Information Sciences 749 (2026) 123506

Contents lists available at ScienceDirect

Information Sciences

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

Adaptive anti-noise feature selection considering correlations between features based on interval-valued dominance relation

Yinliang Liu, Xiaoyan Zhang*

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

ARTICLE INFO

Keywords:
Adaptive anti-noise feature selection
Interval-valued dominance relation
Feature correlations
Noisy data processing
Robust feature selection

ABSTRACT

Feature selection in noisy interval-valued ordered decision systems is challenging due to data uncertainty and noise interference. Existing methods exhibit weak anti-noise ability, poor robustness, and an inability to handle interval-valued data, while neglecting the relationships between features. To address these limitations, this paper proposes an interval-valued dominance relation-based robust anti-noise feature selection framework with interactions consideration (IV-RANFSI), which integrates three synergistic innovations. The framework first employs an elastic anti-noise factor (EAF) to adaptively distinguish noise from boundary samples. Consequently, a multi-threshold parameterization strategy is established to ensure the stability and robustness of the resulting dominance relations across diverse interval-valued scenarios. Building upon these purified relations, a comprehensive evaluation framework is further developed to integrate dominance-based importance with interaction-redundancy analysis. The proposed method achieves superior performance across three state-of-the-art classifiers (SVM, Random Forest, and XGBoost) on twelve UCI datasets with synthetic label noise levels ranging from 0% to 40%, demonstrating significant improvements in classification accuracy under high noise conditions. The IV-RANFSI framework systematically addresses the fundamental challenges of robust feature selection in noisy interval-valued environments by providing adaptive noise rejection, comprehensive feature relationship analysis, and stable performance across varying noise levels.

1. Introduction

Feature selection is widely recognized as a cornerstone in various data-driven fields. In the context of data mining, it has been effectively utilized for discovering conditional functional dependency rules in big data [1], managing dynamic data via neighborhood rough sets [2], and enhancing geographical detectors for spatial data analysis [3]. Within the machine learning domain, feature selection plays a critical role in verifying massive-scale image data systems [4] and optimizing decision tree algorithms like ID3 to handle continuous attributes [5]. These specialized applications collectively demonstrate the necessity of feature selection for managing the complexities of high-dimensional classification tasks. Rough set theory, initially formalized by Pawlak in the early 1980s [6], has emerged as a robust mathematical framework for managing vague and uncertain data in analytical contexts. The theoretical foundations and practical utility of granular computing have been extensively explored across diverse domains. In the realm of knowledge discovery, recent advancements include the development of uncertainty measures like Zentropy to handle heterogeneous data [7]. Within decision support systems, granular frameworks have facilitated effective rule induction for specialized fields such as medical diagnostics [8]. The efficacy of this paradigm in pattern recognition is further evidenced by the integration of overlap functions in fuzzy rough sets for image edge extraction [9] and the use of concept-cognitive learning for dynamic bird song classification

* Corresponding author.
Email addresses: yin2728906486@email.swu.edu.cn (Y. Liu), zhangxiaoyan@swu.edu.cn (X. Zhang).

https://doi.org/10.1016/j.ins.2026.123506
Received 12 September 2025; Received in revised form 7 April 2026; Accepted 11 April 2026
Available online 15 April 2026

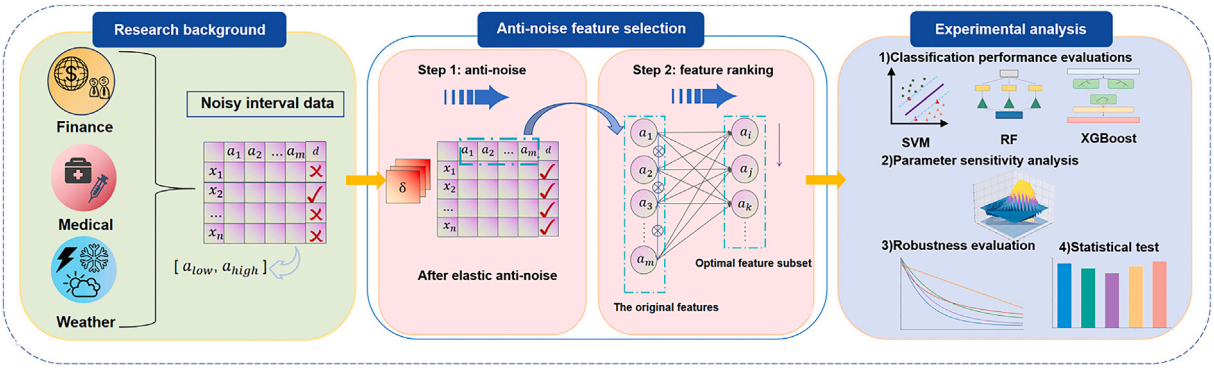

Fig. 1. The structure of this paper.

[10]. Furthermore, granular computing provides robust methodologies for knowledge reduction in formal contexts [11] and offers structured cognitive informatics models for interpreting complex learning processes [12]. As data environments grow increasingly complex, characterized by inherent noise, multi-valued ordered attributes, and intricate feature interdependencies, traditional feature selection methods face unprecedented challenges. The ability to maintain classification performance while ensuring efficient computation in these environments has emerged as a critical requirement for contemporary decision-making systems. This is particularly crucial in real-world applications, ranging from sensor networks to clinical decision support systems, where resilience against data flaws and a comprehensive consideration of feature synergies are essential. A significant limitation of current classification methods is their inadequate handling of noise patterns and the complexities of feature interactions that are inherent to ordered data structures. Our investigation identifies three fundamental limitations of existing feature selection paradigms:

(1) **Noise vulnerability in monotonic classification**: Current robust feature selection methods are insufficient for monotonic classification, as they are susceptible to noise that disrupts the ordinal relationships inherent in strictly monotonic feature-decision dependencies.
(2) **Multi-valued ordered data complexity**: Interval-valued ordered decision systems pose challenges to conventional dominance-based rough set approaches by introducing complexities related to information granularity in the processing of multi-valued attributes.
(3) **Feature synergy neglect**: Static feature selection methods frequently fail to account for the interdependencies and synergistic effects among features. Consequently, it is crucial to systematically incorporate feature interaction analysis to achieve optimal subset selection.

These limitations collectively motivate the development of our proposed interval-valued dominance relation-based robust anti-noise feature selection framework with interactions consideration (IV-RANFSI). The core strength of IV-RANFSI lies in the organic integration of three pivotal contributions that systematically address noise and uncertainty. First, we propose an elastic anti-noise factor (EAF) as the primary defense mechanism. By comparing local density ratios with feature subspace standard deviations, the EAF adaptively distinguishes noise samples from critical boundary samples. Building upon this purified sample space, we establish a multi-threshold parameterization strategy utilizing the Interval Dominance Degree (IDD), Interval Overlap Degree (IOD) and dominance threshold(δ_i). This strategy explicitly models real-world uncertainty and ensures that the constructed interval-valued dominance relations remain stable and robust across diverse noise levels. Finally, these robust relations are synthesized into a comprehensive feature evaluation framework. By integrating dominance-based mutual information (DMI) and conditional mutual information (DCMI), this framework quantifies inter-feature synergies and redundancies, thereby ensuring the selection of optimal feature subsets that maximize collective discriminative power in high-dimensional noisy scenarios. Extensive experiments conducted on twelve UCI datasets, with synthetic label noise levels ranging from 0% to 40%, validate the superiority of our IV-RANFSI approach in terms of noise tolerance, classification accuracy, and feature synergy utilization. These experiments demonstrate significant improvements across three state-of-the-art classifiers: SVM, Random Forest, and XGBoost. Collectively, these contributions establish a rigorous and flexible framework for feature extraction in interval-valued systems, comprehensively considering feature interactions and noise resilience.

The remainder of this paper is systematically organized as follows: First, Section 2 reviews related work in robust feature selection and interval-valued decision systems. To establish the theoretical groundwork, Section 3 presents essential mathematical preliminaries, including interval-valued ordered decision systems, relative density concepts, and information metrics. Building upon these foundations, Section 4 details the IV-RANFSI architecture, which incorporates elastic anti-noise mechanisms, interval-valued dominance relations, and comprehensive feature evaluation metrics. Subsequently, Section 5 presents experimental validation across synthetic and real-world datasets, accompanied by thorough performance comparisons and statistical analyses. Finally, Section 6 concludes with a discussion of research contributions and future directions. Fig. 1 provides an overview of the overall framework of our work.

Table 1
Summary of the current feature selection methods.

Methods	Applying fields	Disadvantages
SIDDM-FDNRS [16], FC [17], IV-DRSA [13]	Interval value processing	Weak anti-noise ability and poor robustness.
GLSFS [15], AE-RAR [14]	Interval value processing	Weak anti-noise ability and poor robustness.
NFRS [20], KNCMI [18]	Anti-noise processing	Cannot handle interval-valued data and does not consider the relationship between features.
DNCFAR [19], PMSNE [21]	Anti-noise processing	Cannot handle interval-valued data and does not consider the relationship between features.
RCDFS [22], FS-RSA [23], SNCMI [24]	Feature relationship consideration	Poor performance in noisy and interval-valued scenarios.

2. Related work

Contemporary research in robust feature selection has made significant strides across three primary dimensions: interval value processing, anti-noise mechanisms, and feature synergy considerations. This section systematically reviews existing approaches within these three categories and identifies critical research gaps that require integrated solutions.

The handling of interval-valued ordered data has seen notable progress through dominance-based rough set approaches. IV-DRSA [13] has pioneered the development of interval dominance degrees (IDD) and overlap metrics (IOD) to construct interval-valued dominance relations, significantly improving classification accuracy. The AE-RAR [14] method introduces an unsupervised reduction approach for interval-valued systems by establishing an α-approximate equivalence relation grounded in fuzzy similarity. Meanwhile, GLSFS [15] presents a feature selection technique for interval-valued data that leverages graph-driven and granular rectangular neighborhood rough set methodologies. This approach constructs a weighted complete graph and employs matrix power series to rank the features. The SIDDM-FDNRS [16] introduces an interval-valued fuzzy dominant neighborhood rough set with an adaptive neighborhood (IV-FDNRS) and an adaptive interval-valued dominant discernibility matrix (SIDDM). Additionally, FC [17] presents the concept of contradictory state sequences and completes the optimal scale selection and feature selection within the multi-scale interval-valued decision table. Traditional single-value dominance models, even when adapted for interval data, lack systematic frameworks to effectively address overlapping intervals and variable information granularity, resulting in suboptimal feature selection in complex interval-valued systems.

Recent advancements in robust feature selection have underscored the importance of noise-resistant mechanisms for general classification tasks. KNCMI [18] integrates k-nearest neighbor rough sets with information entropy to effectively mitigate noise in imbalanced hybrid data. Similarly, DNCFAR [19] introduces a fuzzy attribute reduction method that leverages deep fusion of neurocognitive co-evolution and quantum jump particle swarm optimization. By identifying and adaptively decomposing interdependent fuzzy attribute variables, and employing the nearest neighbor meme complex to enhance the search process, the complexity and inseparability of the data are significantly reduced. Furthermore, NFRS [20] presents a noise-aware fuzzy rough set model. This model designs a weighted fuzzy relationship based on sample density and the local outlier factor, and integrates an evaluation function that considers intra-class aggregation and inter-class dispersion, achieving feature selection that is robust against both attribute noise and class noise. PMSNE [21] consolidates robust label enhancement, sparse reconstruction denoising, and feature selection into a comprehensive end-to-end framework. However, these methods primarily concentrate on generic classification scenarios, often neglecting the specific challenges posed by monotonic classification.

