Adaptive Noise-Resilient Dynamic Feature Extraction for Interval-Valued Data With Evolving Attributes

Yinliang Liu, Xiaoyan Zhang, Yuhua Qian, *Member, IEEE*, and Weihua Xu

***Abstract*—In the context of dynamic feature evolution and intrinsic data noise, the pursuit of robust machine learning has led to develop an adaptive feature selection architecture with both noise reduction and relationship preservation abilities. However, research on robust feature selection tailored to monotonic classification has not been sufficiently investigated. When addressing multivalued ordered data systems, dominance-based rough set approaches show potential but require systematic modification. Practical applications often involve dynamically evolving feature sets shaped by temporal or situational factors, yet existing methodologies lack efficient mechanisms to address such scenarios. In view of these three issues, this study proposes an interval-valued adaptive noise cancellation and dynamic feature selection model. The main contributions of this work are threefold. First, an adaptive k-nearest neighbor framework for density-driven noise suppression is established to improve the robustness of feature extraction. Second, two key metrics, IDD and IOD, are utilized to systematically process interval-valued ordered data systems. Third, a dynamic matrix update strategy is introduced to support both incremental and decremental attribute updates. Experimental results on benchmark datasets demonstrate enhanced noise immunity. The proposed method provides a systematic solution for preserving feature relevance in evolving data environments with inherent measurement uncertainties.**

***Index Terms*—Adaptive feature extraction, evolving attributes handling, interval-valued ordered decision system (IV-ODS), noise-resilient, robust machine learning.**

NOMENCLATURE

Symbol	*Corresponding Explanations*
${Dc_t}^{\preceq}$	Upward union of the decision classes.
$\delta(.,.)$	Interval dominance degree of two intervals.
$\omega(.,.)$	Interval overlap degree of two intervals.
$\mathrm{Dom}_A^{\ddot{\cdot\cdot}}$	Interval-valued dominance relation.
$r_{Dc_t}^A$	Adaptive radius.
$K_{Dc_t}^A$	Adaptive number of neighbors.
$\rho_{Dc_t}^A$	Sample density.
$\mathrm{Dom}_A^{\preceq}$	Dominance relation matrix.
$\mathrm{DiagDom}_A^{\preceq}$	Dominant diagonal matrix.
CDD	Class discrimination degree.
GCDD	Global class discrimination degree.
MSIG	Matrix-based significance.
$\mathrm{MORE}^{\preceq}$	Matrix-based ordered ranking entropy.
$\mathrm{MORMI}^{\preceq}$	Matrix-based ordered ranking mutual information.
$\mathrm{MORCE}^{\preceq}$	Matrix-based ordered ranking conditional entropy.
σ	Feature evaluation index.
θ	Redundancy penalty intensity.

Received 29 April 2025; revised 22 October 2025; accepted 15 November 2025. Date of publication 12 December 2025; date of current version 5 February 2026. This work was supported in part by the National Natural Science Foundation of China under Grant 12371465 and Grant 62376229, and in part by the Natural Science Foundation of Chongqing under Grant CSTB2023NSCQ-MSX1063. Recommended by Associate Editor D. Pelusi. *(Corresponding author: Xiaoyan Zhang.)*

Yinliang Liu, Xiaoyan Zhang, and Weihua Xu are with the College of Artificial Intelligence, Southwest University, Chongqing 400715, China (e-mail: yin2728906486@email.swu.edu.cn; zxy19790915@163.com; chxuwh@gmail.com).

Yuhua Qian is with the Institute of Big Data Science and Industry, Shanxi University, Taiyuan 030006, China (e-mail: jinchengqyh@126.com).

Digital Object Identifier 10.1109/TFUZZ.2025.3636068

I. INTRODUCTION

FEATURE selection remains a cornerstone in data mining and machine learning [1], particularly for handling high-dimensional data in classification tasks. Rough set theory, originally formalized by Pawlak in the early 1980s [2], has emerged as a powerful mathematical approach for managing vague and uncertain data in analysis. This granular computing framework has witnessed substantial theoretical development and practical implementations in diverse fields. Notably, its successful applications span machine learning [3], [4], knowledge discovery in databases [5], [6], decision support systems [7], [8], pattern recognition [9], [10], image processing systems [11], and medical diagnostics [12], [13]. With the growth of intricate data environments characterized by inherent noise [14], multivalued ordered attributes, and dynamic feature spaces, traditional feature selection methods face unprecedented obstacles. The ability to sustain classification performance with efficient computation in such environments has become a critical requirement for modern decision-making systems [15]. This is especially crucial in real-world applications ranging from sensor networks to clinical decision support systems, where stability against data flaws and flexibility to changing feature relevance are vital. A critical limitation of current monotonic classification methods lies in their undifferentiated treatment of noise patterns inherent to ordered data structures. Our investigation reveals the following three fundamental limitations of existing feature selection paradigms (see Fig. 1).

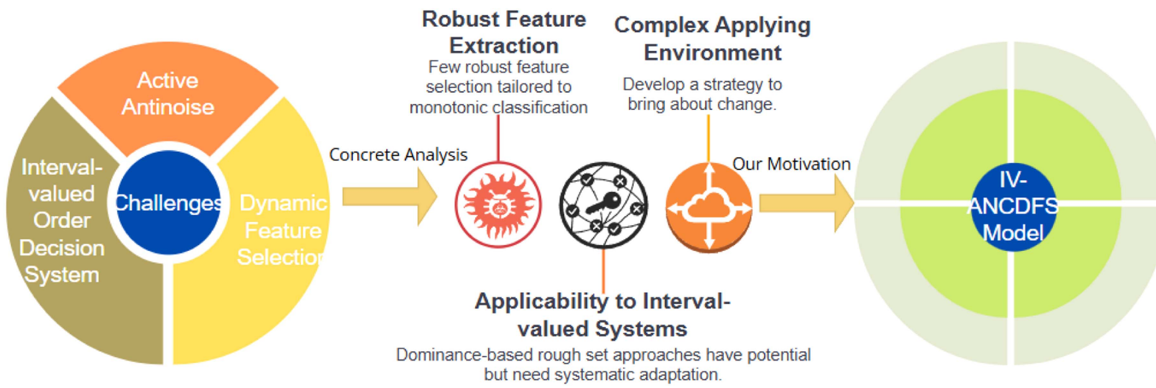

Fig. 1. Motivation of our works.

1) *Noise vulnerability in monotonic classification:* Existing robust feature selection methods prove inadequate for monotonic classification due to noise vulnerability disrupting ordinal relationships in strictly monotonic feature-decision dependencies.
2) *Multivalued ordered data complexity:* Interval-valued ordered decision systems (IV-ODS) challenge conventional dominance-based rough set approaches by introducing information granularity complexities in multivalued attribute processing.
3) *Dynamic feature space adaptation:* Static feature selection methods lack dynamic adaptability for evolving feature spaces, requiring urgent development of incremental updating mechanisms for real-time relevance maintenance.

These limitations collectively motivate the development of our proposed IV-ANCDFS model. This article presents several pivotal contributions to the field of robust feature selection and dynamic data analysis, addressing critical gaps in existing methodologies and advancing practical applications in evolving, noisy environments. First, we propose the IV-ANCDFS framework, a novel architecture that seamlessly integrates adaptive k-nearest neighbors (AKNN) for boundary noise suppression with density-driven sample filtering. Second, we introduce a two-standard-deviation (SD) interval construction method to explicitly model real world. Third, the framework develops matrix-based incremental and decremental dominance relation updates, enabling efficient adaptation to dynamically evolving feature spaces and avoiding complete matrix recalculation. Fourth, we establish class discrimination metrics (e.g., GCDD and MSIG) and redundancy-aware evaluation metrics to systematically balance feature relevance and redundancy, ensuring robust performance in high-dimensional noisy scenarios. Extensive experiments on synthetic and real-world datasets validate the superiority of IV-ANCDFS in noise tolerance, classification accuracy, and scalability, yielding an average accuracy improvement. These contributions collectively establish a rigorous, adaptive framework for feature extraction in dynamic interval-valued systems.

The rest of this article is organized as follows. First, Section II reviews related work in robust feature selection and dynamic decision systems. To establish theoretical groundwork, Section III establishes essential mathematical preliminaries. Building upon these foundations, Section IV details the IV-ANCDFS architecture and its algorithms. Subsequently, Section V presents experimental validation across synthetic and real-world datasets. Finally, Section VI concludes this article with research contributions and future directions.

II. Related Work

Recent advancements in robust feature selection have emphasized noise-resistant mechanisms for general classification tasks. For instance, FC [16] introduced similarity metrics based on fuzzy contradiction state sequences to enhance the robustness of multiscale interval-valued decision tables. KND-UFS [17] presents a novel unsupervised dimensionality reduction framework that synergistically integrates fuzzy C-means clustering with k-nearest neighbor (KNN) rough set theory. Similarly, K-nearest neighborhood conditional mutual information (KNCMI) [18] incorporated KNN rough sets and information entropy to address noise in imbalanced hybrid data. NGM [19] presents a semisupervised feature selection method for partially labeled ordered data, which integrates neighborhood discernibility degree, pseudolabel granular balls, and matrix updating techniques to efficiently select optimal features and improve classification accuracy. Recently, PMSNE [20] unifies robust label enhancement, sparse reconstruction denoising, and feature selection into an end-to-end framework. In addition, FS-RSA [21] 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 [22] introduces a unified evaluation function that quantifies feature correlation, redundancy, complementarity, and synergy. However, these methods primarily focus on generic classification scenarios and overlook the specific challenges posed by monotonic classification. The rigid ordinal dependencies between features and decision attributes in monotonic classification cause traditional noise-handling techniques to falter, as they often struggle to preserve critical ordinal structures under noisy conditions. This limitation underscores the necessity for custom noise-handling methods that adaptively align noise resistance while preserving the ordered structure.

TABLE I
SUMMARIZE THE CURRENT FEATURE SELECTION METHODS

Methods	Applying fields	Disadvantages
FC [16], KND-UFS [17], AAnFSCF [31]	Antinoise rough set	Cannot be applied to interval-valued and dynamic situations.
KNCMI [18], IFS-KNCMI [32], HKCMI [33]	Antinoise rough set	Cannot be applied to interval-valued and dynamic situations.
IV-DRSA [23], IGUFS [24], IFCRL [26]	Interval value	Poor antinoise performance and inefficient in dynamic situations.
PRIFS [25], IFDI [34], FC [16]	Interval value	Poor antinoise performance and inefficient in dynamic situations.
PMLCE-D [29], FACE-IF [35], GMNDRS [30]	Dynamic feature extraction	Poor antinoise performance, unsuitable for interval values.
KNCMI [18], IFS-KNCMI [32], HAR-A/HAR-D [27]	Dynamic feature extraction	Poor antinoise performance, unsuitable for interval values.

The handling of interval-valued ordered data has seen notable progress through dominance-based rough set approaches. Interval-valued dominance-based rough set (IV-DRSA) [23] pioneered interval dominance degrees (IDD) and interval overlap degree (IOD) metrics to construct interval-valued dominance relations, significantly improving classification accuracy. Extensions, such as IGUFS [24], utilized graph theory and matrix operations to reduce computational complexity in interval-valued systems, while PRIFS [25] integrated PROMETHEE for the fusion of multisource interval data. IFCRL [26] model introduces interval-intent fuzzy concepts, a recognition process, and intent similarity-based clustering to enhance both classification accuracy and efficiency in fuzzy environments. Despite these innovations, existing methods continue to struggle with the inherent granularity of multivalued ordered attributes. Traditional single-value dominance models, even when extended to interval data, lack systematic frameworks to address overlapping intervals and variable information granularity, which leads to suboptimal feature selection in complex interval-valued systems.

