4. Heuristic attribute reduction algorithm based on metric fuzzy rough entropy

In this section, the definitions of inner and outer significance functions for attributes are introduced, aimed at evaluating the importance of different attributes. Based on these two types of significance functions, two corresponding algorithms for attribute reduction are proposed.

Proposition 20. *If $F_1 \subseteq F_2 \subseteq C$, then the size relationship between $I(D; F_1)$ and $I(D; F_2)$ cannot be determined.*

Proof. Since $F_1 \subseteq F_2$, by Proposition 6, $R_{F_1} \subseteq R_{F_2}$, then $\forall u_i, u_j \in U, R_{F_1}(u_i, u_j) \le R_{F_2}(u_i, u_j)$. Then $\forall u_i \in U, [u_i]_{R_{F_1}} \subseteq [u_i]_{R_{F_2}}$. So $|[u_i]_{R_{F_1}}| \le |[u_i]_{R_{F_2}}|$, $|[u_i]_{R_{F_1}} \cup [u_i]_D| \le |[u_i]_{R_{F_2}} \cup [u_i]_D|$. Therefore

$$\frac{1}{|[u_i]_{R_{F_1}}|} \ge \frac{1}{|[u_i]_{R_{F_2}}|},$$
$$\frac{1}{|[u_i]_{R_{F_1}} \cup [u_i]_D|} \ge \frac{1}{|[u_i]_{R_{F_2}} \cup [u_i]_D|},$$

by Definition 16, the size relationship between $I(D; F_1)$ and $I(D; F_2)$ cannot be determined.

Algorithm 1 Attribute reduction algorithm using metric fuzzy rough mutual information (MFRMI).

Input: Decision information system $(U, C \cup D)$.
Output: Reduct set red.
Initialize $red \leftarrow \emptyset$;
Fuzzify $(U, C \cup D)$;
for $k = 1$ to $|C|$ **do**
  Compute $\mathrm{sig}_{\mathrm{in}}^{1}(a_k, C, D)$;
  if $\mathrm{sig}_{\mathrm{in}}^{1}(a_k, C, D) > 0$ **then**
    $red \leftarrow red \cup \{a_k\}$;
  end if
end for
while $I(D; red) \neq I(D; C)$ **do**
  for $l = 1$ to $|C - red|$ **do**
    Calculate $\mathrm{sig}_{\mathrm{out}}^{1}(a_l, red, D)$;
  end for
  Select $a' = \max_{a_l \in (C-red)} \{\mathrm{sig}_{\mathrm{out}}^{1}(a_l, red, D)\}$;
  $red \leftarrow red \cup \{a'\}$;
end while
for all $a \in red$ **do**
  if $I(D; red - \{a\}) = I(D; red)$ **then**
    $red \leftarrow red - \{a\}$;
  end if
end for
return red;

Definition 21. Let $F_1 \subseteq C$, we say that F_1 is a reduction of C, if F_1 satisfies

1. $I(D; F_1) = I(D; C)$;
2. $I(D; F_1 - \{a\}) \neq I(D; F_1)$.

Definition 22. Suppose $F_1 \subseteq C$, then

1. inner significance of $\forall a \in F_1$ in F_1 can be defined by $\mathrm{sig}_{\mathrm{in}}^{1}(a, F_1, D) = I(D; F_1) - I(D; F_1 - \{a\})$;
2. outer significance of $\forall a \in (C - F_1)$ to F_1 is defined as $\mathrm{sig}_{\mathrm{out}}^{1}(a, F_1, D) = I(D; F_1 \cup \{a\}) - I(D; F_1)$.

According to Definition 22, for any attribute a within the subset F_1, if a is deemed internally significant relative to F_1, the fuzzy rough mutual information criterion $I(D; F_1) > I(D; F_1 - \{a\})$ is satisfied, thus $\mathrm{sig}_{\mathrm{in}}^{1}(a, F_1, D) > 0$. Conversely, for any attribute a not in F_1 but in the complete set C, if a is externally significant in relation to F_1, $I(D; F_1 \cup \{a\}) > I(D; F_1)$, indicating $\mathrm{sig}_{\mathrm{out}}^{1}(a, F_1, D) > 0$. Using these concepts of inner and outer significance as heuristic functions, a heuristic algorithm for attribute selection is subsequently proposed.

Information entropy, an important index to measure the uncertainty of information, can also be employed for attribute reduction. The goal of this process is to find a minimum subset of attributes whose information entropy is the same as the original attribute set, thereby ensuring that the selected attributes reduce redundancy while maintaining basic information.

Definition 23. Suppose $F_1 \subseteq C$, then

1. inner significance of $\forall a \in F_1$ in F_1 can be defined by $\mathrm{sig}_{\mathrm{in}}^{2}(a, F_1) = E(F_1) - E(F_1 - \{a\})$;
2. outer significance of $\forall a \in (C - F_1)$ to F_1 is defined as $\mathrm{sig}_{\mathrm{out}}^{2}(a, F_1) = E(F_1 \cup \{a\}) - E(F_1)$.

The frameworks for both attribute reduction algorithms are shown in Fig. 2.

In Algorithm 2, focusing on the conditional attribute set C and ignoring the decision attribute D can effectively identify redundant attributes in the conditional attribute set in a shorter time. Since it is not restricted to a specific prediction task, this unsupervised attribute reduction method improves the versatility of the model, making it possible to discover intrinsic patterns in the dataset that are not biased towards any specific result.

The time complexity of the two algorithms can be analyzed as follows. For MFRMI, the inner significance $\mathrm{sig}_{\mathrm{in}}^{1}(a_k, C, D)$ is calculated with time complexity $O(f)$ and looping over the attribute set results in a complexity of $O(|C| \times f)$. In the subsequent *while* loop, the outer significance $\mathrm{sig}_{\mathrm{out}}^{1}(a_l, red, D)$ is computed with time complexity $O(g)$. Since the loop runs m times in the worst case, the total complexity of this part is $O(m \times |C| \times g)$. Lastly, the *for* loop that checks red has time complexity of $O(|C|)$. Therefore, the overall time complexity of MFRMI is given by:

$$T_{\mathrm{MFRMI}} = O(|C| \times f + m \times |C| \times g + |C|).$$

For MFREN, the inner significance $\mathrm{sig}_{\mathrm{in}}^{2}(a_k, C)$ and outer significance $\mathrm{sig}_{\mathrm{out}}^{2}(a_l, red)$ are computed with time complexities $O(f')$ and $O(g')$ respectively. Thus, the total complexity is:

Algorithm 2 Attribute reduction algorithm using metric fuzzy rough entropy (MFREN).

Input: information system (U, C).
Output: Reduct set red.
Initialize $red \leftarrow \emptyset$;
Fuzzify (U, C);
for $k = 1$ to $|C|$ **do**
  Compute $\mathrm{sig}^2_{\mathrm{in}}(a_k, C)$;
  if $\mathrm{sig}^2_{\mathrm{in}}(a_k, C) > 0$ **then**
    $red \leftarrow red \cup \{a_k\}$;
  end if
end for
while $E(red) \neq E(C)$ **do**
  for $l = 1$ to $|C - red|$ **do**
    Calculate $\mathrm{sig}^2_{\mathrm{out}}(a_l, red)$;
  end for
  Select $a' = \max_{a_l \in (C-red)} \{\mathrm{sig}^2_{\mathrm{out}}(a_l, red)\}$;
  $red \leftarrow red \cup \{a'\}$;
end while
for all $a \in red$ **do**
  if $E(red - \{a\}) = E(red)$ **then**
    $red \leftarrow red - \{a\}$;
  end if
end for
return red;

Table 3
Data of example.

Object	a_1	a_2	a_3	a_4	a_5	a_6
u_1	0.8	0.1	0.1	0.5	0.2	0.3
u_2	0.3	0.5	0.2	0.8	0.1	0.1
u_3	0.2	0.2	0.6	0.7	0.3	0.2
u_4	0.6	0.3	0.1	0.2	0.5	0.3
u_5	0.3	0.4	0.3	0.3	0.6	0.1
u_6	0.2	0.3	0.5	0.3	0.5	0.2
u_7	0.3	0.3	0.4	0.2	0.6	0.2
u_8	0.3	0.4	0.3	0.1	0.4	0.5
u_9	0.3	0.2	0.5	0.4	0.4	0.2

$$T_{\mathrm{MFREN}} = O(|C| \times f' + m \times |C| \times g' + |C|).$$

Since $f' < f$ and $g' < g$, MFREN has lower time complexity. In terms of space complexity, both algorithms store the red and use a constant amount of space for intermediate calculations, resulting in the same space complexity for both algorithms:

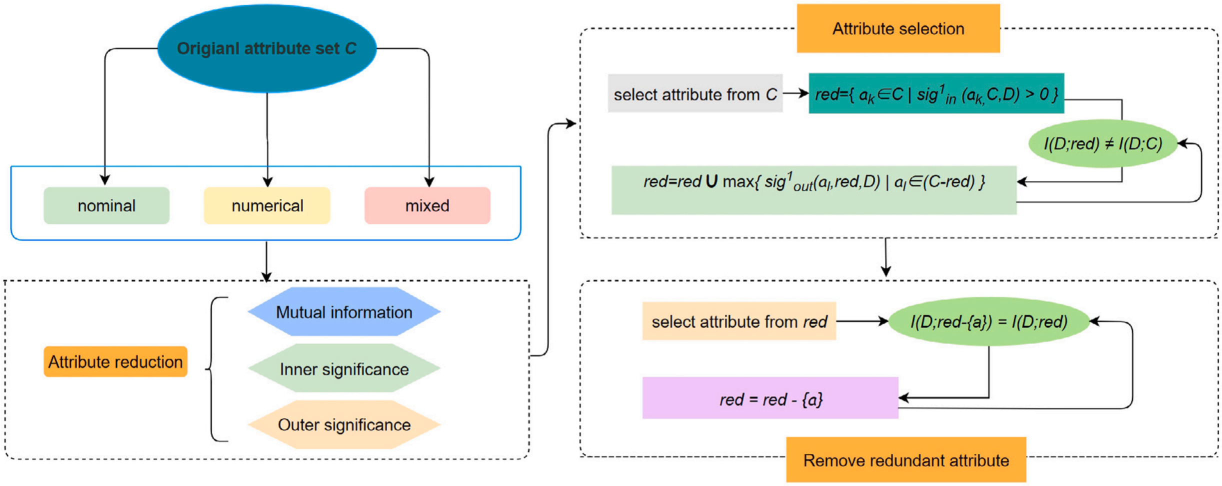
$$S_{\mathrm{MFRMI}} = S_{\mathrm{MFREN}} = O(|C|).$$

The following example demonstrates the calculation process of these two algorithms.

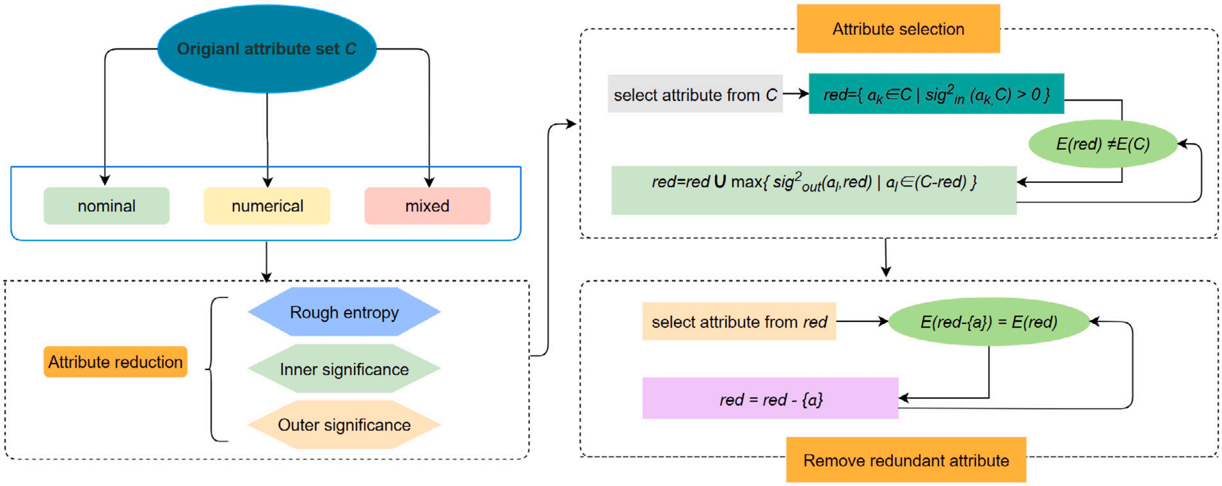
Example 24. Suppose $U = \{u_1, \dots, u_9\}$, $C = \{a_1, \dots, a_6\}$ is a set of fuzzy attributes. D is a partition of U derived with a decision attribute. Let $D = \{D_1, D_2\}$, where $D_1 = \{u_1, u_2, u_4, u_7\}$ and $D_2 = \{u_3, u_5, u_6, u_8, u_9\}$. Then, we have a fuzzy decision system $(U, C \cup D)$. The data characterizing the fuzzy attributes for each element in U are detailed in Table 3.