Feature synergy consideration emphasizes the importance of capturing collaborative interactions among features during the selection process. The RCDFS [22] framework proposes a feature ranking methodology based on complementary-redundant collaborative measurement, enabling the rapid selection of small, robust subsets from high-dimensional, mixed, and imbalanced datasets. Additionally, FS-RSA [23] efficiently identifies high-discriminative and low-redundant feature subsets in ultra-high-dimensional text data through redundant-cooperative subset analysis. Utilizing soft neighborhood rough set and soft interval theory, SNCMI [24] introduces a unified evaluation function that quantifies feature correlation, redundancy, complementarity and synergy. Traditional feature selection methods often assess features independently, neglecting the synergistic effects that emerge when features are analyzed collectively. This limitation is particularly critical in high-dimensional datasets where feature interactions can substantially influence classification performance.

The applications and drawbacks of existing methods are illustrated in Table 1. The current research identifies three major gaps: (1) Limited noise tolerance in monotonic classification due to rigid ordinal assumptions; (2) Inadequate management of data granularity in interval-based ordered systems; (3) Insufficient consideration of feature synergistic interactions. While existing methods, such as IV-DRSA and FC, address individual aspects, their isolated focus overlooks intertwined challenges, including noise, multi-valued data, and feature synergies. This underscores the necessity for an integrated approach that synergistically addresses these issues while maintaining ordinal integrity.

3. Preliminaries

This section mainly introduces some basic knowledge about IV-ODS, relative density, and information metrics in ordered decision systems.

3.1. Interval-valued ordered decision system(IV-ODS) [13]

An IV-ODS is a four-tuple $T = \langle X, AS \bigcup D, V, F \rangle$, where $X = \{x_1, x_2, x_3, \ldots, x_n\}$ represents instances or records in the information system, $AS = \{a_1, a_2, a_3, \ldots, a_n\}$ is a finite and non-empty set of condition attributes, which are the features or variables that describe the samples, D is a nominal decision-making attribute, indicating the categorical target variable to be predicted or decided, and V represents the complete set of the values within the information system. Next, $F = \{F_{a_k} \mid X \to V_{a_k},\ a_k \in \mathrm{AS}\}$, where F_{a_k} is the interval value of a_k on $x \in X$. $Cl_t = \{x_i \in U \mid v(x_i, D) = d_t, t \in \{1, 2, \ldots, T\}\}$. These decision classes follow a preference order based on the decision values:$d_1 < d_2 < \ldots < d_T$, which induces an order among the decision classes:$Cl_1 \prec Cl_2 \prec \ldots \prec Cl_t$. We define upward union:$Cl_t^{\preceq} = \bigcup_{s=t}^{T} Cl_s$.

This approach introduces interval-valued dominance relations in IV-ODS through two essential measures: the Interval Dominance Degree (IDD) and the Interval Overlap Degree (IOD) for comparing intervals.

IDD, denoted as $\alpha(\iota, \kappa)$, quantifies the dominance relation between the two intervals $\iota = [\underline{\iota}, \overline{\iota}]$ and $\kappa = [\underline{\kappa}, \overline{\kappa}]$. The IDD is defined as:

$$\alpha(\iota, \kappa) = \frac{\overline{\kappa} - \overline{\iota}}{\max\{\overline{\iota}, \overline{\kappa}\} - \max\{\underline{\iota}, \underline{\kappa}\}} \tag{1}$$

where the range of α belongs to $[-1, 1]$.

The IDD describes the dominance relation between two intervals. However, we also consider the overlap between intervals, as less overlap often implies a clearer dominance relationship. To address this, we define the IOD, denoted as $\beta(\iota, \kappa)$, which quantifies the level of overlap between two intervals $\iota = [\underline{\iota}, \overline{\iota}]$ and $\kappa = [\underline{\kappa}, \overline{\kappa}]$. Let $l_1 = \overline{\iota} - \underline{\iota}$ and $l_2 = \overline{\kappa} - \underline{\kappa}$ be the lengths of the intervals ι and κ, respectively.

The IOD is defined as:

$$\beta(\iota, \kappa) = \frac{2(\min\{\overline{\iota}, \overline{\kappa}\} - \max\{\underline{\iota}, \underline{\kappa}\})}{l_1 + l_2} \tag{2}$$

where the range of β lies within [0,1].

This definition captures the degree of overlap between the intervals, providing a more nuanced understanding beyond what IDD can offer alone.

For objects $x, y \in X$, along with a subset $A \subseteq AS$, the dominance relation valued by intervals on the set of attributes A is defined as follows:

$$\mathrm{R}_A^{\alpha,\beta} = \{(x, y) \in X^2 \mid \forall a \in A, \alpha(\iota, \kappa) \geq \alpha \wedge \beta(\iota, \kappa) \leq \beta\}$$

If $\alpha(\iota, \kappa) \geq \alpha$ and $\beta(\iota, \kappa) \leq \beta$ hold for all $a \in A$, we say that y dominates x at levels α and β, denoted as $y\mathrm{R}_A^{\alpha,\beta}x$.

The **A-dominating set** of x, denoted as $\mathrm{R}_A^{+,\alpha,\beta}(x)$, is the set of all objects $y \in X$ such that y dominates x on the attribute set A with respect to the thresholds α and β. Mathematically, it is defined as:

$$\mathrm{R}_A^{+,\alpha,\beta}(x) = \{y \in X \mid y\mathrm{R}_A^{\alpha,\beta}x\}$$

The **A-dominated set** of x, denoted as $\mathrm{R}_A^{-,\alpha,\beta}(x)$, is the set of all objects $y \in X$ such that x dominates y on the attribute set A with respect to the thresholds α and β. Mathematically, it is defined as:

$$\mathrm{R}_A^{-,\alpha,\beta}(x) = \{y \in X \mid x\mathrm{R}_A^{\alpha,\beta}y\}$$

These definitions provide a structured way to understand the dominance relationships between objects in an IV-ODS, considering the interval values of attributes and the specified thresholds for dominance.

3.2. Relative density [25]

In this section, we present the foundational concepts of relative density, which serve as a key metric for detecting noise and boundary points in binary classification problems.

Consider a dataset $U \subseteq \mathbb{R}^d$. For an object $p \in U$, let $N_k(p)$ denote the set of its k-nearest neighbors in U (excluding p itself), where the distance between two samples is measured by the distance function *d(p,q)*. The absolute density of p with respect to U is defined as

$$\text{absolute density}(p, U) = \frac{k}{\sum_{q \in N_k(p)} d(p, q)}. \tag{3}$$

This quantity captures the local compactness around p by inversely weighting the distance to its nearest neighbors.

Suppose U is partitioned into two classes: the positive class U^+ and the negative class U^-. For an object $p \in U^+$, the homogeneous k-nearest neighbors, denoted $\mathrm{HON}_k(p)$, comprise the k closest objects to p within U^+. Similarly, the heterogeneous k-nearest neighbors, denoted $\mathrm{HEN}_k(p)$, consist of the k closest objects to p within U^-.

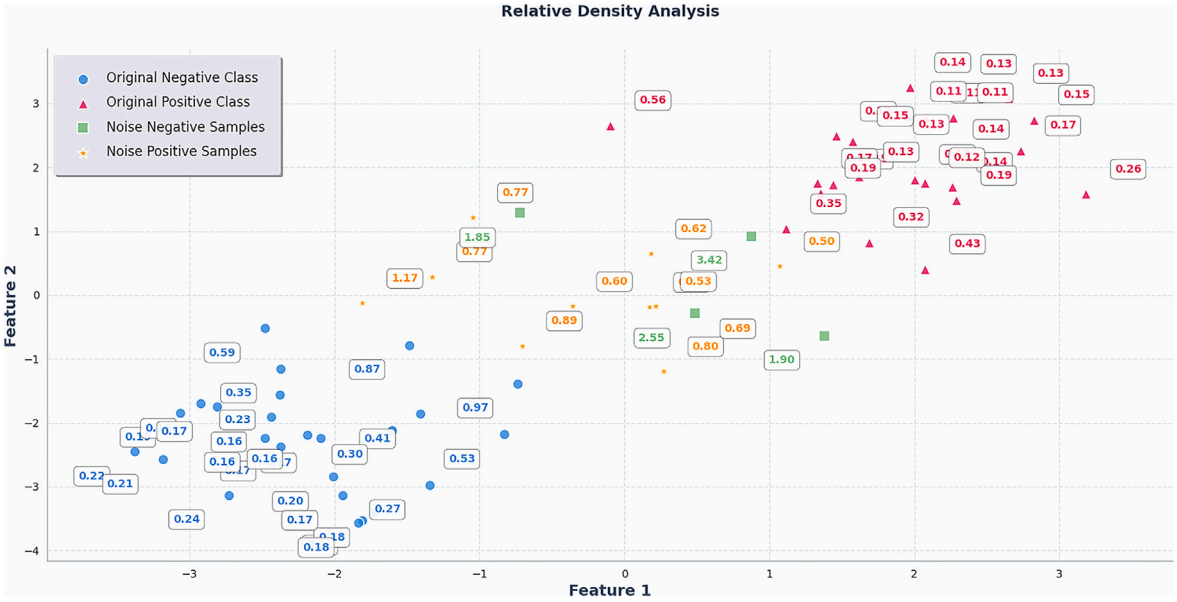

Fig. 2. Scatter plot with relative densities.

For an object $x \in U^+$, the relative density is given by the ratio of its absolute density in the negative class to that in the positive class:

$$\text{relative density}(x) = \frac{\text{absolute density}(x, U^-)}{\text{absolute density}(x, U^+)} = \frac{\sum_{p \in \text{HON}_k(x)} d(x,p)}{\sum_{q \in \text{HEN}_k(x)} d(x,q)}. \tag{4}$$

This metric effectively distinguishes between interior points, boundary points, and noise points, providing insights for our subsequent work on elastic anti-noise strategies. The mechanism of action is illustrated in Fig. 2. The relative densities of the noise samples are higher due to their proximity to the heterogeneous samples.

3.3. Information metrics in ordered decision systems

In ordered decision systems, information metrics derived from information theory are crucial for quantifying uncertainty and interdependencies among features, thereby facilitating effective feature selection. These metrics assess the relevance of attributes to decision-making while accounting for complex interactions, including redundancy and synergy.

A core concept is entropy, which measures the inherent uncertainty within a dataset. In an ordered decision system with universe X and feature subset $A \subseteq F$, the dominance-based entropy is defined as:

$$\text{DH}(A) = -\frac{1}{|X|} \sum_{i=1}^{|X|} \log_2 \frac{|R_A^+(x_i)|}{|X|}. \tag{5}$$

This assesses the average uncertainty based on the dominance classes induced by A, providing insights into how features partition samples in ordered contexts.