Dynamic environments demand feature selection frameworks capable of incrementally updating feature relevance. An attribute reduction algorithm (HAR)-A [27] and HAR-D [27] proposed matrix-based incremental algorithms to handle temporal feature additions and deletions, thereby enhancing the efficiency of ordered datasets. PMLCE [29] further improved scalability via parallel computing and composite entropy metrics, while GMNDRS [30] refined matrix-based updates for multilevel intuitionistic fuzzy data. Nevertheless, these approaches predominantly focus on specific dynamic scenarios (e.g., object insertion/deletion) and lack broader mechanisms for continuous relevance tracking in the evolving feature spaces. Static frameworks fail to adapt to contextual shifts or feature drift, resulting in outdated feature subsets and diminished performance over time.

The applications and drawbacks of existing methods are illustrated in Table I. The current research identifies the following 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, and 3) restricted adaptability to dynamically evolving feature spaces. While existing methods, such as IV-DRSA, HAR-A/D, and FC, address individual aspects, their isolated focus overlooks the intertwined challenges, such as noise, multivalued data, and dynamics. This underscores the necessity for an integrated approach that synergistically addresses these issues while maintaining ordinal integrity.

III. PRELIMINARIES

This section mainly introduces some basic knowledge about IV-ODS, AKNN, and matrix incremental update. Some necessary mathematical symbols and their brief explanations can be seen in Nomenclature.

A. Interval-Valued Ordered Decision System

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 nonempty 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, $F = \{F_{a_k} \mid X \to V_{a_k},\ a_k \in \text{AS}\}$, where F_{a_k} is the interval value of a_k on $x \in X$. $Dc_t = \{x_i \in U \mid v(x_i, D) = d_t, t \in \{1, 2, \ldots, T\}\}$. 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: $Dc_1 \prec Dc_2 \prec \ldots \prec Dc_t$. We define upward union: $Dc_t^{\preceq} = \bigcup_{s=t}^{T} Dc_s$.

This method defines interval-valued dominance relations in IV-ODS using two key metrics: IDD and IOD for interval comparisons.

IDD [23], denoted as $\delta(\iota, \kappa)$, quantifies the dominance relation between the two intervals $\iota = [\underline{\iota}, \overline{\iota}]$ and $\kappa = [\underline{\kappa}, \overline{\kappa}]$. The IDD is defined as follows:

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

where the range of δ 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 $\omega(\iota, \kappa)$, 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 u and v, respectively.

The IOD [23] is defined as

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

where the range of ω belongs to [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 $x, y \in X$ and a subset A of conditional attributes AS, the interval-valued dominance relation on attribute set A is defined as

$$\mathrm{Dom}_A^{\eta,\gamma} = \{(x, y) \in X^2 \mid \delta(\iota, \kappa) \geq \eta \wedge \omega(\iota, \kappa) \leq \gamma\}$$

If $\delta(\iota, \kappa) \geq \eta$ and $\omega(\iota, \kappa) \leq \gamma$ hold for all $a \in A$, then we say that y dominates x at levels η and γ, denoted as $y\mathrm{Dom}_A^{\eta,\gamma}x$.

Let X be the universe of objects and A an attribute set in an IV-ODS. Given dominance thresholds η and γ, write $y\,\mathrm{Dom}_A^{\eta,\gamma}\,x$ to mean that y dominates x on A with respect to (η, γ). The *A-dominating set* of $x \in X$ is

$$\mathrm{Dom}_A^{+,\eta,\gamma}(x) \;=\; \bigl\{\, y \in X \;\bigm|\; y\,\mathrm{Dom}_A^{\eta,\gamma}\,x \,\bigr\}.$$

Conversely, the *A-dominated set* of x is obtained by reversing the dominance relation

$$\mathrm{Dom}_A^{-,\eta,\gamma}(x) \;=\; \bigl\{\, y \in X \;\bigm|\; x\,\mathrm{Dom}_A^{\eta,\gamma}\,y \,\bigr\}.$$

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.

B. Active Antinoise Fuzzy Dominance Rough Sets Using AKNN and Approximation Operators

In classification tasks, the density of samples is crucial for identifying outliers, as higher density samples are more likely to be correctly classified. Traditional KNN strategies involve parameters, such as the number of neighbors (*K*) and the radius (*r*), which are often chosen subjectively and lack adaptability. To address this, we utilize an AKNN strategy that learns these parameters from the data distribution, thereby enhancing its adaptability and generalization.

The adaptive radius $r_{Dc_t}^A$ [31] for the upward union Dc_t under subset A is defined as the average Euclidean distance between all pairs of samples in Dc_t under A

$$r_{Dc_t}^A = \frac{1}{|Dc_t|^2} \sum_{i=1}^{|Dc_t|} \sum_{j=1}^{|Dct|} \lambda_{Dc_t}^A(x_i, x_j) \tag{3}$$

where $\lambda_{Dc_t}^A(x_i, x_j)$ represents the Euclidean distance between samples x_i and x_j under A.

The adaptive number of neighbors $K_{Dc_t}^A$ [31] is determined as the average number of samples within the radius $r_{Dc_t}^A$ for each sample in Dc_t

$$K_{Dc_t}^A = \frac{1}{|Dc_t|} \sum_{i=1}^{|Dc_t|} N_{Dc_t}^A(x_i, r_{Dc_t}^A) \tag{4}$$

where $N_{Dc_t}^A(x_i, r_{Dc_t}^A)$ is the number of samples within the neighborhood of x_i with radius $r_{Dc_t}^A$.

As shown in Fig. 2, the adaptive radius *r* denotes the average distance between the sample pairs within a class, and the adaptive neighbor number *K* denotes the average number of neighbors of each sample within the radius *r*.

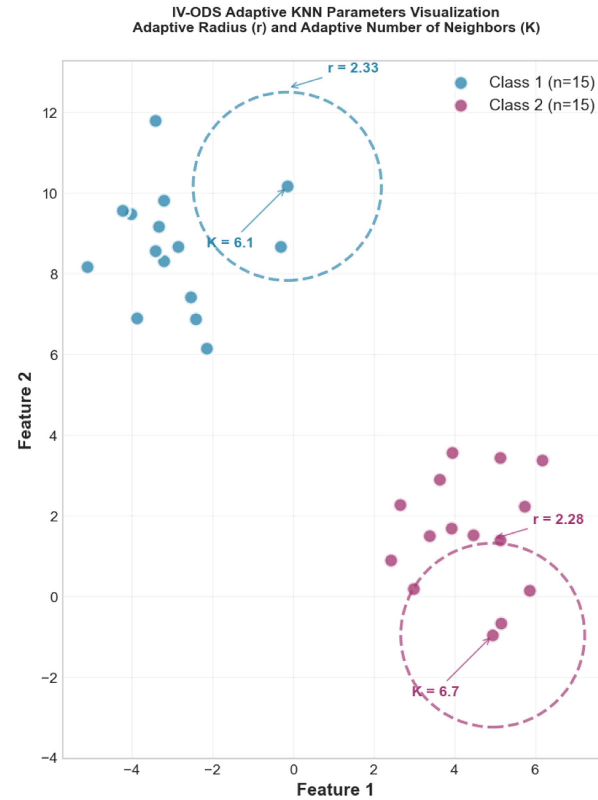

Fig. 2. AKNN parameters visualization.

A sample density x_i in Dc_t under A is calculated by

$$\rho_{Dc_t}^A(x_i) = \frac{N_{Dc_t}^A(x_i, r_{Dc_t}^A)}{K_{Dc_t}^A}. \tag{5}$$

This density measure compares the number of neighbors of x_i to the average number of neighbors in the class, helping identify outliers as samples with lower densities. In order to find the critical value of filtering out the number of sample, we design the critical densities values (CDV) for the lower approximation and the upper approximation, respectively, which are defined as follows.

For density lower approximation and density upper approximation, for all $A \subseteq AS$ and $x_i \in Dc_t^{\preceq}$, the CDV of samples outside the upward union $Dc_t^{\preceq}$ are

$$\rho_{S,\alpha}^{\mathrm{CDV}}(x_i) = \inf_{x_j \in S} \max\Bigl\{1 - \rho_S^A(x_j)$$

$$\mathrm{Dom}_A^{\alpha,\eta,\gamma}(x_i)(x_j),\; \rho_S^A(x_j)\Bigr\}. \tag{6}$$

Specifically, for the density approximations, they are defined as follows:

$$\begin{cases} \text{Density lower approximation:} & \underline{\rho_{(De_t^{\preceq})^c}^{\mathrm{CDV}}}(x_i) = \rho_{(De_t^{\preceq})^c,+}^{\mathrm{CDV}}(x_i) \\ \text{Density upper approximation:} & \overline{\rho_{De_t^{\preceq}}^{\mathrm{CDV}}}(x_i) = \rho_{De_t^{\preceq},-}^{\mathrm{CDV}}(x_i). \end{cases}$$

Based on the above two definitions, we introduce the definition of the upper and lower approximation operators.

For any $A \subseteq AS$, the lower approximation of the upward union $Dc_t^{\preceq}$ under A is

$$\underline{\mathrm{Dom}_A^{\preceq}}(Dc_t^{\preceq})(x_i) = \inf_{x_j \in L_r} \{1 - \mathrm{Dom}_A^{+,\eta,\gamma}(x_i)(x_j)\} \quad (7)$$

where $L_r = \{x_j \mid \rho^A_{(Dc_t^{\preceq})^c}(x_j) \geq \underline{\rho^{\mathrm{CDV}}_{(De_t^{\preceq})^c}}(x_i), x_j \in (Dc_t^{\preceq})^c\}$.

For any $A \subseteq AS$, the upper approximation of the upward union $Dc_t^{\preceq}$ under A is

$$\overline{\mathrm{Dom}_A^{\preceq}}(Dc_t^{\preceq})(x_i) = \sup_{x_j \in U_r} \{\mathrm{Dom}_A^{-,\eta,\gamma}(x_i)(x_j)\} \quad (8)$$

where $U_r = \{x_j \mid \rho^A_{Dc_t^{\preceq}}(x_j) \geq \overline{\rho^{CDV}_{(De_t^{\preceq})^c}}(x_i), x_j \in Dc_t^{\preceq}\}$.

Hence, AKNN establishes an adaptive KNN framework with density-driven noise suppression, formalizing dynamic approximation operators through critical density thresholds.

C. Matrix-Based Incremental Attribute Reduction Under Multifeature Modification [27]

This section presents incremental multiobjective attribute reduction algorithms that dynamically update dominance structures to efficiently handle feature evolution in ordered systems.

Let $S^{\preceq} = (X, AS, V, F)$ denote an ordered information system (OIS) [36]. For any subset $A \subseteq AS$, the dominance relation matrix on X with respect to A is defined as $\mathrm{Dom}_A^{\preceq} = [m^A(i,j)]_{n\times n}$, where

$$m^A(i,j) = \begin{cases} 1, & x_j \succeq_A x_i \\ 0, & \text{otherwise.} \end{cases}$$

The dominant diagonal matrix $\mathbf{DiagDom}_A^{\preceq} = [d^A(i,j)]_{n\times n}$, where $d^A(i,j)$ is defined as

$$d^A(i,j) = \begin{cases} \sum_{l=1}^{n} m^A(i,l), & i = j \\ 0, & i \neq j. \end{cases}$$

Its determinant is $|\mathbf{DiagDom}_A^{\preceq}| = \prod_{i=1}^{n} d^A(i,i)$, and the inverse matrix is $(\mathbf{DiagDom}_A^{\preceq})^{-1} = [\frac{1}{d^A(i,j)}]_{n\times n}$, where

$$\frac{1}{d^A(i,j)} = \begin{cases} \frac{1}{\sum_{l=1}^{n} m^A(i,l)}, & i = j \\ 0, & i \neq j. \end{cases}$$

Given an ordered decision system, the matrix dominance conditional entropy (MDCE) of $A \subseteq AS$ with respect to d is

$$\mathrm{MDH}_{d|A}^{\preceq}(X) = -\frac{\log \left| \mathbf{DiagDom}_{A\cup\{d\}}^{\preceq} \cdot (\mathbf{DiagDom}_A^{\preceq})^{-1} \right|}{|X|}. \quad (9)$$

In OIS [36], dataset dynamics can be categorized into feature addition and deletion. This section proposes two incremental algorithms to optimize complexity by leveraging prior results.