Elements in $\mathbf{M}_C$ are defined as follows:

$$M_{R_{a_1}} = \begin{bmatrix} 0.0 & 0.5 & 0.6 & 0.2 & 0.5 & 0.6 & 0.5 & 0.5 & 0.5 \\ 0.5 & 0.0 & 0.1 & 0.3 & 0.0 & 0.1 & 0.0 & 0.0 & 0.0 \\ 0.6 & 0.1 & 0.0 & 0.4 & 0.1 & 0.0 & 0.1 & 0.1 & 0.1 \\ 0.2 & 0.3 & 0.4 & 0.0 & 0.3 & 0.4 & 0.3 & 0.3 & 0.3 \\ 0.5 & 0.0 & 0.1 & 0.3 & 0.0 & 0.1 & 0.0 & 0.0 & 0.0 \\ 0.6 & 0.1 & 0.0 & 0.4 & 0.1 & 0.0 & 0.1 & 0.1 & 0.1 \\ 0.5 & 0.0 & 0.1 & 0.3 & 0.0 & 0.1 & 0.0 & 0.0 & 0.0 \\ 0.5 & 0.0 & 0.1 & 0.3 & 0.0 & 0.1 & 0.0 & 0.0 & 0.0 \\ 0.5 & 0.0 & 0.1 & 0.3 & 0.0 & 0.1 & 0.0 & 0.0 & 0.0 \end{bmatrix}$$



(a) Flowchart 1: Attribute reduction algorithm based on MFRMI



(b) Flowchart 2: Attribute reduction algorithm based on MFREN

Fig. 2. Flowcharts of two different attribute reduction algorithms.

$$M_{R_{a_2}} = \begin{bmatrix} 0.0 & 0.4 & 0.1 & 0.2 & 0.3 & 0.2 & 0.2 & 0.3 & 0.1 \\ 0.4 & 0.0 & 0.3 & 0.2 & 0.1 & 0.2 & 0.2 & 0.1 & 0.3 \\ 0.1 & 0.3 & 0.0 & 0.1 & 0.2 & 0.1 & 0.1 & 0.2 & 0.0 \\ 0.2 & 0.2 & 0.1 & 0.0 & 0.1 & 0.0 & 0.0 & 0.1 & 0.1 \\ 0.3 & 0.1 & 0.2 & 0.1 & 0.0 & 0.1 & 0.1 & 0.0 & 0.2 \\ 0.2 & 0.2 & 0.1 & 0.0 & 0.1 & 0.0 & 0.0 & 0.1 & 0.1 \\ 0.2 & 0.2 & 0.1 & 0.0 & 0.1 & 0.0 & 0.0 & 0.1 & 0.1 \\ 0.3 & 0.1 & 0.2 & 0.1 & 0.0 & 0.1 & 0.1 & 0.0 & 0.2 \\ 0.1 & 0.3 & 0.0 & 0.1 & 0.2 & 0.1 & 0.1 & 0.2 & 0.0 \end{bmatrix}$$

$$M_{R_{a_3}} = \begin{bmatrix} 0.0 & 0.1 & 0.5 & 0.0 & 0.2 & 0.4 & 0.3 & 0.2 & 0.4 \\ 0.1 & 0.0 & 0.4 & 0.1 & 0.1 & 0.3 & 0.2 & 0.1 & 0.3 \\ 0.5 & 0.4 & 0.0 & 0.5 & 0.3 & 0.1 & 0.2 & 0.3 & 0.1 \\ 0.0 & 0.1 & 0.5 & 0.0 & 0.2 & 0.4 & 0.3 & 0.2 & 0.4 \\ 0.2 & 0.1 & 0.3 & 0.2 & 0.0 & 0.2 & 0.1 & 0.0 & 0.2 \\ 0.4 & 0.3 & 0.1 & 0.4 & 0.2 & 0.0 & 0.1 & 0.2 & 0.0 \\ 0.3 & 0.2 & 0.2 & 0.3 & 0.1 & 0.1 & 0.0 & 0.1 & 0.1 \\ 0.2 & 0.1 & 0.3 & 0.2 & 0.0 & 0.2 & 0.1 & 0.0 & 0.2 \\ 0.4 & 0.3 & 0.1 & 0.4 & 0.2 & 0.0 & 0.1 & 0.2 & 0.0 \end{bmatrix}$$

$$M_{R_{a_4}} = \begin{bmatrix} 0.0 & 0.3 & 0.2 & 0.3 & 0.2 & 0.2 & 0.3 & 0.4 & 0.1 \\ 0.3 & 0.0 & 0.1 & 0.6 & 0.5 & 0.5 & 0.6 & 0.7 & 0.4 \\ 0.2 & 0.1 & 0.0 & 0.5 & 0.4 & 0.4 & 0.5 & 0.6 & 0.3 \\ 0.3 & 0.6 & 0.5 & 0.0 & 0.1 & 0.1 & 0.0 & 0.1 & 0.2 \\ 0.2 & 0.5 & 0.4 & 0.1 & 0.0 & 0.0 & 0.1 & 0.2 & 0.1 \\ 0.2 & 0.5 & 0.4 & 0.1 & 0.0 & 0.0 & 0.1 & 0.2 & 0.1 \\ 0.3 & 0.6 & 0.5 & 0.0 & 0.1 & 0.1 & 0.0 & 0.1 & 0.2 \\ 0.4 & 0.7 & 0.6 & 0.1 & 0.2 & 0.2 & 0.1 & 0.0 & 0.3 \\ 0.1 & 0.4 & 0.3 & 0.2 & 0.1 & 0.1 & 0.2 & 0.3 & 0.0 \end{bmatrix}$$

$$M_{R_{a_5}} = \begin{bmatrix} 0.0 & 0.1 & 0.1 & 0.3 & 0.4 & 0.3 & 0.4 & 0.2 & 0.2 \\ 0.1 & 0.0 & 0.2 & 0.4 & 0.5 & 0.4 & 0.5 & 0.3 & 0.3 \\ 0.1 & 0.2 & 0.0 & 0.2 & 0.3 & 0.2 & 0.3 & 0.1 & 0.1 \\ 0.3 & 0.4 & 0.2 & 0.0 & 0.1 & 0.0 & 0.1 & 0.1 & 0.1 \\ 0.4 & 0.5 & 0.3 & 0.1 & 0.0 & 0.1 & 0.0 & 0.2 & 0.2 \\ 0.3 & 0.4 & 0.2 & 0.0 & 0.1 & 0.0 & 0.1 & 0.1 & 0.1 \\ 0.4 & 0.5 & 0.3 & 0.1 & 0.0 & 0.1 & 0.0 & 0.2 & 0.2 \\ 0.2 & 0.3 & 0.1 & 0.1 & 0.2 & 0.1 & 0.2 & 0.0 & 0.0 \\ 0.2 & 0.3 & 0.1 & 0.1 & 0.2 & 0.1 & 0.2 & 0.0 & 0.0 \end{bmatrix}$$

$$M_{R_{a_6}} = \begin{bmatrix} 0.0 & 0.2 & 0.1 & 0.0 & 0.2 & 0.1 & 0.1 & 0.2 & 0.1 \\ 0.2 & 0.0 & 0.1 & 0.2 & 0.0 & 0.1 & 0.1 & 0.4 & 0.1 \\ 0.1 & 0.1 & 0.0 & 0.1 & 0.1 & 0.0 & 0.0 & 0.3 & 0.0 \\ 0.0 & 0.2 & 0.1 & 0.0 & 0.2 & 0.1 & 0.1 & 0.2 & 0.1 \\ 0.2 & 0.0 & 0.1 & 0.2 & 0.0 & 0.1 & 0.1 & 0.4 & 0.1 \\ 0.1 & 0.1 & 0.0 & 0.1 & 0.1 & 0.0 & 0.0 & 0.3 & 0.0 \\ 0.1 & 0.1 & 0.0 & 0.1 & 0.1 & 0.0 & 0.0 & 0.3 & 0.0 \\ 0.2 & 0.4 & 0.3 & 0.2 & 0.4 & 0.3 & 0.3 & 0.0 & 0.3 \\ 0.1 & 0.1 & 0.0 & 0.1 & 0.1 & 0.0 & 0.0 & 0.3 & 0.0 \end{bmatrix}$$

By MFRMI(RMI) and MFREN(REN), the inner significance of $\forall a \in C$ is calculated as follows.

$$\mathrm{sig}^1_{\mathrm{in}}(a_1, C, D) = 0.0630 > 0, \quad \mathrm{sig}^1_{\mathrm{in}}(a_2, C, D) = 0.0, \quad \mathrm{sig}^1_{\mathrm{in}}(a_3, C, D) = 0.0229 > 0,$$
$$\mathrm{sig}^1_{\mathrm{in}}(a_4, C, D) = 0.1573 > 0, \quad \mathrm{sig}^1_{\mathrm{in}}(a_5, C, D) = 0.0, \quad \mathrm{sig}^1_{\mathrm{in}}(a_6, C, D) = 0.04 > 0,$$
$$\mathrm{sig}^2_{\mathrm{in}}(a_1, C, D) = 0.1032 > 0, \quad \mathrm{sig}^2_{\mathrm{in}}(a_2, C, D) = 0.0, \quad \mathrm{sig}^2_{\mathrm{in}}(a_3, C, D) = 0.0329 > 0,$$
$$\mathrm{sig}^2_{\mathrm{in}}(a_4, C, D) = 0.1932 > 0, \quad \mathrm{sig}^2_{\mathrm{in}}(a_5, C, D) = 0.0, \quad \mathrm{sig}^2_{\mathrm{in}}(a_6, C, D) = 0.0454 > 0.$$

We have

$$\mathrm{red}^1_{\mathbf{RMI}} = \{a_1, a_3, a_4, a_6\}, \mathrm{red}^1_{\mathbf{REN}} = \{a_1, a_3, a_4, a_6\}.$$

Then we find that

$$I(D; \mathrm{red}^1_{\mathbf{RMI}}) = I(D; C) = 1.1451, E(\mathrm{red}^1_{\mathbf{REN}}) = E(C) = 1.6393.$$

Therefore

$$\mathrm{red}^2_{\mathbf{RMI}} = \mathrm{red}^1_{\mathbf{RMI}}, \mathrm{red}^2_{\mathbf{REN}} = \mathrm{red}^1_{\mathbf{REN}}.$$

By Definition 7 and 16, the fuzzy rough entropy and fuzzy rough mutual information are calculated as follows.

$$I(D; \mathrm{red}^2_{\mathbf{RMI}} - \{a_1\}) = 1.0987 \neq I(D; \mathrm{red}^2_{\mathbf{RMI}}), \quad E(\mathrm{red}^2_{\mathbf{REN}} - \{a_1\}) = 1.4903 \neq E(\mathrm{red}^2_{\mathbf{REN}}),$$
$$I(D; \mathrm{red}^2_{\mathbf{RMI}} - \{a_3\}) = 1.0468 \neq I(D; \mathrm{red}^2_{\mathbf{RMI}}), \quad E(\mathrm{red}^2_{\mathbf{REN}} - \{a_3\}) = 1.5756 \neq E(\mathrm{red}^2_{\mathbf{REN}}),$$
$$I(D; \mathrm{red}^2_{\mathbf{RMI}} - \{a_4\}) = 1.1012 \neq I(D; \mathrm{red}^2_{\mathbf{RMI}}), \quad E(\mathrm{red}^2_{\mathbf{REN}} - \{a_4\}) = 1.2794 \neq E(\mathrm{red}^2_{\mathbf{REN}}),$$
$$I(D; \mathrm{red}^2_{\mathbf{RMI}} - \{a_6\}) = 0.8593 \neq I(D; \mathrm{red}^2_{\mathbf{RMI}}), \quad E(\mathrm{red}^2_{\mathbf{REN}} - \{a_6\}) = 1.5819 \neq E(\mathrm{red}^2_{\mathbf{REN}}),$$

so $\mathrm{red}_{\mathbf{RMI}} = \mathrm{red}_{\mathbf{REN}} = \{a_1, a_3, a_4, a_6\}$. We use the attribute reduction based on the $\mathrm{T_M}$-Fuzzy Rough Sets algorithm in Reference [34] to find the core of C is $\{a_3, a_4, a_5, a_6\}$. Comparing the attribute reduction results of the two algorithms, it is noticeable that the number of attributes selected by each algorithm is the same. Among these, three attributes are consistent across the selections. This overlap indicates that they perform with consistent reliability when applied to a given dataset.

We have modified some data in Example 24 to compare and demonstrate the effectiveness of the two algorithms presented in this paper.

Example 25. Suppose $U = \{u_1, \dots, u_9\}$, $C = \{a_1, \dots, a_6\}$ is a set of fuzzy attributes. D is a partition of U derived with a decision attribute. Let $D = \{D_1, D_2\}$, where $D_1 = \{u_1, u_2, u_3, u_4, u_5, u_6, u_7, u_8\}$ and $D_2 = \{u_9\}$. Then, we have a fuzzy decision system $(U, C \cup D)$. The sample data is the same with that in Table 3, which constitutes an imbalanced class dataset.

It is worth mentioning that the traditional calculation of information entropy in fuzzy rough sets mainly considers the similarity between attributes, usually including the operator: $\log_2 |[u_i]_{R_{F_1}} \cap [u_i]_D|$ [35–38]. However, when considering the object u_9, the calculation will encounter a $\log_2 0$ error, which will hinder the algorithm from proceeding to the next step. The two algorithms proposed in this study avoid these calculation defects and successfully solve this problem.

By MFRMI(RMI), we can obtain the following:

$$M_C^{\mathbf{RMI}} = \bigvee_{i=1}^{|C|} R_{a_i}^{\mathbf{RMI}} = \begin{bmatrix} 0.0 & 0.5 & 0.6 & 0.3 & 0.5 & 0.6 & 0.5 & 0.5 & 0.5 \\ 0.5 & 0.0 & 0.4 & 0.6 & 0.5 & 0.5 & 0.6 & 0.7 & 0.4 \\ 0.6 & 0.4 & 0.0 & 0.5 & 0.4 & 0.4 & 0.5 & 0.6 & 0.3 \\ 0.3 & 0.6 & 0.5 & 0.0 & 0.3 & 0.4 & 0.3 & 0.3 & 0.4 \\ 0.5 & 0.5 & 0.4 & 0.3 & 0.0 & 0.2 & 0.1 & 0.4 & 0.2 \\ 0.6 & 0.5 & 0.4 & 0.4 & 0.2 & 0.0 & 0.1 & 0.3 & 0.1 \\ 0.5 & 0.6 & 0.5 & 0.3 & 0.1 & 0.1 & 0.0 & 0.3 & 0.2 \\ 0.5 & 0.7 & 0.6 & 0.3 & 0.4 & 0.3 & 0.3 & 0.0 & 0.3 \\ 0.5 & 0.4 & 0.3 & 0.4 & 0.2 & 0.1 & 0.2 & 0.3 & 0.0 \end{bmatrix}$$

$$M_D^{\mathbf{RMI}} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$([u]_C \cup [u]_D)^{\mathbf{RMI}} = \left[\begin{array}{c|ccccccccc} u_1 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 0.5 \\ u_2 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 0.4 \\ u_3 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 0.3 \\ u_4 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 0.4 \\ u_5 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 0.2 \\ u_6 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 0.1 \\ u_7 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 0.2 \\ u_8 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 0.3 \\ u_9 & 0.5 & 0.4 & 0.3 & 0.4 & 0.2 & 0.1 & 0.2 & 0.3 & 1.0 \end{array}\right]$$

Following the steps of Example 24, we ultimately obtain the reduction result as $\text{red}^{\mathbf{RMI}} = \{a_1, a_3, a_4, a_6\}$, with the fuzzy rough mutual information $I(D; \text{red})^{\mathbf{RMI}} = 1.3961$. In the MFREN($REN$) algorithm, the computation of information entropy is independent of the decision attribute, thus maintaining $E(C) = 1.6393$ and $\text{red}^{\mathbf{REN}} = \{a_1, a_3, a_4, a_6\}$.