Joint entropy extends this to multiple feature subsets, capturing the combined uncertainty: for subsets A and B, it is defined as:

$$\text{DJH}(A, B) = -\frac{1}{|X|} \sum_{i=1}^{|X|} \log_2 \frac{|\text{R}_{A \cup B}^+(x_i)|}{|X|}. \tag{6}$$

Conditional entropy, defined as the difference between joint and marginal entropies, quantifies the residual uncertainty in one subset given observations from another:

$$\text{DCH}(B|A) = -\frac{1}{|X|} \sum_{i=1}^{|X|} \log_2 \frac{\left|\text{R}_{A \cup B}^+(x_i)\right|}{|\text{R}_A^+(x_i)|}. \tag{7}$$

Mutual information further quantifies shared knowledge between subsets, representing the entropy reduction in one due to information from the other:

$$\text{DMI}(A; B) = -\frac{1}{|X|} \sum_{i=1}^{|X|} \log_2 \frac{\left|\text{R}_A^+(x_i)\right| \cdot \left|\text{R}_B^+(x_i)\right|}{\left|\text{R}_{A \cup B}^+(x_i)\right| \cdot |X|}. \tag{8}$$

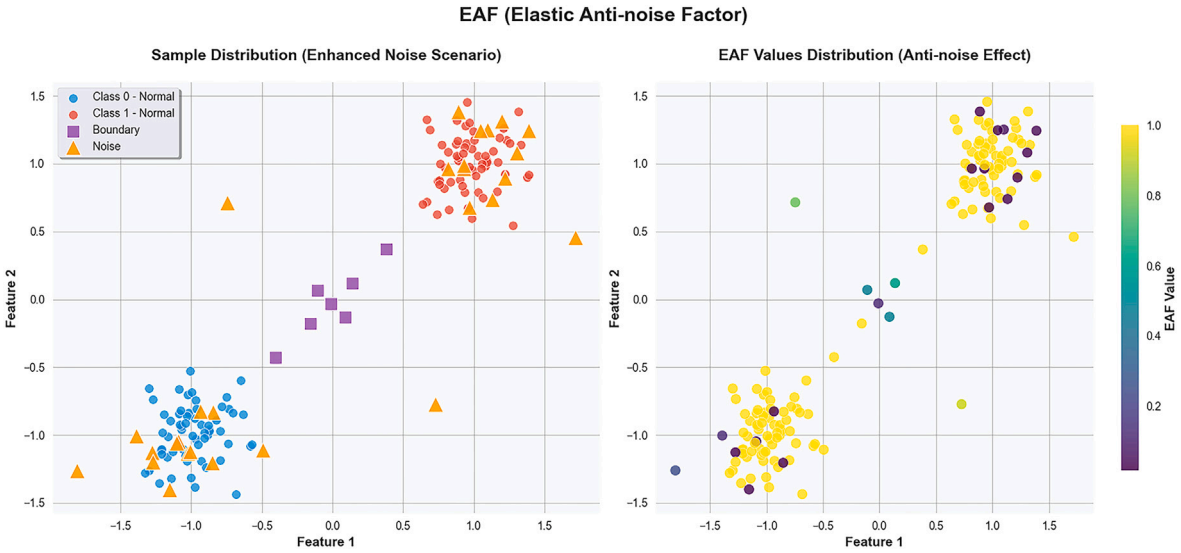

Fig. 3. Schematic diagram of the EAF mechanism.

This metric is pivotal in feature selection, as it gauges the discriminative contribution of a candidate feature relative to the decision attribute.

In summary, these dominance-based information metrics provide a robust theoretical framework for evaluating uncertainty and interdependencies within ordered decision systems. This framework underpins advanced feature selection methods that improve both the accuracy and efficiency of classification tasks by effectively capturing the relevance and relationships of attributes.

4. IV-RANFSI model

The sample distribution in the boundary region, the definition of the dominance relation, and the threshold settings significantly influence the sensitivity of rough sets based on the dominance relation. Drawing inspiration from the concept of relative density discussed in the previous section, we propose an anti-noise method that utilizes the EAF.

4.1. Elastic anti-noise method based on EAF

The sample distribution in the boundary region, the definition of the dominance relation, and the threshold settings significantly influence the sensitivity of the rough set based on the dominance relation. Inspired by the concept of relative density discussed in the previous section, we propose an anti-noise method that leverages the EAF.

Definition 1. Given an ODS $S^{\preceq} = (X, AS \cup \{D\}, V, F)$, for any subset $A \subseteq AS$, let $\mathcal{K}_A^+(x_i)$ denote the first K same-class nearest neighbors of x_i in subspace A, and $\mathcal{K}_A^-(x_i)$ the first K different-class nearest neighbors. Let

$$d_A^+(x_i) = \sum_{z \in \mathcal{K}_A^+(x_i)} D(x_i, z), \qquad d_A^-(x_i) = \sum_{z \in \mathcal{K}_A^-(x_i)} D(x_i, z), \tag{9}$$

where $D(\cdot,\cdot)$ denotes the Euclidean distance restricted to the subspace A. Denote by σ_A the standard deviation of the feature values across all samples in the subspace A.

Then the EAF of sample x_i with respect to A is defined as

$$EAF_A(x_i) = \begin{cases} 1, & \text{if } \dfrac{d_A^-(x_i)}{d_A^+(x_i)} > \sigma_A, \\ \dfrac{d_A^-(x_i)}{d_A^+(x_i)}, & \text{otherwise.} \end{cases} \tag{10}$$

As illustrated in Fig. 3, when the sample exhibits noise, its distance from samples of the same class exceeds its distance from samples of different classes. Consequently, the EAF value should be lower. This mechanism facilitates the automatic elimination of samples; however, it may lead to excessive suppression, causing EAF values for both boundary samples and noise samples to fall below 1. Samples situated in the boundary region typically demonstrate higher uncertainty. Although these samples are generally not noise, they are often mistakenly classified as outliers and discarded. Such excessive removal of noise can significantly impair the model's performance. To address this issue, we propose an elastic mechanism that compares the EAF value with the standard deviation of the feature subspace, enabling adaptive adjustment of the rejection strength for noise samples.

The EAF value of noise samples is typically less than 1, while the EAF value for boundary samples exhibits considerable fluctuation, often remaining around 1. For a single feature, the standard deviation serves as an indicator of the dispersion of samples within the feature space. A larger standard deviation signifies a higher degree of dispersion and a more scattered sample distribution. In this context, the boundary region of boundary samples tends to be sparse, leading to a higher EAF value for these samples. Consequently, this elastic mechanism enhances the capacity to differentiate noise samples from boundary samples across various attribute spaces adaptively.

4.2. Elastic anti-noise rough set based on interval-valued dominance relation

In the context of IV-ODS, it is essential to establish a distance measure between interval-valued data prior to introducing an Elastic Rough Set based on the EAF. It is observed that for two interval-valued data points, greater dominance corresponds to smaller overlap; specifically, this relationship can be articulated as follows: the greater the IDD, the smaller the IOD. Consequently, we can infer that a greater distance exists between the two interval-valued data points.

Definition 2 (**Interval-based distance**). Let $S^{\preceq} = (X, AS \cup \{D\}, V, F)$ be an IV-ODS and let $A \subseteq AS$. For any two objects $x, y \in X$, write $\mathrm{IDD}_A(x, y) = \alpha$ for the Interval Dominance Degree and $\mathrm{IOD}_A(x, y) = \beta$ for the Interval Overlap Degree on A. Empirically, larger dominance (larger α) together with smaller overlap (smaller β) indicates greater separation, so we first set

$$I_A(x, y) = \alpha\,(1 - \beta), \tag{11}$$

where the range of $I_A(x, y)$ belongs to $[-1, 1]$. When $I_A(x, y)$ is greater than 0, it indicates dominance between samples, and vice versa. And through the properties of IDD and IOD, we can easily get $I_A(x, y) = -I_A(y, x)$.

To secure the non-negativity and symmetry required of any metric, we then define

$$D_A(x, y) = \left|I_A(x, y)\right| = \left|\alpha\,(1 - \beta)\right|. \tag{12}$$

To ensure the antisymmetry of the dominance relation, especially when two distinct objects x and y have identical intervals (i.e., $D_A(x, y) = 0$), a hierarchical tie-breaking rule is implemented. First, priority is given to the decision values: the object with a larger decision value is considered superior. If their decision values are also identical, the objects are further distinguished by their indices: the object with a smaller index is considered to be dominated by the one with a larger index (i.e., x is dominated by y if $Index(x) < Index(y)$). This mechanism ensures a unique and consistent dominance order between any two objects. In this paper, we only discuss the dominance relation, so we do not discuss the case where $I_A(x, y)$ is negative, because it belongs to the opposite side of the dominance relation.

Theorem 1 (**Properties of** D_A). *For any objects $x, y, z \in X$, the following properties hold:*

1. ***Non-negativity:*** $D_A(x, y) \geq 0$.
2. ***Symmetry:*** $D_A(x, y) = D_A(y, x)$.
3. ***Asymmetry of dominance:*** *If $x \preceq_A y$ and $y \npreceq_A x$, then $D_A(x, y) \neq D_A(y, x)$ in the dominance sense.*
4. ***Triangle inequality:*** *In general, $D_A(x, z) \nleq D_A(x, y) + D_A(y, z)$.*
5. ***Transitivity of dominance:*** $x \preceq_A y \wedge y \preceq_A z \rightarrow x \preceq_A z$.

For objects $x, y \in X$ and a subset A of conditional attributes AS, the interval-valued elastic anti-noise dominance relation on attribute set A is defined as:

$$\mathrm{R}_A^{\delta_i} = \{(x, y) \in X^2 \mid \forall a \in A, D_a(x, y) \geq EAF_a(x) * \delta_i\}$$

where δ_i refers to the dominance threshold, which is used to control the dominance relationship. Its value ranges from $[0, 1]$. When $\delta_i = 0$, because $D_a(x, y)$ is nonnegative, it shows that y can completely control x under any conditions. On the contrary, when $\delta_i = 1$, the establishment of the dominance relation is most stringent. As δ_i varies from 0 to 1, it can finely control the dominance relationship between interval values.

The **A-dominating set** of x, denoted as $\mathrm{R}_A^+(x)$, is the set of all objects $y \in X$ such that y dominates x on the attribute set A. It is defined as:

$$\mathrm{R}_A^{+,\delta_i}(x) = \{y \in X \mid x \mathrm{R}_A^{\delta_i} y\}$$

The **A-dominated set** of x, denoted as $\mathrm{R}_A^-(x)$, is the set of all objects $y \in X$ such that x dominates y on the attribute set A. Mathematically, it is defined as:

$$\mathrm{R}_A^{-,\delta_i}(x) = \{y \in X \mid y \mathrm{R}_A^{\delta_i} x\}$$

Definition 3. The elastic rough set based on interval-valued dominance relation upper and lower approximations is defined as

$$\underline{R_A^{\delta_i}}(Cl_t^{\preceq}) = \{x \in X \mid R_A^{+,\delta_i}(x) \subseteq Cl_t^{\preceq}\} \tag{13}$$

$$\overline{R_A^{\delta_i}}(Cl_t^{\preceq}) = \{x \in X \mid R_A^{-,\delta_i}(x) \cap Cl_t^{\preceq} \neq \emptyset\} \tag{14}$$

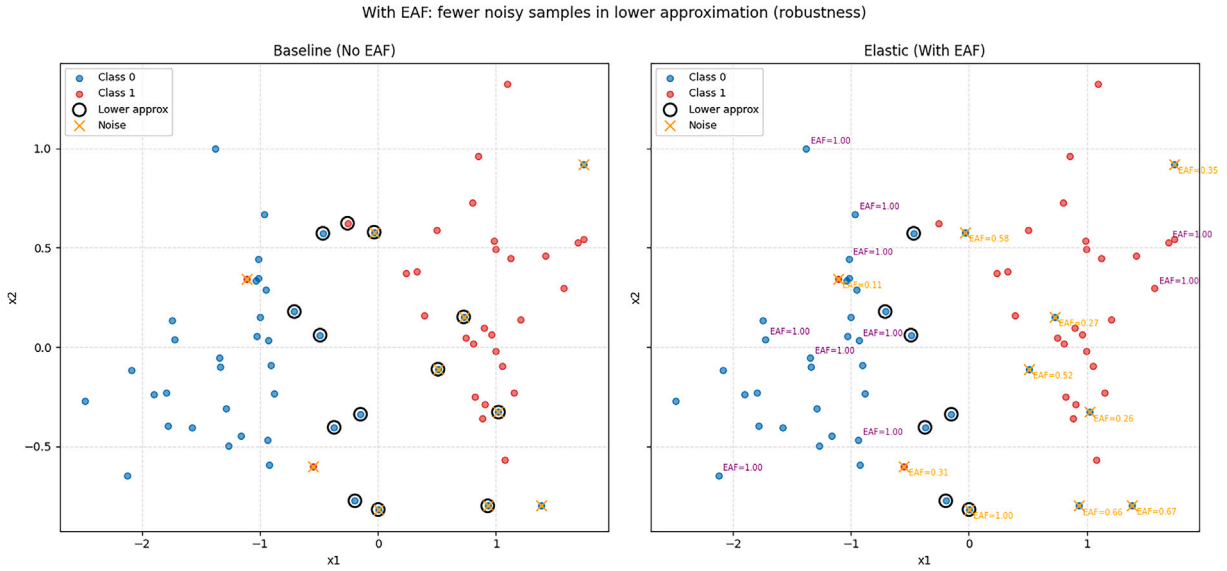

Fig. 4. Rough set with EAF.