1) Incremental Reduction for Feature Addition: This section introduces an incremental MDCE update method for ordered systems using dominance relations and diagonal matrices.

When a new feature set $AS^{+} = \{A_{n+1}, A_{n+2}, \ldots, A_{n+n'}\}$ is added to $S^{\preceq}$, the updated dominance relation matrix $M_{A\cup A^+}^{\preceq} = [m^{A\cup A^+}(i,j)]_{n\times n}$ is defined as

$$m^{A\cup A^+}(i,j) = \begin{cases} 1, & x_j \succeq_{A\cup A^+} x_i \\ 0, & \text{otherwise.} \end{cases}$$

Upon introducing A^+, the updated dominant diagonal matrix $\mathbf{DiagDom}_{A\cup A^+}^{\preceq} = [d^{A\cup A^+}(i,j)]_{n\times n}$ is computed as

$$d^{A\cup A^+}(i,j) = \begin{cases} d^A(i,j) - m^{A\cup A^+}(i,j), & x_j \succeq_{A\cup A^+} x_i \\ d^A(i,j), & x_j \not\succeq_{A\cup A^+} x_i. \end{cases}$$

The incremental update mechanism ensures computational efficiency by reusing prior matrices $M_A^{\preceq}$ and $\mathbf{DiagDom}_A^{\preceq}$, modifying only entries affected by the new features A^+.

2) Incremental Reduction for Feature Deletion: When a feature subset $AS^{-} = \{A_{q_1}, A_{q_2}, \ldots, x_{q_{n'}}\}$ is removed, the updated dominance relation matrix $M_{A\setminus A^-}^{\succeq} = [m^{A\setminus A^-}(i,j)]_{n\times n}$ is defined as

$$m^{A\setminus A^-}(i,j) = \begin{cases} 1, & x_j \succeq_{A\setminus A^-} x_i \\ 0, & \text{otherwise.} \end{cases}$$

After removing A^-, the updated dominant diagonal matrix $\mathbf{DiagDom}_{A\setminus A^-}^{\succeq} = [d^{A\setminus A^-}(i,j)]_{n\times n}$ is computed as

$$d^{A\setminus A^-}(i,j) = \begin{cases} d^A(i,j) + m^{A\setminus A^-}(i,j), & x_j \succeq_{A\setminus A^-} x_i \\ d^A(i,j), & x_j \not\succeq_{A\setminus A^-} x_i. \end{cases}$$

The deletion update mechanism preserves computational efficiency by leveraging the prior matrix structures. Only entries affected by the removed features A^- require modification, avoiding full matrix recomputation.

IV. IV-ANCDFS Model

This section introduces a noise-resistant rough set framework for IV-ODS, combining adaptive approximation, matrix-based dominance relations, and redundancy-aware feature selection.

A. IV-ANCDFS Active Antinoise Method

By deriving r and K from the data's inherent characteristics, the AKNN strategy reduces the need for manual parameter tuning, enhancing its applicability across diverse datasets. Lower density values suggest potential outliers, as these samples possess fewer neighbors than the average. Overall, the AKNN strategy provides a data-driven density calculation method, enhancing classification by adapting to the characteristics of the dataset. In designing approximation operators resilient to noise, the objective is to develop rules that filter noisy boundary samples and select appropriate correct samples—namely, the best external and worst internal samples—for computing the approximations. A critical challenge is determining the number of samples to filter noise, while preventing overfitting.

To address this challenge, CDV and definitions of approximations are introduced for lower and upper approximations. In dominance-based rough sets, Definitions 1–4 establish lower and upper approximation operators using dominance relations and membership functions to handle fuzzy class boundaries in decision analysis.

TABLE II
DYNAMIC UPDATE OF DOMINANCE MATRIX (CASE: ATTRIBUTE ADDITION)

Matrix Type	Matrix Values		
	x_1	x_2	x_3
Original Matrix $M_A^{\eta,\gamma}$	1	1	0
	0	1	1
	1	0	1
Added Attribute Matrix $M_{A+}^{\eta,\gamma}$	1	0	1
	1	1	0
	0	1	1
Updated Matrix $M_{A\cup A+}^{\eta,\gamma}$	1	0#	0
	0	1	0#
	0#	0	1

indicate positions where dominance relation changed due to intersection operation (∩).

TABLE III
DYNAMIC UPDATE OF DOMINANCE MATRIX (CASE: ATTRIBUTE REMOVAL)

Matrix Type	Matrix Values		
	x_1	x_2	x_3
Original Matrix $M_A^{\eta,\gamma}$	1	0	1
	1	1	0
	0	1	1
Removed Attribute Matrix $\neg M_{A-}^{\eta,\gamma}$	1	1	0
	0	1	1
	1	0	1
Updated Matrix $M_{A\setminus A-}^{\eta,\gamma}$	1	1*	1
	1	1	1*
	1*	1	1

* show positions updated by union operation (∪).

Let $f(x_j) = \max\{1-\rho^A_{Dc_t^{\preceq}}(x_j), \rho^A_{Dc_t^{\preceq}}(x_j)\}$ for $x_j \in Dc_t^{\preceq}$.

Definitions 1 and 2: For all $A \subseteq AS$, with conventions $\underline{\rho^{\text{CDV}}_{(Dc_t^{\preceq})^c}}(x_i) = 1$ and $\overline{\rho^{\text{CDV}}_{Dc_t^{\preceq}}}(x_i) = 1$ if sets are empty

$$\rho^{\text{CDV}}_{S,\alpha}(x_i) = \inf\{f(x_j) \mid \text{Dom}_A^{\alpha,\eta,\gamma}(x_i,x_j) = 1\}. \tag{10}$$

Then, we give the definitions of the upper and lower approximation of density

$$\begin{cases} \text{Density lower approximation:} & \underline{\rho^{\text{CDV}}_{(De_t^{\preceq})^c}}(x_i) = \rho^{\text{CDV}}_{(De_t^{\preceq})^c,+}(x_i) \\ \text{Density upper approximation:} & \overline{\rho^{\text{CDV}}_{De_t^{\preceq}}}(x_i) = \rho^{\text{CDV}}_{De_t^{\preceq},-}(x_i). \end{cases}$$

Definition 3: For $x_j \in (Dc_t^{\preceq})^c$:

$$\underline{\text{Dom}_A^{\preceq}}(Dc_t^{\preceq})(x_i) = \begin{cases} 0, & \text{if } \exists\, x_j \in L_r \\ & \quad \text{Dom}_A^{+,\eta,\gamma}(x_i,x_j) = 1 \\ 1, & \text{if } \forall\, x_j \in L_r \\ & \quad \text{Dom}_A^{+,\eta,\gamma}(x_i,x_j) = 0. \end{cases} \tag{11}$$

where $L_r = \{x_j \mid \rho^A_{(Dc_t^{\preceq})^c}(x_j) \geq \underline{\rho^{\text{CDV}}_{(Dc_t^{\preceq})^c}}(x_i)\}$.

Complementing this, the upper decision rule handles positive dominance relations through existential quantification.

Definition 4: For $x_j \in Dc_t^{\preceq}$

$$\overline{\text{Dom}_A^{\preceq}}(Dc_t^{\preceq})(x_i) = \begin{cases} 1, & \text{if } \exists\, x_j \in U_r \\ & \quad \text{Dom}_A^{-,\eta,\gamma}(x_i,x_j) = 1 \\ 0, & \text{if } \forall\, x_j \in U_r \\ & \quad \text{Dom}_A^{-,\eta,\gamma}(x_i,x_j) = 0. \end{cases} \tag{12}$$

where $U_r = \{x_j \mid \rho^A_{Dc_t^{\preceq}}(x_j) \geq \overline{\rho^{\text{CDV}}_{Dc_t^{\preceq}}}(x_i)\}$.

The above definitions establish a dynamic mechanism for determining lower and upper approximations through dominance relations $\text{Dom}_A^{+,\eta,\gamma}$ and $\text{Dom}_A^{-,\eta,\gamma}$, membership functions ρ, and threshold conditions. This framework enables multicriteria ranking, attribute reduction, and rule extraction for complex data with partial order relations.

B. Matrix-Based Dominant Relation Construction Method and Incremental Update Method for IV-ODS

To address the computational redundancy in dominance matrix operations and enhance processing efficiency while accommodating dynamic feature updates, this article proposes a matrix-based method with an incremental update mechanism to improve efficiency.

1) Simulation of Dynamic Matrix Updates: Given an IV-ODS defined as $(X, AS \cup \{D\}, V, F)$ and two thresholds η and γ, let $a_k \in AS$. The matrices of IDD and IOD are derived as

$$\text{Dom}_A^{\eta,\gamma} = \{(x,y) \in X^2 \mid \delta(\iota,\kappa) \geq \eta \wedge \omega(\iota,\kappa) \leq \gamma\}$$

$$M_\eta^{a_k} = [\delta(\iota,\kappa)]_{i\times j}, \quad M_\gamma^{a_k} = [\omega(\iota,\kappa)]_{i\times j}$$

where $\delta(\iota,\kappa)$ and $\omega(\iota,\kappa)$ represent IDD and IOD, respectively, with $\iota = F(x,a_k)$ and $\kappa = F(y,a_k)$.

Definition 5: Given an IV-ODS, for any subset $A \subseteq AS$, the dominance relation Dom_A under A is defined. The dominance relation matrix on X with respect to A is represented as

$$M_A^{\eta,\gamma} = [m_A^{\eta,\gamma}(i,j)]_{n\times n}$$

where

$$m_A^{\eta,\gamma}(i,j) = \begin{cases} 1, & \text{if } x_j \text{Dom}_A^{\eta,\gamma} x_i \\ 0, & \text{otherwise.} \end{cases}$$

To enhance the theoretical foundation, we introduce the following corollaries derived from the aforementioned definitions.

Corollary 1. (Antisymmetry of Single-Attribute Dominance Matrix): For any single conditional attribute $a_k \in AS$, the dominance relation matrix $M_{a_k}^{\eta,\gamma}$ exhibits a complementary antisymmetry with respect to the main diagonal. Formally, the following holds.

1) All diagonal elements satisfy $m_{a_k}^{\eta,\gamma}(i,i) = 1$, reflecting the reflexivity of dominance relations.
2) For off-diagonal elements $(i \neq j)$, $m_{a_k}^{\eta,\gamma}(i,j) = 1 - m_{a_k}^{\eta,\gamma}(j,i)$.

This implies that if x_j dominates x_i under a_k, then x_i cannot dominate x_j, ensuring mutually exclusive dominance directions between the distinct objects.