5. Experiment analysis

5.1. Experimental environment

Both algorithms are implemented in PYTHON 3.9 and executed on an Intel (R) Core (TM) i5-12450 CPU @ 3.60GHz with 16.0GB RAM.

5.2. Algorithms and their configurations for comparison

This section evaluates the performance of the proposed algorithms by comparing them with existing attribute reduction methods, particularly classical methods such as the Fuzzy Neighborhood Rough Set (FNRS) [39], Fuzzy Rough Attribute Reduction (FRAR) [40], attribute reduction based on Fuzzy Rough Entropy (FSFrMI) [35], IARFCIE [19] and MFIGI [41]. The FNRS and FRAR algorithms are based on fuzzy positive regions, while FSFrMI, IARFCIE and MFIGI algorithms use fuzzy rough entropy. The MFREN and MFRMI algorithms proposed in this study specifically use metric fuzzy rough entropy. In addition, the parameters for each attribute reduction algorithm are consistent with those set in the original papers. The specific settings are detailed in Table 4. In this paper, high-dimensional datasets are defined as those with 1000 or more attributes and low-dimensional datasets contain attributes below 1000.

Various classic classification algorithms are introduced for evaluating attribute reduction algorithms, including Decision Tree (DT), RandomForest (RF), AdaBoost (AB), Bagging (BGG), Gradient Boosting (GB), CatBoost (CB), ExtraTrees (ET), Xgboost (XB), Lightgbm (LB), KNeighbors (KNN), SVC, LinearSVC (LSVC), GaussianNB (GN), and Logistic Regression (LR).

Table 4
Parameter Settings for Attribute Reduction Algorithms.

Algorithm	Parameters	Values/Settings
FNRS	λ	0.1 to 0.5 (step size: 0.05)
	α	0.5 to 1.0 (step size: 0.05)
IARFCIE	θ	0 to 1.0 (step size: 0.1)
MFIGI	ω	0.005 for high-dimensional datasets
		0.002 for low-dimensional datasets

Table 5
Dataset Summary (Low-Dimensional and High-Dimensional).

No.	Datasets	Abbreviation	Number of objects	Number of conditional attributes	Decision classes
Low-Dimensional Datasets					
1	Absenteeism_two	Abs	365	19	2
2	bridges	Bri	70	11	3
3	echocardiogram-0	Echo	106	8	2
4	sponge	Spon	76	22	3
5	transfusion	Tran	748	4	2
6	wdbc	Wdbc	569	30	2
7	wpbc	Wpbc	198	33	2
8	australian	Aust	2760	14	2
9	cleveland	Clev	1188	13	5
10	dermatology	Derm	1432	34	6
11	german	Germ	4000	20	2
12	lymphography	Lymp	592	18	4
13	magic	Magi	7608	10	2
14	tic-tac-toe	Tic	3832	9	2
15	vehicle	Vehi	3384	18	4
16	zoo	Z	404	16	7
High-Dimensional Datasets					
17	GSE	GSE	180	14532	5
18	DLBCL	DL	77	5469	2
19	Leukemia	LK	72	7129	2
20	Arcene	AR	200	10000	2

The study applies diverse classification methods with unique advantages and limitations, frequently used for various tasks and datasets. Through comprehensive analysis, the effectiveness of the proposed attribute reduction algorithms is validated.

5.3. Experimental datasets

An empirical investigation is performed using 20 numerical datasets retrieved from the UCI Machine Learning Repository. The particulars of these datasets are enumerated in Table 5. To apply the fuzzy rough set algorithms, all attributes are min-max normalized, adjusting their values within the range of [0,1].

5.4. Outcomes from data reduction

Table 6 shows the number of attributes in the original datasets as well as the number of attributes after reduction by each algorithm. From Table 6, it is evident that the performance of the attribute reduction algorithms varies significantly across different datasets.

In low-dimensional datasets, although the MFRMI and MFREN algorithms do not achieve the highest level of attribute reduction, they effectively remove redundant attributes while retaining features critical to the classification task, excelling in preserving key information. In contrast, while FNRS, FRAR, and MFIGI achieve more significant attribute reduction in some datasets, removing more features. Excessive reduction may lead to information loss, which in turn affects classification performance. The performance of MFRMI and MFREN in low-dimensional datasets strikes a balance between the number of reduced attributes and the retention of crucial information, ensuring that the core information required for classification is preserved, as will be reflected in the subsequent accuracy ranking analysis.

In high-dimensional datasets, the MFRMI and MFREN algorithms show more pronounced advantages. They effectively reduce redundant features while preserving key ones, enhancing the performance of the classification model. Specifically, when applied to complex high-dimensional datasets such as DL, LK, and AR, these two algorithms are able to reduce the data dimensions while maximizing the retention of features critical to the classification task. This results in more stable performance of MFRMI and MFREN in high-dimensional datasets compared to other algorithms, avoiding performance degradation caused by excessive reduction.

Table 6
Sizes of Attribute Subsets (Low-Dimensional and High-Dimensional).

Datasets	Original attributes	FNRS	FRAR	FSFRMI	IARFCIE	MFIGI	MFRMI	MFREN
Low-Dimensional Datasets								
Abs	19	1	2	18	15.5	3	8	8.3
Bri	11	1	3.8	4.5	6	3	3.5	3.5
Echo	8	2.5	2.85	8	7.5	2	4	4
Spon	22	2	9.12	9	4.4	5	9.3	9.3
Tran	4	1.75	1.5	3	4	1	3	3
Wdbc	30	2	4.5	3	30	1	6.5	6.5
Wpbc	33	4.36	4.2	2	33	2	6.5	6.5
Aust	14	1.3	1	14	10.8	4	6	6
Clev	13	5	6	3	9.6	5	13	13
Derm	34	7.5	4.5	34	33	4.5	8	7.6
Germ	20	4.4	1	7.3	12.2	3	11.6	11.6
Lymp	18	2.6	4	8.8	10.3	5	9	9
Magi	10	4.1	4.5	10	10	2	10	10
Tic	9	4.5	1	9	9	4	9	9
Vehi	18	5.6	4	4	15	3	14	14
Z	16	3	1	8	1	5	11	11
Average	17.4	3.3	3.4	9.1	13.2	3.3	8.3	8.3
High-Dimensional Datasets								
GSE	14532	1	/	13732.7	14532	3.5	570	560.3
DL	5469	1	1.5	3445.5	5469	2	78.3	79.6
LK	7129	1	2.5	5892.9	7129	3	65.8	64.5
AR	10000	10000	1.4	10000	4967.8	4.14	199.75	203.6
Average	9282.5	2500.8	1.8	8267.8	8024.5	3.16	228.4	227.0

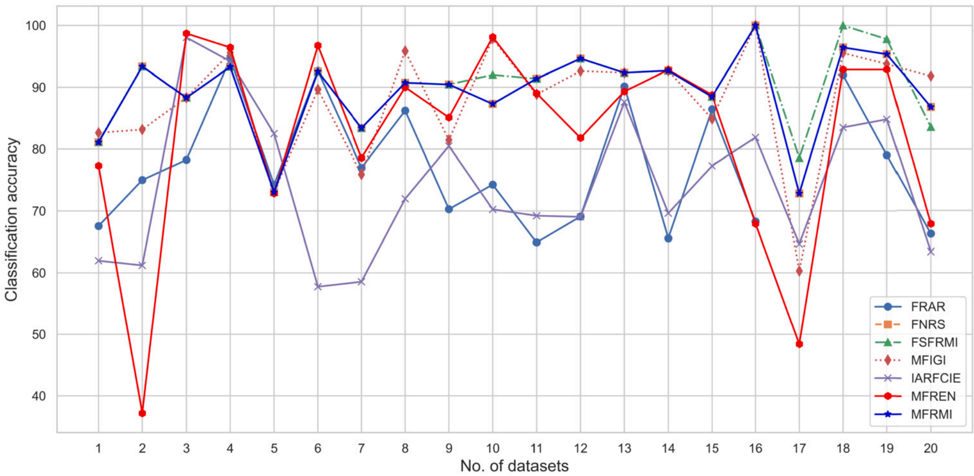


Fig. 3. Comparison of Classification Accuracies of Reduced Data with 14 Models.

5.5. *Classification accuracy comparison of attribute reduction algorithms*

Fig. 3 illustrates the classification accuracies obtained after applying attribute reduction using different algorithms across various datasets. The horizontal axis represents the sequence of 20 datasets, where the number (i) indicates the i-th dataset. The vertical axis shows the classification accuracy. Table 8–13 presents the classification accuracies of the original dataset and those of the datasets reduced by seven different attribute reduction methods, as evaluated across 14 classifier models. Table 7 shows the rank of the attribute reduction algorithms according to their classification accuracies on each dataset, providing a comparative analysis of their performance. Each cell in the table indicates the rank of a specific method for a given dataset, with a rank of 1 representing the highest classification accuracy and higher ranks indicating lower accuracies.

The results from Table 8–13 indicate that the two algorithms proposed in this study generally outperform the other algorithms in terms of average classification accuracy. Although the FRAR and MFIGI algorithms achieve the best attribute reduction rates, their average ranks, as shown in Table 7, are 5.89 and 6.25. Additionally, the FRAR algorithm's running time exceeded six days on the GSE dataset, which is indicated by a slash in tables.

Table 7
Rank of Average Accuracy with 14 Models.

Datasets	Raw data	FRAR	FNRS	FSFRMI	MFIGI	IARFCIE	MFREN	MFRMI
Abs	1	7	6	2	8	5	4	3
Bri	1	6	1	5	7	8	3	3
Echo	3	7	8	3	2	1	3	3
Spon	1	5	3	4	6	2	7	7
Tran	7	2	8	3	1	6	3	3
Wdbc	2	3	7	6	8	1	4	4
Wpbc	3	4	7	6	8	5	1	1
Aust	1	6	8	1	7	5	3	3
Clev	1	7	8	5	6	4	1	1
Derm	2	6	8	2	7	1	5	4
Germ	5	8	6	4	7	3	1	1
Lymp	1	8	7	4	6	5	2	2
Magi	1	5	8	1	7	6	1	1
Tic	2	8	6	2	7	1	2	2
Vehi	3	2	8	6	7	1	3	3
Z	1	6	8	1	5	7	1	1
GSE	5	/	7	4	3	6	2	1
LK	5	7	3	3	8	5	2	1
DL	5	8	3	3	7	5	2	1
AR	1	7	1	1	8	6	4	5
Average	**2.55**	**5.89**	**6.05**	**3.3**	**6.25**	**4.15**	**2.7**	**2.5**

The original dataset demonstrates the highest classification accuracy on 9 datasets, with an overall average ranking of 2.55. The FSFRMI algorithm also performs well, achieving the highest ranking in four datasets and an overall average ranking of 3.3. In contrast, the proposed MFRMI and MFREN algorithms secure the highest classification accuracy rankings in 8 and 5 datasets, respectively. Notably, on high-dimensional datasets such as GSE, LK, and DL, MFRMI and MFREN rank first and second, respectively, highlighting their superior performance in attribute reduction for high-dimensional data.

Specifically, the average ranks of the MFRMI and MFREN algorithms are 2.5 and 2.7, respectively. This demonstrates the advantage of metric methods in handling nonlinear relationships, allowing these algorithms to perform attribute reduction while preserving a high level of information. The use of distance measures further enhances their ability to adapt to the complex distribution of high-dimensional features, avoiding the information loss typically associated with traditional logical connectives. This underscores the strengths of the MFRMI algorithm, which is based on fuzzy rough mutual information, and the MFREN algorithm, which uses fuzzy information entropy, in achieving effective attribute reduction while retaining maximum original information.

5.6. Running time performance comparison of attribute reduction algorithms

Table 14 presents the running times of different algorithms during ten-fold cross-validation on low-dimensional and high-dimensional datasets. Fig. 4 shows the performance of MFREN and MFRMI on these two types of datasets, using dual y-axes to differentiate the running times for low-dimensional and high-dimensional datasets. The left y-axis displays the running times for low-dimensional datasets, ranging from 0 to 500 seconds, while the right y-axis shows the running times for high-dimensional datasets, ranging from 0 to 160,000 seconds. These two y-axes reflect the running times of the two algorithms on low-dimensional and high-dimensional datasets, respectively.

Table 14 and Fig. 4 show that, on most low-dimensional datasets, the running time of MFRMI and MFREN is significantly lower than that of other algorithms. For instance, on datasets such as Bri, Spon, Wdbc, and Wpbc, the running time of these two algorithms is notably shorter than that of the other methods. Although MFRMI and MFREN do not achieve the optimal number of attribute reductions, they effectively remove redundant features while retaining the features critical to the classification task. This indicates that these two algorithms can effectively avoid information loss during the attribute reduction process, preserving key data. In contrast, while FRAR and FNRS achieve stronger attribute reduction on some datasets, they typically require more computation time, and excessive reduction may lead to information loss, which can impact classification performance. The advantage of MFRMI and MFREN lies in their ability to effectively balance the reduction effect with the running time while retaining essential information.