As illustrated in Fig. 4, since noisy samples are typically closer to heterogeneous-class samples, their $\mathrm{EAF}_a(x)$ take smaller values. Consequently, the threshold term $\mathrm{EAF}_a(x) \cdot \delta_i$ decreases, making the condition $D_a(x, y) \geq \mathrm{EAF}_a(x) \cdot \delta_i$ in the relation $\mathrm{R}_A^{\delta_i}$ easier to satisfy. As a result, the **A-dominating set** $R_A^{+,\delta_i}(x)$ of a noisy sample expands significantly.

For the lower approximation $\underline{R_A^{\delta_i}}(Cl_t^{\preceq})$, the membership criterion requires $R_A^{+,\delta_i}(x) \subseteq Cl_t^{\preceq}$. When $R_A^{+,\delta_i}(x)$ inflates due to noise, this condition becomes harder to fulfill (i.e., noisy samples are more likely to be excluded because they now encompass heterogeneous objects). Because the decision class $Cl_t^{\preceq}$ remains unchanged, the number of misclassified noisy samples in the lower approximation is reduced, thereby enhancing the robustness of the model against noise.

Definition 4 (**Approximation Quality**). Given an elastic interval-valued dominance relation R_A and an upward union of decision classes $Cl_t^{\preceq}$, the *approximation quality* of the attribute subset $A \subseteq AS$ is defined as

$$\gamma_A(Cl_t^{\preceq}) = \frac{|\underline{R_A^{\delta_i}}(Cl_t^{\preceq})|}{|X|} \tag{15}$$

where $\underline{R_A^{\delta_i}}(Cl_t^{\preceq})$ is the lower approximation of $Cl_t^{\preceq}$ with respect to A as given in Definition 3, and X is the entire universe of objects.

Definition 5 (**Global Attribute Importance**). Let AS be the complete set of condition attributes and $\mathcal{T} = \{1, \ldots, m\}$ the indices of the decision classes. For any single attribute $a \in AS$, its global importance is defined as the average approximation quality obtained over all upward unions of decision classes.

$$\mathrm{Imp}(a) = \frac{1}{m}\sum_{t=1}^{m}\gamma_a(Cl_t^{\preceq}) = \frac{1}{m}\sum_{t=1}^{m}\frac{\left|\underline{R_a^{\delta_i}}(Cl_t^{\preceq})\right|}{|X|} \tag{16}$$

where $\underline{R_a^{\delta_i}}(Cl_t^{\preceq})$ is the lower approximation of $Cl_t^{\preceq}$ with respect to the singleton attribute set $\{a\}$, and X denotes the entire universe of objects.

4.3. Feature evaluation metrics

Feature interactions are characterized by redundancy and synergy. Redundancy arises when a candidate feature a_c (from $AS-\mathrm{Red}$, where Red is the selected subset) provides overlapping classification information with existing features in Red, measured by their DMI:

$$\mathrm{Rdd}(a_c; a_s) = \mathrm{DMI}(a_c; a_s), \tag{17}$$

which indicates duplicative content that may diminish overall utility.

In contrast, synergy denotes the cooperative enhancement where features jointly offer more informative value than individually. Specifically, for a candidate a_c and selected feature $a_s \in \mathrm{Red}$ with respect to decision d, synergy is formalized as the DCMI:

$$\mathrm{Syn}(a_c, a_s) = \mathrm{DCMI}(a_c; a_s \mid d), \tag{18}$$

where the DCMI was previously defined as:

$$\mathrm{DCMI}(A;B|e) = -\frac{1}{|X|}\sum_{i=1}^{|X|}\log_2\frac{|\mathrm{R}^{+,\delta_i}_{A\cup e}(x_i)|\cdot|\mathrm{R}^{+,\delta_i}_{B\cup e}(x_i)|}{|\mathrm{R}^{+,\delta_i}_{e}(x_i)|\cdot|\mathrm{R}^{+,\delta_i}_{A\cup B\cup e}(x_i)|}. \tag{19}$$

This quantifies emergent discriminative information from their interaction that neither feature supplies alone.

Definition 6. An objective evaluation function considering complementarity and synergism is defined as:

$$J(a_c) = \mathrm{Rdd}(a_c;\, d) - \frac{1}{|\mathrm{Red}|}\sum_{s=1}^{|\mathrm{Red}|}\mathrm{Rdd}(a_c;\, a_s) + \frac{1}{|\mathrm{Red}|}\sum_{s=1}^{|\mathrm{Red}|}\Big(\mathrm{DCMI}(a_c;\, d \mid a_s) + \mathrm{Syn}(a_c, a_s)\Big) \tag{20}$$

First, $\mathrm{Rdd}(a_c; d)$ captures the intrinsic relevance of a_c to the decision attribute d through their direct DMI. Second, $\mathrm{Rdd}(a_c; a_s)$ penalises a_c for any redundancy it introduces with respect to the already selected features $a_s \in \mathrm{Red}$. Third, $\mathrm{DCMI}(a_c; d \mid a_s)$ evaluates the residual information that a_c conveys about d once the knowledge encoded in each a_s has been accounted for, rewarding attributes that maintain predictive value even in the presence of the current subset. Finally, the synergy term $\mathrm{Syn}(a_c, a_s)$ measures the emergent, cooperative information arising from the interaction between a_c and a_s with respect to d.

Definition 7 (Comprehensive Attribute Score). Let $a_c \in AS$ be a candidate condition attribute. Its **comprehensive score** is defined as the sum of the global attribute importance and the objective evaluation function $J(a_c)$:

$$\mathrm{Score}(a_c) = J(a_c) + \mathrm{Imp}(a_c \cup Red) - \mathrm{Imp}(Red) \tag{21}$$

where Red refers to the selected subset of conditional attributes.

4.4. Our IV-RANFSI algorithm and an example

Our complete feature selection process is shown in Algorithm 1. Algorithm 1 presents the pseudo-code for our proposed IV-RANFSI framework. The procedure is structured into four primary phases. First, it initializes by calculating the EAF and constructing an interval-valued dominance relation matrix for each attribute under multiple thresholds. Second, it evaluates $\mathrm{Imp}(a)$ based on $\gamma_A(Cl_t^{\preceq})$ to select the initially most significant feature. Third, a greedy search strategy is employed to iteratively select the remaining features by maximizing $\mathrm{Score}(a_c)$ which balances feature redundancy and synergy (DCMI). Finally, a performance-driven wrapper step uses cross-validation with multiple classifiers (SVM, RF, XGBoost) to determine the optimal feature subset from the ranked sequence.

Next, we provide the computational complexity analysis of our proposed IV-RANFSI algorithm. Let n be the number of samples, m be the number of features, p be the number of dominance thresholds, T be the number of decision classes, and l be the number of selected features ($l \le m$). In terms of time complexity, the algorithm consists of several major steps. First, calculating the EAF values and constructing the interval-based distance matrix for m features requires $O(mn^2)$ time. Second, the multi-threshold processing (p thresholds) involves computing R-matrices and approximation qualities for T decision classes, costing $O(p(mn^2 + Tn^2))$. Third, the feature selection loop iterates at most m times, with each iteration evaluating redundancy, synergy, and DCMI, requiring $O(m^2)$ time in the worst case. Finally, the cross-validation evaluation of l feature subsets using 5-fold cross-validation costs $O(l \cdot T_{cv})$, where T_{cv} denotes the time complexity of the classification algorithm. Therefore, the overall time complexity is $O(pmn^2 + pTn^2) + O(l \cdot T_{cv})$. Since $T \ll n$ and $l \ll m$ in practice, the dominant term is $O(pmn^2)$.

For space complexity, the space complexity is mainly determined by storing the original interval-valued decision system ($O(mn)$), the distance matrix for each feature ($O(mn^2)$), and the R-matrices for different thresholds ($O(pmn^2)$). Thus, the overall space complexity is $O(pmn^2)$.

To demonstrate the complete workflow of our proposed IV-RANFSI framework, we present a comprehensive healthcare case study on heart disease risk prediction. The dataset comprises six patient records characterized by three clinical features: a_1 (Blood Pressure), a_2 (Cholesterol), and a_3 (Age), with binary decision classes $D \in \{0, 1\}$ representing Healthy and Unhealthy (high risk), respectively. Patient x_6 serves as a noise sample with high-risk clinical indicators but is incorrectly labeled as Healthy ($D = 0$), simulating real-world diagnostic errors.

Following min-max normalization, all attributes are transformed to the $[0, 1]$ scale, yielding the normalized dataset in Table 2. The standard deviations are $\sigma_{a_1} = 0.3627$, $\sigma_{a_2} = 0.3558$, and $\sigma_{a_3} = 0.3675$. Using the two-sigma rule, we construct interval-valued representations to capture measurement uncertainty, such as $I_{a_1}(x_1) = [-0.725, 0.725]$.

The EAF is computed with $K = 2$ nearest neighbors. Table 3 shows the complete EAF values for all samples. Normal samples (x_1-x_5) exhibit EAF values around 1.0. Noise sample x_6 shows significantly lower values: $\mathrm{EAF}_{a_1}(x_6) = 0.0552$, $\mathrm{EAF}_{a_2}(x_6) = 0.0495$, and $\mathrm{EAF}_{a_3}(x_6) = 0.1220$. The complete interval-valued decision system is shown in Table 4.

Using the dominance threshold $\delta_i = 0.2$, we construct the dominance relation matrix $R^{\delta_i}_{a_1}$ for blood pressure (Table 5), where entry $(i, j) = 1$ indicates that sample x_j dominates x_i. The global attribute importance scores are $\mathrm{Imp}(a_1) = \mathrm{Imp}(a_3) = 0.6667$ and $\mathrm{Imp}(a_2) = 0.5833$.

Algorithm 1: IV-RANFSI: interval-valued robust anti-noise feature selection with interactions.