Corollary 2. (Monotonicity of Dominance Relations With Thresholds): Let $\eta_1 \geq \eta_2$ and $\gamma_1 \leq \gamma_2$, where $0 < \eta_1, \eta_2, \gamma_1, \gamma_2 < 1$. For any $A \subseteq AS$, the dominance relations satisfy

$$\mathrm{Dom}_A^{\eta_1,\gamma_1}(x) \subseteq \mathrm{Dom}_A^{\eta_2,\gamma_2}(x).$$

Consequently, the dominance matrix $M_A^{\eta_1,\gamma_1}$ contains fewer 1s (and more 0s) than $M_A^{\eta_2,\gamma_2}$. Stricter thresholds (higher η and lower γ) necessitate stronger evidence for dominance, thereby reducing the cardinality of the valid dominance pairs.

For any subsets $A, B \subseteq AS$, the dominance relation matrices on U with respect to A and B are denoted as:

$$M_A^{\eta,\gamma} = [m_A^{\eta,\gamma}(i,j)]_{n\times n} \quad \text{and} \quad M_B^{\eta,\gamma} = [m_B^{\eta,\gamma}(i,j)]_{n\times n}.$$

Definition 6: The intersection operation $\cap$ between $M_U^{\geq A}$ and $M_U^{\geq B}$ is defined as

$$M_A^{\eta,\gamma} \cap M_B^{\eta,\gamma} = [m_A^{\eta,\gamma}(i,j) \times m_B^{\eta,\gamma}(i,j)]_{n\times n}.$$

This operation yields a new dominance relation matrix that simultaneously captures the dominance relations of attribute sets A and B.

2) Properties of Dominance Relation Matrix Updates: Let $M_A^{\eta,\gamma} = [m_A(i,j)]$ denote the dominance relation matrix under attribute set A. The following properties hold for incremental updates.

Corollary 3. (Attribute Addition): When adding a new attribute set A^+, the updated dominance relation matrix $M_{A\cup A^+}^{\eta,\gamma}$ satisfies

$$M_{A\cup A^+}^{\eta,\gamma} = M_A^{\eta,\gamma} \cap M_{A^+}^{\eta,\gamma}$$

where $\cap$ denotes the elementwise logical AND operation. The dynamic update of the dominance matrix when attributes are added is shown in Table II.

Corollary 4. (Attribute Removal): When removing an attribute subset $A^- \subseteq A$, the updated matrix $M_{A\setminus A^-}^{\eta,\gamma}$ satisfies

$$M_{A\setminus A^-}^{\eta,\gamma} = M_A^{\eta,\gamma} \cup \neg M_{A^-}^{\eta,\gamma}$$

where $\cup$ denotes the elementwise logical OR operation and $\neg$ denotes logical negation. The dynamic update of the dominance matrix when the attribute is deleted is shown in Table III.

Corollary 5: $X_j \succeq X_i$ holds under $A \setminus A^-$ if it held under A *or* did not hold under A^-.

C. Feature Evaluation Metrics

This section defines CDD, GCDD, and MSIG based on the IV-ANCDFS model.

In an IV-ODS, we characterize CDD using average approximation results from all class samples. A larger CDD indicates better separability. The class-separability metrics are formalized as follows.

Definition 7: For any $A \subseteq AS$ and $t \in \{2, 3, \ldots, T\}$, the CDD of Dc_t under A is

$$\mathrm{CDD}_A(Dc_t) = \sum_{x_i \in Dc_t} \frac{|\underline{\underline{\mathrm{Dom}_A^{\eta,\gamma}}}(Dc_t)(x_i)|}{|\overline{\overline{\mathrm{Dom}_A^{\eta,\gamma}}}(Dc_t)(x_i)|}. \quad (13)$$

Definition 8: For any $A \subseteq AS$, the GCDD is defined as

$$\mathrm{GCDD}_A(D) = \frac{\sum_{t=2}^{T} \mathrm{CDD}_A^{\eta,\gamma}(Dc_t)}{T-1}. \quad (14)$$

For IV-ODS, classification discernibility positively correlates with both CDD and GCDD.

Definition 9: For $A \subseteq AS$ and $c_k \in AS \setminus A$, the MSIG of c_k to P is

$$\mathrm{MSIG}(c_k, A, D) = \mathrm{GCDD}_{A\cup\{c_k\}}(D) - \mathrm{GCDD}_A(D). \quad (15)$$

The MSIG metric exhibits a monotonic positive relationship with feature importance in ordinal classification.

D. Feature Evaluation Index and Selection Algorithm

We propose a feature evaluation index balancing class-separability and redundancy.

Definition 10: Given an IV-ODS $S^{\preceq} = (X, AS \cup \{D\}, V, F)$, for any subset $A \subseteq AS$, the MORE of X with respect to A is defined as

$$\mathrm{MORE}_X^{\preceq}(A) = -\frac{1}{|X|}\sum_{i=1}^{n} \log \frac{\left|\mathrm{DiagDom}_A^{+,\eta,\gamma}(x_i)\right|}{|X|}. \quad (16)$$

Definition 11: Furthermore, for any $A, B \subseteq AS$, the MORMI of X with respect to A and B is defined as

$$\mathrm{MORMI}_X^{\preceq}(A,B) = -\frac{1}{|X|}\sum_{i=1}^{n} \log \frac{\left|\mathrm{DiagDom}_{A\cup B}^{+,\eta,\gamma}(x_i)\right|}{|X|}. \quad (17)$$

Definition 12: For whatever $A \subseteq AS$, the MORCE of A to d is defined like this

$$\mathrm{MORCE}_X^{\preceq}(d,A) = -\frac{1}{|X|}\sum_{i=1}^{n} \log \frac{\left|\mathrm{DiagDom}_{\{d\}\cup A}^{+,\eta,\gamma}(x_i)\right|}{\left|\mathrm{DiagDom}_A^{+,\eta,\gamma}(x_i)\right|}. \quad (18)$$

Definition 13: For candidate feature c_k and selected subset A, the feature evaluation index

$$\sigma(c_k) = \mathrm{MSIG}(c_k, A, D) - \theta \cdot \mathrm{MORCE}^{\prec}(D, A \cup \{c_k\}) \quad (19)$$

where θ controls the redundancy penalty intensity.

This index enables efficient feature selection by maximizing GCDD gains while minimizing redundant information. The corresponding algorithm iteratively selects features with maximal $\sigma(c_k)$ values until stopping criteria are met.

V. Experiments and Analysis

This section establishes an evaluation framework to assess the IV-ANCDFS algorithm's performance across six dimensions, ensuring rigorous verification and practical applicability.

A. Compared Algorithms

Seven advanced feature selection algorithms have been chosen for comparative analysis.

Algorithm 1: IV-ANCDFS Algorithm.

Input: An IV-ODS $T = \langle X, AS \cup \{D\}, V, F \rangle$, parameters η, γ, θ
Output: Optimal conditional feature subset C_{opt}
1 Initialize $C_{opt} = \emptyset$;
2 **for** *each* $a_k \in AS$ **do**
3 $\quad M(a_k, \eta, \gamma) = \emptyset$
4 **end**
5 **for** *each* $a_k \in AS$ **do**
6 $\quad$ Calculate the dominance matrix $M(a_k, \eta, \gamma)$ by Definitions 5-6
7 $\quad$ **for** $t = 2$ **to** $T - 1$ **do**
8 $\quad\quad$ **for** $x_j \in \left(Dc_t^{\preceq}\right)^c$ **do**
9 $\quad\quad\quad$ Calculate the $\rho^{a_k}_{\left(Dc_t^{\preceq}\right)^c}(x_j)$;
10 $\quad\quad$ **end**
11 $\quad\quad$ **for** $x_j \in Dc_t^{\preceq}$ **do**
12 $\quad\quad\quad$ Calculate the $\rho^{a_k}_{Dc_t^{\preceq}}(x_i)$;
13 $\quad\quad\quad$ Compute the $\rho^{CDV}_{\left(Dc_t^{\preceq}\right)^c}(x_i)$ by Definition 1;
14 $\quad\quad\quad$ Compute the $\overline{\overline{\rho^{CDV}_{\left(Dc_t^{\preceq}\right)}}}(x_i)$ by Definition 2;
15 $\quad\quad$ **end**
16 $\quad\quad$ Compute $CDD_{a_k}(Dc_t)$ by Definition 7;
17 $\quad$ **end**
18 $\quad$ Compute $GCDD_{a_k}(Dc_t)$by Definition 8;
19 $\quad$ Calculate $MSIG(a_k, None, D)$ by Definition 9;
20 **end**
21 Find a_k with the maximum MSIG value;
22 $C_{opt} \leftarrow C_{opt} \cup \{a_k\}$;
23 **for** $s = 1$ **to** $|AS|$ **do**
24 $\quad$ **for** *each* $a_k \in AS$ **do**
25 $\quad\quad$ Calculate the $GCDD_{C_{opt} \cup \{a_k\}}(D)$ by Definition 8 Calculate the $MSIG(a_k, C_{opt}, D)$ by Definition 9;
26 $\quad\quad$ Calculate the $MORCE^{\preceq}(C_{opt}, a_k; D)$ by Definition 12;
27 $\quad\quad$ Calculate the $\sigma(a_k)$ by Definition 13
28 $\quad$ **end**
29 $\quad$ Find $a'_k = \arg\max_{a_k \in AS}\{\sigma(a_k)\}$;
30 $\quad$ $C_{opt} \leftarrow C_{opt} \cup \{a'_k\}$;
31 $\quad$ $AS \leftarrow AS - \{a'_k\}$;
32 $\quad$ $s \leftarrow s + 1$;
33 **end**
34 A feature sequence C_{opt} is obtained, i.e., $C_{opt} = \{a'_1, a'_2, \cdots, a'_l\}$;
35 Let $C_1 = \{c'_1\}, C_2 = \{c'_1, c'_2\}, \cdots, C_l = \{c'_1, c'_2, \cdots, c'_l\}$;
36 **for** $q = 1$ **to** l **do**
37 $\quad$ Use a specified classifier to five fold cross-validate the average classification accuracy $Acc(C_q)$ of the feature subset C_q;
38 **end**
39 $C_{opt} = \arg\max_{C_q \subseteq C_{opt}}\{Acc(C_q), q \in \{1, 2, \cdots, l\}\}$;
40 **return** C_{opt}

Algorithm 2: IV-ANCDFS-A/IV-ANCDFS-D Algorithm.

Input: An IV-ODS $T = \langle X, AS \cup \{D\}, V, F \rangle$, parameters η, γ, θ,dominance matrix$M(a_k, \eta, \gamma)$
Output: Optimal conditional feature subset C_{opt}
1 **Dynamic Update of Conditional Attributes:**
2 **while** *true* **do**
3 $\quad$ **if** *attribute addition* **then**
4 $\quad\quad$ $AS \leftarrow AS \cup AS'$;
5 $\quad\quad$ **for** *each* $a_k \in AS'$ **do**
6 $\quad\quad\quad$ Calculate the dominance matrix $M(a_k, \eta, \gamma)$ by Definitions 5-6;
7 $\quad\quad$ **end**
8 $\quad\quad$ Go back to Algorithm 1 Step 7;
9 $\quad$ **else if** *attribute deletion* **then**
10 $\quad\quad$ $AS \leftarrow AS - AS'$;
11 $\quad\quad$ **for** *each* $a_k \in AS'$ **do**
12 $\quad\quad\quad$ Delete the dominance matrix $M(a_k, \eta, \gamma)$;
13 $\quad\quad$ **end**
14 $\quad\quad$ Go back to Algorithm 1 Step 7;
15 $\quad$ **else**
16 $\quad\quad$ break;
17 **end**
18 **return** C_{opt}

TABLE IV
DATA DESCRIPTION

No.	Data sets	samples	features	classes
1	Period Changer	90	1177	2
2	Wine	178	13	3
3	Ionosphere	351	34	2
4	Dermatology	358	34	6
5	Parkinson	755	754	2
6	German credit	1000	20	2
7	Wireless	2000	7	4
8	Iranian churn	3334	13	2
9	Abalone	4177	8	3
10	Shill Bidding	6321	10	2
11	Electrical Grid Stability	10000	13	2
12	Magic gamma	19019	10	2

1) Original Attributes (Original Data): Direct classification using all conditional attributes without dimensionality reduction.