On high-dimensional datasets, MFRMI and MFREN demonstrate significant advantages. Through precise attribute reduction, these algorithms effectively improve classification accuracy (ranking first and second on high-dimensional datasets like GSE, LK, and DL) while avoiding excessive computational burdens. This suggests that MFRMI and MFREN can achieve ideal attribute reduction results in less computation time, effectively preventing the performance degradation often observed with traditional algorithms on high-dimensional datasets.

5.7. Experiment analysis summary

The following two tables (Table 15 and Table 16) present the key features and properties of the algorithms, as well as their performance, including accuracy ranking and running time across low-dimensional and high-dimensional datasets.

Table 8
Comparison of classification accuracies of reduced data with 14 models.

Dataset	Algorithm	DT	RF	AB	BGG	GB	CB	ET	XB
Abs	MFRMI	86.49±5.41	83.78±5.41	86.49±0.00	90.54±2.70	87.84±2.70	87.84±2.70	81.08±5.41	86.49±0.0
	MFREN	86.49±5.41	83.78±5.41	85.14±2.70	90.54±2.70	87.84±2.70	87.84±2.70	81.08±5.41	86.49±0.00
	IARFCIE	83.78±5.41	82.43±1.35	85.13±4.05	81.08±2.70	79.73±1.35	81.08±0.00	83.78±2.70	82.43±1.35
	MFIGI	56.76±0.0	59.46±0.0	59.46±0.0	56.76±0.0	59.46±0.0	59.46±0.0	56.76±0.0	59.46±0.0
	FSFRMI	86.49±0.0	78.38±0.0	86.49±0.0	86.49±0.0	86.49±0.0	86.49±0.0	67.57±0.0	86.49±0.0
	FNRS	81.08±0.0	81.08±0.0	81.08±0.0	81.08±0.0	81.08±0.0	81.08±0.0	81.08±0.0	81.08±0.0
	FRAR	64.86±5.41	70.27±5.41	67.57±0.0	63.51±8.11	70.27±5.41	68.92±2.70	63.51±8.11	67.57±0.0
	Raw data	86.49±0.0	83.78±0.0	83.78±0.0	83.78±0.0	86.49±0.0	86.49±0.0	70.27±0.0	86.49±0.0
Bri	MFRMI	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0
	MFREN	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0
	IARFCIE	50.0±21.43	50.0±21.43	42.86±28.57	42.86±28.57	50.0±21.43	50.0±21.43	50.0±21.43	42.86±28.57
	MFIGI	57.14±0.0	71.43±0.0	57.14±0.0	71.43±0.0	57.14±0.0	71.43±0.0	57.14±0.0	71.43±0.0
	FSFRMI	85.71±2.86	85.71±2.86	85.71±2.86	85.71±2.86	85.71±2.86	85.71±2.86	85.71±2.86	85.71±2.86
	FNRS	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0
	FRAR	85.71±2.86	85.71±2.86	85.71±2.86	85.71±2.86	78.57±1.43	78.57±1.43	78.57±1.43	78.57±1.43
	Raw data	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0
Echo	MFRMI	72.73±0.0	90.91±0.0	72.73±0.0	90.91±0.0	90.91±0.0	90.91±0.0	90.91±0.0	90.91±0.0
	MFREN	72.73±0.0	90.91±0.0	72.73±0.0	90.91±0.0	90.91±0.0	90.91±0.0	90.91±0.0	90.91±0.0
	IARFCIE	90.91±0.0	100.0±0.0	90.91±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0
	MFIGI	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0
	FSFRMI	72.73±0.0	90.91±0.0	72.73±0.0	90.91±0.0	90.91±0.0	90.91±0.0	90.91±0.0	90.91±0.0
	FNRS	68.18±4.55	72.73±3.64	68.18±4.55	72.73±3.64	72.73±3.64	77.27±2.73	77.27±2.73	77.27±2.73
	FRAR	72.73±3.64	77.27±2.73	68.18±4.55	77.27±2.73	72.73±3.64	81.82±1.82	81.82±1.82	77.27±2.73
	Raw data	72.73±0.0	90.91±0.0	72.73±0.0	90.91±0.0	90.91±0.0	90.91±0.0	90.91±0.0	90.91±0.0
Spon	MFRMI	87.5±0.0	100.0±0.0	93.75±1.25	100.0±0.0	87.5±0.0	100.0±0.0	87.5±0.0	100.0±0.0
	MFREN	87.5±0.0	100.0±0.0	93.75±1.25	100.0±0.0	87.5±0.0	100.0±0.0	87.5±0.0	100.0±0.0
	IARFCIE	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0
	MFIGI	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0
	FSFRMI	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0
	FNRS	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0
	FRAR	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0
	Raw data	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0
Tran	MFRMI	65.33±0.0	73.33±0.0	70.67±0.0	72.00±0.0	77.33±0.0	74.67±0.0	70.67±0.0	76.00±0.0
	MFREN	65.33±0.0	73.33±0.0	70.67±0.0	72.00±0.0	77.33±0.0	74.67±0.0	70.67±0.0	76.00±0.0
	IARFCIE	62.67±0.0	68.0±0.0	64.0±0.0	70.67±0.0	76.0±0.0	69.33±0.0	66.67±0.0	68.0±0.0
	MFIGI	81.33±0.0	81.33±0.0	81.33±0.0	82.67±0.0	81.33±0.0	81.33±0.0	81.33±0.0	82.67±0.0
	FSFRMI	65.33±0.0	73.33±0.0	70.67±0.0	72.00±0.0	77.33±0.0	74.67±0.0	70.67±0.0	76.00±0.0
	FNRS	68.67±12.00	73.33±2.67	70.0±9.33	70.67±8.00	75.33±4.00	72.00±5.33	70.67±8.00	71.33±6.67
	FRAR	72.00±5.33	72.67±4.00	74.00±1.33	74.00±1.33	75.33±1.33	74.67±0.0	73.33±2.67	76.00±2.67
	Raw data	64.00±0.0	73.33±0.0	73.33±0.0	70.67±0.0	77.33±0.0	73.33±0.0	69.33±0.0	76.00±0.0
Wdbc	MFRMI	93.86±1.75	96.49±0.0	93.86±1.75	93.86±1.75	95.61±1.75	96.49±0.0	94.74±0.0	96.49±0.0
	MFREN	93.86±1.75	96.49±0.0	93.86±1.75	93.86±1.75	95.61±1.75	96.49±0.0	94.74±0.0	96.49±0.0
	IARFCIE	96.49±0.0	100.0±0.0	89.47±0.0	100.0±0.0	98.25±0.0	100.0±0.0	100.0±0.0	98.25±0.0
	MFIGI	52.63±0.0	54.39±0.0	52.63±0.0	61.40±0.0	61.40±0.0	54.39±0.0	49.12±0.0	57.89±0.0
	FSFRMI	88.60±1.75	90.35±5.26	89.47±3.51	90.35±8.77	90.35±1.75	89.47±3.51	91.23±3.51	89.47±3.51
	FNRS	78.95±1.40	81.58±1.93	78.95±1.40	80.70±1.40	83.33±1.23	85.09±0.53	78.95±1.75	79.82±1.93
	FRAR	90.35±0.88	92.98±0.70	90.35±0.53	90.35±0.88	93.86±0.53	92.11±0.53	92.98±0.70	92.11±0.88
	Raw data	91.23±0.0	96.49±0.0	94.74±0.0	94.74±0.0	94.74±0.0	96.49±0.0	96.49±0.0	96.49±0.0
Wpbc	MFRMI	70.0±10.0	87.5±5.0	72.5±15.0	87.5±5.0	77.5±5.0	90.0±0.0	90.0±0.0	82.5±5.0
	MFREN	70.0±10.0	87.5±5.0	72.5±15.0	87.5±5.0	77.5±5.0	90.0±0.0	90.0±0.0	82.5±5.0
	IARFCIE	65.0±0.0	90.0±0.0	70.0±0.0	85.0±0.0	80.0±0.0	90.0±0.0	90.0±0.0	85.0±0.0
	MFIGI	55.0±0.0	60.0±0.0	55.0±0.0	60.0±0.0	55.0±0.0	60.0±0.0	60.0±0.0	55.0±0.0
	FSFRMI	62.5±5.0	75.0±10.0	62.5±5.0	80.0±10.0	70.0±10.0	72.5±15.0	85.0±10.0	75.0±20.0
	FNRS	65.0±5.0	70.0±5.0	65.0±5.0	70.0±5.0	67.5±4.5	70.0±4.0	70.0±5.0	65.0±5.0
	FRAR	67.5±3.5	77.5±1.5	70.0±3.0	75.0±2.0	75.0±3.0	85.0±1.0	80.0±2.0	75.0±2.0
	Raw data	65.0±0.0	90.0±0.0	70.0±0.0	85.0±0.0	80.0±0.0	90.0±0.0	90.0±0.0	85.0±0.0

Analyzing data from Tables 6-16 and Fig. 3-4, we can draw the following conclusions about the proposed MFREN and MFRMI Algorithms:

1. Among the 20 datasets used for attribute reduction, the MFREN algorithm and the MFRMI algorithm achieve the best classification accuracy rankings on 8 datasets and 5 datasets, respectively. The final average accuracy ranks of the two algorithms are the

Table 9
Comparison of classification accuracies of reduced data with 14 models.

Dataset	Algorithm	DT	RF	AB	BGG	GB	CB	ET	XB
Aust	MFRMI	100.0±0.0	100.0±0.0	100.0±0.0	98.91±0.0	92.03±0.0	100.0±0.0	100.0±0.0	100.0±0.0
	MFREN	100.0±0.0	100.0±0.0	100.0±0.0	98.91±0.0	92.03±0.0	100.0±0.0	100.0±0.0	100.0±0.0
	IARFCIE	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	91.3±3.62	100.0±0.0	100.0±0.0	100.0±0.0
	MFIGI	72.10±0.0	72.10±0.0	72.10±0.0	72.10±0.0	72.46±0.0	72.10±0.0	72.10±0.0	72.46±0.0
	FSFRMI	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	95.65±0.0	100.0±0.0	100.0±0.0	100.0±0.0
	FNRS	70.47±23.55	70.47±23.55	70.47±23.55	70.65±23.91	66.49±15.58	70.83±24.28	70.47±23.55	70.65±23.91
	FRAR	86.23±0.0	86.23±0.0	86.23±0.0	86.23±0.0	86.23±0.0	86.23±0.0	86.23±0.0	86.23±0.0
	Raw data	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	95.65±0.0	100.0±0.0	100.0±0.0	100.0±0.0
Clev	MFRMI	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0
	MFREN	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0
	IARFCIE	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	98.32±1.68	100.0±0.0	100.0±0.0	100.0±0.0
	MFIGI	90.76±0.0	90.76±0.0	90.76±0.0	90.76±0.0	82.35±0.0	90.76±0.0	90.76±0.0	90.76±0.0
	FSFRMI	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	85.71±0.0	100.0±0.0	100.0±0.0	100.0±0.0
	FNRS	74.79±5.42	74.79±5.42	74.79±5.42	74.79±5.42	73.95±4.87	74.79±5.42	74.79±5.42	74.79±5.42
	FRAR	74.79±0.0	74.79±0.0	74.79±0.0	74.79±0.0	72.27±0.0	74.79±0.0	74.79±0.0	74.79±0.0
	Raw data	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0
Derm	MFRMI	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	96.88±0.63	100.0±0.0	100.0±0.0	100.0±0.0
	MFREN	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	98.61±0.28	100.0±0.0	100.0±0.0	100.0±0.0
	IARFCIE	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0
	MFIGI	71.53±5.56	71.53±5.56	71.53±5.56	71.53±5.56	71.53±5.56	71.53±5.56	71.53±5.56	71.53±5.56
	FSFRMI	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0
	FNRS	65.97±6.81	65.97±6.81	65.97±6.81	65.97±6.81	65.97±6.81	65.97±6.81	65.97±6.81	65.97±6.81
	FRAR	75.35±2.08	75.35±2.08	75.35±2.08	75.35±2.08	75.35±2.08	75.35±2.08	75.35±2.08	75.35±2.08
	Raw data	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0
Germ	MFRMI	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	85.25±2.5	100.0±0.0	100.0±0.0	100.0±0.0
	MFREN	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	85.25±2.5	100.0±0.0	100.0±0.0	100.0±0.0
	IARFCIE	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	83.0±9.5	99.63±0.37	100.0±0.0	100.0±0.0
	MFIGI	69.50±0.0	69.50±0.0	69.50±0.0	69.50±0.0	69.50±0.0	69.50±0.0	69.50±0.0	69.50±0.0
	FSFRMI	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	79.5±3.0	99.63±0.37	100.0±0.0	99.5±0.5
	FNRS	82.75±3.45	82.75±3.45	82.75±3.45	82.75±3.45	75.75±2.15	82.75±3.45	82.75±3.45	82.75±3.45
	FRAR	65.5±0.0	65.5±0.0	65.5±0.0	65.5±0.0	65.5±0.0	65.5±0.0	65.5±0.0	65.5±0.0
	Raw data	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	87.5±0.0	100.0±0.0	100.0±0.0	100.0±0.0
Lymp	MFRMI	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0
	MFREN	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0
	IARFCIE	85.0±15.0	85.0±15.0	85.0±15.0	86.67±13.13	85.0±15.0	85.0±15.0	85.0±15.0	85.0±15.0
	MFIGI	71.67±0.0	71.67±0.0	71.67±0.0	71.67±0.0	71.67±0.0	71.67±0.0	71.67±0.0	71.67±0.0
	FSFRMI	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	96.67±3.33	100.0±0.0	100.0±0.0	100.0±0.0
	FNRS	74.17±0.42	74.17±0.42	74.17±0.42	74.17±0.42	73.33±0.4	74.17±0.42	74.17±0.42	74.17±0.42
	FRAR	72.5±2.5	72.5±2.5	72.5±2.5	72.5±2.5	72.5±2.5	72.5±2.5	72.5±2.5	72.5±2.5
	Raw data	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0
Magi	MFRMI	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	88.96±0.0	99.61±0.0	100.0±0.0	100.0±0.0
	MFREN	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	88.96±0.0	99.61±0.0	100.0±0.0	100.0±0.0
	IARFCIE	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	88.96±0.0	99.61±0.0	100.0±0.0	100.0±0.0
	MFIGI	100.0±0.0	100.0±0.0	100.0±0.0	99.74±0.0	82.13±0.0	100.0±0.0	100.0±0.0	94.74±0.0
	FSFRMI	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	88.96±0.0	99.61±0.0	100.0±0.0	100.0±0.0
	FNRS	83.84±3.23	83.38±3.23	83.84±3.23	84.10±3.23	74.05±1.98	81.87±3.55	83.84±3.23	83.84±3.23
	FRAR	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	83.51±5.91	97.57±1.5	100.0±0.0	98.95±0.79
	Raw data	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	88.96±0.0	99.61±0.0	100.0±0.0	100.0±0.0
Tic	MFRMI	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	91.93±0.0	100.0±0.0	100.0±0.0	100.0±0.0
	MFREN	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	91.93±0.0	100.0±0.0	100.0±0.0	100.0±0.0
	IARFCIE	94.01±0.0	100.0±0.0	100.0±0.0	100.0±0.0	93.49±0.0	100.0±0.0	100.0±0.0	100.0±0.0
	MFIGI	71.22±0.26	71.22±0.26	70.70±0.78	71.22±0.78	71.48±0.26	71.22±0.78	71.22±0.26	70.70±0.78
	FSFRMI	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	91.93±0.0	100.0±0.0	100.0±0.0	100.0±0.0
	FNRS	82.03±3.59	82.42±3.52	82.42±3.52	82.42±3.52	77.99±3.52	82.42±3.52	82.03±3.59	82.42±3.52
	FRAR	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	83.51±5.91	97.57±1.64	100.0±0.0	98.95±0.79
	Raw data	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	91.93±0.0	100.0±0.0	100.0±0.0	100.0±0.0

top two highest, excluding the original dataset, indicating that they have the best performance in attribute classification after reduction.