Input: Interval-valued Ordered Decision System $S^{\preceq} = \langle X, AS \cup \{D\}, V, F \rangle$; Relative density parameter K; Dominance threshold set $\delta_{set} = \{\delta_1, \ldots, \delta_p\}$

Output: Optimal conditional feature subset P_{opt}

; // Stage 1: Multi-threshold Anti-noise Dominance Relation Construction

Initialize selected feature subset $P_{opt} \leftarrow \emptyset$;

for *each threshold* $\delta_i \in \delta_{set}$ **do**

  for *each candidate feature* $a_k \in AS$ **do**

    Calculate Elastic Anti-noise Factor $EAF_{a_k}(x_i)$ for each sample $x_i \in X$ using Definition 1;

    // Identifies noise samples based on local density

    Compute pairwise interval distance matrix D_{a_k} using Definition 2;

    Construct elastic anti-noise interval dominance relation matrix $R_{a_k}^{\delta_i}$;

  end

end

; // Stage 2: Global Attribute Importance Evaluation

for *each threshold* $\delta_i \in \delta_{set}$ **do**

  for *each feature* $a_k \in AS$ **do**

    Generate dominance matrix $R_{a_k}^{\delta_i}$ based on current threshold δ_i;

    for *each decision class* $t \in \{1, 2, \ldots, T\}$ **do**

      Compute upward union of decision classes $Cl_t^{\preceq}$;

      Calculate Lower Approximation $\underline{R_{a_k}^{\delta_i}}(Cl_t^{\preceq})$ and Upper Approximation $\overline{R_{a_k}^{\delta_i}}(Cl_t^{\preceq})$;

      // Rough set approximations for uncertainty handling

      Compute Approximation Quality $\gamma_{a_k}^{\delta_i}(Cl_t^{\preceq})$ using Definition 4;

    end

    Calculate Global Attribute Importance $Imp(a_k)$ using Definition 5;

  end

end

Select feature $a_{best} = \arg\max_{a_k \in AS}\{Imp(a_k)\}$; $P_{opt} \leftarrow P_{opt} \cup \{a_{best}\}$, $AS \leftarrow AS \setminus \{a_{best}\}$;

; // Stage 3: Interaction-based Greedy Feature Selection

while $AS \neq \emptyset$ **do**

  for *each candidate feature* $a_c \in AS$ **do**

    for *each selected feature* $a_s \in P_{opt}$ **do**

      Calculate Redundancy $Rdd(a_c, a_s)$ using Eq. (17);

      Calculate Synergy $Syn(a_c, a_s)$ using Eq. (18);

      Compute Conditional Mutual Information $DCMI(a_c; a_s \mid D)$;

    end

    Calculate Objective Evaluation Function $J(a_c)$ (Definition 6);

    // Integrates redundancy and synergy terms

    Compute Final Comprehensive Score $Score(a_c)$ (Definition 7);

  end

  Select feature $a'_c = \arg\max_{a_c \in AS}\{Score(a_c)\}$; $P_{opt} \leftarrow P_{opt} \cup \{a'_c\}$, $AS \leftarrow AS \setminus \{a'_c\}$;

end

; // Stage 4: Optimal Subset Determination via Cross-Validation

Obtain ranked feature sequence $P_{opt} = \{a'_1, a'_2, \ldots, a'_l\}$;

for $q = 1$ **to** l **do**

  Let $P_q = \{a'_1, \ldots, a'_q\}$; Use SVM, RF, and XGBoost with 5-fold cross-validation to compute $Acc(P_q)$;

end

$P_{opt} = \arg\max_{P_q \subseteq P_{opt}}\{Acc(P_q), q \in \{1, \ldots, l\}\}$;

return P_{opt}

The feature selection phase employs greedy forward selection optimizing the comprehensive score (Definition 7), balancing relevance, redundancy, and synergy. In iteration 1 with Red = $\{a_1\}$, we evaluate a_2 (Score = −0.1972) and a_3 (Score = 0.0253), selecting a_3. In iteration 2, evaluating a_2 against $\{a_1, a_3\}$ yields Score = −0.1268, leading to the final ranking $\{a_1, a_3, a_2\}$.

Table 2
Normalized healthcare dataset.

Sample	a_1 (BP)	a_2 (CHO)	a_3 (Age)	Decision
x_1	0.0000	0.0000	0.0000	0
x_2	0.0909	0.1429	0.0909	0
x_3	0.4545	0.4286	0.4545	1
x_4	0.6364	0.7143	0.5455	1
x_5	1.0000	1.0000	0.8727	1
x_6	**0.8182**	**0.7857**	**1.0000**	**0**

Table 3
Elastic anti-noise factor (EAF) values ($K = 2$).

Sample	EAF_{a_1} (BP)	EAF_{a_2} (CHO)	EAF_{a_3} (Age)
x_1	1.0000	1.0000	0.5000
x_2	0.8000	1.0000	0.4059
x_3	0.8000	0.5125	1.0000
x_4	1.0000	1.0000	1.0000
x_5	1.0000	1.0000	1.0000
x_6	**0.0552**	**0.0495**	**0.1220**

Table 4
Final interval-valued decision system (IV-ODS).

Sample	a_1 (BP)	a_2 (CHO)	a_3 (Age)	Decision
x_1	[−0.725, 0.725]	[−0.712, 0.712]	[−0.735, 0.735]	0
x_2	[−0.634, 0.816]	[−0.569, 0.854]	[−0.644, 0.826]	0
x_3	[−0.271, 1.180]	[−0.283, 1.140]	[−0.280, 1.189]	1
x_4	[−0.089, 1.362]	[0.003, 1.426]	[−0.189, 1.280]	1
x_5	[0.275, 1.725]	[0.288, 1.712]	[0.138, 1.608]	1
x_6	**[0.093, 1.544]**	**[0.074, 1.497]**	**[0.265, 1.735]**	**0**

Table 5
Dominance matrix for blood pressure (a_1, $\delta_i = 0.2$).

	x_1	x_2	x_3	x_4	x_5	x_6
x_1	0	0	0	0	1	1
x_2	0	0	0	0	1	1
x_3	0	0	0	0	0	0
x_4	0	0	0	0	0	0
x_5	1	1	0	0	0	0
x_6	1	1	1	1	1	0

5. Experiments and analysis

In this section, the performance evaluation framework of the IV-RANFSI algorithm is established, which ensures strict verification and practical application.

5.1. Compared algorithms

(1) ***Original Attributes (Original Data):*** Direct classification using all conditional attributes without dimensionality reduction.
(2) ***Interval-Valued Dominance-based Rough Set (IV-DRSA)*** [13]: This method extracts features in interval-valued ordered systems using interval-based dominance rough approximations which measure overlap and dominance relationships.
(3) ***Incremental feature selection: Parallel approach with local neighborhood rough sets and composite entropy(PMLCE)*** [26]: A parallel incremental feature selection algorithm combining local neighborhood rough sets and composite entropy is proposed, which can effectively deal with the challenge of the simultaneous growth of samples and features in large-scale dynamic data.
(4) ***Unsupervised Attribute Reduction Based on α-approximate Equal Relation in Interval-valued Information Systems (AE-RAR)*** [14]: AE-RAR proposes an unsupervised reduction method for interval-valued systems by defining an α-approximate equal relation based on fuzzy similarity and utilizing information entropy for attribute evaluation.
(5) ***Feature selection considering synergy between features based on soft neighborhood rough sets (SNCMI)*** [24]: The SNCMI algorithm improves feature selection for noisy datasets in data mining and pattern recognition, enhancing classification accuracy through optimized feature synergy.
(6) ***An optimal feature selection method for text classification through redundancy and synergy analysis(FS-RSA)*** [23]: An interaction-information-driven feature selector that retains high-unique or highly synergistic terms to curb redundancy without accuracy loss, excelling in content-heavy Chinese-text classification tasks.

Table 6
Data description.

No.	Data sets	samples	features	classes
1	Period Changer	90	1177	2
2	Dermatology	358	34	6
3	Parkinson's Disease Classification	755	754	2
4	German credit	1000	20	2
5	Wireless Indoor Localization	2000	7	4
6	Iranian churn	3334	13	2
7	Abalone	4177	8	3
8	Optical Recognition of Handwritten Digits	5620	64	10
9	Shill Bidding	6321	10	2
10	Electrical Grid Stability	10,000	13	2
11	Online Shoppers Purchasing	12,330	17	2
12	Magic gamma	19,019	10	2

(7) ***Feature Selection for Unbalanced Distribution Hybrid Data (KNCMI)*** [18]: This method combines KNN and δ-neighborhood rough sets with information entropy to evaluate feature importance in imbalanced hybrid data, effectively addressing data distribution challenges.
(8) ***Hyperspectral Band Selection Based On Rough Set(HKCMI)*** [27]: This rough set-based hyperspectral band selection method optimizes computational efficiency while selecting informative bands through relevance and significance analysis.

To ensure a comprehensive and fair evaluation, we have carefully selected seven advanced feature selection algorithms based on the following criteria: Firstly, we choose two algorithms IV-DRSA and AE-RAR applied to interval values to compare the feature selection effect of our algorithm in dealing with interval value scenarios. Then, both SNCMI and FS-RSA are feature selection algorithms that consider the relationship between features to verify the effectiveness of our algorithm in considering feature synergy. Finally, PMLCE, KNCMI and HKCMI are all algorithms based on neighborhood rough set, which are used to verify the adaptive anti-noise advantage of our algorithm. Therefore, we use these seven comparison algorithms to verify the three important innovations proposed in our article. This diverse selection enables a thorough evaluation across multiple data characteristics and methodological paradigms.

5.2. Experimental data preprocessing

The methodological framework begins with twelve datasets listed in Table 6, sourced from the UCI database. Initially, we employ the min-max normalization technique to standardize the conditional attribute values within the datasets. Subsequently, all datasets are considered as clean data (D_{clean}) with no label noise. To simulate varying degrees of label noise, we generate five versions of the dataset by introducing adversarial label perturbations at noise rates of 0, 0.1, 0.2, 0.3, and 0.4. For each noise level $\eta \in \{0, 0.1, 0.2, 0.3, 0.4\}$, a proportion η of the samples is randomly selected, and their class labels are exchanged with those from different classes ($y_i \leftrightarrow y_j,\ y_i \neq y_j$). To enhance uncertainty quantification, interval-valued representations are constructed for each class-attribute pair using the two-sigma principle:

$$I_c = [x_c - 2\sigma_c, x_c + 2\sigma_c] \tag{22}$$

where x_c and σ_c denote the normalized conditional attribute value and standard deviation, respectively.

5.3. Classification performance evaluations of algorithm

This section validates the efficacy of the proposed IV-RANFSI framework.

To rigorously substantiate the efficacy of the proposed IV-RANFSI framework, we conducted a systematic comparative study against seven state-of-the-art feature-selection algorithms. Classification performance was quantified via three well-established machine-learning models—support-vector machine (SVM), random forest (RF), and extreme gradient boosting (XGBoost). A five-fold cross-validation protocol was consistently applied. As shown in Tables 7–9, the aggregated results under 40% label noise are reported in the accompanying table; each cell presents the mean classification accuracy (left of "±") together with its standard deviation (right of "±"), and the highest accuracy per configuration is highlighted in bold.