2) Interval-Valued Dominance-based Rough Set [23]: This method extracts features in interval-valued ordered systems using interval-based dominance rough approximations that measure overlap and dominance relationships.

3) *Active Antinoise Fuzzy Dominance Rough Feature Selection(AAnFSCF)* [31]:A noise-resistant method integrating AKNN denoising and fuzzy dominance rough sets to optimize features through fuzzy evaluation of class separability and redundancy.

4) *Matrix-Based Feature Selection Approach Using Conditional Entropy (HAR-A/HAR-D)* [27]:This approach dynamically adjusts features using matrix-based conditional entropy, with two efficient variants (addition and deletion) balancing accuracy and computational performance.

5) Infinite Feature Selection: A Graph-Based Feature Filtering Approach (INF-FS) [28]: This framework proposes a graph-based feature selection method, modeling feature subsets as graph paths and evaluating their relevance/redundancy via matrix or probabilistic computations for efficient ranking.

6) 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 handling data distribution challenges.

7) Incremental Reduction of Imbalanced Distributed Mixed Data Based on KNN Rough Set (IFS-KNCMI) [32]: This approach dynamically updates feature subsets in imbalanced

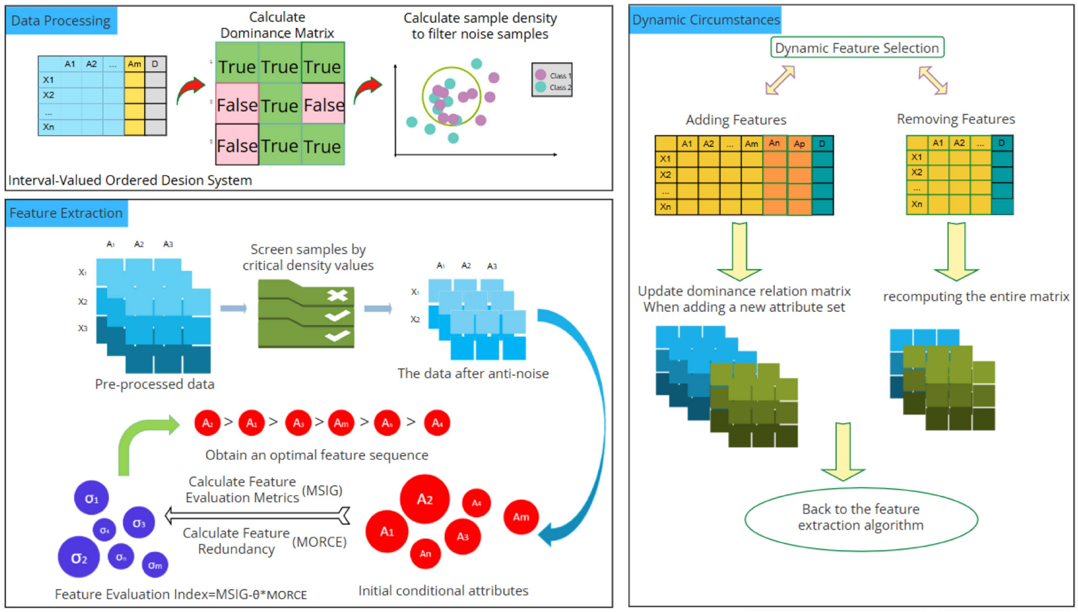

Fig. 3. Dynamic feature selection process for IV-ODS.

environments by integrating mutual information with max-relevance/min-redundancy criteria and k-NN rough approximations for neighborhood assessment.

8) Hyperspectral Band Selection Based On Rough Set (HKCMI) [33]: This rough set-based hyperspectral band selection method optimizes computational efficiency while selecting informative bands through relevance and significance analysis.

9) L1 Logistic Regression for Feature Selection (L1): This method employs L1 regularization to achieve sparsity in feature coefficients, effectively selecting relevant features by minimizing the loss function with an L1 penalty, which drives insignificant feature weights to zero.

10) Attention-Based Feature Selection (Attention): This approach utilizes attention mechanisms to dynamically assign importance weights to features, enabling context-aware feature selection based on learned attention scores that highlight contributive features.

The algorithms address distinct challenges: IV-DRSA handles interval-valued data via intersection/inclusion dominance for multisource fusion. AAnFSCF combats label noise through adaptive neighborhoods, while IFS-KNCMI addresses class imbalance via fuzzy dominance. HAR series and INF-FS optimize incremental feature evaluation using matrix entropy and path-weighted graphs, balancing accuracy–efficiency in streaming data. KNCMI/HKCMI tackle hybrid data with k-NN neighborhoods and granular-ball approximations. These dominance-based, entropy-driven methods adapt to data dynamics and labeling quality, offering context-specific solutions for classification systems. To enhance the comparative analysis, we incorporated additional baseline methods: an attention mechanism, as a deep learning-based approach, and L1-regularized logistic regression, a classical feature selection technique.

B. Experimental Data Preprocessing

The method framework starts with 12 benchmark datasets obtained from the UCI database (as shown in Table IV). The visualization of the dynamic feature selection process of IV-ANCDFS can be seen in Fig. 3. First, we assume that all initial datasets constitute clean data ($\mathcal{D}_{\text{clean}}$) without noise contamination. Then, we introduce 40% adversarial instances through label perturbation, specifically by random label exchanges between different classes: $(y_i \leftrightarrow y_j$ where $y_i \neq y_j)$

$$\mathcal{D}_{\text{noisy}} = \mathcal{D}_{\text{clean}} \cup \{x_i | x_i \in \text{ Every Class Values}\}_{i=1}^{0.4|\mathcal{D}|}.$$

To evaluate algorithmic robustness under dynamic feature spaces, each dataset underwent random stratified partitioning as follows:

$$\mathcal{F} = \mathcal{F}_{\text{core}} \cup \mathcal{F}_{\text{aug}}, \quad \mathcal{F}_{\text{core}} \cap \mathcal{F}_{\text{aug}} = \emptyset$$

where $\mathcal{F}_{\text{core}}$ represents the core feature subset (50% of total features) and $\mathcal{F}_{\text{aug}}$ denotes the incremental augmentation subset (the remaining 50%). This protocol simulates real-world scenarios where feature spaces expand incrementally, necessitating adaptive re-evaluation of feature relevance. To enhance uncertainty quantification, interval-valued representations were constructed for each class–attribute pair using the two-sigma principle

$$I_c = [\mu_c - 2\sigma_c, \mu_c + 2\sigma_c] \tag{20}$$

where μ_c and σ_c denote the original value of the conditional attribute and the SD of the class, respectively. An example using IV-ANCDFS algorithm is illustrated in Appendix 1 in the Supplementary Material (conditional attribute deletion is the opposite).

TABLE V
COMPARISON OF CLASSIFICATION PERFORMANCE OF DIFFERENT FEATURE SELECTION ALGORITHMS ON CLASSIFIER NB (%)

Datasets	Original Data	IV-DRSA	AAnFSCF	HAR	INF-FS	KNCMI	IFS-KNCMI	HKCMI	L1	Attention	IV-ANCDFS
period	55.56±9.30	52.98±11.70	65.56±4.16	56.67±10.77	60.00±11.86	**71.11±6.48**	46.67±23.57	54.44±5.44	62.22±10.77	67.77±9.55	66.66±8.66
wine	74.16±7.15	61.79±4.20	74.19±7.65	74.13±6.79	65.21±7.92	66.83±6.17	70.78±3.44	64.03±4.60	**74.73±6.78**	71.85±6.10	74.71±3.55
Iono	69.52±1.32	63.60±5.01	62.63±4.89	61.26±5.44	55.84±8.14	54.98±3.37	61.67±3.30	54.41±4.29	64.96±3.14	64.09±4.62	**69.81±3.69**
dermatology	79.63±6.44	48.35±2.72	30.42±3.91	80.71±3.96	87.97±2.70	88.00±2.83	80.55±2.02	81.30±3.93	79.62±6.15	**88.56±3.18**	81.86±4.83
parkinson	61.80±2.35	**73.26±1.44**	50.79±2.17	70.20±5.25	63.31±5.72	64.90±4.25	65.67±8.58	64.24±1.11	38.80±5.64	67.54±4.53	64.63±2.91
german	**70.00±0.00**	67.53±2.71	65.07±1.88	65.80±3.20	58.80±2.89	57.50±3.45	60.90±1.16	58.70±1.78	68.09±2.88	55.19±2.24	**70.00±0.00**
wireless	78.00±1.37	77.45±1.60	76.06±1.11	78.00±1.37	74.50±1.64	73.50±1.35	76.60±1.59	70.25±2.47	78.00±1.36	74.60±2.35	**78.40±1.55**
churn	74.60±0.41	79.82±2.68	70.68±0.50	76.06±2.25	79.49±0.51	79.30±0.84	84.06±0.77	79.30±1.67	84.06±0.76	79.33±1.05	**84.25±0.55**
Abalone	43.88±0.43	43.99±0.48	44.07±0.85	43.88±0.43	38.93±2.17	38.23±1.21	43.96±0.54	38.07±1.85	43.88±0.43	**45.36±1.53**	44.31±0.52
shill	88.18±0.41	87.24±0.54	72.97±0.85	87.96±0.43	85.33±0.65	85.29±0.61	85.30±0.10	84.89±0.37	**90.28±0.63**	87.96±0.55	**90.28±0.63**
Electrical	69.24±1.26	64.02±0.19	64.02±0.19	63.80±0.00	69.24±1.26	59.02±1.02	69.40±0.54	64.14±0.33	69.31±0.92	66.91±0.61	**69.37±1.21**
magic	65.78±0.59	64.39±0.33	65.75±0.58	65.31±0.54	57.84±0.42	57.94±0.37	65.27±0.53	56.60±0.40	65.70±0.52	**66.53±0.34**	65.57±0.68
Avg.	68.87±2.60	65.12±2.78	61.02±2.41	68.31±3.35	67.02±3.00	66.42±2.66	68.22±3.81	64.22±2.89	68.39±3.17	69.72±2.79	**71.65±2.45**

TABLE VI
COMPARISON OF CLASSIFICATION PERFORMANCE OF DIFFERENT FEATURE SELECTION ALGORITHMS ON CLASSIFIER SVM (%)