2. In terms of running time, both the MFREN algorithm and the MFRMI algorithm can significantly shorten the running time while ensuring classification accuracy. Considering both running time and classification accuracy, the two algorithms perform better than the others.

Table 10
Comparison of classification accuracies of reduced data with 14 models.

Dataset	Algorithm	DT	RF	AB	BGG	GB	CB	ET	XB
Vehi	MFRMI	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	95.58±0.0	100.0±0.0	100.0±0.0	100.0±0.0
	MFREN	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	95.58±0.0	100.0±0.0	100.0±0.0	100.0±0.0
	IARFCIE	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	95.58±0.0	100.0±0.0	100.0±0.0	100.0±0.0
	MFIGI	95.28±0.0	95.28±0.0	95.28±0.0	94.69±0.0	72.27±0.0	95.28±0.0	95.28±0.0	94.69±0.0
	FSFRMI	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	87.32±0.0	100.0±0.0	100.0±0.0	100.0±0.0
	FNRS	68.88±6.22	69.03±6.22	68.88±6.22	68.44±6.31	66.37±5.84	69.03±6.19	68.88±6.22	68.88±6.22
	FRAR	98.38±0.21	98.38±0.21	98.38±0.21	98.67±0.15	84.22±0.21	97.57±0.18	98.38±0.21	97.64±0.30
	Raw data	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	95.58±0.0	100.0±0.0	100.0±0.0	100.0±0.0
Z	MFRMI	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0
	MFREN	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0
	IARFCIE	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0
	MFIGI	81.71±2.44	81.71±2.44	81.71±2.44	81.71±2.44	81.71±2.44	81.71±2.44	81.71±2.44	81.71±2.44
	FSFRMI	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0
	FNRS	65.85±5.37	65.85±5.37	65.85±5.37	65.85±5.37	65.85±5.37	65.85±5.37	65.85±5.37	65.85±5.37
	FRAR	68.29±0.0	68.29±0.0	68.29±0.0	68.29±0.0	68.29±0.0	68.29±0.0	68.29±0.0	68.29±0.0
	Raw data	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0
GSE	MFRMI	69.44±27.78	83.33±11.11	72.22±33.33	77.78±11.11	88.89±11.11	91.67±5.56	88.89±11.11	91.67±5.56
	MFREN	66.67±5.56	75.00±2.78	61.11±5.56	72.22±0.00	75.00±2.78	77.78±0.00	80.56±2.78	75.00±2.78
	IARFCIE	55.56±0.00	66.67±0.00	61.11±0.00	55.56±0.00	66.67±0.00	72.22±0.00	72.22±0.00	72.22±0.00
	MFIGI	41.67±38.89	52.78±27.78	38.89±33.33	44.44±22.22	47.22±27.78	47.22±27.78	50.00±33.33	52.78±27.78
	FSFRMI	55.56±11.11	61.11±11.11	63.89±5.56	58.33±16.67	63.89±5.56	66.67±0.00	58.33±5.56	58.33±5.56
	FNRS	38.89±0.00	38.89±0.00	38.89±0.00	44.44±0.00	38.89±0.00	44.44±0.00	38.89±0.00	38.89±0.00
	FRAR	/	/	/	/	/	/	/	/
	Raw data	55.56±0.00	55.56±0.00	55.56±0.00	50.00±0.00	61.11±0.00	72.22±0.00	55.56±0.00	66.67±0.00
LK	MFRMI	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0
	MFREN	87.5±12.5	93.75±6.25	87.5±12.5	93.75±6.25	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0
	IARFCIE	75.0±0.0	100.0±0.0	62.5±0.0	100.0±0.0	62.5±0.0	100.0±0.0	100.0±0.0	100.0±0.0
	MFIGI	87.5±12.5	87.5±12.5	87.5±12.5	87.5±12.5	87.5±12.5	81.25±18.75	87.5±12.5	87.5±12.5
	FSFRMI	81.25±12.5	100.0±0.0	81.25±12.5	100.0±0.0	75.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0
	FNRS	81.25±12.5	100.0±0.0	81.25±12.5	100.0±0.0	75.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0
	FRAR	81.25±18.75	93.75±6.25	87.5±12.5	93.75±6.25	87.5±12.5	93.75±6.25	93.75±6.25	87.5±12.5
	Raw data	75.0±0.0	100.0±0.0	62.5±0.0	100.0±0.0	62.5±0.0	100.0±0.0	100.0±0.0	100.0±0.0
DL	MFRMI	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	93.75±6.25	100.0±0.0
	MFREN	68.55±18.75	81.25±6.25	62.3±25.0	75.0±12.5	68.72±6.25	93.75±6.25	87.5±12.5	93.75±6.25
	IARFCIE	87.5±0.0	100.0±0.0	75.0±0.0	75.0±0.0	87.5±0.0	100.0±0.0	100.0±0.0	100.0±0.0
	MFIGI	87.5±0.0	93.75±6.25	81.25±18.75	81.25±18.75	100.0±0.0	75.0±0.0	93.75±6.25	81.25±6.25
	FSFRMI	75.0±0.0	100.0±0.0	68.75±12.5	87.5±12.5	87.5±12.5	100.0±0.0	100.0±0.0	100.0±0.0
	FNRS	75.0±0.0	100.0±0.0	68.75±12.5	87.5±12.5	87.5±12.5	100.0±0.0	100.0±0.0	100.0±0.0
	FRAR	75.0±12.5	81.25±6.25	75.0±12.5	75.0±12.5	75.0±12.5	81.25±6.25	75.0±12.5	93.75±6.25
	Raw data	87.5±0.0	100.0±0.0	75.0±0.0	75.0±0.0	87.5±0.0	100.0±0.0	100.0±0.0	100.0±0.0
AR	MFRMI	72.50±15.00	85.00±10.00	75.00±10.00	87.50±7.5	77.50±5.00	85.00±10.00	92.50±5.00	82.50±5.00
	MFREN	77.50±12.50	92.50±7.50	75.00±5.00	90.00±5.00	87.50±2.50	87.50±2.50	95.00±5.00	87.50±2.50
	IARFCIE	60.00±5.00	77.50±17.50	57.50±2.50	62.50±2.50	70.00±10.00	75.00±15.00	77.50±17.50	67.50±7.50
	MFIGI	60.00±10.00	65.00±20.00	57.50±15.00	65.00±20.00	67.50±25.00	65.00±20.00	65.00±20.00	65.00±10.00
	FSFRMI	80.00±0.00	95.00±0.00	80.00±0.00	95.00±0.00	95.00±0.00	90.00±0.00	95.00±0.00	95.00±0.00
	FNRS	80.00±0.00	95.00±0.00	80.00±0.00	95.00±0.00	95.00±0.00	90.00±0.00	95.00±0.00	95.00±0.00
	FRAR	72.50±7.50	70.00±10.00	72.50±7.50	70.00±10.00	70.00±10.00	70.00±10.00	70.00±10.00	67.50±7.50
	Raw data	80.00±0.00	95.00±0.00	80.00±0.00	95.00±0.00	95.00±0.00	90.00±0.00	95.00±0.00	95.00±0.00

3. The running time of the MFREN algorithm is shorter than that of the MFRMI algorithm. This result is primarily due to the operating mechanisms of the MFRMI and MFREN algorithms. The MFREN algorithm, which does not consider decision attributes during the reduction process, is not influenced by any specific predictive tasks, resulting in faster operation. In contrast, the MFRMI algorithm incorporates decision attributes. Although it takes slightly longer to run, it enhances the performance of attribute reduction compared to the MFREN algorithm.

5.8. Statistical tests

To assess the statistical significance of differences in classification accuracy among various attribute reduction methods, two statistical tests—the Friedman test and the Bonferroni-Dunn test—are employed. These tests help validate the proposed methods and determine whether significant differences exist between them.

Table 11
Comparison of classification accuracies of reduced data with 14 models.

Dataset	Algorithm	LB	KNN	SVC	LSVC	GN	LR	Average
Abs	MFRMI	86.49±0.0	56.76±0.0	64.86±0.0	64.86±0.0	85.14±2.70	87.84±2.70	81.18±2.12
	MFREN	86.49±0.0	56.76±0.0	64.86±0.0	64.86±0.0	85.14±2.70	87.84±2.70	81.08±2.32
	IARFCIE	81.08±0.0	71.62±1.35	56.76±0.00	56.76±0.00	75.68±0.0	81.08 ±	77.32±1.64
	MFIGI	59.46±0.0	64.86±0.0	70.27±0.0	70.27±0.0	67.57±0.0	67.57±0.0	61.97±0.0
	FSFRMI	86.49±0.0	75.68±0.0	75.68±0.0	89.19±0.0	78.38±0.0	86.49±0.0	82.63±0.0
	FNRS	81.08±0.0	81.08±0.0	67.57±0.0	67.57±0.0	67.57±0.0	67.57±0.0	77.22±0.0
	FRAR	70.27±5.41	70.27±10.81	68.92±2.70	68.92±2.70	62.16±10.81	68.92±2.70	67.57±5.02
	Raw data	86.49±0.0	75.68±0.0	81.08±0.0	86.49±0.0	78.38±0.0	89.19±0.0	83.20±0.0
Bri	MFRMI	100.0±0.0	57.14±0.0	71.43±0.0	78.57±14.29	100.0±0.0	100.0±0.0	93.37±1.02
	MFREN	100.0±0.0	57.14±0.0	71.43±0.0	78.57±14.29	100.0±0.0	100.0±0.0	93.37±1.02
	IARFCIE	14.29±0.0	28.57±0.0	28.57±0.0	14.29±14.29	35.71±7.14	21.43±7.14	37.24±15.82
	MFIGI	42.86±0.0	57.14±0.0	71.43±0.0	71.43±0.0	28.57±0.0	71.43±0.0	61.22±0.0
	FSFRMI	85.71±2.86	85.71±2.86	85.71±2.86	78.57±4.29	64.29±4.29	78.57±4.29	83.16±3.16
	FNRS	100.0±0.0	100.0±0.0	85.71±0.0	57.14±0.0	100.0±0.0	71.43±0.0	93.88±0.0
	FRAR	71.43±0.0	71.43±2.86	78.57±1.43	64.29±1.43	42.86±5.71	64.29±1.43	75.00±2.14
	Raw data	100.0±0.0	57.14±0.0	71.43±0.0	85.71±0.0	100.0±0.0	100.0±0.0	93.88±0.0
Echo	MFRMI	90.91±0.0	90.91±0.0	90.91±0.0	90.91±0.0	90.91±0.0	90.91±0.0	88.31±0.0
	MFREN	90.91±0.0	90.91±0.0	90.91±0.0	90.91±0.0	90.91±0.0	90.91±0.0	88.31±0.0
	IARFCIE	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	98.70±0.0
	MFIGI	100.0±0.0	100.0±0.0	100.0±0.0	90.91±0.0	90.91±0.0	90.91±0.0	98.05±0.0
	FSFRMI	90.91±0.0	90.91±0.0	90.91±0.0	90.91±0.0	90.91±0.0	90.91±0.0	88.31±0.0
	FNRS	81.82±1.82	68.18±4.55	77.27±2.73	81.82±1.82	81.82±1.82	72.73±0.0	75.00±2.92
	FRAR	81.82±1.82	77.27±2.73	81.82±1.82	81.82±1.82	81.82±1.82	81.82±1.82	78.25±2.53
	Raw data	90.91±0.0	90.91±0.0	90.91±0.0	90.91±0.0	90.91±0.0	90.91±0.0	88.31±0.0
Spon	MFRMI	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	50.0±0.0	100.0±0.0	93.30±0.09
	MFREN	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	50.0±0.0	100.0±0.0	93.30±0.09
	IARFCIE	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	50.0±50.0	100.0±0.0	96.43±3.57
	MFIGI	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	18.75±12.5	100.0±0.0	94.20±0.89
	FSFRMI	100.0±0.0	100.0±0.0	100.0±0.0	93.75±6.25	50.0±25.00	100.0±0.0	95.98±2.68
	FNRS	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	50.0±10.0	100.0±0.0	96.43±0.71
	FRAR	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	25.00±25.00	100.0±0.0	94.64±1.79
	Raw data	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	62.50±0.0	100.0±0.0	97.32±0.0
Tran	MFRMI	74.67±0.0	74.67±0.0	74.67±0.0	74.67±0.0	74.67±0.0	74.67±0.0	73.05±0.0
	MFREN	74.67±0.0	74.67±0.0	74.67±0.0	74.67±0.0	74.67±0.0	74.67±0.0	73.05±0.0
	IARFCIE	70.67±0.0	78.67±0.0	81.33±0.0	81.33±0.0	80.0±0.0	82.67±0.0	72.86±0.0
	MFIGI	82.67±0.0	82.67±0.0	84.0±0.0	84.0±0.0	84.0±0.0	84.0±0.0	82.48±0.0
	FSFRMI	74.67±0.0	74.67±0.0	74.67±0.0	74.67±0.0	74.67±0.0	74.67±0.0	73.05±0.0
	FNRS	73.33±2.67	73.33±5.33	74.00±1.33	74.67±0.0	73.33±2.67	74.67±0.0	72.52±4.86
	FRAR	74.67±0.0	74.00±4.00	74.67±0.0	74.67±0.0	74.67±0.0	74.67±0.0	74.24±1.62
	Raw data	73.33±0.0	69.33±0.0	74.67±0.0	72.00±0.0	74.67±0.0	72.00±0.0	72.29±0.0
Wdbc	MFRMI	96.49±0.0	91.23±0.0	91.23±0.0	75.44±28.07	91.23±0.0	90.35±1.75	92.54±2.63
	MFREN	96.49±0.0	91.23±0.0	91.23±0.0	75.44±28.07	91.23±0.0	90.35±1.75	92.54±2.63
	IARFCIE	100.0±0.0	92.98±0.0	92.98±0.0	94.74±0.0	94.74±0.0	96.49±0.0	96.74±0.0
	MFIGI	57.89±0.0	54.39±0.0	63.16±0.0	63.16±0.0	63.16±0.0	63.16±0.0	57.77±0.0
	FSFRMI	91.23±7.02	88.60±1.75	91.23±0.0	81.58±15.79	90.35±1.75	92.98±3.51	89.66±4.39
	FNRS	82.46±1.75	81.58±1.23	86.84±0.88	60.53±1.58	85.09±0.53	52.63±0.0	78.32±1.25
	FRAR	93.86±0.53	93.86±0.53	93.86±0.53	94.74±0.35	91.23±1.05	94.74±0.35	92.67±0.64
	Raw data	94.74±0.0	96.49±0.0	96.49±0.0	96.49±0.0	94.74±0.0	96.49±0.0	95.36±0.0
Wpbc	MFRMI	87.5±5.0	85.0±0.0	85.0±0.0	85.0±0.0	80.0±0.0	87.5±5.0	83.39±3.93
	MFREN	87.5±5.0	85.0±0.0	85.0±0.0	85.0±0.0	80.0±0.0	87.5±5.0	83.39±3.93
	IARFCIE	90.0±0.0	75.0±0.0	80.0±0.0	80.0±0.0	60.0±0.0	90.0±0.0	78.57±0.0
	MFIGI	60.0±0.0	60.0±0.0	65.0±0.0	60.0±0.0	60.0±0.0	60.0±0.0	58.57±0.0
	FSFRMI	75.0±10.0	82.5±5.0	85.0±0.0	70.0±30.0	82.5±5.0	85.0±0.0	75.89±9.64
	FNRS	72.5±3.50	75.0±3.0	80.0±1.0	82.5±0.5	77.5±1.50	85.0±0.0	72.50±3.43
	FRAR	75.0±3.0	77.5±1.50	85.0±0.0	80.0±1.0	75.0±1.0	80.0±1.0	76.96±1.82
	Raw data	90.0±0.0	85.0±0.0	85.0±0.0	85.0±0.0	82.5±5.0	87.5±5.0	81.07±0.0