Analysis of the experimental results in Tables 7–9 reveals several key insights regarding the efficacy of the proposed IV-RANFSI framework. Most notably, IV-RANFSI demonstrates remarkable robustness to high-intensity noise; under a 40% label noise level, it achieves the highest average accuracy for both SVM (0.639 ± 0.016) and RF (0.634 ± 0.023) classifiers. This superior performance relative to the "Original" baseline suggests that the proposed feature selection mechanism effectively identifies and filters out noise-sensitive attributes that would otherwise compromise the integrity of decision boundaries. Furthermore, the statistical significance of these gains is rigorously confirmed by the Wilcoxon signed-rank test, which shows that IV-RANFSI maintains a significant lead ($p < 0.05$) over all seven comparative algorithms in the RF evaluation. Although AE-RAR exhibits a marginal lead in mean accuracy for the XGBoost classifier, the associated p-value of 0.7473 indicates no statistically significant difference, thereby establishing IV-RANFSI as a consistently top-tier performer across diverse learning paradigms. Such resilience is particularly evident in challenging datasets like *recognition* and *dermatology*, where traditional methods such as PMLCE and SNCMI suffer from severe performance degradation. The stability of IV-RANFSI in these scenarios is largely attributed to the interval-valued reduction strategy, which provides a more robust

Table 7
Classification performance comparison with SVM classifier (40% noise).

Dataset	Original	IV-RANFSI	IV-DRSA	AE-RAR	PMLCE	SNCMI	KNCMI	HKCMI	FS-RSA
abalone	0.438±0.008	**0.442±0.006**	0.438±0.007	0.438±0.008	0.434±0.012	0.441±0.007	0.429±0.008	0.438±0.011	0.438±0.008
churn	0.713±0.006	**0.714±0.008**	0.713±0.006	0.712±0.005	0.712±0.006	0.711±0.010	**0.714±0.008**	0.707±0.002	0.713±0.006
dermatology	0.656±0.064	**0.657±0.064**	0.463±0.079	0.636±0.086	0.469±0.061	0.656±0.064	0.447±0.045	0.606±0.069	0.597±0.060
german	0.642±0.014	0.656±0.017	0.653±0.022	0.637±0.016	0.652±0.002	0.651±0.006	**0.661±0.018**	0.643±0.023	0.637±0.018
shoppers	0.705±0.004	0.706±0.004	0.705±0.000	0.705±0.005	0.704±0.000	0.704±0.000	**0.707±0.006**	0.704±0.000	0.705±0.004
magic	0.657±0.008	0.657±0.007	0.653±0.006	0.655±0.007	**0.658±0.006**	0.647±0.000	0.655±0.007	0.649±0.005	0.657±0.008
parkinson	0.551±0.029	**0.574±0.022**	0.519±0.031	0.552±0.017	0.517±0.025	0.568±0.036	0.539±0.025	0.560±0.026	0.531±0.023
shill	0.732±0.004	**0.735±0.005**	0.723±0.004	0.733±0.006	0.732±0.006	0.733±0.006	0.732±0.006	0.720±0.004	0.732±0.005
period	**0.644±0.027**	**0.644±0.027**	**0.644±0.027**	0.600±0.042	0.578±0.067	0.611±0.050	0.622±0.022	0.611±0.061	0.600±0.082
wireless	0.614±0.016	0.614±0.016	0.603±0.014	0.614±0.016	0.511±0.027	0.609±0.021	**0.619±0.018**	0.573±0.011	0.614±0.016
recognition	0.600±0.011	**0.602±0.012**	0.464±0.009	0.594±0.011	0.311±0.011	0.256±0.009	0.383±0.009	0.592±0.009	0.557±0.009
Electrical	0.670±0.007	0.670±0.007	0.634±0.003	0.630±0.000	**0.687±0.009**	0.640±0.006	0.684±0.007	0.681±0.007	0.673±0.007
Average	0.635±0.017	**0.639±0.016**	0.601±0.017	0.626±0.018	0.580±0.019	0.602±0.018	0.599±0.015	0.624±0.019	0.621±0.021

Note: p-values of Wilcoxon test (IV-RANFSI vs. others): Original (0.0039), IV-DRSA (0.0005), AE-RAR (0.0005), PMLCE (0.0056), SNCMI (0.0002), KNCMI (0.0840†), HKCMI (0.0034), FS-RSA (0.0059). † indicates no significant difference ($p > 0.05$).

Table 8
Classification performance comparison with RF classifier (40% noise).

Dataset	Original	IV-RANFSI	IV-DRSA	AE-RAR	PMLCE	SNCMI	KNCMI	HKCMI	FS-RSA
abalone	0.404±0.011	**0.415±0.016**	0.410±0.008	0.412±0.012	0.387±0.028	0.401±0.011	0.397±0.014	0.397±0.018	0.411±0.012
churn	0.687±0.014	0.709±0.007	0.682±0.011	0.688±0.010	0.690±0.008	**0.711±0.010**	0.645±0.010	0.676±0.009	0.689±0.011
dermatology	0.648±0.071	**0.650±0.068**	0.399±0.034	0.589±0.072	0.463±0.059	**0.650±0.074**	0.447±0.051	0.575±0.043	0.533±0.041
german	0.613±0.022	0.654±0.010	0.611±0.014	0.589±0.029	0.623±0.020	0.654±0.010	**0.666±0.019**	0.587±0.015	0.600±0.028
shoppers	0.696±0.002	0.703±0.001	0.691±0.003	0.689±0.008	0.684±0.004	0.689±0.005	**0.704±0.007**	0.665±0.003	0.680±0.007
magic	0.652±0.003	0.652±0.003	0.644±0.004	0.643±0.005	0.622±0.006	0.643±0.002	**0.655±0.005**	0.602±0.008	0.652±0.003
parkinson	**0.589±0.021**	**0.589±0.021**	0.475±0.019	0.564±0.013	0.495±0.028	0.560±0.022	0.576±0.021	0.548±0.039	0.521±0.032
shill	0.715±0.005	0.724±0.005	0.701±0.008	0.704±0.005	0.730±0.004	**0.733±0.006**	0.684±0.005	0.689±0.004	0.715±0.005
period	0.556±0.070	0.633±0.108	0.611±0.086	0.600±0.108	**0.667±0.050**	0.600±0.065	0.467±0.097	0.578±0.067	0.611±0.070
wireless	**0.598±0.018**	0.596±0.018	0.586±0.022	0.589±0.018	0.516±0.034	0.516±0.014	0.578±0.017	0.473±0.012	0.597±0.016
recognition	0.594±0.009	**0.595±0.010**	0.423±0.013	0.590±0.008	0.248±0.007	0.240±0.013	0.298±0.009	0.586±0.009	0.532±0.010
Electrical	0.688±0.008	0.686±0.009	0.600±0.008	0.625±0.005	0.660±0.011	0.617±0.006	**0.694±0.010**	0.674±0.006	0.685±0.007
Average	0.620±0.021	**0.634±0.023**	0.569±0.019	0.607±0.024	0.565±0.022	0.584±0.020	0.568±0.022	0.588±0.019	0.602±0.020

Note:p-values of the Wilcoxon test (IV-RANFSI vs. others): Original (0.0107), IV-DRSA (0.0002), AE-RAR (0.0002), PMLCE (0.0081), SNCMI (0.0059), KNCMI (0.0105), HKCMI (0.0002), FS-RSA (0.0015). All comparisons show significant differences ($p < 0.05$).

Table 9
Classification performance comparison with XGBoost classifier (40% noise).

Dataset	Original	IV-RANFSI	IV-DRSA	AE-RAR	PMLCE	SNCMI	KNCMI	HKCMI	FS-RSA
abalone	0.402±0.012	0.414±0.014	0.413±0.010	**0.422±0.015**	0.409±0.024	0.406±0.012	0.397±0.013	**0.422±0.012**	0.402±0.012
churn	0.669±0.010	0.709±0.008	0.672±0.006	0.707±0.004	0.696±0.006	**0.711±0.010**	0.678±0.009	0.702±0.008	0.677±0.011
dermatology	0.567±0.051	0.568±0.068	0.377±0.032	**0.589±0.065**	0.480±0.055	0.558±0.068	0.461±0.049	0.561±0.050	0.519±0.045
german	0.584±0.030	0.656±0.013	0.577±0.015	0.605±0.022	0.624±0.021	0.651±0.012	**0.665±0.018**	0.590±0.014	0.604±0.023
shoppers	0.690±0.006	0.703±0.001	0.677±0.005	**0.704±0.005**	0.694±0.002	0.699±0.001	0.699±0.005	0.702±0.001	0.692±0.006
magic	0.652±0.006	0.651±0.004	0.645±0.007	**0.662±0.004**	0.647±0.006	0.645±0.001	0.643±0.004	0.647±0.005	0.652±0.006
parkinson	0.522±0.038	0.565±0.036	0.509±0.020	0.568±0.030	0.489±0.026	0.543±0.014	**0.572±0.018**	0.531±0.052	0.498±0.035
shill	0.696±0.004	0.725±0.003	0.682±0.008	0.732±0.005	0.726±0.004	**0.733±0.006**	0.715±0.007	0.712±0.006	0.696±0.004
period	0.533±0.083	0.622±0.054	0.511±0.155	0.611±0.050	**0.678±0.089**	0.533±0.067	0.478±0.067	0.533±0.075	0.467±0.057
wireless	0.552±0.015	0.553±0.017	0.543±0.009	**0.599±0.015**	0.511±0.030	0.567±0.026	0.546±0.013	0.568±0.018	0.552±0.015
recognition	0.563±0.008	0.566±0.010	0.393±0.006	**0.582±0.009**	0.261±0.015	0.269±0.012	0.331±0.010	0.579±0.009	0.494±0.010
Electrical	0.661±0.013	0.661±0.013	0.588±0.008	0.627±0.005	0.670±0.008	0.606±0.011	0.664±0.012	**0.691±0.009**	0.664±0.004
Average	0.591±0.023	0.616±0.020	0.549±0.023	**0.617±0.019**	0.574±0.024	0.577±0.020	0.571±0.019	0.603±0.022	0.576±0.019

Note: p-values of the Wilcoxon test (IV-RANFSI vs. others): Original (0.0020), IV-DRSA (0.0002), AE-RAR (0.7473†), PMLCE (0.0276), SNCMI (0.0320), KNCMI (0.0112), HKCMI (0.2175†), FS-RSA (0.0020). † indicates no significant difference ($p > 0.05$).

granular representation of uncertainty compared to conventional single-valued approaches. In summary, the empirical evidence substantiates that IV-RANFSI not only enhances classification precision but also offers a stable and versatile preprocessing solution for high-noise environments.

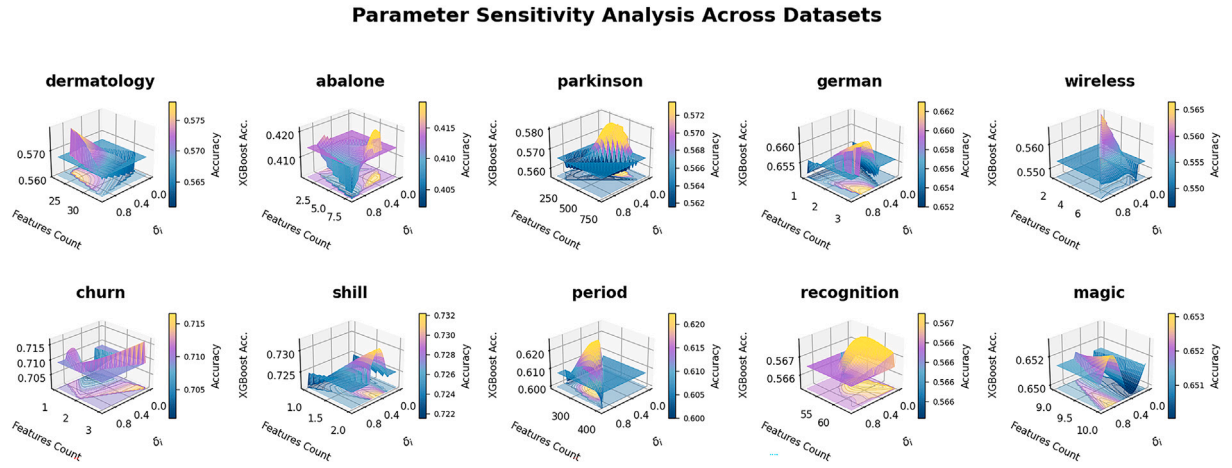

Fig. 5. Parameter sensitivity analysis visualization.