Datasets	Original Data	IV-DRSA	AAnFSCF	HAR	INF-FS	KNCMI	IFS-KNCMI	HKCMI	L1	Attention	IV-ANCDFS
period	71.22±4.97	69.49±0.72	64.44±2.72	71.11±5.44	60.00±8.31	66.67±7.54	70.00±7.20	64.44±3.14	72.22±7.85	64.44±6.66	**73.33±4.15**
wine	76.97±3.66	64.00±4.62	77.54±3.86	76.41±4.10	76.97±3.66	65.68±4.15	70.76±4.40	68.02±8.92	75.82±3.86	73.60±2.74	**78.66±3.70**
Iono	68.10±2.87	67.66±2.71	65.52±3.65	65.25±3.39	67.24±2.28	65.53±3.23	65.00±2.62	65.81±3.94	67.52±2.33	63.52±3.24	**69.53±4.09**
dermatology	85.62±1.55	65.22±4.64	51.88±5.79	89.66±0.76	91.34±1.86	58.06±12.70	89.11±1.59	85.21±2.68	91.05±0.74	89.38±1.14	**91.61±1.54**
parkinson	74.44±0.53	74.62±0.20	53.97±1.55	74.30±0.50	74.17±0.84	60.13±4.17	**75.67±0.47**	70.99±1.94	74.43±0.79	69.66±2.56	67.15±3.61
german	70.00±0.00	69.32±0.82	66.47±0.78	69.30±1.03	69.30±1.03	66.20±2.44	**70.50±0.55**	69.00±0.55	68.70±1.12	61.70±3.85	69.70±1.02
wireless	78.30±1.95	77.85±1.86	**90.40±1.74**	78.30±1.95	78.25±1.71	62.25±1.70	76.65±1.79	74.30±0.70	78.30±1.94	78.30±2.16	78.50±1.62
churn	84.76±0.66	84.66±0.64	70.99±0.58	84.51±0.66	84.51±0.78	70.89±0.46	84.63±0.77	84.70±0.41	84.47±0.36	**85.36±0.76**	84.98±0.61
Abalone	44.72±0.62	44.70±0.63	45.07±0.71	44.72±0.62	43.98±0.57	43.43±0.45	44.77±1.01	44.17±0.89	44.72±0.62	**45.72±1.48**	45.34±0.71
shill	90.06±0.63	89.46±0.20	73.43±0.70	90.06±0.63	90.06±0.63	73.49±0.73	90.04±0.26	88.02±0.41	90.28±0.63	90.04±0.53	**90.34±0.42**
Electrical	68.25±0.81	63.98±0.12	63.98±0.12	63.80±0.00	68.25±0.81	**70.87±0.27**	70.32±0.53	64.00±0.27	68.39±0.67	67.20±0.23	69.59±0.98
magic	66.95±0.53	64.76±0.37	**67.29±0.53**	66.95±0.53	66.95±0.53	65.91±0.62	64.65±0.09	65.79±0.58	67.01±0.54	67.28±0.41	66.94±0.53
Avg.	73.29±1.57	69.65±1.47	65.92±1.81	71.16±1.64	72.67±1.87	64.10±3.23	70.97±1.74	70.37±1.69	73.41±1.92	70.70±2.05	**73.81±1.92**

TABLE VII
COMPARISON OF CLASSIFICATION PERFORMANCE OF DIFFERENT FEATURE SELECTION ALGORITHMS ON CLASSIFIER CART (%)

Datasets	Original Data	IV-DRSA	AAnFSCF	HAR	INF-FS	KNCMI	IFS-KNCMI	HKCMI	L1	Attention	IV-ANCDFS
period	65.56±13.33	63.27±3.18	65.93±5.05	64.44±5.44	41.11±8.31	**73.33±6.48**	66.67±4.97	54.44±5.44	60.00±10.77	62.22±6.47	67.77±8.88
wine	70.21±6.19	58.46±5.49	70.73±5.11	65.27±6.52	**76.97±3.66**	49.38±4.26	70.76±4.40	74.46±6.79	65.69±5.98	70.20±2.99	70.25±3.59
Iono	63.25±1.28	56.09±5.48	53.84±1.65	61.24±3.39	67.24±2.28	53.55±3.05	65.00±2.62	64.68±3.85	57.82±9.78	55.54±3.61	**63.53±3.88**
dermatology	89.39±2.86	58.62±3.55	62.50±5.57	**89.66±0.76**	91.34±1.86	56.11±7.40	88.55±1.83	85.21±2.68	88.27±3.24	88.53±2.74	**89.66±2.91**
parkinson	67.15±3.62	63.73±2.09	59.60±0.53	**74.30±0.50**	74.17±0.84	58.81±3.39	69.00±3.27	74.04±1.35	63.97±4.63	68.74±3.68	61.45±3.33
german	69.60±0.80	59.14±2.69	65.20±0.00	58.80±2.89	58.80±2.89	65.90±1.32	68.20±1.36	69.00±0.55	66.29±3.17	61.10±1.15	**70.00±0.31**
wireless	73.80±1.39	73.32±1.53	**80.19±2.49**	74.50±1.64	74.50±1.64	50.45±1.04	71.15±2.10	74.30±0.70	73.94±1.76	73.60±0.95	74.49±1.85
churn	84.03±0.29	79.07±1.08	70.70±0.45	83.62±0.86	83.62±0.86	70.89±0.52	83.62±0.86	**84.70±0.65**	84.47±0.36	79.07±0.70	84.47±0.36
Abalone	40.10±1.76	38.66±1.65	41.70±1.35	39.21±0.62	39.21±0.62	43.62±0.79	39.21±1.68	**44.79±0.78**	38.90±2.31	38.90±0.66	40.19±1.87
shill	85.76±0.92	82.79±0.77	73.34±0.81	86.61±0.55	86.61±0.55	73.53±0.61	**90.76±0.53**	90.00±0.64	90.28±0.63	85.88±0.38	89.70±0.36
Electrical	59.00±1.17	54.32±0.79	54.32±0.79	54.61±0.89	**68.25±0.81**	87.02±0.34	60.28±0.78	55.44±0.87	59.53±0.58	60.15±0.47	59.70±0.70
magic	58.18±0.55	55.81±0.50	**58.43±0.33**	57.91±0.35	57.91±0.35	57.07±0.59	55.08±0.40	56.60±0.40	57.22±0.52	57.94±1.08	58.17±0.20
Avg.	67.98±2.81	61.61±2.45	62.99±1.88	67.56±1.95	68.31±1.96	61.64±2.48	68.99±1.97	68.94±2.01	67.45±3.58	66.49±2.20	**69.12±2.35**

C. *Classification Performance Evaluations of Algorithm*

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

The model's performance is critically dependent on the η and γ thresholds. To quantitatively demonstrate this relationship, three representative interval configurations were employed to visualize the impact of variations in η and γ on classification outcomes. The values of η and γ were sampled from 0 to 0.6 with a step size of 0.1. Furthermore, parameter optimization for θ was conducted using grid search within the interval $[0, 1]$ using a step size of 0.2. This systematic exploration ensures the identification of feature subsets that maximize discriminative power while maintaining computational efficiency.

The proposed methodology was rigorously evaluated against seven state-of-the-art feature selection algorithms. The classification performance was assessed using three established machine learning models: Naive Bayes (NB), support vector machine (SVM), and classification and regression trees (CART). Standard parameter configurations were maintained for all classifiers to ensure baseline comparability. A thorough fivefold cross-validation procedure was implemented. Comprehensive performance metrics across the three classifier configurations (NB, SVM, and CART) with 40% noise are systematically presented in Tables V–VII. Each table entry contains two evaluation metrics: mean classification accuracy (to the left of the "±" symbol) and SD (to the right of the "±" symbol). Top accuracy values per dataset are highlighted in *boldface notation*.

A systematic comparison of the proposed IV-ANCDFS against existing baseline methods across four critical dimensions is summarized in Table VIII. This comparative analysis highlights the unique and comprehensive advantages of our framework. Specifically, IV-ANCDFS is the only method that integrates all four key capabilities: robust *noise handling* through its neighborhood-based approximation, inherent *interval-value adaptability* for modeling uncertainty, a *dynamic update mechanism* for efficient feature space evolution, and manageable *computational complexity*. In contrast, while other baselines may excel in one or two aspects, they all exhibit significant limitations in others. For instance, IV-DRSA, although handling interval-valued data, lacks dynamic updating and is sensitive to noise. Methods, such as AAnFSCF and IFS-KNCMI, support

TABLE VIII
HORIZONTAL COMPARISON OF IV-ANCDFS AND BASELINE METHODS ACROSS FOUR KEY DIMENSIONS

Method	Noise Handling	Interval-Value Adaptability	Dynamic Update Mechanism	Computational Complexity
IV-ANCDFS	✓	✓	✓	$O(\|m\|^3 \cdot \|n\|^2)$
IV-DRSA	×	✓	×	$O(\|m\|^3 \cdot \|n\|^2)$
AAnFSCF	✓	×	✓	$O(\|m\|^2 \cdot \|n\|^2)$
HAR	×	×	✓	$O(\|m\|^3 \cdot \|n\|^2)$
INF-FS	✓	×	×	$O(n^2 + m^3(1+n))$
KNCMI	✓	×	×	$O(\|m\| \cdot \|n\|^2)$
IFS-KNCMI	✓	×	✓	$O(\|m\| \cdot \|n\| \cdot \|p\|)$
HKCMI	✓	×	×	$O(\|m\| \cdot \|n\|^2)$

TABLE IX
COMPARISON OF COMPUTATIONAL EFFICIENCY OF DIFFERENT FEATURE SELECTION ALGORITHMS (RUNNING TIME IN SECONDS)

Datasets	IV-DRSA	AAnFSCF	HAR	INF-FS	KNCMI	IFS-KNCMI	HKCMI	L1	Attention	IV-ANCDFS
period	25.2860	100.0877	14.5166	14.5515	130.8185	508.4884	7.0357	27.7065	8.8701	13.2936
wine	0.8824	3.9828	2.1653	0.1070	4.0903	18.9663	0.5887	6.1254	6.2488	0.4238
Iono	2.5879	44.4697	4.5902	0.2946	9.3275	20.4924	19.3595	9.9401	8.5745	0.4338
dermatology	3.6576	144.2309	6.2008	0.2585	8.8898	21.4321	36.0593	13.9812	8.3139	5.9970
parkinson	49.8680	2192.5819	20.3476	23.3050	450.0702	299.7887	37.1394	127.7922	24.0500	22.5490
german	3.6507	125.4126	2.2471	0.2706	12.9504	20.4134	7.2393	5.6914	17.5190	9.7756
wireless	1.0015	130.8502	1.3030	0.2066	14.5188	19.7639	1.4745	7.3562	31.3088	0.9122
churn	1.7687	130.8502	7.1196	0.4741	46.9442	20.9924	7.1515	6.0039	53.9402	4.2187
Abalone	77.3134	245.8867	4.5472	2.0325	79.6847	23.6700	6.0281	9.0186	74.4069	4.0372
shill	31.4774	315.6156	15.6816	0.5460	131.6749	20.6001	10.8328	6.2420	101.1772	3.4529
Electrical	88.9833	159.5250	29.2680	2.1653	227.2040	137.0908	19.6747	10.2873	178.9787	11.7008
magic	337.0755	484.3922	84.3822	18.7744	1540.8283	114.3518	40.4996	18.7765	360.2903	58.9803
Avg. Time	52.0464	336.4904	16.0801	**5.2655**	222.0002	119.5042	16.0076	20.8268	72.8065	**11.3145**

The bold values represent the first two algorithms with the shortest average running time among the compared methods.

dynamic updates but are fundamentally designed for crisp-valued data, lacking the formalism to process interval-valued information directly. This holistic integration of capabilities within IV-ANCDFS establishes a more versatile and powerful framework, providing a principled solution for complex feature selection tasks involving uncertain, noisy, and dynamically changing data environments.

D. Analysis of Time Complexity

To empirically validate the computational efficiency of the proposed IV-ANCDFS framework, we conducted extensive experiments measuring the actual running times across 12 diverse datasets, with the comparative results detailed in Table IX. The empirical data strongly corroborate the theoretical complexity advantages. IV-ANCDFS demonstrated robust and highly efficient performance, achieving an average running time of 11.3145 s. This solidly positions it as the second-fastest algorithm overall, being significantly more efficient than most competitors—notably, it is approximately 29.8 times faster than AAnFSCF (336.4904 s) and nearly 78% faster than its theoretical precursor, IV-DRSA (52.0464 s).

The algorithm's efficiency is particularly pronounced on complex datasets. For instance, on the "parkinson" dataset, IV-ANCDFS (22.5490 s) drastically outperformed AAnFSCF (2192.5819 s), showcasing its superior scalability. Furthermore, the consistently low running times across datasets of varying sizes and complexities validate the practical benefits of IV-ANCDFS. These results conclusively establish IV-ANCDFS not merely as a theoretically sound model but as a computationally feasible and highly efficient solution for practical feature selection tasks.