First, the Friedman test is used to test the null hypothesis that all methods perform equally well across all datasets. The test is appropriate because it is a non-parametric method that compares multiple methods based on ranked data, which is particularly useful in situations where assumptions of normality cannot be made. The Friedman test statistic is calculated as follows [42]:

$$\chi_F^2 = \frac{12N}{k(k+1)}\left(\sum_{j=1}^{k} R_j^2 - \frac{k(k+1)^2}{4}\right), \quad F_F = \frac{(N-1)\chi_F^2}{N(k-1)-\chi_F^2}$$

Table 12
Comparison of classification accuracies of reduced data with 14 models.

Dataset	Algorithm	LB	KNN	SVC	LSVC	GN	LR	Average
Aust	MFRMI	100.0±0.0	87.68±0.0	69.20±0.0	67.03±0.0	69.20±0.0	86.23±0.0	90.74±0.0
	MFREN	100.0±0.0	87.68±0.0	69.20±0.0	67.03±0.0	69.20±0.0	86.23±0.0	90.74±0.0
	IARFCIE	100.0±0.0	89.86±0.0	66.67±0.0	60.87±31.16	75.91±5.43	75.54±4.71	90.01±3.21
	MFIGI	72.83±0.0	71.01±0.0	71.74±0.0	72.10±0.0	70.65±0.0	71.74±0.0	71.97±0.0
	FSFRMI	100.0±0.0	97.10±0.0	89.13±0.0	88.04±0.0	83.70±0.0	88.41±0.0	95.86±0.0
	FNRS	69.02±20.65	67.75±18.12	59.60±1.81	59.24±1.09	59.96±2.54	59.06±0.72	66.80±16.20
	FRAR	86.23±0.0	86.23±0.0	86.23±0.0	86.23±0.0	86.23±0.0	86.23±0.0	86.23±0.0
	Raw data	100.0±0.0	97.10±0.0	89.13±0.0	88.04±0.0	83.70±0.0	88.41±0.0	95.86±0.0
Clev	MFRMI	100.0±0.0	90.76±0.0	88.24±0.0	60.50±0.0	62.18±0.0	64.71±0.0	90.46±0.0
	MFREN	100.0±0.0	90.76±0.0	88.24±0.0	60.50±0.0	62.18±0.0	64.71±0.0	90.46±0.0
	IARFCIE	100.0±0.0	79.83±0.0	52.10±0.0	39.08±56.30	60.92±10.92	61.34±5.04	85.11±5.40
	MFIGI	90.76±0.0	86.55±0.0	56.30±0.0	57.98±0.0	57.98±0.0	59.66±0.0	80.49±0.0
	FSFRMI	100.0±0.0	82.35±0.0	49.58±0.0	15.97±0.0	54.62±0.0	52.10±0.0	81.45±0.0
	FNRS	74.79±5.42	68.49±4.29	63.87±2.86	53.36±0.92	52.52±0.92	53.78±1.01	68.88±3.94
	FRAR	74.79±0.0	70.59±0.0	70.59±0.0	59.66±0.0	53.78±0.0	58.82±0.0	70.29±0.0
	Raw data	100.0±0.0	90.76±0.0	88.24±0.0	60.50±0.0	62.18±0.0	64.71±0.0	90.46±0.0
Derm	MFRMI	100.0±0.0	90.28±2.78	85.07±4.86	79.86±5.56	55.90±10.42	80.21±2.08	92.01±1.88
	MFREN	100.0±0.0	88.20±4.17	54.86±9.72	62.15±20.14	46.88±14.58	71.53±8.33	87.30±4.09
	IARFCIE	100.0±0.0	95.14±1.39	89.58±0.0	100.0±0.0	89.58±1.39	99.31±1.39	98.12±0.30
	MFIGI	71.53±5.56	69.44±5.56	71.53±9.72	70.49±7.64	58.68±10.42	69.79±6.25	70.26±6.40
	FSFRMI	100.0±0.0	91.67±0.0	97.22±0.0	100.0±0.0	81.25±0.0	100.0±0.0	97.87±0.0
	FNRS	65.97±34.03	57.64±7.64	63.54±6.32	64.93±6.04	56.94±5.00	63.54±5.90	64.31±6.58
	FRAR	75.35±2.08	65.63±2.08	75.35±2.08	75.35±2.08	69.79±13.19	75.35±2.08	74.26±2.88
	Raw data	100.0±0.0	91.67±0.0	97.22±0.0	100.0±0.0	81.25±0.0	100.0±0.0	97.87±0.0
Germ	MFRMI	99.75±0.25	89.38±2.25	88.12±2.75	71.88±1.25	73.00±2.00	72.13±1.25	91.39±0.89
	MFREN	99.75±0.25	89.38±2.25	88.12±2.75	71.88±1.25	73.00±2.00	72.13±1.25	91.39±0.89
	IARFCIE	97.5±2.5	90.25±0.5	68.25±0.0	59.0±16.0	73.5±9.0	75.13±9.75	89.02±3.61
	MFIGI	69.5±0.0	65.75±0.0	69.5±0.0	69.5±0.0	69.5±0.0	69.5±0.0	69.23±0.0
	FSFRMI	94.50±2.00	92.75±2.00	67.38±3.75	66.13±10.25	73.13±0.75	71.50±1.50	88.86±1.93
	FNRS	82.12±3.33	67.38±4.38	74.00±1.70	67.63±0.63	67.50±1.20	66.75±0.40	77.17±2.71
	FRAR	65.50±0.0	61.50±0.0	65.50±0.0	65.50±0.0	61.50±0.0	65.50±0.0	64.93±0.0
	Raw data	100.0±0.0	89.25±0.0	66.75±0.0	54.25±0.0	72.25±0.0	73.00±0.0	88.79±0.0
Lymp	MFRMI	100.0±0.0	100.0±0.0	91.67±5.0	78.33±6.67	75.83±1.67	80.0±10.0	94.64±2.14
	MFREN	100.0±0.0	100.0±0.0	91.67±10.0	78.33±6.67	75.83±1.67	80.0±10.0	94.64±2.14
	IARFCIE	85.0±15.0	83.33±23.33	77.5±15.0	71.67±16.67	74.17±15.0	71.67±20.0	81.79±12.74
	MFIGI	71.67±0.0	63.33±0.0	71.67±0.0	71.67±0.0	46.67±0.0	68.33±0.0	69.05±0.0
	FSFRMI	100.0±0.0	93.33±6.67	89.17±5.84	77.50±18.33	65.00±16.67	75.00±23.33	92.62±5.95
	FNRS	74.17±4.17	65.00±4.67	70.0±3.33	67.50±3.50	34.17±6.50	69.17±2.83	69.46±4.15
	FRAR	72.50±2.50	62.50±4.83	72.50±2.50	71.67±2.33	36.67±7.00	71.67±2.33	69.11±2.96
	Raw data	100.0±0.0	100.0±0.0	96.67±0.0	86.67±0.0	85.00±0.0	88.33±0.0	96.43±0.0
Magi	MFRMI	99.21±0.0	93.43±0.0	84.23±0.0	76.87±0.0	73.06±0.0	77.40±0.0	92.34±0.0
	MFREN	99.21±0.0	93.43±0.0	84.23±0.0	76.87±0.0	73.06±0.0	77.40±0.0	92.34±0.0
	IARFCIE	99.87±0.0	91.20±0.0	77.66±0.0	47.96±0.0	70.70±0.0	74.38±0.0	89.33±0.0
	MFIGI	87.12±0.0	92.12±0.0	77.79±0.0	76.22±0.0	62.02±0.0	62.02±0.0	87.61±0.0
	FSFRMI	99.21±0.0	93.43±0.0	84.23±0.0	76.87±0.0	73.06±0.0	77.40±0.0	92.34±0.0
	FNRS	80.62±3.61	78.58±2.76	71.75±1.58	68.53±0.93	67.54±0.87	68.86±1.00	78.19±2.55
	FRAR	94.88±4.21	91.13±2.50	76.61±7.10	73.52±5.65	72.21±4.07	73.52±5.39	90.14±2.78
	Raw data	99.21±0.0	93.43±0.0	84.23±0.0	76.87±0.0	73.06±0.0	77.40±0.0	92.34±0.0
Tic	MFRMI	100.0±0.0	94.27±0.0	100.0±0.0	70.05±0.0	72.14±0.0	69.27±0.0	92.69±0.0
	MFREN	100.0±0.0	94.27±0.0	100.0±0.0	70.05±0.0	72.14±0.0	69.27±0.0	92.69±0.0
	IARFCIE	100.0±0.0	94.01±0.0	99.48±0.0	70.57±0.0	71.35±0.0	70.31±0.0	92.80±0.0
	MFIGI	71.22±0.26	67.45±0.52	70.18±0.78	65.89±0.52	65.89±1.04	65.89±0.52	69.64±0.56
	FSFRMI	100.0±0.0	94.27±0.0	100.0±0.0	70.05±0.0	72.14±0.0	69.27±0.0	92.69±0.0
	FNRS	82.42±3.52	77.86±3.28	82.03±3.59	65.49±0.91	67.45±0.99	65.76±0.34	78.23±2.87
	FRAR	66.41±0.0	66.41±0.0	66.41±0.0	64.84±0.0	64.84±0.0	58.07±0.0	65.59±0.0
	Raw data	100.0±0.0	94.27±0.0	100.0±0.0	70.05±0.0	72.14±0.0	69.27±0.0	92.69±0.0

where N represents the number of datasets, k is the number of methods, and F_F follows an F-distribution with $(k-1)$ and $(k-1)(N-1)$ degrees of freedom. R_j is the average ranking of the j-th method across all datasets, defined as $R_j = \sum_{i=1}^{N} r_{ji}/N$, where r_{ji} denotes the ranking of the j-th method on the i-th dataset.

In this paper, two algorithms are proposed: MFREN and MFRMI. Each of these algorithms is compared with the other five algorithms through separate Friedman tests, meaning two Friedman tests are conducted, involving six algorithms in total. Based on the

Table 13
Comparison of classification accuracies of reduced data with 14 models.