5.4. Parameter sensitivity analysis of algorithm

This section investigates the sensitivity of the dominance threshold parameter δ_i on the performance of the XGBoost classifier within the IV-RANFSI algorithm across ten diverse datasets. The parameter δ_i is used to control the strictness of the dominance relation in feature selection and thereby affect the calculation of rough set approximation. To systematically evaluate the parameter's influence, we tested six different δ_i values in the range $[0, 1]$, specifically $\delta_i \in \{0.0, 0.2, 0.4, 0.6, 0.8, 1.0\}$, across all benchmark datasets. For each parameter setting, we measured XGBoost classification accuracy and tracked the corresponding feature counts. The sensitivity was quantified using the coefficient of variation (CV) and accuracy range metrics. The sensitivity analysis reveals remarkable robustness of the IV-RANFSI algorithm to parameter variations. The mean coefficient of variation across all datasets is merely 0.83%, indicating high performance stability across different δ_i values. Furthermore, the average accuracy range across all parameter settings is only 0.0129, demonstrating that the algorithm's performance is relatively insensitive to parameter variations. This stability suggests that the algorithm can maintain consistent classification performance even when the threshold parameter deviates from its optimal value.

Regarding parameter selection, the performance of each δ_i value across all datasets is as follows: $\delta_i = 0.0$ (mean accuracy: 0.6015 ± 0.0859), $\delta_i = 0.2$ (0.5981 ± 0.0841), $\delta_i = 0.4$ (0.6038 ± 0.0847), $\delta_i = 0.6$ (0.6047 ± 0.0866), $\delta_i = 0.8$ (0.6020 ± 0.0861), and $\delta_i = 1.0$ (0.5986 ± 0.0882). Based on these results, we recommend $\delta_i = 0.6$ as the optimal parameter setting, achieving the best overall performance with a mean accuracy of 0.6047 across all datasets. Despite the overall stability, the optimal δ_i value varies across datasets, with different datasets exhibiting different preferences for threshold settings. For instance, dermatology achieves its best performance at $\delta_i = 0.8$ (0.5779), while abalone and parkinson both peak at $\delta_i = 0.4$ with accuracies of 0.4192 and 0.5735, respectively. The overall classification performance ranges from 0.4192 (abalone) to 0.7322 (shill), reflecting the inherent difficulty and characteristics of different datasets. Notably, the recognition dataset demonstrates exceptional stability with a CV of only 0.03%, while period shows the highest sensitivity at 1.72%.

Fig. 5 presents a comprehensive 3D visualization of the parameter sensitivity analysis across all ten datasets. Each subplot illustrates the relationship between the parameter δ_i, the number of selected features, and the XGBoost classification accuracy. This visual evidence further substantiates the robustness of the algorithm and justifies the selection of $\delta_i = 0.6$ as a parameter setting that generalizes effectively across diverse datasets and classification tasks.

5.5. Robustness evaluations of algorithm

This section investigates the algorithmic robustness by analyzing trends in classification accuracy across various noise levels. Comprehensive quantitative results and the corresponding performance trajectories are detailed in Table 10 and Fig. 6, respectively.

To rigorously evaluate the resilience of each algorithm against data perturbations, we utilized the twelve benchmark datasets presented in Table 6 to synthesize five corrupted variants. These variants incorporate increasing noise ratios ranging from 0% to 40%. The primary objective is to assess how varying noise levels affect the stability of feature scoring and the subsequent classification accuracy. For each algorithm, the average accuracy across these noise intervals under XGBoost is calculated and visualized as a line chart to facilitate comparative performance analysis.

As observed in Table 10, the proposed IV-RANFSI algorithm exhibits superior performance compared to the baseline methods across the majority of datasets. Specifically, IV-RANFSI achieves the highest average accuracy (**0.7373**), outperforming the second-best algorithm, AE-RAR (**0.7293**), and significantly exceeding standard approaches like IV-DRSA (**0.6425**). Notably, in challenging datasets such as *Period* and *Parkinson*, IV-RANFSI maintains a substantial lead, demonstrating its capability to extract salient features even under severe data corruption.

The visual evidence presented in Fig. 6 further underscores the stability of IV-RANFSI. While most algorithms experience a rapid deterioration in classification performance as the noise ratio escalates, the accuracy curve of IV-RANFSI remains consistently elevated

Table 10
Robustness evaluation of feature scores for different algorithms.

Dataset	FS-RSA	HKCMI	IV-DRSA	IV-RANFSI	PMLCE	AE-RAR	SNCMI	KNCMI
German	0.6124	0.6217	0.6110	**0.6684**	0.6478	0.6278	0.6606	0.6488
Dermatology	0.6335	0.7228	0.4535	0.7496	0.5010	**0.7591**	0.7367	0.6140
Period	0.5911	0.6089	0.5622	**0.6733**	0.6600	0.6467	0.6000	0.6111
Parkinson	0.6384	0.7097	0.6527	**0.7279**	0.6220	0.7152	0.6670	0.6289
Electrical	0.8189	**0.8361**	0.6475	0.8206	0.8223	0.7552	0.6722	0.8207
Churn	0.7908	0.7864	0.7689	0.8098	0.7846	**0.8116**	0.7995	0.7832
Shill	0.8288	0.8215	0.7782	0.8433	0.8382	**0.8476**	0.8430	0.8275
Abalone	0.4693	0.4748	0.4656	0.4786	0.4564	**0.4856**	0.4665	0.4590
Shoppers	0.7677	0.7702	0.7617	0.7821	0.7517	**0.7829**	0.7547	0.7733
Magic	0.7484	0.7267	0.7048	0.7498	0.6960	**0.7523**	0.7461	0.7284
Wireless	0.7621	0.7506	0.7428	0.7693	0.6456	**0.7892**	0.7645	0.7552
Recognition	0.6906	0.7745	0.5607	0.7747	0.3644	**0.7788**	0.3405	0.4591
Avg	0.6959	0.7170	0.6425	**0.7373**	0.6492	0.7293	0.6710	0.6758

Note: p-values of the Wilcoxon test (IV-RANFSI vs. others): FS-RSA (0.0002), HKCMI (0.0017), IV-DRSA (0.0002), PMLCE (0.0005), AE-RAR (0.4548†), SNCMI (0.0002), KNCMI (0.0005). † indicates no significant difference ($p > 0.05$).

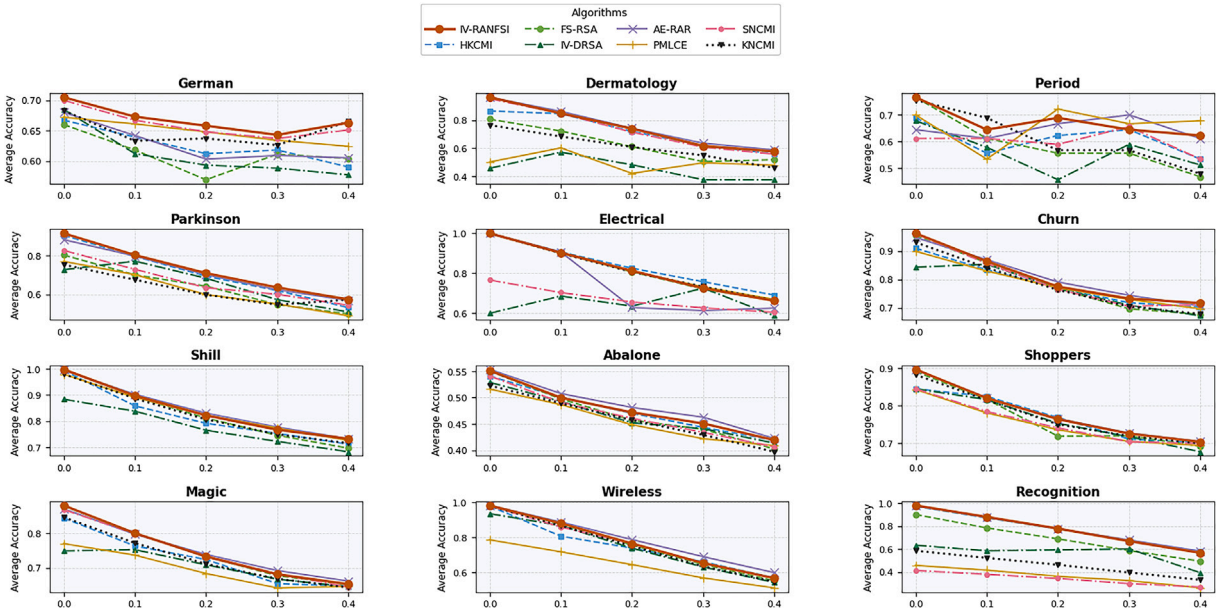

Fig. 6. Robustness evaluation of different algorithms under varying noise levels.

and exhibits a more gradual decline. Consequently, the proposed method offers more reliable quantification of feature importance, ensuring robust performance in practical applications where data contamination is inevitable.

5.6. Statistical test

To rigorously evaluate the performance of the proposed IV-RANFSI algorithm, we implemented a comprehensive two-stage statistical analysis framework that compares it against seven state-of-the-art feature selection methods. This framework includes the Iman-Davenport test for global hypothesis testing and the Wilcoxon signed-rank test for pairwise post-hoc comparisons. The analysis was conducted across twelve benchmark datasets and three distinct learning paradigms: Random Forest (RF), Support Vector Machines (SVM), and XGBoost.

First, the Iman-Davenport test was applied to determine whether significant performance disparities exist among the eight evaluated algorithms. Setting the significance level at $\alpha = 0.05$ with a critical value of $F_{crit} = 2.045$, the results are summarized in Table 11. For all three classifiers, the computed F-statistics ($F_{RF} = 11.026$, $F_{SVM} = 9.286$, and $F_{XGB} = 10.057$) significantly exceed the critical value, with all p-values well below the 0.05 threshold ($p < 10^{-7}$). Consequently, the null hypothesis H_0, which posits equivalent performance across all algorithms, is rejected for every classifier configuration. This confirms the existence of substantial performance heterogeneity among the evaluated methods.

Following the rejection of the global null hypothesis, we conducted the one-sided Wilcoxon signed-rank test to assess specific pairwise differences between IV-RANFSI and the other seven competitors. The significance level was maintained at 0.05, with the alternative hypothesis H_1 suggesting that IV-RANFSI yields superior classification performance. As detailed in Table 12, IV-RANFSI

Table 11
Global statistical significance results using the Iman-Davenport test ($\alpha = 0.05$, $F_{crit} = 2.045$).