E. Parameter Sensitivity Analysis of Algorithm

This section investigates the parametric sensitivity of η, γ, and θ on the performance of the SVM classifier in Algorithm IV-ANCDFS, as demonstrated in Figs. 4 and 5.

For η and γ, we first know that the range of η is [−1, 1], and the range of γ is [0, 1]. These two parameters together determine the dominance relationship between interval values. When $\eta = 1$, there is a complete dominance relationship between the two intervals, while when $\eta = 0$, the dominance relationship is not obvious. Therefore, a smaller η value means more stringent restrictions, while an excessively large η value may lead to insignificant differences between classes. As far as γ is concerned, when $\gamma = 0$, it means that there is no overlap between the two intervals, and when $\gamma = 1$, the degree of overlap reaches the maximum. It also has an optimal value with the highest classification accuracy. In order to find the optimal value, we iterate the values of η and γ from 0.1 to 0.6, with a step size of 0.1, and then calculate the one-time reduction of the dataset and the classification accuracy of the reduced data.

For θ, the value range of θ is [0, 1], which is used to control the redundancy of the feature evaluation index in Definition 13. $\theta = 0$ means that there is no redundant penalty, and only those features that can improve class discrimination are selected; $\theta = 1$ means full redundancy penalty, selecting features that can improve class discrimination but do not increase too much

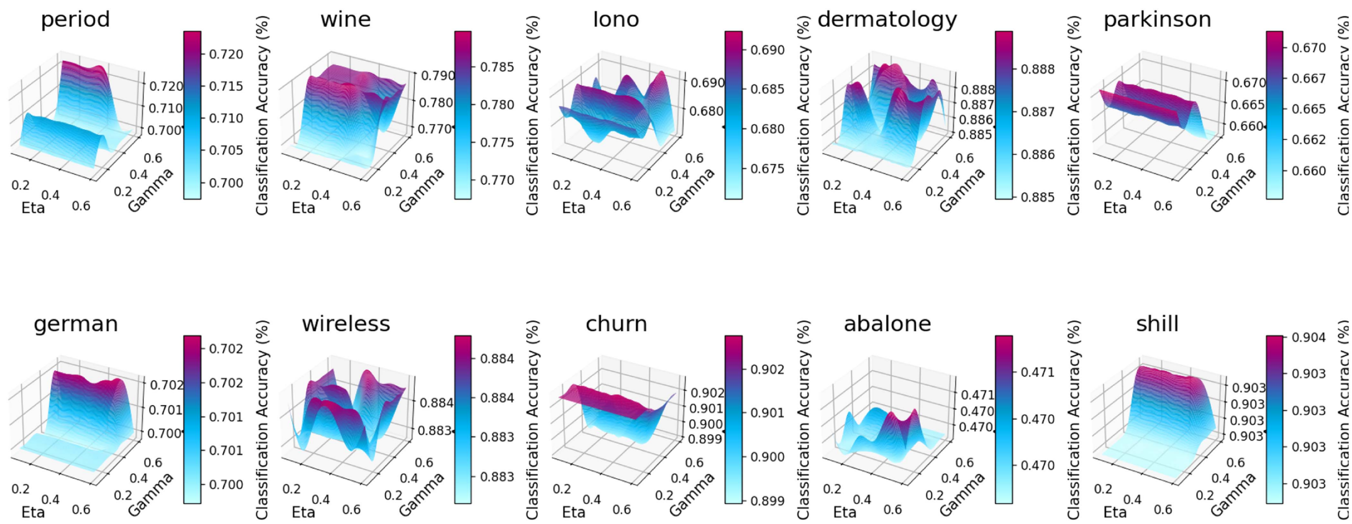

Fig. 4. Effect of parameters η and γ and the classfication accuracy on the classification accuracy tested by classifier SVM.

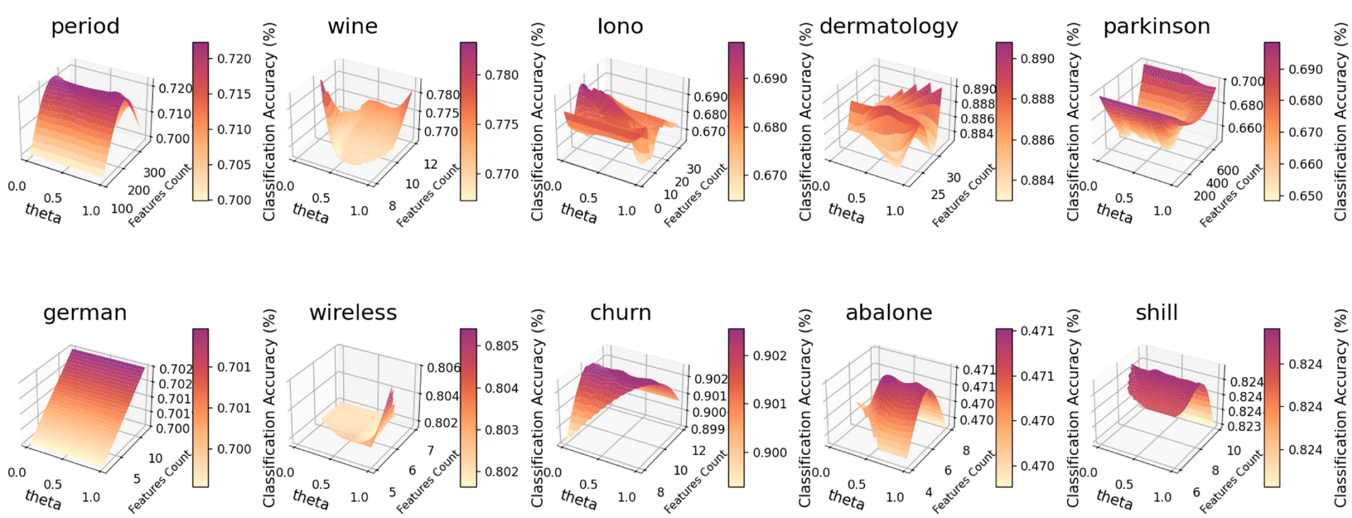

Fig. 5. Effect of parameter θ and the features count on the classification accuracy tested by classifier SVM.

uncertainty after adding. By adjusting θ, the behavior of the feature selection algorithm can be controlled to achieve different tradeoffs. In order to find the optimal value of θ, we iterate the value of θ from 0 to 1 with a step size of 0.2, and then calculate the one-time reduction of the dataset and the classification accuracy of the reduced data.

The experimental observations indicate that different datasets may have different optimal parameter values, and each dataset may also have multiple optimal values, which can be obtained by training on a specific dataset.

F. Robustness Evaluations of Algorithm

This section presents a methodological framework for assessing algorithmic robustness through feature sequence consistency and subset stability under varying noise conditions. The results are illustrated presented in Table X and Fig. 6.

For eight datasets listed in Table IV, we generated five synthetically corrupted variants with increasing noise ratios (0%–40% synthetic noise). To evaluate the robustness of feature scoring algorithms, we performed a comparative analysis of their stability across different noise levels. For each algorithm, we calculated the SD of feature importance values as the evaluation metric under conditions of incremental noise (0%–40% with 10% intervals). Specifically, given a feature f and its importance scores $\{s_0, s_1, \ldots, s_5\}$ at noise levels $\{0\%, 10\%, \ldots, 40\%\}$, the robustness metric is defined as

$$\sigma(f) = \text{SD}(s_0, s_1, \ldots, s_5) \tag{21}$$

where lower σ indicates better noise resistance.

Results from eight benchmark datasets were visualized through radial charts (see Fig. 6), where each axis represents $\sigma(f)$ for distinct features. The comparative analysis reveals that the IV-ANCDFS algorithm achieves the minimal coverage area across most subfigures, particularly notable in high-noise

TABLE X
ROBUSTNESS EVALUATION OF FEATURE SCORES FOR DIFFERENT ALGORITHMS UNDER DIFFERENT NOISE LEVELS (BASED ON AVERAGE PRECISION)

Datasets	IV-DRSA	HAR	INF-FS	AAnFSCF	KNCMI	IFS-KNCMI	HKCMI	L1	Attention	IV-ANCDFS
wine	0.792	0.816	0.821	0.826	0.820	0.775	0.762	0.875	0.858	**0.885**
period	0.682	0.638	0.638	0.684	0.727	0.658	0.638	0.730	0.615	**0.731**
Iono	0.774	0.784	0.782	0.794	0.776	0.740	0.778	0.796	0.767	**0.826**
dermatology	0.665	0.785	0.794	0.798	0.742	0.695	0.734	0.941	0.932	**0.947**
parkinson	0.676	0.660	0.700	0.787	0.702	0.630	0.692	0.705	**0.787**	0.720
german	0.665	0.659	0.659	0.670	0.667	0.670	0.661	**0.705**	0.630	**0.705**
wireless	0.799	0.799	0.800	0.798	0.793	0.786	0.754	0.882	0.881	**0.884**
churn	0.794	0.799	0.799	0.801	0.790	0.790	0.791	0.883	**0.903**	0.885
shill	0.838	0.841	0.841	0.848	0.849	0.837	0.831	0.938	**0.945**	**0.945**
abalone	0.498	0.496	0.491	0.498	0.480	0.490	0.493	0.499	0.503	**0.504**
Avg.	0.718	0.728	0.733	0.750	0.735	0.707	0.713	0.795	0.782	**0.803**

The bold values represent the algorithm with the highest robustness score under each data set.

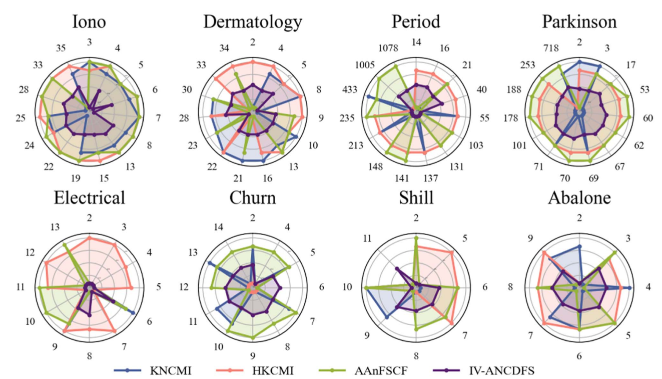

Fig. 6. Robustness evaluation of feature scores for different algorithms under different noise levels.

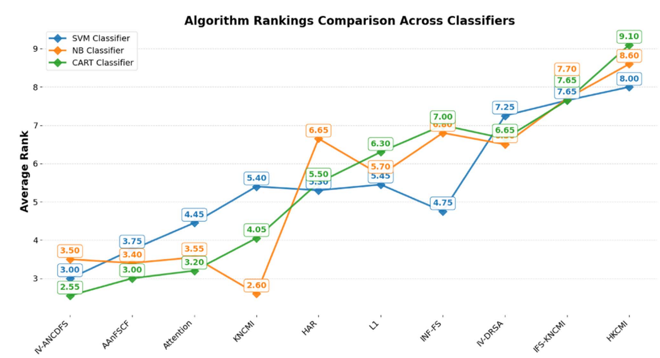

Fig. 7. Friedman test with Nemenyi posthoc analysis (α=0.05) on SVM, NB, and CART classifiers.

scenarios. This demonstrates that IV-ANCDFS maintains superior stability in feature importance quantification compared to other methods when handling contaminated data. The reduced chart coverage area directly correlates with enhanced algorithmic robustness, as it reflects lower variance in feature score distributions under noise perturbations.