Dataset	Algorithm	LB	KNN	SVC	LSVC	GN	LR	Average
Vehi	MFRMI	100.0±0.0	89.97±0.0	59.59±0.0	77.58±0.0	46.02±0.0	69.62±0.0	88.45±0.0
	MFREN	100.0±0.0	89.97±0.0	59.59±0.0	77.58±0.0	46.02±0.0	69.62±0.0	88.45±0.0
	IARFCIE	100.0±0.0	92.33±0.0	57.82±0.0	72.57±0.0	50.15±0.0	72.27±0.0	88.77±0.0
	MFIGI	94.10±0.0	84.07±0.0	47.20±0.0	42.48±0.0	42.48±0.0	42.48±0.0	77.29±0.0
	FSFRMI	100.0±0.0	91.74±0.0	61.36±0.0	38.05±0.0	49.26±0.0	60.47±0.0	84.87±0.0
	FNRS	69.03±6.19	62.09±6.28	55.75±5.31	46.17±3.27	42.33±2.63	46.46±3.16	62.16±5.45
	FRAR	97.64±0.30	92.77±0.15	76.70±0.24	61.21±0.88	48.53±0.21	61.95±0.47	86.46±0.22
	Raw data	100.0±0.0	89.97±0.0	59.59±0.0	77.58±0.0	46.02±0.0	69.62±0.0	88.45±0.0
Z	MFRMI	100.0±0.0	100.10±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0
	MFREN	100.0±0.0	100.10±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0
	IARFCIE	73.17±0.0	73.17±0.0	73.17±0.0	51.22±0.0	41.46±0.0	53.66±0.0	67.94±0.0
	MFIGI	81.71±2.44	89.02±2.44	81.71±2.44	81.71±2.44	79.27±7.32	79.27±7.32	81.88±3.14
	FSFRMI	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0
	FNRS	65.85±5.37	62.20±4.63	64.63±5.12	58.54±3.90	56.10±6.34	57.32±3.66	63.68±5.14
	FRAR	68.29±0.0	68.29±0.0	68.29±0.0	68.29±0.0	68.29±0.0	68.29±0.0	68.29±0.0
	Raw data	100.0±0.0	100.10±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0
GSE	MFRMI	88.89±11.11	69.44±5.56	83.33±0.00	58.33±5.56	80.56±5.56	58.33±5.56	78.57±10.32
	MFREN	75.00±2.78	80.56±2.78	77.78±0.00	69.44±2.78	55.56±5.56	75.00±2.78	72.82±2.98
	MFIGI	47.22±27.78	50.00±11.11	55.56±11.11	36.11±27.78	50.00±22.22	55.56±22.22	48.41±26.19
	IARFCIE	66.67±0.00	55.56±0.00	66.67±0.00	66.67±0.00	72.22±0.00	55.56±0.00	64.68±0.00
	FSFRMI	63.89±5.56	63.89±5.56	61.11±0.00	55.56±0.00	55.56±0.00	55.56±0.00	60.32±4.76
	FNRS	38.89±0.00	44.44±0.00	50.00±0.00	50.00±0.00	50.00±0.00	50.00±0.00	44.05±0.00
	FRAR	/	/	/	/	/	/	/
	Raw data	61.11±0.00	61.11±0.00	55.56±0.00	61.11±0.00	50.00±0.00	61.11±0.00	58.73±0.00
LK	MFRMI	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0
	MFREN	100.0±0.0	87.5±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	96.43±2.68
	IARFCIE	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	92.86±0.0
	MFIGI	87.5±12.5	75.0±12.5	75.0±12.5	62.5±12.5	87.5±12.5	87.5±12.5	83.5±13.84
	FSFRMI	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	95.54±1.79
	FNRS	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	95.54±1.79
	FRAR	87.5±12.5	93.75±6.25	93.75±6.25	100.0±0.0	93.75±6.25	100.0±0.0	91.97±8.04
	Raw data	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	92.86±0.0
DL	MFRMI	100.0±0.0	87.5±0.0	87.5±0.0	100.0±0.0	100.0±0.0	100.0±0.0	97.77±0.90
	MFREN	100.0±0.0	86.5±0.0	86.5±0.0	100.0±0.0	100.0±0.0	100.0±0.0	95.34±0.90
	IARFCIE	100.0±0.0	87.5±0.0	87.5±0.0	100.0±0.0	100.0±0.0	100.0±0.0	92.86±0.0
	MFIGI	100.0±0.0	100.0±0.0	100.0±0.0	31.25±31.25	81.25±12.5	81.25±18.75	84.82±8.04
	FSFRMI	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	93.75±6.25	100.0±0.0	93.75±5.4
	FNRS	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	93.75±6.25	100.0±0.0	93.75±5.4
	FRAR	100.0±0.0	81.25±18.75	75.0±12.5	68.75±18.75	81.25±6.25	75.0±12.5	79.02±11.16
	Raw data	100.0±0.0	87.5±0.0	87.5±0.0	100.0±0.0	100.0±0.0	100.0±0.0	92.86±0.0
AR	MFRMI	87.5±5.0	85.0±10.0	82.5±5.0	92.5±5.0	75.0±0.0	90.0±10.0	83.57±7.86
	MFREN	85.0±0.0	95.0±0.0	95.0±0.0	82.5±7.5	75.0±0.0	90.0±0.0	86.79±3.57
	IARFCIE	70.0±10.0	62.5±22.5	65.0±5.0	77.5±17.5	50.0±10.0	77.5±17.5	67.86±11.43
	MFIGI	70.0±20.0	62.5±15.0	62.5±15.0	55.0±20.0	62.5±5.0	65.0±20.0	63.39±16.79
	FSFRMI	95.0±0.0	95.0±0.0	100.0±0.0	90.0±0.0	85.0±0.0	95.0±0.0	91.79±0.00
	FNRS	95.0±0.0	95.0±0.0	100.0±0.0	90.0±0.0	85.0±0.0	95.0±0.0	91.79±0.00
	FRAR	65.0±5.0	35.0±5.0	70.0±5.0	70.0±0.0	55.0±15.0	70.0±0.0	66.25±7.32
	Raw data	95.0±0.0	95.0±0.0	100.0±0.0	90.0±0.0	85.0±0.0	95.0±0.0	91.79±0.00

classification accuracy of the 20 datasets presented in Tables 8–13, the average rankings of the six methods in two groups across 14 classification models are calculated. The calculation results for the values of χ_F^2 and F_F are shown in Table 17.

The critical value for $F(5, 5\times 19)$ is 2.310. Since the computed Friedman test statistic exceeds this threshold, we reject the null hypothesis, indicating that significant differences exist between the methods.

Subsequently, the Bonferroni-Dunn test is applied for pairwise comparisons of methods to identify significant differences between them. The critical difference (CD) is calculated using the formula [43]:

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

where q_α is the critical value for the chosen significance level ($\alpha = 0.05$, where $q_{0.05} = 1.96$), and N is the number of datasets. For $\alpha = 0.05$, the corresponding CD value is 1.339.

Fig. 5 presents a comparison of the average accuracies of 14 methods using the CD diagram proposed by Demsar and Schuurmans, clearly illustrating the performance differences among the algorithms. The horizontal axis represents the mean ranking of each

Table 14
Comparison of Running Time of Different Methods (Low-Dimensional and High-Dimensional).

Datasets	FRAR	FNRS	FSFRMI	MFIGI	IARFCIE	MFREN	MFRMI
Low-Dimensional Datasets							
Abs	2.82	81.12	8.97	4.3	58.8	2.14	2.82
Bri	1.3	2.59	0.14	2.3	38	0.02	0.04
Echo	0.28	0.31	0.63	2	11	0.41	0.47
Spon	1076.89	7.02	1.66	5.7	49.4	0.08	0.1
Tran	0.66	96.93	3.15	4.2	28.7	0.87	1.21
Wdbc	24786	1260.2	87.53	10	184.7	11.07	12
Wpbc	2430	85.21	14.76	3.5	54.6	1.59	1.62
Aust	6.78	3360.05	358.73	156.2	511.6	74.64	81.86
Clev	24300	1728.65	59.99	40.1	377.6	12.77	13.99
Derm	19440	4916.46	80.54	124.7	1311.1	148.96	174.51
Germ	12.62	12647.03	1725.59	330.4	1652.4	281.19	325.87
Lymp	19800	420.13	4.19	14.7	174.8	12.38	13.73
Magi	17.31	48795.67	625.38	451.5	614.5	491.98	557.88
Tic-	5.14	4923.4	30.56	163.6	453.1	205.79	253.46
Vehi	42387.07	21543.49	909.37	215.9	1399.9	267.02	313.7
Z	2.04	132.13	2.33	11	109	3.29	3.56
Average	8391.81	6250.02	244.6	96.26	439.33	94.64	109.80
High-Dimensional Datasets							
GSE	>6d	9471.8	99031.8	1313.4	62703.7	18319.45	27607.8
DL	663.3	559.3	26440.6	88.6	44120.6	13416.6	15476.7
LK	1350.1	627.2	73095.8	180	68580.2	27222.3	31400.1
AR	2597.4	3126.8	384033.2	777.3	418916.2	149014.8	163654.68
Average	130752.7	3446.3	145650.4	589.8	148580.2	51993.3	59534.8

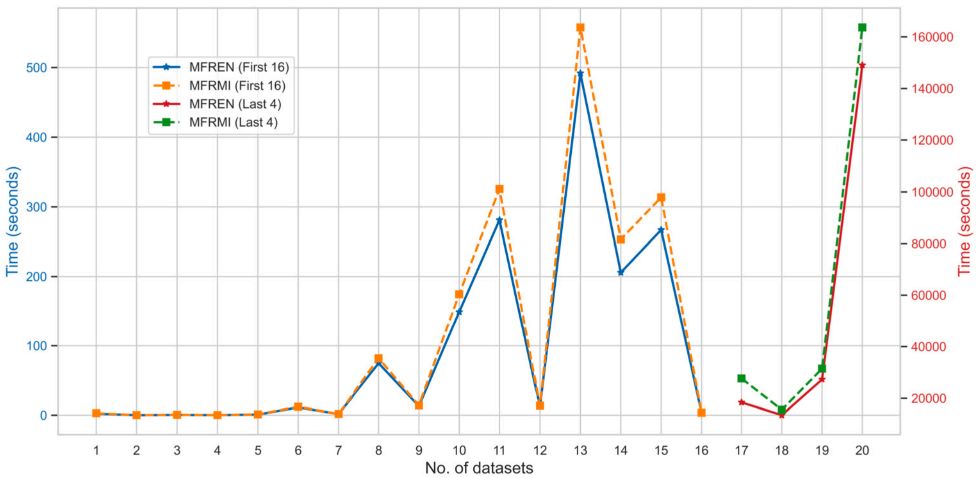


Fig. 4. Running Time Performance Comparison of Attribute Reduction.

Table 15
Summary of Key Features and Properties of Attribute Reduction Algorithms.

Alg.	Key Feature	Properties (Advantages & Disadvantages)
FRAR MFIGI	Fuzzy rough set (positive regions) Fuzzy implication granularity information	Efficient reduction, but may lead to information loss from over-reduction.
FNRS	Fuzzy neighborhood rough set	Efficient and handles fuzziness, but struggles with extreme fuzziness.
FSFRMI	Fuzzy rough entropy	Better performance on low-dimensional data, but relatively worse on high-dimensional data
IARFCIE	Fuzzy conditional entropy	Effective reduction in both low-dimensional and high-dimensional data, but cannot achieve further improvement.
MFREN MFRMI	Metric fuzzy rough entropy Metric mutual information	Both have fast runtime and strong classification capabilities, effectively balancing accuracy and efficiency. However, MFRMI relies on decision attributes to function properly.

Table 16
Average of Accuracy Ranking, Running Time and Reduction Size of Attribute Reduction Algorithms.

Alg.	FRAR	FNRS	FSFRMI	MFIGI	IARFCIE	MFREN	MFRMI
Accuracy Ranking	5.89	6.05	3.3	6.25	4.15	2.7	2.5
Low Dim. Running Time (s)	8391.81	6250.02	244.6	96.26	439.33	94.64	109.80
High Dim. Running Time (s)	130752.7	3446.3	145650.4	589.8	148580.2	51993.3	59534.8
Low Dim. Reduction Size	3.4	3.3	9.1	3.3	13.2	8.3	8.3
High Dim. Reduction Size	1.8	2500.8	8267.8	3.16	8024.5	227.0	228.4

Table 17
The Statistical Test of 6 Attribute Reduction Methods Under Average Effect of 14 Models.

Classifiers	Mean ranking						χ_F^2	F_F
	FRAR	FNRS	FSFRMI	MFIGI	IARFCIE	MFRMI(EN)		
Average	4.15	4.65	2.35	4.6	2.75	2.05	12.8049	3.5971

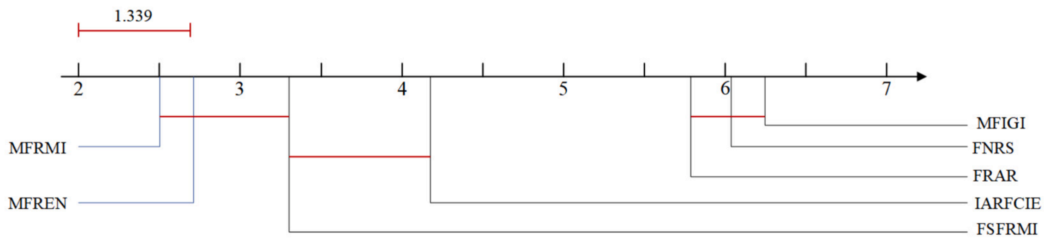


Fig. 5. Comparisons of accuracies with the average effect of 14 methods.

method, with lower values indicating better performance. When the CD threshold (marked by the red line at the top) is exceeded, the performance differences between methods become statistically significant. Methods that are not connected by horizontal lines show statistically significant differences, while those connected by horizontal lines do not show significant differences.

It is evident that MFRMI and MFREN have the lowest average rankings, with no significant difference between them, indicating their similar and superior performance. Further analysis reveals that the MFRMI and MFREN algorithms significantly outperform the other methods, except for FSFRMI, demonstrating their advantage in high accuracy.

6. Conclusion

Fuzzy rough set theory has proven to be an effective method for addressing issues of uncertainty in data, particularly when dealing with large-scale and high-dimensional datasets. However, traditional techniques often focus on extracting uncertainty information from the lower approximations of fuzzy rough sets, without sufficiently considering the critical measurement aspects required for attribute reduction. Furthermore, conventional methods that use logical conjunctions for calculating fuzzy similarity relations can suffer from information loss, leading to inaccuracies in the handling of large attribute sets.