Classifier	Friedman value	Iman–Davenport F	P value
RF	41.532	11.026	1.44×10^{-9}
SVM	38.647	9.286	2.76×10^{-8}
XGBoost	40.110	10.057	7.27×10^{-9}

Table 12
Wilcoxon signed-rank test p-values comparing IV-RANFSI with competitor algorithms across three classifiers.

Classifier	AE-RAR	FS-RSA	SNCMI	HKCMI	KNCMI	PMLCE	IV-DRSA
RF	0.0005	0.0010	< 0.001	0.0005	0.0005	0.0005	< 0.001
SVM	0.0134	0.0068	0.0012	0.0061	0.0012	0.0005	< 0.001
XGBoost	0.0171	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001

Note: All tests are one-sided (H_1 : IV-RANFSI > others). P-values < 0.05 indicate that IV-RANFSI performs significantly better than the corresponding competitor.

achieved statistically significant improvements over all competitors across all classifier paradigms. Notably, even when compared against strong competitors such as AE-RAR and FS-RSA, the p-values remained below 0.05 (e.g., $p = 0.0171$ for AE-RAR under XGBoost). In most cases, such as comparisons against SNCMI, HKCMI, and KNCMI, the p-values were lower than 0.001, statistically substantiating that the superiority of IV-RANFSI is robust and not attributable to random variation.

The comprehensive ranking analysis further corroborates these findings, identifying three distinct performance tiers. The superior tier is led by **IV-RANFSI**, which consistently secures the first rank across all classifiers (Mean Ranks: RF = 1.29, SVM = 1.79, XGBoost = 1.71), followed by **AE-RAR** and **FS-RSA**. The intermediate tier comprises **SNCMI**, **HKCMI**, and **KNCMI**, while the inferior tier consists of **PMLCE** and **IV-DRSA**. The consistent top-ranking of IV-RANFSI across diverse learning environments validates its exceptional stability and substantiates its efficacy as a robust feature selection algorithm.

6. Conclusion

In summary, this research addresses the challenge of robust feature selection in complex data environments through the IV-RANFSI framework. By integrating noise-resistance mechanisms with uncertainty modeling, the proposed approach offers a reliable solution for interval-valued ordered decision systems. The core contributions, practical utility, and potential research directions are elaborated upon in the subsequent subsections.

6.1. Summary of contributions

This paper introduces the IV-RANFSI framework, a comprehensive solution designed to tackle the significant challenges of robust feature selection in noisy interval-valued ordered decision systems. The proposed approach integrates three key innovations: an elastic anti-noise mechanism that adaptively distinguishes noise from boundary samples through relative density analysis; interval-valued dominance relations that effectively model uncertainty using IDD and IOD metrics; and a synergistic feature evaluation methodology that balances individual relevance with collective interactions. Extensive experimental validation across twelve UCI datasets demonstrates that IV-RANFSI significantly outperforms existing methods in terms of classification accuracy, noise tolerance, and algorithmic stability across multiple classifier paradigms.

6.2. Practical implications

The practical implications of the IV-RANFSI framework are substantial, especially in high-stakes decision-making contexts characterized by data uncertainty. In medical diagnostics, such as the detection of Parkinson's disease and heart disease, this framework effectively manages the inherent variability in clinical measurements (intervals) while robustly filtering out mislabeled records (noise). This capability significantly reduces the risk of misdiagnosis in borderline cases. Similarly, in financial risk assessment, the framework's ability to manage ordered data ensures that credit scoring models remain reliable, even when the training data contains subjective errors or inconsistencies. By selecting feature subsets that are both robust and synergistic, IV-RANFSI enhances the interpretability and reliability of decision support systems in these real-world environments.

6.3. Future directions

Future research directions will focus on broadening the applicability and technical depth of the proposed framework. First, we aim to extend the proposed method to accommodate multi-modal data types, particularly by integrating interval-valued structural data with unstructured text and image features. Second, we will explore the integration of IV-RANFSI with deep learning architectures, examining how the elastic anti-noise mechanism can function as a preprocessing or attention layer within neural networks. Finally, we intend to investigate adaptive parameter optimization strategies utilizing evolutionary algorithms and explore their applications in emerging distributed environments, such as federated learning and edge computing, where noise resilience and feature efficiency are of paramount importance.

CRediT authorship contribution statement

Yinliang Liu: Writing – review & editing, Writing – original draft, Visualization, Software, Methodology, Investigation, Formal analysis, Data curation. **Xiaoyan Zhang:** Validation, Supervision, Project administration, Methodology, Investigation, Funding acquisition, Conceptualization.

Declaration of competing interests

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

Acknowledgement

The authors would like to thank the Associate Editor and the reviewers for their insightful comments and suggestions.

This work was supported by the National Natural Science Foundation of China (Grant No. 12371465) and the Chongqing Natural Science Foundation of China (Grant No. CSTB2023NSCQ-MSX1063).

Data availability

The link to all data used for the research described has been shared in the article.

References

[1] M. Li, H. Wang, J. Li, Mining conditional functional dependency rules on big data, Big Data Min. Anal. 3 (2020) 68–84, https://doi.org/10.26599/BDMA.2019.9020019
[2] J. Zhang, T. Li, D. Ruan, D. Liu, Neighborhood rough sets for dynamic data mining, Int. J. Intell. Syst. 27 (2012) 317–342, https://doi.org/10.1002/int.21523
[3] H. Bai, D. Li, Y. Ge, J. Wang, F. Cao, Spatial rough set-based geographical detectors for nominal target variables, Inf. Sci. 586 (2022) 525–539, https://doi.org/10.1016/j.ins.2021.12.019
[4] J. Ding, X. Hu, V. Gudivada, A machine learning based framework for verification and validation of massive scale image data, IEEE Trans. Big Data 7 (2021) 451–467, https://doi.org/10.1109/TBDATA.2017.2680460
[5] C. Liu, J. Lai, B. Lin, D. Miao, An improved ID3 algorithm based on variable precision neighborhood rough sets, Appl. Intell. 53 (2023) 23641–23654, https://doi.org/10.1007/s10489-023-04779-y
[6] Z. Pawlak, Rough sets, Int. J. Comput. Inf. Sci. 11 (1982) 341–356, https://doi.org/10.1007/BF01001956
[7] K. Yuan, D. Miao, W. Pedrycz, W. Ding, H. Zhang, Ze-HFS: zentropy-based uncertainty measure for heterogeneous feature selection and knowledge discovery, IEEE Trans. Knowl. Data Eng. 36 (2024) 7326–7339, https://doi.org/10.1109/TKDE.2024.3419215
[8] T.L. Tseng, C.C. Huang, K. Fraser, H.W. Ting, Rough set based rule induction in decision making using credible classification and preference from medical application perspective, Comput. Methods Programs Biomed. 127 (2016) 273–289, https://doi.org/10.1016/j.cmpb.2015.12.015
[9] X. Zhang, M. Li, S. Shao, J. Wang, (I, O)-fuzzy rough sets based on overlap functions with their applications to feature selection and image edge extraction, IEEE Trans. Fuzzy Syst. 32 (2024) 1796–1809, https://doi.org/10.1109/TFUZZ.2023.3335108
[10] J. Lin, W. Wen, J. Liao, A novel concept-cognitive learning method for bird song classification, Mathematics 11 (2023) 4298, https://doi.org/10.3390/math11204298
[11] W.Z. Wu, Y. Leung, J.S. Mi, Granular computing and knowledge reduction in formal contexts, IEEE Trans. Knowl. Data Eng. 21 (2009) 1461–1474, https://doi.org/10.1109/TKDE.2008.223
[12] Y. Yao, Interpreting concept learning in cognitive informatics and granular computing, IEEE Trans. Syst. Man Cybern. B Cybern. 39 (2009) 855–866, https://doi.org/10.1109/TSMCB.2009.2013334
[13] W. Li, H. Zhou, W. Xu, X.Z. Wang, W. Pedrycz, Interval dominance-based feature selection for interval-valued ordered data, IEEE Trans. Neural Netw. Learn. Syst. 33 (2022) 6898–6912, https://doi.org/10.1109/TNNLS.2022.3184120
[14] X. Liu, J. Dai, J. Chen, C. Zhang, Unsupervised attribute reduction based on α-approximate equal relation in interval-valued information systems, Int. J. Mach. Learn. Cybern. 11 (2020) 1693–1707, https://doi.org/10.1007/s13042-020-01091-w
[15] X. Zhang, X. Shen, Graph-driven feature selection via granular-rectangular neighborhood rough sets for interval-valued data sets, Appl. Soft Comput. 169 (2025) 112716, https://doi.org/10.1016/j.asoc.2025.112716
[16] J. Chen, H. Su, Z. Li, J. Zhai, Self-adaptive interval dominance-based feature selection for monotonic classification of interval-valued attributes, Int. J. Mach. Learn. Cybern. 15 (2024) 2209–2228, https://doi.org/10.1007/s13042-023-02024-z
[17] X. Zhang, Z. Feng, Feature selection based on contradictory state sequence for multi-scale interval-valued decision table, Inf. Sci. 677 (2024) 120926, https://doi.org/10.1016/j.ins.2024.120926
[18] W. Xu, Z. Yuan, Z. Liu, Feature selection for unbalanced distribution hybrid data based on k-nearest neighborhood rough set, IEEE Trans. Artif. Intell. 5 (2024) 229–243, https://doi.org/10.1109/TAI.2023.3237203
[19] W. Ding, C. Lin, Z. Cao, Deep neuro-cognitive co-evolution for fuzzy attribute reduction by quantum leaping PSO with nearest-neighbor memeplexes, IEEE Trans. Cybern. 48 (2018) 2521–2534, https://doi.org/10.1109/tcyb.2018.2834390
[20] X. Yang, H. Chen, T. Li, C. Luo, A noise-aware fuzzy rough set approach for feature selection, Knowl. Based Syst. 250 (2022) 109092, https://doi.org/10.1016/j.knosys.2022.109092
[21] Y. Wu, P. Li, Y. Zou, Partial multi-label feature selection with feature noise, Pattern Recognit. 162 (2025) 111310, https://doi.org/10.1016/j.patcog.2024.111310
[22] Z. Chen, C. Wu, Y. Zhang, Z. Huang, B. Ran, M. Zhong, N. Lyu, Feature selection with redundancy-complementariness dispersion, Knowl.-based Syst. 89 (2015) 203–217, https://doi.org/10.1016/j.knosys.2015.07.004
[23] L. Farek, A. Benaidja, An optimal feature selection method for text classification through redundancy and synergy analysis, Multimed. Tools Appl. 84 (2025) 16397–16423, https://doi.org/10.1007/s11042-024-19736-1
[24] L. Chen, J. Chen, Y. Lin, Feature selection considering synergy between features based on soft neighborhood rough sets, Eng. Appl. Artif. Intell. 150 (2025) 110553, https://doi.org/10.1016/j.engappai.2025.110553
[25] S. Xia, Z. Xiong, Y. Luo, L. Dong, C. Xing, Relative density based support vector machine, Neurocomputing 149 (2015) 1424–1432, https://doi.org/10.1016/j.neucom.2014.08.053
[26] W. Xu, W. Ye, Incremental feature selection: parallel approach with local neighborhood rough sets and composite entropy, Pattern Recognit. 159 (2025) 111141, https://doi.org/10.1016/j.patcog.2024.111141
[27] S. Patra, P. Modi, L. Bruzzone, Hyperspectral band selection based on rough set, IEEE Trans. Geosci. Remote Sens. 53 (2015) 5495–5503, https://doi.org/10.1109/TGRS.2015.2424236