G. Statistical Test

A multistage statistical evaluation process was implemented to assess eight feature selection algorithms across three distinct classifiers. The analysis combines nonparametric hypothesis testing with posthoc comparisons, supported by comprehensive ranking visualizations presented in Fig. 7.

The verification process employed an enhanced Friedman test with a significance level of $\alpha = 0.05$, addressing the limitations of conventional Friedman testing by improving the variance estimation. For $k = 8$ algorithms and $N = 10$ datasets, the hypothesis testing framework proceeded as follows.

1) *Null Hypothesis (H_0):* Equivalent classification performance across all the algorithms.
2) *Alternative Hypothesis (H_1):* Significant performance differences exist.
3) *Test Execution:* Rank-based evaluation utilizing maximum classification accuracy for each noise-free dataset.

The computed statistics revealed significant interalgorithm variations across all classifiers.

1) *Support vector machines:* $F_{\text{SVM}} = 12.34 \quad (p < 0.0001)$.
2) *Naive Bayes:* $F_{\text{NB}} = 8.91 \quad (p = 0.0028)$.
3) *CART:* $F_{\text{CART}} = 11.67 \quad (p < 0.0001)$.

Following hypothesis rejection, Nemenyi posthoc testing with critical difference $\text{CD} = 2.15$ identified specific performance disparities. The multiclassifier ranking visualization in Fig. 7 reveals the following.

1) *SVM Classifier:* IV-ANCDFS achieved superior ranking (1.32 ± 0.18) versus IV-DRSA's suboptimal performance (5.41 ± 0.22).
2) *Naive Bayes:* KNCMI dominated (1.55 ± 0.15) while IV-DRSA trailed (5.28 ± 0.19).
3) *CART:* IV-ANCDFS maintained leadership (1.44 ± 0.21) contrasting with IV-DRSA's inferior results (5.53 ± 0.25).

The proposed IV-ANCDFS demonstrated exceptional stability with minimal rank variance $(\Delta R = 0.42)$ across classifier paradigms, significantly outperforming IV-DRSA's maximum variability $(\Delta R = 3.89)$. The comprehensive rejection of H_0 ($p < 0.003$ for all classifiers) confirms statistically significant performance differences at 95% confidence level

$$\Delta R_{\mathcal{A}} = \max_{c_i \in \mathcal{C}} R_{\mathcal{A},c_i} - \min_{c_j \in \mathcal{C}} R_{\mathcal{A},c_j} \tag{22}$$

where $\mathcal{C} = \{\text{SVM}, \text{NB}, \text{CART}\}$ represents classifier set and $R_{\mathcal{A},c_i}$ denotes algorithm $\mathcal{A}$'s average rank with classifier c_i.

The statistical evidence substantiates the superior classification performance preservation of the IV-ANCDFS model across diverse learning paradigms.

VI. Conclusion

In conclusion, the IV-ANCDFS model presented in this study offers a robust and efficient solution for feature selection in dynamic, interval-valued data environments. By leveraging AKNN for noise suppression and employing matrix-based incremental updates, the model significantly enhances computational efficiency and noise immunity. The experimental results across multiple benchmark datasets demonstrate the superior performance of IV-ANCDFS in maintaining feature relevance and classification accuracy under varying attribute conditions. The proposed feature evaluation metrics and selection algorithms provide a systematic approach to balancing class separability and redundancy, making the model particularly suitable for real-world applications where data evolve over time. Future work may focus on extending the model to handle more complex data types and further optimizing the parameter tuning process to adapt to diverse operational scenarios.

References

[1] W. Zheng, S. Chen, Z. Fu, F. Zhu, H. Yan, and J. Yang, "Feature selection boosted by unselected features," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 9, pp. 4562–4574, Sep. 2022.

[2] Z. Pawlak, "Rough sets," *Int. J. Comput. Inf. Sci.*, vol. 11, no. 5, pp. 341–356, 1982.

[3] C. Liu, J. Lai, B. Lin, and D. Miao, "An improved ID3 algorithm based on variable precision neighborhood rough sets," *Appl. Intell.*, vol. 53, no. 20, pp. 23641–23654, 2023.

[4] W. Xu, D. Guo, J. S. Mi, J. Xian, K. Zheng, and W. Ding, "Two-way concept-cognitive learning via concept movement viewpoint," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 34, no. 10, pp. 6798–6812, Oct. 2023.

[5] M. Li, H. Wang, and J. Li, "Mining conditional functional dependency rules on Big Data," *Big Data Min. Anal.*, vol. 3, no. 1, pp. 68–84, 2020.

[6] J. Zhang, T. Li, D. Ruan, and D. Liu, "Neighborhood rough sets for dynamic data mining," *Int. J. Intell. Syst.*, vol. 27, no. 4, pp. 317–342, 2012.

[7] T. Shaheen, B. Mian, M. Shabir, and F. Feng, "A novel approach to decision analysis using dominance-based soft rough sets," *Int. J. Fuzzy Syst.*, vol. 21, no. 3, pp. 954–962, 2019.

[8] T. Tseng, C. Huang, K. Fraser, and H. Ting, "Rough set based rule induction in decision making using credible classification and preference from medical application perspective," *Comput. Methods Programs Biomed.*, vol. 127, pp. 273–289, 2016.

[9] X. Zhang and J. Hou, "A novel rough set method based on adjustable-perspective dominance relations in intuitionistic fuzzy ordered decision tables," *Int. J. Approx. Reason.*, vol. 154, pp. 218–241, 2023.

[10] J. Lin, W. Wen, and J. Liao, "A novel concept-cognitive learning method for bird song classification," *Mathematics*, vol. 11, 2023, Art. no. 4298.

[11] C. Lee and G. G. Lee, "Information gain and divergence-based feature selection for machine learning-based text categorization," *Inf. Process. Manage.*, vol. 42, no. 1, pp. 155–165, 2006.

[12] X. Zhao and W. Xu, "NFMPAtt-Unet: Neighborhood fuzzy C-means multi-scale pyramid hybrid attention Unet for medical image segmentation," *Neural Netw.*, vol. 178, 2024, Art. no. 106489.

[13] D. Guru, B. Kiranagi, and P. Nagabhushan, "Multivalued type proximity measure and concept of mutual similarity value useful for clustering symbolic patterns," *Pattern Recognit. Lett.*, vol. 25, no. 10, pp. 1203–1213, 2004.

[14] Q. H. Hu, L. Zhang, A. Shuang, D. Zhang, and D. R. Yu, "On robust fuzzy rough set models," *IEEE Trans. Fuzzy Syst.*, vol. 20, no. 4, pp. 636–651, Aug. 2012.

[15] X. D. Wu and X. Q. Zhu, "Mining with noise knowledge: Error-aware data mining," *IEEE Trans. Syst., Man, Cybern. A., Syst. Humans*, vol. 38, no. 4, pp. 917–932, Jul. 2008.

[16] X. Zhang and Z. Feng, "Feature selection based on contradictory state sequence for multi-scale interval valued decision table," *Inf. Sci.*, vol. 677, 2024, Art. no. 120926.

[17] P. P. Li, H. X. Zhang, X. G. Hu, and X. D. Wu, "High-dimensional multi-label data stream classification with concept drifting detection," *IEEE Trans. Knowl. Data Eng.*, vol. 35, no. 5, pp. 8085–8099, Aug. 2023.

[18] W. Xu, Z. Yuan, and Z. Liu, "Feature selection for unbalanced distribution hybrid data based on *k*-nearest neighborhood rough set," *IEEE Trans. Artif. Intell.*, vol. 5, no. 1, pp. 229–243, Jan. 2024.

[19] W. Xu and J. Li, "Granular-ball-matrix-based incremental semi-supervised feature selection approach to high-dimensional variation using neighbourhood discernibility degree for ordered partially labelled dataset," *Appl. Intell.*, vol. 55, no. 4, 2025, Art. no. 268.

[20] Y. Wu, P. Li, and Y. Zou, "Partial multi-label feature selection with feature noise," *Pattern Recognit.*, vol. 162, pp. 16397–16423, 2025.

[21] L. Farek and A. Benaidja, "An optimal feature selection method for text classification through redundancy and synergy analysis," *Multimedia Tools Appl.*, vol. 84, no. 16, pp. 16397–16423, 2025.

[22] L. Chen, J. Chen, and Y. Lin, "Feature selection considering synergy between features based on soft neighborhood rough sets," *Eng. Appl. Artif. Intell.*, vol. 150, 2025, Art. no. 110553.

[23] W. Li, H. Zhou, W. Xu, X.-Z. Wang, and W. Pedrycz, "Interval dominance-based feature selection for interval-valued ordered data," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 34, no. 10, pp. 6898–6912, Oct. 2023.

[24] W. Xu, M. Huang, Z. Jiang, and Y. Qian, "Graph-based unsupervised feature selection for interval-valued information system," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 35, no. 9, pp. 12576–12589, Sep. 2024.

[25] W. Xu and Z. Tian, "Feature selection and information fusion based on preference ranking organization method in interval-valued multi-source decision-making information systems," *Inf. Sci.*, vol. 700, 2025, Art. no. 121860.

[26] D. Guo and W. Xu, "Fuzzy-based concept-cognitive learning: An investigation of novel approach to tumor diagnosis analysis," *Inf. Sci.*, vol. 639, 2023, Art. no. 118998.

[27] W. Xu and Y. Yang, "Matrix-based feature selection approach using conditional entropy for ordered data set with time-evolving features," *Knowl.-Based Syst.*, vol. 279, 2023, Art. no. 110947.

[28] G. Roffo, S. Melzi, U. Castellani, A. Vinciarelli, and M. Cristani, "Infinite feature selection: A graph-based feature filtering approach," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 43, no. 12, pp. 4396–4410, Dec. 2021.

[29] W. Xu and W. Ye, "Incremental feature selection: Parallel approach with local neighborhood rough sets and composite entropy," *Pattern Recognit.*, vol. 159, 2025, Art. no. 111141.

[30] X. Zhang, J. Wang, and J. Hou, "Matrix-based approximation dynamic update approach to multigranulation neighborhood rough sets for intuitionistic fuzzy ordered datasets," *Appl. Soft Comput.*, vol. 163, 2024, Art. no. 111915.

[31] B. Sang, W. Xu, H. Chen, and T. Li, "Active antinoise fuzzy dominance rough feature selection using adaptive K-nearest neighbors," *IEEE Trans. Fuzzy Syst.*, vol. 31, no. 11, pp. 3944–3958, Nov. 2023.

[32] W. Xu and C. Liu, "Incremental reduction of imbalanced distributed mixed data based on *k*-nearest neighbor rough set," *Int. J. Approx. Reason.*, vol. 172, 2024, Art. no. 109218.

[33] S. Patra, P. Modi, and L. Bruzzone, "Hyperspectral band selection based on rough set," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 10, pp. 5495–5503, Oct. 2015.

[34] W. Xu, Y. F. Lin, and N. Wang, "A novel multi-source information fusion method based on dependency interval," *IEEE Trans. Emerg. Top. Comput. Intell.*, vol. 8, no. 4, pp. 3180–3194, Aug. 2024.

[35] K. Cai and W. Xu, "An efficient multi-source information fusion approach for dynamic interval-valued data via fuzzy approximate conditional entropy," *Int. J. Mach. Learn. Cybern.*, vol. 15, no. 9, pp. 3619–3645, 2024.

[36] W. Xu, Y. Liu, and T. Li, "Intuitionistic fuzzy ordered information system," *Int. J. Uncertain. Fuzziness Knowl.-Based Syst.*, vol. 21, no. 3, pp. 367–390, 2013.