In this paper, we propose a new method for calculating entropy using distance measures within the fuzzy rough set framework. This method helps mitigate information loss by improving the accuracy and stability of entropy computations, thus enhancing the precision of information granulation. Additionally, we present new concepts, such as fuzzy joint rough entropy, fuzzy conditional rough entropy, and fuzzy rough mutual information, to better evaluate attribute correlation and uncertainty. These innovations not only boost feature selection and classification performance but also demonstrate their robustness and effectiveness in real-world fuzzy information systems, particularly in complex datasets.

While the proposed attribute reduction methods excel in static datasets, their applicability to dynamic, real-time data is limited. Given the rapidly changing nature of data in real-time applications, recalculating entropy values and attribute importance frequently may introduce substantial computational overhead. Future research will focus on improving the efficiency of entropy computation for dynamic datasets. We aim to develop adaptive entropy calculation methods to handle real-time data streams more effectively, thereby minimizing recalculation costs. Integrating real-time data processing with adaptive learning systems will make the proposed methods more suitable for environments like financial forecasting, autonomous driving, and health monitoring, where data is continuously generated, and immediate decision-making is essential.

Further work will also aim to extend this framework by incorporating unsupervised learning and multi-view learning techniques, enabling broader applications in data mining, pattern recognition, and intelligent decision support systems. The roadmap for future research, as illustrated in Fig. 6, outlines key stages including dynamic attribute reduction, real-time decision making, adaptive learning based on feedback, and the integration of online learning. These developments are crucial for enhancing model adaptability to rapidly changing environments. By expanding the scope of application domains and optimizing for multi-scenario use, the proposed methods can be customized for specific needs in healthcare, transportation, finance, and more. In summary, while the current research

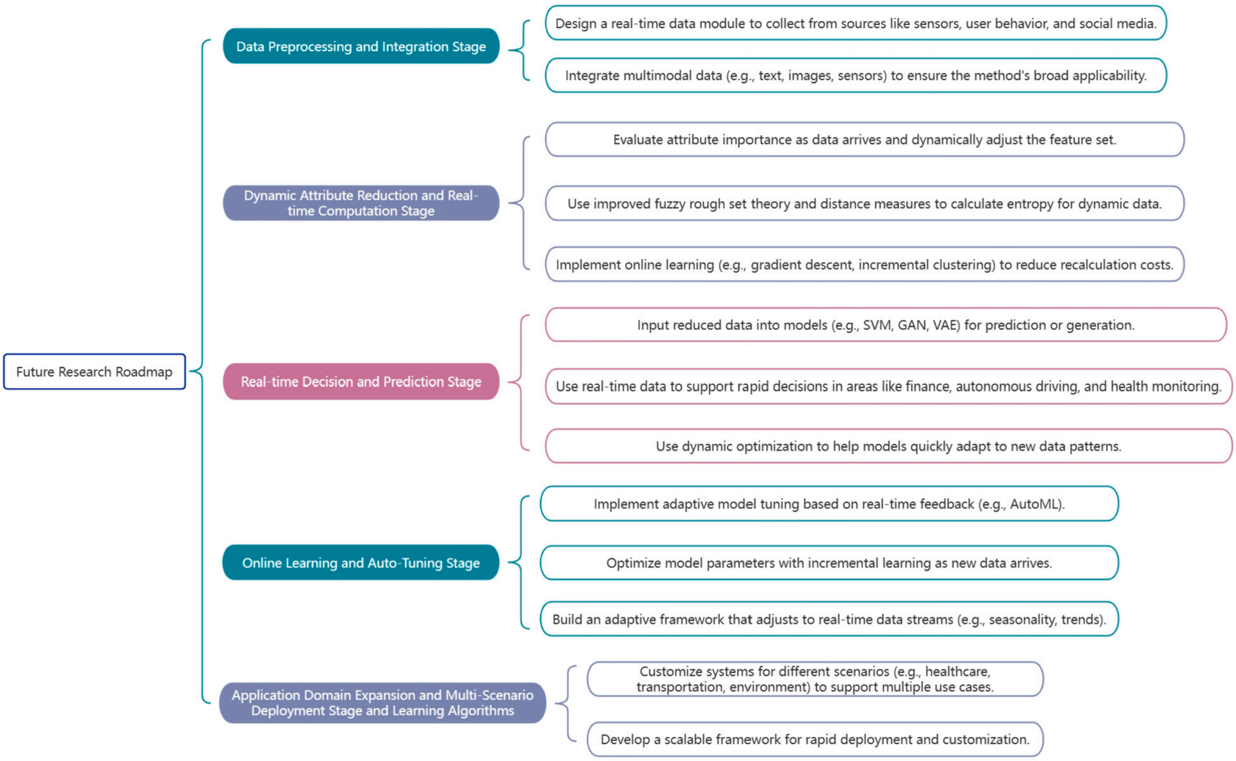


Fig. 6. Directions of Future Research.

presents promising results for static data, adapting the methods for dynamic, real-time environments and improving computational efficiency will be pivotal for ensuring their success in fast-paced, data-driven industries.

CRediT authorship contribution statement

Guirong Peng: Writing – original draft, Visualization, Software, Methodology, Data curation, Conceptualization. **Fei Li:** Supervision, Investigation. **Wei Yao:** Writing – review & editing.

Funding

This work was supported by the National Natural Science Foundation of China (Nos. 12371462, 12111540250, 12231007) and the Fundamental Research Funds for the Central Universities (No. 2019zy20).

Declaration of competing interest

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

Data availability

Data will be made available on request.

References

[1] W.B. Feng, T.T. Sun, A dynamic attribute reduction algorithm based on relative neighborhood discernibility degree, Sci. Rep. 14 (1) (2024) 15637.
[2] J.H. Dai, Z.L. Zhu, M. Li, X.T. Zou, C.C. Zhang, Attribute reduction for heterogeneous data based on monotonic relative neighborhood granularity, Int. J. Approx. Reason. 170 (2024) 109210.
[3] B.Z. Jin, X.J. Xu, Price forecasting through neural networks for crude oil, heating oil, and natural gas, Meas. Energy 1 (1) (2024) 100001.
[4] Z. Pawlak, Rough sets, Int. J. Comput. Inf. Sci. 11 (1982) 341–356.
[5] X.N. Song, Y.L. Song, V. Stojanovic, S. Song, Improved dynamic event-triggered security control for T–S fuzzy lpv-pde systems via pointwise measurements and point control, Int. J. Fuzzy Syst. 25 (8) (2023) 3177–3192.
[6] Z.B. Du, X.P. Xie, Z.F. Qu, Y.Y. Hu, V. Stojanovic, Dynamic event-triggered consensus control for interval type-2 fuzzy multi-agent systems, IEEE Trans. Circuits Syst. I, Regul. Pap. 71 (8) (2024) 3857–3866.

[7] R.M.K. Ullah, U. Qamar, M.S. Raza, J.A. Erkoyuncu, An incremental approach for calculating dominance-based rough set dependency, Soft Comput. 28 (5) (2024) 3757–3781.
[8] Y.Y. Fan, X.H. Zhang, J.Q. Wang, Overlap function-based fuzzy β-covering relations and fuzzy β-covering rough set models, Int. J. Approx. Reason. 169 (2024) 109164.
[9] S.Z. Wu, L.T. Wang, S.Y. Ge, Z.W. Hao, Y.L. Liu, Neighborhood rough set with neighborhood equivalence relation for feature selection, Knowl. Inf. Syst. 66 (3) (2024) 1833–1859.
[10] A. Theerens, C. Cornelis, On the granular representation of fuzzy quantifier-based fuzzy rough sets, Inf. Sci. 665 (2024) 120385.
[11] D. Dubois, H. Prade, Putting rough sets and fuzzy sets together, in: Intelligent Decision Support: Handbook of Applications and Advances of the Rough Sets Theory, Springer, 1992, pp. 203–232.
[12] Z.W. Li, S.X. Wei, S.P. Liu, Outlier detection using conditional information entropy and rough set theory, J. Intell. Fuzzy Syst. 46 (Preprint) (2024) 1899–1918.
[13] Z.W. Li, Q.L. Zhang, S.P. Liu, Y.C. Peng, L.L. Li, Information fusion and attribute reduction for multi-source incomplete mixed data via conditional information entropy and ds evidence theory, Appl. Soft Comput. 151 (2024) 111149.
[14] K.L. Qu, P. Gao, Q. Dai, Y.H. Sun, X. Hua, Attribute reduction using self-information uncertainty measures in optimistic neighborhood extreme-granulation rough set, Inf. Sci. 686 (2025) 121340.
[15] Q. Wang, X.Y. Zhang, Feature selection using three-stage heuristic measures based on mutual fuzzy granularities, Appl. Intell. 54 (2) (2024) 1445–1473.
[16] Y.Y. Yang, D.G. Chen, Z.Y. Ji, X. Zhang, L.J. Dong, A two-way accelerator for feature selection using a monotonic fuzzy conditional entropy, Fuzzy Sets Syst. 483 (2024) 108916.
[17] K. Cai, W.H. Xu, An efficient multi-source information fusion approach for dynamic interval-valued data via fuzzy approximate conditional entropy, Int. J. Mach. Learn. Cybern. 15 (9) (2024) 3619–3645.
[18] J.H. Dai, Q. Liu, W.X. Chen, C.C. Zhang, Multi-label feature selection based on fuzzy mutual information and orthogonal regression, IEEE Trans. Fuzzy Syst. 32 (9) (2024) 5136–5148.
[19] X.Q. Ma, J. Wang, W.C. Yu, Q.L. Zhang, Attribute reduction of hybrid decision information systems based on fuzzy conditional information entropy, Comput. Mater. Continua 79 (2) (2024) 2063–2083.
[20] J.H. Dai, J.L. Chen, Feature selection via normative fuzzy information weight with application into tumor classification, Appl. Soft Comput. 92 (2020) 106299.
[21] L. Sun, L.Y. Wang, W.P. Ding, Y.H. Qian, J.C. Xu, Feature selection using fuzzy neighborhood entropy-based uncertainty measures for fuzzy neighborhood multigranulation rough sets, IEEE Trans. Fuzzy Syst. 29 (1) (2020) 19–33.
[22] A. Theerens, C. Cornelis, Fuzzy rough Choquet distances for classification, arXiv:2403.11843, 2024, https://doi.org/10.48550/arXiv.2403.11843.
[23] 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.
[24] J. Hu, W. Pedrycz, G.Y. Wang, A roughness measure of fuzzy sets from the perspective of distance, Int. J. Gen. Syst. 45 (3) (2016) 352–367.
[25] Y.C. Wang, X. Tang, Y.T. Pan, M.Q. Cai, Chebyshev distance entropy combined with todim method in lq-rofs multiattribute group decision making, Discrete Dyn. Nat. Soc. 2023 (1) (2023) 4993116.
[26] W.H. Xu, K. Cai, D.D. Wang, A novel information fusion method using improved entropy measure in multi-source incomplete interval-valued datasets, Int. J. Approx. Reason. 164 (2024) 109081.
[27] J.C. Xu, X.R. Meng, K.L. Qu, Y.H. Sun, Q.C. Hou, Feature selection using relative dependency complement mutual information in fitting fuzzy rough set model, Appl. Intell. 53 (15) (2023) 18239–18262.
[28] W. Yao, Y.H. She, L.X. Lu, Metric-based L-fuzzy rough sets: approximation operators and definable sets, Knowl.-Based Syst. 163 (2019) 91–102.
[29] W. Yao, G.X. Zhang, C.J. Zhou, Real-valued hemimetric-based fuzzy rough sets and an application to contour extraction of digital surfaces, Fuzzy Sets Syst. 459 (2023) 201–219.
[30] G.X. Zhang, W. Yao, Fuzzy rough sets based on modular hemimetrics, Soft Comput. 27 (8) (2023) 4393–4401.
[31] Y.Y. Yao, Relational interpretations of neighborhood operators and rough set approximation operators, Inf. Sci. 111 (1–4) (1998) 239–259.
[32] T. Beaubouef, F.E. Petry, Fuzzy rough set techniques for uncertainty processing in a relational database, Int. J. Intell. Syst. 15 (5) (2000) 389–424.
[33] Y.H. Qian, J.Y. Liang, Z.W. Wei-zhi, C.Y. Dang, Information granularity in fuzzy binary grc model, IEEE Trans. Fuzzy Syst. 19 (2) (2010) 253–264.
[34] D.G. Chen, Attribute Reduction Based on T_M-Fuzzy Rough Sets, Science Press, Beijing, ISBN 978-7-03-039017-2, 2013, pp. 110–115.
[35] Z.H. Wang, H.M. Chen, Z. Yuan, X.L. Yang, P.F. Zhang, T.R. Li, Exploiting fuzzy rough mutual information for feature selection, Appl. Soft Comput. 131 (2022) 109769.
[36] 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.
[37] L.L. Xie, G.P. Lin, J.J. Li, Y.D. Lin, A novel fuzzy-rough attribute reduction approach via local information entropy, Fuzzy Sets Syst. 473 (2023) 108733.
[38] W. Ahmed, M.M. Sufyan Beg, T. Ahmad, Entropy based feature selection for fuzzy set-valued information systems, 3D Res. 9 (2) (2018) 1–17.
[39] C.Z. Wang, M.W. Shao, Q. He, Y.H. Qian, Y.L. Qi, Feature subset selection based on fuzzy neighborhood rough sets, Knowl.-Based Syst. 111 (2016) 173–179.
[40] R. Jensen, Q. Shen, Semantics-preserving dimensionality reduction: rough and fuzzy-rough-based approaches, IEEE Trans. Knowl. Data Eng. 16 (12) (2004) 1457–1471.
[41] J.H. Dai, Z.L. Zhu, X.T. Zou, Fuzzy rough attribute reduction based on fuzzy implication granularity information, IEEE Trans. Fuzzy Syst. 32 (6) (2024) 3741–3752.
[42] M. Friedman, A comparison of alternative tests of significance for the problem of m rankings, Ann. Math. Stat. 11 (1) (1940) 86–92.
[43] J. Demšar, Statistical comparisons of classifiers over multiple data sets, J. Mach. Learn. Res. 7 (2006) 1–30.